

El modelo Copo de Nieve



Universidad
Tecnológica
del Perú

Contenido:

1. Introducción
2. Logro
3. Transformación
4. Cierre

1. Introducción:

- Los modelos de Data Mart

2. Logro de la sesión

Al finalizar la sesión, el estudiante podrá entender los conceptos básicos del modelo copo de nieve para un Data mart

3. Transformación

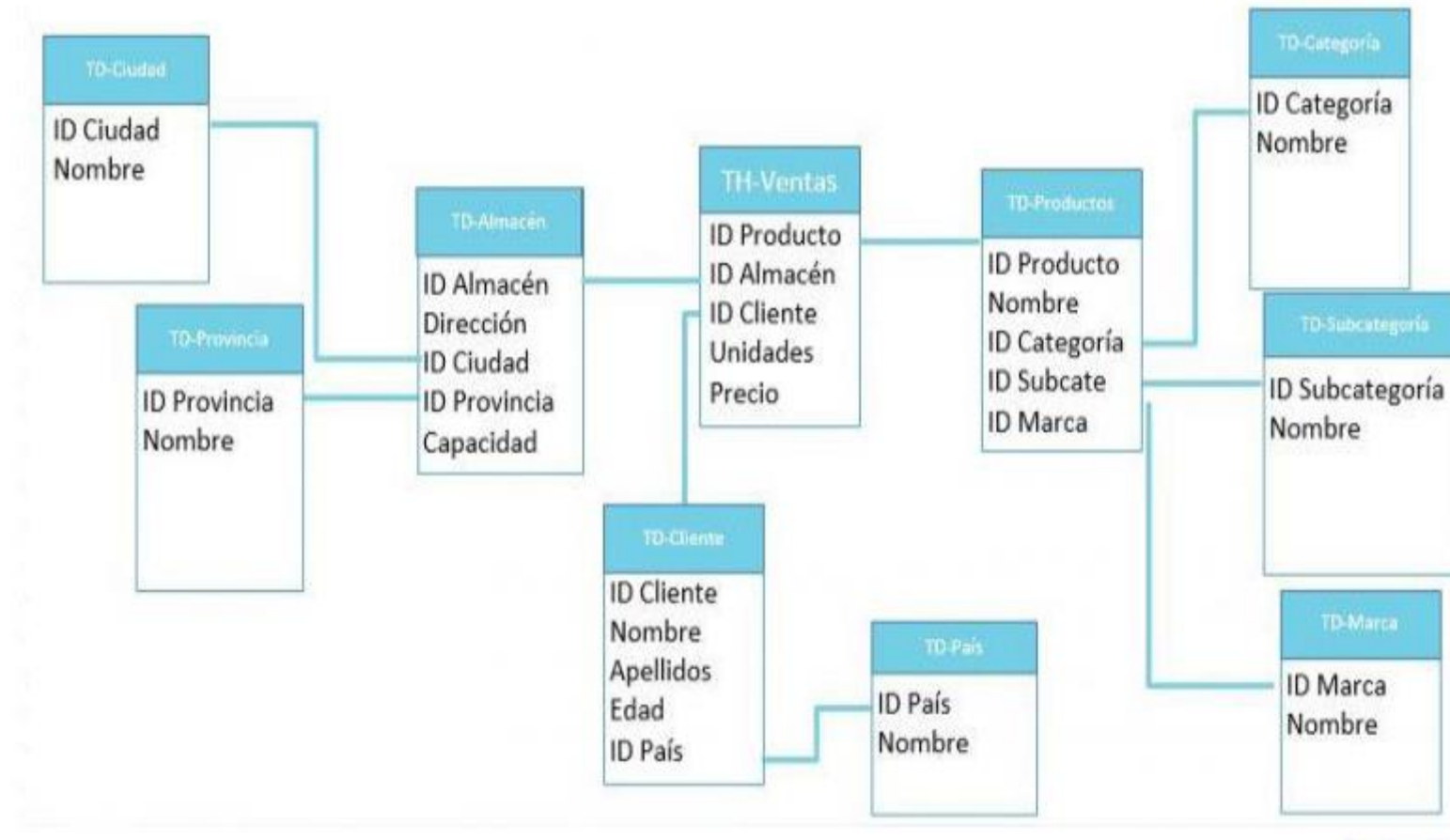
Esquema de copo de nieve

Un esquema de copo de nieve es una estructura más compleja que el esquema de estrella. Se da cuando alguna de las dimensiones se implementa con más de una tabla de datos.

El objetivo es normalizar estas tablas y reducir el espacio de almacenamiento al eliminar la redundancia. Se representa como una tabla de hechos conectada con dimensiones anidadas. Al normalizar por completo las dimensiones, el resultado parece un copo de nieve..

