





Programa de especialización

DATA DRIVEN DECISIONS 2026

Nivel Intermedio

SESIÓN 5: Introducción al modelado de datos

TRANSFORMAR AHORA,
trascender siempre

Sesión 5

Introducción al modelado de datos

Logro de aprendizaje

Al finalizar la sesión, lograrás:

Comprender los conceptos básicos del modelado de datos.

Identificar y establecer relaciones entre tablas.

Aplicar buenas prácticas en el modelado de datos.

Desarrollar un modelo de datos básico.

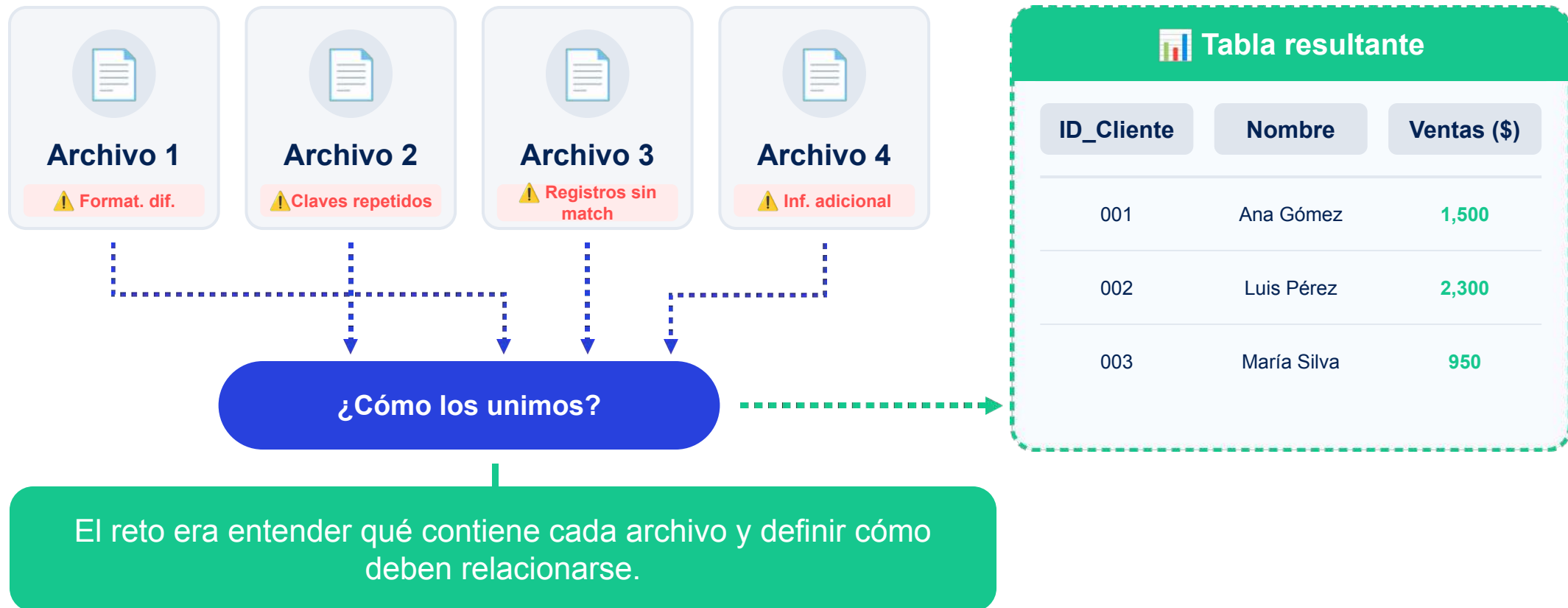
Contenido:

- 1. Caso práctico: avanzado
- 2. ¿Qué es un modelo de datos?
- 3. Conceptos básicos de modelado de datos
- 4. OLAP vs. OLTP

Caso práctico: avanzado

¿Qué debíamos resolver?

Caso práctico: de archivos con problemas a una tabla lista para analizar



La estrategia antes de tocar Power Query

Antes de unir: primero entendemos y diseñamos

01 · Diagnosticar 🔍

Con Gemini:

- ¿Qué representa una fila de cada archivo?
- ¿Cuál es su identificador?
- ¿Necesitamos una clave compuesta?
- ¿Qué problemas debemos corregir antes de unir?



02 · Diseñar la unión 🧩

- **APPEND**
¿Qué archivos agregan nuevas filas?
- **MERGE / JOIN**
¿Qué información permite agregar nuevas columnas?
- **COINCIDENCIAS**
¿Qué registros debemos conservar cuando no existe correspondencia?



03 · Resultado 🎯

Antes de ejecutar, debemos poder responder:

¿Qué esperamos obtener después de combinar los archivos?

⚠️ No empezamos por “hacer un Merge”. Empezamos por entender qué relación existe entre los datos.

De la estrategia a la ejecución

Ahora sí: llevamos la estrategia a Power Query

Lo que le pedimos a Gemini

No decir simplemente:

✗ **“Une estos archivos.”**

Sino entregarle un Diagnóstico con:

- Nivel de detalle
- Claves
- Problemas encontrados
- Estrategia de unión
- Resultado esperado

Lo que ejecutamos en Power Query

1. Importar archivos
↓
2. Corregir y transformar datos
↓
3. Crear/verificar claves
↓
4. Aplicar Append / Merge
↓
5. Validar el resultado
↓
6. Cargar los datos

¿Cómo sabemos que quedó bien?

- ¿El número de filas cambió como esperábamos?
- ¿Qué registros no encontraron coincidencia?
- ¿Aparecieron combinaciones inesperadas?
- ¿Qué representa ahora cada fila?
- ¿El resultado corresponde con lo que habíamos definido?

 El éxito en Power Query no depende de dar clics, sino de la claridad de las instrucciones que diseñamos previamente.

1.¿Qué es un modelo de datos?

¿Qué es un modelo de datos?

Definición:

Proceso de estructuración

Proceso de estructuración y organización de datos para su análisis. Involucra la creación de tablas y la definición de relaciones entre ellas.

Importancia

Facilita la comprensión y el análisis de grandes volúmenes de datos. Mejora la precisión y eficiencia de los informes y dashboards.

Objetivos del Modelado de Datos

Organización

Crear una estructura lógica y coherente para los datos.

Integridad

Asegurar la consistencia y exactitud de los datos.

Eficiencia

Optimizar el rendimiento de consultas y cálculos en los informes.



2. Conceptos básicos de modelado de datos

Tablas

Tablas de Hechos

Registran eventos que queremos analizar.
Ejemplo: una fila por cada matrícula de un estudiante en un curso.

Tablas de Dimensiones

Describen el contexto del hecho: estudiante, curso, periodo, docente, facultad, entre otros.

HECHO = ¿qué ocurrió? → una matrícula

DIMENSIONES = ¿quién, qué, cuándo o bajo qué contexto ocurrió? → estudiante, curso, periodo, docente

Tablas (ejemplo)

ID_Matricula	ID_Estudiante	ID_Curso	ID_Periodo	ID_Docente	Créditos	Nota_Final	Asistencia_%
M001	E023	C145	2026-1	D018	4	15.8	92%
M002	E104	C145	2026-1	D018	4	12.5	81%
M003	E087	C210	2026-1	D044	3	9.2	68%

Cada fila representa **la matrícula de un estudiante en un curso durante un periodo académico.**

Tablas (ejemplo)

Dim_Estudiantes

ID_Estudiante	Programa	Facultad	Cohorte	Nivel
E023	Economía	Ciencias Sociales	2024	Pregrado
E104	Derecho	Derecho	2023	Pregrado
E087	Psicología	Psicología	2024	Pregrado

Dim_Cursos

ID_Curso	Curso	Departamento	Créditos	Modalidad
C145	Estadística Aplicada	Economía	4	Presencial
C210	Investigación I	Psicología	3	Híbrida

Dim_Periodos

ID_Periodo	Año	Semestre
2026-1	2026	1
2026-2	2026	2

Dim_Docentes

ID_Docente	Departamento	Categoría
D018	Economía	Ordinario
D044	Psicología	Contratado

Relaciones - Definición

Definición

Un vínculo establecido entre dos tablas que define cómo los datos de una tabla se conectan con los datos de otra tabla.

