

DATA WAREHOUSE KIMBALL

Eduardo Aranda Arteaga,

0238814

Jorge Barba Pérez,

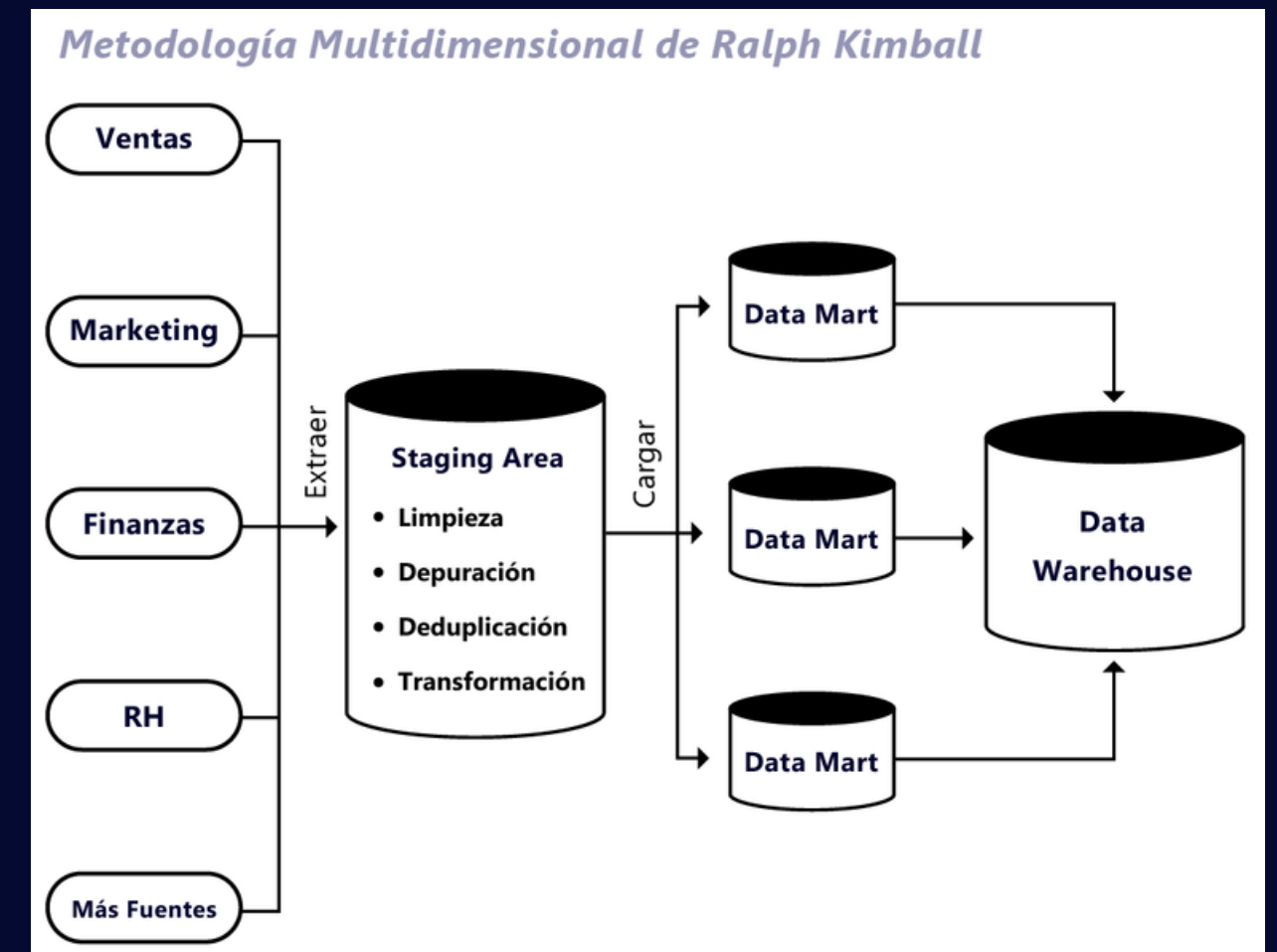
0237328

Laudiel Vinalay Ataxca

0245560

MODELADO DIMENSIONAL

Según Kimball, un Data Warehouse es la copia de los datos transaccionales específicamente estructurados para consultas analíticas e informes, con el fin de apoyar la toma de decisiones. Con esta metodología, al crear primero los Data Marts se proporcionan capacidades analíticas de informes para procesos específicos de negocio y funcionales.



