Target SQL Business Case Study – Project Report

This project report presents the SQL-based analysis of Target's operations in Brazil between 2016 and 2018. The dataset contains 100,000 orders with information on customers, orders, payments, products, sellers, geolocation, and reviews. The analysis is performed in BigQuery using eight CSV files loaded as separate tables in the dataset `target` under the project `Ecommerce`. The objective is to extract valuable insights from the dataset to improve operational efficiency, pricing strategies, shipping performance, and customer satisfaction.

1. Data Loading & Setup

- Created dataset `target` in BigQuery under project `Ecommerce`. - Uploaded the eight CSVs as separate tables: customers, geolocation, order_items, payments, reviews, orders, products, sellers. - Verified schema and converted date/time columns to TIMESTAMP where required using PARSE_TIMESTAMP.

2. SQL Queries & Analysis

```
The following sections outline the main SQL queries used to answer the business case
questions.

I. Exploratory Analysis:
A. Data type of all columns in customers table:
SELECT column_name, data_type
FROM `Ecommerce.target.INFORMATION_SCHEMA.COLUMNS`
WHERE table_name = 'customers';

B. Time range of orders:
SELECT MIN(order_purchase_timestamp) AS first_order, MAX(order_purchase_timestamp)
AS last_order
FROM `Ecommerce.target.orders`;

C. Count unique cities & states:
SELECT COUNT(DISTINCT customer_city) AS unique_cities, COUNT(DISTINCT
customer_state) AS unique_states
FROM `Ecommerce.target.customers`;

II. In-depth Exploration:
A. Growing trend in orders: Extract year & month from purchase timestamp, count
orders, group & order.
B. Monthly seasonality: Group only by month across all years.
C. Time of day categories: Use CASE with EXTRACT(HOUR FROM order_purchase_timestamp)
to classify into Dawn, Morning, Afternoon, Night.

III. Evolution of Orders:
A. Month-on-month orders per state: Join orders & customers, group by
state/year/month.
B. Customer distribution: COUNT(DISTINCT customer_id) grouped by state.

IV. Economic Impact:
A. % increase from Jan–Aug 2017 to Jan–Aug 2018: SUM(payment_value) grouped by year
and month filter.
B. Total & average price per state: SUM & AVG(price) joining order_items, orders,
customers.
C. Total & average freight per state: SUM & AVG(freight_value) similarly.

V. Sales, Freight & Delivery:
A. Delivery time & difference: DATE_DIFF(order_delivered_customer_date,
order_purchase_timestamp, DAY) and DATE_DIFF(order_estimated_delivery_date,
order_delivered_customer_date, DAY).
B. Top/bottom 5 avg freight: AVG(freight_value) grouped by state, ordered.
C. Top/bottom 5 avg delivery time: Similar to freight but using delivery time.
D. Fastest vs estimated delivery: Use diff_estimated_delivery.

VI. Payments:
```

```
A. Month-on-month orders per payment type: COUNT orders grouped by year, month,
payment_type.
B. Orders by installment count: COUNT orders where payment_installments >= 1 grouped
by installment count.
```

3. Key Insights

1. Orders increased steadily from 2016 to 2018, with noticeable peaks in November (Black Friday period). 2. Most orders were placed in the Afternoon and Night time windows. 3. São Paulo (SP) dominated order volume, but states like RR and AP had higher freight costs. 4. Freight charges varied significantly by region, impacting total order cost. 5. Average delivery times were higher in distant states; some states had faster deliveries than estimated dates. 6. Credit card was the most common payment method; most orders had no installments or up to 3 installments.

4. Recommendations

- Optimize freight logistics in high-cost regions to reduce overall delivery expenses. - Leverage seasonal peaks (e.g., November) with targeted promotions. - Focus marketing efforts in Afternoon and Night slots where customer activity is highest. - Improve delivery efficiency in slower states through better route planning. - Encourage higher installment options in regions with lower purchase frequency to improve sales.

End of Report