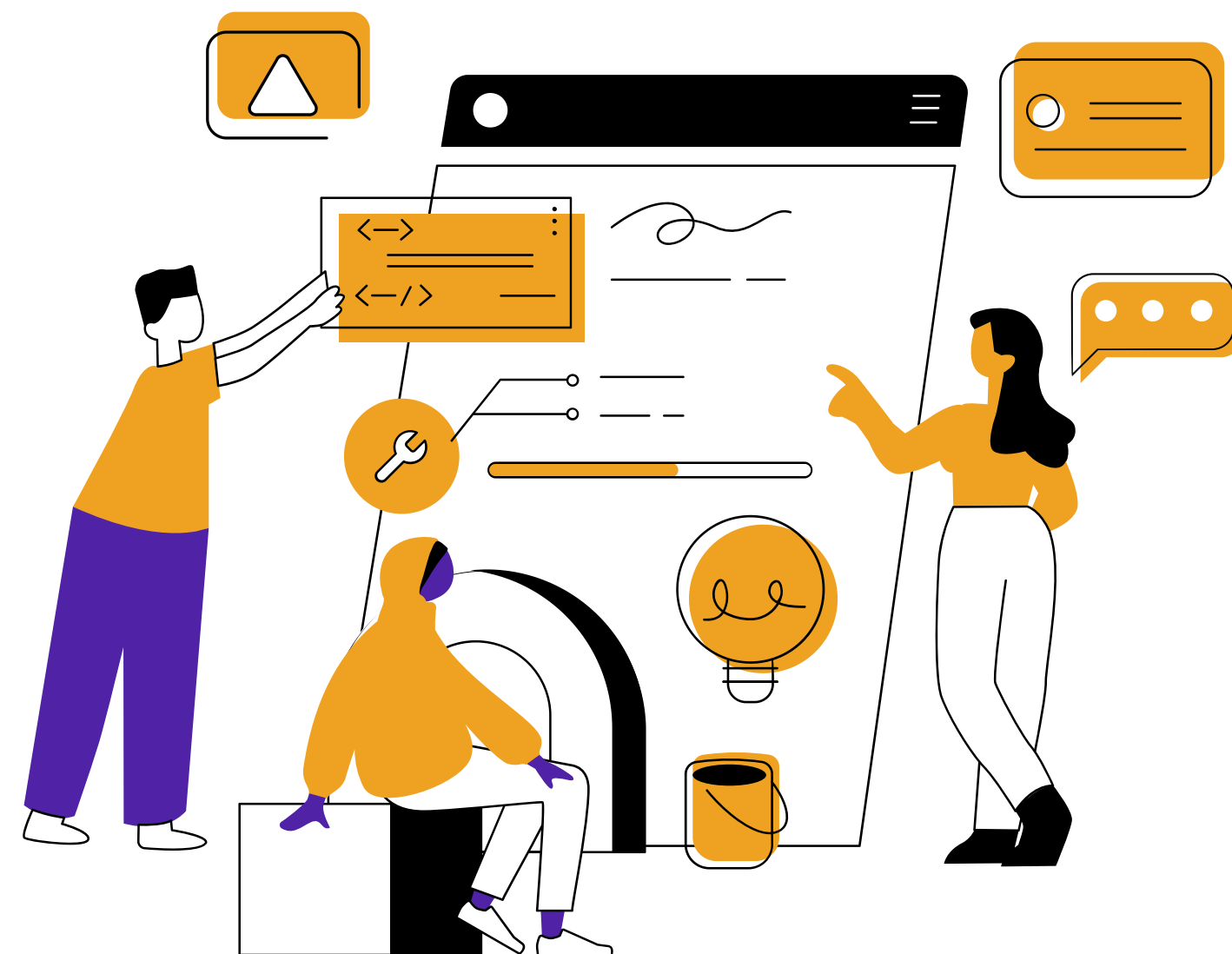
