

# RETAIL COMPANY ANALYSIS

Fabiana Arca Cruz Tortorelli

# INTRODUCTION

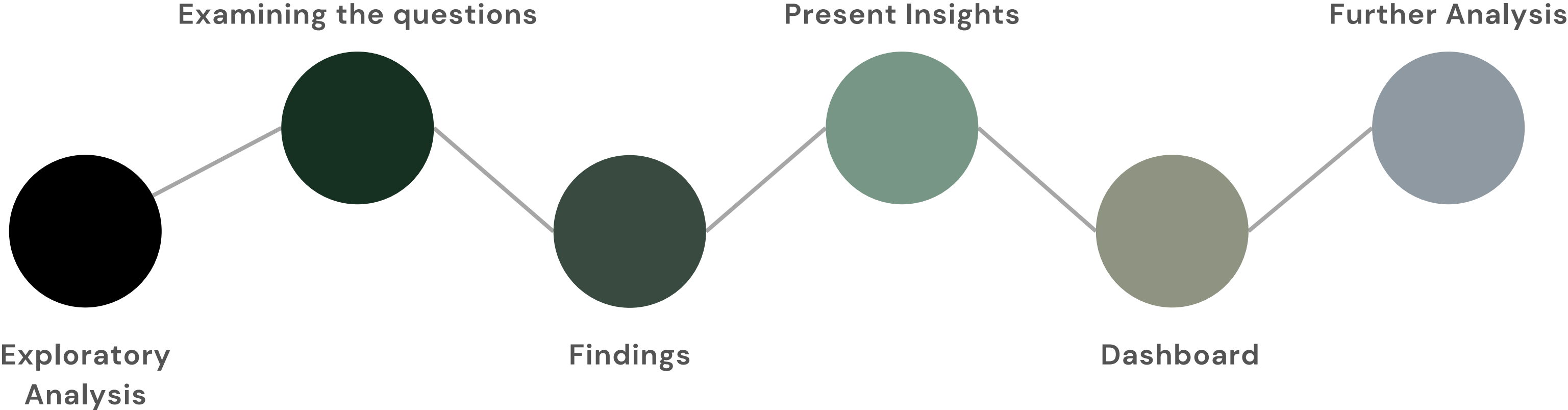


Retail companies are businesses that sell goods and services directly to consumers. They play a crucial role in the supply chain by providing a range of products to the end user, often through physical stores or online platforms. Retailers can range from small, specialized boutiques to large, multinational chains. Their primary functions include purchasing goods from manufacturers or wholesalers, managing inventory, and selling to consumers at various price points.

The **goal** of this work is to answer the following 10 business questions about a dataset of a Retail Company located in EUA.

- Which City has the Highest Sales Value for Products in the 'Office Supplies' Category?
- What is the Total Sales by Order Date?
- What is the Total Sales by State?
- What are the Top 10 Cities with the Highest Total Sales?
- Which Segment Had the Highest Total Sales?
- What is the Total Sales by Segment and by Year?
- What is the Average Sales by Segment, by Year, and by Month?
- How many Sales would receive a 15% discount based on the following rule?
  - If the Sales Value is greater than 1000, it receives a 15% discount.
  - If the Sales Value is less than 1000, it receives a 10% discount.
- Assume the company decides to grant a 15% discount from the previous item. What would be the average Sales Value before and after the discount?
- What is the Total Sales by Category and SubCategory, Considering Only the Top 12 SubCategories?

# PROJECT TIMELINE



# EXPLORATORY ANALYSIS

# DATA SUMMARY

The client dataset contains 9,700 observations and 11 features.

All rows are complete, with no missing or duplicated values.

The data types of each feature are consistent with what they represent.

