



FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA PROFESIONAL DE INGENIERÍA DE COMPUTACIÓN Y SISTEMAS

**ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE DATAMART PARA
LA SECRETARÍA DE PLANIFICACIÓN ESTRATÉGICA MINEDU**

PRESENTADA POR
**MARYURY GARCÍA ANTICONA
KARLA JIMÉNEZ GARCÍA**

**TESIS PARA OPTAR EL TÍTULO PROFESIONAL DE
INGENIERO DE COMPUTACIÓN Y SISTEMAS**

LIMA – PERÚ

2015



**Reconocimiento - No comercial - Compartir igual
CC BY-NC-SA**

El autor permite transformar (traducir, adaptar o compilar) a partir de esta obra con fines no comerciales, siempre y cuando se reconozca la autoría y las nuevas creaciones estén bajo una licencia con los mismos términos.

<http://creativecommons.org/licenses/by-nc-sa/4.0/>



USMP
UNIVERSIDAD DE
SAN MARTÍN DE PORRES

**FACULTAD DE
INGENIERÍA Y ARQUITECTURA**

ESCUELA PROFESIONAL DE INGENIERÍA DE COMPUTACIÓN Y SISTEMAS

**ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE DATAMART
PARA LA SECRETARÍA DE PLANIFICACIÓN ESTRATÉGICA
MINEDU**

TESIS

**PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERA DE
COMPUTACIÓN Y SISTEMAS**

PRESENTADO POR

**GARCÍA ANTICONA, MARYURY
JIMÉNEZ GARCÍA, KARLA**

LIMA - PERÚ

2015

Agradecer a Dios por permitirme cumplir con mis metas. A mis padres, que siempre me han dado su apoyo incondicional y a quienes debo este triunfo profesional.

A todos los docentes, por entregar parte de su tiempo en nuestro desarrollo profesional y a todos los que luchan por sus sueños.

Maryury García Anticona.

Este trabajo va dedicado a mi familia y a todas las personas que me brindan su apoyo incondicional y creen en mí.

Karla Jiménez García

ÍNDICE

	Página
RESUMEN	ix
ABSTRACT	x
INTRODUCCIÓN	xi
CAPÍTULO I. MARCO TEÓRICO	1
1.1 Antecedentes de la investigación	1
1.2 Bases teóricas	5
1.3 Definición de términos básicos	23
CAPÍTULO II. METODOLOGÍA	26
2.1 Materiales	27
2.2 Métodos	31
CAPÍTULO III. DESARROLLO DEL PROYECTO	40
3.1 Gestión del Proyecto	40
3.2 Etapa de Planeamiento	61
3.3 Etapa de análisis	62
3.4 Etapa de diseño	78
3.5 Etapa de construcción	86
CAPÍTULO IV. PRUEBAS Y RESULTADOS	104
4.1 Pruebas	104
4.2 Resultados	105

CAPÍTULO V. DISCUSIÓN Y APLICACIONES	115
5.1 Discusión	115
5.2 Aplicaciones	119
CONCLUSIONES	120
RECOMENDACIONES	122
FUENTES DE INFORMACIÓN	123
ANEXOS	127

TABLAS

Tabla 1.1	Diferencia entre Sistema Tradicional y Data Warehouse	8
Tabla 1.2	Diferencias entre Data Warehouse y Datamart	14
Tabla 1.3	Diferencias entre Olap y Oltp	15
Tabla 1.4	Diferencias entre Molap y Rolap	17
Tabla 1.5	Herramientas ETL	21
Tabla 1.6	Comparación de herramientas de reporte	22
Tabla 2.1	Roles Asumidos para el desarrollo del proyecto	27
Tabla 2.2	Requerimiento de hardware	27
Tabla 2.3	Requerimiento de software	28
Tabla 2.4	Presupuesto del personal por hora	29
Tabla 2.5	Gastos de instalación del software	29
Tabla 2.6	Gastos del servicio de implementación de dispositivos de seguridad	29
Tabla 2.7	Gastos de instalación de hardware	30
Tabla 2.8	Gastos de dispositivos de seguridad	30
Tabla 2.9	Presupuesto del proyecto	30
Tabla 2.10	Definición de características de las metodologías	33
Tabla 2.11	Cuadro de criterios para elegir la metodología	34
Tabla 2.12	Cuadro comparativo de metodologías BI	35
Tabla 3.1	Requisitos del proyecto	45
Tabla 3.2	Fases requeridas	46
Tabla 3.3	Cronograma resumen de proyecto	48

Tabla 3.4	Línea base de costos	50
Tabla 3.5	Roles del proyecto	52
Tabla 3.6	Requisitos de comunicación de los interesados	53
Tabla 3.7	Matriz de comunicaciones	54
Tabla 3.8	Riesgos del proyecto	56
Tabla 3.9	Análisis de hacer o comprar	58
Tabla 3.10	Formato de criterios de evaluación y ponderación	60
Tabla 3.11	Registro de interesados	61
Tabla 3.12	Estados de matrícula	64
Tabla 3.13	Estados de nómina	66
Tabla 3.14	Definición de indicadores	68
Tabla 3.15	Cálculo de indicadores	69
Tabla 3.16	Variables de análisis	70
Tabla 3.17	Matriz de variables de análisis versus indicadores	71
Tabla 3.18	Variables de análisis	72
Tabla 3.19	Dimensiones con variables de análisis	73
Tabla 3.20	Dimensiones del Datamart	73
Tabla 3.21	Facts versus dimensiones	74
Tabla 3.22	Definición de reportes	77
Tabla 4.1	Cuadre de conteo de registros Fact matrícula	106
Tabla 4.2	Cuadre de conteo de registros Fact formato	106
Tabla 4.3	Resumen de resultados prueba funcionalidad	108
Tabla 5.1	Objetivos versus resultados	116
Tabla 5.2	Evaluación de tiempo proceso manual	117
Tabla 5.3	Evaluación de tiempo usando datamart	118

FIGURAS

Figura 1.1	Organigrama de la oficina de Informática del Ministerio de Educación	2
Figura 1.2	Implementación tradicional en cascada	7
Figura 1.3	Despliegue BI	7
Figura 1.4	Uso de las herramientas ETL en el proceso ETL	9
Figura 1.5	Tipos de consulta en diferentes niveles de los datos	11
Figura 2.1	Metodología del proyecto adaptada del Business Intelligence Roadmap	36
Figura 3.1	Flujo de control de cambios	43
Figura 3.2	Organigrama del proyecto	52
Figura 3.3	Estados de la matrícula	65
Figura 3.4	Estados de la nómina	67
Figura 3.5	Vista prototipo de matrícula	75
Figura 3.6	Vista prototipo de nóminas	75
Figura 3.7	Vista prototipo de estudiantes	76
Figura 3.8	Vista principal modelo de datos de Fact matrícula	78
Figura 3.9	Vista principal modelo de datos de Fact formato	79
Figura 3.10	Arquitectura de la solución	79
Figura 3.11	Diagrama de red	80
Figura 3.12	Configuración Pentaho con SQL	87
Figura 3.13	Configuración-Tableau con SQL	88

Figura 3.14	Carga de Facts en Datamart	89
Figura 3.15	Carga de registros al archivo intermedio Institución Educativa	90
Figura 3.16	Carga de dimensión Institución Educativa	90
Figura 3.17	Carga de registros al archivo intermedio estudiante	91
Figura 3.18	Carga de dimensión estudiante	91
Figura 3.19	Carga de registros al archivo intermedio ubicación	92
Figura 3.20	Carga de dimensión ubicación	92
Figura 3.21	Carga de registros al archivo modalidad Institución Educativa	93
Figura 3.22	Carga de dimensión modalidad Institución Educativa	93
Figura 3.23	Carga de registros al archivo estado matrícula	93
Figura 3.24	Carga de dimensión modalidad educativa	94
Figura 3.25	Carga de registros al archivo nivel educativo	94
Figura 3.26	Carga de dimensión nivel educativo	95
Figura 3.27	Carga de registros al archivo estado estudiante	95
Figura 3.28	Carga de dimensión estado estudiante	95
Figura 3.29	Carga de dimensión tiempo	96
Figura 3.30	Carga de registros al archivo detalle de formato	97
Figura 3.31	Carga de dimensión detalle de formato	97
Figura 3.32	Carga de registros al archivo estado formato	98
Figura 3.33	Carga de dimensión estado formato	98
Figura 3.34	Carga de registros al archivo tipos de formato	99
Figura 3.35	Carga de dimensión tipos de formato	99
Figura 3.36	Carga de los archivos intermedio STG	99
Figura 3.37	Fact Matrícula en Tableau	100
Figura 3.38	Fact Formato en Tableau	100
Figura 3.39	Dashboard matrícula	101
Figura 3.40	Dashboard estudiante	102
Figura 3.41	Dashboard nómina	103
Figura 4.1	Consulta TOP5 D_UBICACION	107
Figura 4.2	Pruebas: Dashboard estudiante	108

Figura 4.3	Reporte cantidad estudiantes	109
Figura 4.4	Reporte de estudiantes por Nivel Educativo del departamento de Amazonas	109
Figura 4.5	Resultado de carga de Fact matrícula	110
Figura 4.6	ETL Fact matrícula	111
Figura 4.7	Log de carga de Fact matrícula	112
Figura 4.8	Resultado de carga de Fact formato	112
Figura 4.9	ETL Fact formato	113
Figura 4.10	Log de carga de Fact formato	114
Figura 5.1	Muestra del Excel nómina de matrícula	117
Figura 5.2	Muestra del Excel de matrícula	117

RESUMEN

La educación, en la sociedad peruana, es uno de los soportes para superar con éxito sus desafíos más importantes. El Ministerio de Educación (MINEDU) es el ente encargado de velar por el desarrollo y progreso de la educación. Con la incorporación de las tecnologías de la información en las actividades educativas propicia que los volúmenes de datos sea cada vez creciente y constituye un activo clave para la explotación de la misma. En este contexto, la Oficina de Informática (OFIN) del MINEDU maneja el sistema transaccional de información de alumnos, matrículas, docentes, actas de evaluación y nóminas que da soporte a la Secretaría de Planificación Estratégica a través del proceso manual de generación de reportes. El desarrollo de esta tesis se centra en diseñar, construir e implementar un Datamart para la Secretaría de Planificación Estratégica, que brinda información útil para la toma de decisiones, maneja adecuadamente la información, reduce el tiempo carga y da soporte a las necesidades del área. Mientras más información se tenga, y esta sea más confiable y precisa, el proceso de toma de decisiones precisará más certeramente el éxito.

Palabras clave: Sistema Transaccional, Inteligencia de Negocios, Calidad Educativa, MINEDU.

ABSTRACT

Education in Peruvian society is one of the supports to successfully overcome their biggest challenges. The Ministry of Education (MINEDU) is responsible for overseeing the development and progress of education. With the incorporation of information technology in educational activities makes data volumes increase and growing and is a key asset to exploit it. In this context the Office of Computing (OFIN) handles transactional Student Information System, Enrollment, Teachers, Proceedings of Evaluation and Payroll that supports the Secretariat of Strategic Planning which will provide useful information for decision-making, properly manage information, reduce the time burden and support needs of the area. The more information you have, and this is more reliable and accurate, the process of decision making will target more accurately success.

Keywords: Transactional System, Business Intelligence, Education Quality, MINEDU.

INTRODUCCIÓN

La educación en la sociedad peruana es uno de los soportes para superar con éxito sus desafíos más importantes. El Ministerio de Educación (MINEDU) cumple uno de los papeles más importantes de garantizar a todos los niños, niñas, adolescentes y jóvenes del Perú una educación de calidad.

Con el avance de las tecnologías de información, los volúmenes de datos es cada vez creciente y constituyen un activo del MINEDU. Para obtener mayor conocimiento en la organización los gerentes y tomadores de decisiones requieren de un acceso rápido, fácil a información útil y valiosa para entender e identificar oportunidades, así como problemas potenciales.

MINEDU cuenta con varios sistemas transaccionales, por esta razón, debido a la necesidad que tiene la Oficina de Informática (OFIN) requiere implementar una solución de Inteligencia de Negocios para que apoye en la toma de decisiones para el proceso de Matrícula de las Instituciones Educativas. Motivadas por esta situación, la presente tesis propone brindar el análisis, diseño e implementación de un Datamart para la OFIN de MINEDU que contemple el Proceso de Matricula y Proceso de Nóminas.

Este trabajo ha sido dividido en cinco (5) capítulos, además aborda las conclusiones, recomendaciones y anexos.

En el capítulo I, se abordan los aspectos teóricos del proyecto así como las soluciones exitosas encontradas aplicadas a la mejora de la Calidad Educativa en otras partes del mundo, además se describen las terminologías y conceptos aplicados a lo largo del desarrollo de la tesis.

El capítulo II comprende los materiales y métodos que fueron usados para el desarrollo del proyecto. La importancia de realizar la selección correcta de la metodología y los implementos a utilizar implican un factor importante en el proyecto.

En el capítulo III, se explica el desarrollo del proyecto de acuerdo con el cronograma de trabajo, en el cual se verán las fases de la metodología aplicadas a la realización del Datamart propuesto. En el cuarto, se presentan las pruebas realizadas y los resultados obtenidos. En el quinto, se presentan las discusiones e interpretaciones de resultados además de la aplicación de la solución en MINEDU en la actualidad.

El problema es el engorroso proceso manual para disponer de información en los procesos de matrícula y nómina generando un debilitado apoyo a la toma de decisiones de la Secretaría de Planificación Estratégica del Ministerio de Educación.

Como objetivo general se propone mejorar el apoyo a la toma de decisiones en los procesos de matrículas y nóminas en la Secretaría de Planificación estratégica mediante la implementación del Datamart.

Los objetivos específicos son: Implementar un Datamart para la toma de decisiones de la Secretaría de Planificación Estratégica del MINEDU, identificar y definir los indicadores de los procesos de matrícula y nómina, construir los procesos ETL para la Extracción, Transformación y Carga de datos correspondientes al proceso de matrículas y de nóminas desde el SIAGIE.

La justificación teórica implica que la Inteligencia de Negocios o también conocido como Business Intelligence (BI) dispone de una gama muy amplia en lo que a herramientas y funciones se refiere, se complementan

con la integración y el control de calidad de datos para consolidar de forma precisa y coherente información procedente de fuentes diversas. Los cuadros de mando y otras técnicas de visualización permiten a los usuarios comprender rápidamente los resultados del análisis, lo que supone un componente crucial de las soluciones de Inteligencia de Negocios.

Al elegir un producto de Inteligencia de Negocios, además de las características específicas del producto, es importante tener en cuenta las ventajas que presenta, como la facilidad de uso, la facilidad de implementación, administración, la escalabilidad, la mejor visibilidad de hacia dónde debe ir la organización, el mejor uso de recursos, la cultura de análisis, las opciones de la interfaz de usuario y la integración en la plataforma actual y futura de su empresa.

Como justificación práctica tenemos que la toma de decisiones constituye una de las tareas principales de los cargos directivos de MINEDU. Si puede mejorar la calidad educativa en los procesos de toma de decisiones, podrá mejorar no sólo la efectividad en toda la organización, sino también la eficiencia en el mismo. La Inteligencia de Negocios ayudará a tomar mejores decisiones y se podrá gestionar de forma más eficiente y efectiva.

La Inteligencia de Negocios permite a la organización analizar los planes y medir las instituciones educativas y proporciona una visión de los aspectos que funcionan de forma satisfactoria e identifica las áreas para emprender acciones correctivas con la finalidad de facilitar las herramientas necesarias para llevar a cabo con total autonomía procesos de evaluación y a partir de estos contar con la información oportuna que permitan a la organización avanzar hacia la mejora continua.

La justificación social es la implementación de una solución de Inteligencia de Negocios apoyará a tomar mejores decisiones a través de los reportes de gestión en las Instituciones Educativas de educación básica regular de la información a los usuarios finales de MINEDU ya que al ser una entidad del estado influye en el progreso del país para alcanzar su potencial

y contribuir al desarrollo de la mejora de las Instituciones Educativas de manera descentralizada.

La justificación educativa plantea que el Ministerio de Educación, a través de la secretaría de Planificación Estratégica, necesita disponer de información consolidada como detallada de las nóminas de matrículas de las Instituciones educativas, para la toma de decisiones proactivas que ayuden a una mejor administración de los recursos en el sector educativo.

CAPÍTULO I

MARCO TEÓRICO

En este capítulo, se describe la situación actual del Ministerio de Educación (MINEDU) y las soluciones aplicadas de Inteligencia de Negocios en las instituciones educativas a nivel Mundial.

Asimismo, se mencionan los conceptos relacionados a la solución de Inteligencia de Negocios que se pretende brindar en la organización es decir se describe los conceptos, modelos, herramientas y en particular de la Calidad Educativa que se utiliza en la base teórica del proyecto que se viene realizando, culminando con la terminología del área de MINEDU.

1.1 Antecedentes de la investigación

El Ministerio de Educación es el órgano rector de las políticas educativas nacionales y ejerce su rectoría a través de una coordinación y articulación intergubernamental con los Gobiernos Regionales y Locales, propiciando mecanismos de diálogo y participación.

El Ministerio de Educación (MINEDU) fue fundado el 4 de Febrero de 1837 y cuenta en su estructura organizacional con la Secretaría de Planificación Estratégica.

La Secretaría de Planificación Estratégica es responsable de coordinar, integrar, formular, monitorear y evaluar la política, objetivos y estrategias del sector educación, cultura, deporte y recreación. Esta área tiene a su cargo a la Oficina de Informática (OFIN) encargada de establecer las políticas, normas y estándares, así como conducir el uso de los recursos informáticos en el Sector Educación (SPE, s.f.).

La Oficina de Informática (OFIN) cuenta con el sistema SIAGIE (Sistema de Información de Apoyo a la Gestión de la Institución Educativa), que es un aplicativo web informático que el Ministerio de Educación pone a disposición de las instituciones educativas públicas y privadas a nivel nacional a fin que gestionen la información de los procesos de matrícula, asistencia y evaluación de estudiantes (OFIN, s.f.). En la Figura 1.1, se muestra el Organigrama de MINEDU:

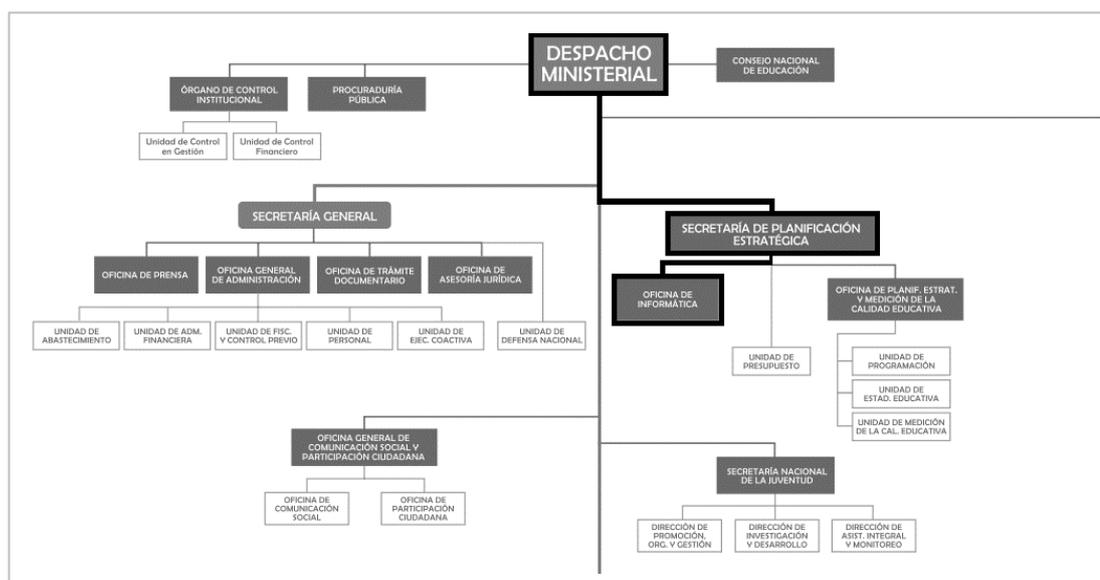


Figura 1.1 Organigrama de la Oficina de Informática del Ministerio de Educación
Fuente: Ministerio De Educación (OFIN, s.f.)

SIAGIE permite el aseguramiento de la aplicación estandarizada del marco normativo que regula los procesos de matrícula y evaluación de estudiantes, cuenta con una base de datos nacional el cual permite disponer de indicadores de gestión que apoyen la toma de decisiones a nivel de Instituciones Educativas, UGEL, DRE, Ministerio de Educación y otras entidades (SIAGIE, s.f.).

La capacidad para tomar decisiones en los negocios se ha convertido en una de las claves que permiten a las organizaciones ser exitosas y permanecer en el tiempo. En otras palabras, se trata de acceder a la información correcta en un formato apropiado y en un tiempo adecuado para optimizar los procesos internos de la organización.

Según la Oficina Regional de Educación para América Latina y el Caribe. (2013) menciona:

La incorporación de las nuevas tendencias tecnológicas juega un papel importante en la medición de los resultados educativos en las instituciones que se ha transformado con frecuencia en un problema para las autoridades educativas en todo el mundo. Una de las exigencias de los sistemas de evaluación educativa es que cada actor cuente con la información apropiada para las decisiones que debe tomar de acuerdo con su nivel, esto implica que autoridades, directivos escolares, docentes, estudiantes y familias deben tener acceso a la información pertinente de los resultados, de manera que apoye sus decisiones y sus responsabilidades, y sobre todo, que fortalezca el trabajo conjunto entre todos los actores para introducir los cambios que sean necesarios. Durante mucho tiempo los sistemas de medición han estado fuertemente limitados, en primer lugar, por los altos costos asociados a su implementación. Por otro lado, por la resistencia de los sistemas educativos y sus actores, asociados al uso que se ha hecho de la medición educativa. Actualmente los sistemas de medición han puesto una señal de alerta a las organizaciones sobre los verdaderos alcances y limitaciones de las mediciones de la calidad educativa.

Con los efectos de globalización, se ha ampliado el abanico de alternativas educativas para alumnos, profesores y patrocinadores en la búsqueda de las mejores opciones, tanto la educación pública como la privada se sitúan en este contexto, lo que a su vez ha generado una mayor competencia por el talento y los recursos financieros.

En países donde existe una gran demanda laboral, también existe una gran demanda de educación de calidad. En este contexto, desde hace algún

tiempo se está dando una nueva tendencia de la Inteligencia de Negocios en el sector de la educación, donde actualmente las universidades como Glamorgan en Reino Unido menciona uno de los beneficios “La implementación de una solución de Inteligencia de Negocios permitirá que los estudiantes tengan las mejores habilidades y conocimientos para hacerlo competitivos” (BI-Spain, 2010).

La Fundación Universitaria Konrad Lorenz (FUKL) de Colombia, para mejorar la calidad de la información y la integración en un solo sitio utilizó herramientas libres como propietarias como respuesta a la problemática de la información dividida sin relación tecnológica alguna, debido a esta situación se creó el proyecto de Desarrollo de un Sistema de Información con Inteligencia de Negocios para la Oficina de Egresados de la FUKL, permitiendo la consulta de información verídica, integrada y actualizada generando informes que respondan y soporten la toma de decisiones a nivel táctico como estratégico. “Les debe permitir generar “información” para la toma de decisiones, aplicando procesos de análisis de datos a través de una solución de Inteligencia de Negocios” (Forero, F., 2009).

La Universidad South Leicestershire College (SLC) en Reino Unido utilizó dashboards (cuadros de mando) intuitivos que permitan medir los indicadores de rendimiento en tiempo real. Suki Rai, Jefe de Planificación, Información y Financiación en South Leicestershire College, comenta: “Nos propusimos capacitar a nuestro personal con la herramienta adecuada para darles el control total sobre la información que afecta a su facultad o departamento. Necesitábamos una solución fácil, usar panel de control que ofrece un único punto de acceso a la información más relevante para la toma de decisiones más eficaz. Dashboards intuitivos es fundamental nos permite mejorar nuestro desempeño e identificar los aspectos operativos de la escuela, que exigen una acción inmediata” (Virtual Strategy, 2012).

El caso de la organización Columbus City Schools ubicado en EE.UU utiliza la tecnología de Information Builders en un entorno de Inteligencia de Negocios llamado Learning Circle SM (Círculo de Aprendizaje) que permite

examinar el historial de las evaluaciones cuatrimestrales, a través de reportes y cuadros de mando interactivos, para ayudar al personal administrativo a interpretar y comunicar el desempeño académico. La solución permite monitorear el estado académico de escuelas y estudiantes al utilizar los datos de logros disponibles (BI-Spain, 2012).

A nivel nacional no se han registrado casos de éxito en el sector de educación acerca de herramientas BI implementadas. En conclusión, la situación de las nuevas tendencias tecnológicas en Inteligencia de Negocios va creciendo exponencialmente en el sector educativo a nivel mundial brindando los medios para mejorar, aprovechar la información y satisfacer los niveles de educación que se requiere ante la era del conocimiento.

1.2 Bases teóricas

1.2.1 Inteligencia de Negocios vs Sistemas Tradicionales

Los ambientes competitivos de las organizaciones obligan a la alta dirección a buscar soluciones y estrategias para la generación de ventaja competitiva. Existen muchas definiciones pero aún no se ha llegado a un consenso de la misma.

La aparición de los ordenadores en la década de 1970 popularizó la utilización de las computadoras personales en los procesos administrativos, en ese momento se guardaron datos en la organización, creándose bases de datos históricos. Una de las formas es utilizar un conjunto de estrategias y herramientas enfocadas en la administración y creación de conocimiento mediante el análisis de estos datos llamada "Inteligencia de Negocios", también conocida como Business Intelligence (Gamarra, D., 2011).

Inteligencia de Negocios es un proceso interactivo que aplica métodos y conceptos para la explotación de la información en conocimiento

mediante la accesibilidad de la información con el fin de conducir de manera eficaz la organización a través del apoyo a la toma de decisiones a niveles estratégicos, tácticos y operacionales orientados al usuario final.

Los sistemas tradicionales no fueron diseñados o construidos pensando en integración. Cada sistema tiene un principio y un final, y cada sistema diseñado es para resolver un solo problema aislado de un conjunto de personas de una línea de negocio. Sin embargo, no son muy apropiados para las iniciativas integradas de Inteligencia de Negocios debido a que las viejas prácticas no incluyen ninguna de las actividades de la organización necesaria para mantener un entorno de apoyo a la toma de decisiones en toda organización (Moss, L. y Atre, S., 2003).

El desarrollo de un sistema no integrado de metodología en cascada, proporciona suficiente orientación para la planeación, elaboración, ejecución y sistemas independientes. Sin embargo, estas metodologías tradicionales no cubren la planificación estratégica, análisis de negocios de toda la organización o la evaluación de las nuevas tecnologías con cada proyecto; ni tampoco aceptar el concepto de lanzamientos de aplicaciones. Las metodologías tradicionales suelen comenzar con una necesidad de negocio funcional, entonces se concentran en el diseño y el desarrollo, y finalmente terminan en el mantenimiento. En la Figura 1.2, se muestra el flujo de la Implementación de un sistema tradicional:

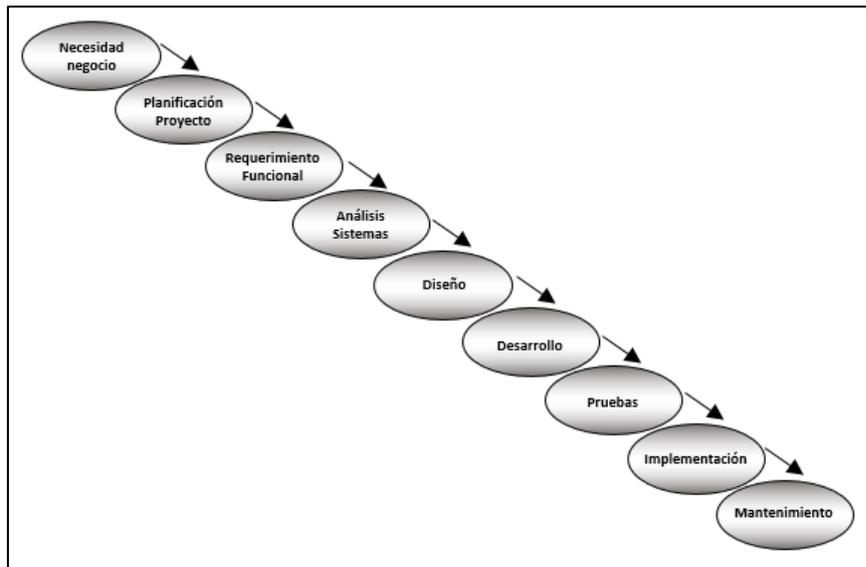


Figura 1.2 Implementación tradicional en cascada
 Fuente: Adaptado de the complete project lifecycle for decision (De Moss, L. y Atre, S.)

A diferencia de los sistemas independientes, un ambiente de BI es dinámico e integrado y da soporte a la toma de decisiones. Los datos y la funcionalidad deben ser versiones iterativas y cada despliegue desencadena nuevos requisitos para la próxima versión. En la Figura 1.3, se muestra el flujo de una solución BI:

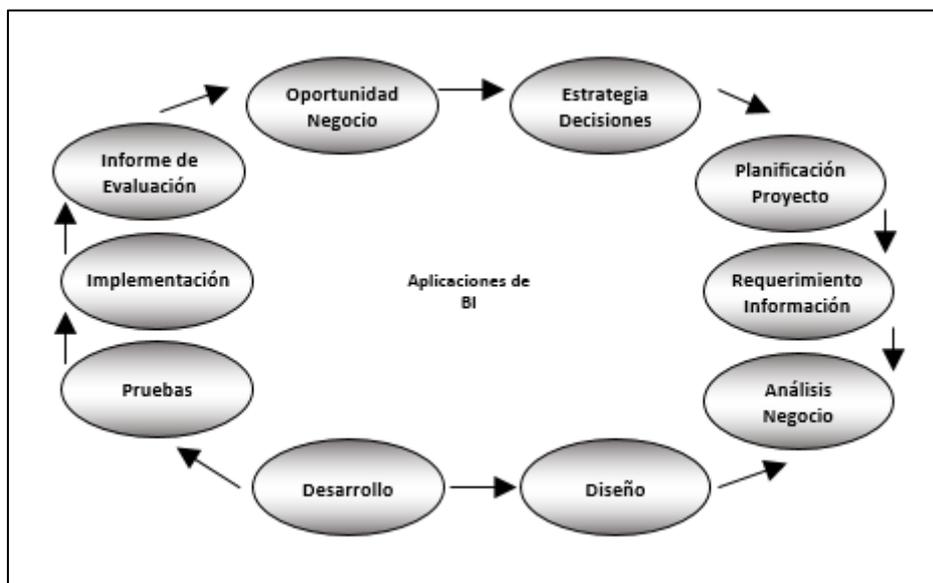


Figura 1.3 Despliegue Bi
 Fuente: Adaptado de the complete project lifecycle for decision (De Moss, L. y Atre, S.)

En la Tabla 1.1, se presentan las diferencias entre los sistemas tradicionales y datawarehouse:

Tabla 1.1 Diferencia entre Sistema Tradicional y Data Warehouse

SISTEMA TRADICIONAL	DATA WAREHOUSE
Predomina la actualización	Predomina la consulta
La actividad más importante es de tipo operativo (día a día)	La actividad más importante es el análisis y la decisión estratégica
Predomina el proceso puntual	Predomina el proceso masivo
Mayor importancia en la estabilidad	Mayor importancia al dinamismo
Datos en general desagregados	Datos en distintos niveles de detalle y agregación
Importancia en el dato actual	Importancia en el dato histórico
Estructura relacional	Visión multidimensional
Usuarios de perfiles medios bajos	Usuarios de perfiles altos
Explotación de la información relacionada con la operación de cada aplicación	Explotación de toda la información interna y externa relacionada con el negocio

Fuente: The complete project lifecycle for decision (De Moss, L. y Atre, S.)

1.2.2 Componentes de la Inteligencia de Negocios

Los componentes de la Inteligencia de Negocio son los siguientes:

1.2.2.1 Sistemas Fuente

Son los sistemas transaccionales diseñados para la recolección, almacenamiento y recuperación de todo tipo de información que se genera por las transacciones que produce la organización, estos sistemas se dirigen principalmente para el soporte de las operaciones del negocio como: Compras, Ventas, Almacenes, Contabilidad etc. En la mayoría de casos, en las organizaciones sus aplicaciones han sido modificadas a lo largo del tiempo y no integrada, parte de la gestión de Inteligencia de Negocios es lograr lo mencionado (Moss, L. y Atre S., 2003).

1.2.2.2 OLTP

Acrónimo de On Line Analytical Processing son aplicaciones que definen el comportamiento habitual de un entorno operacional de gestión y ejecutan las operaciones del día a día, tiene como

finalidad almacenar toda la información relativa al negocio y permitir el acceso on-line a todos los datos disponibles. Los sistemas OLTP se utilizan para soportar las tareas del día a día. Esta herramienta es buena para almacenar información y acceder rápidamente a ella, pero no es capaz de ofrecer rápidamente resúmenes de datos ni informes, ni tampoco almacena datos históricos (Wolff, G., 2002).

1.2.2.3 ETL

Acrónimo de Extract, Transform and Load. Se menciona las funciones como indica su acrónimo:

- Extracción de los datos de las distintas fuentes de información.
- Transformación de los datos al formato definido por los estándares de la organización.
- Cargar la información de las base de datos operacionales hacia las base de datos dimensionales (p.e. Data Warehouse).

La mayoría de las organizaciones prefieren utilizar una herramienta ETL para todos o algunos de los procesos ETL, especialmente para los procesos de transformación y extracción. Debido a la complejidad de diseño y desarrollo (Moss, L. y Atre S., 2003). En la Figura 1.4 se muestra el uso de las herramientas ETL en el proceso ETL:

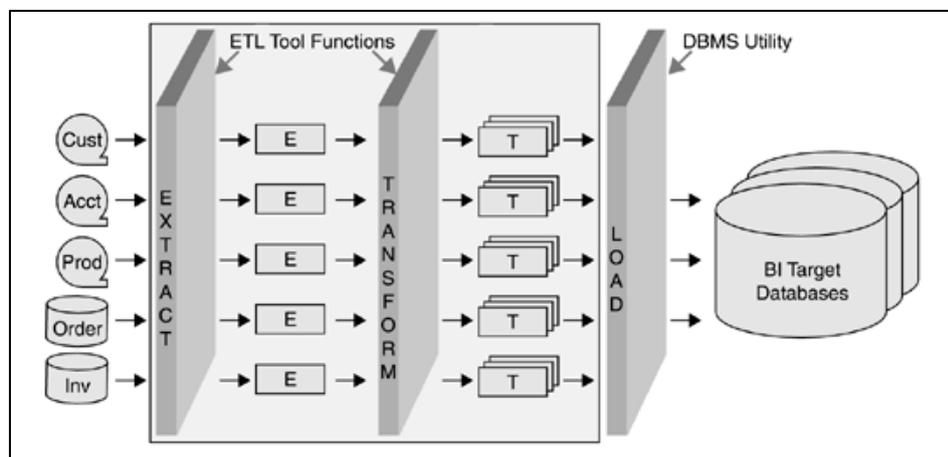


Figura 1.4 Uso de las herramientas ETL en el proceso ETL
Fuente: The complete project lifecycle for decision (De Moss, L. y Atre, S.)

1.2.2.4 Data Warehouse (DWH)

En los años 90, almacenamiento de datos fue propuesto como una solución adecuada al problema general de satisfacer las necesidades de información de gestión organizativa.

El primer almacén de datos es el Data Warehouse. El papel principal del almacén de datos es servir como un repositorio de datos que almacena los datos de diferentes fuentes, por lo que es accesible a otro conjunto de almacenes de datos - Datamart (Imhoff, C. et al, 2003).

Data Warehouse o Almacenamiento de datos es un gran almacén que contiene datos integrados, datos primitivos o históricos que transforma los datos operacionales en una herramienta competitiva, los usuarios finales en el entorno de Data Warehouse acceden directamente a los datos, utilizando herramientas de consulta amigable de usuario en lugar de basarse en informes generados por especialistas. Esto ayuda a reducir la dependencia del usuario en personal de TI para satisfacer las necesidades de información (Moss, L. y Atre, S., 2003).

EL estado de la arquitectura de Data Warehouse evoluciona de forma natural, la mayoría de negocios la arquitectura no es suficientemente robusta para satisfacer las necesidades futuras, es ahí donde interviene Data Warehouse, a continuación presentamos en la Figura 1.5 las diferencias de un Data Warehouse y Datamart con los sistemas tradicionales:

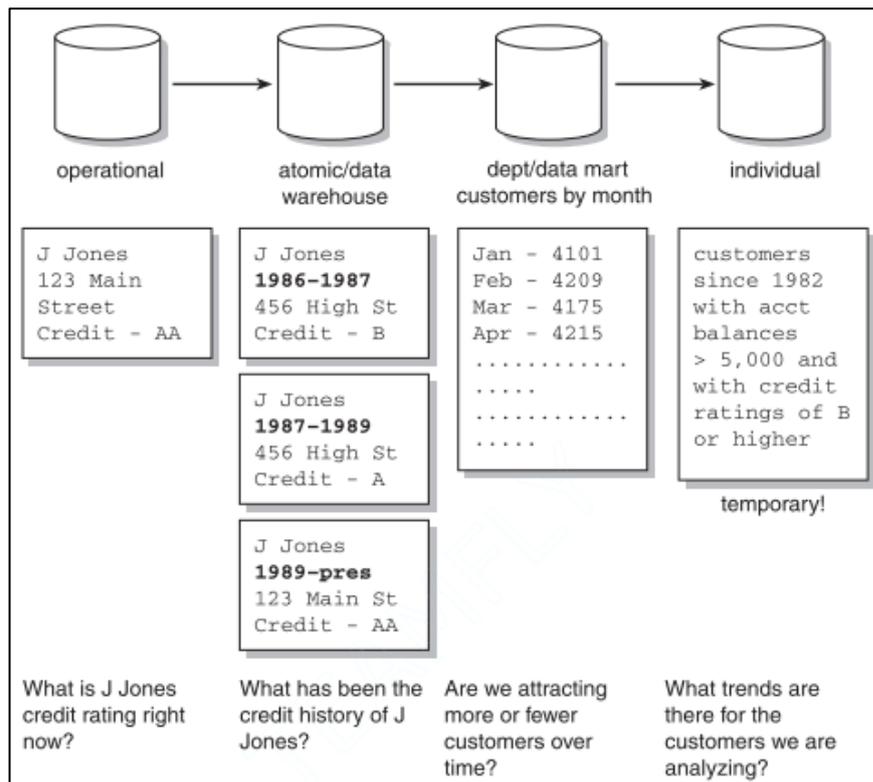


Figura 1.5 Tipos de consulta en diferentes niveles de los datos
Fuente: The complete project lifecycle for decision (De Moss, L. Y Atre, S.)

1.2.2.5 Datamart

Datamart son un subconjunto de almacén de datos y son donde la mayoría de las actividades analítica en el entorno de BI tienen lugar. Los datos en cada Datamart son por lo general a medida para una capacidad o función particular, como el análisis de la rentabilidad del producto, análisis KPI, análisis demográfico cliente, y así sucesivamente. Cada Datamart específico no es necesariamente válida para otros usos (Imhoff, C. et al, 2003).

Según el artículo Castillo, J. y Palomino, L. (2012) menciona la clasificación de los datamarts en tres tipos:

- **Datamart Dependiente:** Los Datamart dependientes son aquellos que reciben los datos desde un Data warehouse corporativo.

- **Datamart Independiente:** Los Datamart independientes son aquellos que reciben los datos directamente de los orígenes de información.
- **Datamart Híbrido:** Los Datamart híbridos permiten combinar las fuentes de datos de un Data warehouse corporativo con otras fuentes de datos tales como sistemas transaccionales y/o operacionales.

1.2.2.6 ODS

Acrónimo de Operational Data Storage. Es un sistema de procesamiento de transacciones, destinado a liberar a los sistemas operacionales con el fin de unificar la plataforma de almacenamiento donde se copien los datos de todas las aplicaciones en un solo sistema. El ambiente ODS es usado para la integración de los aplicativos fuentes en un modelo válido para toda la organización y se encuentra orientada por temas (clientes, productos, inventarios, etc.). ODS es la fuente los datos integrados en tiempo real donde se puede acceder desde cualquier parte de la organización (Imhoff, C. et al, 2003).

1.2.2.7 ODSSTG

Acrónimo de Stage Operational Data Storage que viene a ser una etapa previa al ambiente ODS. Es un área temporal donde se recogen los datos necesarios del sistema origen necesarios para la carga, se aplica la limpieza de datos, normalización y el mínimo de transformación de la información para su carga para el ambiente ODS. Es decir es el punto de comunicación de los sistemas origen y el DWH (Imhoff, C. et al, 2003). El diseño de esta capa sirve como input para la implementación de la herramienta OLAP.

1.2.2.8 BDS

Acrónimo de Business Data Storage. Forma parte de las buenas prácticas dentro de un Data warehouse. Este ambiente es usado

para contener la información resumida y agregada de lo que se tiene en ODS, realiza la exploración analítica de lo que ocurra en el negocio ya procesada y resumida de la información operacional y sus estructuras de datos se encuentran preparadas para procesar consultas con un mayor performance que un modelo relacional (Imhoff, C. et al, 2003).

1.2.2.9 Tecnologías OLAP

Acrónimo de On Line Analytical Processing son aplicaciones que se encargan de analizar datos del negocio para generar información táctica y estratégica que sirve de soporte para la toma de decisiones. Mientras que las transacciones OLTP utilizan Bases de Datos Relacionales u otro tipo de archivos, OLAP logra su máxima eficiencia y flexibilidad operando sobre Bases de datos Multidimensionales (Mendez, A. et al., 2003).

1.2.3 Data Warehouse vs Datamart

Datamart son almacenes de datos especializados, diseñados para soportar necesidades de análisis específicas para un único departamento o área funcional de la empresa, por ejemplo marketing, finanzas, producción, etc. Estos almacenes soportan menos usuarios y menos cantidades de datos que un Data warehouse centralizado, y por lo tanto pueden ser optimizados para cargar y recuperar la información de forma más rápida y eficaz que un Data Warehouse (T. Moss y Atre Shaku, 2003). En el caso de MINEDU se adecúa a las necesidades específicas para responder a los objetivos. A continuación en la Tabla 1.2 se presenta las diferencias que existen entre Data Warehouse y Datamart:

Tabla 1.2 Diferencias entre Data Warehouse y Datamart

	DATA WAREHOUSE	DATAMART
Alcance	Construido para satisfacer las necesidades de información de toda la organización.	Construido para satisfacer las necesidades de un área de negocios específica.
Objetivo	Diseñado para optimizar la integración y la administración de los datos fuente.	Diseñado para optimizar la entrega de la información de soporte de decisiones.
Características de los datos	Administra grandes cantidades de datos históricos a nivel atómico.	Se concentra en administrar resúmenes y/o datos totalizados.
Pertenencia	Pertenece a toda la organización.	Pertenece al área de negocio al cual está orientado.
Administración	Es administrado por la unidad de sistema de la organización.	Es administrado por el personal de sistema de la unidad propietaria del Datamart.

Fuente: Implementación de un Datamart como una solución de Inteligencia de Negocios para el área de Logística de T-Impulso (De Castillo, J., Palomino, I., 2012)

1.2.4 OLAP vs OLTP

Mientras que las aplicaciones OLTP se caracterizan por estar actualizadas constantemente por varios usuarios a través de transacciones operacionales sobre datos individuales, las aplicaciones OLAP son utilizadas por personal de niveles ejecutivos que requieren datos con alto grado de agregación y desde distintas perspectivas (dimensiones), por ejemplo: totales de venta por región, por producto, por período de tiempo, etc. (Mendez, A. et al., 2003). A continuación en la Tabla 1.3 se presenta las diferencias entre OLAP y OLTP:

Tabla 1.3 Diferencias entre Olap y Oltp

OLTP	OLAP
Atomizado	Sumarizado
Datos Históricos	Datos Actuales
Un registro a la vez	Muchos registros a la vez
Orientado a la información operativa	Orientado a la información estratégica
Datos relacionales	Datos Multidimensionales
Consultas simples predefinidas	Consultas ad-hoc
Volumen de datos acotados	Grandes volúmenes de datos

Fuente: The Complete Project Lifecycle For Decision (De Mendez, A. et al., 2003).

1.2.5 Técnicas y metodologías

1.2.5.1 Metodología de Desarrollo ROADMAP

Es una guía para el desarrollo de una aplicación eficaz de Inteligencia de Negocios (BI) empresarial (Moss, L. y Atre, S., 2003). RoadMap cuenta con un conjunto de estándares de TI y mejores prácticas de TI que administra la organización y control de la estructura, diseño, almacenamiento, movimiento, seguridad y calidad de la información.

1.2.5.2 Metodología de gestión del proyecto PMBOK

El PMBOK fue desarrollado por el Project Management Institute (PMI) y constituye un estándar para la gestión de proyectos, es un conjunto de procesos y cuenta con diez (10) áreas del conocimiento estas áreas corresponden a la Gestión de Integración, Gestión de Alcance, Gestión del Tiempo, Gestión de Costos, Gestión de Calidad,

Gestión de Comunicaciones, Gestión de Recursos Humanos, Gestión de Riesgos, Gestión de Adquisiciones, Gestión de los Interesados (Guía de conocimientos de la gestión de proyectos., 2013).

1.2.5.3 Técnicas de diseño

En el diseño del Datamart existen técnicas para estructurar la información destinada a la explotación, tradicionalmente los sistemas OLAP se clasifican según las siguientes categorías: MOLAP, ROLAP, HOLAP. En la Tabla 1.4 se presenta las diferencias entre MOLAP, ROLAP y HOLAP. A continuación, una breve descripción de cada una de ellas:

ROLAP: Acrónimo de Relational Online Analytical Processing, almacena los datos en un motor relacional logrando una mejor flexibilidad mediante los tipos de análisis disponibles, tener menor tiempo de respuesta para la elaboración de reportes, análisis de una enorme cantidad de datos. Se implementa sobre tablas físicas diseñadas siguiendo un modelo en estrella o copo de nieve (Tamayo, M. y Moreno F., 2006).

MOLAP: Acrónimo de Multidimensional Online Analytical Processing, almacena los datos de una base de datos multidimensional para la optimización de los tiempos de respuesta con estructuras optimizadas para acceso multidimensional, las matrices multidimensionales no admiten la ampliación dinámica o desbordamiento de la matriz lo cual lo hace poco dinámico pero a su vez con una gran capacidad de respuesta (Tamayo, M. y Moreno F., 2006).

HOLAP: Acrónimo Hybrid Online Analytical Process, almacena datos con las dos técnicas anteriores, utilizando MOLAP que ofrece análisis sobre los datos agregados, métricas o indicadores precalculados y ROLAP que ofrece escalabilidad, cálculo en tiempo real de reportes requeridos por usuarios,

conurrencia y administración madura de los datos (Tamayo, M. y Moreno F., 2006).

Tabla1.4 Diferencias entre Molap y Rolap

	MOLAP	ROLAP	HOLAP
Datos	Detalle y precalculados (agregados)	Detalle y agregados	Detalle y agregados
Estructura	Matrices comprimidas	Tablas relacionales	Multidimensional
Administración	Especialista en BDMD	Administrador BD	Administrador BD
Acceso	Lenguaje Especializado	SQL	SQL

Fuente: Análisis Del Modelo De Almacenamiento Molap Frente Al Modelo De Almacenamiento Rolap (De Tamayo, M. y Moreno, F., 2006)

1.2.6 Herramientas de ETL

InfoSphere DataStage (DS)

InfoSphere DataStage es una herramienta ETL que integra los datos a través de múltiples sistemas que utilizan un marco paralelo de alto rendimiento, y es compatible con la gestión de metadatos extendido y conectividad empresarial. La plataforma escalable proporciona una integración más flexible de todos los tipos de datos, incluyendo grandes datos en reposo o en movimiento, en plataformas distribuidas y de mainframe.

