

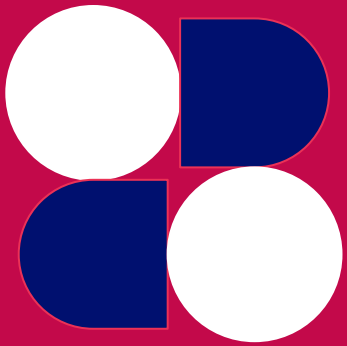


Programa PUCP DATA DRIVEN

Formación tecnológica

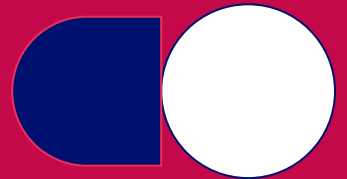
Dirección Académica de Planeamiento y Evaluación
Dirección de Gestión del Talento Humano





SESIÓN 12:

Modelos de datos tipo estrella y copo de nieve





Al finalizar la sesión, lograrás:

- Comprender la estructura y componentes de un modelo de datos tipo estrella
- Comprender la estructura y componentes de un modelo de datos tipo copo de nieve
- Entender la diferencia entre los dos tipos de modelos y sus casos de aplicación



Contenido:



1

Introducción

2

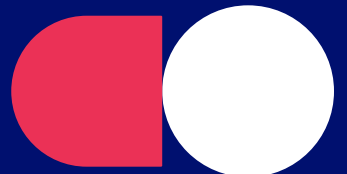
Modelo de datos tipo estrella

3

Modelo de datos tipo copo de nieve



1. Introducción





1. ¿Qué es un modelo de datos?

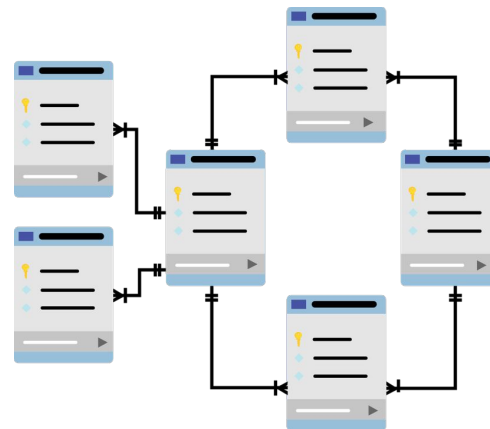


Definición:

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

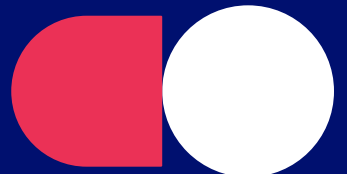
Importancia:

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





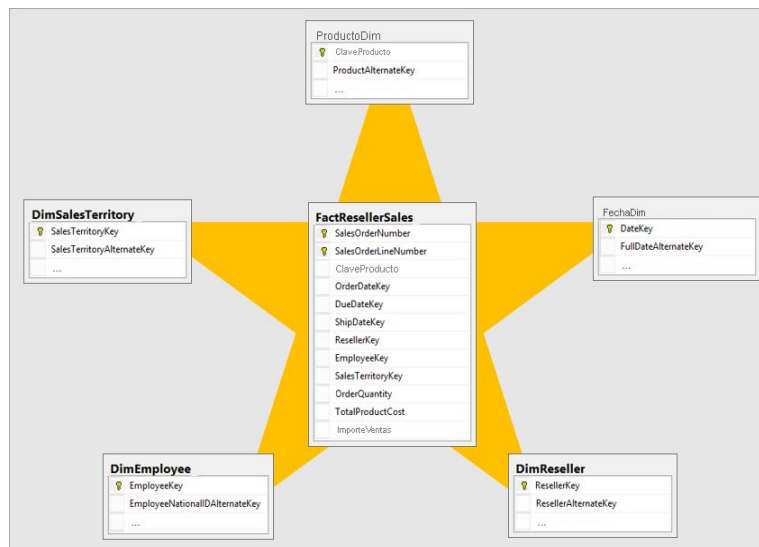
2. Modelo de datos tipo estrella





2. ¿Qué es un Modelo Estrella?

Un modelo estrella es un esquema de base de datos utilizado en Data Warehousing que organiza los datos en una estructura central (tabla de hechos) rodeada por varias tablas auxiliares (tablas de dimensiones).

