



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

**IMPLEMENTACIÓN DE UN AZURE DATA WAREHOUSE  
PARA LA TOMA DE DECISIONES ORGANIZACIONALES  
EN EL INSTITUTO CULTURAL PERUANO  
NORTEAMERICANO**

**PRESENTADO POR  
WILDER LEONCIO CORREA SANCHEZ**

**ASESOR  
GENER VÍCTOR ZAMBRANO LOLI**

**TRABAJO DE SUFICIENCIA PROFESIONAL  
PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO DE  
COMPUTACIÓN Y SISTEMAS**

**LIMA – PERÚ  
2024**



**CC BY-NC-ND**

**Reconocimiento – No comercial – Sin obra derivada**

El autor sólo permite que se pueda descargar esta obra y compartirla con otras personas, siempre que se reconozca su autoría, pero no se puede cambiar de ninguna manera ni se puede utilizar comercialmente.

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



**FACULTAD DE INGENIERÍA Y ARQUITECTURA**

**ESCUELA PROFESIONAL DE INGENIERÍA DE COMPUTACIÓN Y SISTEMAS**

**IMPLEMENTACIÓN DE UN AZURE DATA WAREHOUSE PARA LA  
TOMA DE DECISIONES ORGANIZACIONALES EN EL INSTITUTO  
CULTURAL PERUANO NORTEAMERICANO**

**TRABAJO DE SUFICIENCIA PROFESIONAL PARA OPTAR EL TÍTULO  
PROFESIONAL DE INGENIERO DE COMPUTACIÓN Y SISTEMAS**

**PRESENTADO POR:**

**CORREA SANCHEZ WILDER LEONCIO**

**ASESOR**

**ZAMBRANO LOLI, GENER VÍCTOR**

**Lima, PERÚ**

**2024**

## Dedicatoria

A Isaac y Blanca, mis padres, por su abnegado apoyo y la inspiración de vida.

A Elvia y Ana, mis hermanas, por su ejemplo y por su empuje a que consiga mis objetivos.

A Marco, por el apoyo y paciencia y por ayudarme a concretar una a una las tareas.

## Agradecimiento

A los profesores de la Universidad por su dedicación con los cursos y por actualizar los conceptos necesarios para lograr la realización del curso en general.

## Resumen

El Instituto Cultural Peruano Norteamericano (ICPNA) se destaca como una institución líder en la enseñanza del idioma inglés en Perú, ofreciendo una amplia gama de cursos para niños, jóvenes y adultos en diversos niveles, así como preparación para exámenes internacionales. Sin embargo, enfrenta el desafío de tener datos dispersos en diferentes aplicaciones y bases de datos, dificultando la toma de decisiones estratégicas.

Para abordar este problema, se propone el siguiente trabajo de diseño, desarrollo e implementación de un Azure Data Warehouse. Los objetivos principales incluyen la centralización de datos y la mejora de la toma de decisiones. Se utilizará una metodología híbrida que combina la metodología tradicional en cascada con la metodología ágil Medallion para garantizar una implementación eficiente y adaptativa.

La solución final será un Azure Data Warehouse que centraliza la información de las áreas académica, comercial y financiera del ICPNA. Esto permitirá a los tomadores de decisiones acceder a información valiosa, concisa, confiable e integrada, facilitando la creación de informes y análisis para respaldar la formulación de estrategias y la toma de decisiones en toda la organización

**Palabras Claves:** DataWarehouse, proceso de toma de decisiones, Azure, dimensiones, tablas de hechos.

## **Abstrat**

The Peruvian North American Cultural Institute (ICPNA) stands out as a leading institution in English language teaching in Peru, offering a wide range of courses for children, youth, and adults at various levels, as well as preparation for international exams. However, it faces the challenge of having dispersed data across different applications and databases, hindering strategic decision-making.

To address this issue, the proposes the design, development, and implementation of an Azure Data Warehouse. The main objectives include data centralization and improvement of decision-making. A hybrid methodology will be used, combining the traditional waterfall methodology with the agile Medallion methodology to ensure efficient and adaptive implementation.

The final solution will be an Azure Data Warehouse that centralizes information from the academic, commercial, and financial areas of the ICPNA. This will allow decision-makers to access valuable, concise, reliable, and integrated information, facilitating the creation of reports and analysis to support strategy formulation and decision-making throughout the organization.

**Keywords:** Data Warehouse, decision-making process, Azure, dimensions, fact tables.

## INTRODUCCIÓN

El presente informe se centra en el diseño, desarrollo e implementación de un Data Warehouse utilizando la plataforma Azure, con el objetivo de mejorar el proceso de toma de decisiones en el Instituto Cultural Peruano Norteamericano (ICPNA).

Se describe la metodología utilizada para llevar a cabo el proyecto, incluyendo la combinación de enfoques tradicionales (Cascada) y ágiles (Medallion) para garantizar una implementación eficiente y adaptativa.

Se detallará el proceso de diseño, desarrollo e implementación del Azure Data Warehouse, incluyendo la integración de datos dispersos de las áreas académica, comercial y financiera del ICPNA, así como la configuración de dimensiones y tablas de hechos para facilitar el análisis de la información.

Finalmente, se presentarán las conclusiones obtenidas del proyecto, destacando los logros alcanzados. Además, se ofrecerán recomendaciones para futuras mejoras y optimizaciones en el proceso de toma de decisiones organizacionales en el ICPNA.



## Índice de Contenido

CAPÍTULO I: Experiencia Profesional.....	8
1.1. Trayectoria Profesional.....	8
1.2. Contexto donde se desarrolló la experiencia Profesional .....	1
1.2.1. Descripción de la empresa donde desarrollo la experiencia profesional .....	1
1.2.2. Visión y Misión.....	1
1.2.3. Organigrama .....	2
1.2.4. Descripción del cargo o puesto .....	3
1.2.5. Productos y/o servicios .....	3
1.3. Contribución a la experiencia profesional .....	4
1.3.1. Contribución en la solución de problemas que se hayan presentado durante su estancia en la empresa.....	4
1.3.2. Contribución en términos de las competencias y habilidades adquiridas durante su formación profesional .....	4
1.4. Reflexión crítica de la experiencia .....	5
1.4.1. Aportes y beneficios obtenidos por el centro laboral, como parte de su contribución en la solución de las situaciones problemáticas. ....	5
1.4.2. Planteamiento de mejoras propuestas a la organización. ....	6
CAPÍTULO II: INFORME DEL PROYECTO DE ESPECIALIDAD .....	8
2. Informe Del Proyecto De Especialidad .....	8
2.1. Metodología .....	8
2.1.1. Descripción de la Metodología .....	8
2.1.2. Organización de la metodología .....	10
2.1.2.1. Estructura desglosable del trabajo – EDT.....	10
2.1.2.2. Cronograma de actividades.....	11
2.2. Desarrollo .....	12
2.2.1. Situación problemática o encargatura de labor.....	12
2.2.1.1. Contextualización del desarrollo de la experiencia profesional.....	12
2.2.1.2. Problema de la solicitud de trabajo a realizar .....	13
2.2.1.3. Objetivos.....	15
2.2.1.4. Alcance .....	16
2.2.1.5. Limitaciones.....	16
2.2.1.6. Justificación .....	16
2.2.2. Diseño de la solución .....	16
2.2.2.1. Inicio y planificación del proyecto (Hito 1).....	17

2.2.2.2.	Diseño de Arquitectura (Hito2).	18
2.2.2.3.	Diseño de dimensiones y tablas de hechos (Hito3).	20
2.2.2.4.	Desarrollo Azure Data Warehouse (Hito4).	23
2.2.2.5.	Pruebas funcionales y entrega (Hito5).	29
2.3.	Conclusiones	34
2.4.	Recomendaciones	35
GLOSARIO DE TERMINOS		36
REFERENCIAS		37

## Índice de Figuras

Figura 1: Organigrama Instituto Cultural Peruano Norteamericano .....	2
Figura 2: Estructura Desglosable del Trabajo .....	10
Figura 3: Cronograma de actividades .....	11
Figura 4: Diagrama de Ishikawa .....	15
Figura 5: Mapeo de Campos Tabla Source .....	19
Figura 6: Mapeo de Campos Tabla Target .....	19
Figura 7: Diagrama de Arquitectura.....	20
Figura 8: Tablas de hechos y dimensiones Área Gestión Académica .....	22
Figura 9: Creación de los esquemas y tabla para la ingesta de datos.....	24
Figura 10: Ingesta área Gestión Académica. ....	26
Figura 11: Ingesta área Finanzas. ....	27
Figura 12: Capa bronze .....	28
Figura 13: Capa silver .....	29
Figura 14: Capa gold.....	29
Figura 15: Matriz de pruebas .....	30
Figura 16: Reporte de diferencias por sede en el periodo 202305. ....	31
Figura 17: Reporte de diferencias por curso en el periodo 202305.....	31
Figura 18: Reporte de diferencias por aula en el periodo 202209.....	31

## Índice de Tablas

Tabla 1: Resumen de trayectoria profesional .....	1
Tabla 2: Equipos por áreas .....	18
Tabla 3: Componentes Azure .....	23
Tabla 4: Recursos Azure Data Factory .....	24

## Índice de Anexos

Anexo 01: ICPNA - Documento de Diseño de Arquitectura v3 .....	38
Anexo 02: ICPNA - Documento de tablas de hechos y dimensiones - Área Gestión Académica .....	59
Anexo 03: ICPNA - Documento de tablas de hechos y dimensiones - Área Contabilidad y Finanzas .....	69
Anexo 04: ICPNA - Documento de tablas de hechos y dimensiones - Área Comercial.....	82
Anexo 05: ICPNA - Mapeo de Campos .....	92
Anexo 06: ICPNA – Evidencias Azure Data Factory .....	97

# Reporte Turnitin

Reporte de similitud

NOMBRE DEL TRABAJO

**IMPLEMENTACIÓN DE UN AZURE DATA WAREHOUSE PARA LA TOMA DE DECISIONES ORGANIZACIONALES EN EL INSTITUT**

AUTOR

**WILDER LEONCIO CORREA SANCHEZ**

RECUENTO DE PALABRAS

**14262 Words**

RECUENTO DE CARACTERES

**82838 Characters**

RECUENTO DE PÁGINAS

**135 Pages**

TAMAÑO DEL ARCHIVO

**7.7MB**

FECHA DE ENTREGA

**Aug 9, 2024 10:13 AM GMT-5**

FECHA DEL INFORME

**Aug 9, 2024 10:15 AM GMT-5**

## ● 7% de similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para cada base de datos.

- 6% Base de datos de Internet
- Base de datos de Crossref
- 4% Base de datos de trabajos entregados
- 0% Base de datos de publicaciones
- Base de datos de contenido publicado de Crossref

## ● Excluir del Reporte de Similitud

- Material bibliográfico
- Material citado
- Coincidencia baja (menos de 10 palabras)



**Biblioteca FIA**






María Vásquez Claros  
Bibliotecóloga

## **CAPÍTULO I: Experiencia Profesional**





### **1.1. Trayectoria Profesional**

Soy un apasionado por las soluciones basadas en datos. Me formé como bachiller en Ingeniería de Computación y Sistemas y realicé el Programa de Especialización Analítica Data Engineer. Tengo más de 5 años de experiencia profesional, he colaborado en proyectos de diferentes sectores (industrial, agro-exportador, financiero y educativo). He participado en el diseño, desarrollo e implementación de Data Warehouse utilizando Oracle, SQL Server, Azure Data Factory y Databricks. También estuve involucrado en la creación y mantenimiento de Jobs, Funcions, Procedures, Cursores y Triggers. Me gusta trabajar en equipo, aportar a la sociedad y aprender constantemente. Mi objetivo es seguir desarrollando mis competencias en Data Engineering y contribuir a la innovación y el crecimiento de las organizaciones y el país. A continuación, dejo un cuadro que resumen mi paso por diferentes empresas en el territorio nacional.

**Tabla 1: Resumen de trayectoria profesional**

Logo	Empresa	Cargo	Inicio	Fin	Logros	Aprendizaje
	Indra	Data Engineer	dic-23	Actualidad	Estandarización de Assets en el proceso de pre-migración ADA - User Migration	Git, Python, Metodologías ágiles, Bases de datos, SQL, PySpark, Bitbucket, Big data
	Dígito	Data Engineer	may-23	oct-23	Desarrollo e implementación de un Azure Data Warehouse para ICPNA	Resolución de problemas, Modelo de datos, DataStage, Modelado de datos, Azure Data Factory, Ingeniería, SQL, Modelado de datos
	Sofftek	Analista Programador SQL	abr-22	abr-23	Pase a producción del sistema QRM – Solución ALMR para el área de Tesorería y Liquidez del Banco de Crédito del Perú	Resolución de problemas, Modelo de datos, Análisis de datos, Desarrollo de software, Metodologías ágiles, Microsoft Azure, Bases de datos
	CSTI	Consultor Oracle Developer	ago-21	abr-22	Desarrollo e implementación de reportes y corrección de formularios para el sistema AxcelX para la empresa Rímac Seguros	Análisis de datos, Desarrollo de software, PL/SQL, Bases de datos
	Cooperativa de Ahorro y Crédito Pacífico	Analista Programador PL/SQL	ene-21	jul-21	Desarrollo de reportes y corrección de formularios para el sistema AxcelX	Análisis de datos, Desarrollo de software, PL/SQL, Bases de datos



	Mebol	Implementation Team Leader	ago-20	nov-20	Implementación del sistema ERP Smart Production 4.0	Capacidad de análisis, Liderazgo, Trabajo en equipo, Metodologías ágiles, Pruebas de integración, Toma de decisiones
	Metalpren	Asistente Sistemas	jun-19	ago-20	Administración de la base de datos del ERP empresarial y la encargatura de toda el área de sistemas en pandemia por un período de 3 meses.	Modelo de datos, Capacidad de análisis, PL/SQL, ISO 27001, Administración de bases de datos
	Ricoh	Practicante Profesional Auditoría	sep-17	ene-19	Auditoría de los equipos a través de hardware especializado en todos los clientes. Incremento del control de equipos auditados del 40% al 98%.	Resolución de problemas, Investigación, Trabajo en equipo, Microsoft Excel, Comunicación
	USMP	Practicante Soporte Técnico	sep-16	jul-17	Atención y solución oportuna de incidentes relacionados con las computadoras y redes de la Facultad de Ingeniería y Arquitectura	Cumplimiento, Orden de trabajo, Skills técnicos

Fuente: Elaboración propia, 2024

## **1.2. Contexto donde se desarrolló la experiencia Profesional**

### **1.2.1. Descripción de la empresa donde desarrollo la experiencia profesional**

El Instituto Cultural Peruano Norteamericano (ICPNA) es una institución líder en la promoción del intercambio cultural entre Perú y Estados Unidos. Pertenece al sector Educación y Cultura. Fundado en 1946, el ICPNA ha sido un pionero en la enseñanza del idioma inglés y la difusión de la cultura norteamericana en Perú. A lo largo de los años, ha expandido su oferta educativa para incluir cursos de inglés, programas académicos, actividades culturales y servicios de biblioteca, convirtiéndose en un referente en educación y cultura en el país. El ICPNA ha mantenido una presencia destacada en Lima y otras regiones a nivel nacional, sirviendo como un puente cultural entre Perú y Estados Unidos y promoviendo la comprensión intercultural y el desarrollo educativo en la región. Actualmente la sede central se ubica en la Av. Angamos Oeste 120, en el distrito de Miraflores, pero cuenta con más sedes en todo Lima: San Miguel, San Isidro, Surco, La Molina y Lima Centro, cuenta también con sedes en las regiones: Arequipa, Cusco, Piura, Chiclayo e Iquitos.

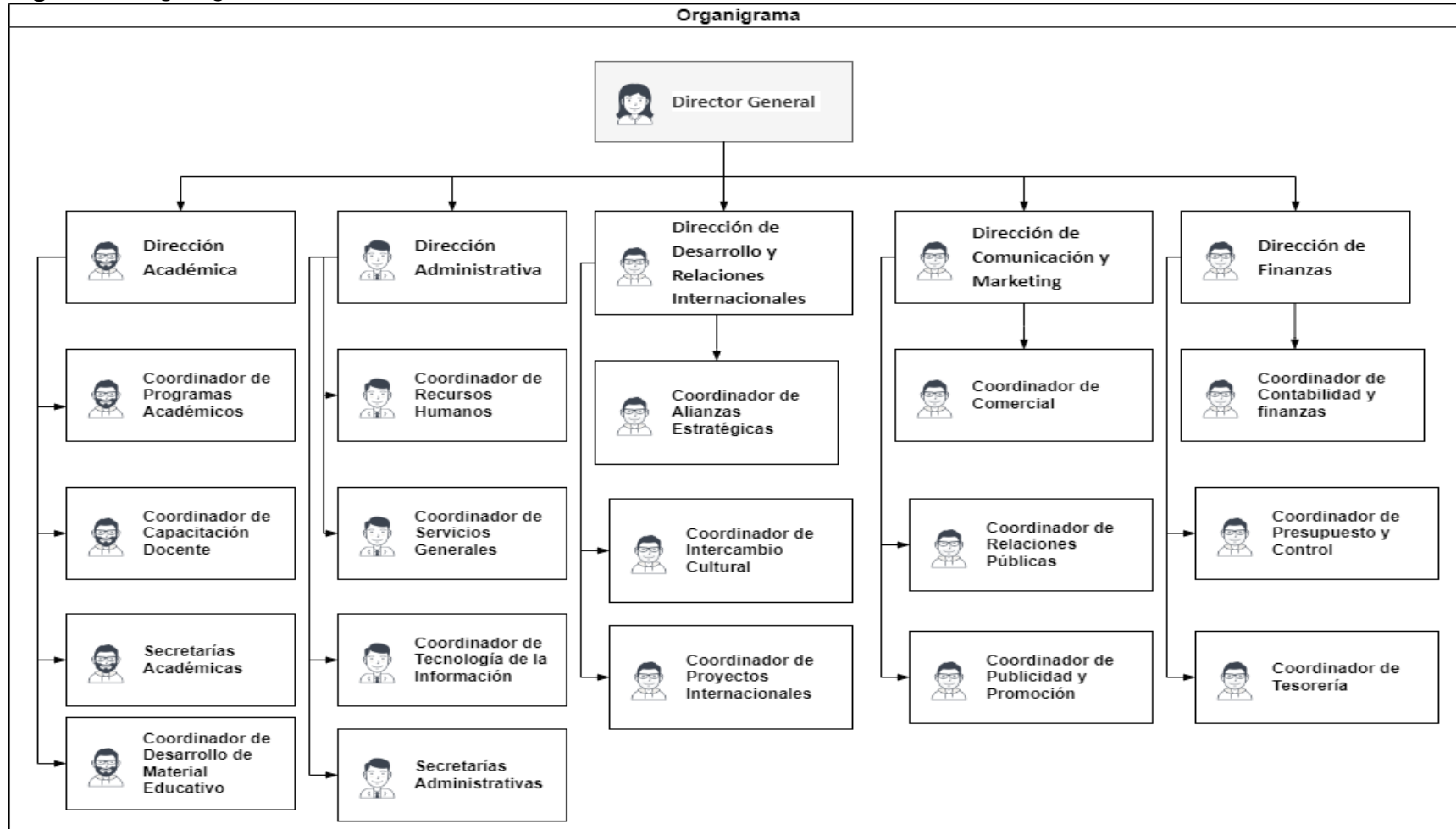
### **1.2.2. Visión y Misión**

**Visión:** “Ser el centro binacional líder a nivel global brindando servicios culturales y educativos a la comunidad con excelencia, innovación y calidez, comprometido con el desarrollo de nuestros estudiantes y colaboradores”.

**Misión:** Contribuir al desarrollo de la comunidad, promoviendo el intercambio cultural entre Perú y los Estados Unidos de América a través de la enseñanza de idiomas, servicios internacionales, red de bibliotecas y la difusión de actividades culturales, con servicios de calidad y el compromiso de los colaboradores hacia la satisfacción del cliente, la innovación y la eficiencia de los procesos”.

### 1.2.3. Organigrama

Figura 1: Organigrama Instituto Cultural Peruano Norteamericano



Fuente: Elaboración Propia, 2024

#### **1.2.4. Descripción del cargo o puesto**

Debido a que el ICPNA no cuenta con un área específica de desarrollo, se contrató a una consultora para la ejecución total del proyecto. La empresa seleccionada fue Dígito, elegida por su experiencia y especialización en la gestión, transformación y presentación de datos. Fui contratado por Dígito y participé activamente en todo el proceso del proyecto. Esto incluyó reuniones entre el cliente y nuestro equipo técnico, compuesto principalmente por el Gerente General (Business Analytics), la Project Manager y el Arquitecto de Datos..

**Cargo:** Data Engineer.

**Funciones:**

- Levantamiento de la información funcional para el modelamiento dimensional y de tablas de hechos y dimensiones.
- Realización del mapeo de procesos para la gestión académica, campañas del área comercial e informes del área financiera.
- Desarrollo de Datasets, Data Flow, Linked Services y Pipelines en Data Factory de Azure. Bajo la metodología Medallion.
- Elaboración de ETL y Modelado de datos (Data Warehouse / Data Lake)
- Documentación de cada fase y entregable.

#### **1.2.5. Productos y/o servicios**

El Instituto Cultural Peruano Norteamericano ofrece los siguientes servicios:

- **Cursos de Inglés:** Ofrecen una amplia gama de cursos de inglés para niños, adolescentes y adultos, adaptados a diferentes niveles de habilidad y necesidades específicas.
- **Programas Académicos:** Diplomados, certificaciones y programas especiales en áreas como traducción, negocios, cultura y literatura.

- **Actividades Culturales:** Conciertos, exposiciones, conferencias y proyecciones de cine para promover el intercambio cultural entre Perú y Estados Unidos.
- **Servicios de Biblioteca:** Colección de libros en inglés y español, recursos multimedia y acceso a bases de datos para estudiantes e investigadores.
- **Eventos Especiales:** Organización de eventos especiales como ferias educativas, ferias del libro y festivales culturales para la comunidad.

### **1.3. Contribución a la experiencia profesional**

#### **1.3.1. *Contribución en la solución de problemas que se hayan presentado durante su estancia en la empresa.***

Dentro de las contribuciones alcanzadas podemos resaltar la importancia de una comunicación efectiva y una colaboración estrecha entre los diferentes equipos involucrados en el proyecto.

A lo largo del proceso, fue crucial mantener una comunicación clara y constante con los stakeholders de cada una de las áreas del ICPNA para comprender sus necesidades y expectativas, así como con el equipo técnico para garantizar una implementación exitosa del Data Warehouse.

Esta experiencia subraya la importancia de la gestión de proyectos y el trabajo en equipo en el éxito de iniciativas tecnológicas complejas.

#### **1.3.2. *Contribución en términos de las competencias y habilidades adquiridas durante su formación profesional***

La implementación del proyecto del Data Warehouse ha sido un hito importante en mi carrera profesional, brindándome la oportunidad de aplicar mis conocimientos en ingeniería de datos en un contexto real y desafiante. Sin embargo, esta experiencia también ha revelado áreas de mejora y aprendizaje que son cruciales para futuros proyectos similares.

Uno de los principales desafíos encontrados durante el desarrollo del proyecto fue la integración de datos dispersos de múltiples fuentes en el Data Warehouse. Aunque logramos centralizar la información de las áreas académica, comercial y financiera del ICPNA, enfrentamos dificultades en la estandarización y limpieza de datos, lo que afectó la calidad y precisión de los análisis posteriores. Esta experiencia resalta la importancia de establecer procesos robustos de gestión de datos desde el inicio del proyecto.

#### **1.4. Reflexión crítica de la experiencia**

##### **1.4.1. *Aportes y beneficios obtenidos por el centro laboral, como parte de su contribución en la solución de las situaciones problemáticas.***

Podemos indicar los siguientes aportes y beneficios:

- **Integración de datos dispersos:** Lideré el proceso de integración de datos procedentes de diversas fuentes, incluyendo reportes en Excel, Dynamics 360 y bases de datos SQL (Académica y Spring), garantizando la centralización y coherencia de la información en el Data Warehouse del ICPNA.
- **Diseño y desarrollo del Data Warehouse:** Colaboré estrechamente con el equipo técnico en el diseño y desarrollo del esquema de datos del Data Warehouse, incluyendo la definición de dimensiones y tablas de hechos, lo que permitió facilitar el análisis y la generación de informes para las áreas académica, comercial y financiera del ICPNA.
- **Implementación de metodologías híbridas:** Contribuí en la aplicación de una metodología híbrida que combinó enfoques tradicionales y ágiles, como la metodología Cascada y la metodología Medallion, lo que garantizó una implementación eficiente y adaptativa del proyecto.

- **Comunicación y colaboración:** Mantuvo una comunicación clara y constante con los stakeholders del ICPNA y el equipo técnico durante todo el proyecto, asegurando una comprensión completa de los requisitos y expectativas, y promoviendo una colaboración efectiva entre todos los involucrados.

#### **1.4.2. Planteamiento de mejoras propuestas a la organización.**

Podría indicar las siguientes mejoras para que el ICPNA pueda mantener los resultados obtenidos por el proyecto:

- **Capacitación del Personal:** Desarrollar un programa de capacitación continua para el personal en el uso de las nuevas herramientas y plataformas de Azure. Esto asegurará que todos los stakeholders estén al tanto de las mejores prácticas y puedan utilizar eficientemente los nuevos sistemas implementados.
- **Integración de Sistemas:** Fomentar la integración de sistemas informáticos para asegurar que todas las áreas del ICPNA puedan compartir y acceder a la información de manera fluida. Esta integración evitará la duplicación de esfuerzos y garantizará que todos los departamentos trabajen con la misma base de datos actualizada.
- **Mejora en la Infraestructura Tecnológica:** Actualizar y optimizar la infraestructura tecnológica para soportar la nueva carga de trabajo y garantizar un rendimiento óptimo de los sistemas. Esto incluye la actualización de servidores, almacenamiento, memoria, procesadores y la adquisición de las licencias y software necesarios.
- **Monitoreo y Mantenimiento Continuo:** Establecer un sistema de monitoreo continuo para identificar y resolver problemas rápidamente. Además, planificar mantenimientos periódicos para asegurar que los sistemas funcionen sin

interrupciones y se adapten a las nuevas necesidades de la organización.



## CAPÍTULO II: INFORME DEL PROYECTO DE ESPECIALIDAD

### 2. Informe Del Proyecto De Especialidad

#### 2.1. Metodología

##### 2.1.1. Descripción de la Metodología

El enfoque metodológico seleccionado para la implementación del proyecto de desarrollo e implementación del Data Warehouse para el Instituto Cultural Peruano Norteamericano (ICPNA) fue la metodología **Cascada**. Este enfoque se caracteriza por su naturaleza secuencial y lineal, donde cada fase del proyecto se organiza de manera jerárquica y progresiva, requiriendo que cada fase se complete antes de avanzar a la siguiente (Pressman, 2014) .

Las fases de la metodología Cascada se estructuraron de la siguiente manera considerando que para cada fase le correspondería un Hito que sería la entrega de documentación y/o artefactos según corresponda:

#### **Fase 1: Inicio y Planificación del Proyecto**

- **Hito 1:** Inicio y planificación del proyecto

#### **Fase 2: Diseño**

- **Hito 2:** Diseño de Arquitectura
- **Hito 3:** Diseño de dimensiones y tablas de hechos

#### **Fase 3: Desarrollo e Implementación**

- **Hito 4:** Desarrollo Azure Data Warehouse

**Nota:** Se adoptó la metodología **Medallion** para la construcción de artefactos en Azure, dividiendo esta fase en tres capas: Bronze, Silver y Gold (Microsoft, 2020) .

#### **Fase 4: Entrega**

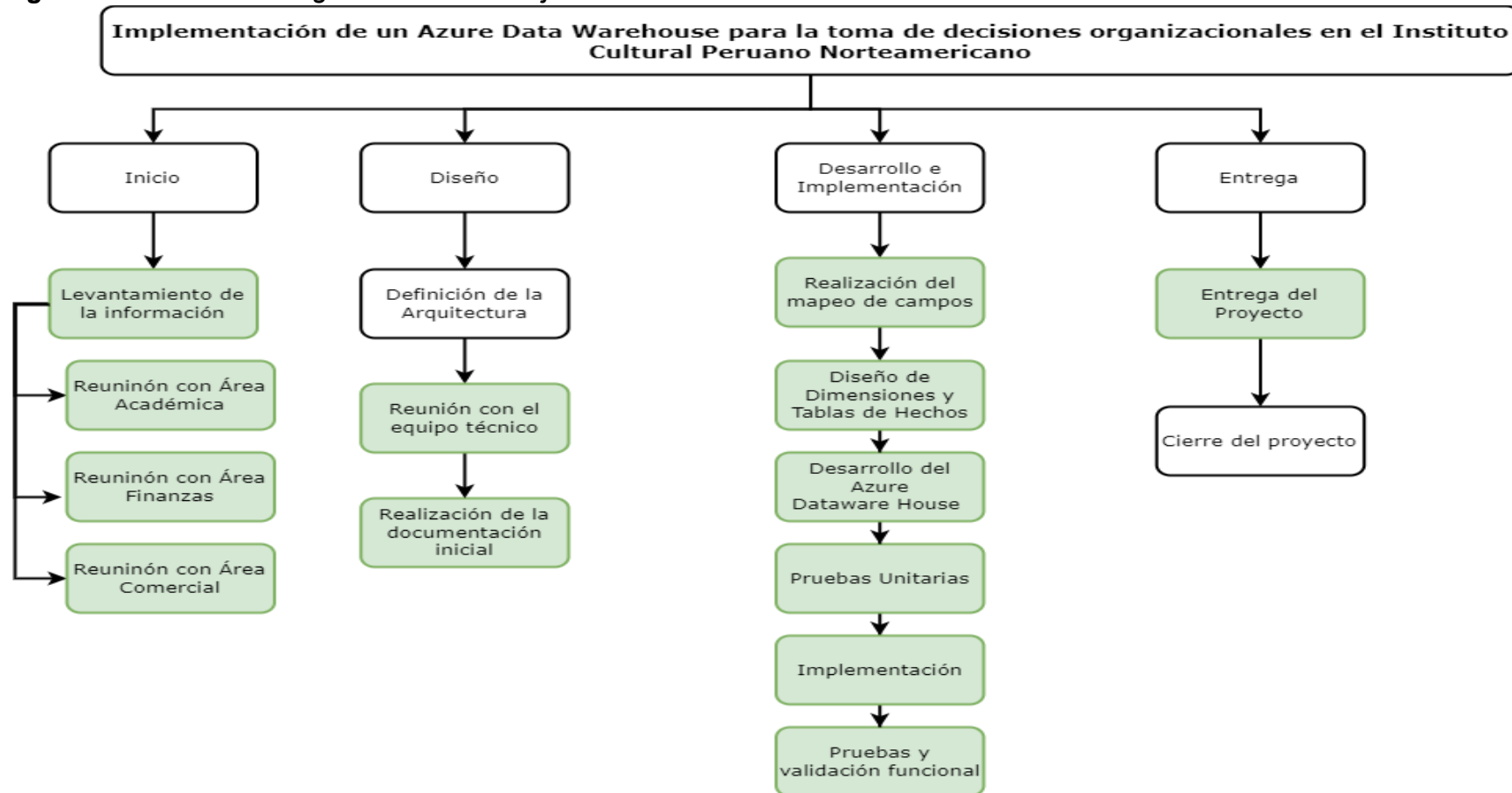
- **Hito 5:** Pruebas funcionales y entrega al cliente

Esta estructura permitió una gestión eficiente del proyecto, con hitos claramente definidos y entregables específicos en cada etapa del proceso. La metodología Cascada se adaptó a las necesidades del proyecto, proporcionando un marco sólido para garantizar la calidad y el éxito de la implementación del Data Warehouse en el ICPNA.

