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(webscraping using beautiful-soop)
- Seaborn For visualizations.

Prerequisites

Install Requirnment.txt file using Pip.

Installation

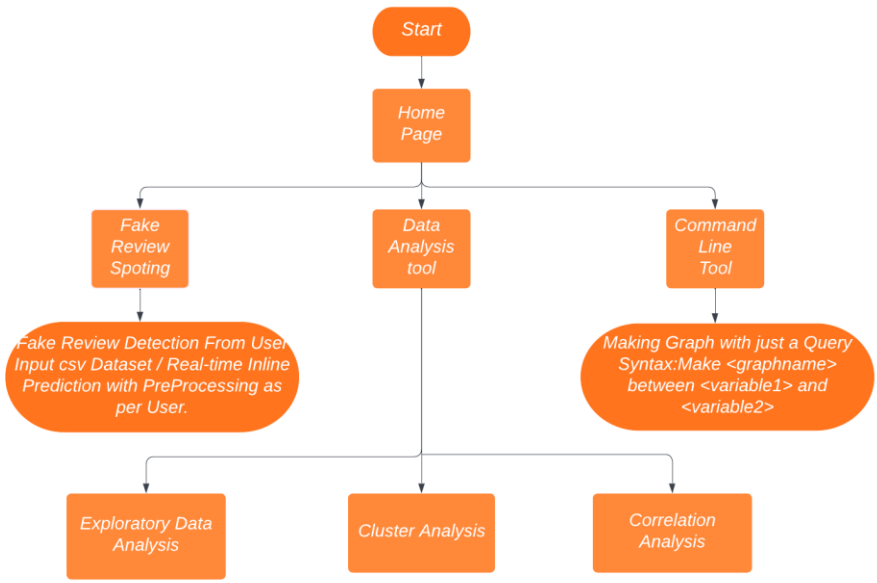
1. Clone the repo
2. Cd (check where manage.py must be present)
3. Use python manage.py makemigrations

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.

