

Used Car Analysis

Capstone Project

Author : Carla Kirby



This Photo by Unknown Author is licensed under [CC BY-NC](#)



Discussion Points

- Business Problems data aims to solve
- Data
- Process Methods Used
- Visualizations
- Results
- Summarize Conclusions



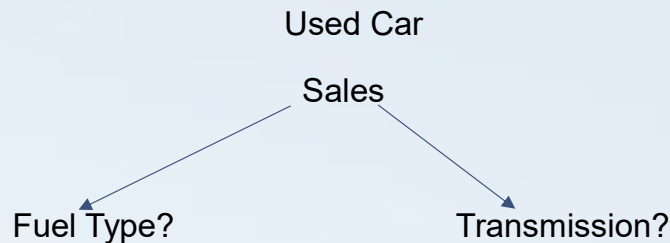
Business Problem

Used Car Lot™ is a car sales company running for 2 years as a local business in India. The General Manager has requested some information about second hand car sales in the local regions. With fuel prices increasing and economic uncertainty important questions were asked so that business can maintain profitability.

1. What regions in India sold the highest number of used cars?
2. Does fuel type affect the car sale price?
3. Do electric cars make an impact on a sales?
4. What type of transmission sold the most?
5. What are some other factors that share a connection to sales prices?

From this analysis we intend on building a multiple linear regression model to understand the relationships of the variables to the sale price.

What leads to higher used car sales? 1. Fuel type or 2. Transmission?



Data

Large Dataset size is over 7200 line - Source of used cars sold in India.

The dataset is comprised of values such as mileage, year built, power, seats, engine type, kilometers driven and the price.

