GROUP PROJECT
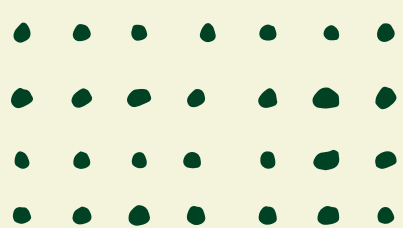
Spreadsheet Sales Analysis (Python)

Presenter: Gina Ghouri

February, 2025

PROJECT OBJECTIVES

Key Tasks

1. Read the data from the spreadsheet (sales.cvs file)

2. Collect all of the sales from each month into a single list

3. Output the total sales across all months

Extended Tasks

1. Output a summary of the results to a spreadsheet
2. Calculate the following:
a. Monthly changes as a percentage
b. The average
c. Months with the highest and lowest sales
3. Use a data source from a different spreadsheet

Visualisation

Using Python Pandas&Seaborn:
1. Bar Chart: monthly sales distribution to identify trends and patterns

2. Line Chart: percentage change in sales to understand fluctuations and growth rates

3. Dot chart: car sales advanced data

```python
49  def plot_sales(data): #Task4:generate bar plot for monthly sales
55      plt.ylabel('Sales')
56      plt.xticks(rotation=45)
57      plt.tight_layout()
58      plt.show()
59
60  def plot_percentage_changes(data):#Task5:generate line plot for percenta
61      percentages = changes_percentage(data)
62      months = [row['month'] for row in data][1:]
63      sns.lineplot(x=months, y=percentages, marker='o', color='b')
64      plt.title('Monthly Percentage Changes in Sales')
65      plt.xlabel('Month')
66      plt.ylabel('Percentage Change')
67      plt.xticks(rotation=45)
68      plt.tight_layout()
69      plt.show()
70
```

📊 Sales Analysis Dashboard

This dashboard provides analysis and visualization of sales data across different months. Use the sidebar to navigate through different analyses
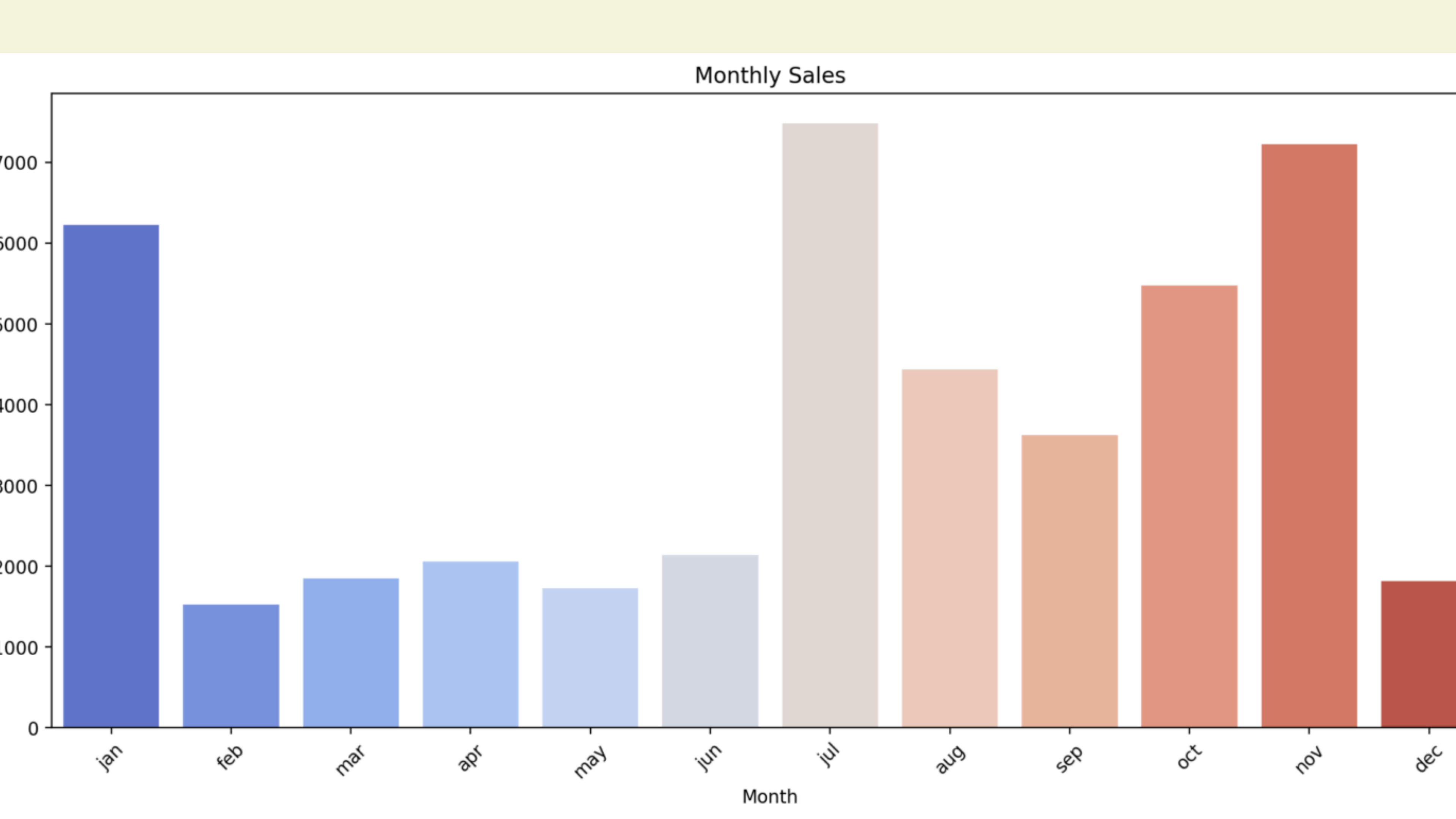
Total Sales	Average Monthly Sales	Total Months
$45,542.00	$3,795.17	12

