

FINAL PROJECT

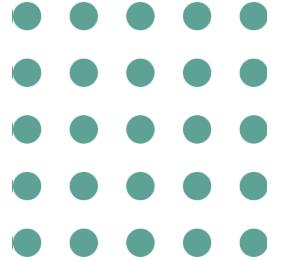
DATA ANALYST

Batch #31

1

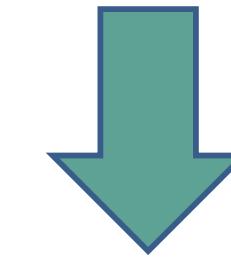
RINDI PASMAWI SYA'BAN

BEGINNER LEVEL



OLIST

MARKETPLACE BESAR di Brazil



menghubungkan bisnis kecil dan menengah
dengan platform e-commerce besar.



Objective



Customer Demographic

Mengelompokkan pelanggan



Order Behaviour

Menentukan pola pemesanan



Review Customer

Mengukur tingkat kepuasan pelanggan berdasarkan ulasan yang diberikan



Time Series Analysis

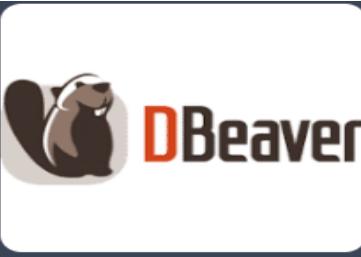
Menentukan tren penjualan berdasarkan periode waktu tertentu

Data Explanation



Sumber utama

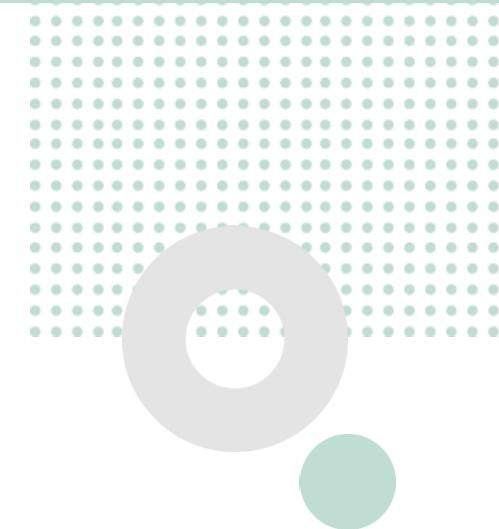
Dataset
OLIST



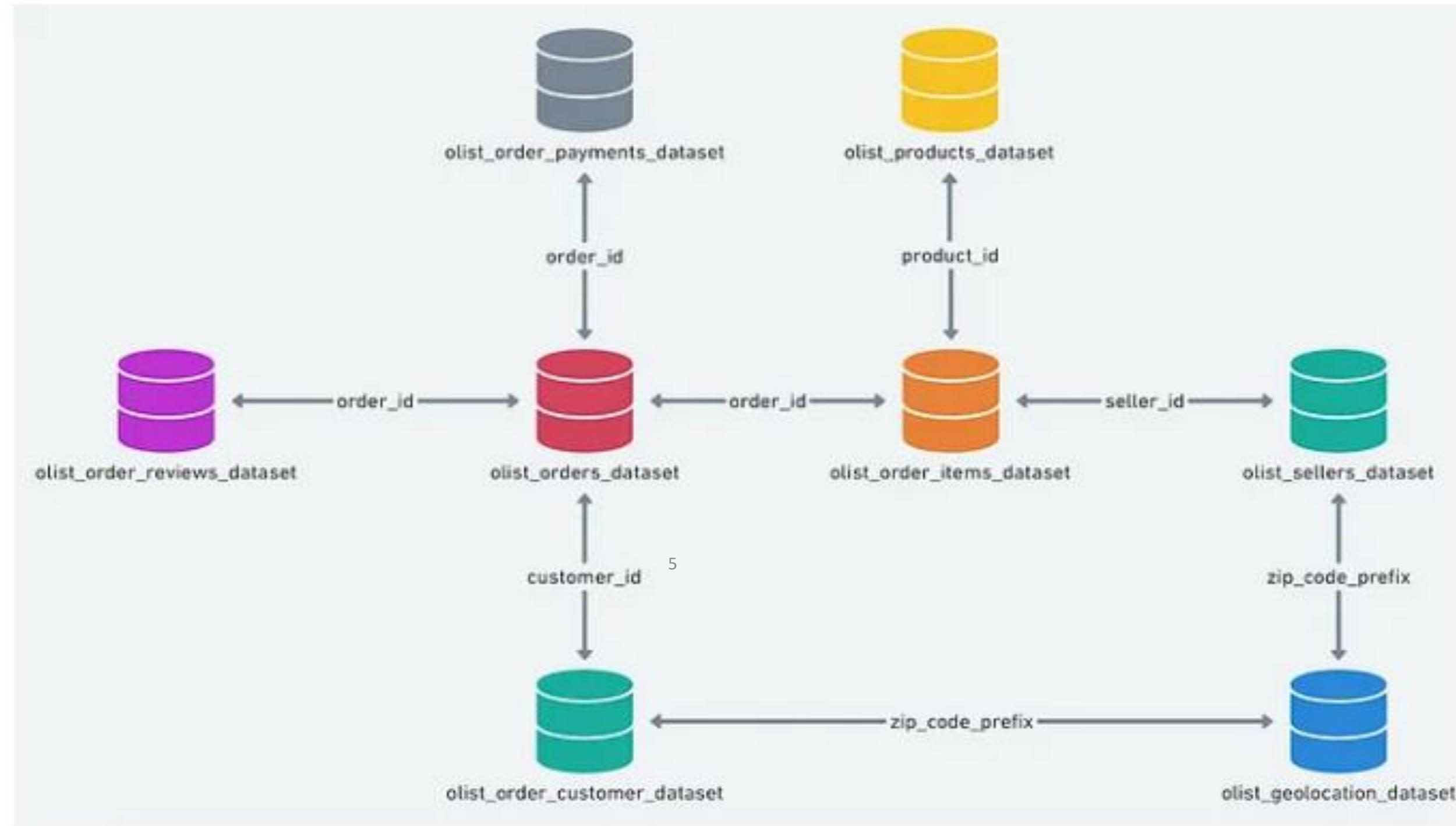
Tools



Library



Struktur Dataset OLIST



```
# Cek kolom
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 119143 entries, 0 to 119142
Data columns (total 40 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   order_id         119143 non-null   object 
 1   customer_id      119143 non-null   object 
 2   order_status      119143 non-null   object 
 3   order_purchase_timestamp  119143 non-null   object 
 4   order_approved_at 119143 non-null   object 
 5   order_delivered_carrier_date  119143 non-null   object 
 6   order_delivered_customer_date 119143 non-null   object 
 7   order_estimated_delivery_date 119143 non-null   object 
 8   customer_unique_id 119143 non-null   object 
 9   customer_zip_code_prefix 119143 non-null   object 
 10  customer_city     119143 non-null   object 
 11  customer_state    119143 non-null   object 
 12  order_item_id     118310 non-null   object 
 13  product_id        118310 non-null   object 
 14  seller_id         118310 non-null   object 
 15  shipping_limit_date 118310 non-null   object 
 16  price              118310 non-null   object 
 17  freight_value     118310 non-null   object 
 18  product_category_name 118310 non-null   object 
 19  product_name_lenght 118310 non-null   object 
 20  product_description_lenght 118310 non-null   object 
 21  product_photos_qty 118310 non-null   object 
 22  product_weight_g  118310 non-null   object 
 23  product_length_cm 118310 non-null   object 
 24  product_height_cm 118310 non-null   object 
 25  product_width_cm  118310 non-null   object 
 26  product_category_name_english 116576 non-null   object 
 27  seller_zip_code_prefix 118310 non-null   object 
 28  seller_city        118310 non-null   object 
 29  seller_state       118310 non-null   object 
 30  payment_sequential 119140 non-null   object 
 31  payment_type       119140 non-null   object 
 32  payment_installments 119140 non-null   object 
 33  payment_value      119140 non-null6  object 
 34  review_id          118146 non-null   object 
 35  review_score        118146 non-null   object 
 36  review_comment_title 118146 non-null   object 
 37  review_comment_message 118146 non-null   object 
 38  review_creation_date 118146 non-null   object 
 39  review_answer_timestamp 118146 non-null   object 
```