## 2.1.2. Organización de la metodología

### 2.1.2.1. Estructura desglosable del trabajo – EDT.

Figura 2: Estructura Desglosable del Trabajo

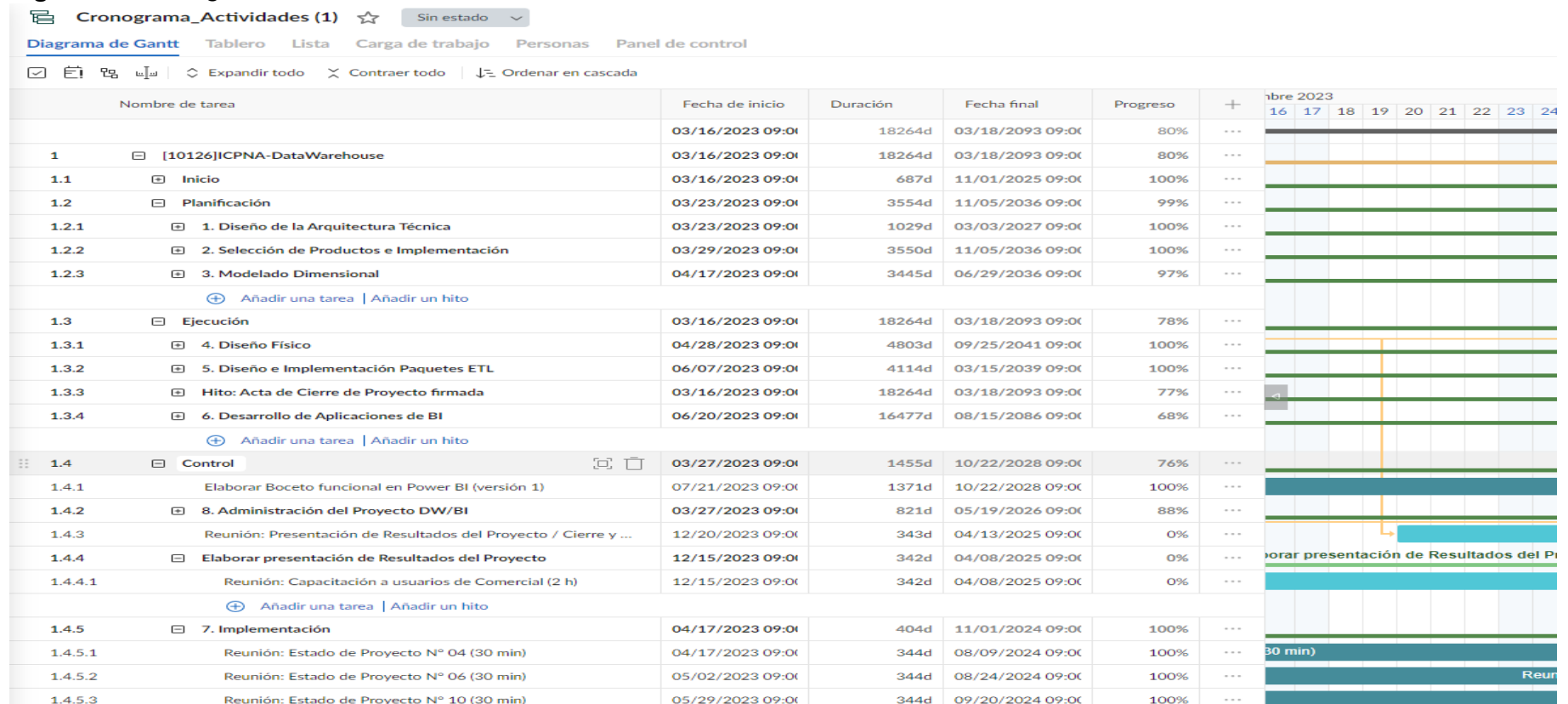


Nota: Las actividades resaltadas en verde representan las áreas donde he participado de manera activa.

### 2.1.2.2. Cronograma de actividades.

A continuación de muestra, el cronograma de actividades que se siguió a fin de completar cada una de las fases descritas en la metodología.

**Figura 3: Cronograma de actividades**



Fuente: Dígito para ICPNA, 2023.

## **2.2. Desarrollo**

### **2.2.1. Situación problemática o encargatura de labor**

#### **2.2.1.1. Contextualización del desarrollo de la experiencia profesional**

La implementación del proyecto del Azure Data Warehouse para el Instituto Cultural Peruano Norteamericano (ICPNA) surge en respuesta a una serie de desafíos y necesidades identificadas en la organización.

En el contexto del ICPNA, se observó que los datos relevantes para la toma de decisiones estaban dispersos en diferentes aplicaciones y sistemas (La matrícula se almacenaba en la Base de Datos Académica, los datos financieros en la base de datos Spring y SAP y la obtención de cursos, aulas y docentes disponibles se trabajaban en el Aplicativo Incio, el área comercial trabaja con Dynamic 360 y además se usan documentos en excel que están en el Sharepoint), lo que dificultaba la obtención de una visión integral y precisa de las operaciones de la institución. Además, la falta de integración entre estas fuentes de datos generaba procesos manuales y duplicados, ralentizando la generación de informes y análisis cruciales para la gestión eficaz de la organización.

Esta situación se agravaba por la complejidad y diversidad de las áreas de negocio del ICPNA, que abarcaban desde la gestión académica hasta aspectos comerciales y financieros. La necesidad de contar con una plataforma centralizada y unificada para la gestión de datos se volvió imperativa para optimizar la toma de decisiones y mejorar la eficiencia operativa en todas las áreas de la organización.

Además, se identificaron retos adicionales relacionados con la infraestructura tecnológica existente y la resistencia al cambio por parte del personal. La falta de integración entre los sistemas existentes y la utilización de herramientas manuales generaban ineficiencias y dificultades operativas (reportes que se trabajaban en excel), lo que impactaba negativamente en la experiencia tanto del personal como de los estudiantes y clientes del ICPNA.

En este contexto, el proyecto de desarrollo e implementación del Data Warehouse se presenta como una solución integral y estratégica para abordar estos desafíos y transformar la forma en que el ICPNA gestiona y

utiliza sus datos. La centralización de la información, la mejora en los procesos de toma de decisiones y la optimización de la eficiencia operativa son algunos de los beneficios esperados de esta iniciativa, que busca impulsar el crecimiento y la excelencia en la gestión del Instituto Cultural Peruano Norteamericano.

#### **2.2.1.2. Problema de la solicitud de trabajo a realizar**

La falta de centralización de la información y la existencia de sistemas informáticos independientes, cada uno manejando su propia base de datos, plantea un desafío significativo para el proyecto. Esta dispersión de datos dificulta el cruce de información entre las diferentes áreas de la organización y genera discrepancias debido a las diferencias entre las bases de datos utilizadas.

Además, el tiempo requerido para la entrega de informes, especialmente los relacionados con matrícula y asignación docente, es excesivo, con un proceso que puede tomar hasta dos semanas. Esta demora impacta negativamente en la capacidad del ICPNA para tomar decisiones oportunas y eficientes, retrasando la respuesta a las necesidades del instituto y sus alumnos.

El uso del Diagrama de Ishikawa nos permite identificar las causas raíz de un problema mediante la visualización de diferentes factores que podrían contribuir a su aparición. En el caso del ICPNA, el problema principal es la falta de centralización de la información y la demora en la generación de informes clave para la toma de decisiones organizacionales.

#### **Causas Potenciales:**

- **Tecnología y Sistemas Informáticos:**
  - Falta de integración entre los sistemas utilizados en diferentes áreas (BD Académica, BD Spring, SAP, BD Inicio, Dynamic 360).
  - Dependencia de sistemas obsoletos que dificultan el intercambio de datos como es el caso del aplicativo Inicio que es el sistema de escritorio on-premise que siguen utilizando

para generar reportes de disponibilidad de cursos, aulas y docentes.

- **Procesos Internos:**

- Procesos manuales y redundantes para recopilar y consolidar datos para el cruce de la información de las bases de datos Académica con el aplicativo Inicio.
- Falta de estandarización en los procesos de generación de informes, ya que cada área trabaja sus propios informes y al tratar de unificarlos para presentarlos a la dirección general tienen que realizar varios “ajustes” para llegar a la presentación sin tantas inconsistencias.

- **Recursos Humanos:**

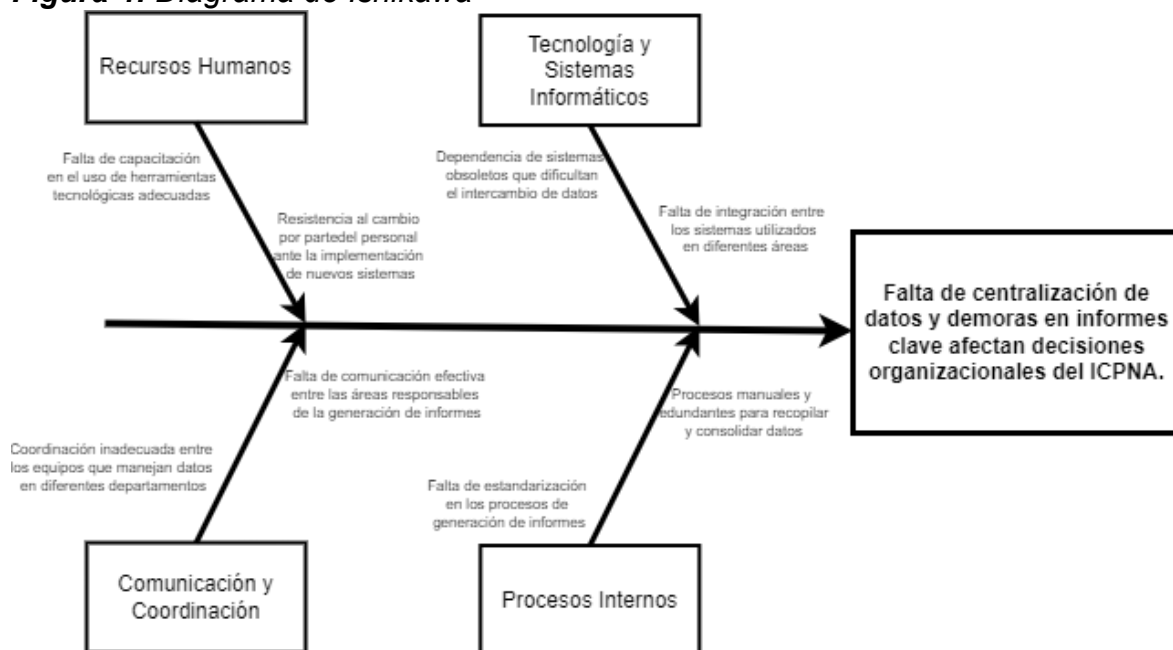
- Falta de capacitación en el uso de herramientas tecnológicas adecuadas.
- Resistencia al cambio por parte del personal ante la implementación de nuevos sistemas. La negación de utilizar otros sistemas más eficientes como PowerBI en reemplazo de Excel.

- **Comunicación y Coordinación:**

- Falta de comunicación efectiva entre las áreas responsables de la generación de informes.
- Coordinación inadecuada entre los equipos que manejan datos en diferentes departamentos.

El uso del Diagrama de Ishikawa nos ayuda a comprender la complejidad del problema y a identificar áreas clave que requieren atención para mejorar la eficiencia y la efectividad en la gestión de la información en el ICPNA.

**Figura 4: Diagrama de Ishikawa**



Fuente: Elaboración propia, 2024

### 2.2.1.3. Objetivos

Dada la actual coyuntura de los diferentes centros de datos y su modernización en lagos de datos es necesario desarrollar una solución acorde con la actualidad de la gestión de datos, que integre por completo todos los repositorios de datos (BD Académica, BD Spring) de todas las aplicaciones disponibles y necesarias para la generación de reportes por lo que se propone la creación de Data Lake que contenga la totalidad de datos y al mismo tiempo lograr que se elaboren reportes en el menor tiempo posible y evitar los errores generados de la manipulación manual de los datos; para lograrlo es necesario establecer los siguientes objetivos específicos:

- Centralizar la información de todas las áreas del ICPNA en un repositorio único y accesible (DataLake).
- Mejorar la precisión y calidad de los informes generados, al eliminar los errores asociados con la manipulación manual de datos dispersos.



#### **2.2.1.4. Alcance**

- Integrar datos dispersos de las áreas académica, comercial y finanzas del ICPNA en un repositorio centralizado.
- Desarrollar y poner en funcionamiento un Data Warehouse utilizando Azure para almacenar y gestionar los datos de manera eficiente.
- Capacitar al personal técnico del ICPNA en el uso efectivo del nuevo repositorio (datalake) y los procesos asociados.

#### **2.2.1.5. Limitaciones**

- Limitaciones de tiempo y recursos pueden afectar la velocidad y la extensión del proyecto.
- La resistencia al cambio por parte del personal podría retrasar la adopción y el uso efectivo del nuevo sistema.
- La disponibilidad de datos históricos y la calidad de los mismos pueden influir en la precisión de los informes generados.

#### **2.2.1.6. Justificación**

La implementación de un Data Warehouse en el ICPNA es crucial para mejorar la eficiencia operativa y la toma de decisiones en la organización. Al centralizar la información y reducir los tiempos de generación de informes, el instituto podrá responder de manera más rápida y precisa a las necesidades de sus alumnos y mejorar la calidad de sus servicios educativos.

### **2.2.2. Diseño de la solución**

El proyecto consiste en implementar un Azure Data Warehouse en el ICPNA. Se planifica comenzar con la fase de diseño y desarrollo del sistema para la centralización de datos, siguiendo la metodología Medallion para la creación de los diferentes artefactos en Azure Data Factory en cada una de las capas (Bronze, Silver y Gold). Se establece un cronograma de ejecución que contempla la implementación inicial en Lima desde el 16 de marzo hasta el 31 de mayo.

Durante la fase de implementación en Lima, se enfocará en abordar posibles desafíos identificados durante las reuniones semanales y quincenales, como problemas de acceso, resistencia al cambio y errores funcionales. Se recopilará retroalimentación detallada para optimizar el proceso de implementación y garantizar su éxito.

### **2.2.2.1. Inicio y planificación del proyecto (Hito 1).**

#### **2.2.2.1.1. Inicio.**

Se tiene que cumplir con la entrega de la Acta de Kick off aprobada a fin que el Hito 1 se dé por finalizado, para ello se debe agendar la reunión para la elaboración de la acta que debe contemplar el resumen de los requerimientos funcionales que se indicaron inicialmente a nivel contractual. La project manager del proyecto se encargará de la realización del cronograma de actividades que contemple los tiempos para cada una de las fases. Los artefactos de esta fase son :

- Reunión Kick Off
- Elaborar cronograma detallado
- Elaborar acta de Kick Off (Alcance, Responsable, Horarios)
- Acta de Kick Off aprobada.

#### **2.2.2.1.2. Reuniones de requerimientos funcionales.**

Se deben planificar reuniones con las tres áreas que involucran al proyecto: académica, comercial y finanzas, a fin de determinar los requerimientos funcionales, se revisarán los indicadores que maneja cada área, así como, las reglas de negocio que se toman en cuenta para hallar los KPI's.

Se agendarán reuniones con el equipo técnico para el entendimiento de los procesos de generación de informes que se utilizan en cada una de las áreas para determinar la arquitectura requerida en el Azure Data Warehouse, así como las diferentes fuentes que se utilizan en la elaboración de los informes.

Las reuniones se van a estructurar por equipos a fin de poder organizar las agendas:

**Tabla 2: Equipos por áreas**

Área	Equipo
Académica	Equipo 1: Programación. Equipo 2: Asignación Docente
Comercial	Equipo 3: Campus y App. Inicio. Equipo 4: Campañas. Equipo 5: Ventas. Equipo 6: CRM.
Fianzas	Equipo 7: Gestión Financiera Equipo 8: Gestión Contable. Equipo 9: Spring y SAP

Fuente: Elaboración propia, 2024

### **2.2.2.2. Diseño de Arquitectura (Hito2).**

#### **2.2.2.2.1. Diseño de Arquitectura técnica y revisión de fuentes de datos actuales.**

En esta fase se va a diseñar la arquitectura que soportará la solución Azure Data Warehouse, para ello se debe tener los requisitos funcionales completos de las tres áreas comprometidas y de todos los equipos establecidos.

#### **2.2.2.2.2. Realizar el inventario de fuentes de datos (mapeo de campos).**

Luego de establecer los requerimientos funcionales, se va a elaborar el inventario de datos, los mismos que van a describir las tablas inputs (source) que pertenecen a las distintas fuentes de datos. Se deberán elaborar cuadros con la especificación de cada uno de los atributos de la tabla. En excel se deben considerar las siguientes columnas para describir a cada uno de los atributos: Source System, Source Schema, Source Table, Source Field Name, Source Datatype, ETL Rules, Comments

**Figura 5: Mapeo de Campos Tabla Source**

Source						
Source System	Source Schema	Source Table	Source Field Name	Source Datatype	ETL Rules	Comments
AcademicaUID	dbo	Curso	uidIdCurso	varchar(100)	Llave para buscar en DimCurso.IdCurso	
AcademicaUID	dbo	Curso	chrCursoCodigo	char(20)		
AcademicaUID	dbo	Curso	vchNombreCurso	varchar(100)	Select del nombre del curso	
AcademicaUID	dbo	Curso	vchCursoAbreviado	varchar(50)	Select del curso abreviado	
AcademicaUID	dbo	Curso	vchDescripcionCurs	varchar(50)		
AcademicaUID	dbo	Curso	bitVigencia	bit		
AcademicaUID	dbo	Nivel	chrCodNivel	varchar(50)		
AcademicaUID	dbo	Curso	tinMesesDuracion	int	Select de los meses de duración	
AcademicaUID	dbo	Curso	vchTipoCurso	varchar(10)		
AcademicaUID	dbo	Fase	uidIdFase	uid	Llave utilizada para la desnormalización	Este campo no se incluirá en la Dimensión
AcademicaUID	dbo	Fase	vchFaseCodigo	varchar(10)		
AcademicaUID	dbo	Fase	vchNombreFase	varchar(100)		
AcademicaUID	dbo	Fase	vchFaseAbreviado	varchar(50)		
AcademicaUID	dbo	Frecuencia	uidIdFrecuencia	uid	Llave utilizada para la desnormalización	Este campo no se incluirá en la Dimensión
AcademicaUID	dbo	Frecuencia	vchFrecuenciaNombr	varchar(50)		
AcademicaUID	dbo	Intensidad	uidIdIntensidad	uid	Llave utilizada para la desnormalización	Este campo no se incluirá en la Dimensión
AcademicaUID	dbo	Intensidad	vchIntensidadNombr	varchar(20)		
AcademicaUID	dbo	Modalidad	uidIdModalidad	uid	Llave utilizada para la desnormalización	Este campo no se incluirá en la Dimensión
AcademicaUID	dbo	Modalidad	vchNombreModalidad	varchar(20)		
AcademicaUID	dbo	Curso	tinCurso	tinyint	Carga Inicial: tinCurso=1 Carga Incremental: Si tinCurso>1, tinCurso=1	
AcademicaUID	dbo	Nivel	vchNombre	varchar(50)		
AcademicaUID	dbo	Programa	vchNombreProg	varchar(50)		

Fuente: Elaboración propia, 2024

Adicionalmente se debe realizar el diccionario de datos de las dimensiones, éste trabajo se realizará en paralelo conforme se vaya afinando el diseño de las dimensiones y tablas de hecho. Al igual que con las tablas anteriores, ahora se van a mapear campos de las tablas dimensionales (target)

**Figura 6: Mapeo de Campos Tabla Target**

Target									
Column Name	Description	Attribute Group	Datatype	Size	Key?	FK To	NULL?	Example Values	
CursoKey	Llave subrogada de la Dimension	Key	int		PK		N		
IdCurso	Identificador del curso	Identificador	varchar	100			N	1, 2, 3, 4 ...	
CodigoCurso	El código del curso	Name	varchar	100			N	B11A1	
NombreCurso	El nombre del curso	Name	varchar	100			N	(INT) Intermediate	
CursoAbreviado	El nombre del curso abreviado	Name	varchar	50			N	Int. Conv.	
DescripcionCurso	Se describe el curso	Name	varchar	50			N		
Vigencia	Muestra la vigencia del curso	Vigencia	varchar	50			N		
CodigoNivel	Código del nivel	Código	varchar	50			N		
MesesDuracion	Cantidad de duracion del curso en meses	Name	int				N		
TipoCurso	El nombre del tipo de curso	Name	varchar	10			N		
uidIDFase	Código indificador de la Fase	Identificador	uid		FK	dbo.Fase	N		
CodigoFase	EL codigo del la fase	Name	varchar	10			N		
NombreFase	El nombre de la fase	Name	varchar	100			N		
FaseAbreviada	El nombre de la fase abreviado	Name	varchar	50			N		
uidIDFrecuencia	Código indificador de la Frecuencia	Identificador	uid		FK	dbo.Frecuencia	N		
FrecuenciaNombre	El nombre de la frecuencia del curso	Name	varchar	50			N		
uidIDIntensidad	Código indificador de la Intensidad	Identificador	uid		FK	dbo.Intensidad	N		
IntensidadNombre	El nombre de la intensidad del curso	Name	varchar	20			N		
uidIDModalidad	Código indificador de la Modalidad	Identificador	uid		FK	dbo.Fase/dbo.Modalida	N		
NombreModalidad	El nombre de la modalidad del curso	Name	varchar	20			N		
EstadoCurso	Indica que el curso activo o inactivo	Estado	bit	2			N		
NombreNivel	Nombre del nivel	Name	varchar	50			N		
NombrePrograma	Nombre del programa académico	Name	varchar	50			N		

Fuente: Elaboración propia, 2024

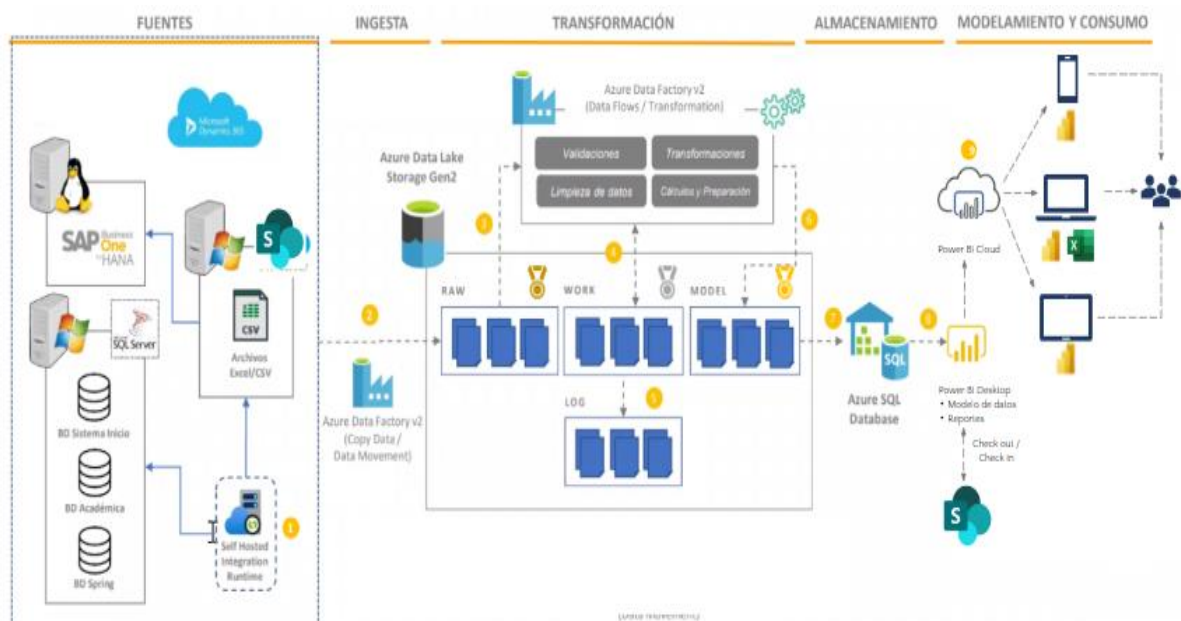
El documento que contenga todo mapeo de campos se puede revisar en los Anexos del presente documento ([Anexo 05: ICPNA – Mapeo de Campos](#))

### 2.2.2.3. Elaborar el diagrama de arquitectura de la solución.

Esta actividad será ejecutada por el Arquitecto de Datos, una vez que haya culminado deberá agendar una reunión para discutir el diagrama propuesto con el equipo técnico a fin de aceptar o regularizar observaciones.

Se deberá realizar también un documento que contenga el detalle de lo que involucraría cada una de las partes del diagrama. Dicho documento se puede revisar en los Anexos (Anexo 01: ICPNA – Documento de Diseño de Arquitectura).

**Figura 7: Diagrama de Arquitectura**



Fuente: Dígito, 2023

### 2.2.2.3. Diseño de dimensiones y tablas de hechos (Hito3).

El diseño de dimensiones y tablas de hechos es una fase crítica en la implementación de un Azure Data Warehouse para el ICPNA. Esta etapa establece la estructura y la relación entre los datos que se almacenarán en el data lake, garantizando su coherencia y accesibilidad para su posterior análisis y generación de informes.

#### 2.2.2.3.1. Definir tablas de hechos y dimensiones / Reuniones técnicas.

Las tablas de hechos representan eventos o transacciones que se registran en el sistema y son el centro de un esquema de estrella o copo de nieve. Por otro lado, las dimensiones son atributos que proporcionan contexto a

las medidas almacenadas en las tablas de hechos, permitiendo un análisis más detallado y enriquecido de los datos.

Para el ICPNA, las tablas de hechos están definidas por: matrículas reserva, programación. Mientras que las dimensiones son: curso, alumno, profesor, aula, sede, horario. Para el área de Gestión Académica. Y se está trabajando con un centro de un esquema estrella.

En la parte de Anexos se puede revisar las tablas de hechos y dimensiones que se están contemplando para cada una de las áreas.

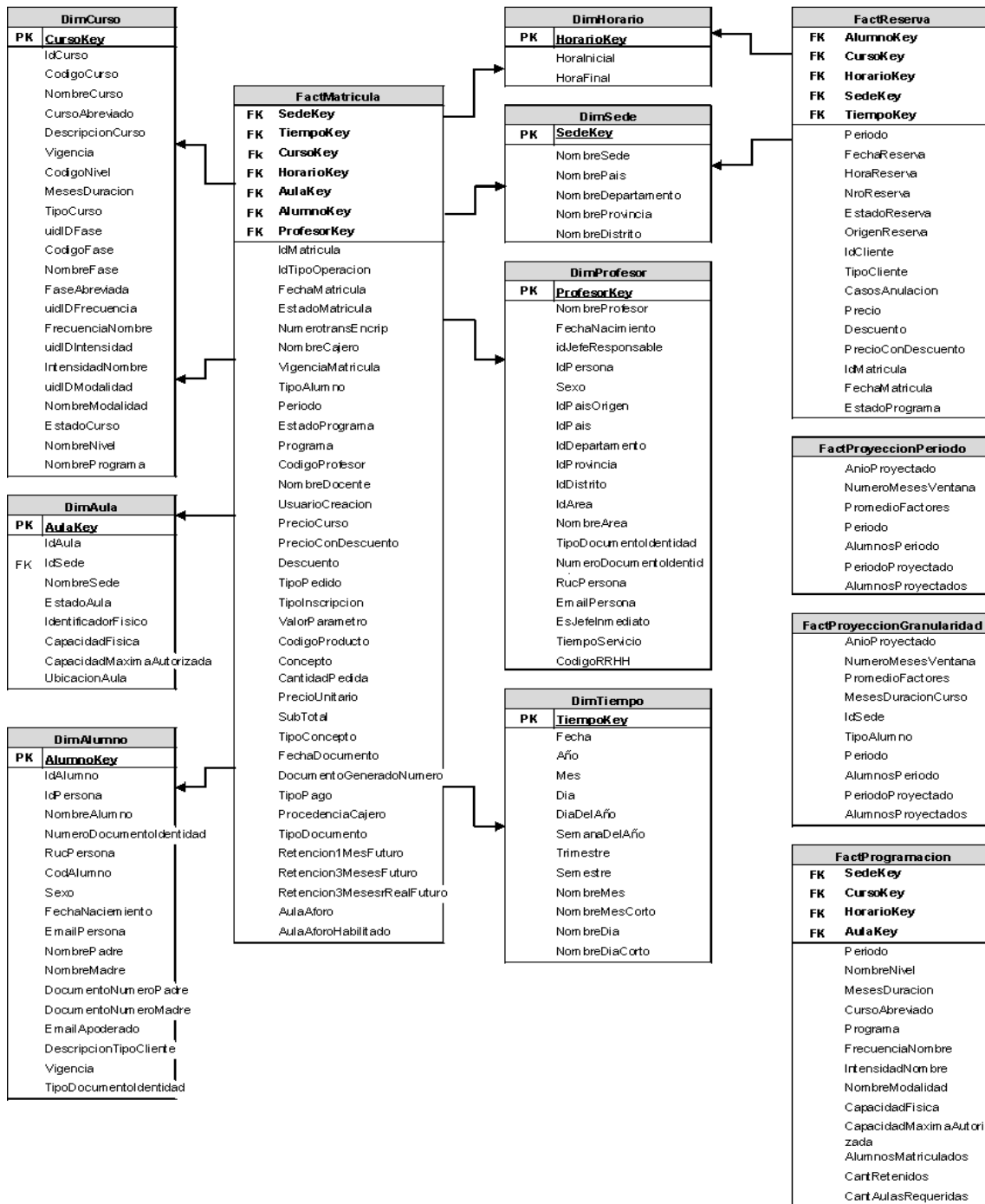
Anexo 02: Documento de tablas de hechos y dimensiones - Área Gestión Académica.

