

Rappi Data Engineer Challenge

El desafío

Dados 3 días de datos anonimizados de transacciones, tu tarea es implementar un ETL que permita ingresar los datos necesarios y poder generar un dataset (o los que consideres necesarios) que respondan las necesidades de negocio.

Datasets

Se provee el dataset Order products (*challenge_orders.csv*) con la siguiente definición:

- *order_id*: identificador de órdenes.
- *product_id*: identificador de productos.
- *obsf_user_id*: identificador de usuarios (ofuscado).
- *payment_method*: método de pago: credit, cash, etc.
- *total_value*: valor monetario de la orden, en Pesos Colombianos.
- *created_at*: tiempo registrado de la creación de la orden.
- *store_to_user_distance*: distancia en metros del store al usuario.
- *obsf_store_type*: tipo de tienda (obfuscated)

Utilizar algún servicio Cloud para obtener la cotización del Peso Colombiano. Un ejemplo podría ser <https://openexchangerates.org/>.

Necesidades de Negocio

1. Ranking por tipo de tienda que más dólares generaron.
2. En el ranking anterior, indicar la proporción por método de pago.
3. Incluir al ranking el AOV (average order value) del tipo de tienda.
4. Algún otro insight que el candidato considere pertinente.

Sugerencias Técnicas

- Diseñar e implementar una solución escalable y robusta.
- Seguir las buenas prácticas de desarrollo, en el lenguaje de programación elegido.
- Considerar un buen modelado de datos.
- Asegurar la calidad del código.
- Enviar la solución en un repositorio privado. Incluyendo todo lo necesario para poder reproducir la solución.
- Proveer formas de visualizar los resultados.

Entrega de Resultados

1. El entregable principal será una presentación.
 - a. Contendrá las respuestas a las necesidades de negocio.
 - b. Explicación de la solución técnica.
2. Puedes utilizar cualquier herramienta, lenguaje y servicio Cloud que necesites.
 - a. La implementación deberá ser entregada en un repositorio privado, compartiendo el repositorio a ariel.wolfmann@rappi.com .
3. Tendrás 30 minutos para compartir la presentación.
4. Esperamos recibir tu trabajo dentro de los próximos 15 días.
 - a. Cualquier comunicación durante este tiempo es altamente agradecida. La colaboración para lograr objetivos es clave en nuestros equipos.

Documentación

Te alentamos que escribas cada paso que realices, decisiones de negocio, referencias que utilizaste, dificultades, insights y resultados parciales. Nos interesa mucho entender tu modo de razonar.