InfoSphere DataStage ofrece estas características y beneficios:

- ETL escalable plataforma apoya la recolección, integración y transformación de grandes volúmenes de datos, con estructuras de datos que van de lo simple a lo complejo.
- Cerca de integración en tiempo real-como datos así como la conectividad entre las fuentes de datos y aplicaciones.

- Carga de trabajo y las reglas de negocio de gestión-le ayuda a optimizar la utilización del hardware y priorizar las tareas de misión crítica.
- Facilidad de uso-ayuda a mejorar la velocidad, la flexibilidad y la eficacia para construir, desplegar, actualizar y administrar su infraestructura de integración de datos (DataStage, s.f.).

Server Integration Services (SIS)

SIS es una plataforma para la creación de soluciones empresariales de transformaciones de datos e integración de datos. Integration Services sirve para resolver complejos problemas empresariales mediante la copia o descarga de archivos, el envío de mensajes de correo electrónico como respuesta a eventos, la actualización de almacenamientos de datos, la limpieza y minería de datos, y la administración de objetos y datos de SQL Server.

SIS ofrece estas características y beneficios:

- El Motor de base de datos proporciona acceso controlado y procesamiento rápido de transacciones para cumplir los requisitos de las aplicaciones consumidoras de datos más exigentes de su empresa.
- Integration Services es una plataforma para generar soluciones de integración de datos de alto rendimiento, lo que incluye paquetes que proporcionan procesamiento de extracción, transformación y carga (ETL) para almacenamiento de datos.
- Master Data Services, se crea un repositorio central de los datos maestros y se mantiene un registro auditable y protegible de los mismos a medida que van cambiando con el tiempo.

- Reporting Services ofrece funcionalidad empresarial de poder crear informes que extraigan contenido a partir de diversos orígenes de datos, publicar informes con distintos formatos y administrar la seguridad y las suscripciones de forma centralizada (SIS, s.f.).

Microstrategy

Una completa infraestructura de servidor diseñada para conectarse a múltiples fuentes de datos, garantizando la compatibilidad con todos los estilos de analítica (desde distribución de informes hasta apps basadas en información, pasando por la visualización de datos autoservicio). Capaz de dar soporte a cientos de miles de usuarios y usado para el análisis de datos, informes y cuadros de mando, inteligencias móviles, minería de datos, la previsión, gestión de operaciones y toma de decisiones ejecutivas. También incluye herramientas de administración y monitorización para gestionar eficazmente la implementación de las analíticas; permitiendo de esta forma mejorar y predecir el comportamiento del negocio.

Microstrategy ofrece estas características y beneficios:

- Almacenes de datos relacional con terabytes a los mercados de datos más pequeños.
- Compatibilidad con fuentes de datos.
- Garantizar que cuadros de mando, informes y análisis de ejecutar lo más rápido posible (Microstrategy, s.f.).

IBM Cognos

De la inteligencia empresarial al rendimiento financiero y de la gestión de estrategias a las aplicaciones analíticas, el software IBM Cognos proporciona a su organización todo lo necesario para basarse en la analítica y lograr el máximo rendimiento. Con productos a nivel individual, de grupo de trabajo, departamento, medianas empresas o grandes organizaciones, el software de

Cognos está diseñado para ayudar a todas las personas de su organización a tomar decisiones enfocadas a incrementar los resultados de negocio, para el presente y el futuro (Cognos, s.f.).

Pentaho

Es una herramienta de código abierto y de uso libre muy completa, pues incluye elaboración de reportes, cubos, data mining, ETL y una plataforma BI.

Pentaho ofrece estas características y beneficios:

- Extracción, transformación y carga gráfica (ETL) herramienta para cargar y fuentes de datos grandes de proceso en forma familiar.
- Extraer y combinar los datos existentes y diversos para producir uniformes y de alta calidad listos para analizar los datos.
- Una mayor flexibilidad, reducción del riesgo, y el aislamiento de los cambios en el gran ecosistema de datos.
- Perfil de los datos y garantizar la calidad de datos con capacidades integrales para los administradores de datos (Pentaho, s.f.).

Para seleccionar la herramienta a utilizar en el desarrollo del proyecto en la Tabla 1.5 se presenta un cuadro comparativo de las herramientas para el proceso ETL, para la comparación se han tomado siete criterios, como resultado de la evaluación en base a los criterios se ha seleccionado la herramienta Pentaho debido a que es una herramienta Open Source lo que permitirá ahorrar costos para la Oficina de Informática además soporta múltiples plataformas Windows y Linux o ambos, aunque es una herramienta Open Source no cuenta con soporte; sin embargo, hay una comunidad donde brinda ayuda de la herramienta.

Tabla 1.5 Herramientas ETL

HERRAMIENTA	1 PERMITE ETL	2 OPEN SOURCE	3 SOPORTE DE LA HERRAMIENTA	4 SOPORTE DE MÚLTIPLES PLATAFORMAS	5 ELABORACION DE REPORTE	6 ANALISIS DE INFORMACION EN TODAS LAS DIMENSIONES	7 EXPORTACIÓN A ARCHIVOS
DATASTAGE	SI	NO	SI	SI	NO	SI	NO
SIS	SI	NO	SI	SI	SI	SI	SI
MICROSTRATEGY	SI	NO	SI	SI	SI	SI	SI
IBM COGNOS	NO	NO	SI	SI	SI	SI	SI
PENTAHO	SI	SI	NO	SI	SI	SI	SI

Fuente: Análisis, Diseño e Implementación de Datamarts para las áreas de ventas y recursos humanos de una empresa dedicada a la exportación e importación de productos alimenticios (De Moreno, H., 2013)

1.2.7 Herramientas de Explotación

Business Objects

Simplifican la manipulación de datos, lo que permite a los usuarios acceder, navegar, analizar, dar formato y compartir información a través de un entorno corporativo. Permiten una amplia gama de procesos, desde la búsqueda y navegación para el análisis avanzado, consulta la empresa, elaboración de informes y análisis, cuadros de mando y visualización, y la gestión de la infraestructura de información (Business Objects, s.f.).

Qlikview

La plataforma Business Discovery de QlikView aporta un verdadero BI a demanda que capacita a los usuarios de negocio para que tomen decisiones de forma innovadora, facilita la vida de analistas de negocio y desarrolladores de Business Intelligence al crear aplicaciones de BI muy avanzadas y específicas que los usuarios pueden ampliar y mejorar a medida que van cambiando las necesidades del negocio (Qlikview, s.f.).

Tableau

Tableau Software (NYSE: DATA) ayuda a las personas a ver y comprender datos. Tableau ofrece análisis rápidos e inteligencia de negocios de alta velocidad, crea visualizaciones y controles de mandos en minutos. El resultado es obtener respuestas rápidamente a partir de los datos, sin necesidad de programación (Tableau, s.f.).

Para seleccionar la herramienta de explotación a utilizar en el proyecto en la Tabla 1.6 se presenta un cuadro comparativo de las herramientas de explotación, para la comparación se han tomado el mayor puntaje, como resultado de la se ha seleccionado la herramienta Tableau, el cual utilizaremos para la solución que se está proponiendo en este proyecto.

Tabla 1.6 Comparación de herramientas de reporte

Nro.	Criterio	Business Objects	Qlikview	Tableau
1	Herramienta de Desarrollo	4	5	5
2	Administración de Metadatos	5	3	3
3	Mantenimiento	4	3	3
4	Robustez	5	3	4
5	Colaboración	5	4	4
6	Tableros de Control	3	4	4
7	Reportes Analíticos	4	4	4
8	Capacidad de Exploración	3	5	5
9	Integración con Office	4	4	4
10	OLAP	4	3	2
11	Visualización Interactiva	3	4	5
12	Costo	2	4	5
13	Total Puntaje	55	54	56

Elaboración: Las autoras

1.2.8 ISO/IEC 9126-1:2001

Es la norma internacional para la evaluación de la calidad del Software. Su objetivo fundamental es abordar algunos de los prejuicios humanos que pueden afectar negativamente a la entrega y percepción de un proyecto de desarrollo de software. El uso de esta norma en la presente tesis

es para seleccionar los criterios de evaluación que permita seleccionar la metodología para el desarrollo del proyecto (ISO/ IEC 9126, s.f).

1.2.9 PMBOK

Es la Guía de los Fundamentos para la Dirección de Proyectos, proporciona pautas para la dirección de proyectos individuales y define conceptos relacionados con la dirección de proyectos. Describe asimismo el ciclo de vida de la dirección de proyectos y los procesos relacionados, así como el ciclo de vida del proyecto (PMBOK, 2013).

1.3 Definición de términos básicos

Agregación: Son resúmenes de datos precalculados que mejoran el tiempo de respuesta de consultas al tener preparadas las respuestas antes de que se planteen las preguntas.

Cubo: Estructura de datos multidimensional que presenta la intersección de una combinación de dimensiones.

Datamart: Conjunto de datos estructurados que contiene información de un área en particular.

Datamining: Análisis de los datos para descubrir relaciones, patrones, o asociaciones desconocidas.

Diccionario de datos: Un compendio de definiciones y especificaciones para las categorías de datos y sus relaciones.

Dimensión: Una dimensión es un objeto de negocio a través del cual la información es agrupada con algún fin. Una dimensión viene a ser una

variable con la que se realiza algún tipo de análisis (productos, clientes, departamento, etc.).

Esquema en estrella: Organización física de los Datamarts que facilita el acceso a los datos y al análisis. Se caracteriza por tener una tabla central de hechos rodeada por tablas de dimensiones que contienen información desnormalizada de los hechos.

ETL (Extraction, Transformation and Loading): herramientas dedicadas a la extracción de los datos desde las fuentes donde estos se encuentren a los Datamarts.

Evaluación: Proceso por el cual los esfuerzos del programa, actividades o resultados son analizados y valorados en relación a los objetivos trazados por el proyecto educativo, utilizando un conjunto de indicadores para su medición, que permitirá tomar decisiones basadas en evidencias y no decisiones intuitivas.

Fact Table: Una fact o tabla de hechos representa un arreglo de datos pre calculado alrededor de la cual se realiza el análisis. Una fact tiene el mayor nivel de información que es sintetizada a través de diferentes dimensiones.

Foreign keys: Una foreign key es la representación de la primary key de una dimensión en la tabla de hechos. Las foreign keys son tomadas de las primary keys para cada dimensión de la tabla. La combinación de estas claves es el identificador para cada registro de la tabla de hechos.

Multidimensional: Indicadores que conforman una base de datos y que se analizan en función de varios criterios, las dimensiones.

Tablero o Cuadro de Mando (Dashboards): Sistema de gestión que permite a las organizaciones clarificar sus visiones y estrategias para trasladarlas a acciones.

OLAP (On-line Analytical Processing): Conjunto de principios que proveen una ambiente de trabajo dimensional para soporte decisional.

OLTP (On-line Transaction Processing): Sistema transaccional diario (o en detalle) que mantiene los datos operacionales del negocio.

SSL (Secure Sockets Layer): Es un protocolo que hace uso de certificados digitales para establecer comunicaciones seguras a través de Internet predecesor del TLS (Transport Layer Security).

Llave Pública: Usada para cifrar la información, haciendo una analogía, es como la llave utilizada para cerrar una puerta y mantener fuera a cualquier persona.

Llave Privada: Usada para cifrar la información, haciendo una analogía, la llave que abre la puerta y sólo la posee la persona autorizada, por lo tanto esta debe mantenerse en secreto.

CA (Autoridad Certificadora): Es una entidad de confianza del emisor y del receptor de una comunicación. Esta confianza de ambos en una 'tercera parte confiable' (trusted third party) permite que cualquiera de los dos confíe a su vez en los documentos firmados por la Autoridad Certificadora, en particular, en los certificados que identifican ambos extremos.

CAPÍTULO II METODOLOGÍA

En el presente proyecto, se ha dado uso de la Investigación Aplicada y la Investigación Documental, dado que nos permite identificar y evaluar alternativas de mejora de la calidad educativa, sustentada en los conocimientos adquiridos en la carrera profesional, además de tomar como base información de diferentes fuentes, todos ellos referidos a la Inteligencia de Negocios, Administración de la Información y Gestión de toma de decisiones, que validan la realización del presente proyecto.

La investigación documentaria se pudo realizar con una recopilación y clasificación de información acerca de las soluciones para mejorar la medición de la calidad educativa y con ello se pudo definir el producto final que se desea obtener. Los documentos que han sido revisados fueron proveídos por la oficina de Informática del MINEDU, los cuales muestran su operatividad actual. A continuación, se presenta la lista de los documentos revisados:

- Archivo Fuente de SIAGIE (2014)
- Boletín : PISA 2015
- Reglamento de la Ley N° 28044, Ley General de Educación (aprobado por Decreto Supremo N° 011-2012-ED)
- Planes operativos y estratégicos del sector educativo.

2.1 Materiales

2.1.1 Recursos humanos

En la Tabla 2.1, se presentan los roles que participaron para el desarrollo del proyecto:

Tabla 2.1 Roles asumidos para el desarrollo del proyecto

ROL	ABREV	RESPONSABLE
Gestor de Proyecto	GP	Maryury García.
Analista Funcional	AF	Karla Jiménez.
Diseñador del modelo	DM	Maryury García.
Analista Técnico *	AT	Rubén Rojas.
Certificador	CT	Karla Jiménez.

*Las funciones de Analista Programador son realizadas por el Analista Técnico.

Elaboración: Las autoras

2.1.2 Hardware

En la Tabla 2.2, se presentan los requerimientos mínimos para el óptimo funcionamiento de la solución a implementar:

Tabla 2.2 Requerimiento de Hardware

EQUIPO	CARACTERÍSTICAS	CANTIDAD
CPU	INTEL CORE DUO- HDD 250MB	2
MONITOR	LCD-COLOR 32 BITS	2
SERVIDOR DE BD	INTEL CORE DUO- 3.0GHZ- RAM 2GB- HDD 1TB	1

Elaboración: Las autoras

2.1.3 Software

En la Tabla 2.3, se presenta la lista de requerimiento de software:

Tabla 2.3 Requerimiento de Software

ANÁLISIS Y CONSTRUCCIÓN			
Nº	SOFTWARE	VERSION	LICENCIA
1	MICROSOFT WINDOWS	7.0	Microsoft
2	MS PROJECT 2007	12.0.4518.1014	Microsoft
3	MS VISIO 2007	12.0.4518.1014	Microsoft
IMPLEMENTACION			
1	SQL Server	R8	Software Libre
2	PENTAHO	5.2.0	Software Libre
3	INTERNET EXPLORER	7.0	Software Libre
4	TABLEAU DESKTOP	8.2	Tableau

Elaboración: Las autoras

2.1.4 Cronograma del Proyecto

El cronograma se representa el análisis de las tareas, la secuencia de las mismas, la duración y los recursos requeridos. Se considera, a su vez, que el total de Horas de Trabajo es de 437 horas, el detalle de las actividades y tareas por etapas se encuentran en el Anexo 01 Cronograma del Proyecto.

2.1.5 Presupuesto

En la Tabla 2.4, se presenta el presupuesto del personal por hora designado para el desarrollo del proyecto es el siguiente.

Tabla 2.4 Presupuesto del personal por hora

ROL	COSTO / HORA (S/.)	CANTIDAD (Hrs)	TOTAL (S/.)
Gestor de Proyecto	120	89.7h	S/. 10,764.42
Analista Funcional	100	65.78h	S/. 6,577.99
Diseñador del modelo	100	94.33h	S/. 9,433.33
Analista Técnico	100	12.83h	S/. 1,283.33
Analista Programador	100	193.54h	S/. 19,354.23
Certificador	100	36h	S/. 3,600.00
Total (S/.)			S/. 51,013.30

Elaboración: Las autoras

En la Tabla 2.5, se presentan los gastos de instalación de software:

Tabla 2.5 Gastos de instalación del software

Inversión	Soles (S/.)
Software	
SQL Server *	17,760.00
Pentaho	Software Libre
Tableau	6,079.96
TOTAL (S/.)	23,839.96
*Posee licencia	

Elaboración: Las autoras

En la Tabla 2.6, se presentan los gastos del Servicio de Implementación de dispositivos de seguridad:

Tabla 2.6 Gastos del Servicio de Implementación de dispositivos de seguridad

CARACTERISTICAS	COSTO (S/.)
IMPLEMENTACIÓN DE FIREWALL y RED DMZ	S/. 2,360.00
IMPLEMENTACIÓN CERTIFICADO	S/. 2,655.00
TOTAL (S/.)	S/. 5,015.00

Elaboración: las autoras

En la Tabla 2.7, se presentan los gastos de instalación de hardware:

Tabla 2.7 Gastos de instalación de hardware

EQUIPO	CARACTERÍSTICAS	COSTO (S/.)
CPU	INTEL CORE DUO- HDD 250MB	S/. 1,600.00
MONITOR	LCD-COLOR 32 BITS	S/. 860.00
SERVIDOR DE BD	INTEL CORE DUO- 3.0GHZ- RAM 2GB- HDD 1TB	S/. 2,000.00
TOTAL (S/.)		S/. 4,460.00
*En el caso de Gasto de Hardware son con los que cuentan el área actualmente.		

Elaboración: Las autoras

En la Tabla 2.8, se presentan los gastos de dispositivos de seguridad:

Tabla 2.8 Gastos de dispositivos de seguridad

EQUIPO	CARACTERÍSTICAS	COSTO (S/.)
CISCO ASA 5515- X Firewall	CISCO ASA 5515-X WITH SW, 6GE DATA	S/. 9,696.00
SMARTNET	SMARTNET 8X5XNBD ASA 5515-X	S/. 2253.00
TOTAL (S/.)		S/. 11,949.00

Elaboración: Las autoras

En la Tabla 2.9, se presenta el presupuesto total que abarca Recursos Humanos, Software y Hardware entre otros:

Tabla 2.9 Presupuesto del Proyecto

DESCRIPCIÓN	COSTO (S/.)
Recursos Humanos	S/. 51,013.30
Software	S/. 23,839.96
Hardware	S/. 4,460.00
Dispositivos de seguridad	S/. 11,949.00
Servicio de Implementación de dispositivos de seguridad	S/. 5,015.00
TOTAL (S/.)	S/. 96,277.26

Elaboración: Las autoras

2.2 Métodos

Como hemos mencionado anteriormente, se ha elegido PMBOK como Metodología de gestión.

2.2.1 Metodología de Gestión de Proyectos: PMBOK

Las diez (10) áreas del conocimiento del PMBOK relacionadas al siguiente proyecto se encuentran desarrolladas en el capítulo III Desarrollo del Proyecto. A continuación, se menciona cada área:

- **Gestión de integración:** El presente proyecto se desarrolla el acta de constitución del proyecto, control de cambios, se dirige y gestiona el trabajo del proyecto. Se encuentra en la sección 3.1.1.
- **Gestión de alcance:** El presente proyecto se define la estructura de Desglose de Trabajo - EDT, límites, supuestos que se contemplaran en la tesis. Se encuentra en la sección 3.1.2.
- **Gestión de tiempo:** El presente proyecto está conformado por etapas de desarrollo las que comprenden espacios de tiempo según la planificación en el cronograma. Se encuentra en la sección 3.1.3.
- **Gestión de costos:** El presente proyecto está conformado por la estimación de presupuesto, este tema se refleja en el punto 2.1.5 de la tesis. Se encuentra en la sección 3.1.4.
- **Gestión de la calidad del Proyecto:** En el presente proyecto se realiza el aseguramiento de calidad. Se encuentra en la sección 3.1.5.

- Gestión de Recursos humanos: El presente proyecto está conformado por la planificación y asignación de los recursos humanos. Se encuentra en la sección 3.1.6.
- Gestión de comunicaciones: El presente proyecto se planifica las comunicaciones, se distribuye la información. Se encuentra en la sección 3.1.7.
- Gestión de riesgos: El presente proyecto está conformado por la Identificación de Riesgos, análisis cuantitativo y cualitativo de riesgos. Se encuentra en la sección 3.1.8.
- Gestión de adquisiciones: El presente proyecto se planifica la gestión de las adquisiciones del proyecto. Se encuentra en la sección 3.1.9.
- Gestión de los interesados: El presente proyecto incluye la planificación de la gestión e los interesados, la gestión y control. Se encuentra en la sección 3.1.10.

Para poder realizar la elección de la Metodología para el desarrollo de la solución BI se realizará una selección de características, se establecerán los criterios de selección y luego se listarán las metodologías existentes afines al tema tratado para poder elegir la más adecuada para el desarrollo de la propuesta.

2.2.2 Selección de Metodología de Desarrollo

Para poder elegir una metodología para implementar un Datamart se ha adaptado un método utilizado para realizar una selección adecuada.

2.2.2.1 Definición de las características de la Metodología

En este punto, se definen las características que se establecen para los criterios de selección, finalmente elegiremos la metodología a implementar. A continuación, se muestra en la Tabla 2.10 las características que deseamos obtener de la metodología:

Tabla 2.10 Definición de características de las metodologías

ID	ASPECTO	DEFINICIÓN	Basado en:
1	Portabilidad y Adaptabilidad	Aplicable a múltiples plataformas, facilidad de ser transferida de un entorno a otro.	ADAPTADO ISO 9126
2	Portabilidad y Coexistencia	La metodología puede coexistir con otras metodologías del entorno.	ADAPTADO ISO 9126
3	Ciclo de vida	Abarca desde el planeamiento hasta el despliegue.	NO APLICA
4	Planeación estratégica	Cubre la planeación estratégica de la organización.	NO APLICA
5	Etapas detalladas	Sumándolas dan como resultado la completitud del proyecto de BI.	NO APLICA
6	Actividades por Etapa	Las actividades contenidas ayudan a lograr su completitud.	NO APLICA
7	Aprendizaje	Fácil comprensión del Proyecto.	ADAPTADO ISO 9126
8	Popularidad	Aceptación de la metodología.	NO APLICA
9	Fácil de investigar	La información de la metodología es fácil de ubicar.	NO APLICA

Fuente: Adaptado de Method for selecting a reference model for software process deployment (Bayona, S. et al., 2012)

2.2.2.2 Criterios de selección para la comparación de Metodologías.

Para evaluar cada criterio de selección definiremos los valores ideales y valores posibles. El valor ideal son las características que debe tener la metodología a usar acorde con las necesidades del usuario y el valor posible son las alternativas a lo esperado. A continuación, se muestra en la Tabla 2.11 los criterios que tomaremos en cuenta para la selección de la metodología:

Tabla 2.11 Cuadro de criterios para elegir la metodología

ID	CRITERIO	DESCRIPCIÓN	VALOR IDEAL	VALORES POSIBLES
C1	Portabilidad y Adaptabilidad	Aplicable a múltiples plataformas, facilidad de ser transferida de un entorno a otro.	SÍ	NO
C2	Portabilidad y Coexistencia	La metodología puede coexistir con otras metodologías del entorno.	SÍ	NO
C3	Ciclo de vida	Abarca desde el planeamiento hasta el despliegue.	SÍ	NO
C4	Planeación estratégica	Cubre la planeación estratégica de la organización.	SÍ	NO
C5	Etapas detalladas	Sumándolas dan como resultado la completitud del proyecto de BI.	SÍ	NO
C6	Actividades por Etapa	Las actividades contenidas ayudan a lograr su completitud.	SÍ	NO
C7	Aprendizaje	Fácil comprensión del proyecto.	SÍ	NO
C8	Popularidad	Aceptación de la metodología.	SÍ	NO
C9	Fácil de investigar	La información de la metodología es fácil de ubicar.	SÍ	NO

Fuente: Adaptado de Method for selecting a reference model for software process deployment (Bayona, S. et al., 2012)

2.2.2.3 Metodologías más usadas.

Se describen las metodologías principales para el desarrollo de una solución de BI:

Business Intelligence Roadmap: En términos generales de BI, especifica el camino y la dirección que deben seguir las aplicaciones, estructuras, herramientas y personas que intervienen en un proyecto de este tipo. Un mapa de ruta (roadmap) para Inteligencia de Negocios es una guía que provee un esqueleto para proyectos de BI con entradas flexibles; esto significa que una organización puede iniciar el desarrollo en cualquier paso del ciclo siempre y cuando se cumplan los prerrequisitos de ese paso. (Moss, L. y Atre, S., 2003).

Metodología Kimball: Es muy amplia la manera de abordar los elementos para las etapas de desarrollo, y deja claro qué se debe hacer, pero no cómo lograrlo, lo que provoca demoras en los resultados. Esta metodología no

detalla la manera en que se deben diseñar los modelos de datos ni la forma de obtener las variables para lograr la correspondencia con los datos fuentes. Además, no es contemplada como un ciclo completo de desarrollo sino que abarca solo la etapa de diseño del almacén de datos. (Leonard, E. y Castro, Y., 2013).

Metodología Data Warehouse Engineering Process (DWEP): Es una metodología basada en RUP y en la herramienta UML para desarrollar un Data Warehouse o DataMart llamada Data Warehouse Engineering Process (DWEP). Esta metodología fue propuesta por Sergio Luján-Mora y Juan Trujillo en el año 2006. Es una metodología que contempla el ciclo completo de desarrollo de un almacén de datos, en este punto, se considera que se manejan muchos artefactos. (Leonard B, Eric y Castro B. Yudi, 2013).

2.2.2.4 Matriz de Criterios y Metodologías

En la Tabla 2.12, se presenta la calificación con los resultados de la investigación realizada previamente:

Tabla 2.12 Cuadro comparativo de metodologías BI

Criterios Metodología	ADAPTABILIDAD	COEXISTENCIA	CICLO DE VIDA	PLANEACIÓN ESTRATÉGICA	ETAPAS DETALLADAS	ACTIVIDADES POR ETAPA	ENTENDIBILIDAD	POPULARIDAD	FÁCIL DE INVESTIGAR	Resultado
ROADMAP	SI	SI	SI	SI	SI	SI	SI	SI	SI	OK
KIMBALL	SI	SI	SI	SI	NO	SI	SI	SI	SI	NO
DWEP	SI	SI	NO	NO	SI	SI	SI	NO	NO	NO
METODOLOGÍA ESPERADA	SI	SI	SI	SI	SI	SI	SI	SI	SI	

Elaboración: Las autoras

Basándonos en los resultados de la comparación realizada, las metodologías a través de criterios selectivos, concluimos que Roadmap

posee todas las características que se desea, por ello esta metodología va acorde a nuestro proyecto.

El Business Intelligence Roadmap es una guía de buenas prácticas que ha sido descrita en el marco teórico. En la Figura 2.1, se representa el ciclo de vida para el desarrollo de un sistema de soporte de decisiones según Roadmap.

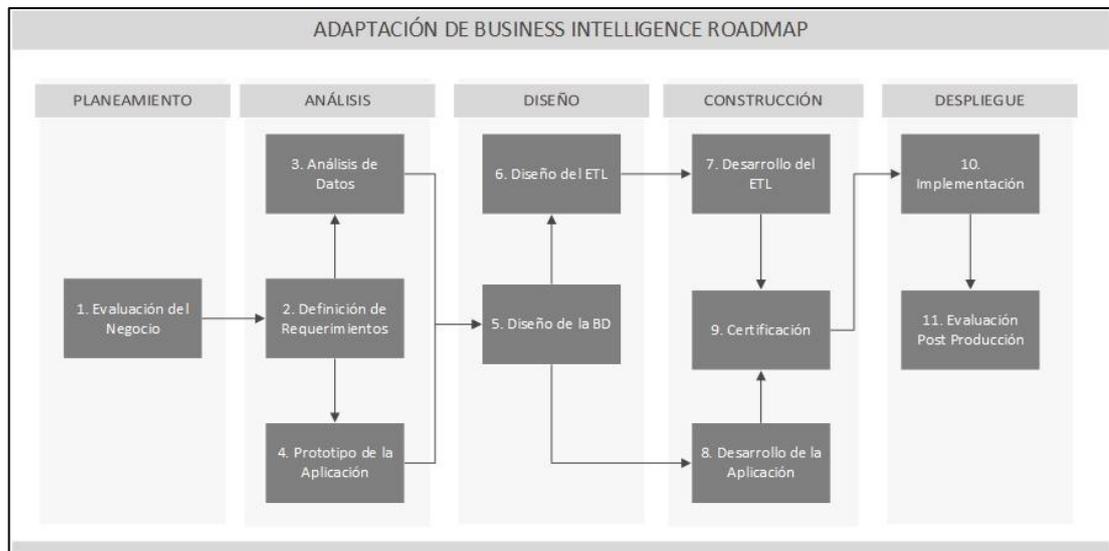


Figura 2.1 Metodología del proyecto adaptada del Business Intelligence Roadmap
Fuente: the complete project lifecycle for decision (De Moss, L. y Atre, S.)

A continuación, pasaremos a explicar cada fase del ciclo de la Metodología RoadMap:

Etapa de planeamiento

Fase Evaluación del negocio:

La evaluación del negocio ayuda plantear una solución que resuelva el problema de la organización, especificando las desventajas y beneficios que esta tendría. En esta etapa se realizó la revisión de los Objetivos de MINEDU.

Etapa de análisis

Fase Definición de requerimientos

En esta etapa se debe elegir los requerimientos de acuerdo a las posibilidades de cada entregable, sin crear falsas expectativas. Esta etapa consta de las siguientes actividades:

- Revisión de definición y cálculo de indicadores.
- Definición de reportes básicos por indicadores.

Fase Análisis de datos:

En esta etapa se evalúan los datos fuentes con la finalidad de obtener una mejor calidad de estos. Esta etapa consta de las siguientes actividades:

- Análisis dimensional.
- Elaboración de matriz de dimensiones de análisis.

Fase Prototipo de aplicación

Esta etapa permite a los desarrolladores y a los involucrados ver el potencial y las limitaciones de la tecnología, y también brinda la oportunidad de ajustar los requerimientos del proyecto, y las expectativas del mismo. Esta etapa consta de las siguientes actividades:

- Definición de Reportes.
- Prototipo de Reportes.

Etapa de diseño

Fase Diseño de la Base de Datos

El diseño de la base de datos debe estar acorde con los requerimientos planteados para acceder a la información con la que cuenta la organización. Esta etapa consta de las siguientes actividades:

- Elaboración de Modelo de datos.
- Validación del Modelo de datos.

Fase Diseño del ETL

En esta etapa, se realizaron las especificaciones del ETL y se fusionaron los datos que vienen de diversas plataformas en un formato para el Datawarehouse. Esta etapa consta de las siguientes actividades:

- Elaboración de mapeo técnico de datos.
- Validación del Mapeo técnico de datos.

Etapa de construcción

Fase Desarrollo del ETL

En esta etapa, se construye el producto en un margen de tiempo predeterminado. Esta etapa consta de las siguientes actividades:

- Construcción de los componentes en ETL.
- Construcción de scripts.
- Construcción de los Jobs.

Fase Desarrollo de la aplicación

No se contempla el desarrollo de la aplicación porque se hizo uso de la herramienta de explotación Tableau.

Fase Certificación

Se verifica que las funcionalidades implementadas y la disponibilidad de información correspondan con las especificaciones iniciales de los requerimientos del usuario. Esta etapa consta de las siguientes actividades:

- Pruebas funcionales
- Pruebas de integridad.
- Cuadre de carga.

Etapa de despliegue

Fase de Implementación

Se comienza a instalar el motor de la base de datos y la plataforma web. Se programa entrenamiento para los usuarios. Esta etapa consta de las siguientes actividades:

- Implementación de los Componentes (mantener la base de datos, programar y correr los Jobs ETL, monitorear el comportamiento del sistema y afinar la base de datos)

Fase de Evaluación posproducción

Esta etapa consta de las siguientes actividades:

- Seguimiento postproducción.
- Elaboración de Informe de cierre de Proyecto.

CAPÍTULO III DESARROLLO DEL PROYECTO

3.1 Gestión del Proyecto

3.1.1 Gestión de Integración

En este punto, incluye los procesos necesarios que se requieren para completar el proyecto con éxito. Se indican las actividades para la gestión de los interesados.

En el caso del presente proyecto, se consideraron la aplicación de tres procesos:

3.1.1.1 Desarrollar el Acta de Constitución del Proyecto

Se utilizó el acta de Constitución de Proyecto, usando la técnica:

Juicio de Expertos: Se solicitó a los interesados que aporten sus experiencias y formen parte de la gestión de control de cambios, que se aplicó en cualquier detalle técnico y de gestión.

A través del Acta de Constitución del Proyecto se autoriza la formalidad de la ejecución del proyecto. Ver Anexo 09 Acta de Constitución del Proyecto.

3.1.1.2 Control de Cambios del Proyecto

En el control de cambios del proyecto para la mitigación del problema de cambios en el alcance, se implementarán las siguientes acciones:

- Definir y planificar el alcance del proyecto, un alcance adecuado permite una correcta estimación de tiempos y costos para el proyecto.
- Crear y Definir EDT, a mayor nivel de detalle, mejor definido se encuentra el alcance.
- Generar solicitudes de cambio, estas solicitudes de cambio contarán con un formato para la gestión de la misma.

Para la ejecución del control de cambios, se utilizaron las siguientes técnicas y herramientas:

Juicio de Expertos: Se solicitó a los interesados que aporten sus experiencias y formen parte de la gestión de control de cambios, que se aplicará en cualquier detalle técnico y de gestión.

El comité de Control de cambios está conformado por los responsables para para la aprobación de los cambios en el alcance:

- El Representante de la Secretaría de Planificación Estratégica.
- El Jefe de OFIN
- El Gestor de Proyecto.
- El Equipo del Proyecto.

Reuniones: Estas reuniones se suelen nominar reuniones de control de cambios. Se realizan los acuerdos a través de actas de reuniones para la realización de las acciones en el avance del proyecto. Las características de

las reuniones deben de contemplar la siguiente agenda de gestión de cambios:

- Revisar e influir en los factores que eluden el control integrado de cambios, de tal forma que solo se implementen cambios aprobados.
- Revisar, analizar y aprobar las solicitudes de cambio de forma rápida, esto es esencial, ya que una decisión tardía podría influenciar negativamente en el tiempo, el costo o la viabilidad de un cambio.

Para la aprobación de los cambios será de la siguiente manera:

- Si los cambios son clasificados como de pequeño impacto, estos serán aprobados por el Jefe del Proyecto.
- Si los cambios son clasificados como de pequeño y grande impacto, se requerirá la aprobación del Comité de Control de Cambios.

Para ver el detalle de Flujo del Control de Cambios. Ver Figura 3.1 Flujo de Control de Cambios. El resultado del uso de las herramientas y técnicas a través del formato de la solicitud de cambios que serán almacenados en el banco de información del proyecto. Ver Anexo 10 Solicitud de Control de Cambios.

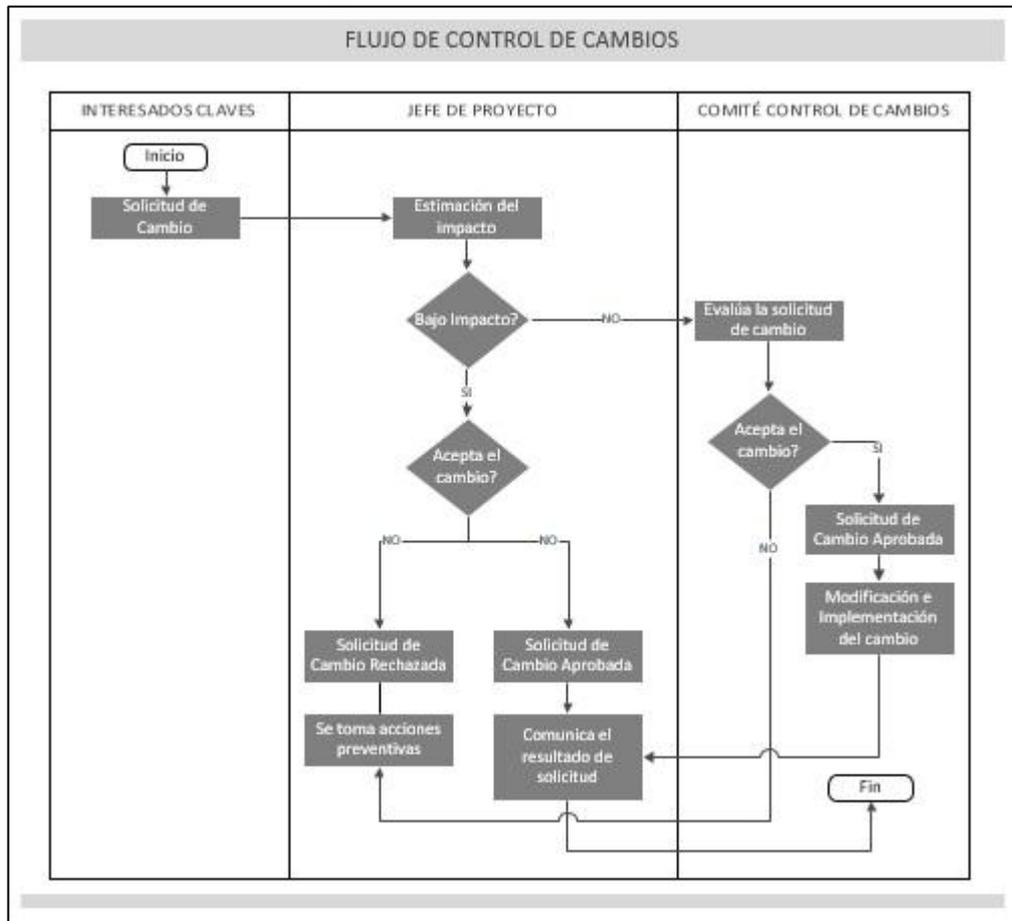


Figura 3.1 Flujo de control de cambios
Elaboración: las autoras

3.1.1.3 Cerrar el Proyecto

Se realiza la transferencia de archivos del Proyecto relacionados al alcance, costo, cronograma, riesgos del proyecto, análisis dimensional, las dimensiones y tablas hecho, entre otros.

3.1.2 Gestión de Alcance

En este rubro, se gestiona el alcance del proyecto enfocándose en definir y controlar lo que si incluye y no el proyecto. Se realizará los siguientes procesos:

3.1.2.1 Recopilar requisitos

En este punto se define y documenta las necesidades de los interesados a fin de cumplir con los objetivos del proyecto. Se ha tomado como entrada:

Acta de Constitución del Proyecto: El acta de constitución del proyecto se usa para proporcionar los requisitos de alto nivel del proyecto, así como una descripción de alto nivel del producto del proyecto, se establecen los requisitos detallados del producto.

Registro de Interesados: Se usa para identificar a los interesados que pueden proporcionar información acerca de los requisitos detallados del proyecto y del producto.

Para la recopilación de requisitos se utilizó las herramientas y técnicas:

Entrevistas: Es una manera formal o informal de obtener información acerca de los interesados, a través de un diálogo directo con ellos

En la tabla 3.1, se describe la documentación de requisitos:

Tabla 3.1 Requisitos del Proyecto

REQUISITOS	TIPO	PRIORIDAD	INTERESADOS
El Datamart debe ser diseñado para que permita ofrecer la escalabilidad y optimización de la información de los procesos de matrículas y nóminas a través de la extracción, transformación y carga de datos.	PRODUCTO	ALTA	USUARIO FINAL
El precio que estaría dispuestos a pagar la mayoría de interesados por este producto es de entre un 5% y un 10% menor del precio total que ofrece el mercado de un Datamart realizado por un proveedor especialista. Muy pocos estarían dispuestos a pagar más del 95%.	PRODUCTO	ALTA	USUARIO FINAL
Facilidad de uso del Datamart.	PRODUCTO	ALTA	USUARIO FINAL
El Datamart debe estar parametrizado y contar con un diccionario de datos para posteriores implementaciones o modificaciones.	PRODUCTO	ALTA	USUARIO FINAL
El Datamart permita la comunicación con el SIAGIE y la posibilidad de comunicación con otros Datamart.	PRODUCTO	ALTA	USUARIO FINAL
Diseño/estética, que sea voluminoso, que llame la atención, que sea clara y se pueda tomar una decisión rápidamente mientras se visualice la información.	PRODUCTO	ALTA	USUARIO FINAL
Para el diseño del Datamart se debe realizar prototipos.	PRODUCTO	MEDIA	USUARIO FINAL
Proyecto del Datamart debe seguir la guía de buenas prácticas PMBOK 5ta. Edición.	PRODUCTO	ALTA	USUARIO FINAL
El Datamart se ha desarrollado con éxito cumpliendo con el alcance planificado, tiempo y costos con los clientes.	PRODUCTO	ALTA	USUARIO FINAL
Fecha de finalización del Datamart antes del 30 de Octubre del 2014.	PRODUCTO	ALTA	USUARIO FINAL

Elaboración: Las autoras

3.1.2.2 Definir el alcance

Se describen los límites del proyecto, se seleccionaron los requisitos definidos del proyecto. Se ha tomado como entrada el Acta de Constitución del Proyecto que proporciona un alto nivel de descripción del proyecto también utilizaremos la técnica:

Juicio de Expertos: Se solicitó a los interesados que aporten sus experiencias y formen parte de la gestión de control de cambios, que se aplicará en cualquier detalle técnico y de gestión.

A continuación, se realiza la definición del alcance del proyecto:

Alcance del Proyecto

Análisis, Diseño e Implementación de un Datamart al área de Secretaría de Planificación Estratégica.

Descripción del Alcance del Proyecto

En la Tabla 3.2, se muestra el resumen de las principales fases que se realizará para el desarrollo del proyecto a ejecutar. Las horas de trabajo se distribuyen en las siguientes fases del proyecto:

Tabla 3.2 Fases Requeridas

Fases de Proyecto	Tiempo
Gestión del Proyecto	85 horas
Evaluación del Negocio	4 horas
Definición de Requerimientos	16 horas
Análisis de Datos	34 horas
Prototipo de la Aplicación	12 horas
Diseño de base de datos	48 horas
Diseño del ETL	34 horas
Desarrollo de ETL	118 horas
Desarrollo de la aplicación	52 horas
Certificación	19 horas
Implementación	12 horas
Evaluación Postproducción	3 horas
TOTAL	437 horas

Elaboración: Las autoras

Objetivo general:

Mejorar el apoyo a la toma de decisiones en los procesos de matrículas y nóminas en la Secretaría de Planificación Estratégica mediante la implementación del Datamart.

Limitaciones del Proyecto:

No se realizará cambios después de haberse realizado la fase de Certificación.

3.1.2.3 Crear EDT

Es el proceso se subdivide en entregables y el trabajo del proyecto en componentes más pequeños.

Como técnica en la Gestión del Alcance se creó el EDT (Estructura Descomposición del Trabajo).

Se solicitó a los interesados que aporten sus experiencias y formen parte de la gestión de control de cambios, que se aplicará en cualquier detalle técnico y de gestión.

En el Anexo 02 EDT del Proyecto, se muestran las diversas fases, en cada una de ellas se han definido las siguientes actividades.

3.1.2.4 Control de alcance

Las modificaciones del Alcance, deberán ser a través del proceso de Control de Cambios. Los interesados clave para la solicitud de cambios en el alcance son:

- El Representante de la Secretaría de Planificación Estratégica.
- El Jefe de OFIN
- El Gestor de Proyecto.
- El Equipo del Proyecto.

3.1.3 Gestión del tiempo

En ello, se incluyen los procesos requeridos para gestionar la terminación del plazo del proyecto.

3.1.3.1 Definir las actividades

Para la definición de las actividades se utilizó la técnica “Juicio de Expertos” a través de reuniones de coordinación.

3.1.3.2 Desarrollar el Cronograma

Se analizó el orden de las actividades, la secuencia de las actividades y las restricciones para la creación del cronograma utilizando el software Microsoft Project que genera un modelo de programación con fechas para completar las actividades del proyecto incorporando actividades, duración, recursos; se ha utilizado el método de la ruta crítica para estimar la duración del proyecto, calculando las fechas de inicio y finalización del proyecto. La siguiente tabla 3.3 muestra el cronograma resumen del proyecto.

Tabla 3.3 Cronograma Resumen de Proyecto

EDT	Nombre de tarea	Trabajo	Comienzo	Fin	Periodo de Tiempo											
					Agosto				Setiembre				Octubre			
					S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4
1	Proyecto BI DM SPE	437 horas	sáb 09/08/14	sáb 25/10/14												
1.1	Gestión del Proyecto	85 horas	sáb 09/08/14	vie 24/10/14												
1.2	Evaluación del Negocio	4 horas	mar 12/08/14	mié 13/08/14												
1.3	Definición de requerimientos	16 horas	mié 13/08/14	jue 14/08/14												
1.4	Análisis de Datos	34 horas	jue 14/08/14	vie 22/08/14												
1.5	Prototipo de aplicación	12 horas	vie 22/08/14	mar 26/08/14												
1.6	Diseño de base de datos	48 horas	mié 27/08/14	mar 09/09/14												
1.7	Diseño del ETL	34 horas	mié 10/09/14	vie 19/09/14												
1.8	Desarrollo del ETL	118 horas	vie 19/09/14	vie 17/10/14												
1.9	Desarrollo de la aplicación	52 horas	mié 15/10/14	mar 21/10/14												
1.1	Certificación	19 horas	mar 21/10/14	jue 23/10/14												
1.11	Implementación	12 horas	jue 23/10/14	vie 24/10/14												
1.12	Evaluación postproducción	3 horas	jue 23/10/14	sáb 25/10/14												

Elaboración: Las autoras

3.1.3.3 Controlar el Cronograma

Para el control del cronograma, se utilizó el método de ruta crítica para comparar el avance del estado de cronograma del proyecto. La variación tuvo un impacto en la finalización del proyecto.

3.1.4 Gestión de Costos

En este punto, se incluyen los procesos relacionados con la estimación de los costos y el control de los costos del proyecto.

3.1.4.1 Estimación de costos

Se desarrolla una estimación aproximada de los recursos monetarios necesarios para completar las actividades del proyecto expresadas en nuevos soles. Para la estimación de los costos hemos utilizado como entrada el Cronograma del Proyecto así como también el EDT que contiene las actividades de cada paquete de trabajo.

Se elaboró la línea base de desempeño de costos (presupuesto). La estimación de costos se realizó utilizando la técnica "Juicio de Expertos". En la siguiente tabla 3.4, se muestra la línea base de Costos.