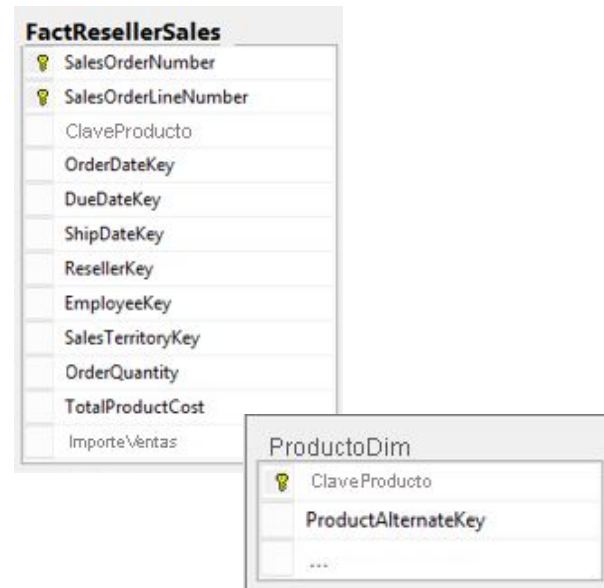




2. Componentes del Modelo Estrella



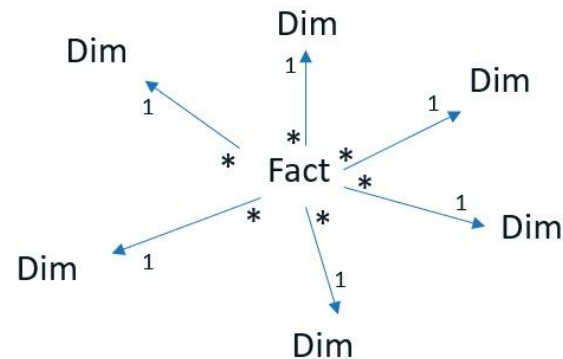
- **Tabla de Hechos:**
Contiene métricas numéricas que se analizan (como ventas, ingresos, cantidad vendida, etc.). Suele estar relacionada con tablas de dimensiones a través de claves foráneas.
- **Tablas de Dimensiones:**
Contienen atributos descriptivos que proporcionan contexto a los datos de la tabla de hechos (como producto, cliente, tiempo, etc.).





2. Tabla de Hechos

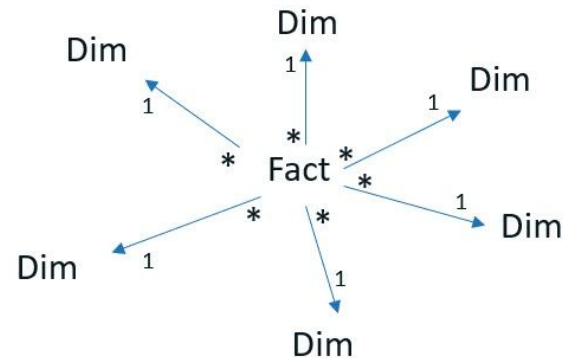
- La tabla de hechos es el corazón del modelo estrella.
- Almacena datos numéricos o métricas de negocio que se analizan (ventas, ingresos, cantidades, etc.).
- Suele estar conectada con tablas de dimensiones a través de claves foráneas para proporcionar contexto a las métricas almacenadas.





2. Tabla de Dimensiones

- Las tablas de dimensiones son tablas auxiliares que contienen atributos descriptivos.
- Proporcionan contexto a los datos numéricos almacenados en la tabla de hechos (por ejemplo, información sobre productos, clientes, tiempo, ubicaciones, etc.).
- Estas tablas suelen tener relaciones uno a muchos con la tabla de hechos.





2. Normalización Vs. Desnormalización



Normalización es el término que se usa para describir los datos almacenados de una manera que reduce los datos repetitivos.

Una tabla de ventas se considera normalizada cuando almacena solo claves, como la clave de producto. En la imagen siguiente, observe que sólo la columna ProductKey registra el producto.

	SalesOrderNumber	OrderDate	ProductKey	ResellerKey	SalesAmount
1	SO69561	2020-05-31	594	546	226.00
2	SO69560	2020-05-30	513	100	218.45
3	SO69560	2020-05-30	594	100	113.00
4	SO69539	2020-05-28	243	529	858.90
5	SO69539	2020-05-28	378	529	1466.01
6	SO69541	2020-05-28	594	661	113.00
7	SO69542	2020-05-28	243	317	1717.80
8	SO69544	2020-05-28	243	666	3435.60
9	SO69545	2020-05-28	378	436	5864.04
10	SO69532	2020-05-27	594	312	113.00
11	SO69532	2020-05-27	513	312	436.90
12	SO69533	2020-05-27	594	476	226.00



2. Normalización Vs. Desnormalización



Pero si la tabla de ventas almacena detalles de los productos más allá de la clave, se considera desnormalizada. En la imagen siguiente, observe que ProductKey y otras columnas relacionadas con los productos registran el producto.

