Southern Methodist University

Data Science Bootcamp (09/2023)

Project 3: TITLE

Team 2 Members: Misha Borunda, Damarje Brown, Kimberly Childers, Jason Cisneros, and Raheem Yusuff

Executive Summary

Project 3 encompasses a data visualization and data engineering track for students to showcase their skills. Team 2 has found an open-source dataset to manipulate and answer Project 3 requirements. A story will be revealed by using visualizations and implementing ETL skills to design a robust database.

Background

Craigslist is America’s go to for classifieds as newspaper popularity dwindles with the onset of the internet. It is an American company that is privately held and was founded in 1995. For anyone in the data science field, this website offers an abundance of nationwide raw data to draw on for analysis. Specifically for this project, Team 2 plans to investigate the world of used vehicles to explore data based on manufacturer, model, odometer, drive, size, type, and other data points not listed here.

Research Questions

1. What state has the highest selling price for used cars?
2. What State has the most amount of used cars for sale?
3. What manufacturer has the most amount of used cars for sale in the market?
4. What is the most popular type of car for sale in Texas vs New York?

Dashboard Design

\*Maybe we can create an image on power point instead of trying to write out a layout? \*

Roles and Responsibilities

Data Visualization Track:

Data Engineering Track:

Administrative (stand-up/proposal/PP/report/meeting invites/QC of requirements/convert final products to pdf):

Misha Borunda

Damarje Brown

Kimberly Childers

Jason Cisneros

Raheem Yusuff

Conclusion/Future Work

This research can assist consumers with determining the best used car to meet their needs and educate a potential seller on the value of their car. The location, car model, odometer reading, and other factors explored may assist the consumer and seller of the best deal they should expect for a used vehicle. Car companies may even benefit from this study to determine if the field is saturated with a certain model and how well they retain their value.

References

**Used Cars Dataset**

Reese, Austin

[**https://www.kaggle.com/datasets/austinreese/craigslist-carstrucks-data**](https://www.kaggle.com/datasets/austinreese/craigslist-carstrucks-data)(CCO: Public Domain)

**AustinReese/UsedVehicleSearch: Filter vehicles from Craigslist**

Reese, Austin

[**https://github.com/AustinReese/UsedVehicleSearch/tree/master**](https://github.com/AustinReese/UsedVehicleSearch/tree/master)