Predictive Model for Pricing Used Vehicles

a presentation by Kelly Goforth

Project Goals

- Clean, analyze, visualize, and refine the data in order to identify patterns and trends to build a dependable model.
- Develop a sound and reliable predictive method to accurately estimate pricing for used vehicles.

Model Data

- 5068 usable rows of data
- 12 meaningful columns of information
  - Make
  - Year
  - Mileage
  - Seats
  - Fuel Type
  - # of Owners
  - Transmission
  - Kilo's Driven
  - Engine
  - Power
  - Location
  - Price

Model Data

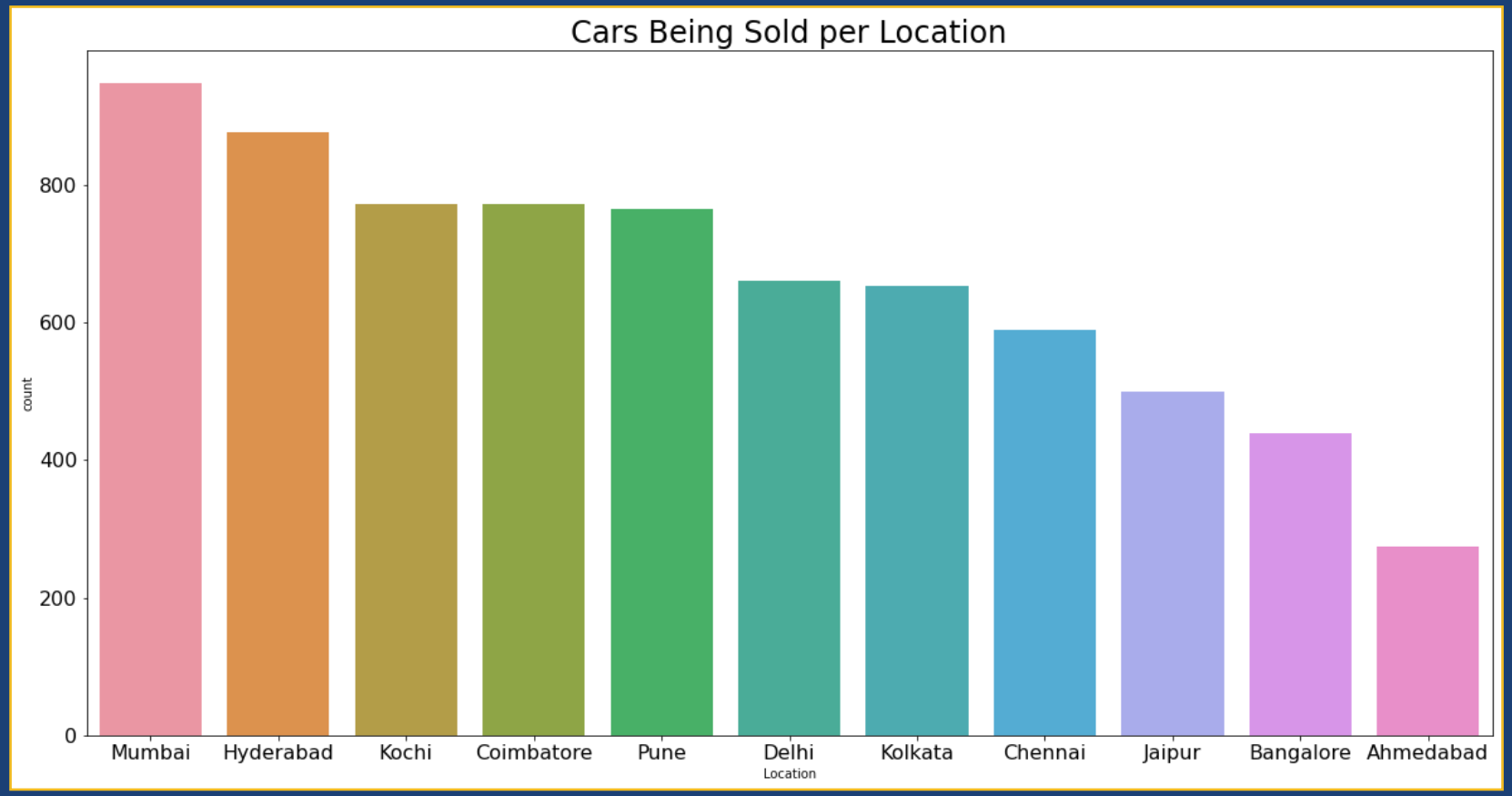
Important manipulations performed:

- Outliers removed from data set – model may not be able to predict "unique" vehicles (such as Lamborghinis and electric cars).
- Missing values removed from data set or investigated and filled to ensure model reliability and accuracy.

