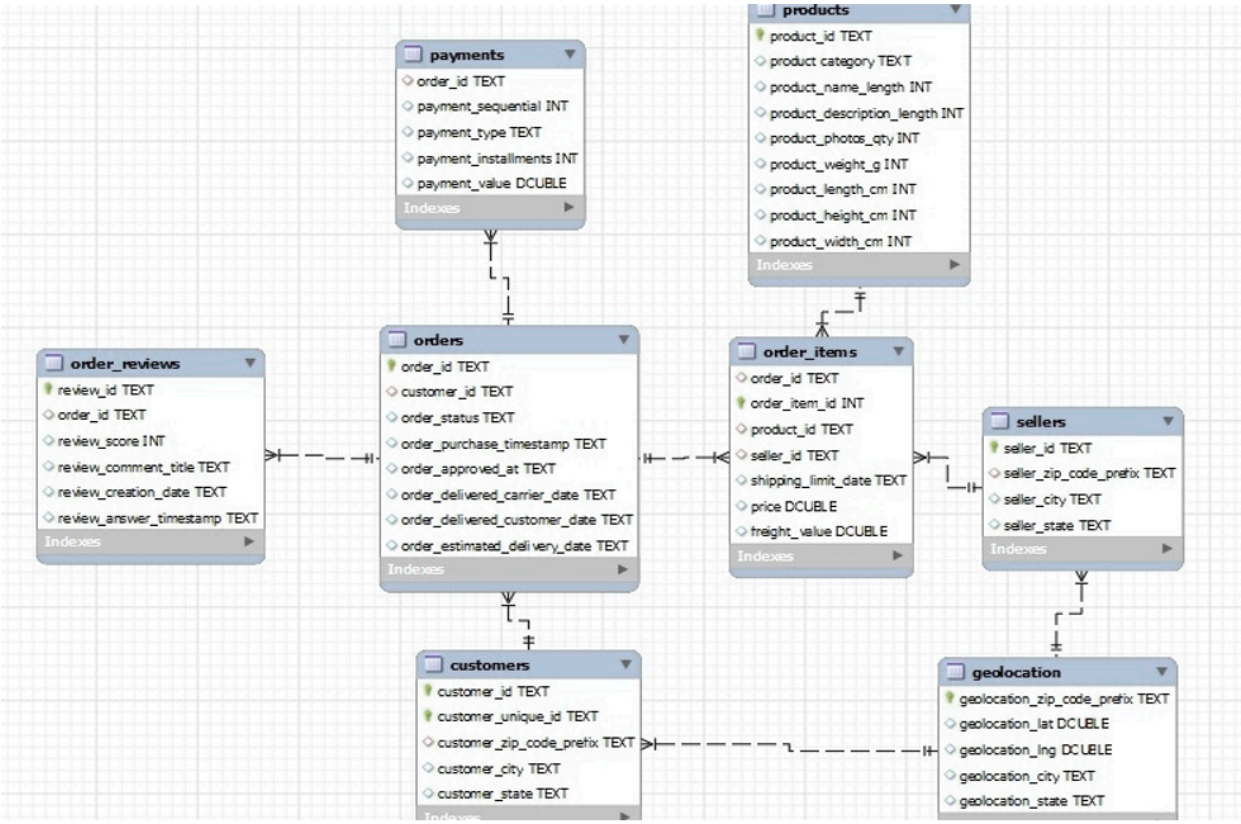
SQL PROJECT

Target is a globally renowned brand and a prominent retailer in the United States. 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 particular business case focuses on the operations of Target in Brazil and provides insightful information about 100,000 orders placed between 2016 and 2018. The dataset offers a comprehensive view of various dimensions including the order status, price, payment and freight performance, customer location, product attributes, and customer reviews.

By analyzing this extensive dataset, it becomes possible to gain valuable insights into Target's operations in Brazil. The information can shed light on various aspects of the business, such as order processing, pricing strategies, payment and shipping efficiency, customer demographics, product characteristics, and customer satisfaction levels.

Here is the Dataset Schema:



A. **Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset:**

Time range between which the orders were placed.

	start time	end time	time period
0	2016-09-04 21:15:19	2018-10-17 17:30:18	2

Total cities as well as States with unique customers from CUSTOMERS DATASET.

	total cities	total states	total_unique_customers	total_unique_zipcodes
0	4119	27	96096	14994

Total unique order and unique customers with customer order status FROM ORDERS DATASET.

	count(distinct(order_id))	count(distinct(customer_id))	count(distinct(order_status))
0	99441	99441	8

Total product id's with total sellers as well as total revenue and average of revenue as well as freight from ORDER ITEMS