Tabla 3.4 Línea base de costos

EDT	Nombre de tarea	Trabajo	Comienzo	Fin	Costo Soles	Detalle	Agosto			Setiembre			Octubre		
1	Proyecto BI DM SPE	437 horas	sáb 09/08/14	sáb 25/10/14	Sl. 51,013.30	Costo	Sl. 84.00	Sl. 8,570.08	Sl. 6,303.85	Sl. 3,435.38	Sl. 5,963.82	Sl. 3,764.07	Sl. 7,210.26	Sl. 10,257.69	Sl. 5,444.15
1.1	Gestión del Proyecto	85 horas	sáb 09/08/14	vie 24/10/14	Sl. 11,143.63	Costo	Sl. 84.00	Sl. 2,520.08	Sl. 1,453.85	Sl. 1,315.38	Sl. 1,453.82	Sl. 1,130.73	Sl. 1,476.93	Sl. 1,107.69	Sl. 601.15
1.2	Evaluación del Negocio	4 horas	mar 12/08/14	mié 13/08/14	Sl. 500.00	Costo		Sl. 500.00							
1.3	Definición de requerimientos	16 horas	mié 13/08/14	jue 14/08/14	Sl. 2,100.00	Costo		Sl. 2,100.00							
1.4	Análisis de Datos	34 horas	jue 14/08/14	vie 22/08/14	Sl. 3,700.00	Costo		Sl. 3,450.00	Sl. 250.00						
1.5	Prototipo de aplicación	12 horas	vie 22/08/14	mar 26/08/14	Sl. 1,300.00	Costo			Sl. 1,300.00						
1.6	Diseño de base de datos	48 horas	mié 27/08/14	mar 09/09/14	Sl. 5,200.00	Costo			Sl. 3,300.00	Sl. 1,900.00					
1.7	Diseño del ETL	34 horas	mié 10/09/14	vie 19/09/14	Sl. 3,810.00	Costo				Sl. 200.00	Sl. 3,610.00				
1.8	Desarrollo del ETL	118 horas	vie 19/09/14	vie 17/10/14	Sl. 12,816.67	Costo					Sl. 900.00	Sl. 2,633.34	Sl. 5,733.33	Sl. 3,550.00	
1.9	Desarrollo de la aplicación	52 horas	mié 15/10/14	mar 21/10/14	Sl. 6,400.00	Costo								Sl. 5,600.00	Sl. 800.00
1.10	Certificación	19 horas	mar 21/10/14	jue 23/10/14	Sl. 2,200.00	Costo									Sl. 2,200.00
1.11	Implementación	12 horas	jue 23/10/14	vie 24/10/14	Sl. 1,300.00	Costo									Sl. 1,300.00
1.12	Evaluación postproducción	3 horas	jue 23/10/14	sáb 25/10/14	Sl. 543.00	Costo									Sl. 543.00
	Adquisición de Hardware				Sl. 4,460.00	Costo		Sl. 4,460.00							
	Adquisición de Software				Sl. 23,839.96	Costo									Sl. 23,839.96
	Adquisición de dispositivos Seguridad				Sl. 11,949.00	Costo									Sl. 11,949.00
	Implementación de dispositivos de Seguridad				Sl. 5,015.00	Costo									Sl. 5,015.00
	SubTotal						Sl. 84.00	Sl. 13,030.08	Sl. 6,303.85	Sl. 3,415.38	Sl. 5,963.82	Sl. 3,764.07	Sl. 7,210.26	Sl. 10,257.69	Sl. 46,248.11
	Total Mes						Sl.		19,417.93	Sl.		13,143.27	Sl.		63,716.06
	Costo Acumulado						Sl.		19,417.93	Sl.		32,561.20	Sl.		96,277.26
	% Acumulado								20%			34%			100%

Elaboración: Las autoras

3.1.4.2 Controlar los costos

Se monitorea el estado del proyecto para actualizar los cambios en la línea base del proyecto, estos cambios se realizarán en el software Ms Project, estas se ejecutan a través del control integrado de cambios.

3.1.5 Gestión de calidad

El proceso de aseguramiento de control de calidad se encuentra detallado en el capítulo IV: Pruebas y Resultados.

3.1.6 Gestión de Recursos humanos

Se menciona al equipo que está compuesto por las personas a las que se ha asignado roles y responsabilidades para completar el proyecto, incluye los siguientes procesos:

3.1.6.1 Planificar la Gestión de Recursos humanos

Es el proceso de identificar y documentar los roles dentro del proyecto y responsabilidades, nos hemos basado usando las técnicas de "Organigrama y Descripciones de Cargos".

El equipo de proyecto está conformado de la siguiente manera, ver Figura 3.2:

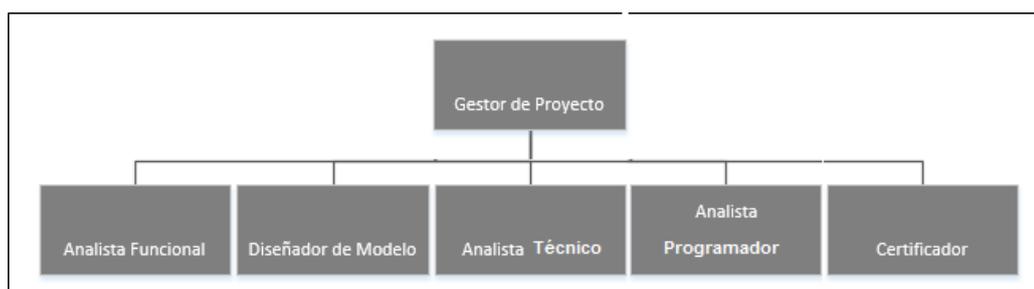


Figura 3.2 Organigrama del proyecto
Elaboración: las autoras

A continuación, se menciona la descripción de los cargos del proyecto ver Tabla 3.5:

Tabla 3.5 Roles del proyecto

Roles	Funciones
Gestor de Proyecto	Encargado que realizará el seguimiento del proyecto, los documentos que se requieren.
Analista Funcional	Encargado que realizará el levantamiento de información y requisitos.
Diseñador del Modelo	Encargado que realizará el modelo relacional, las reglas de negocio mediante las cuales se rige la Información.
Analista Técnico	Encargado que realizará la evaluación y soporte.
Analista Programador	Encargado que realizará la codificación de los requerimientos solicitados de acuerdo a las definiciones funcionales y técnicas alcanzadas.
Certificador	Encargado que realizará los casos de prueba para validad que lo construido este de acuerdo a las definiciones funcionales.

Fuentes: Las autoras

3.1.7 Gestión de comunicaciones

La comunicación es un medio que ayuda a las personas a la solución de problemas. Para iniciar la gestión de comunicaciones se identificó las necesidades de los involucrados (stakeholders), seguido a ello, se determinó los medios necesarios para satisfacer dichas necesidades es un factor importante para garantizar el éxito del proyecto.

Luego se aprecian los requisitos de comunicaciones de los interesados, ver Tabla 3.6 y la matriz de comunicaciones que se llevó a cabo con los interesados, ver Tabla 3.7.

Tabla 3.6 Requisitos de comunicación de los interesados

Interesados	Requisitos de Comunicación	Interés e Influencia en el proyecto
Ing Manuel Cok Aparcana Jefe de Informática MINEDU	Coordinar las reuniones con una semana de anticipación	Alta
Manuel Vargas Alegría Coordinador SIAGIE-OFIN MINEDU	Coordinar las reuniones con una semana de anticipación Realizar un Acta para cada reunión.	Alta
Maryury García Gestor de Proyecto	Registrar en un Acta los acuerdos al finalizar las reuniones.	Alta
Karla Jiménez Analista	Registrar en un Acta los acuerdos al finalizar las reuniones.	Alta

Fuentes: Las autoras

Tabla 3.7 Matriz de comunicaciones

Comunicación	Objetivo	Contenido	Formato	Medio	Plazo para confirmar recepción	Responsable	Aprobador	Audiencia / Receptor
Inicio Proyecto	Solicitar la aprobación de la propuesta del Proyecto.	Propuesta de solución para que sea aprobada por la Oficina de Informática del MINEDU	Impreso	Oficio	15 días	Maryury García Karla Jiménez	Ing. Manuel Cok Aparcana Manuel Vargas.	Ing. Manuel Cok Aparcana Manuel Vargas.
Reunión N°1	Reunión Kick off	Se realizó la presentación del Proyecto con las personas interesadas.	PPT	Presencial	1 día	Maryury García Karla Jiménez	Manuel Vargas- Coordinador Del SIAGIE OFIN	Manuel Vargas- Coordinador Del SIAGIE OFIN
Solicitud Reunión N°2	Solicitar Información de la Organización	Se solicitó una reunión para que nos proporcionen la información del Proceso de Matrícula- Proceso de Registro Nóminas	Escrito	Mail	1 día	Maryury García Karla Jiménez	Manuel Vargas- Coordinador Del SIAGIE OFIN	Maryury García Karla Jiménez
Reunión N°2	Brindar Información solicitada de Proceso de Matrícula y Proceso de Registro de Actas	Se solicitó la información de los procesos que se llevan a cabo con el SIAGIE: Proceso de Matrícula- Proceso de Registro Nóminas	Impreso	Presencial	5 días	Manuel Vargas	Maryury García Karla Jiménez	Maryury García Karla Jiménez

Comunicación	Objetivo	Contenido	Formato	Medio	Plazo para confirmar recepción	Responsable	Aprobador	Audiencia / Receptor
Solicitud Reunión N°3	Solicitar Data de Muestra	Solicitud para que nos brinden una muestra de su BD para poder realizar la carga del Datamart.	Escrito	Mail	1 día	Maryury García Karla Jiménez	Manuel Vargas Persona Encargada de la BD de Siagie	
Reunión N°3	Brindar la Data de Muestra para poder realizar el piloto del producto.	Se realizó la recepción de la data de muestra.	Txt	Presencial	5 días	Manuel Vargas	Maryury García Karla Jiménez	Maryury García Karla Jiménez
Reunión N°4	Presentación y conformidad del proyecto realizado.	S presentó el producto finalizado.	Ppt. Tableau	Presencial	3 días	Maryury García Karla Jiménez	Ing. Manuel Cok Aparcana	Ing. Manuel Cok Aparcana

Fuentes: Las autoras

3.1.8 Gestión de Riesgos

En la Tabla 3.8 se presenta los Riesgos del Proyecto.

Tabla 3.8 Riesgos del Proyecto

Cód	Riesgo	Disparador	Prob	Impacto	Puntaje	Tipo de Riesgo	Respuestas Planificadas	Tipo Respuesta
1	No se realice un adecuado levantamiento de información.	Falta de interés en el levantamiento de información.	0.3	0.8	0.81	Alto	Evaluar que la información recogida pueda satisfacer los requerimientos	MI
2	Pérdida de entregables del proyecto.	Falta de réplicas del proyecto y manejo de versiones.	0.1	0.8	0.83	Alto	Implementar mecanismos de versiones y réplicas de proyecto.	MI
3	Datamart no cumpla con los requerimientos establecidos.	Inadecuado análisis de requerimientos.	0.3	0.8	0.85	Alto	Realizar un adecuado análisis de los requerimientos, tener un feedback constante con los usuarios finales.	EV
4	El proyecto dura más tiempo de lo planificado.	Retraso en ejecución de las actividades del proyecto.	0.5	0.4	0.75	Alto	Realizar un control en la ejecución de cada actividad.	MI
5	Elección de una mala herramienta de explotación.	La herramienta no se adapta al Datamart fácilmente.	0.3	0.2	0.45	Moderado	Evaluar las herramientas de explotación a través de diversos criterios.	MI
6	No se cuente con un equipo de trabajo eficiente	Aumento de esfuerzo en las actividades del proyecto.	0.5	0.4	0.75	Alto	Realizar una entrevista de conocimientos por cada rol.	MI

Probabilidad	Valor	Impacto	Valor	Puntaje	Tipo de riesgo	Tipo de rpta	Abreviatura
Muy Improbable	0.1	Muy Bajo	0.05	Mayor a 0.5	Alto	Evitar	EV
Relativamente	0.3	Bajo	0.1	Menores a 0.3	Moderado	Mitigar	MI
Probable	0.5	Moderado	0.2	Menor a 0.1	Bajo	Aceptar	AC
Muy Probable	0.7	Alto	0.4				
Casi Certeza	0.9	Muy Alto	0.8				

3.1.9 Gestión de Adquisición

En este punto se detalla el proceso de efectuar las adquisiciones.

3.1.9.1 Planificar las adquisiciones

Se realizó la planificación de la adquisición del proyecto. Se prepararon los siguientes documentos de adquisición según lo siguiente:

- RFP (Requerimiento de Propuesta): Es un documento que el proyecto empresa emite para solicitar propuestas de posibles proveedores de productos o servicios.
- Formato de Contrato para la adquisición del producto.

Se utilizó la técnica “Análisis de Hacer o Comprar” que consiste en determinar si un trabajo en particular puede ser realizado satisfactoriamente por el equipo o debe ser adquirido por fuentes externas, para ello se ha realizado las comparaciones del costo para llevar a cabo las actividades, ver Tabla 3.9.

A base de este análisis, se ha determinado que el Análisis, diseño e implementación del Datamart realizado por el equipo de proyecto tendría un costo de S/. 79,313.26 mientras adquirir el producto del Datamart a terceros tendría un costo de S/. 80,105.91 en vista a este resultado e implicando que contamos con personal capacitado para el proyecto se tomará la decisión de realizar el producto. En el punto de Servicios de Implementación de Seguridad realizado por el equipo tendría un costo S/.18,660.00 sin embargo el equipo no se encuentra capacitado para este tipo de línea la cuál incurriría a una capacitación que extendería el plazo de tiempo planificado del proyecto, mientras que adquirir el servicio con terceros tendría un costo de S/. 16,964.00 tomando la decisión de comprar el servicio a través de la técnica Juicio de Expertos.

Tabla3.9 Análisis de hacer o comprar

ANÁLISIS DE HACER O COMPRAR				
NOMBRE DEL PROYECTO		ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE DATAMART PARA LA SECRETARÍA PLANIFICACIÓN ESTRATÉGICA MINEDU		
PREPARADO POR		Maryury García Anticona.		
FECHA		25/10/2014		
		COSTO HACER	COSTO COMPRAR	
ITEM#	DESCRIPCIÓN DEL COSTO	COSTO TOTAL	COSTO TOTAL	NOTAS ó RECOMENDACIONES
1	Análisis, Diseño e Implementación del Datamart.	S/. 79,313.26	S/. 80,105.91	Se cuenta con personal capacitado para el Datamart.
2	Servicio de Implementación de dispositivos de seguridad	S/. 18,660.00	S/. 16,964.00	Se notifica que el HACER afectaría el tiempo por la capacitación del personal.
Análisis, Diseño e Implementación del Datamart.				
Detalle de Hacer				
N°	DESCRIPCIÓN DEL COSTO	COSTO		
1	Recursos Humanos	S/. 51,013.30		
2	Software	S/. 23,839.96		
3	Hardware	S/. 4,460.00		
SubTotal		S/. 79,313.26		
Recursos Humanos –Hacer				
ROL	COSTO / HORA (S/.)	CANTIDAD (Hrs)	TOTAL (S/.)	
Gestor de Proyecto	120	89.7h	S/. 10,764.42	
Analista Funcional	100	65.78h	S/. 6,577.99	
Diseñador del modelo	100	94.33h	S/. 9,433.33	
Analista Técnico	110	12.83h	S/. 1,283.33	
Analista Programador	100	193.54h	S/. 19,354.00	
Certificador	110	36h	S/. 3,600.00	
SubTotal			S/. 51,013.30	
Detalle de Comprar				
N°	DESCRIPCIÓN DEL COSTO	COSTO		
1	Recursos Humanos	S/. 52,134.95		
2	Software	S/. 23,170.96		
3	Hardware	S/. 4,800.00		
SubTotal		S/. 80,105.91		

Recursos Humanos – Comprar				
N°	ROL	COSTO / HORA (S/.)	CANTIDAD (Hrs)	TOTAL (S/.)
1	Gestor de Proyecto	125	89.7h	S/. 11,212.50
2	Analista Funcional	110	65.78h	S/. 7,235.80
3	Diseñador del modelo	95	94.33h	S/. 8,961.35
4	Analista Técnico	110	12.83h	S/. 1,411.30
5	Analista Programador	100	193.54h	S/. 19,354.00
6	Certificador	110	36h	S/. 3,960.00
SubTotal				S/. 52,134.95
Servicio de Implementación de dispositivos de seguridad				
Detalle de Hacer				
N°	DESCRIPCIÓN DEL COSTO	COSTO		
1	1 Capacitación Firewall y Red DMZ	S/. 3,500.00		
2	1 Capacitación Implementación Certificado	S/. 2,500.00		
3	CISCO ASA 5515-X Firewall	S/. 10,100.00		
4	SMARTNET	S/. 2,560.00		
SubTotal		S/. 18,660.00		
Detalle de Comprar				
N°	DESCRIPCIÓN DEL COSTO	COSTO		
1	Configuración del Firewall y Red DMZ	S/. 2,360.00		
2	Implementación Certificado	S/. 2,655.00		
3	CISCO ASA 5515-X Firewall	S/. 9,696.00		
4	SMARTNET	S/. 2,253.00		
SubTotal		S/. 16,964.00		

Elaboración: Las autoras

3.1.9.2 Efectuar las adquisiciones

Para la ejecución de la adquisición, se utilizó la técnica de “Juicio de Expertos” y “Criterios de Evaluación de Propuestas”. En la tabla 3.10, se muestra el formato que se toma en cuenta para criterios de evaluación y ponderación.

Tabla 3.10 Formato de criterios de evaluación y ponderación.

Criterios de Evaluación	Puntaje Máximo	Proveedor 1	Proveedor 2	Proveedor 3
Costo				
Plazo				
Calidad del Producto				
Satisfacción				
Plazo de Entrega				

Elaboración: Las autoras

3.1.9.3 Control y Cierre de Adquisiciones

Para el control de las adquisiciones, se tomó las solicitudes de cambio al plan de dirección del proyecto. Para el cierre de adquisiciones se aborda el contrato aplicable al proyecto asegurando los acuerdos contractuales sean completados o terminados.

3.1.10 Gestión de los Interesados del Proyecto

Al inicio del proyecto, se identificaron a los interesados, es decir, a las personas que su interés y participación impactan en el éxito del proyecto. Estas personas están a lo largo del ciclo de vida del proyecto, su participación consta en brindar sus necesidades y expectativas que se desean satisfacer con la elaboración del producto final. En la tabla 3.11, se muestra el registro de los interesados del proyecto.

Factores críticos de éxito

El éxito del proyecto consiste en tener el producto a tiempo, en costo y expectativas de ambas partes, con el cliente satisfecho por el alcance y funcionalidad. Para esto definimos los cuatro (4) factores implicados en la implementación de Datamart, que son los siguientes:

- a) Negociación del proyecto: Desde aquí se pudo identificar rápidamente si el proyecto tiene los argumentos para ser exitoso o está en riesgo el alcance y las expectativas del usuario final. Con respecto al presupuesto, este puede ser aprobado de la manera tradicional, presentando la propuesta económica de otro proveedor
- b) Tecnología: La infraestructura de la organización define también la tecnología a usarse. La selección de la tecnología seleccionada es vital para el costo del proyecto.
- c) Metodología: Seleccionar una metodología orientada hacia el resultado.
- d) Recursos: El último factor del cual depende el éxito de nuestro proyecto son los recursos que estarán involucrados, es decir, las personas y sus respectivos perfiles de conocimientos y experiencias en el tipo de proyecto, metodología de trabajo y la tecnología.

3.3 Etapa de análisis

3.3.1 Definición de requerimiento

En esta definición, se cubren los objetivos del área, la lógica de negocio y los indicadores necesarios para realizar los reportes requeridos para la Secretaría de Planificación Estratégica. A continuación, se expone el resumen de lo obtenido para cada uno de ellos:

Objetivos del Área

Los objetivos del área de la Secretaría de Planificación Estratégica se pudieron recopilar a través de reuniones que se han tenido acerca del proyecto en las cuales se entrevistaron al representante del Área de Secretaría de Planificación Estratégica y el Jefe de la Oficina de Informática.

Los objetivos estratégicos que se han identificado a base de las reuniones sostenidas fueron:

- Garantizar la prestación de servicio de enseñanza de las niñas, niños y adolescentes en las Instituciones educativas de Educación Básica Regular.
- Garantizar los registros de las nóminas realizadas por las Instituciones educativas de Educación Básica Regular.
- Garantizar la participación de estudiantes en el sistema educativo.

Reglas de Negocio SIAGIE

En la presente sección, se indica la regla de negocio que se incluirá dentro de este proyecto para realizar nuestros reportes. Se detalla el siguiente modelo:

- Modelo de Estados de Matrícula
- Modelo de Estados de Nómina

La matrícula es el documento técnico pedagógico que contiene la relación de un estudiante con datos del estudiante, datos del nivel, datos del grado y datos de la sección en la IE. La matrícula es elaborada por la IE para ser validada por el SIAGIE los datos del estudiante y aprobada por Resolución del Director de la misma cuya copia se remite a la Unidad de Gestión Educativa Local o Dirección Regional de Educación para conocimiento y fines estadísticos. En la Tabla 3.12, se presentan los estados de matrícula:

Tabla 3.12 Estados de matrícula

Estado Matrícula	Estado Consolidador	Descripción	Desencadenador al siguiente Estado
Prematrícula	Prematrícula	IE realiza la inscripción de las materias que cursará el estudiante antiguo.	Realizada la inscripción de la matrícula se actualiza el estado en "Estado Proceso"
En Proceso	Matrícula en Proceso	IE carga la matrícula del estudiante	Realizada la carga de la matrícula se actualiza el estado en "Estado Proceso"
En Definitivo	Matrícula en Estado Definitivo	Analista de UGEL a través del SIAGIE aprueba la matrícula.	Realizada la aprobación de la matrícula se actualiza el estado en "En Definitivo"
Anular Matrícula	Matrícula en Estado Anulado	IE procede a guardar grado y sección del estudiante	Realizada el cambio de matrícula del estudiante se actualiza el estado en "Anulado"

Elaboración: Las autoras

En la Figura 3.3 se muestra el flujo del estado de las matrículas:

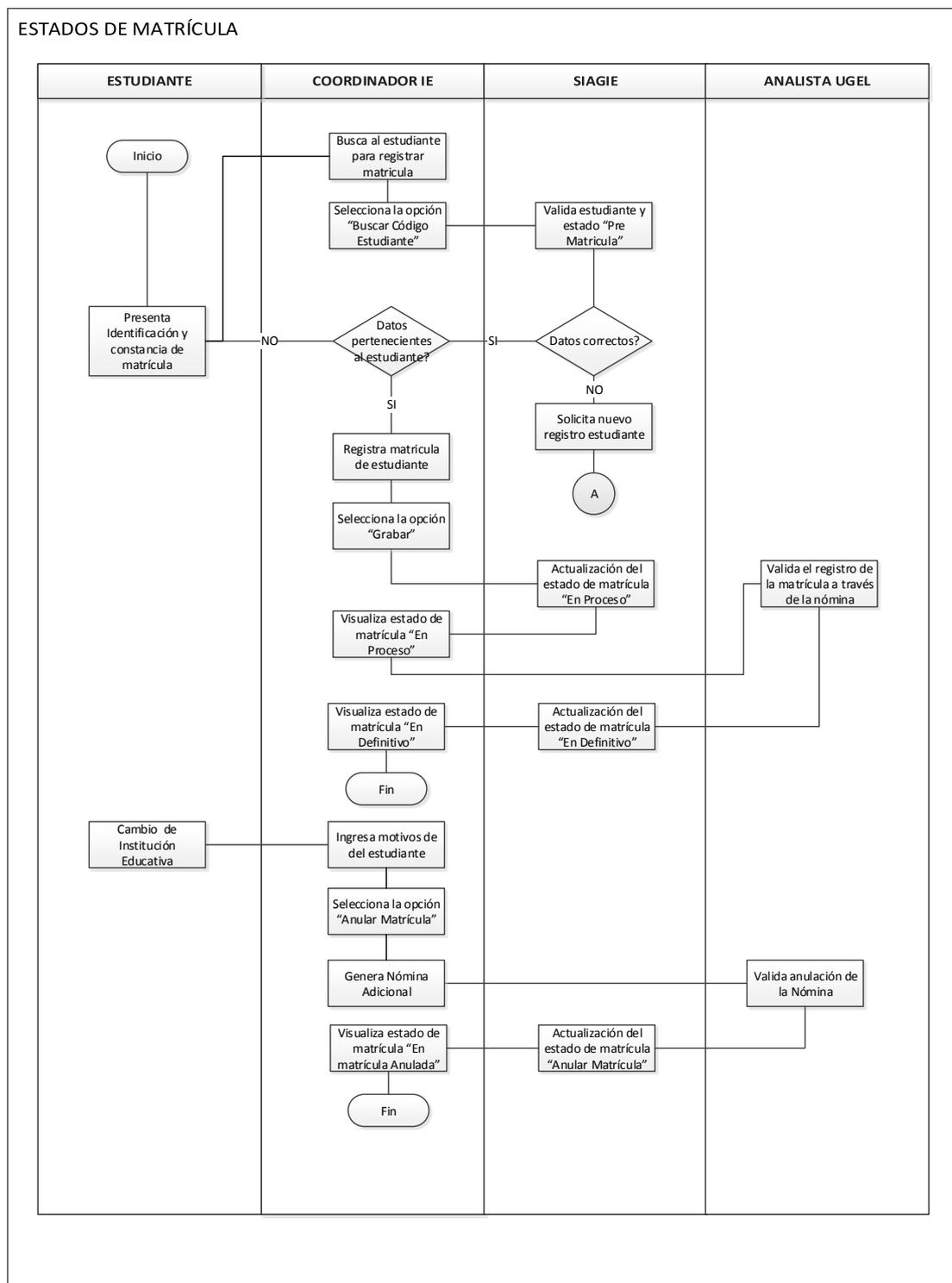


Figura 3.3 Estados de la matrícula
Elaboración: Las autoras

Modelo de Estados de nómina

La nómina es el documento técnico pedagógico que contiene la relación de estudiantes matriculados por sección en la IE. Es elaborada por la IE y aprobada por Resolución del Director de la misma, cuya copia se remite a la Unidad de Gestión Educativa Local o Dirección Regional de Educación para conocimiento y fines estadísticos. En la Tabla 3.13, se describen los estados de las nóminas:

Tabla 3.13 Estados de nómina

Estado Nómina	Estado Consolidador	Descripción	Desencadenador al siguiente Estado
Nómina Remitida	Nómina por Remitir	IE cuya nómina se realiza el envío al SIAGIE	Realizada la remisión de la nómina se actualiza el estado en "Nomina Generada"
Nómina Generada	Nómina por Generar	IE cuya nómina se realiza el registro de matrículas relación de estudiantes	Fecha de matrícula sea mayor que la fecha de cierre de matrícula de inicio de año. Dentro del rango de 45 días posteriores a la fecha de inicio del año escolar
Nómina Aprobada	Nómina por Aprobar	Analista de la OFIN a través del SIAGIE cuya nómina cumple con los requisitos de registro	Contiene todos los requisitos de la nómina se actualiza el estado en "Nomina Aprobada"
Nómina Rechazada	Nómina por Rechazar	Analista de la OFIN a través del SIAGIE cuya nómina ha sido anulada por el SIAGIE.	Falta uno de los requisitos de la nómina se actualiza el estado en "Nomina Rechazada"

Elaboración: Las autoras

En la Figura 3.4, se muestra el siguiente gráfico del flujo del estado de nóminas:

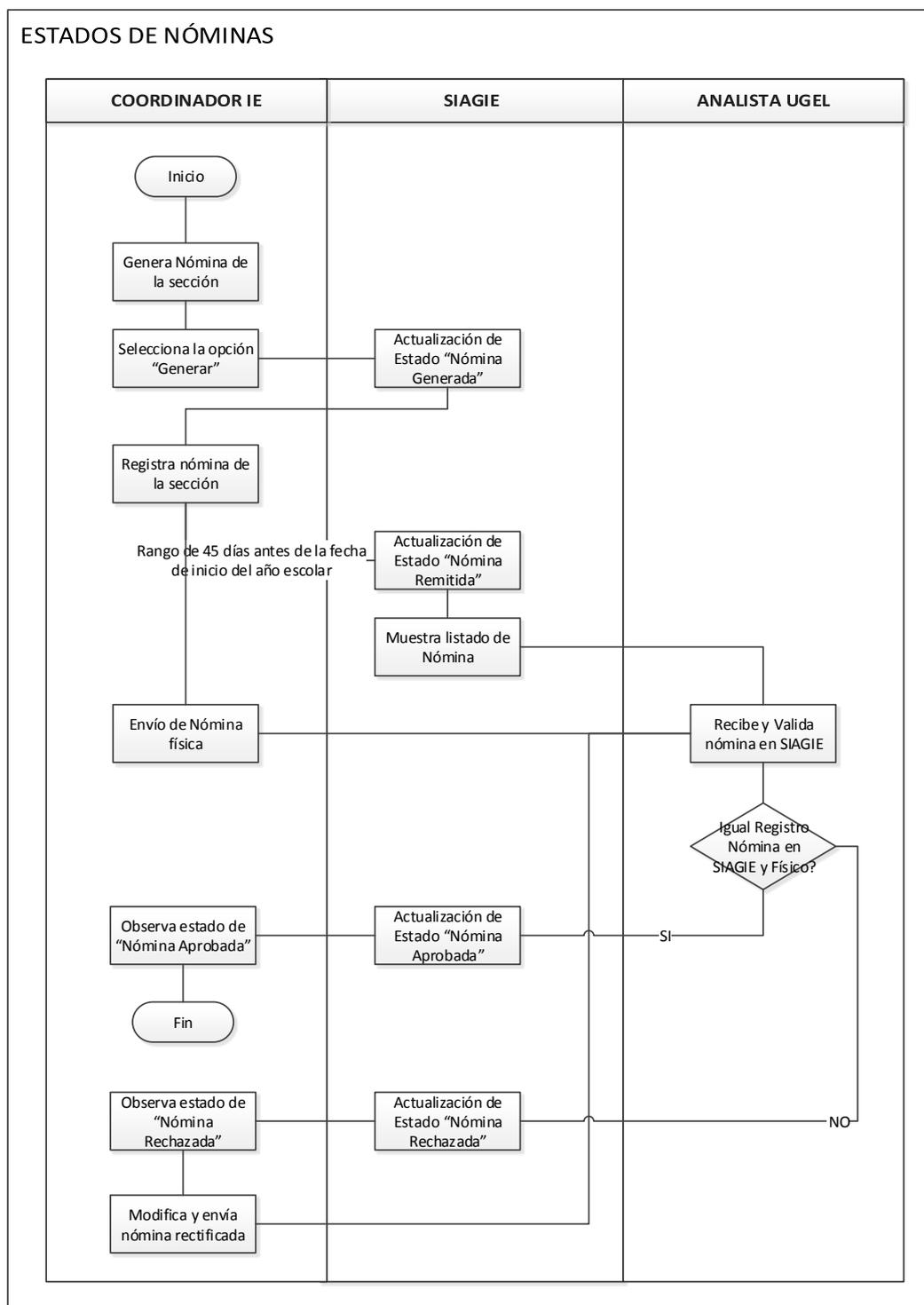


Figura 3.4 Estados de la nómina
Elaboración: Las autoras

Para el proyecto se tiene el apoyo del SIAGIE (Sistema de Información de Apoyo a la Gestión de la Instituciones Educativa) que cuenta con los procesos de Matrícula, Nóminas, Estudiantes, Aula; de estos procesos se enfocan, según alcance definido, por los procesos de matrículas y nóminas.

3.3.2 Análisis de datos

En este punto, después de haber realizado la definición de los indicadores a base de los requerimientos del área se procede a realizar el análisis de datos. En la tabla 3.14, se puede revisar la relación existente entre los indicadores planteados y los objetivos del área.

Tabla 3.14 Definición de Indicadores

Objetivo Estratégico de SPE	Indicador	Definición
Garantizar la prestación del servicio de enseñanza de las niñas, niños y adolescentes en las Instituciones Educativas de Educación Básica Regular.	% pre matriculados en Departamento	Mide la proporción de pre matriculados sobre el total de estudiantes en el sistema educativo en el Departamento
	% matrícula en proceso por Departamento	Mide la proporción estado matrícula en proceso sobre el total de estudiantes en el sistema educativo en el Departamento
	% matrícula definitiva en Departamento	Mide la proporción estado matriculada Definitivo sobre el total de estudiantes en el sistema educativo en el Departamento
	% matrícula anulada en Departamento	Mide la proporción estado matrícula anulada sobre el total de estudiantes en el sistema educativo en el Departamento
	% pre matriculados en el Distrito	Mide la proporción de pre matriculados sobre el total de estudiantes en el sistema educativo en el Distrito
	% matrícula en proceso en Distrito	Mide la proporción estado matrícula en proceso sobre el total de estudiantes en el sistema educativo en el Distrito
	% matrícula definitiva en Distrito	Mide la proporción estado matriculada Definitivo sobre el total de estudiantes en el sistema educativo en Distrito
	% matrícula anulada en Distrito	Mide la proporción estado matrícula anulada sobre el total de estudiantes en el sistema educativo en Distrito
	% pre matriculados en del IE	Mide la proporción de pre matriculados sobre el total de estudiantes en el sistema educativo en el IE
	% matrícula en proceso por IE	Mide la proporción estado matrícula en proceso sobre el total de estudiantes en el sistema educativo en el IE
	% matrícula definitiva en IE	Mide la proporción estado matriculada Definitivo sobre el total de estudiantes en el sistema educativo en IE
% matrícula anulada en IE	Mide la proporción estado matrícula anulada sobre el total de estudiantes en el sistema educativo en IE	
Garantizar los registros de las nóminas realizadas por las Instituciones Educativas de Educación Básica Regular.	% nómina Generada	Mide la cantidad de nóminas que están en estado Generada
	% nómina Remitida	Mide la cantidad de nóminas que están en estado Remitida
	% nómina Aprobada	Mide la cantidad de nóminas que están en estado Aprobada
	% nómina Rechazada	Mide la cantidad de nóminas que están en estado Rechazada

Objetivo Estratégico de SPE	Indicador	Definición
Garantizar la participación de Estudiantes en el sistema educativo.	% de matrícula según género	Mide la proporción de matriculados sobre el total de estudiantes por género del estudiante
	% estado de matrícula por grado	Mide la proporción de estado de matrícula por grado sobre el total de estudiantes en el sistema educativo
	% estado del estudiante	Mide la proporción de estado de estudiante sobre el total de estudiantes en el sistema educativo
	% de estudiantes por nivel educativo	Mide a los estudiantes que están por niveles educativos por
	Cantidad de estudiantes	Número de estudiantes en el sistema educativo

Elaboración: Las autoras

Se tiene como indicador principal el porcentaje de alumnos matriculados por IE porque a través de este cálculo se puede realizar una gestión adecuada de los profesores para contratarlos para el año escolar. En la Tabla 3.15 se puede apreciar el Cálculo de Indicadores.

Tabla 3.15 Cálculo de Indicadores

Indicador	Cálculo
% pre matriculados en Departamento	Σ Estado Pre Matriculado por Departamento / Total Sistema Educativo Departamento * 100%
% matrícula en proceso por Departamento	Σ Estado Matrícula en proceso por Departamento / Total Sistema Educativo Departamento * 100%
% matrícula definitiva en Departamento	Σ Estado Matriculada Definitivo por Departamento / Total Sistema Educativo Departamento * 100%
% matrícula anulada en Departamento	Σ Estado Matrícula anulada por Departamento / Total Sistema Educativo Departamento * 100%
% pre matriculados en el Distrito	Σ Estado Pre Matriculado por Distrito / Total Matriculados * 100%
% matrícula en proceso en Distrito	Σ Estado Matrícula en proceso por Distrito / Total Sistema Educativo Distrito * 100%
% matrícula definitiva en Distrito	Σ Estado Matriculada Definitivo por Distrito / Total Sistema Educativo Distrito * 100%
% matrícula anulada en Distrito	Σ Estado Matrícula anulada por Distrito / Total Sistema Educativo Distrito * 100%
% pre matriculados en del IE	Σ Estado Pre Matriculado por IE / Total Matriculados * 100%
% matrícula en proceso por IE	Σ Estado Matrícula en proceso por IE / Total Sistema Educativo IE * 100%
% matrícula definitiva en IE	Σ Estado Matriculada Definitivo por IE / Total Sistema Educativo IE * 100%
% matrícula anulada en IE	Σ Estado Matrícula anulada por IE / Total Sistema Educativo IE * 100%
% nómina Generada	Total nóminas con indicador de nóminas "Generada" / Total de Nóminas
% nómina Remitida	Total nóminas con indicador de nóminas "Remitida" / Total de Nóminas
% nómina Aprobada	Total nóminas con indicador de nóminas "Aprobada" / Total de Nóminas
% nómina Rechazada	Total nóminas con indicador de nóminas "Rechazada" / Total de Nóminas
% de matrícula según género	Σ Matriculado por género por Departamento / Total Matriculados * 100%
% estado de matrícula por grado	Σ Estado estudiante Matrícula por grado/ Total Sistema Educativo Departamento * 100%
% estado del estudiante	Σ Estado estudiante/ Total Sistema Educativo Departamento * 100%
% de estudiantes por nivel educativo	Σ Estudiantes por nivel educativo/ Total Estudiantes * 100%
Cantidad de estudiantes	Σ Estudiantes en el sistema educativo

Elaboración: Las autoras

*

Luego se realiza la definición de las variables de análisis. En la Tabla 3.16, se muestra las variables de análisis:

Tabla 3.16 Variables de análisis

Variables de Análisis	Descripción	Ejemplos
Identificación de Institución Educativa	Datos de la Institución Educativa	Identificación de la IE
Código de estudiante	Identificación del Estudiante	Código Estudiante
Género de Estudiante	Género del estudiante	Masculino, Femenino
Ubicación	Localización de las Instituciones educativas	Departamento, Provincia, Distrito
Modalidad IE	Clasificación de las modalidades	Educación Básica Regular, Educación Básica Alternativa, Educación Básica Especial
Estado de Matrícula	Estado de matrícula	Pre Matriculado, En Proceso, Definitivo, Anulado
Estado Estudiante	Clasificación de estudiantes	Aprobado, Desaprobado
Estado Grado Educativo	Clasificación de grado por IE	Primero, Segundo, Tercero, Cuarto
Nivel Educativo	Cantidad de grupos dentro del nivel educativo	Cuna, Inicial, Primaria, Secundaria
Estado de Nómina	Clasificación de los estados de nómina	Generada, Remitida, Aprobada, Rechazada

Elaboración: Las autoras

En la Tabla 3.17 se muestra el detalle de las variables de análisis que aplican por cada indicador:

Tabla 3.17 Matriz de variables de análisis versus indicadores

Variable de Análisis	% pre matriculados en Departamento	% pre matrícula en proceso en Departamento	% pre matrícula en Definitiva en Departamento	% pre matrícula en anulada en Departamento	% pre matriculados en Distrito	% pre matrícula en proceso en Distrito	% pre matrícula en Definitiva en Distrito	% pre matrícula en anulada en Distrito	% pre matriculados en IE	% pre matrícula en proceso en IE	% pre matrícula en Definitiva en IE	% pre matrícula en anulada en IE	% de matriculados según género	% estado de matrícula por grado	% estado del estudiante	% matrícula según género	Número de estudiantes por aula	% nómina Generada	% nómina Remitida	% nómina Aprobada	% nómina Rechazada	% de estudiantes por nivel educativo
Identificación Institución Educativa									X	X	X	X										X
Código de estudiante	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X					X
Género de Estudiante													X	X		X	X	X				
Ubicación	X	X	X	X	X	X	X	X					X	X	X	X	X	X	X	X	X	X
Modalidad IE	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Estado de Matrícula	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X						
Estado Estudiante															X							
Estado Grado Educativo																			X	X	X	
Nivel Educativo																						X
Estado de Nómina																			X	X	X	

Elaboración: Las autoras

3.3.2.1 Análisis dimensional

Las dimensiones se determinaron a base de las variables de análisis, los usuarios de la secretaría de Planificación Estratégica realizan sus análisis. Para concretar las dimensiones con el que va a contar el Datamart se empezó identificando las variables de análisis por las cuales el usuario realiza en el trabajo del día realiza sus reportes. En la Tabla 3.18, se presentan las variables de análisis:

Tabla 3.18 Variables de análisis

Nro.	Variables de Análisis
1	Identificación institución educativa
2	Código de estudiante
3	Género de estudiante
4	Ubicación
5	Modalidad IE
6	Estado de matrícula
7	Estado estudiante
8	Estado grado educativo
9	Nivel educativo
10	Estado de nómina

Elaboración: Las autoras

Luego realizamos la agrupación de las variables de análisis por afinidad entre ellas, cada una son atributos de una entidad importante la cual decidiremos para que pueda ser una dimensión. En la Tabla 3.19, se presenta la agrupación de las variables de análisis con las dimensiones propuestas:

Tabla 3.19 Dimensiones con variables de análisis

Dimensión	Variabes de Análisis
Institución Educativa	Identificación de Institución Educativa, Modalidad de IE
Estudiante	Código de Estudiante, Género de Estudiante
Ubicación	Ubicación Geográfica
Estado Matrícula, Tiempo	Estado de matrícula, Año, Mes, Día
Modalidad Inst. Educativa	Modalidad IE
Grado Educativo	Estado de Grado Educativo
Nivel Educativo	Modalidad de IE
Estado Estudiante	Descripción Estado de Estudiante
Tipo Formato, Estado Formato, Sección	Estado Nómina

Elaboración: Las autoras

Luego del análisis anterior se concluye que las dimensiones que conforman el Datamart son. En la Tabla 3.20, se presentan las dimensiones del Datamart:

Tabla 3.20 Dimensiones del Datamart

Nro.	Dimensión
1	Identificación de institución educativa
2	Código de estudiante
3	Ubicación
4	Tiempo
5	Modalidad institución educativa
6	Estado matrícula
7	Grado educativo
8	Nivel educativo
9	Estado estudiante
10	Estado sección
11	Tipo de formato
12	Estado de formato

Elaboración: Las autoras

Las dimensiones mencionadas son explicadas con mayor detalle, la definición y jerarquías de las dimensiones en el Anexo 04 Análisis Dimensional.

Fact

Las Facts contienen las dimensiones que definen su nivel de detalle que conllevan al tema de análisis que corresponde. La Fact de matrículas contienen la información del proceso de matrículas y la Fact de Formato tiene el alcance sólo al Proceso de Nóminas que será considerado para este proyecto. En la Tabla 3.21, presentamos la relación de las Facts y las dimensiones a utilizar en el Datamart:

Tabla 3.21 Facts versus Dimensiones

Dimensiones	FACT	
	Fact de Matrículas	Fact de Formato
Dimensión Institución Educativa	X	X
Dimensión Estudiante	X	
Dimensión Ubicación	X	X
Dimensión Tiempo	X	X
Dimensión Modalidad Institución Educativa	X	X
Dimensión Estado Matrícula	X	
Dimensión Nivel Educativo	X	X
Dimensión Estado Estudiante	X	
Dimensión Detalle de Formato		X
Dimensión Tipo de Formato		X
Dimensión Estado de Formato		X

Elaboración: Las autoras

3.3.3 Prototipo de la aplicación

El prototipo de reportes se realizó acorde con los indicadores que el usuario final desea medir del Proceso de Matrículas y Nóminas, en esta primera fase. En la Figura 3.5, se puede visualizar el prototipo de matrícula, en la Figura 3.6 se puede visualizar el prototipo nóminas y en la Figura 3.7, se puede visualizar el prototipo de estudiantes.

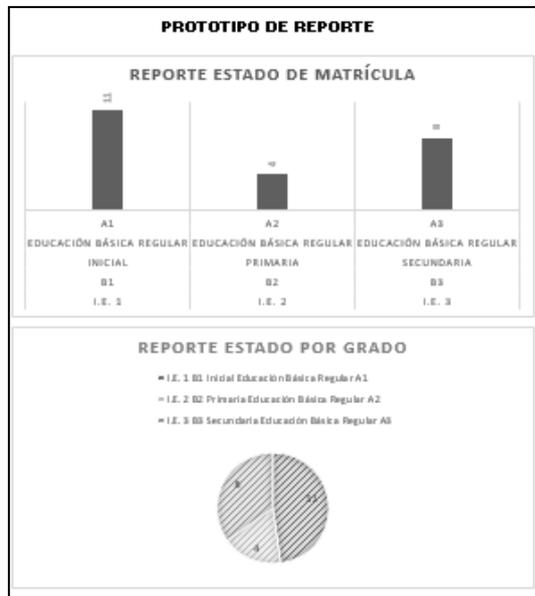


Figura 3.5 Vista Prototipo de Matrícula
Elaboración: Las autoras

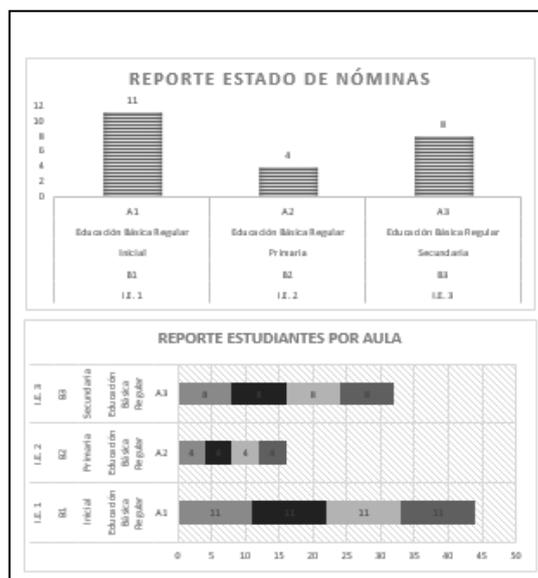


Figura 3.6 Vista Prototipo de Nóminas
Elaboración: Las autoras

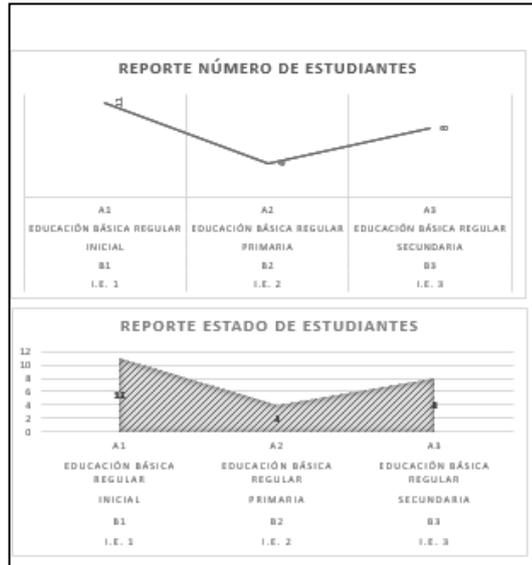


Figura 3.7 Vista Prototipo de Estudiantes
Elaboración: Las autoras

Definición de reporte

Se han definido los reportes teniendo en cuenta la información proporcionada por la Secretaría de Planificación Estratégica. Se ha incluido, en el análisis, los reportes que actualmente se emiten y se han agregado reportes propuestos con sus respectivos detalles. Ver Tabla 3.22.

Tabla 3.22 Definición de reportes

	Indicador	Nombre del Reporte	Descripción del contenido	Información contenida	Filtros Especiales
REPORTES EXISTENTES	% pre matriculados % matrícula en proceso % matrícula definitiva % matrícula anulada % matrícula según genero	Reporte de estado matrícula	Contendrá datos de los estados de la matrícula.	-Nombre de Institución Educativa, Ubicación IE, Tipo de matrícula, Modalidad IE	-Nombre de Institución Educativa, Estado Matrícula, Ubicación IE, Modalidad IE con "Educación Básica Regular"
	% nóminas Generada % nóminas Remitida % nóminas Aprobada % nóminas Rechazada	Reporte de Nómina	Contendrá datos de los estados de nóminas.	-Estado de Nómina, Modalidad IE, Nivel IE	-Estado de Nómina remitida, Estado de Nómina aprobada, Estado de Nómina Generada, Estado de Nómina Rechazada, Modalidad IE con "Educación Básica Regular", Niveles IE
REPORTES AGREGADOS	% Estado del estudiante	Reporte de estado del estudiante	Contendrá % de la situación de los estudiantes.	-Nombre de Institución Educativa, Ubicación IE, Tipo de matrícula, Modalidad IE, Estudiante	-Nombre de Institución Educativa, Estado Matrícula, Ubicación IE, Modalidad IE con "Educación Básica Regular"
	% Estudiantes por Nivel Educativo.	Reporte de estudiantes por nivel educativo	Contendrá % estudiantes que están por niveles educativos por nivel educativo	-Tipo de Matrícula -Modalidad de IE -Nivel de IE	-Todas Instituciones Educativas -Modalidad IE con "Educación Básica Regular", Niveles IE
	Número de estudiantes en el sistema educativo	Reporte de Participación de estudiantes por género	Contendrá el número de los estudiantes por género	-Tipo de Matrícula -Modalidad de IE -Nivel de IE -Género	-Todas Instituciones Educativas -Modalidad IE con "Educación Básica Regular", Niveles IE
	% estado de matrícula por grado	Reporte de estado por Grado	Contendrá % estudiantes que están por grado educativos	-Nombre de Institución Educativa, Ubicación IE, Tipo de matrícula, Modalidad IE, Grado	-Nombre de Institución Educativa, Estado Matrícula, Ubicación IE, Modalidad IE con "Educación Básica Regular"

Elaboración: Las autoras

3.4 Etapa de diseño

3.4.1 Diseño de la Base de datos

En la Figura 3.8, se muestran los componentes más representativos del modelo, se puede apreciar que se cuenta con un Fact Matrícula central que contiene los indicadores y a su alrededor se encuentran las variables de análisis más frecuentes: estudiante, institución educativa, ubicación, estado matrícula, nivel educativo, grado educativo, estado del estudiante.

