

CarSamjho.com

About The Project

The automobile industry is flooded with options, it is a dream for a middle-class family to buy a car, and with this incentive, the company is doing everything possible to give a smooth ride utilising analytics. As a result, I'm attempting to develop an analytical tool to provide Data-Analysis for User given Dataset for the automotive sector as per Manufacturing Industry Employees as a user to take informed decisions.

X-factor

As the manufacturing industry must make selections based on car reviews. The dataset is being harmed by FAKE/Misleading Reviews, which is a big problem. It must be identified and extracted from the data in order to make excellent business decisions.

- Industry Employees are the User for the Project to Take informed Decisions by this tool.

Hence my Project will also solve this Problem by doing Real-time Fake Review Detection.

Built With

- Django
- Html/css/javascript
- Dbsqlite Database
- bootstrap
- jupyter-notebook(web scraping using beautiful-soop)
- Seaborn For visualizations.

Prerequisites

Install Requirement.txt file using Pip.

Installation

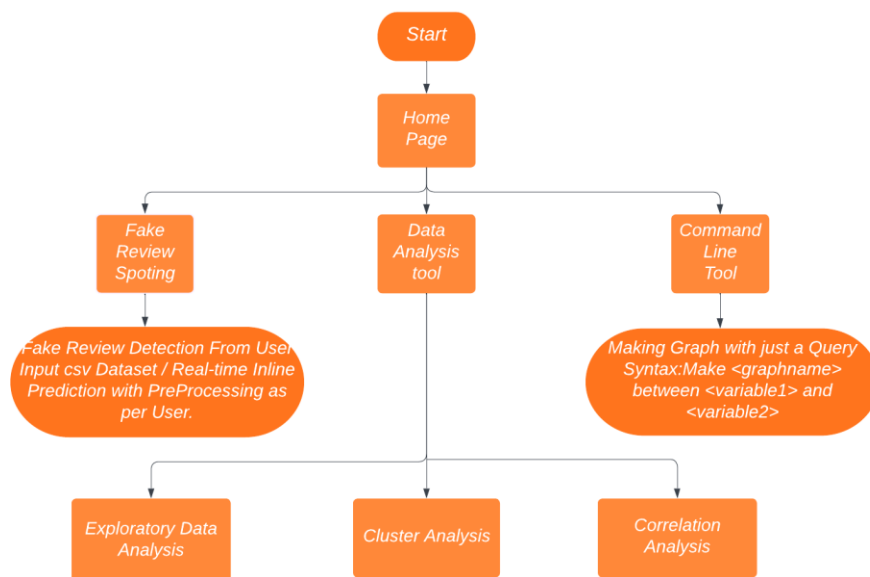
1. Clone the repo
2. Cd (check where manage.py must be present)
3. Use python manage.py makemigrations

4. Followed by python manage.py migrate
5. The project setup is completed and ready to start. Use python manage.py runserver to Start the project in local Host.