40 Kolom dan
119143 Baris

Preprocessing Data

**Proses
1**

Delete Unnecesary
Columns

**Proses
2**

Change Dtypes

**Proses
3**

Cek statistik
deskriptif

**Proses
4**

Cek Data Duplikat

**Proses
5**

Cek Data
Missing Value

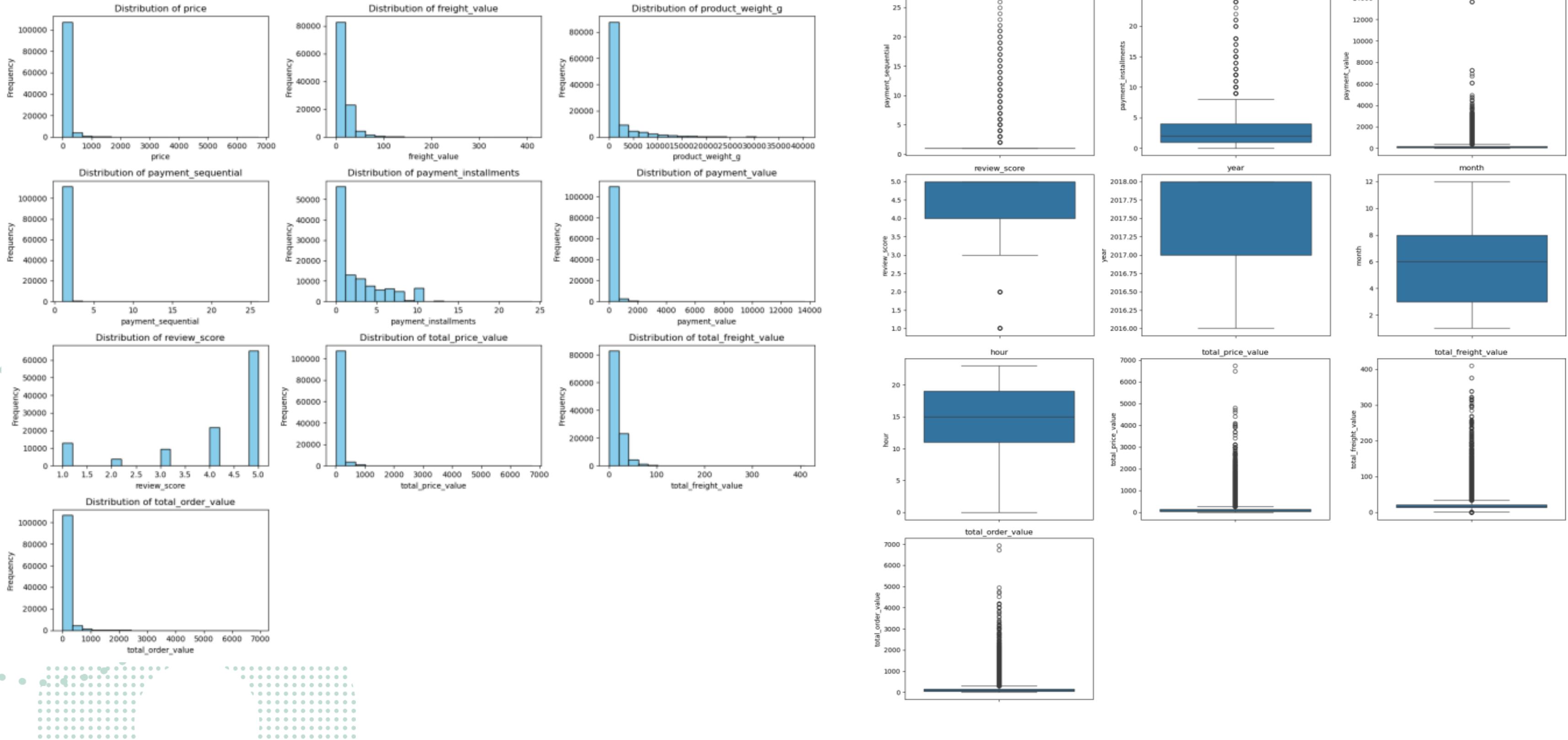
**Proses
6**

Drop Missing
Values

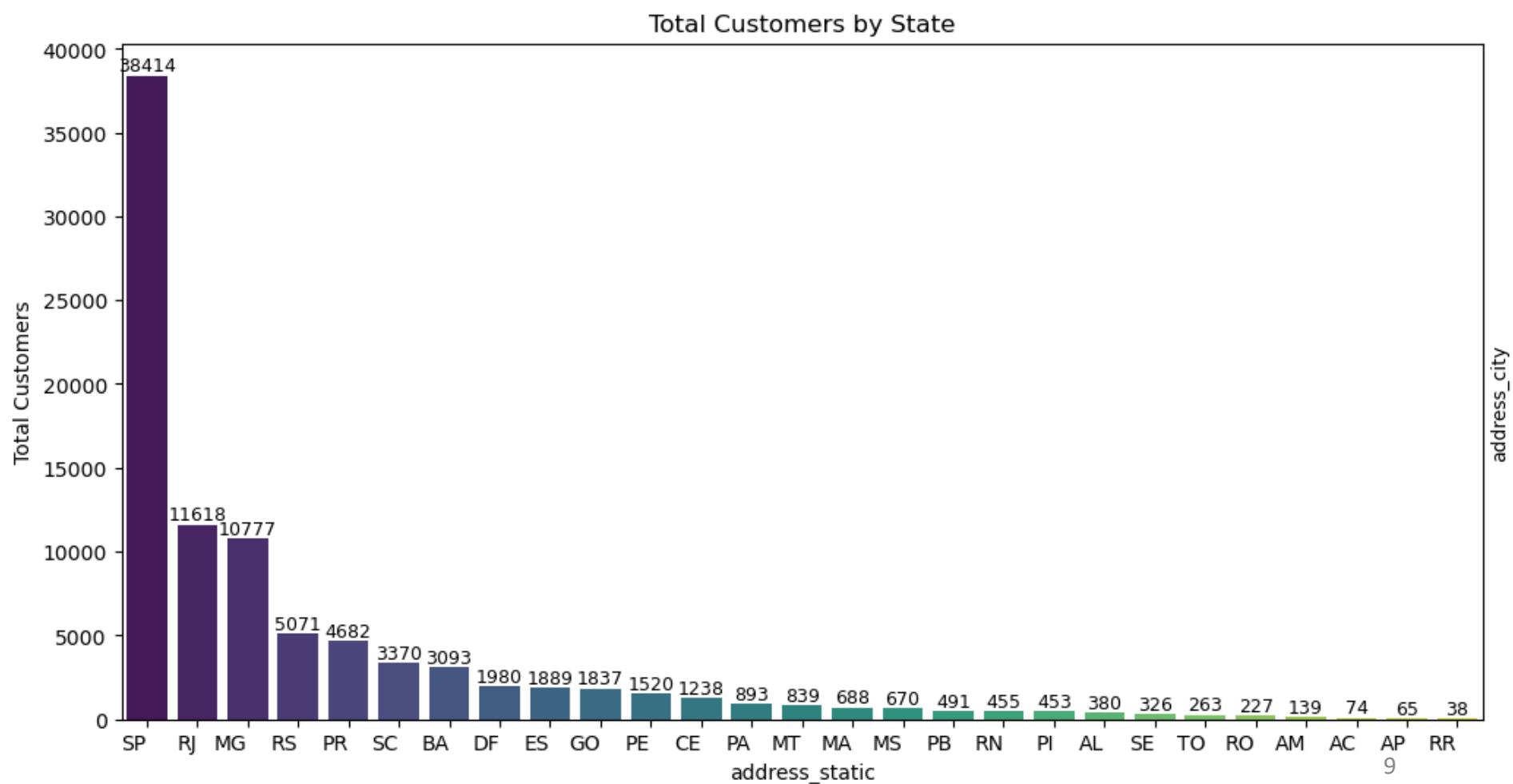
**Proses
7**