	SalesOrderNumber	OrderDate	ProductKey	Product	Category	Color	Size	ResellerKey	SalesAmount
1	SO69561	2020-05-31	594	Mountain-500 Silver, 48	Bikes	Silver	48	546	226.00
2	SO69560	2020-05-30	513	ML Mountain Frame-W - Silver, 46	Components	Silver	46	100	218.45
3	SO69560	2020-05-30	594	Mountain-500 Silver, 48	Bikes	Silver	48	100	113.00
4	SO69539	2020-05-28	243	HL Road Frame - Red, 44	Components	Red	44	529	858.90
5	SO69539	2020-05-28	378	Road-250 Black, 52	Bikes	Black	52	529	1466.01
6	SO69541	2020-05-28	594	Mountain-500 Silver, 48	Bikes	Silver	48	661	113.00
7	SO69542	2020-05-28	243	HL Road Frame - Red, 44	Components	Red	44	317	1717.80
8	SO69544	2020-05-28	243	HL Road Frame - Red, 44	Components	Red	44	666	3435.60
9	SO69545	2020-05-28	378	Road-250 Black, 52	Bikes	Black	52	436	5864.04
10	SO69532	2020-05-27	594	Mountain-500 Silver, 48	Bikes	Silver	48	312	113.00
11	SO69532	2020-05-27	513	ML Mountain Frame-W - Silver, 46	Components	Silver	46	312	436.90
12	SO69533	2020-05-27	594	Mountain-500 Silver, 48	Bikes	Silver	48	476	226.00

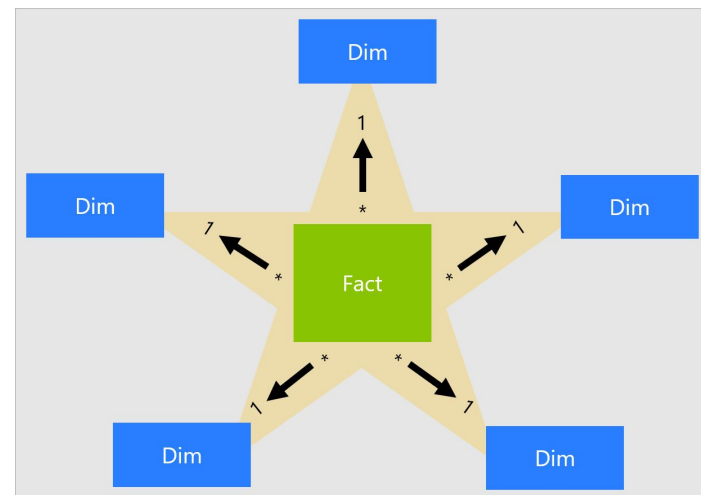


2. Relevancia del Modelo Estrella en Power BI



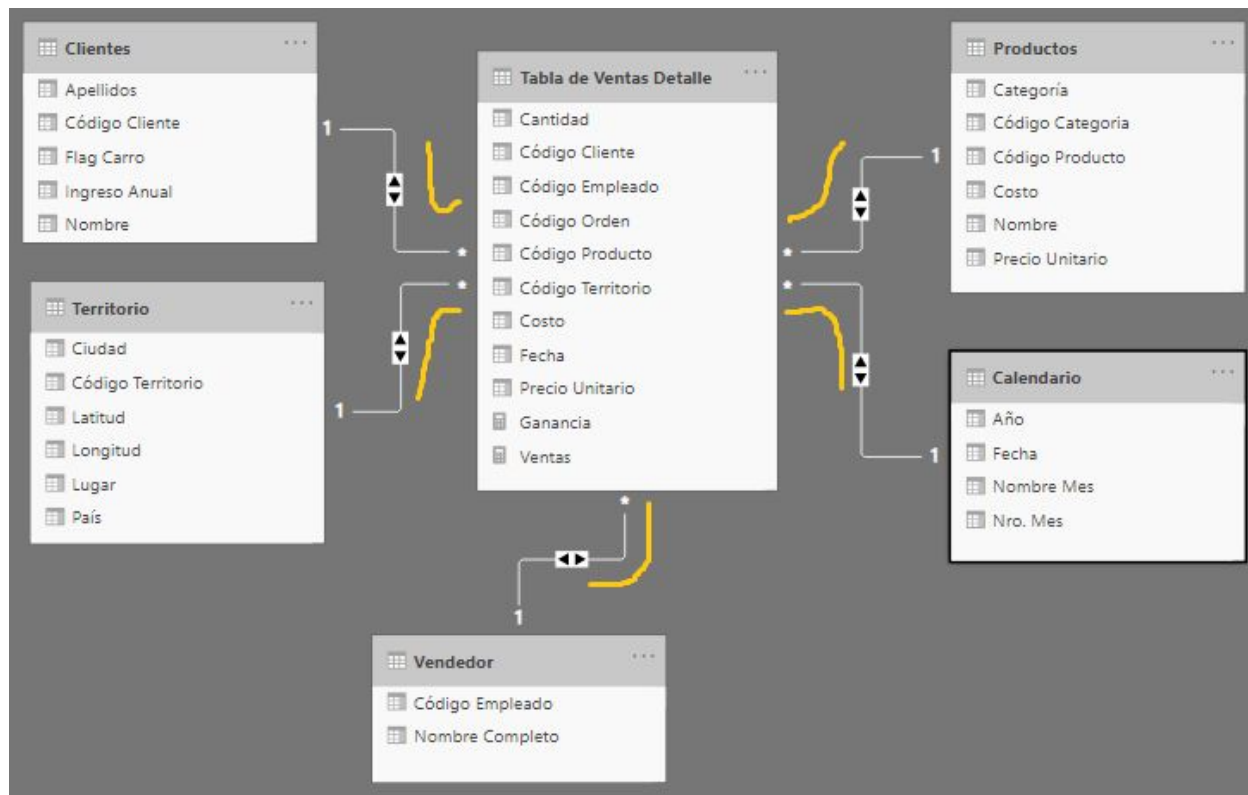
Tenga en cuenta que cada objeto visual del informe de Power BI genera una consulta que se envía al modelo de Power BI. Estas consultas se usan para filtrar, agrupar y resumir los datos del modelo. Por tanto, un modelo bien diseñado es aquel que proporciona tablas para filtrar y agrupar y tablas para resumir. Este diseño se ajusta bien a los principios de los esquemas de estrella:

- Las tablas de dimensiones admiten el filtrado y la agrupación
- Las tablas de hechos admiten el resumen





2. Ejemplo de Modelo Estrella - Power BI





2. Modelo Estrella - Ejercicio práctico



Contexto: Una universidad necesita diseñar una base de datos para gestionar la información sobre estudiantes, cursos, profesores, periodos académicos y matrículas. Los estudiantes deben diseñar las tablas necesarias, definir las llaves primarias y foráneas, y establecer las relaciones entre las tablas.

Instrucciones

1. Leer el enunciado del problema.
2. Diseñar las tablas necesarias.
3. Identifique tablas de hechos y dimensiones.
4. Definir las llaves primarias para cada tabla.
5. Definir las llaves foráneas para establecer relaciones entre las tablas.
6. Dibujar el modelo tipo estrella correspondiente al modelo



2. Modelo Estrella - Ejercicio práctico



Contexto: Una universidad necesita diseñar una base de datos para gestionar la información sobre estudiantes, cursos, profesores, periodos académicos y matrículas. Los estudiantes deben diseñar las tablas necesarias, definir las llaves primarias y foráneas, y establecer las relaciones entre las tablas, así como diseñar el modelo correspondiente.

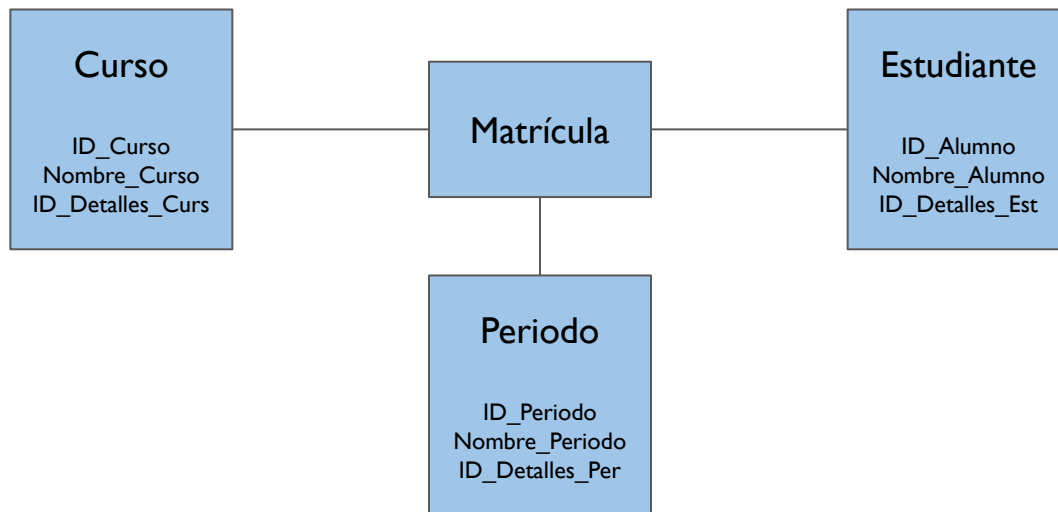
Instrucciones

1. Se dividirán en grupos de 4 personas en salas pequeñas, se darán 10 minutos para resolver el ejercicio propuesto
2. Leer el enunciado del problema.
3. Diseñar las tablas necesarias.
4. Identifique tablas de hechos y dimensiones.
5. Definir las llaves primarias para cada tabla.
6. Definir las llaves foráneas para establecer relaciones entre las tablas.
7. Dibujar el modelo tipo estrella correspondiente al modelo
8. Se socializarán las soluciones de 3 equipos elegidos al azar