## NUMERICAL FEATURE



Sales Value

## CATEGORICAL FEATURES



Order ID



Order Date



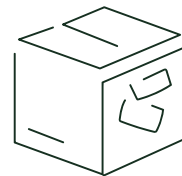
Customer ID



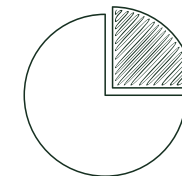
Country

State

City



Product ID



Segment



Category

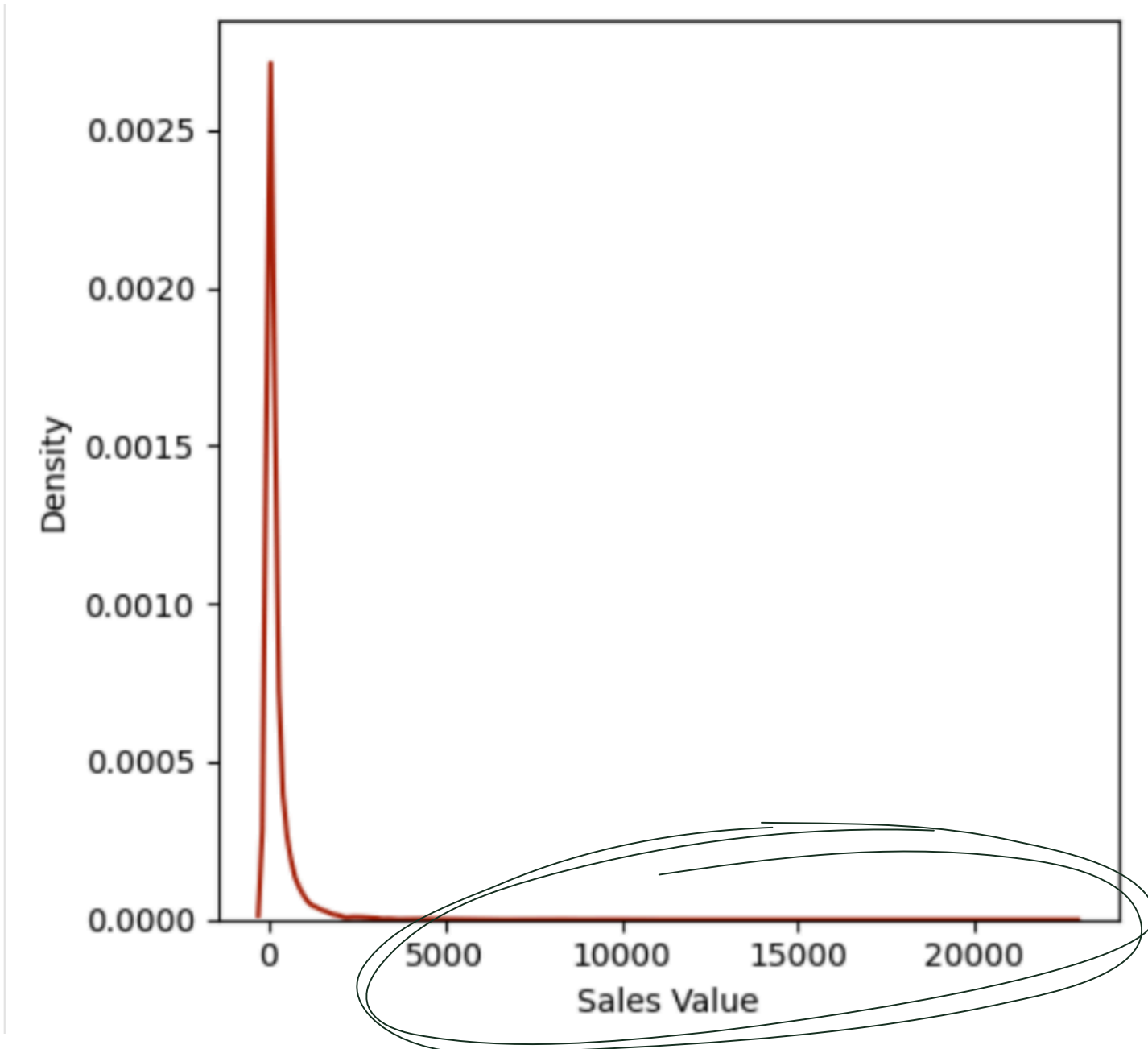


Sub Category

## NUMERICAL FEATURE

# DATA SUMMARY

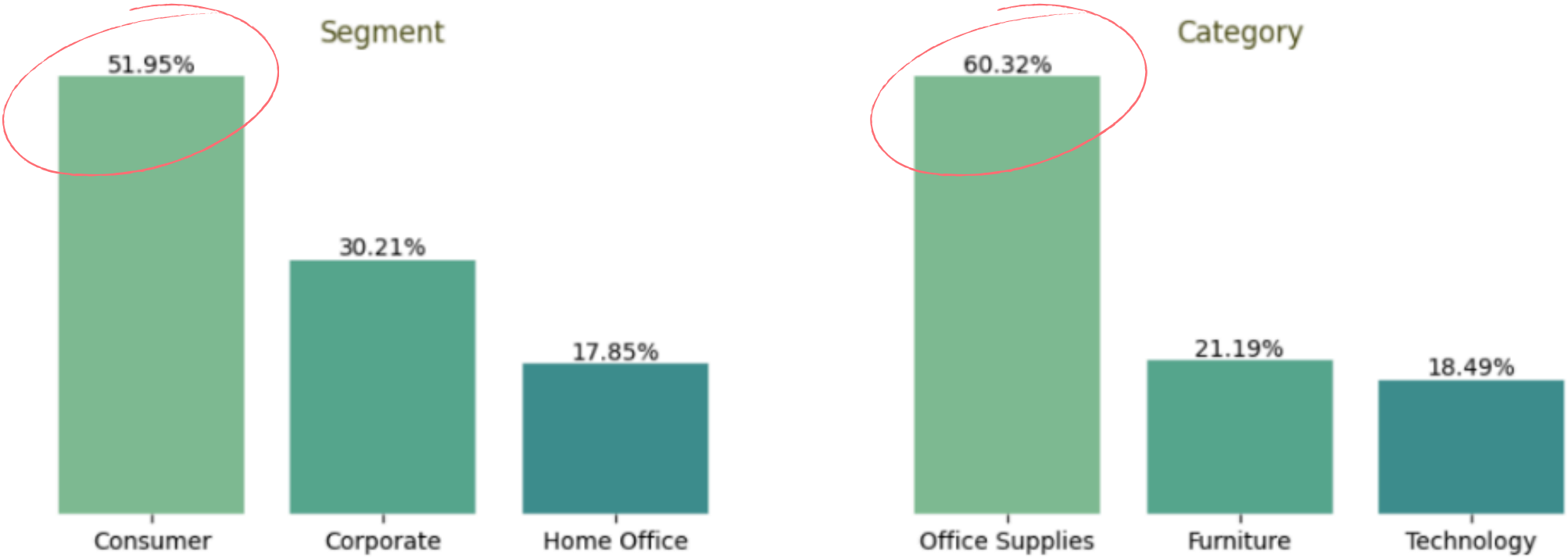
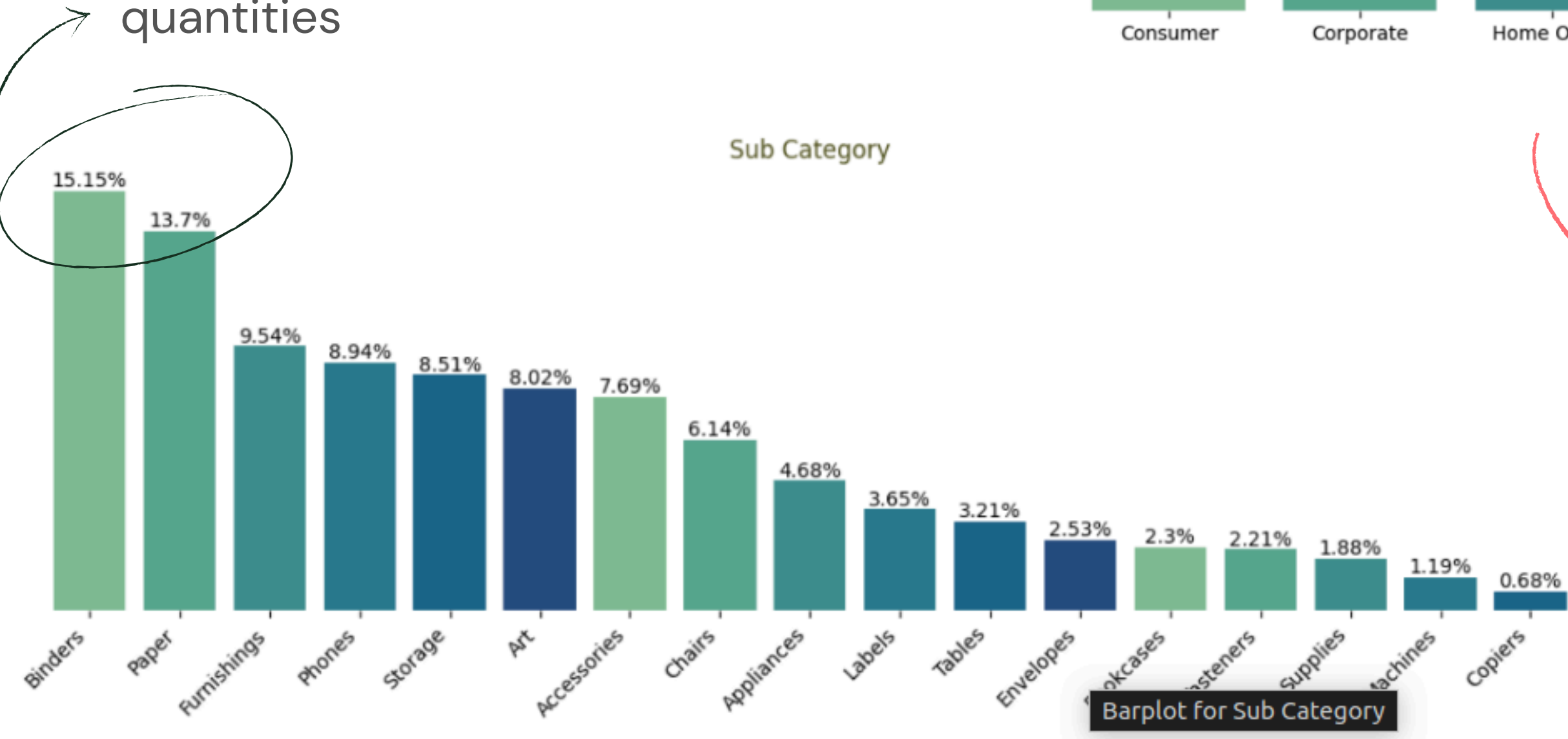
- The distribution is **right-skewed**, indicating that most observations are concentrated on the left, with lower values
- The **mean** is 230, while the **median** is 54. The **minimum** value is 0.44, and the **maximum** is 22,638.48. The **standard deviation** is 627.50
- Upon examining the higher values, they seem **reasonable**, as some customers may be purchasing large quantities of specific products



# CATEGORICAL FEATURES

## DATA SUMMARY

- Within the **Sub Category**, Binders and Paper have the highest sales quantities

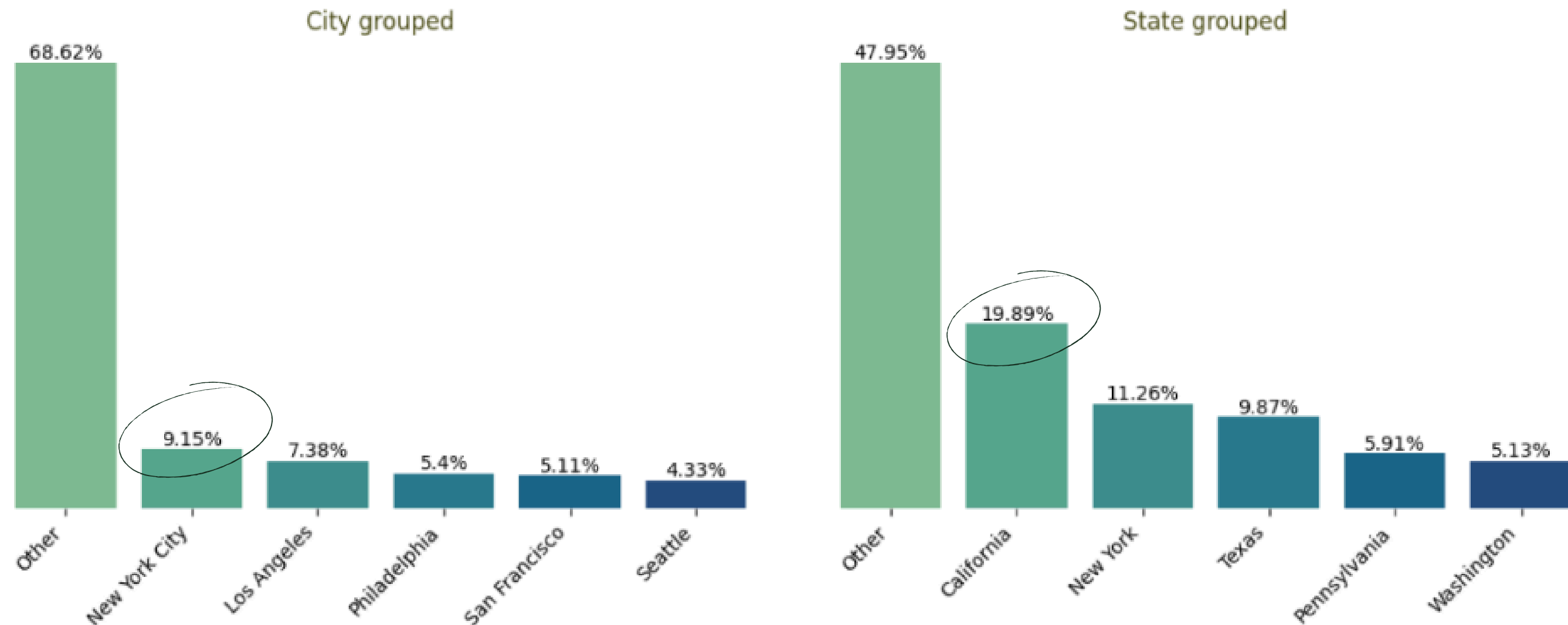


- Among the **segments**, Consumer has the highest sales, while Office Supplies is the leading **category**