Monthly Sales Data

	year	month	sales	expenditure
0	2018	jan	6226	3808
1	2018	feb	1521	3373
2	2018	mar	1842	3965
3	2018	apr	2051	1098
4	2018	may	1728	3046
5	2018	jun	2138	2258
6	2018	jul	7479	2084

Monthly Sales

| Month | jan | feb | mar | apr | may | jun | jul | aug | sep | oct | nov | dec |

Monthly Sales Percentage Changes
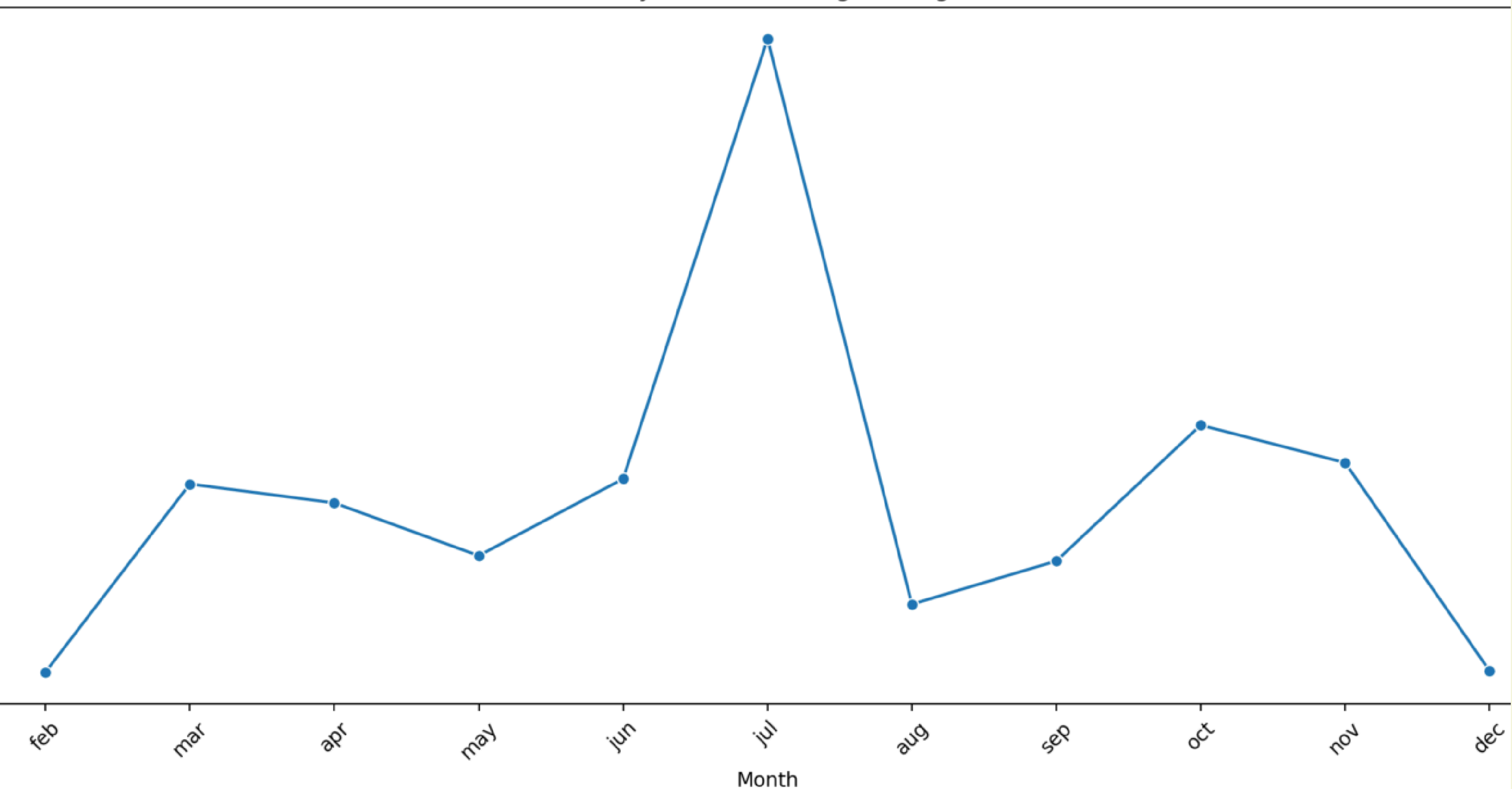
kaggle.com/datasets/syedanwarafridi/vehicle-sales-data?resource=download