En la Figura 3.9, se muestra que se cuenta con un Fact Formato central que contiene los indicadores y a su alrededor se encuentran las variables de análisis más frecuentes: sección, institución educativa, ubicación, grado educativo, modalidad de la institución educativa.

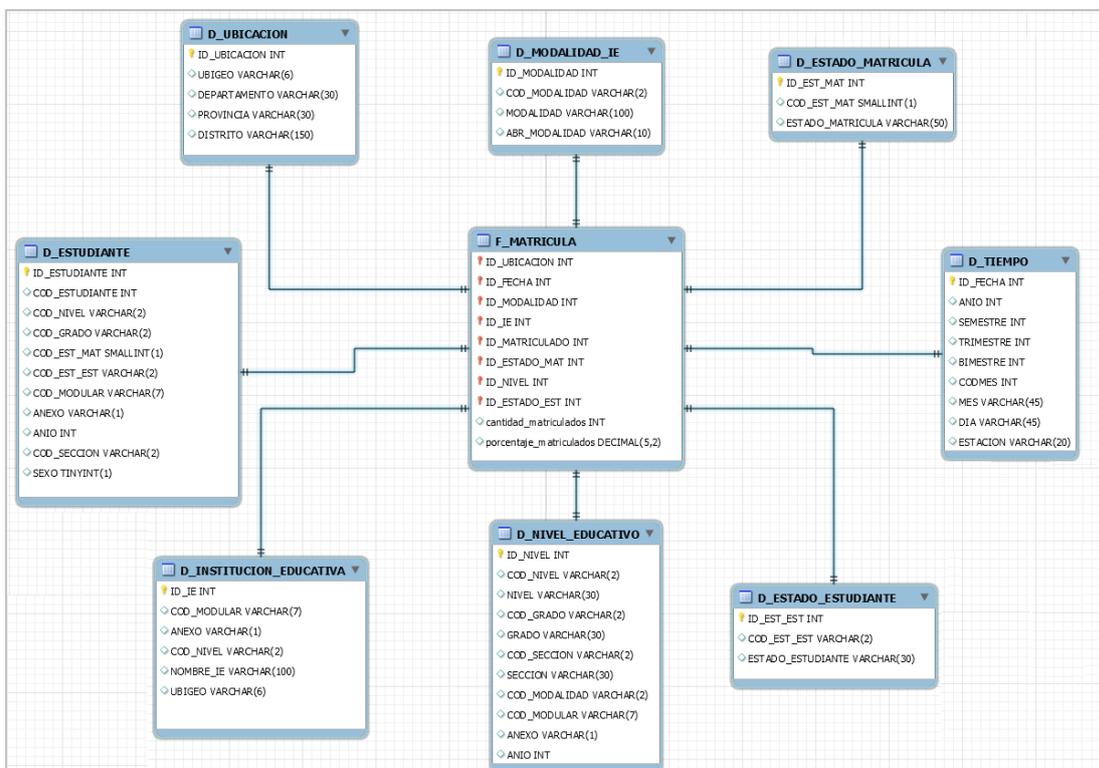


Figura 3.8 Vista Principal Modelo de Datos de Fact Matrícula
Elaboración: Las autoras

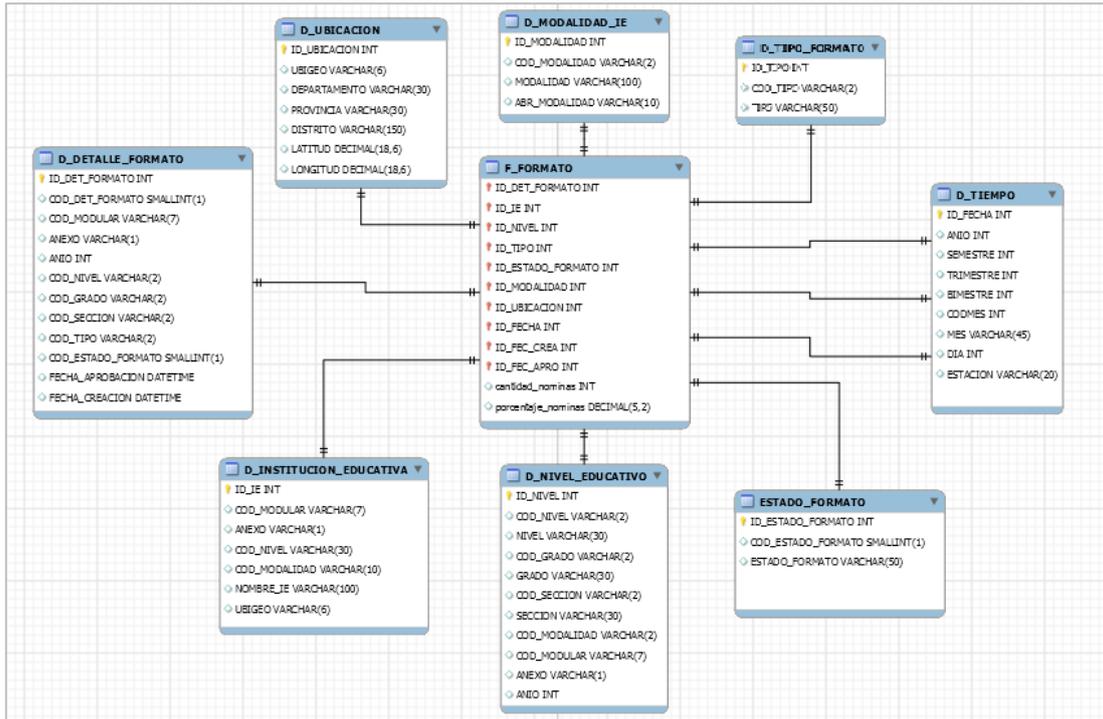


Figura 3.9 Vista Principal Modelo de Datos de Fact Formato

Elaboración: Las autoras

Arquitectura de la solución

En la figura 3.10, se presenta el Datamart de OFIN cuenta con la siguiente arquitectura:

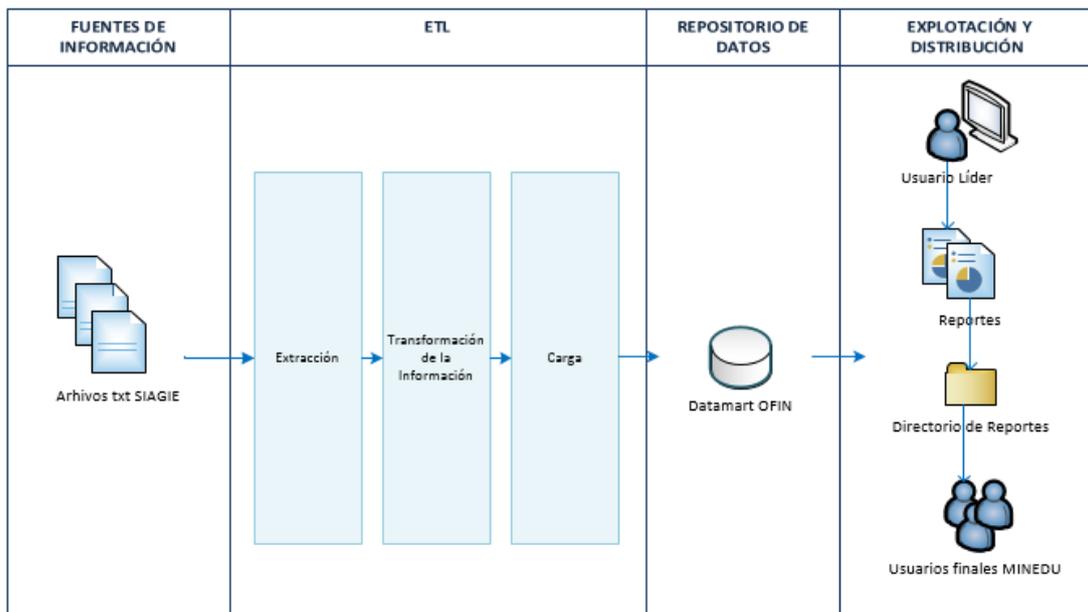


Figura 3.10 Arquitectura de la Solución

Elaboración: Las autoras

La extracción de los datos se realiza desde la base de datos de la propia organización y estos pasarán por el proceso ETL para luego cargar el Datamart para luego ser entregada la información a los usuarios mediante la herramienta de explotación.

Seguridad de la información

La seguridad en el Datamart es un factor crítico debido a que se cuenta con información real que abarca la gestión Educativa que es captado a través del SIAGIE es por ello que representa información sensible. La herramienta ETL que se ha propuesto por ser código abierto no se consideraran gastos de licencia, sin embargo, debemos tomar medidas para la seguridad de la información, por ello se requiere un servidor de aplicaciones, un servidor de base de datos y un dominio. A continuación, en la Figura 3.11, se propone al siguiente diagrama de red.

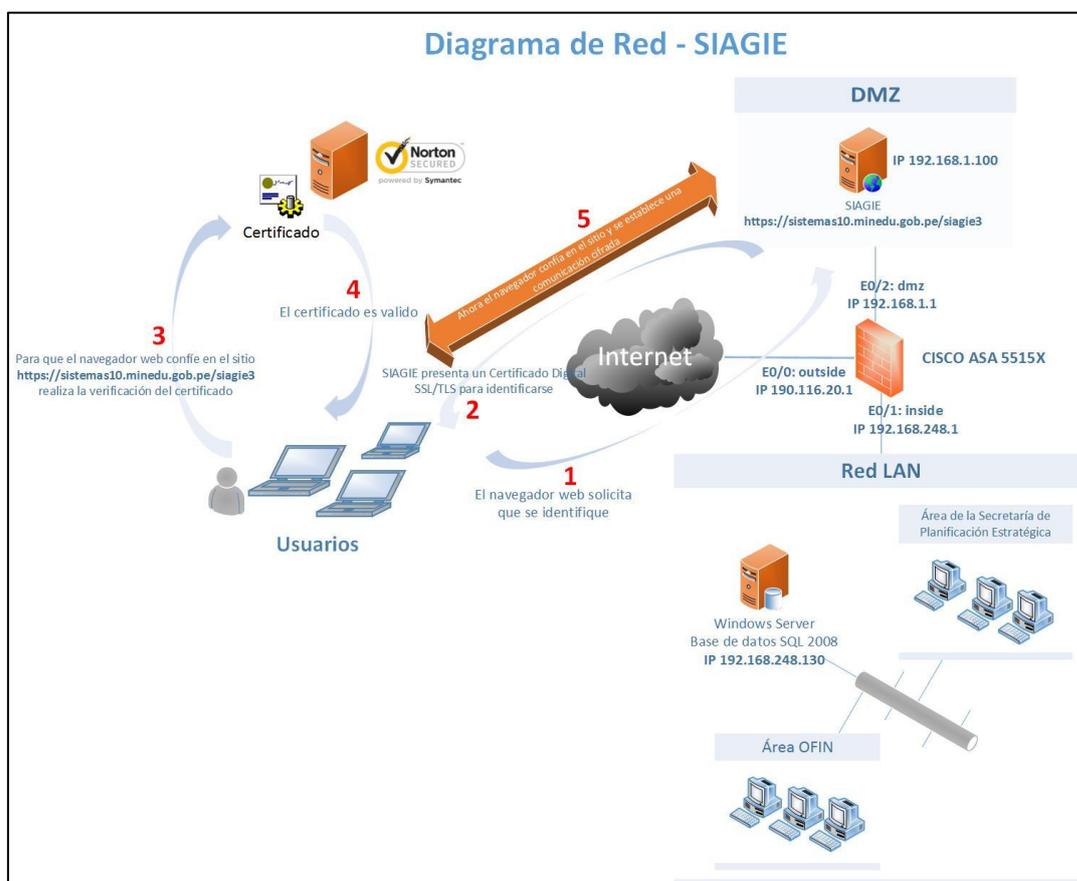


Figura 3.11 Diagrama de Red
Elaboración: Las autoras

El diagrama de red propuesto consta de los siguientes elementos:

Certificado digital SSL/TLS

Para garantizar la vinculación entre la entidad y la llave pública. En ella se registrará la información del propietario como nombre, correo electrónico, la organización a la que pertenece y su llave pública.

Firewall Cisco

Firewall CISCO permitirá validar solo los puertos necesarios para la comunicación entre el Servidor de Base de Datos y el de Servidor de Aplicaciones.

DMZ (Red Perimetral)

Permitirá que las conexiones desde la red interna puedan ser accesibles a la zona desde el exterior sin correr el riesgo de comprometer la seguridad de la organización, actuando como un filtro entre la conexión de internet y la red de equipos (host).

En el diagrama de red propuesto, se detallan los cinco pasos a realizarse:

Paso 1

- Cuando el navegador hace una petición al sitio seguro de SIAGIE, este envía un mensaje donde indica que quiere establecer una conexión segura y envía datos sobre la versión del protocolo SSL/TLS.

Paso 2

- A base de esta información enviada por el navegador, el servidor web de SIAGIE responde con un mensaje informando que está de acuerdo en establecer la conexión segura con los datos de SSL/TLS proporcionados.

- Una vez que ambos conocen los parámetros de conexión, el sitio del SIAGIE presenta su certificado digital al navegador web para identificarse como un sitio confiable.

Paso 3

El navegador cuando cuenta con el certificado del sitio web de SIAGIE, realiza algunas verificaciones antes de confiar en el sitio:

- Integridad del certificado: Verifica que el certificado se encuentre íntegro, esto lo hace descifrando la firma digital incluida en él mediante la llave pública de la AC y comparándola con una firma del certificado generada en ese momento, si ambas son iguales entonces el certificado es válido.
- Vigencia del certificado: Revisa el periodo de validez del certificado, es decir, la fecha de emisión y la fecha de expiración incluidos en él.
- Verifica emisor del certificado: Hace uso de una lista de Certificados raíz almacenados en tu computadora y que contienen las llaves públicas de las ACs conocidas y de confianza.

Paso 4

- A base de esta lista, el navegador revisa que la AC del certificado sea de confianza, de no serlo, el navegador mostrará una advertencia indicando que el certificado fue emitido por una entidad en la cual no confía.

Paso 5

- Una vez que el certificado cumplió con todas las pruebas del navegador, se establece la conexión segura al sitio de SIAGIE.

Acceso de información

Para el acceso, visualización de información y elaboración de los reportes se deben definir los perfiles de usuarios en función de los siguientes pasos:

Seguridad de acceso a los reportes

En este punto, se proponen los tipos de roles para el acceso:

Rol 1 – Lectura

- Solo consulta reportes
- No puede crear reportes
- No puede modificar reportes

Rol 2 – Lectura/ Escritura

- Consultar reportes
- Crear reporte
- Modificar reportes

Seguridad de Acceso al Proceso ETL

En este punto, se proponen los tipos de roles para el acceso:

Rol 1 - Lectura

- Solo consulta al proceso ETL
- No puede crear al proceso ETL
- No puede modificar al proceso ETL

Rol 2 – Lectura/ Escritura

- Consultar al proceso ETL
- Crear al proceso ETL
- Modificar al proceso ETL

La Oficina de Informática cuenta con el procedimiento para registrar el acceso de los usuarios a las transacciones de las aplicaciones.

Con el fin de minimizar restricción de la información impartida por el Ministerio de Educación estará regida por la Ley N°27927, Ley que modifica la Ley N°27806, Ley de transparencia y acceso a la información pública.

3.4.2 Diseño del ETL

En este punto, se expone el diseño propuesto de solución que conformará la solución, así como los procesos de ETL (Extracción, Transformación y Carga) involucrados, se mostrará el proceso de extracción, transformación y carga del Datamart de OFIN. Luego se presenta a modo de ejemplo, la dimensión institución educativa. Mayor detalle de las dimensiones se presenta en el Anexo 05 - Diseño de ETL

Carga de la Dimensión institución educativa

a) Descripción

El presente proceso se realiza la extracción del universo de institución educativa con el que contará el Datamart, contiene la relación de la información de las instituciones educativas.

b) Descripción de tablas Fuentes

Tipo de Fuente	Nombre de la Tabla	Descripción
TXT	INSTITUCION_EDUCATIVA	La tabla contiene información acerca de las Instituciones Educativa que se encuentran registrados en el SIAGIE
TXT	NIVEL_EDUCATIVO	La tabla contiene información acerca de los niveles educativos de las Instituciones Educativas.
TXT	MODALIDAD	La tabla contiene información acerca de las modalidades de las Instituciones Educativas.
TXT	PATRON	La tabla contiene información acerca de las ubicaciones de las Instituciones Educativas.

c) Estandarización de Datos y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
COD_MOD	PK	VARCHAR(7)	TEXTO	NO NULL	NO TIENE
ANEXO		VARCHAR(1)	TEXTO		NO TIENE
NIVEL		VARCHAR(30)	TEXTO		NO TIENE
ABR_MODALIDAD		VARCHAR(10)	TEXTO		NO TIENE
NOMBRE_IE		VARCHAR(100)	TEXTO		NO TIENE
UBIGEO		VARCHAR(6)	TEXTO		NO TIENE

d) Fuente de Datos

Tablas: INSTITUCION_EDUCATIVA, NIVEL_EDUCATIVO, MODALIDAD, PATRON					
Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
COD_MOD	PK	CHAR(7)	TEXTO	NO NULL	NO TIENE
ANEXO		CHAR (1)	TEXTO		NO TIENE
DSC_NIVEL		CHAR(30)	TEXTO		NO TIENE
ABR_MODALIDAD		VARCHAR(10)	TEXTO		NO TIENE
CEN_EDU		VARCHAR (100)	TEXTO		NO TIENE
CODGEO		CHAR(6)	TEXTO		NO TIENE

e) Tabla Destino

Tabla	D_INSTITUCION_EDUCATIVA	
Campo	Tipo	Mapeo
COD_MOD	CHAR(7)	INSTITUCION_EDUCATIVA. COD_MOD
ANEXO	CHAR (1)	INSTITUCION_EDUCATIVA. ANEXO
NIVEL	CHAR (2)	NIVEL_EDUCATIVO. DSC_NIVEL
ABR_MODALIDAD	VARCHAR(10)	MODALIDAD. ABR_MODALIDAD
NOMBRE_IE	VARCHAR (100)	INSTITUCION_EDUCATIVA. CEN_EDU
UBIGEO	VARCHAR(6)	PATRON. CODGEO

3.5 Etapa de construcción

En esta etapa, se contemplaron los pasos principales desde la configuración de la herramienta hasta su construcción tomando como base el Análisis y Diseño puntos mencionados anteriormente en este proyecto.

▪ **Configuración del Software**

Las configuraciones fueron realizadas en el Sistema Operativo Windows 7, adicional se muestra las configuraciones de la Base de Datos, Pentaho BI (Kettle) y Tableau para la construcción del ETL y explotación de reportes.

- Base de Datos: SQL Server 2008.
- Entorno de Desarrollo: SQL Server 2008.
- Herramienta de ETL: Pentaho Data Integration (Kettle) 5.2.0.
- Herramienta de Explotación: Tableau Desktop 8.2
- Software de oficina para la documentación: Microsoft Office 2007.

Se Utilizará la Base de datos SQL 2008 Server, tanto para las transaccionales y repositorio del ETL.

▪ **Configuración de SQL Server 2008**

Para realizar la conexión de SQL Server 2008 con Pentaho hay que seguir lo siguiente:

- Primero descargar el driver JDBC para SQL Server 2008, luego de haber realizado la descarga del driver se coloca en la siguiente ruta:
Administration- console > jdbc.jar
- Antes de realizar el Test de Conexión, se solicita los siguientes datos:

Nombre	Se Ingresa el nombre del servidor
Driver Class	com.microsoft.sqlserver.jdbc.SQLServerDriver
User Name	Nombre de Usuario
Password	Clave del usuario
URL	Se ingresa la URL del servidor, en este caso se realizará de forma local (localhost)

▪ Configuración de Pentaho

Para la instalación de Pentaho previamente se debe descargar 7.1 JDK (Java Development Kit) un kit de desarrollo oficial del lenguaje de programación Java, luego se realiza la descarga en la página oficial de Pentaho en el sistema operativo que se va a trabajar.

- Para iniciar la instalación solo es necesario hacer clic en el ícono denominado Spoon del directorio "C:\pentaho\pdi-ce-5.2.0-stable\data-integration".
- Antes de realizar el Test de Conexión, se solicita los siguientes datos, ver Figura 3.12

Connection Name: Se ingresa nombre de la conexión.
Host Name: Se ingresa nombre de Host, en este caso local (LocalHost).
Database Name: Se ingresa el nombre de la Base de Datos.
Instance Name: Nombre de Instancia de SQL Server, en este caso dejaremos en blanco por defecto.
Port Number: Se ingresa número de puerto, en este caso 1433.
User Name: Se ingresa nombre de usuario.
Password: Se ingresa password del usuario.
Connection Type: Nombre de la conexión, en este caso MS SQL Server.
Access: Se ingresa el acceso JDBC, el nativo.

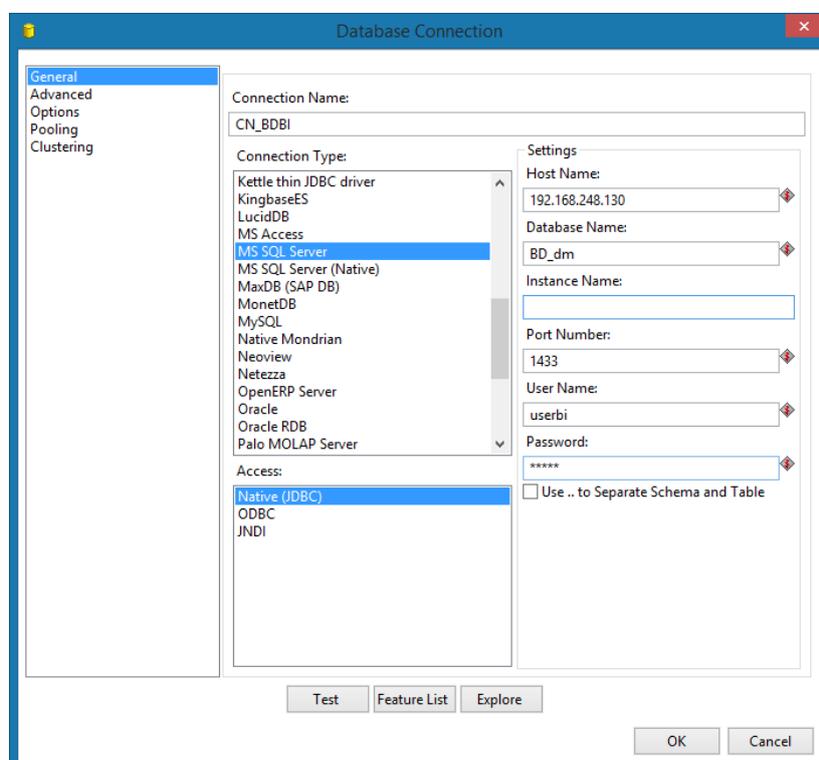


Figura 3.12 Configuración Pentaho con SQL
Elaboración: Las autoras

▪ Configuración de Tableau Desktop

Para la configuración de Tableau Desktop se realiza lo siguiente:

- Ingresar a Menu Datos, seleccionar Conectar Datos.
- Se elige el Servidor (SQL Server) establecido.
- En el caso de ser local, se coloca Localhost y el nombre con la contraseña respectiva. Ver Figura 3.13.

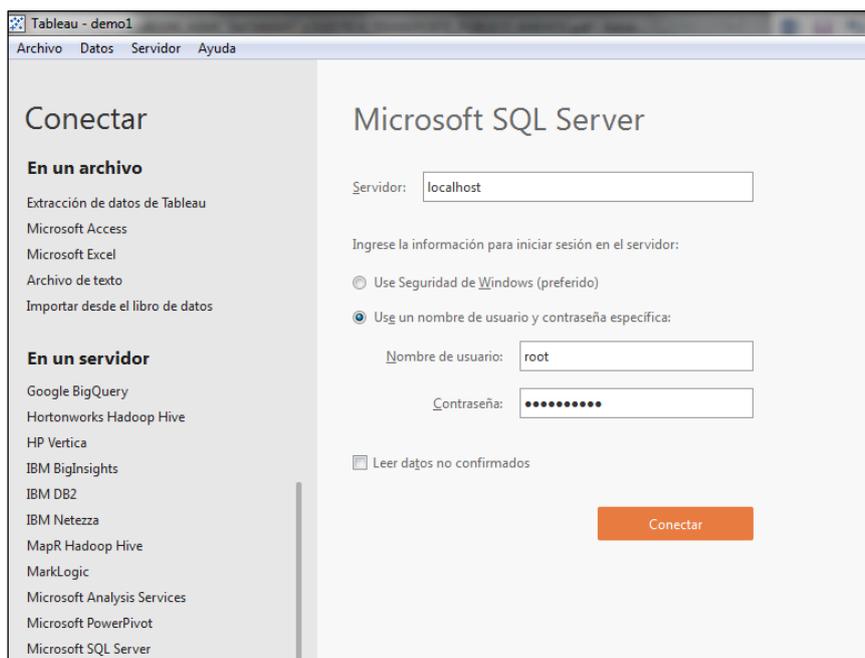


Figura 3.13 Configuración-Tableau con SQL
Elaboración: Las autoras

▪ Configuración del Firewall CISCO ASA

Para ayudar a mantener la información libre de toda alteración mal intencionado que dañe su integridad, se debe realizar la configuración de la Plataforma de Seguridad de la Información. Ver Anexo 12 Configuración de Firewall CISCO ASA.

3.5.1 Desarrollo del ETL

En la Figura 3.14 se muestra la carga de la Fact Matrícula y de la Fact Formato para el Datamart.



Figura 3.14 Carga de Facts en Datamart

Elaboración: Las autoras

3.5.1.1 Carga de dimensiones

En este punto, se presentaran los procesos de carga de las tablas de dimensiones y hechos, previamente se han debido realizar la instalación y configuraciones respectivas.

Para poder realizar la carga primero debemos realizar la extracción de la información para luego proceder a la transformación conformada por una serie de pasos (steps).

Proceso de Carga de la Dimensión institución educativa

i. Carga de registros en un archivo intermedio

Se extrajo los datos del archivo "TXT_INSTITUCION_EDUCATIVA", "TXT_PATRON" de acuerdo al mapeo y se cargó a al archivo intermedio "STG_D_INSTITUCION_EDUCATIVA".

ii. Carga de la dimensión

Se extrajo los valores del archivo "STG_D_INSTITUCION_EDUCATIVA" y se carga en la dimensión "D_INSTITUCION_EDUCATIVA". En caso los registros hagan referencia a nuevas Instituciones Educativas se insertan nuevas filas en la tabla; caso contrario, si los registros hacen referencia a antiguas Instituciones Educativas se actualiza los campos.

iii. Borrar el archivo intermedio

Se borró el archivo intermedio “STG_D_INSTITUCION_EDUCATIVA”.

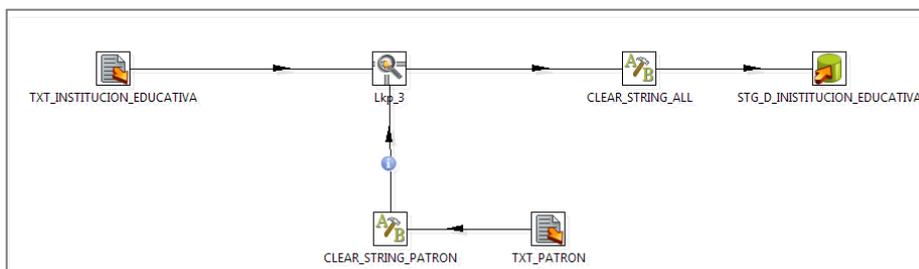


Figura 3.15 Carga de registros al archivo intermedio Institución Educativa
Elaboración: Las autoras



Figura 3.16 Carga de dimensión Institución Educativa
Elaboración: Las autoras

Proceso de Carga de la Dimensión estudiante

i. Carga de registros en un archivo intermedio

Se extrajo los datos del archivo “TXT_MATRICULA”, “TXT_NIVEL_EDUCATIVO”, “TXT_ESTADO_MATRICULA”, “TXT_PERSONA”, “TXT_GRADO”, “TXT_ESTADO_POR_GRADO”, “TXT_INSTITUCION_EDUCATIVA”, de acuerdo al mapeo y se cargó a al archivo intermedio “STG_D_ESTUDIANTE”.

ii. Carga de la dimensión

Se extrajo los valores del archivo “STG_D_ESTUDIANTE” y se carga en la dimensión “D_ESTUDIANTE”. En caso los registros se pueden agregar registros mas no actualizar los datos del estudiante.

iii. Borrar el archivo intermedio

Se borró el archivo intermedio “STG_D_ESTUDIANTE”.

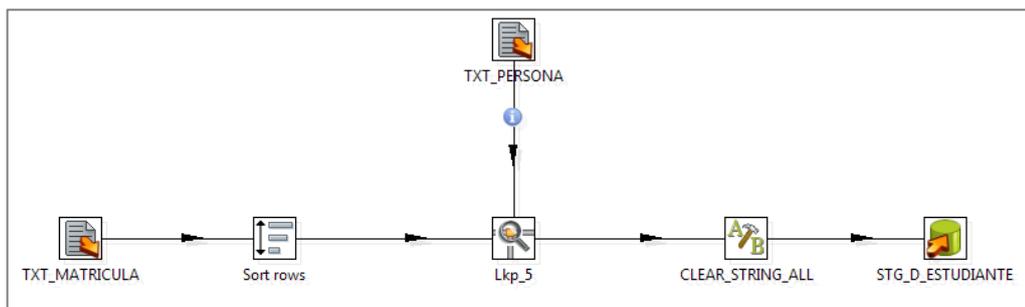


Figura 3.17 Carga de registros al archivo intermedio estudiante
Elaboración: Las autoras



Figura 3.18 Carga de dimensión estudiante
Elaboración: Las autoras

Proceso de Carga de la Dimensión ubicación

i. Carga de registros en un archivo intermedio

Se extrajeron los datos del archivo "TXT_UBIGEO", "TXT_PADRON", de acuerdo al mapeo y se cargó a al archivo intermedio "STG_D_UBICACION".

ii. Carga de la dimensión

Se extrajeron los valores del archivo "STG_D_UBICACION" y se carga en la dimensión "D_UBICACION". En caso los registros hagan referencia a una nueva ubicación se insertan nuevas filas en la tabla; caso contrario, si los registros hacen referencia a antiguas Ubicaciones se actualiza los campos.

iii. Borrar el archivo intermedio

Se borró el archivo intermedio "STG_D_UBICACION".

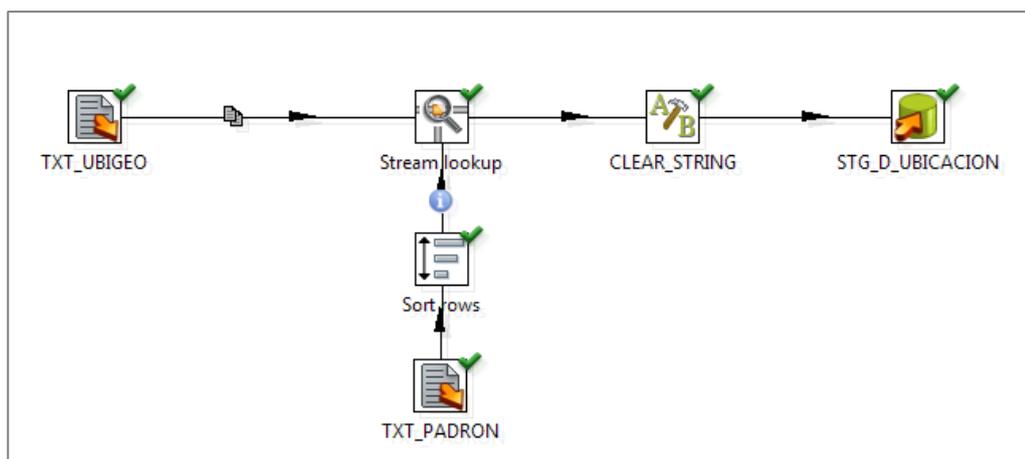


Figura 3.19 Carga de registros al archivo intermedio Ubicación
Elaboración: Las autoras

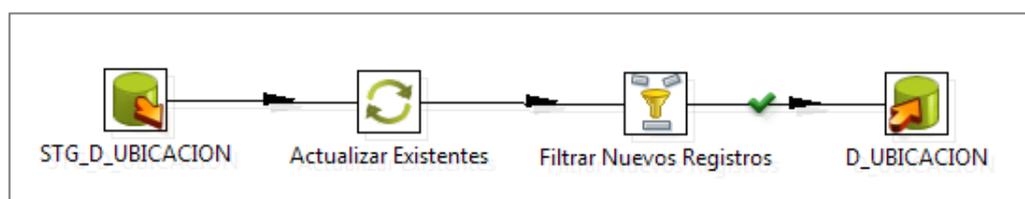


Figura 3.20 Carga de dimensión ubicación
Elaboración: Las autoras

Proceso de Carga de la Dimensión Modalidad institución educativa

i. Carga de registros en un archivo intermedio

Se extrajo los datos del archivo “TXT_MODALIDAD”, de acuerdo al mapeo y se cargó a al archivo intermedio “STG_D_MODALIDAD_IE”.

ii. Carga de la dimensión

Se extrajo los valores del archivo “STG_D_MODALIDAD_IE” y se carga en la dimensión “D_MODALIDAD_IE”. En caso los registros hagan referencia a nuevas Instituciones Educativas se insertan nuevas filas en la tabla; caso contrario, si los registros hacen referencia a antiguas Instituciones Educativas se actualiza los campos.

iii. Borrar el archivo intermedio

Se borró el archivo intermedio “STG_D_MODALIDAD_IE”.

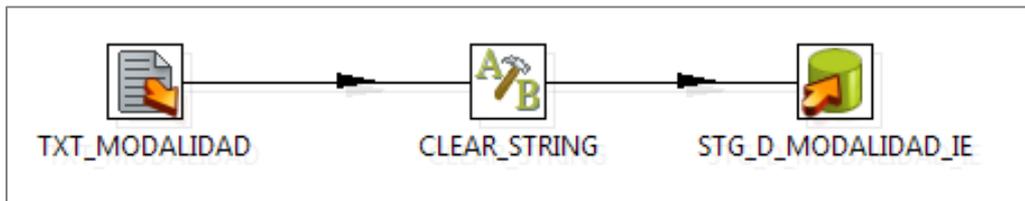


Figura 3.21 Carga de registros al archivo Modalidad Institución Educativa
Elaboración: Las autoras



Figura 3.22 Carga de dimensión Modalidad Institución Educativa
Elaboración: Las autoras

Proceso de Carga de la Dimensión estado matrícula

i. Carga de registros en un archivo intermedio

Se extrajo los datos del archivo “TXT_ESTADO_MATRICULA”, de acuerdo al mapeo y se cargó a al archivo intermedio “STG_D_ESTADO_MATRICULA”.

ii. Carga de la dimensión

Se extrajo los valores del archivo “STG_D_ESTADO_MATRICULA” y se carga en la dimensión “D_ESTADO_MATRICULA”. En caso los registros hagan referencia a nuevas Instituciones Educativas se insertan nuevas filas en la tabla; caso contrario, si los registros hacen referencia a antiguas Instituciones Educativas se actualiza los campos.

iii. Borrar el archivo intermedio

Se borró el archivo intermedio “STG_D_ESTADO_MATRICULA”.

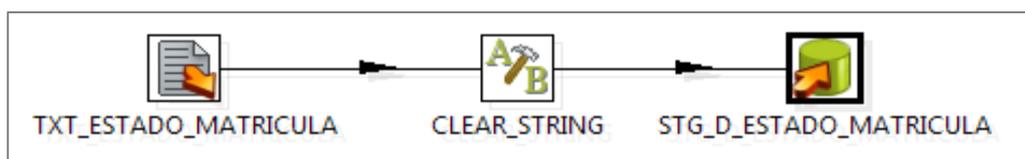


Figura 3.23 Carga de registros al archivo estado matrícula
Elaboración: Las autoras



Figura 3.24 Carga de dimensión modalidad educativa
Elaboración: Las autoras

Proceso de Carga de la Dimensión nivel educativo

i. Carga de registros en un archivo intermedio

Se extrajeron los datos del archivo “TXT_NIVEL_ EDUCATIVO”, “TXT_GRADO”, de acuerdo al mapeo y se cargó a al archivo intermedio “STG_D_NIVEL_EDUCATIVO”.

ii. Carga de la dimensión

Se extrajeron los valores del archivo “STG_D_NIVEL_EDUCATIVO”. y se carga en la dimensión “D_NIVEL_EDUCATIVO”. En caso los registros hagan referencia a nuevas Instituciones Educativas se insertan nuevas filas en la tabla; caso contrario, si los registros hacen referencia a antiguas Instituciones Educativas se actualiza los campos.

iii. Borrar el archivo intermedio

Se borró el archivo “STG_D_NIVEL_EDUCATIVO”.

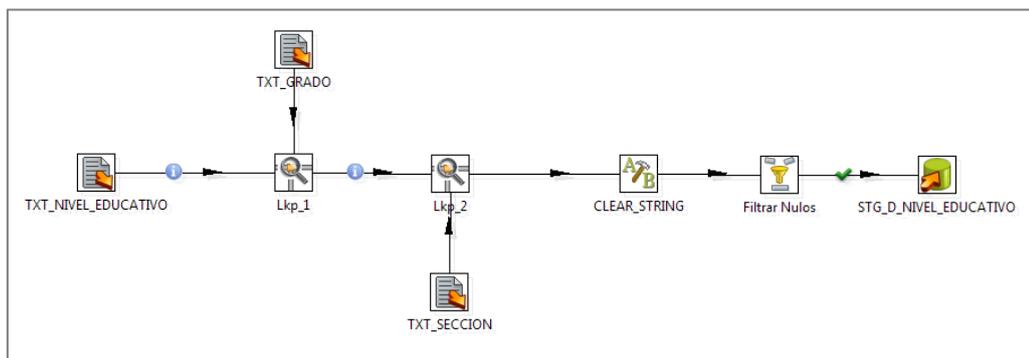


Figura 3.25 Carga de registros al archivo Nivel Educativo
Elaboración: Las autoras



Figura 3.26 Carga de Dimensión Nivel Educativo
Elaboración: Las autoras

Proceso de Carga de la Dimensión estado estudiante

i. Carga de registros en un archivo intermedio

Se extrajo los datos del archivo “TXT_ESTADO_POR_GRADO”, de acuerdo al mapeo y se cargó a al archivo intermedio “STG_D_ESTADO_ESTUDIANTE”.

ii. Carga de la dimensión

Se extrajo los valores del archivo “STG_D_ESTADO_ESTUDIANTE” y se carga en la dimensión “D_ESTADO_ESTUDIANTE”. En caso los registros hagan referencia a nuevas Instituciones Educativas se insertan nuevas filas en la tabla; caso contrario, si los registros hacen referencia a antiguas Instituciones Educativas se actualiza los campos.

iii. Borrar el archivo intermedio

Se borró el archivo “STG_D_ESTADO_ESTUDIANTE”.

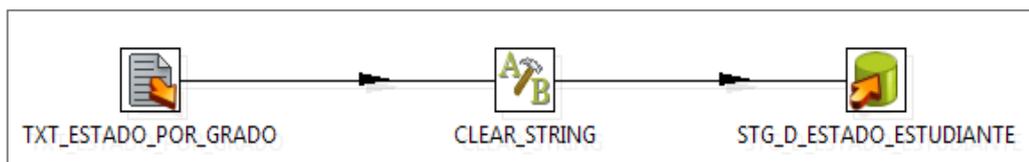


Figura 3.27 Carga de registros al archivo Estado Estudiante
Elaboración: Las autoras



Figura 3.28 Carga de dimensión Estado Estudiante
Elaboración: Las autoras

Proceso de Carga de la Dimensión tiempo

i. Creación de la Tabla “D_TIEMPO” en SQL

Se realiza el procedimiento para cargar la Tabla “D_TIEMPO” con todos los días del año a partir de Enero del 2014 en adelante.

ii. Carga de registros en un archivo intermedio

Se extrajo los datos de la tabla “D_TIEMPO” de acuerdo al mapeo y se cargó en un archivo de texto intermedio “tiempo”

iii. Carga de la dimensión

Se extrajo los valores del archivo “tiempo” y se cargó en la dimensión “D_TIEMPO” de acuerdo al mapeo. Este proceso se realiza mediante un truncate a la tabla “D_TIEMPO”; en el caso existan datos en la tabla entonces se borrarán para proceder a nueva carga.

iv. Borrar el archivo intermedio

Se borró el archivo intermedio “Tiempo”.

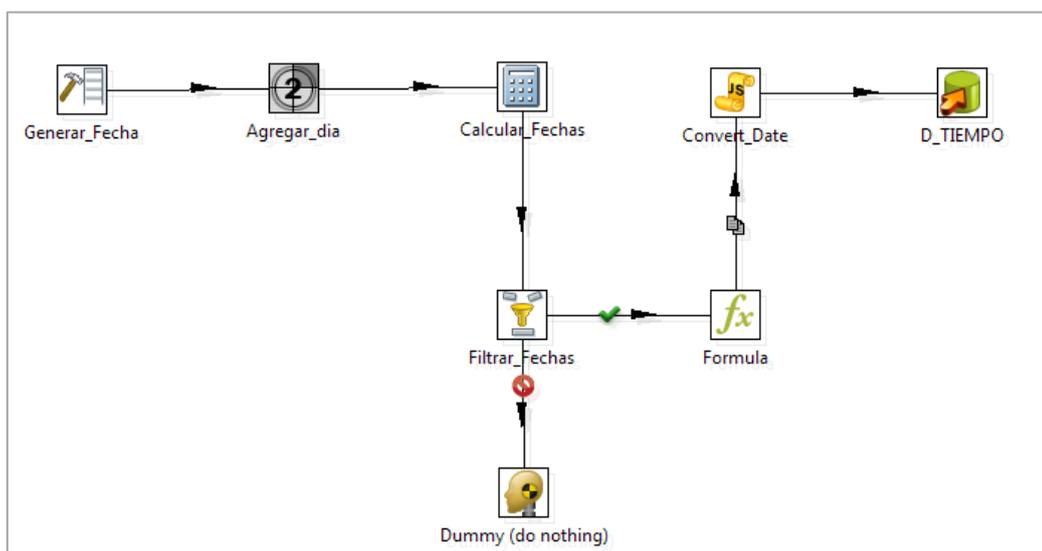


Figura 3.29 Carga de dimensión tiempo

Elaboración: Las autoras

Proceso de Carga de la Dimensión detalle de formato

i. Carga de registros en un archivo intermedio

Se extrajo los datos del archivo “TXT_FORMATOS_OFICIALES_SECCION” de acuerdo al mapeo y se cargó a al archivo intermedio “STG_D_DETALLE_FORMATO”.

ii. Carga de la dimensión

Se extrajo los valores del archivo “STG_D_DETALLE_FORMATO” y se carga en la dimensión “D_DETALLE_FORMATO”. En caso los registros hagan referencia a nuevos detalles de formato se insertan nuevas filas en la tabla; caso contrario, si los registros hacen referencia a antiguos detalles de formato se actualiza los campos.

iii. Borrar el archivo intermedio

Se borró el archivo intermedio “STG_D_DETALLE_FORMATO”.

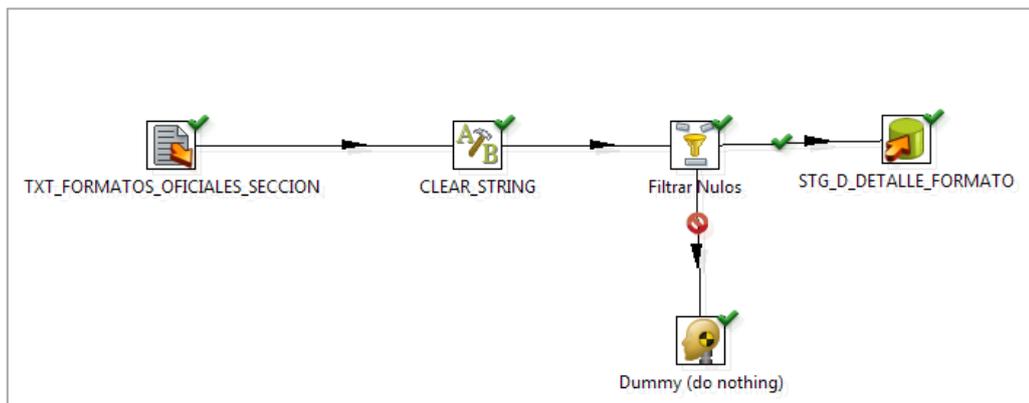


Figura 3.30 Carga de registros al archivo detalle de formato
Elaboración: Las autoras



Figura 3.31 Carga de dimensión detalle de formato
Elaboración: Las autoras

Proceso de Carga de la Dimensión Estado de formato

i. Carga de registros en un archivo intermedio

Se extrajo los datos del archivo “TXT_ ESTADO_FORMATO” de acuerdo al mapeo y se cargó a al archivo intermedio “STG_D_ESTADO_FORMATO”.

ii. Carga de la dimensión

Se extrajo los valores del archivo “STG_D_ ESTADO_FORMATO” y se carga en la dimensión “D_ESTADO_FORMATO”. En caso los registros hagan referencia a nuevos estados de formato se insertan nuevas filas en la tabla; caso contrario, si los registros hacen referencia a antiguos estados de formato se actualiza los campos.

iii. Borrar el archivo intermedio

Se borró el archivo intermedio “STG_D_ESTADO_FORMATO”.

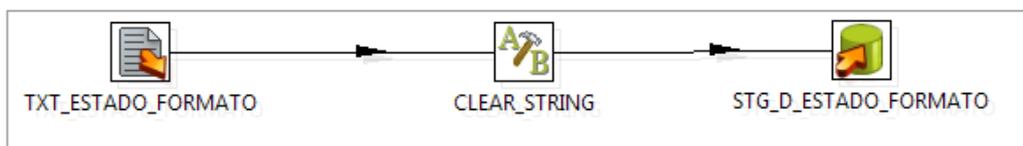


Figura 3.32 Carga de registros al archivo estado formato
Elaboración: Las autoras

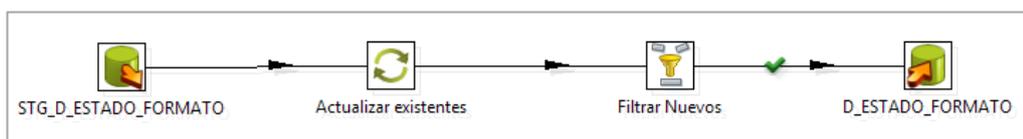


Figura 3.33 Carga de dimensión estado formato
Elaboración: Las autoras

Proceso de Carga de la Dimensión tipo de formato

i. Carga de registros en un archivo intermedio

Se extrajo los datos del archivo “TXT_ TIPO_FORMATO” de acuerdo al mapeo y se cargó a al archivo intermedio “STG_D_ TIPO_FORMATO”.

ii. Carga de la dimensión

Se extrajo los valores del archivo “STG_D_ TIPO_FORMATO” y se carga en la dimensión “D_ TIPO_FORMATO”. En caso los registros hagan referencia a nuevos tipos de formato se insertan nuevas filas

en la tabla; caso contrario, si los registros hacen referencia a antiguos tipos de formato se actualiza los campos.

iii. Borrar el archivo intermedio

Se borró el archivo intermedio “STG_D_TIPO_FORMATO”.

