

### **Descripción General del Proyecto**

FleetLogix es una empresa logística que busca modernizar su ecosistema de datos y optimizar la toma de decisiones a través de análisis históricos, modelos dimensionales y procesamiento en tiempo real.

El Proyecto Integrador se compone de cuatro avances que cubren todo el ciclo de vida de datos: desde consultas operativas hasta una arquitectura cloud serverless.

Este documento resume de forma integrada todo el trabajo realizado.

---

### **Avance 1 – Análisis Operacional y Consultas SQL**

**Objetivo:** comprender la base operativa y responder preguntas del negocio utilizando SQL.

#### **Tareas realizadas:**

- Exploración de tablas principales: deliveries, trips, vehicles, drivers y routes.
- Limpieza y estandarización básica de datos.
- Construcción de consultas SQL para obtener:
  - distancia recorrida
  - eficiencia de combustible
  - horas trabajadas por conductor
  - entregas por viaje
  - tiempos de entrega y retrasos
- Optimización de JOINS y filtros para mejorar tiempos de respuesta.

**Tecnologías:** PostgreSQL, SQL, DBeaver.

---

### **Avance 2 – Diseño del Modelo Dimensional**

**Objetivo:** transformar la base operativa en un modelo analítico apto para Business Intelligence.

#### **Tareas realizadas:**

- Identificación de métricas y dimensiones clave.
- Diseño del modelo estrella (star schema) con una tabla de hechos central.
- Definición del primer diagrama dimensional.
- Evaluación de modelos: estrella, snowflake e híbrido.
- Selección final de modelo híbrido (vehículos y conductores con historización).

**Tecnologías:** SQL, modelado dimensional, diseño de diagramas.

---

### **Avance 3 – Implementación del Data Warehouse + ETL Automatizado**

**Objetivo:** construir un Data Warehouse real en Snowflake y un pipeline ETL desde PostgreSQL.

#### **Data Warehouse en Snowflake**

Se crearon:

- **Base:** FLEETLOGIX\_DW
- **Warehouse:** con AUTO\_RESUME y AUTO\_SUSPEND
- **Schema:** ANALYTICS
- **Time Travel:** 1 día de retención
- **Dimensiones implementadas:**
  - dim\_date
  - dim\_time
  - dim\_driver (SCD Tipo 2)
  - dim\_vehicle (SCD Tipo 2)
  - dim\_customer
  - dim\_route

- **Tabla de hechos:** fact\_deliveries  
Incluye métricas avanzadas como: tiempo de entrega, retrasos, eficiencia de combustible, entregas por hora, costo estimado y revenue por entrega.

### **Pipeline ETL Automatizado**

Desarrollado en Python, contiene:

- Extracción diaria desde PostgreSQL
- Transformación con cálculos avanzados
- Validaciones de calidad
- Carga incremental
- Manejo de historización con SCD Tipo 2
- Cálculo de totales diarios (tabla daily\_totals)
- Logging completo y manejo de errores
- Automatización diaria a las 02:00 AM

**Tecnologías:** Snowflake, PostgreSQL, Python (pandas, psycopg2, schedule, snowflake-connector).

---

### **Avance 4 – Arquitectura Cloud en AWS**

**Objetivo:** migrar la operación a la nube y permitir procesamiento en tiempo real.

#### **Componentes diseñados**

##### **API Gateway**

Punto de entrada para recibir datos desde las apps móviles de los conductores.

##### **AWS Lambda**

Se desarrollaron tres funciones:

1. Verificar si una entrega fue completada
2. Calcular ETA
3. Enviar alertas por desvío de ruta (distancia Haversine + SNS)

##### **DynamoDB**

- Tabla de estado actual de entregas
- Tabla de historial de alertas

#### **RDS PostgreSQL**

Migración de la base transaccional.  
Activa backups automáticos de 7 días.

#### **S3**

Almacenamiento histórico de datos crudos.

#### **Script IaC**

Se desarrolló infraestructura como código en Python/boto3 para aprovisionar:

- S3
- DynamoDB
- RDS

#### **Diagrama Arquitectónico AWS**

Incluye:

API Gateway → Lambda → DynamoDB/SNS → S3 → RDS

**Tecnologías:** AWS (Lambda, API Gateway, DynamoDB, SNS, RDS, S3), Python, boto3.

---

#### **Conclusión General**

El Proyecto Integrador de FleetLogix implementa una solución completa que cubre:

- análisis operativos,
- diseño dimensional,
- construcción de un Data Warehouse en Snowflake,
- pipeline ETL profesional,
- y una arquitectura cloud serverless en AWS.

Con este desarrollo, FleetLogix puede comenzar a generar análisis históricos, tableros ejecutivos, alertas en tiempo real y futuros modelos de machine learning.