Para qué sirven

Son fundamentales para integrar y analizar datos provenientes de diferentes fuentes o tablas dentro de un mismo modelo de datos.



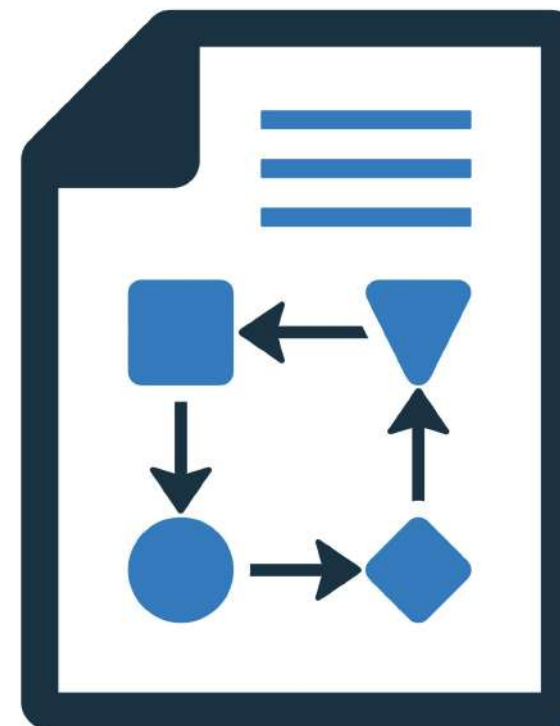
Relaciones - Importancia

Análisis integrado

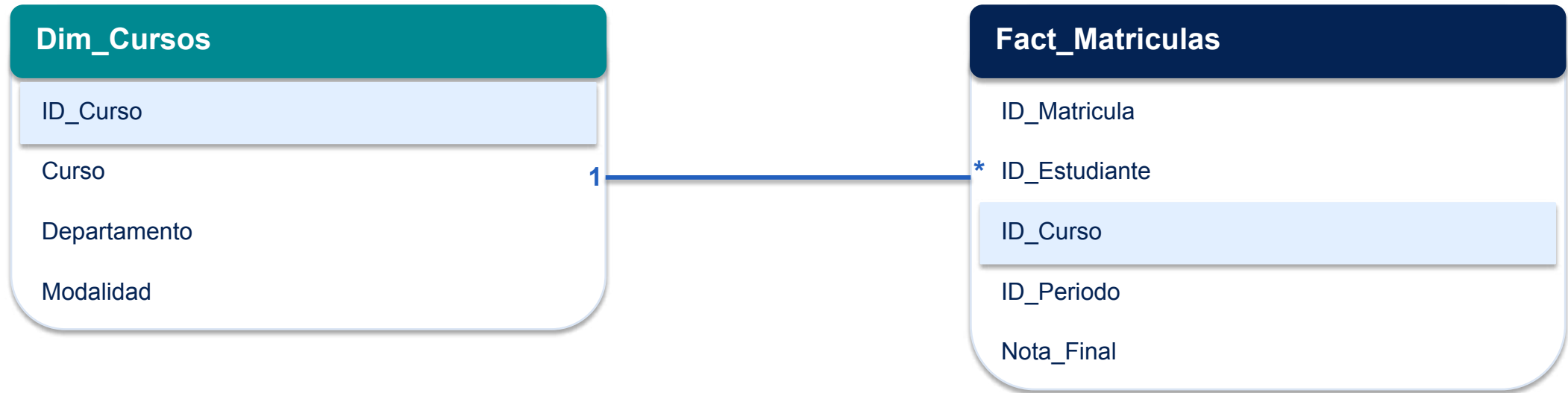
Las relaciones permiten a los analistas y a herramientas de BI, como Power BI, combinar y comparar datos de manera efectiva. Sin relaciones, cada tabla se trataría de manera aislada.

Propagación de filtros

Una relación de modelo propaga los filtros aplicados en la columna de una tabla a otra tabla. Los filtros se propagan siempre que haya una ruta de relación que seguir.

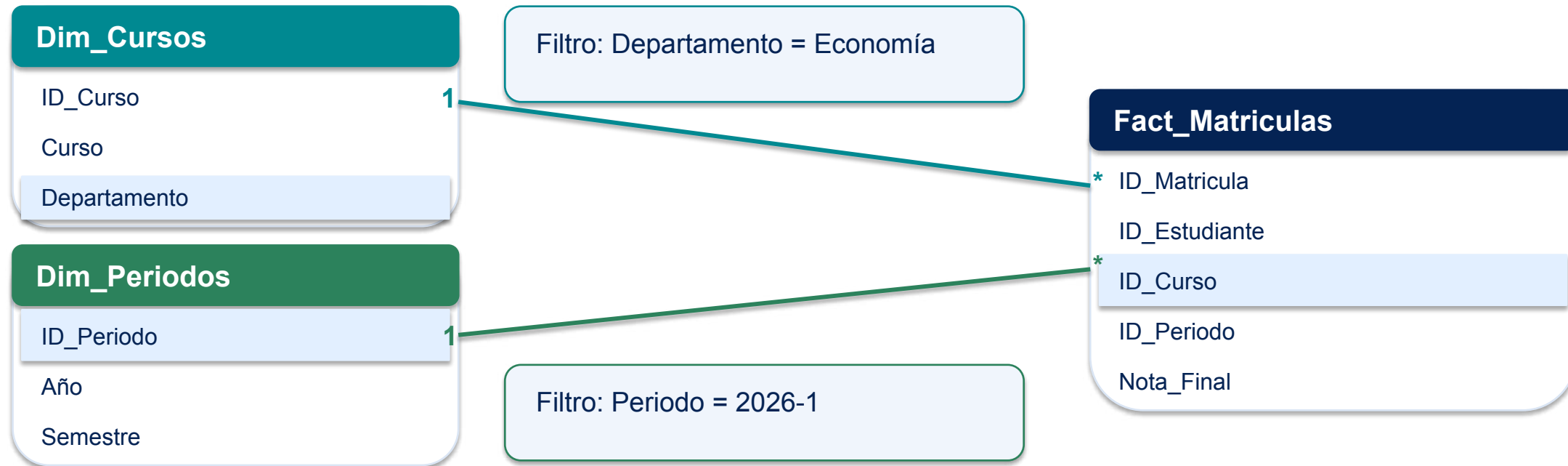


Relaciones - Importancia



Gracias a esta relación podemos analizar las matrículas por curso, departamento o modalidad sin duplicar esos datos en cada registro.

Relaciones - Importancia (ejemplo)



Ejemplo: el filtro se propaga desde Curso y Periodo hacia Fact_Matriculas para analizar únicamente las matrículas de Economía en 2026-1.

Llaves primarias y foráneas

Llave Primaria

Un identificador único para cada fila. Ejemplo: ID_Curso identifica de forma única cada curso en Dim_Cursos.

Dim_Cursos

ID_Curso 🔑 PK

Curso

Departamento

Llave Foránea

Un campo que hace referencia a la llave primaria de otra tabla. Ejemplo: ID_Curso en Fact_Matriculas indica en qué curso ocurrió la matrícula.

