Sports Car Feature Importance Analysis

Files in Folder:

* Sports\_car\_Part\_1.ipynb is a notebook file that contains code, comments, markdowns and data analysis. This code dives into ‘Sport\_car\_price.csv’ and handles data cleaning, feature engineering to expand the dataset, and data analysis with visual representations.
* Sports\_car part 2.ipynb is a notebook file that contains code, comments, markdowns and data analysis. This code runs the dataset 2, cleaning, preprocessing, and integration into visualization while also comparing dataset 1 models.
* Sports Car Analysis Write-Up is a txt file that contains the explanation of our project's purpose followed with the support of our plots, in depth data analysis, and conclusions.
* Sports Car Data Analysis Presentation is a presentation containing slides that were presented to the class. These highlighted the important details leading to our project’s conclusion.

Table of Contents

* General Information
* Setup
* Usage
* Project Status
* Acknowledgements
* Contact

General Information

* What makes a sports car a sports car?
* Our feature helps to identify which sports car features or variables have the most impact on a target variable such as price or choice. This analysis provides insights into the factors that contribute to the outcome in order to get a better understanding of the decision-making process. The purpose of this feature can help car dealerships, market analysts, car manufacturers, or an individual looking to buy a sports car.

Setup

Reference the links below to download the necessary csv files if for any reason, the csv file within this folder is not loading:.

* <https://www.kaggle.com/datasets/rkiattisak/sports-car-prices-dataset>
* <https://www.kaggle.com/datasets/vspencer88/sports-car-choice-data?select=sportscar_choice_long.csv>
* When downloading the csv files for the respective notebooks, put it in a path that you want the code to read from. Through notebooks, there are no installing dependencies. Notebooks already have access to the import platforms.
* Several python libraries and programs were utilized in the code, Colab Notebook should include all pre-installed Python libraries.

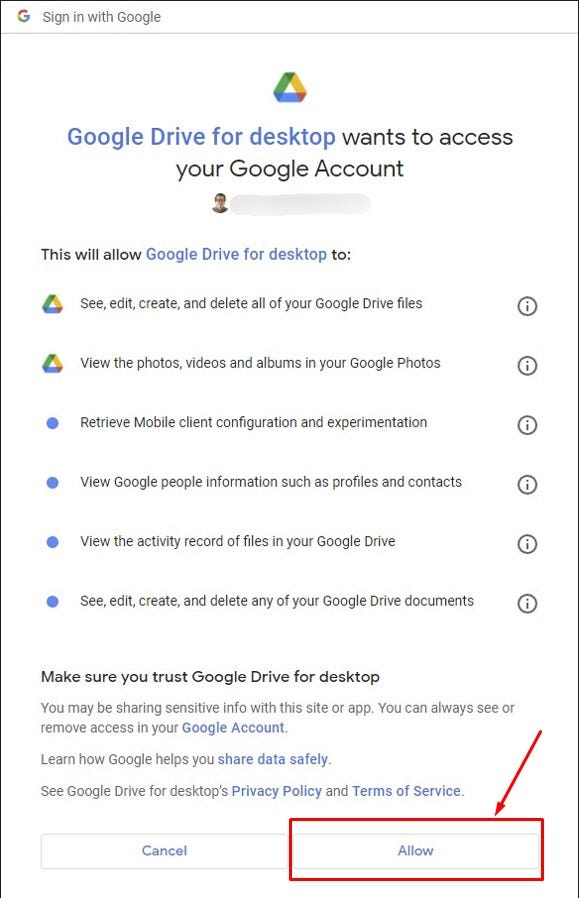
Usage

How does one go about using it?

Make sure to have this part of the code input in the beginning of running the code:

For running Sports\_Car\_Part\_1:

1. Download ‘Sport\_car\_price.csv’
2. This message should come up, please allow access to the file from your google drive in order to run the codes:



from re import X

from google.colab import drive

drive.mount('/content/drive')

