PROJECT REPORT

Trip Based Modeling of Fuel Consumption in Modern Fleet Vehicles Using Machine Learning

TEAM ID: PNT2022TMID49664

TEAM LEAD: ANSHA SHOBA.R

TEAM MEMBER: INDUMATHI.P

TEAM MEMBER: MAHA LAKSHMI.M

TEAM MEMBER:PACKIYALAKSHMI.S

TABLE OF CONTENTS

5.3 User Stories

6. PROJECT PLANNING & SCHEDULING

7. CODING & SOLUTIONING (Explain the features added in the project along with code)

8. ADVANTAGES & DISADVANTAGES

9. CONCLUSION

10.FUTURE SCOPE

Source Code

GitHub & Project Demo Link

1. INTRODUCTION

1.1 Project Overview

Fuel consumption models for vehicles are of interest to manufacturers, regulators, and consumers. They are needed across all the phases of the vehicle life-cycle. In this paper, we focus on modeling average fuel consumption for heavy vehicles during the operation and maintenance phase. In general, techniques used to develop models for fuel consumption fall under three main categories:

- Physics-based models, which are derived from an indepth understanding of the physical system. These models describe the dynamics of the components of the vehicle at each time step using detailed mathematical equations .
- Machine learning models, which are data-driven and represent an abstract mapping from an input space consisting of a selected set of predictors to an output space that represents the target output, in this case average fuel consumption
- Statistical models, which are also data-driven and establish a mapping between the probability distribution of a selected set of predictors and the target outcome. Trade-offs among the above techniques are primarily with respect to cost and accuracy as per the requirements of the intended application. In this paper, a model that can be easily developed for individual heavy vehicles in a large fleet is proposed.

1.2 Purpose

A model that can be easily developed for individual heavy vehicles in a large fleet is proposed. Relying on accurate models of all of the vehicles in a fleet, a fleet manager can optimize the route planning for all of the vehicles based on each unique vehicle predicted fuel consumption thereby ensuring the route assignments are aligned to minimize overall fleet fuel consumption. These types of fleets exist in various sectors including, road transportation of goods, public transportation, construction trucks and refuse trucks . For each fleet, the methodology must apply and adapt to many different vehicle technologies (including future ones) and configurations without detailed knowledge of the vehicles specific physical characteristics and measurements. These requirements make machine learning the technique of choice when taking into consideration the desired accuracy versus the cost of the development and adaptation of an individualized model for each vehicle in the fleet.

2. LITERATURE SURVEY

2.1 Existing problem

Previously proposed machine learning models for average fuelconsumption use a set of predictors that are collected over a time period to predict the corresponding fuel consumption in terms of either gallons per mile or liters per kilometer. While still focusing on average fuel consumption, our proposed approach differs from that used in previous models because the input space of the predictors is quantized with respect to a fixed distance as opposed to a fixed time period. In the proposed model, all the predictors are aggregated with respect to a fixed window that represents the distance traveled by the vehicle thereby providing a better mapping from the input space to the output space of the model. In contrast, previous machine learning models must not only learn the patterns in the input data but also perform a conversion from the timebased scale of the input domain to the distance-based scale of the output domain.

2.2 References

https://www.mdpi.com/1996-1073/14/24/8592/htm.

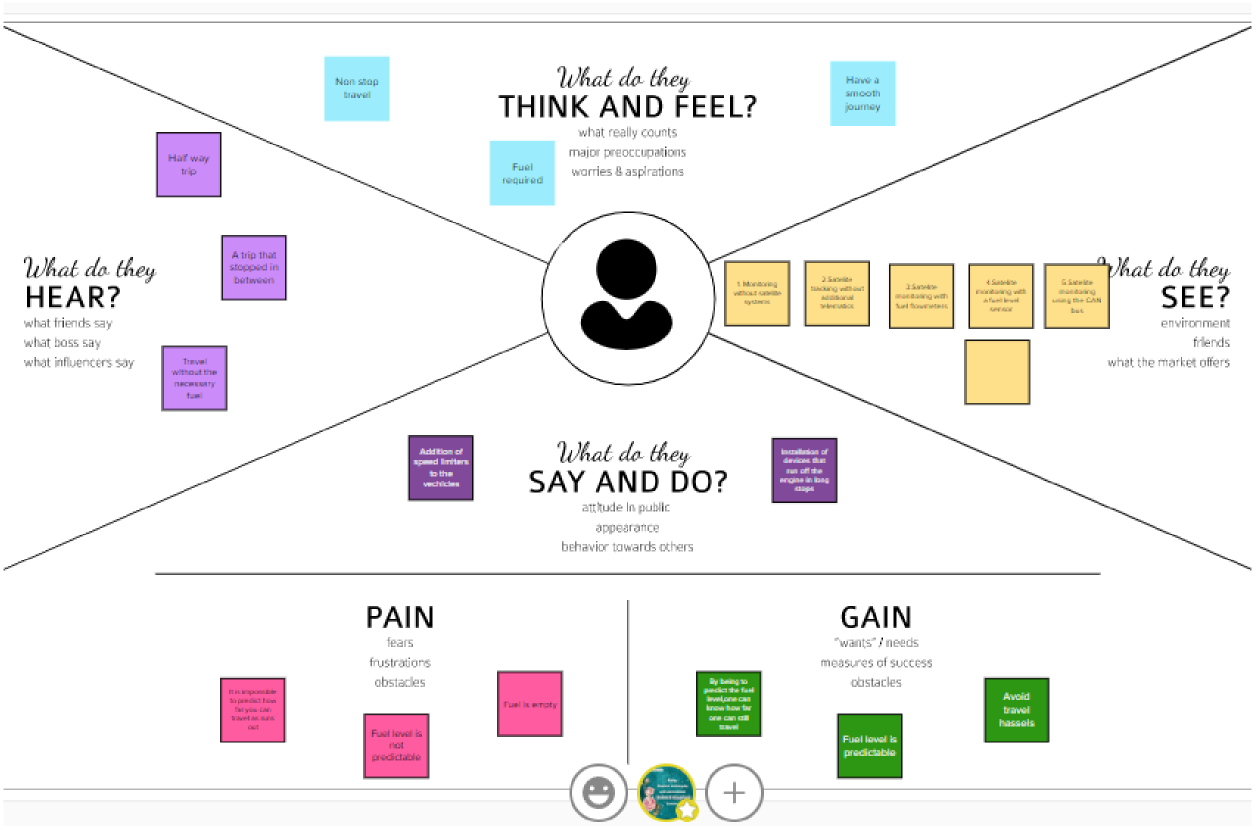
https://core.ac.uk/download/pdf/199435647.pdf.

