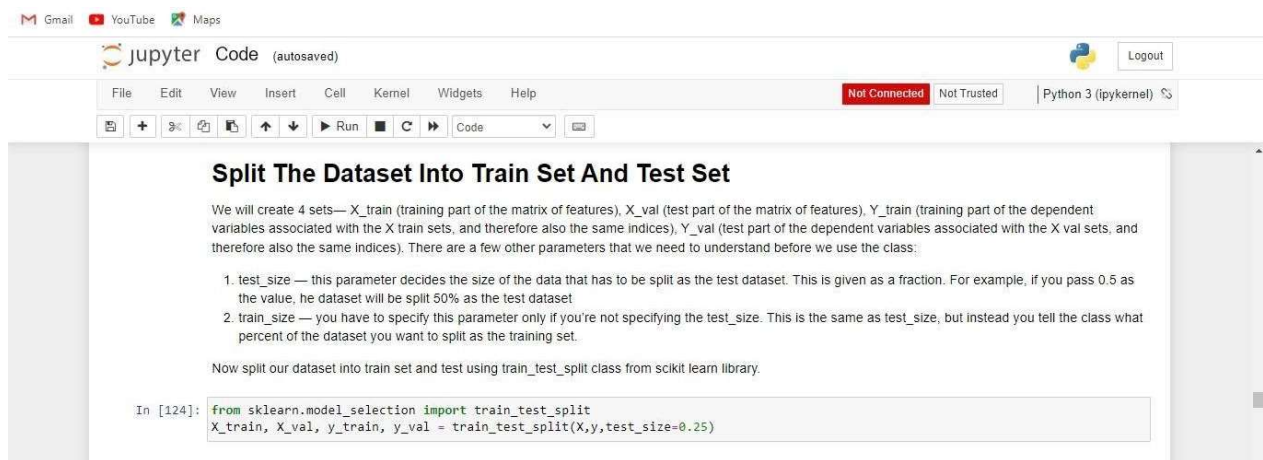


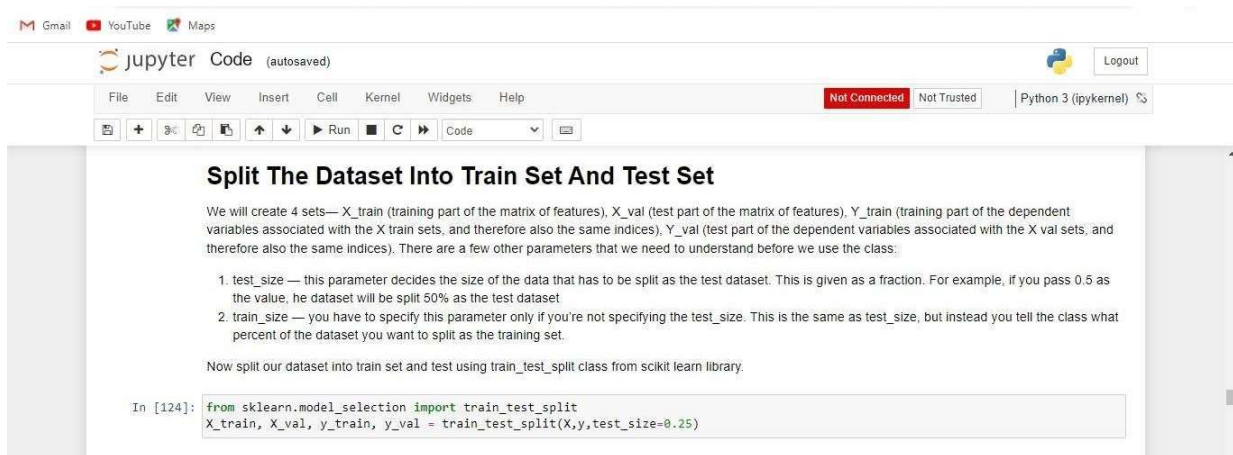
TEAM ID: PNT2022TMID32006

PROJECT NAME: DemandEst -AI powered FoodDemandForecaster

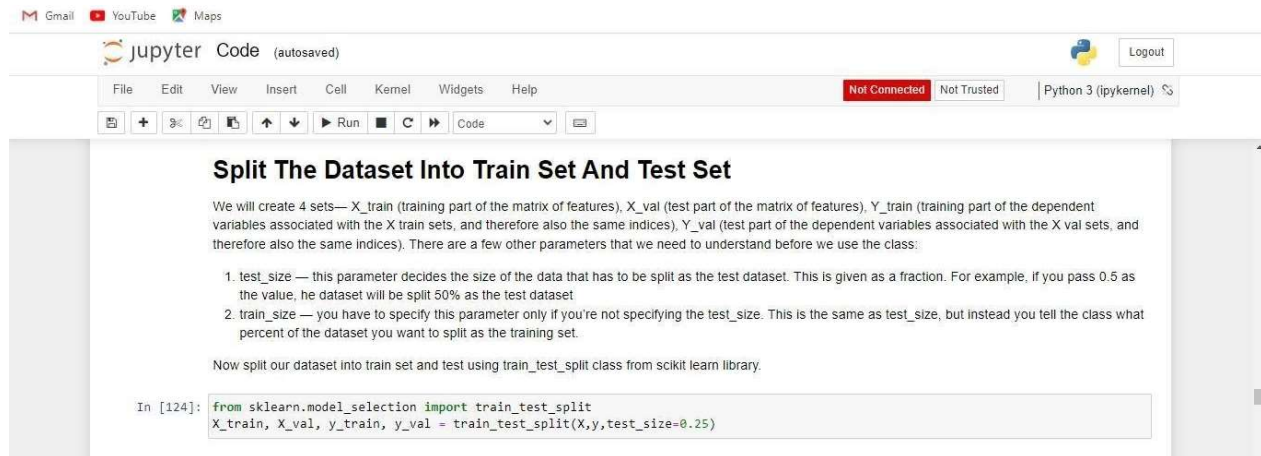
Team Leader



Team Member 1



Team Member 2



The screenshot shows a Jupyter Notebook interface. At the top, there are links for Gmail, YouTube, and Maps. The notebook title is "jupyter Code (autosaved)". The menu bar includes File, Edit, View, Insert, Cell, Kernel, Widgets, and Help. The status bar shows "Not Connected", "Not Trusted", and "Python 3 (ipykernel)".

Split The Dataset Into Train Set And Test Set

We will create 4 sets— X_{train} (training part of the matrix of features), X_{val} (test part of the matrix of features), Y_{train} (training part of the dependent variables associated with the X train sets, and therefore also the same indices), Y_{val} (test part of the dependent variables associated with the X val sets, and therefore also the same indices). There are a few other parameters that we need to understand before we use the class:

1. `test_size` — this parameter decides the size of the data that has to be split as the test dataset. This is given as a fraction. For example, if you pass 0.5 as the value, the dataset will be split 50% as the test dataset
2. `train_size` — you have to specify this parameter only if you're not specifying the `test_size`. This is the same as `test_size`, but instead you tell the class what percent of the dataset you want to split as the training set.

Now split our dataset into train set and test using `train_test_split` class from scikit learn library.

```
