

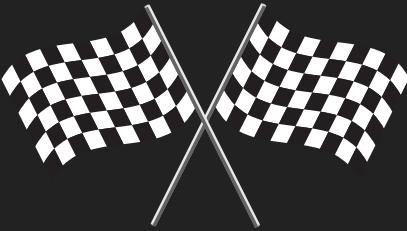
YEAGER AUTOWERKS

...

01

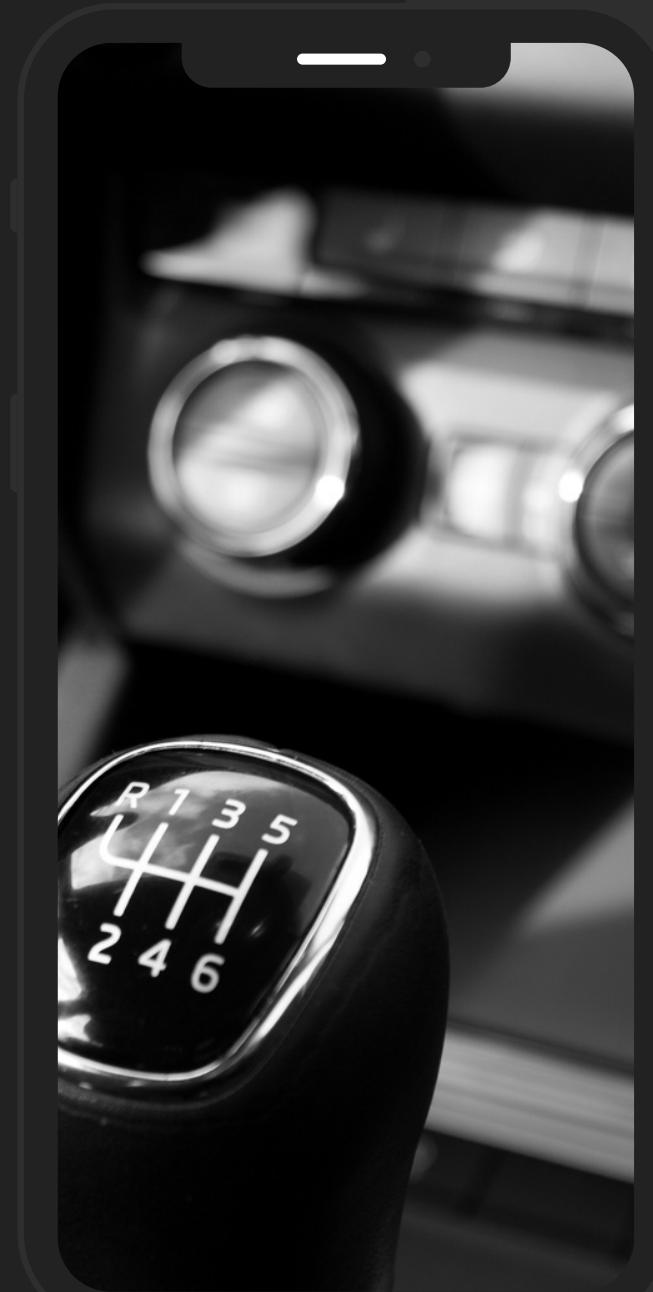
UNDERSTANDING USED VEHICLE MARKET PRICES





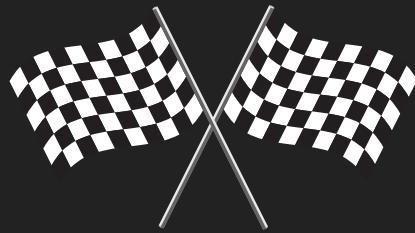
02

Overview & Goals



Obtaining the proper vehicles to maximize profit in the used auto market can prove difficult. Our analysis aims to make this into less of a game of chance. Our firm has obtained a dataset which includes the used vehicle values from all Craigslist used car listings during the current year, in an attempt to identify those features which have the greatest impact on automotive resale value. In addition, the data was fitted to a linear regression model to allow for predictive pricing capabilities for possible purchases.





03

\$11,733.27

the average price of a used car in our region.

9

the average age of a used car in our region.

104,294

the average mileage of a used car in our region.

