

Caso IATA

Autor

Martín Vladimir Alonso Sierra Galvis

Gestión de Datos, Maestría en Ciencia de Datos
Pontificia Universidad Javeriana Cali

Versión 2.0
Santiago de Cali, marzo de 2023

Tabla de contenido

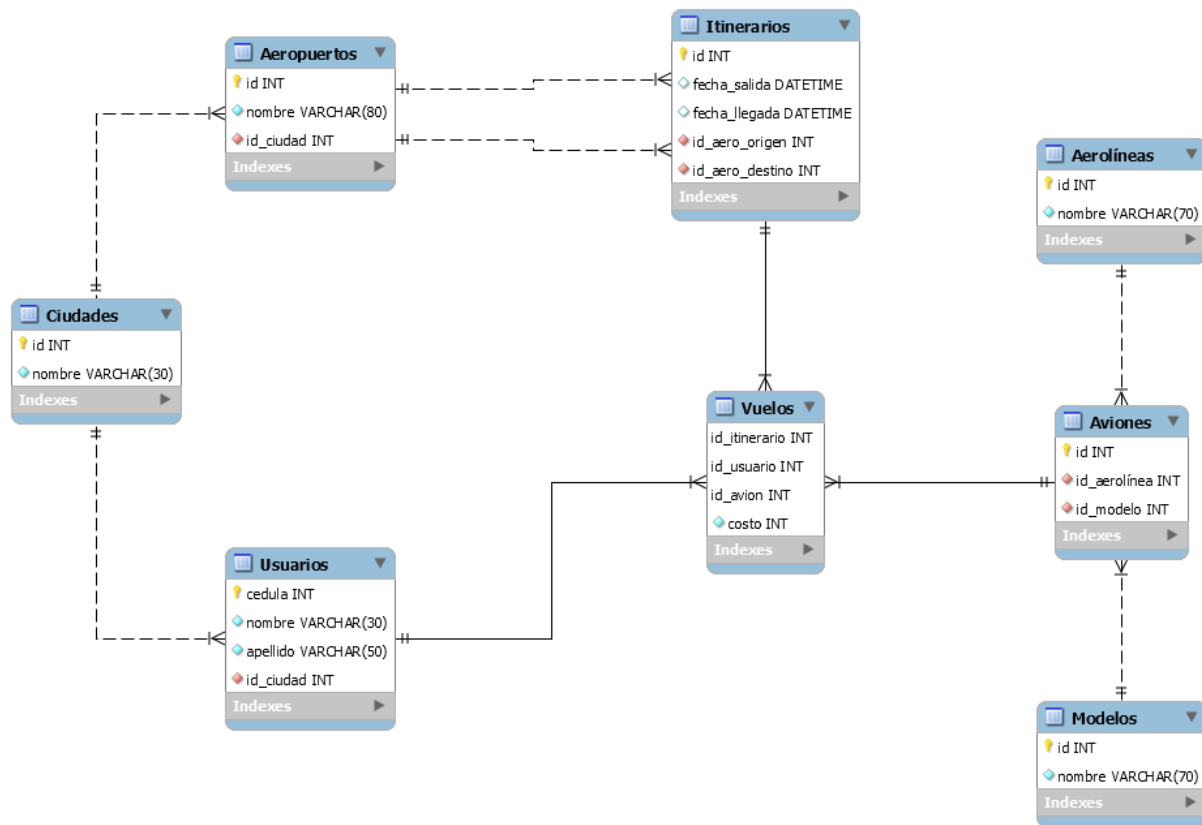
Contexto	3
Caso	4
Alcance	5
Recursos	6

Contexto

Uno de los métodos de transporte más utilizados por los seres humanos es el transporte aéreo. Aunque el número de vuelos diarios se vio afectado por la pandemia de la COVID-19, es bien sabido que son muchos los aviones que diariamente se movilizan de un sitio a otro llevando pasajeros. Es por ello que, para tener un control de estos movimientos, se han creado diversas organizaciones internacionales que llevan un registro de aerolíneas, aviones, itinerarios, equipajes, entre otras cosas. Una de ellas es la Asociación Internacional de Transporte Aéreo, **IATA** por sus siglas en inglés. Esta asociación surge en el año 1919 en los Países Bajos y estaba conformada por 57 miembros de 31 países pertenecientes, principalmente a Europa y Norteamérica. Años más tarde, en 1945, fue relanzada en Cuba. Actualmente, incluye 290 aerolíneas de 120 países. Sus misiones principales comprenden la promoción de la seguridad, la protección y la fiabilidad en el transporte aéreo y el cuidado del medio ambiente. Todo esto en pro del beneficio económico de sus accionistas. Adicional a esto la IATA representa a las 290 aerolíneas que la conforman, simplifica los procesos y reduce los costos que aumentan su flujo financiero, y actúa como un puente para asegurar el movimiento de las personas alrededor del mundo.

Caso

La Asociación Internacional de Transporte Aéreo, IATA, posee una base de datos donde tiene almacenada la información recolectada de todos los vuelos hechos a nivel global a partir del año 1945. El Modelo Relacional que la IATA utilizó para la construcción de su Base de Datos Relacional fue el siguiente:



En este Modelo Relacional se encuentran almacenados los siguientes datos:

1. Los usuarios y su respectiva ciudad de origen.
2. Los aeropuertos y la ciudad donde se encuentran ubicados.
3. Los aviones, sus modelos y la aerolínea a la que pertenece cada uno.
4. Los itinerarios, incluyendo fecha de salida, fecha de llegada, ciudad de origen del vuelo y ciudad de destino del vuelo.
5. Los vuelos, incluyendo itinerarios de cada vuelo, aviones implicados en cada vuelo, usuarios que tomaron determinado vuelo y costo de los diferentes vuelos registrados.