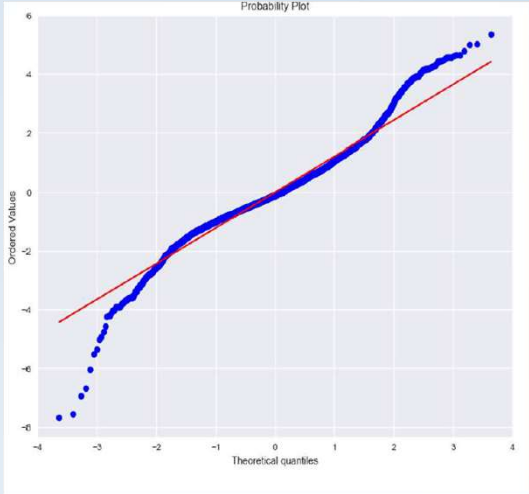
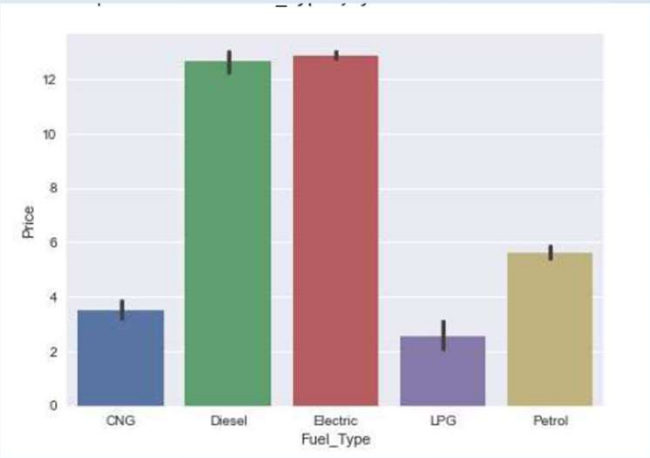
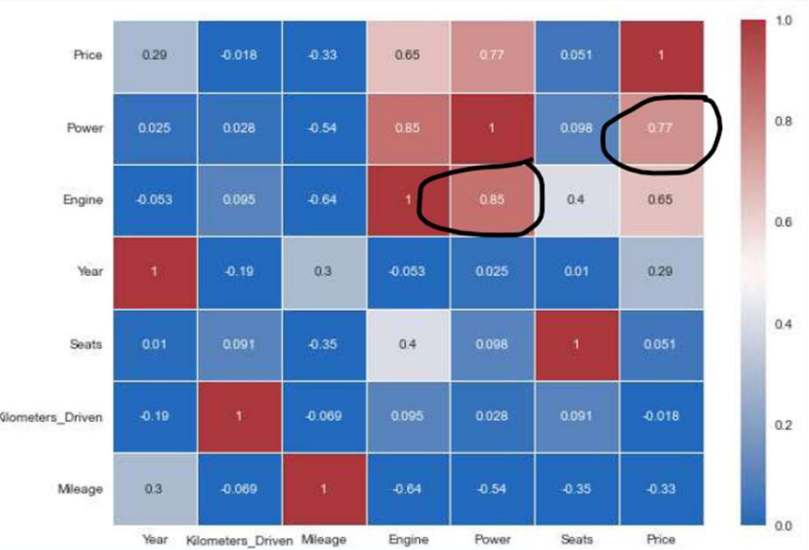
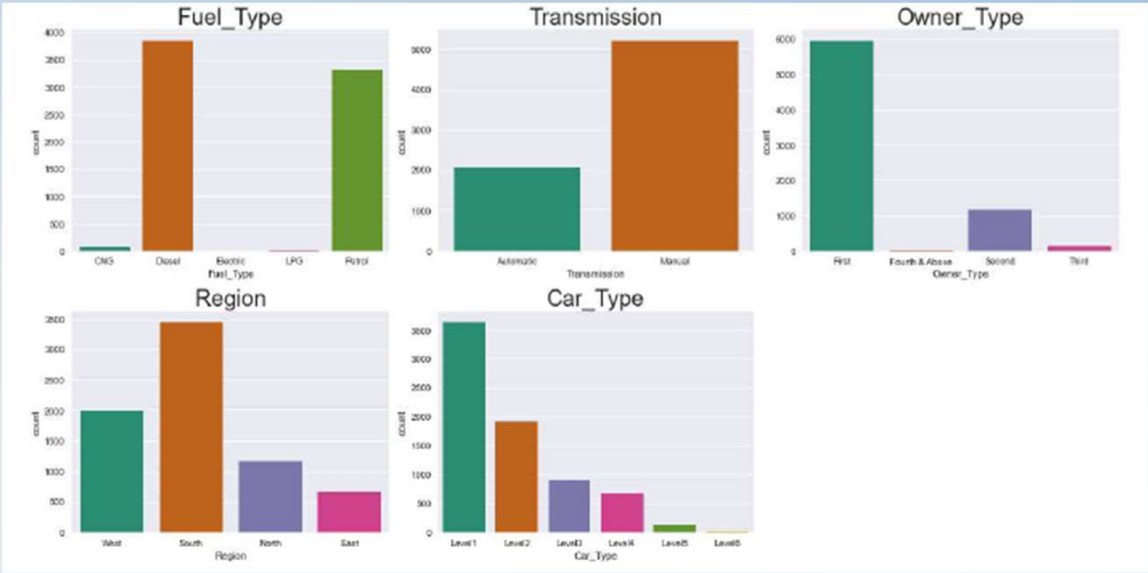
Sourced from data.gov, a reliable data source online.

Dataset required much categorical refinement in order to prepare for exploratory analysis.