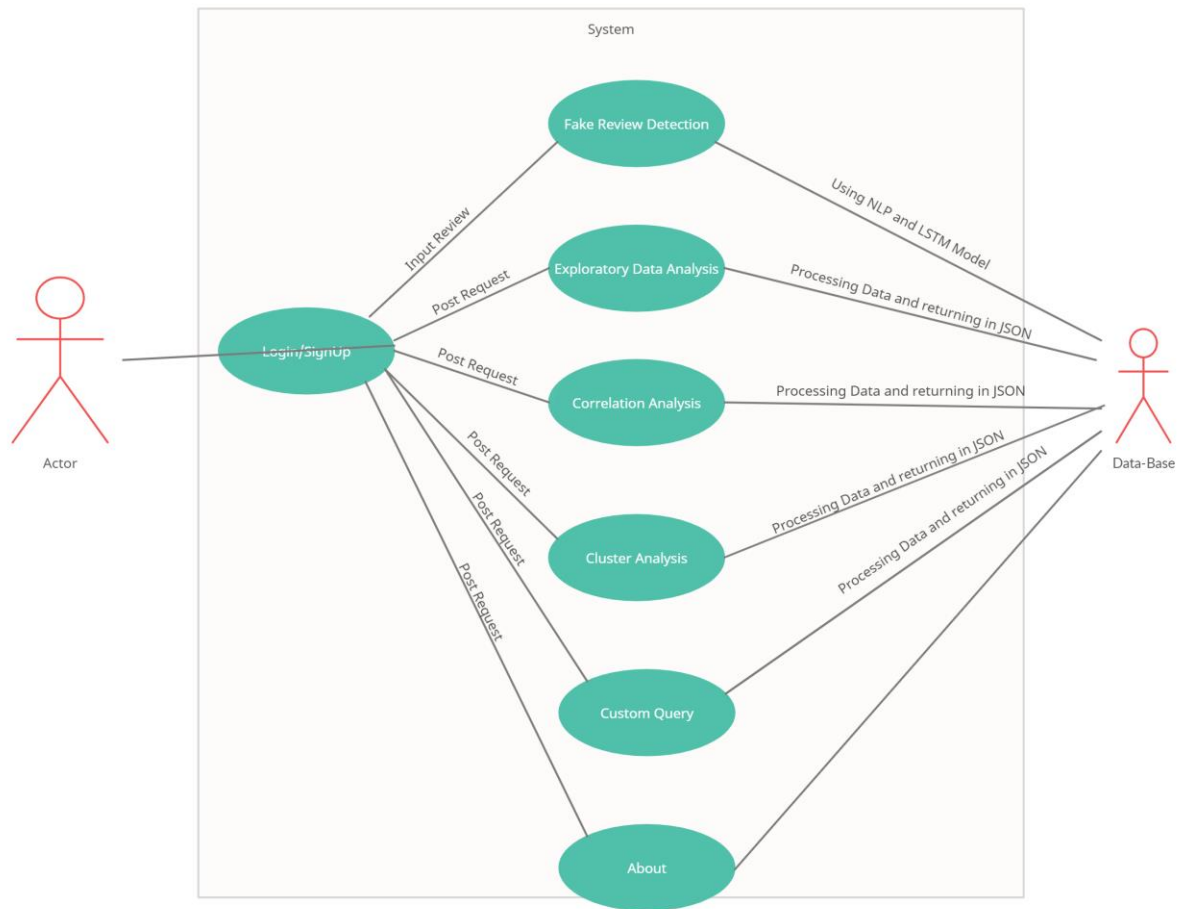
Features

- Home Page - DashBoard
- Fake-Review Detection (Real-time)
 - Web Scrapping From Amazon review to Train model.
- Data Analysis Tool (for custom dataset)
 - Exploratory Data-Analysis
 - Cluster Analysis
 - Correlation Analysis
- Command-line Query for Generating graphs
- SignUp/SignIn (for particular user)