Fact_Matriculas

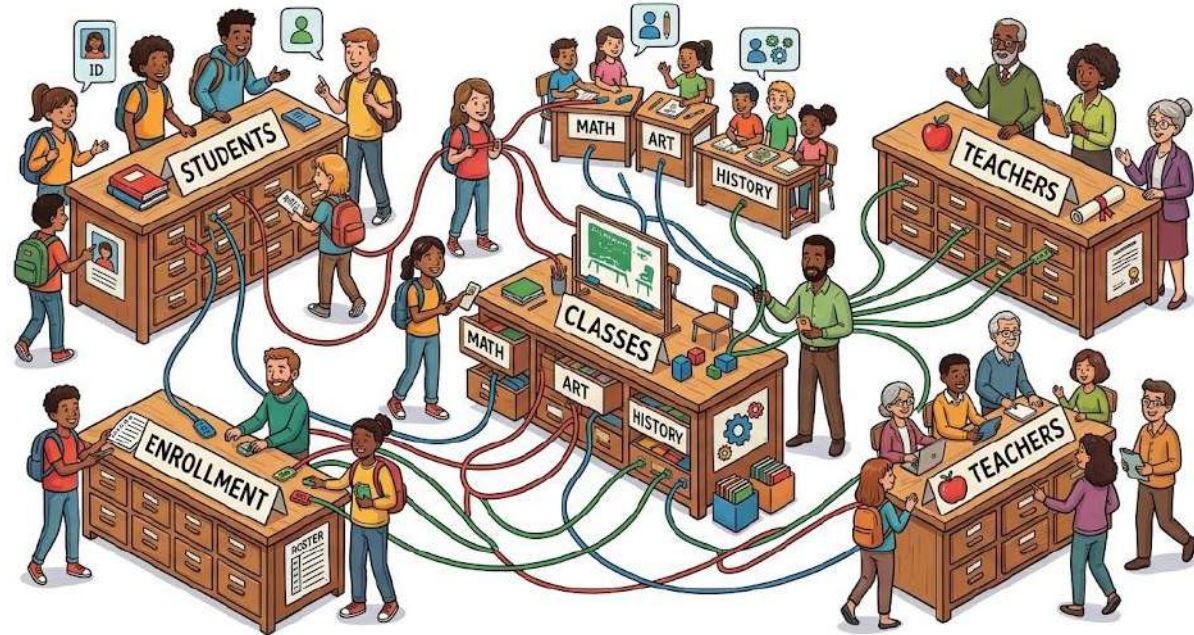
ID_Matricula 🔑 PK

ID_Curso ↗ FK

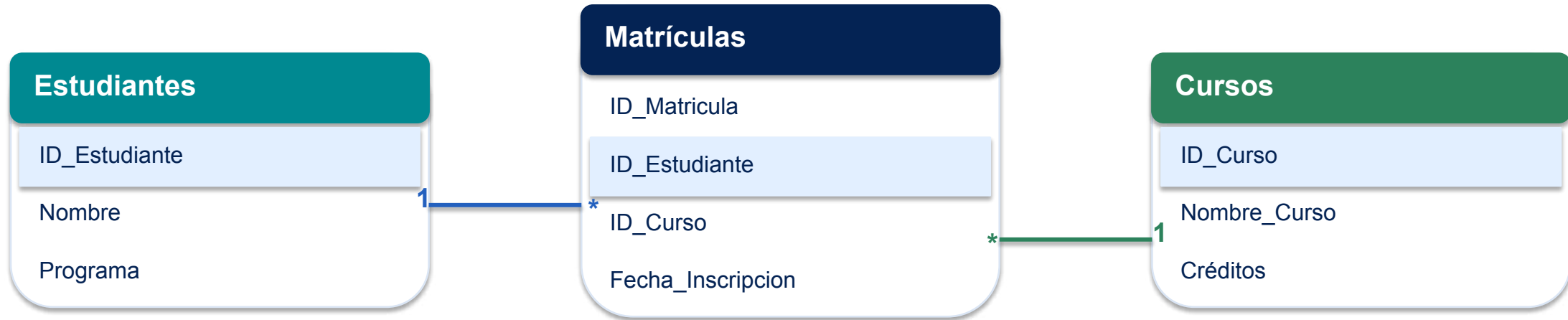
ID_Estudiente ↗ FK

Cardinalidad

- Uno a Uno (1:1)
- Uno a Muchos (1:*)
- Muchos a Muchos (*:*)



Llaves primarias y foráneas (ejercicio)



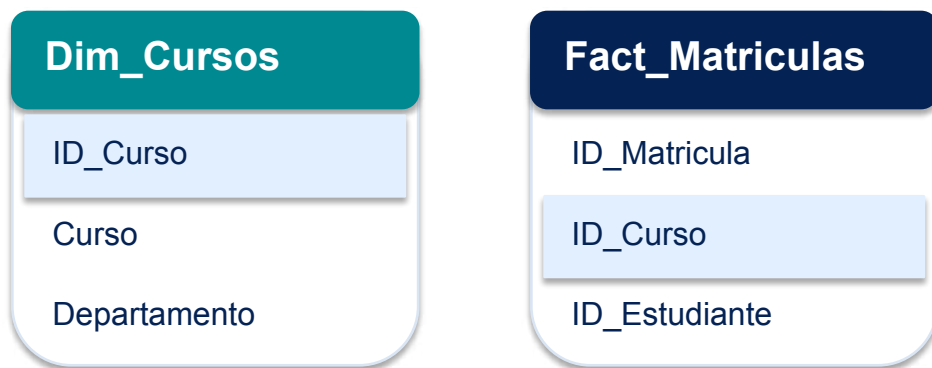
Pregunta: ¿cuáles son las llaves primarias y foráneas? ¿Qué relaciones permiten resolver que un estudiante curse varias asignaturas y un curso tenga varios estudiantes?



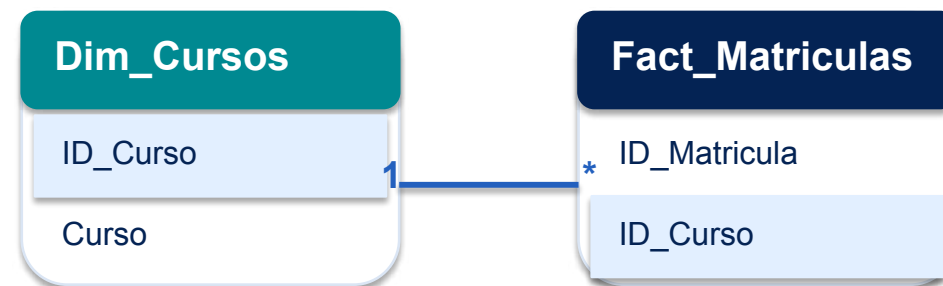
La tabla Matriculas actúa como puente entre Estudiantes y Cursos y registra el evento que queremos analizar.

Ejemplo de relación - Power BI

1. Identificar el campo compartido



2. Crear la relación



Izquierda: ID_Curso aparece en ambas tablas. Derecha: se establece una relación 1 a muchos entre Dim_Cursos y Fact_Matriculas.

Ejemplo de relación - Power BI

Vista de modelo

Dim_Cursos

ID_Curso

Curso

Fact_Matriculas

ID_Matricula

ID_Curso

Clic derecho sobre la línea → Editar relación

Editar relación

Tabla	Dim_Cursos
Columna	ID_Curso
Tabla relacionada	Fact_Matriculas
Cardinalidad	Uno a varios (1:*)

Clic derecho sobre la relación, opción de propiedades para editarla y revisar las tablas, columnas y cardinalidad utilizadas.

Ejemplo de relación - Power BI

Vista de modelo

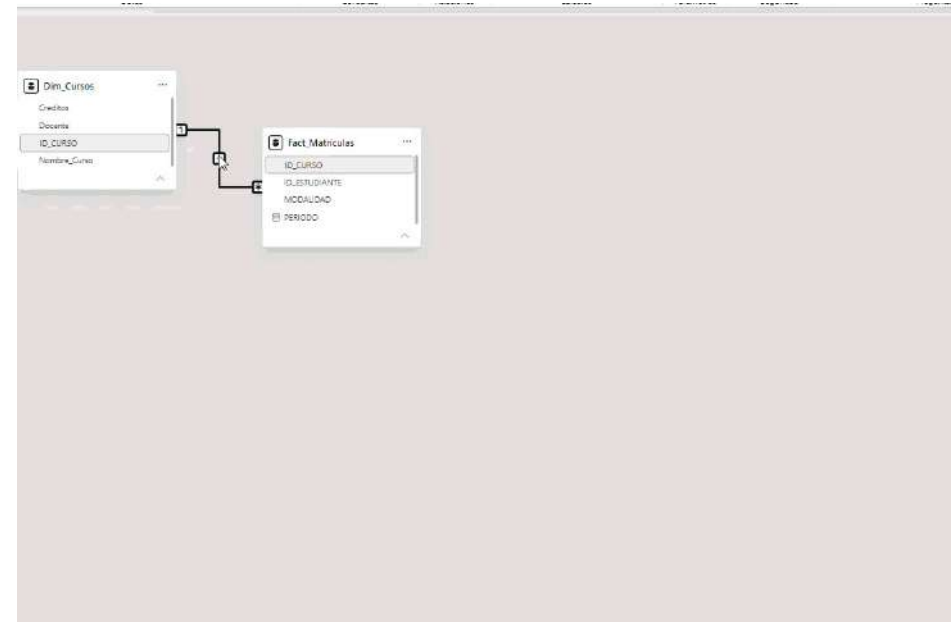


Clic derecho sobre la línea → Editar relación

Ejemplo de relación - Power BI

Editar relación

Tabla	Dim_Cursos
Columna	ID_Curso
Tabla relacionada	Fact_Matriculas
Cardinalidad	Uno a varios (1:*)



Clic derecho sobre la relación, opción de propiedades para editarla y revisar las tablas, columnas y cardinalidad utilizadas.