Data Trends – Location

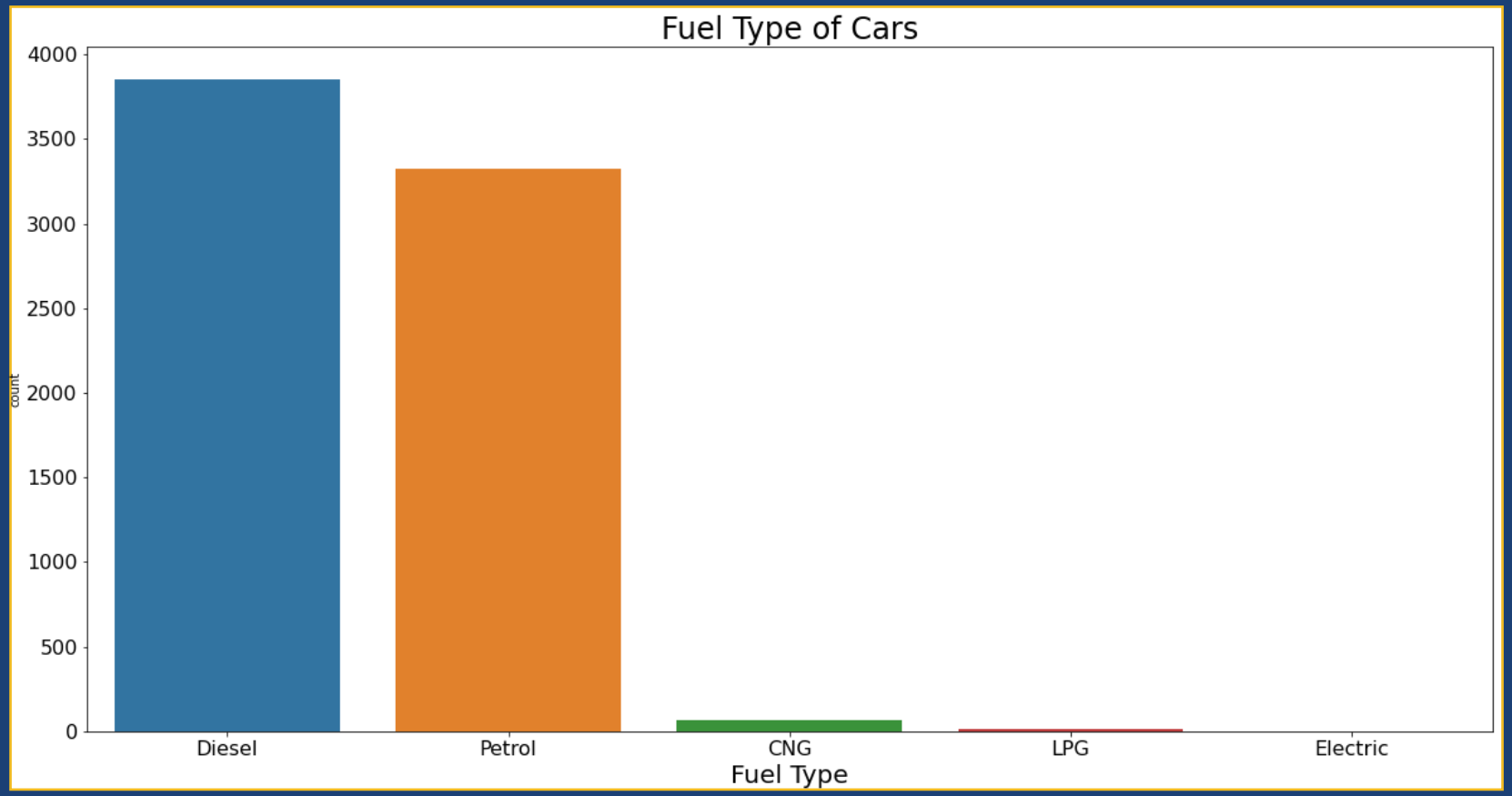
Top selling cities:

1) Mumbai
2) Hyderabad
3) Kochi



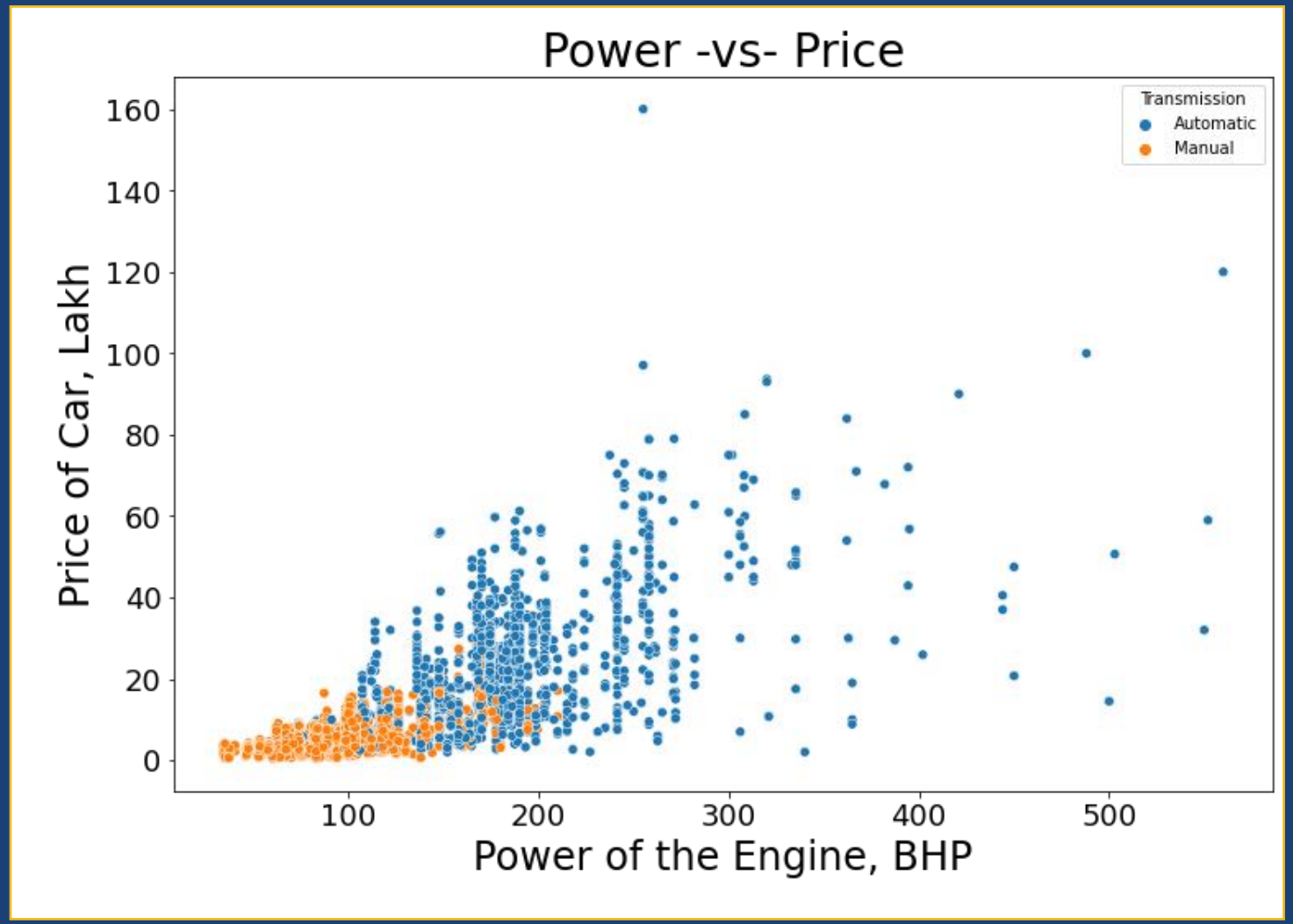
Data Trends – Fuel Type

Diesel and Petrol dominate the used car market.