Con los datos almacenados en su base de datos relacional, la IATA quiere responder algunos requerimientos de análisis relacionados con los vuelos.

Alcance

La IATA quiere analizar en qué medida la pandemia de COVID-19 afectó el transporte aéreo de pasajeros durante el año 2020. Los **requerimientos de análisis** son los siguientes:

1. ¿Cuál aerolínea realizó el mayor número de vuelos a la ciudad de Roma en el año 2019 y cuál en el año 2020?
2. Total de dinero recaudado por vuelos de cada aerolínea en el primer semestre del año 2019 y en el primer semestre del año 2020.
3. ¿Cuál modelo de avión realizó el mayor número de vuelos en el año 2019 y cuál en el año 2020?
4. ¿Cuál fue la ciudad cuyos habitantes viajaron más en el año 2019 y cuál en el año 2020?

Para este fin, la IATA contrata a un grupo de científicos de datos, expertos en la gestión de datos en bases de datos analíticas, para que implemente, a partir de su base de datos relacional, un Modelo Estrella que permita dar respuesta a los requerimientos de análisis. Para responder estos requerimientos y dar completa solución al caso deben tener en cuenta los siguientes puntos:

1. Realizar un análisis del caso, descrito en el Momento de Análisis Crítico-Contextual, de forma que les permita documentar todos los pasos requeridos para construir el Modelo Estrella.
2. Construir la estructura física del Data Mart a partir del Modelo Estrella creado en el primer apartado. Esto lo deben hacer usando Oracle Database.
3. Construir un Cubo OLAP a partir del Modelo Multidimensional creado en el primer apartado. Para este punto deben usar Pentaho Schema Workbench.
4. Implementar un proceso ETL que permita poblar el Data Mart a partir de los datos almacenados en la base de datos relacional de la IATA. Para este punto deben usar Pentaho Data-Integration.
5. Visualizar los datos y obtener los resultados de los requerimientos de análisis, utilizando para ello el Data Mart, el Cubo OLAP y LinceBI.
6. Un escrito reflexivo breve, descrito en el Momento de Conclusión.

Recursos

La estructura física base de datos relacional de la IATA, junto con los datos almacenados en ella, se pueden encontrar como insumo en el Momento de Definición y Delimitación del Caso, acompañando a este documento. Este recurso fue preparado por el docente con el propósito de resolver el caso propuesto, por lo que el volumen de datos está pensado solamente para un contexto académico.