

**UNIVERSIDAD DE EL SALVADOR**  
**FACULTAD DE INGENIERIA Y ARQUITECTURA**  
**ESCUELA DE INGENIERIA DE SISTEMAS INFORMATICOS**  
**ESPECIALIZACIÓN EN INGENIERÍA DE DATOS**



**Parcial N° 3**

**Catedrático:** Ing. René Fabricio Quintanilla Gómez

**Estudiante:** Samael Arnulfo Agustín Alvayero N° Carné: AA18087

Ciudad Universitaria 25 de octubre 2024

# Introducción

En este documento se presenta el análisis del requerimiento de un Data Warehouse (DW) para la empresa aeronáutica “Para Volar” y su sistema de vuelos “Gamatron”.

La empresa “Para Volar” actualmente opera el sistema “Gamatron”, el cual gestiona los vuelos comerciales realizados hacia diferentes aeropuertos alrededor del mundo, incluyendo la administración de reservaciones y abordajes. Tras varios años de operación, la compañía ha decidido avanzar en su estrategia de Big Data, para lo cual me ha contratado para diseñar e implementar un repositorio donde se pueda almacenar y acceder a información clave de manera ágil y eficiente.

Como parte del proyecto, se ha compartido un conjunto de datos descargados directamente desde los sistemas de la empresa, que muestra el estado de los vuelos. Esta información se encuentra disponible en el siguiente enlace: [Flight Status Prediction](#)

La compañía busca, en particular, obtener una proyección de vuelos cancelados, retrasados y diferidos, así como analizar las causas detrás de estos incidentes, que pueden incluir factores como clima, fallas en el equipo (llantas), seguridad del viaje y llegadas tardías. Este análisis servirá como base para el desarrollo de estrategias de optimización y mejora en la gestión de vuelos y reservas.

# Modelo dimensional propuesto

**Proceso:** Análisis de vuelos.

**Granularidad:** Cada fila de datos representa un vuelo específico, con detalles asociados a la aerolínea, origen, destino, y fecha.

## Dimensiones:

- a) DimFecha
  - TiempoID
  - Year
  - Quarter
  - Month
  - DayofMonth
  - DayOfWeek
  - FlightDate
  
- b) DimAeropuertoOrigen
  - AeropuertoOrigenID
  - OriginAirportID
  - Origin
  - OriginCityName
  - OriginState
  - OriginStateName
  - OriginWac
  
- c) DimAeropuertoDestino
  - AeropuertoDestinoID
  - DestAirportID
  - Dest
  - DestCityName
  - DestState
  - DestStateName
  - DestWac

