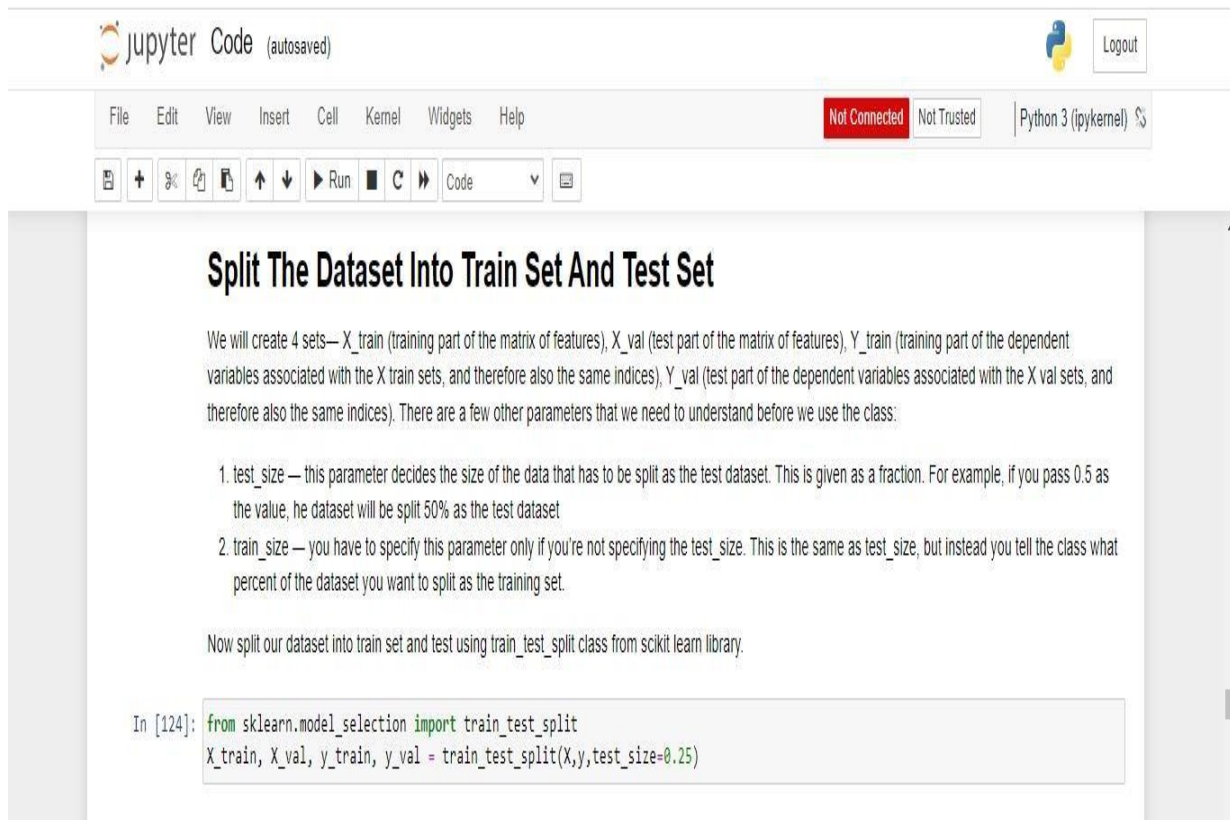


**TEAM ID: PNT2022TMID51052**

**PROJECT NAME: DemandEst - AI powered Food Demand Forecasting**

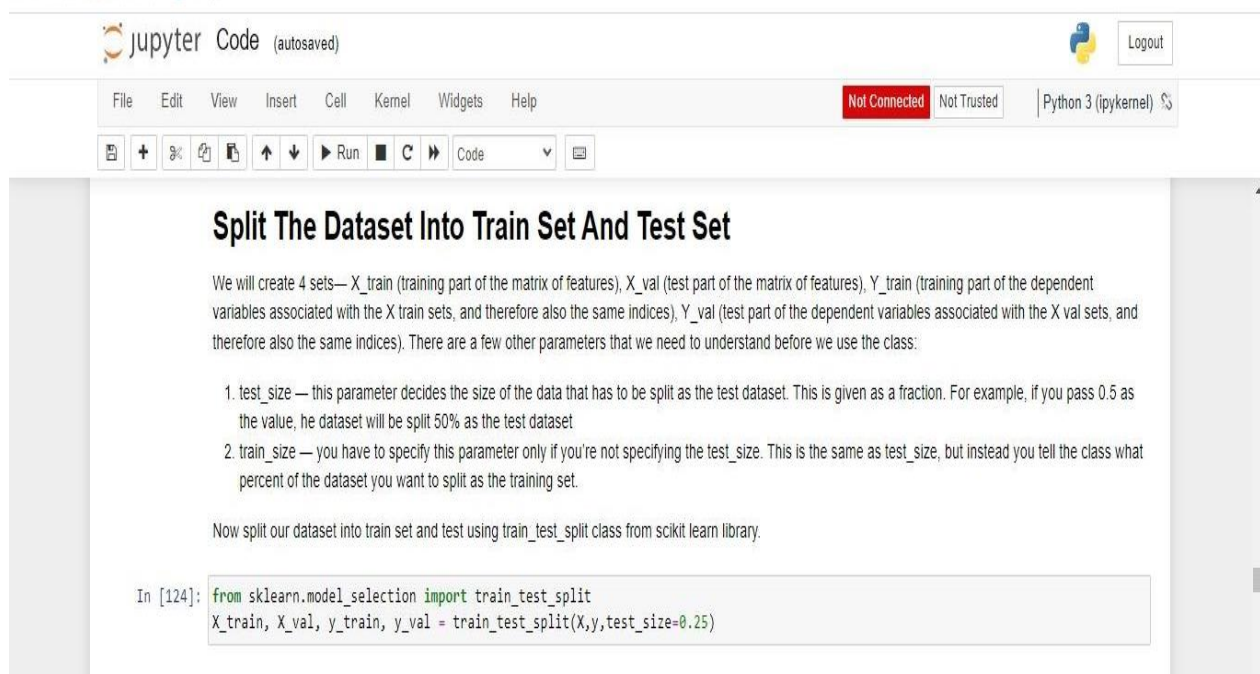
Team Member 1



The screenshot shows a Jupyter Notebook interface. At the top, the title bar says "jupyter Code (autosaved)" with a "Logout" button on the right. Below the title bar is a menu bar with "File", "Edit", "View", "Insert", "Cell", "Kernel", "Widgets", and "Help". To the right of the menu bar are status indicators: "Not Connected" (in a red box), "Not Trusted", and "Python 3 (ipykernel)". Below the menu bar is a toolbar with icons for saving, adding, deleting, and running cells, as well as a "Code" dropdown menu. The main content area of the notebook has a title "Split The Dataset Into Train Set And Test Set" in bold. Below the title is a paragraph explaining the creation of four sets:  $X_{train}$ ,  $X_{val}$ ,  $Y_{train}$ , and  $Y_{val}$ . It then lists two parameters: `test_size` and `train_size`. Finally, it states that the dataset will be split using the `train_test_split` class from the `sklearn` library. Below this text is a code cell with the following Python code:

```
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 2



The screenshot shows a Jupyter Notebook interface. At the top, the title bar says "jupyter Code (autosaved)" with a Python logo and a "Logout" button. Below this is a menu bar with "File", "Edit", "View", "Insert", "Cell", "Kernel", "Widgets", and "Help". To the right of the menu bar are status indicators: "Not Connected" (in a red box), "Not Trusted", and "Python 3 (ipykernel)". Below the menu bar is a toolbar with icons for saving, adding, deleting, and running cells, as well as a dropdown menu currently set to "Code".

The main content area has a title "Split The Dataset Into Train Set And Test Set". Below the title is a paragraph of text explaining the purpose of the code: "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:

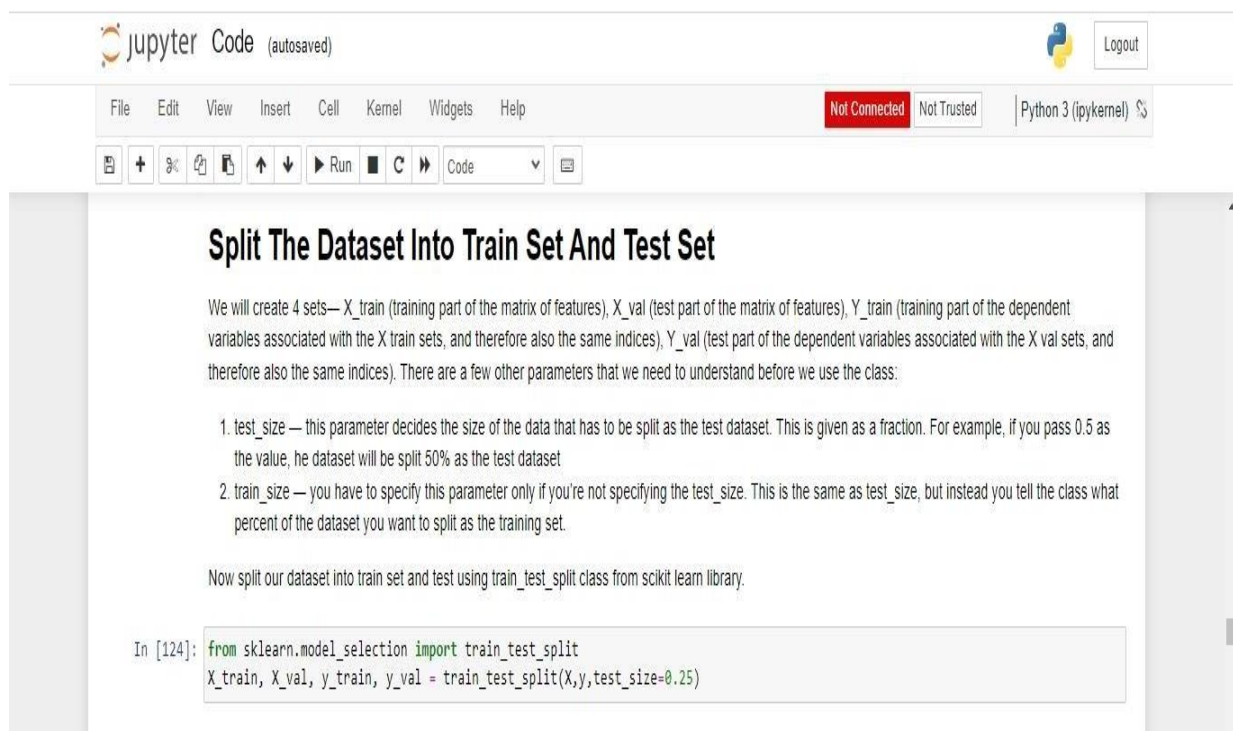
1. test\_size — this parameter decides the size of the data that has to be split as the test dataset. This is given as a fraction. For example, if you pass 0.5 as the value, the dataset will be split 50% as the test dataset
2. train\_size — you have to specify this parameter only if you're not specifying 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.

Below the list is another paragraph: "Now split our dataset into train set and test using train\_test\_split class from scikit learn library."

At the bottom, there is a code cell with the following code:

```
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, identical in layout to the one above. The title bar says "jupyter Code (autosaved)" with a Python logo and a "Logout" button. The menu bar includes "File", "Edit", "View", "Insert", "Cell", "Kernel", "Widgets", and "Help". The status indicators show "Not Connected", "Not Trusted", and "Python 3 (ipykernel)". The toolbar is the same as in the previous screenshot.

The main content area has the same title "Split The Dataset Into Train Set And Test Set". The explanatory text and list of parameters are identical to the previous screenshot.

Below the list is the same paragraph: "Now split our dataset into train set and test using train\_test\_split class from scikit learn library."

At the bottom, there is a code cell with the following code:

```
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)
```