kaggle

Create

Datasets

Models

Code

Discussions

Learn

More

Your Work

VIEWED

Vehicle Sales Data

🔍 Search

Vehicle Sales Data

▲ 553   <> Code   ⬇ Download

Data Card   Code (95)   Discussion (10)   Suggestions (0)

Dataset Description:

The "Vehicle Sales and Market Trends Dataset" provides a comprehensive collection of information pertaining to the sales transactions of various vehicles. This dataset encompasses details such as the year, make, model, trim, body type, transmission type, VIN (Vehicle Identification Number), state of registration, condition rating, odometer reading, exterior and interior colors, seller information, Manheim Market Report (MMR) values, selling prices, and sale dates.

Key Features:

Vehicle Details: Includes specific information about each vehicle, such as its make, model, trim, and manufacturing year.

Transaction Information: Provides insights into the sales transactions, including selling prices and sale dates.

Market Trends: MMR values offer an estimate of the market value of each vehicle, allowing for analysis of market trends and fluctuations.

License

MIT

Expected update frequency

Annually

Tags

Tabular   Transportation

Data Analytics   Travel

Data Visualization

🐍 spreadsheet_entire_project.py 3   ▦ car_prices.csv   🐍 spreadsheet_entire_project_cars.py 3 ✕   ▦ sales.csv

C: › Users › HP › OneDrive › Documents › CFG › Intro to Pyton › Spreadsheet Project Group 1 › 🐍 spreadsheet_entire_project_cars.py › ⬡ main

```python
34    def calculate_statistics(data): #average, highest/lowest sales and percentage changes
35        if data.empty:
36            return 0, 0, 0    # return defaults if empty
37
38        data['sellingprice'] = pd.to_numeric(data['sellingprice'], errors='coerce')
39        data.dropna(subset=['sellingprice'], inplace=True)
40
41        data['Monthly Change (%)'] = data['sellingprice'].pct_change() * 100
42        avg_sales = data['sellingprice'].mean() #mean means average
43        highest_sale = data['sellingprice'].max()
44        lowest_sale = data['sellingprice'].min()
45
46        return avg_sales, highest_sale, lowest_sale
47
48    def generate_summary(data, output_file='vehicle_sales_summary.csv'): #write the summary results to a new CSV file
49        if data is not None and not data.empty:
50            data.to_csv(output_file, index=False)
51            print(f"Summary saved to {output_file}")
52        else:
53            print("No valid data to save.")
54
55    def visualise_sales(data): #creates a sales trend graph based on sale dates, handling bad dates
56        if data is None or data.empty:
57            print("No data available for visualisation.")
58            return
```