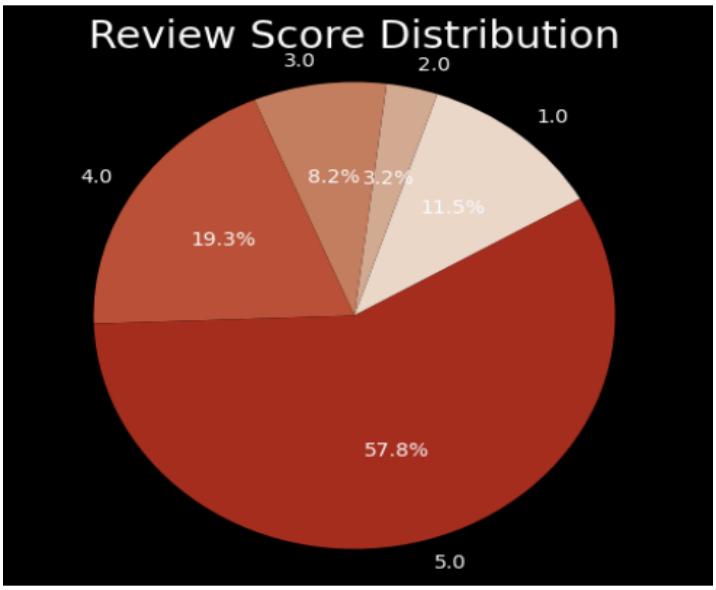
	total orders	total product ids	total sellers	total revenue	average_price	average_freight
0	98666	32951	3095	1.584355e+07	120.653739	19.99032

Total unique seller ID as well as seller cities and states

	count(distinct(seller_id))	count(distinct(seller_city))	count(distinct(seller_state))
0	3095	610	23

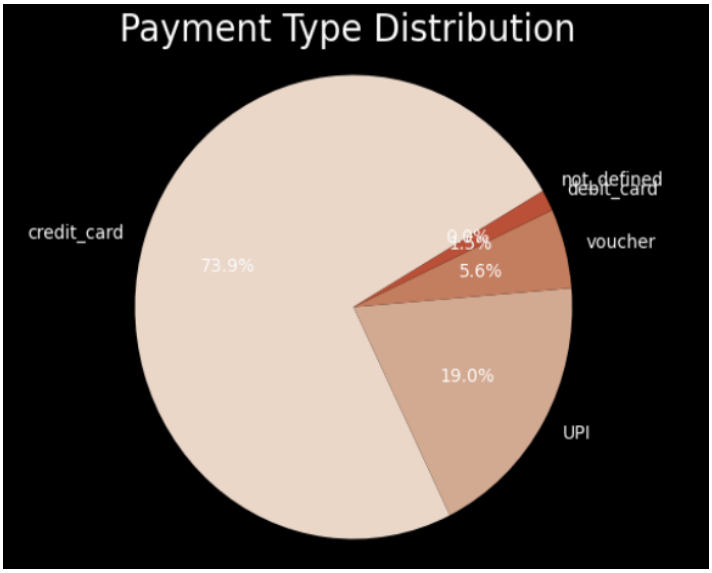
Reviews core distribution FROM REVIEWS DATASET