2. Modelo Estrella - Ejemplo de solución





2. Modelo Estrella - Ejemplo de solución



Tabla de Hechos: Tabla de Matrículas:

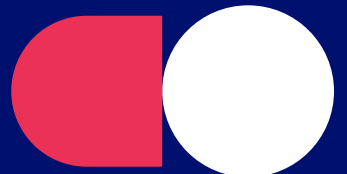
- Campos:
 - ID_Matricula (clave principal)
 - ID_Alumno (clave foránea relacionada con la tabla de dimensiones de Alumnos)
 - ID_Curso (clave foránea relacionada con la tabla de dimensiones de Cursos)
 - Fecha_Matricula
 - Monto_Pagado
 - Estado_Matricula (por ejemplo, "Activa", "Cancelada", etc.)
 - ID_Periodo (clave foránea relacionada con la tabla de dimensiones de Periodos Académicos)

Tablas de Dimensiones:

- Tabla de Alumnos:
 - Campos:
 - ID_Alumno (clave principal)
 - Nombre_Alumno
 - Edad
 - Género
 - Otros atributos descriptivos del alumno.
- Tabla de Cursos:
 - Campos:
 - ID_Curso (clave principal)
 - Nombre_Curso
 - Departamento
 - Créditos
 - Otros atributos descriptivos del curso.
- Tabla de Periodos Académicos:
 - Campos:
 - ID_Periodo (clave principal)
 - Fecha_Inicio
 - Fecha_Fin
 - Otros atributos descriptivos del periodo académico.



3. Modelo tipo Copo de Nieve

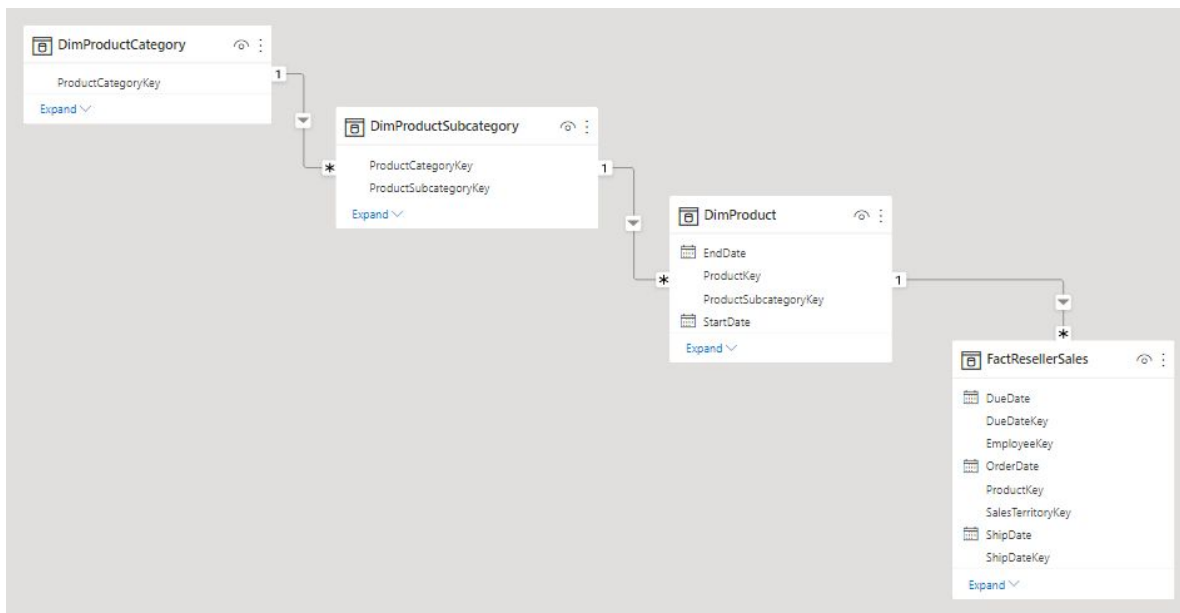




3. ¿Qué es un Modelo tipo Copo de Nieve?

