

mkiganane20 /
phase_three_project-

<> Code

Issues

Pull requests

Actions

Projects

Wiki

Security

Insights

Setting

Project Phase 3

☆ 0 stars

🔗 0 forks

👁 1 watching

🌿 Branches

📈 Activity

📁 Tags

🌐 Public repository

🔗

🌿 1 Branch

📁 0 Tags

🔗

📁

🔍 Go to file

t








Go to file

+



Add file ▾

Code

⋮

	mkiganane20 Automobile Dataset	1d90444 · now	
	Automobile_data.csv	Automobile Dataset	now
	Jupyter notebook.pdf.pdf	Jupyter notebook.pdf	now
	MYRA PROJECT JUPYTER NO...	jupyter notebook ipynb file	9 minutes ago
	Presentation.pdf.pdf	Project Presentation	12 minutes ago
	README.md	README.md	20 minutes ago

📖 README



phase_three_project-

Project Phase 3

1. Introduction

1.1 Project Overview

This project presents an exploratory data analysis (EDA) of an automobile dataset to uncover trends and relationships among various car features and their influence on pricing, insurance costs, and consumer preferences.

1.1.1 Business Understanding

The automobile industry is highly competitive, and understanding what factors influence car prices, performance preferences, and insurance risk can give manufacturers and dealers a strategic advantage.

1.1.2 Business Problem

Stakeholders need actionable insights on what drives vehicle pricing, insurance cost estimation, and customer demand to make data-driven decisions for production, marketing, and pricing strategies.

1.1.3 Key Business Questions

How do engine size and horsepower influence car pricing?

Does drive configuration impact vehicle cost and buyer preferences?

Are certain body types more likely to incur higher insurance symboling values?

What recommendations can be derived to maximize profitability?

2. Data Understanding

2.1 Data Preprocessing

2.1.1 The Data

The dataset used in this project contains various automobile attributes such as:

Engine size, horsepower

Drive configuration (FWD, RWD, 4WD)

Body type, fuel system, curb weight

Price and insurance symboling

Each row represents a unique vehicle, and the features allow for in-depth analysis of pricing and risk factors.

3. Technologies Used

Programming Language: Python

Libraries: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn

Tools: Jupyter Notebook for data exploration and visualization

4. Usage

Clone the repository to your local machine.

Open the Jupyter Notebook environment.

Navigate to the .ipynb notebook in the /notebooks/ directory.

Run the cells sequentially to explore data cleaning, visualization, and insights.

Review visual findings and recommendations for business strategy.

5. Project Structure

/data/ - Contains the raw and cleaned automobile dataset

/notebooks/ - Jupyter notebooks for data exploration and visualization

/scripts/ - Python scripts for data preprocessing and modeling

README.md - Project documentation

6. Key Insights & Recommendations

Vehicles with larger engine sizes and higher horsepower show a strong positive correlation with price. Manufacturers should invest in efficient, high-performance engines to justify premium pricing.

Rear-Wheel Drive (RWD) and 4WD vehicles command higher average prices and are preferred by performance-conscious consumers, especially in premium segments.

Body type influences insurance symboling, with some types linked to higher risk. Collaborating with insurance firms to transparently communicate pricing based on body-type risk categories can improve customer trust and satisfaction.

7. Conclusion

Data-driven decision-making is essential for optimizing vehicle design, pricing, and marketing strategies. • Future steps: Conduct deeper analysis of engine performance, drive configurations, and insurance-related factors. • Action: Implement these insights to enhance product value, attract target buyers, and improve overall profitability.



Releases

No releases published

[Create a new release](#)

Packages

No packages published

[Publish your first package](#)

Languages

● Jupyter Notebook 100.0%