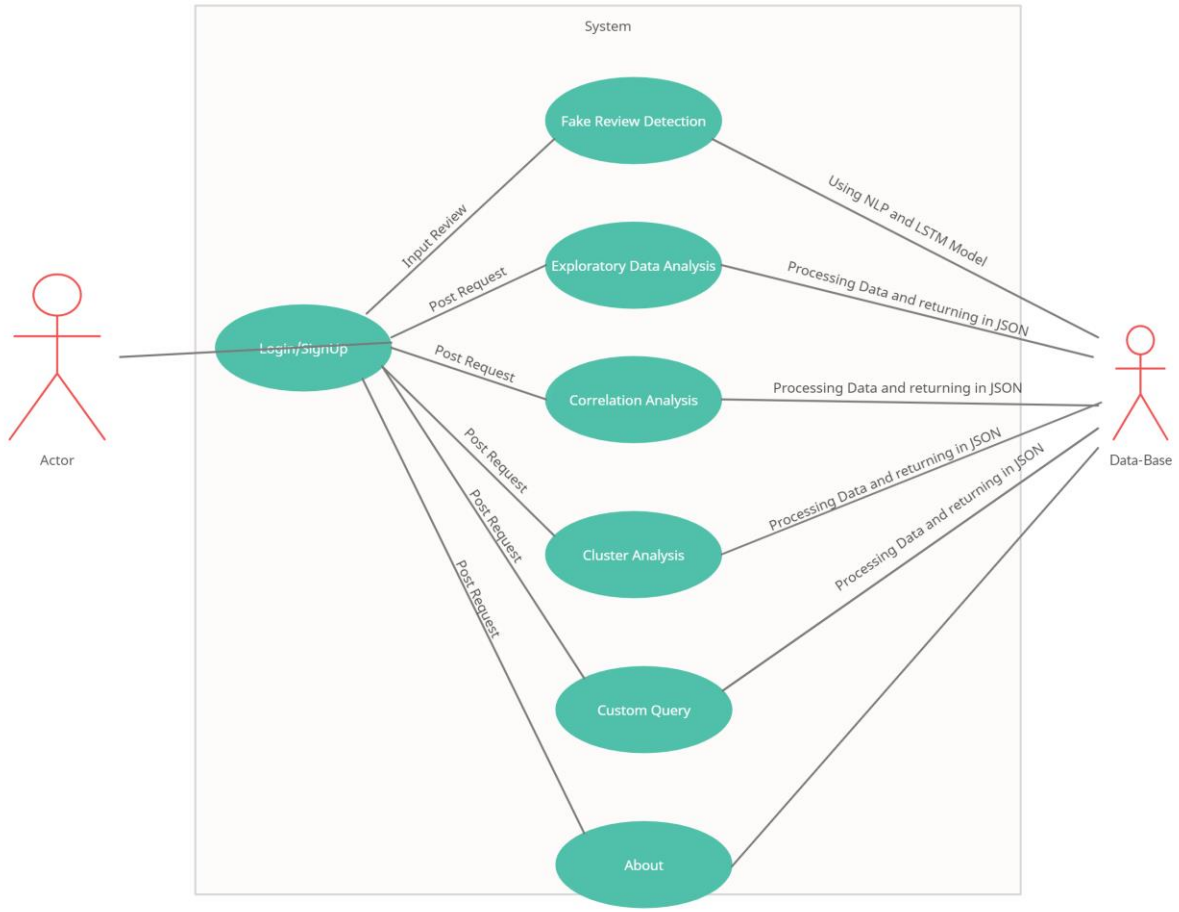
Features

- Home Page - DashBoard
- Fake-Review Detection (Real-time)
    - Web Scraping 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	
project	27-05-2022 10:07	File folder	
staticfiles	27-05-2022 12:01	File folder	
templates	27-05-2022 10:07	File folder	
webscrapting	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
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
nltk	27-05-2022 10:07	Text Document	1 KB
Procfile	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
tfidfvect2.pkl	27-05-2022 10:07	PKL File	74,739 KB