Anexo 03: Documento de tablas de hechos y dimensiones - Área Contabilidad y Finanzas.

Anexo 04: Documento de tablas de hechos y dimensiones - Área Comercial.

A continuación, la figura muestra uno de los diagramas.

**Figura 8: Tablas de hechos y dimensiones Área Gestión Académica**



Fuente: Dígito, 2023

### 2.2.2.3.2. Habilitar Servicios de Azure (Data Lake)

La habilitación de servicios de Azure, específicamente la configuración de un Azure Data Lake, es un componente crucial en el proceso de implementación del Azure Data Warehouse para el ICPNA. Este servicio

proporcionará un almacenamiento escalable y seguro para los datos que alimentarán el almacén de datos, permitiendo un acceso rápido y eficiente a grandes volúmenes de información.

A modo de resumen se puede indicar la siguiente tabla:

**Tabla 3: Componentes Azure**

Fase	Componente
Integración	Integration Runtime (IR) / Sharepoint 365
Movimiento de datos	Azure data Factory (ADF) / Azure Data Lake (ADLS)
Validaciones	Azure data Factory (ADF) / Azure Data Lake (ADLS)
Limpieza y transformación	Azure data Factory (ADF) / Azure Data Lake (ADLS)
Modelado	Azure data Factory (ADF) / Azure Data Lake (ADLS)
Datawarehousing	Azure SQL Database
Entrega	Azure data Factory (ADF) / Azure Data Lake (ADLS)

Fuente: *Dígito*, 2023

#### **2.2.2.4. Desarrollo Azure Data Warehouse (Hito4).**

El desarrollo del Azure Data Warehouse para el Instituto Cultural Peruano Norteamericano (ICPNA) es un paso crítico en el proceso de implementación de su solución de inteligencia de negocios. Esta etapa implicará el diseño físico del almacén de datos, así como la implementación de paquetes ETL utilizando la metodología Medallion para garantizar la extracción, transformación y carga eficientes de los datos desde las fuentes hasta el almacén de datos en Azure.

##### **2.2.2.4.1. Diseño Físico.**

El diseño físico del Azure Data Warehouse se llevará a cabo considerando los requisitos de almacenamiento, rendimiento y escalabilidad del ICPNA. Esto incluirá la definición de esquemas de bases de datos, tablas, índices y particiones para optimizar el rendimiento de las consultas y la eficiencia en el acceso a los datos. Además, se configurarán las opciones de almacenamiento y recursos de cómputo en Azure para garantizar un funcionamiento óptimo del almacén de datos.

Al comienzo del proceso de Ingesta de datos se realizó el desarrollo para la creación de esquemas y la tabla que serviría para invocar desde un



pipeline de Azure Data Factory para la ingesta de tablas que se detalla en los siguientes scripts:

**Figura 9: Creación de los esquemas y tabla para la ingesta de datos**

```
--CREACIÓN DE ESQUEMAS
CREATE SCHEMA intg; --ESQUEMA PARA TABLAS AUXILIARES DE INTEGRACION
go
CREATE SCHEMA stg; --ESQUEMA DE STAGE [priorizar uso de capa bronce en ADLS]
GO
CREATE SCHEMA dw; --ESQUEMA DE DATAWAREHOUSE DIM + FACT [Capa Oro]
GO

--TABLA AUXILIAR PARA CANALIZACION DE INGESTA
CREATE TABLE intg.TablasIngesta (
  IdTabla int NOT NULL identity(1,1),
  Grupo varchar(255) NOT NULL,
  NombreEsquema varchar(255) NOT NULL,
  NombreTabla varchar(255) NOT NULL,
  Estado bit,
  UsuarioCreacion varchar(255) NOT NULL,
  FechaCreacion date NOT NULL);
GO

--LAVE PRIMARIA DE LA TABLA
ALTER TABLE intg.TablasIngesta
ADD CONSTRAINT PK_intg_TablasIngesta
PRIMARY KEY (IdTabla);
GO

--VALOR POR DEFAULT PARA EL CAMPO UsuarioCreacion
ALTER TABLE intg.TablasIngesta
ADD CONSTRAINT DF_intg_TablasIngesta_UsuarioCreacion
DEFAULT suser_name() FOR UsuarioCreacion;
GO

--VALOR POR DEFAULT PARA EL CAMPO FechaCreacion
ALTER TABLE intg.TablasIngesta
ADD CONSTRAINT DF_intg_TablasIngesta_FechaCreacion
DEFAULT getdate() FOR FechaCreacion;
GO

--VALOR POR DEFAULT PARA EL CAMPO Estado
ALTER TABLE intg.TablasIngesta
ADD CONSTRAINT DF_intg_TablasIngesta_Estado
DEFAULT 1 FOR Estado;
GO

--INDICE UNICO PARA EVITAR DUPLICIDAD DE TABLAS
CREATE UNIQUE NONCLUSTERED INDEX IDX_UNC_TablasIngesta_Grupo_NombreEsquema_NombreTabla
ON intg.TablasIngesta (Grupo,NombreEsquema,NombreTabla);
GO
```

Fuente: Dígito, 2023

**Tabla 4: Recursos Azure Data Factory**

Tipo de recurso	Storage Account
Suscripción Azure	lcpnaedu

Nombre del recurso	adlsazurebidw (adlsazurebidwdev)
Versión	StorageV2 (general purpose v2)
Grupo de recursos	rg_azurebi_dev
Contenedores	Raw - work - log

Fuente: *Dígito*, 2023

#### 2.2.2.4.2. Diseño e Implementación Paquetes ETL (Metodología

##### Medallion).

La metodología Medallion será utilizada para el diseño e implementación de los paquetes ETL, asegurando la calidad y consistencia de los datos en todo el proceso de extracción, transformación y carga. Esto implicará las siguientes actividades:

- **Habilitar Servicios de Azure (Azure Data Factory, SQL Database):** Se configurarán y provisionarán los servicios de Azure necesarios, incluyendo Azure Data Factory para la orquestación de los flujos de datos y SQL Database para el almacenamiento de metadatos y procesamiento de consultas.
- **Actualizar Documento Técnico:** Se revisará y actualizará el documento técnico del proyecto, especificando la nomenclatura, storage accounts y almacenamiento requeridos para el Azure Data Warehouse. Esto garantizará una implementación coherente y documentada del almacén de datos.
- **Creación de directorios en Azure Data Factory:** Se crearán directorios en Azure Data Factory para organizar y gestionar los flujos de datos desde las fuentes hasta el almacén de datos.
- **Ingesta de datos desde las fuentes hacia Azure Data Factory:** Se configurarán los pipelines de datos en Azure Data Factory para la ingestión de datos desde las fuentes de origen hacia el almacén de datos en Azure, esto posterior a la ejecución del script mostrado en la Figura 9.

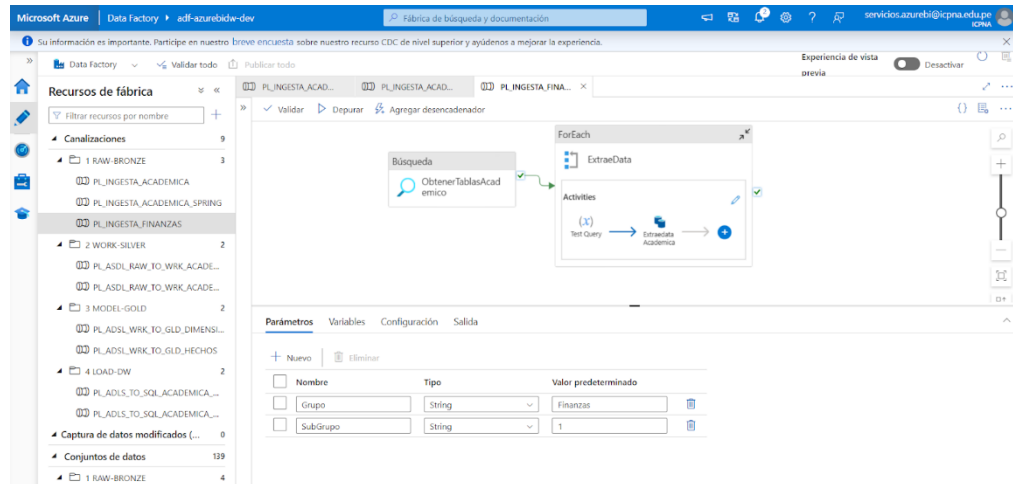
**Figura 10: Ingesta área Gestión Académica.**

The screenshot displays the Microsoft Azure Data Factory interface. On the left, the 'Recursos de fábrica' (Factory Resources) pane shows a tree view of pipelines, with 'PL\_INGESTA\_ACAD...' selected under the 'ACADEMICA' group. The main workspace shows a pipeline diagram with a 'Búsqueda' (Search) activity containing the text 'ObtenerTablasAcademico' connected to an 'ExtraeData' (Extract Data) activity. The 'ExtraeData' activity is nested within a 'ForEach' loop. Below the diagram, the 'Parámetros' (Parameters) tab is active, showing a table with columns for 'Nombre' (Name), 'Tipo' (Type), and 'Valor predeterminado' (Default Value).

Nombre	Tipo	Valor predeterminado
Grupo	String	Académica
SubGrupo	String	2

Fuente: Azure Data Factory ICPNA, 2023

**Figura 11:** Ingesta área Finanzas.



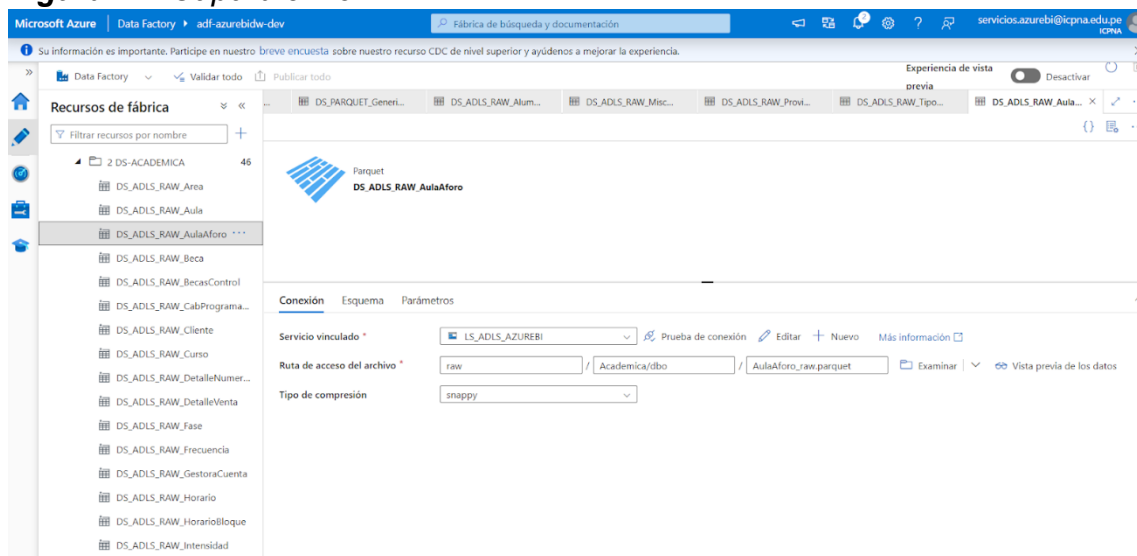
Fuente: Azure Data Factory ICPNA, 2023

- **Diseño e Implementación Paquetes ETL:** Se diseñarán y desarrollarán los paquetes ETL utilizando la metodología Medallion, abordando las siguientes áreas funcionales del ICPNA:
  - **Gestión Académica:** Se crearán datasets en la capa Bronce para la gestión académica, se realizará la limpieza y transformación de datos en la capa Plata, y se elaborarán y presentarán dimensiones y tablas de hechos en la capa Oro.
  - **Comercial:** Se seguirá un proceso similar al área de gestión académica para la creación de datasets, limpieza y transformación de datos, y elaboración de dimensiones y tablas de hechos en la capa Bronce, Plata y Oro respectivamente.
  - **Finanzas:** Se replicará el proceso para el área de finanzas, garantizando la creación, limpieza y transformación de datos, así como la elaboración de dimensiones y tablas de hechos en todas las capas del almacén de datos.

El desarrollo del Azure Data Warehouse utilizando la metodología Medallion asegurará una implementación eficiente y robusta del almacén de datos, proporcionando al ICPNA una plataforma sólida para el análisis y la toma de decisiones basadas en datos en todas sus áreas funcionales.

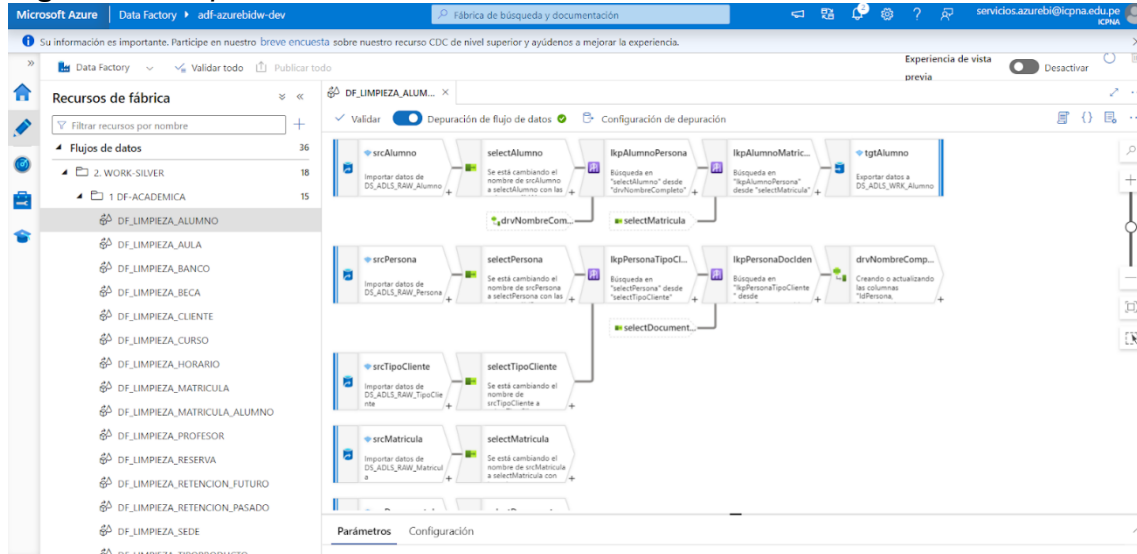
En el Anexo 06: ICPNA – Evidencias Azure Data Factory se puede observar el detalle del trabajo realizado en las tres capas de desarrollo.

**Figura 12: Capa bronce**



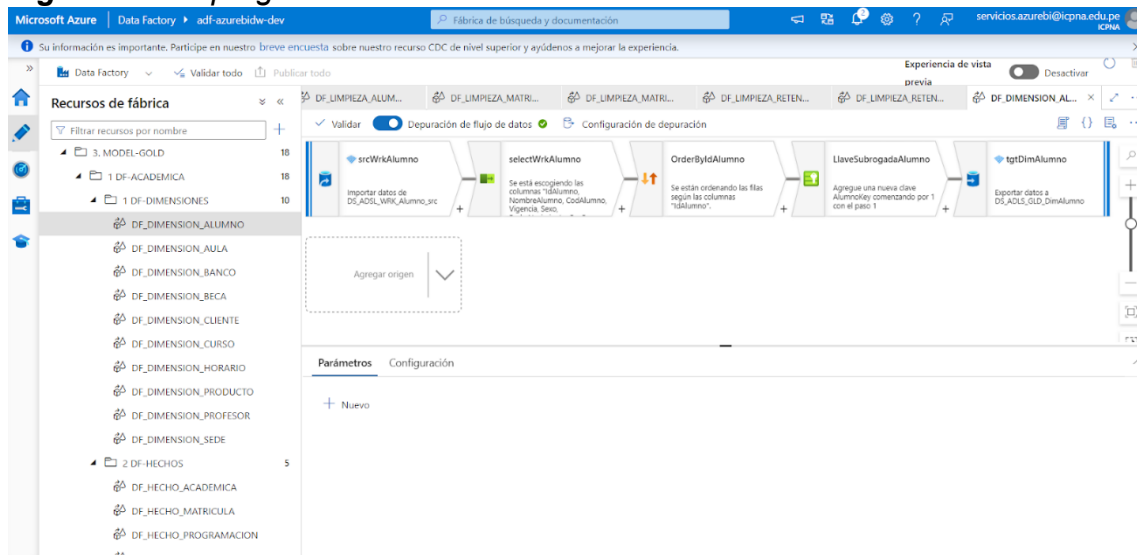
Fuente: Azure Data Factory ICPNA, 2023

**Figura 13: Capa silver**



Fuente: Azure Data Factory ICPNA, 2023

**Figura 14: Capa gold**



Fuente: Azure Data Factory ICPNA, 2023

### 2.2.2.5. Pruebas funcionales y entrega (Hito5).

En esta etapa del proyecto de implementación del Azure Data Warehouse para el Instituto Cultural Peruano Norteamericano (ICPNA), se llevaron a cabo las pruebas funcionales y la entrega del sistema instalado y configurado.

### 2.2.2.5.1. Pruebas Funcionales.

Las pruebas funcionales se realizaron para garantizar que el Azure Data Warehouse cumpliera con los requisitos de negocio y funcionara según lo esperado. Se llevaron a cabo las siguientes actividades:

- Planificación de las Pruebas:** Se planificaron las pruebas funcionales detallando los casos de prueba (reportes por sede, curso, aula), y los criterios de aceptación (revisión del encargado de cada área).
- Ejecución de Pruebas:** Se ejecutaron los casos de prueba diseñados, verificando la funcionalidad de cada componente en Azure DataFactory en diferentes escenarios y condiciones.
- Registro de Resultados:** Se registraron los resultados de las pruebas, identificando cualquier defecto o anomalía encontrada durante la ejecución de las pruebas.

Se elaboró una matriz de pruebas para comparar si el diseño dimensional cumplía con lo implementado en el Azure DataWarehouse.

**Figura 15: Matriz de pruebas**

N°	Fecha	Revisor	Prueba	TABLAS	COLUMNAS (Mapeo de Campos)	COLUMNAS (SQL Azure DW)	Resultado esperado (opcional)	Resultado	Observaciones
1	5/21/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	IdAlumno	AlumnoKey	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
2	5/22/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	Nombre	NombreAlumno	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
3	5/23/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	CodAlumno	CodAlumno	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
4	5/24/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	Sexo	Sexo	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
5	5/25/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	RUC	RucPersona	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
6	5/26/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	FechaNacimiento	FechaNacimiento	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
7	5/27/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	Correo	EmailPersona	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
8	5/28/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	NombrePadre	NombrePadre	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
9	5/29/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	NombreMadre	NombreMadre	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
10	5/30/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	DocumentoNumeroPadre	DocumentoNumeroPadre	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
11	5/31/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	DocumentoNumeroMadre	DocumentoNumeroMadre	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
12	6/1/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	CorreoApoderado	EmailApoderado	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
13	6/2/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	Tipo	DescripcionTipoCliente	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
14	6/3/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	Ninguna	Vigencia	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
15	6/4/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimesion Alumno	DNI	Pendiente	La columna no se encuentra en SQL Azure DW	Pendiente	Se definirá próximamente
16	6/5/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimension Curso	IdCurso	CursoKey	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
17	6/6/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimension Curso	NombreCurso	NombreCurso	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
18	6/7/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimension Curso	CursoAbreviado	CursoAbreviado	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
19	6/8/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimension Curso	FrecuenciaNombre	FrecuenciaNombre	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna
20	6/9/2023	Equipo ICNPA - TI	Comparación Modelo Dimensional - SQL Azure	Dimension Curso	IntensidadNombre	IntensidadNombre	La columna se encuentra en SQL Azure DW	Exitoso	Ninguna

Fuente: Dígito, 2023

Así mismo se realizó una comparación para hallar diferencias comparando reportes elaborados por el ICPNA y los reportes obtenidos del Azure DataWarehouse

**Figura 16: Reporte de diferencias por sede en el periodo 202305.**

Periodo	IdSede	IdSede2	NombreSede	DIGITO	ICPNA	DIFERENCIA
202305	8BA65461-60B5-4716-BEB3-22BC7B71BC09	8BA65461-60B5-4716-BEB3-22BC7B71BC09	La Molina	3113	3109	4
202305	8F67B2B3-667E-4EBC-8605-766D2FF71B55	8F67B2B3-667E-4EBC-8605-766D2FF71B55	Lima Centro	2701	2699	2
202305	257AD17D-91F7-48C8-81D4-71EBD35A4E5C	257AD17D-91F7-48C8-81D4-71EBD35A4E5C	Lima Norte	6931	6928	3
202305	26894D3F-E9E1-476C-9314-764DC0BCD003	26894D3F-E9E1-476C-9314-764DC0BCD003	Lima Norte Satélite	272	271	1
202305	7379A7D3-6DC5-42CA-9ED4-97367519F1D9	7379A7D3-6DC5-42CA-9ED4-97367519F1D9	Miraflores	1890	1890	0
202305	1CE8E5F9-56D2-4C35-B3C1-ED0D91C3D4B1	1CE8E5F9-56D2-4C35-B3C1-ED0D91C3D4B1	Provincias - Chimbote	190	189	1
202305	CE894D3F-E9E1-476C-9314-764DC0BCD003	CE894D3F-E9E1-476C-9314-764DC0BCD003	Provincias - Chinchá	63	63	0
202305	6944EC08-1CA5-40D2-9576-850DE912DBEF	6944EC08-1CA5-40D2-9576-850DE912DBEF	Provincias - Ica	626	626	0
202305	EF7BF999-E359-40C1-A712-BCA0450888F4	EF7BF999-E359-40C1-A712-BCA0450888F4	Provincias - Iquitos	858	855	3
202305	30DE73B6-8203-4F81-96C8-3B27977BB924	30DE73B6-8203-4F81-96C8-3B27977BB924	San Miguel	3194	3193	1
202305	FE3AABDB-531D-4601-A3D4-E4F697335806	FE3AABDB-531D-4601-A3D4-E4F697335806	Surco	4001	4000	1
202305	27894D3F-E9E1-476C-9314-764DC0BCD003	27894D3F-E9E1-476C-9314-764DC0BCD003	VECOR	172	172	0
202305	28894D3F-E9E1-476C-9314-764DC0BCD003	28894D3F-E9E1-476C-9314-764DC0BCD003	Virtual	35296	35262	34

Fuente: Dígito, 2023

**Figura 17: Reporte de diferencias por curso en el periodo 202305.**

Periodo	IdCurso	uidIdCurso	NombreCurso	DIGITO	ICPNA	DIFERENCIA
202305	8B9485DD-FAA2-4CCF-8B60-D8D0648A8841	8B9485DD-FAA2-4CCF-8B60-D8D0648A8841	(AG1RV) Advanced Grammar 1 Diario - CLASES A DISTANCIA	727	725	2
202305	8B9485DD-FAA2-4CCF-8B60-D8D0648A8847	8B9485DD-FAA2-4CCF-8B60-D8D0648A8847	(AR1RV) Advanced Reading & Writing 1 Diario - CLASES A DISTANCIA	542	540	2
202305	0666A544-5A17-4F86-8791-7736EE865EE1	0666A544-5A17-4F86-8791-7736EE865EE1	(B00) Introducción al inglés - PRESENCIAL	1438	1436	2
202305	8B9485DD-FAA2-4CCF-8B60-D8D0648A8802	8B9485DD-FAA2-4CCF-8B60-D8D0648A8802	((OSRV) Intermediate Five Diario - CLASES A DISTANCIA	799	797	2
202305	74DC63DF-DAF9-4CC0-84A8-04ECSA702FA0	74DC63DF-DAF9-4CC0-84A8-04ECSA702FA0	(JRO1S) Juniors 1 Saturday - PRESENCIAL	415	413	2

Fuente: Dígito, 2023

**Figura 18: Reporte de diferencias por aula en el periodo 202209.**

Periodo	IdAula	uidIdAula	Aula	DIGITO	ICPNA	DIFERENCIA
202209	B5153E1E-35B8-4179-B7EF-AC64F6EC03F0	B5153E1E-35B8-4179-B7EF-AC64F6EC03F0	1003	20	20	0
202209	03D01A0C-ECE1-4CE4-9ECE-69688BF0E32A	03D01A0C-ECE1-4CE4-9ECE-69688BF0E32A	101	101	101	0
202209	8E273401-8147-43AF-ADDC-FC25A387A0E9	8E273401-8147-43AF-ADDC-FC25A387A0E9	101	40	40	0
202209	E6C10832-218E-4FA4-8FB1-14957E933BAB	E6C10832-218E-4FA4-8FB1-14957E933BAB	101	32	32	0
202209	00908B28-3B9E-4700-91FC-FOC3420D037D	00908B28-3B9E-4700-91FC-FOC3420D037D	102	19	19	0
202209	718BDC34-E66F-402A-A5E6-688ACE6B08C1	718BDC34-E66F-402A-A5E6-688ACE6B08C1	102	29	29	0
202209	ADF30994-833A-4B75-A752-937F986D0ABB	ADF30994-833A-4B75-A752-937F986D0ABB	102	109	109	0
202209	37A030AE-7B07-42ED-A5EF-A770C4BD9C98	37A030AE-7B07-42ED-A5EF-A770C4BD9C98	103	21	21	0
202209	B9ACD9B6-13A1-451E-9B01-9AEB68A2B6B5	B9ACD9B6-13A1-451E-9B01-9AEB68A2B6B5	103	133	133	0
202209	C6EE47F6-D928-4B92-B52E-6D95783AEF4F	C6EE47F6-D928-4B92-B52E-6D95783AEF4F	103	22	22	0
202209	14A33ECF-C217-4192-A9A3-E0057C768CF4	14A33ECF-C217-4192-A9A3-E0057C768CF4	104	12	12	0
202209	3318F679-4EC7-4F34-86A9-C5120BD0BE72	3318F679-4EC7-4F34-86A9-C5120BD0BE72	104	19	19	0
202209	DE3C5A90-BA76-42CA-B85E-496FB7CCD934	DE3C5A90-BA76-42CA-B85E-496FB7CCD934	104	60	60	0
202209	454F25E7-C8CF-4526-A70F-8AD07A27C9AA	454F25E7-C8CF-4526-A70F-8AD07A27C9AA	105	80	80	0
202209	4C20D5CA-4E76-43F7-8E15-ADB2072B0D03	4C20D5CA-4E76-43F7-8E15-ADB2072B0D03	106	75	75	0
202209	128771FA-493E-4EDD-AF93-3094115C6187	128771FA-493E-4EDD-AF93-3094115C6187	107	54	54	0
202209	C68CD825-D38A-4086-A966-2A7D8E8CC99D	C68CD825-D38A-4086-A966-2A7D8E8CC99D	107	27	27	0
202209	291A2C4D-A4E3-4305-916F-C46ACA44ADC8	291A2C4D-A4E3-4305-916F-C46ACA44ADC8	108	100	100	0
202209	2358E099-687C-40F1-96A5-D3B62A233007	2358E099-687C-40F1-96A5-D3B62A233007	109	78	78	0
202209	DFB3026B-E498-4DC9-BD40-711D95415BC6	DFB3026B-E498-4DC9-BD40-711D95415BC6	110	70	70	0
202209	104CFE8B-097A-4188-B68A-2161CD9E83E5	104CFE8B-097A-4188-B68A-2161CD9E83E5	201	79	79	0
202209	201C919E-CA29-4EC7-B362-E8C4D14D84EE	201C919E-CA29-4EC7-B362-E8C4D14D84EE	201	58	58	0
202209	6AF47C10-7828-4094-B3DA-5C96B9215001	6AF47C10-7828-4094-B3DA-5C96B9215001	201	53	53	0
202209	90525A49-4B94-4F3A-88C6-A99FE0578E70	90525A49-4B94-4F3A-88C6-A99FE0578E70	201	21	21	0
202209	9E05B1ED-68FF-4697-8180-1DC2E46475D5	9E05B1ED-68FF-4697-8180-1DC2E46475D5	201	14	14	0
202209	202C919E-CA29-4EC7-B362-E8C4D14D84EE	202C919E-CA29-4EC7-B362-E8C4D14D84EE	202	12	12	0
202209	6AF47C10-7828-4094-B3DA-5C96B9215002	6AF47C10-7828-4094-B3DA-5C96B9215002	202	36	36	0
202209	779445CD-6D1F-466A-B9B8-225CD5E85AD6	779445CD-6D1F-466A-B9B8-225CD5E85AD6	202	107	107	0
202209	203C919E-CA29-4EC7-B362-E8C4D14D84EE	203C919E-CA29-4EC7-B362-E8C4D14D84EE	203	13	13	0
202209	6AF47C10-7828-4094-B3DA-5C96B9215003	6AF47C10-7828-4094-B3DA-5C96B9215003	203	22	22	0
202209	F69CD166-68CB-45E1-B2F1-52527729734B	F69CD166-68CB-45E1-B2F1-52527729734B	203	68	68	0
202209	6AF47C10-7828-4094-B3DA-5C96B9215004	6AF47C10-7828-4094-B3DA-5C96B9215004	204	78	78	0
202209	A6DD1AAF-97BE-409A-8F0E-8BAD81A5701E	A6DD1AAF-97BE-409A-8F0E-8BAD81A5701E	204	76	76	0
202209	65F8B23A-A44A-4C34-9732-BCA9CC9B4D9A	65F8B23A-A44A-4C34-9732-BCA9CC9B4D9A	205	108	108	0
202209	6AF47C10-7828-4094-B3DA-5C96B9215005	6AF47C10-7828-4094-B3DA-5C96B9215005	205	99	99	0
202209	6AF47C10-7828-4094-B3DA-5C96B9215006	6AF47C10-7828-4094-B3DA-5C96B9215006	206	54	54	0
202209	A5BB7B0B-6DD1-43A9-93FE-4B4F0643C972	A5BB7B0B-6DD1-43A9-93FE-4B4F0643C972	206	145	145	0

Fuente: Dígito, 2023



- d. **Corrección de Defectos:** Se corrigieron los defectos identificados durante las pruebas, realizando pruebas adicionales para verificar que las correcciones se implementaron correctamente.
- e. **Aprobación de Pruebas:** Una vez completadas las pruebas y corregidos los defectos, se obtuvo la aprobación de las pruebas funcionales por parte del equipo gerencial del ICPNA y los stakeholders de cada área.