2.3 Problem Statement Definition

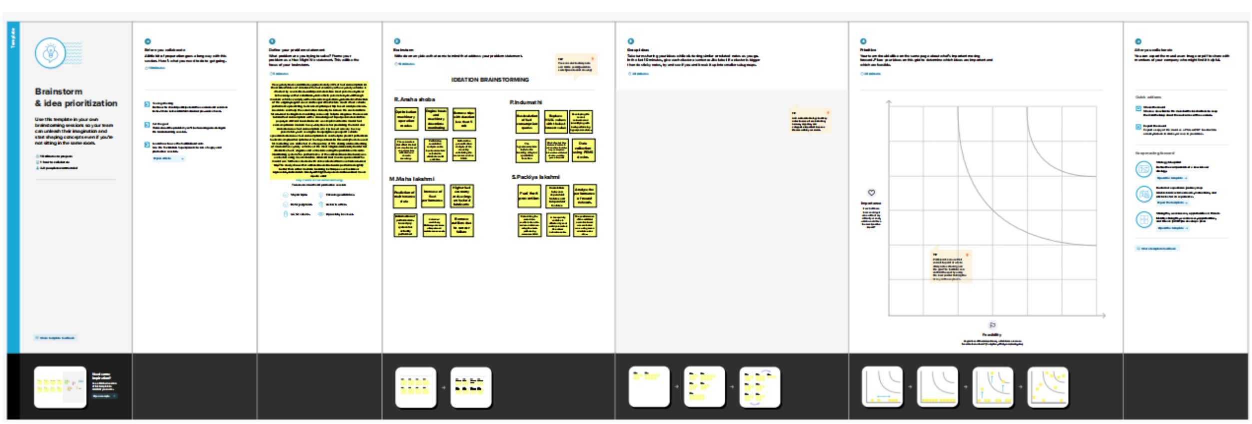
This project a data summarization approach based on distance rather than the traditional time period when developing individualized machine learning models for fuel consumption. This approach is used in conjunction with seven predictors derived from vehicle speed and road grade to produce a highly predictive neural network model for average fuel consumption in heavy vehicles. The proposed model can easily be developed and deployed for each individual vehicle in a fleet in order to optimize fuel consumption over the entire fleet. The predictors of the model are aggregated over fixed window sizes of distance traveled. Different window sizes are evaluated and the results show that a 1 km window is able to predict fuel consumption with a 0.91 coefficient of determination and mean absolute peak-to-peak percent error less than 4% for routes that include both city and highway duty cycle segments.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas



3.2 Ideation&Brainstorming



3.3 Proposed Solution

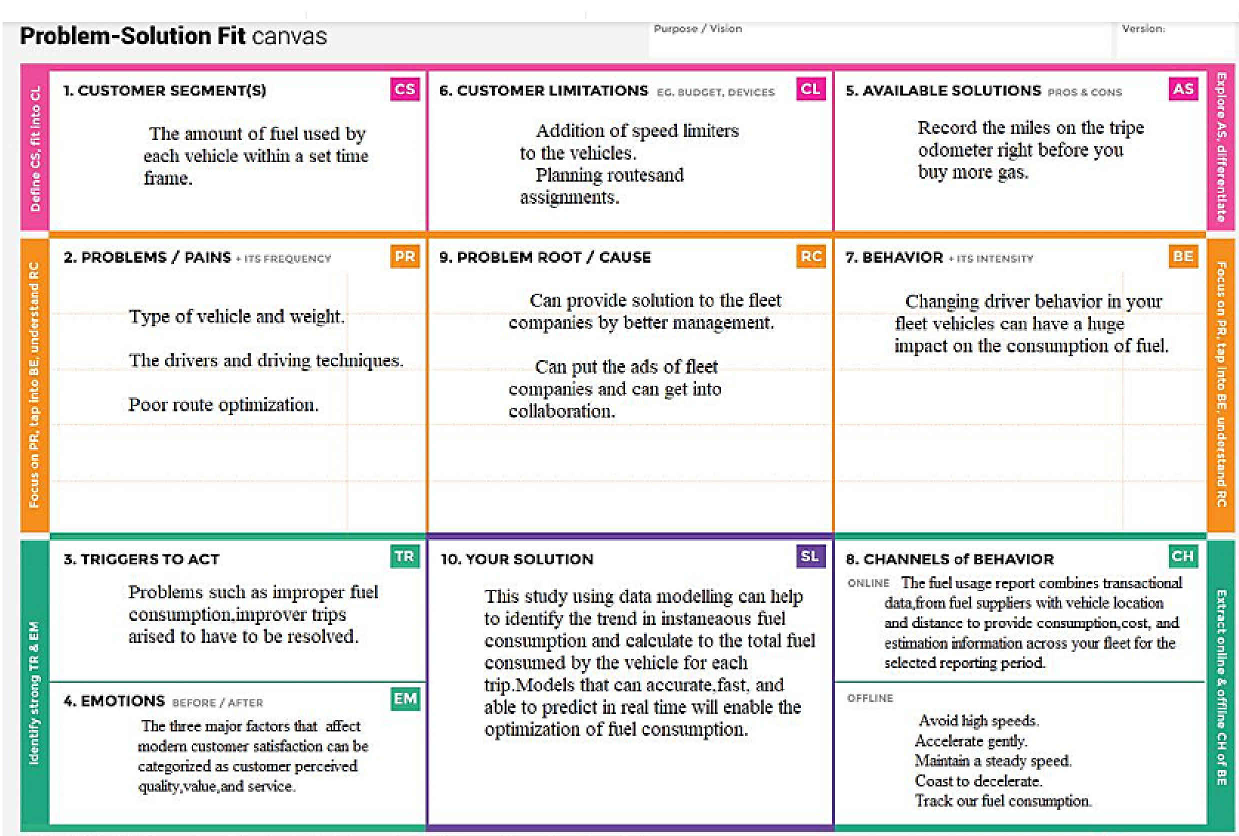
Proposed Solution Template

S.NO	PARAMETER	DESCRIPTION
1.	Problemstatement (Problem to be solved)	Enables you stay on schedule and complete trips trips on-time.Also retrace a vehicle's route,including the location of alerts and warnings
		triggered along the way.

2.	**Idea/solution description**	**Record the miles on the tripe odometer right before you buy more gas.**
3.	**Novelty/Uniqueness**	**The study of the novelty numerical method has been addressed in this research to decrease the fuel consumption of diesel engine and restrict the exhaust gases emission from the operational activities .**
4.	**SocialImpact/Customer Satisfaction**	**The three major factors that affect modern customer satisfaction can be categorized as customer perceived quality,value,and service**

5.	BusinessModel(Revenue Model)	Some simple strategies,like checking tire pressure and replacing oxygen sensors,can help your business

3.4 Problem Solution Fit



4. REQUIREMENT ANALYSIS

4.1 Functional requirement

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIN
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	User dashboard	Enter the use details Change the password
FR-4	User select The category	Select Display Option click IBM,Doctorconsultant,Logout ,Messanger.