	review_score	frequency
0	1.0	11424
1	2.0	3151
2	3.0	8179
3	4.0	19142
4	5.0	57328

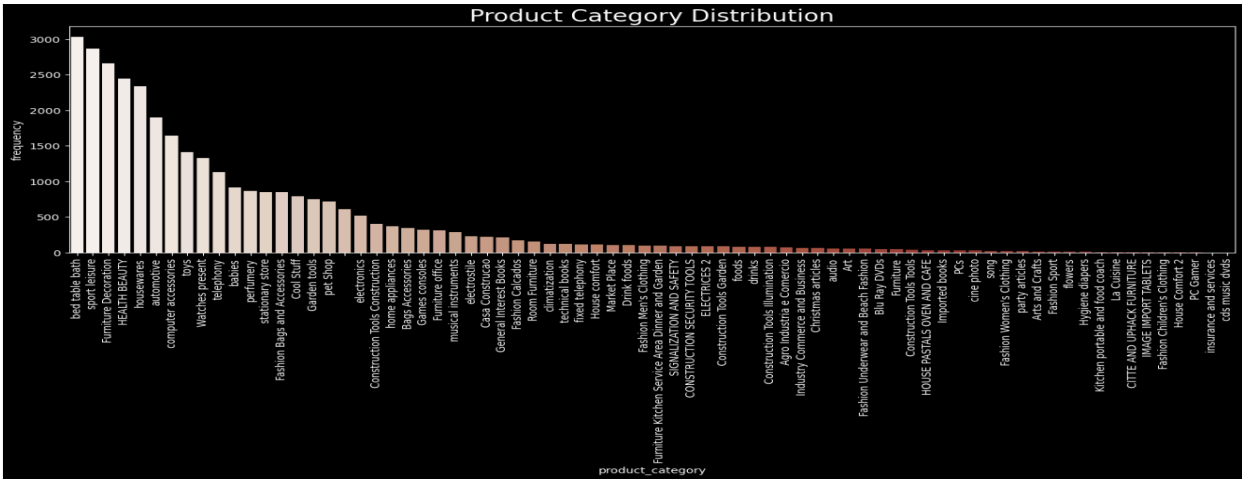


Payment-type distribution FROM PAYMENTS DATASET

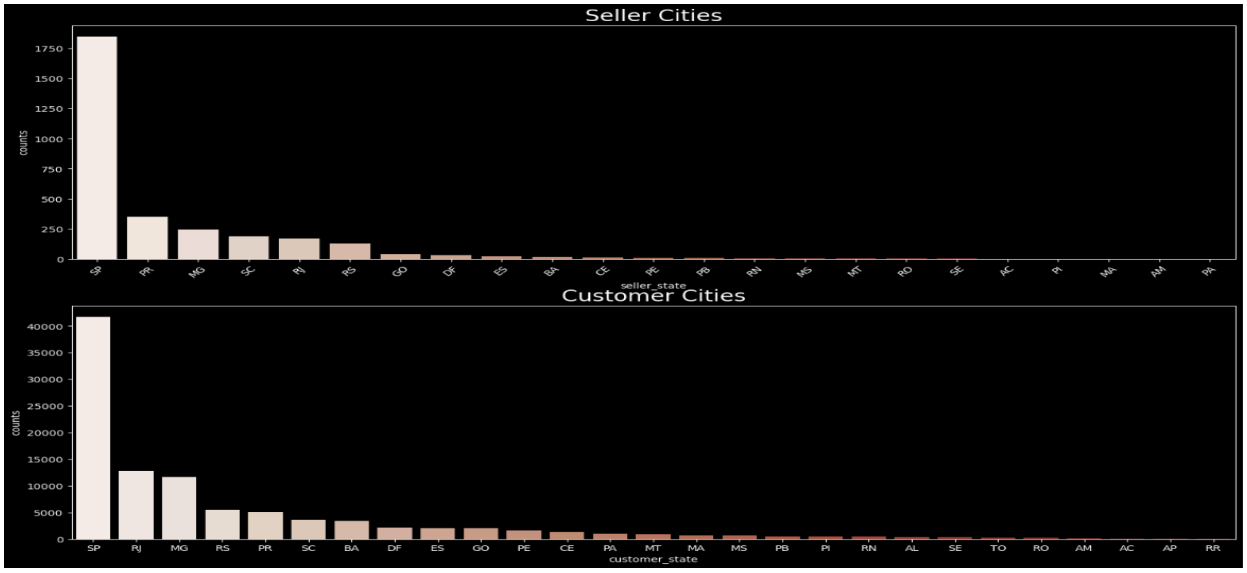
	payment_type	total counts
0	credit_card	76795
1	UPI	19784
2	voucher	5775
3	debit_card	1529
4	not_defined	3



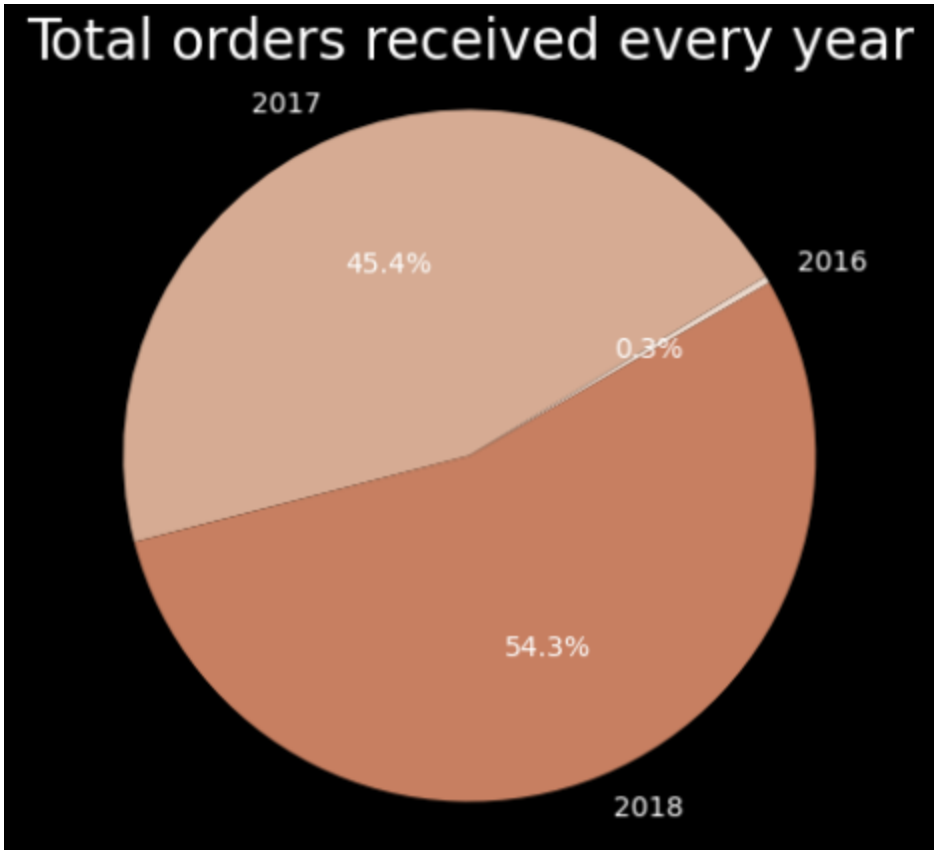
Product Categories' Distribution FROM PRODUCT DATASET.