Ejemplo de relación - Power BI

Crear nueva relación

Tabla	Dim_Cursos
Columna	ID_Curso
Tabla relacionada	Fact_Matriculas
Columna relacionada	ID_Curso
Cardinalidad	Uno a varios (1:*)
Dirección de filtro	Único

Ejemplo de configuración de una relación entre Dim_Cursos[ID_Curso] y Fact_Matriculas[ID_Curso].

Ejemplo de relación - Power BI

Dirección de filtro cruzado según el tipo de cardinalidad:

Uno a Varios

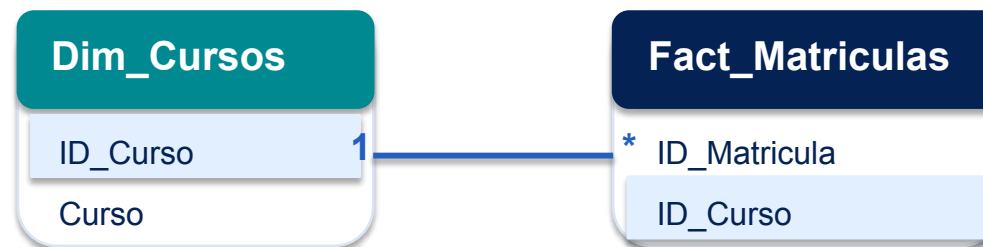
Opciones de filtro cruzado: Único, Ambos.

Uno a Uno

Opciones de filtro cruzado: Ambos.

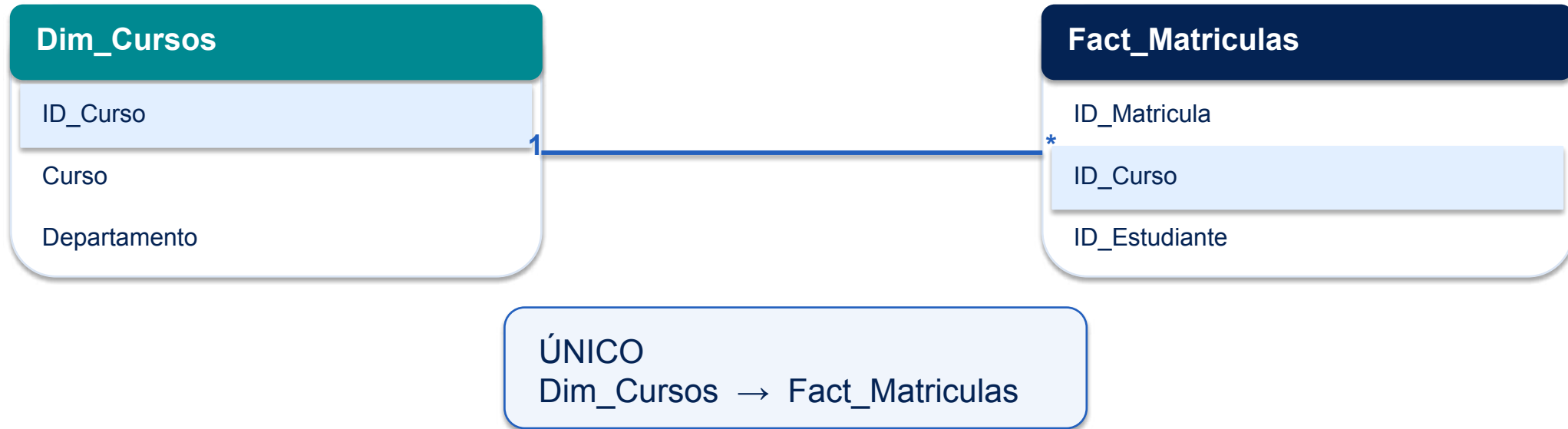
Varios a Varios

Opciones de filtro cruzado: Único (definiendo qué tabla filtra a la otra), Ambos.



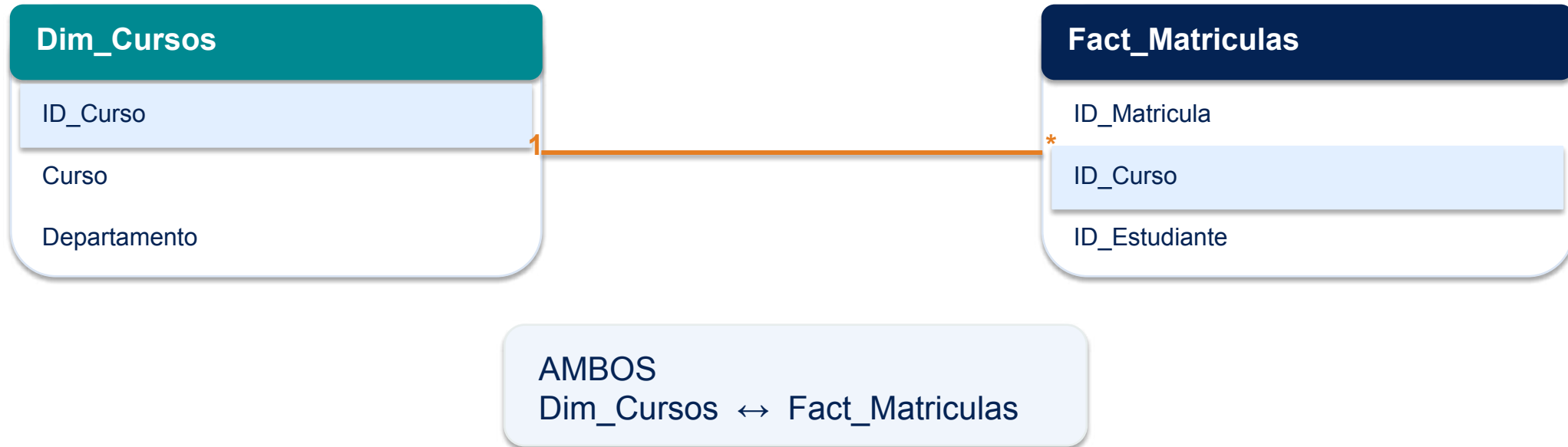
En un modelo estrella, lo habitual es que el filtro viaje desde la dimensión hacia la tabla de hechos.