3. Transformación

Esquema de copo de nieve



Observamos en la imagen 2 como se dividen las dimensiones de TD-Almacén, TD-Producto y TD-Cliente en sub-dimensiones normalizadas.

3. Transformación

Esquema de copo de nieve

Las principales **ventajas** del esquema de copo de nieve son:

- Algunas herramientas de modelado de bases de datos multidimensional OLAP se optimizan
- La normalización de los atributos reduce el almacenamiento de datos

Las principales **desventajas** del esquema de copo de nieve son:

- Queries complejas debido a la normalización (implica un mayor número de cruces)
- Bajo rendimiento debido a la normalización

3. Transformación

ITEM	ESTRELLA	COPO DE NIEVE
Mantenimiento	Tiene redundancia, peor mantenimiento	No hay redundancia, fácil mantenimiento
Facilidad de uso	Queries menos complejas, fácil uso	Queries complejas, difícil de entender
Rendimiento de las queries	Ejecuciones más rápidas	Más tiempo de ejecución debido a los cruces
Tipo de DWH	Data Mart	Data Warehouse
Joins	Bajo número de Joins	Alto número de Joins
Tablas de dimensión	Una tabla por cada dimensión	Más de una tabla por cada dimensión
Cuándo usarlo	Cuando las tablas de dimensión tienen pocas filas	Cuando las tablas de dimensión tienen un tamaño bastante elevado
Normalización / Denormalización	Tablas de dimensiones y de hechos denormalizadas	Tablas de dimensión normalizadas. Tablas de hechos denormalizadas
Modelo de datos	Top Down	Bottom Up

3. Transformación

Esquema en copo de nieve con Tableau

La primera tabla ("*Ventas*") es la tabla de hechos, conteniendo información sobre ventas:

Ventas

	A	B
1	Id. Producto	Importe
2	1	10
3	2	15
4	1	10

La segunda tabla ("*Productos*") -relacionada con la anterior a través del campo Id. Producto- contiene información sobre los productos:

Productos

	A	B	C
1	Id. Producto	Nombre	Id. Categoría
2	1	Ordenador	1
3	2	Windows 10	2

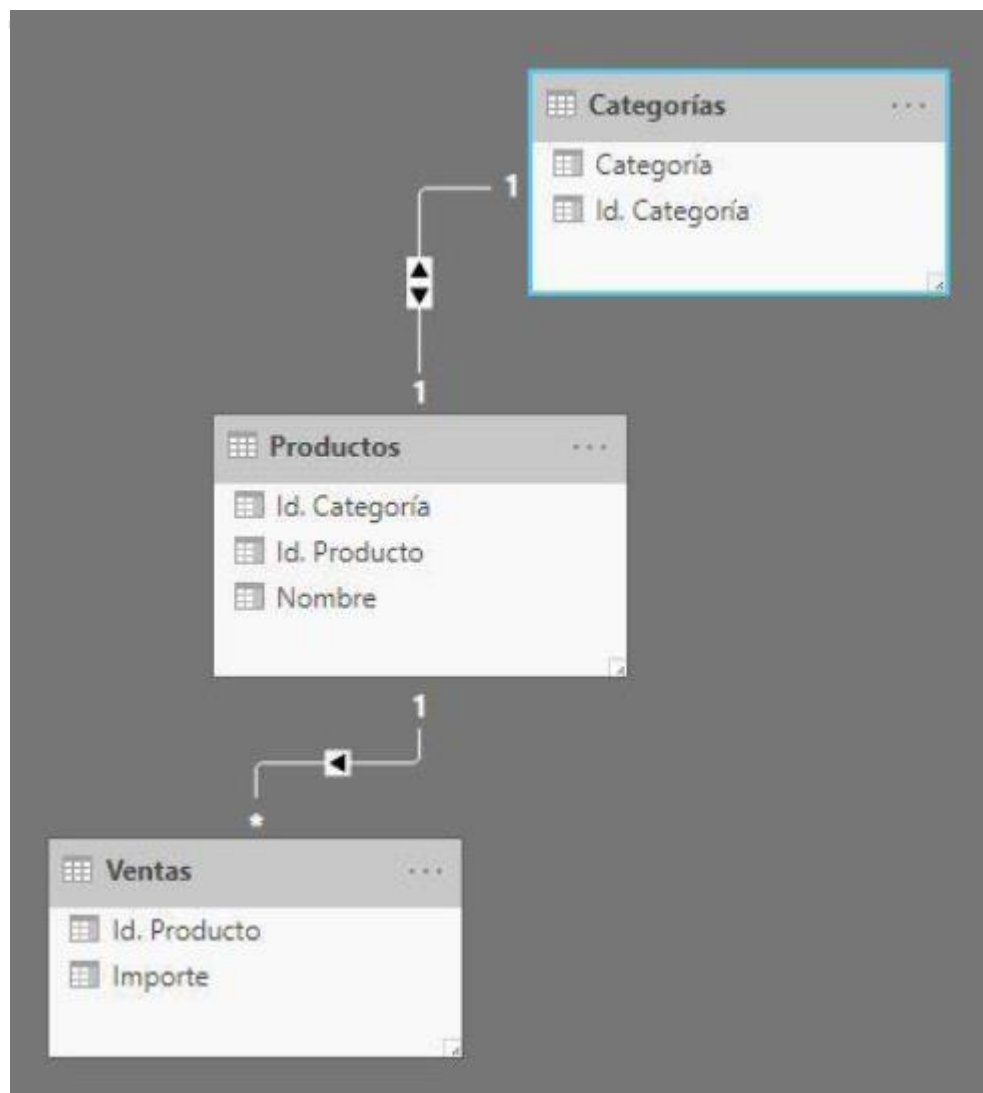
Y, por último, la tercera tabla ("*Categorías*") -relacionada a su vez con la de productos a través del campo Id. Categoría- contiene información sobre la categoría a la que pertenece cada producto:

Categorías

	A	B
1	Id. Categoría	Categoría
2	1	Hardware
3	2	Software

3. Transformación

El esquema de este
modelo de datos es
el siguiente:



de lo que te limita

3. Transformación

Si cargamos este archivo Excel en Tableau y llevamos la tabla de hechos ("Ventas") al área ya conocida con el texto "*Arrastrar hojas aquí*", la interfaz de Tableau nos muestra lo siguiente:

The screenshot shows the Tableau interface. On the left, the 'Conexiones' pane lists 'Ventas-Productos-Categorías' as a Microsoft Excel connection. Below it, the 'Hojas' pane lists 'Categorías', 'Productos', 'Ventas', and 'Nueva unión'. The 'Ventas' sheet is selected. On the right, the 'Ventas' table is displayed with columns 'Id. Producto' and 'Importe'. The table contains three rows of data.

Ventas Id. Producto	Ventas Importe
1	10
2	15
1	10

Desaprende lo que te limita

3. Transformación

Si ahora arrastramos la tabla de Productos, nos muestra lo siguiente:

The screenshot shows the Microsoft Access interface. On the left, the 'Conexiones' pane shows 'Ventas-Productos-Categorías' as a Microsoft Excel connection. Below it, the 'Hojas' pane lists 'Categorías', 'Productos', 'Ventas', and 'Nueva unión'. The main area shows a query design view for a query named 'Ventas'. The design grid shows a join between the 'Ventas' table and the 'Productos' table. The resulting query results are displayed below the design grid.

#	Abc	#	#	#
Productos	Productos	Productos	Ventas	Ventas
Id. Producto (Prod...	Nombre	Id. Categoría	Id. Producto	Importe
1	Ordenador	1	1	10
2	Windows 10	2	2	15
1	Ordenador	1	1	10

Desaprende lo que te limita

3. Transformación

Por último, si arrastramos la tabla de *Categorías*, tenemos el siguiente resultado:

The screenshot shows the Tableau interface. On the left, the 'Conexiones' (Connections) pane lists 'Ventas-Productos-Categorías' as a Microsoft Excel file. Below it, the 'Hojas' (Sheets) pane shows 'Categorías', 'Productos', 'Ventas', and 'Nueva unión' (New Union). The main workspace displays a diagram of three tables: 'Ventas', 'Productos', and 'Categorías', connected by double-headed arrows. Below the diagram, a data table is shown with columns for 'Categorías', 'Productos', and 'Ventas'. The table contains three rows of data.

#	Id. Categoría (Cat...)	Nombre	Id. Producto (Prod...)	Nombre	#	Id. Categoría	Id. Producto	Importe
1	Hardware		1	Ordenador	1	1	1	10
2	Software		2	Windows 10	2	2	2	15
1	Hardware		1	Ordenador	1	1	1	10

Es decir, aun cuando la fuente de datos (la tabla desnormalizada) no lo refleje de un modo explícito, Tableau ha combinado de la forma correcta las tres tablas: *Productos* y *Categorías* por un lado, y el resultado de dicha combinación con *Ventas*.

4. Cierre

Test:

- Cómo es el modelo copo de nieve?
- Ventajas y desventajas del modelo copo de nieve



**Universidad
Tecnológica
del Perú**