TestData-FakeReview / CarProduction or Engage Dataset

Django Project Files

HTML Templates

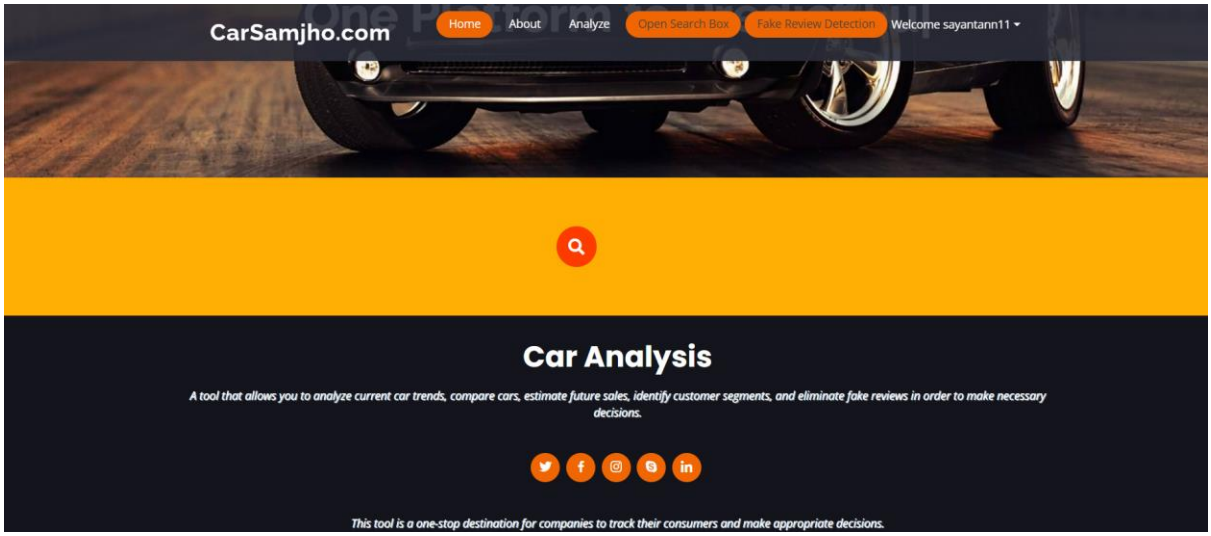
Amazon Product Commnet by WebScraping

ML Trained Model in pkl format

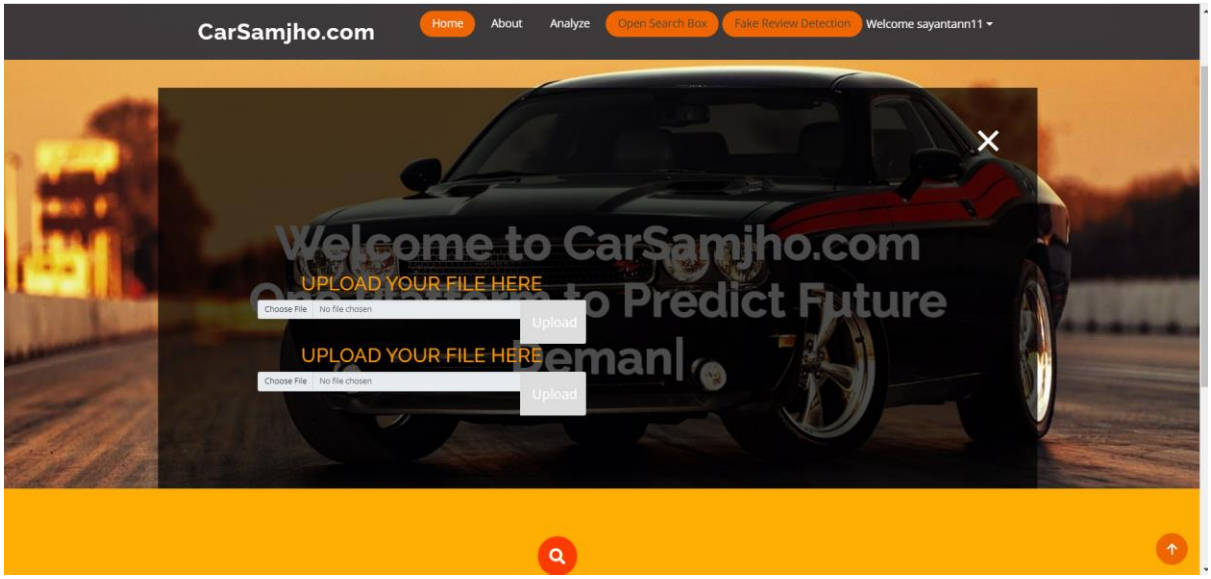
Text to Vector using Vectoriser

Home Page - DashBoard

Footer:



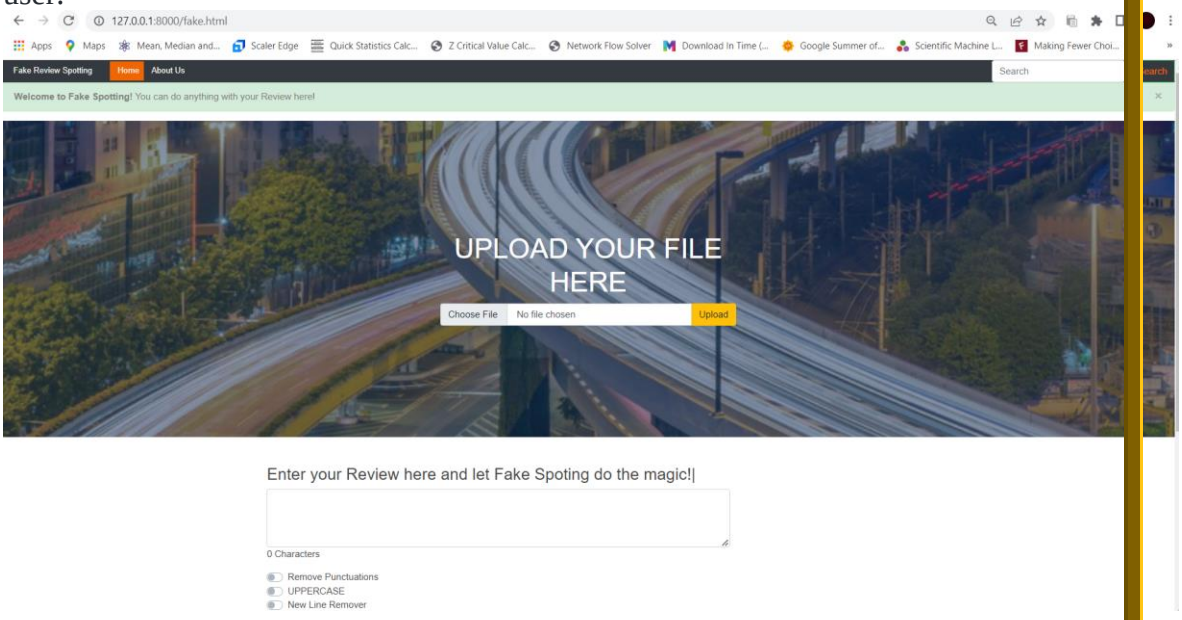
Hover Section (User can insert Dataset)



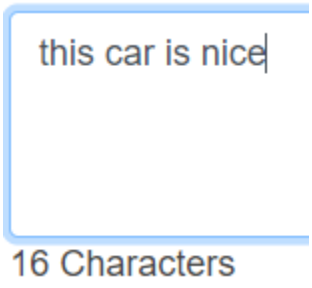
Fake Reviews Detection

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.