Flow-Chart



USE-CASE diagram



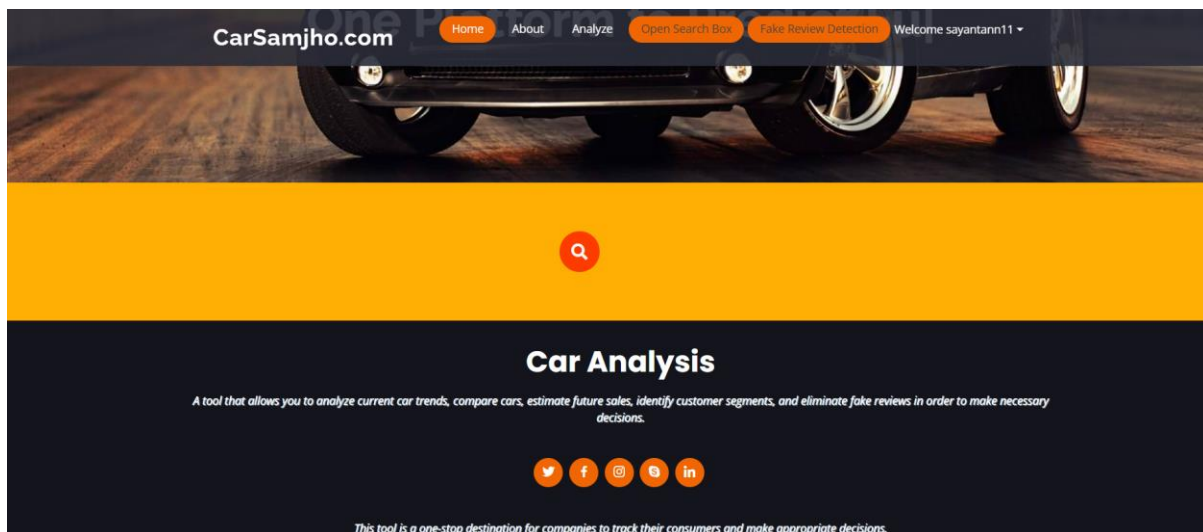
File-Structure

Name	Date modified	Type	Size	
.github	27-05-2022 10:07	File folder		
asset	27-05-2022 10:07	File folder		
assets	27-05-2022 10:07	File folder		
dataset	27-05-2022 10:07	File folder		TestData-FakeReview / CarProduction or Engage Dataset
project	27-05-2022 10:07	File folder		Django Project Files
staticfiles	27-05-2022 12:01	File folder		
templates	27-05-2022 10:07	File folder		HTML Templates
webscraping	27-05-2022 12:23	File folder		
ABC.html	27-05-2022 10:07	Text Document	0 KB	
cars_engage_2022	27-05-2022 10:07	Microsoft Excel Co...	1,320 KB	Amazon Product Commnet by WebScraping
db.sqlite3	27-05-2022 13:19	SQLITE3 File	128 KB	
manage	27-05-2022 10:07	PY File	1 KB	
model2.pkl	27-05-2022 10:07	PKL File	41 KB	ML Trained Model in pkl format
nlTK	27-05-2022 10:07	Text Document	1 KB	
Profile	27-05-2022 10:07	File	1 KB	
README	27-05-2022 10:07	Markdown Source ...	6 KB	
requirements	27-05-2022 10:07	Text Document	1 KB	
testdata	27-05-2022 10:07	Microsoft Excel Co...	1 KB	
tfidfvec2.pkl	27-05-2022 10:07	PKL File	74,739 KB	Text to Vector using Vectoriser