The Data

The dataset used in this analysis is the "Kaggle Used Cars Dataset: Vehicles listings from Craigslist.org". The data is made up of all content, as scraped from the Craigslist used vehicle listing section (nationally), during January 2020. It consists of 509,577 records, with 25 features. Of these, our study uses only those from our sales region (MD, VA, PA, and WV), as well as those records which are the most robust in terms of feature content; resulting in a filtered and formatted dataset of roughly 10,000 records. All records containing null values were discarded, as well as those records with values lying over 2 standard deviations from their respective means.

~But What do these numbers tell us?

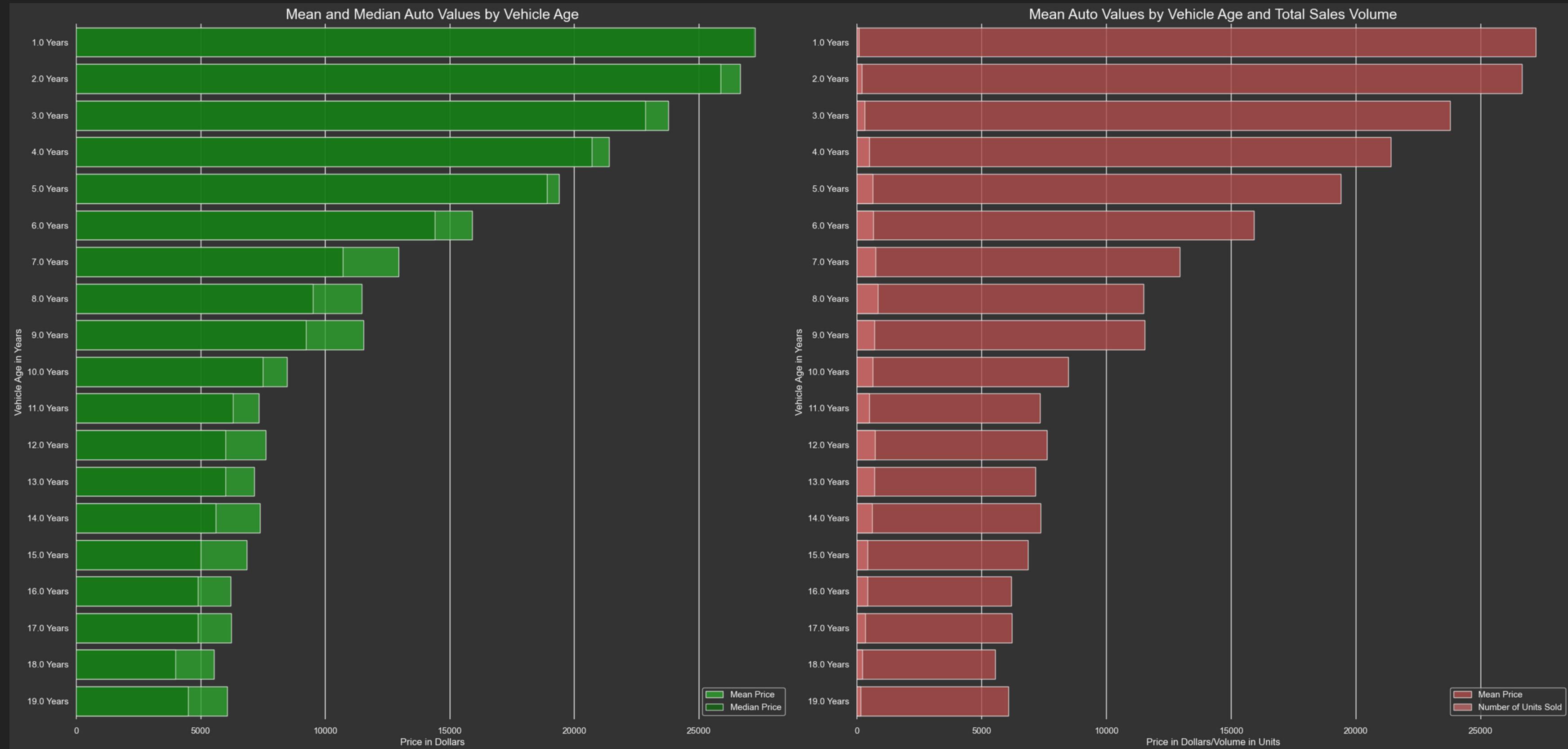
•••

SIGNIFICANT FEATURES: AGE

According to our data, a vehicle loses, on average \$20,000 in value by reaching 10 years of age. This represents a majority of the value it will lose over its lifetime. With that in mind, the volume of all vehicles sold in this market which are over 10 years old is 43%. Younger car = higher profit, but potential additional volatility between age cutoffs