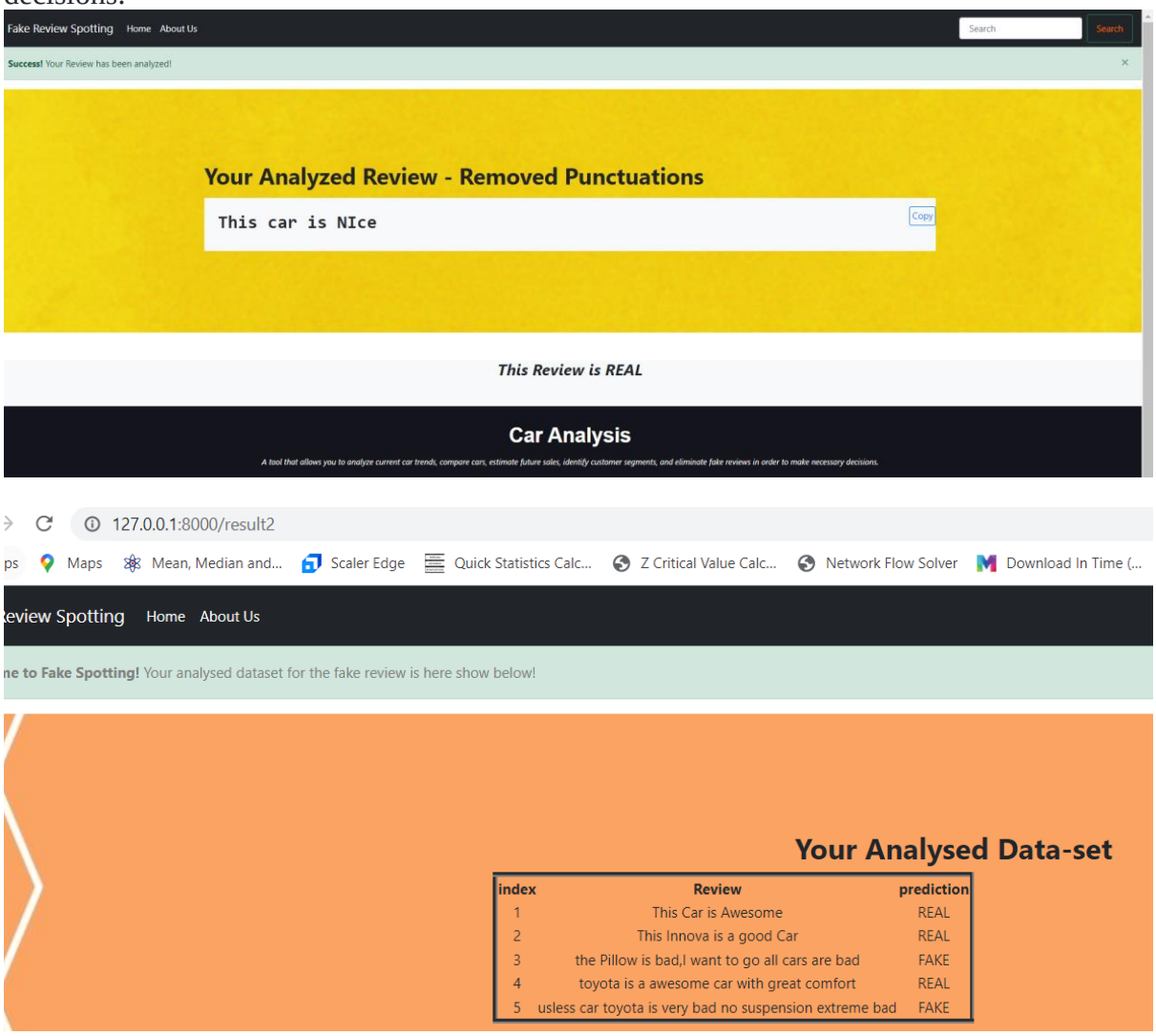
- TEXT box where user can Write its Query Whether it is Fake or Not / also can insert Fake Review excel dataset
    - Pre-processing of Text as per user.