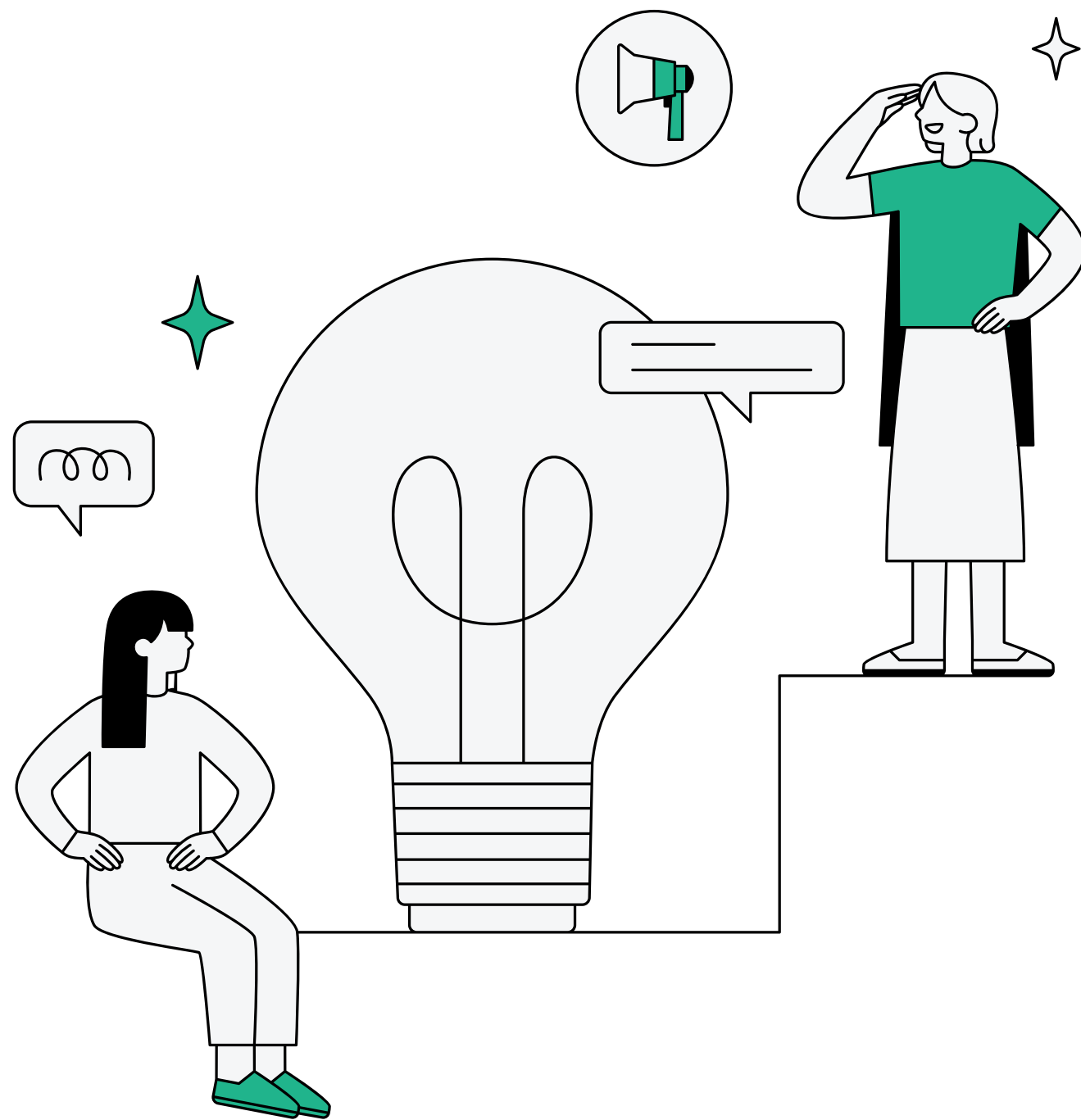