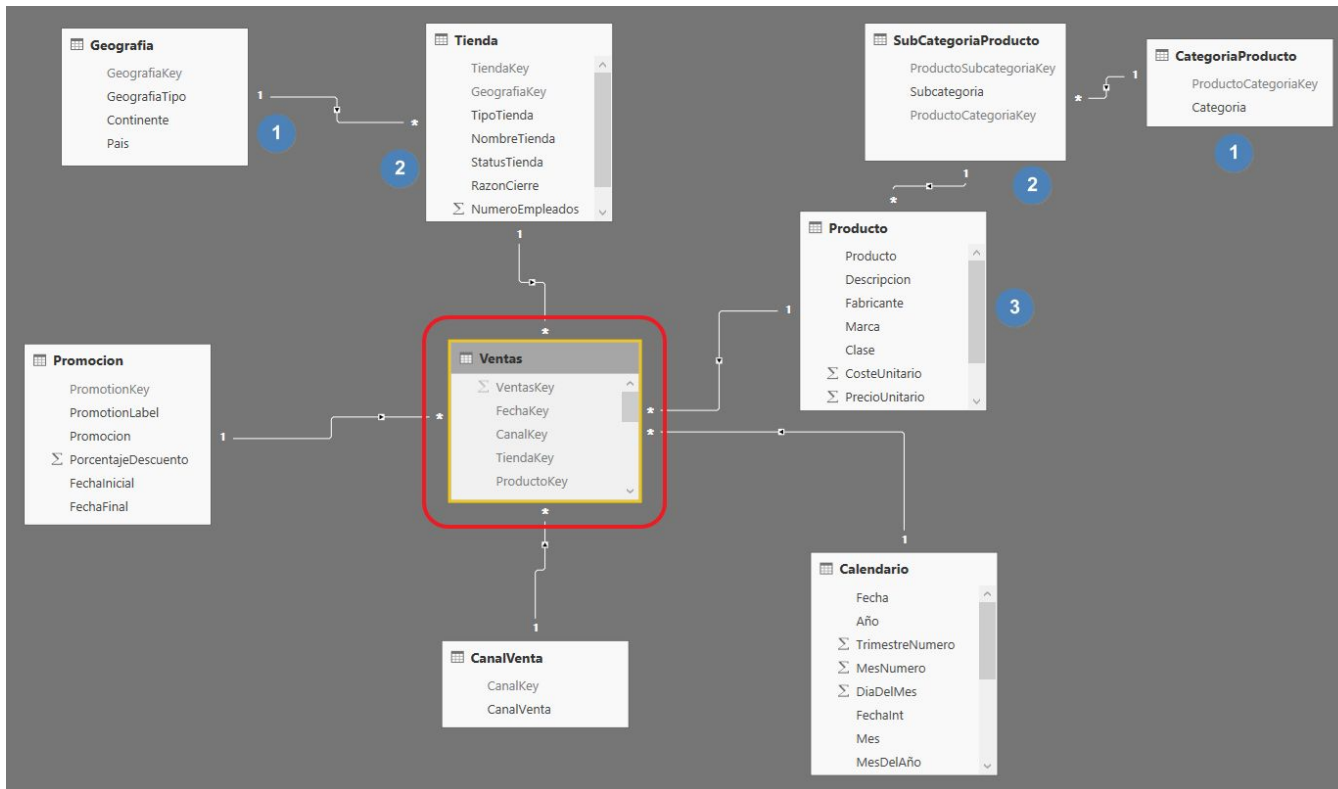


Un modelo copo de nieve es una variante del modelo estrella donde las tablas de dimensión se normalizan en múltiples tablas relacionadas.





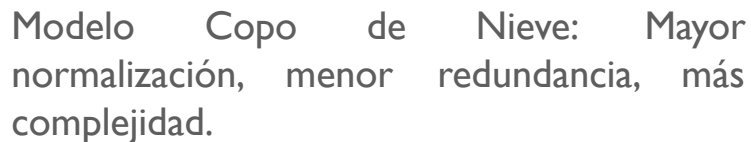
3. Componentes de modelo tipo Copo de Nieve



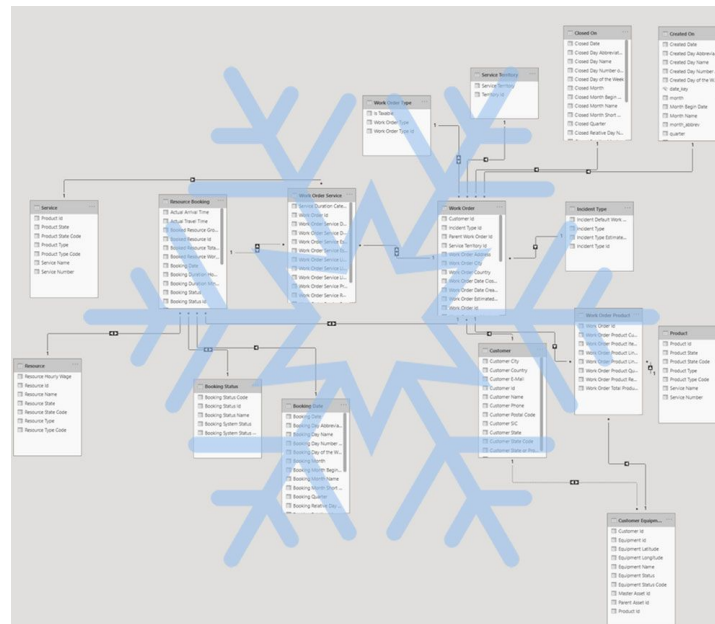


3. Comparación con el modelo estrella





Modelo Estrella: Menor normalización, más redundancia, menos complejidad.





