Case Study: Analyzing Target's Operations in Brazil (2016-2018)

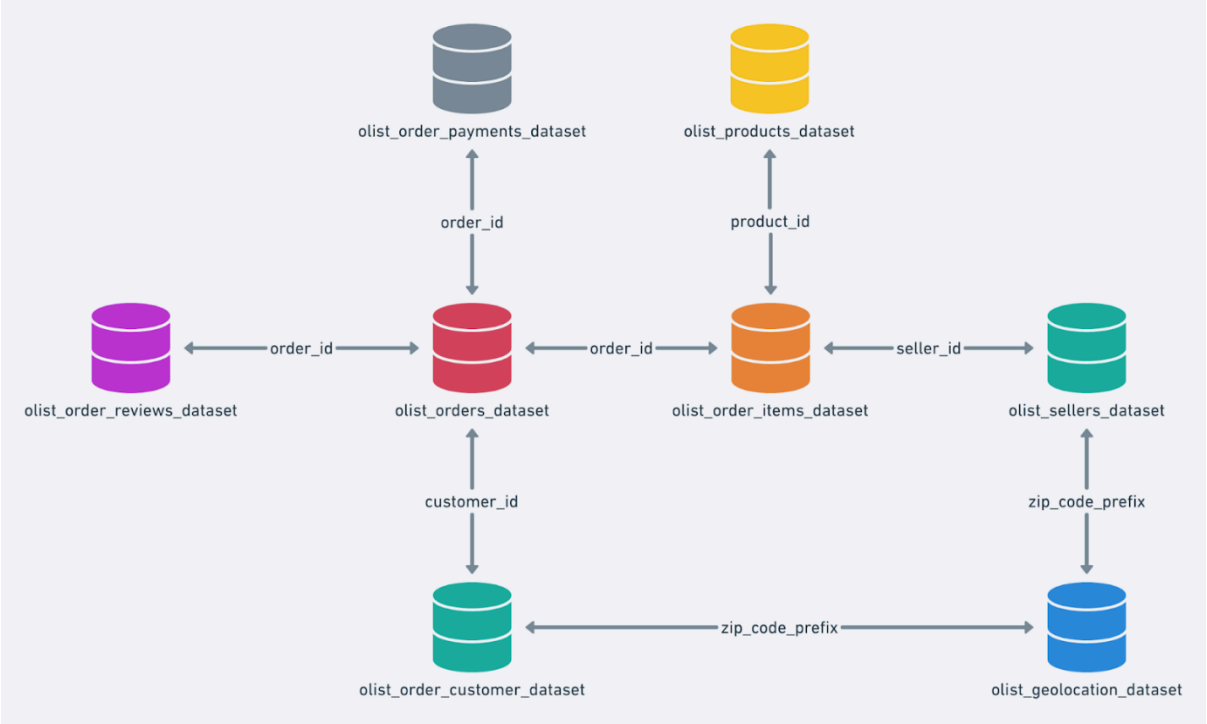
Introduction

Target is a globally renowned brand and a prominent retailer in the United States. Known for its outstanding value, innovation, and exceptional guest experience, Target aims to be a preferred shopping destination. This case study focuses on Target's operations in Brazil, analyzing a dataset of 100,000 orders placed between 2016 and 2018. The dataset includes various dimensions such as order status, pricing, payment and freight performance, customer location, product attributes, and customer reviews. By exploring this data, we can gain valuable insights into Target's Brazilian operations and provide actionable recommendations for improvement.

Dataset Description

https://drive.google.com/drive/folders/1oZOQW_jnaz-yhm5HAQJhIeOqezjl6BRf?usp=sharing

Schema



The dataset is available in 8 CSV files:

- **customers.csv**: Information about the customers.
- **sellers.csv**: Information about the sellers.

- **order_items.csv**: Details of the order items.
- **geolocation.csv**: Geographical information.
- **payments.csv**: Payment details.
- **reviews.csv**: Customer reviews.
- **orders.csv**: Order details.
- **products.csv**: Product details.