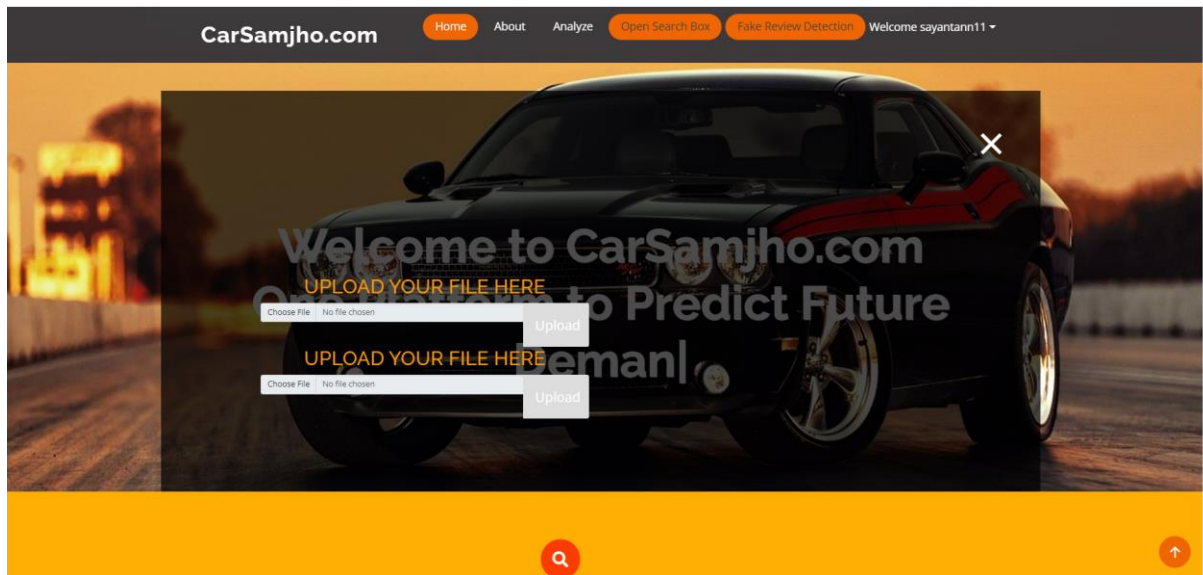
Home Page - DashBoard



Footer:



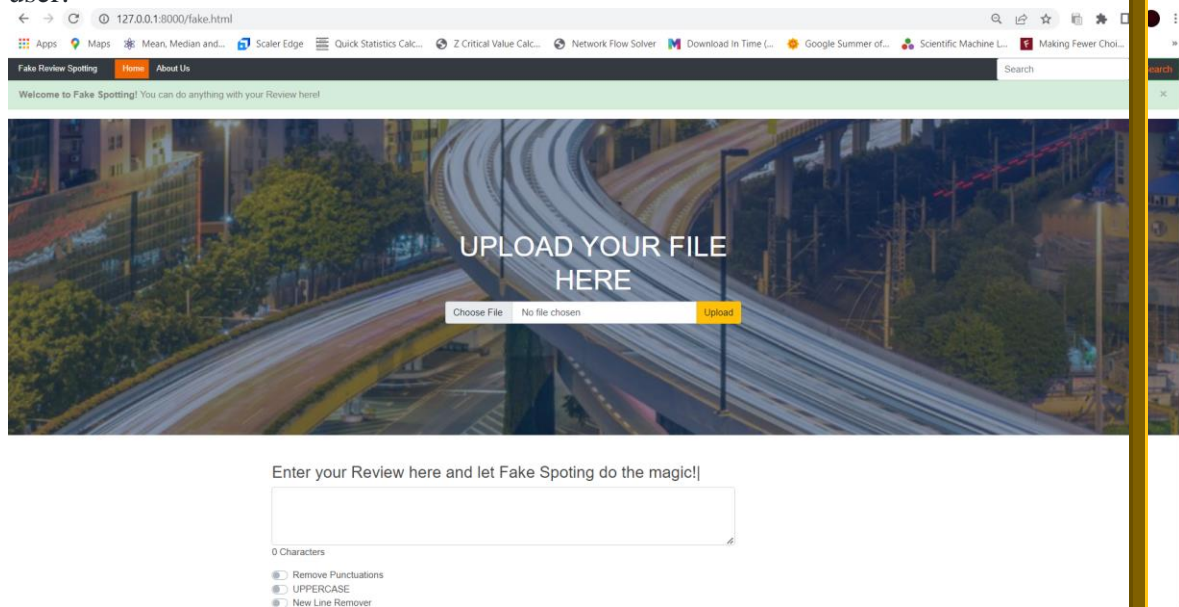
Hover Section (User can insert Dataset)



Fake Reviews Detection

Fake reviews make it extremely difficult for manufacturers to make informed judgments, therefore I decided to write a function to detect and remove fake reviews from the dataset for accurate demand and feature forecasts.

- TEXT box where user can Write its Query Whether it is Fake or Not / also can insert Fake Review excel dataset
 - Pre-processing of Text as per user.