4.2 Non Functional requirements

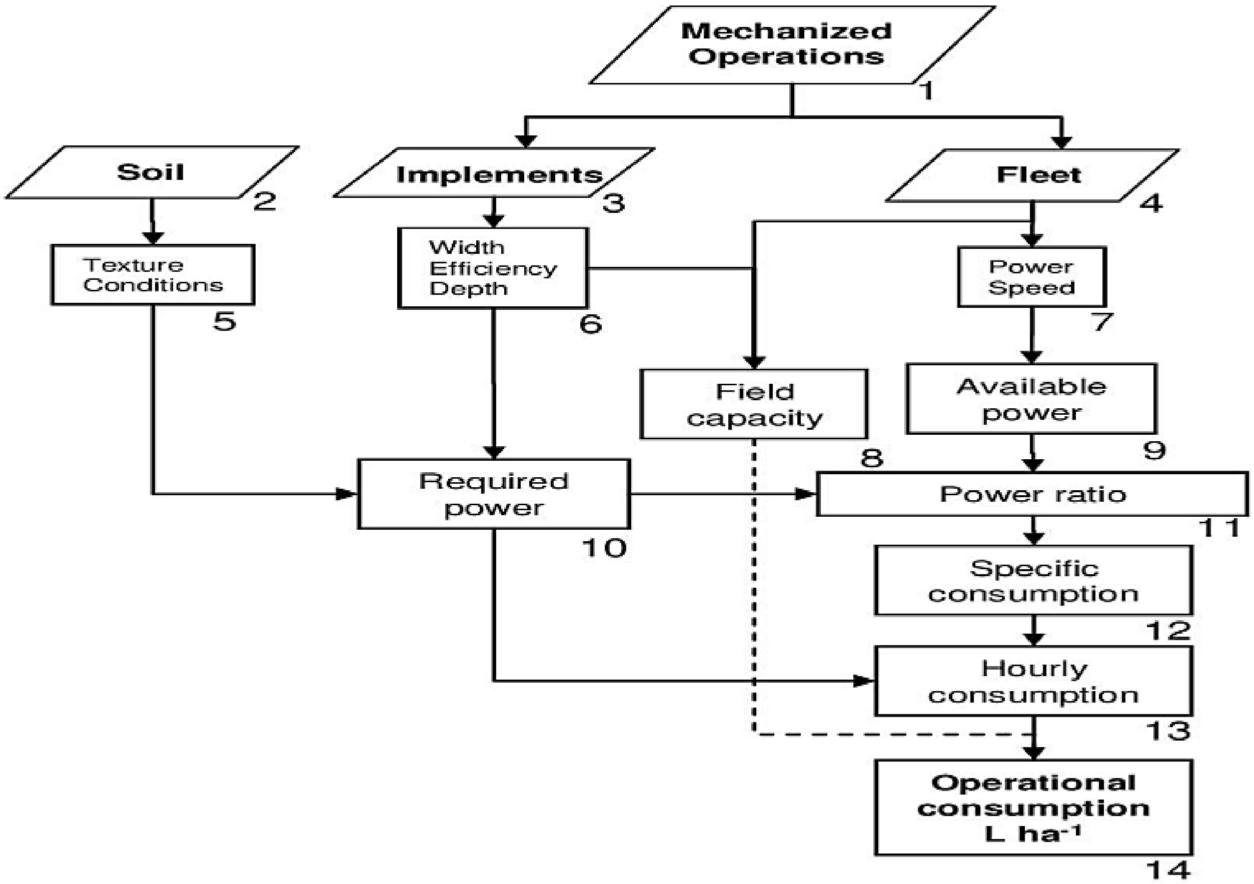
FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	1. Based on a formal analysis of user needs,ninerecommended prototype FEDI component sets(FEDI-CS)werw created. 2. Thirteen participants completed three usabilityevaluation task.
NFR-2	**Security**	1. Fill up your gas tank and note the distance asindicated on the dashboard odometer . 2. The next time you fill up note the amount of fuel required,as listed on the pump display.

NFR-3	**Reliability**	1. Personally ,reliability is not really a big factorwhen choosing a car. 2. After all besides the obvious rust buckets most new cars have pretty much the same reliability. 3. It all boils down to how you take care of it.
NFR-4	**Performance**	1. Fuel economy is the distance travelled per unit volume of fuel used. 2. The higher values the more economic a vehicle is the more distance it can travel with a certain volume of fuel.
NFR-5	**Availability**	1. Initial usability and preference are importantfor a good display. 2. However other aspects such as longer term motivation,engagement,with the display and potential.
NFR-6	**Scalability**	1. Participants preferred representational formsof fuel economy information as compared to text,although text may improve comprehensive.

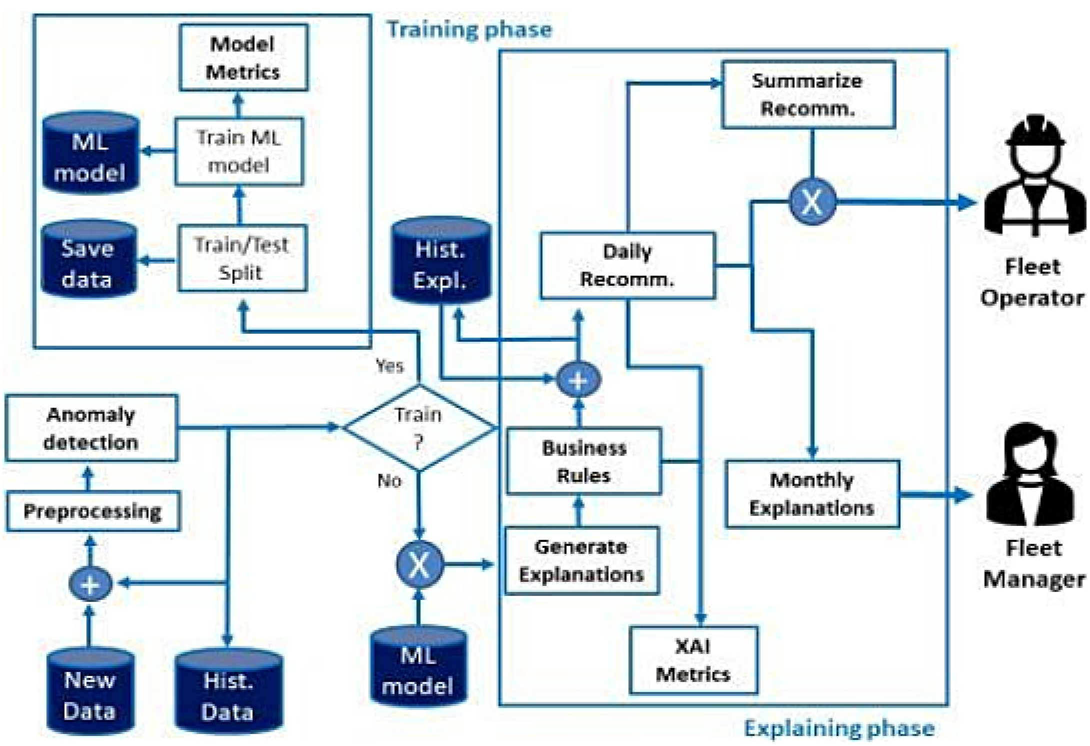
5. PROJECT DESIGN

5.1 Data Flow Diagrams

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirementgraphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