#### 2.2.2.6. Entrega y Cierre

La fase de entrega y cierre del proyecto de implementación del Azure Data Warehouse para el ICPNA se realizó con éxito, abordando las siguientes actividades:

**Capacitaciones sobre el Uso de Azure:** Se llevaron a cabo sesiones de capacitación para el personal del ICPNA (para el encargado de DataLake) sobre el uso de Azure y la actualización de los componentes del Data Warehouse. Estas capacitaciones se diseñaron para proporcionar a los usuarios las habilidades y el conocimiento necesarios para administrar y utilizar eficazmente el entorno de Azure, así como para realizar actualizaciones y mantenimiento del Data Warehouse.

Durante estas sesiones, se cubrieron los siguientes temas:

- a. **Introducción a Azure:** Se proporcionó una visión general de los servicios de Azure, destacando su importancia y beneficios para el ICPNA.
- b. **Gestión del Azure Data Warehouse:** Se enseñaron las técnicas y mejores prácticas para gestionar y administrar el Data Warehouse en Azure, incluyendo la configuración de recursos, la monitorización del rendimiento y la optimización de la carga de trabajo.
- c. **Actualización de Componentes:** Se detallaron los pasos necesarios para realizar actualizaciones y mantenimiento de los componentes del Data Warehouse, asegurando que el personal estuviera preparado para llevar a cabo estas tareas de manera efectiva y segura.

**Cierre de Proyecto:** Una vez completadas las capacitaciones y asegurado que el personal del ICPNA estaba equipado con las habilidades necesarias, se procedió al cierre oficial del proyecto. Esta etapa incluyó:

- a. **Documentación Final:** Se preparó y entregó toda la documentación final del proyecto, incluyendo los documentos que se muestran en los anexos, guías de referencia y cualquier otra documentación relevante para el uso y mantenimiento del Azure Data Warehouse.
- b. **Evaluación del Proyecto:** Se llevó a cabo una evaluación exhaustiva del proyecto para identificar lecciones aprendidas, éxitos y áreas de mejora. Esta retroalimentación se utilizó para informar futuros proyectos (la utilización de modelos de machine learning) y mejorar las prácticas de implementación.
- c. **Agradecimientos y Reconocimientos:** Se expresaron agradecimientos y reconocimientos a todos los involucrados en el proyecto, incluyendo al equipo de proyecto, al personal del ICPNA y al equipo de Dígito que contribuyó al éxito del proyecto.

Con la entrega de las capacitaciones sobre el uso de Azure y el cierre oficial del proyecto, se finalizó con éxito la implementación del Azure Data Warehouse en el ICPNA, dejando a la organización una poderosa herramienta para la gestión y análisis de datos.

### 2.3. Conclusiones

- El primer objetivo específico se logró de manera efectiva. La implementación del Azure Data Lake permitió integrar datos de diversas fuentes (SQL Server, SAP, Dynamics 360, Sharepoint) en una única ubicación, facilitando el acceso y la gestión centralizada de la información. Este avance no solo mejoró la visibilidad y transparencia de los datos, sino que también simplificó los procesos de búsqueda y recuperación de información tal como se muestra en las **figuras 12 - 14.**
- El segundo objetivo específico se consiguió al eliminar los errores asociados con la manipulación manual de datos dispersos y automatizar el proceso de generación de informes, se han reducido significativamente las inexactitudes y se ha incrementado la confiabilidad de la información presentada. Los usuarios ahora confían en la integridad de los datos para la toma de decisiones estratégicas tal como se mostró en las pruebas realizadas en el **punto 3.2.5.**

## 2.4. Recomendaciones

- Para mantener la efectividad del Azure Datawarehouse, es crucial realizar un mantenimiento regular y asegurarse de que todas las fuentes de datos estén actualizadas y sincronizadas. Se recomienda establecer un calendario de mantenimiento periódico y asignar personal dedicado para supervisar este proceso.
- Se recomienda implementar herramientas de monitoreo y análisis de rendimiento para el Data Warehouse. Esto permitirá identificar y resolver posibles cuellos de botella y optimizar el uso de recursos. El monitoreo continuo asegurará que el sistema funcione de manera eficiente y que cualquier problema sea detectado y resuelto rápidamente.
- Con el éxito de la centralización de datos y la mejora en la generación de informes, se sugiere explorar el desarrollo de nuevas funcionalidades que puedan agregar valor adicional. Esto podría incluir la implementación de herramientas de análisis avanzadas, como inteligencia artificial y machine learning, para obtener insights más profundos y predecir tendencias futuras.
- Establecer un sistema de feedback regular donde los usuarios puedan reportar problemas, sugerir mejoras y compartir sus experiencias con el sistema. Este feedback es valioso para realizar mejoras continuas y adaptar el sistema a las necesidades cambiantes de la organización.

## GLOSARIO DE TERMINOS

**Azure Data Factory (ADF):** Servicio de integración de datos basado en la nube que permite crear, programar y orquestar flujos de trabajo de datos a gran escala en Azure.

**Data Lake:** Repositorio centralizado que permite almacenar grandes volúmenes de datos en su formato original, estructurados y no estructurados.

**Data Warehouse (DWH):** Sistema de almacenamiento de datos diseñado para consultas y análisis de datos a gran escala, centralizando información de múltiples fuentes.

**ETL (Extract, Transform, Load):** Proceso de extracción de datos de múltiples fuentes, transformación según las necesidades de análisis y carga en un sistema de almacenamiento como un Data Warehouse.

**Medallion Architecture:** Metodología de diseño de Data Lake que organiza los datos en tres capas (Bronce, Plata y Oro) para gestionar y procesar datos de manera eficiente.

- Bronce: Capa de datos crudos sin procesar.
- Plata: Capa de datos limpiados y transformados.
- Oro: Capa de datos optimizados y preparados para análisis y reporting.

**Paquetes ETL:** Conjuntos de procesos y herramientas diseñados para realizar las operaciones de ETL, manejando grandes volúmenes de datos desde su origen hasta su destino final.

**Pipeline:** Flujo de trabajo que define una secuencia de actividades en Azure Data Factory para mover y transformar datos.

**Linked Services:** Conexiones configuradas en Azure Data Factory que permiten a los pipelines conectarse a diversas fuentes de datos y servicios.

**Dataset:** Representación de datos en Azure Data Factory que define la estructura y ubicación de los datos utilizados en los pipelines.

## REFERENCIAS

Codd, E. F. (1970). A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, 13(6), 377-387. doi:10.1145/362384.362685

Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling* (3rd ed.). Wiley.

Randstad Fundación. (2023). *Las fases y características de la metodología de cascada*.

Sadalage, P. & Fowler, M. (2012). *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence*. Addison-Wesley.

Snowflake Inc. (2021). *Medallion Architecture: A New Paradigm for Data Management*.

(<https://community.snowflake.com/s/question/0D5VI00000CVzQG0A1/medallion-architecture-in-snowflake>)

Pressman, R. S. (2014). *Software Engineering: A Practitioner's Approach* (8th ed.). McGraw-Hill Education.

Microsoft. (2020). *Azure Data Factory Documentation*. Retrieved from <https://docs.microsoft.com/en-us/azure/data-factory/>

Microsoft Azure. (2024). *Azure Data Factory documentation*. (<https://learn.microsoft.com/en-us/azure/data-factory/>)

## ANEXOS

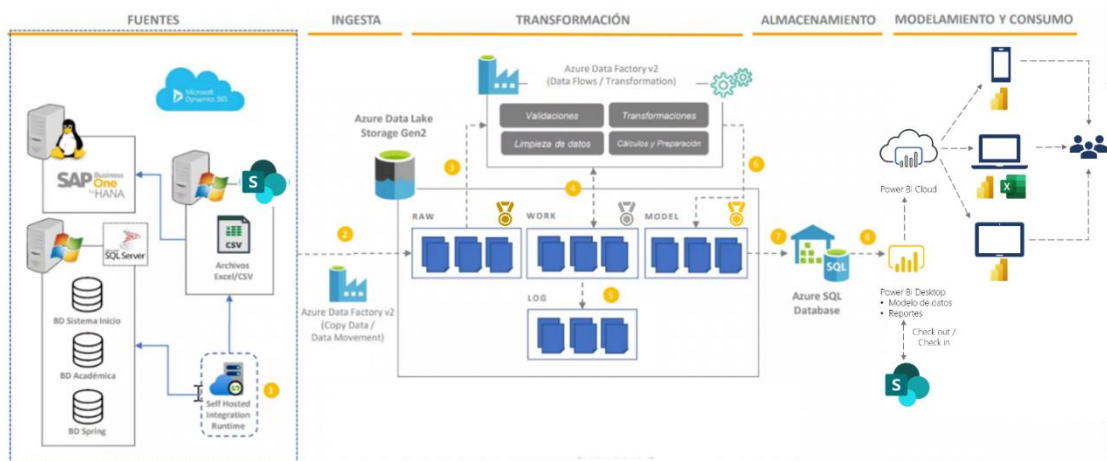
### Anexo 01: ICPNA - Documento de Diseño de Arquitectura v3

#### Objetivo principal

Identificar y describir la arquitectura propuesta para el proyecto Azure Datawarehouse, que incluye las fuentes de datos y los componentes de la solución. En este documento se listan y se explican cada uno de los elementos de la solución, para una comprensión profunda de cada uno de ellos. Este documento servirá de base para la implementación física de la solución.

#### Diagrama de arquitectura propuesto

A continuación, se muestra la arquitectura propuesta por el equipo de Dígito



Fases vs. Componentes

A continuación, se muestra una tabla con la relación entre las fases del proceso y los componentes de la solución a ser utilizados y que son detallados más adelante:

Fase	Componente
------	------------

Integración	Integration Runtime (IR) / Sharepoint 365
Movimiento de datos / Ingesta	Azure data Factory (ADF) / Azure Data Lake (ADLS)
Validaciones	Azure data Factory (ADF) / Azure Data Lake (ADLS)
Limpieza y transformación	Azure data Factory (ADF) / Azure Data Lake (ADLS)
Registros	Azure data Factory (ADF) / Azure Data Lake (ADLS)
Modelado	Azure data Factory (ADF) / Azure Data Lake (ADLS)
Datawarehousing	Azure SQL Database
Desarrollo de reportes	Power BI Desktop / Sharepoint 365
Entrega	Power BI Service

### **Fases del proceso:**

#### **Integración**

Establece el enlace entre las fuentes de datos On-Premises con los servicios de nube.

#### **Movimientos de datos**

Implementación de canalizaciones de datos, para la lectura de las fuentes y traslado de datos al repositorio correspondiente a los datos de tipo RAW (datos crudos).

#### **Validaciones**

Implementación de flujos de datos en los que se aplican las primeras reglas, correspondientes a la validación, para garantizar que los datos sean correctos y consistentes.



## **Limpieza y transformación**

Implementación de flujos de datos donde se aplican las reglas definitivas de limpieza y transformación. Se obtienen los datos listos para ser modelados (WORK).

## **Registros**

Los flujos de datos de validación, limpieza y transformación cuentan con procesos de registro que permiten mantener una historia y conocer el estado de los procesos, de modo que se pueda monitorear para mantener una historia y detectar posibles fallas en el flujo.

## **Modelado**

Implementación de los flujos que cargan los datos en el modelo final (MODEL), previo a la carga en el Datawarehouse.

## **Datawarehousing**

Almacenamiento del modelo de datos definitivo. Es la fase final en el procesamiento de los datos y en el que se obtiene la “única versión de la verdad”.

Desarrollo de reportes

Creación del contenido visual e interactivo que serán publicados en el servicio para ser consumido y compartido por los usuarios.

## **Entrega**

Fase de implementación de los espacios de trabajo y roles para el consumo de los reportes por parte de los usuarios.

Inventario de fuentes de datos

A continuación, se muestra una tabla detallando las fuentes de datos que se han mapeado.

N.	Nombre	Sistema	Tema	Tipo	Plataforma
1	dynamics	Microsoft Dynamics 365	Comercial	Cloud	Dynamics
2	sap	SAP Business One	Finanzas	On premise	SAP Hana
3	inicio	Aplicativo inicio	Académico/Operaciones	On premise	SQL server
4	academica	Sistema académico	Académico	On premise	SQL server
5	spring	Sistema Spring	Financiero (Legacy)	On premise	SQL server
6	Check list ppto alumnos	n.a.	Operaciones	Cloud / Sharepoint	Excel
7	Clases canceladas	n.a.	Operaciones	Cloud / Sharepoint	Excel
8	GAP Oferta y capacidad docente	n. .a.	Operaciones	Cloud / Sharepoint	Excel
9	nuevos_proy_23 v1	n.a.	Operaciones	Cloud / Sharepoint	Excel
10	rentencion_alum nos	n.a.	Académico	Cloud / Sharepoint	Excel

11	clases_desagregadas	n.a.	Académico	Cloud / Sharepoint	Excel
13	00.Distribución.xlsx	n.a.	Académica	Cloud / Sharepoint	Excel
14	Proyeccion anual 2023_Aprobado.xlsx	n.a.	Académica	Cloud / Sharepoint	Excel
15	ppto_alumnos_nuevos	n.a.	Comercial	Cloud / Sharepoint	Excel
16	asesoras	n.a.	Comercial	Cloud / Sharepoint	Excel
17	Plantilla de GyP – Agrupación de Sedes	n.a.	Finanzas	Cloud / Sharepoint	Excel
18	Plantilla de GyP – Agrupación de Centros de Costo	n.a.	Finanzas	Cloud / Sharepoint	Excel
19	Plantilla de GyP – Agrupación de Gerencia	n.a.	Finanzas	Cloud / Sharepoint	Excel
20	Plantilla de GyP – Agrupación de Cuenta Contables (incluye equivalencia SAP-SPRING)	n.a.	Finanzas	Cloud / Sharepoint	Excel
21	Plantilla de GyP – Agrupación de Niveles por cuenta contable	n.a.	Finanzas	Cloud / Sharepoint	Excel

22	valorizacion_des agregada	n.a.	Finanzas	Cloud / Sharepoint	Excel
23	1.0 PRESUPUESTO VENTAS MODALIDAD BLENDED	n.a.	Finanzas	Cloud / Sharepoint	Excel
24	1.0 PRESUPUESTO VENTAS MODALIDAD VIRTUAL	n.a.	Finanzas	Cloud / Sharepoint	Excel
25	1.0. PRESUPUESTO VENTAS MODALIDAD PRESENCIAL	n.a.	Finanzas	Cloud / Sharepoint	Excel
26	1.1 PRESUPUESTO BECAS	n.a.	Finanzas	Cloud / Sharepoint	Excel
27	1.1 PRESUPUESTO CONVENIOS y CAMPAÑAS	n.a.	Finanzas	Cloud / Sharepoint	Excel
28	2.1 PRESUPUESTO VENTAS EXÁMENES	n.a.	Finanzas	Cloud / Sharepoint	Excel
29	2.1 PRESUPUESTO VENTAS 2023 -	n.a.	Finanzas	Cloud / Sharepoint	Excel

	TRADUCCIONES				
30	2.2 PRESUPUESTO VENTAS 2023 - BIBLIOTECA	n.a.	Finanzas	Cloud / Sharepoint	Excel
31	PRESUPUESTO VENTAS 2023 - CULTURAL	n.a.	Finanzas	Cloud / Sharepoint	Excel
32	2.4 PRESUPUESTO VENTAS 2023 - SERVICIOS ADMINISTRATIVOS	n.a.	Finanzas	Cloud / Sharepoint	Excel
33	2.5 PRESUPUESTO VENTAS 2023 - CONCESIONARIOS	n.a.	Finanzas	Cloud / Sharepoint	Excel
34	2.6 PRESUPUESTO VENTAS 2023 - ACPNA	n.a.	Finanzas	Cloud / Sharepoint	Excel
35	2.7 PRESUPUESTO VENTAS 2023 - VENTAS CORPORATIVO	n.a.	Finanzas	Cloud / Sharepoint	Excel
36	3.1 PRESUPUESTO	n.a.	Finanzas	Cloud / Sharepoint	Excel

	2023 - COSTO DOCENTES				
37	4.1 PRESUPUESTO 2023 - PERSONAL ADMINISTRATI VO	n.a.	Finanzas	Cloud / Sharepoint	Excel
38	5.1 PRESUPUESTO 2023 - GASTOS DE GESTIÓN - GERENCIA GENERAL	n.a.	Finanzas	Cloud / Sharepoint	Excel
39	5.2 PRESUPUESTO 2023 - GASTOS DE GESTIÓN - GERENCIA ACADEMICA	n.a.	Finanzas	Cloud / Sharepoint	Excel
40	5.3.1 PRESUPUESTO GASTOS DE GESTIÓN - GERENCIA DE ADMINISTRACI ON Y FINANZAS	n.a.	Finanzas	Cloud / Sharepoint	Excel
41	5.4 PRESUPUESTO GASTOS DE GESTIÓN - GERENCIA DE	n.a.	Finanzas	Cloud / Sharepoint	Excel

	TECNOLOGIA Y COMUNICACIÓN				
42	5.5 PRESUPUESTO GASTOS DE GESTIÓN - GERENCIA DE OPERACIONES	n.a.	Finanzas	Cloud / Sharepoint	Excel
43	5.6 PRESUPUESTO GASTOS DE GESTIÓN - GERENCIA DE COMERCIAL Y MARKETING	n.a.	Finanzas	Cloud / Sharepoint	Excel
44	5.7 PRESUPUESTO GASTOS DE GESTIÓN - GERENCIA DE BIBLIOTECA	n.a.	Finanzas	Cloud / Sharepoint	Excel
45	5.8 PRESUPUESTO GASTOS DE GESTIÓN - GERENCIA RRHH	n.a.	Finanzas	Cloud / Sharepoint	Excel
46	5.9 PRESUPUESTO GASTOS DE GESTIÓN -	n.a.	Finanzas	Cloud / Sharepoint	Excel

	GERENCIA LEGAL				
47	5.10 PRESUPUESTO GASTOS DE GESTIÓN - GERENCIA CULTURAL	n.a.	Finanzas	Cloud / Sharepoint	Excel
48	5.11 PRESUPUESTO GASTOS DE GESTIÓN - AUDITORIA INTERNA	n.a.	Finanzas	Cloud / Sharepoint	Excel
49	4.2 PRESUPUESTO SERVICIOS GENERALES	n.a.	Finanzas	Cloud / Sharepoint	Excel
50	4.3 PRESUPUESTO GASTOS DE PUBLICIDAD	n.a.	Finanzas	Cloud / Sharepoint	Excel
51	4.4 PRESUPUESTO SEGURIDAD INTEGRAL	n.a.	Finanzas	Cloud / Sharepoint	Excel
52	4.5 PRESUPUESTO SERVICIO DE RECAUDACIÓN	n.a.	Finanzas	Cloud / Sharepoint	Excel



53	4.6	n.a.	Finanzas	Cloud / Sharepoint	Excel
	PRESUPUESTO				
	SEGUROS				
	GASTOS BANCARIOS				
54	4.7	n.a.	Finanzas	Cloud / Sharepoint	Excel
	PRESUPUESTO				
	GASTOS COMUNES TIC				
55	7.2	n.a.	Finanzas	Cloud / Sharepoint	Excel
	PRESUPUESTO				
	INTERESES DE QUINQUENIO				

Nota: En el caso de las fuentes de datos de Excel, en la etapa de desarrollo se determinará si el archivo se utilizará como fuente o como plantilla para procesos y reportes

Componentes de la solución

Sharepoint 365

Es un servicio basado en la nube, hospedado y gestionado por Microsoft que va a cumplir diferentes propósitos en el diseño de arquitectura propuesto:

### **Manejo de fuentes de información**

Se va a usar como un lugar seguro donde almacenar, organizar y acceder a la información semiestructurada, disponible en archivos (Excel, texto, etc.).

### **Rutina de integración (Integration Runtime)**

Es la infraestructura de proceso que usan las canalizaciones de Azure Data Factory para integrar fuentes On-Premises con los servicios en la nube de Azure.

Integration Runtime se instala como una puerta de enlace en un equipo que esté conectado a la red local (LAN) de la organización y que tenga visibilidad de las fuentes de datos On-Premise, desde donde se van a extraer los datos que serán utilizados en los siguientes pasos del flujo.

### **Azure Data Lake**

Azure Data Lake es un servicio gestionado en Azure que se utiliza como repositorio para cargas de trabajo de análisis de macrodatos. Azure Data Lake permite almacenar datos de cualquier tamaño y tipo de ingesta en un único lugar para, realizar análisis exploratorios y operativos. Esta implementación permite el escalado flexible para gestionar las cargas de trabajo en el tiempo, además de garantizar una alta disponibilidad de acuerdo con los niveles de servicio implementados. Se usará Azure Data Lake para almacenar toda la información requerida en el diseño del Datawarehouse y estará conformado por contenedores de información de tipo RAW (datos crudos), WORK (datos trabajados) y MODEL (datos modelados), así como los tipos de datos LOG (datos de registro) que representan el detalle e historia de los procesos ejecutados.

### **Azure Data Factory**

Azure Data Factory es una solución de integración de datos que contiene una serie de sistemas interconectados que proporcionan una plataforma completa de un extremo a otro para los ingenieros de datos. Azure Data Factory es un servicio gestionado en la nube, creado para implementar complejos proyectos híbridos de extracción, transformación y carga (ETL), extracción, carga y transformación (ELT) e integración de datos. Se usará ADF para el movimiento de los datos desde la fuente hacia el datawarehouse y los dos componentes claves a utilizar son:

## **Canalizaciones**

Las canalizaciones de ADF es una agrupación lógica de actividades para realizar una unidad de trabajo de tipo extracción y carga. Las canalizaciones se utilizan en la fase de ingesta de datos y permitirá acceder a las fuentes y mover los datos al destino de almacenamiento, en los contenedores de datos de tipo RAW (datos crudos) en un formato adecuado para ser utilizados en los siguientes pasos del flujo.

## **Flujo de datos**

El flujo de datos de ADF contempla un entorno gráfico para construir lógica de transformación de datos que puede usar para limpiar y transformar datos de cualquier tipo y tamaño. Los flujos de datos se utilizan en las etapas de Validaciones, limpieza y transformación, con el objetivo de promover los datos almacenados a los contenedores de datos de tipo WORK (datos trabajados) y de tipo MODEL (datos modelados). Adicionalmente, los flujos de datos también contemplan el almacenamiento de datos de tipo LOG (datos de registro) para tener un registro completo del historial de actividades ejecutadas en el proceso.

## **Azure SQL Database**

Opción de implementación PaaS de Azure SQL. Permite obtener una base de datos y consumirla como un servicio de datos en la nube, hospedado y gestionado por Azure. Abstrae tanto el sistema operativo como la instancia de SQL Server. Esta implementación permite el escalado flexible para gestionar las cargas de trabajo en el tiempo, además de garantizar una alta disponibilidad de acuerdo con los niveles de servicio implementados. Azure SQL Database será

utilizada para almacenar el Datawarehouse que representa el modelo de datos optimizado para el análisis.

### Power BI

Power BI es una solución de análisis empresarial basado en la nube, que permite unir diferentes fuentes de datos, analizarlos y presentar un análisis de estos a través de informes y paneles. Con Power BI se logra de manera fácil el acceso a datos dentro y fuera de la organización, casi desde cualquier dispositivo. Power BI se utilizar como herramienta de desarrollo de reportes visuales y tableros de mando. Los dos componentes clave de Power BI son:

### Power BI Desktop

Power BI Desktop es una aplicación gratuita que se instala en el equipo local de los desarrolladores de contenido. Permite conectarse a los datos para modelarlos y visualizarlos. Power BI Desktop será utilizado en la fase de desarrollo de reportes y con la herramienta se crearán los archivos PBIX que luego serán publicados para la entrega y consumo.

### Power BI Service

Power BI Service es un servicio gestionado de nube que permite a los desarrolladores publicar los reportes desarrollados en Power BI Desktop. Permite crear y gestionar espacios de trabajo (Workspaces) para modularizar el acceso al contenido mediante diferentes espacios y roles de usuario. Además, el servicio permite a los usuarios acceder a los reportes desde cualquier dispositivo (desktop, laptop, móvil). El servicio cuenta con un portal de administración que permite a los administradores la gestión del funcionamiento del servicio y el

gobierno del contenido mediante permisos y capacidades para los usuarios o grupos de usuarios.

### **Implementación de componentes**

Integración

Servidor de integración

El servidor de integración opera como Gateway para integrar las canalizaciones de ingesta de datos con las fuentes on premise:

Tipo de recurso	Máquina virtual de Cirion
Sistema Operativo	Windows Server 2019 Datacenter
Nombre	CLBISHIR01
Dominio	lcpnace
CPU	1 CPU / 2 CORE VIRTUALES
RAM	8 GB RAM
ALMACENAMIENTO	100 GB

### **Rutina de integración**

A rutina de integración se instala en el servidor que opera como Gateway. Es el componente de Azure Data Factory que permite integrar las canalizaciones de ingesta de datos con las fuentes on premise:

Tipo de recurso	Self hosted integration runtime
Suscripción Azure	icpnaedu

Nombre de ADF	adf-azurebidw (adf-azurebidw-dev)
Nombre de la rutina	itgruntime-azurebi-onpremises
Versión	5.27.8466.1 (20/05/2023)
Nodos instalados	CLBISHIR01

## Sharepoint

Sharepoint opera como repositorio de fuentes no estructuradas y semiestructuradas para ser consumidas por las canalizaciones de Azure data Factory

Tipo de recurso	Sitio de Sharepoint
Nombre del sitio	AzureBI
URL del sitio	<a href="https://icpnaedu.sharepoint.com/sites/AzureBI">https://icpnaedu.sharepoint.com/sites/AzureBI</a>
Carpeta de fuentes	Fuentes BI
Subcarpetas de fuentes	Académica Comercial Finanzas

## Movimiento de datos / Ingesta

### Azure data Factory (ADF)

#### Solución ETL del proyecto:

Tipo de recurso	Data Factory
-----------------	--------------

Suscripción Azure	lcpnaedu
Nombre del recurso	adf-azurebidw (adf-azurebidw-dev)
Versión	V2
Grupo de recursos	rg_azurebi_dev

### Azure data Factory (ADF) Linked Services

	Fuente	Nombre del servicio vinculado
1	Microsoft Dynamics 365	LS_MSD365_CRM
2	SAP Business One	LS_HANA_SAPBO
3	Aplicativo inicio	LS_SQL_INICIO
4	Sistema académico	LS_SQL_ACADEMICO
5	Sistema Spring	LS_SQL_SPRING
6	Sharepoint AzureBI	LS_MSSP_AZUREBI
7	Azure Data Lake	LS_ABLB_AZUREBI / LS_ADLS_AZUREBI

### Azure data Factory (ADF) Dataset

#### Dataset Conformados (Comunes)

	Entidad	Dataset
1	Tiempo	DS_FUNC_Tiempo
2	Modalidad	DS_MD365_Modalidad

3	Escenario	DS_MSSP_XLS_Escenario
4	Curso	DS_SQL_Curso
5	DetalleNroPedidos	DS_SQL_DetalleNroPedidos
6	Duracion	DS_SQL_Duracion
7	Producto	DS_SQL_Producto
8	Programa	DS_SQL_Programa
9	Sede	DS_SQL_Sede
10	Tipoalumno	DS_SQL_Tipoalumno

### **Dataset Académica/Operaciones**

Nro	Entidad	Dataset
1	Alumno	DS_SQL_Alumno
2	Aula	DS_SQL_Aula
3	CasosReserva	DS_SQL_CasosReserva
4	Detalle	DS_SQL_Detalle
5	Docente	DS_SQL_Docente
6	Horario	DS_SQL_Horario
7	PerfilDocente	DS_SQL_PerfilDocente
8	Programación Académica	DS_SQL_Programación Académica
9	Proyeccion	DS_SQL_Proyeccion
10	RangosTiempo	DS_SQL_RangosTiempo
11	Reserva	DS_SQL_Reserva



## Dataset Comercial

Nro	Entidad	Dataset
1	Asesor	DS_MD365_Asesor
2	Campaña	DS_MD365_Campaña
3	Canal	DS_MD365_Canal
4	Contacto	DS_MD365_Contacto
5	EquipoAsesor	DS_MD365_EquipoAsesor
6	GrupoCanal	DS_MD365_GrupoCanal
7	ProgramaInteres	DS_MD365_ProgramaInteres
8	CampañaVentas	DS_MSSP_XLS_CampañaVentas
9	Metas	DS_MSSP_XLS_Metas

## Dataset Contabilidad y Finanzas

Nro	Entidad	Dataset
1	CentroCosto	DS_HANA_CentroCosto
2	CuentaContable	DS_HANA_CuentaContable
3	SaldoContable	DS_HANA_SaldoContable
4	Casos	DS_MSSP_XLS_Casos

5	Concepto	DS_MSSP_XLS_Concepto
6	CtaContableNivel	DS_MSSP_XLS_CtaContableNivel
7	Gerencia	DS_MSSP_XLS_Gerencia
8	PlanificacionPresupuesto	DS_MSSP_XLS_PlanificacionPresupuesto
9	Proyecto	DS_MSSP_XLS_Proyecto
10	VentasPresupuesto	DS_MSSP_XLS_VentasPresupuesto
11	VentasValoradas	DS_MSSP_XLS_VentasValoradas

### Azure Data Lake (ADLS)

Solución de almacenamiento del proyecto:

Tipo de recurso	Storage Account
Suscripción Azure	Icpnaedu
Nombre del recurso	adlsazurebidw (adlsazurebidwdev)
Versión	StorageV2 (general purpose v2)
Grupo de recursos	rg_azurebi_dev
Contenedores	raw work log

### Anexo 1 – Conexiones a las fuentes estructuradas

NOMBRE	SISTEMA	PLATAFORMA	SERVIDOR	P	BASE DE DATOS	DATOS ADICIONALES
--------	---------	------------	----------	---	---------------	-------------------

dynamics	Microsoft Dynamics 365	Dynamics	Por Definir			
sap	SAP Business One	SAP Hana		0.100.10 3.2	1	INSTANCIA 01
inicio	Aplicativo inicio	SQL server	QAPREBAS	92.168.1 62.19	1	Academ
academica	Sistema académico	SQL server	E01			icaUID
spring	Sistema Spring	SQL server	QAPREBAS E02	92.168.1 62.20	1	SPRIN G

## Anexo 2 – Niveles de servicio por componente

COMPONENTE	NIVEL DE SERVICIO	NOTAS
Storage Account	Data Lake Storage Gen 2	Capa de acceso = Hot; Estructura plana de archivo
Azure Data Factory	Azure Data Factory V2	Tipo de Servicio = Data Pipeline
Azure SQL Database	Base de datos única de uso general	Capa = provisionada; vCore = 4
Power BI Desktop	Aplicación gratuita	La versión se actualiza periódicamente
Power BI Service	Power BI PRO	

Notas: Se favorecerá la activación de servicios en la zona Sur central de USA

Los servicios serán activados según sean requeridos, de acuerdo a la etapa del proyecto, para reducir los gastos

## Anexo 02: ICPNA - Documento de tablas de hechos y dimensiones - Área Gestión Académica

### Objetivo principal

El objetivo es poder diseñar un modelo conceptual, que nos especifique el indicador, las lógicas usadas dentro de los procesos, las tablas de hecho, dimensiones y auxiliares, para posteriormente elaborar el diseño físico, que nos aportará valor para la elaboración de nuestro ETL.

