

## Prueba Técnica para vacante de Data Engineer

### Proyecto: Creación de un Data Warehouse y Dashboard para análisis de datos

#### Descripción del proyecto:

Tu objetivo es diseñar y desarrollar un Data Warehouse para almacenar y analizar datos de conectividad de algunas sim cards. Además, deberás construir un dashboard interactivo que muestre visualizaciones y métricas relevantes para el análisis del proyecto. El proyecto constará de las siguientes tareas:

- Diseño del esquema de la base de datos: Deberás diseñar el esquema de la base de datos del Data Warehouse, teniendo en cuenta las entidades y relaciones relevantes para el análisis. Considera aspectos como la granularidad de los datos, las dimensiones y las medidas necesarias para el análisis.
- Extracción, transformación y carga (ETL): Deberás desarrollar un proceso de ETL para extraer datos de fuentes diversas (por ejemplo, archivos CSV), transformarlos según las necesidades del esquema del Data Warehouse y cargarlos en las tablas correspondientes.
- Creación de tablas y modelos: Utilizando un lenguaje de consultas adecuado (por ejemplo, SQL), deberás crear las tablas necesarias en el Data Warehouse y definir las relaciones entre ellas. También podrás aplicar técnicas como particionamiento y optimización para mejorar el rendimiento de las consultas.
- Desarrollo del dashboard: Utilizando una herramienta de visualización de datos (Power BI), deberás crear un dashboard interactivo que permita explorar y analizar los datos.

#### Requisitos técnicos:

- Conocimiento y experiencia en diseño y creación de Data Warehouses.
- Dominio de herramientas y lenguajes de consulta, como SQL.
- Experiencia en el desarrollo de procesos ETL para la extracción, transformación y carga de datos.
- Habilidades en el uso de herramientas de visualización de datos, Power BI.

- Capacidad para optimizar el rendimiento del Data Warehouse y las consultas.

**Entregables:**

- Esquema del Data Warehouse diseñado.
- Código y scripts del proceso de ETL.
- Tablas y modelos creados en el Data Warehouse.
- Dashboard interactivo con visualizaciones y métricas relevantes.

**Evaluación:**

- Se evaluará la calidad del diseño del Data Warehouse, la eficiencia del proceso de ETL, la estructura y optimización de las consultas, la claridad y usabilidad del dashboard.

**Consideraciones:**

- Columna ICCID: Es el identificador único de una simCard, no puede existir una misma simCard asignada a dos clientes al tiempo.
- Columna CDR\_ID: Identificador de CDR, es un id único e irrepitable que identifica una conexión realizada por una simCard.
- Columna Type: Tipo de uso realizado por la simCard, puede ser SMS o DATA
- Columna CONNECT\_TIME y CLOSE\_TIME: fecha de inicio y fin de la conexión de la simCard en UTC
- Columna Duration: bytes consumidos por la simCard en la conexión, o cantidad de SMS enviados durante la conexión.
- COUNTRY\_NAME: Nombre del país donde se realiza la conexión.
- MCC: Código del país donde se hizo la conexión
- MNC: Código de la red donde se hizo la conexión.
- IMSI\_ID: identificador internacional de suscriptor de red
- INVENTORY\_ID: Código del cliente
- COMPANY\_NAME: Nombre del Cliente