Add Columns

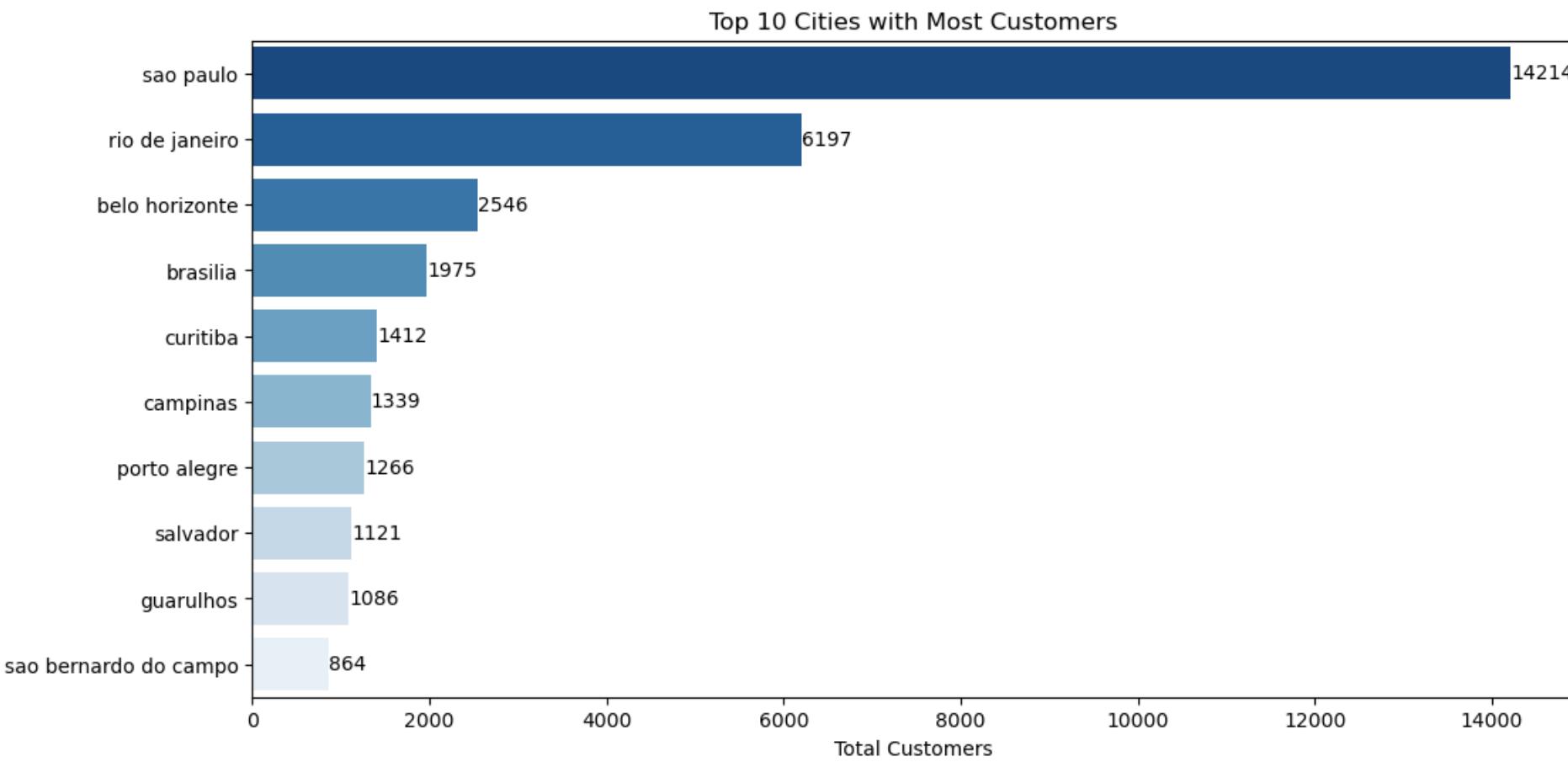
Exploratory Data Analysis



Customer Demographic

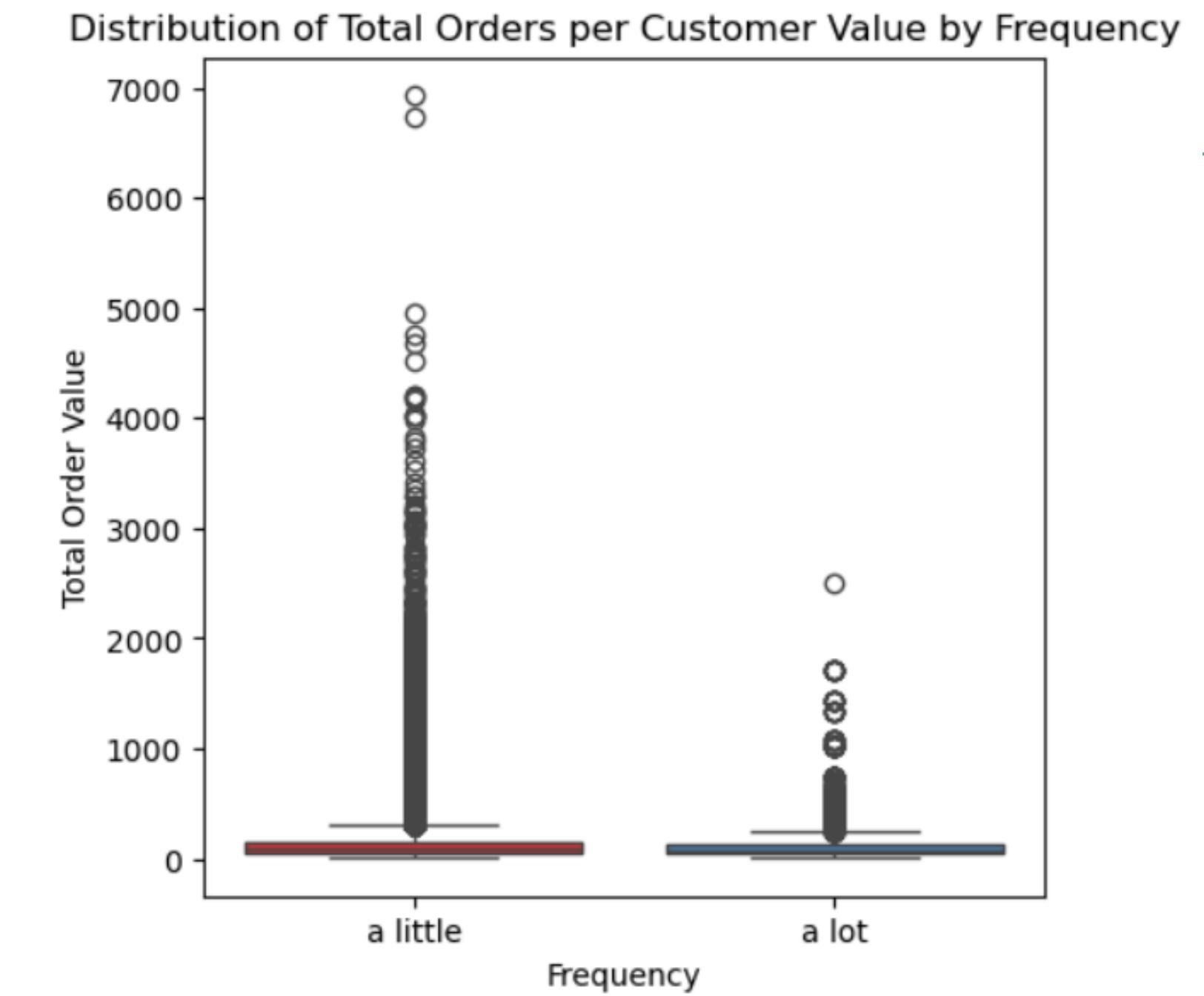
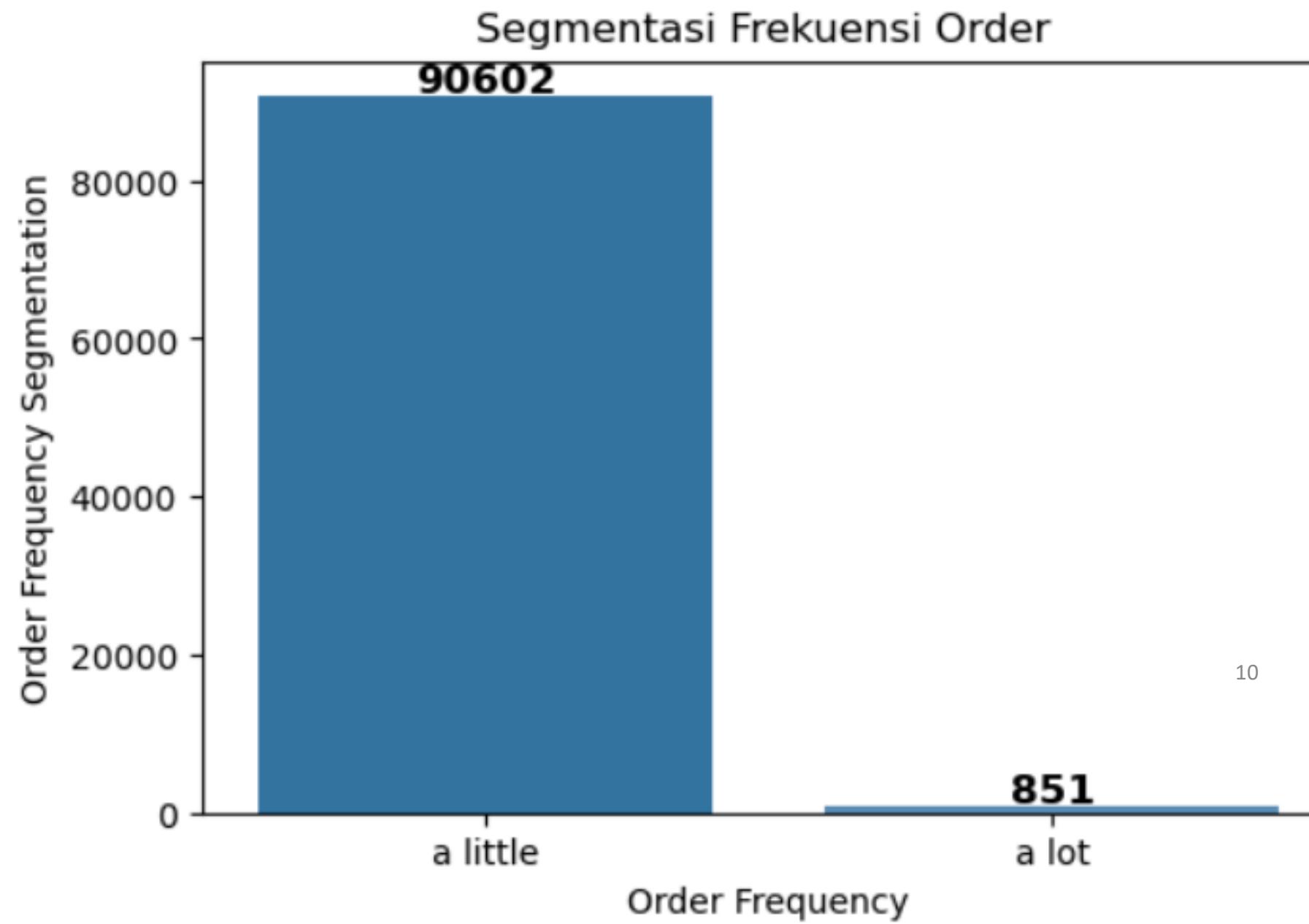