    - Inline Character as well as Word Count feature



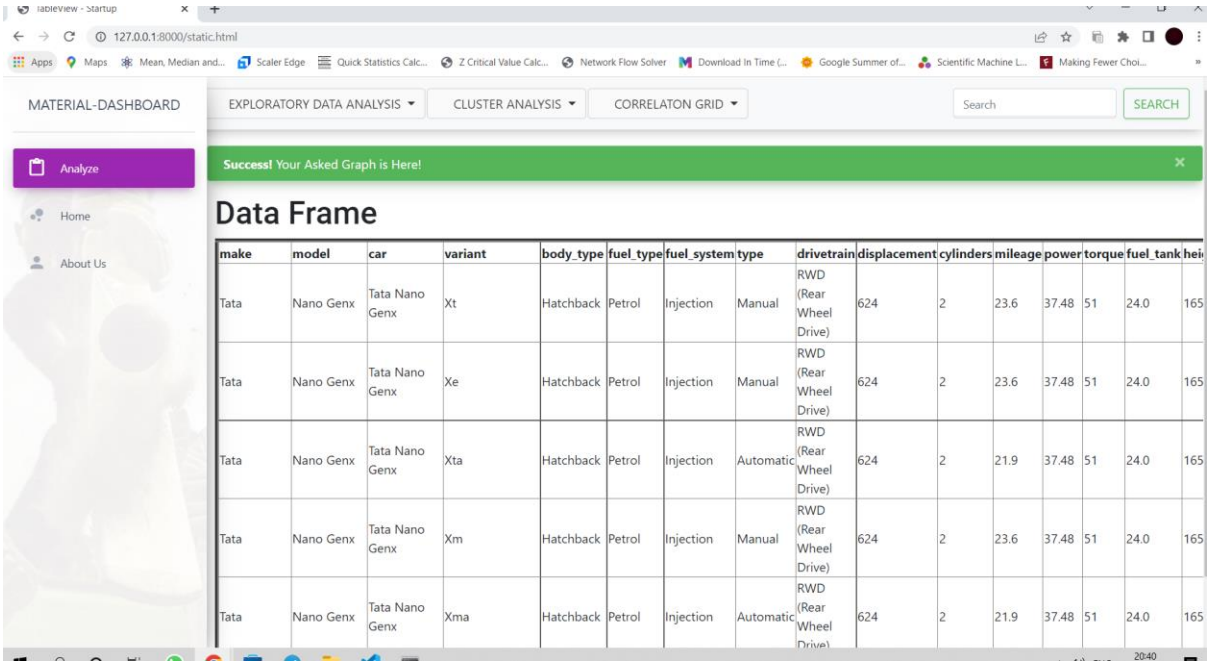
- Result Page with decisions:



Data-Analysis-Tool

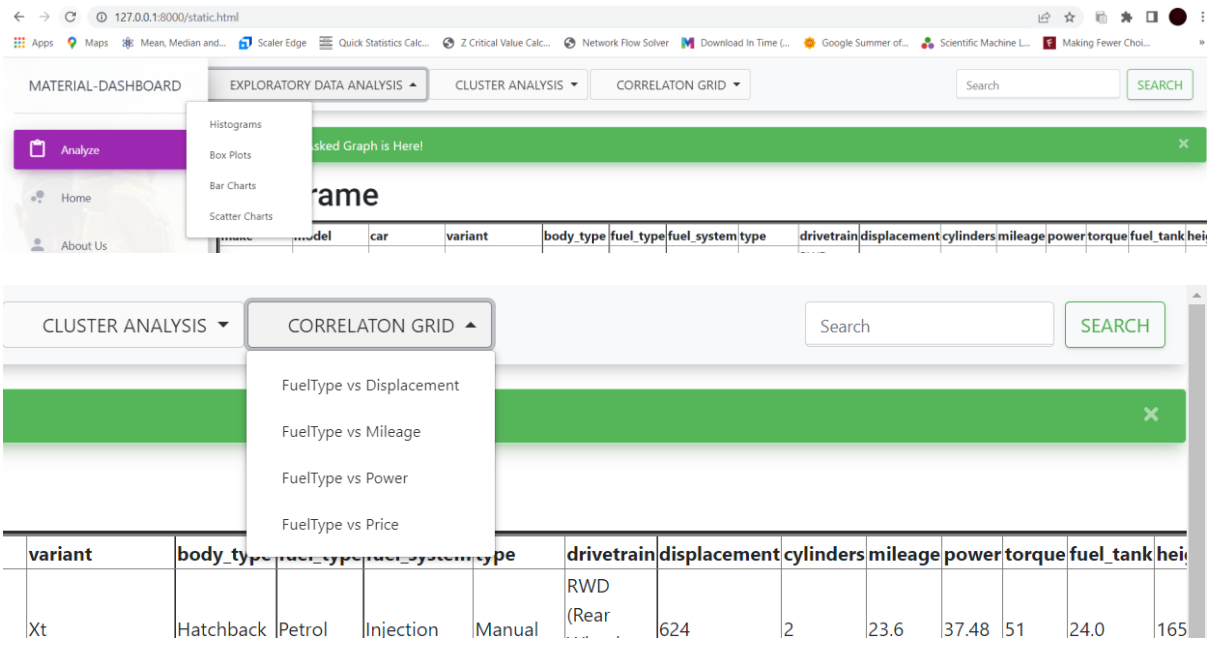
- In addition, I will provide a default analysis of the given dataset, including client groups, the most popular automobile specification combinations (engine type, fuel, mileage, and so on), the ideal time to introduce a new car, and so on. as it is capable of:

- 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.



make	model	car	variant	body_type	fuel_type	fuel_system	type	drivetrain	displacement	cylinders	mileage	power	torque	fuel_tank	hei
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



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

Exploratory Data analysis Tab

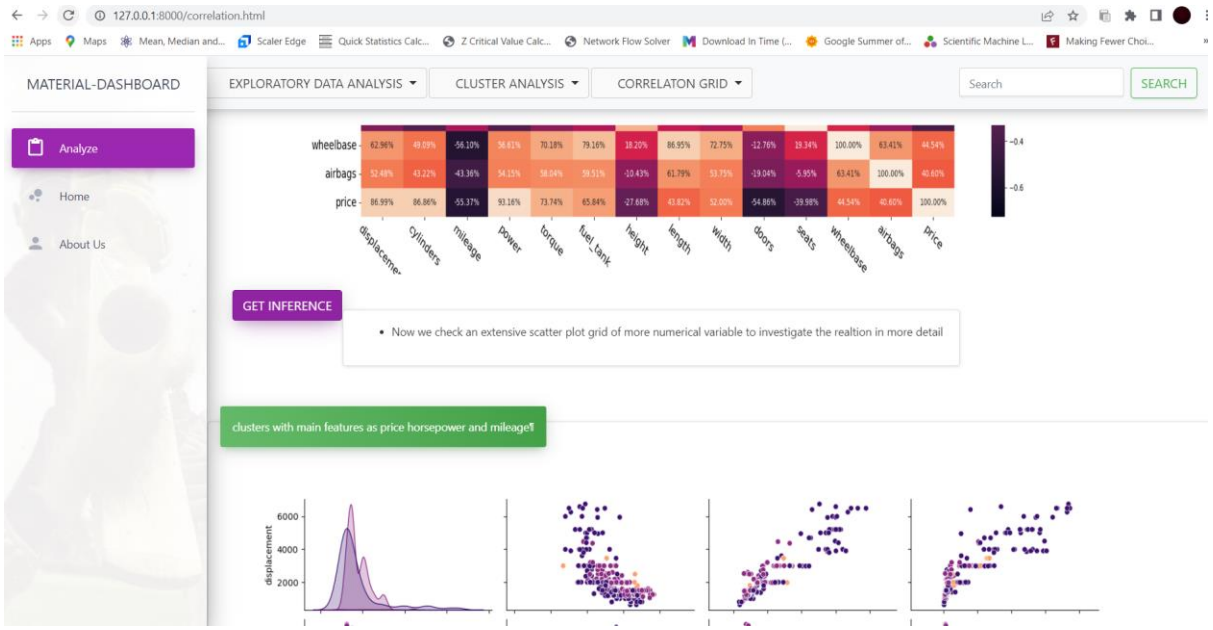
- Histogram of Price
- Dominating car BodyType
- BoxPlot for Price (Outlier analysis)
- engine size comparision
- Relationship for Price and Power