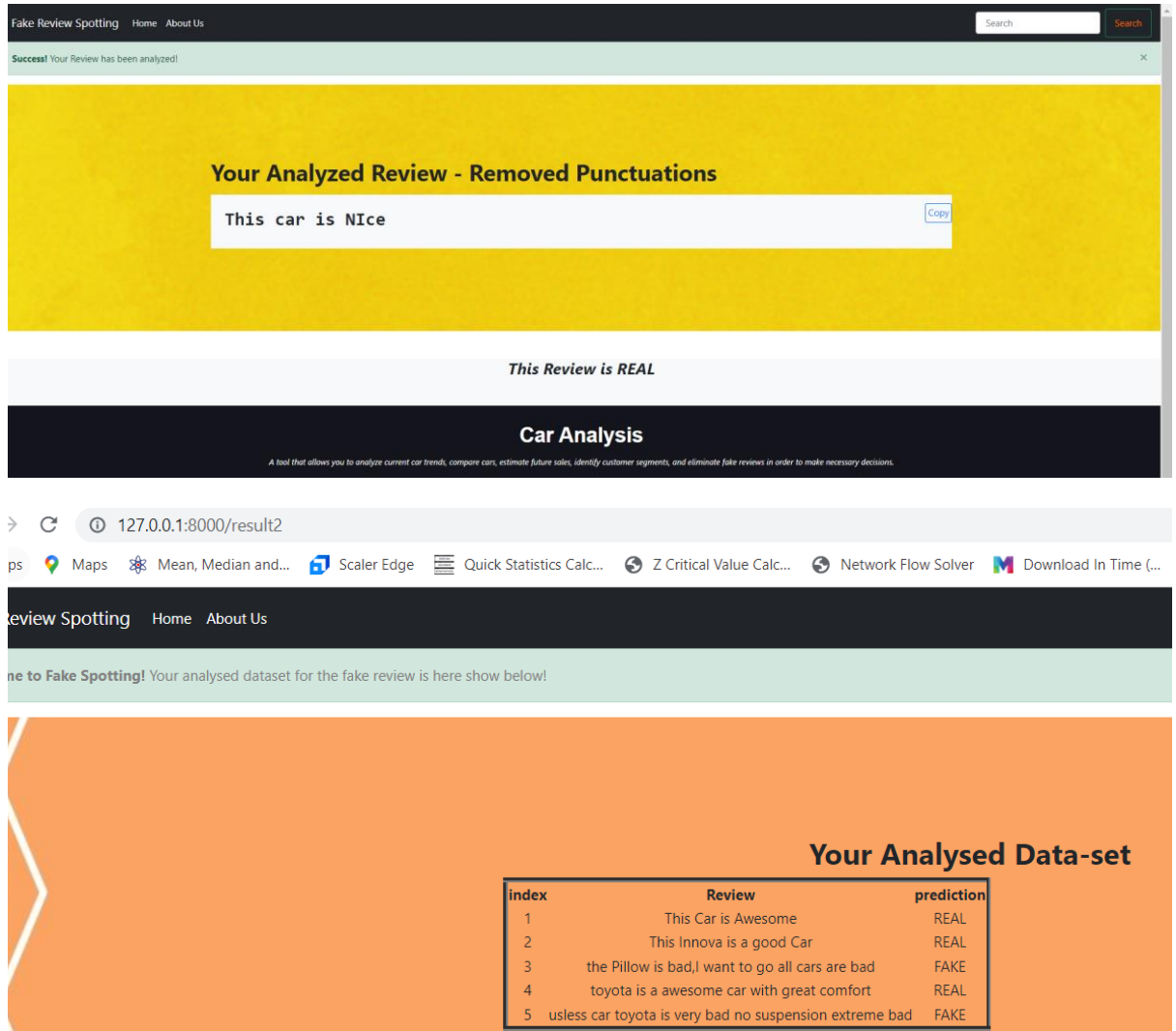


- Inline Character as well as Word Count feature

this car is nice

16 Characters

- Result Page with decisions:



Success! Your Review has been analyzed!

Your Analyzed Review - Removed Punctuations

This car is NICE

This Review is REAL

Car Analysis

A tool that allows you to analyze current car trends, compare cars, estimate future sales, identify customer segments, and eliminate fake reviews in order to make necessary decisions.

127.0.0.1:8000/result2

ps Maps Mean, Median and... Scaler Edge Quick Statistics Calc... Z Critical Value Calc... Network Flow Solver Download In Time (...)

review Spotting Home About Us

ne to Fake Spotting! Your analysed dataset for the fake review is here show below!