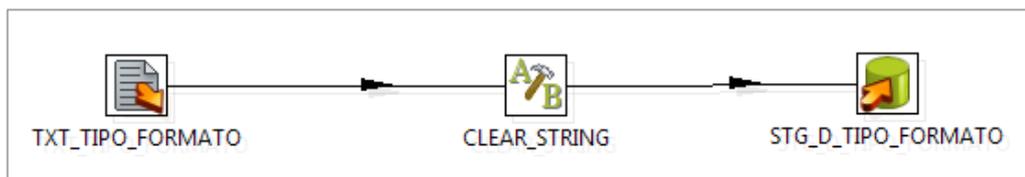


Figura 3.34 Carga de registros al archivo Tipos de Formato
Elaboración: Las autoras

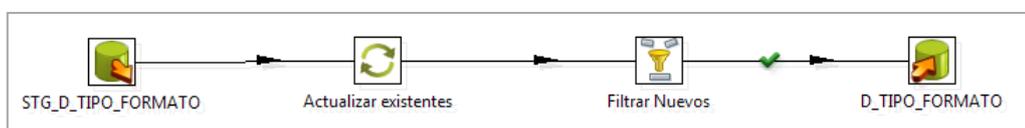


Figura 3.35 Carga de dimensión Tipos de Formato
Elaboración: Las autoras

3.5.1.2 Construcción de la Fact

Matrícula

En la Figura 3.37 se muestra la carga de todos los archivos intermedios en la capa Staging (STG).

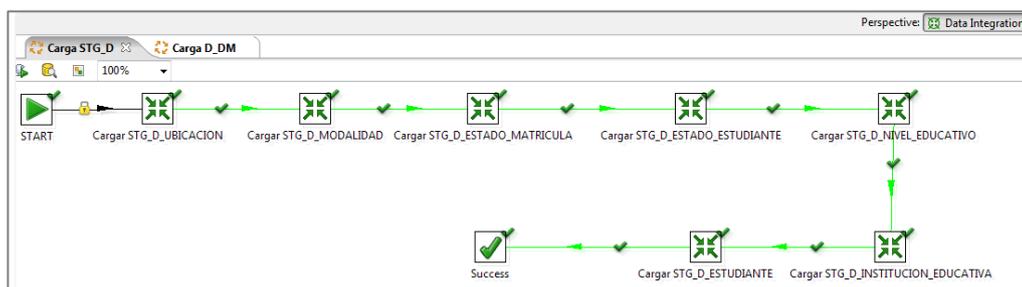


Figura 3.36 Carga de los archivos intermedio STG
Elaboración: Las autoras

Formato

En la Figura 3.36, se muestra la carga de todos los archivos intermedios en la capa Staging (STG).

3.5.2 Desarrollo de aplicación

La explotación de la información se realizará a través de la herramienta Tableau. En esta sección se mostrará reportes de estado de matrícula, nóminas, cantidad de secciones que se elaborarán como parte de este proyecto.

Para poder realizar los reportes se en Tableau se realizó la carga del Fact Matrícula (Ver Figura 3.32) y Fact Formato (Ver Figura 3.33).

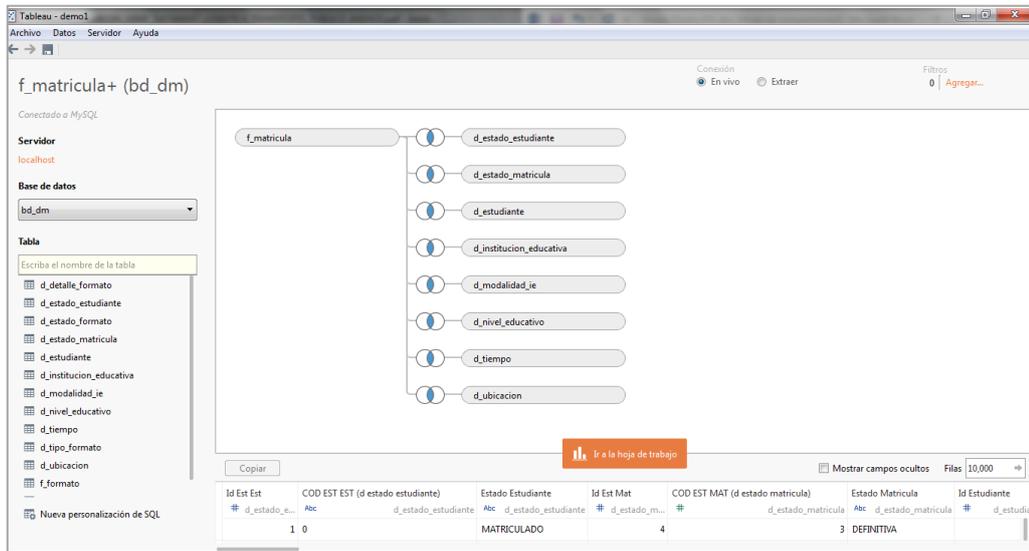


Figura 3.37 Fact Matrícula en Tableau
Elaboración: Las autoras

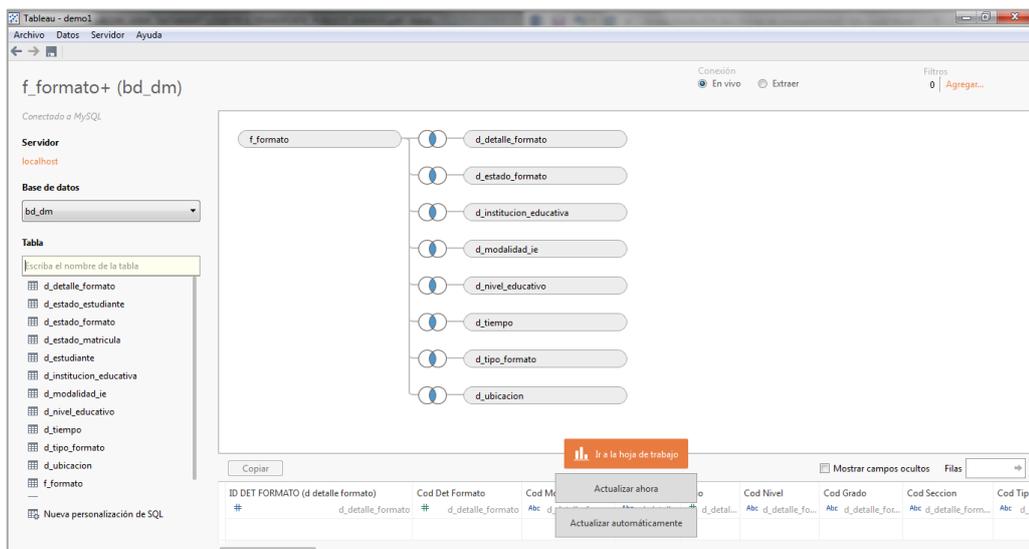


Figura 3.38 Fact formato en Tableau
Elaboración: Las autoras

3.5.2.1 Configuración del reporte

- Tamaño de Hoja: A4
- Orientación: Vertical u Horizontal (Depende del tipo de reporte)
- Márgenes (cm): izquierdo 12.7cm, derecho 12.7cm, superior 12.7cm, inferior 12.7cm.
- Tamaño de Dashboard: Estandar (1000 x 800)
- Título Dashboard: Letra Trebuchet MS –Tamaño 12- Color Verde Olivo- Negrita.
- Título Reporte: Letra Trebuchet MS –Tamaño 10- Color Verde - Negrita.

A continuación, se presentan los Dashboards generados.

Dashboard matrícula

Contiene los siguientes reportes:

- Reporte de Estado de Matrícula en Instituciones Educativas
- Reporte de Estado de Matrícula por el Nivel Educativo
- Reporte de la Modalidad de Matriculados por Grado

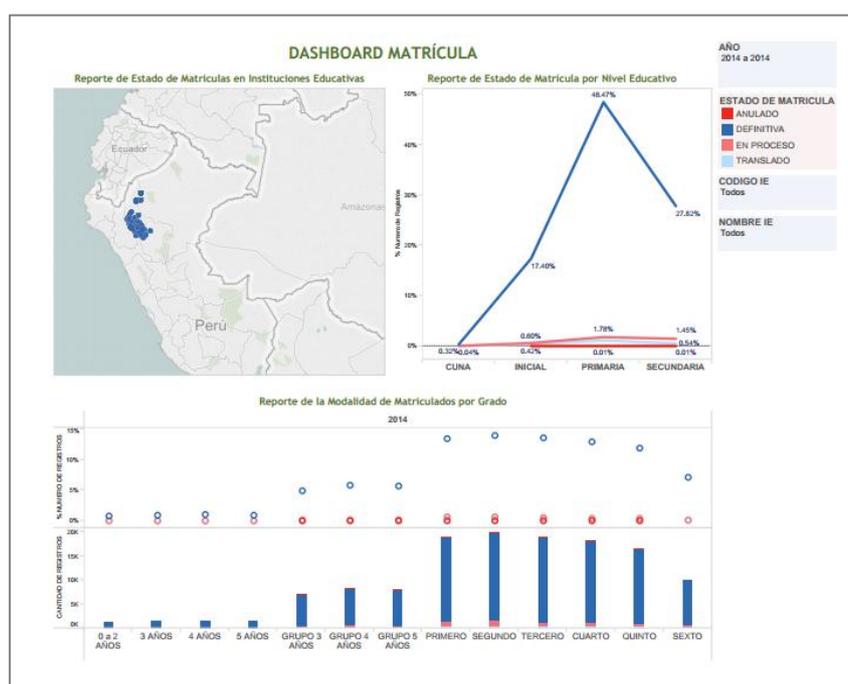


Figura 3.39 Dashboard Matrícula
Elaboración: Las autoras

Dashboard estudiante

Contiene los siguientes reportes:

- Reporte Cantidad de Estudiantes
- Reporte Estado Estudiante
- Reporte Estudiantes por Nivel Educativo
- Reporte Estudiantes por Grado

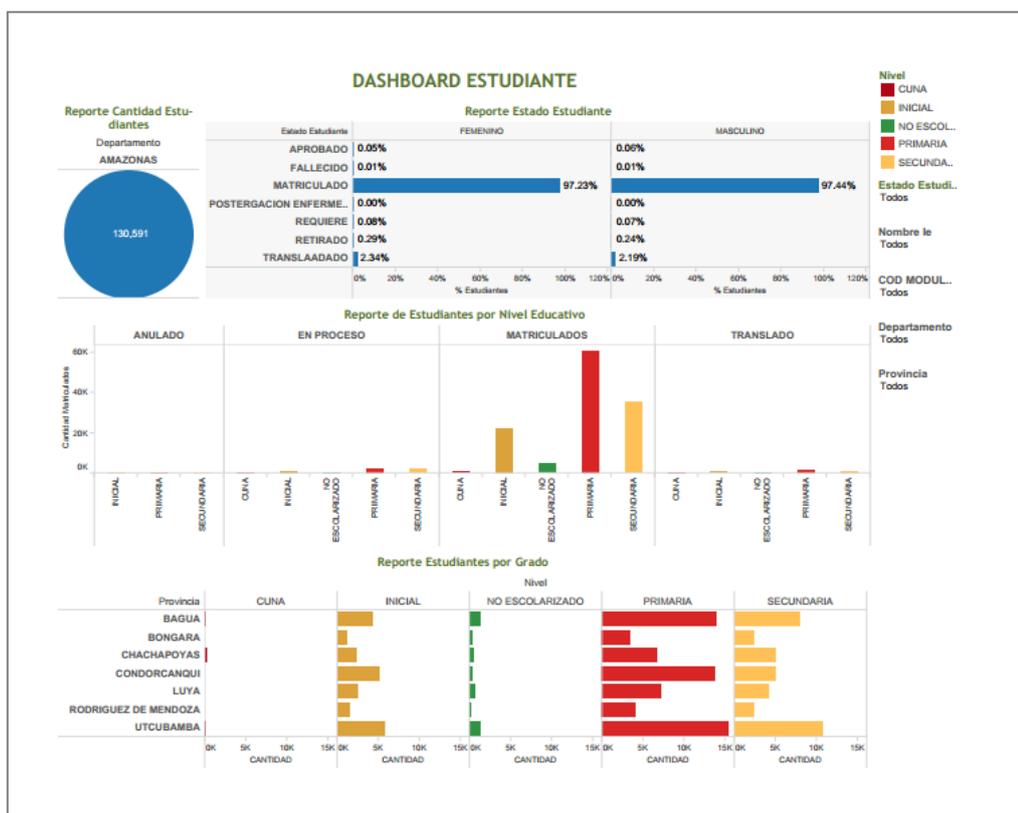


Figura 3.40 Dashboard estudiante
Elaboración: Las autoras

Dashboard nómina

Contiene los siguientes reportes:

- Reporte Cantidad de Estudiantes por Nómina.
- Reporte Estado de Formato tipo Nómina Mensual.
- Reporte Estado de Formato tipo Nómina por Provincia.
- Reporte de Estado de Nóminas tipo Nómina por Trimestre.

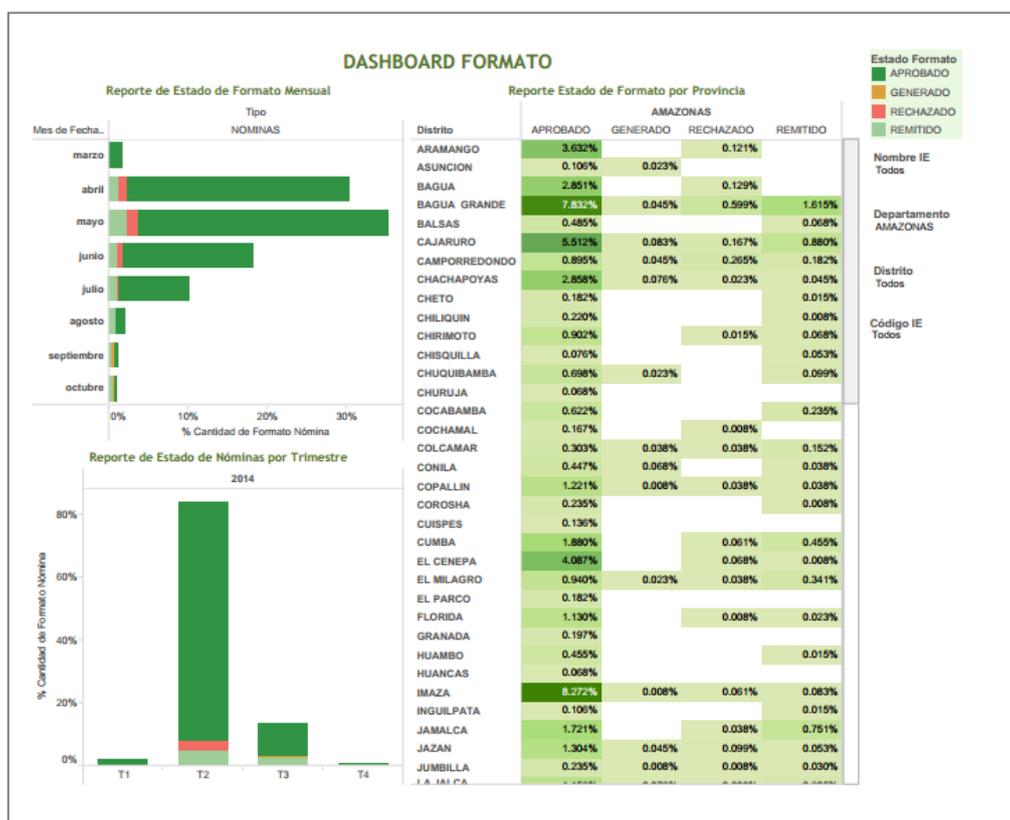


Figura 3.41 Dashboard nómina
Elaboración: Las autoras

3.5.3 Certificación

El desarrollo de la Certificación está explicado en el Capítulo IV Pruebas y Resultados.

CAPÍTULO IV

PRUEBAS Y RESULTADOS

4.1 Pruebas

La siguiente sección muestra los tipos de pruebas y su contenido respectivo de manera que se pueda comprobar que los datos ingresados se vean reflejados, correctamente, en los informes elaborados. La realización de las pruebas se definió en una matriz de casos de prueba (Ver Anexo 06 Estimación Preliminar de Pruebas) y el tiempo estimado para la certificación se encuentra detallado en el cronograma del proyecto (Ver Anexo 01 Cronograma del Proyecto). Las pruebas que se realizaron son Pruebas de Funcionalidad y Pruebas de Integración, adicional a ello se realizó una prueba solicitada por el usuario final llamada Cuadre de Carga, consiste en evaluar la carga a un nivel técnico mediante consultas comparativas. Las pruebas las detallamos a continuación.

4.1.1 Cuadre de carga

El cuadro de carga ayuda a verificar que la información llevada desde la fuente hacia el destino no ha sufrido ninguna alteración adicional a las ya esperadas luego a los cambios del ETL

Entre los principales casos se tienen validaciones de tipo:

- Cantidad de registros procesados.
- Cantidad de registros procesados bajo ciertos criterios (filtros).
- Operaciones aritméticas para la consolidación de montos. (Suma, promedio mínimo y máximo valor).

Este tipo de cuadros se realiza de manera interna, la conformidad es dada por el usuario final (administrador de repositorio de datos).

4.1.2 Pruebas de funcionalidad

Las pruebas de funcionalidad del interface del Datamart son pruebas específicas y concretas que se realizan, donde si los resultados son los esperados se da conformidad y se realiza el pase a producción. La ejecución de pruebas se realizó de forma manual por el Analista de Pruebas en conjunto con el usuario final.

En estas pruebas, se realiza la verificación del buen funcionamiento de los filtros, la realización de las consultas y los resultados reflejados en los reportes, las cuales se brinda a los usuarios finales quienes hacen el análisis respectivo para poder dar su aprobación.

4.1.3 Pruebas de integridad

Estas pruebas consisten en verificar que el Datamart y el SIAGIE (Sistema de OFIN- MINEDU) puedan trabajar en conjunto sin problemas. Esta prueba es importante ya que el Datamart recepcionará la data a través del archivo plano (TXT) que se obtiene del SIAGIE sin problemas o inconvenientes en su proceso.

4.2 Resultados

A continuación, se muestran los resultados obtenidos, luego de ser realizadas las pruebas a la aplicación, que se citaron anteriormente.

4.2.1 Resultados de Cuadre de carga

Conteo de registros

El conteo de registros ayuda a identificar si la carga fue exitosa cuantitativamente, es recomendable identificar que filtros sirven para realizar un mejor análisis dependiendo de las restricciones seguidas en los componentes ETL. Ver Tabla 4.1, Tabla 4.2

Tabla 4.1 Cuadre de conteo de registros Fact matrícula

N	Tabla Origen	N° Regis - Tros	Tabla Destino	N° Registros	Conforme SI/NO
1	STG_D_ESTADO_ESTUDIANTE	9	D_ESTADO_ESTUDIANTE	9	SI
2	STG_D_ESTADO_MATRICULA	6	D_ESTADO_MATRICULA	6	SI
3	STG_D_ESTUDIANTE	130591	D_ESTUDIANTE	130591	SI
4	STG_D_MODALIDAD_IE	1	D_MODALIDAD_IE	1	SI
5	STG_D_UBICACION	85	D_UBICACION	85	SI
6	STG_D_INSTITUCION_EDUCATIVA	3277	D_INSTITUCION_EDUCATIVA	3277	SI
7	STG_D_NIVEL_EDUCATIVO	13384	D_NIVEL_EDUCATIVO	13384	SI

Elaboración: Las autoras

Tabla 4.2 Cuadre de conteo de registros Fact formato

N	Tabla Origen	N° Regis - Tros	Tabla Destino	N° Registros	Conforme SI/NO
1	STG_D_DETALLE_FORMATO	14347	D_DETALLE_FORMATO	14347	SI
2	STG_D_ESTADO_FORMATO	5	D_ESTADO_FORMATO	5	SI
3	STG_D_TIPO_FORMATO	4	D_TIPO_FORMATO	4	SI
4	STG_D_MODALIDAD_IE	1	D_MODALIDAD_IE	1	SI
5	STG_D_UBICACION	85	D_UBICACION	85	SI
6	STG_D_INSTITUCION_EDUCATIVA	3277	D_INSTITUCION_EDUCATIVA	3277	SI
7	STG_D_NIVEL_EDUCATIVO	13384	D_NIVEL_EDUCATIVO	13384	SI

Elaboración: Las autoras

Para ver los detalles de esta prueba ver Anexo 07 Registros de Cuadre de Conteo.

En la Figura 4.1, podemos ver la consulta de la Tabla Destino D_UBICACION.



ID_UBICACION	UBIGEO	DEPARTAMENTO	PROVINCIA	DISTRITO	LONGITUD	LATITUD
1	010101	AMAZONAS	CHACHAPOYAS	CHACHAPOYAS	-77.818380	-6.256220
2	010102	AMAZONAS	CHACHAPOYAS	ASUNCION	-77.710510	-6.032190
3	010103	AMAZONAS	CHACHAPOYAS	BALSAS	-77.919130	-6.787170
4	010104	AMAZONAS	CHACHAPOYAS	CHETO	-77.694420	-6.289510
5	010105	AMAZONAS	CHACHAPOYAS	CHILIQVIN	-77.724080	-6.114030

Figura 4.1 Consulta TOP5 D_UBICACION
Elaboración: Las autoras

En el Anexo 08 Consulta de Dimensiones, se pueden ver las evidencias de las pruebas realizadas para verificar que la información no se vio afectaba a través del proceso de Carga.

4.2.2 Resultados de Pruebas de Funcionales

Los resultados de las Pruebas Funcionales realizadas se agruparon de acuerdo al tipo de caso de prueba, los cuales se muestran en la Tabla 4.3.

Tabla 4.3 Resumen de resultados prueba funcionalidad

Tipo Caso Prueba	Fecha de Ejecución	Resultado Esperado	Resultado Real
Carga de Datamart	27/10/2014	La data se cargó satisfactoriamente	Sí cumple
Generar el Reporte de estado matrícula	27/10/2014	El reporte se generó satisfactoriamente.	Sí cumple
Generar Reporte de Nóminas	27/10/2014	El reporte se generó satisfactoriamente.	Sí cumple
Generar Reporte de estado de estudiante	27/10/2014	El reporte se generó satisfactoriamente.	Sí cumple
Generar Reporte de estudiantes por nivel educativo.	27/10/2014	El reporte se generó satisfactoriamente.	Sí cumple
Generar Reporte de número de estudiantes	27/10/2014	El reporte se generó satisfactoriamente.	Sí cumple
Generar Reporte de participación de estudiantes por género.	27/10/2014	El reporte se generó satisfactoriamente.	Sí cumple
Generar Reporte de estado por grado.	27/10/2014	Los filtros funcionan correctamente	Sí cumple

Elaboración: Las autoras

Para poder corroborar que las cantidades que se muestran en los Dashboard, hemos seleccionado el Dashboard Estudiante para poder realizar la comparación de las cantidades reflejadas en los reportes: Reporte Cantidad Estudiantes y Reporte de Estudiantes por Nivel Educativo. Ver Figura 4.2.

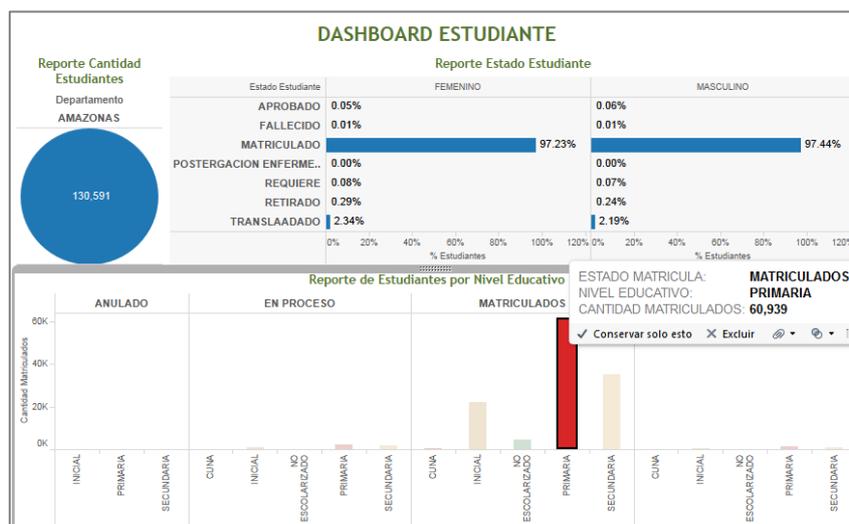


Figura 4.2 Pruebas: Dashboard Estudiante
Elaboración: Las autoras

En la Figura 4.3, podemos ver que la cantidad de estudiantes en el Departamento de Amazonas es igual a 130,591 estudiantes.



Figura 4.3 Reporte cantidad estudiantes
Elaboración: Las autoras

En la Figura 4.4, podemos ver que la cantidad de estudiantes matriculados en el nivel Primaria del Departamento de Amazonas es igual a 60,939 estudiantes.

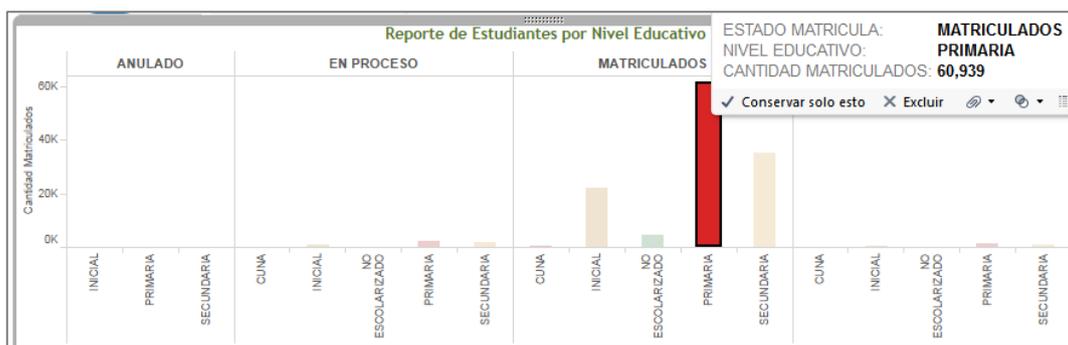


Figura 4.4 Reporte de estudiantes por Nivel Educativo del departamento de Amazonas
Elaboración: Las autoras

Los resultados obtenidos están de acuerdo con el rango de tiempo aceptable, es por ello que se puede confirmar que las pruebas de funcionalidad del proyecto fueron exitosas.

4.2.3 Resultados de Pruebas de Integridad

Se realizaron las Pruebas de Integridad para la verificación de la integridad del sistema SIAGIE con el Datamart propuesto, obteniendo como resultados lo siguiente:

- La carga de data origen SIAGIE con destino Datamart se realizó de forma satisfactoria.
- Los campos de las tablas de la BD del SIAGIE y del Datamart son compatibles. Estos incluyen los campos de tipo Fecha, Texto, y Numéricos.

El proceso de carga Fact Matrícula se puede apreciar en las siguientes Figuras.

Execution Results											
Execution History Logging Step Metrics Performance Graph Metrics Preview data											
#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time
32	Agregar Campos	0	130591	130591	0	0	0	0	0	Finished	12.6s
29	Convert Fecha	0	130591	130591	0	0	0	0	0	Finished	9.7s
6	D_ESTADO_ESTUDIANTE	0	0	9	9	0	0	0	0	Finished	0.1s
7	D_ESTADO_MATRICULA	0	0	6	6	0	0	0	0	Finished	0.1s
4	D_ESTUDIANTE	0	0	130591	130591	0	0	0	0	Finished	1.2s
3	D_INSTITUCION_EDUCATIVA	0	0	3277	3277	0	0	0	0	Finished	0.1s
10	D_MODALIDAD_JE	0	0	4	4	0	0	0	0	Finished	0.1s
2	D_NIVEL_EDUCATIVO	0	0	13384	13384	0	0	0	0	Finished	0.1s
1	D_TIEMPO	0	0	10001	10001	0	0	0	0	Finished	0.1s
5	D_UBICACION	0	0	85	85	0	0	0	0	Finished	0.1s
37	F_MATRICULA	0	130591	130591	0	130591	0	0	0	Finished	40.8s
17	F_MATRICULA Existentes	0	0	0	0	0	0	0	0	Finished	0.1s
36	Filtrar Nuevos	0	130591	130591	0	0	0	0	0	Finished	39.5s
14	Merge Join	0	143975	130591	0	0	0	0	0	Finished	2.0s
16	Merge Join 2	0	130597	130591	0	0	0	0	0	Finished	4.0s
22	Merge Join 3	0	130600	130591	0	0	0	0	0	Finished	6.0s
24	Merge Join 4	0	133868	130591	0	0	0	0	0	Finished	6.6s
26	Merge Join 5	0	130595	130591	0	0	0	0	0	Finished	8.9s
28	Merge Join 6	0	130676	130591	0	0	0	0	0	Finished	9.7s
31	Merge Join 7	0	140592	130591	0	0	0	0	0	Finished	12.6s
35	Merge Join 8	0	130591	130591	0	0	0	0	0	Finished	38.0s

Figura 4.5 Resultado de carga de Fact matrícula
Elaboración: Las autoras

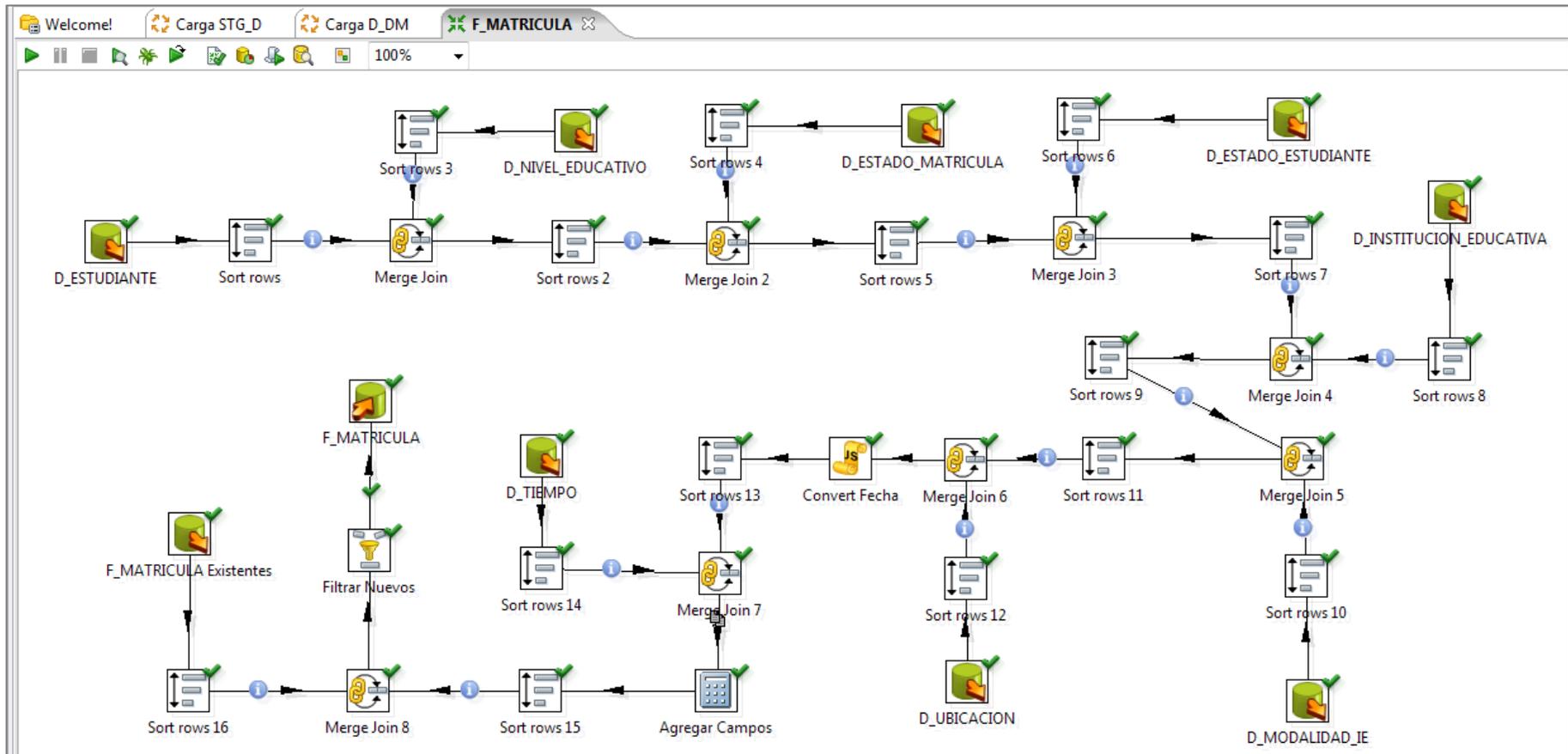


Figura 4.6 ETL Fact matrícula
Elaboración: Las autoras

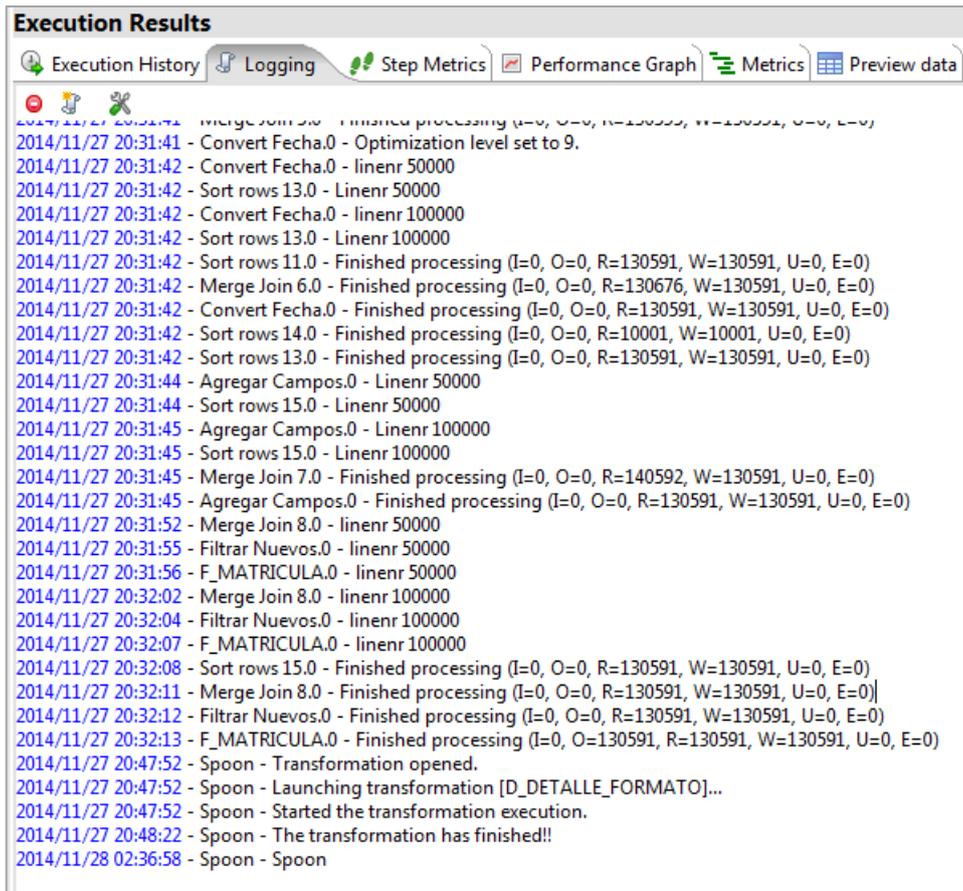


Figura 4.7 Log de carga de Fact matrícula
Elaboración: las autoras

El proceso de carga Fact Formato se puede apreciar en las siguientes Figuras.

Execution Results

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time
1	D_ESTADO_FORMATO	0	0	5	5	0	0	0	0	Finished	0.0s
2	D_INSTITUCION_EDUCATIVA	0	0	3277	3277	0	0	0	0	Finished	0.0s
3	D_NIVEL_EDUCATIVO	0	0	13384	13384	0	0	0	0	Finished	0.1s
4	D_TIPO_FORMATO	0	0	1	1	0	0	0	0	Finished	0.0s
5	D_TIEMPO	0	0	30003	10001	0	0	0	0	Finished	0.0s
6	D_UBICACION	0	0	85	85	0	0	0	0	Finished	0.0s
7	D_DETALLE_FORMATO	0	0	0	0	0	0	0	0	Finished	0.0s
8	Sort rows	0	0	0	0	0	0	0	0	Finished	0.0s
9	F_FORMATO Existentes	0	0	0	0	0	0	0	0	Finished	0.0s
10	Sort rows 4	0	5	5	0	0	0	0	0	Finished	0.0s
11	Sort rows 3	0	13384	13384	0	0	0	0	0	Finished	0.2s
12	D_MODALIDAD_IE	0	0	4	4	0	0	0	0	Finished	0.0s
13	Sort rows 8	0	3277	3277	0	0	0	0	0	Finished	0.0s
14	Merge Join	0	13384	0	0	0	0	0	0	Finished	0.4s
15	Sort rows 6	0	1	1	0	0	0	0	0	Finished	0.0s
16	Sort rows 2	0	0	0	0	0	0	0	0	Finished	0.4s
17	Merge Join 2	0	5	0	0	0	0	0	0	Finished	0.7s
18	Sort rows 5	0	0	0	0	0	0	0	0	Finished	0.7s
19	Merge Join 3	0	1	0	0	0	0	0	0	Finished	1.0s
20	Sort rows 7	0	0	0	0	0	0	0	0	Finished	1.1s
21	Sort rows 14	0	10001	10001	0	0	0	0	0	Finished	2.3s
22	Sort rows 17	0	10001	10001	0	0	0	0	0	Finished	2.6s
23	Sort rows 12	0	85	85	0	0	0	0	0	Finished	0.0s
24	Sort rows 10	0	4	4	0	0	0	0	0	Finished	0.0s

Figura 4.8 Resultado de carga de Fact formato
Elaboración: Las autoras

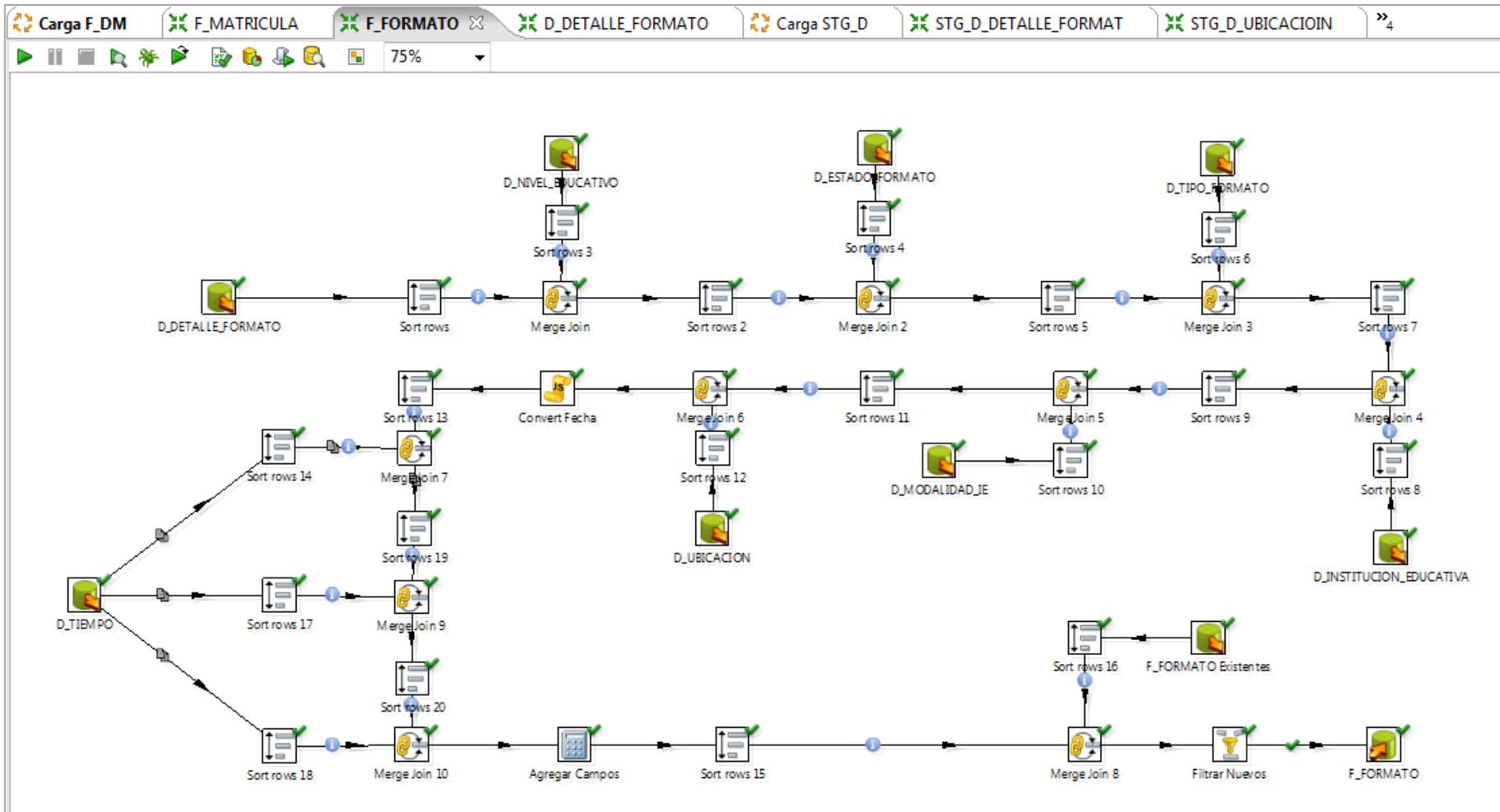


Figura 4.9 ETL Fact formato
Elaboración: Las autoras

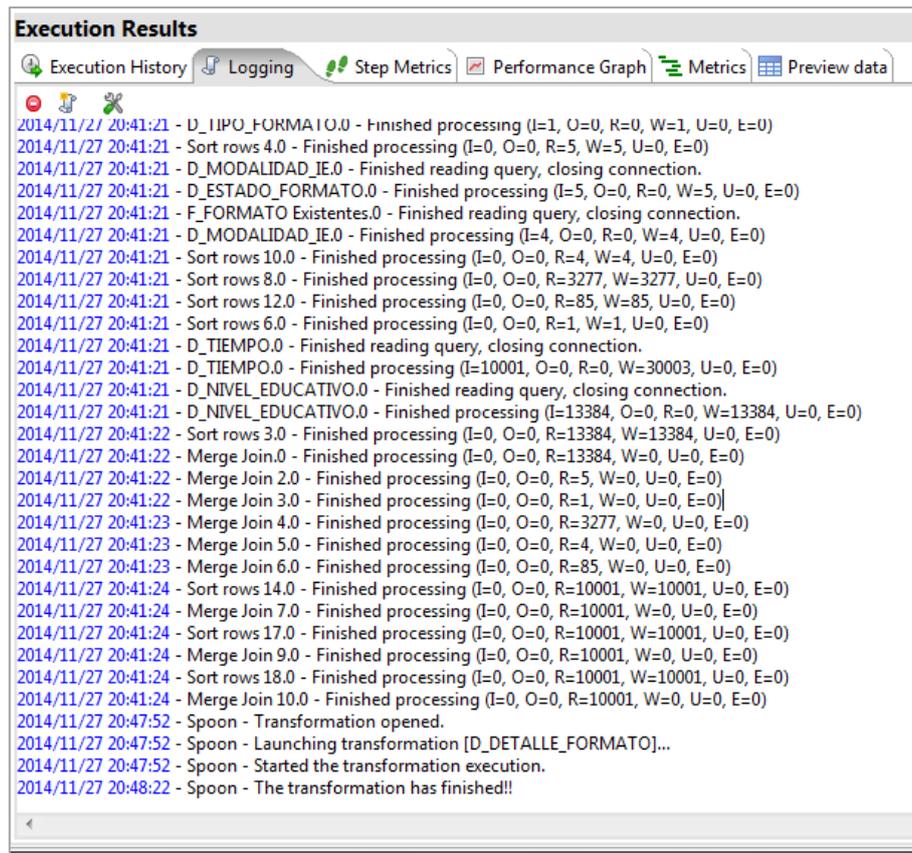


Figura 4.10 Log de carga de Fact formato
Elaboración: Las autoras

Por ello se puede concluir que las Pruebas de Integridad culminaron con resultados satisfactorios.

Para visualizar relación de causalidad con el propósito del proyecto, ver Anexo 03 Matriz de Consistencia.

CAPÍTULO V

DISCUSIÓN Y APLICACIONES

5.1 Discusión

Teniendo el marco referencial del uso del Datamart, después de las pruebas realizadas con éxito en la Oficina de Informática del MINEDU, se pudo realizar comparaciones del tiempo ahorrado en cada reporte generado, contrastándolo con el tiempo dedicado anteriormente para los reportes realizados de manera manual. Para el análisis se tomó como base entrevistas que se hicieron a las personas encargadas de la Oficina de Informática (OFIN) de MINEDU y al representante de la Secretaria de Planificación Estratégica, junto con las respuestas dadas a través de la Encuesta de Satisfacción, ver Anexo 11 Encuesta de Satisfacción.

Para la captura de requerimientos y necesidades se solicitó la planificación de entrevistas así se obtuvieron datos reales para iniciar el análisis y diseño del Datamart que permita facilitar la generación de reportes que los usuarios finales requieren.

El beneficio obtenido cubrirán los objetivos específicos trazados en el inicio de la Tesis, se puede ver en la Tabla 5.1.

Tabla 5.1 Objetivos versus resultados

Objetivos del Proyecto (General y Específicos)	Antes de la Implementación del Datamart	Después de la Implementación del Datamart	Satisfacción del Cliente (Encuesta)
Mejorar el apoyo a la toma de decisiones en los procesos de matrículas y nominas en la Secretaría de Planificación estratégica mediante la implementación del Datamart.	La Secretaría de Planificación Estratégica del MINEDU estaba afectada por la falta de una solución tecnológica que apoye la toma de decisiones que permitan dar seguimiento a las Instituciones Educativas.	La Secretaría de Planificación Estratégica del MINEDU cuenta con un mejor apoyo para la toma de decisiones que permitan dar seguimiento a las Instituciones Educativas de educación básica regular.	Muy Satisfecho
Implementar un Datamart para la toma de decisiones de la Secretaría de Planificación Estratégica del MINEDU.	Cuenta con un proceso manual para realizar las consultas, reportes y métricas. La información para la toma de decisiones de la Secretaría de Planificación Estratégica del MINEDU no es oportuna.	Ausencia de la generación manual de reportes. Obtiene la información oportuna Reportes personalizados e intuitivos que apoyan a la toma decisiones para la Secretaría de Planificación Estratégica.	Satisfecho
Identificar y definir los indicadores de los procesos de matrícula y nómina.	Los indicadores con los que cuentan para realizar los reportes no se dan abasto para mostrar la información que desean consultar.	Se cuenta con una matriz de indicadores definidos y establecidos para cada reporte de interés para la Secretaría de Planificación Estratégica.	Muy Satisfecho
Construir los procesos ETL para la Extracción, Transformación y Carga de datos correspondientes al proceso de matrículas y de nóminas desde el SIAGIE	El proceso ETL es realizado manualmente.	El proceso ETL es realizado de forma automatizada.	Satisfecho

Elaboración: Las autoras

Para evidenciar con exactitud el éxito de la solución ofrecida fue necesario medir el tiempo que se utiliza para la extracción de la información del SIAGIE, la transformación y la carga de cada reporte, dando uso del Datamart y haciendo una comparación de tiempo invertido antes de ello. (Ver Tabla 5.2)

Tabla 5.2 Evaluación de tiempo proceso manual

N°	Reporte	Tiempo de Realización -Mínimo	Tiempo de Realización-Máximo
1	Reporte de estado matrícula	1.5 hrs.	3 hrs.
2	Reporte de Nómina	1.5 hrs.	3 hrs.
3	Reporte de estado del estudiante	1.5 hrs.	3 hrs.
4	Reporte de estudiantes por nivel educativo	1.5 hrs.	3 hrs.
5	Reporte de Participación de estudiantes por género	1.5 hrs.	3 hrs.
6	Reporte de estado por Grado	1.5 hrs.	3 hrs.
TOTAL		9 hrs.	18 hrs.