Column Descriptions

customers.csv:

- `customer_id`: ID of the consumer who made the purchase.
- `customer_unique_id`: Unique ID of the consumer.
- `customer_zip_code_prefix`: Zip Code of the consumer's location.
- `customer_city`: Name of the city from where the order was made.
- `customer_state`: State code from where the order was made.

sellers.csv:

- `seller_id`: Unique ID of the seller.
- `seller_zip_code_prefix`: Zip Code of the seller's location.
- `seller_city`: Name of the city of the seller.
- `seller_state`: State code.

order_items.csv:

- `order_id`: Unique ID of the order.
- `order_item_id`: Unique ID of each item ordered.
- `product_id`: Unique ID of each product.
- `seller_id`: Unique ID of the seller.
- `shipping_limit_date`: Date before which the product must be shipped.
- `price`: Price of the products ordered.
- `freight_value`: Freight cost for delivery.

geolocations.csv:

- `geolocation_zip_code_prefix`: First 5 digits of Zip Code.
- `geolocation_lat`: Latitude.
- `geolocation_lng`: Longitude.
- `geolocation_city`: City.
- `geolocation_state`: State.

payments.csv:

- `order_id`: Unique ID of the order.
- `payment_sequential`: Sequences of payments in case of EMI.
- `payment_type`: Mode of payment used.
- `payment_installments`: Number of installments for EMI purchase.
- `payment_value`: Total amount paid for the purchase order.

orders.csv:

- `order_id`: Unique ID of the order.
- `customer_id`: ID of the consumer who made the purchase.
- `order_status`: Status of the order.
- `order_purchase_timestamp`: Timestamp of the purchase.
- `order_delivered_carrier_date`: Delivery date at which carrier made the delivery.
- `order_delivered_customer_date`: Date at which customer received the product.
- `order_estimated_delivery_date`: Estimated delivery date.

reviews.csv:

- `review_id`: ID of the review.
- `order_id`: Unique ID of the order.
- `review_score`: Review score given by the customer.
- `review_comment_title`: Title of the review.
- `review_comment_message`: Review comments.
- `review_creation_date`: Timestamp of the review creation.
- `review_answer_timestamp`: Timestamp of the review answered.

products.csv:

- `product_id`: Unique identifier for the product.
- `product_category_name`: Name of the product category.
- `product_name_length`: Length of the product name.
- `product_description_length`: Length of the product description.
- `product_photos_qty`: Number of photos of each product.
- `product_weight_g`: Weight of the product in grams.
- `product_length_cm`: Length of the product in centimeters.
- `product_height_cm`: Height of the product in centimeters.
- `product_width_cm`: Width of the product in centimeters.

Problem Statement

As a data analyst/scientist at Target, your task is to analyze the dataset to extract valuable insights and provide actionable recommendations.

Evaluation Criteria (100 points)

1. **Initial Exploration (15 points)**
    - Check the structure & characteristics of the dataset.
    - Data type of all columns in the "customers" table.
    - Get the time range between which the orders were placed.
    - Count the cities & states of customers who ordered during the given period.
2. **In-depth Exploration (15 points)**
    - Identify trends in the number of orders placed over the years.
    - Detect any monthly seasonality in the number of orders.

    - Determine the time of day when Brazilian customers mostly place orders (Dawn, Morning, Afternoon, or Night).
3. **Evolution of E-commerce Orders in Brazil (10 points)**
    - Month-on-month number of orders placed in each state.
    - Distribution of customers across all states.
4. **Impact on Economy (20 points)**
    - Analyze money movement by looking at order prices, freight, and other factors.
    - Calculate the percentage increase in the cost of orders from 2017 to 2018 (Jan-Aug).
    - Calculate the total & average value of order prices and freight for each state.
5. **Analysis on Sales, Freight, and Delivery Time (20 points)**
    - Calculate the delivery time and the difference between estimated and actual delivery dates.
    - Identify the top 5 states with the highest & lowest average freight values.
    - Identify the top 5 states with the highest & lowest average delivery times.
    - Identify the top 5 states where delivery is faster than the estimated date.
6. **Analysis Based on Payments (10 points)**
    - Month-on-month number of orders placed using different payment types.
    - Number of orders based on payment installments.
7. **Actionable Insights & Recommendations (10 points)**
    - Provide insights and recommendations based on the analysis.