SALES DATA ANALYSIS

*Sailing through data cleaning straight
to clever business questions!*





INTRODUCTION

- **Objective:** The goal of the project, which is to analyze monthly sales data for an online retail business to derive actionable insights for business improvement.
- **Data Source:** Downloaded from an online retail business website for the year 2019 which is merged into a single CSV for analysis.

TOOLS AND LIBRARIES USED



- Platform



- For Data Manipulation

matplotlib

- For Data Visualization



Drop rows of NAN

```
# Find NAN
nan_df = all_data[all_data.isna().any(axis=1)]
display(nan_df.head())

all_data = all_data.dropna(how='all')
all_data.head()
```

Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address
1	NaN	NaN	NaN	NaN	NaN

Get rid of text in order date column

```
all_data = all_data[all_data['Order Date'].str[0:2]!='Or']
```

Make columns correct type

```
all_data['Quantity Ordered'] = pd.to_numeric(all_data['Quantity Ordered'])
all_data['Price Each'] = pd.to_numeric(all_data['Price Each'])
```

DATA PREPERATION

- Merge Data from Each Month into One CSV
- Dropping rows with missing values (NaN) using **dropna()**. (above image)
- Removing text from the 'Order Date' column that starts with 'Or'. (Below image)
- Converting columns to the correct data types (e.g., 'Quantity Ordered' to numeric).