04

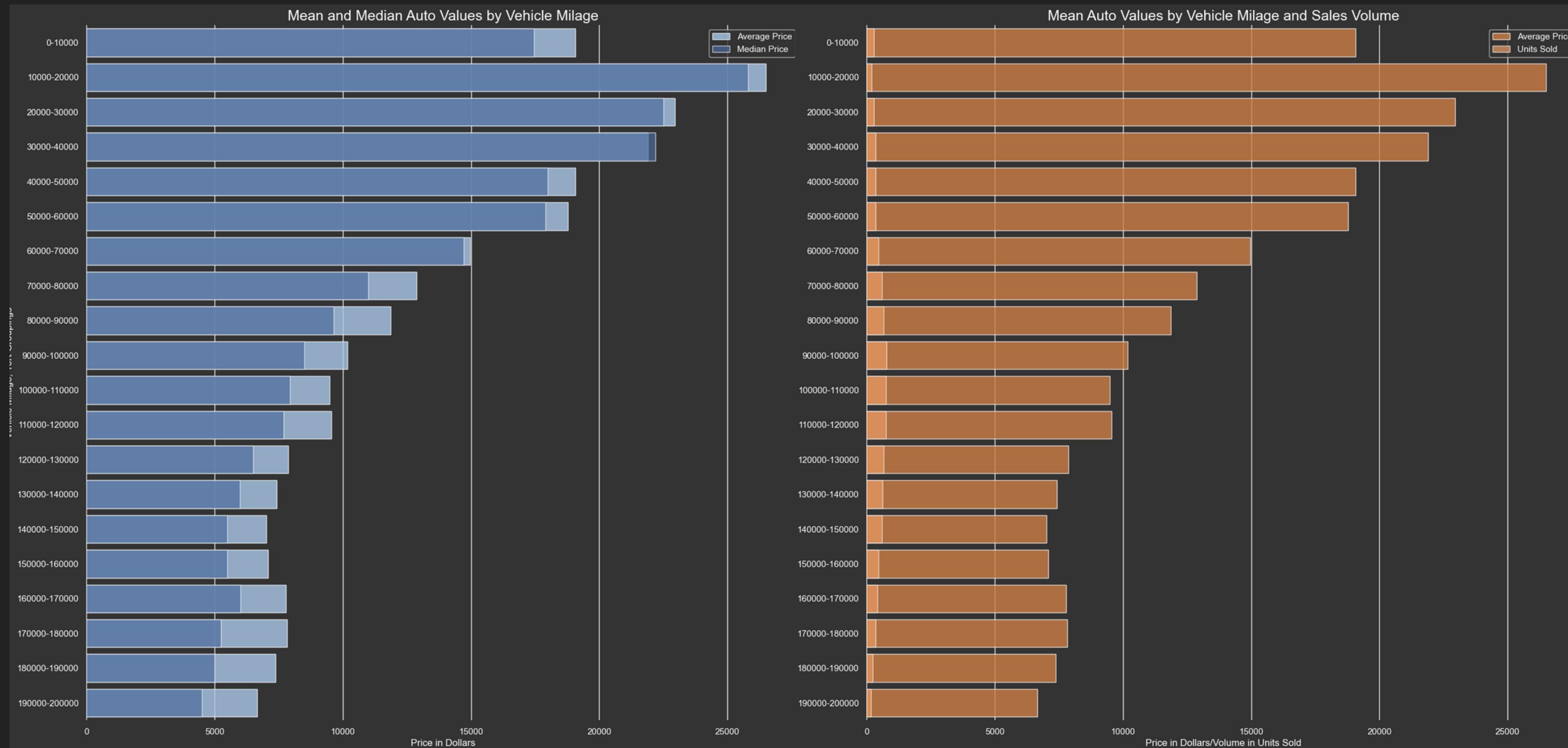


SIGNIFICANT FEATURES: MILEAGE

Similarly, a vehicle loses, on average \$17,000 of value by reaching 100k miles, again a majority of the value it will lose over its lifetime. The volume of all vehicles sold in this market which have 100k+ miles is 54%. < 100k Miles = higher profit, but potential additional volatility between mileage ranges.