Elaboración: Las autoras

El tiempo de realización de reporte antes de la implementación es el que abarca todo el proceso manual de generación de reportes se realiza desde la extracción de la información contenida en la BD de SIAGIE, para ello el especialista de BD realiza los scripts y los ejecuta manualmente, luego se realizan tablas en Excel, ver Figura 5.1 y 5.2 para luego utilizarlas como base para la generación de gráficos. Todo ello tiene un tiempo promedio ubicado en el rango de 1.5hrs -3hrs, lo cual multiplicado por seis reportes se obtiene el total de tiempo invertido promedio de 18 horas por la generación de los reportes.

NOMBREIE	COD_NIVEL	DSC_NIVEL	DSCMODALIDAD	CODTIPOIE	DESCTIPOIE
11005	B0	Primaria	Educación Básica Regular	A1	Pública - Sector Educación
LOS CLAVELES	A5	Primaria	Educación Básica Regular	A1	Pública - Sector Educación
050	A2	Secundaria	Educación Básica Regular	A1	Pública - Sector Educación

Figura 5.1 Muestra del Excel nómina de matrícula

Elaboración: Las autoras

NOMBREIE	COD_NIVEL	NRO ESTUDIANTE	COD ESTUDIANTE	SEXO	AP. PATERNO	APE. MATERNO	NOMBRES	FOTO	PAIS NACIM.	FECHA NACIM.	LUGAR NACIM.
11005	B0	12550	289701	F	Laos	Morán	Maria	img01.gif	Perú	01/12/2010	Bagua
11005	B0	14790	289702	M	Lopez	Insil	Jorge	img02.gif	Perú	14/06/2010	Bagua
11005	B0	18550	289703	F	Mendez	Luna	Ely	img03.gif	Perú	03/10/2011	Bagua

Figura 5.2 Muestra del Excel de matrícula

Elaboración: Las autoras

En el Datamart se alojará la información que se extraerá a través de la comunicación con la BD del SIAGIE la cual pasará por un proceso ETL (Extracción, Transformación y Carga), el rango de tiempo promedio de su generación es de 0 a 5 minutos como lo podemos observar en la siguiente Tabla 5.3

Tabla 5.3 Evaluación de tiempo usando Datamart

N°	Reporte	Tiempo de Realización -Mínimo	Tiempo de Realización- Máximo
1	Reporte de estado matrícula	<1min	5 min.
2	Reporte de Nómina	<1min	5 min.
3	Reporte de estado del estudiante	<1min	5 min.
4	Reporte de estudiantes por nivel educativo	<1min	5 min.
5	Reporte de Participación de estudiantes por género	<1min	5 min.
6	Reporte de estado por Grado	<1min	5 min.
TOTAL		6min	30 min

Elaboración: Las autoras

Cálculo de Porcentaje de Reducción de Tiempo:

$$\% \text{ Reducción de Tiempo} = 100 - \frac{\text{Tiempo de Proceso actual} \times 100\%}{\text{Tiempo de Proceso anterior}}$$

$$\% \text{ Reducción Tiempo} = 100 - (30 \times 100) / 1080$$

$$\% \text{ Reducción de Tiempo} = 97.22\%$$

El tiempo invertido ha dado como resultado una reducción de 1050 minutos equivalente a 97.22 % (Intervalo mayor), estos tiempos son de

importancia en la toma de decisiones del Ministerio de Educación para realizar las diversas gestiones que impactan al sector Educativo.

5.2 Aplicaciones

El presente proyecto tuvo como alcance realizar un análisis, diseño e implementación de Datamart para la generación de reportes de los Procesos de Matrícula y Nómina para la Secretaría de Planificación Estratégica del Ministerio de Educación, quienes consultan los reportes con el fin de apoyarse para la toma de decisiones estratégicas. El alcance de este tipo de soluciones se puede aplicar en diversas áreas de la organización. El uso de la información que se obtiene desde el Datamart es de gran ayuda en distintos aspectos, uno de ellos es ser difundida para diversos proyectos y planes educacionales en donde se busque mejorar la calidad educativa, un ejemplo de ello es optando por realizar un seguimiento constante en aquellas Instituciones Educativas de un determinado Departamento del Perú que se encuentran con un índice alto de estudiantes que no continúan con la educación básica secundaria, ya que solo se quedan en el nivel de primaria, esto con el fin de realizar un plan para disminuir este índice negativo. El comparar resultados a través de los reportes puede ayudar también el brindan apoyo a los escolares hallados en las zonas donde abunda la pobreza y pobreza extrema para ayudarlos a seguir educándose.

La solución se desarrolló bajo una arquitectura deseada, escalable y adaptable lo cual permite la incorporación de datos de otros procesos de interés en el sector educativo, ya que en esta primera fase sólo se ha realizado la generación de reportes provenientes a los procesos de Matrícula y Nómina con la carga de data de la Dirección Regional de Educación de Amazonas.

CONCLUSIONES

1. Es indispensable para la Secretaría de Planificación Estratégica del MINEDU, contar con una herramienta tecnológica que apoye a la toma de decisiones para el seguimiento en las Instituciones Educativas de educación básica regular en los Procesos de Matrícula y Nóminas.
2. El Datamart constituirá una herramienta válida para el análisis cuantitativo y cualitativo de los procesos en mención.
3. La identificación de los principales indicadores son de gran importancia ayuda obtener una visión general y una visión detallada del estado de las Instituciones Educativas a nivel Departamental.
4. La selección de herramientas para la extracción y explotación de datos mejoró el proceso de obtención de la información, evitando colapsar el sistema transaccional, SIAGIE.
5. El Datamart permite el ahorro de tiempo promedio de 97%, anteriormente implicaba sobrecarga de trabajo para la Oficina de Informática del MINEDU.

6. La definición de reportes debe contar con la participación de todos los usuarios involucrados y debe ser difundido por un medio formal hacia todos los usuarios finales.
7. La construcción de los procesos ETL permite realizar consultas a partir de información previamente procesada con lo cual se da flexibilidad al usuario a realizar diferentes consultas pre elaboradas. Es importante resaltar la relación que existe entre el desempeño de la consulta y la flexibilidad del análisis.
8. La capa de datos BDS perteneciente a la arquitectura del presente proyecto aísla el entorno operacional para solo enfocarse en información consolidada que ayude a la toma de decisiones.
9. La capa de datos ODS perteneciente a la arquitectura del presente proyecto integró de manera exitosa información proveniente de un sistema transaccional SIAGIE en un repositorio que cuenta con un modelo de datos del negocio que soporta consultas de la información a un nivel transaccional.

RECOMENDACIONES

1. En futuras implementaciones de soluciones de inteligencia de negocios para otras áreas de la organización se considera conveniente realizar un exhaustivo análisis de las necesidades del negocio, ya que un erróneo entendimiento de estas podría terminar en el fracaso del proyecto.
2. Se recomienda que en una segunda fase se incorporen nuevos procesos, para la generación de nuevos reportes de gestión. En el presente proyecto las bases están construidas para seguir extendiendo el número de indicadores en el Datamart.
3. Seguir respetando la arquitectura iniciada con lo cual se garantiza la integración y reutilización de la información. Esto permitirá seguir centralizando toda la información de la empresa y las reglas de negocio de modo que la información se pueda distribuir en el mismo contexto a todas las áreas de la empresa.
4. Actualmente hay ausencia de una Plataforma Tecnológica de Seguridad de la Información, se recomienda para el despliegue del proyecto se considere los aspectos de la Seguridad de Información propuesto.

FUENTES DE INFORMACIÓN

Bibliográficas:

1. **Bayona, S., Calvo-Manzano, J., Cuevas, G. & San Feliu, T. (2012)** Method for selecting a reference model for software process deployment [Artículo]. Madrid. Universidad Politécnica de Madrid.
2. **Castillo, J. & Palomino, L. (2012)** Implementación de un Datamart como una solución de Inteligencia de Negocios para el área de logística de T-Impulso [Artículo] Universidad Nacional Mayor de San Marcos Facultad de Ingeniería de Sistemas e Informática.
3. **Espejo, A. (2013)** Aplicación De La Extensión Para La Construcción De La Guía Del PMBOK - Tercera Edición, En La Gerencia De Proyecto De Una Presa De Relaves En La Unidad Operativa Arcata-Arequipa. [Tesis]. Perú. Pontificia Universidad La Católica.
4. **Garita, O. (2010).** Plan de gestión de proyecto para la construcción del Oficentro de la iglesia misión carismática internacional san José. [Tesis]. Costa Rica. Universidad para la Cooperación Internacional.
5. **Imhoff, C., Galemno N. & Geiger, J. (2003)** Mastering Data Warehouse Design.

6. **Leonard, E. & Castro, Y. (2013).** Metodologías para desarrollar Almacén de Datos. Revista de Arquitectura e Ingeniería.
7. **Moreno, H. (2013)** Análisis, Diseño e Implementación de Datamarts para las áreas de ventas y Recursos Humanos de una empresa dedicada a la exportación e importación de productos alimentación. [Tesis]. Perú. Pontificia Universidad La Católica.
8. **Moss, L. & Atre, S. (2003).** Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications.
9. **PMBOK (2013)** A Guide to the project management body of knowledge, Fifth Edition. Project Management Institute, Inc. Pennsylvania. ISBN 978-1-933890-51-7

Electrónicas:

1. **Aguerrondo, I. (s.f.)** La Calidad de la Educación. Recuperado 21 de Agosto del 2014, <http://www.oei.es/calidad2/aguerrondo.htm>
2. **BI-Spain. (2012)** Learning Circle mejora el desempeño de los estudiantes de educación básica con la tecnología de BI y de integración de Information Builders. Recuperado el 16 de Agosto 2014, de <http://www.bi-spain.com/articulo/72123/information-builders/educacion-y-formacion/learning-circle-mejora-el-desempeno-de-los-estudiantes-de-educacion-basica-con-la-tecnologia-de-bi-y-de-integracion-de-information-builders>.
3. **Business Object (s.f.) SAP Business Objects Business Intelligence (BI).** Recuperado el 25 de Agosto del 2014, <http://www.saphana.com/docs/DOC-3765>
4. **Cecchini, S. (2005)** Indicadores sociales de América Latina y el Caribe. Recuperado el 20 de Agosto del 2014, <http://www.cepal.org/publicaciones/xml/0/23000/lcl2383e.pdf>

5. **Cognos (s.f.)** Software Cognos - Gestión de rendimiento y business intelligence. Recuperado el 25 de Agosto del 2014, <http://www-01.ibm.com/software/pe/analytics/cognos/>
6. **DataStage (s.f.)** IBM® InfoSphere® DataStage®. Recuperado el 22 de Agosto del 2014, <http://www-03.ibm.com/software/products/en/ibminfodata>
7. **ISO/ IEC 9126 (s.f)** Software Engineering. Recuperado el 02 de Noviembre del 2014, de http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=22749
8. **Mendez, A., Mártire, A., Britos, P. & Garcia-Martínez, R. (2003).** Fundamentos de Data Warehouse. Recuperado el 11 de Agosto del 2014 de <http://artemisa.unicauca.edu.co/~ecaldon/docs/bd/fundamentosdedatawarehouse.pdf>
9. **MicroStrategy (s.f.)** MicroStrategy Server. Recuperado el 25 de Agosto del 2014, <http://www.microstrategy.com/es/acerca-de-nosotros/experiencia>
10. **Oficina Regional de Educación para América Latina y el Caribe. (2013)** Enfoques Estratégicos sobre las TICs en Educación en América Latina y el Caribe. Recuperado el 15 de Agosto 2014, de <http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Santiago/images/ticsesp.pdf>.
11. **OFIN (s.f.)** Oficina de Informática del Ministerio de Educación. Recuperado el 23 de Agosto del 2014, <http://www.minedu.gob.pe/ofin/>

12. **Page, A. (1993)** Los Indicadores de Calidad de Educación. Recuperado el 21 de Agosto del 2014, <http://dialnet.unirioja.es/descarga/articulo/2383871.pdf>
13. **Pentaho (s.f.)** Pentaho Data Integration. Recuperado el 25 de Agosto del 2014, <http://www.pentaho.com/product/data-integration>
14. **SIAGIE (s.f.)** Sistema de Información de Apoyo a la Gestión de la Institución Educativa. Recuperado el 22 de Agosto del 2014, http://siagie.minedu.gob.pe/archivos/brochure_pbp.pdf
15. **SIS (s.f.)** SQL Server Integration Services. Recuperado el 25 de Agosto del 2014, <http://msdn.microsoft.com/es-pe/library/ms141026.aspx>
16. **SPE (s.f.)** Secretaría de Planificación Estratégica del Ministerio de Educación. Recuperado el 23 de Agosto del 2014, www.minedu.gob.pe/planificacionestrategica/
17. **Tamayo, M. & Moreno, F. (2006)** Análisis del modelo de almacenamiento MOLAP frente al modelo de almacenamiento ROLAP. Recuperado el 18 de Agosto 2014, de <http://www.redalyc.org/pdf/643/64326317.pdf>
18. **Virtual Strategy (2012)** South Leicestershire College Enhances Performance Management. Recuperado el 16 de Agosto 2014, de <http://www.virtual-strategy.com/2012/03/29/south-leicestershire-college-enhances-performance-management#axzz3CqlllqIF>.
19. **Wolff, G. (2002).** La Tecnología Datawarehousing. Recuperado el 18, de Agosto del 2014, de <http://www.inf.udec.cl/~revista/ediciones/edicion3/cwolff.PDF>.

ANEXOS

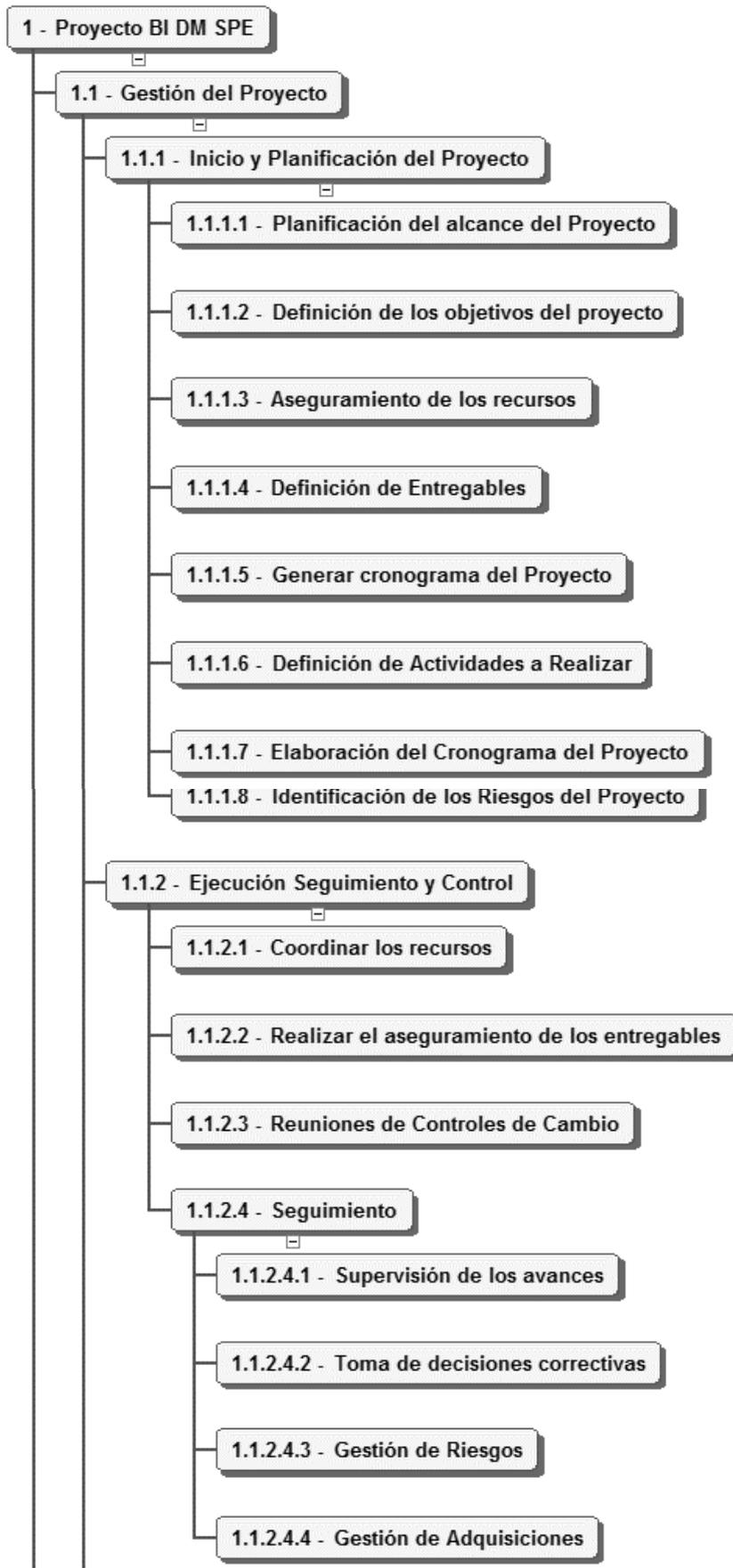
1. Cronograma del Proyecto
2. EDT del Proyecto
3. Matriz de Consistencia
4. Análisis Dimensional
5. Diseño de ETL
6. Estimación preliminar de Pruebas
7. Registros del Cuadre de Conteo
8. Consulta de Dimensiones
9. Acta de Constitución del Proyecto
10. Solicitud de Control de Cambios
11. Encuesta de Satisfacción
12. Configuración de Firewall CISCO ASA

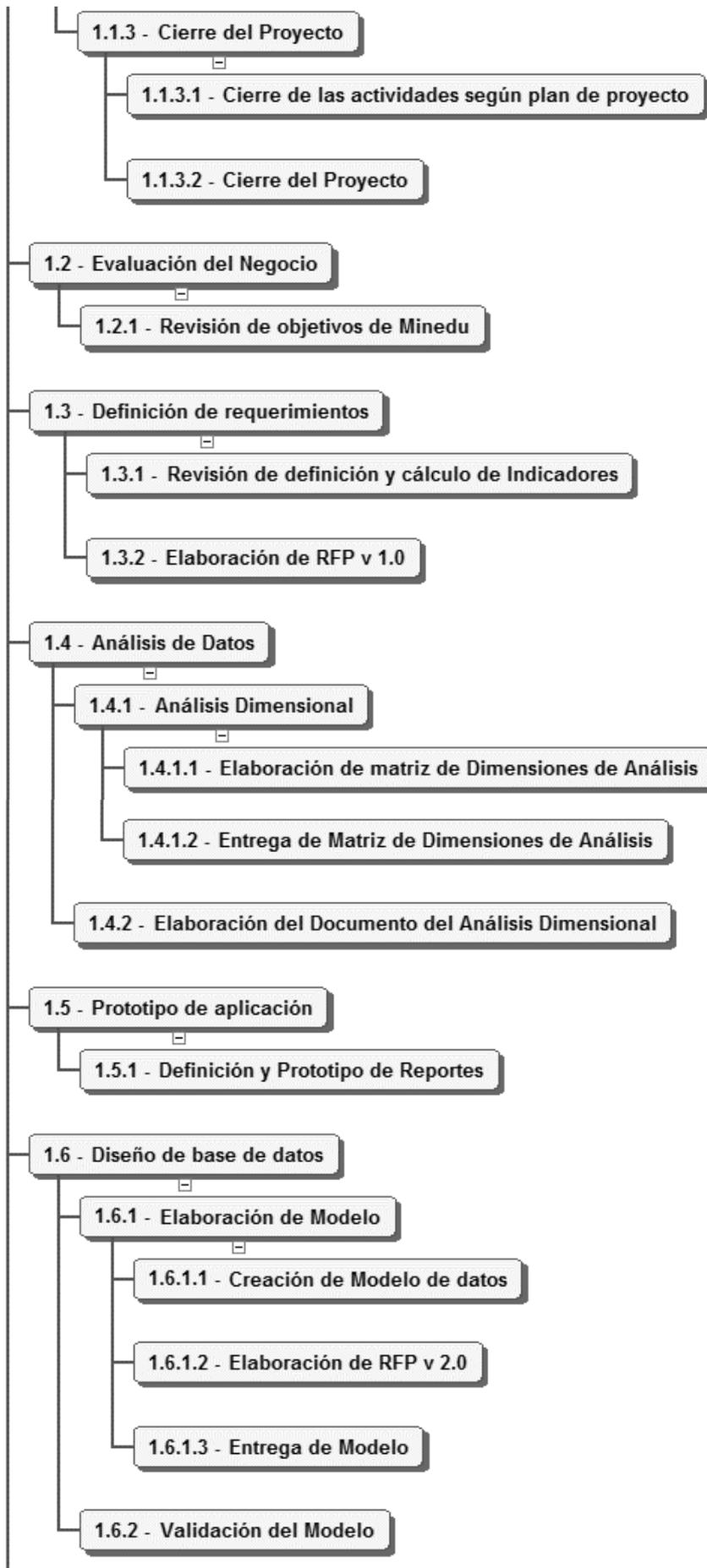
ANEXO 01 CRONOGRAMA DEL PROYECTO

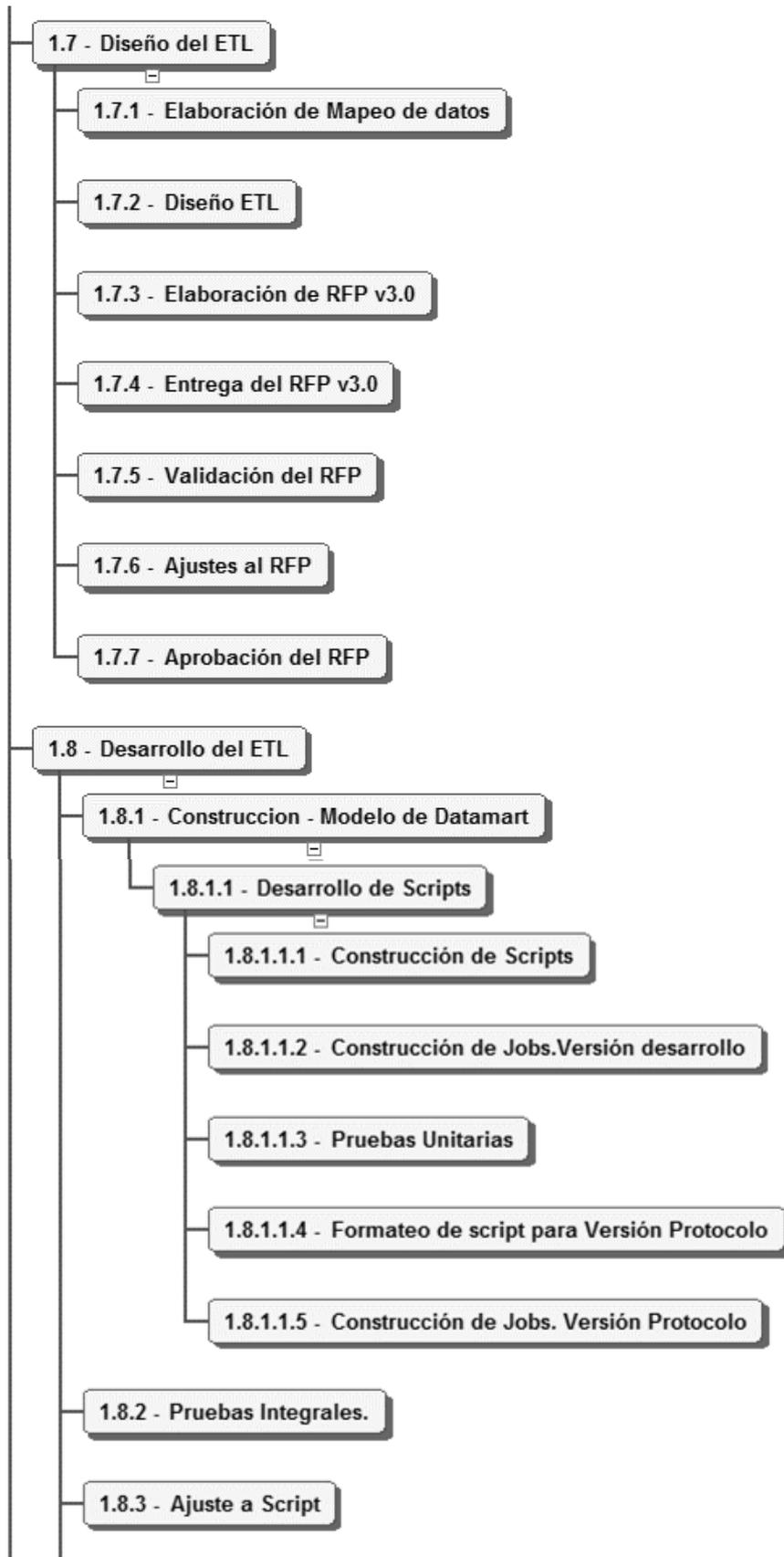
EDT	Nombre de tarea	Trabajo	Comienzo	Fin
1	Proyecto BI DM SPE	437 horas	sáb 09/08/14	sáb 25/10/14
1.1	Gestión del Proyecto	85 horas	sáb 09/08/14	vie 24/10/14
1.1.1	Inicio y Planificación del Proyecto	4 horas	sáb 09/08/14	mié 13/08/14
1.1.1.1	Planificación del alcance del Proyecto	0.5 horas	lun 11/08/14	lun 11/08/14
1.1.1.2	Definición de los objetivos del proyecto	0.5 horas	sáb 09/08/14	lun 11/08/14
1.1.1.3	Aseguramiento de los recursos	0.5 horas	lun 11/08/14	lun 11/08/14
1.1.1.4	Definición de Entregables	0.5 horas	lun 11/08/14	lun 11/08/14
1.1.1.5	Generar cronograma del Proyecto	0.5 horas	lun 11/08/14	mar 12/08/14
1.1.1.6	Definición de Actividades a Realizar	0.5 horas	mar 12/08/14	mié 13/08/14
1.1.1.7	Elaboración del Cronograma del Proyecto	0.5 horas	lun 11/08/14	mar 12/08/14
1.1.1.8	Identificación de los Riesgos del Proyecto	0.5 horas	lun 11/08/14	lun 11/08/14
1.1.2	Ejecución Seguimiento y Control	80 horas	lun 11/08/14	mié 22/10/14
1.1.2.1	Coordinar los recursos	0 horas	lun 11/08/14	lun 11/08/14
1.1.2.2	Realizar el aseguramiento de los entregables	0 horas	lun 11/08/14	lun 11/08/14
1.1.2.3	Reuniones de Controles de Cambio	0 horas	lun 11/08/14	lun 11/08/14
1.1.2.4	Seguimiento	80 horas	lun 11/08/14	mié 22/10/14
1.1.2.4.1	Supervisión de los avances	20 horas	lun 11/08/14	mié 22/10/14
1.1.2.4.2	Toma de decisiones correctivas	20 horas	lun 11/08/14	mié 22/10/14
1.1.2.4.3	Gestión de Riesgos	20 horas	lun 11/08/14	mié 22/10/14
1.1.2.4.4	Gestión de Adquisiciones	20 horas	lun 11/08/14	mié 22/10/14
1.1.3	Cierre del Proyecto	1 hora	jue 23/10/14	vie 24/10/14
1.1.3.1	Cierre de las actividades según plan de proyecto	1 hora	jue 23/10/14	vie 24/10/14
1.1.3.2	Cierre del Proyecto	0 horas	vie 24/10/14	vie 24/10/14
1.2	Evaluación del Negocio	4 horas	lun 11/08/14	mar 12/08/14
1.2.1	Revisión de objetivos de Minedu	4 horas	lun 11/08/14	mar 12/08/14
1.3	Definición de requerimientos	16 horas	mar 12/08/14	jue 14/08/14
1.3.1	Revisión de definición y cálculo de Indicadores	8 horas	mar 12/08/14	mar 12/08/14
1.3.2	Elaboración de RFP v 1.0	8 horas	mié 13/08/14	jue 14/08/14
1.4	Análisis de Datos	34 horas	jue 14/08/14	vie 22/08/14
1.4.1	Análisis Dimensional	32.5 horas	jue 14/08/14	jue 21/08/14
1.4.1.1	Elaboración de matriz de Dimensiones de Análisis	32 horas	jue 14/08/14	mié 20/08/14
1.4.1.2	Entrega de Matriz de Dimensiones de Análisis	0.5 horas	mié 20/08/14	mié 20/08/14
1.4.2	Elaboración del Documento del Análisis Dimensional	1.5 horas	jue 21/08/14	vie 22/08/14
1.5	Prototipo de aplicación	12 horas	vie 22/08/14	mar 26/08/14
1.5.1	Definición y Prototipo de Reportes	12 horas	vie 22/08/14	mar 26/08/14
1.6	Diseño de base de datos	48 horas	mié 27/08/14	jue 04/09/14
1.6.1	Elaboración de Modelo	36 horas	mié 27/08/14	mar 02/09/14
1.6.1.1	Creación de Modelo de datos	30 horas	mié 27/08/14	lun 01/09/14
1.6.1.2	Elaboración de RFP v 2.0	5 horas	lun 01/09/14	mar 02/09/14
1.6.1.3	Entrega de Modelo	1 hora	vie 29/08/14	sáb 30/08/14
1.6.2	Validación del Modelo	12 horas	mar 02/09/14	jue 04/09/14
1.7	Diseño del ETL	34 horas	mié 10/09/14	vie 19/09/14
1.7.1	Elaboración de Mapeo de datos	30 horas	mié 10/09/14	mar 16/09/14
1.7.2	Diseño ETL	0.5 horas	jue 18/09/14	jue 18/09/14
1.7.3	Elaboración de RFP v3.0	0.5 horas	jue 18/09/14	vie 19/09/14
1.7.4	Entrega del RFP v3.0	0.5 horas	vie 19/09/14	vie 19/09/14
1.7.5	Validación del RFP	1 hora	vie 19/09/14	vie 19/09/14
1.7.6	Ajustes al RFP	1 hora	vie 19/09/14	vie 19/09/14
1.7.7	Aprobación del RFP	0.5 horas	jue 18/09/14	jue 18/09/14
1.8	Desarrollo del ETL	118 horas	vie 19/09/14	jue 16/10/14
1.8.1	Construcción - Modelo de Datamart	53 horas	vie 19/09/14	mié 08/10/14
1.8.1.1	Desarrollo de Scripts	53 horas	vie 19/09/14	mié 08/10/14
1.8.1.1.1	Construcción de Scripts	30 horas	vie 19/09/14	jue 25/09/14
1.8.1.1.2	Construcción de Jobs.Versión desarrollo	9 horas	jue 25/09/14	sáb 04/10/14

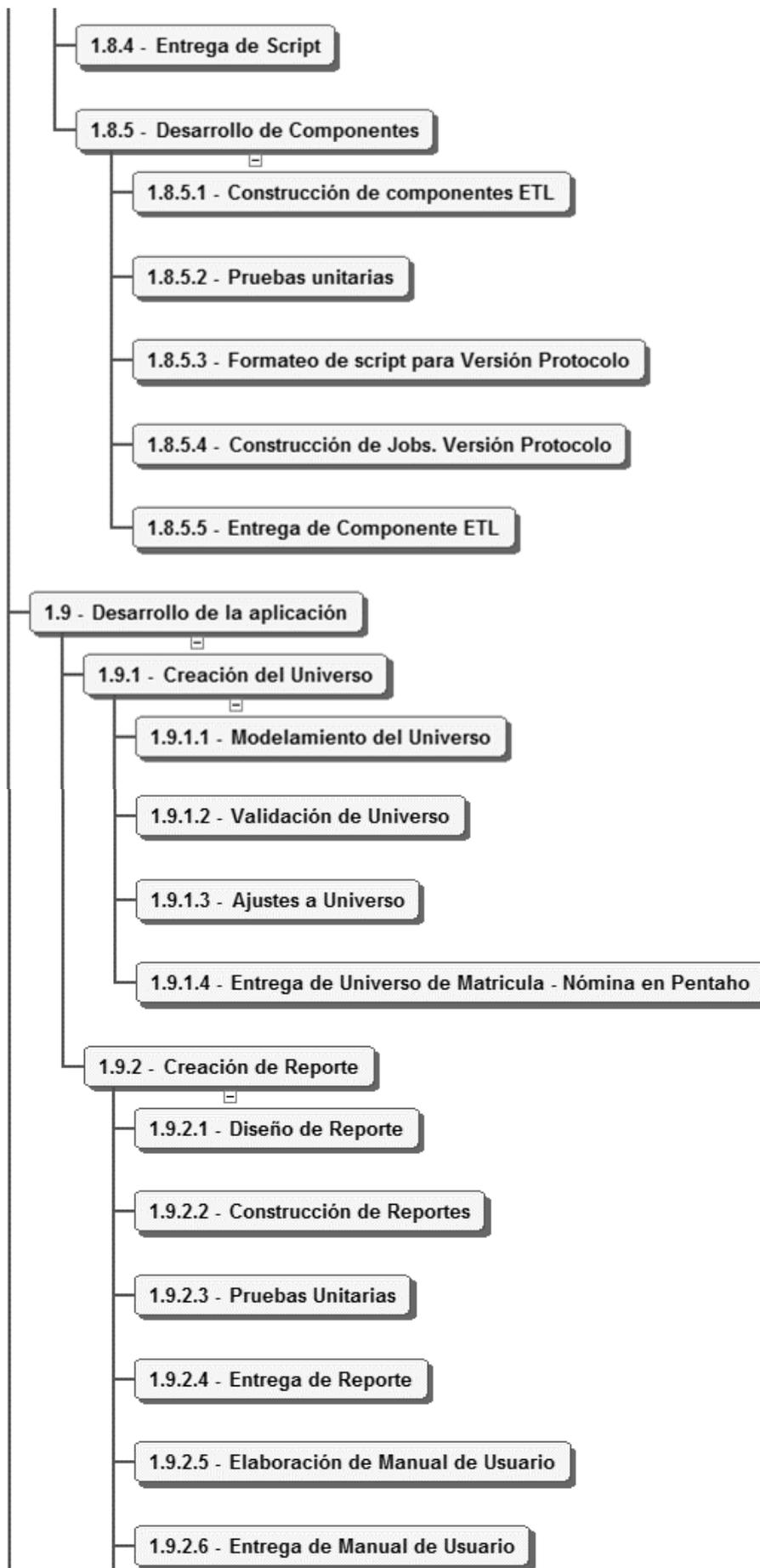
EDT	Nombre de tarea	Trabajo	Comienzo	Fin
1.8.1.1.3	Pruebas Unitarias	8 horas	sáb 04/10/14	mar 07/10/14
1.8.1.1.4	Formateo de script para Versión Protocolo	3 horas	sáb 04/10/14	mar 07/10/14
1.8.1.1.5	Construcción de Jobs. Versión Protocolo	3 horas	mar 07/10/14	mié 08/10/14
1.8.2	Pruebas Integrales.	8 horas	mar 07/10/14	mié 08/10/14
1.8.3	Ajuste a Script	1 hora	mié 08/10/14	mié 08/10/14
1.8.4	Entrega de Script	0.5 horas	mié 01/10/14	mié 01/10/14
1.8.5	Desarrollo de Componentes	55.5 horas	mié 08/10/14	jue 16/10/14
1.8.5.1	Construcción de componentes ETL	48 horas	mié 08/10/14	jue 16/10/14
1.8.5.2	Pruebas unitarias	3 horas	jue 16/10/14	jue 16/10/14
1.8.5.3	Formateo de script para Versión Protocolo	1 hora	jue 16/10/14	jue 16/10/14
1.8.5.4	Construcción de Jobs. Versión Protocolo	3 horas	jue 16/10/14	jue 16/10/14
1.8.5.5	Entrega de Componente ETL	0.5 horas	lun 13/10/14	lun 13/10/14
1.9	Desarrollo de la aplicación	52 horas	mié 15/10/14	mar 21/10/14
1.9.1	Creación del Universo	23.5 horas	jue 16/10/14	mar 21/10/14
1.9.1.1	Modelamiento del Universo	20 horas	jue 16/10/14	mar 21/10/14
1.9.1.2	Validación de Universo	1 hora	mar 21/10/14	mar 21/10/14
1.9.1.3	Ajustes a Universo	2 horas	mar 21/10/14	mar 21/10/14
1.9.1.4	Entrega de Universo de Matricula - Nómina en Pentaho	0.5 horas	vie 17/10/14	vie 17/10/14
1.9.2	Creación de Reporte	28.5 horas	mié 15/10/14	mar 21/10/14
1.9.2.1	Diseño de Reporte	15 horas	jue 16/10/14	jue 16/10/14
1.9.2.2	Construcción de Reportes	4 horas	jue 16/10/14	jue 16/10/14
1.9.2.3	Pruebas Unitarias	2 horas	jue 16/10/14	jue 16/10/14
1.9.2.4	Entrega de Reporte	0.5 horas	mié 15/10/14	mié 15/10/14
1.9.2.5	Elaboración de Manual de Usuario	3 horas	mié 15/10/14	jue 16/10/14
1.9.2.6	Entrega de Manual de Usuario	4 horas	lun 20/10/14	mar 21/10/14
1.9.2.7	Pase a Certificación	0 horas	lun 20/10/14	lun 20/10/14
1.10	Certificación	19 horas	mar 21/10/14	mié 22/10/14
1.10.1	Pruebas Funcionales y Pruebas de Estres	8 horas	mar 21/10/14	mié 22/10/14
1.10.2	Pruebas de Certificación	8 horas	mar 21/10/14	mié 22/10/14
1.10.3	Ajustes a los procesos	3 horas	mié 22/10/14	mié 22/10/14
1.11	Implementación	12 horas	mié 22/10/14	vie 24/10/14
1.11.1	Implementación de los Componentes	12 horas	mié 22/10/14	vie 24/10/14
1.12	Evaluación postproducción	3 horas	jue 23/10/14	sáb 25/10/14
1.12.1	Seguimiento	1 hora	jue 23/10/14	vie 24/10/14
1.12.2	Informe de cierre de proyecto	1 hora	jue 23/10/14	vie 24/10/14
1.12.3	Entrega de informe de Cierre de proyecto	1 hora	vie 24/10/14	sáb 25/10/14
1.12.4	Cierre del Proyecto	0 horas	sáb 25/10/14	sáb 25/10/14

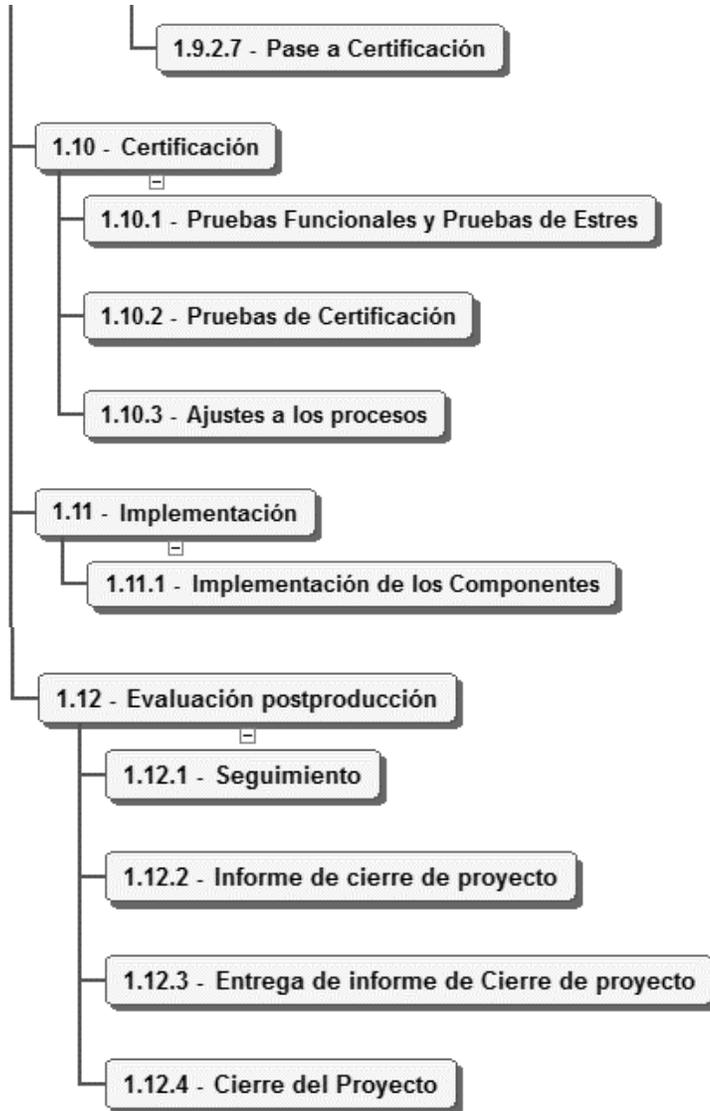
ANEXO 02 EDT DEL PROYECTO











ANEXO 03 MATRIZ DE CONSISTENCIA

TITULO	Problema	Objetivo General y Especifico	Metodología	Resultados	Conclusiones
<p style="text-align: center;">ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE DATAMART PARA LA GENERACION DE REPORTES SECRETARÍA PLANIFICACIÓN ESTRATÉGICA MINEDU</p>	<p>Problema Engorroso proceso manual para disponer de información en los procesos de matrícula y nómina generando un debilitado apoyo a la toma de decisiones de la Secretaría de Planificación Estratégica del Ministerio de Educación.</p> <p>Problema Específico</p> <ul style="list-style-type: none"> -Falta de un aplicativo que procese la información oportuna para la Secretaría de Planificación Estratégica. -Información inconsistente y compleja para el usuario final generando inadecuada toma de decisiones. -Obtención de la información manualmente de OFIN para el área usuaria de manera inadecuada. 	<p>Mejorar el apoyo a la toma de decisiones en los procesos de matrículas y nominas en la Secretaría de Planificación estratégica mediante la implementación del Datamart.</p> <p>Objetivos Específicos:</p> <ul style="list-style-type: none"> -Implementar un Datamart para la toma de decisiones de la Secretaría de Planificación Estratégica del MINEDU. -Identificar y definir los indicadores de los procesos de matrícula y nómina. -Construir los procesos ETL para la Extracción, Transformación y Carga de datos correspondientes al proceso de matrículas y de nóminas desde el SIAGIE 	<p>Tipo de Investigación Investigación Aplicada Investigación Documental</p> <p>Software: SQL Server 2008 Pentaho Tableau</p> <p>Metodología de Gestión PMBOK 5.0</p> <p>Metodología de Desarrollo Roadmap</p>	<p>La Secretaría de Planificación Estratégica del MINEDU cuenta con un mejor apoyo para la toma de decisiones que permitan dar seguimiento a las Instituciones Educativas de educación básica regular.</p> <p>-Ausencia de la generación manual de reportes. Obtiene la información oportuna Reportes personalizados e intuitivos que apoyan a la toma de decisiones para la Secretaría de Planificación Estratégica.</p> <p>-Se cuenta con una matriz de indicadores definidos y establecidos para cada reporte de interés para la Secretaría de Planificación Estratégica.</p> <p>-El proceso ETL es realizado de forma automatizada.</p>	<p>o Es indispensable para la Secretaría de Planificación Estratégica del MINEDU, contar con una herramienta tecnológica que apoye a la toma de decisiones para el seguimiento en las Instituciones Educativas de educación básica regular en los Procesos de Matrícula y Nóminas.</p> <p>o El Datamart constituirá una herramienta válida para el análisis cuantitativo y cualitativo de los procesos en mención.</p> <p>o La identificación de los principales indicadores son de gran importancia ayuda obtener una visión general y una visión detalla del estado de las Instituciones Educativas a nivel Departamental.</p> <p>o La selección de herramientas para la extracción y explotación de datos mejoró el proceso de obtención de la información, evitando colapsar el sistema transaccional, SIAGIE.</p>

ANEXO 04 ANÁLISIS DIMENSIONAL

a) Dimensión Institución Educativa

i. Descripción

Esta dimensión abarca a la IE sus principales atributos:
Identificación de IE, Modalidad de IE.

ii. Atributos

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_IE	Código numérico de la Institución Educativa. Es un correlativo.	INTEGER	NINGUNO
COD_MOD	Código que identifica a la Institución Educativa	VARCHAR(7)	NINGUNO
ANEXO	Número de sede que cuenta la IE. 0 si es única Institución Principal, 1 Si es una Institución principal y una sede, N.	VARCHAR(1)	NINGUNO
NIVEL	Nivel de la Institución Educativa pertenece.	VARCHAR(30)	NINGUNO
ABR_MODALIDAD	Abreviación de la Modalidad de la Institución Educativa. EBR(Educación Básica Regular)	VARCHAR(10)	NINGUNO
NOMBRE_IE	Nombre completo de la Institución Educativa	VARCHAR(100)	NINGUNO
UBIGEO	Código de Ubicación Geográfica de la Institución Educativa	VARCHAR(6)	NINGUNO

iii. Jerarquías

Esta jerarquía agrupa la modalidad que asume las Instituciones Educativas hacia un tipo en particular

Nivel	Atributos
Nivel 1	Identificación de IE
Nivel 2	Modalidad de IE
Nivel 3	Nivel de IE
Nivel 4	Nombre de Institución Educativa
Nivel 5	Anexo
Nivel 6	Ubigeo

b) Dimensión Estudiantes

i. Descripción

Esta dimensión abarca la situación de los estudiantes sus principales atributos: Código de Estudiante, Nivel de Estudiante.

ii. Atributos

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_ESTUDIANTE	Código numérico del estudiante. Es un correlativo.	INTEGER	NINGUNO
COD_ESTUDIANTE	Código que identifica al estudiante	INTEGER	NINGUNO
ANIO	Año que cursa el estudiante	INTEGER	NINGUNO
NIVEL	Nivel del estudiante	VARCHAR(30)	NINGUNO
GRADO	Grado del estudiante	VARCHAR(30)	NINGUNO
SECCION	Sección matriculada del estudiante	VARCHAR(30)	NINGUNO
ESTADO_MATRICULA	Estado de la matrícula	VARCHAR(50)	NINGUNO
ESTADO_ESTUDIANTE	Estado del Estudiante	VARCHAR(30)	NINGUNO
NOMBRE_IE	Nombre de la Institución Educativa	VARCHAR(100)	NINGUNO
SEXO	Género del estudiante	SMALLINT	NINGUNO

iii. Jerarquías

Esta jerarquía agrupa los datos del estudiante:

Nivel	Atributos
Nivel 1	Código de Estudiante
Nivel 2	Año del Estudiante
Nivel 3	Nivel de Estudiante
Nivel 4	Grado de Estudiante
Nivel 5	Sección de Estudiante
Nivel 6	Estado de Matrícula
Nivel 7	Estado del Estudiante
Nivel 8	Nombre de la Institución Educativa
Nivel 9	Género de Estudiante

c) Dimensión Ubicación

i. Descripción

Esta dimensión abarca la ubicación geográfica de las Instituciones Educativas sus principales atributos: Ubigeo, Departamento, Provincia, Distrito.

ii. Atributos

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_UBICACIÓN	Código numérico de la ubicación	INTEGER	NINGUNO
UBIGEO	Código de ubicación geográfica	INTEGER	NINGUNO
DEPARTAMENTO	Nombre Departamento	VARCHAR(45)	NINGUNO
PROVINCIA	Nombre Provincia	VARCHAR(45)	NINGUNO
DISTRITO	Nombre Distrito	INTEGER	NINGUNO

iii. Jerarquías

Esta jerarquía agrupa la ubicación geográfica de las Instituciones Educativas:

Nivel	Atributos
Nivel 1	Ubigeo
Nivel 2	Departamento
Nivel 3	Provincia
Nivel 4	Distrito
Nivel 5	Latitud
Nivel 6	Longitud

d) Dimensión Tiempo

i. Descripción

Esta dimensión abarca los periodos de análisis que se suelen usar.

ii. Atributos

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_FECHA	Código numérico del tiempo	INTEGER	NINGUNO
ANIO	Año	INTEGER	NINGUNO
SEMESTRE	Número de Semestre	INTEGER	NINGUNO
TRIMESTRE	Número de Trimestre	INTEGER	NINGUNO
BIMESTRE	Número de Bimestre	INTEGER	NINGUNO
ANIMES	Año y Mes. Ej. 201402	INTEGER	NINGUNO
CODMES	Código de Mes (1 a 12)	INTEGER	NINGUNO
MES	Nombre del Mes	VARCHAR(45)	NINGUNO
DIA	Número Día	VARCHAR(45)	NINGUNO
ESTACION	Género del estudiante	VARCHAR(20)	NINGUNO

iii. Jerarquías

Esta jerarquía contiene la clasificación de por Año, Semestre, Trimestre, Bimestre, etc.