## Fact Table (Tabla de Hechos):

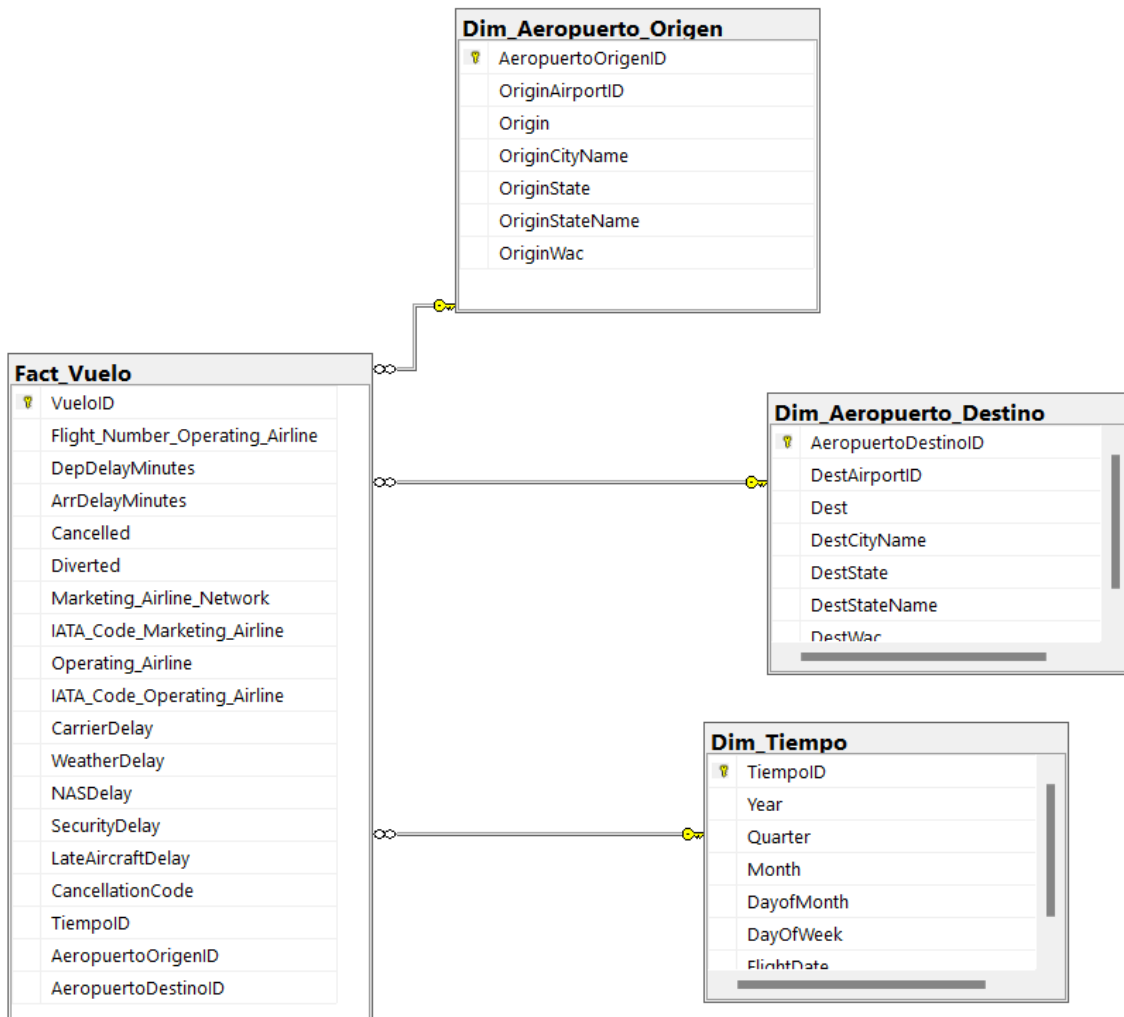
- a) Fact\_Vuelo
  - VueloID
  - Flight\_Number\_Operating\_Airline: Número de vuelo de la aerolínea operadora.
  - FlightDate: Fecha del vuelo.
  - DepDelayMinutes: Minutos de retraso en la salida.
  - ArrDelayMinutes: Minutos de retraso en la llegada.

- Cancelled: Indicador de vuelo cancelado.
- Diverted: Indicador de vuelo desviado.
- Marketing\_Airline\_Network: Código de la aerolínea de marketing.
- IATA\_Code\_Marketing\_Airline: Código IATA de la aerolínea de marketing.
- Operating\_Airline: Código de la aerolínea operadora.
- IATA\_Code\_Operating\_Airline: Código IATA de la aerolínea operadora.
- CarrierDelay: Minutos de retraso por problemas de la aerolínea.
- WeatherDelay: Minutos de retraso por condiciones meteorológicas.
- NASDelay: Minutos de retraso por el Sistema Nacional del Espacio Aéreo (NAS).
- SecurityDelay: Minutos de retraso por razones de seguridad.
- LateAircraftDelay: Minutos de retraso por llegada tardía de la aeronave.
- CancellationCode: Código de cancelación.
- Métricas:
- **Número de Vuelos Exitosos:** Vuelos realizados sin ninguna incidencia (sin retrasos ni cancelaciones).
- **Vuelos Cancelados:** Total de vuelos con el indicador Cancelled marcado.
- **Vuelos Atrasados:** Total de vuelos con demoras, utilizando campos como DepDelayMinutes y ArrDelayMinutes.
- **Vuelos Diferidos:** Vuelos desviados de su ruta original, identificados con el indicador Diverted marcado.

#### Métricas:

1. **Vuelos Cancelados:** Total de vuelos con el indicador Cancelled marcado.
2. **Vuelos Atrasados:** Total de vuelos con demoras, utilizando campos como DepDelayMinutes y ArrDelayMinutes.
3. **Vuelos Diferidos:** Vuelos desviados de su ruta original, identificados con el indicador Diverted marcado.

## Diagrama dimensional para el DW



# Script de creación de modelo dimensional

El script para la creación de las tablas se encuentra detallado en el archivo con nombre "DW\_P3\_AA18087.sql" ubicado en la carpeta raíz del proyecto.

## Conjunto de Datos

Dado el gran tamaño de la información utilizada para cargar las dimensiones, no se incluyó en el repositorio. Sin embargo, puedes descargarla desde el siguiente enlace:

<https://drive.google.com/file/d/1rWauw-GkeeMjtNBeQxiNGqEbrDSuvy5L/view?usp=sharing>

## Carga de datos exitosa

Para documentar el éxito en la carga de datos, se tomaron capturas como comprobante. Estas imágenes se encuentran en la carpeta raíz del proyecto, dentro de la carpeta "imagesDW".