05



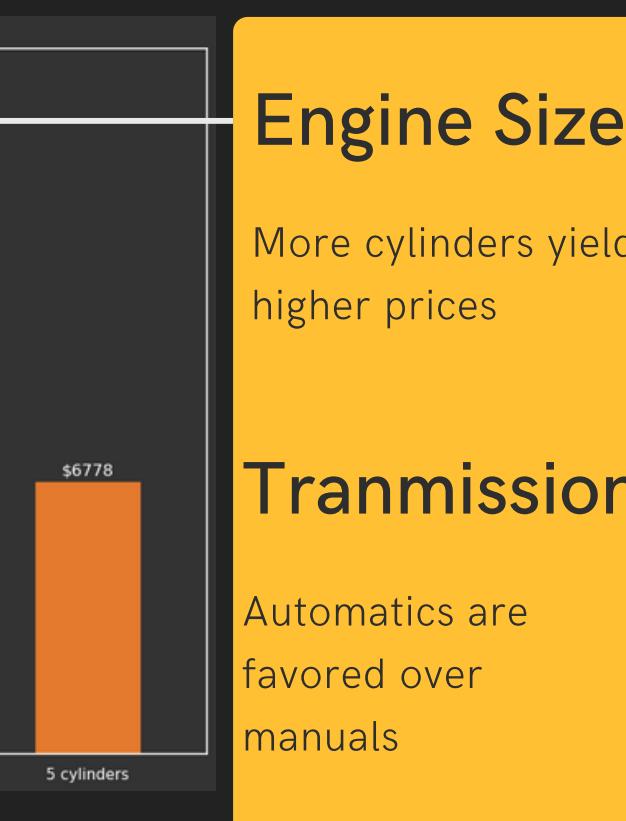
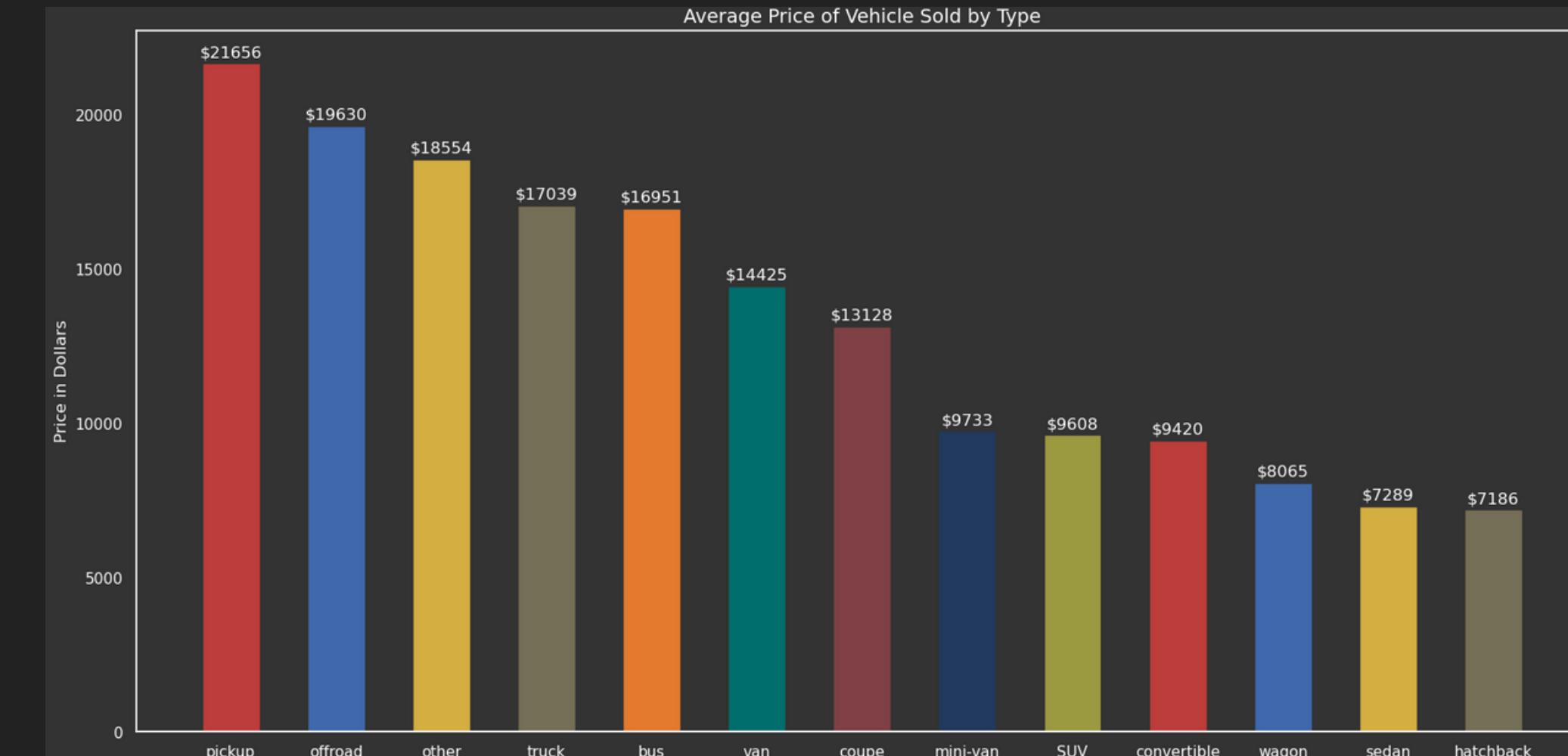