3. Características del modelo Copo de Nieve



Características:

- Menor redundancia de datos.
- Estructura más compleja y más joins en las consultas.

Ventajas: Menor redundancia, mejor integridad de datos.

Desventajas: Consultas más complejas y lentas debido a múltiples joins.

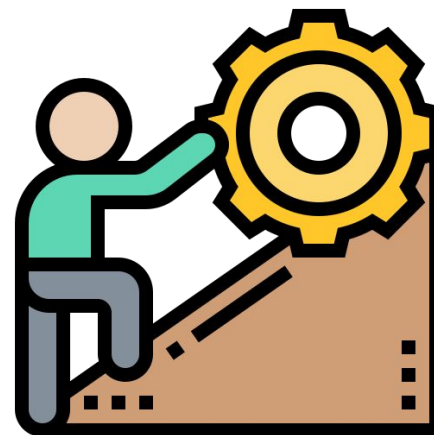




3. Desafíos del modelo Copo de Nieve en Power BI



- **Carga de Más Tablas:** Power BI carga más tablas, lo que resulta menos eficaz en almacenamiento y rendimiento.
- **Columnas Adicionales:** Las tablas deben incluir columnas para admitir relaciones, aumentando el tamaño del modelo.
- **Propagación de Filtros:** Las cadenas de propagación de filtros son más largas y menos eficaces.
- **Experiencia del Usuario:** El panel Campos muestra más tablas, haciendo la experiencia menos intuitiva.
- **Creación de Jerarquías:** No es posible crear jerarquías que abarquen múltiples tablas.

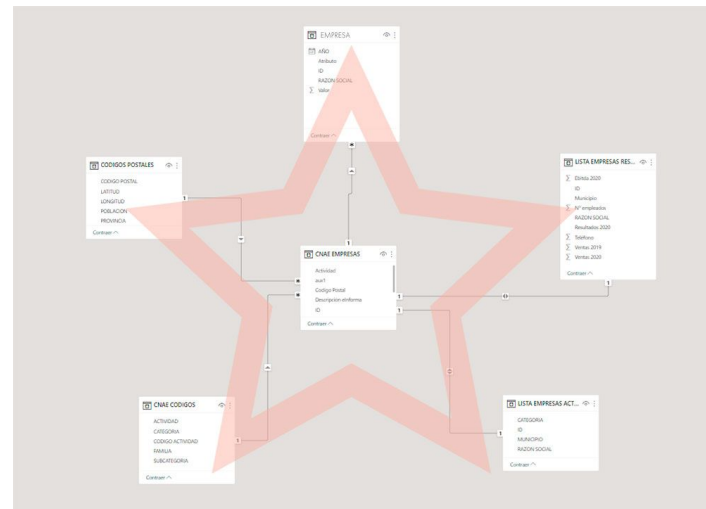




3. Recomendaciones y mejores prácticas



- Simplificación del Modelo: Considerar el uso del modelo estrella.
- Optimización de Relaciones: Reducir la cantidad de tablas y relaciones necesarias.
- Uso de Jerarquías Internas: Crear jerarquías dentro de una sola tabla





3. Modelo copo de nieve - Ejercicio práctico



Contexto: Una universidad necesita diseñar una base de datos para gestionar la información sobre estudiantes, cursos, profesores, periodos académicos y matrículas. Los estudiantes deben diseñar las tablas necesarias, definir las llaves primarias y foráneas, y establecer las relaciones entre las tablas, así como diseñar el modelo correspondiente.

