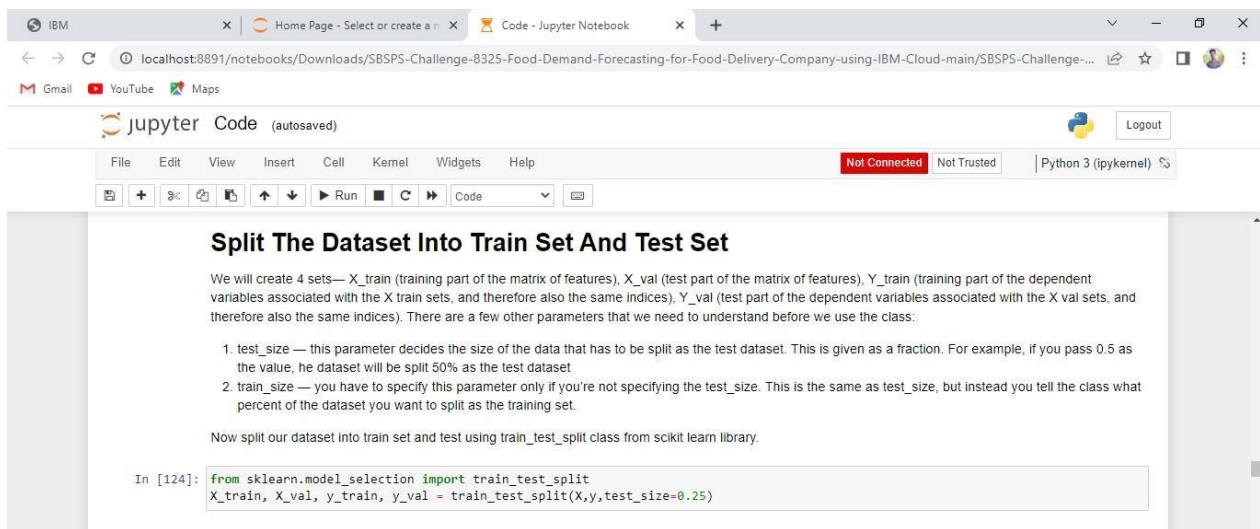


TEAM ID: PNT2022TMID14136

PROJECT NAME: DemandEst - AI powered Food Demand Forecaster

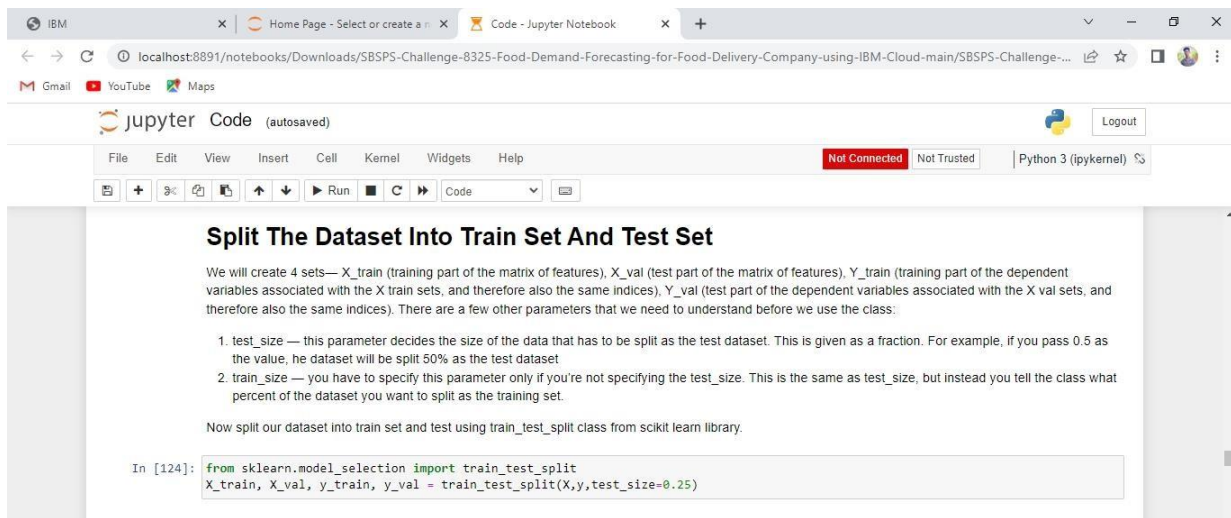
Team Leader



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 the URL: '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 top bar with 'jupyter Code (autosaved)' and a 'Logout' button. Below this is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Cell', 'Kernel', 'Widgets', and 'Help'. A status bar indicates 'Not Connected', 'Not Trusted', and 'Python 3 (ipykernel)'. The main content area has a title 'Split The Dataset Into Train Set And Test Set'. The text explains that four sets will be created: X_train, X_val, Y_train, and Y_val. It lists two parameters: 'test_size' (a fraction, e.g., 0.5 for 50%) and 'train_size' (a percentage of the dataset). It then states: 'Now split our dataset into train set and test using train_test_split class from scikit learn library.' Below this 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)
```

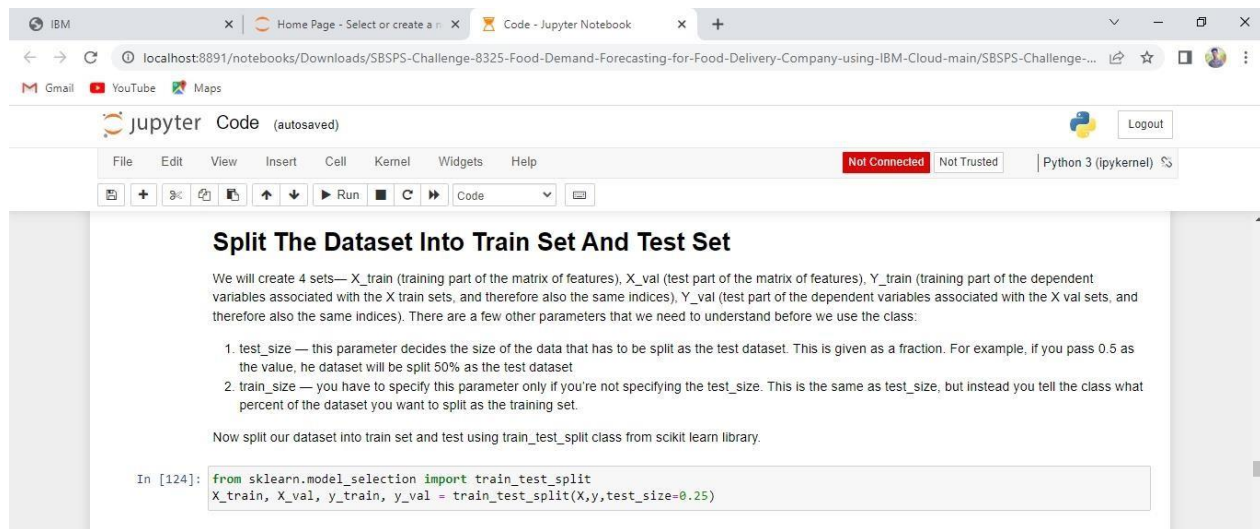
Team Member 1



The screenshot shows a Jupyter Notebook interface in a web browser, identical to the one above. The browser tabs include 'IBM', 'Home Page - Select or create a n...', and 'Code - Jupyter Notebook'. The address bar shows the URL: '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 top bar with 'jupyter Code (autosaved)' and a 'Logout' button. Below this is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Cell', 'Kernel', 'Widgets', and 'Help'. A status bar indicates 'Not Connected', 'Not Trusted', and 'Python 3 (ipykernel)'. The main content area has a title 'Split The Dataset