OTHER FEATURES TO CONSIDER WHEN MAXIMIZING PROFIT

06

By Vehicle Type

Pickup Trucks, Offroad focused vehicles, Vans, and Coupes tend to bring higher prices.

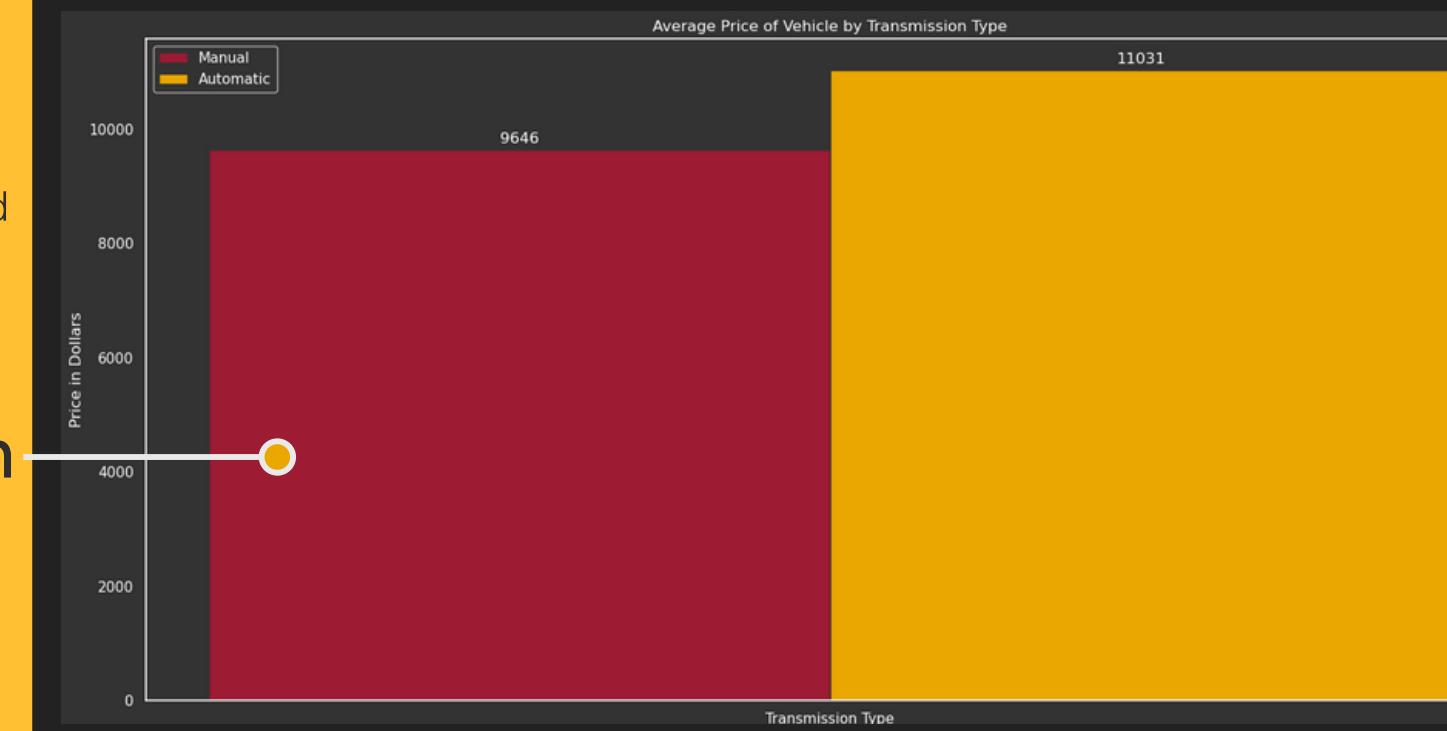


Engine Size

More cylinders yield higher prices

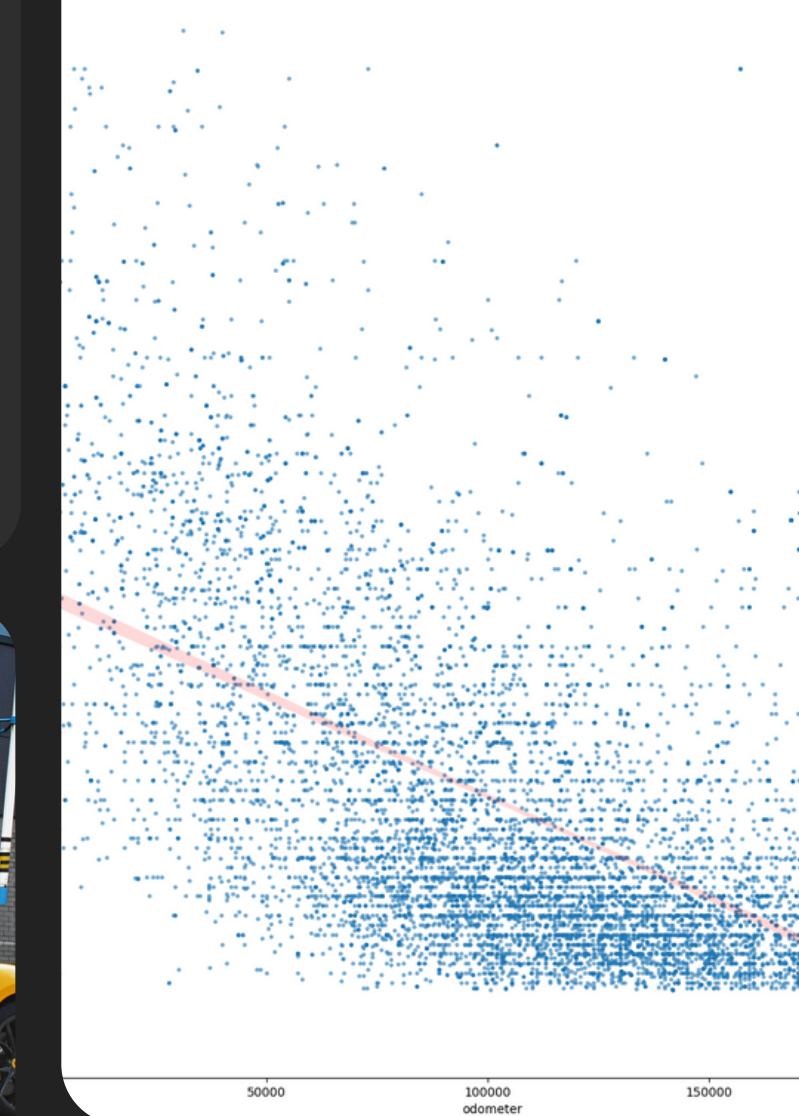
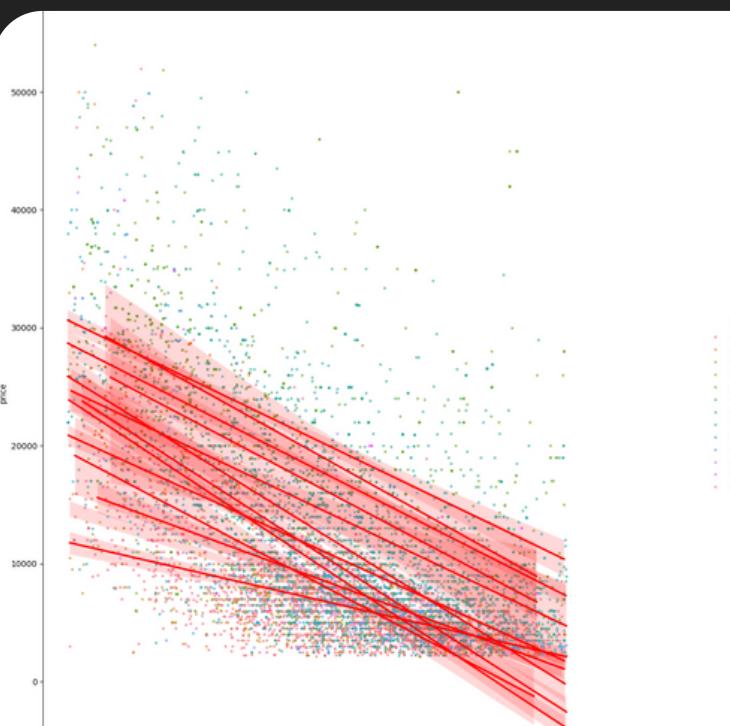
Transmission