PROBLEMS 6   OUTPUT   DEBUG CONSOLE   TERMINAL   PORTS   ⟩ python   + ⌄ ⊟ 🗑 ··· ⌃

```
PS C:\Users\HP\OneDrive\Documents\CFG\Intro to Pyton\Spreadsheet Project Group 1> python .\spreadsheet_entire_project_cars.py
Total Sales Revenue: $7606367587.00
Average Sale Price: $13611.36
Total Sales Revenue: $7606367587.00
Average Sale Price: $13611.36
Average Sale Price: $13611.36
Highest Sale Price: $230000.0
Lowest Sale Price: $1.0
```
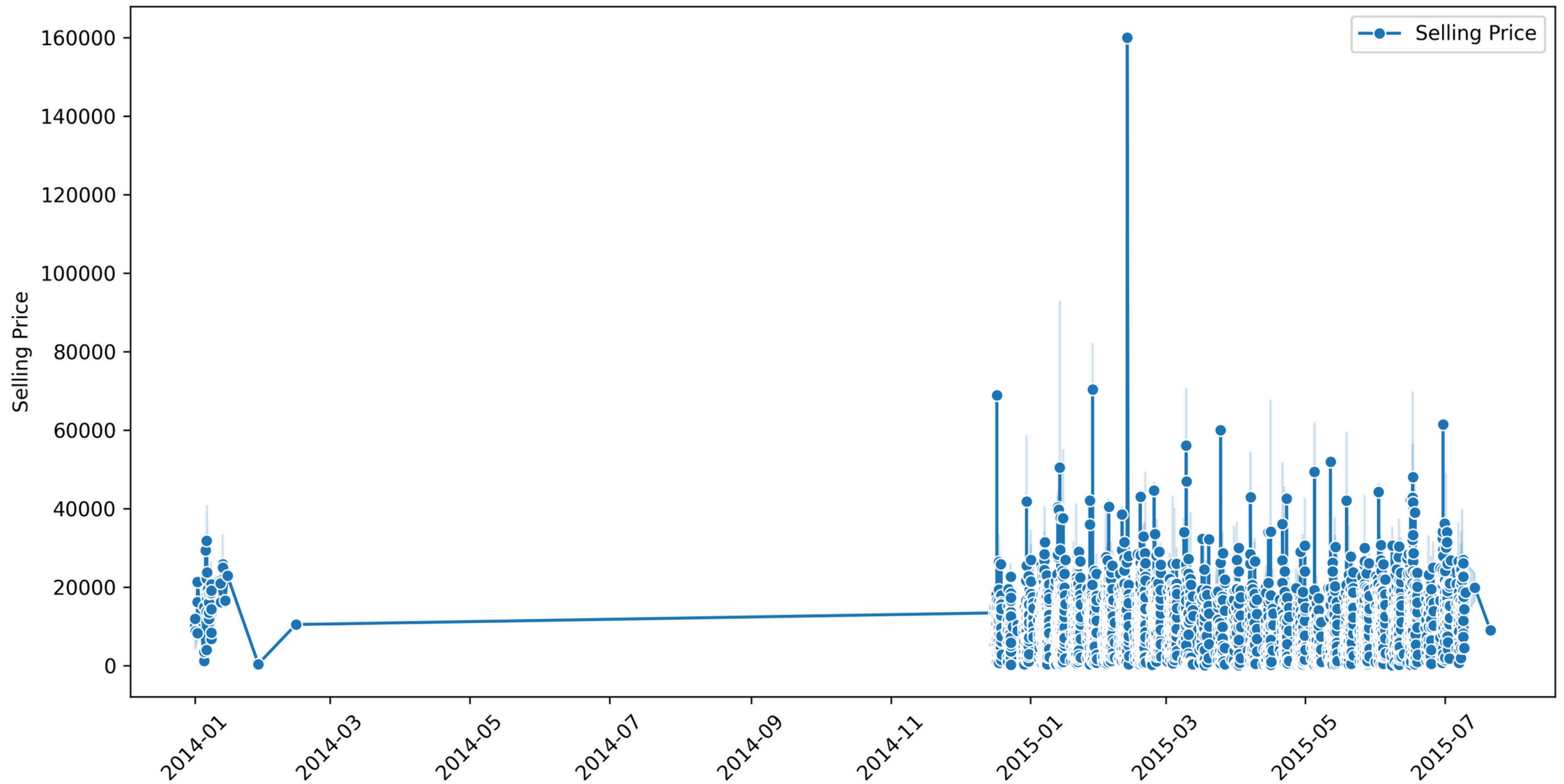
Vehicle Sales Price Trend Over Time

CHALLENGES FACED

Group Communication Issues:

- Project was not divided into smaller sections
- Live calls and task assignment was missing

Different Data Sheet:

- Kaggle car sales data cvs file was too large
- Visualisation graph required advanced Object Oriented Programming

Suggestions for Improvement:

- Foster stronger relationships with customers through personalised engagement strategies: retention rates and loyalty program
- Detailed guidelines to use software frameworks like jira or agile and moderated breakout rooms and meeting discussions

CONCLUSION

By implementing spreadsheet analysis and data visualisation, the project goal is to boost immediate sales figures and also to establish a sustainable framework for future growth🐱🕶️

Question and Answers🙌

https://github.com/ginaghouri/Data-Analysis