Ejemplo de relación - Power BI



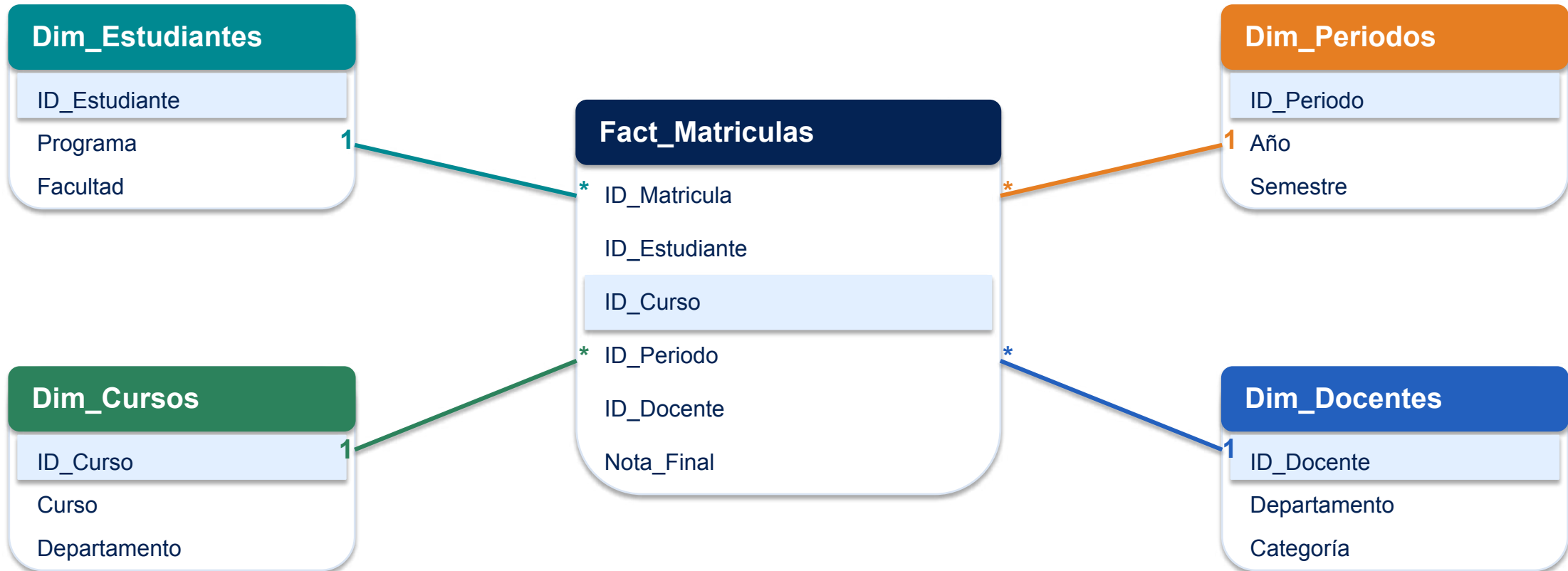
Único: el filtro viaja desde la dimensión hacia la tabla de hechos. Es la dirección más habitual en un modelo estrella.

Ejemplo de relación - Power BI



Ambos: el filtro puede viajar en los dos sentidos. Puede ser útil en casos específicos, pero debe utilizarse con criterio para evitar ambigüedades.

Ejemplo de relación - Power BI

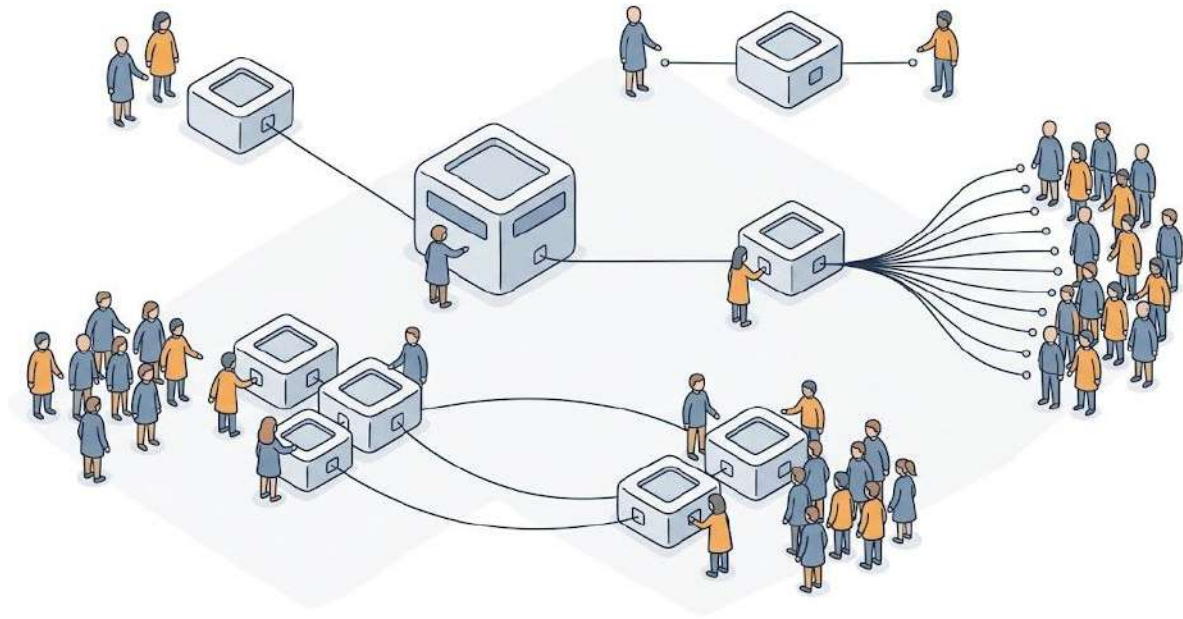


Fact_Matriculas al centro y dimensiones de Estudiantes, Cursos, Periodos y Docentes alrededor.

Ejercicio práctico - Tablas y Relaciones

Contexto

Una universidad quiere analizar sus matrículas y notas finales: cuántos estudiantes se matriculan, en qué cursos, en qué periodos y con qué resultados. El reto es diseñar un modelo que permita responder estas preguntas sin duplicar información.



Ejercicio práctico - Tablas y Relaciones

Instrucciones:

01

Leer

Leer el enunciado del problema.

02

Diseñar

Diseñar las tablas necesarias.

03

Identificar

Identificar tablas de hechos y dimensiones.

04

Llaves primarias

Definir las llaves primarias para cada tabla.

05

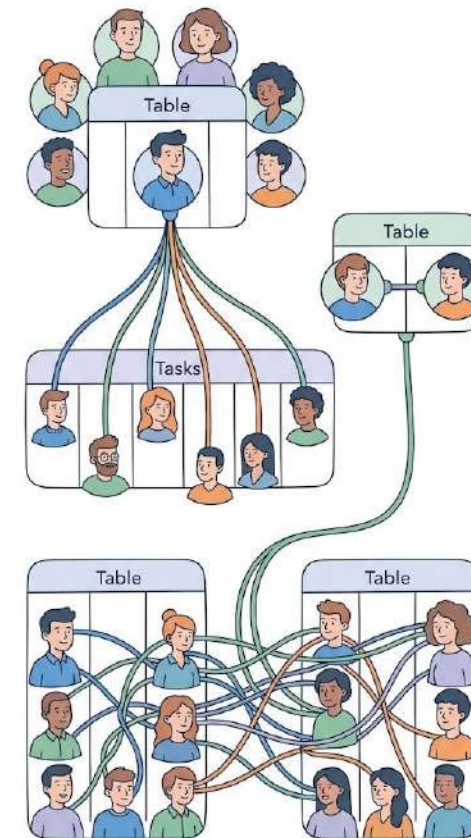
Llaves foráneas

Definir las llaves foráneas para establecer relaciones entre las tablas.

06

Diagramar

Dibujar un diagrama simple que muestre las tablas y sus relaciones.



Ejercicio práctico - Requisitos del sistema

Estudiantes

Cada estudiante tiene un ID único, programa, facultad y cohorte.

Cursos

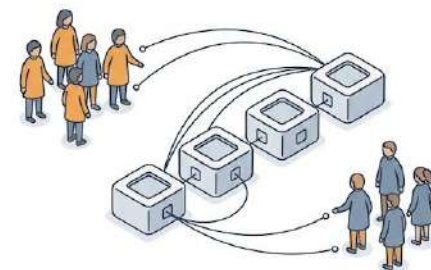
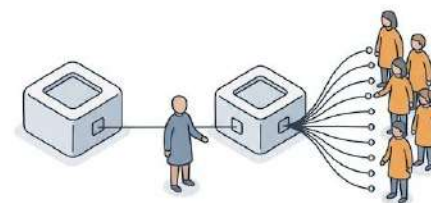
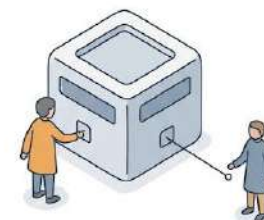
Cada curso tiene un ID único, nombre, departamento, número de créditos y modalidad.

Docentes

Cada docente tiene un ID único, nombre y departamento académico.

Matrículas

Cada registro representa a un estudiante matriculado en un curso durante un periodo. Incluye fecha de inscripción, nota final y porcentaje de asistencia.



Ejercicio práctico - Ejemplo de solución

Dim_Estudiantes

ID_Estudiante (PK), Programa, Facultad, Cohorte.