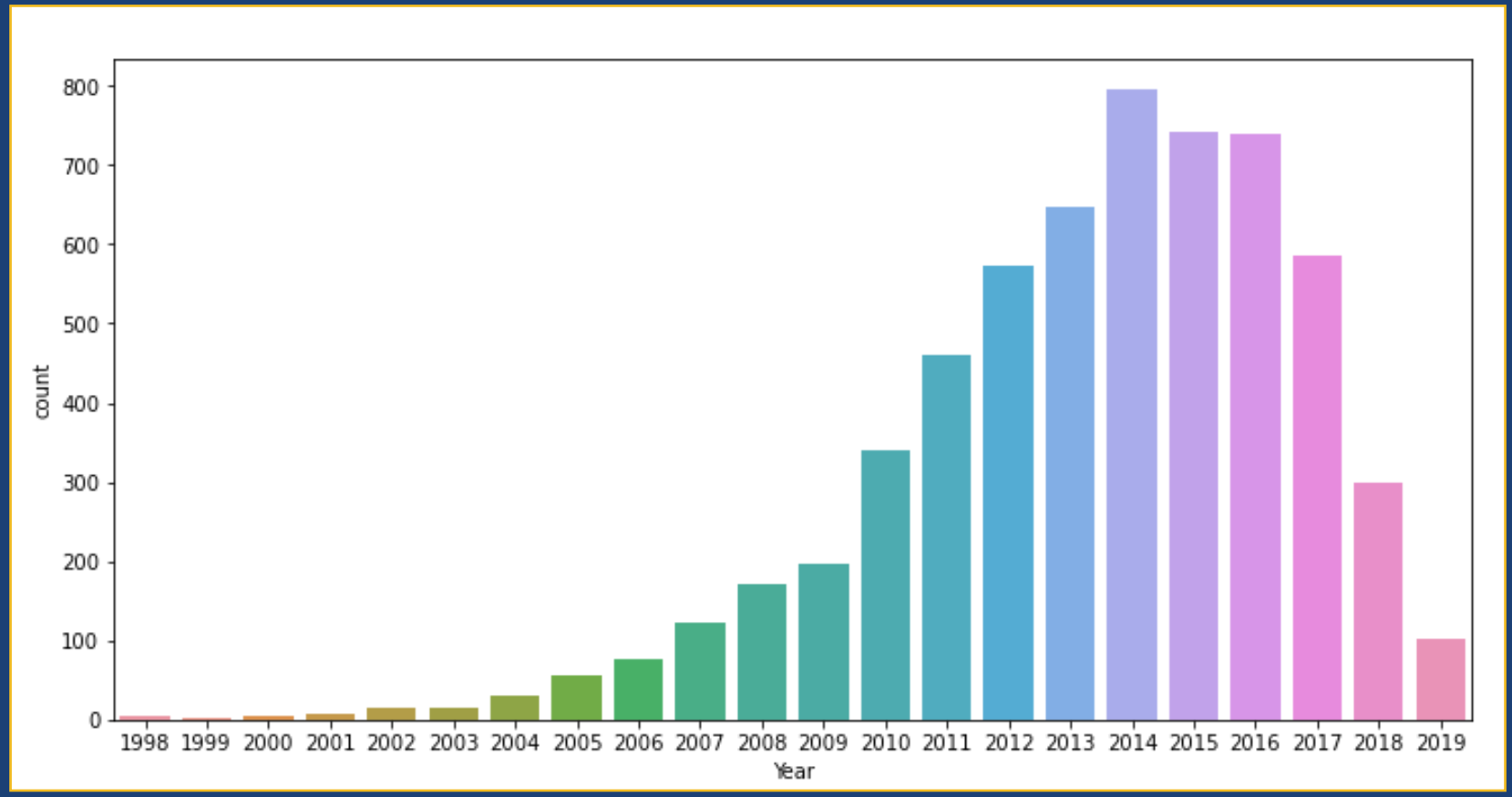
Data Trends – Power

Manual transmission cars sell for much less and are less powerful than Automatics.



Data Trends – Year

Most used cars on the market now were made after 2010.



Pricing Model

- Machine learning techniques were used to develop and refine a pricing model based on the trends in the data.

- Variance Inflation Factors and p-values were minimized in order to create the most accurate model possible.

Pricing Model

In a nutshell:

- You tell me the following details about the vehicle, and I'll tell you what price to sell the vehicle for:

Make, Kilometers Driven, Year Manufactured, Mileage, Power, Location of Sale, Number of Owners, Number of Seats, Transmission

Key Pricing Indicators

- Power has a strong positive impact on the price of the vehicle.
- The more power the engine has, the more we can sell the vehicle for.

Key Pricing Indicators

- The age of the vehicle has a strong impact on the price of the vehicle.
- The newest vehicles sell for the most, while older vehicles sell for less.

Key Pricing Indicators

- Some car makes increase the price of the vehicle as you would expect (such as Mini, Land Rover, BMW, and Audi), but it is not as a rule.

- For example, Porsche has a high negative coefficient, indicating that the model predicts that that make lowers the predicted price.

Overall Take-Aways

- Machine learning can find patterns that humans cannot.
- This pricing model will allow us to accurately price our used vehicles, leading to higher sales, and therefore, higher profits.

Room for Improvement?

- More data is required in order to build a reliable model for pricing vehicles with high Power and alternative fuel types.