Automatics are favored over manuals



c

07

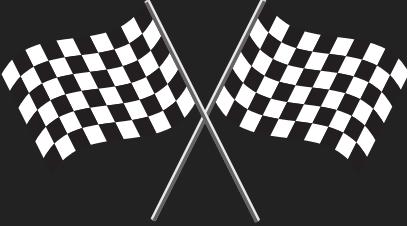


Our Model

The linear regression model created by our team takes all of the above features into consideration and is able to account for 65% of the variance in price based on this information.

As you can see to the left, the most significant features of age and mileage were not displaced or changed when taking other factors into account. They remain the leading factors.

...



08

Use Case

The screenshot shows a mobile application interface for valuing a car. At the top, it displays "My Car's Value" for a "2015 BMW 6 Series 640i Coupe 2D" located "near Lancaster, PA 17601". It shows a "Mileage: 30,000" and a "4.7 ★ (15 Ratings)". Below this, there are two notifications: "Recalls: 1 Recall Found" and "Next Service: 35,000 miles". The main section is titled "1 Compare Your Values" and includes tabs for "Instant Cash Offer", "Trade-in" (which is selected), "Private Party", and "Donate Your Car". Under the "Trade-in" tab, it shows a "Trade-in Range \$22,571 - \$26,188" with a "Trade-in Value \$24,380". A green progress bar indicates the value is "Good". To the right, there is a "Trade-In Value" section with a description, "Average Time to Trade-In: 1-7 DAYS", "How Much to Expect: \$\$\$", and a "Level of Effort: Medium". At the bottom, it says "Valid for ZIP Code 17601 through 09/20/2020" and shows a "4.7 ★ (15 Ratings)" rating.

Example

After completing the model, an theoretical vehicle evaluation was run and a price was predicted for this.

The Car

2015 BMW 6 Series with 30k miles.

The Outcome

Our model predicts a price of approximately \$23,000. To validate this, we ran this same information through the Kelly Blue book, a trusted vehicle valuation tool. The value range according to KBB is \$22-26K. The operative value of our tool being that large data sets (Auctions perhaps) can be run at scale via the model, as opposed to single, time intensive lookup.

The Next Steps

09

Larger Dataset

Obtain data spanning additional years.

Additional Datasets

Acquire aggregate data for historical vehicle pricing trends. Perhaps join data with additional features and more structured variables/definitions

More Specific Geospatial Analysis

CL data includes Lat Long values. This feature fell outside the immediate scope of this analysis, but could be used to further refine future iterations.

Further Refining Data

Employ RegEx in an attempt to better standardize free text inputs, such as vehicle model.

•••

OUR TEAM:
CHRISTOPHER YEAGER

Thank You

Developing insightful and practicable analysis and visualization products, to impact business in a meaningful way via data driven policy development and strategies.

YEAGER
AUTOWERKS
AUTO SALES CONSULTING