Dim_Cursos

ID_Curso (PK), Nombre_Curso, Departamento, Créditos, Modalidad.

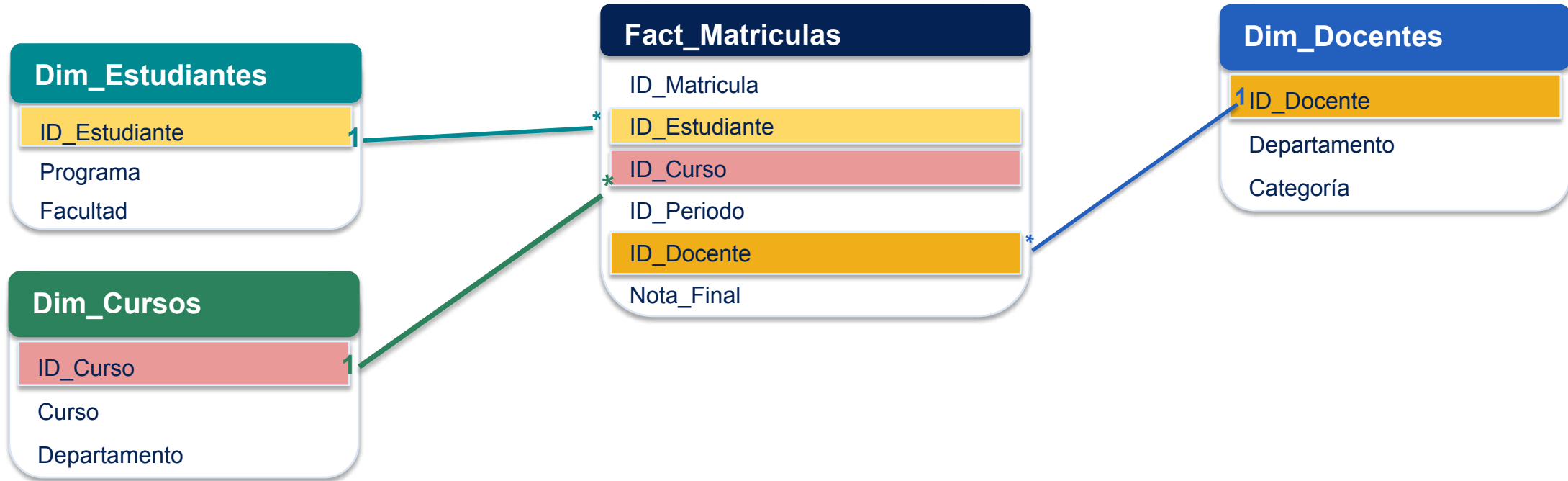
Dim_Docentes

ID_Docente (PK), Nombre, Departamento.

Fact_Matriculas

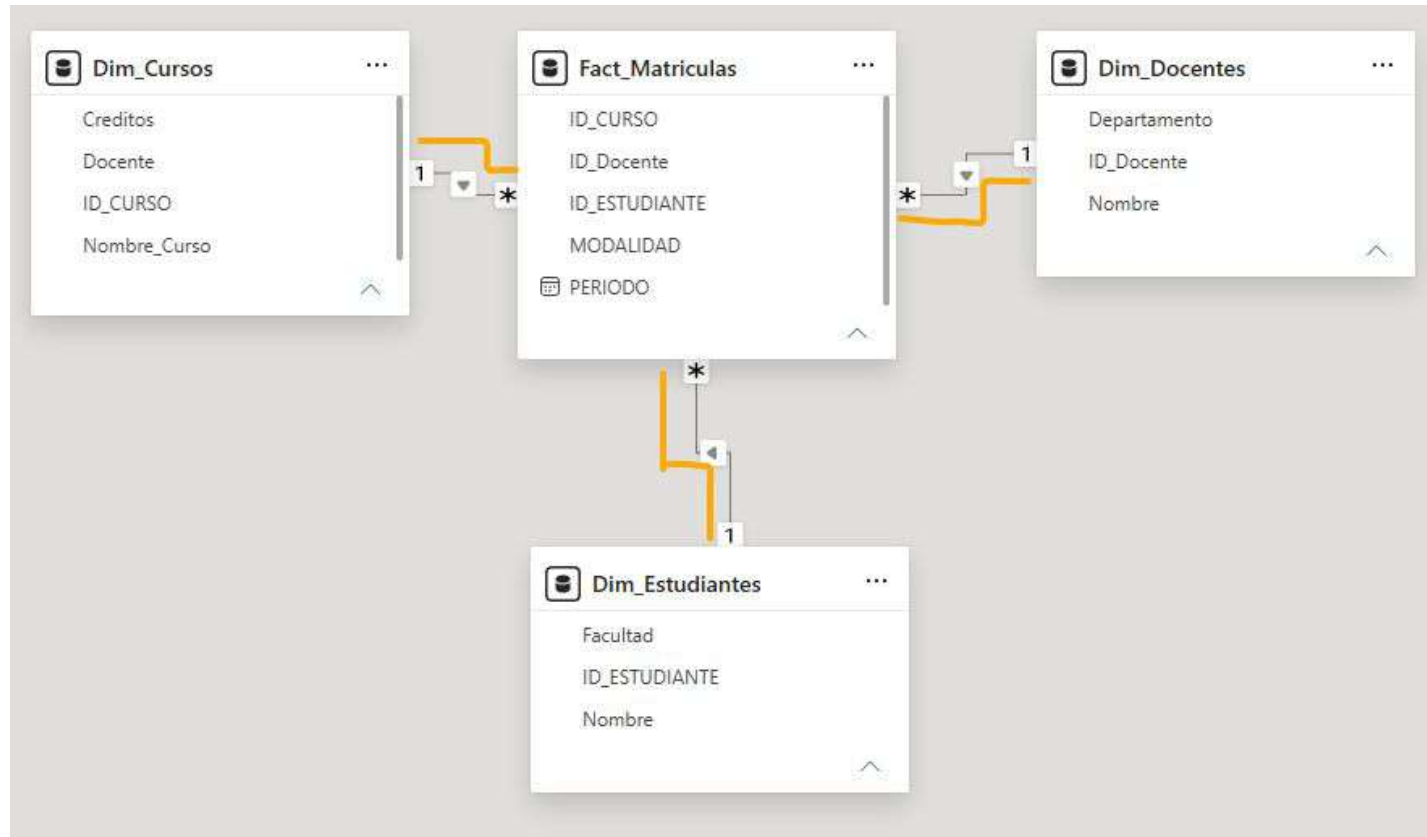
ID_Matricula (PK), ID_Estudiante (FK), ID_Curso (FK), ID_Periodo (FK), ID_Docente (FK), Nota_Final, Asistencia_% .

Ejemplo de relación - Power BI



Fact_Matriculas al centro y dimensiones de Estudiantes, Cursos, Periodos y Docentes alrededor.

Ejemplo de relación - Power BI



Ejercicio práctico - Uso de Gemini como auditor

Actúa como revisor de un modelo de datos analítico para una universidad.

Queremos analizar matrículas, rendimiento académico y evolución de estudiantes por periodo.

Hemos diseñado las siguientes tablas:

[PEGAR AQUÍ LAS TABLAS PROPUESTAS POR EL GRUPO]

Revisa nuestro modelo, pero **no lo rediseñes todavía**.

Evalúa específicamente:

- si identificamos correctamente la tabla de hechos y las dimensiones;
- si las llaves primarias son realmente únicas;
- si las llaves foráneas están correctamente ubicadas;
- si las relaciones 1:* tienen sentido;
- si existe alguna relación muchos a muchos que pueda generar problemas;
- si estamos duplicando información innecesariamente.

Para cada problema encontrado, explica **por qué podría afectar el análisis en Power BI**.

Finalmente, haznos máximo 3 preguntas que deberíamos responder antes de modificar el modelo.

Ejercicio práctico - Uso de Gemini

En el mismo chat:

Ahora, teniendo en cuenta nuestras respuestas y las correcciones que aceptamos, propón una versión mejorada del modelo.

Preséntala indicando:

1. tabla de hechos;
2. dimensiones;
3. llave primaria de cada tabla;
4. llaves foráneas;
5. cardinalidad de cada relación.