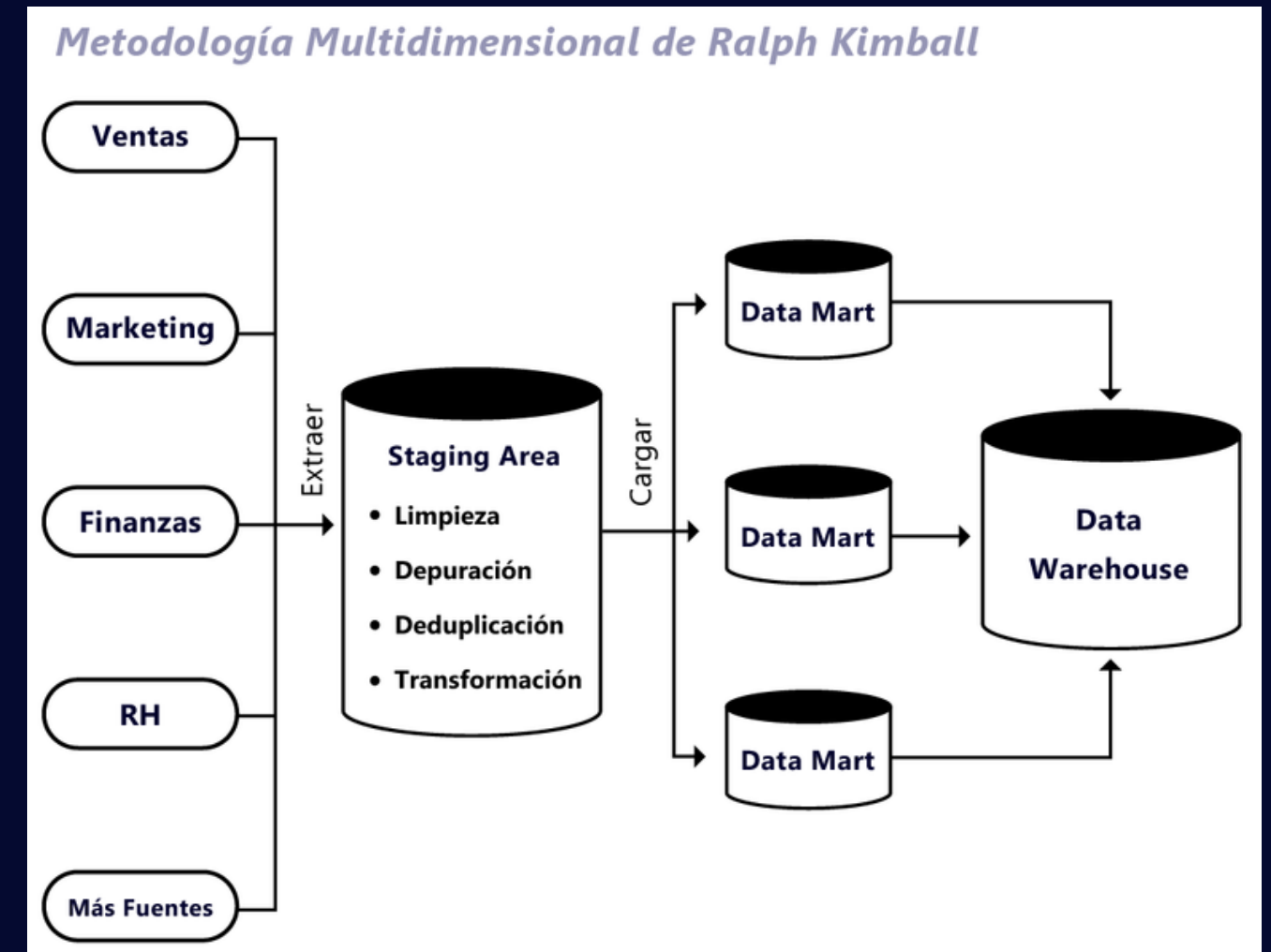


KIMBALL

La metodología multidimensional o metodología de Ralph Kimball, sigue un enfoque de abajo hacia arriba para almacenamiento de datos diseño de arquitectura en el que los data marts se forman primero en función de los requisitos comerciales.

