

Prueba Técnica Vivvidero - Data Junior

Dataset: Transacciones de Ventas en Línea

Objetivo

Evaluar tus habilidades en diseño de sistemas de datos, optimización de consultas, integración de datos y documentación técnica utilizando un dataset de transacciones de ventas en línea.

Parte 1:

Diseño y Desarrollo de Sistemas de Datos

Tarea: Crear un Pipeline de Datos Simple

Descripción:

Utilizando Python, crea un pipeline de datos que lea los datos del archivo CSV proporcionado, realice una transformación simple (por ejemplo, calcular el precio total si no está presente como columna) y guarde los datos transformados en una base de datos PostgreSQL.

Herramientas: Python, PostgreSQL.

Entrega: Un script Python que realice el pipeline y una breve explicación del código.

Parte 2:

Análisis y Optimización de Consultas

Tarea 1:

Análisis de Ventas

Descripción: Utilizando PostgreSQL, escribe una consulta para obtener el total de ventas (sumatoria del precio total) por categoría y región.

Entrega: La consulta SQL y los resultados obtenidos.

Tarea 2:

Optimización de Consulta

Descripción: Proporcionamos una consulta SQL que obtenga las ventas totales por categoría pero que es ineficiente. Optimiza esta consulta.

Consulta Ineficiente:

sql

Copy code

```
SELECT category, SUM(total_price)
```

FROM sales

WHERE total_price IS NOT NULL

GROUP BY category;

Herramientas: PostgreSQL.

Entrega: La consulta optimizada y una breve explicación de los cambios realizados.

Parte 3:

Documentación Técnica

Tarea: Crear Documentación Técnica

Descripción:

Documenta el pipeline de datos creado en la Parte 1 y el proceso ETL de la Parte 3, incluyendo una descripción del problema, la solución implementada y cómo se puede mantener.

Herramientas:

Cualquier herramienta de documentación (Markdown, Word, etc.).

Entrega: Documento técnico detallado.

Parte 4:

Análisis de Datos y Visualización

Tarea:

Analizar Tendencias de Ventas

Descripción: Utilizando herramientas de visualización de datos como Tableau o Power BI, crea visualizaciones que muestren las tendencias de ventas a lo largo del tiempo y la popularidad de diferentes categorías de productos en distintas regiones.

Entrega: Capturas de pantalla de las visualizaciones y una breve interpretación de los resultados.

Dataset:

https://drive.google.com/file/d/1ZFdORj_dCDAI7fDc6Gz3lBY_HldY7GT/view?usp=sharing