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 stretegic groups

- Cars body type with each cluster

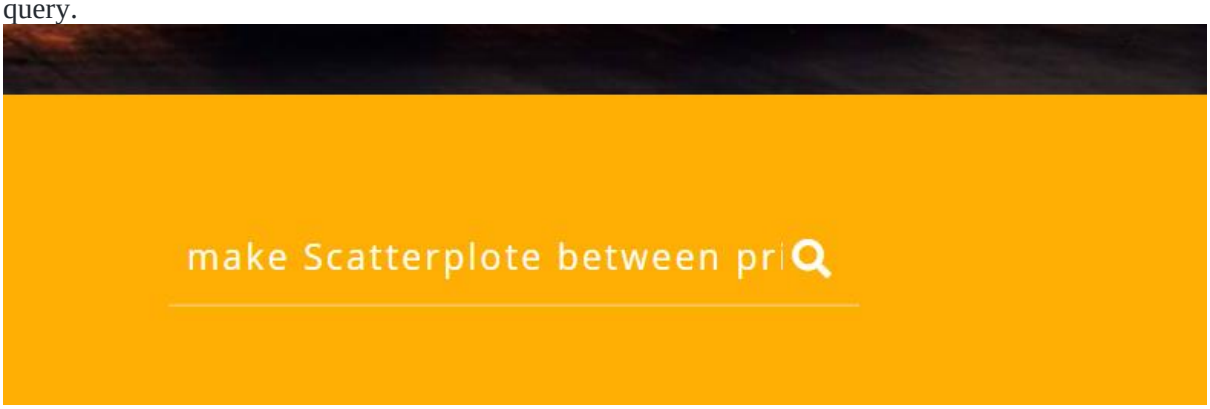
Correlation Grid

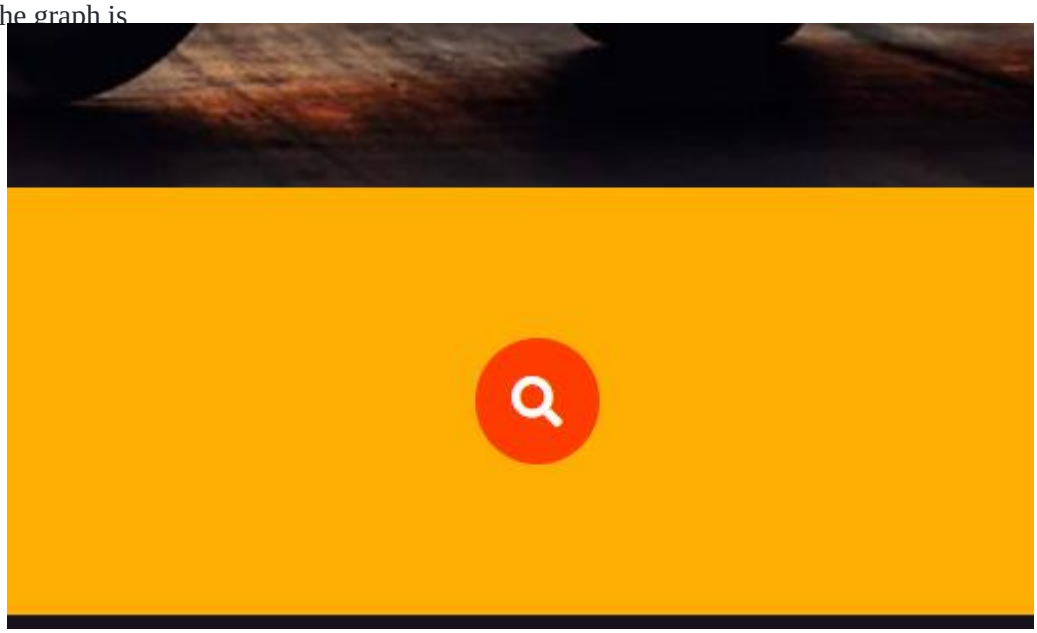


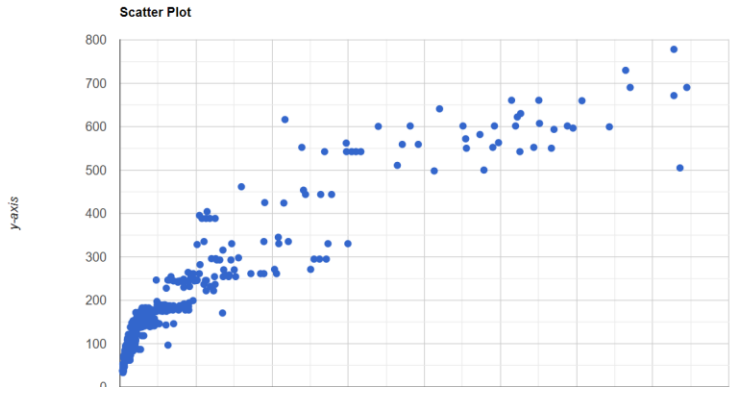
- Correlation Matrix (to know which features all strongly correlated)