Se evalúan las fuentes de datos primarias y se usa una herramienta de extracción, transformación y carga (ETL) para obtener datos de varias fuentes y cargarlos en un área de preparación del servidor de base de datos relacional.

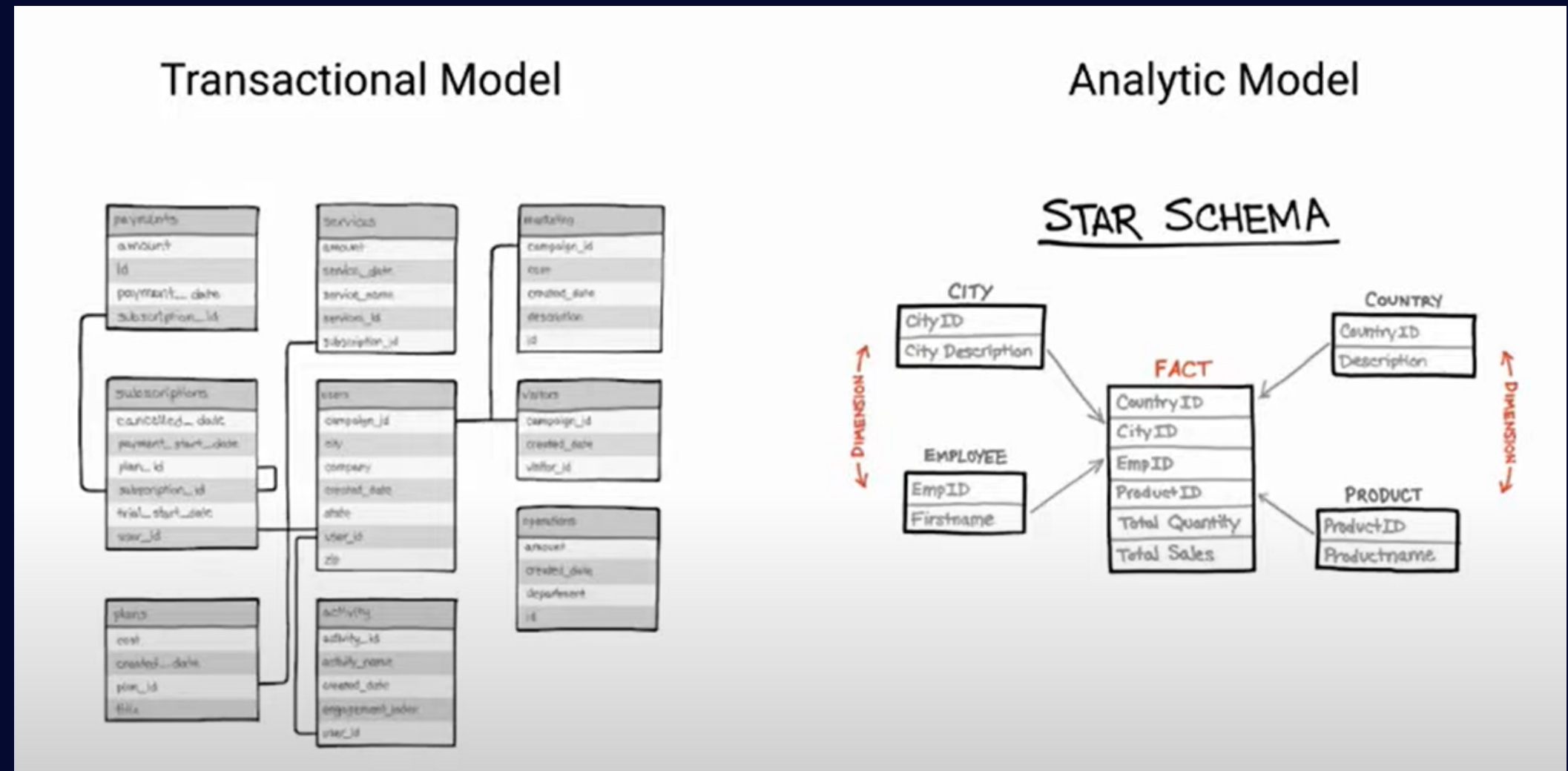
Una vez que los datos se cargan en el área de ensayo del almacén de datos, la siguiente fase incluye la carga de datos en un modelo de almacén de datos dimensional que no está normalizado por naturaleza. Este modelo divide los datos en la tabla de hechos, que son datos transaccionales numéricos o tabla de dimensiones, que es la información de referencia que respalda los hechos.