## CATEGORICAL FEATURES

# DATA SUMMARY

- The feature **City** contains 528 unique values, with the highest sales recorded in New York City (9.15%)
- The feature **State** contains 49 unique values, with the highest sales recorded in California (19.89%)

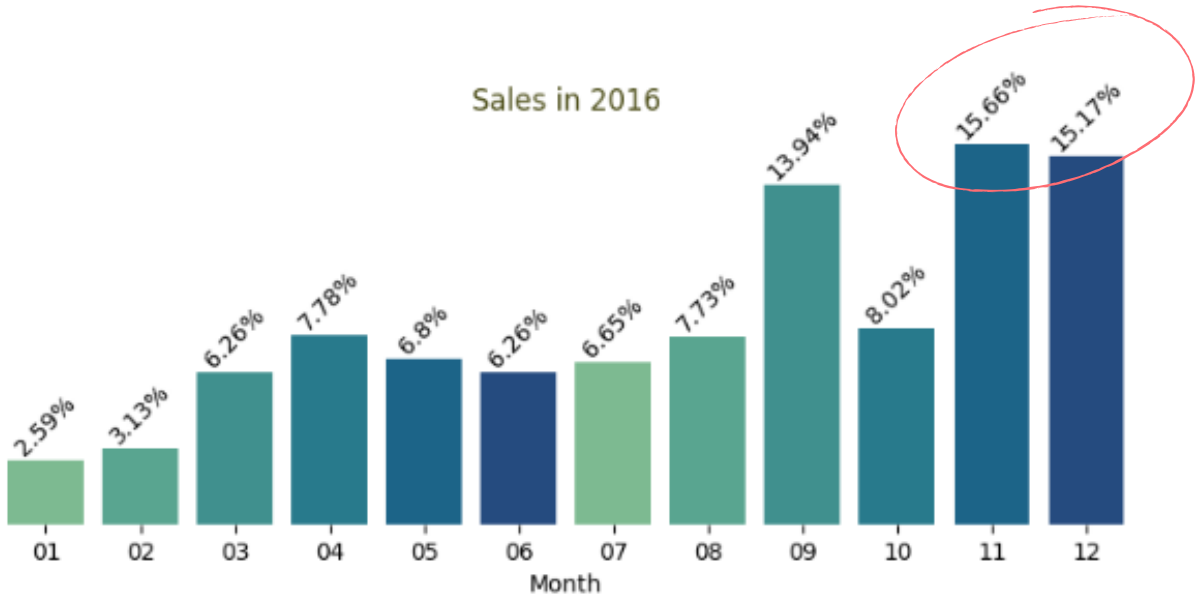
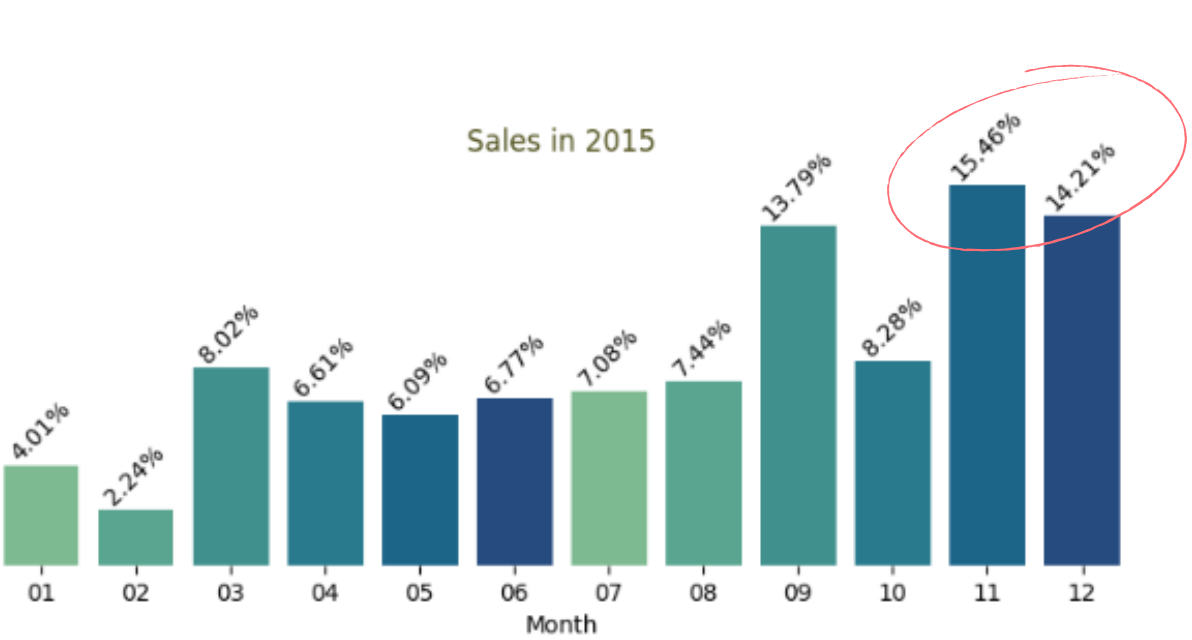
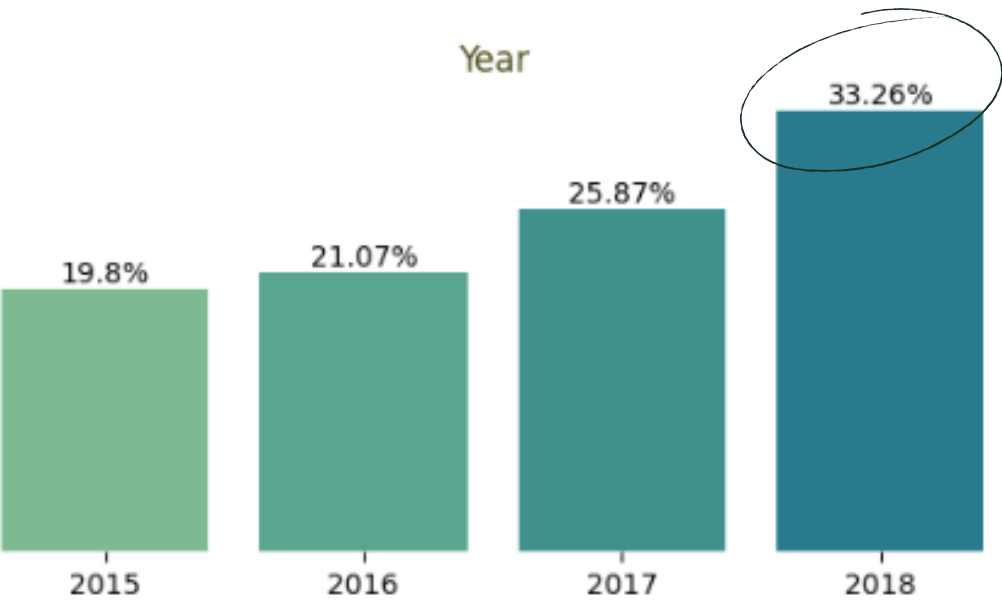




## CATEGORICAL FEATURES

# DATA SUMMARY

- For the date range from 2015 to 2018, the highest sales were recorded in 2018
- Across all years, the highest sales occur in November and December
- Sales were between 2% and 8% until August, with a notable increase to 13% in October.



# CATEGORICAL FEATURES VS SALES VALUE

Test for Significant Differences in Sales Value Means Between Groups:

We used the One-Way ANOVA method, which is employed to compare the means of three or more groups.

**H0:** The mean Sales Value is not significantly different between groups

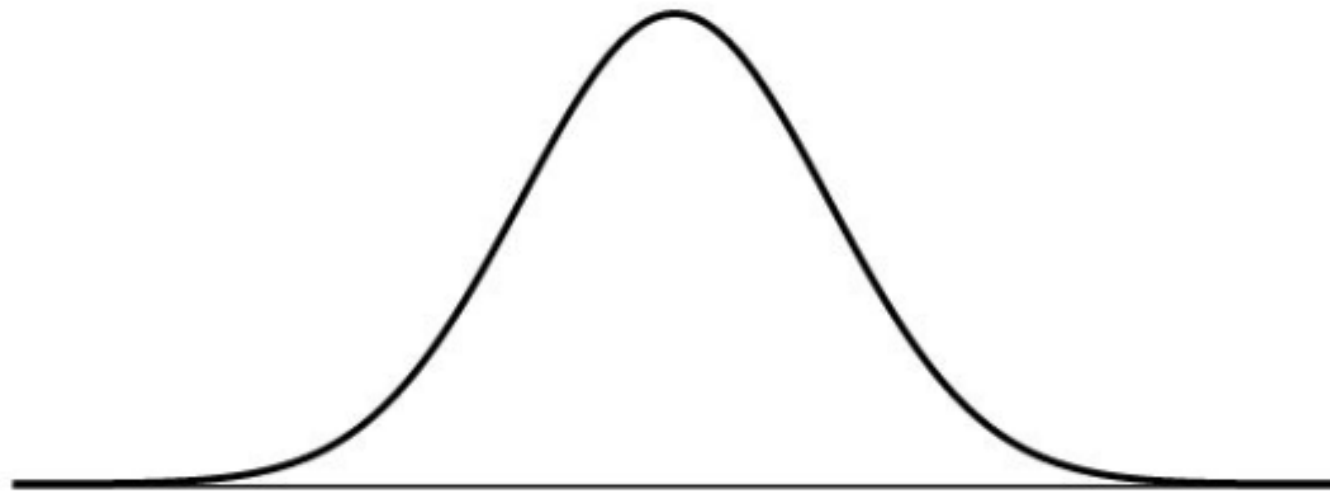
**H1:** The mean Sales Value is significantly different between groups