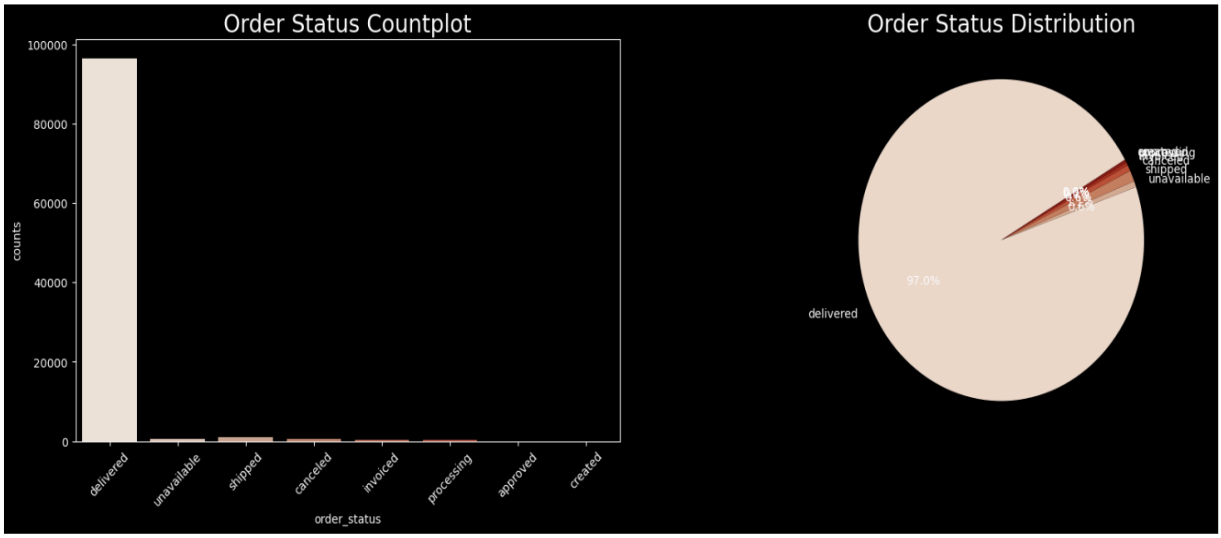
Seller Cities and Customer Cities



Yearly Distribution in the Dataset.



Order Status Distribution FROM ORDER DATASET.



DATA ANALYTICS STARTS HERE

A. Orders made on the basis of

YEAR

MONTH

WEEK

```
SELECT YEAR(order_purchase_timestamp) AS 'year',
       MONTH(order_purchase_timestamp) AS 'month',
       WEEK(order_purchase_timestamp) AS 'week',
       COUNT(order_id) as'total_orders',
        SUM(COUNT(order_id)) OVER (
        PARTITION BY YEAR(order_purchase_timestamp)
        ORDER BY YEAR(order_purchase_timestamp),
MONTH(order_purchase_timestamp)
        ROWS BETWEEN UNBOUNDED PRECEDING AND CURRENT ROW
    ) AS rolling_total_orders
FROM target.orders
WHERE  YEAR(order_purchase_timestamp) >2016 and order_status =
'delivered' and MONTH(order_purchase_timestamp) between 1 and 8
GROUP BY YEAR(order_purchase_timestamp),
MONTH(order_purchase_timestamp),
WEEK(order_purchase_timestamp)
ORDER BY YEAR(order_purchase_timestamp),
MONTH(order_purchase_timestamp),
WEEK(order_purchase_timestamp)"""
```