In [124]: from sklearn.model_selection import train_test_split
X_train, X_val, y_train, y_val = train_test_split(X,y,test_size=0.25)
```

Team Member 3



The screenshot shows a Jupyter Notebook interface. At the top, there are links for Gmail, YouTube, and Maps. The notebook title is "jupyter Code (autosaved)". The menu bar includes File, Edit, View, Insert, Cell, Kernel, Widgets, and Help. The status bar shows "Not Connected", "Not Trusted", and "Python 3 (ipykernel)".

Split The Dataset Into Train Set And Test Set

We will create 4 sets— X_{train} (training part of the matrix of features), X_{val} (test part of the matrix of features), Y_{train} (training part of the dependent variables associated with the X train sets, and therefore also the same indices), Y_{val} (test part of the dependent variables associated with the X val sets, and therefore also the same indices). There are a few other parameters that we need to understand before we use the class:

1. `test_size` — this parameter decides the size of the data that has to be split as the test dataset. This is given as a fraction. For example, if you pass 0.5 as the value, the dataset will be split 50% as the test dataset
2. `train_size` — you have to specify this parameter only if you're not specifying the `test_size`. This is the same as `test_size`, but instead you tell the class what percent of the dataset you want to split as the training set.

Now split our dataset into train set and test using `train_test_split` class from scikit learn library.

```
In [124]: from sklearn.model_selection import train_test_split
X_train, X_val, y_train, y_val = train_test_split(X,y,test_size=0.25)
```