- The mean sales value is nearly the same across each segment
- When grouped by city, the mean sales value is almost uniform across cities
- When grouped by state, the mean sales value is similarly consistent across states
- The mean sales value for the Office Supplies category is lower than that for Furniture and Technology
- Within sub categories, the mean sales value varies, with Fasteners having the lowest and Copiers the highest



**For all tests, we do not have enough evidence to conclude that the groups are significantly different**

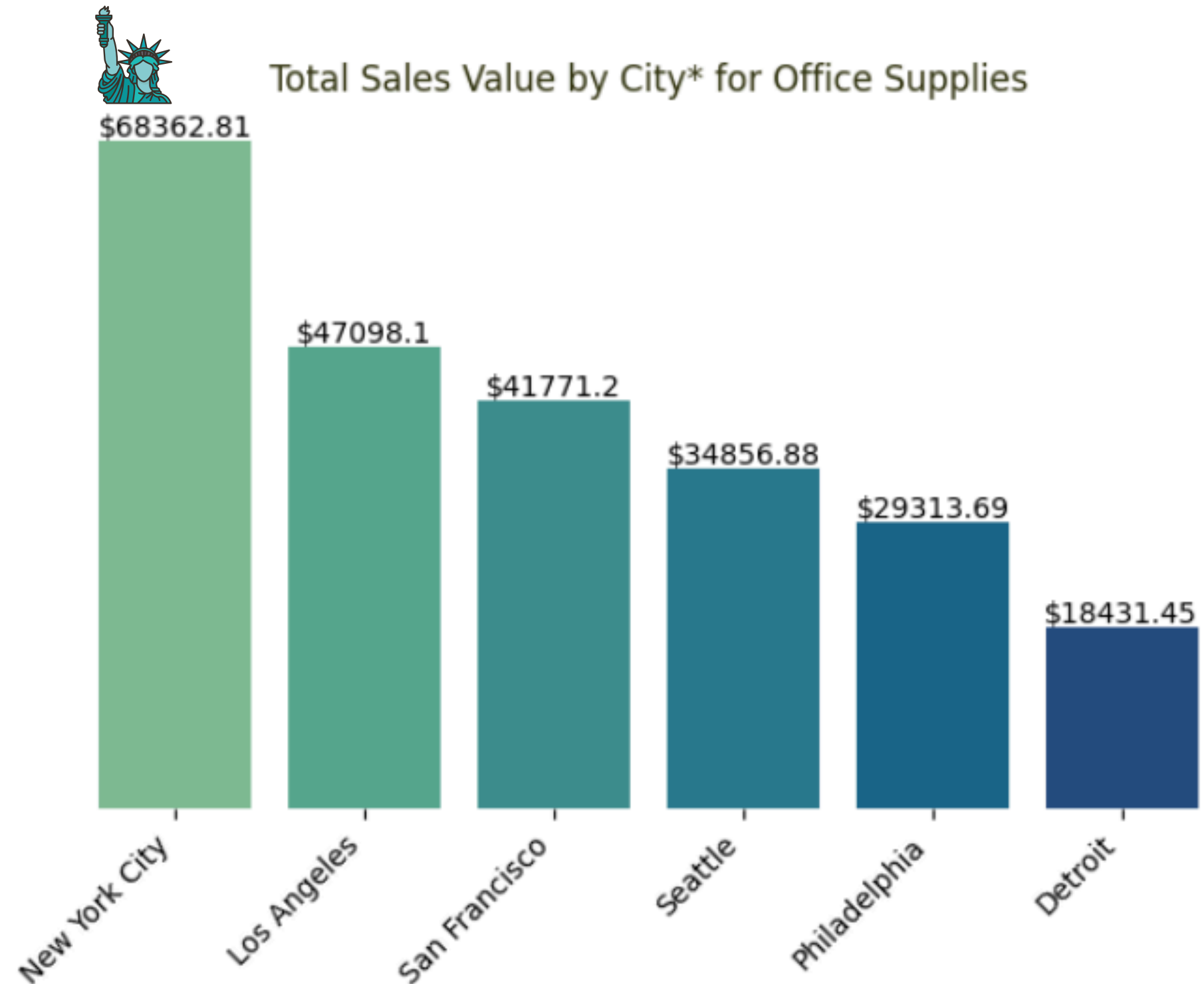
## DATA SUMMARY



# **EXAMINING THE QUESTIONS**

# WHICH CITY HAS THE HIGHEST SALES VALUE FOR PRODUCTS IN THE 'OFFICE SUPPLIES' CATEGORY?

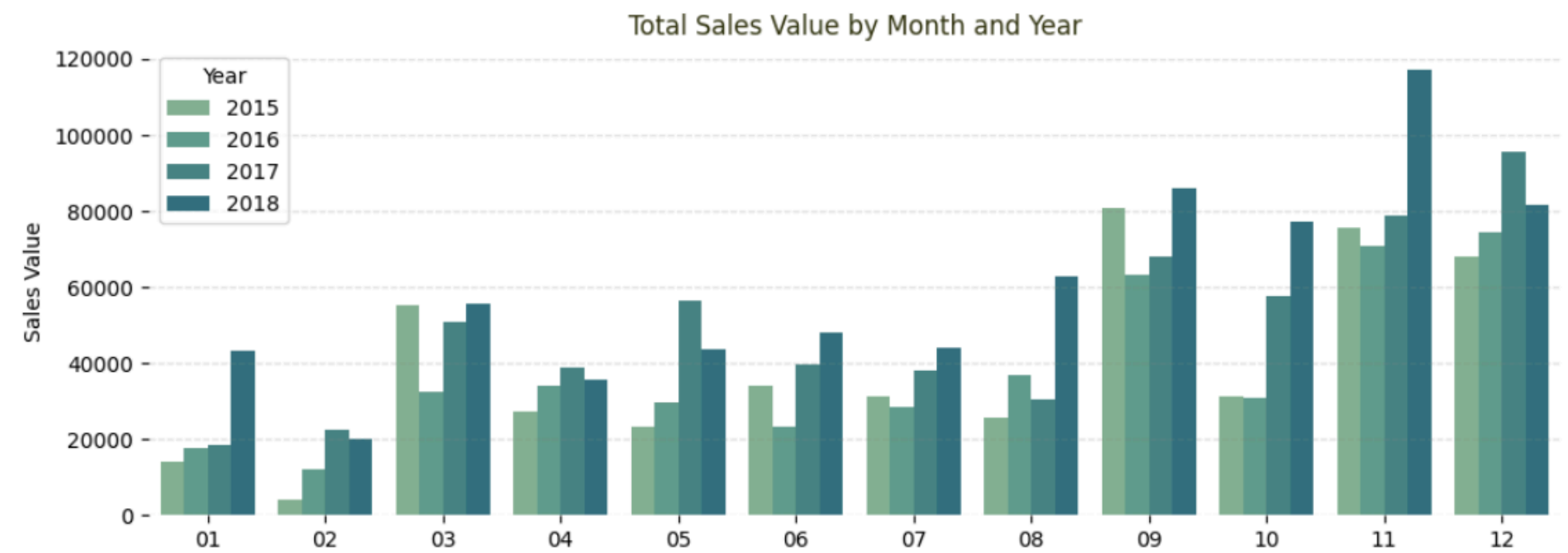
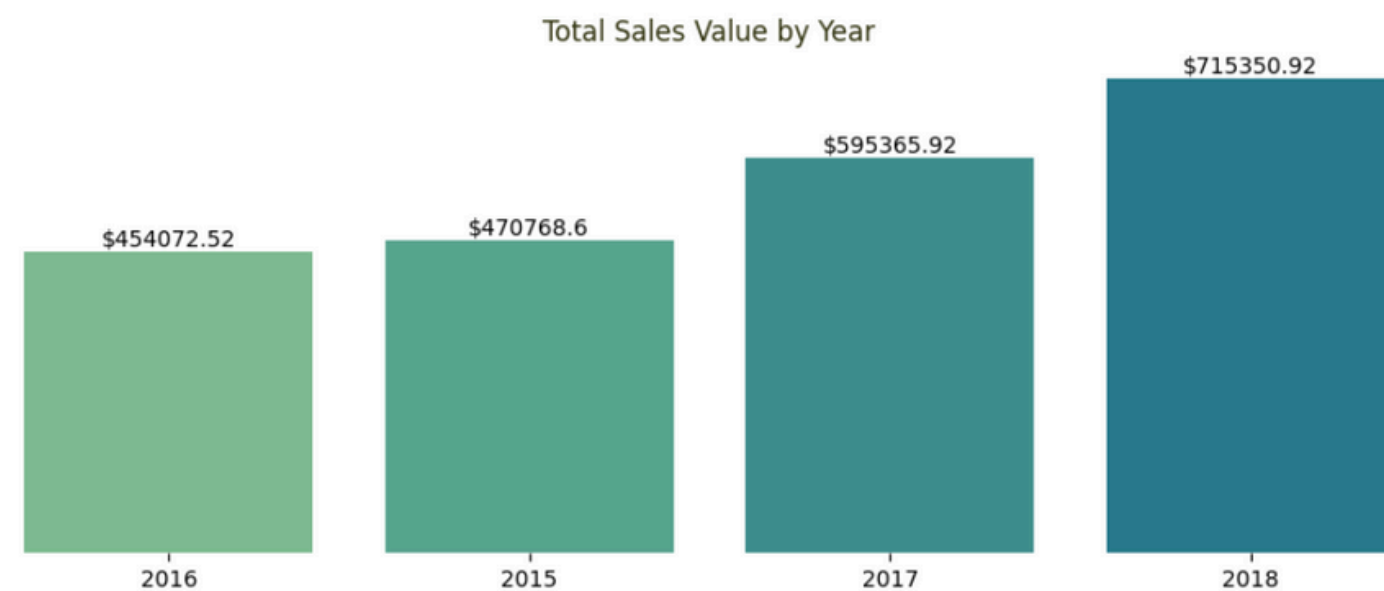
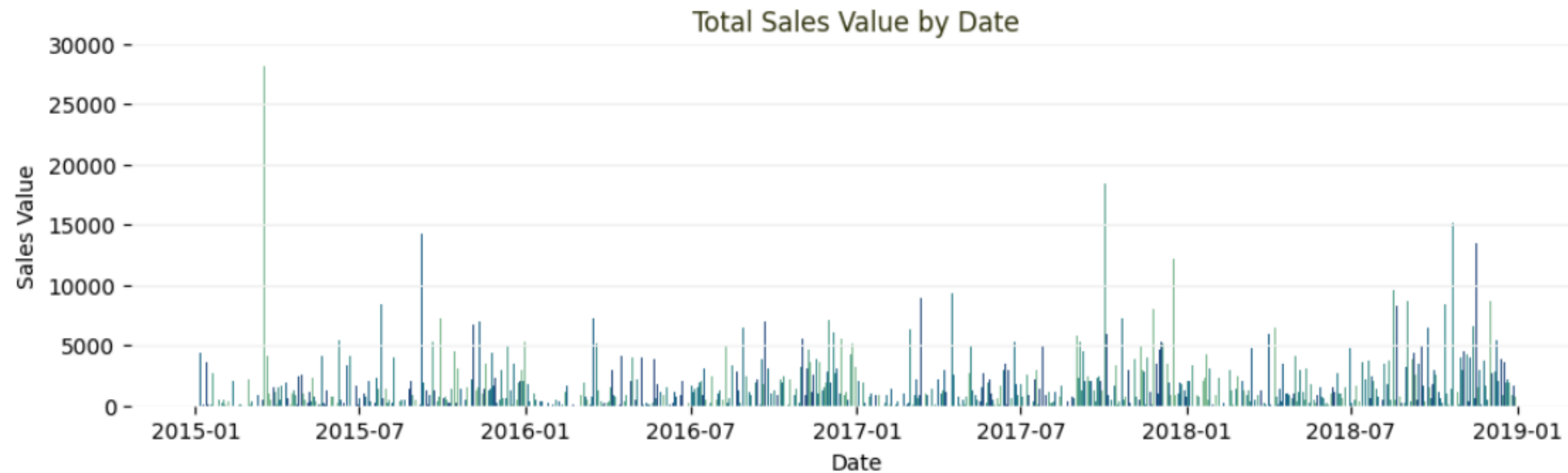
- New York City has the highest sales value for products in the Office Supplies category