5.2 Solution&Technical Architecture



5.3 User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer	Fuel filter	USN-1	Fuel is sent from the tank to the engine by suction generated by the fuel pump.	It consider fuel compatibility, re usability,position in your fuel system.	High	Sprint-1

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
			or fuel injector of the internal combustion engine.			
	Fuel injection pump	USN-4	It compresses the fuel to high pressure where cam the plunger and the sent it to to the injector.	Full control over rate of fuel injection.Proper spray pattern to ensure mixing of air and fuel.		
Customer (Web user)	Atomize	USN-5	The purpose of a fuel atomizer in	It is to convert the analyst to a reproducible	Medium	Spirit-2

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
			automobile engineering is to atomize or break the oil into fine particles.	amount of gaseous atoms that appropriately represent the sample test.		
Customer Care Executive	Fuel feed pump	USN-6	To supply fuel from main fuel storage tank to the injection system.	It moves a fluid such as a fuel at at controlled rate.	Medium	Sprint-3
Administrator	Carburettor	USN-7	It is used to supplying a spark-ignition	Cannot provide a perfect air-fuel ratio	Medium	Sprint-4

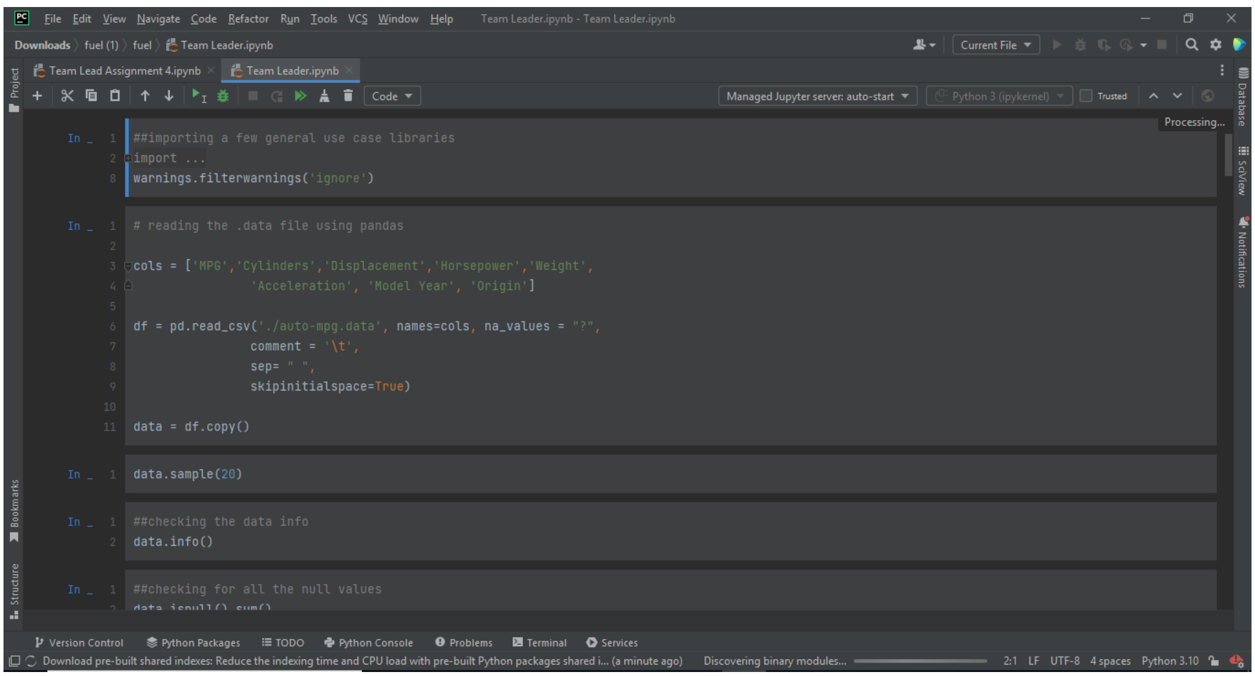
User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
			engine with a mixture of fuel and air.	consistently.		