### Modelo conceptual

Indicadores y lógicas de cálculo

### Asignación docente

Código	Indicador	Lógica	Dimensiones	Atributos	Tablas	Sistema
GA01	Clases Ofertadas (programación y AD)	Total de clases ofertadas y que no estén canceladas: ordenadas por sede, modalidad y frecuencia.	Curso, Sede, Calendario, Docente	IdProgramas, IdSede, Modalidad, Frecuencia, Estado.	Programa Academico	AcademicaUID
GA02	Clases Usadas (programación y AD)	Conteo de las filas que estado no sea cancelado y cuyas vacantes usadas > 1 y para la data que viene de prev()	Curso, Sede, Docente	IdProgramas, IdSede, Estado, vacantes usadas	Programa Academico	AcademicaUID

		siempre será una clase usada				
GA03	Clases Matriculadas	Conteo de las filas cuyos matriculados >1 y estado no sea cancelado	Curso, Sede, Calendario, Alumno	IdProgramas, Estado, IdSede, IdAlumno, Matriculados	Programa Academico Alumno	AcademicaUID
GA04	Clases estudiando	Conteo de las filas cuyo spring comience con "S" pero que no tenga "SI" en la tabla Datacurt (Revisar Spring)	Curso, Sede, Calendario, Alumno, Excel Distribucion	IdProgramas, IdAlumno, IdSede, spring	Programa Academico Alumno	AcademicaUID 00.Distribución.xlsx
GA05	Clases Vacias	Conteo de las filas cuyos matriculados sea <1 y estado no sea cancelado	Curso, Sede, Calendario, Alumno	IdProgramas, Estado, IdSede, IdAlumno, Matriculados	Programa Academico Alumno	AcademicaUID
GA06	Clases Potenciales a Cancelar <= minimo	Conteo de las filas cuyos matriculados pase el mínimo necesario indicador en la tabla Datacurt	Curso, Sede, Calendario, Alumno, Excel Distribucion	IdProgramas, IdAlumno, IdSede, Matriculados	Programa Academico Alumno Datacurt	AcademicaUID 00.Distribución.xlsx
GA07	Clases Con/Sin Docente	Conteo de las filas que profesor no esté en blanco y que el estado no sea cancelado.	Curso, Sede, Calendario, Docente	IdProgramas, IdDocente, Estado, IdSede,	Programa Academico	AcademicaUID

GA08	Clases Canceladas con Docente	Conteo de las filas cuyo estado sea Clases Canceladas y profesor no sea blanco	Curso, Sede, Calendario, Docente	IdProgramas, IdDocente, Estado, IdSede,	Programa Academico	Academic aUID
GA09	Clases Canceladas Dictadas	Conteo de las filas cuyo estado sea Clases Canceladas y situación sea Dictada	Curso, Sede, Calendario, Docente	IdProgramas, IdDocente, Estado, IdSede,	Programa Academico Empleado Mast	Academic aUID Spring
GA10	Clases Canceladas No Dictadas	Conteo de las filas cuyo estado sea Clases Canceladas y situación sea No Dictada	Curso, Sede, Calendario, Docente	IdProgramas, IdDocente, Estado, IdSede,	Programa Academico Empleado Mast	Academic aUID S pring
GA11	Clases virtuales de qué sede son	Conteo de las filas cuya modalidad se "Virtual" y estado no sea cancelado.	Curso, Sede, Calendario	IdProgramas, Estado, IdSede, Modalidad	Programa Academico	Academic aUID
GA12	Clases reassignadas	Conteo de las filas cuyo flag tenga valor "1" identificar el campo estado de la clase	Curso, Sede, Calendario	IdProgramas, IdSede,	Programa Academico	Academic aUID
GA13	% Balance	Total de clases vs total de docentes	Curso, Sede, Calendario, Docente	IdProgramas, IdDocente, IdSede, Sede	Programa Academico	Academic aUID

GA14	Cantidad de horas del docente	Horas transcurridas en cada horario por la cantidad de clases	Curso, Sede, Horario	chrHoraInicio, chrHoraFin, idProgramas, Estado, IdSede	Programa Academico	AcademicaUID
------	-------------------------------	---	----------------------	--	--------------------	--------------

## Programación

Código	Indicador	Lógicas	Dimensión	Atributos	Tablas	Sistema
GA15	Número de retención de alumnos	Si el estudiante sigue estudiando en el siguiente periodo en el mismo curso (4 casos)	Curso, Sede, Alumno	Periodo, IdCurso, IdAlumno	ProgramaAcademico Alumno	AcademicaUID
GA16	Cantidad de alumnos proyectados	En base al histórico de alumnos se multiplica con el %retención, %duración y %composición teniendo en cuenta si es mensual o bimensual (Considerar el % de crecimiento)	Curso, Sede, Alumno, Excel Proyección	IdProgramas Retención, Duración y Composición	ProgramaAcademico Alumno	AcademicaUID Proyección anual 2023_Aprobado.xlsx

GA17	Cantidad de clases proyectadas	En base al histórico de clases se multiplica con el %retención, %duración y %composición teniendo en cuenta si es mensual o bimensual	Curso, Sede, Excel Proyección	IdProgramas Retención, Duración y Composición	ProgramaAcademico	AcademicaUID Proyección anual 2023_Aprobado.xlsx
GA18	Cantidad de Vacantes Ofertadas	Suma de la columna de Aforo Habilitados	Curso, Sede, Aula	IdCurso, IdSede, IdAula	ProgramaAcademico	AcademicaUID AulaAforo
GA19	Cantidad de Vacantes Usadas	Suma de la columna vacantes usadas	Curso, Sede, Horario, Alumno	IdCurso, IdSede, IdAlumno, Matriculados	ProgramaAcademico Alumno	AcademicaUID
GA20	Cantidad de Alumnos matriculados	Suma de la columna vacantes usadas que son las personas que ya pagaron la matrícula	Curso, Sede, Aula, Alumno	Vacantes usadas, Fech_Pago_Matric, IdAula, IdAlumno, tinEstado, bitVigencia	ProgramaAcademico Alumno PagoMatricula Matricula	AcademicaUID
GA21	Cantidad de Reservas activas	Suma de la diferencia de la columna matrícula y vacantes usadas	Curso, Sede, Aula, Alumno	Vacantes usadas, Matriculados, tinEstado,	ProgramaAcademico Numero Pedidos	AcademicaUID



				bitVigencia,		
GA22	Cantidad de Vacantes disponibles	Suma de la columna vacantes disponible (aula aforo – vacantes usadas)	Curso, Sede, Aula	AulaAforo Vacantes usadas	ProgramaAcademico PagoMatricula Matricula	AcademicaUID
GA23	Cantidad de clases	Suma de la cantidad de clases	Curso	IdProgramas	ProgramaAcademico	AcademicaUID
GA24	Cantidad de alumnos sobre el aforo	Suma de la cantidad de alumnos que sobre pase la columna de Aforo	Curso, Sede, Aula, Alumno	IdCurso, IdSede, AulaAforo Matriculados x	ProgramaAcademico	AcademicaUID
GA25	Cantidad de aulas duplicadas	Conteo de los duplicados de la columna concatenada Aula y Curso	Aula, Curso, Horario	IdAula IdCurso	ProgramaAcademico	AcademicaUID
GA26	Porcentaje Utilización de Aulas	Cantidad de Aulas Ocupadas / Cantidad Total que dispone la sede	Aula, Sede	IdAula IdSede	ProgramaAcademico	AcademicaUID
GA27	Porcentaje Ocupabilidad de aulas utilizadas	Clases Usadas / Capacidad de aula	Aula	AulaAforo IdProgramas	ProgramaAcademico	AcademicaUID

## Tabla de Hechos y Dimensiones

### Tabla de Hechos

Hechos	Descripción
Matrícula	Esta Fact se alimenta de las dimensiones: Tiempo, Curso, Horario, Sede, Aula, Docente y Alumno. En la cual se tiene una lista del programa académico mostrando el tipo de programa, fecha de inicio y fin de clases, estado, matriculados, vacantes usadas y disponibles, nombre de banco, monto de la matrícula (regular, dscto y total), estado de pago y fecha del pago
ProyeccionPeriodo	Es una Fact que muestra los cálculos para la proyección de un año específico por periodo
ProyeccionGranularidad	Es una Fact que muestra los cálculos para la proyección de un año específico por periodo, meses de duración, sede y tipo de alumno.
Reservas	Es una Fact que se alimenta de las dimensiones: Alumno, Curso, Horario, sede y tiempo, además muestra el estado.
Programación	Es una Facr que se alimenta de las dimensiones Sede, curso, horario y aula para mostrar datos de la programación de aulas por horario.

### Tabla de Dimensiones

Dimensiones	Descripción
Tiempo	Es un maestro que contiene año, mes, día, trimestre y semestre.

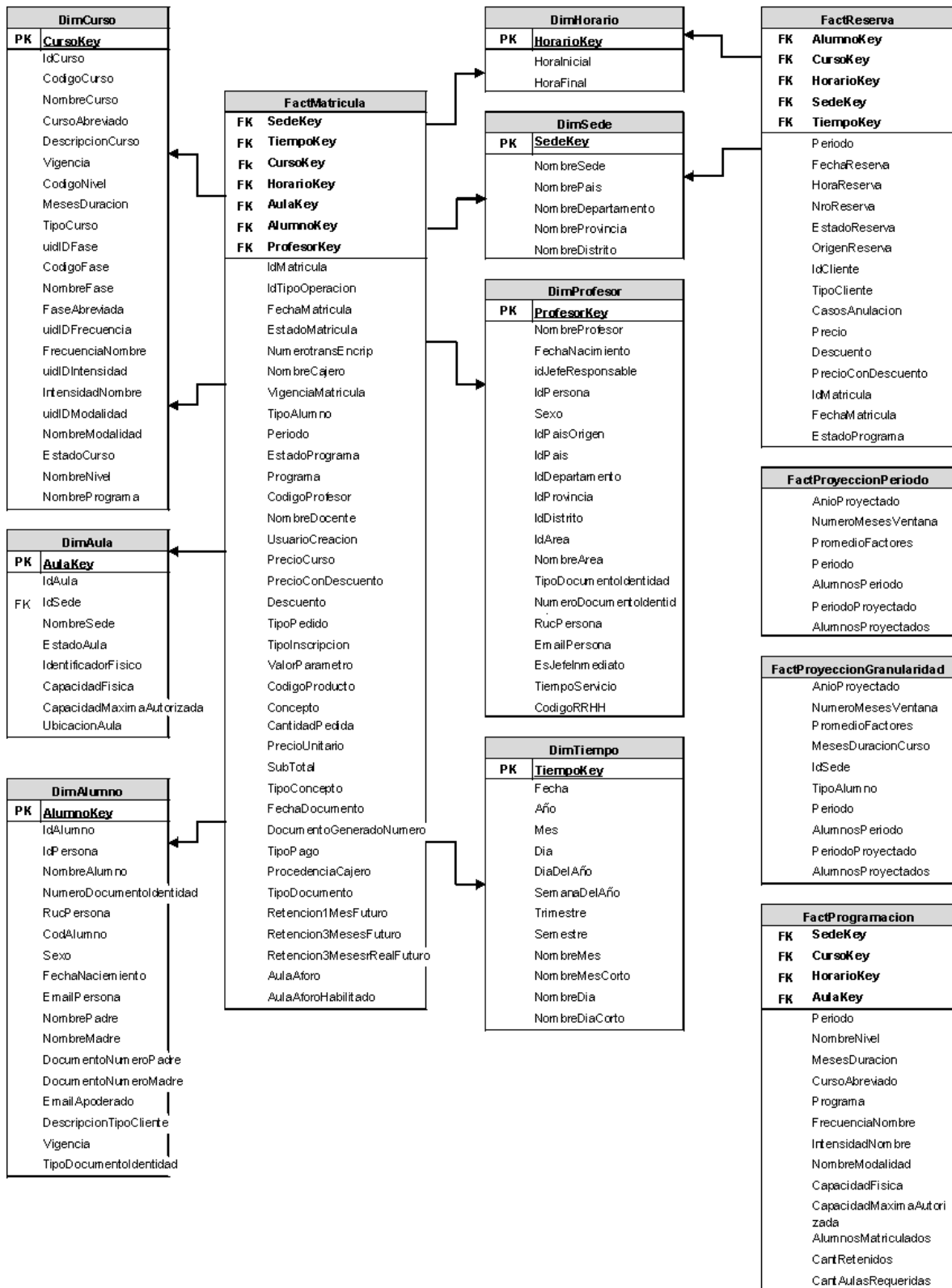
Curso	Es un maestro que nos indicará CursoKey, nombre, curso abreviado, IdFrecuencia, IdIntensidad, meses de duración, Tipo, fase código, nombre fase, fase abreviada, IdModalidad, nombre frecuencia y nombre Modalidad
Aula	Es un maestro que nos indicará AulaKey, estado, aforo, aforo habilitado.
Sede	Es un maestro que nos indicará a qué sede, distrito o departamento pertenece el alumno o el docente, la cual se relaciona por el SedeKey
Profesor	Es un maestro que nos brinda toda la información relacionada a los docentes de la organización como nombre completo, sexo, área a la que pertenece, la condición del empleado, la cual se relaciona por el ProfesorKey
Alumno	Es un maestro que nos brinda toda la información relacionada a los alumnos como DNI, nombre completo, tipo, sexo, fecha de nacimiento, teléfono, correo y se relaciona por el AlumnoKey
Horario	Es un maestro que nos brinda la hora inicial, hora final y se relaciona con el HorarioKey.

### Tabla de Dimensiones – Auxiliares

Auxiliares	Dimensiones
Intensidad, Frecuencia, Nivel, Fase, Situación, modalidad	Curso
AulaAforo, Programa, modalidad, Nivel	Aula
Pais, Departamento, Provincia, Distrito, Capacidad sede	Sede
EmpleadoTipo, Persona, Area, EmpleadoMast, Empleado	Profesor

Persona, TipoCliente, Alumno	Alumno
------------------------------	--------

## Esquema Estrella.



## Anexo 03: ICPNA - Documento de tablas de hechos y dimensiones - Área Contabilidad y Finanzas

### Objetivo principal

El objetivo es poder diseñar un modelo conceptual, que nos especifique el indicador, las lógicas usadas dentro de los procesos, las tablas de hecho, dimensiones y auxiliares, para posteriormente elaborar el diseño físico, que nos aportará valor para la elaboración de nuestro ETL.

### Modelo conceptual

#### Indicadores y Lógicas

#### Financieros

Código	Indicador	Lógica	Dimensiones	Atributos	Tablas	Sistema
GF01	Margen Bruto mes y acumulado	Suma total de las ventas menos el costo de los servicios mensualmente.	Gerencia CuentaContable Sede Tiempo CentroCosto Concepto Programa Duracion	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP Agrup Publicidad	Ventas TIOSPRO DUCTO DetalleNumerosPedidos	SPRING SAP AcademicaUID
GF02	Margen Operativo	Suma total de los ingresos netos divididos	Gerencia	Cta Mayor SAP Codigo CeCo	Ventas	SPRING SAP

	mes y acumulado	entre el total de ingresos en ventas.	CuentaContable Sede Tiempo CentroCosto Concepto Programa Duracion	NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP Agrup Publicidad	TIOSPRO DUCTO DetalleNumerosPedidos	AcademicaUID
GF03	Margen Neto mes y acumulado	La resta de la utilidad buta menos los gastos e impuestos, divididos entre los ingresos totales	Gerencia CuentaContable Sede Tiempo CentroCosto Concepto Programa Duracion	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP Responsable Gastos	Ventas TIOSPRO DUCTO DetalleNumerosPedidos	SPRING SAP AcademicaUID
GF04	Margen EBITDA mes y acumulado	Suma de la utilidad operativa y la depreciación. Dividido entre el total de las ventas.	Gerencia CuentaContable Sede Tiempo CentroCosto Concepto Programa	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP Agrup Publicidad	Ventas TIOSPRO DUCTO DetalleNumerosPedidos	SPRING SAP AcademicaUID

			Duracion			
GF05	Gasto Operativo / Ventas	La suma total del gato operativo dividido entre el total de las ventas.	Gerencia CuentaContable Sede Tiempo CentroCosto Concepto Programa Duracion	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP Agrup Publicidad Responsible Gastos	Ventas TIOSPRO DUCTO DetalleNumerosPedidos	SPRING SAP AcademicaUID
GF06	ROA	La división de la utilidad neta sobre los activos totales expresados en porcentaje.	Gerencia CuentaContable Sede Tiempo CentroCosto Concepto Programa Duracion	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP Agrup Publicidad voucherline	Ventas TIOSPRO DUCTO DetalleNumerosPedidos	SPRING SAP AcademicaUID
GF07	ROE	La división de la utilidad neta sobre el patrimonio total	Gerencia CuentaContable Sede	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP	Ventas TIOSPRO DUCTO DetalleNumerosPedidos	SPRING SAP AcademicaUID



		expresados en porcentaje.	Tiempo CentroCosto Concepto Programa Duracion	NIVEL 3 GYP NIVEL 4 GYP voucherline		
GF09	Ratio de Liquidez	La división del activo corriente entre el pasivo corriente.	Gerencia CuentaContable Sede Tiempo CentroCosto Concepto Programa Duracion	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP Agrup Publicidad Responsable Gastos voucherline	Ventas TIOSPRO DUCTO DetalleNumerosPedidos	SPRING SAP AcademicaUID
GF10	Capital de trabajo	Es la resta de los activos corrientes menos los pasivos corrientes	Gerencia CuentaContable Sede Tiempo CentroCosto Concepto Programa Duracion	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP Agrup Publicidad	Ventas TIOSPRO DUCTO DetalleNumerosPedidos	SPRING SAP AcademicaUID

				Responsable		
				Gastos		
				voucherline		
GF11	Ratio de Apalancamiento Financiero Total	Es la suma de la inversión dividido entre la suma de los recursos propios invertidos	Gerencia CuentaContable Sede Tiempo CentroCosto Concepto Programa Duracion	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP voucherline	Ventas TIOSPRO DUCTO DetalleNumerosPedidos	SPRING SAP AcademicaUID
GF12	Ratio de Apalancamiento Financiero LP	Es la suma de la inversión dividido entre la suma de los recursos propios invertidos	Gerencia CuentaContable Sede Tiempo CentroCosto Concepto Programa Duracion	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP voucherline	Ventas TIOSPRO DUCTO DetalleNumerosPedidos	SPRING SAP AcademicaUID
GF13	Pasivo / Patrimonio	División del pasivo sobre el patrimonio	Gerencia CuentaContable Sede	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP	Ventas TIOSPRO DUCTO	SPRING SAP AcademicaUID

			Tiempo	NIVEL 3 GYP	DetalleNumerosPedidos	
			CentroCosto	NIVEL 4 GYP		
			Concepto	Responsable		
			Programa	Gastos		
			Duracion	voucherline		
GF14	Costo de Ventas sobre la Venta vs (% de Descuento, % de Docentes, % de Textos, % de Exámenes, % de Licencias, % ACPNA)	Resta total de las ventas menos el total de los costos por: (descuentos, docentes, textos, exámenes, licencias)	Gerencia	Cta Mayor SAP	Ventas	SPRING
			CuentaContable	Codigo CeCo	TIOSPRO	SAP
			Sede	NIVEL 1 GYP	DUCTO	AcademicaUID
			Tiempo	NIVEL 2 GYP		
			CentroCosto	NIVEL 3 GYP		
			Concepto	NIVEL 4 GYP		
			Programa	Agrup	DetalleNumerosPedidos	
			Duracion	Publicidad		
				Responsable		
				Gastos		
				Voucherline		
				CODIGOPRO		
				DUCTO		
GF15	Variación (%) de ventas vs mes anterior y PPTO	Es la cantidad real vendida menos la cantidad presupuestada vendida multiplicada por el margen	Gerencia	Cta Mayor SAP	Ventas	SPRING
			CuentaContable	Codigo CeCo	TIOSPRO	SAP
			Sede	NIVEL 1 GYP	DUCTO	AcademicaUID
			Tiempo	NIVEL 2 GYP		
			CentroCosto	NIVEL 3 GYP		
			Concepto	NIVEL 4 GYP		
				voucherline		

		de contribución presupuestado	Programa Duracion			
GF16	Variación (%) de costos vs mes anterior y PPTO	Es el costo total proyectado menos el costo real. Expresado en porcentaje	Gerencia CuentaCont able Sede Tiempo CentroCosto Concepto Programa Duracion	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP Responsable Gastos voucherline	Ventas TIOSPRO DUCTO DetalleNum erosPedidos	SPRING SAP AcademicaUI D
GF17	Variación (%) de gasto operativo vs mes anterior y PPTO	Es el gato operativo total proyectado menos el gasto operativo real. Expresado en porcentaje	Gerencia CuentaCont able Sede Tiempo CentroCosto Concepto Programa Duracion	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP Responsable Gastos voucherline	Ventas TIOSPRO DUCTO DetalleNum erosPedidos	SPRING SAP AcademicaUI D

## Gestión

Código	Indicador	Variables Negocio	Lógicas	Dimension e	Atributos	Tablas	Siste ma
--------	-----------	----------------------	---------	----------------	-----------	--------	-------------

GF18	Tarifa promedio mensualizado	Modalidad Sede Nivel Programa Tipo Alumno Libros Exámenes	El promedio de la suma de las tarifas de los cursos ofrecidos ordenado por cada variable del negocio	Gerencia CuentaContable Sede Tiempo CentroCos Concepto Programa Duracion Fase Modalidad	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP CODIGOFASE	Ventas TIPOS PRODU CTO Detalle NumerosPedidos	SPRING SAP AcademicaUID
GF19	Costo promedio mensualizado	Aula Docente Tipo Alumno Modalidad Sede Programa DocentexPrograma	El promedio de los costos generados por cada variable del negocio	Gerencia CuentaContable Sede Tiempo CentroCos Concepto Programa Duracion Fase Modalidad	Responsable Gastos Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP CODIGOFASE NIVEL FRECUENCIA	Ventas TIPOS PRODU CTO Detalle NumerosPedidos	SPRING SAP AcademicaUID
GF20	Cantidad de	Modalidad	Conteo de las filas de	Gerencia	codigoalumno	Ventas	SPRING

	Alumnos estudiando mensualmente	Sede Nivel Programa Tipo Alumno	los alumnos matriculados ordenados por cada variable del negocio	CuentaContable Sede Tiempo CentroCosto Concepto Programa Duracion Fase Modalidad	Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP CODIGOFASE NIVEL FRECUENCIA	TIPOS PRODU CTO Detalle Numeros Pedidos	SAP Academica UID
GF21	Cantidad de Alumnos estudiando mensualmente vs mes anterior	Modalidad Sede Nivel Programa Tipo Alumno	Comparación del conteo de las filas de los alumnos matriculados con el conteo de los alumnos matriculados el mes anterior ordenados por cada variable del negocio	Gerencia CuentaContable Sede Tiempo CentroCosto Concepto Programa Duracion Fase Modalidad	codigoalumno Cta Mayor SAP Codigo CeCo NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP CODIGOFASE NIVEL FRECUENCIA	Ventas TIPOS PRODU CTO Detalle Numeros Pedidos	SPRING SAP Academica UID
GF22	Margen Bruto	Modalidad Programa	Suma total de las ventas menos el	Gerencia	Cta Mayor SAP Codigo CeCo	Ventas	SPRING

	mensualizado	Sede	costo de los servicios mensualmente y ordenado por cada variable del negocio	CuentaContable Sede Tiempo CentroCosto Concepto Programa Duracion Fase Modalidad	NIVEL 1 GYP NIVEL 2 GYP NIVEL 3 GYP NIVEL 4 GYP CODIGOFASE NIVEL FRECUENCIA Sede	TIPOS PRODU CTO Detalle NumerosPedidos	SAP AcademicaUID
--	--------------	------	--	--	---	--	---------------------

## Tabla de Hechos y Dimensiones

### Tabla de Hechos

Hechos	Descripción
SaldoContable	Se alimenta de los sistemas Spring y SAP para los reportes de GyP
PlanificacionPresupuesto	Se alimenta de los sistemas Spring y SAP para los reportes de finanzas
VentasPresupuesto	Se alimenta de los sistemas Spring y SAP para los reportes de finanzas
VentasValoradas	Se alimenta de Spring y SAP para los reportes de ventas

## Tabla de Dimensiones

Dimensiones	Descripción
Gerencia	Es un maestro con el IdGerencia de cada una de las gerencias existentes con su correspondiente nombre
CuentaContable	Es un maestro que nos muestra la Cuenta Mayor de SAP y Spring con su correspondiente nombre
Sede	Es un maestro que nos muestra el nombre de la sede, la zona y el ubigeo
Tiempo	Es un maestro que nos indica el mes, año, trimestre y semestre
CentroCosto	Es un maestro que nos muestra el nombre del centro de costo, la gerencia responsable y la modalidad
Proyecto	Es un maestro que nos indica el nombre del proyecto.
Fase	Es un maestro que nos muestra las subfamilias de un curso
Modalidad	Es un maestro que nos muestra el IdModalidad y nombre de las modalidades de un curso (virtual, presencial)
TipoAlumno	Es un maestro que nos muestra el IdTipoAlumno y descripción que tiene cada alumno. (Nuevo, reingresante, regular)
Duración	Es un maestro de que muestro el IdDuracion y la descripción de la duración de un curso. (mensual, bimensual)
Concepto	Es un maestro que nos muestra IdConcepto, nombre y tipo de concepto que tiene cada una de las ventas.