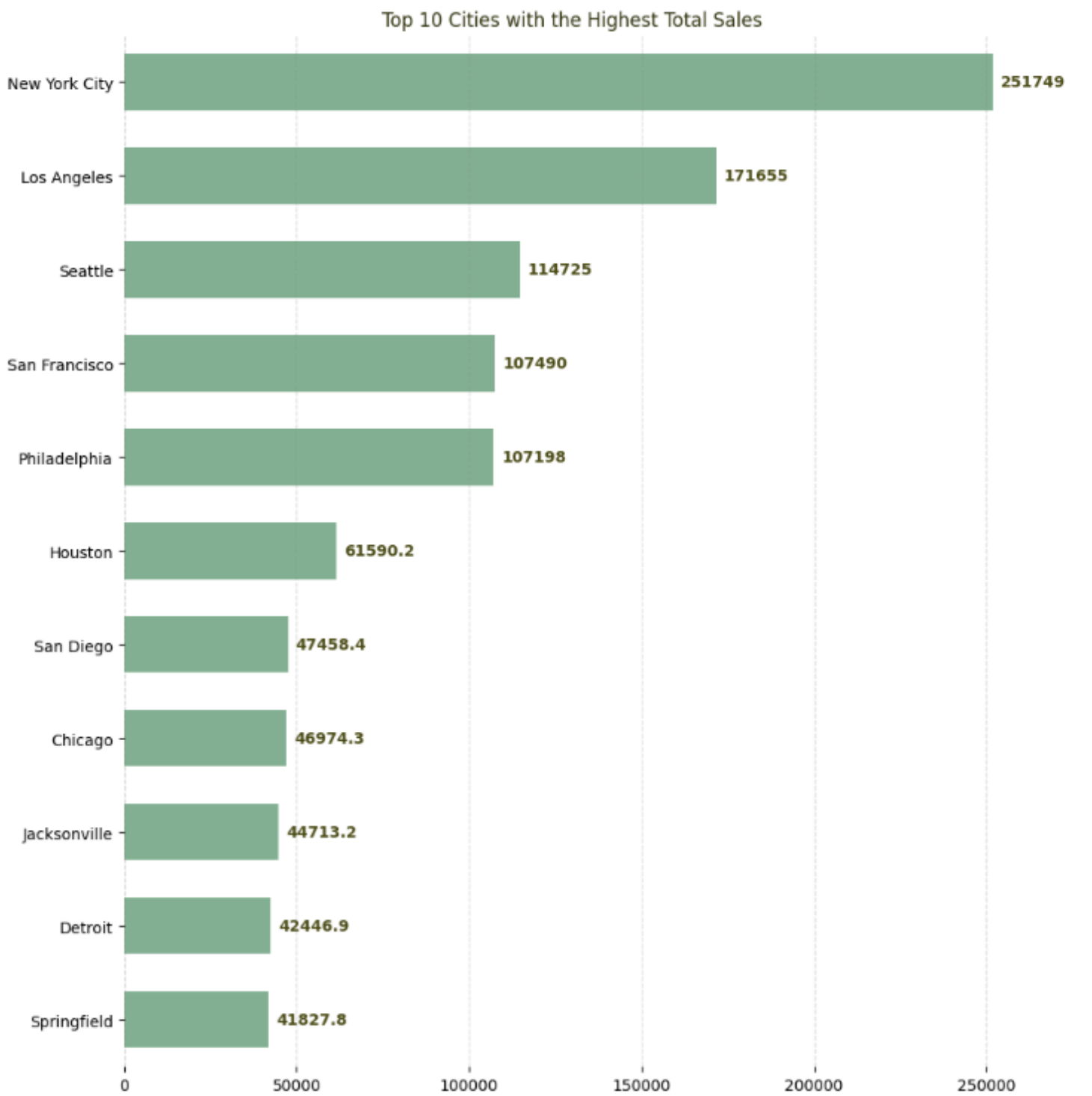
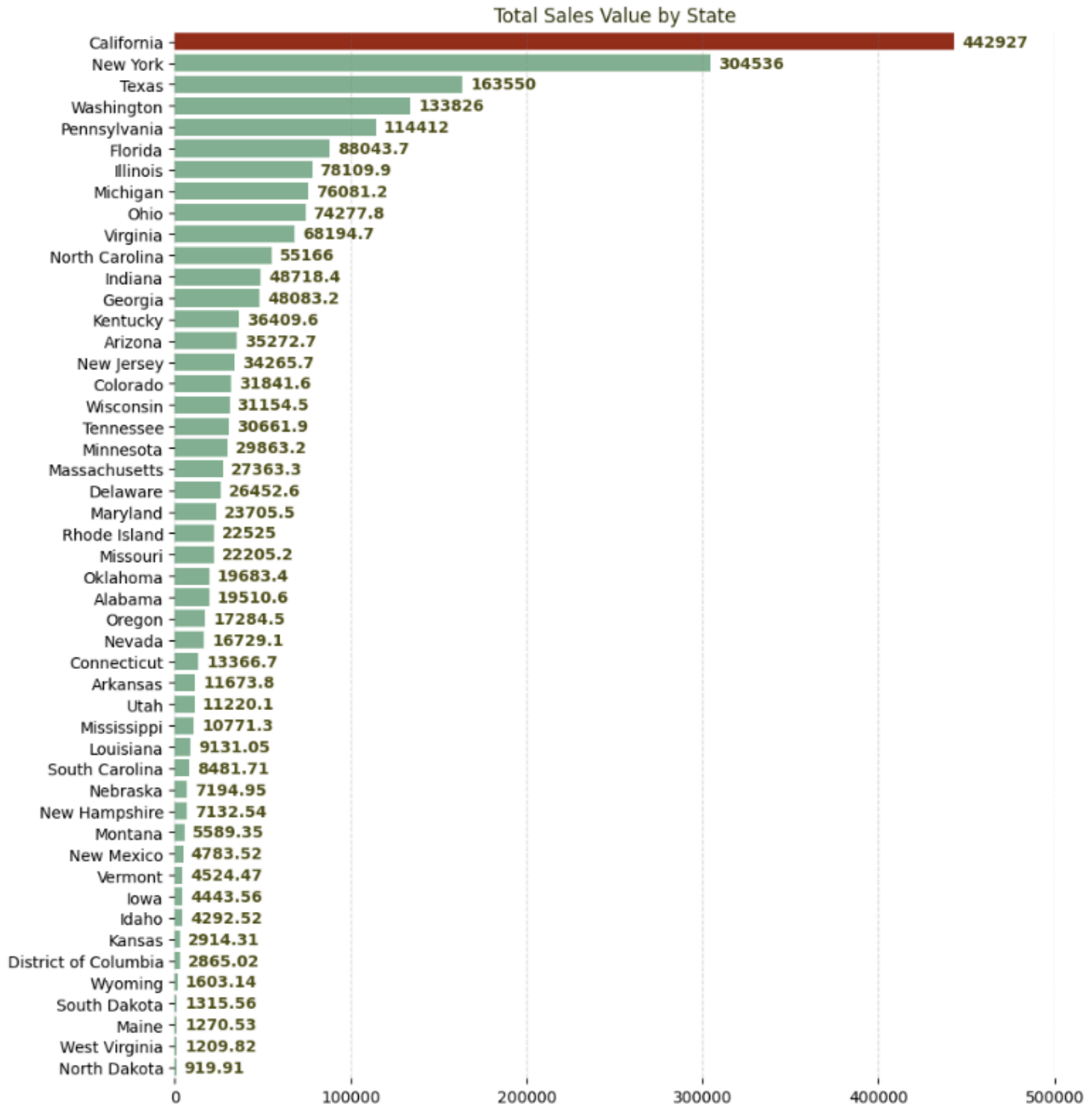
\*The graph shows the six cities with the highest sales values