Nivel	Atributos
Nivel 1	Año
Nivel 2	Semestre
Nivel 3	Trimestre
Nivel 4	Bimestre
Nivel 5	Año Mes
Nivel 4	Mes
Nivel 6	Día
Nivel 7	Estación

e) Dimensión Modalidad Institución Educativa

i. Descripción

Esta dimensión abarca las modalidades de las Instituciones Educativas sus principales atributos: Código de Modalidad, descripción de modalidad.

ii. Atributos

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_MODALIDAD	Código numérico de la Modalidad. Es un número correlativo.	INTEGER	NINGUNO
COD_MODALIDAD	Código Modalidad de la Institución Educativa	VARCHAR(2)	NINGUNO
MODALIDAD	Nombre de la Modalidad de la Institución Educativa	VARCHAR(100)	NINGUNO
ABR_MODALIDAD	Abreviatura de la Modalidad de la Institución Educativa	VARCHAR(10)	NINGUNO

iii. Jerarquías

Esta jerarquía contiene todo el detalle de la clasificación:

Nivel	Atributos
Nivel 1	Código Modalidad
Nivel 2	Nombre de Modalidad
Nivel 3	Abreviación de Modalidad

f) Dimensión Estado de Matrícula

i. Descripción

Esta dimensión abarca el estado de las matrículas sus principales atributos: Código Estado de Matrícula, Estado de matrícula

ii. Atributos

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_EST_MAT	Código numérico de la Matricula. Es un número correlativo.	INTEGER	NINGUNO
COD_EST_MAT	Código Modalidad de la Matricula	SMALLINT	NINGUNO
ESTADO_MATRICULA	Nombre de la Modalidad de la Matricula	VARCHAR(50)	NINGUNO

iii. Jerarquías

Esta jerarquía contiene todo el detalle de la clasificación:

Nivel	Atributos
Nivel 1	Código Estado de Matrícula
Nivel 2	Estado de Matrícula

g) Dimensión Grado Educativo

i. Descripción

Esta dimensión abarca el detalle del estado de Grados Educativos sus principales atributos: Código de Grado, Descripción de Grado

ii. Atributos

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_GRADO	Código numérico del grado de matrícula. Es un número correlativo.	INTEGER	NINGUNO
COD_GRADO	Código Grado de la Matricula	CHAR(2)	NINGUNO
GRADO	Nombre del Grado de la Matricula	CHAR(30)	NINGUNO

iii. Jerarquías

Esta jerarquía contiene todo el detalle de la clasificación:

Nivel	Atributos
Nivel 1	Código Grado
Nivel 2	Descripción de Grado

h) Dimensión Nivel Educativo

i. Descripción

Esta dimensión abarca los niveles educativos sus principales atributos: Código Nivel, Descripción de Nivel

ii. Atributos

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_NIVEL	Código numérico de Nivel de educativo. Es un número correlativo.	CHAR(2)	NINGUNO
COD_NIVEL	Código Nivel Educativo	CHAR(2)	NINGUNO
NIVEL	Nombre del Nivel Educativo	CHAR(2)	NINGUNO

iii. Jerarquías

Esta jerarquía contiene todo el detalle de la clasificación:

Nivel	Atributos
Nivel 1	Código Nivel Educativo
Nivel 2	Descripción de Nivel Educativo

i) Dimensión Estado Estudiante

i. Descripción

Esta dimensión abarca el detalle del estado de estudiantes sus principales atributos: Código de Estudiante, Descripción de Estudiante

ii. Atributos

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_EST_EST	Código numérico de Estado de estudiante. Es un número correlativo.	INTEGER	NINGUNO
COD_EST_EST	Código de estado de Estudiante	CHAR(2)	NINGUNO
ESTADO_ESTUDIANTE	Nombre de Estado de Estudiante	VARCHAR(100)	NINGUNO

iii. *Jerarquías*

Esta jerarquía contiene todo el detalle de la clasificación:

Nivel	Atributos
Nivel 1	Código Estudiante
Nivel 2	Estado de Estudiante

j) **Dimensión Nivel Educativo**

iv. *Descripción*

Esta dimensión abarca los niveles educativos sus principales atributos: Código Nivel, Descripción de Nivel

v. *Atributos*

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_NIVEL	Código numérico de Nivel de educativo. Es un número correlativo.	CHAR(2)	NINGUNO
COD_NIVEL	Código Nivel Educativo	CHAR(2)	NINGUNO
NIVEL	Nombre del Nivel Educativo	CHAR(2)	NINGUNO

vi. *Jerarquías*

Esta jerarquía contiene todo el detalle de la clasificación:

Nivel	Atributos
Nivel 1	Código Nivel Educativo
Nivel 2	Descripción de Nivel Educativo

k) **Dimensión Sección**

i. *Descripción*

Esta dimensión abarca el detalle de las secciones sus principales atributos: Código de Sección, Sección

ii. Atributos

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_SECCION	Código numérico de la sección. Es un número correlativo.	INTEGER	NINGUNO
ANIO	Año del Estudiante	CHAR(2)	NINGUNO
NIVEL	Nivel del Estudiante	VARCHAR(30)	NINGUNO
GRADO	Grado del Estudiante	VARCHAR(30)	NINGUNO
SECCION	Sección del Estudiante	VARCHAR(30)	NINGUNO
TIPO_FORMATO	Tipo de Formato Nómina	SMALLINT	NINGUNO
NOMBRE_IE	Nombre de la IE	VARCHAR(100)	NINGUNO
SEXO	Sexo del estudiante	SMALLINT	NINGUNO
ESTADO_FORMATO	Estado de Formato Nómina	CHAR(50)	NINGUNO

iii. Jerarquías

Esta jerarquía contiene todo el detalle de la clasificación de la dimensión sección:

Nivel	Atributos
Nivel 1	Código Sección
Nivel 2	Año
Nivel 3	Nivel
Nivel 4	Grado

Nivel 5	Sección
Nivel 6	Tipo Formato
Nivel 7	Nombre Institución Educativa
Nivel 8	Sexo
Nivel 9	Estado Formato

I) Dimensión Tipo Formato

i. Descripción

Esta dimensión abarca el detalle del tipo de formato sus principales atributos: Código de Tipo Formato, Tipo Formato

ii. Atributos

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_TIPO	Código numérico del tipo de formato. Es un número correlativo.	INTEGER	NINGUNO
COD_TIPO	Código Tipo Formato	CHAR(2)	NINGUNO
TIPO	Tipo Formato	CHAR(50)	NINGUNO

iii. Jerarquías

Esta jerarquía contiene todo el detalle de la clasificación de la dimensión tipo formato:

Nivel	Atributos
Nivel 1	Código Tipo Formato
Nivel 2	Tipo Formato

m) Dimensión Estado Formato

i. Descripción

Esta dimensión abarca el detalle del estado de formato sus principales atributos: Código de Estado Formato, Tipo Estado

ii. Atributos

Nombre del Atributo	Contenido		
	Descripción	Formato	Valor por Defecto
ID_ESTADO	Código numérico del tipo de formato. Es un número correlativo.	INTEGER	NINGUNO
COD_ESTADO	Código Estado Formato	SMALLINT	NINGUNO
ESTADO_FORMATO	Estado Formato	CHAR(50)	NINGUNO

iii. Jerarquías

Esta jerarquía contiene todo el detalle de la clasificación de la dimensión tipo formato:

Nivel	Atributos
Nivel 1	Código Estado Formato
Nivel 2	Tipo Estado Formato

Fact de Matrícula

i. Descripción

Involucra todas las matrículas realizadas. La Fact permitirá cumplir con la elaboración de los reportes.

ii. Granularidad

N°	Nombre de la Dimensión	Descripción	Llave Primaria
1	Institución Educativa	Esta dimensión abarca toda la información acerca de la IE, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Código de IE, Código Modalidad IE, Anexo, Nivel, Descripción Modalidad, Nombre de IE, Ubigeo	SI
2	Estudiante	Esta dimensión abarca toda la información acerca de los estudiantes, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Identificación Estudiante, Código Estudiante, Nombre de Estudiante, Año , Nivel, Grado, Sección, Sexo	SI
3	Ubicación	Esta dimensión abarca toda la información acerca de la ubicación, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Ubigeo, Departamento, Provincia, Distrito	SI
4	Tiempo	Esta dimensión abarca toda la información acerca la organización del tiempo en los periodos de análisis, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Año, Semestre, Trimestre, Bimestre, Año Mes, Código Mes, Mes, Día, Estación	SI
5	Modalidad IE	Esta dimensión abarca toda la información acerca de la modalidad de IE, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Código Modalidad IE, Descripción de modalidad	SI
6	Estado Matrícula	Esta dimensión abarca toda la información acerca estado de matrícula, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Código Estado de Matrícula, Descripción de Matricula	SI

N°	Nombre de la Dimensión	Descripción	Llave Primaria
7	Grado Educativo	Esta dimensión abarca toda la información acerca de Grado Educativo, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: <ul style="list-style-type: none"> - Código Grado Educativo - Descripción de Grado Educativo 	
8	Nivel Educativo	Esta dimensión abarca toda la información acerca del Nivel Educativo, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: <ul style="list-style-type: none"> - Código Nivel Educativo - Descripción Nivel Educativo 	
9	Estado Estudiante	Esta dimensión abarca toda la información acerca del Estado del estudiante, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: <ul style="list-style-type: none"> - Código Estado de Estudiante - Descripción de Estado de Estudiante 	

iii. Medidas

N°	Nombre	Descripción
1	% pre matriculados	Mide la proporción de pre matriculados sobre el total de estudiantes en el sistema educativo en el Departamento
2	% matrícula en proceso	Mide la proporción estado matrícula en proceso sobre el total de estudiantes en el sistema educativo en el Departamento
3	% matrícula definitiva	Mide la proporción estado matriculada Definitivo sobre el total de estudiantes en el sistema educativo en el Departamento

N°	Nombre	Descripción
4	% matrícula anulada	Mide la proporción estado matrícula anulada sobre el total de estudiantes en el sistema educativo en el Departamento
5	% de matrícula según género	Mide la proporción de matriculados sobre el total de estudiantes por género del estudiante
6	% estado de matrícula por grado	Mide la proporción de estado de matrícula por grado sobre el total de estudiantes en el sistema educativo
7	% estado del estudiante	Mide la proporción de estado de estudiante sobre el total de estudiantes en el sistema educativo
8	Número de estudiantes por aula	Mide la eficacia de cumplimiento de distribución de estudiantes por aula
9	% de estudiantes por nivel educativo	Mide a los estudiantes que están por niveles educativos por

Fact de Formato

i. Descripción

Esta Fact involucra todas las matrículas realizadas. La Fact permitirá cumplir con la elaboración de los reportes.

ii. Granularidad

N°	Nombre de la Dimensión	Descripción	Llave Primaria
1	Institución Educativa	Esta dimensión abarca toda la información acerca de la IE, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Código de IE, Código Modalidad IE, Anexo, Nivel, Descripción Modalidad, Nombre de IE, Ubigeo	SI
2	Ubicación	Esta dimensión abarca toda la información acerca de la ubicación, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Ubigeo, Departamento, Provincia, Distrito	SI
3	Tiempo	Esta dimensión abarca toda la información acerca la organización del tiempo en los periodos de análisis, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Año, Semestre, Trimestre, Bimestre, Año Mes, Código Mes, Mes, Día, Estación	SI
4	Modalidad IE	Esta dimensión abarca toda la información acerca de la modalidad de IE, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Código, Modalidad IE, Descripción de modalidad	SI
5	Grado Educativo	Esta dimensión abarca toda la información acerca de Grado Educativo, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Código Grado Educativo, Descripción de Grado Educativo	SI

N°	Nombre de la Dimensión	Descripción	Llave Primaria
6	Nivel Educativo	Esta dimensión abarca toda la información acerca del Nivel Educativo, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Código Nivel Educativo, Descripción Nivel Educativo	SI
7	Sección	Esta dimensión abarca toda la información acerca de las secciones, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Código Sección, Sección	SI
8	Estado Formato	Esta dimensión abarca toda la información acerca del Estado de Formato, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Código Tipo Formato, Tipo Formato	SI
9	Tipo Formato	Esta dimensión abarca toda la información acerca del Tipo Formato, normalizando sus principales atributos. De esta dimensión toma los siguientes atributos: Código Estado Formato, Estado Formato	SI

iii. *Medidas*

N°	Nombre	Descripción
1	% nóminas generadas	Mide la cantidad de nóminas que están en estado generada
2	% nóminas remitida	Mide la cantidad de nóminas que están en estado remitida
3	% nóminas aprobada	Mide la cantidad de nóminas que están en estado aprobado
4	% nóminas Rechazada	Mide la cantidad de nóminas que están en estado rechazada

Como resultado del Análisis dimensional el Datamart cuenta con la siguiente estructura en cuanto a las dimensiones:

Nro.	Campos
1	Código de IE
2	Código Modalidad IE
3	Anexo
4	Nivel
5	Descripción Modalidad
6	Nombre de IE
7	Ubigeo

Nro.	Campos
1	Identificación Estudiante
2	Código Estudiante
3	Nombre de Estudiante
4	Año
5	Nivel
6	Grado
7	Sección
8	Sexo
9	Identificación Estudiante
10	Código Estudiante
11	Nombre de Estudiante

Nro.	Campos
1	Ubigeo
2	Departamento
3	Provincia
4	Distrito

Nro.	Campos
1	Año
2	Semestre
3	Trimestre
4	Bimestre
5	Año Mes
6	Código Mes
7	Mes
8	Día
9	Estación

Nro.	Campos
1	Código Modalidad IE
2	Descripción de modalidad

Nro.	Campos
1	Código Estado de Matrícula
2	Descripción de Matricula

Nro.	Campos
1	Código Grado Educativo
2	Descripción Grado Educativo

Nro.	Campos
1	Código Nivel Educativo
2	Descripción Nivel Educativo

Nro.	Campos
1	Código Estado Estudiante
2	Descripción Estado Estudiante

Nro.	Campos
1	Código Sección
2	Año
3	Nivel
4	Grado
5	Sección
6	Tipo Formato
7	Nombre Institución Educativa
8	Sexo
9	Estado Formato

Nro.	Campos
1	Código Tipo Formato
2	Tipo de Formato

ANEXO 05 DISEÑO DE ETL

Carga de la Dimensión Institución Educativa

f) Descripción

El presente proceso se realiza la extracción del universo de Institución Educativa con el que contará el Datamart, contiene la relación de la información de las Instituciones Educativas.

g) Descripción de los Archivos Fuentes

Tipo de Fuente	Nombre de la Tabla	Descripción
TXT	INSTITUCION_EDUCATIVA	La tabla contiene información acerca de las Instituciones Educativa que se encuentran registrados en el SIAGIE
TXT	PATRON	La tabla contiene información acerca de las ubicaciones de las Instituciones Educativas.

h) Descripción del Archivo Intermedio

Tipo de Fuente	Nombre de la Tabla	Descripción
BD(SQL)	STG_D_INSTITUCION_EDUCATIVA	La tabla contiene información acerca de las Instituciones Educativas que se encuentran registrados en el SIAGIE

i) Estandarización de Datos y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_IE	PK	INTEGER	NUMERICO	NO NULL	NO TIENE
COD_MODULAR		VARCHAR(7)	TEXTO		NO TIENE
ANEXO		VARCHAR(1)	TEXTO		NO TIENE
COD_NIVEL		VARCHAR (2)	TEXTO		NO TIENE
NOMBRE_IE		VARCHAR(100)	TEXTO		NO TIENE
UBIGEO		VARCHAR(6)	TEXTO		NO TIENE

j) Fuente de Datos

Tablas: INSTITUCION EDUCATIVA, PATRON					
Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
COD_MOD	PK	CHAR(7)	TEXTO	NO NULL	NO TIENE
ANEXO	PK	CHAR (1)	TEXTO	NO NULL	NO TIENE
ID_NIVEL		CHAR (2)	TEXTO		NO TIENE
CEN_EDU		VARCHAR (100)	TEXTO		NO TIENE
CODIGEL		CHAR(6)	TEXTO		NO TIENE

k) Tabla Destino Intermedio

Tabla	STG_D_INSTITUCION EDUCATIVA	
Campo	Tipo	Mapeo
ID_IE	INTEGER	AUTOGENERADO
COD_MODULAR	VARCHAR(7)	INSTITUCION EDUCATIVA. COD_MOD, PATRON. COD_MOD
ANEXO	VARCHAR (1)	INSTITUCION EDUCATIVA. ANEXO
COD_NIVEL	CHAR (2)	INSTITUCION EDUCATIVA. ID_NIVEL
NOMBRE_IE	VARCHAR (100)	PATRON. CEN_EDU
UBIGEO	VARCHAR(6)	PATRON. CODIGEL

l) Tabla Destino Final

Tabla	D_INSTITUCION EDUCATIVA	
Campo	Tipo	Mapeo
ID_IE	INTEGER	STG_D_ INSTITUCION EDUCATIVA. ID_IE
COD_MODULAR	VARCHAR(7)	STG_D_ INSTITUCION EDUCATIVA. COD_MODULAR
ANEXO	VARCHAR (1)	STG_D_ INSTITUCION EDUCATIVA. ANEXO
COD_NIVEL	CHAR (2)	STG_D_ INSTITUCION EDUCATIVA. COD_NIVEL
NOMBRE_IE	VARCHAR (100)	STG_D_ INSTITUCION EDUCATIVA. NOMBRE_IE
UBIGEO	VARCHAR(6)	STG_D_ INSTITUCION EDUCATIVA. UBIGEO

Carga de la Dimensión Estudiante

a) Descripción

El presente proceso se realiza la extracción del universo de los estudiantes con el que contará el Datamart, contiene la relación de la información de los estudiantes.

b) Descripción de los Archivos Fuentes

Tipo de Fuente	Nombre de la Tabla	Descripción
TXT	MATRICULA	La tabla contiene información acerca de las matriculas que se encuentran registrados en el SIAGIE
TXT	PERSONA	La tabla contiene información acerca de los estudiantes que se encuentran registrados en el SIAGIE

c) Descripción del Archivo Intermedio

Tipo de Fuente	Nombre de la Tabla	Descripción
BD(SQL)	STG_D_ESTUDIANTE	La tabla contiene información acerca de los estudiantes de las Instituciones Educativas

d) Estandarización de Datos y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_ESTUDIANTE	PK	INTEGER	NUMERICO	NO NULL	NO TIENE
COD_ESTUDIANTE		INTEGER	NUMÉRICO		NO TIENE
COD_NIVEL		VARCHAR(2)	TEXTO		NO TIENE
COD_GRADO		VARCHAR(2)	TEXTO		NO TIENE
COD_EST_MAT		SMALLINT(1)	TEXTO		NO TIENE
COD_EST_EST		VARCHAR(2)	TEXTO		NO TIENE
COD_MODULAR		VARCHAR(7)	TEXTO		NO TIENE
ANEXO		VARCHAR(1)	TEXTO		NO TIENE
ANIO		INTEGER	NUMERICO		NO TIENE
COD_SECCION		VARCHAR(2)	TEXTO		NO TIENE
SEXO		SMALLINT	ENTERO		NO TIENE

e) Fuente de Datos

Tablas: MATRICULA, PERSONA					
Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_PERSONA		INTEGER	NUMERICO		NO TIENE
ID_NIVEL		CHAR(2)	TEXTO		NO TIENE
ID_GRADO		CHAR(2)	TEXTO		NO TIENE
ESTADO_MATRICULA		SMALLINT	ENTERO		NO TIENE
ESTADO_POR_GRADO		VARCHAR(2)	TEXTO		NO TIENE
COD_MOD	PK	CHAR(7)	TEXTO	NO NULL	NO TIENE
ANEXO	PK	CHAR(1)	TEXTO	NO NULL	NO TIENE
ID_ANIO	PK	SMALLINT	ENTERO	NO NULL	NO TIENE
ID_SECCION		CHAR(2)	TEXTO		NO TIENE
SEXO		SMALLINT (1)	ENTERO		NO TIENE

f) Tabla Destino Intermedio

Tabla	STG_D_ESTUDIANTE	
Campo	Tipo	Mapeo
ID_ESTUDIANTE	INTEGER	AUTOGENERADO
COD_ESTUDIANTE	INTEGER	PERSONA.ID_PERSONA, MATRICULA.ID_PERSONA
COD_NIVEL	VARCHAR(2)	MATRICULA.ID_ANIO
COD_GRADO	VARCHAR(2)	NIVEL_EDUCATIVO.DSC_NIVEL
COD_EST_MAT	SMALLINT(1)	MATRICULA.ESTADO_MATRICULA
COD_EST_EST	VARCHAR(2)	MATRICULA.ESTADO_POR_GRADO
COD_MODULAR	VARCHAR(7)	MATRICULA.COD_MOD
ANEXO	VARCHAR(1)	MATRICULA.ANEXO
ANIO	INTEGER	MATRICULA.ID_ANIO
COD_SECCION	VARCHAR(2)	MATRICULA.ID_SECCION
SEXO	SMALLINT	PERSONA.SEXO

g) Tabla Destino Final

Tabla	D_ESTUDIANTE	
Campo	Tipo	Mapeo
ID_ESTUDIANTE	INTEGER	STG_D_ESTUDIANTE.ID_ESTUDIANTE
COD_ESTUDIANTE	INTEGER	STG_D_ESTUDIANTE.COD_ESTUDIANTE
COD_NIVEL	VARCHAR(2)	STG_D_ESTUDIANTE.COD_NIVEL
COD_GRADO	VARCHAR(2)	STG_D_ESTUDIANTE.COD_GRADO
COD_EST_MAT	SMALLINT(1)	STG_D_ESTUDIANTE.COD_EST_MAT
COD_EST_EST	VARCHAR(2)	STG_D_ESTUDIANTE.COD_EST_EST
COD_MODULAR	VARCHAR(7)	STG_D_ESTUDIANTE.COD_MODULAR
ANEXO	VARCHAR(1)	STG_D_ESTUDIANTE.ANEXO
ANIO	INTEGER	STG_D_ESTUDIANTE.ANIO
COD_SECCION	VARCHAR(2)	STG_D_ESTUDIANTE.COD_SECCION
SEXO	SMALLINT	STG_D_ESTUDIANTE.SEXO

Carga de la Dimensión Ubicación

a) Descripción

El presente proceso se realiza la extracción del universo de Institución Educativa con el que contará el Datamart, contiene la relación de la información de las localizaciones.

b) Descripción de los Archivos Fuentes

Tipo de Fuente	Nombre de la Tabla	Descripción
TXT	UBIGEO	La tabla contiene información acerca de la locación geográfica de las Instituciones Educativas
TXT	PADRON	La tabla contiene la localización de las Instituciones Educativas (longitud, latitud)

c) Descripción del Archivo Intermedio

Tipo de Fuente	Nombre de la Tabla	Descripción
BD(SQL)	STG_D_UBICACION	La tabla contiene la ubicación de las Instituciones Educativas

d) Estandarización de Datos y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_UBICACION	PK	INTEGER	NUMERICO	NO NULL	NO TIENE
UBIGEO		INTEGER	NUMÉRICO		NO TIENE
DEPARTAMENTO		VARCHAR(30)	TEXTO		NO TIENE
PROVINCIA		VARCHAR(30)	TEXTO		NO TIENE
DISTRITO		VARCHAR(150)	TEXTO		NO TIENE
LONGITUD		DECIMAL(18,6)	DECIMAL		NO TIENE
LATITUD		DECIMAL(18,6)	DECIMAL		NO TIENE

e) Fuente de Datos

Tabla: UBIGEO, PADRON					
Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
UBIGEO		CHAR(6)	TEXTO		NO TIENE
DEPARTAMENTO		CHAR(30)	TEXTO		NO TIENE
PROVINCIA		CHAR(30)	TEXTO		NO TIENE
DISTRITO		CHAR(150)	TEXTO		NO TIENE
NLAT_IE		DECIMAL(18,6)	DECIMAL		NO TIENE
NLONG_IE		DECIMAL(18,6)	DECIMAL		NO TIENE

f) Tabla Destino Intermedio

Tabla	STG_D_UBICACION	
Campo	Tipo	Mapeo
ID_UBICACION	INTEGER	AUTOGENERADO
UBIGEO	INTEGER	UBIGEO. UBIGEO, PADRON. CODIGEL
DEPARTAMENTO	VARCHAR(30)	UBIGEO. DEPARTAMENTO
PROVINCIA	VARCHAR(30)	UBIGEO. PROVINCIA
DISTRITO	VARCHAR(150)	UBIGEO. DISTRITO
LATITUD	DECIMAL(18,6)	PADRON. NLAT_IE
LONGITUD	DECIMAL(18,6)	PADRON. NLONG_IE

g) Tabla Destino Final

Tabla	D_UBICACION	
Campo	Tipo	Mapeo
ID_UBICACION	INTEGER	STG_D_UBICACION. ID_UBICACION
UBIGEO	INTEGER	STG_D_UBICACION. UBIGEO
DEPARTAMENTO	VARCHAR(30)	STG_D_UBICACION. DEPARTAMENTO
PROVINCIA	VARCHAR(30)	STG_D_UBICACION. PROVINCIA
DISTRITO	VARCHAR(150)	STG_D_UBICACION. DISTRITO
LATITUD	DECIMAL(18,6)	STG_D_UBICACION. LATITUD
LONGITUD	DECIMAL(18,6)	STG_D_UBICACION. LONGITUD

Carga de la Dimensión Tiempo

a) Descripción

El presente proceso se realiza la extracción del universo de los tiempos con el que contará el Datamart. Se carga a partir en la base de datos SQL.

b) Descripción de los Archivos Fuentes

Tipo de Fuente	Nombre de la Tabla	Descripción
Base de Datos SQL	D_TIEMPO	Tabla de tiempo es generada mediante un procedimiento en SQL. Posee todos los días a partir del año 2014.

c) Estandarización de Datos y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_FECHA	PK	INTEGER	NUMÉRICO	NO NULL	NO TIENE
ANIO		INTEGER	NUMÉRICO		NO TIENE
SEMESTRE		INTEGER	NUMÉRICO		NO TIENE
TRIMESTRE		INTEGER	NUMÉRICO		NO TIENE
BIMESTRE		INTEGER	NUMÉRICO		NO TIENE
ANIOMES		INTEGER	NUMÉRICO		NO TIENE
CODMES		INTEGER	NUMÉRICO		NO TIENE
MES		VARCHAR(45)	TEXTO		NO TIENE
DIA		VARCHAR(45)	TEXTO		NO TIENE
ESTACION		VARCHAR(20)	TEXTO		NO TIENE

d) Fuente de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_FECHA	PK	INTEGER	NUMÉRICO	NO NULL	NO TIENE
ANIO		INTEGER	NUMÉRICO		NO TIENE
SEMESTRE		INTEGER	NUMÉRICO		NO TIENE
TRIMESTRE		INTEGER	NUMÉRICO		NO TIENE
BIMESTRE		INTEGER	NUMÉRICO		NO TIENE
ANIOMES		INTEGER	NUMÉRICO		NO TIENE
CODMES		INTEGER	NUMÉRICO		NO TIENE
MES		VARCHAR(45)	TEXTO		NO TIENE
DIA		VARCHAR(45)	TEXTO		NO TIENE
ESTACION		VARCHAR(20)	TEXTO		NO TIENE

e) *Tabla Destino*

Tabla	D_TIEMPO	
Campo	Tipo	Mapeo
ID_FECHA	INTEGER	D_TIEMPO.ID_FECHA
ANIO	INTEGER	D_TIEMPO.ANIO
SEMESTRE	INTEGER	D_TIEMPO.SEMESTRE
TRIMESTRE	INTEGER	D_TIEMPO.TRIMESTRE
BIMESTRE	INTEGER	D_TIEMPO.BIMESTRE
ANIOMES	INTEGER	D_TIEMPO.ANIOMES
CODMES	INTEGER	D_TIEMPO.CODMES
MES	VARCHAR(45)	D_TIEMPO.MES
DIA	VARCHAR(45)	D_TIEMPO.DIA
ESTACION	VARCHAR(20)	D_TIEMPO.ESTACION

Carga de la Dimensión Modalidad Institución Educativa

a) Descripción

El presente proceso se realiza la extracción del universo de la modalidad de Institución Educativa con el que contará el Datamart, contiene la relación de la modalidad de la IE.

b) Descripción de los Archivos Fuentes

Tipo de Fuente	Nombre de la Tabla	Descripción
TXT	MODALIDAD	La tabla contiene información acerca de la modalidad de las Instituciones Educativas

c) Descripción del Archivo Intermedio

Tipo de Fuente	Nombre de la Tabla	Descripción
BD(SQL)	STG_D_MODALIDAD_IE	La tabla contiene información acerca de las modalidades que se encuentran en las Instituciones E en el SIAGIE

d) Estandarización de Datos y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_MODALIDAD	PK	INTEGER	NUMERICO	NO NULL	NO TIENE
COD_MODALIDAD		VARCHAR(2)	TEXTO		NO TIENE
MODALIDAD		VARCHAR(100)	TEXTO		NO TIENE
ABR_MODALIDAD		VARCHAR(10)	TEXTO		NO TIENE

e) Fuente de Datos

Tabla: MODALIDAD					
Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_MODALIDAD	PK	CHAR(2)	TEXTO	NO NULL	NO TIENE
DSC_MODALIDAD		CHAR (30)	TEXTO		NO TIENE
ABR_MODALIDAD		VARCHAR(10)	TEXTO		NO TIENE

f) *Tabla Destino Intermedio*

Tabla	STG_D_MODALIDAD_IE	
Campo	Tipo	Mapeo
ID_MODALIDAD	INTEGER	AUTOGENERADO
COD_MODALIDAD	CHAR(2)	MODALIDAD.ID_MODALIDAD
MODALIDAD	CHAR (30)	MODALIDAD.DSC_MODALIDAD
ABR_MODALIDAD	VARCHAR(10)	MODALIDAD.ABR_MODALIDAD

g) *Tabla Destino Final*

Tabla	D_MODALIDAD_IE	
Campo	Tipo	Mapeo
ID_MODALIDAD	INTEGER	STG_D_MODALIDAD_IE.ID_MODALIDAD
COD_MODALIDAD	CHAR(2)	STG_D_MODALIDAD_IE.COD_MODALIDAD
MODALIDAD	CHAR (30)	STG_D_MODALIDAD_IE.MODALIDAD
ABR_MODALIDAD	VARCHAR(10)	STG_D_MODALIDAD_IE.ABR_MODALIDAD

Carga de la Dimensión Estado Matrícula

a) Descripción

El presente proceso se realiza la extracción del universo del estado de matrícula con el que contará el Datamart, contiene la relación de los estados de matrícula.

b) Descripción de los Archivos Fuentes

Tipo de Fuente	Nombre de la Tabla	Descripción
TXT	ESTADO_MATRICULA	La tabla contiene información acerca del estado de matrícula de las Instituciones Educativas

c) Descripción del Archivo Intermedio

Tipo de Fuente	Nombre de la Tabla	Descripción
BD(SQL)	STG_D_ESTADO_MATRICULA	La tabla contiene información acerca de estado de matrículas en las Instituciones Educativas

d) Estandarización de Datos y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_EST_MAT	PK	INTEGER	NUMERICO	NO NULL	NO TIENE
COD_EST_MAT		SMALLINT	ENTERO		NO TIENE
ESTADO_MATRICULA		VARCHAR(50)	TEXTO		NO TIENE

e) Fuente de Datos

Tabla: ESTADO_MATRICULA					
Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ESTADO_MATRICULA	PK	SMALLINT	ENTERO	NO NULL	NO TIENE
DESC_EM		VARCHAR(50)	TEXTO		NO TIENE

f) *Tabla Destino Intermedio*

Tabla	STG_D_ESTADO_MATRICULA	
Campo	Tipo	Mapeo
ID_EST_MAT	INTEGER	AUTOGENERADO
COD_EST_MAT	SMALLINT	ESTADO_MATRICULA. ESTADO_MATRICULA
ESTADO_MATRICULA	VARCHAR(50)	ESTADO_MATRICULA. DESC_EM

g) *Tabla Destino Final*

Tabla	D_ESTADO_MATRICULA	
Campo	Tipo	Mapeo
ID_EST_MAT	INTEGER	STG_D_ESTUDIANTE. ID_EST_MAT
COD_EST_MAT	SMALLINT	STG_D_ESTUDIANTE. COD_EST_MAT
ESTADO_MATRICULA	VARCHAR(50)	STG_D_ESTUDIANTE. ESTADO_MATRICULA

Carga de la Dimensión Nivel Educativo

a) Descripción

El presente proceso se realiza la extracción del universo de los niveles educativos el que contará el Datamart, contiene la relación de los niveles educativos.

b) Descripción de los Archivos Fuentes

Tipo de Fuente	Nombre de la Tabla	Descripción
TXT	NIVEL_EDUCATIVO	La tabla contiene información acerca de los niveles educativos.
TXT	GRADO	La tabla contiene información de grados de las Instituciones Educativas
TXT	SECCION	La tabla contiene información acerca de las secciones en las instituciones educativas.

c) Descripción del Archivo Intermedio

Tipo de Fuente	Nombre de la Tabla	Descripción
BD(SQL)	STG_D_NIVEL_EDUCATIVO	La tabla contiene información acerca de los niveles educativos en las Instituciones Educativas

d) Estandarización de Datos y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_NIVEL	PK	INTEGER	NUMERICO	NO NULL	NO TIENE
COD_NIVEL		VARCHAR(2)	TEXTO		NO TIENE
NIVEL		VARCHAR(30)	TEXTO		NO TIENE
COD_GRADO		VARCHAR(2)	TEXTO		NO TIENE
GRADO		VARCHAR(30)	TEXTO		NO TIENE
COD_SECCION		VARCHAR(30)	TEXTO		NO TIENE
SECCION		VARCHAR(30)	TEXTO		NO TIENE
COD_MODALIDAD		VARCHAR(2)	TEXTO		NO TIENE
COD_MODALAR		VARCHAR(7)	TEXTO		NO TIENE
ANEXO		VARCHAR(1)	TEXTO		NO TIENE
ANIO		INTEGER	NUMERICO		NO TIENE

e) Fuente de Datos

Tabla: NIVEL EDUCATIVO					
Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_NIVEL	PK	CHAR(2)	TEXTO	NO NULL	NO TIENE
ID_GRADO	PK	CHAR(2)	TEXTO		NO TIENE
DSC_NIVEL	PK	CHAR(30)	TEXTO		NO TIENE
DSC_GRADO		CHAR(30)	TEXTO		NO TIENE
ID_SECCION	PK	CHAR(2)	TEXTO		NO TIENE
DSC_SECCION		CHAR(30)	TEXTO		NO TIENE
COD_MOD		CHAR(7)	TEXTO		NO TIENE
ID_MODALIDAD		CHAR(2)	TEXTO		NO TIENE
ANEXO		CHAR(1)	TEXTO		NO TIENE
ID_ANIO		SMALLINT	ENTERO		NO TIENE

f) Tabla Destino Intermedio

Tabla	STG_D_NIVEL EDUCATIVO	
Campo	Tipo	Mapeo
ID_NIVEL	INTEGER	AUTOGENERADO
COD_NIVEL	CHAR(2)	NIVEL EDUCATIVO. ID_NIVEL,
NIVEL	VARCHAR(30)	NIVEL EDUCATIVO. DSC_NIVEL
COD_GRADO	VARCHAR(2)	GRADO. ID_GRADO
GRADO	VARCHAR(30)	GRADO. DSC_GRADO
COD_SECCION	VARCHAR(30)	SECCION. ID_SECCION
SECCION	VARCHAR(30)	SECCION. DSC_SECCION
COD_MODALIDAD	VARCHAR(2)	NIVEL EDUCATIVO. ID_MODALIDAD
COD_MODULAR	VARCHAR(7)	SECCION. COD_MOD
ANEXO	VARCHAR(1)	SECCION. ANEXO
ANIO	INTEGER	SECCION. ID_ANIO

g) Tabla Destino

Tabla	D_NIVEL EDUCATIVO	
Campo	Tipo	Mapeo
ID_NIVEL	INTEGER	STG_D_NIVEL EDUCATIVO. ID_NIVEL
COD_NIVEL	CHAR(2)	STG_D_NIVEL EDUCATIVO. COD_NIVEL
NIVEL	VARCHAR(30)	STG_D_NIVEL EDUCATIVO. NIVEL
COD_GRADO	VARCHAR(2)	STG_D_NIVEL EDUCATIVO. COD_GRADO
GRADO	VARCHAR(30)	STG_D_NIVEL EDUCATIVO. GRADO
COD_SECCION	VARCHAR(30)	STG_D_NIVEL EDUCATIVO. COD_SECCION
SECCION	VARCHAR(30)	STG_D_NIVEL EDUCATIVO. SECCION
COD_MODALIDAD	VARCHAR(2)	STG_D_NIVEL EDUCATIVO. COD_MODALIDAD
COD_MODULAR	VARCHAR(7)	STG_D_NIVEL EDUCATIVO. COD_MODULAR
ANEXO	VARCHAR(1)	STG_D_NIVEL EDUCATIVO. ANEXO
ANIO	INTEGER	STG_D_NIVEL EDUCATIVO. ANIO

Carga de la Dimensión Estado Estudiante

a) Descripción

El presente proceso se realiza la extracción del universo de estados de estudiantes el que contará el Datamart, contiene la relación de los estados de estudiantes.

b) Descripción de los Archivos Fuentes

Tipo de Fuente	Nombre de la Tabla	Descripción
TXT	ESTADO_POR_GRADO	La tabla contiene información acerca de los estados de estudiantes.

c) Descripción del Archivo Intermedio

Tipo de Fuente	Nombre de la Tabla	Descripción
BD(SQL)	STG_D_ESTADO_ESTUDIANTE	La tabla contiene información acerca del estado del estudiante en las Instituciones Educativas

d) Estandarización de Datos y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_EST_EST	PK	INTEGER	NUMERICO	NO NULL	NO TIENE
COD_EST_EST		CHAR(2)	TEXTO		NO TIENE
ESTADO_ESTUDIANTE		VARCHAR(100)			NO TIENE

e) Fuente de Datos

Tabla: ESTADO_POR_GRADO					
Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ESTADO_POR_GRADO	PK	CHAR(2)	TEXTO	NO NULL	NO TIENE
DESC_EPG		VARCHAR(100)	TEXTO		NO TIENE

f) *Tabla Destino Intermedio*

Tabla	STG_D_ESTADO_ESTUDIANTE	
Campo	Tipo	Mapeo
ID_EST_EST	INTEGER	AUTOGENERADO
COD_EST_EST	CHAR(2)	ESTADO_POR_GRADO. ESTADO_POR_GRADO
ESTADO_ESTUDIANTE	VARCHAR(100)	ESTADO_POR_GRADO. DESC_EPG

g) *Tabla Destino*

Tabla	D_ESTADO_ESTUDIANTE	
Campo	Tipo	Mapeo
ID_EST_EST	INTEGER	STG_D_ESTADO_ESTUDIANTE. ID_EST_EST
COD_EST_EST	CHAR(2)	STG_D_ESTADO_ESTUDIANTE. COD_EST_EST
ESTADO_ESTUDIANTE	VARCHAR(100)	STG_D_ESTADO_ESTUDIANTE. ESTADO_ESTUDIANTE

Carga de la Dimensión Tipo Formato

a) Descripción

El presente proceso se realiza la extracción del universo datos de tipo formato el que contará el Datamart, contiene la relación de los tipos de formato de las Instituciones Educativas.

b) Descripción de los Archivos Fuentes

Tipo de Fuente	Nombre de la Tabla	Descripción
TXT	FORMATO	La tabla contiene información acerca de los tipos de formato

c) Descripción del Archivo Intermedio

Tipo de Fuente	Nombre de la Tabla	Descripción
BD(SQL)	STG_D_TIPO_FORMATO	La tabla contiene información acerca de los tipos de formato de las Instituciones Educativas

d) Estandarización de Datos y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_TIPO	PK	INTEGER	NUMERICO	NO NULL	NO TIENE
COD_TIPO		VARCHAR(2)	TEXTO		NO TIENE
TIPO		VARCHAR(50)	TEXTO		NO TIENE

e) Fuente de Datos

Tabla: FORMATO					
Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_FORMATO	PK	CHAR(2)	TEXTO	NO NULL	NO TIENE
DESC_FORMATO		CHAR(30)	TEXTO		NO TIENE

f) *Tabla Destino Intermedio*

Tabla	STG_D_TIPO_FORMATO	
Campo	Tipo	Mapeo
ID_TIPO	INTEGER	AUTOGENERADO
COD_TIPO	CHAR(2)	ESTADO_FORMATO_SECCION.ID_FORMATO
TIPO	CHAR(30)	ESTADO_FORMATO_SECCION.DESC_FORMATO

g) *Tabla Destino Final*

Tabla	D_TIPO_FORMATO	
Campo	Tipo	Mapeo
ID_TIPO	INTEGER	STG_D_TIPO_FORMATO.ID_TIPO
COD_TIPO	CHAR(2)	STG_D_TIPO_FORMATO.COD_TIPO
TIPO	CHAR(30)	STG_D_TIPO_FORMATO.TIPO

Carga de la Dimensión Estado Formato

a) Descripción

El presente proceso se realiza la extracción del universo datos de estado de formato que contará el Datamart, contiene la relación de los estados de formatos de las Instituciones Educativas.

b) Descripción de tablas Fuentes

Tipo de Fuente	Nombre de la Tabla	Descripción
TXT	ESTADO_FORMATO_SECCION	La tabla contiene información acerca del formato de las secciones

c) Descripción del Archivo Intermedio

Tipo de Fuente	Nombre de la Tabla	Descripción
BD(SQL)	STG_D_ESTADO_FORMATO	La tabla contiene información acerca de los estados de formato de las Instituciones Educativas

d) Estandarización de Datos y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_ESTADO_FORMATO	PK	INTEGER	NUMERICO	NO NULL	NO TIENE
COD_ESTADO		SMALLINT	TEXTO		NO TIENE
ESTADO		VARCHAR(50)	TEXTO		NO TIENE

e) Fuente de Datos

Tabla: ESTADO_FORMATO_SECCION					
Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ESTADO_FORMATO_SECCION	PK	SMALLINT	TEXTO	NO NULL	NO TIENE
DESC_EFS		VARCHAR(50)	TEXTO		NO TIENE

f) *Tabla Destino Intermedio*

Tabla	STG_D_ESTADO_FORMATO	
Campo	Tipo	Mapeo
ID_ESTADO_FORMATO	INTEGER	AUTOGENERADO
COD_ESTADO	SMALLINT	ESTADO_FORMATO_SECCION, ESTADO_FORMATO_SECCION
ESTADO	VARCHAR(50)	ESTADO_FORMATO_SECCION.DESC_EFS

g) *Tabla Destino Final*

Tabla	D_ESTADO_FORMATO	
Campo	Tipo	Mapeo
ID_ESTADO_FORMATO	INTEGER	STG_D_ESTADO_FORMATO.ID_ESTADO_FORMATO
COD_ESTADO	SMALLINT	STG_D_ESTADO_FORMATO.COD_ESTADO
ESTADO	VARCHAR(50)	STG_D_ESTADO_FORMATO.ESTADO

Carga de la Dimensión Detalle Formato

a) Descripción

El presente proceso se realiza la extracción del universo datos del detalle de formato el que contará el Datamart, contiene la relación del detalle de formato de las Instituciones Educativas.

b) Descripción de los Archivos Fuentes

Tipo de Fuente	Nombre de la Tabla	Descripción
CSV	FORMATOS_OFICIALES_SECCION	La tabla contiene información acerca de los formatos oficiales de sección

c) Descripción del Archivo Intermedio

Tipo de Fuente	Nombre de la Tabla	Descripción
BD(SQL)	STG_D_DETALLE_FORMATO	La tabla contiene información acerca del detalle de formato de las Instituciones Educativas

d) Estandarización de Datos y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_DET_FORMATO	PK	INTEGER	NUMERICO	NO NULL	NO TIENE
COD_DET_FORMATO		SMALLINT(1)	ENTERO		NO TIENE
COD_MODULAR		VARCHAR(7)	TEXTO		NO TIENE
ANEXO		VARHCHAR(1)	TEXTO		NO TIENE
ANIO		INTEGER	NUMERICO		NO TIENE
COD_NIVEL		VARCHAR(2)	ENTERO		NO TIENE
COD_GRADO		VARCHAR(2)	TEXTO		NO TIENE
COD_SECCION		VARCHAR(2)	TEXTO		NO TIENE
COD_TIPO		VARCHAR(2)	TEXTO		NO TIENE
COD_ESTADO_FORMATO		SMALLINT(1)	ENTERO		NO TIENE
FECHA_APROBACION		DATETIME	TIEMPO		NO TIENE
FECHA_CREACION		DATETIME	TIEMPO		NO TIENE

e) Fuente de Datos

Tabla: FORMATOS_OFICIALES_SECCION					
Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_FORMATO_SECCION	PK	SMALLINT	ENTERO		NO TIENE
COD_MOD	PK	CHAR(7)	TEXTO		NO TIENE
ANEXO	PK	CHAR(1)	TEXTO		NO TIENE
ID_ANIO	PK	SMALLINT	ENTERO		NO TIENE
ID_NIVEL	PK	CHAR(2)	TEXTO		NO TIENE
ID_GRADO	PK	CHAR(2)	TEXTO		NO TIENE
ID_SECCION	PK	CHAR(2)	TEXTO		NO TIENE
ID_FORMATO	PK	CHAR(2)	TEXTO		NO TIENE
ID_FORMATO_SECCION	PK	SMALLINT	ENTERO		NO TIENE
ESTADO_FORMATO_SECCION		SMALLINT	ENTERO		NO TIENE

f) Tabla Destino Intermedio

Tabla	STG_D_DETALLE_FORMATO	
Campo	Tipo	Mapeo
ID_DET_FORMATO	INTEGER	AUTOGENERADO
COD_DET_FORMATO	SMALLINT(1)	ESTADO_FORMATO_SECCION.ID_FORMATO_SECCION
COD_MODULAR	VARCHAR(7)	ESTADO_FORMATO_SECCION.COD_MOD
ANEXO	VARCHAR(1)	ESTADO_FORMATO_SECCION.ANEXO
ANIO	INTEGER	ESTADO_FORMATO_SECCION.ID_ANIO
COD_NIVEL	VARCHAR(2)	ESTADO_FORMATO_SECCION.ID_NIVEL
COD_GRADO	VARCHAR(2)	ESTADO_FORMATO_SECCION.ID_GRADO
COD_SECCION	VARCHAR(2)	ESTADO_FORMATO_SECCION.ID_SECCION
COD_TIPO	VARCHAR(2)	ESTADO_FORMATO_SECCION.ID_FORMATO
COD_ESTADO_FORMATO	SMALLINT(1)	ESTADO_FORMATO_SECCION.ESTADO_FORMATO_SECCION
FECHA_APROBACION	DATETIME	SYSDATE
FECHA_CREACION	DATETIME	SYSDATE

g) Tabla Destino Final

Tabla	D_DETALLE_FORMATO	
Campo	Tipo	Mapeo
ID_DET_FORMATO	INTEGER	STG_DETALLE_FORMATO.ID_DET_FORMATO
COD_DET_FORMATO	SMALLINT(1)	STG_DETALLE_FORMATO.COD_DET_FORMATO
COD_MODULAR	VARCHAR(7)	STG_DETALLE_FORMATO.COD_MODULAR
ANEXO	INTEGER	STG_DETALLE_FORMATO.ANEXO
ANIO	VARCHAR(1)	STG_DETALLE_FORMATO.ANIO
COD_NIVEL	VARCHAR(2)	STG_DETALLE_FORMATO.COD_NIVEL
COD_GRADO	VARCHAR(2)	STG_DETALLE_FORMATO.COD_GRADO
COD_SECCION	VARCHAR(2)	STG_DETALLE_FORMATO.COD_SECCION
COD_TIPO	VARCHAR(2)	STG_DETALLE_FORMATO.COD_TIPO
COD_ESTADO_FORMATO	SMALLINT(1)	STG_DETALLE_FORMATO.COD_ESTADO_FORMATO
FECHA_APROBACION	DATETIME	STG_DETALLE_FORMATO.FECHA_APROBACION
FECHA_CREACION	DATETIME	STG_DETALLE_FORMATO.FECHA_CREACION