Costo, Rendimiento, Entendimiento

El esquema en estrella es el elemento fundamental del modelo de almacén de datos dimensional.

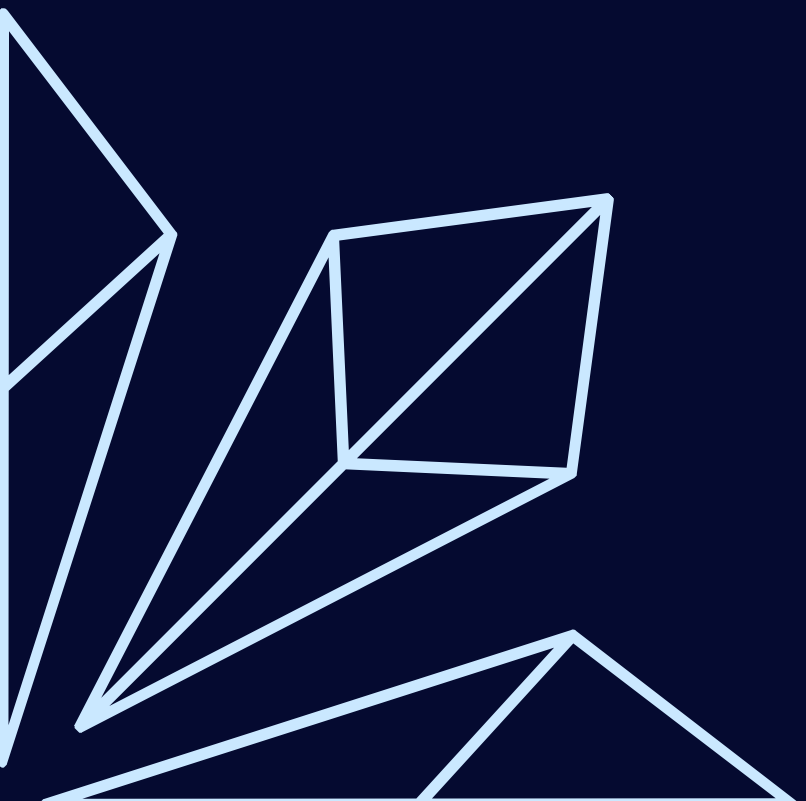
El modelado dimensional de Kimball permite a los usuarios construir varios esquemas en estrella para satisfacer diversas necesidades de generación de informes.