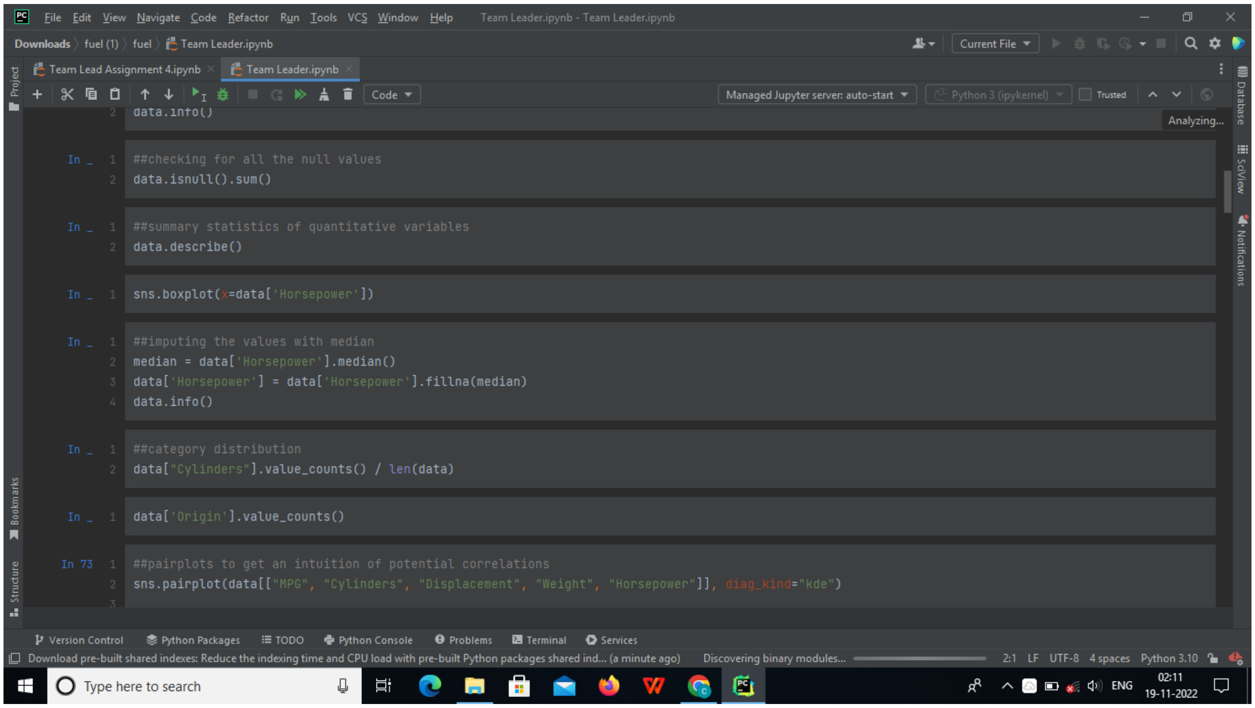
6.PROJECT PLANNING&SCHEDULING

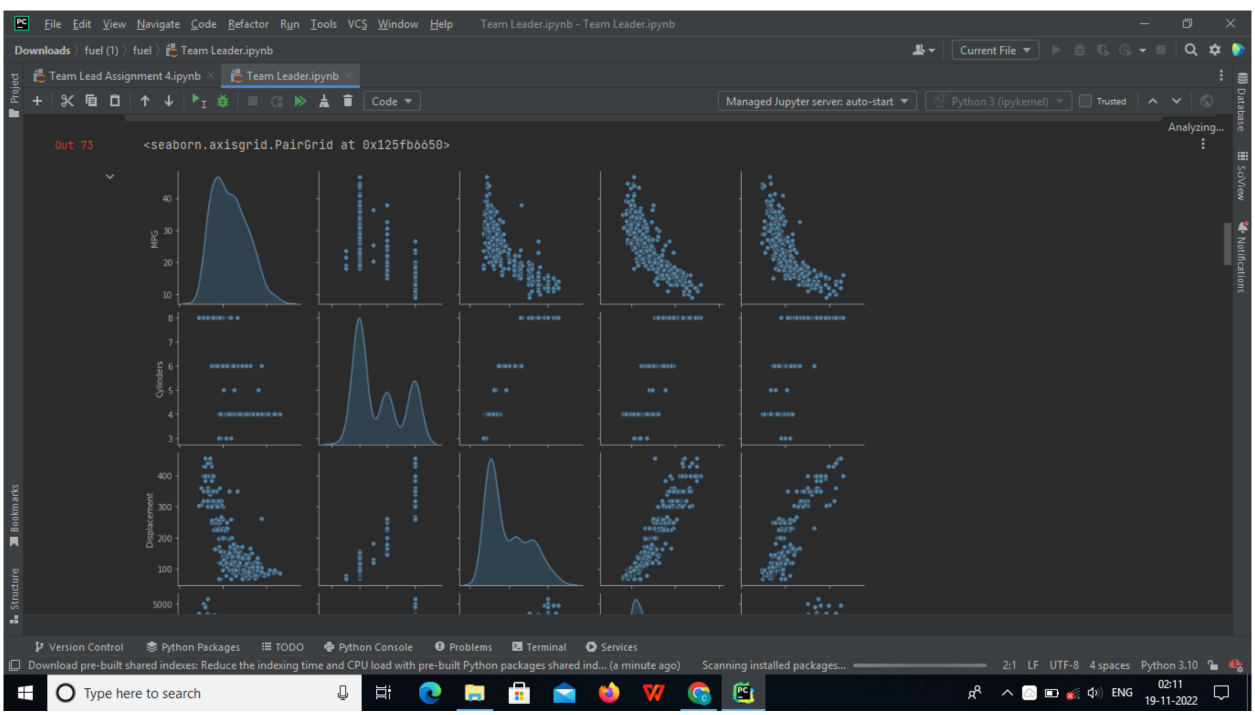
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN - 1	As a user, I can register using email and password	4	High	Ansha Shoba
Sprint-2		USN - 2	As a user, I can register using Gmail	2	Medium	Indumathi
Sprint-1		USN - 3	As a user, I will receive confirmation email once I have registered for the application	1	Low	Maha lakshmi
	Login	USN - 4	As a user, I can login to my dashboard through email id and password	2	High	Ansha Shoba
	Dashboard	USN - 5	I can access my account details on dashboard	1	Low	Packiyalakshmi
Sprint-2	Prediction Model	USN - 6	Once I enter the dashboard I can input values for a single sample prediction	8	High	Ansha Shoba

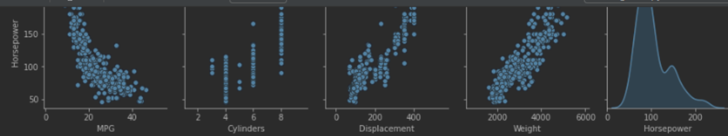
Sprint-3		USN - 7	I can input values via excel sheet for multiple sample prediction as per the template and perform prediction	6	Medium	Indumathi
		USN - 8	As a user I can get visual representation of the prediction	4	Medium	Indumathi
	Report Generation	USN - 9	As a user I can view the detailed report of my prediction	3	High	Ansha Shoba
Sprint-4	RestAPI	USN - 10	As a developer, I can use API Token to send request to the server	3	Low	Maha Lakshmi
	Documentation	USN - 11	As a user I can refer to the documentation and user manual for support and guidance	4	High	Packiyalaskshmi
		USN - 12	As a developer, I can refer to technical Documentation for understanding the application flow	6	Medium	Maha Lakshmi

7. CODING AND SOLUTIONS




data.info()
##checking for all the null values
data.isnull().sum()
##summary statistics of quantitative variables
data.describe()
sns.boxplot(x=data['Horsepower'])
##imputing the values with median
median = data['Horsepower'].median()
data['Horsepower'] = data['Horsepower'].fillna(median)
data.info()
##category distribution
data["Cylinders"].value_counts() / len(data)
data['Origin'].value_counts()
In 73
##pairplots to get an intuition of potential correlations
sns.pairplot(data[["MPG", "Cylinders", "Displacement", "Weight", "Horsepower"]], diag_kind="kde")


Out 73
<seaborn.axisgrid.PairGrid at 0x125fb6650>
MPG
Cylinders
Displacement



In 75
```python
# set aside the test data
from sklearn.model_selection import train_test_split
train_set, test_set = train_test_split(data, test_size=0.2, random_state=42)

test_set.shape
```

Out 75
```
(80, 8)
```

In 76
```python
train_set['Cylinders'].value_counts() / len(train_set)
```

Out 76
```
4    0.509434
8    0.248428
6    0.220126
3    0.012579
5    0.009434
Name: Cylinders, dtype: float64
```

In 77
```python
test_set["Cylinders"].value_counts() / len(test_set)
```

Out 77
```
4    0.525
8    0.300
6    0.175
Name: Cylinders, dtype: float64
```

In 78
```python
from sklearn.model_selection import StratifiedShuffleSplit

split = StratifiedShuffleSplit(n_splits=1, test_size=0.2, random_state=42)
for train_index, test_index in split.split(data, data["Cylinders"]):
    strat_train_set = data.loc[train_index]
    strat_test_set = data.loc[test_index]
```

In 79
```python
strat_test_set.shape
```

Out 79
```
(80, 8)
```

In 80
```python
##checking for cylinder category distribution in training set
strat_train_set['Cylinders'].value_counts() / len(strat_train_set)
```

