Used Cars

By: Alex Moore, Andrew Wick, Jessica Price, and Samrawit Basazinew

Table of Contents

Purpose

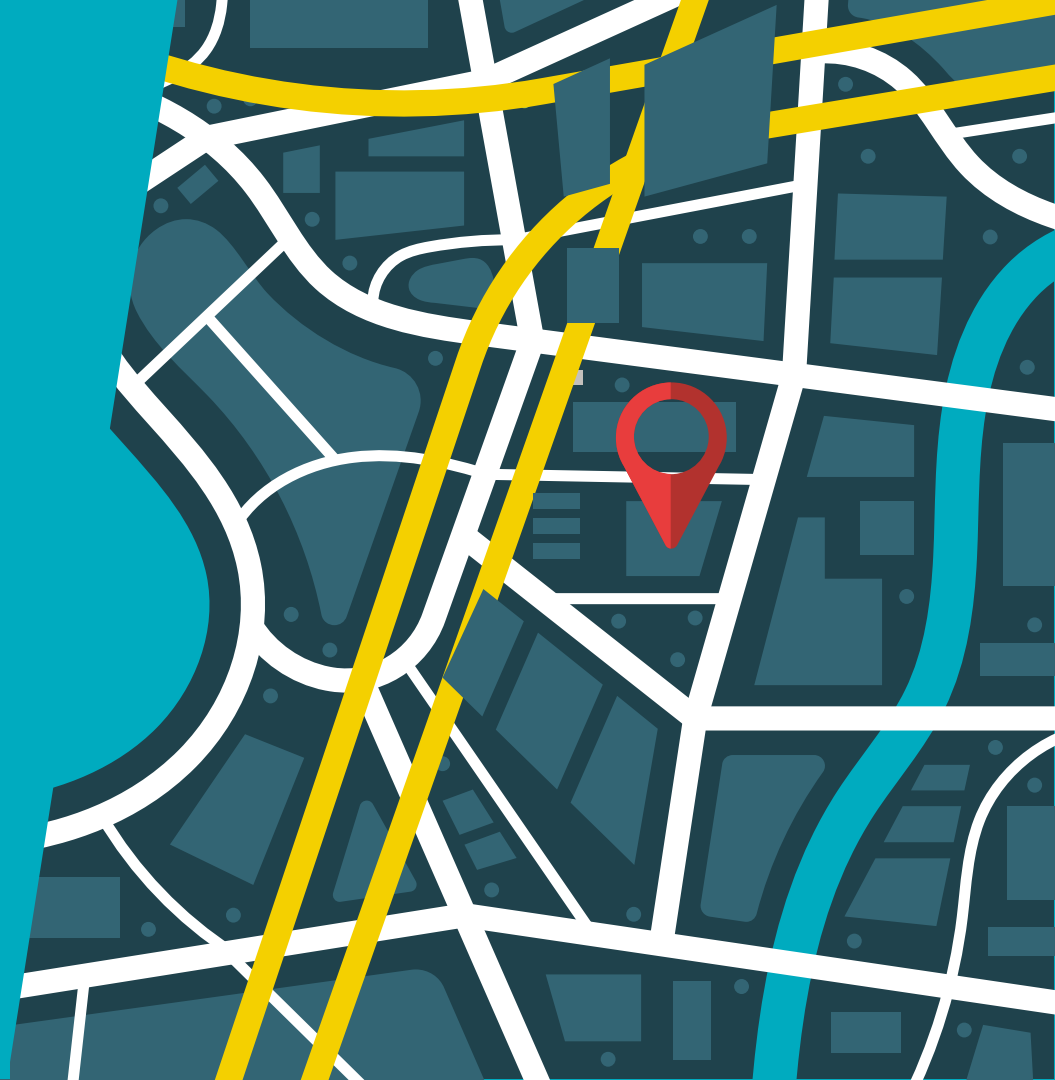
The purpose of this project was to determine the price of a used car based on various features. This tool can assist consumers looking to sell their used car by providing a fair and reasonable price estimate, based on comparable vehicles with similar features in the market.

Inspiration

- The reason we choose this dataset was because we knew that with the number of rows it contained, we would be able to create a great machine learning model. Even after cleaning and removing null values we still had over 600k rows, which we then lowered to 150k rows.
- We also decided on this dataset because of the smaller number of features within the dataset. This would help with keeping a smaller amount of features to be needed in the model and reduce the time to run the model.
- Finally, we also wanted a csv file so that it would be easier to load into a jupyter notebook file to build the model as well as load into Tableau.

Design Concepts

For the color design, we chose a simple, visually pleasing palette that's easy to read and navigate for users of all ages. We ensured sufficient contrast for accessibility and used a consistent color scheme across dashboards to highlight key patterns and outliers. Our goal was to balance aesthetics with functionality, making the tool both informative and engaging.



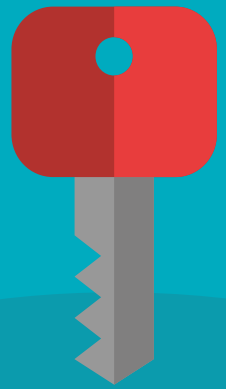
Research Questions

How can a used car's price be predicted from its features?

Which states have the highest used car prices by make, model, and body type?

Which makes and models are the most commonly used?

Live Demonstration



Limitation/Bias

There is a limited amount of data available for cars with higher price

Preprocessing the data took some time before we could run the model

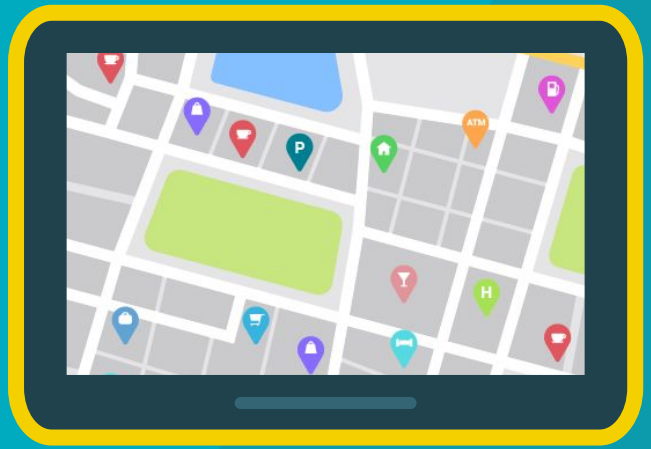
The dataset does not include the accident history of the cars, which is a key factor that can significantly affect a vehicle's value

Future Work

- Continue on preprocessing and look for better ways to improve the models performance.

Conclusion

In conclusion we built a linear regression model to predict what the price of a used car could go for on the open market depending on its various features. Although we would have liked more time to improve the model with more preprocessing and using our preferred model type, our linear regression model still gives a great predicted price of the vehicle in question.



Questions?

Resources

Kaggle