## WHAT IS THE TOTAL SALES BY ORDER DATE?

- We observed that sales values have increased over the years
- March 2015 experienced a peak in sales, but this trend did not recur in subsequent years

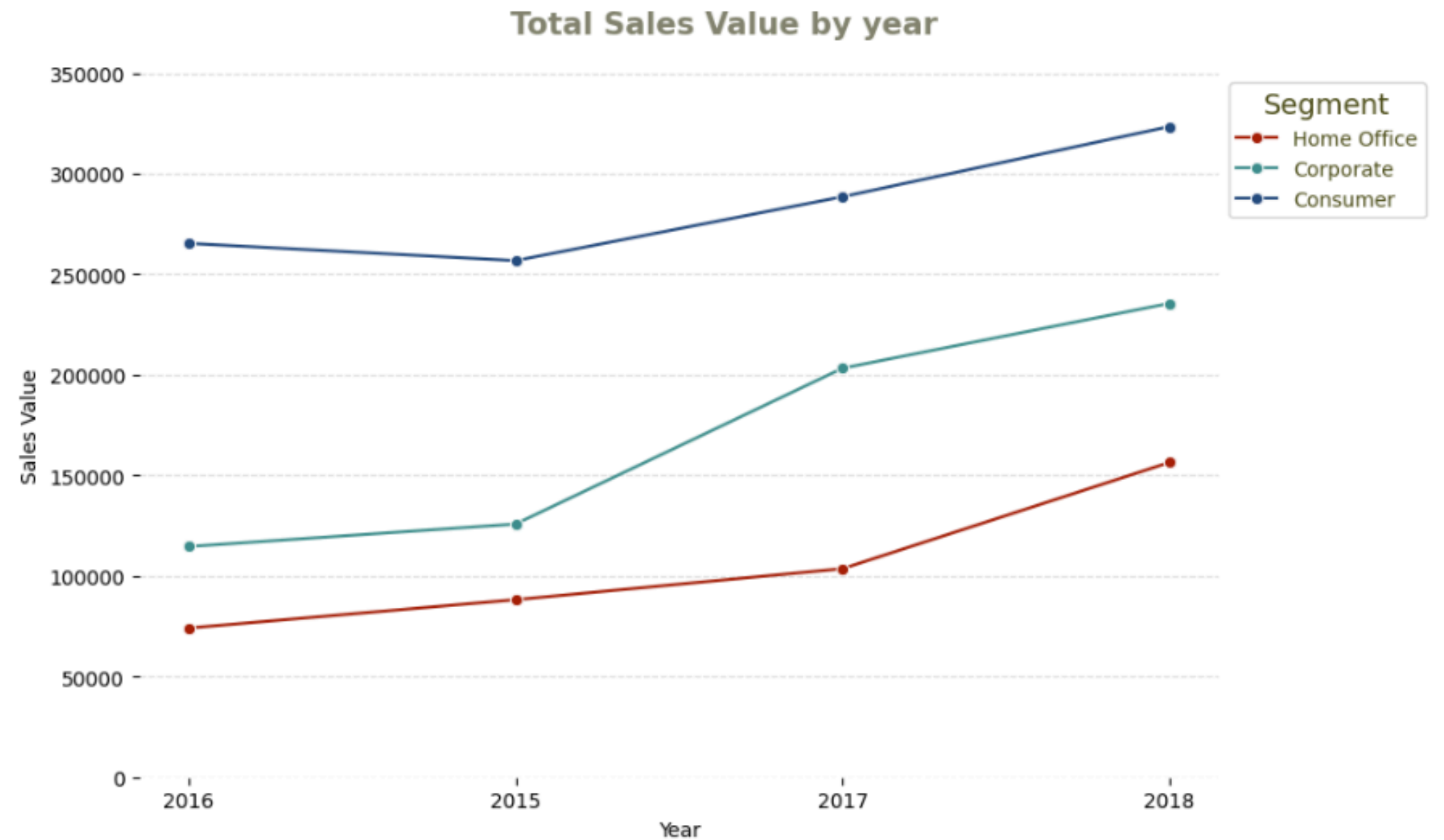
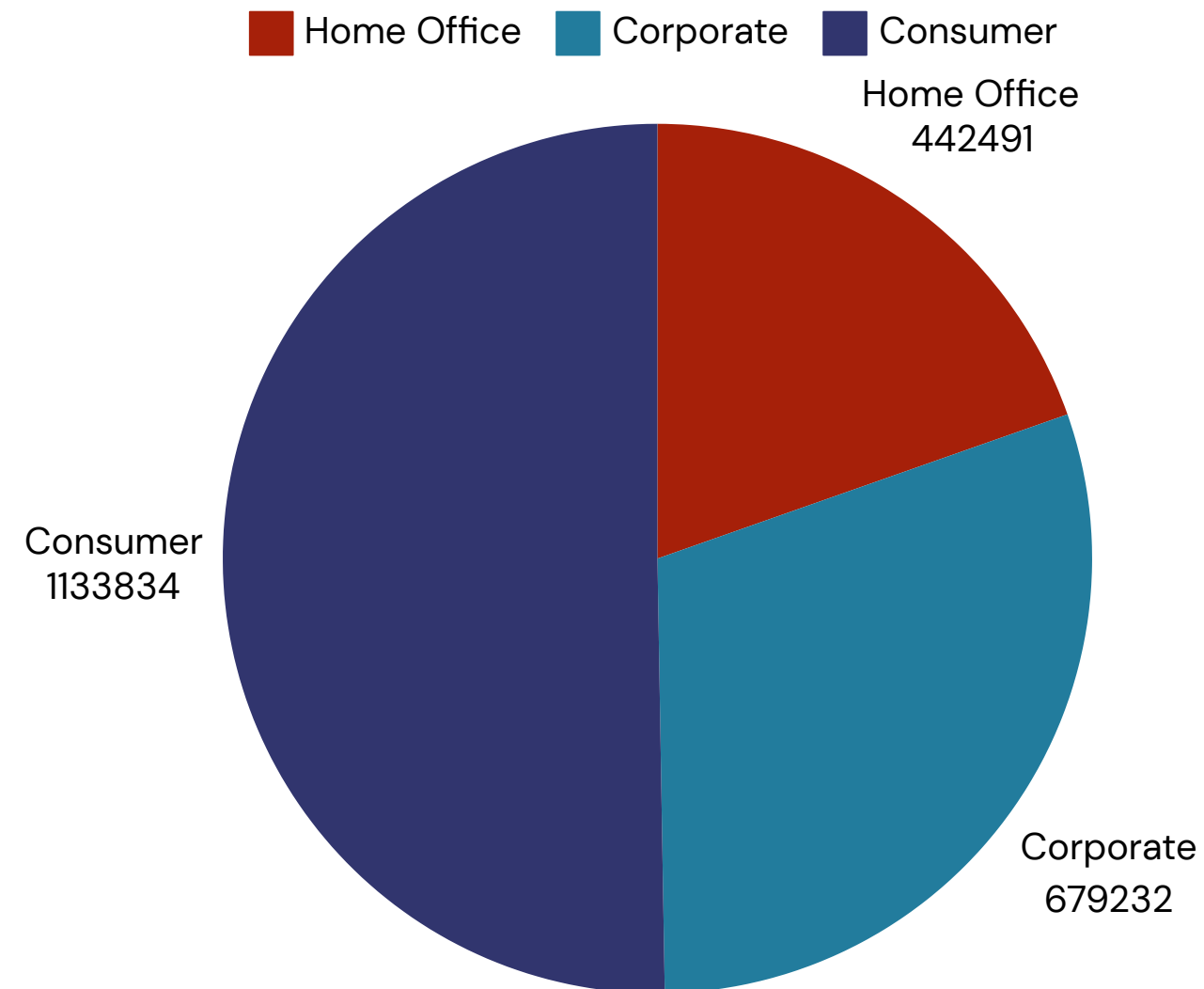