Out 80
```
4    0.512579
```

```
In 79:
strat_test_set.shape
```

```
Out 79:
(80, 8)
```

```
In 80:
##checking for cylinder category distribution in training set
strat_train_set['Cylinders'].value_counts() / len(strat_train_set)
```

```
Out 80:
4    0.512579
8    0.257862
6    0.210692
5    0.009434
3    0.009434
Name: Cylinders, dtype: float64
```

```
In 81:
##checking for cylinder category distribution in testing set
strat_test_set["Cylinders"].value_counts() / len(strat_test_set)
```

```
Out 81:
4    0.5125
8    0.2625
6    0.2125
3    0.0125
Name: Cylinders, dtype: float64
```

```
In 82:
##converting integer classes to countries in Origin column
train_set['Origin'] = train_set['Origin'].map({1: 'India', 2: 'USA', 3 : 'Germany'})
train_set.sample(10)
```

Out 82:

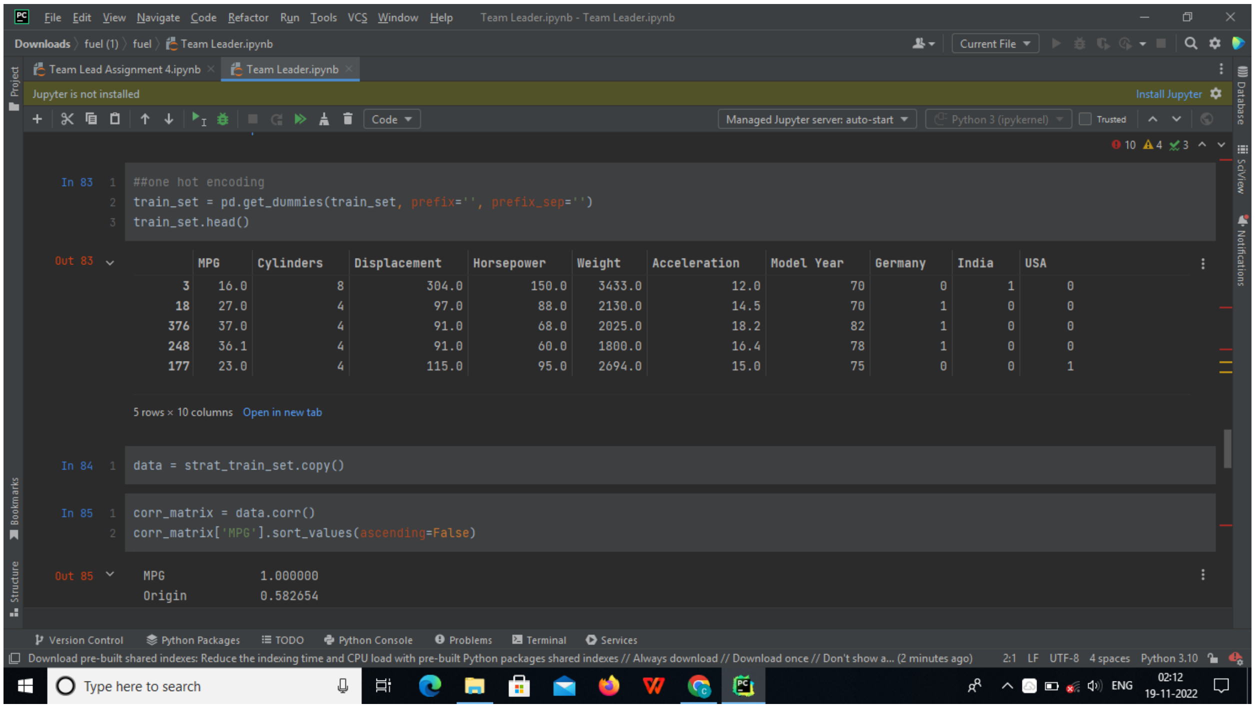
	MPG	Cylinders	Displacement	Horsepower	Weight	Acceleration	Model Year	Origin
253	20.5	6	200.0	95.0	3155.0	18.2	78	India
379	36.0	4	98.0	70.0	2125.0	17.3	82	India
193	24.0	6	200.0	81.0	3012.0	17.6	76	India
326	43.4	4	90.0	48.0	2335.0	23.7	80	USA
10	15.0	8	383.0	170.0	3563.0	10.0	70	India
371	29.0	4	135.0	84.0	2525.0	16.0	82	India
177	23.0	4	115.0	95.0	2694.0	15.0	75	USA

10 rows × 8 columns Open in new tab

```
In 83:
##one hot encoding
train_set = pd.get_dummies(train_set, prefix='', prefix_sep='')
train_set.head()
```

Out 83:

	MPG	Cylinders	Displacement	Horsepower	Weight	Acceleration	Model Year	Germany	India	USA
3	16.0	8	304.0	150.0	3433.0	12.0	70	0	1	0



```
##one hot encoding
train_set = pd.get_dummies(train_set, prefix='', prefix_sep='')
train_set.head()
```

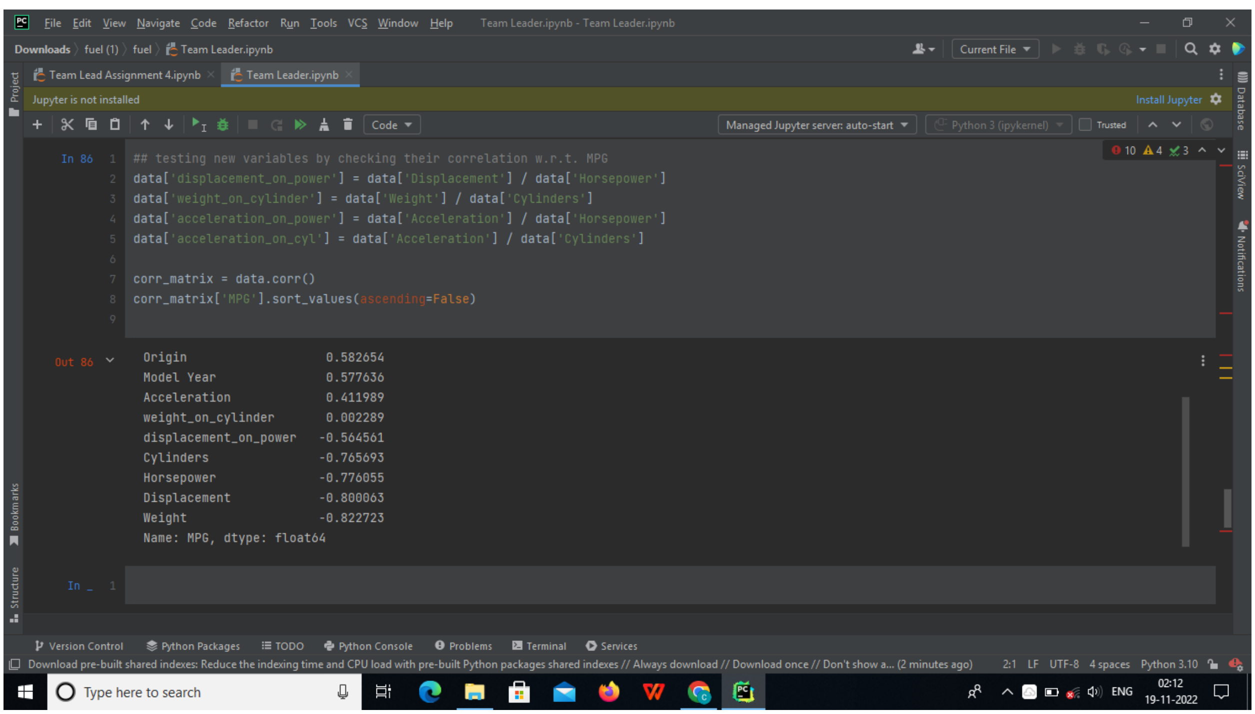
	MPG	Cylinders	Displacement	Horsepower	Weight	Acceleration	Model Year	Germany	India	USA
3	16.0	8	304.0	150.0	3433.0	12.0	70	0	1	0
18	27.0	4	97.0	88.0	2130.0	14.5	70	1	0	0
376	37.0	4	91.0	68.0	2025.0	18.2	82	1	0	0
248	36.1	4	91.0	60.0	1800.0	16.4	78	1	0	0
177	23.0	4	115.0	95.0	2694.0	15.0	75	0	0	1