Top State with Highest Customers



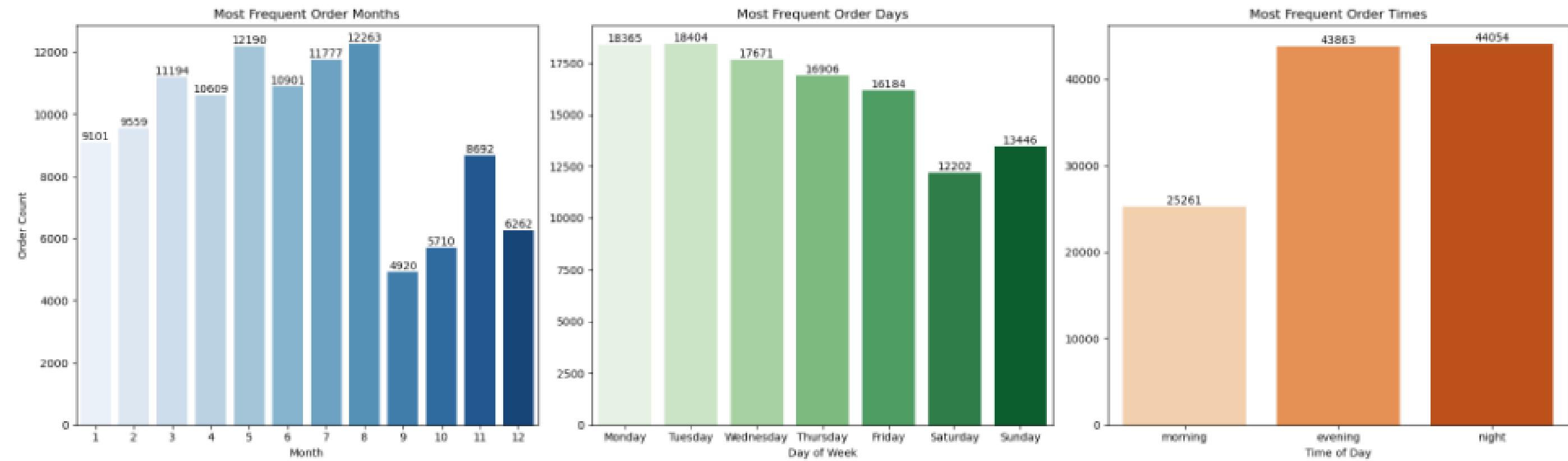
Top City with Highest Customers

Customer Frequency Segmentation

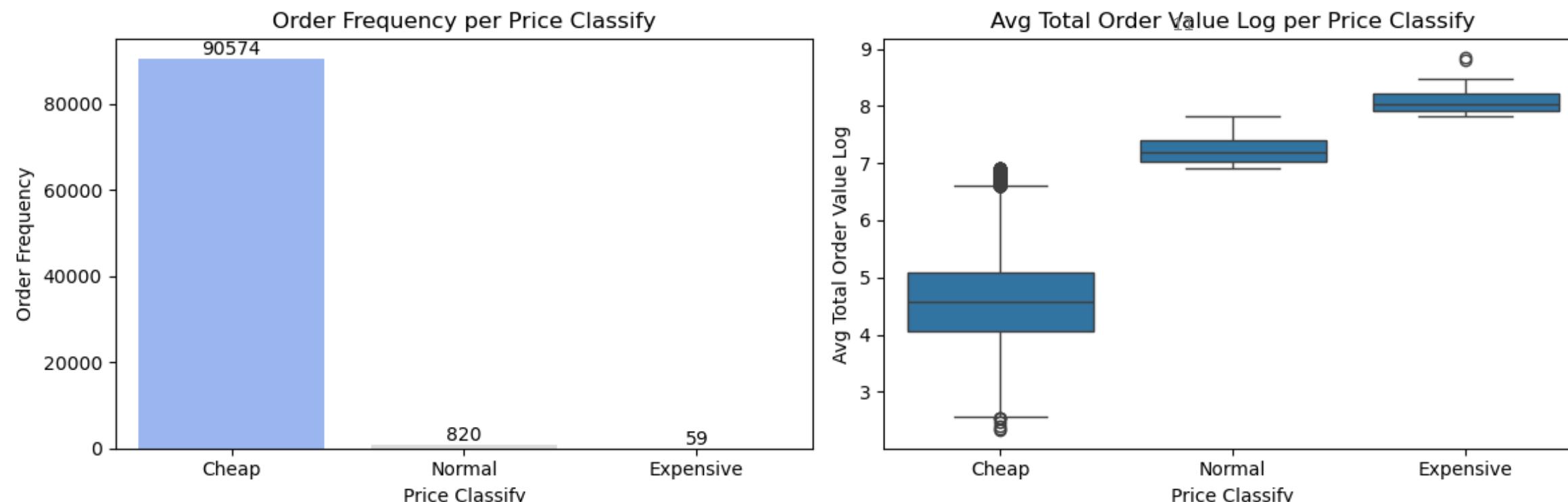


Order Behavior

The Most Frequent Order Times



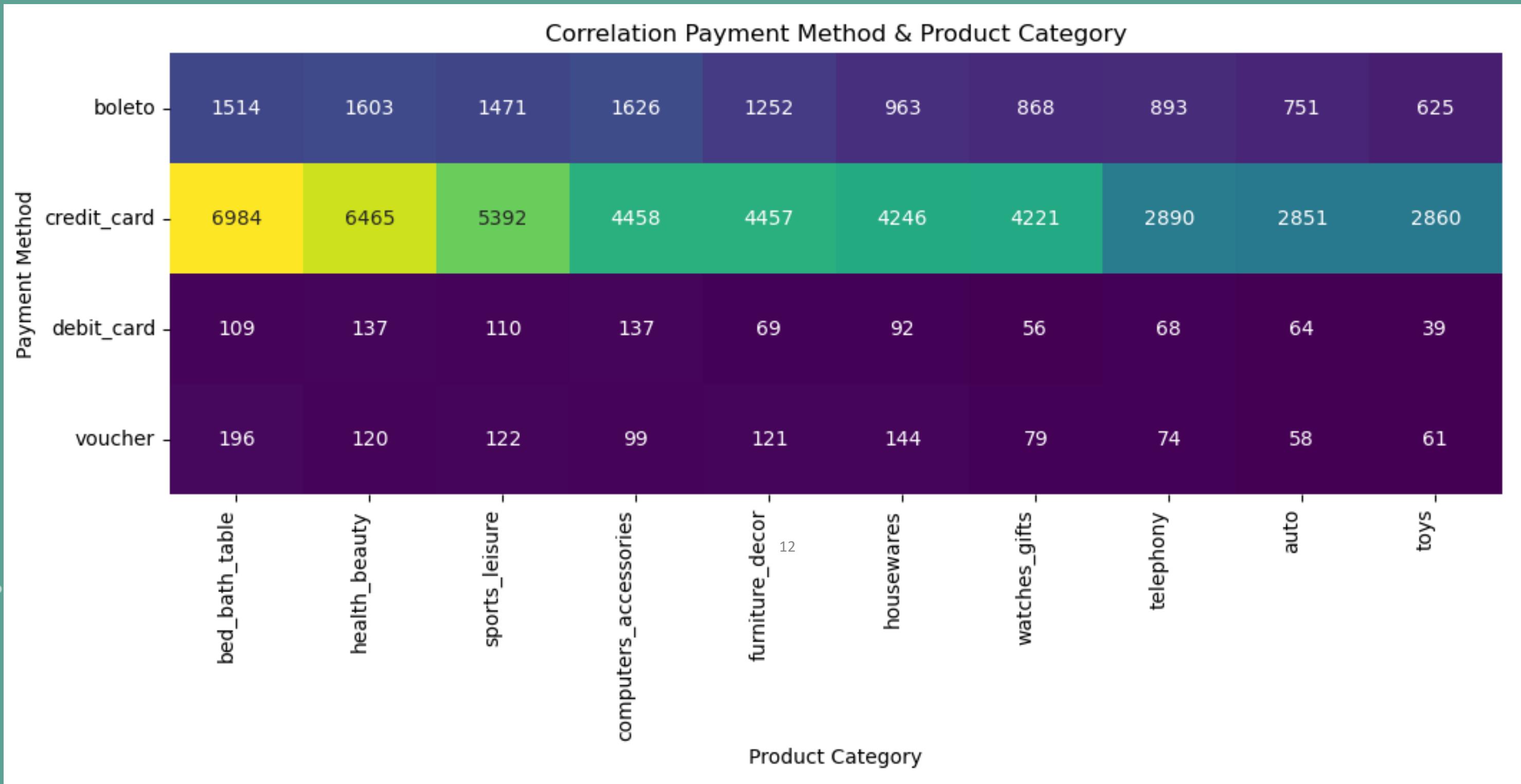
Distribution Order Frequency by Price Classify