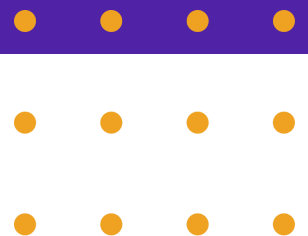




Augment Data with Additional Columns

This part adds additional columns to the DataFrame:

- 'Month': Extracts the month from the 'Order Date'.
- 'Month 2': Alternative method to extract the month using `pd.to_datetime`.
- 'City': Extracts the city and state from the 'Purchase Address'.

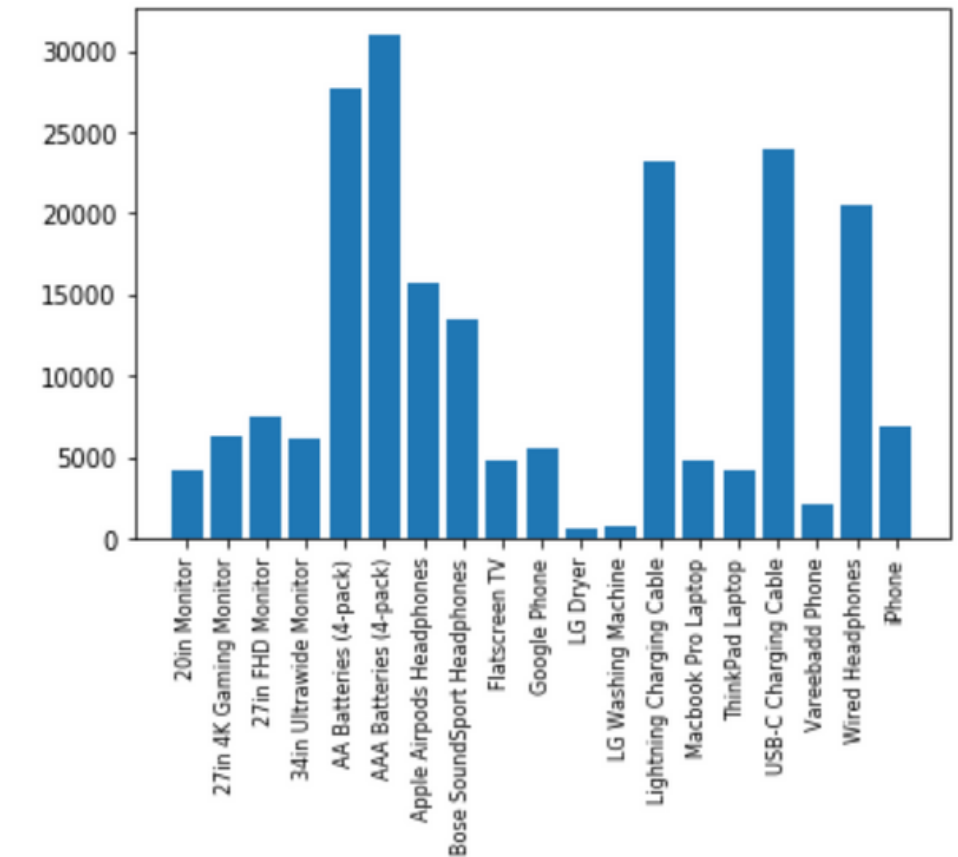
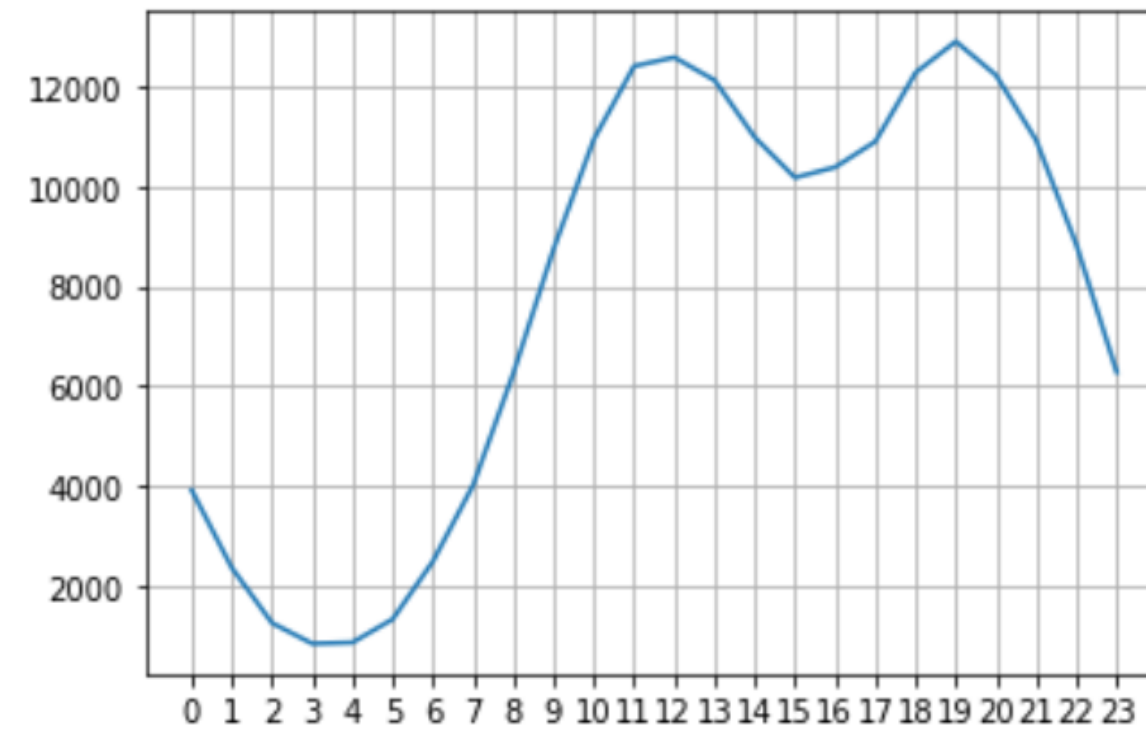
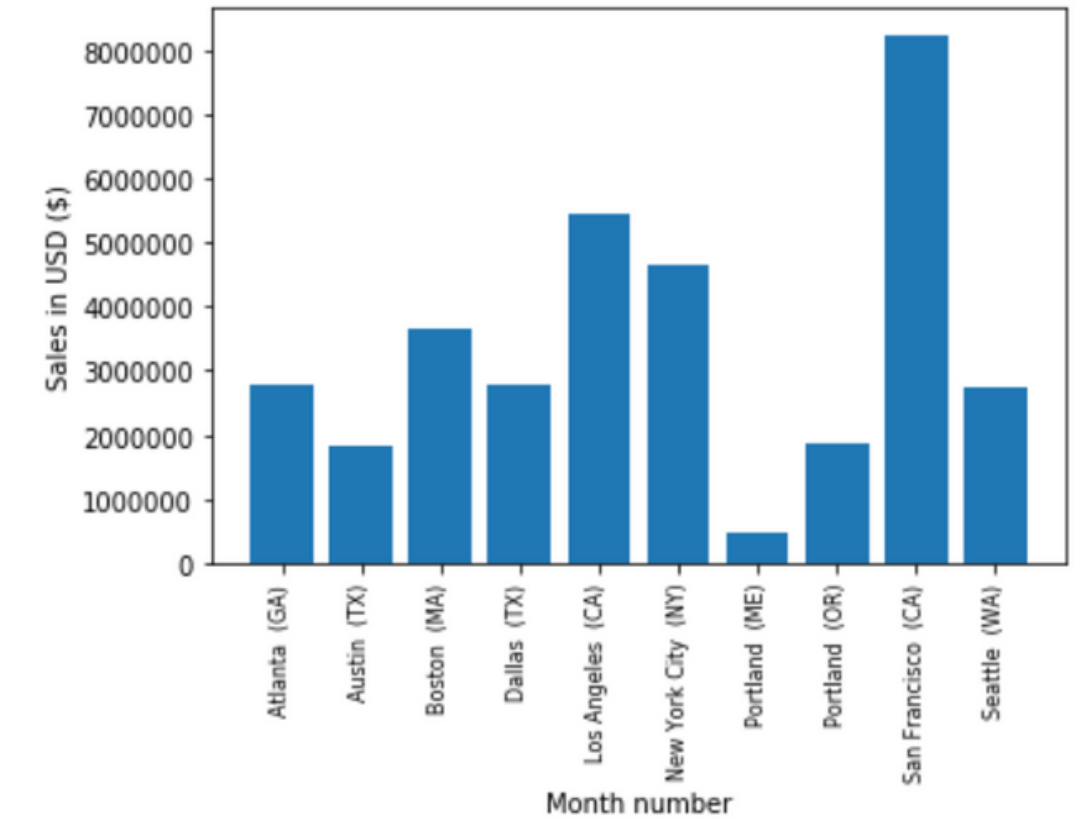
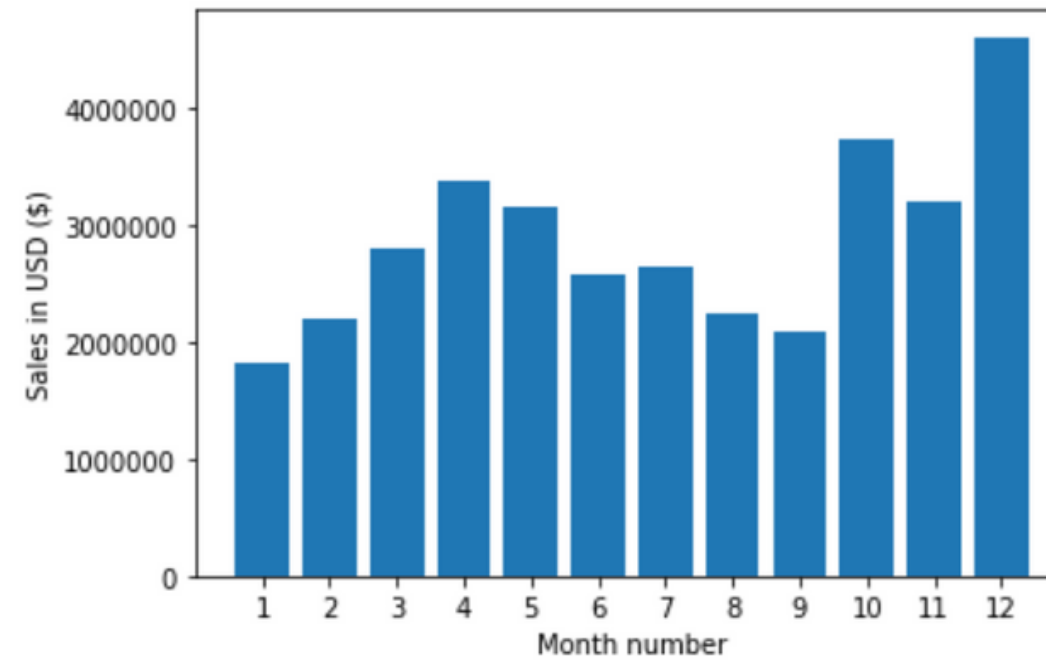