- Extensive scatter plot grid of more numerical variable to investigate the realtion in more detail

Command-line Query for Generating graphs

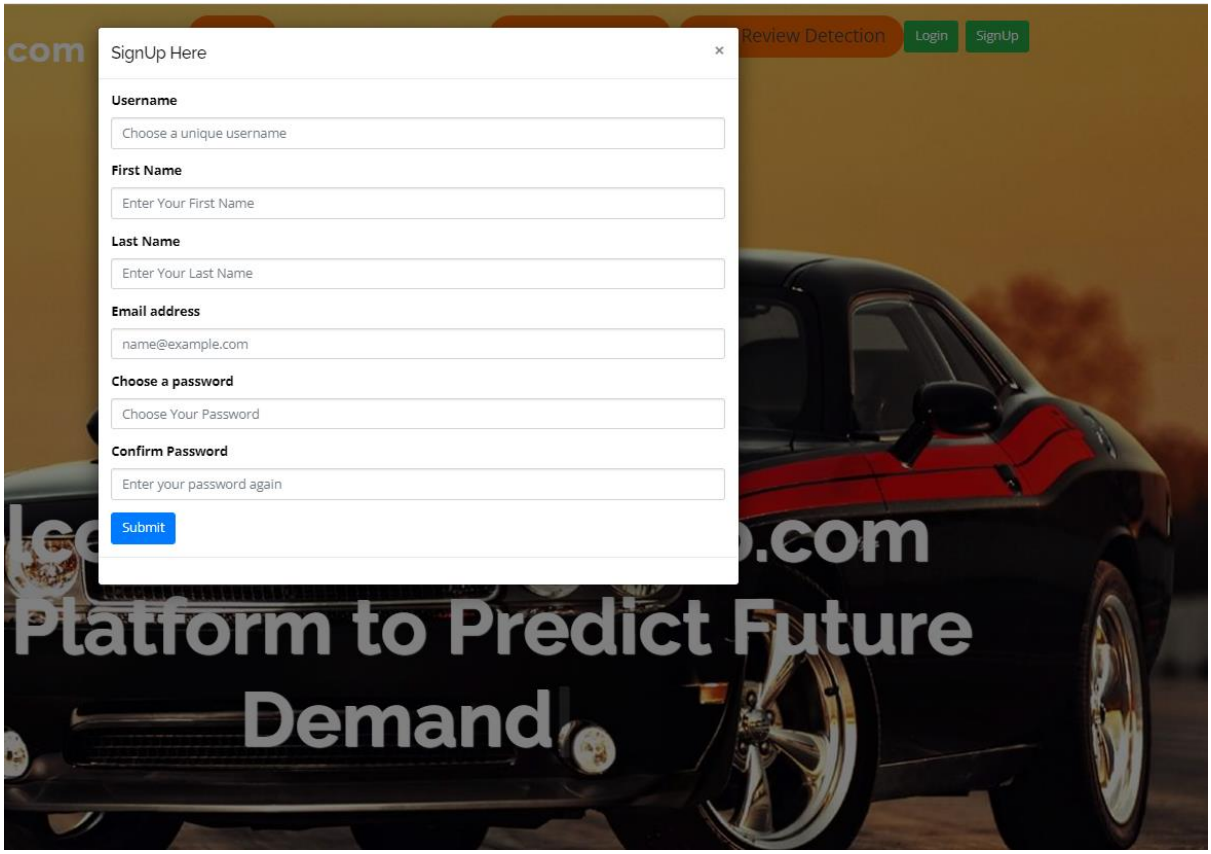
- A command line tool where the user can dynamically enter a query.



- 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.



Login with a success message