INTEGRACION DE DATOS

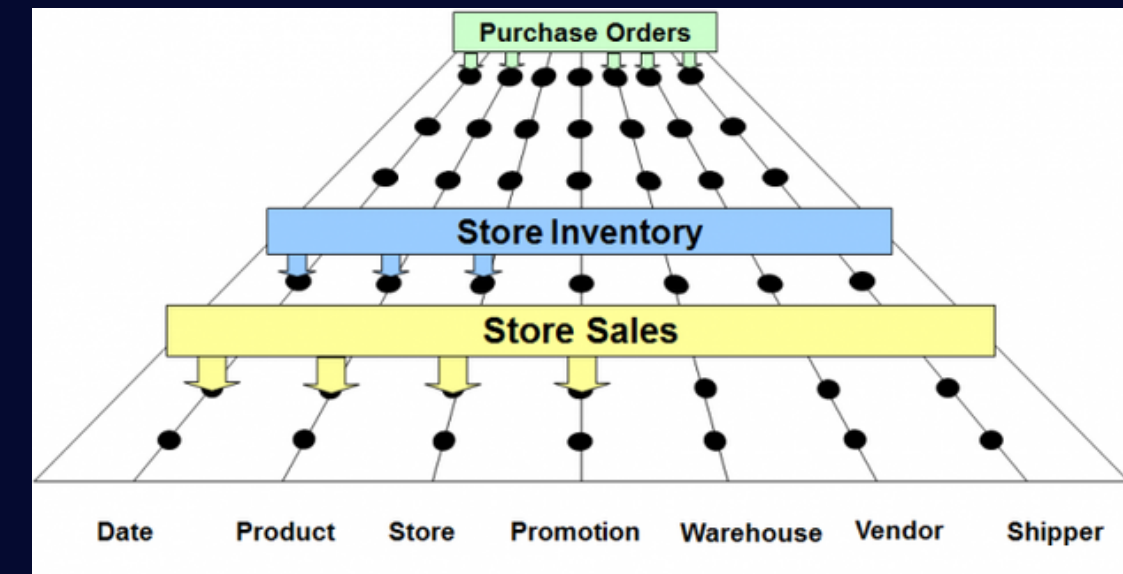


Para integrar datos, el enfoque de Kimball para el ciclo de vida de Data Warehouse sugiere la idea de dimensiones de datos conformadas. Existe como una tabla de dimensiones básica compartida entre diferentes tablas de hechos (como cliente y producto) dentro de un almacén de datos o como las mismas tablas de dimensiones en varios data marts de Kimball. Esto garantiza que un solo elemento de datos se utilice de manera similar en todos los hechos.



MATRIZ BUS KIMBALL

Una herramienta de diseño importante en la metodología de almacenamiento de datos de Ralph Kimball es la matriz de bus empresarial o la arquitectura de bus de Kimball que registra verticalmente los hechos y registra horizontalmente las dimensiones conformadas. La matriz de Kimball, que forma parte de la arquitectura del bus, muestra cómo se construyen los esquemas en estrella. Los equipos de gestión empresarial lo utilizan como entrada para priorizar qué fila de la matriz de Kimball se debe implementar primero.



VENTAJAS

- El modelado dimensional de Kimball es rápido de construir ya que no implica normalización, lo que significa una rápida ejecución de la fase inicial del almacenamiento de datos de procesos.
- Una ventaja del esquema en estrella es que la mayoría de los operadores de datos pueden comprenderlo fácilmente debido a su estructura desnormalizada, que simplifica las consultas y el análisis.
- La huella del sistema de almacenamiento de datos es trivial porque se centra en áreas y procesos comerciales individuales en lugar de en toda la empresa. Por lo tanto, ocupa menos espacio en la base de datos, lo que simplifica la administración del sistema.
- Permite la recuperación rápida de datos del almacén de datos, ya que los datos se segregan en tablas de hechos y dimensiones. Por ejemplo, la tabla de hechos y dimensiones para la industria de seguros incluiría transacciones de pólizas y transacciones de reclamos.
- Un equipo más pequeño de diseñadores y planificadores es suficiente para la gestión del almacén de datos porque los sistemas de origen de datos son estables y el almacén de datos está orientado a procesos. Además, la optimización de consultas es sencilla, predecible y controlable.
- Estructura dimensional conformada para marco de calidad de datos. El enfoque de Kimball para el ciclo de vida del almacén de datos también se conoce como el enfoque de estilo de vida dimensional empresarial porque permite que las herramientas de inteligencia empresarial profundicen en varios esquemas en estrella y genera información confiable.

DESVENTAJAS

- Los datos no están completamente integrados antes de la presentación de informes; la idea de una "fuente única de verdad se pierde".
- Pueden ocurrir irregularidades cuando los datos se actualizan en la arquitectura Kimball DW. Esto se debe a que, en la técnica de desnormalización, se agregan datos redundantes a las tablas de la base de datos.
- En la arquitectura Kimball DW, pueden ocurrir problemas de rendimiento debido a la adición de columnas en la tabla de hechos, ya que estas tablas son bastante detalladas. La adición de nuevas columnas puede expandir las dimensiones de la tabla de hechos, lo que afecta su rendimiento. Además, el modelo de almacén de datos dimensional se vuelve difícil de modificar con cualquier cambio en las necesidades comerciales.
- Como el modelo de Kimball está orientado a los procesos comerciales, en lugar de centrarse en la empresa en su conjunto, no puede manejar todos los requisitos de informes de BI.
- El proceso de incorporar grandes cantidades de datos heredados en el almacén de datos es complejo.

Preguntas

¿Qué esquema es fundamental para el modelo de almacén de datos dimensional?

¿Cómo es el enfoque de Kimball para el almacenamiento de datos?

¿Por qué es importante la técnica de modelado dimensional en el Kimball Data Warehouse?

Una ventaja y una desventaja del modelo kimball