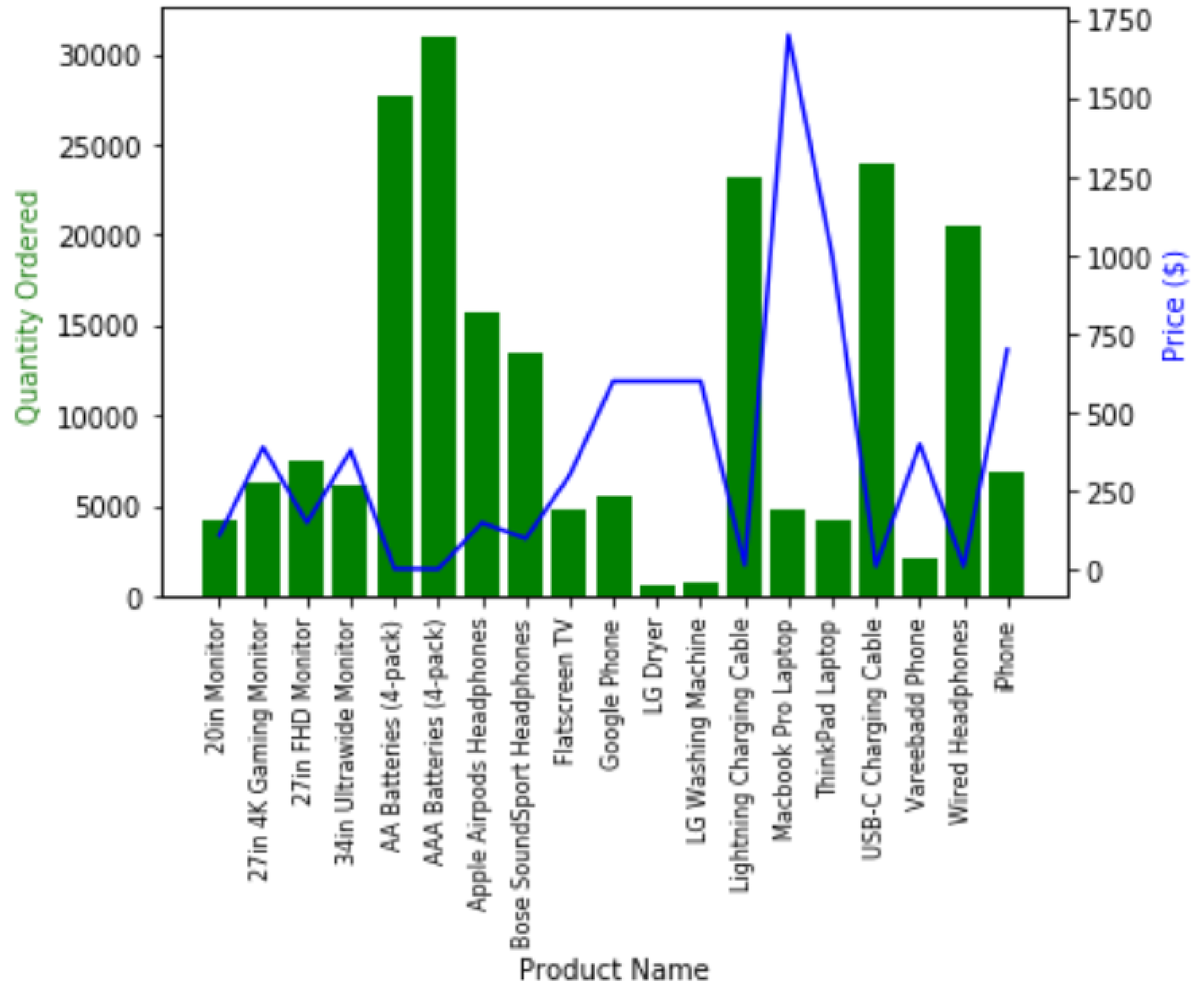
Data Exploration

This section explores the data by answering several questions:

- Identifying the best month for sales and the amount earned in that month.
- Determining which city sold the most products.
- Finding the best time to display advertisements to maximize the likelihood of customers buying products.
- Identifying which products are often sold together using combination analysis.



PRICE AND QUANTITY ORDERED CORELATION



Sales Analysis Insights and Recommendations



58%

December as
Highest-Grossing
Month

December emerged as the best month for sales, indicating a strong correlation with holiday shopping and suggesting potential targeted marketing strategies for high-sales periods.

1.2x

San Francisco, CA,
Leads in Sales

San Francisco, CA, recorded the highest number of sales, underscoring the significance of tailored marketing efforts and localized strategies in high-sales cities.

400K

Optimal
Advertisement
Timing

Ads displayed just before 11 AM or around 7 PM were identified as potential optimal times for higher conversion rates, aligning advertisement strategies with customer purchase patterns.

74Mn

Identified Product
Bundling
Opportunities

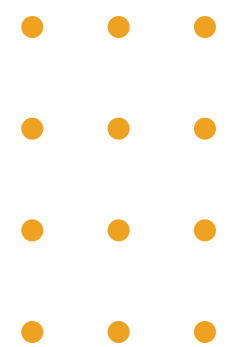
Identification of frequently bought together items offers valuable insights for promoting product bundles, potentially boosting sales and enhancing marketing strategies.

34Bn

Continuous Sales
Analysis for
Improvement

Emphasizing the importance of ongoing sales analysis to identify evolving trends, optimize strategies, and drive continuous improvement in sales performance.





Thank You

- Explore the full project on my GitHub repository for detailed code and analysis.
- Dive into various Tableau visualizations on my Tableau Public profile.

Github:

<https://github.com/ksundeeponne/Sales-analysis-with-EDA/tree/main>

Tableau Public:

<https://public.tableau.com/app/profile/sundeeponne.kukkala/vizzes>