Explica cada relación en lenguaje sencillo para una persona que está aprendiendo modelado de datos.

3.OLAP vs. OLTP

¿Qué es OLAP?

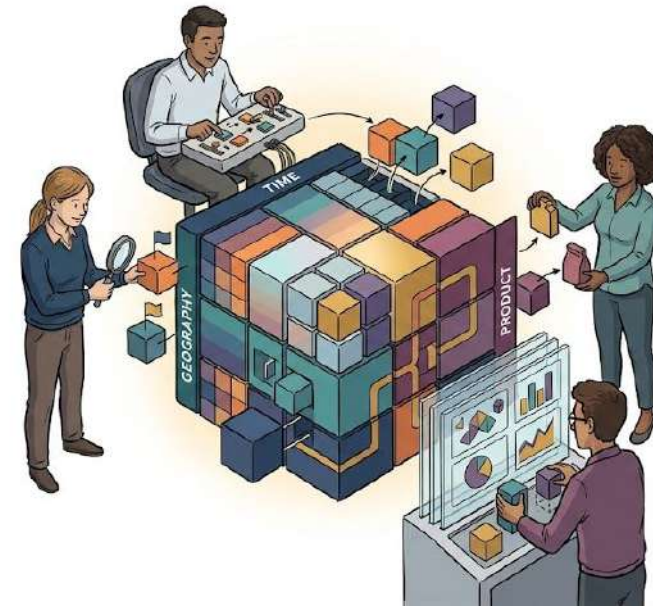
OLAP es un tipo de sistema de base de datos diseñado para:

Análisis de Datos

Facilita la consulta y análisis de grandes volúmenes de datos históricos.

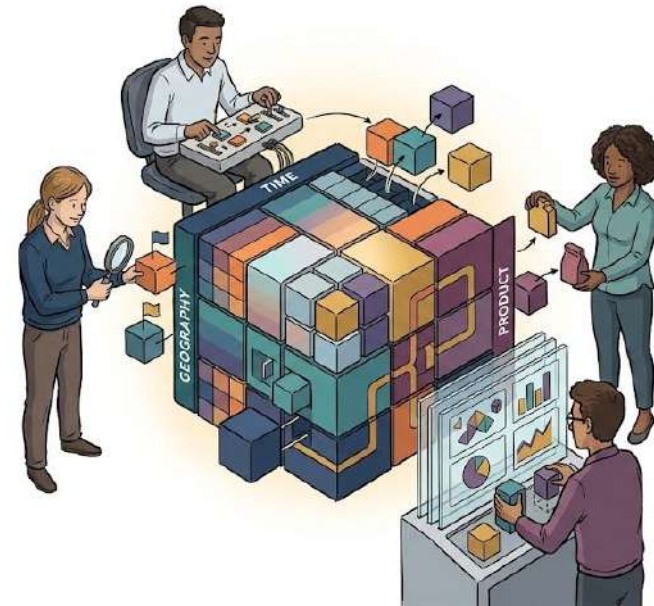
Toma de Decisiones

Ayuda a los gerentes y analistas a tomar decisiones informadas basadas en datos.



OLAP - Características clave

- Consultas Complejas
- Datos Históricos
- Usuarios
- Ejemplos
 - Sistemas de inteligencia de negocios (BI) como data warehouses.



¿Qué es OLTP?

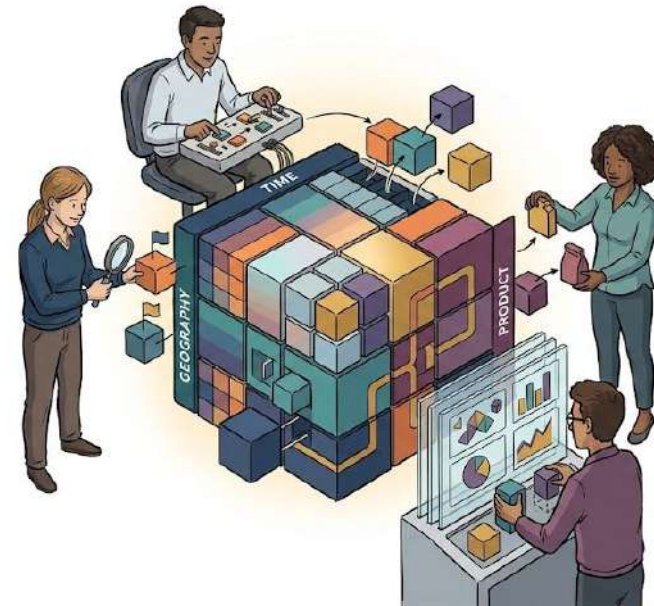
OLTP es un tipo de sistema de base de datos diseñado para:

- **Gestión de Transacciones**
- **Alta Velocidad y Precisión**



OLTP - Características clave

- Consultas Simples
- Datos Actuales
- Usuarios
- Ejemplos
 - Sistemas de punto de venta (POS), sistemas de reservas en línea.



OLAP vs. OLTP

OLAP

Maneja todos los datos históricos.

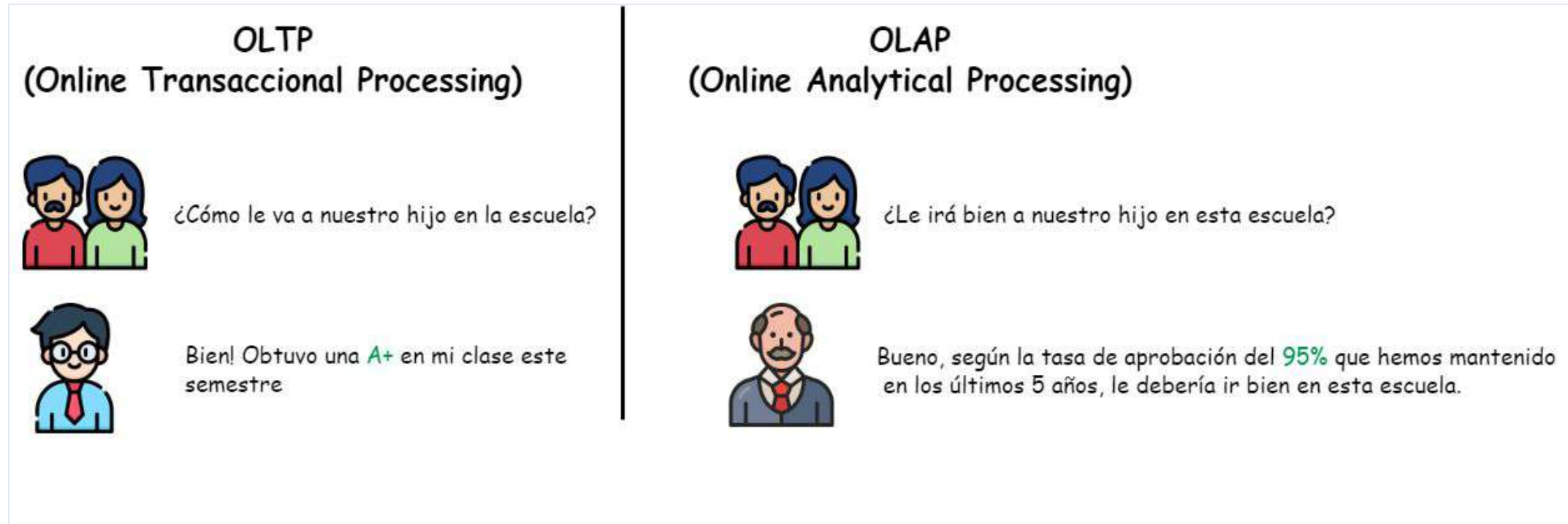
- **Objetivo:** Tomar decisiones a partir de grandes fuentes de datos
- **Consultas:** Complejas.
- **Procesamiento:** Menor velocidad; solo requiere operaciones de lectura.

OLTP

Maneja datos operacionales recientes.