# WHAT IS THE TOTAL SALES BY STATE? WHAT ARE THE TOP 10 CITIES WITH THE HIGHEST TOTAL SALES?



## WHICH SEGMENT HAD THE HIGHEST TOTAL SALES? WHAT IS THE TOTAL SALES BY SEGMENT AND BY YEAR?

- The Consumer segment had the highest total sales
- Over the years, this segment consistently recorded the highest total sales



## DISCOUNT ANALYSIS?

- The company managers are considering offering different discount tiers and would like to run a simulation based on the following rule:
  - If the Sales Value is greater than 1000, it receives a 15% discount.
  - If the Sales Value is less than 1000, it receives a 10% discount.

How many Sales would receive a 15% discount?



**457 (4,72%) Sales would receive 15% discount**

- Assume the company decides to grant a 15% discount from the previous item. What would be the average Sales Value before and after the discount?





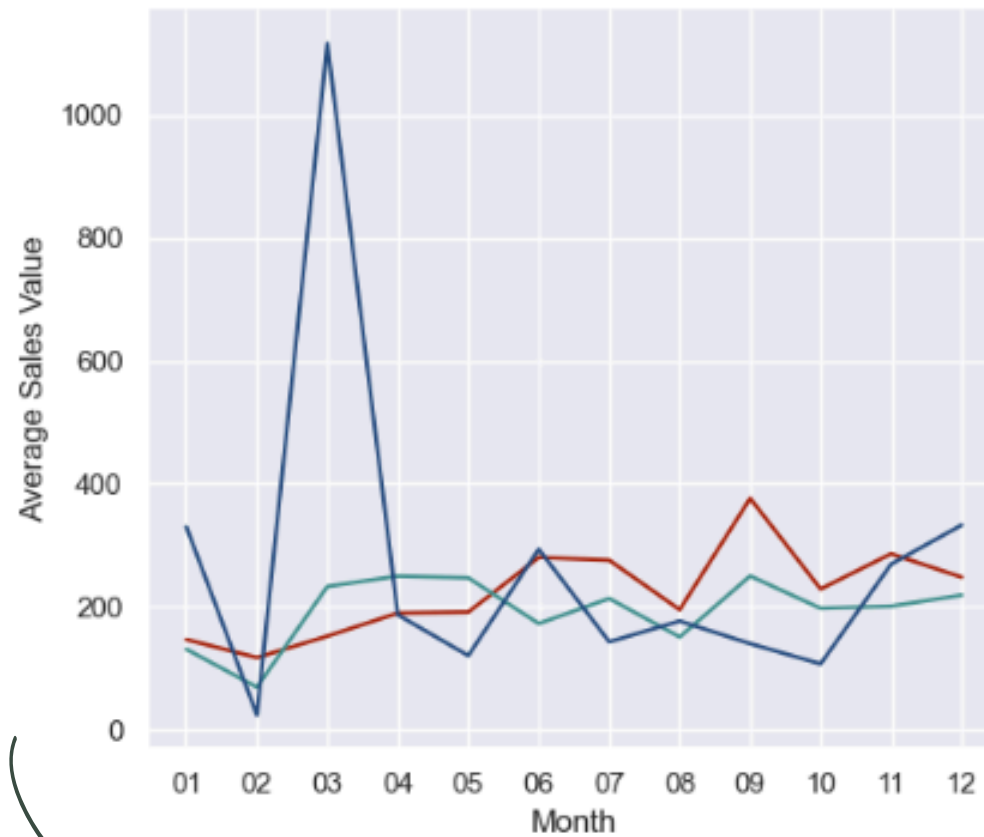
# WHAT IS THE AVERAGE SALES BY SEGMENT, BY YEAR, AND BY MONTH?