Fact Matricula

i. Descripción

Representa la Extracción, transformación y Carga de la Fact Matricula.

ii. Descripción de Tablas Fuente

Tipo de Fuente	Nombre de Tabla	Descripción
BD(SQL)	D_UBICACION	La tabla contiene la localización de las Instituciones Educativas (longitud, latitud)
BD(SQL)	D_MODALIDAD_IE	La tabla contiene información acerca de las modalidades que se encuentran en las Instituciones E en el SIAGIE
BD(SQL)	D_INSTITUCION_EDUCATIVA	La tabla contiene información acerca de las Instituciones Educativas que se encuentran registrados en el SIAGIE
BD(SQL)	D_ESTADO_MATRICULA	La tabla contiene información acerca de estado de matrículas en las Instituciones Educativas
BD(SQL)	D_NIVEL_EDUCATIVO	La tabla contiene información acerca de los niveles educativos en las Instituciones Educativas
BD(SQL)	D_ESTADO_ESTUDIANTE	La tabla contiene información acerca del estado del estudiante en las Instituciones Educativas
BD(SQL)	D_ESTUDIANTE	La tabla contiene información acerca de los estudiantes de las Instituciones Educativas
BD(SQL)	D_TIEMPO	La tabla contiene información de los periodos de análisis que abarca las Instituciones Educativas.

iii. Estandarización y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_UBICACION	PK	INTEGER	NUMÉRICO	No debe ser nulos	NULL
ID_MODALIDAD	PK	INTEGER	NUMÉRICO	No debe ser nulos	NULL
ID_IE	PK	INTEGER	NUMÉRICO	No debe ser nulos	NULL
ID_MATRICULADO	PK	INTEGER	NUMÉRICO	No debe ser nulos	NULL
ID_ESTADO_MAT	PK	INTEGER	NUMÉRICO	No debe ser nulos	NULL
ID_NIVEL	PK	INTEGER	NUMÉRICO	No debe ser nulos	NULL
ID_ESTADO_EST	PK	INTEGER	NUMÉRICO	No debe ser nulos	NULL
CANTIDAD_NOMINAS		INTEGER	NUMÉRICO	No debe ser nulos	NULL
PORCENTAJE_NOMINAS		DECIMAL	DECIMAL	No debe ser nulos	NULL

iv. Fuente de Datos

Tabla: D_UBICACION, D_MODALIDAD_IE, D_INSTITUCION_EDUCATIVA, D_ESTADO_MATRICULA, D_NIVEL_EDUCATIVO, D_ESTADO_ESTUDIANTE, D_ESTUDIANTE				
Nombre	Llave	Tipo	Formato	Consideración Importante
ID_UBICACION	PK	INTEGER	NUMÉRICO	NO TIENE
ID_MODALIDAD	PK	INTEGER	NUMÉRICO	NO TIENE
ID_IE	PK	INTEGER	NUMÉRICO	NO TIENE
ID_MATRICULADO	PK	INTEGER	NUMÉRICO	NO TIENE
ID_ESTADO_MAT	PK	INTEGER	NUMÉRICO	NO TIENE
ID_NIVEL	PK	INTEGER	NUMÉRICO	NO TIENE
ID_ESTADO_EST	PK	INTEGER	NUMÉRICO	NO TIENE

v. Tablas Destino

Tabla	F_ MATRICULA	
Campo	Tipo	Mapeo
ID_UBICACION	INTEGER	D_UBICACION.ID_UBICACION
ID_MODALIDAD	INTEGER	D_MODALIDAD.ID_MODALIDAD
ID_IE	INTEGER	D_INSTITUCION_EDUCATIVA.ID_IE
ID_MATRICULADO	INTEGER	D_ESTUDIANTE.ID_MATRICULADO
ID_ESTADO_MAT	INTEGER	D_ESTADO_MATRICULA.ID_ESTADO_MAT
ID_NIVEL	INTEGER	D_NIVEL_EDUCATIVO.ID_NIVEL
ID_ESTADO_EST	INTEGER	D_ESTADO_ESTUDIANTE.ID_ESTADO_EST
CANTIDAD_NOMINAS	INTEGER	NO APLICA
PORCENTAJE_NOMINAS	DECIMAL	NO APLICA

Facts de Formato

i. Descripción

Representa la Extracción, transformación y Carga de la Fact Matricula.

ii. Descripción de las Tablas Fuentes

Tipo de Fuente	Nombre de Tabla	Descripción
BD(SQL)	D_UBICACION	La tabla contiene la ubicación de las Instituciones Educativas
BD(SQL)	D_MODALIDAD_IE	La tabla contiene información acerca de las modalidades que se encuentran en las Instituciones E en el SIAGIE
BD(SQL)	D_INSTITUCION_EDUCATIVA	La tabla contiene información acerca de las Instituciones Educativas que se encuentran registrados en el SIAGIE
BD(SQL)	D_NIVEL_EDUCATIVO	La tabla contiene información acerca de los niveles educativos en las Instituciones Educativas
BD(SQL)	D_TIPO_FORMATO	La tabla contiene información acerca de los tipos de formato de las Instituciones Educativas
BD(SQL)	D_ESTADO_FORMATO	La tabla contiene información acerca de los estados de formato de las Instituciones Educativas
BD(SQL)	D_DETALLE_FORMATO	La tabla contiene información acerca del detalle de formato de las Instituciones Educativas
BD(SQL)	D_TIEMPO	La tabla contiene información de los periodos de análisis que abarca las Instituciones Educativas.

iii. Estandarización y Limpieza de Datos

Nombre	Llave	Tipo	Formato	Limpieza	Valor por Defecto
ID_DET_FORMATO	PK	INTEGER	NUMERICO	No debe ser nulos	NULL
ID_IE	PK	INTEGER	NUMERICO	No debe ser nulos	NULL
ID_NIVEL	PK	INTEGER	NUMERICO	No debe ser nulos	NULL
ID_TIPO	PK	INTEGER	NUMERICO	No debe ser nulos	NULL
ID_ESTADO_FORMATO	PK	INTEGER	NUMERICO	No debe ser nulos	NULL
ID_MODALIDAD	PK	INTEGER	NUMERICO	No debe ser nulos	NULL
ID_UBICACION	PK	INTEGER	NUMERICO	No debe ser nulos	NULL
ID_FECHA	PK	INTEGER	NUMERICO	No debe ser nulos	NULL
ID_FEC_CREA	PK	INTEGER	NUMERICO	No debe ser nulos	NULL
ID_FEC_APRO	PK	INTEGER	NUMERICO	No debe ser nulos	NULL
CANTIDAD_NOMINAS		INTEGER	NUMERICO	No debe ser nulos	NULL
PORCENTAJE_NOMINAS		DECIMAL	DECIMAL	No debe ser nulos	NULL

iv. Fuente de Datos

Tabla: D_UBICACION, D_MODALIDAD_IE, D_INSTITUCION_EDUCATIVA, D_NIVEL_EDUCATIVO, D_TIPO_FORMATO, D_ESTADO_FORMATO, D_DETALLE_FORMATO				
Nombre	Llave	Tipo	Formato	Consideración Importante
ID_DET_FORMATO	PK	INTEGER	NUMERICO	NO TIENE
ID_IE	PK	INTEGER	NUMERICO	NO TIENE
ID_NIVEL	PK	INTEGER	NUMERICO	NO TIENE
ID_TIPO	PK	INTEGER	NUMERICO	NO TIENE
ID_ESTADO_FORMATO	PK	INTEGER	NUMERICO	NO TIENE
ID_MODALIDAD	PK	INTEGER	NUMERICO	NO TIENE
ID_UBICACION	PK	INTEGER	NUMERICO	NO TIENE
ID_FECHA	PK	INTEGER	NUMERICO	NO TIENE
ID_FEC_CREA	PK	INTEGER	NUMERICO	NO TIENE
ID_FEC_APRO	PK	INTEGER	NUMERICO	NO TIENE

v. Tablas Destino

Tabla	F_FORMATO	
Campo	Tipo	Mapeo
ID_DET_FORMATO	INTEGER	D_DETALLE_FORMATO.ID_DET_FORMATO
ID_IE	INTEGER	D_INSTITUCION_EDUCATIVA.ID_IE
ID_NIVEL	INTEGER	D_NIVEL_EDUCATIVO.ID_NIVEL
ID_TIPO	INTEGER	D_TIPO_FORMATO.ID_TIPO
ID_ESTADO_FORMATO	INTEGER	D_ESTADO_FORMATO.ID_ESTADO_FORMATO
ID_MODALIDAD	INTEGER	D_MODALIDAD_IE.ID_MODALIDAD
ID_UBICACION	INTEGER	D_UBICACION_EDUCATIVA.ID_UBICACION
ID_FECHA	INTEGER	D_TIEMPO.ID_FECHA
ID_FEC_CREA	INTEGER	D_TIEMPO.ID_FEC_CREA
ID_FEC_APRO	INTEGER	D_TIEMPO.ID_FEC_APRO
CANTIDAD_NOMINAS	INTEGER	NO APLICA
PORCENTAJE_NOMINAS	DECIMAL	NO APLICA

ANEXO 06 ESTIMACIÓN PRELIMINAR DE PRUEBAS

Proyecto Implementación Datamart- SPE MINEDU Proceso de Matrícula

Rol asignado: Analista de Pruebas

Secuencia	Actividad	Duración Fracción de Hora
1	Realizar la carga de datos a través del archivo fuente- SIAGIE	4.50
2	Generar los reportes del Proceso de Matrícula.	2.00
3	Validar los Indicadores del Proceso de Matrícula.	3.00
4	Elaborar el resumen de las operaciones realizadas.	0.50

Total Horas	10.00
--------------------	--------------

Consideraciones:

La Oficina de Informática (OFIN) de MINEDU gestionará la entrega del archivo fuente del SIAGIE.
Se considera verificar que las transacciones culminen correctamente.
Se considera ejecutar las operaciones indicadas en la Matriz de Casos de Prueba.

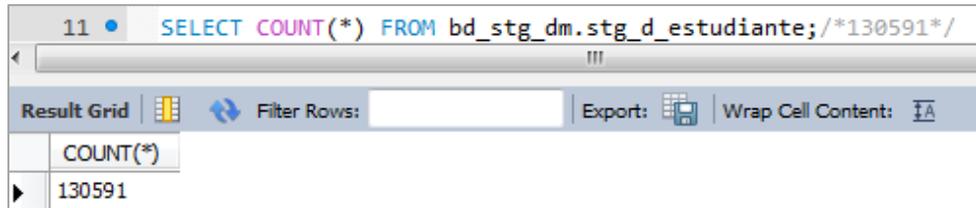
Matriz de Casos de Prueba

CP	Ambiente	Caso de Prueba	Descripcion del Caso	Indicadores	Resultado Esperado	Fecha	Hora	Usuario	Ejecutado
1	Datamart	Carga de Datamart	1. Realizar la carga de la data.		-La carga de data se realiza satisfactoriamente en	27/10/2014	08:30	AP	OK
2	Tableau- Dashboard	Generar el Reporte de estado matricula	1. Ingresar al aplicativo Tableau- Dashboard. 2. Ingresar los parámetros para generar el reporte. 3. Generar el reporte de Matricula en educación básica.	% pre matriculados % matricula en proceso % matricula definitiva % matricula anulada % matricula según genero	-La generación del reporte es satisfactoria. -El tiempo de generación del reporte está en el rango de tiempo aceptable.	27/10/2014	13:03	AP	OK
3	Tableau- Dashboard	Generar Reporte de Nóminas	1. Ingresar al aplicativo Tableau- Dashboard. 2. Ingresar los parámetros para generar el reporte. 3. Generar el reporte de Matricula en educación básica.	% nóminas Generada % nóminas Remitida % nóminas Aprobada % nóminas Rechazada	-La generación del reporte es satisfactoria. -El tiempo de generación del reporte está en el rango de tiempo aceptable.	27/10/2014	13:21	AP	OK
4	Tableau- Dashboard	Generar Reporte de estado de estudiante	1. Ingresar al aplicativo Tableau- Dashboard. 2. Ingresar los parámetros para generar el reporte. 3. Generar el reporte de Matricula en educación básica.	% estado del estudiante	-La generación del reporte es satisfactoria. -El tiempo de generación del reporte está en el rango de tiempo aceptable.	27/10/2014	13:37	AP	OK
5	Tableau- Dashboard	Generar Reporte de estudiantes por nivel educativo.	1. Ingresar al aplicativo Tableau- Dashboard. 2. Ingresar los parámetros para generar el reporte. 3. Generar el reporte de Matricula en educación básica.	% Estudiantes por Nivel Educativo.	-La generación del reporte es satisfactoria. -El tiempo de generación del reporte está en el rango de tiempo aceptable.	27/10/2014	13:46	AP	OK
6	Tableau- Dashboard	Generar Reporte de número de estudiantes	1. Ingresar al aplicativo Tableau- Dashboard. 2. Ingresar los parámetros para generar el reporte. 3. Generar el reporte de Matricula en educación básica.	Número de Estudiantes por Aula	-La generación del reporte es satisfactoria. -El tiempo de generación del reporte está en el rango de tiempo aceptable.	27/10/2014	14:05	AP	OK
7	Tableau- Dashboard	Generar Reporte de participación de estudiantes por	1. Ingresar al aplicativo Tableau- Dashboard. 2. Ingresar los parámetros para generar el reporte. 3. Generar el reporte de Matricula en educación básica.	% participación de estudiantes por género en el sistema educativo	-La generación del reporte es satisfactoria. -El tiempo de generación del reporte está en el rango de tiempo aceptable.	27/10/2014	14:05	AP	OK
8	Tableau- Dashboard	Generar Reporte de estado por grado.	1. Ingresar al aplicativo Tableau- Dashboard. 2. Ingresar los parámetros para generar el reporte. 3. Generar el reporte de Matricula en educación básica.	% estado de matricula por grado	-La generación del reporte es satisfactoria. -El tiempo de generación del reporte está en el rango de tiempo aceptable.	27/10/2014	14:05	AP	OK

ANEXO 07 REGISTROS DEL CUADRE DE CONTEO

A continuación se muestra la carga de la Fact Matrícula:

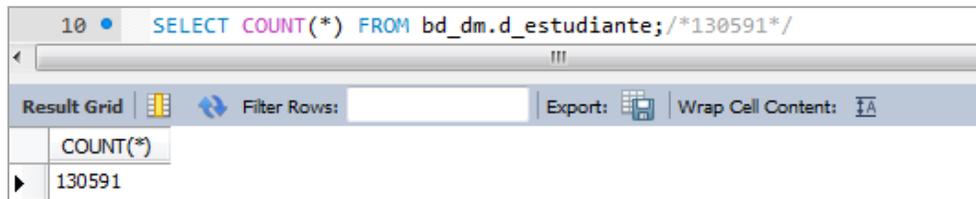
- Cantidad de Registros en Tabla fuente: stg_d_estudiante



The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_stg_dm.stg_d_estudiante; /*130591*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '130591'.

COUNT(*)
130591

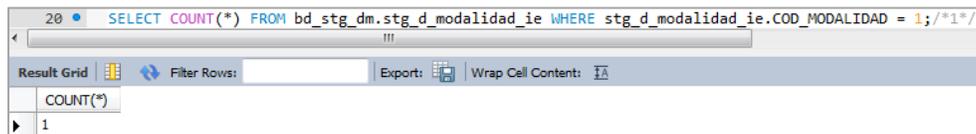
- Cantidad de Registros en Tabla destino: d_estudiante



The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_dm.d_estudiante; /*130591*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '130591'.

COUNT(*)
130591

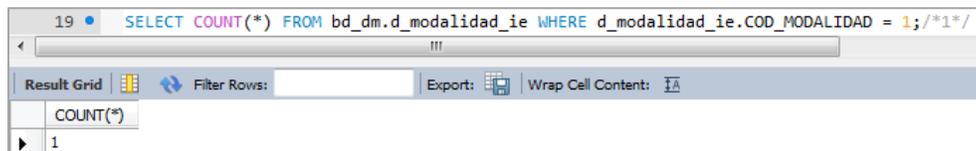
- Cantidad de Registros en Tabla fuente: stg_d_modalidad_ie



The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_stg_dm.stg_d_modalidad_ie WHERE stg_d_modalidad_ie.COD_MODALIDAD = 1; /*1*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '1'.

COUNT(*)
1

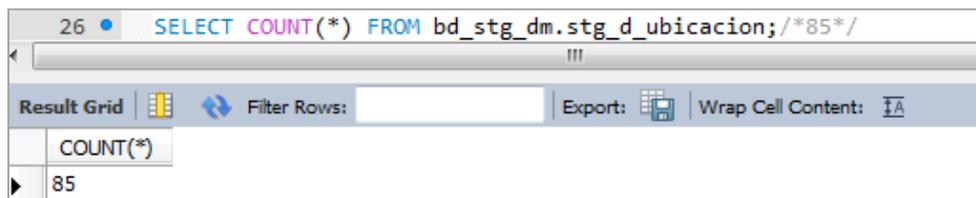
- Cantidad de Registros en Tabla destino: d_modalidad_ie



The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_dm.d_modalidad_ie WHERE d_modalidad_ie.COD_MODALIDAD = 1; /*1*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '1'.

COUNT(*)
1

- Cantidad de Registros en Tabla fuente: stg_d_ubicacion



The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_stg_dm.stg_d_ubicacion; /*85*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '85'.

COUNT(*)
85

- Cantidad de Registros en Tabla destino: d_ubicación

The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_dm.d_ubicacion; /*85*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '85'.

COUNT(*)
85

- Cantidad de Registros en Tabla fuente: stg_d_institucion_educativa

The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_stg_dm.stg_d_institucion_educativa; /*3277*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '3277'.

COUNT(*)
3277

- Cantidad de Registros en Tabla destino: d_institucion_educativa

The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_dm.d_institucion_educativa; /*3277*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '3277'.

COUNT(*)
3277

- Cantidad de Registros en Tabla fuente: stg_d_estado_matricula

The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_stg_dm.stg_d_estado_matricula; /*6*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '6'.

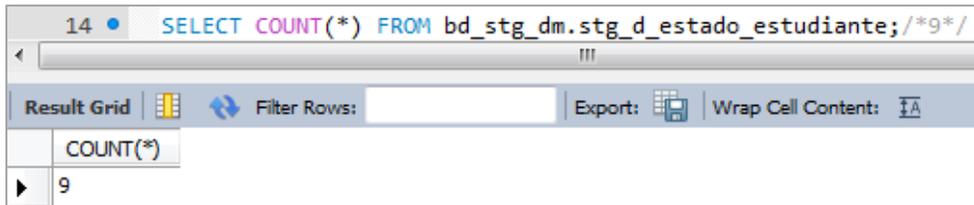
COUNT(*)
6

- Cantidad de Registros en Tabla destino: d_estado_matricula

The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_dm.d_estado_matricula; /*6*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '6'.

COUNT(*)
6

- Cantidad de Registros en Tabla fuente: stg_d_estado_estudiante

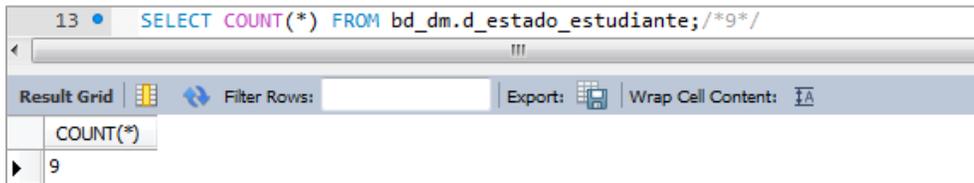


14 • SELECT COUNT(*) FROM bd_stg_dm.stg_d_estado_estudiante; /*9*/

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

COUNT(*)
9

- Cantidad de Registros en Tabla destino: d_estado_estudiante

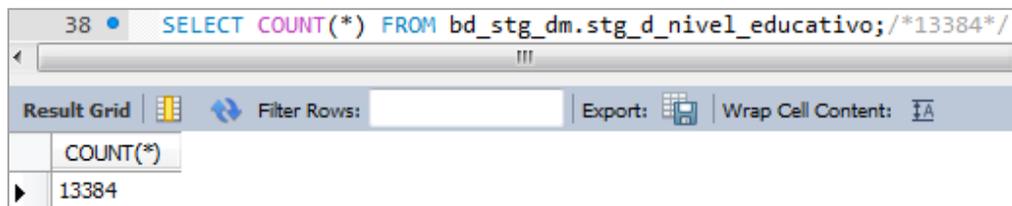


13 • SELECT COUNT(*) FROM bd_dm.d_estado_estudiante; /*9*/

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

COUNT(*)
9

- Cantidad de Registros en Tabla fuente: stg_d_nivel_educativo

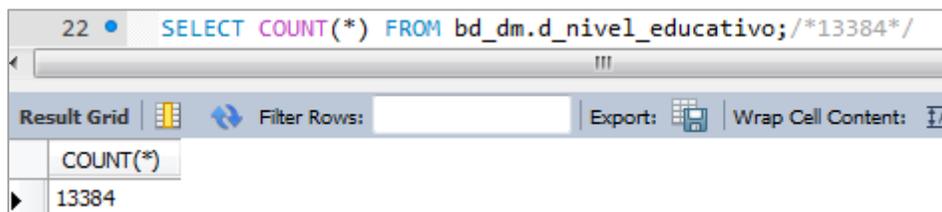


38 • SELECT COUNT(*) FROM bd_stg_dm.stg_d_nivel_educativo; /*13384*/

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

COUNT(*)
13384

- Cantidad de Registros en Tabla destino: d_nivel_educativo



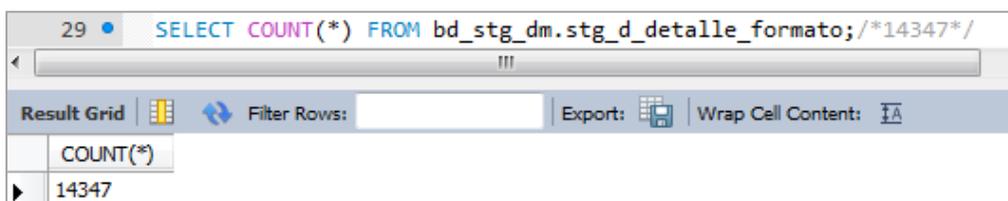
22 • SELECT COUNT(*) FROM bd_dm.d_nivel_educativo; /*13384*/

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

COUNT(*)
13384

A continuación se muestra la carga de la Fact Formato:

- Cantidad de Registros en Tabla fuente: stg_d_detalle_formato



29 • SELECT COUNT(*) FROM bd_stg_dm.stg_d_detalle_formato; /*14347*/

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

COUNT(*)
14347

- Cantidad de Registros en Tabla destino: d_detalle_formato

The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_dm.d_detalle_formato; /*14347*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '14347'.

COUNT(*)
14347

- Cantidad de Registros en Tabla fuente: stg_d_estado_formato

The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_stg_dm.stg_d_estado_formato; /*5*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '5'.

COUNT(*)
5

- Cantidad de Registros en Tabla destino: d_estado_formato

The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_dm.d_estado_formato; /*5*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '5'.

COUNT(*)
5

- Cantidad de Registros en Tabla fuente: stg_d_tipo_formato

The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_stg_dm.stg_d_tipo_formato; /*4*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '4'.

COUNT(*)
4

- Cantidad de Registros en Tabla destino: d_tipo_formato

The screenshot shows a SQL query window with the following text: `SELECT COUNT(*) FROM bd_dm.d_tipo_formato; /*4*/`. Below the query, there is a toolbar with 'Result Grid', 'Filter Rows', 'Export', and 'Wrap Cell Content' options. The result grid displays a single row with the column header 'COUNT(*)' and the value '4'.

COUNT(*)
4

ANEXO 08 CONSULTA DE DIMENSIONES

Luego de haber creado las Dimensiones se realizaron consultas para poder verificar que la data no tuvo alteraciones en el proceso de carga.

- Consulta TOP5 Dimensión D_ESTADO_ESTUDIANTE

Result Grid			
Filter Rows: <input type="text"/>			
	ID_EST_EST	COD_EST_EST	ESTADO_ESTUDIANTE
▶	1	0	MATRICULADO
	2	1	APROBADO
	3	2	DESAPROBADO
	4	3	REQUIERE
	5	4	RETIRADO

- Consulta TOP5 Dimensión D_ESTADO_MATRICULA

Result Grid			
Filter Rows: <input type="text"/>			
	ID_EST_MAT	COD_EST_MAT	ESTADO_MATRICULA
▶	1	0	PRE MATRICULA
	2	1	EN PROCESO
	3	2	ANULADO
	4	3	DEFINITIVA
	5	4	CON RETORNO

- Consulta TOP5 Dimensión D_ESTUDIANTE

Result Grid											
Filter Rows: <input type="text"/>											
Edit: <input type="text"/>											
Export/Import: <input type="text"/>											
Wrap Cell Content: <input type="text"/>											
Fetch rows: <input type="text"/>											
	ID_ESTUDIANTE	COD_ESTUDIANTE	COD_NIVEL	COD_GRADO	COD_EST_MAT	COD_EST_EST	COD_MODULAR	ANEXO	ANIO	COD_SECCION	SEXO
▶	1	17757	F0	12	3	0	0926725	0	2014	03	1
	2	19075	F0	11	1	4	0522672	0	2014	01	1
	3	82268	F0	10	1	4	0913913	0	2014	01	1
	4	84781	F0	13	3	0	0926725	0	2014	01	1
	5	90391	F0	14	3	0	0504662	0	2014	02	1

- Consulta TOP5 Dimensión D_IE

ID_IE	COD_MODALAR	ANEXO	COD_NIVEL	NOMBRE_IE	UBIGEO
1	0223362	0	B0	16581	010701
2	0223370	0	B0	16582	010702
3	0223396	0	B0	16584	010202
4	0223412	0	B0	16586	010202
5	0223438	0	B0	16588	010202

- Consulta TOP5 Dimensión D_MODALIDAD

ID_MODALIDAD	COD_MODALIDAD	MODALIDAD	ABR_MODALIDAD
1	01	Educación Básica Regular	EBR
2	02	Educación Básica	EB
3	03	EDUCACIÓN BÁSICA ALTERNATIVA	EBA
4	04	Educación Básica Especial	EBE

- Consulta TOP5 Dimensión D_NIVEL_EDUCATIVO

ID_NIVEL	COD_NIVEL	NIVEL	COD_GRADO	GRADO	COD_SECCION	SECCION	COD_MODALIDAD	COD_MODALAR	ANEXO	ANIO
1	B0	Primaria	04	PRIMERO	01	A	01	0223362	0	2014
2	B0	Primaria	05	SEGUNDO	01	A	01	0223362	0	2014
3	B0	Primaria	06	TERCERO	01	A	01	0223362	0	2014
4	B0	Primaria	07	CUARTO	01	A	01	0223362	0	2014
5	B0	Primaria	08	QUINTO	01	A	01	0223362	0	2014

- Consulta TOP5 Dimensión D_UBICACION

ID_UBICACION	UBIGEO	DEPARTAMENTO	PROVINCIA	DISTRITO	LONGITUD	LATITUD
1	010101	AMAZONAS	CHACHAPOYAS	CHACHAPOYAS	-77.818380	-6.256220
2	010102	AMAZONAS	CHACHAPOYAS	ASUNCION	-77.710510	-6.032190
3	010103	AMAZONAS	CHACHAPOYAS	BALSAS	-77.919130	-6.787170
4	010104	AMAZONAS	CHACHAPOYAS	CHETO	-77.694420	-6.289510
5	010105	AMAZONAS	CHACHAPOYAS	CHILIQVIN	-77.724080	-6.114030

- Consulta TOP5 Dimensión D_DETALLE_FORMATO

COD_DET_FORMATO	COD_MODULAR	ANEXO	ANIO	COD_NIVEL	COD_GRADO	COD_SECCION	COD_TIPO	COD_ESTADO_FORMATO	FECHA_APROBACION	FECHA_CREACION
1	0223362	0	2014	B0	04	01	01	1	2014-06-04 15:34:53	2014-05-28 05:31:34
1	0223362	0	2014	B0	05	01	01	1	2014-06-04 15:36:24	2014-05-28 05:32:16
1	0223362	0	2014	B0	06	01	01	1	2014-06-04 15:36:51	2014-05-28 05:32:50
1	0223362	0	2014	B0	07	01	01	1	2014-06-04 15:37:22	2014-05-28 05:33:22
1	0223362	0	2014	B0	08	01	01	1	2014-06-04 15:38:05	2014-05-28 05:34:00

- Consulta TOP5 Dimensión D_ESTADO_FORMATO

COD_ESTADO_FORMATO	ESTADO_FORMATO
0	GENERADO
1	APROBADO
2	RECHAZADO
3	ANULADO
4	REMITIDO

- Consulta TOP5 Dimensión D_MODALIDAD_IE

COD_MODALIDAD	MODALIDAD	ABR_MODALIDAD
01	Educación Básica Regular	EBR
02	Educación Básica	EB
03	EDUCACIÓN BÁSICA ALTERNATIVA	EBA
04	Educación Básica Especial	EBE

- Consulta TOP5 Dimensión D_NIVEL_EDUCATIVO

COD_NIVEL	NIVEL	COD_GRADO	GRADO	COD_SECCION	SECCION	COD_MODALIDAD	COD_MODULAR	ANEXO	ANIO
B0	Primaria	04	PRIMERO	01	A	01	0223362	0	2014
B0	Primaria	05	SEGUNDO	01	A	01	0223362	0	2014
B0	Primaria	06	TERCERO	01	A	01	0223362	0	2014
B0	Primaria	07	CUARTO	01	A	01	0223362	0	2014
B0	Primaria	08	QUINTO	01	A	01	0223362	0	2014

- Consulta TOP5 Dimensión D_TIPO_FORMATO

COD_TIPO	TIPO
01	NOMINAS
02	NOMINAS ADICIONAL
03	FORMATO OFICIAL DE ACTA
04	RECUPERACION

ANEXO 09 ACTA DE CONSTITUCIÓN DEL PROYECTO

Acta de Constitución del Proyecto (Project Charter)

A. Información General

Nombre del Proyecto	Análisis, Diseño e Implementación de Datamart para la generación de reportes secretaría planificación estratégica MINEDU.	Fecha de Preparación	11/08/2014
Patrocinador:	Jefe Secretaría de Planificación Estratégica.	Fecha de Modificación:	13/08/2014
Preparado por:	Maryury García A. Karla Jiménez G.	Autorizado por:	Gestor Proyecto

B. Descripción del producto o servicio del Proyecto

El proyecto consiste en generar los reportes de gestión en las Instituciones Educativas de educación básica regular que apoyan la toma de decisiones de la Secretaría de Planificación Estratégica del MINEDU correspondientes al proceso de matrícula y el proceso de nóminas a nivel departamental.

C. Alineamiento del Proyecto

Objetivos de la Organización <small>(Objetivos que tiene la Organización o del cliente)</small>	Propósitos del Proyecto
Mejorar la toma de decisiones del proceso de matrículas y nóminas en las instituciones educativas de Educación Básica Regular.	Se brindará reportes intuitivos que apoye a la toma de decisiones.
	Se implementara una solución tecnológica que brindará reportes en fracciones de tiempos.

D. Objetivos del Proyecto

Objetivos del Proyecto
<p>Objetivo general</p> <p>Disponer en forma oportuna los reportes de gestión en las Instituciones Educativas de educación básica regular que apoyan la toma de decisiones de la Secretaría de Planificación Estratégica del MINEDU correspondientes al proceso de matrícula y el proceso de nóminas a nivel departamental.</p>
<p>Objetivos específicos</p> <ul style="list-style-type: none">-Implementar un Datamart para generar los reportes de gestión oportuna para la toma de decisiones de la Secretaría de Planificación Estratégica del MINEDU.-Identificar y seleccionar las herramientas existentes para la extracción y explotación de datos.-Construir los procesos ETL para la Extracción, Transformación y Carga de datos de nóminas de matrículas desde el SIAGIE.

E. Alcance y Extensión del Proyecto

<p>Fases del Proyecto:</p> <ul style="list-style-type: none">- Gestión del Proyecto- Evaluación del Negocio- Definición de Requerimientos- Análisis de Datos- Prototipo de la Aplicación- Diseño de base de datos- Diseño del ETL- Desarrollo de ETL- Desarrollo de la aplicación- Certificación- Implementación- Evaluación Postproducción
<p>Limitaciones del Proyecto</p> <p>-No se realizará cambios después de haberse realizado la fase de Certificación.</p>
<p>Supuestos del Proyecto</p> <ul style="list-style-type: none">-Contamos con el apoyo de la OFIN.-Contaremos con recursos humanos multidisciplinarios.

F. Cronograma del Proyecto

Se muestra el diagrama de actividades para el presente proyecto. La ampliación de esta imagen se encuentra en el Anexo 01 Cronograma.

EDT	Nombre de tarea	Trabajo	Comienzo	Fin
1	<input type="checkbox"/> Proyecto BI DM MINEDU	437 horas	sáb 09/08/14	sáb 25/10/14
1.1	<input type="checkbox"/> Gestión del Proyecto	85 horas	sáb 09/08/14	vie 24/10/14
1.1.1	<input type="checkbox"/> Inicio y Planificación del Proyecto	4 horas	sáb 09/08/14	mié 13/08/14
1.1.1.1	Planificación del alcance del Proyecto	0.5 horas	lun 11/08/14	mar 12/08/14
1.1.1.2	Definición de los objetivos del proyecto	0.5 horas	sáb 09/08/14	lun 11/08/14
1.1.1.3	Aseguramiento de los recursos	0.5 horas	lun 11/08/14	mar 12/08/14
1.1.1.4	Definición de Entregables	0.5 horas	mar 12/08/14	mar 12/08/14
1.1.1.5	Cronograma del Proyecto	0.5 horas	mar 12/08/14	mié 13/08/14
1.1.1.6	Definición de Actividades a Realizar	0.5 horas	mié 13/08/14	mié 13/08/14
1.1.1.7	Elaboración del Cronograma del Proyecto	0.5 horas	lun 11/08/14	mar 12/08/14
1.1.1.8	Identificación de los Riesgos del Proyecto	0.5 horas	lun 11/08/14	mar 12/08/14
1.1.2	<input type="checkbox"/> Ejecución Seguimiento y Control	80 horas	lun 11/08/14	mié 22/10/14
1.1.2.1	Coordinar los recursos	0 horas	mar 21/10/14	mié 22/10/14
1.1.2.2	Realizar el aseguramiento de los entregables	0 horas	mar 21/10/14	mié 22/10/14
1.1.2.3	Reuniones de Controles de Cambio	0 horas	mar 21/10/14	mié 22/10/14
1.1.2.4	<input checked="" type="checkbox"/> Seguimiento	80 horas	lun 11/08/14	mié 22/10/14
1.1.3	<input type="checkbox"/> Cierre del Proyecto	1 hora	jue 23/10/14	vie 24/10/14
1.1.3.1	Cierre de las actividades según plan de proyecto	1 hora	jue 23/10/14	vie 24/10/14
1.1.3.2	Cierre del Proyecto	0 horas	vie 24/10/14	vie 24/10/14

Elaboración: las autoras

G. Planeamiento Inicial del Proyecto

Estimación de recursos requeridos:

- 1 Gestor de Proyecto
- 1 Analista Funcional
- 1 Diseñador del modelo
- 1 Analista Técnico
- 1 Certificador

Costo Estimado del Proyecto:

62,553.26 PEN como importe estimado. (Valor del servicio que se ofrece al cliente). Completar satisfactoriamente el proyecto. Concluir el proyecto en el tiempo de la fecha establecida: 25 de Octubre del 2014.

Beneficios Esperados:

- Lograr que el cliente desee nuevamente nuestros servicios.
- Realizar negociaciones con futuros clientes.
- Optimizar e incrementar el desarrollo de contratos y convenios, para mejorar el desarrollo del rubro.
- Maximizar el tiempo de producción para obtener así un stock.

Estimación de Fechas a Programar:

- Fecha de inicio del proyecto: Se inicia el 09 de Agosto del 2014.
- Fecha de término: Se estima concluir en 25 de Octubre del 2014.

H. Autoridad del Proyecto

- **Autorización**

Ing. Manuel Cok Aparcana – Gerente de Oficina de Informática
Jorge Mesinas – Representante de la Secretaría de Planificación Estratégica

- **Gestor de Proyecto**

Maryury García Anticona – Gestor de Proyecto

- **Comité de Seguimiento (Dirección)**

Ing. Manuel Cok Aparcana – Gerente de Oficina de Informática.

I. Integrantes del equipo del proyecto, Roles y Responsabilidades

EQUIPO BASE

Gestor de Proyecto:

Encargado que realizará el seguimiento del proyecto, los documentos que se requieren.

Analista Funcional:

Encargado que realizará el levantamiento de información y requisitos.

Diseñador del Modelo:

Encargado que realizará el modelo relacional, las reglas de negocio mediante las cuales se rige la Información.

Analista Técnico:

Encargado que realizará la evaluación y soporte.

Analista Programador:

Encargado que realizará la codificación de los requerimientos solicitados de acuerdo a las definiciones funcionales y técnicas alcanzadas.

Certificador:

Encargado que realizará los casos de prueba para validad que lo construido este de acuerdo a las definiciones funcionales.

Líder Usuario:

Provee las necesidades, los detalles y alcances de las mismas.

ANEXO 10 SOLICITUD DE CONTROL DE CAMBIOS

CONTROL DE CAMBIOS DEL ALCANCE DEL PROYECTO			
NOMBRE DEL PROYECTO: Análisis, Diseño e Implementación de Datamart para la generación de reportes de la Secretaría de Planificación Estratégica del MINEDU			
FECHA	16/10/2014	REQUERIMIENTOS DE CAMBIOS No.	01
INFORMACIÓN GENERAL DEL PROYECTO			
NOMBRE DE LA (S) PERSONA (S) QUE SOLICITA EL CAMBIO		Jorge Mesinas	
NOMBRE DE LA (S) PERSONA (S) QUE RECIBE EL CAMBIO			
ÁREA SOLICITANTE		Secretaría de Planificación Estratégica.	
CAMBIO PROPUESTO			
DESCRIPCIÓN DEL CAMBIO: Añadir un (1) reporte de "Estado de Estudiante".			
JUSTIFICACIÓN DEL CAMBIO: Se necesita añadir el reporte propuesto.			
REGISTRO DE IMPACTO			
DESCRIPCIÓN DEL IMPACTO TÉCNICO: Este cambio afecta en la construcción del reporte.			
DESCRIPCIÓN DEL IMPACTO EN CRONOGRAMA O TIEMPO: No hay impacto en el cronograma.			
DESCRIPCIÓN DEL IMPACTO EN PRESUPUESTO O COSTOS: El costo adicional por el cambio efectuado será de S/ 350.00 Nuevos Soles.			
DESCRIPCIÓN DEL IMPACTO EN RECURSOS: Se incrementará el tiempo de trabajo del Analista en sobretiempo.			
REGISTRO DE OTROS IMPACTOS:			
RESOLUCIÓN DEL CAMBIO			
ACEPTADO X	RECHAZADO	ACEPTADO CON CONDICIONES	
RAZONES: Por medio de la reunión con el stakeholder se vio necesario el cambio solicitado y solicitar el sobretiempo para no atrasar las actividades planificadas.			
OBSERVACIONES:			
FIRMA DE RESPONSABLES			
GESTOR DE PROYECTO Maryury García Anticona.		LÍDER USUARIO Ing. Manuel Cok Aparcana	

Elaboración: las autoras

CONTROL DE CAMBIOS DEL ALCANCE DEL PROYECTO			
NOMBRE DEL PROYECTO: Análisis, Diseño e Implementación de Datamart para la generación de reportes de la Secretaría de Planificación Estratégica del MINEDU			
FECHA	10/08/2014	REQUERIMIENTOS DE CAMBIOS No.	02
INFORMACIÓN GENERAL DEL PROYECTO			
NOMBRE DE LA (S) PERSONA (S) QUE SOLICITA EL CAMBIO		Jorge Mesinas	
NOMBRE DE LA (S) PERSONA (S) QUE RECIBE EL CAMBIO			
ÁREA SOLICITANTE		Secretaría de Planificación Estratégica.	
CAMBIO PROPUESTO			
DESCRIPCIÓN DEL CAMBIO: Añadir en el proceso de matrícula el estado "En Definitivo".			
JUSTIFICACIÓN DEL CAMBIO: Se necesitan visualizar el estado de la Nómina en el diagrama de flujo del Proceso de Nóminas.			
REGISTRO DE IMPACTO			
DESCRIPCIÓN DEL IMPACTO TÉCNICO: Este cambio afecta en el proceso de nómina.			
DESCRIPCIÓN DEL IMPACTO EN CRONOGRAMA O TIEMPO: El cronograma no se verá afectado en la fase de análisis.			
DESCRIPCIÓN DEL IMPACTO EN PRESUPUESTO O COSTOS: Al no tener impacto no tiene impacto en el presupuesto.			
DESCRIPCIÓN DEL IMPACTO EN RECURSOS: No tiene impacto.			
REGISTRO DE OTROS IMPACTOS:			
RESOLUCIÓN DEL CAMBIO			
ACEPTADO X	RECHAZADO	ACEPTADO CON CONDICIONES	
RAZONES: Por medio de la reunión con el stakeholder se vio necesario el cambio solicitado.			
OBSERVACIONES:			
FIRMA DE RESPONSABLES			
GESTOR DE PROYECTO Maryury García Anticona.		LÍDER USUARIO Ing. Manuel Cok Aparcana	

Elaboración: las autoras

ANEXO 11 ENCUESTA DE SATISFACCIÓN

Nombre y Apellido:

Cargo:

Área/ Departamento:

Se da a su disposición la siguiente encuesta con la cual puede mostrar su grado de satisfacción con el producto brindado. Su opinión es de máxima importancia para poder definir los resultados obtenidos. Muchas gracias por su tiempo.

Coloque dentro del recuadro el valor de su puntuación según su punto de vista en representación del MINEDU.

Se tiene la siguiente escala:



Donde 1 es "Muy insatisfecho y 5 "Muy Satisfecho".

Items	Puntuación:
1. ¿Se dispone en forma oportuna los reportes del proceso de matrícula de las instituciones educativas a nivel departamental?	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
2. ¿Se dispone en forma oportuna los reportes del proceso de nóminas de las instituciones educativas a nivel departamental?	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
3. ¿Esta Ud. Conforme con la implementación del Datamart para la generación de reportes?	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
4. ¿Los indicadores propuestos van acorde a la información que desea plasmarse en los reportes?	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
5. ¿Esta Ud. Conforme con que la extracción, transformación y carga de datos se realice con la data contenida en el SIAGIE?	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
6. ¿Qué tan importante son los reportes generados con la ayuda del Datamart para la secretaria de planificación estratégica?	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
7. ¿El tiempo invertido en la generación de reportes de los procesos de nóminas y matrícula se ha reducido considerablemente?	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5

Si Ud. Tiene algún comentario adicional que quiera compartir con nosotros escríbalo en este espacio.

ANEXO 12 CONFIGURACIÓN DE FIREWALL CISCO ASA

A continuación se muestra el paso a paso de la configuración del Firewall:

1. Se asigna la VLAN 2 para el segmento Outside, interface Ethernet0/0

```
interface Ethernet0/0
description Connected to Outside Segment
switchport access vlan 2
```

2. Se asigna la VLAN 1 para el segmento Inside, interface Ethernet0/1

```
interface Ethernet0/1
description Connected to Inside Segment
switchport access vlan 1
```

3. Se asigna la VLAN 3 para el segmento DMZ, interface Ethernet0/2

```
interface Ethernet0/2
description Connected to DMZ Segment
switchport access vlan 3
```

4. Fijamos la IP WAN del segmento Outside

```
interface Vlan2
nameif outside
security-level 0
ip address 190.116.20.1 255.255.255.0
```

5. Fijamos la IP LAN del segmento Inside

```
interface Vlan1
nameif inside
security-level 100
ip address 192.168.248.1 255.255.255.0
```

6. Fijamos la IP DMZ del segment DMZ

```
object network inside-subnet
subnet 192.168.248.0 255.255.255.0
```

7. Creamos el Objeto inside-subnet para la Red LAN

```
interface Vlan3
 nameif dmz
 security-level 50
 ip address 1 92.168.1.1 255.255.255.0
```

8. Creamos el Objeto dmz-subnet para la Red DMZ

```
object network dmz-subnet
 subnet 192.168.1.0 255.255.255.0
```

9. Creamos el Objeto SIAGIE

```
object network SIAGIE
 host 192.168.1.100
```

10. Creamos el Objeto SIAGIE-external-ip

```
object network SIAGIE-external-ip
 host 190.116.20.64
```

11. Creamos la lista de acceso outside_acl (permitimos el acceso a los puertos 80 y 443)

```
access-list outside_acl extended permit tcp any object SIAGIE eq www
access-list outside_acl extended permit tcp any object SIAGIE eq https
```

12. Creamos la lista de acceso dmz_acl (restringimos el acceso a la Red DMZ)

```
access-list dmz_acl extended deny ip any object inside-subnet
access-list dmz_acl extended permit ip any any
```

13. Creamos el Objeto inside-subnet y habilitamos el acceso de la Red Inside hacia la Red Outside

```
object network inside-subnet
 nat (inside,outside) dynamic interface
```

14. Creamos el Objeto dmz-subnet y habilitamos el acceso de la Red DMZ hacia la Red Outside

```
object network dmz-subnet
 nat (dmz,outside) dynamic interface
```

15. Creamos el Objeto SIAGIE y habilitamos el acceso de la Red DMZ hacia la Red Outside para la IP 190.116.20.64 y el puerto 80

```
object network SIAGIE
  nat (dmz,outside) static SIAGIE-external-ip service tcp www www
```

16. Creamos el Objeto SIAGIE-https y habilitamos el acceso de la Red DMZ hacia la Red Outside para la IP 190.116.20.64 y el puerto 443

```
object network SIAGIE-https
  nat (dmz,outside) static SIAGIE-external-ip service tcp https https
```

17. Creamos los grupos de acceso outside_acl y dmz_acl

```
access-group outside_acl in interface outside
access-group dmz_acl in interface dmz
```

18. Creamos la ruta estática para el acceso a Internet

```
route outside 0.0.0.0 0.0.0.0 190.116.20.2 1
```