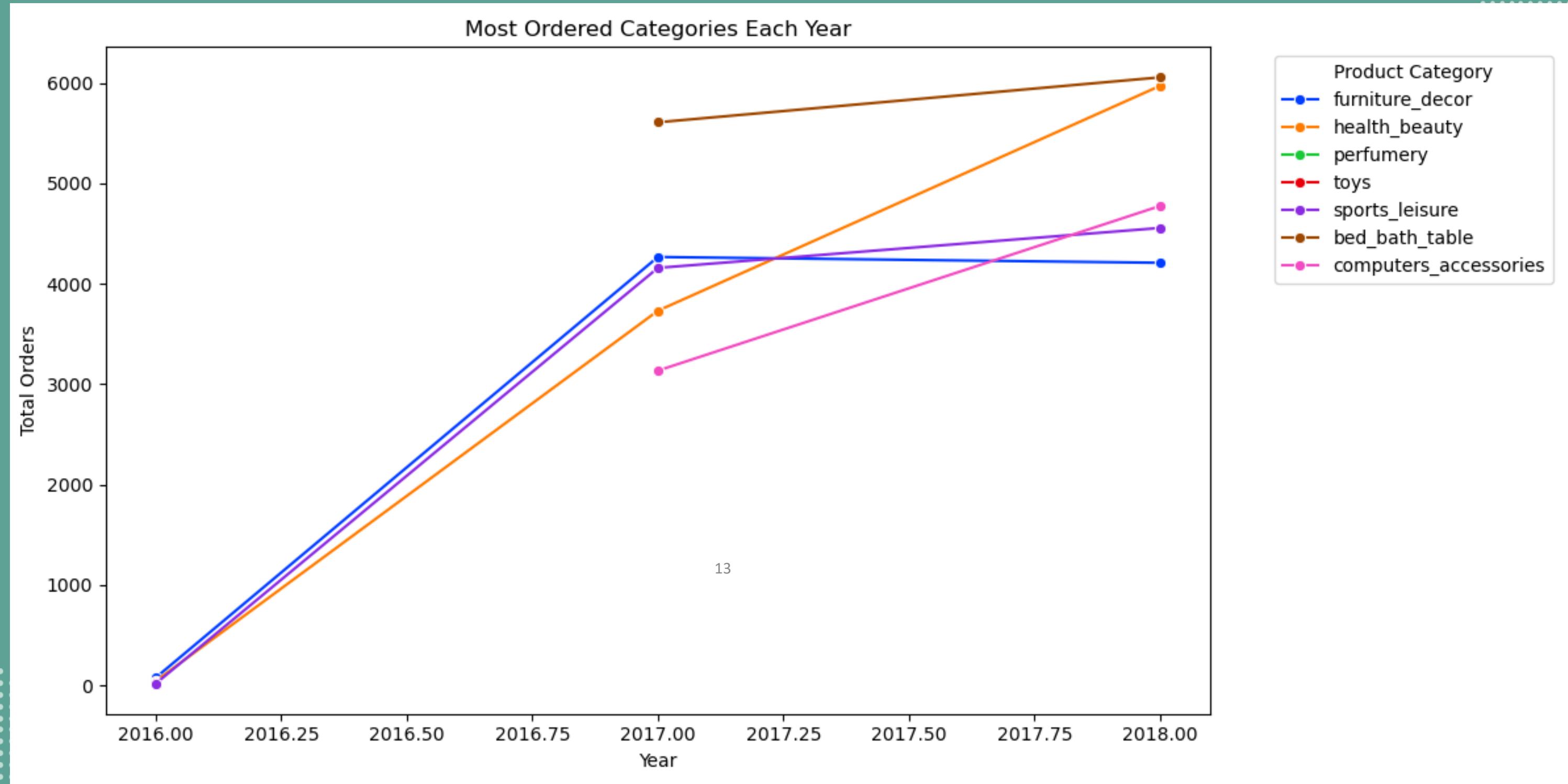


price_classify	132.271654	124.797070	95.220	9.341429	997.37
Cheap	3325.032627	874.499093	3042.360	2512.530000	6929.31
Normal	1420.627555	362.094795	1323.545	1002.710000	2467.33

