

**UNIVERSIDAD DE EL SALVADOR**  
**Facultad de Ingeniería y Arquitectura**  
**Escuela de Ingeniería de Sistemas Informáticos**  
**Especialización Ingeniería de Datos**



**Solución Parcial 3**

**Procesamiento de datos – Empresa para Volar**

**DOCENTE:**

**Mcs. Fabricio Quintanilla**

**Estudiante:**

**ML17018 - Mendoza López, Claudia Roxana**

***26 de octubre de 2024, Ciudad Universitaria, San Salvador***

# Contenido

---

1. ENUNCIADO del problema.....	3
2. SOLUCION .....	4
1.    Diseñar una base de datos para almacenar el datawarehouse con el que se analizará esta información.....	4
2.    Cree una estructura de carpetas para procesar la información .....	4
3.    Cree una base de datos en SQL Server siguiendo el diseño de la base de datos elaborada en el punto 1.....	5
4.    Construya procesos ETL para la carga de información en la carpeta designada y su procesamiento para prepararlo para su carga en SQL Server. ....	5

# 1. ENUNCIADO DEL PROBLEMA

---

La empresa aeronautica “Para Volar” tiene en operación el sistema de vuelos “Gamatron”, el cual se encarga de operativizar los vuelos comerciales realizados por la empresa hacia diferentes aeropuertos en el mundo a través de reservaciones de vuelos y de abordajes por vuelos realizados.

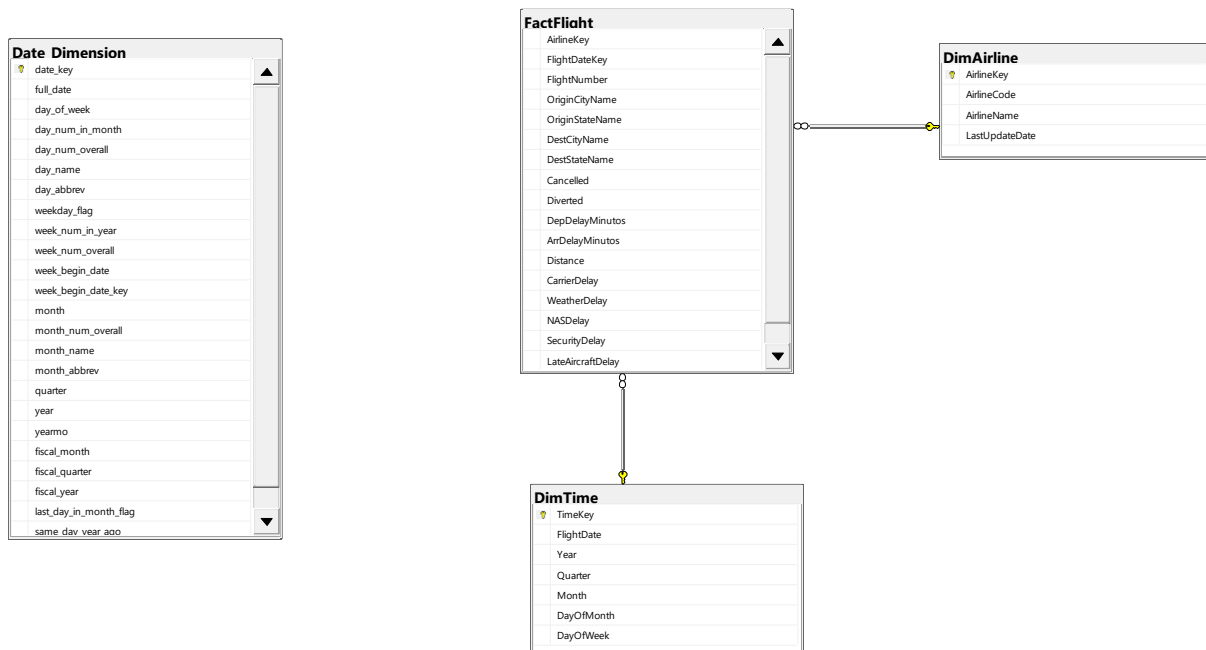
Después de varios años operando, la compañía ha decidido contratarlo para diseñar su estrategia de Big Data, para lo cual le comparte la información que se descarga de sus sistemas donde se muestran datos sobre estados de vuelos y se busca que esta sea cargada en un repositorio donde se pueda exponer dicha información de forma ágil. Esta información se puede descargar de este enlace:

<https://www.kaggle.com/datasets/robikscube/flight-delay-dataset-20182022>

Para comenzar a implementar su estrategia de Big Data, el negocio requiere hacer una proyección de vuelos cancelados, retrasados y diferidos, así como las causas del posible incidente (clima, llantas, seguridad del viaje, llegadas tardías, etc).

## 2. SOLUCION

1. Diseñar una base de datos para almacenar el datawarehouse con el que se analizará esta información



2. Cree una estructura de carpetas para procesar la información

Nombre	Fecha de Modificación	Tipo	Tamaño
Database design	26/10/2024 9:47	Carpeta de archivos	
Parcial_ML17108	19/10/2024 18:33	Carpeta de archivos	
RAW	26/10/2024 9:44	Carpeta de archivos	
SQL	26/10/2024 9:46	Carpeta de archivos	
.gitignore	19/10/2024 18:29	Archivo de origen Git...	7 KB
README	19/10/2024 18:18	Archivo de origen M...	1 KB

En la cuales se encuentra

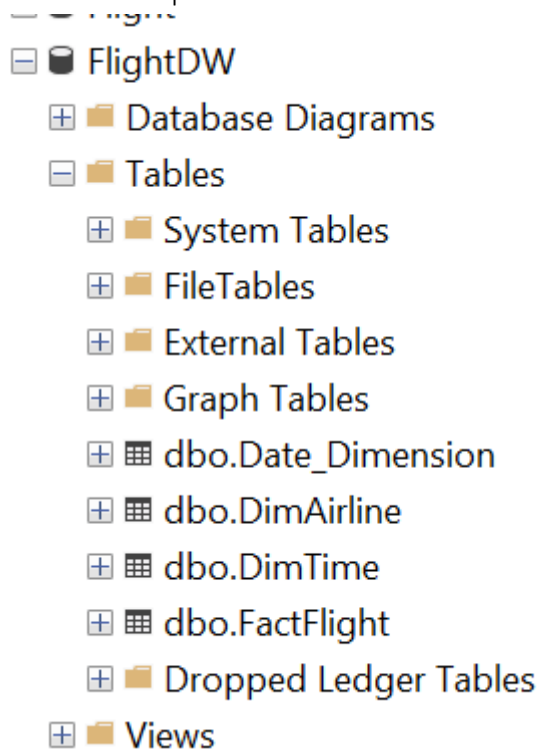
Database Design: Diagrama del diseño del data warehouse

Parcial\_ML17018: Solución de visual Studio 2022

RAW: CSV usados para poblar el DW

SQL: Script para crear base de datos y reseteo de llave de dimensiones y tabla de hecho

3. Cree una base de datos en SQL Server siguiendo el diseño de la base de datos elaborada en el punto 1.



4. Construya procesos ETL para la carga de información en la carpeta designada y su procesamiento para prepararlo para su carga en SQL Server.