Consumer

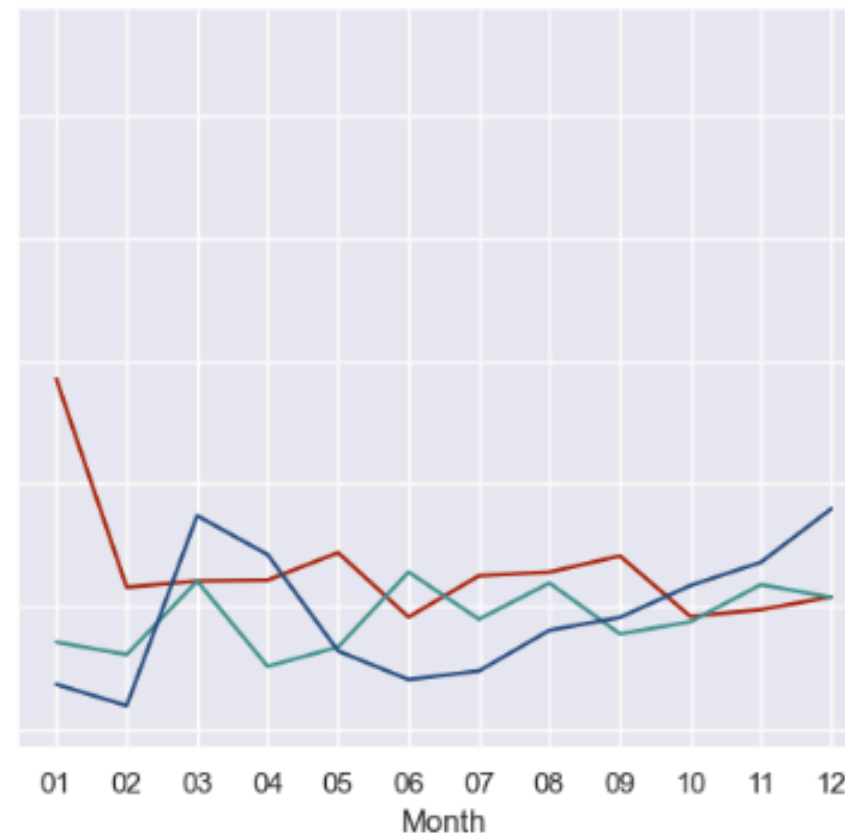
Corporate

Home Office

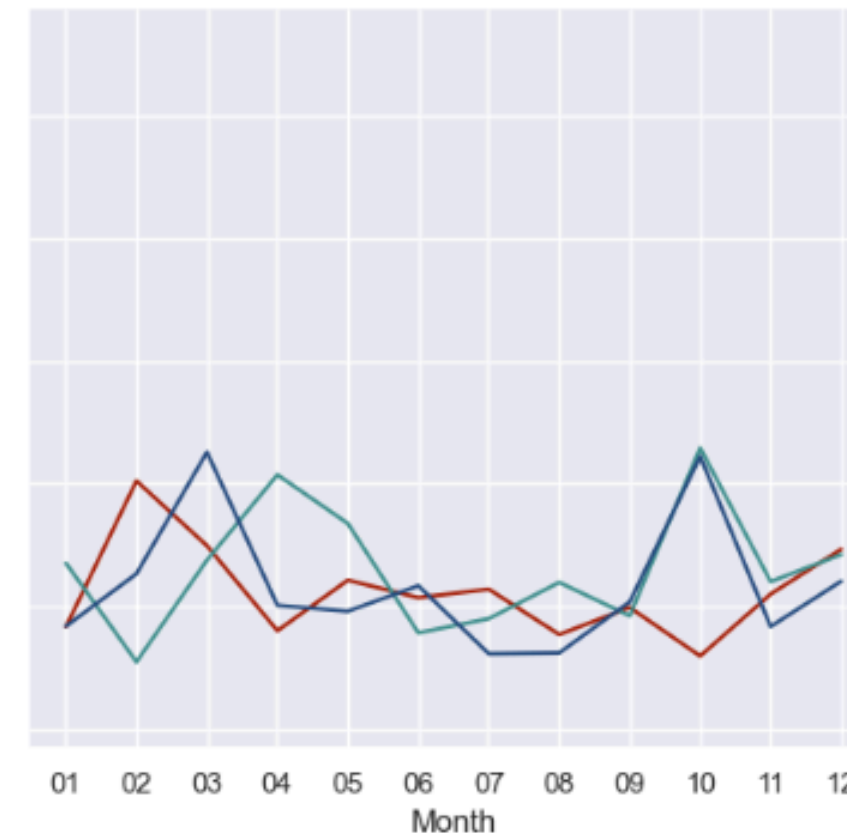
Year: 2015



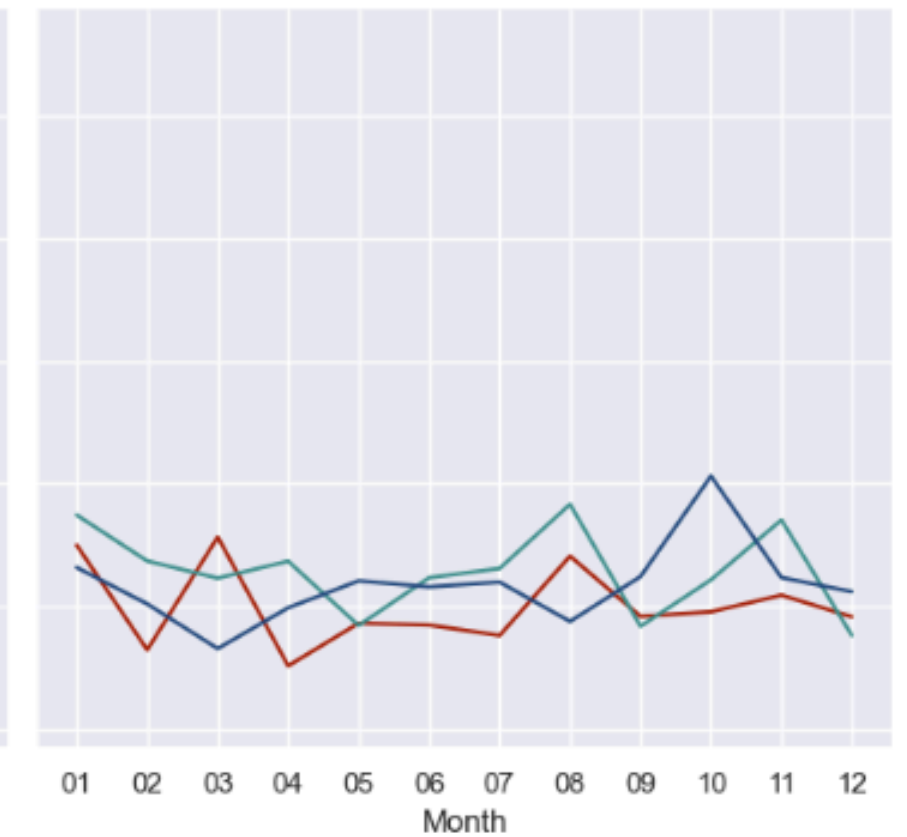
Year: 2016



Year: 2017



Year: 2018

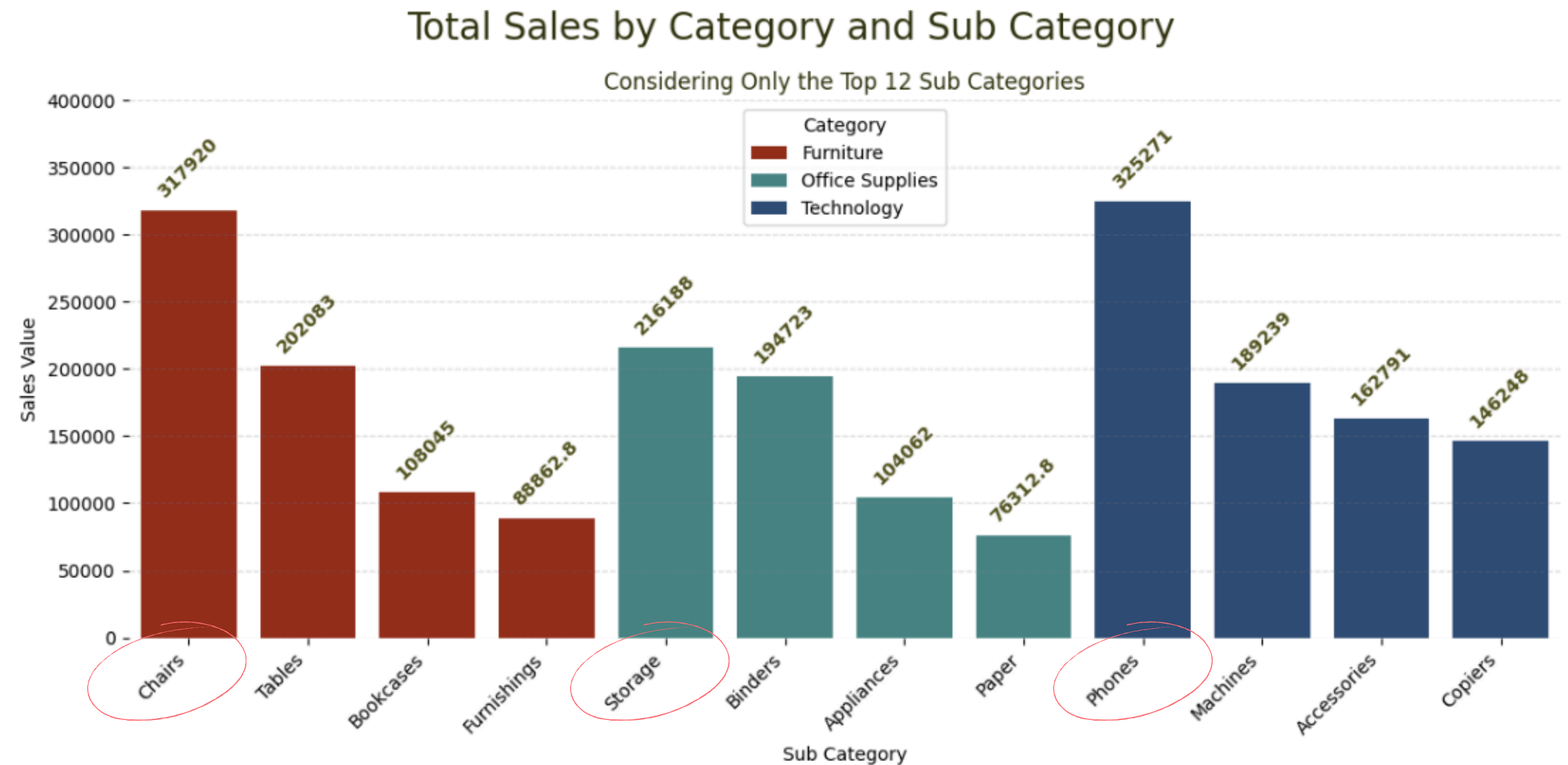


- In March **2015**, the Home Office segment had the highest average sales
- In **2016**, except for January when the Consumer segment had the highest average, all segments generally varied within a similar range

- For **2017** and **2018**, all segments had similar ranges of average sales

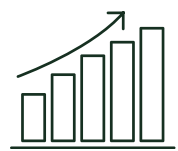
## WHAT IS THE TOTAL SALES BY CATEGORY AND SUBCATEGORY, CONSIDERING ONLY THE TOP 12 SUBCATEGORIES?

- For the **Furniture** category, Chairs had the highest sales value
- For the **Office Supplies** category, Storage had the highest sales value
- For the **Technology** category, Phones had the highest sales value



# INSIGHTS

# SUMMARY REPORT



Sales have **increased** over the years



Nearly **37%** of the sales come from the Technology category



**New York City** has the highest sales overall, while **California** has the highest total sales by state



Over the years, the **Consumer** segment has consistently achieved the highest sales, whereas the Home Office segment has had the lowest sales



The 4th quarter consistently experiences the highest sales, likely due to events such as Black Friday or Christmas



# DASHBOARD

# SALES REPORT

Dashboard allows sales teams and management monitor performance dynamically and react to market changes, customer behavior, or performance issues immediately



- This report includes total sales, number of customers, and number of orders
- You can filter the data by period (2015, 2016, 2017, and 2018), segment, and state
- The graphs represent quarterly sales, total sales by month for each segment, total sales for the top 5 cities, and sales explain

# **FURTHER ANALYSIS**

# FUTURE ANALYSIS OPPORTUNITIES

To complement the current analysis, the following areas can be explored in future assessments:

- **Customer Retention Analysis:** Investigate customer retention rates to determine the proportion of repeat versus new customers and assess overall customer loyalty
- **Purchase Frequency Analysis:** Analyze average customer purchase frequency and seasonal buying patterns to identify low-demand periods for targeted promotions
- **Customer Lifetime Value (CLV) Analysis:** Estimate the lifetime value of customers across different segments to prioritize high-potential customer groups
- **Delivery Time Analysis:** Assess average delivery times across regions and products to identify areas for improving customer satisfaction
- **Sales Channel Segmentation Analysis:** Analyze performance across different sales channels (online, physical) to determine which is performing better and whether strategy adjustments are needed
- **Low-Performing Products Analysis:** Identify products with low sales to decide on actions like promotion, price adjustments, or discontinuation