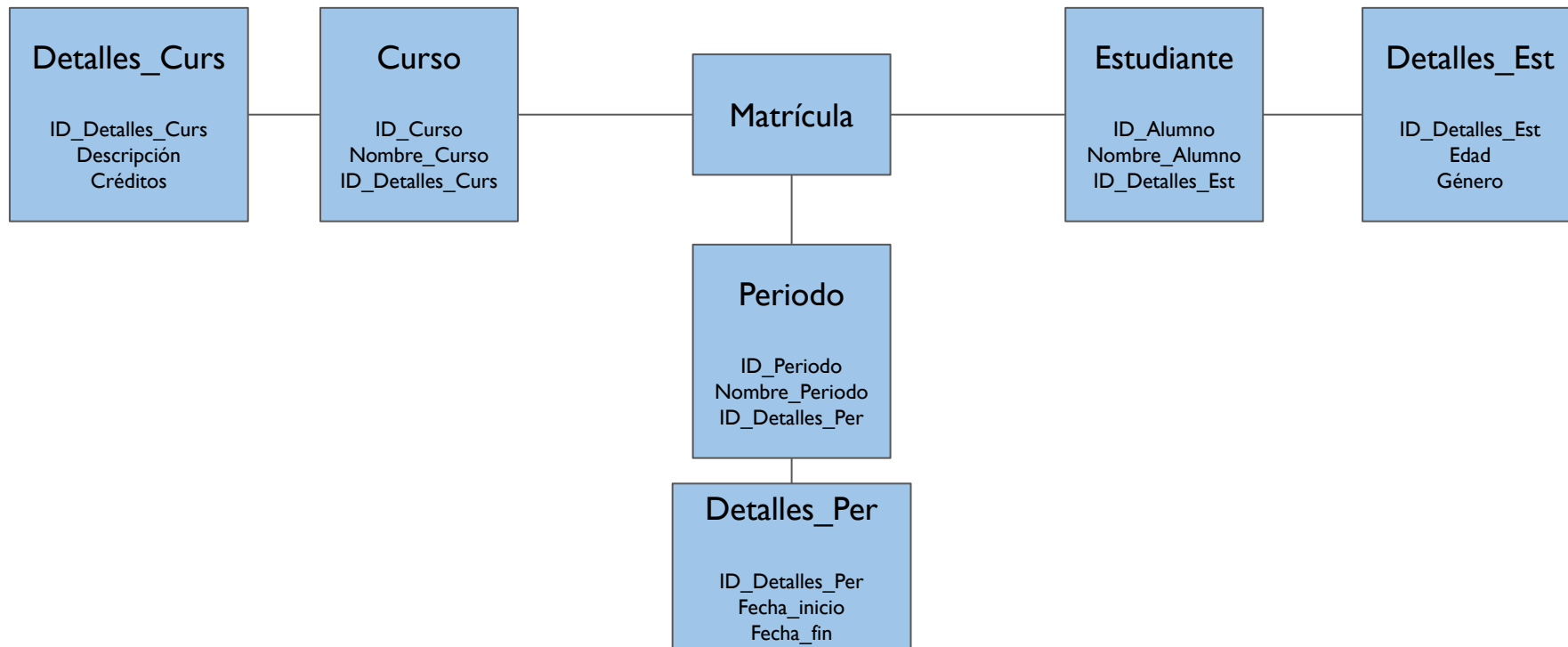
Instrucciones

1. Se dividirán en grupos de 4 personas en salas pequeñas, se darán 10 minutos para resolver el ejercicio propuesto
2. Leer el enunciado del problema.
3. Transformar el modelo tipo estrella obtenido del ejercicio anterior en un modelo copo de nieve, realizando el proceso de normalización de las tablas de dimensiones
4. Se socializarán las soluciones propuestas por 3 grupos elegidos al azar



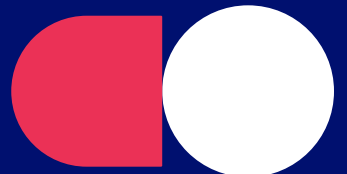


3. Modelo copo de nieve - Ejemplo de solución





4. Actividad grupal



- slideswith.com/PREPASOS12





Respuesta



- ¿Qué es un modelo estrella en bases de datos?

R// Un esquema de base de datos donde una tabla central (hechos) se conecta directamente con varias tablas de dimensiones

- En el modelo estrella, ¿qué es una tabla de hechos?

R// Una tabla central que contiene los datos cuantitativos (mediciones)

- En el modelo estrella, ¿qué es una tabla de dimensiones?

R// Una tabla que describe los contextos de las medidas en la tabla de hechos

- ¿Qué es un modelo copo de nieve en bases de datos?

R// Un esquema de base de datos donde las tablas de dimensiones están normalizadas en múltiples tablas relacionadas



Respuesta



- ¿Cuál es la principal diferencia entre un modelo estrella y un modelo copo de nieve?

R// El modelo estrella tiene tablas de dimensiones desnormalizadas, mientras que el copo de nieve tiene tablas de dimensiones normalizadas

- ¿Cuál es una ventaja del modelo estrella?

R// Consultas más rápidas y simples debido a su estructura desnormalizada

- ¿Cuál es una ventaja del modelo copo de nieve?

R// Menor redundancia de datos y mejor integridad



Respuesta



- En el contexto de una universidad, ¿qué podría ser una tabla de hechos en un modelo estrella?

R// Una tabla de calificaciones de los estudiantes

- En el contexto de una universidad, ¿qué podría ser una tabla de dimensiones en un modelo estrella?

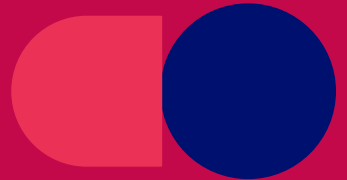
R// Una tabla con información de los estudiantes, como nombres y direcciones

- ¿Cuál es una posible desventaja del modelo copo de nieve?

R// Consultas más complejas debido a la normalización



Gracias





PUCP