# Data Science Capstone

Zach Peagler

# Executive Summary

- Data collection and cleaning
- Exploratory data analysis
- Model creation
- Visualization and dashboard creation

# Introduction

I’m completing the MIT Data Science Professional Certificate to advance my career and show employers I mean business. Real business.

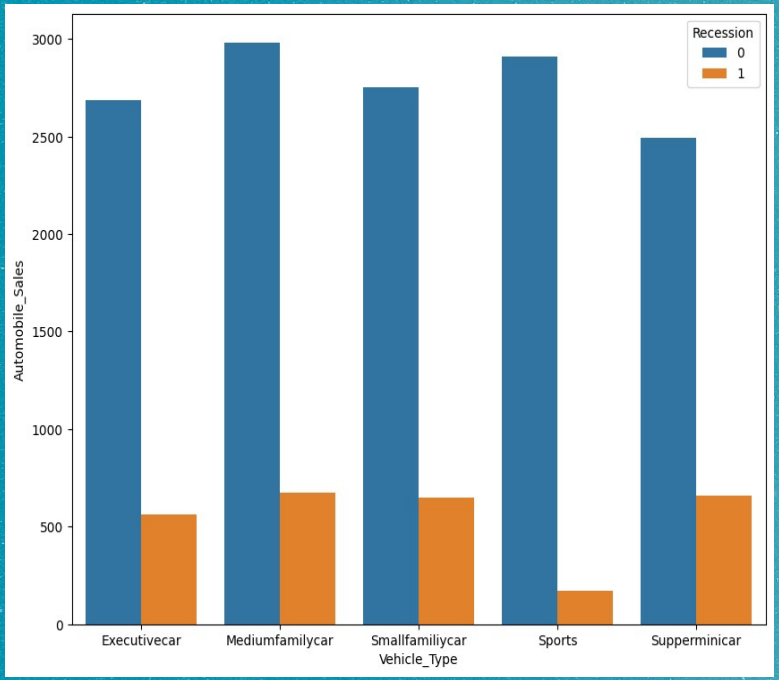
# Exploratory Data Analysis Methodology

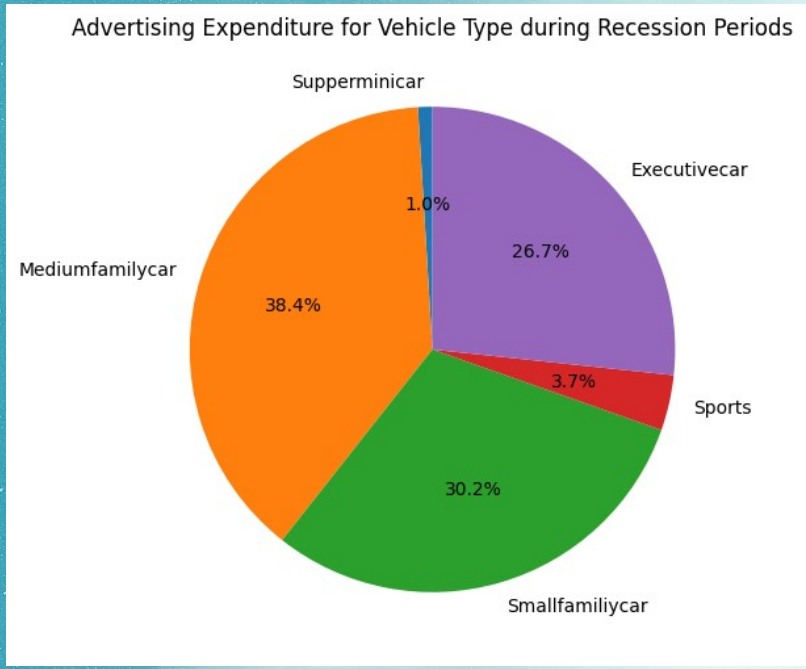
- Histograms, bar charts, and scatter plots, oh my!

# Predictive Analysis Methodology

- Linear regressions, principal component analyses, and machine learning.

# EDA with Visualizations



# EDA with SQL

- SQL in Python can be an excellent way to query databases without leaving your preferred IDE
- “magic” SQL functions enhance this, and tools such as SQLite make it possible

# Folium

- I don’t have a screenshot of this
- Folium is a tool for displaying interactive maps in Python, and is an excellent tool for dashboard creation

# Plotly Dash

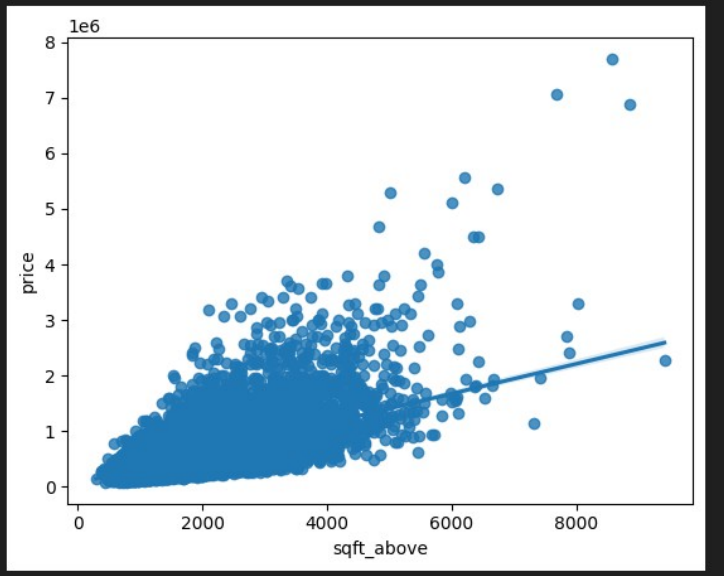
Average Automobile Sales fluctuation over Recession Period
Automobile_Sales
Year
700
650
600
550
1980
1990
2000
2010
2020
Total Advertising Expenditure by Vehicle Type During Recession Periods
Mediumfamilycar
Smallfamiliycar
Supperminicar
Sports
Executivecar
30.2%
38.4%
26.7%
3.7%
0.954%
Average Automobile Sales by Vehicle Type
Automobile_Sales
Vehicle_Type
Executivecar
Mediumfamilycar
Smallfamiliycar
Sports
Supperminicar
Effect of Unemployment Rate on Vehicle Type and Sales
Average Automobile Sales
Unemployment Rate

# Predictive analysis results

```
sns.regplot(x=df['sqft_above'], y=df['price'])
```

[36] ✓ 0.5s

```
<Axes: xlabel='sqft_above', ylabel='price'>
```



# Conclusion

- I’m highly competent with many data science techniques.
- Look at graph, Michael.