5 rows × 10 columns Open in new tab

```
data = strat_train_set.copy()
```

```
corr_matrix = data.corr()
corr_matrix['MPG'].sort_values(ascending=False)
```

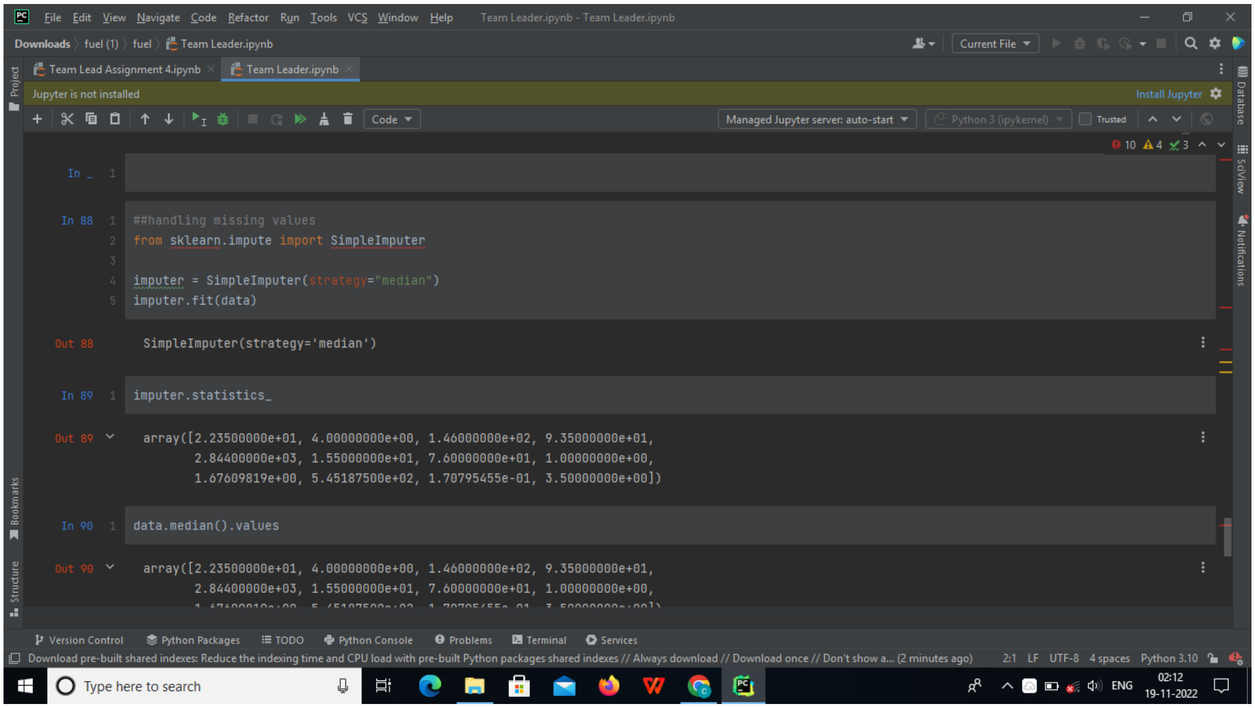
```
MPG        1.000000
Origin     0.582654
```

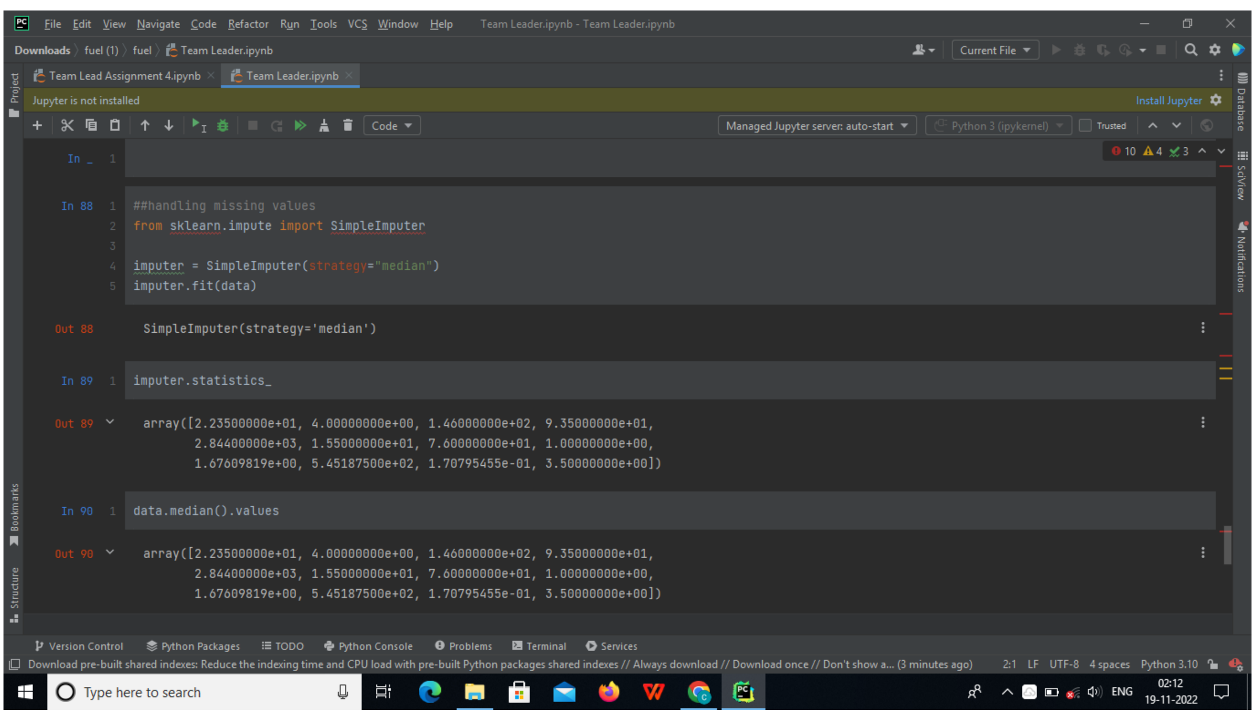


```
## testing new variables by checking their correlation w.r.t. MPG
data['displacement_on_power'] = data['Displacement'] / data['Horsepower']
data['weight_on_cylinder'] = data['Weight'] / data['Cylinders']
data['acceleration_on_power'] = data['Acceleration'] / data['Horsepower']
data['acceleration_on_cyl'] = data['Acceleration'] / data['Cylinders']

corr_matrix = data.corr()
corr_matrix['MPG'].sort_values(ascending=False)
```

```
Origin                   0.582654
Model Year               0.577636
Acceleration             0.411989
weight_on_cylinder       0.002289
displacement_on_power   -0.564561
Cylinders               -0.765693
Horsepower              -0.776055
Displacement            -0.800063
Weight                  -0.822723
Name: MPG, dtype: float64
```