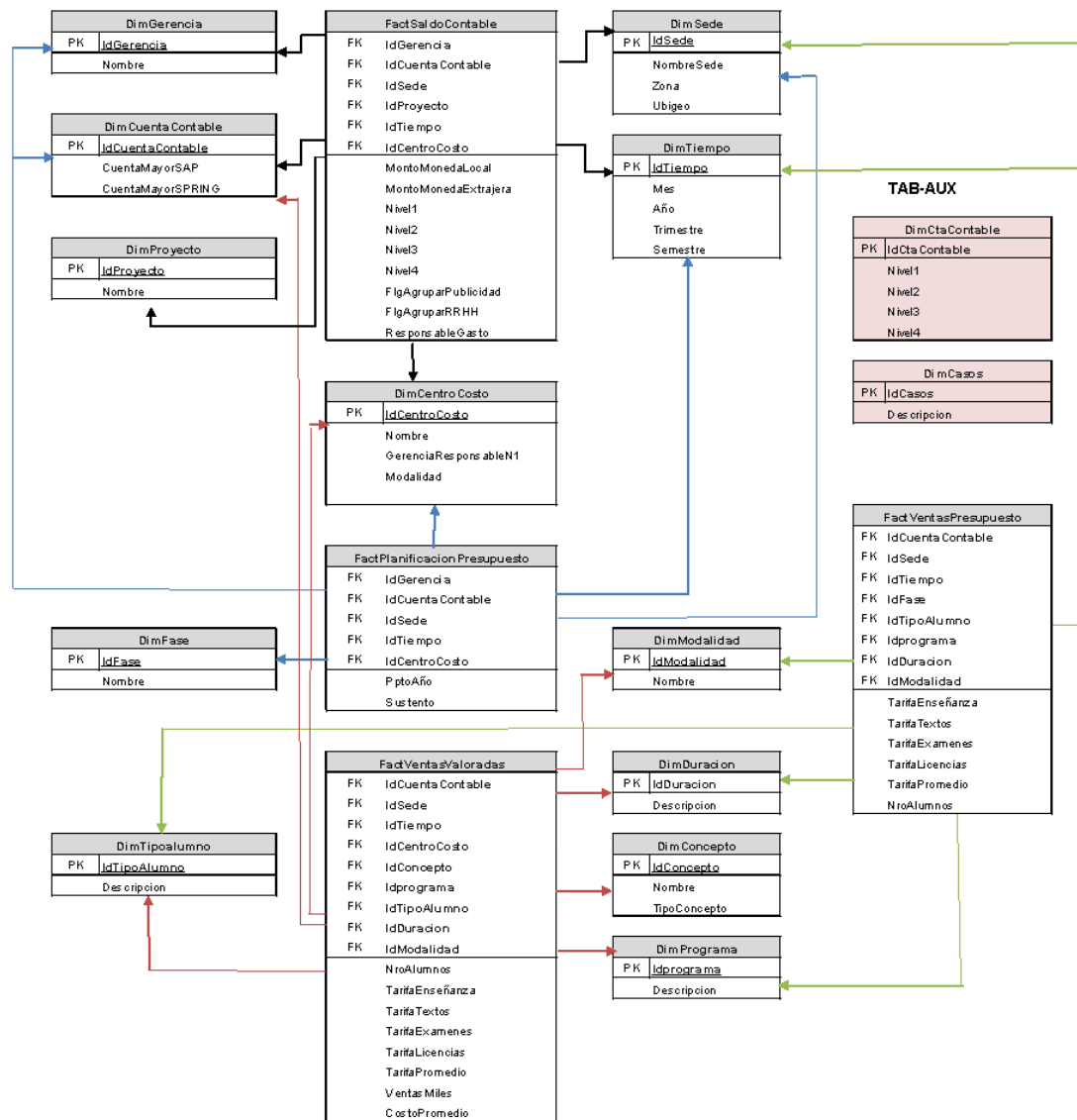


Programa	Es un maestro que nos muestra el IdPrograma y la descripción de los programas existentes. (Programa de Inglés para Niños, Conversation Program, Corporativo)
----------	--

### Tabla de Dimensiones – Auxiliares

Auxiliares	Dimensiones
TIOSPRODUCTO, DetalleNumerosPedidos, numerospedidos	Concepto
Modalidad, fase, horario, aula	Curso
Alumno, persona, matricula, ProgramaAcademico, sct_detalleventa	TipoAlumno
CtaContable, Casos	CuentaContable

### Esquema Copo de Nieve.



## Anexo 04: ICPNA - Documento de tablas de hechos y dimensiones - Área Comercial

### Objetivo principal

El objetivo es poder diseñar un modelo conceptual, que nos especifique el indicador, las lógicas usadas dentro de los procesos, las tablas de hecho, dimensiones y auxiliares, para posteriormente elaborar el diseño físico, que nos aportará valor para la elaboración de nuestro ETL.

### Modelo conceptual

#### Indicadores y Lógicas

#### Campañas

Código	Indicador	Lógica	Dimensiones	Atributos	Tablas	Sistema
GC01	Cantidad de leads por campaña	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés)	Contacto, Tiempo	New_periododecampananame fullname New_numero_documento New_programad einteresname	LEAD icpna_periodo decampana	Dynamic 365
GC02	Cantidad de leads por campaña y canal de contacto	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés), considerando la	Contacto, Tiempo, Grupo_canal	New_periododecampananame fullname New_numero_documento	LEAD icpna_periodo decampana Icpna_canalde contacto	Dynamic 365

		columna Canal de contacto		New_programad einteresname Icpna_name	contact	
GC03	Cantidad de leads por campaña y programa de interés	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés), considerando la columna Programa de interes	Contacto, Tiempo	New_periododec ampananame fullname New_numerodo cumento New_programad einteresname	LEAD Icpna_periodes decampana	Dynamic 365
GC04	Cantidad de leads por campaña y fuente	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés) considerando la columna Fuente (LANDING, WEB)	Contacto, Tiempo. Grupo_canal	New_periododec ampananame fullname New_numerodo cumento New_programad einteresname fuente	LEAD Icpna_periodes decampana	Dynamic 365
GC05	Cantidad de leads por campaña y origen (fb, Google)	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés), considerando la columna utmsource (Definir la	Contacto, Tiempo. Grupo_canal	New_periododec ampananame fullname New_numerodo cumento New_programad einteresname Utm_source	LEAD Icpna_periodes decampana	Dynamic 365

		columna en el CRM)				
GC06	Cantidad de leads por campaña y Programa complementario	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés), considerando la columna Programa complementario	Contacto, Tiempo. Grupo_canal	New_periododecampananame fullname New_numero_documento New_programadeinteresname Programa_complementario	LEAD icpna_periodescampana	Dynamic 365
GC07	Cantidad de leads matriculados por campaña, canal de contacto	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés), considerando la columna Canal de contacto y matriculados no sea null	Contacto, Tiempo. Grupo_canal Matricula	New_periododecampananame fullname New_numero_documento New_programadeinteresname Icpna_name	LEAD icpna_periodescampana contact	Dynamic 365 AcademicaUID
GC08	Cantidad de leads matriculados por campaña y programa de interés	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés), considerando la columna Programa de interes y	Contacto, Tiempo. Grupo_canal Matricula	New_periododecampananame fullname New_numero_documento New_programadeinteresname	LEAD icpna_periodescampana	Dynamic 365 AcademicaUID

		matriculados no sea null				
GC09	Cantidad de leads matriculados por campaña y fuente	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés), considerando la columna Fuente y matriculados no sea null	Contacto, Tiempo. Grupo_canal Matricula	New_periododecampananame fullname New_numerodocumento New_programad einteresname Fuente IdAlumno	LEAD icpna_periodes decampana Matricula	Dynamic 365 Academic aUID
GC10	Cantidad de leads matriculados por campaña y origen (fb, Google)	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés), considerando la columna utmsource(Definir la columna en el CRM) y matriculados no sea null	Contacto, Tiempo. Grupo_canal Matricula	New_periododecampananame fullname New_numerodocumento New_programad einteresname Utm_source IdAlumno	LEAD icpna_periodes decampana Matricula	Dynamic 365 Academic aUID
GC11	Cantidad de leads matriculados por campaña y Programa	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés), considerando la columna	Contacto, Tiempo. Grupo_canal Matricula	New_periododecampananame fullname New_numerodocumento	LEAD icpna_periodes decampana Matricula	Dynamic 365 Academic aUID

	complementario	Programa complementario y matriculados no sea null		New_programad einteresname  Programa_comp lementario  IdAlumno		
--	----------------	--	--	--	--	--

## Ventas

Código	Indicador	Lógicas	Dimensiones	Atributos	Tablas	Sistema
GC12	Porcentaje del alcance del logro vs la Meta de alumnos	Suma total de los alumnos matriculados dividida entre la meta mensual esperada. Se va a mostrar por tipo de alumno (Nuevo, Reingresante y Regular)	Contacto, Tiempo, Grupo_canal, Matricula, sede, Tiempo, Excel Metas	New_periodo decampanan ame  fullname  New_numero documento	LEAD  icpna_period odecampana	Dynamic 365  AcademicaU ID  Seguimiento Proyecciones.xlsx
GC13	Cantidad matriculados por Canal	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés), considerando la columna Canal Final y matriculados no sea null. Se va a mostrar por tipo de canal (Chat web,	Contacto, Tiempo, Grupo_canal, Matricula, sede, Tiempo	icpna_name  New_periodo decampanan ame  fullname  New_numero documento	LEAD  icpna_period odecampana  icpna_canald econtacto	Dynamic 365

		Counter, Digital, Llamada y Whats app)		New_programeinteres name		
GC14	Porcentaje matriculados por Producto vs Meta Mensual	Conteo distintivo de la columna RAZ SOCIAL, considerando la columna TIPO DE PRODUCTO. Se va a mostrar por tipo de producto (Compra de matrículas, Convenios, Exámenes, InHouse y Traducción)	Contacto, Tiempo, Matrícula, Tiempo, Producto	NombreProducto New_periododecampananame fullname New_numero documento	LEAD icpna_periododecampana icpna_canald econtacto	Dynamic 365 Seguimiento Proyecciones.xlsx
GC15	Cantidad matriculados por Programa de Interes	Conteo distintivo de la columna Concat(Nombre + DNI + programa de interés), considerando la columna Programa de interés. Se va a mostrar por tipo de curso (Adultos, Niños, Conversation, MET y TEK)	Contacto, Tiempo, Matrícula, Tiempo	New_periododecampananame fullname New_numero documento New_programeinteres name	LEAD icpna_periododecampana	Dynamic 365 Crm_digital.xlsx

## Tabla de Hechos y Dimensiones



## Tabla de Hechos

Hechos	Descripción
CampañaVentas	Se alimenta del sistema Dynamic 365 por el área comercial y del cual se van a extraer los reportes de Leads y Ventas
Metas	Es una Fact que se alimenta de las dimensiones: Producto, Escenario, Tipo de Alumno, Contacto, Sede, Tiempo y Asesor para mostrar las metas e indicadores. Contiene también la meta mensual y el mes estimado de cierre

## Tabla de Dimensiones

Dimensiones	Descripción
Contacto	Es un maestro que contiene los leads y contiene las siguientes columnas: DNI/RUC, Nombre, Tipo, Fuente, Origen, razón social, teléfono y correo
Tiempo	Es un maestro que contiene el día, semana, mes, trimestre y semestre.
Sede	Es un maestro que nos indicará a qué sede, distrito o departamento pertenece el alumno o el docente, la cual se relaciona por el IdSede
Canal	Es un maestro que contiene el nombre y tipo del canal de contacto.
GrupoCanal	Es un maestro que nos indicará las agrupaciones de los canales de contacto
Escenario	Es un maestro que nos mostrará la descripción y el porcentaje de cada uno de los escenarios.

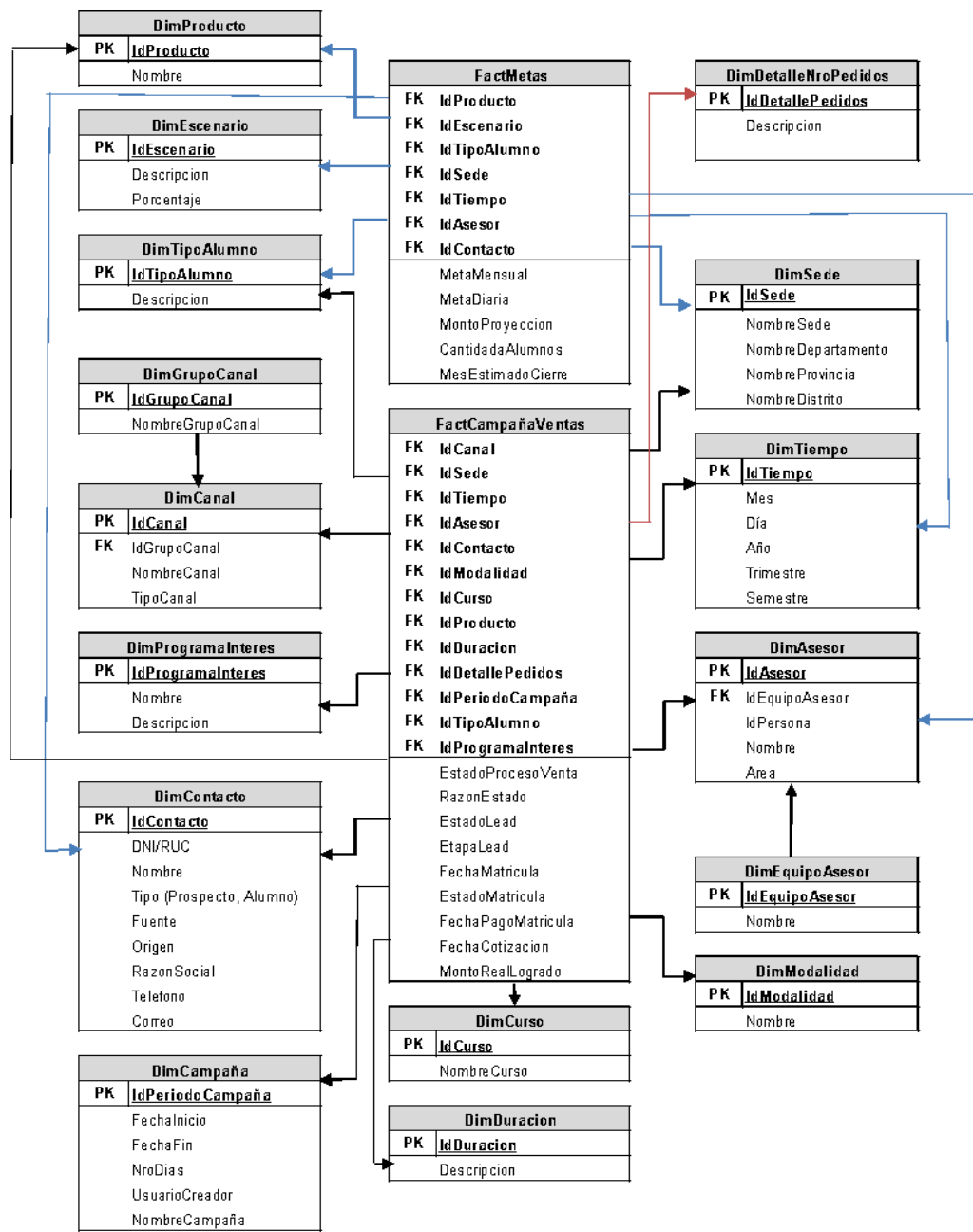
Asesor	Es un maestro que contine información del Asesor de ventas: IdPersona, Nombre, Area.
EquipoAsesor	Es un maestreo que indica la agrupación de los asesores (Individual, corporativo).
Producto	Es un maestro que nos muestra el nombre de cada uno de los productos que maneja el área de ventas.
DetalleNroPedidos	Es una dimensión de la cual vamos a ver la relación entre cada curso y su correspondiente matrícula asociada a un asesor.
Modalidad	Es un maestro que nos muestra el IdModalidad y nombre de la modalidad.
Curso	Es un maestro que nos indicará IdCurso y su correspondiente nombre.
Campaña	Es un maestro que nos indica el periodo de la campaña, Fecha inicio y fin, número de días, usuario creador y el nombre de la campaña
Duración	Es un maestro que indica la duración de la campaña

### Tabla de Dimensiones – Auxiliares

Auxiliares	Dimensiones
Pais, Departamento, Provincia, Distrito, Capacidad_sede, Geografía	Sede
Alumno, Persona	Contacto
Persona	Asesor

SCT_Beca, CB_TipoBeca, sct_TipoControlBeca	DetalleNroPedidos
--	-------------------

**Esquema Copo de Nieve.**



## Anexo 05: ICPNA - Mapeo de Campos

### Mapeo para la dimensión Curso Source

Source						
Source System	Source Schema	Source Table	Source Field Name	Source Datatype	ETL Rules	Comments
AcademicaUID	dbo	Curso	uidIdCurso	varchar(100)	Llave para buscar en DimCurso.IdCurso	
AcademicaUID	dbo	Curso	chrCursoCodigo	char(20)		
AcademicaUID	dbo	Curso	vchNombreCurso	varchar(100)	Select del nombre del curso	
AcademicaUID	dbo	Curso	vchCursoAbreviado	varchar(50)	Select del curso abreviado	
AcademicaUID	dbo	Curso	vchDescripcionCurs	varchar(50)		
AcademicaUID	dbo	Curso	bitVigencia	bit		
AcademicaUID	dbo	Nivel	chrCodNivel	varchar(50)		
AcademicaUID	dbo	Curso	tinMesesDuracion	int	Select de los meses de duración	
AcademicaUID	dbo	Curso	vchTipoCurso	varchar(10)		
AcademicaUID	dbo	Fase	uidIdFase	uid	Llave utilizada para la desnormalización	Este campo no se incluirá en la Dimensión
AcademicaUID	dbo	Fase	vchFaseCodigo	varchar(10)		
AcademicaUID	dbo	Fase	vchNombreFase	varchar(100)		
AcademicaUID	dbo	Fase	vchFaseAbreviado	varchar(50)		
AcademicaUID	dbo	Frecuencia	uidIdFrecuencia	uid	Llave utilizada para la desnormalización	Este campo no se incluirá en la Dimensión
AcademicaUID	dbo	Frecuencia	vchFrecuenciaNombre	varchar(50)		
AcademicaUID	dbo	Intensidad	uidIdIntensidad	uid	Llave utilizada para la desnormalización	Este campo no se incluirá en la Dimensión
AcademicaUID	dbo	Intensidad	vchIntensidadNombre	varchar(20)		
AcademicaUID	dbo	Modalidad	uidIdModalidad	uid	Llave utilizada para la desnormalización	Este campo no se incluirá en la Dimensión
AcademicaUID	dbo	Modalidad	vchNombreModalidad	varchar(20)		
AcademicaUID	dbo	Curso	tinCurso	tinyint	Carga Inicial: tinCurso=1 Carga Incremental: Si tinCurso>1, tinCurso=1	
AcademicaUID	dbo	Nivel	vchNombre	varchar(50)		
AcademicaUID	dbo	Programa	vchNombreProg	varchar(50)		

### Mapeo para la dimensión Curso Target

Target								
Column Name	Description	Attribute Group	Datatype	Size	Key?	FK To	NULL?	Example Values
CursoKey	Llave subrogada de la Dimension	Key	int		PK		N	
IdCurso	Identificador del curso	Identificador	varchar	100			N	1, 2, 3, 4 ...
CodigoCurso	El código del curso	Name	varchar	100			N	B11A1
NombreCurso	El nombre del curso	Name	varchar	100			N	(INT) Intermediate
CursoAbreviado	El nombre del curso abreviado	Name	varchar	50			N	Int. Conv.
DescripcionCurso	Se describe el curso	Name	varchar	50			N	
Vigencia	Muestra la vigencia del curso	Vigencia	varchar	50			N	
CodigoNivel	Código del nivel	Código	varchar	50			N	
MesesDuracion	Cantidad de duración del curso en meses	Name	int				N	
TipoCurso	El nombre del tipo de curso	Name	varchar	10			N	
uidIDFase	Código indicador de la Fase	Identificador	uid			FK	dbo.Fase	N
CodigoFase	EL codigo del la fase	Name	varchar	10			N	
NombreFase	El nombre de la fase	Name	varchar	100			N	
FaseAbreviada	El nombre de la fase abreviado	Name	varchar	50			N	
uidIDFrecuencia	Código indicador de la Frecuencia	Identificador	uid			FK	dbo.Frecuencia	N
FrecuenciaNombre	El nombre de la frecuencia del curso	Name	varchar	50			N	
uidIDIntensidad	Código indicador de la Intensidad	Identificador	uid			FK	dbo.Intensidad	N
IntensidadNombre	El nombre de la intensidad del curso	Name	varchar	20			N	
uidIDModalidad	Código indicador de la Modalidad	Identificador	uid			FK	dbo.Fase/dbo.Modalida	N
NombreModalidad	El nombre de la modalidad del curso	Name	varchar	20			N	
EstadoCurso	Indica que el curso activo o inactivo	Estado	bit	2			N	
NombreNivel	Nombre del nivel	Name	varchar	50			N	
NombrePrograma	Nombre del programa académico	Name	varchar	50			N	

## Mapeo para la dimensión Profesor Source

Source						
Source System	Source Schema	Source Table	Source Field Name	Source Datatype	ETL Rules	Comments
AcademicaUID	dbo	Persona	NombreProfesor	varchar(25)		
AcademicaUID	dbo	Persona	sdlFechaNacimiento	date		
Spring	dbo	EmpleadoMast	JEFERESPONSABLE	int		
AcademicaUID	dbo	Persona	uidIdPersona	varchar(100)	Llave utilizada para la desnormalización	Este campo no se incluirá en la Dimensión
AcademicaUID	dbo	Persona	chrSexo	char(2)	Select del sexo de los docentes	
AcademicaUID	dbo	Persona	uidIdPaisOrigen	int		
AcademicaUID	dbo	Persona	uidIdPais	int		
AcademicaUID	dbo	Persona	uidIdDepartamento	int	Select del departamento	
AcademicaUID	dbo	Persona	uidIdProvincia	int	Select del provincia	
AcademicaUID	dbo	Persona	uidIdDistrito	int	Select del distrito	
Spring	dbo	Empleado	uidIdArea	int	Llave utilizada para la desnormalización	Este campo no se incluirá en la Dimensión
AcademicaUID	dbo	Area	vchNombreArea	varchar(50)		
AcademicaUID	dbo	DocumentoIdentidad	vchNombreDocumento	varchar(100)		
AcademicaUID	dbo	Persona	vchDocumentoNumero	varchar(50)		
AcademicaUID	dbo	Persona	vchRucPersona	varchar(50)		
AcademicaUID	dbo	Persona	vchEmailPersona	varchar(50)		
AcademicaUID	dbo	EmpleadoMast	idJefeResponsable	int		
AcademicaUID	dbo	EmpleadoMast	TIEMPO_SERVICIO	varchar(50)		
AcademicaUID	dbo	Empleado	vchCodigoRRHH	varchar(50)		

## Mapeo para la dimensión Profesor Target

Target								
Column Name	Description	Attribute Group	Datatype	Size	Key?	FK To	NULL?	Example Values
ProfesorKey	Llave subrogada de la dimensión	Key	int		PK		N	
NombreProfesor	Nombre del empleado	Nombre	varchar	25			N	
FechaNacimiento	Fecha nacimiento	Fecha	Datatype				N	
idJefeResponsable	Código del jefe responsable	Código	varchar	25			N	
IdPersona	Llave para relacionar con la tabla Persona	Identificador	varchar	50	FK	dbo.Persona	N	
Sexo	Sexo del empleado	Nombre	char	2			N	
IdPaisOrigen	Identificador del país de origen del empleado	Identificador	int		FK	dbo.Persona	N	
IdPais	Identificador del país del empleado	Identificador	int		FK	dbo.Persona	N	
IdDepartamento	Identificador del departamento	Identificador	int		FK	dbo.Persona	N	
IdProvincia	Identificador de la provincia	Identificador	int		FK	dbo.Persona	N	
IdDistrito	Identificador del distrito	Identificador	int		FK	dbo.Persona	N	
IdArea	Llave para relacionar con la tabla Área	Identificador	int		FK	dbo.Area	N	Reseller, Individual
NombreArea	Nombre del área	Nombre	varchar	50			N	
TipoDocumentoIdentidad	Nombre con el tipo del documento de identidad	Nombre	varchar	50			N	DNI
NumeroDocumentoIdentidad	Número del documento de identidad	Código	int				N	
RucPersona	El número del RUC del empleado	Código	int				N	
EmailPersona	El email del empleado	Nombre	varchar	50			N	
EsJefeInmediato	Un boolean indicando si el empleado es o no jefe	Bit	char	2			N	
TiempoServicio	Es un contador que indica la cantidad de años, meses y días tiene el empleado	Fecha	Datatype				N	28a 06m 11d
CodigoRRHH	Código del empleado para RRHH	Código	int				N	

## Mapeo para la dimensión Aula Source

Source						
Source System	Source Schema	Source Table	Source Field Name	Source Datatype	ETL Rules	Comments
AcademicaUID	dbo	Aula	uidIdAula	int		
AcademicaUID	dbo	Sede	uidIdSede	varchar(100)	Llave para buscar en DimSede.IdSede	
AcademicaUID	dbo	Sede	vchNombreSede	varchar(100)		
AcademicaUID	dbo	Aula	tinEstado	int	Select del estado	
AcademicaUID	dbo	Aula	vchIdentificadorFisico	int		
AcademicaUID	dbo	Aula	tinCapacidadFisica	int		
AcademicaUID	dbo	Aula	tinCapacidadMaxima	int		
AcademicaUID	dbo	Aula	vchUbicacion	varchar(50)		

## Mapeo para la dimensión Aula Target

Target								
Column Name	Description	Attribute Group	Datatype	Size	Key?	FK To	NULL?	Example Values
AulaKey	Llave subrogada de la dimensión	Key	int		PK		N	
IdAula	Identificador de la tabla aula	Identificador	varchar	100	FK		N	
IdSede	Identificador de la tabla sede	Identificador	varchar	100	FK	DimSede.uidIdSede	N	AW00000001
NombreSede	Nombre de la sede del aula	Nombre	varchar				N	
EstadoAula	Un flag que muestra si el estado es activo o inactivo	Identificador	int				N	Reseller, Individual
IdentificadorFisico	Número del aula	Nombre	int				N	
CapacidadFisica	Número de aforo físico en el aula	Número	int				N	
CapacidadMaximaAutorizada	Número de aforo autorizado en el aula	Número	int				N	
UbicacionAula	Nombre de la ubicación del aula	Nombre	varchar	50			N	

## Mapeo para la dimensión Alumno Source

Source						
Source System	Source Schema	Source Table	Source Field Name	Source Datatype	ETL Rules	Comments
AcademicaUID	dbo	Alumno	uidIdAlumno	int		
AcademicaUID	dbo	Persona	uidIdPersona	int	Llave utilizada para la desnormalización	
AcademicaUID	dbo	Persona	vchPrimerNombre, vchSegundoNombre,	varchar(100)		
AcademicaUID	dbo	Persona	vchDocumentoNumero	varchar(20)		
AcademicaUID	dbo	Persona	vchRucPersona	varchar(50)		
AcademicaUID	dbo	Alumno	vchCodal	varchar(15)		
AcademicaUID	dbo	Persona	chrSexo	char(2)		
AcademicaUID	dbo	Persona	sdtFechaNacimiento	smalldatetime		
AcademicaUID	dbo	Persona	vchEmailPersona	varchar(25)		
AcademicaUID	dbo	Persona	strNombrePadre	varchar(60)		
AcademicaUID	dbo	Persona	strNombreMadre	varchar(60)		
AcademicaUID	dbo	Persona	vchDocumentoNumeroPadre	varchar(20)		
AcademicaUID	dbo	Persona	vchDocumentoNumeroMadre	varchar(20)		
AcademicaUID	dbo	Persona	vchEmailApoderado	varchar(25)		
AcademicaUID	dbo	TipoCliente	Descripcion	varchar(25)		
AcademicaUID	dbo	Alumno	bitVigencia	bit		
AcademicaUID	dbo	DocumentoIdentidad	vchNombreDocumento	varchar(50)		

## Mapeo para la dimensión Alumno Target

Target								
Column Name	Description	Attribute Group	Datatype	Size	Key	FK To	NULL	Example Values
AlumnoKey	Llave subrogada de la dimensión	Key	int		PK		N	
IdAlumno	Identificador de la Dimensión	Identificador	int		FK		N	
IdPersona	Llave para relacionar con la tabla Persona	Identificador	varchar	50	FK	uidIdPersona	N	
NombreAlumno	Nombre del alumno	Nombre	varchar	25			N	
NumeroDocumentoidentida	Número del documento de identidad	Número	varchar	24			N	
RucPersona	Número del RUC	Número	varchar	25			N	
CodAlumno	Código único del alumno	Número	varchar	25			N	
Sexo	Sexo del alumno	Nombre	char	2			N	
FechaNacimiento	FechaNacimiento	Fecha	date				N	
EmailPersona	Correo del alumno	Nombre	varchar	25			N	
NombrePadre	Nombre del Padre del Alumno	Nombre	varchar	60			N	
NombreMadre	Nombre del Madre del Alumno	Nombre	varchar	60			N	
DocumentoNumeroPadre	Documento del Padre del Alumno	Número	varchar	20			N	
DocumentoNumeroMadre	Documento del Madre del Alumno	Número	varchar	20			N	
EmailApoderado	Correo del apoderado del alumno	Nombre	varchar	20			N	
DescripcionTipoCliente	Descripcion del tipo de cliente que es Alumno	Nombre	varchar	20			N	
Vigencia	Si está vigente o no el alumno	Nombre	varchar	20			N	
TipoDocumentoidentidad	Nombre del documento de identidad	Nombre	varchar	20			N	

## Mapeo para la dimensión Sede Source

Source						
Source System	Source Schema	Source Table	Source Field Name	Source Datatype	ETL Rules	Comments
AcademicaUID	dbo	Sede	uidIdSede	int		
AcademicaUID	dbo	Sede	vchNombreSede	varchar(50)		
AcademicaUID	dbo	Sede	vcrIdPais	varchar(100)	Llave utilizada para la	
AcademicaUID	dbo	Sede	uidIdDepartamento	varchar(100)	Llave utilizada para la	
AcademicaUID	dbo	Sede	uidIdProvincia	varchar(100)	Llave utilizada para la desnormalización	
AcademicaUID	dbo	Sede	uidIdDistrito	varchar(100)	Llave utilizada para la desnormalización	
AcademicaUID	dbo	Pais	vchNombrePais	varchar(50)		
AcademicaUID	dbo	Departam	vchNombreDepartame	varchar(50)		
AcademicaUID	dbo	Provincia	vchNombreProvincia	varchar(50)		
AcademicaUID	dbo	Distrito	vchNombreDistrito	varchar(50)		

## Mapeo para la dimensión Sede Target

Target								
Column Name	Description	Attribute Group	Datatype	Size	Key?	FK To	NULL?	Example Values
SedeKey	Llave subrogada de la dimensión	Key	int		PK		N	
IdSede	Identificador de la tabla Sede	Identificador	varchar	100	FK		N	
NombreSede	Nombre de la sede	Nombre	varchar	50			N	
IdPais	Llave para relacionar con la tabla Pais	Identificador	varchar	100	FK	DimPais.charIdPais	N	
IdDepartamento	Llave para relacionar con la tabla Departamento	Identificador	varchar	100	FK	DimDepartamento.uidIdDepartamen	N	
IdProvincia	Llave para relacionar con la tabla Provincia	Identificador	varchar	100	FK	DimProvincia.uidIdProvincia	N	
IdDistrito	Llave para relacionar con la tabla Distrito	Identificador	varchar	100	FK	DimDistrito.uidIdDistrito	N	
NombrePais	Nombre del país	Nombre	varchar	50			N	
NombreDepartamento	Nombre del departamento	Nombre	varchar	50			N	
NombreProvincia	Nombre de la provincia	Nombre	varchar	50			N	
NombreDistrito	Nombre del distrito	Nombre	varchar	50			N	



## Mapeo para la dimensión Horario Source

Source						
Source System	Source Schema	Source Table	Source Field Name	Source Datatype	ETL Rules	Comments
AcademicaUID	dbo	Horario	uidIdHorario	varchar(100)		
AcademicaUID	dbo	Horario	chrHoraInicial	varchar(100)		
AcademicaUID	dbo	Horario	chrHoraFinal	varchar(100)		