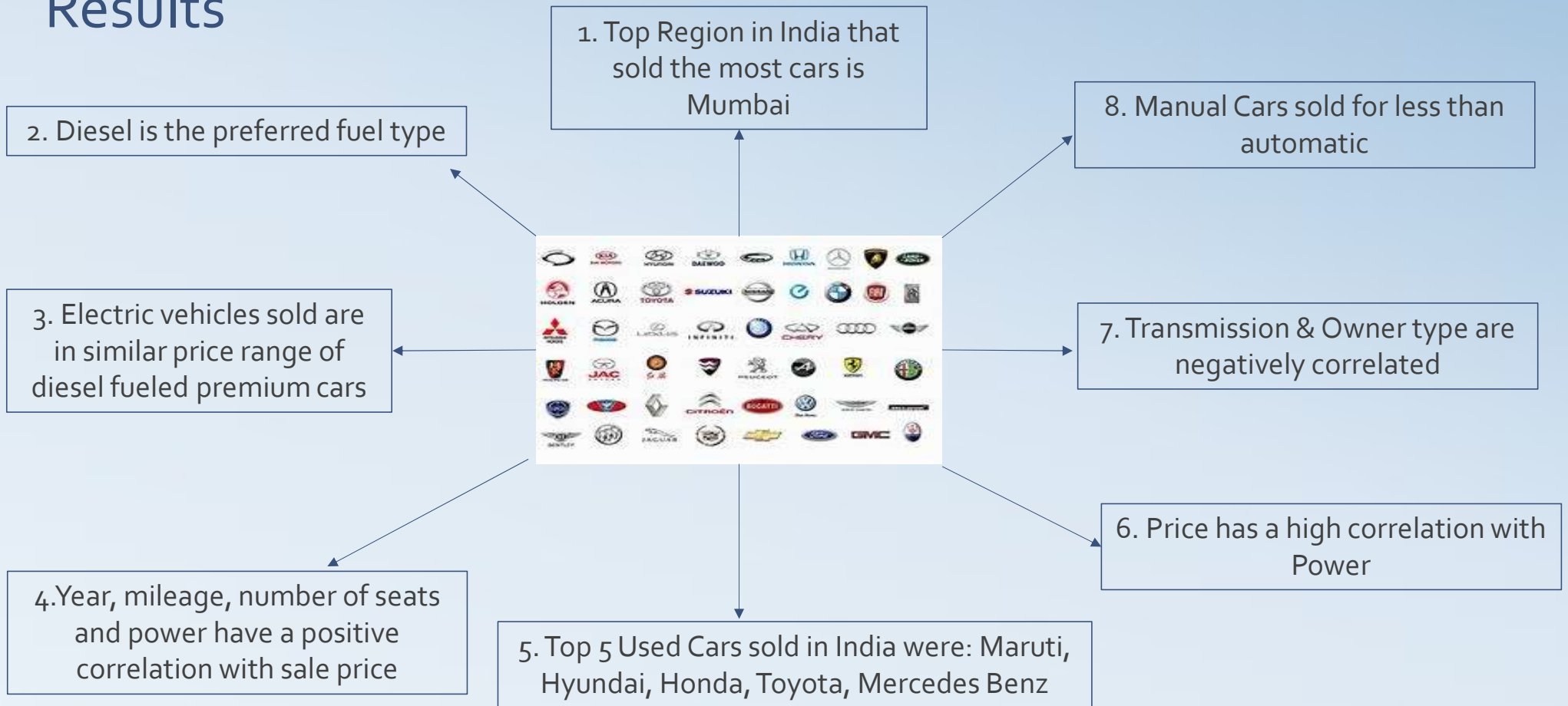
Data analysis packages consisted of matplotlib, numPy, pandas, statsmodels, and scikitlearn.

Models & Visualizations

	count	unique		top	freq
Name	7253	2041	Mahindra XUV500 W8 2WD		55
Location	7253	11	Mumbai		949
Fuel_Type	7253	5	Diesel		3852
Transmission	7253	2	Manual		5204
Owner_Type	7253	4	First		5952



Results



1. What regions in India sold the highest number of used cars? :
2. Does fuel type affect the car sale price?
3. Do electric cars make an impact on a sales
4. What are some other factors that share a connection to sales prices

- Best Sellers Maruti, Hyundai, Honda, Toyota, Mercedes Benz
- Price has a high correlation with Power
- Transmission & Owner type are negatively correlated
- Manual Cars sold for less than automatic

Conclusions

- Electric Fuel Price & Diesel scored the highest based on the data set directly affecting the cost of the car. Petrol, LPG, and CNG vehicles were sold for less.
- Transmission and owner type have a negative correlation to Price
- Conclude that the OLS₃ is a good model for inference and prediction with an adjusted R² value of 95.1% - refer to Jupyter notebook
- Electric car impact directly aligns with price with slight multicollinearity to fuel consumption biases.

Questions?