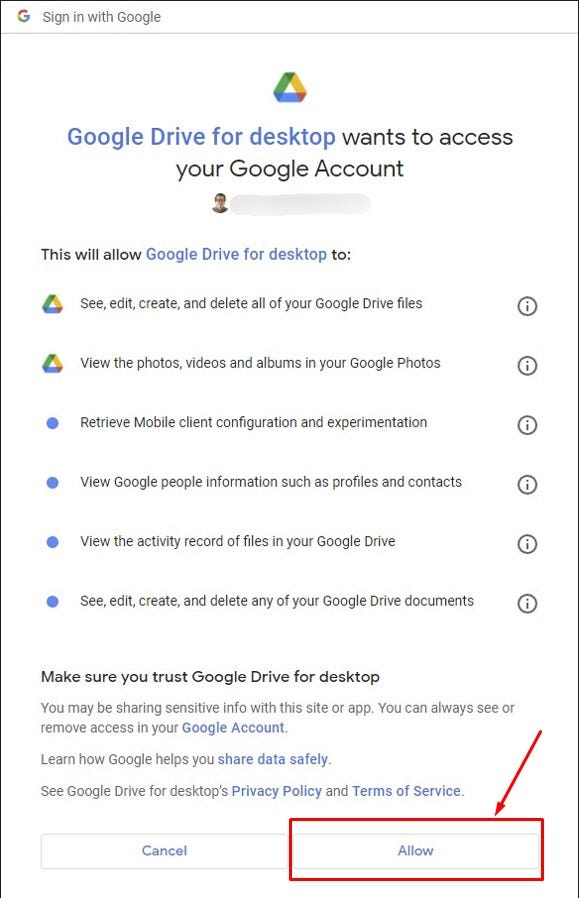
# Edit Path to wherever you store the csv file at

sports\_car\_df = pd.read\_csv('drive/MyDrive/Sport\_car\_price.csv')

Note: Edit the path to wherever you store the csv file to. Make sure all names of path and file match.

Sports\_car Part 2:

1. Download ‘Sport\_car\_price.csv’ and ‘sportscar\_choice.csv’ onto your google drive.
2. This message should come up, please allow access to the file from your google drive in order to run the codes:



1. Please ensure the file names are saved exactly like shown below ‘Sport\_car\_prices.csv’ and ‘sportscar\_choice\_long.csv’ if the file names are saved under a different name the code will not run.
2. Make sure to mount google drive into the colab environment before starting using drive.mount(‘/content/drive’)

from re import X

from google.colab import drive

drive.mount('/content/drive')

# !ls "/content/drive/MyDrive/Colab\_Notebooks"

# Edit Path to wherever you store the csv file at

sports\_car\_df = pd.read\_csv('drive/MyDrive/Sport\_car\_price.csv')

file\_path = 'drive/MyDrive/Sport\_car\_price.csv'

df = pd.read\_csv(file\_path)

### second dataset ####

from re import X

from google.colab import drive

drive.mount('/content/drive')

# Edit Path to wherever you store the csv file at

purchase\_data = pd.read\_csv('drive/MyDrive/sportscar\_choice\_long.csv')

filepath = 'drive/MyDrive/sportscar\_choice\_long.csv'

purchase\_data = pd.read\_csv(filepath)

Project Status

Project is: Complete

Acknowledgements

Give credit here:

* Dr. Patrick Seyed
* Chetana Shinde
* This project was based on several data analysis examples:
* Dataset #1: (<https://www.kaggle.com/datasets/rkiattisak/sports-car-prices-dataset>)
* Dataset # 2(<https://www.kaggle.com/datasets/vspencer88/sports-car-choice-data>)
* (<https://www.kaggle.com/code/kerneler/starter-sports-car-choice-data-488e6e2e-f>)
* (<https://www.kaggle.com/code/borispelichek/exploratory-data-analysis-car-csv#8.-Conclusion>)
* (<https://www.kaggle.com/code/shreyasnaphad/sports-car-price-prediction>)
* (<https://github.com/spensorflow/Marketing-Analytics---Choice-Modeling-Sports-Car-Sales/blob/master/Choice_Modeling_Sports_Car_Sales.md>)
* (<https://medium.com/@Nivitus./car-price-prediction-using-machine-learning-6d7678cdbf94>)
* (<https://github.com/ritaly/README-cheatsheet>).
* We considered the notes and comments made by Dr. Patrick Seyed mentioned that our dataset was too small, so we thought of ideas on how we can expand the data that can potentially lead to a more interesting and proper analysis and conclusions.

Contact

Created by James Davis ([jdavid07@umd.edu](mailto:jdavid07@umd.edu)) and Zannate Malik ([zmalik1@umd.edu](mailto:zmalik1@umd.edu))