## Mapeo para la dimensión Horario Target

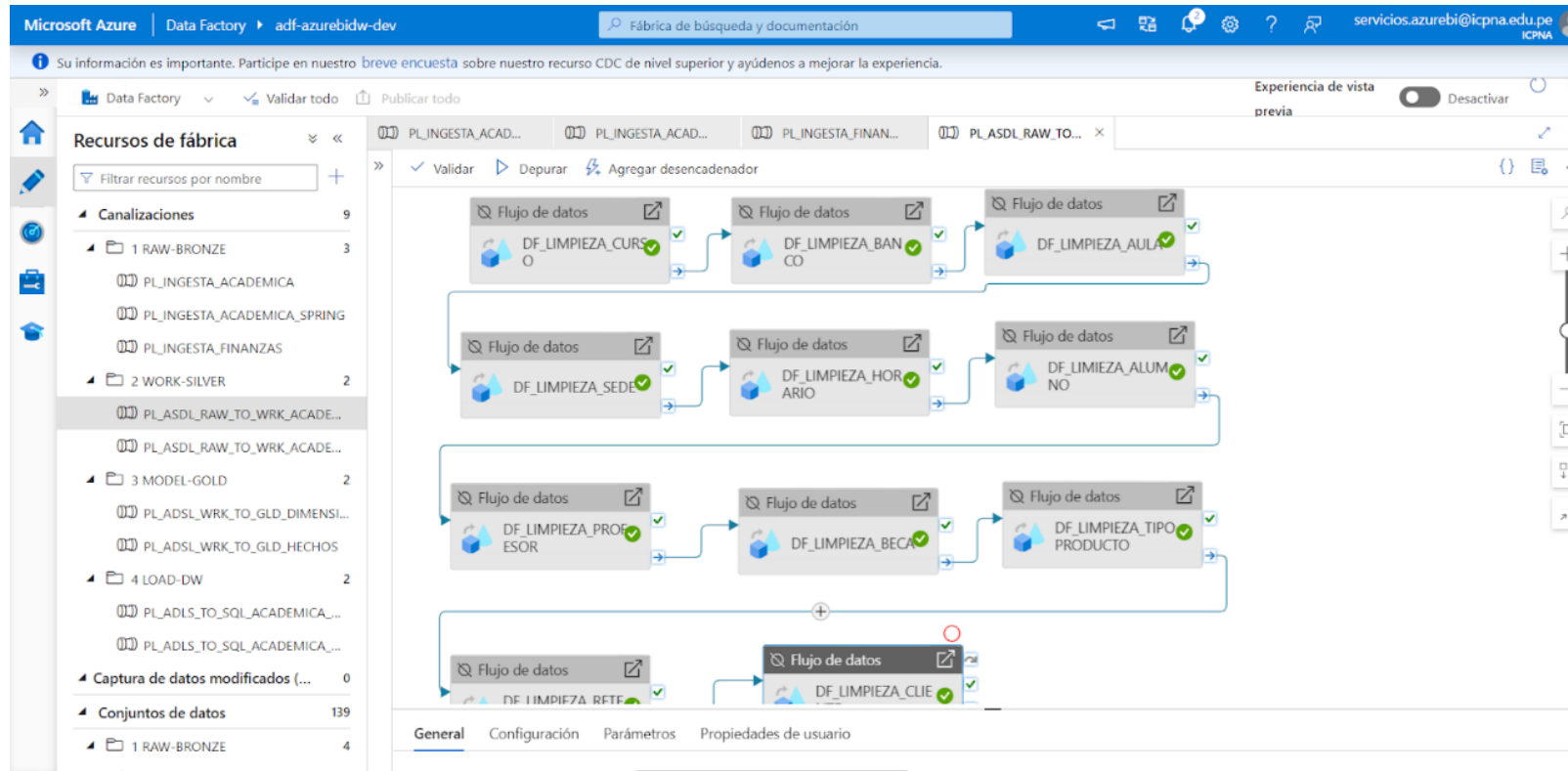
Target									
Column Name	Description	Attribute Group	Datatype	Size	Key?	FK To	NULL?	Default Value	Example Values
HorarioKey	Llave subrogada de la dimensión	Key	int		PK		N		
IdHorario	Identificador de la tabla Horario	Identificador	varchar	100	FK		N		1, 2, 3, 4 ...
HoraInicial	La hora en la que inicia el curso	Número	char	5			N		
HoraFinal	La hora en la que finaliza el curso	Número	char	5			N		

## Mapeo para la dimensión Tiempo Target (sólo existe en el Azure DW)

Target								
Column Name	Description	Attribute Group	Datatype	Size	Key?	FK To	NULL?	Example Values
FechaKey	Llave de la dimensión	Tiempo	int		PK		N	
Fecha	La fecha completa	Tiempo	Datatype				N	
Año	El año de la fecha	Tiempo	varchar	50			N	
Mes	El mes de la fecha	Tiempo	varchar	50			N	
Dia	El día de la fecha	Tiempo	varchar	50			N	
DiaDelAño	El día del año de la fecha	Tiempo	varchar	50			N	
SemanaDelAño	La semana del año de la fecha	Tiempo	varchar	50			N	
Trimestre	El trimestre de la fecha	Tiempo	varchar	50			N	
Semestre	El semestre de la fecha	Tiempo	varchar	50			N	
NombreMes	El nombre del mes de la fecha	Tiempo	varchar	50			N	
NombreMesCorto	El nombre corto del mes de la fecha	Tiempo	varchar	50			N	
NombreDia	El nombre del día de la fecha	Tiempo	varchar	50			N	
NombreDiaCorto	El nombre corto del día de la fecha	Tiempo	varchar	50			N	

## Anexo 06: ICPNA – Evidencias Azure Data Factory

Los pipelines que se elaboraron para la ejecución de los datasets de limpieza (capa silver)



Microsoft Azure | Data Factory | adf-azurebidw-dev | fábrica de búsqueda y documentación | servicios.azurebi@icpna.edu.pe

Su información es importante. Participe en nuestro breve encuesta sobre nuestro recurso CDC de nivel superior y ayúdenos a mejorar la experiencia.

Experiencia de vista: Desactivar

Validar | Depurar | Agregar desencadenador | Depuración de flujo de datos

### Recursos de fábrica

Filtrar recursos por nombre

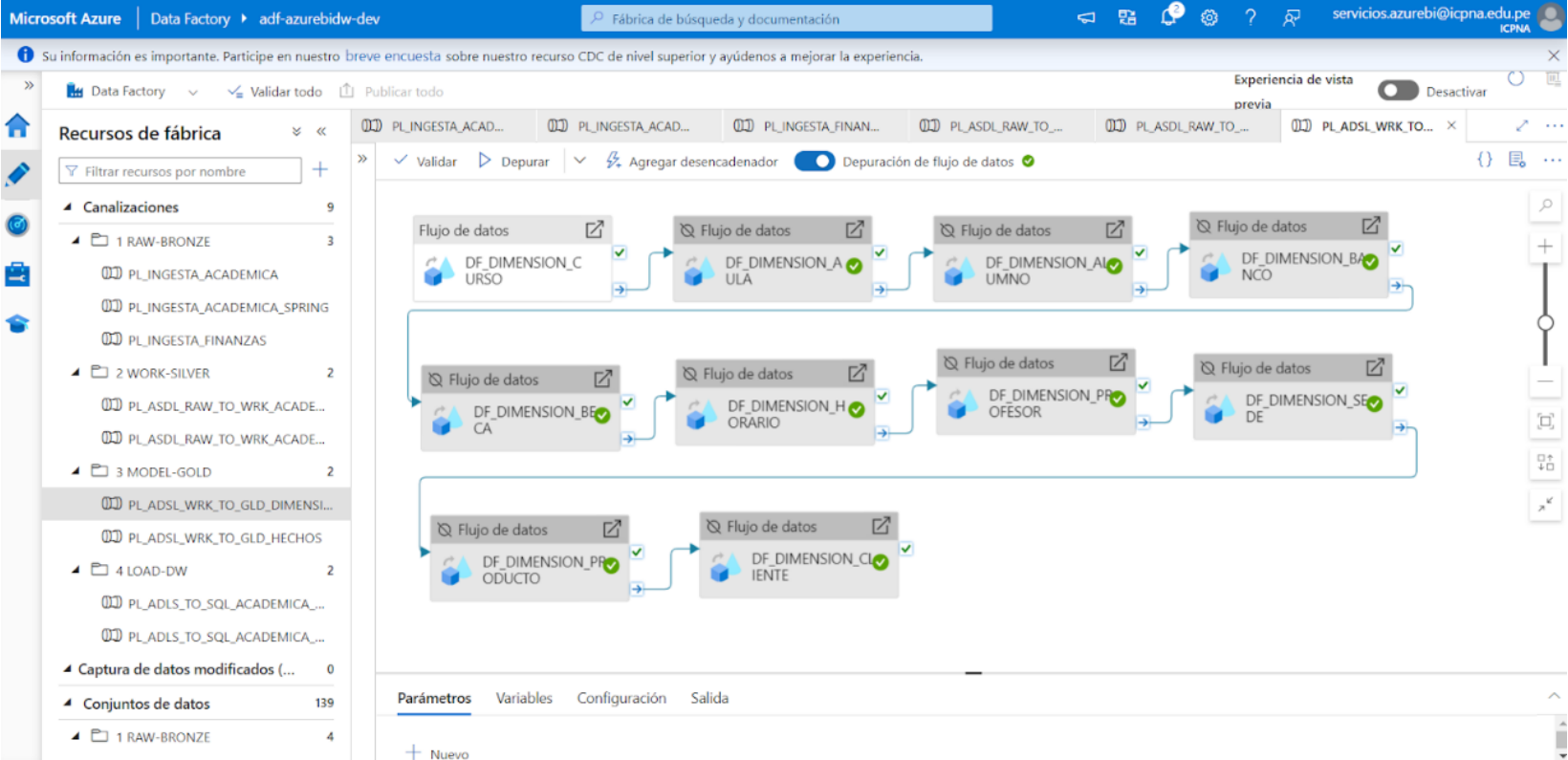
- Canalizaciones (9)
  - 1 RAW-BRONZE (3)
    - PL\_INGESTA\_ACADEMICA
    - PL\_INGESTA\_ACADEMICA\_SPRING
    - PL\_INGESTA\_FINANZAS
  - 2 WORK-SILVER (2)
    - PL\_AS DL\_RAW\_TO\_WRK\_ACADE...
    - PL\_AS DL\_RAW\_TO\_WRK\_ACADE...
  - 3 MODEL-GOLD (2)
    - PL\_ADSL\_WRK\_TO\_GLD\_DIMENSI...
    - PL\_ADSL\_WRK\_TO\_GLD\_HECHOS
  - 4 LOAD-DW (2)
    - PL\_ADLS\_TO\_SQL\_ACADEMICA...
    - PL\_ADLS\_TO\_SQL\_ACADEMICA...
- Captura de datos modificados (...): 0
- Conjuntos de datos (139)
  - 1 RAW-BRONZE (4)

Flujo de datos: DF\_LIMPIEZA\_MAT RÍCULA → DF\_LIMPIEZA\_MAT RÍCULA ALUMNO → DF\_LIMPIEZA\_RESE RVA

Parámetros: Variables | Configuración | Salida

+ Nuevo

Los pipelines que se elaboraron para la ejecución de los data flows (capa Oro)



Microsoft Azure | Data Factory | adf-azurebidw-dev | Fábrika de búsqueda y documentación | servicios.azurebi@icpna.edu.pe

Su información es importante. Participe en nuestro [breve encuesta](#) sobre nuestro recurso CDC de nivel superior y ayúdenos a mejorar la experiencia.

Experiencia de vista:  Desactivar

Validar | Depurar | Agregar desencadenador | Depuración de flujo de datos

### Recursos de fábrica

Filtrar recursos por nombre

- Canalizaciones (9)
  - RAW-BRONZE (3)
    - PL\_INGESTA\_ACADEMICA
    - PL\_INGESTA\_ACADEMICA\_SPRING
    - PL\_INGESTA\_FINANZAS
  - WORK-SILVER (2)
    - PL\_AS DL\_RAW\_TO\_WRK\_ACADE...
    - PL\_AS DL\_RAW\_TO\_WRK\_ACADE...
  - MODEL-GOLD (2)
    - PL\_AD SL\_WRK\_TO\_GLD\_DIMENSI...
    - PL\_AD SL\_WRK\_TO\_GLD\_HECHOS
  - LOAD-DW (2)
    - PL\_AD SL\_TO\_SQL\_ACADEMICA...
    - PL\_AD SL\_TO\_SQL\_ACADEMICA...
- Captura de datos modificados (...): 0
- Conjuntos de datos (139)
  - RAW-BRONZE (4)

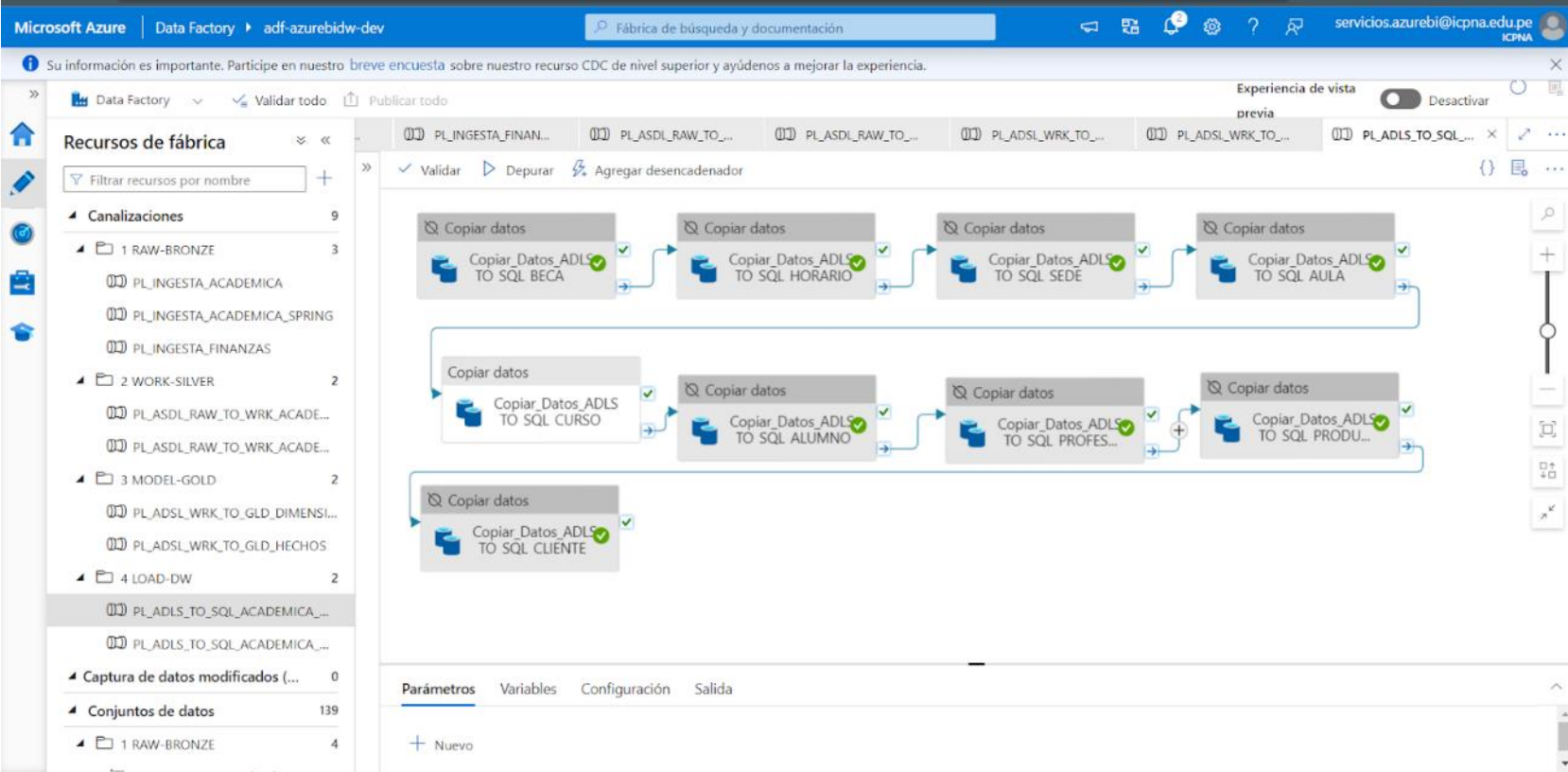
Id. de ejecución de canalización: 96643de5-e5f6-47c1-bad9-aae428d4ae5b | Estado de la canalización: ✔ Correcto

All status | Supervisión en métricas de Azure | Exportar a CSV

Mostrando elementos del 1 al 4 de un total de 4

Nombre de actividad	Estado de actividad	Tipo de actividad	Inicio de la ejecución	Duración	Entorno de ejecución	Propiedades de usuario	Id. de
DF_HECHO_PROGRMACION	Inactive	Flujo de datos	9/29/2023, 11:32:17 AM	Menos de 1 s	Desconocido		3ef535
DF_HECHO_PROYECCION	Inactive	Flujo de datos	9/29/2023, 11:32:17 AM	Menos de 1 s	Desconocido		363751
DF_HECHO_RESERVA	Inactive	Flujo de datos	9/29/2023, 11:32:16 AM	Menos de 1 s	Desconocido		1d44e

Los pipelines que se elaboraron para cargar las dimensiones y tablas de hechos al Azure DW



Microsoft Azure | Data Factory | adf-azurebidw-dev | Fábrica de búsqueda y documentación | servicios.azurebi@icpna.edu.pe

Su información es importante. Participe en nuestro [breve encuesta](#) sobre nuestro recurso CDC de nivel superior y ayúdenos a mejorar la experiencia.

Experiencia de vista:  Desactivar

Validar | Depurar | Agregar desencadenador

### Recursos de fábrica

Filtrar recursos por nombre

- Canalizaciones (9)
  - 1 RAW-BRONZE (3)
    - PL\_INGESTA\_ACADEMICA
    - PL\_INGESTA\_ACADEMICA\_SPRING
    - PL\_INGESTA\_FINANZAS
  - 2 WORK-SILVER (2)
    - PL\_AS DL\_RAW\_TO\_WRK\_ACADE...
    - PL\_AS DL\_RAW\_TO\_WRK\_ACADE...
  - 3 MODEL-GOLD (2)
    - PL\_AS DL\_WRK\_TO\_GLD\_DIMENSI...
    - PL\_AS DL\_WRK\_TO\_GLD\_HECHOS
  - 4 LOAD-DW (2)
    - PL\_ADLS\_TO\_SQL\_ACADEMICA...
    - PL\_ADLS\_TO\_SQL\_ACADEMICA...
- Captura de datos modificados (...): 0
- Conjuntos de datos: 139
  - 1 RAW-BRONZE: 4

Parámetros | Variables | Configuración | **Salida**

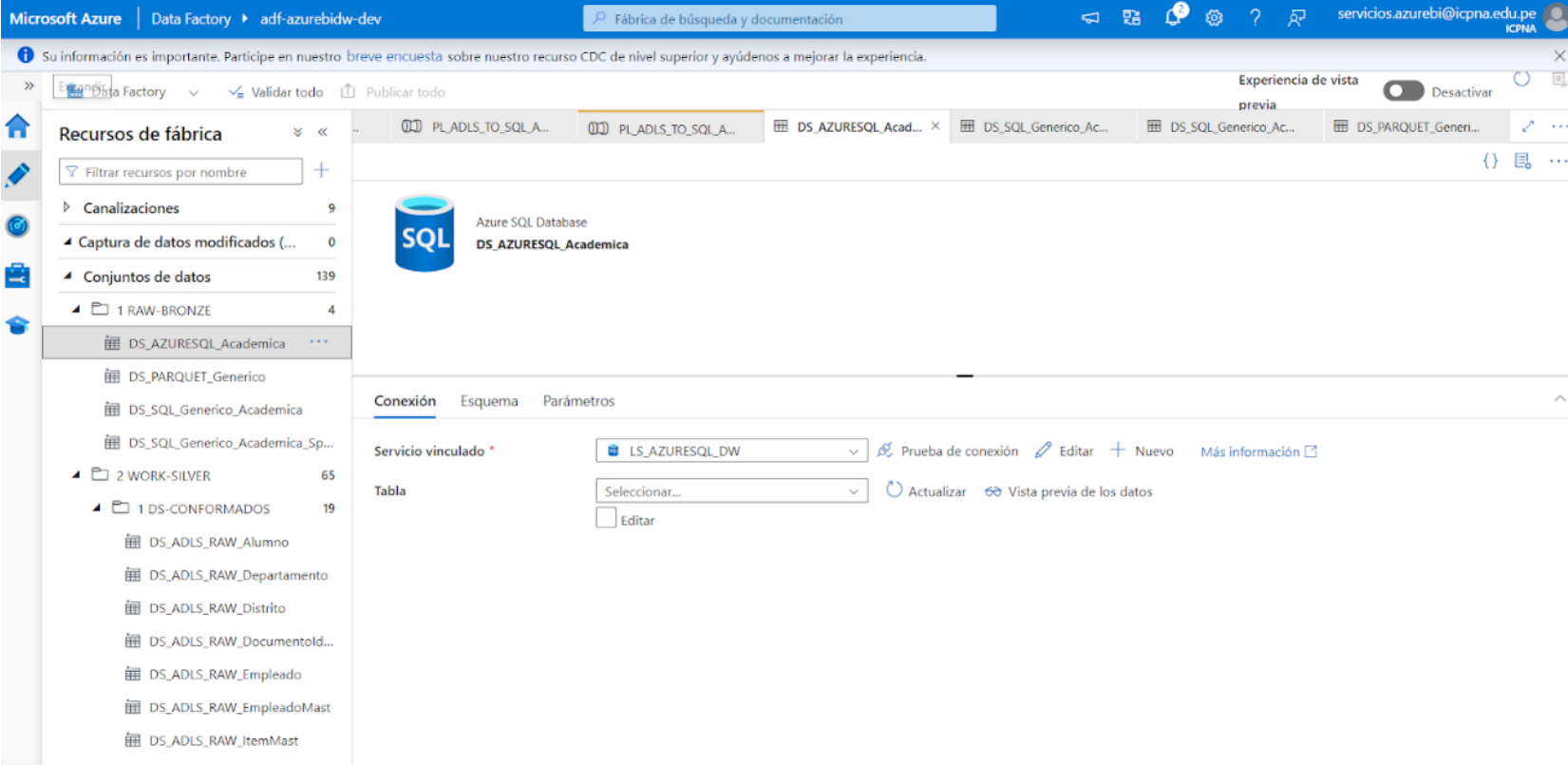
**Id. de ejecución de canalización:** 367816c5-5f60-4d2c-8b85-1fcd65563ca4 | **Estado de la canalización:** ✔ Correcto

All status Supervisión en métricas de Azure  Exportar a CSV

Mostrando elementos del 1 al 5 de un total de 5

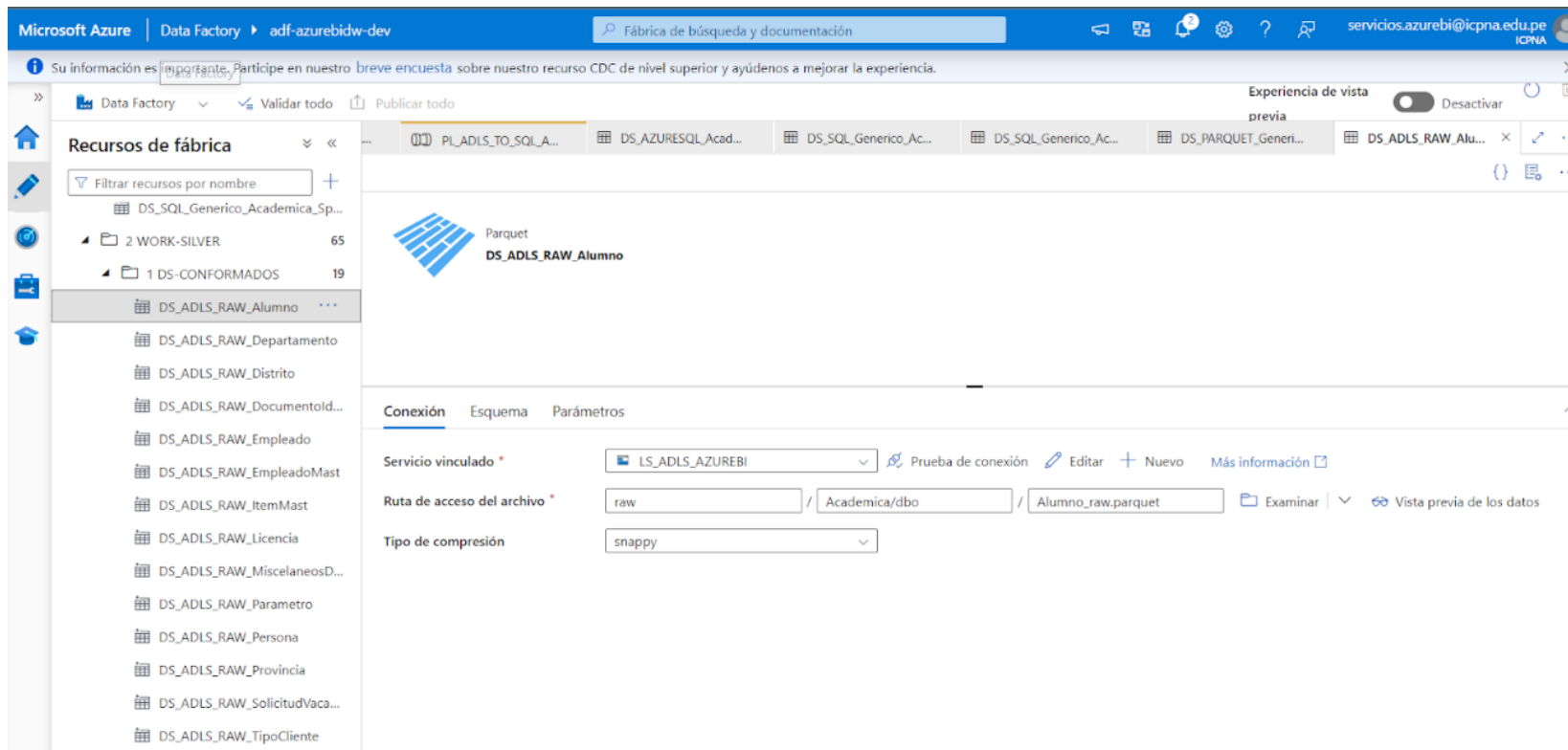
Nombre de actividad	Estado de actividad	Tipo de actividad	Inicio de la ejecución	Duración	Entorno de ejecución	Propiedades de usuari	Id. de
Copiar_Datos_ADLS_TO_SQL...	Inactivo	Copiar datos	9/29/2023, 12:04:34 PM	Menos de 1 s	Desconocido		38649
Copiar_Datos_ADLS_TO_SQL...	Inactivo	Copiar datos	9/29/2023, 12:04:34 PM	Menos de 1 s	Desconocido		cf9657
Copiar_Datos_ADLS_TO_SQL...	Inactivo	Copiar datos	9/29/2023, 12:04:33 PM	Menos de 1 s	Desconocido		955a8!

# Datasets capa raw-bronze



# Datasets capa work-silver





Datasets capa model-gold

Microsoft Azure | Data Factory | adf-azurebidw-dev | Fábrica de búsqueda y documentación | servicios.azurebi@icpna.edu.pe

Su información es importante. Participe en nuestro breve encuesta sobre nuestro recurso CDC de nivel superior y ayúdenos a mejorar la experiencia.

Experiencia de vista previa Desactivar

### Recursos de fábrica

Filtrar recursos por nombre

- DS\_ADLS\_RAW\_TipoContrato
- DS\_ADLS\_RAW\_TipoDocume...
- DS\_ADLS\_RAW\_TipoLicencia
- DS\_ADLS\_RAW\_TipoOperacion
- DS\_ADLS\_RAW\_TipoProducto
- 2 DS-ACADEMICA 46
- 3 MODEL-GOLD 39
- 1 DS-ACADEMICA 39
- 1 DS-DIMENSIONES ...
  - DS\_ADLS\_GLD\_DimAlumno
  - DS\_ADLS\_GLD\_DimAula
  - DS\_ADLS\_GLD\_DimBanco
  - DS\_ADLS\_GLD\_DimBeca
  - DS\_ADLS\_GLD\_DimCliente
  - DS\_ADLS\_GLD\_DimCurso
  - DS\_ADLS\_GLD\_DimHorario
  - DS\_ADLS\_GLD\_DimProdu...

Parquet  
DS\_ADLS\_RAW\_TipoContrato

#### Conexión

Esquema Parámetros

Servicio vinculado \* LS\_ADLS\_AZUREBI Prueba de conexión Editar + Nuevo Más información

Ruta de acceso del archivo \* raw / Academica/dbo / HR\_TipoContrato\_raw.parquet Examinar Vista previa de los datos

Tipo de compresión snappy

Datasets capa model-gold

Microsoft Azure | Data Factory | adf-azurebidw-dev | Fábrika de búsqueda y documentación | servicios.azurebi@icpna.edu.pe

Su información es importante. Participe en nuestro [breve encuesta](#) sobre nuestro recurso CDC de nivel superior y ayúdenos a mejorar la experiencia.