File Edit View Navigate Code Refactor Run Tools VCS Window Help Team Leader.ipynb - Team Leader.ipynb
Downloads fuel (1) fuel Team Leader.ipynb
Team Lead Assignment 4.ipynb
Team Leader.ipynb
Jupyter is not installed
Install Jupyter
Managed Jupyter server: auto-start
Python 3 (ipykernel)
Trusted
In 88
##handling missing values
from sklearn.impute import SimpleImputer
imputer = SimpleImputer(strategy="median")
imputer.fit(data)
Out 88
SimpleImputer(strategy='median')
In 89
imputer.statistics_
Out 89
array([2.23500000e+01, 4.00000000e+00, 1.46000000e+02, 9.35000000e+01,
2.84400000e+03, 1.55000000e+01, 7.60000000e+01, 1.00000000e+00,
1.67609819e+00, 5.45187500e+02, 1.70795455e-01, 3.50000000e+00])
In 90
data.median().values
Out 90
array([2.23500000e+01, 4.00000000e+00, 1.46000000e+02, 9.35000000e+01,
2.84400000e+03, 1.55000000e+01, 7.60000000e+01, 1.00000000e+00,
Version Control Python Packages TODO Python Console Problems Terminal Services
Download pre-built shared indexes: Reduce the indexing time and CPU load with pre-built Python packages shared indexes // Always download // Download once // Don't show a... (2 minutes ago)
2:1 LF UTF-8 4 spaces Python 3.10
Type here to search
ENG
02:12
19-11-2022


File Edit View Navigate Code Refactor Run Tools VCS Window Help Team Leader.ipynb - Team Leader.ipynb
Downloads fuel (1) fuel Team Leader.ipynb
Team Lead Assignment 4.ipynb
Team Leader.ipynb
Jupyter is not installed
Install Jupyter
Managed Jupyter server: auto-start
Python 3 (ipykernel)
Trusted
In 88
##handling missing values
from sklearn.impute import SimpleImputer
imputer = SimpleImputer(strategy="median")
imputer.fit(data)
Out 88
SimpleImputer(strategy='median')
In 89
imputer.statistics_
Out 89
array([2.23500000e+01, 4.00000000e+00, 1.46000000e+02, 9.35000000e+01,
2.84400000e+03, 1.55000000e+01, 7.60000000e+01, 1.00000000e+00,
1.67609819e+00, 5.45187500e+02, 1.70795455e-01, 3.50000000e+00])
In 90
data.median().values
Out 90
array([2.23500000e+01, 4.00000000e+00, 1.46000000e+02, 9.35000000e+01,
2.84400000e+03, 1.55000000e+01, 7.60000000e+01, 1.00000000e+00,
1.67609819e+00, 5.45187500e+02, 1.70795455e-01, 3.50000000e+00])
Version Control Python Packages TODO Python Console Problems Terminal Services
Download pre-built shared indexes: Reduce the indexing time and CPU load with pre-built Python packages shared indexes // Always download // Download once // Don't show a... (3 minutes ago)
2:1 LF UTF-8 4 spaces Python 3.10
Type here to search
ENG
02:12
19-11-2022

8. ADVANTAGES AND DISADVANTAGES

Advantages:

- Among the factors of energy consumption by transportation, vehicle fuel efficiency plays a significant role. With an increase in fuel efficiency (in miles per gallon), marginal fuel consumption decreases. The most important fuel consumption benefits are achieved in the lower ranges of improvements. For instance, an improvement from 10 to 20 miles per gallon reduces fuel consumption by 50%, while an improvement from 20 to 30 miles per gallon will further reduce fuel consumption by 33%. Thus, vehicle-wise a significant fuel economy is reached if a consumer switches from a Sport Utility Vehicle (15 miles per gallon) to a regular car (25 miles per gallon)
- Although switching to a more fuel-efficient vehicle such as a hybrid (35 miles per gallon) results in fuel economy gains, they are not marginally that significant for an individual consumer, but much more at the aggregate level (fuel consumption by the society). This is particularly the case if the higher price of a more energy-efficient vehicle does not compensate for the gain in fuel efficiency, then it is not a rational choice from an economic standpoint. Therefore, for fuel efficiency to be beneficial for society, the price of the vehicle should remain similar as its fuel efficiency increases, or at least its fuel efficiency should compensate for its higher price. Significant gains in fuel efficiency are also achieved when vehicles operate in conditions with less congestion.
- It can save you money. Driving a fuel-efficient car reduces the running costs you will have to pay throughout the lifetime of your vehicle. There are many different savings associated with fuel efficiency, and together they add up to a significant amount of money.

Disadvantages :

- lack of reliable information on transport operations
- The efficiency of fuel consumption control low
- Probability of fuel fraud: high
- Where can be used: vehicles with tanks not adapted to the installation of third-party sensors.
- High installation and maintenance costs, no traceability of fuel drains from the tank
- Where can be used: special equipment with tanks that are not adapted to the installation of remote control units complicated installation on non-standard tanks
- This bad habit is threefold – driving too fast, accelerating too quickly, and stopping too suddenly. All three of these actions lead to high fuel consumption. Where possible, you should accelerate slowly and drive with the speed of traffic.

9.CONCLUSION

Construct a machine learning model for each of the fleet's heavies in a short amount of time.The seven predictorsincluded in the model are as follows: the number ofstops, the duration of each halt, the average speed, the characteristic acceleration,the aerodynamic speed squared, the change in kinetic energy, and the change in potential energy. In this study, we introduce the final two predictors to better represent the typical dynamic behaviour of the vehicle. The model's predictors are all computed using data collected on vehicle velocity and road gradient.Such data is easily accessible via the telematics devices increasingly found in modern automobiles. Not only that, but the predictors may be quickly calculated on-board based on these first two variables.

10.FUTURE SCOPE

Average fuel consumption for heavy vehicles is predicted using Machine Learning Algorithms like ANN, which are described in this study (ArtificialNeural Networks). The author has derived 7 variables from a dataset of heavy vehicles to forecast fuel usage.

DEMO LINK