Your Analysed Data-set

index	Review	prediction
1	This Car is Awesome	REAL
2	This Innova is a good Car	REAL
3	the Pillow is bad,I want to go all cars are bad	FAKE
4	toyota is a awesome car with great comfort	REAL
5	usless car toyota is very bad no suspension extreme bad	FAKE

Data-Analysis-Tool

- In addition, I will provide a default analysis of the given dataset, including client groups, the most popular automobile specification combinations (engine type, fuel, mileage, and so on), the ideal time to introduce a new car, and so on. as it is capable of:
- After that, the user must Insert Dataset. It will take the user to the next page, where they can view the dataset and its features.

Material-Dashboard

EXPLORARY DATA ANALYSIS | CLUSTER ANALYSIS | CORRELATON GRID

Success! Your Asked Graph is Here!

Data Frame

make	model	car	variant	body_type	fuel_type	fuel_system	type	drivetrain	displacement	cylinders	mileage	power	torque	fuel_tank	height
Tata	Nano Genx	Tata Nano Genx	Xt	Hatchback	Petrol	Injection	Manual	RWD (Rear Wheel Drive)	624	2	23.6	37.48	51	24.0	165
Tata	Nano Genx	Tata Nano Genx	Xe	Hatchback	Petrol	Injection	Manual	RWD (Rear Wheel Drive)	624	2	23.6	37.48	51	24.0	165
Tata	Nano Genx	Tata Nano Genx	Xta	Hatchback	Petrol	Injection	Automatic	RWD (Rear Wheel Drive)	624	2	21.9	37.48	51	24.0	165
Tata	Nano Genx	Tata Nano Genx	Xm	Hatchback	Petrol	Injection	Manual	RWD (Rear Wheel Drive)	624	2	23.6	37.48	51	24.0	165
Tata	Nano Genx	Tata Nano Genx	Xma	Hatchback	Petrol	Injection	Automatic	RWD (Rear Wheel Drive)	624	2	21.9	37.48	51	24.0	165

- Three options are available in the navigation bar. This will traverse according to the user's actions

Material-Dashboard

EXPLORARY DATA ANALYSIS | CLUSTER ANALYSIS | CORRELATON GRID

Success! Your Asked Graph is Here!

Data Frame

make	model	car	variant	body_type	fuel_type	fuel_system	type	drivetrain	displacement	cylinders	mileage	power	torque	fuel_tank	height
Tata	Nano Genx	Tata Nano Genx	Xt	Hatchback	Petrol	Injection	Manual	RWD (Rear Wheel Drive)	624	2	23.6	37.48	51	24.0	165

CLUSTER ANALYSIS | CORRELATON GRID

FuelType vs Displacement
FuelType vs Mileage
FuelType vs Power
FuelType vs Price

variant	body_type	fuel_type	fuel_system	type	drivetrain	displacement	cylinders	mileage	power	torque	fuel_tank	height
Xt	Hatchback	Petrol	Injection	Manual	RWD (Rear Wheel Drive)	624	2	23.6	37.48	51	24.0	165

Exploratory Data analysis Tab



- Histogram of Price
- Dominating car BodyType
- BoxPlot for Price (Outlier analysis)
- engine size comparison
- Relationship for Price and Power