Experiencia de vista: **previa** | Desactivar

### Recursos de fábrica

Filtrar recursos por nombre

- 2 DS-ACADEMICA (46)
  - DS\_ADLS\_RAW\_Area
  - DS\_ADLS\_RAW\_Aula
  - DS\_ADLS\_RAW\_AulaAforo
  - DS\_ADLS\_RAW\_Beca
  - DS\_ADLS\_RAW\_BecasControl
  - DS\_ADLS\_RAW\_CabPrograma...
  - DS\_ADLS\_RAW\_Cliente
  - DS\_ADLS\_RAW\_Curso
  - DS\_ADLS\_RAW\_DetalleNumer...
  - DS\_ADLS\_RAW\_DetalleVenta
  - DS\_ADLS\_RAW\_Fase
  - DS\_ADLS\_RAW\_Frecuencia
  - DS\_ADLS\_RAW\_GestoraCuenta
  - DS\_ADLS\_RAW\_Horario**
  - DS\_ADLS\_RAW\_HorarioBloque
  - DS\_ADLS\_RAW\_Intensidad

Parquet  
**DS\_ADLS\_RAW\_Horario**

**Conexión** | Esquema | Parámetros

Servicio vinculado \*  [Prueba de conexión](#) [Editar](#) [+ Nuevo](#) [Más información](#)

Ruta de acceso del archivo \*  /  /  [Examinar](#) | [Vista previa de los datos](#)

Tipo de compresión

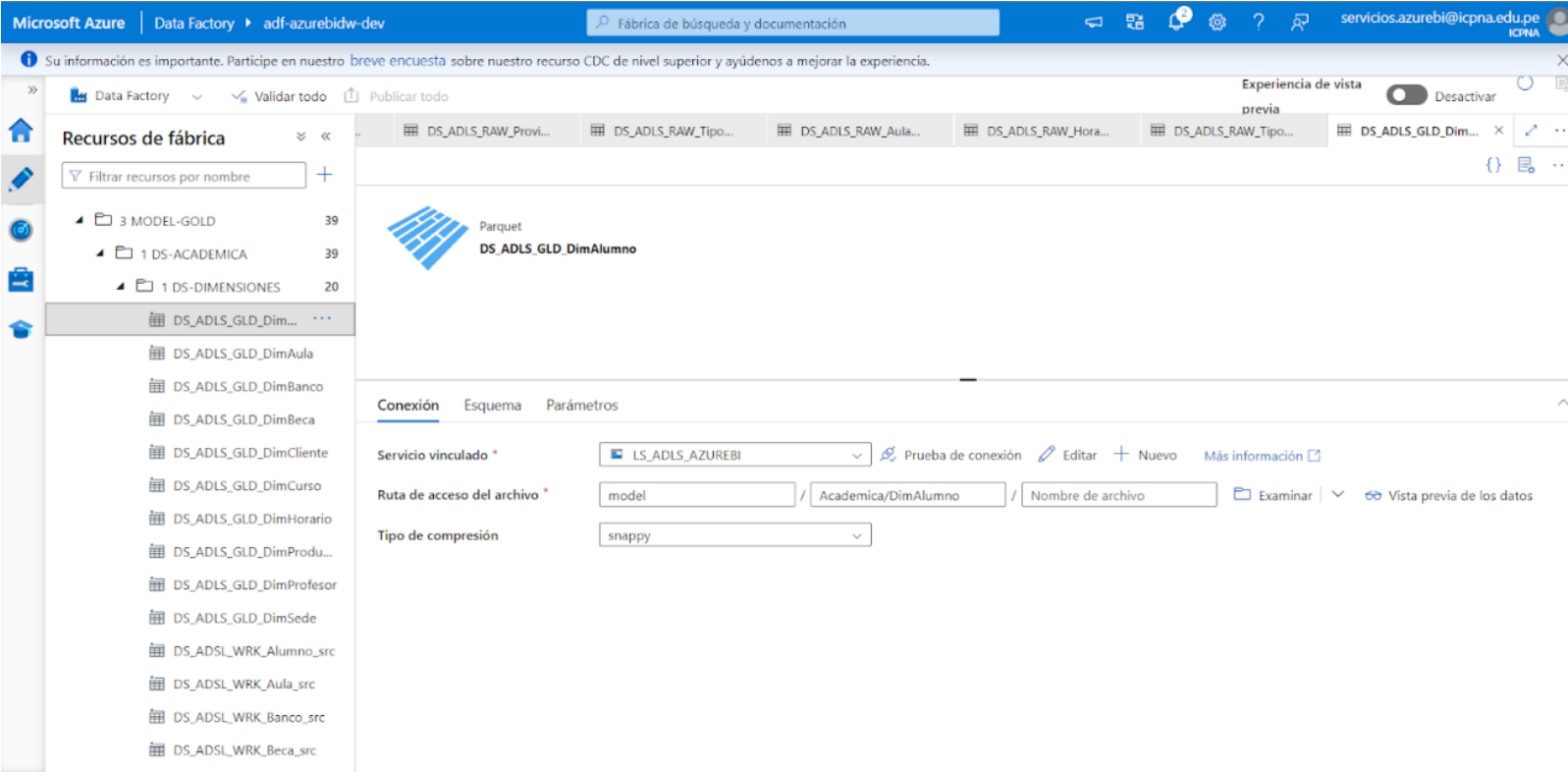
# Datasets capa model-gold

The screenshot shows the Microsoft Azure Data Factory interface. The top navigation bar includes 'Microsoft Azure', 'Data Factory', and the environment name 'adf-azurebidw-dev'. A search bar contains 'Fábrica de búsqueda y documentación'. The user's email 'servicios.azurebi@icpna.edu.pe' is visible in the top right. A notification banner at the top states: 'Su información es importante. Participe en nuestro breve encuesta sobre nuestro recurso CDC de nivel superior y ayúdenos a mejorar la experiencia.'

The main interface is divided into a left sidebar and a main workspace. The sidebar, titled 'Recursos de fábrica', contains a search box 'Filtrar recursos por nombre' and a list of datasets. The dataset 'DS\_ADLS\_RAW\_TipoBeca' is selected and highlighted. The main workspace displays the configuration for this dataset, which is a Parquet file. The configuration is organized into three tabs: 'Conexión', 'Esquema', and 'Parámetros'. The 'Conexión' tab is active and shows the following settings:

- Servicio vinculado \***: A dropdown menu showing 'LS\_ADLS\_AZUREBI' with options for 'Prueba de conexión', 'Editar', '+ Nuevo', and 'Más información'.
- Ruta de acceso del archivo \***: A text input field containing 'raw / Academica/dbo / CB\_TipoBeca\_raw.parquet' with an 'Examinar' button and a 'Vista previa de los datos' link.
- Tipo de compresión**: A dropdown menu showing 'snappy'.

# Datasets capa model-gold



# Datasets capa model-gold (Dimensiones)

Microsoft Azure | Data Factory | adf-azurebidw-dev | Fábrica de búsqueda y documentación | servicios.azurebi@icpna.edu.pe

Su información es importante. Participe en nuestro [breve encuesta](#) sobre nuestro recurso CDC de nivel superior y ayúdenos a mejorar la experiencia.

Experiencia de vista previa  Desactivar

Recursos de fábrica

Filtrar recursos por nombre

- 3 MODEL-GOLD 39
  - 1 DS-ACADEMICA 39
    - 1 DS-DIMENSIONES 20
      - DS\_ADLS\_GLD\_DimAlumno
      - DS\_ADLS\_GLD\_DimAula
      - DS\_ADLS\_GLD\_DimBanco
      - DS\_ADLS\_GLD\_DimBeca
      - DS\_ADLS\_GLD\_DimCliente
      - DS\_ADLS\_GLD\_DimCurso
      - DS\_ADLS\_GLD\_DimHorario
      - DS\_ADLS\_GLD\_DimProdu...
      - DS\_ADLS\_GLD\_DimProfesor
      - DS\_ADLS\_GLD\_DimS...**
      - DS\_ADLS\_WRK\_Alumno\_src
      - DS\_ADLS\_WRK\_Aula\_src
      - DS\_ADLS\_WRK\_Banco\_src
      - DS\_ADLS\_WRK\_Beca\_src

Parquet  
DS\_ADLS\_GLD\_DimSede

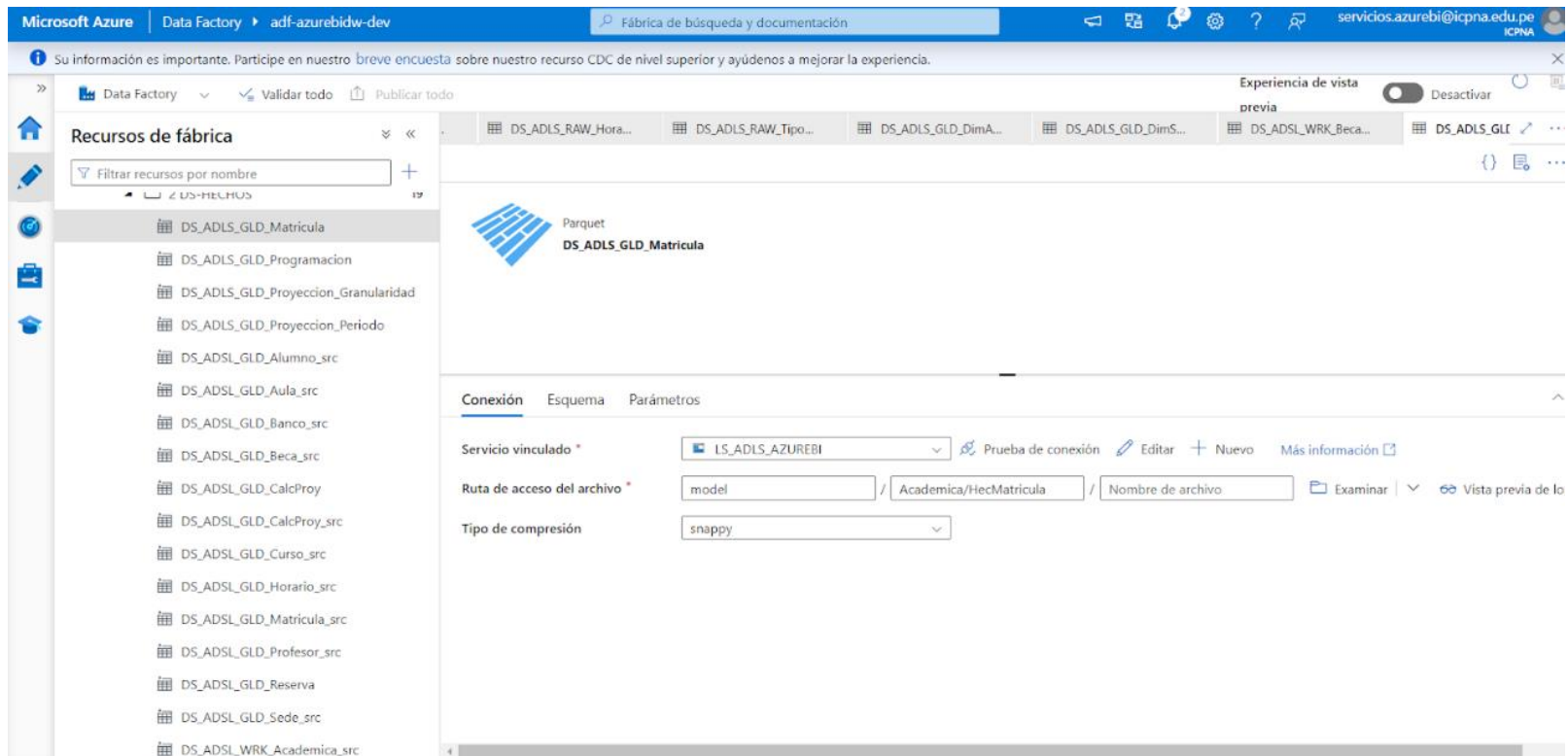
Conexión Esquema Parámetros

Servicio vinculado \*  Prueba de conexión Editar + Nuevo Más información

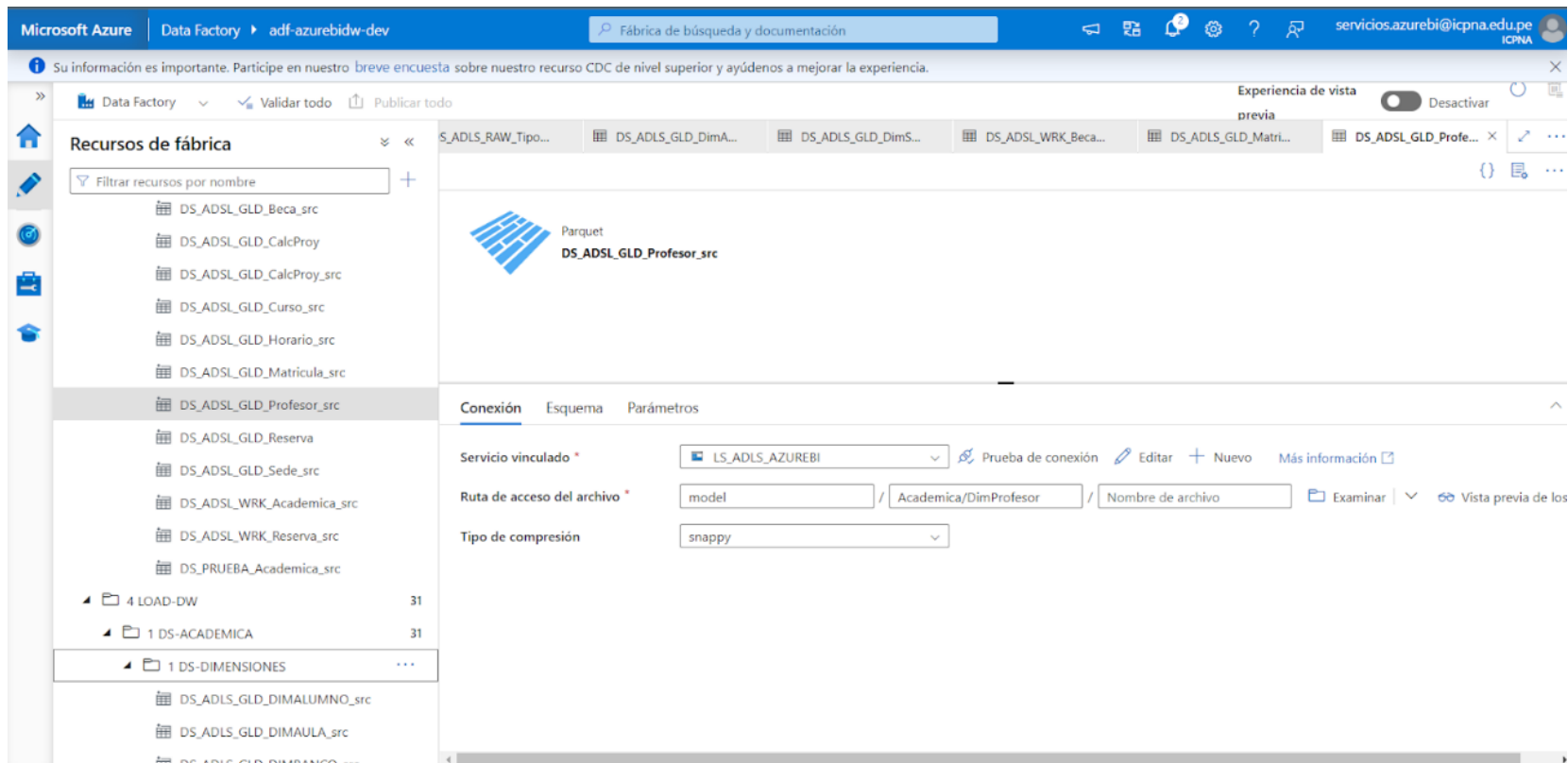
Ruta de acceso del archivo \*  /  /  Examinar Vista previa de los datos

Tipo de compresión

Datasets capa model-gold (Tabla de Hechos)



Datasets carga datawarehouse



Data flows para la limpieza y transformación de datos (capa work-silver)



Microsoft Azure | Data Factory | adf-azurebidw-dev | Fábrica de búsqueda y documentación | servicios.azurebi@icpna.edu.pe

Su información es importante. Participe en nuestro breve encuesta sobre nuestro recurso CDC de nivel superior y ayúdenos a mejorar la experiencia.

Data Factory | Validar todo | Publicar todo | Experiencia de vista: previa | Desactivar

Recursos de fábrica

Filtrar recursos por nombre

Flujos de datos: 36

2. WORK-SILVER: 18

1 DF-ACADEMICA: 15

- DF\_LIMPIEZA\_ALUMNO
- DF\_LIMPIEZA\_AULA
- DF\_LIMPIEZA\_BANCO
- DF\_LIMPIEZA\_BECA
- DF\_LIMPIEZA\_CLIENTE
- DF\_LIMPIEZA\_CURSO
- DF\_LIMPIEZA\_HORARIO
- DF\_LIMPIEZA\_MATRICULA
- DF\_LIMPIEZA\_MATRICULA\_ALUMNO
- DF\_LIMPIEZA\_PROFESOR
- DF\_LIMPIEZA\_RESERVA
- DF\_LIMPIEZA\_RETENCION\_FUTURO
- DF\_LIMPIEZA\_RETENCION\_PASADO
- DF\_LIMPIEZA\_SEDE

DF\_LIMPIEZA\_ALUMNO

Validar | Depuración de flujo de datos | Configuración de depuración

srcAlumno: Importar datos de DS\_ADLS\_RAW\_Alumno

selectAlumno: Se está cambiando el nombre de srcAlumno a selectAlumno con las

lkpAlumnoPersona: Búsqueda en "selectAlumno" desde "drvNombreCompleto"

lkpAlumnoMatric...: Búsqueda en "lkpAlumnoPersona" desde "selectMatricula"

tgtAlumno: Exportar datos a DS\_ADLS\_WRK\_Alumno

srcPersona: Importar datos de DS\_ADLS\_RAW\_Persona

selectPersona: Se está cambiando el nombre de srcPersona a selectPersona con las

lkpPersonaTipoCl...: Búsqueda en "selectPersona" desde "selectTipoCliente"

lkpPersonaDocIden: Búsqueda en "lkpPersonaTipoCliente" desde

drvNombreComp...: Creando o actualizando las columnas "IdPersona"

srcTipoCliente: Importar datos de DS\_ADLS\_RAW\_TipoCliente

selectTipoCliente: Se está cambiando el nombre de srcTipoCliente a

srcMatricula: Importar datos de DS\_ADLS\_RAW\_Matricula

selectMatricula: Se está cambiando el nombre de srcMatricula a selectMatricula con

Parámetros | Configuración

Data flows para la limpieza y transformación de datos (capa work-silver)

Microsoft Azure | Data Factory | adf-azurebidw-dev | Fábrica de búsqueda y documentación | servicios.azurebi@icpna.edu.pe

Su información es importante. Participe en nuestro [breve encuesta](#) sobre nuestro recurso CDC de nivel superior y ayúdenos a mejorar la experiencia.

Data Factory | Validar todo | Publicar todo | Experiencia de vista previa | Desactivar

### Recursos de fábrica

Filtrar recursos por nombre

- Flujos de datos: 36
  - 2. WORK-SILVER: 18
    - 1 DF-ACADEMICA: 15
      - DF\_LIMPIEZA\_ALUMNO
      - DF\_LIMPIEZA\_AULA
      - DF\_LIMPIEZA\_BANCO
      - DF\_LIMPIEZA\_BECA
      - DF\_LIMPIEZA\_CLIENTE
      - DF\_LIMPIEZA\_CURSO
      - DF\_LIMPIEZA\_HORARIO
      - DF\_LIMPIEZA\_MATRICULA**
      - DF\_LIMPIEZA\_MATRICULA\_ALUMNO
      - DF\_LIMPIEZA\_PROFESOR
      - DF\_LIMPIEZA\_RESERVA
      - DF\_LIMPIEZA\_RETENCION\_FUTURO
      - DF\_LIMPIEZA\_RETENCION\_PASADO
      - DF\_LIMPIEZA\_SEDE

DF\_LIMPIEZA\_ALUM... | DF\_LIMPIEZA\_MATR... x

Validar | Depuración de flujo de datos | Configuración de depuración

Parámetros | Configuración

Data flows para la limpieza y transformación de datos (capa model-gold)

Microsoft Azure | Data Factory | adf-azurebidw-dev | Fábrica de búsqueda y documentación | servicios.azurebi@icpna.edu.pe

Su información es importante. Participe en nuestro [breve encuesta](#) sobre nuestro recurso CDC de nivel superior y ayúdenos a mejorar la experiencia.

Data Factory | Validar todo | Publicar todo | Experiencia de vista:  Desactivar

Recursos de fábrica

Filtrar recursos por nombre

- 3. MODEL-GOLD (18)
  - 1 DF-ACADEMICA (18)
    - 1 DF-DIMENSIONES (10)
      - DF\_DIMENSION\_ALUMNO
      - DF\_DIMENSION\_AULA
      - DF\_DIMENSION\_BANCO
      - DF\_DIMENSION\_BECA
      - DF\_DIMENSION\_CLIENTE
      - DF\_DIMENSION\_CURSO
      - DF\_DIMENSION\_HORARIO
      - DF\_DIMENSION\_PRODUCTO
      - DF\_DIMENSION\_PROFESOR
      - DF\_DIMENSION\_SEDE
    - 2 DF-HECHOS (5)
      - DF\_HECHO\_ACADEMICA
      - DF\_HECHO\_MATRICULA
      - DF\_HECHO\_PROGRAMACION
      - DF\_HECHO\_PROVECCION

DF\_LIMPIEZA\_ALUM... | DF\_LIMPIEZA\_MATRI... | DF\_LIMPIEZA\_MATRI... | DF\_LIMPIEZA\_RETEN... | DF\_LIMPIEZA\_RETEN... | DF\_DIMENSION AL...

Validar | Depuración de flujo de datos | Configuración de depuración

srcWrkAlumno | selectWrkAlumno | OrderByIdAlumno | LlaveSubrogadaAlumno | tgtDimAlumno

Importar datos de DS\_ADSL\_WRK\_Alumno\_src

Se está escogiendo las columnas "IdAlumno, NombreAlumno, CodAlumno, Vigencia, Sexo".

Se están ordenando las filas según las columnas "IdAlumno".

Agregue una nueva clave AlumnoKey comenzando por 1 con el paso 1

Exportar datos a DS\_ADSL\_GLD\_DimAlumno

Agregar origen

Parámetros | Configuración

+ Nuevo

Data flows para la limpieza y transformación de datos (capa model-gold)

The screenshot displays the Microsoft Azure Data Factory interface. The top navigation bar shows 'Microsoft Azure', 'Data Factory', and the specific instance 'adf-azurebidw-dev'. The main workspace is titled 'Fábrica de búsqueda y documentación'. On the left, the 'Recursos de fábrica' (Factory Resources) pane shows a hierarchical view of the data lake structure, including folders like '3. MODEL-GOLD' and '1 DF-ACADEMICA', and various dimension and fact data flows such as 'DF\_DIMENSION\_PRODUCTO'. The central canvas shows a data flow pipeline with five tasks: 1. 'srcWrkProducto' (Importar datos de DS\_ADLS\_WRK\_TipoProducto\_s rc), 2. 'selectWrkProducto' (Se está cambiando el nombre de srcWrkProducto a selectWrkProducto...), 3. 'OrderByIdProducto' (Se están ordenando las filas según las columnas "IdTipoProducto"), 4. 'LlaveSubrogadaProducto' (Agregue una nueva clave ProductKey...), and 5. 'tgtDimProducto' (Exportar datos a DS\_ADLS\_GLD\_DimProducto). Below the pipeline, there are tabs for 'Parámetros' and 'Configuración', with a '+ Nuevo' button visible. The right side of the interface includes a search bar and a vertical toolbar with various icons for editing and viewing the pipeline.

Data flows para la limpieza y transformación de datos (capa model-gold)

Microsoft Azure | Data Factory | adf-azurebidw-dev | Fábrica de búsqueda y documentación | servicios.azurebi@icpna.edu.pe

Su información es importante. Participe en nuestro [breve encuesta](#) sobre nuestro recurso CDC de nivel superior y ayúdenos a mejorar la experiencia.

Validar todo | Publicar todo | Experiencia de vista previa | Desactivar

### Recursos de fábrica

Filtrar recursos por nombre

- DF\_DIMENSION\_ALUMNO
- DF\_DIMENSION\_AULA
- DF\_DIMENSION\_BANCO
- DF\_DIMENSION\_BECA
- DF\_DIMENSION\_CLIENTE
- DF\_DIMENSION\_CURSO
- DF\_DIMENSION\_HORARIO
- DF\_DIMENSION\_PRODUCTO
- DF\_DIMENSION\_PROFESOR
- DF\_DIMENSION\_SEDE
- 2 DF-HECHOS
- DF\_HECHO\_ACADEMICA
- DF\_HECHO\_MATRICULA
- DF\_HECHO\_PROGRAMACION
- DF\_HECHO\_PROYECCION
- DF\_HECHO\_RESERVA
- 3 SANDBOX

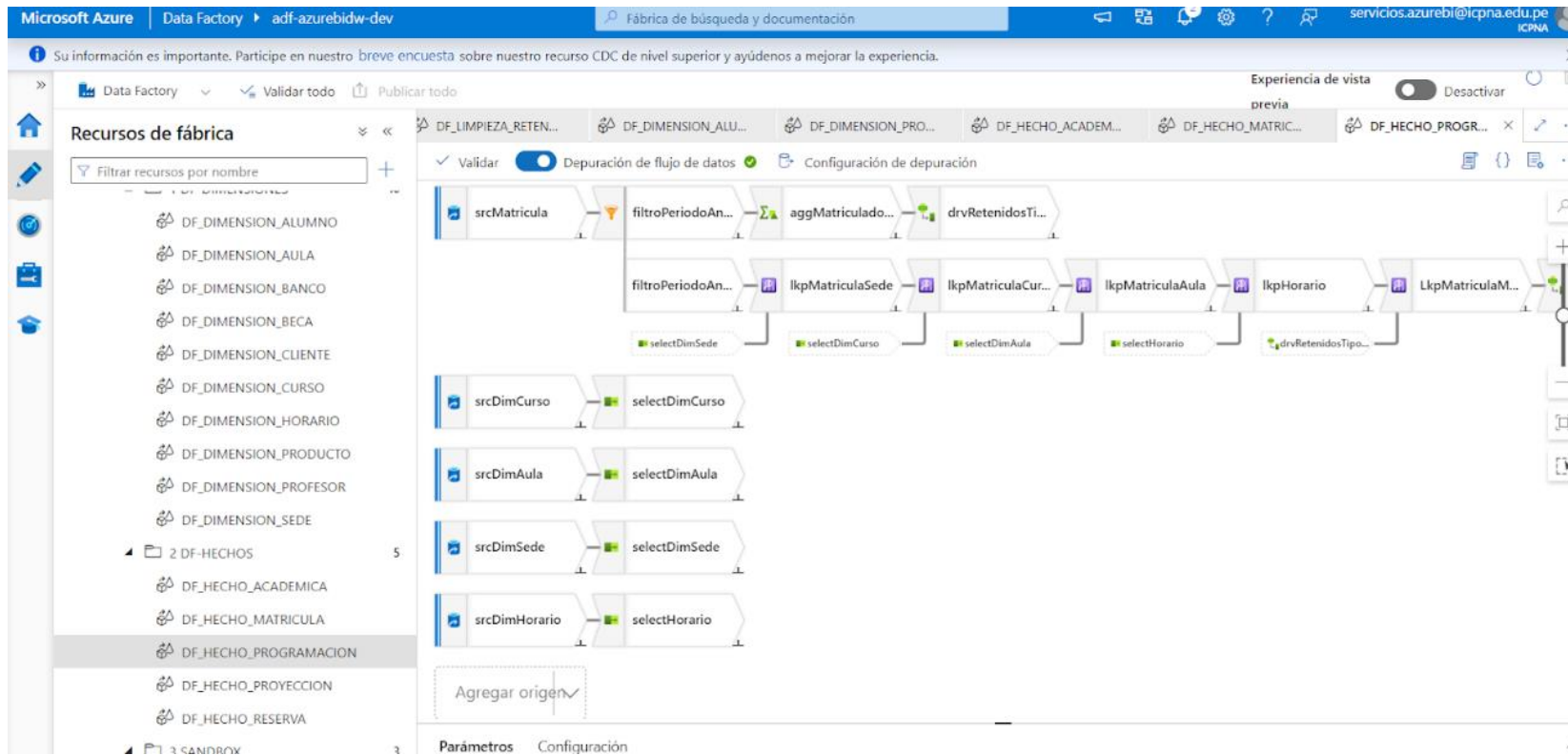
Validar | Depuración de flujo de datos | Configuración de depuración

srcMatricula → drvFechaMatric... → JoinMatriculaDI... → JoinMatriculaSe... → JoinMatriculaCu... → JoinMatriculaH... → JoinMatriculaAu... → JoinMatriculaPr... → JoinMatricula

srcDimAlumno, srcDimSede, srcDimCurso, srcDimHorario, srcDimAula, srcDimProfesor, srcDimTiempo

Parámetros | Configuración

Data flows para la limpieza y transformación de datos (capa model-gold)



Data flows para la limpieza y transformación de datos (capa model-gold)

Microsoft Azure | Data Factory | adf-azurebidw-dev | Fábrika de búsqueda y documentación | servicios.azurebi@icpna.edu.pe

Su información es importante. Participe en nuestro breve encuesta sobre nuestro recurso CDC de nivel superior y ayúdenos a mejorar la experiencia.

Validar | Depuración de flujo de datos | Configuración de depuración

Recursos de fábrica

- DF\_DIMENSION\_CLIENTE
- DF\_DIMENSION\_CURSO
- DF\_DIMENSION\_HORARIO
- DF\_DIMENSION\_PRODUCTO
- DF\_DIMENSION\_PROFESOR
- DF\_DIMENSION\_SEDE
- 2 DF-HECHOS
  - DF\_HECHO\_ACADEMICA
  - DF\_HECHO\_MATRICULA
  - DF\_HECHO\_PROGRAMACION
  - DF\_HECHO\_PROYECCION
  - DF\_HECHO\_RESERVA
- 3 SANDBOX
  - DF\_HECHO\_PROYECCION\_V1
  - DF\_HECHO\_PROYECCION\_V2
  - DF\_HECHO\_PROYECCION\_V3

Power Query

DF\_DIMENSION\_ALU... | DF\_DIMENSION\_PRO... | DF\_HECHO\_ACADEM... | DF\_HECHO\_MATRIC... | DF\_HECHO\_PROGRA... | DF\_HECHO\_RESERVA

srcReserva | drvFechaReservaE... | joinReservaAlumno | joinReservaSede | joinReservaCurso | joinReservaHorario | joinReserva...

srcDimAlumno | srcDimSede | srcDimCurso | srcDimHorario

Parámetros | Configuración