FINAL PROJECT

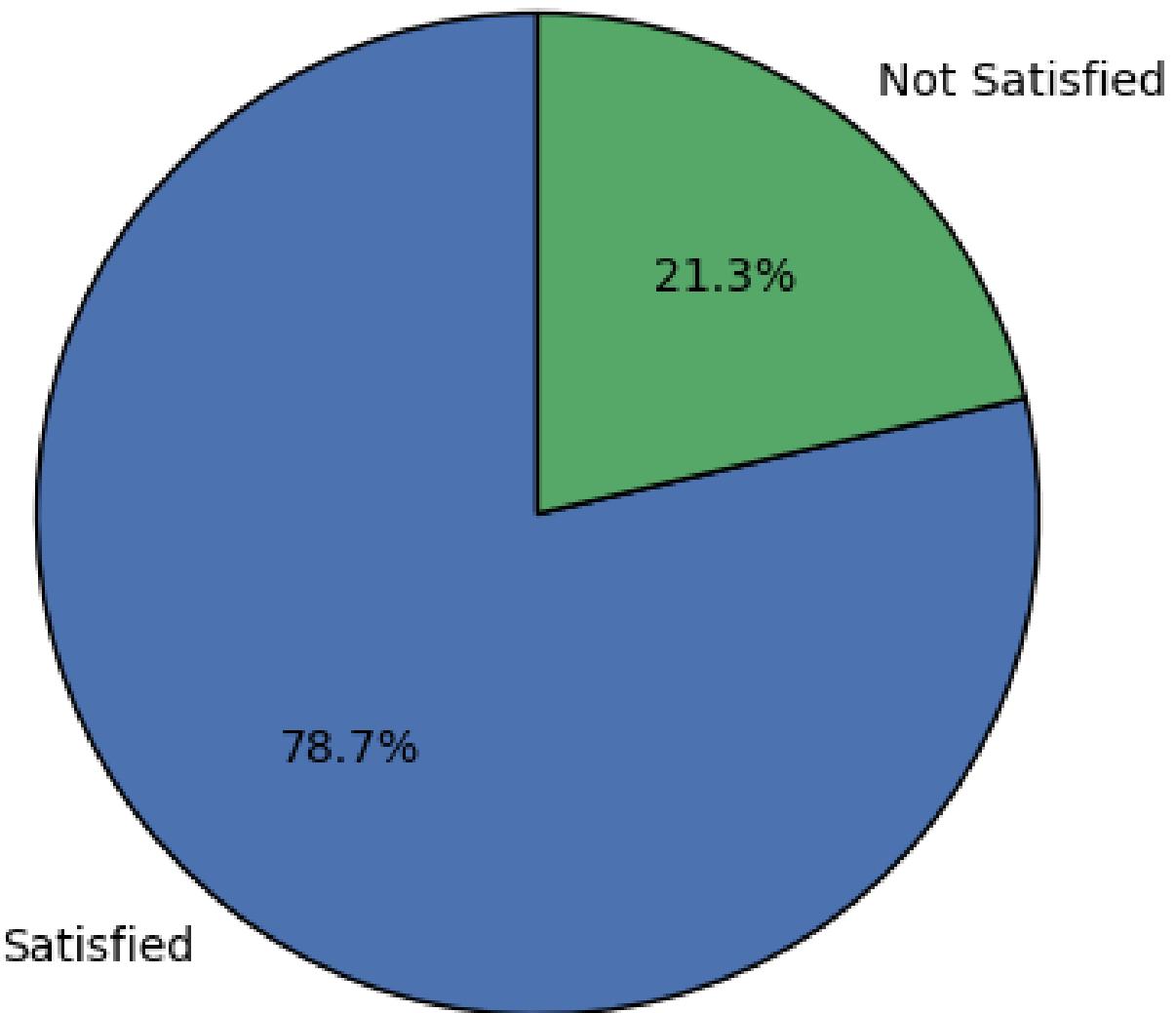
10





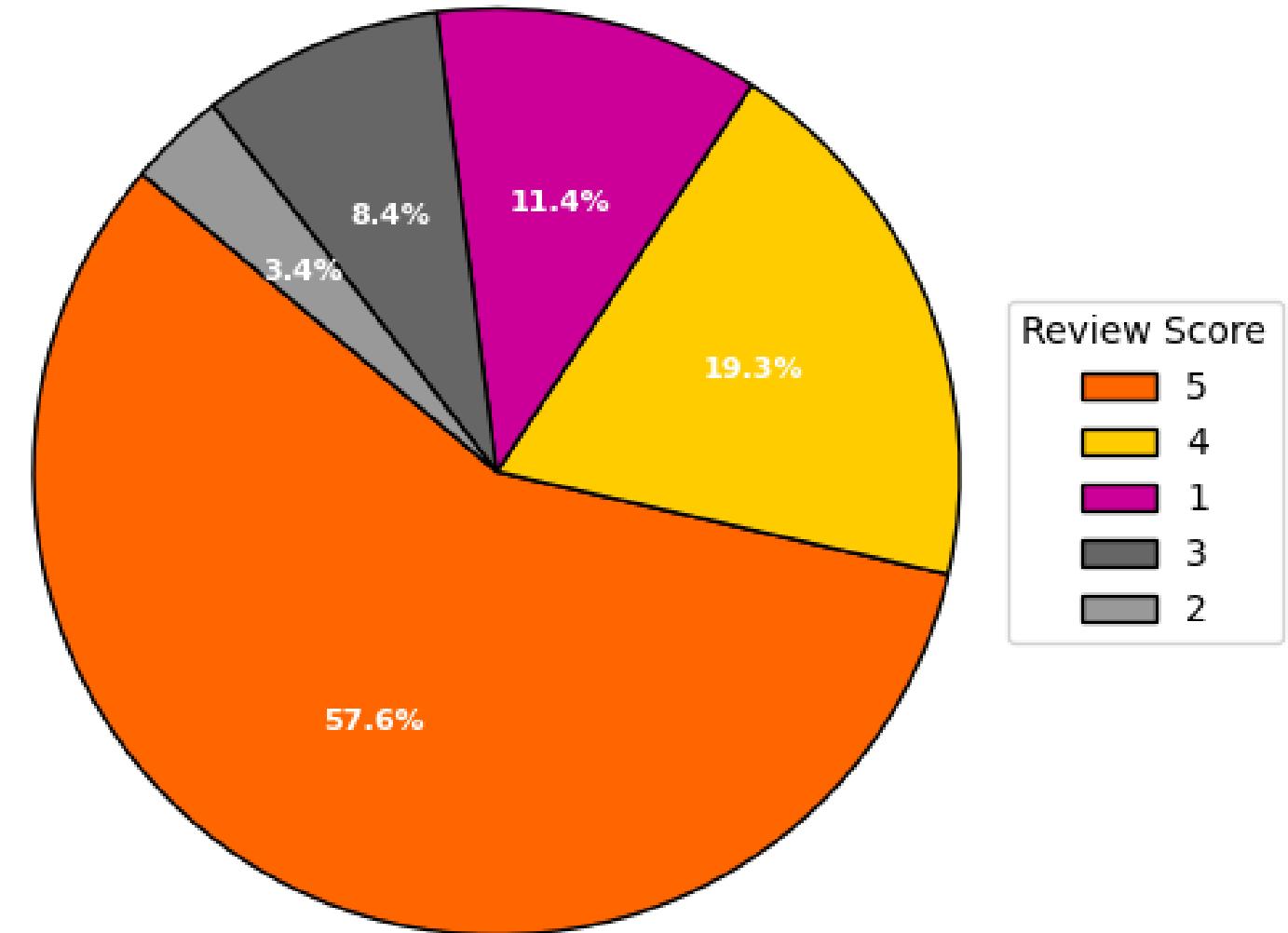
Review Customer

Customer Satisfaction Percentage

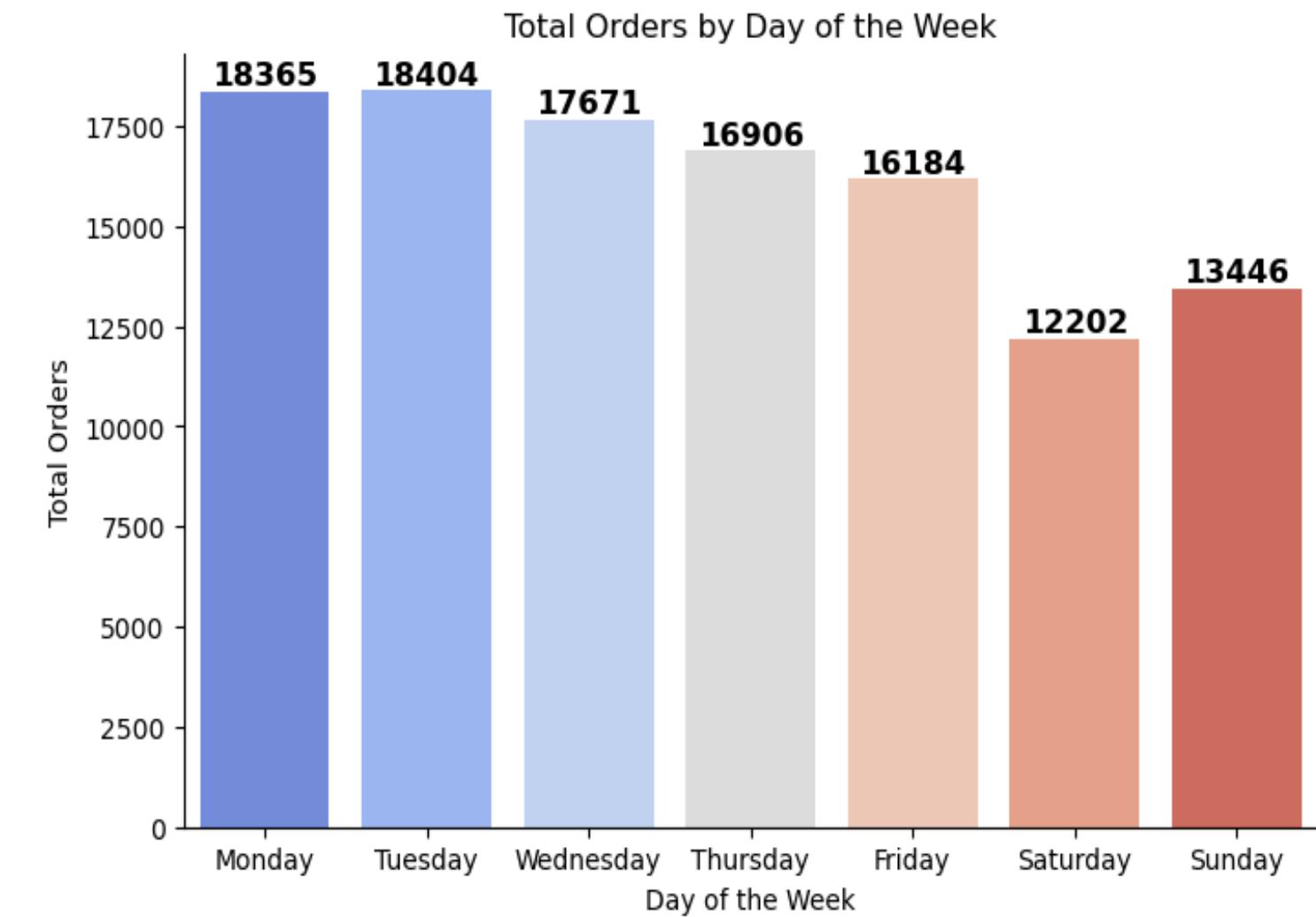
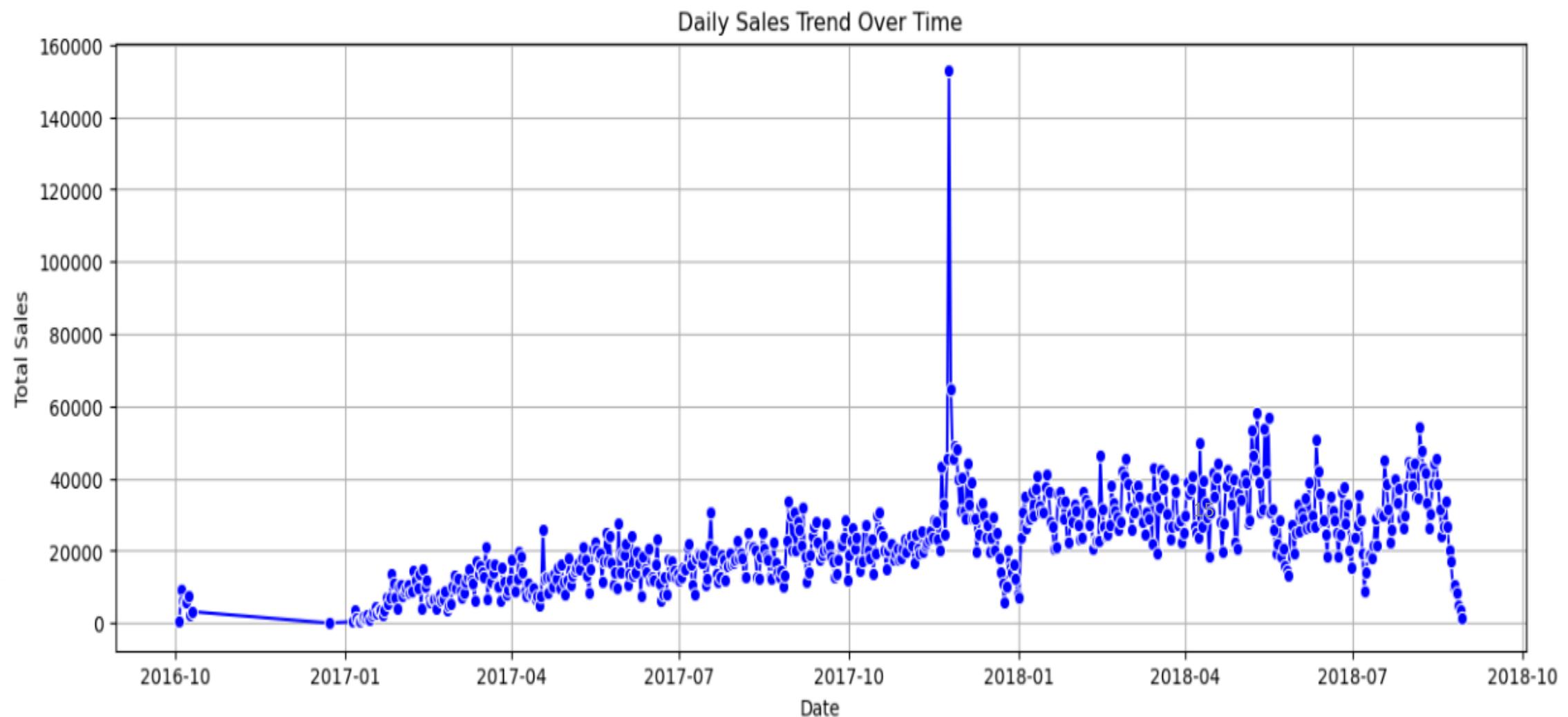


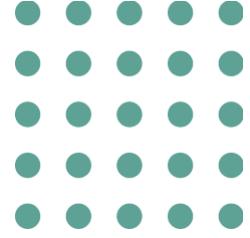
14

Persentase Review Score



Time Series Analysis





KESIMPULAN



01. Customer Demografi

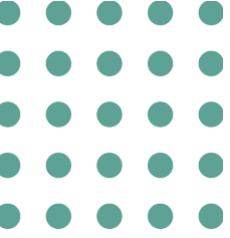
- Mayoritas customer dari kota besar: Sao Paulo, Rio de Janeiro
- Persebaran belum merata secara geografis

02. Perilaku Pemesanan

- Sebagian besar customer hanya melakukan 1 kali pembelian.
- Waktu pembelian tertinggi: bulan Agustus, Senin–Rabu, dan Sore hari – Malam hari.

16 03. Kualitas Produk & Layanan

- Customer puas: skor review 4 dan 5 mendominasi.
- Sekitar 12% customer memberi skor 1-2, menandakan ada area untuk perbaikan.



KESIMPULAN



04. Kategori & Harga Produk

- Produk murah lebih laris, menunjukkan customer sensitif harga.
- Kategori yang meningkat tiap tahun: `bad_bath_table`, `health_beauty`, `furniture_decor`.

05. Metode Pembayaran

- Kartu kredit adalah metode paling populer.¹⁷
- Menunjukkan preferensi pada transaksi cepat dan fleksibel.

Recommendations/ Suggestions

Retensi Customer

- Pancing Repeat Order ke customer 1x transaksi.
- Buat program loyalitas/reward untuk customer aktif.

Prioritas Target Pasar

- Fokus ke wilayah dengan banyak customer.

¹⁸Optimalisasi Momen Musiman

- Maksimalkan momen seperti Adanya Lonjakan Pada Akhir Tahun 2017.
- Persiapkan stok & logistik lebih awal.

Recommendations/ Suggestions

Evaluasi Layanan

- Analisis ulasan buruk untuk deteksi masalah produk/penjual.
- Tingkatkan kontrol kualitas & layanan pelanggan.

Segmentasi Harga

- Tambahkan produk terjangkau namun tetap berkualitas.
- Fokus pada kategori populer
- Tawarkan **bundling** atau **upsell** untuk tingkatkan transaksi.

19

Operasional & Pembayaran

- Fokus pada jam/hari sibuk: Senin-Rabu, Sore hari – Malam hari.
- Sediakan beragam metode pembayaran



Terima Kasih