Into Train Set And Test Set'. The text explains that four sets will be created: X_train, X_val, Y_train, and Y_val. It lists two parameters: 'test_size' (a fraction, e.g., 0.5 for 50%) and 'train_size' (a percentage of the dataset). It then states: 'Now split our dataset into train set and test using train_test_split class from scikit learn library.' Below this 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 shows 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 menu bar (File, Edit, View, Insert, Cell, Kernel, Widgets, Help) and a toolbar with icons for file operations, running code, and viewing output. The notebook is titled "Code (autosaved)" and shows a "Not Connected" status. The main content area displays a code cell with the following text:

Split The Dataset Into Train Set And Test Set

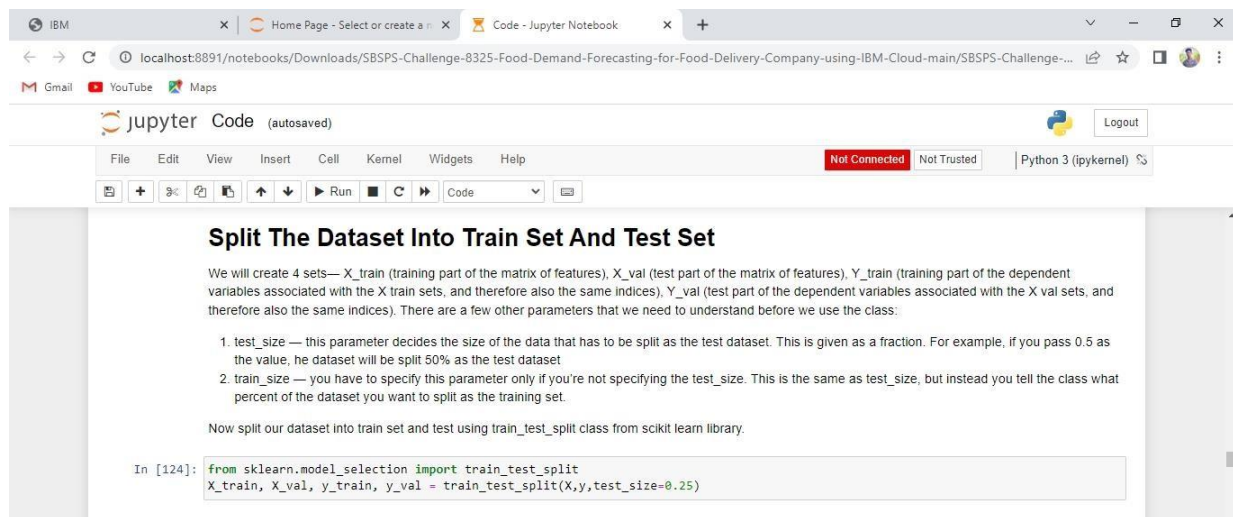
We will create 4 sets— `X_train` (training part of the matrix of features), `X_val` (test part of the matrix of features), `Y_train` (training part of the dependent variables associated with the `X` train sets, and therefore also the same indices), `Y_val` (test part of the dependent variables associated 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



The screenshot shows a Jupyter Notebook interface in a web browser. The browser's address bar shows 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 menu bar (File, Edit, View, Insert, Cell, Kernel, Widgets, Help) and a toolbar with icons for file operations, running code, and viewing output. The notebook is titled "Code (autosaved)" and shows a "Not Connected" status. The main content area 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)
```