Business Case: Target SQL

Target is one of the world's most recognized brands and one of America's leading retailers. Target makes itself a preferred shopping destination by offering outstanding value, inspiration, innovation and an exceptional guest experience that no other retailer can deliver.

This business case has information of 100k orders from 2016 to 2018 made at Target in Brazil. Its features allows viewing an order from multiple dimensions: from order status, price, payment and freight performance to customer location, product attributes and finally reviews written by customers.

Data is available in 8 csv files:

1. customers.csv

2. geolocation.csv

3. order_items.csv

4. payments.csv

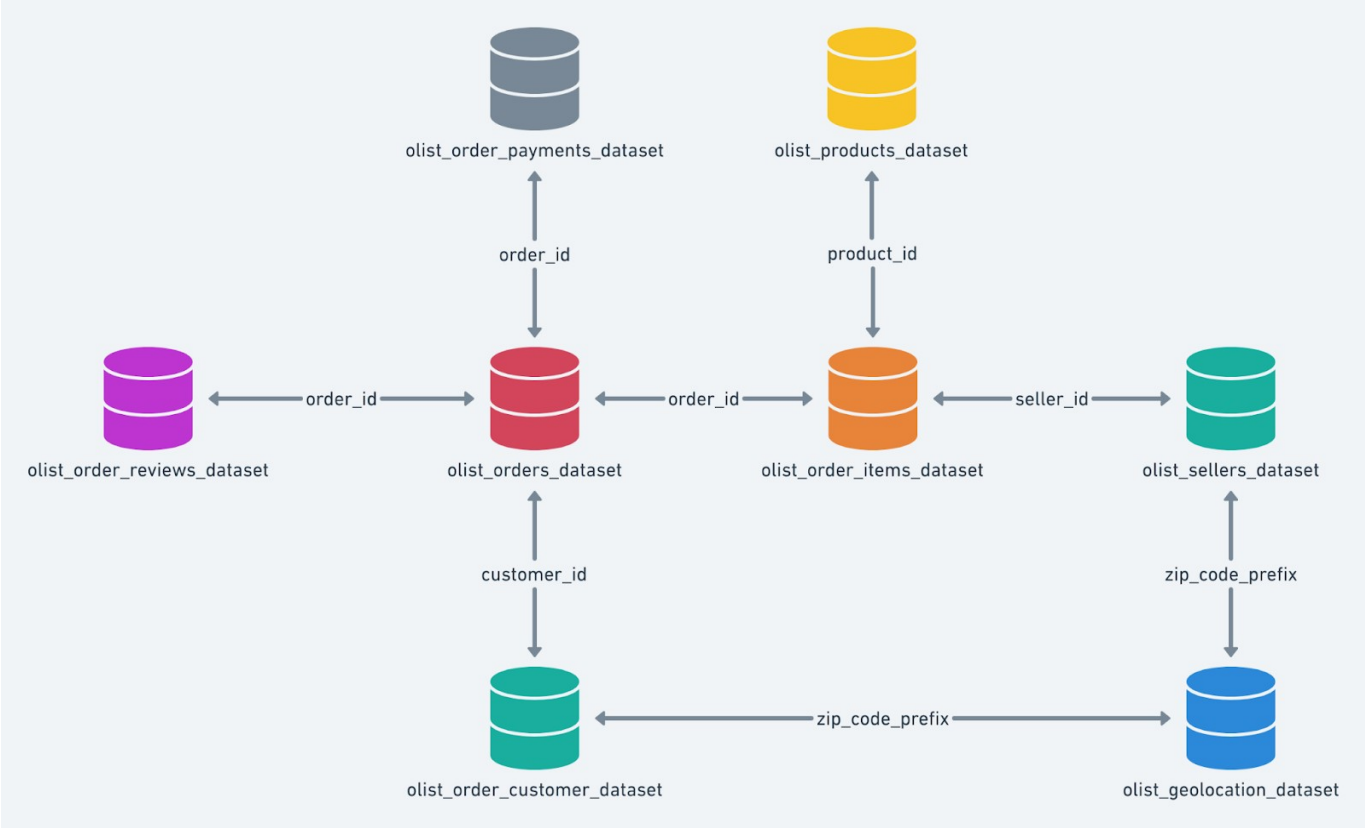
5. reviews.csv

6. orders.csv

7. products.csv

8. sellers.csv

High level overview of relationship between datasets:



Questions needed to be answered:

Assume you are a data scientist at Target, and are given this data to analyze and provide some insights and recommendations from it.

1. Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset
    1. Get number of rows in the data
    2. Number of null or missing values in a column
    3. Data type of columns in a table
    4. Get the time period for which the data is given
    5. Number of cities in our dataset
    6. Number of states in our dataset

2. In-depth Exploration:
    1. How many orders do we have for each order status?
    2. Is there a growing trend on e-commerce in Brazil? How can we describe a complete scenario?
    3. On what day of week brazilians customers tend to do online purchasing?
    4. What time do Brazilian customers tend to buy (Dawn, Morning, Afternoon or Night)?
    5. Feature Extraction: Through order_purchase_timestamp in “orders” dataset extract
        1. order_purchase_year
        2. order_purchase_month
        3. order_purchase_date
        4. order_purchase_day
        5. order_purchase_dayofweek
        6. order_purchase_dayofweek_name
        7. order_purchase_hour
        8. order_purchase_time_day
3. Evolution of E-commerce orders in the Brazil region:
    1. Get month on month orders by region
    2. Total of customer orders by state
    3. Top 10 brazilian cities most no. of orders
    4. How are customers distributed in Brazil
    5. City wise number of unique customers
4. Impact on Economy: Analyze the money movemented by e-commerce by looking at order prices, freight and others.

Answer the following questions:

1. Total amount sold in 2017 between Jan to August

2. Total amount sold in 2018 between Jan to august

3. % increase from 2017 to 2018

4. Mean & Sum of price by customer state

5. Mean & Sum of freight value by customer state

6. Analysis on sales, freight and delivery time

7. Top 5 states with highest/lowest average freight value

8. Top 5 states with highest/lowest average time to delivery

9. Top 5 states where delivery is really fast/ not so fast compared to estimated date

10. Count of orders for different payment types

11. Distribution of payment installments and count of orders

12. Count of orders for different payment types Month over Month

Business Case: Target SQL

Target is one of the world's most recognized brands and one of America's leading retailers. Target makes itself a preferred shopping destination by offering outstanding value, inspiration, innovation and an exceptional guest experience that no other retailer can deliver.

This business case has information of 100k orders from 2016 to 2018 made at Target in Brazil. Its features allows viewing an order from multiple dimensions: from order status, price, payment and freight performance to customer location, product attributes and finally reviews written by customers.

We have orders purchased from 4th Sep 2016 to 17th Oct 2018.

I have created a word document which consists of all the queries for questions asked in this business problem. It also contains answers in the form of tables corresponding to those queries. I have also found few missing values in some of the columns in the tables provided.

Given below are few of my insights and recommendations from the answers I got by writing queries:

- In the dataset given, we can see that there are 625 cancelled and 609 unavailable orders. We have to bring these values down. Otherwise, our customers will shift to a different e-commerce platform.
- We can clearly see a growing e-commerce trend in Brazil. The no of orders at Target are increasing every year and every month. The order count has increased by approx 10,000 from 2017 to 2018.
- Most number of purchases are made on weekdays as compared to weekends. Since, during weekends, people are mostly in front of televisions or mobiles, we can advertise more on the famous video streaming platforms like YouTube to bring in more orders.
- On a daily basis, people tend to purchase more during evenings which is from 17:00 to 22:59. Considering 17:00 as office end time, it is obvious that people do not get enough free time to browse the target e-commerce website. Hence, we can provide recommendations (food, stationary etc.) in the form of messages during busy hours to increase the number of orders.
- SP state and sao paulo city in brazil tops the chart with most number of customers. We need to create more awareness by advertising among other states and cities to increase the number of customers and orders.
- Between Jan to august in the years 2017 and 2018, there is an increase of 137.26 % of total amount sold. This shows huge growth. We can clearly see the demand. We need to keep the momentum going by adding more products, advertising etc.
- In the states or cities where there are less number of orders, we can offer more discounts to bring in more customers. Also, for these cities or states, we can see that the freight prices are higher. This might be the reason behind the less number of orders. We have start to building more warehouses etc. in these places to decrease the freight cost. Along, with the freight charges, we can also see that the average time of delivery is far beyond the estimated delivery date. This will clearly frustrate the customer and avoid them to rebuy again.
- Very less people tend to pay in full. Also, most of them pay the full amount within 10 months. To bring in more orders, we can go for no cost EMI or decrease the interest rates beyond 10 months.

To whoever reads this, I hope my insights from this case study were meaningful.

Thank you,
Krishna