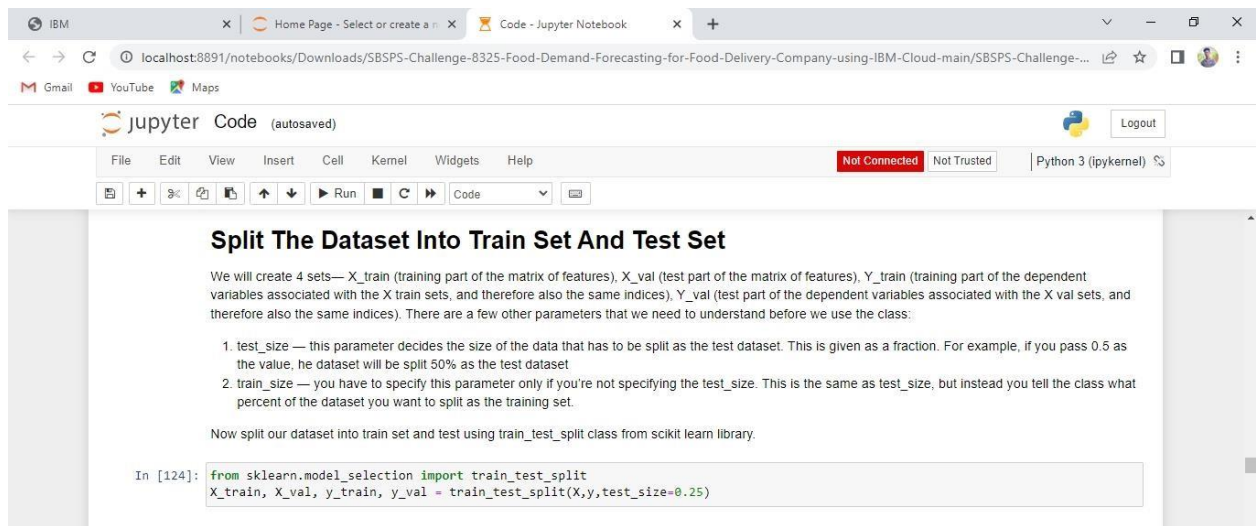


TEAM ID: PNT2022TMID28841

PROJECT NAME: DemandEst -
DemandForecaster

AI powered Food

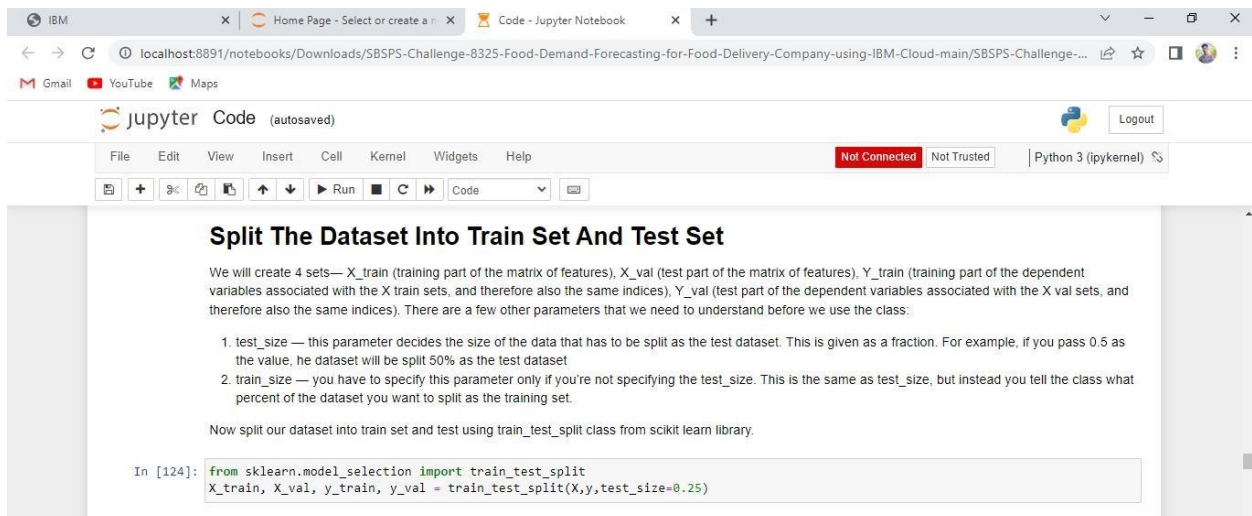
Team Leader



The screenshot shows a Jupyter Notebook interface in a web browser. The browser tabs include 'IBM', 'Home Page - Select or create a n...', and 'Code - Jupyter Notebook'. The address bar shows 'localhost:8891/notebooks/Downloads/SBSPS-Challenge-8325-Food-Demand-Forecasting-for-Food-Delivery-Company-using-IBM-Cloud-main/SBSPS-Challenge-...'. The Jupyter interface has a menu bar (File, Edit, View, Insert, Cell, Kernel, Widgets, Help) and a toolbar with icons for file operations, running, and code execution. The main content area is titled 'Split The Dataset Into Train Set And Test Set'. It contains a paragraph explaining the creation of four sets: X_train, X_val, Y_train, and Y_val. Below this, there are two numbered points explaining the 'test_size' and 'train_size' parameters. At the bottom, there 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 1



The screenshot shows a Jupyter Notebook interface in a web browser, similar to the one above. The browser tabs include 'IBM', 'Home Page - Select or create a n...', and 'Code - Jupyter Notebook'. The address bar shows 'localhost:8891/notebooks/Downloads/SBSPS-Challenge-8325-Food-Demand-Forecasting-for-Food-Delivery-Company-using-IBM-Cloud-main/SBSPS-Challenge-...'. The Jupyter interface has a menu bar (File, Edit, View, Insert, Cell, Kernel, Widgets, Help) and a toolbar with icons for file operations, running, and code execution. The main content area is titled 'Split The Dataset Into Train Set And Test Set'. It contains a paragraph explaining the creation of four sets: X_train, X_val, Y_train, and Y_val. Below this, there are two numbered points explaining the 'test_size' and 'train_size' parameters. At the bottom, there 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 in a web browser. The browser's address bar displays the URL: `localhost:8891/notebooks/Downloads/SBSPS-Challenge-8325-Food-Demand-Forecasting-for-Food-Delivery-Company-using-IBM-Cloud-main/SBSPS-Challenge-...`. The Jupyter Notebook interface includes a top bar with the Jupyter logo, the text "jupyter Code (autosaved)", and a "Logout" button. Below this is a menu bar with options: File, Edit, View, Insert, Cell, Kernel, Widgets, and Help. To the right of the menu bar are status indicators: "Not Connected", "Not Trusted", and "Python 3 (ipykernel)". A toolbar with various icons for file operations and execution is located below the menu bar. The main content area of the notebook displays a code cell with the following text:

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

This screenshot is identical to the one above, showing the same Jupyter Notebook interface and code cell. The code cell contains the following text:

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)
```