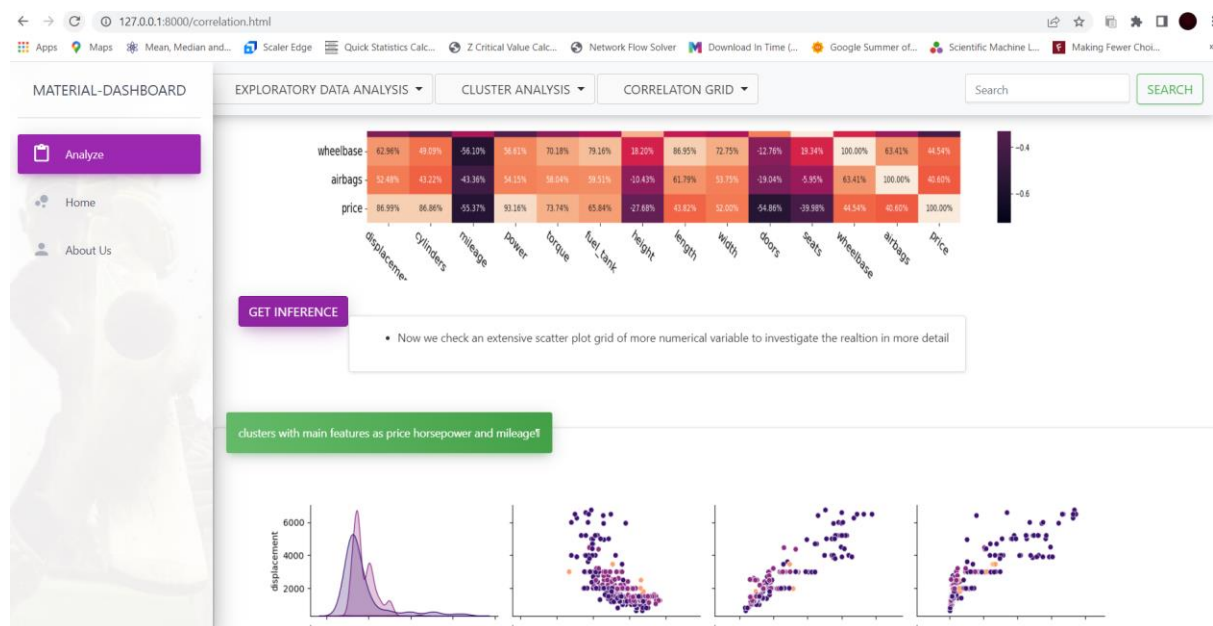
Cluster analysis Tab



- Cluster the cars types and cars using k-means algorithm
- Price and horse power with cluster price
- Power and Mileage after clustering
- Engine size with Fuel tanks
- Average price with each cluster
- Finding potential strategic groups