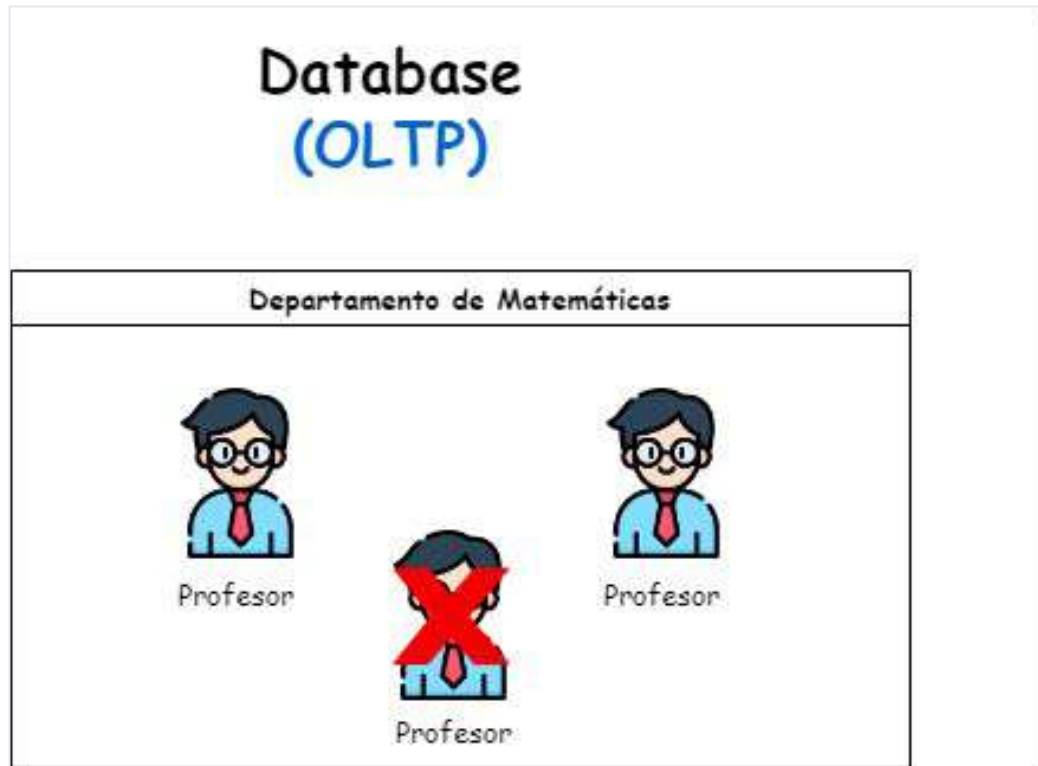
- **Objetivo:** Realizar las operaciones diarias.
- **Consultas:** Simples.
- **Procesamiento:** Mayor velocidad; requiere operaciones de lectura y escritura.

Ejemplo OLAP vs. OLTP



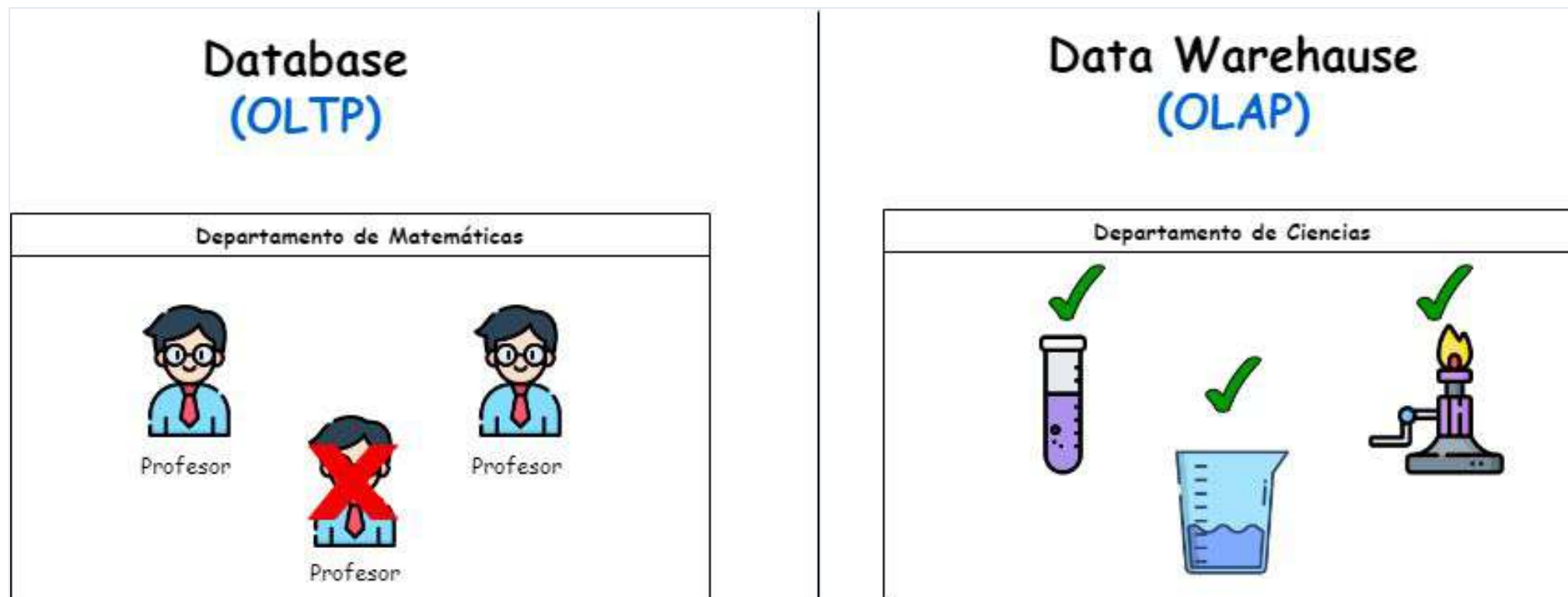
Debido a que el tipo de datos almacenados en una base de datos es diferente a un almacén de datos, el tipo de información que cada uno puede proporcionar también difiere.

Ejemplo OLAP vs. OLTP



La escuela recorta fondos y el director pregunta a la base de datos: “¿De qué maestro debería deshacerme según la cantidad de quejas?” (consulta operacional, tipo OLTP).

Ejemplo OLAP vs. OLTP



La escuela recibe una subvención y debe elegir en qué departamento invertir. El almacén de datos está diseñado para responder exactamente ese tipo de pregunta (consulta analítica, tipo OLAP).

Repaso

Es hora de poner a prueba lo aprendido. Únete a la actividad interactiva ingresando al enlace desde tu computadora o dispositivo móvil.



Respuestas (1/3)

01

Sistema de inscripciones

P: El sistema de gestión de inscripciones de la universidad registra cada vez que un estudiante se inscribe en un curso. ¿Qué tipo de sistema es este? R: OLTP.

02

Análisis de calificaciones

P: El departamento de análisis académico quiere analizar las calificaciones de los últimos diez años para identificar tendencias. ¿Qué tipo de sistema es más adecuado? R: OLAP.

Respuestas (2/3)

03

Análisis complejo

P: El departamento de TI organiza los datos de inscripciones, calificaciones y asistencia para análisis complejos. ¿Qué tipo de sistema usan? R: OLAP.

04

ID_Estudiante

P: Cada estudiante tiene un identificador único llamado ID_Estudiante. ¿Qué es en este contexto? R: Una llave primaria.

05

ID_Curso en Inscripciones

P: La tabla Inscripciones contiene el campo ID_Curso, que se refiere al ID_Curso en la tabla Cursos. ¿Qué es ID_Curso en la tabla Inscripciones? R: Una llave foránea.

Respuestas (3/3)

06

Relación 1:M

P: En una relación uno a muchos (1:M) entre Profesores y Cursos, ¿cuántos cursos puede enseñar un profesor? R: Varios.

07

Relación M:M

P: Cada estudiante puede inscribirse en varios cursos, y cada curso puede tener varios estudiantes. ¿Qué tipo de relación es esta? R: Muchos a muchos (M:M).

Evaluación de la Sesión



<https://docs.google.com/forms/d/e/1FAIpQLSefDWVWtRfqMuFD1mJMU3xsePBLH3zG-sga1nLQpqBigGBwyA/viewform?usp=dialog>

Fin Sesión 5



La *transformación digital*
es parte de nuestro

Plan
Estratégico
Institucional

23
27



Conoce más ↗