- Cars body type with each cluster

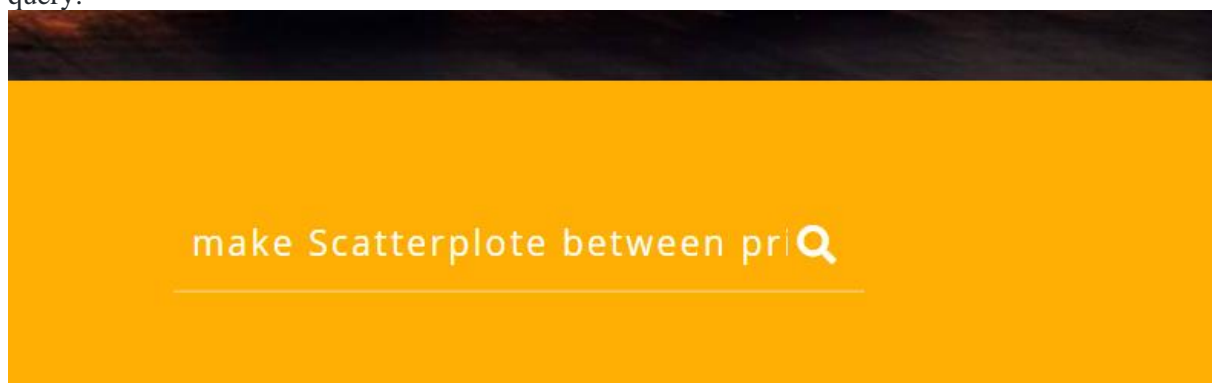
Correlation Grid



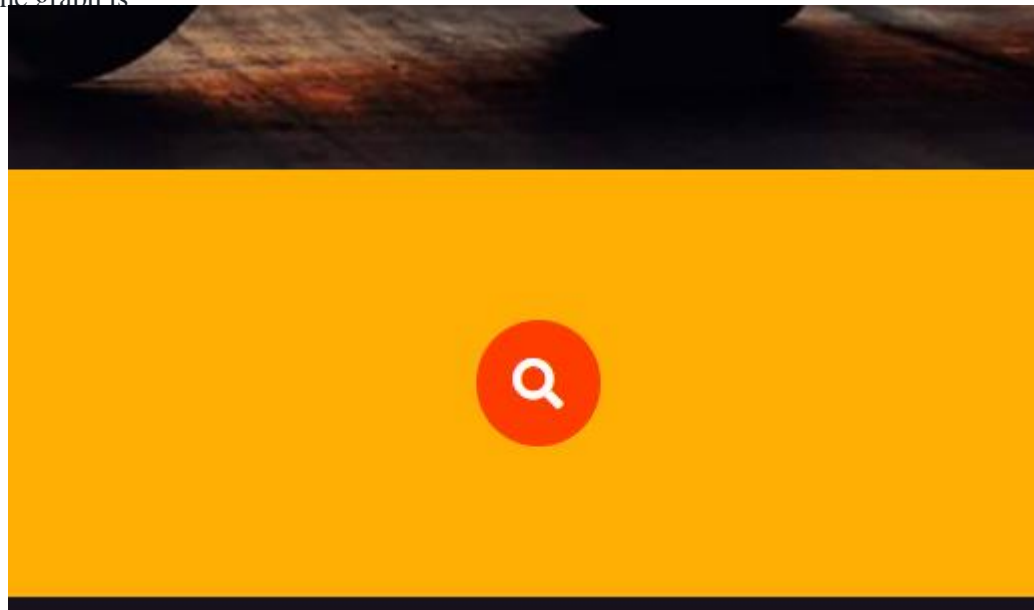
- Correlation Matrix (to know which features all strongly correlated)
- Extensive scatter plot grid of more numerical variable to investigate the relation in more detail

Command-line Query for Generating graphs

- A command line tool where the user can dynamically enter a query.

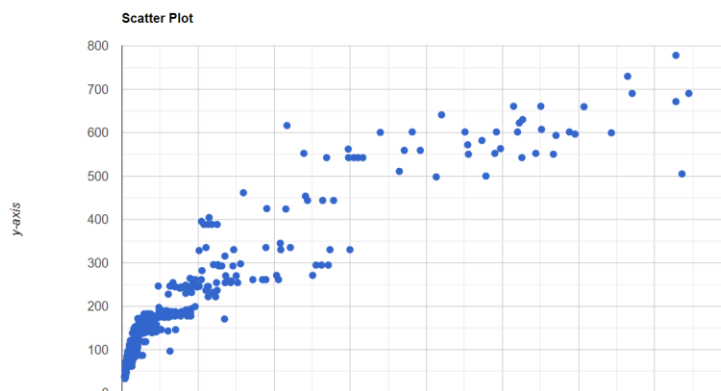
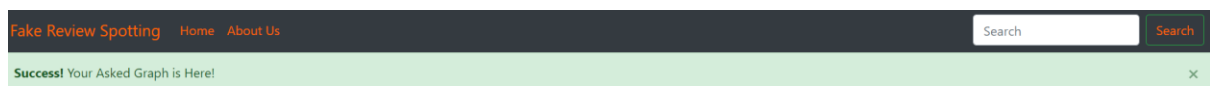


- the required visualisation will be generated with its inference, allowing the user to understand what the graph is



doing.

- Currently which option is limited to only 3 graphs as per prototype(line,scatter,bar)



Login/Logout- to have profiles based on user and track their work.

com

Review Detection Login SignUp

SignUp Here x

Username

Choose a unique username

First Name

Enter Your First Name

Last Name

Enter Your Last Name

Email address

name@example.com

Choose a password

Choose Your Password

Confirm Password

Enter your password again

Submit

Platform to Predict Future Demand!

Login with a success message

Welcome sayantann11 ▾

Message : Successfully Logged In x