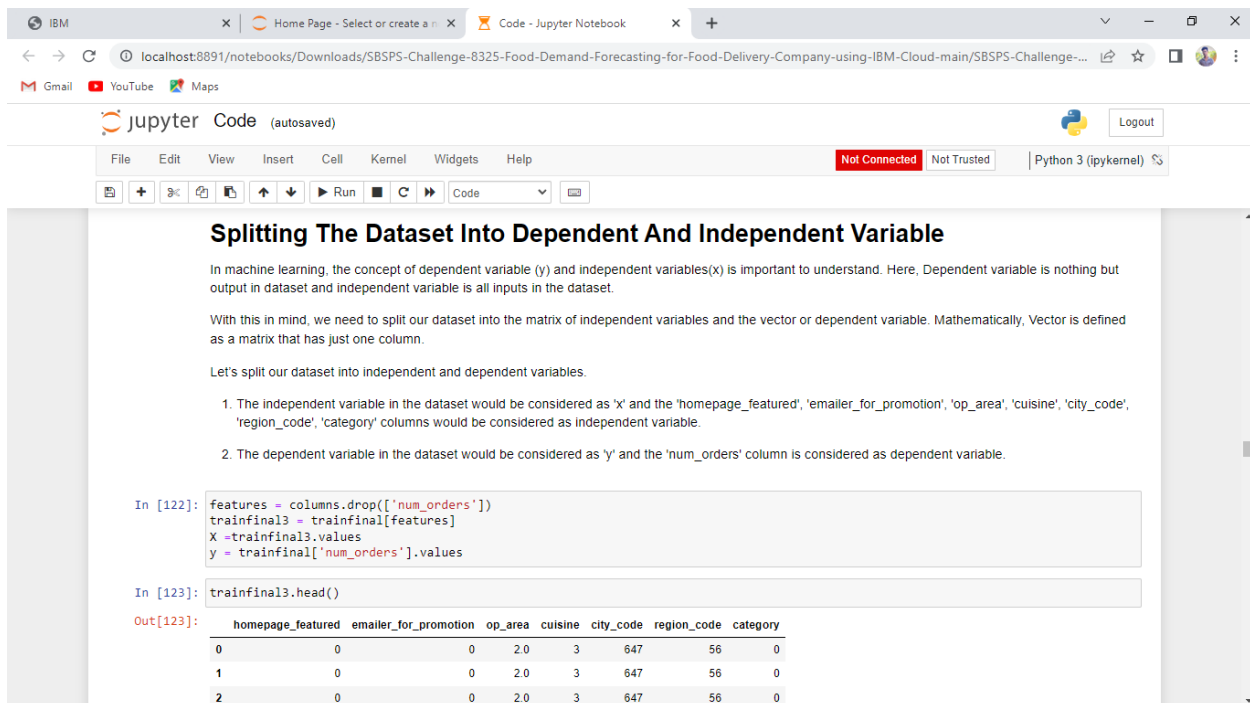


TEAM ID: PNT2022TMID14136

## PROJECT NAME: DemandEst - AI powered Food Demand Forecaster

### Team Leader



The screenshot shows a Jupyter Notebook interface with a title bar indicating it's running on IBM Cloud. The notebook is titled "Splitting The Dataset Into Dependent And Independent Variable". The content includes a brief explanation of dependent and independent variables in machine learning, followed by a list of two points: 1. The independent variable in the dataset would be considered as 'x' and the 'homepage\_featured', 'emailer\_for\_promotion', 'op\_area', 'cuisine', 'city\_code', 'region\_code', 'category' columns would be considered as independent variable. 2. The dependent variable in the dataset would be considered as 'y' and the 'num\_orders' column is considered as dependent variable. Below the text, there are two code cells. The first cell (In [122]) contains the following code: 

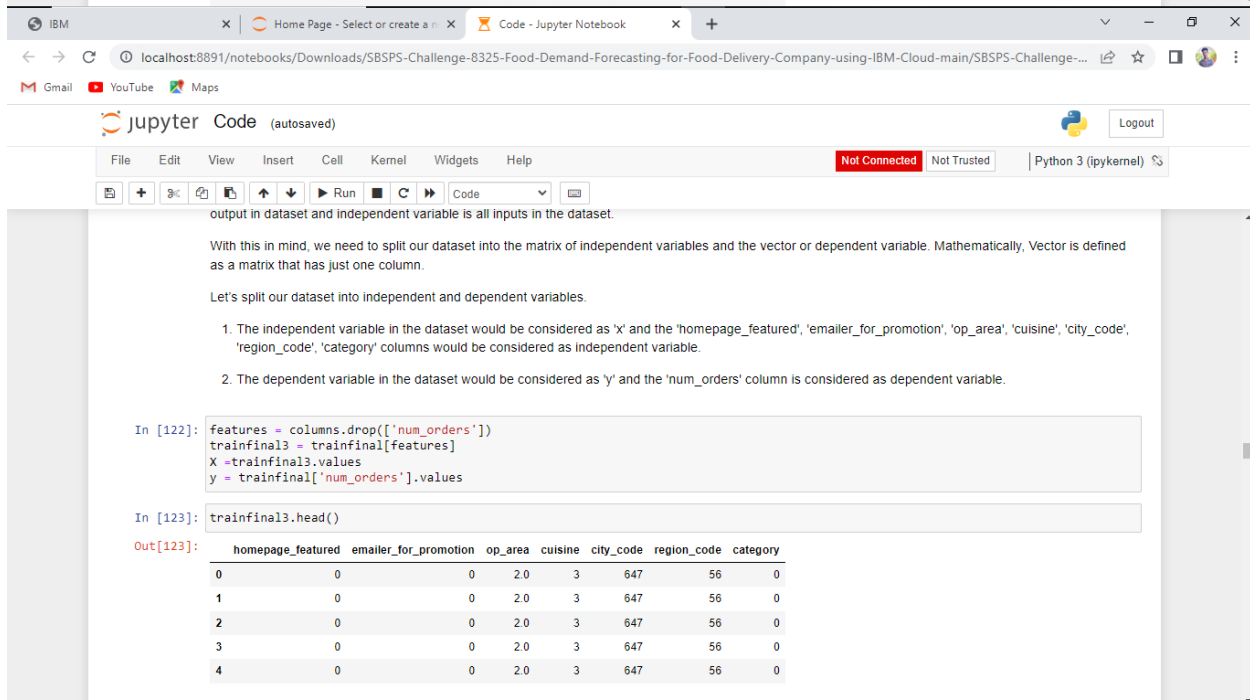
```
features = columns.drop(['num_orders'])
trainfinal3 = trainfinal[features]
X = trainfinal3.values
y = trainfinal['num_orders'].values
```

 The second cell (In [123]) contains the code: 

```
trainfinal3.head()
```

 The output of the second cell (Out[123]) shows a table with 8 columns: homepage\_featured, emailer\_for\_promotion, op\_area, cuisine, city\_code, region\_code, category, and num\_orders. The first three rows of data are displayed.

	homepage_featured	emailer_for_promotion	op_area	cuisine	city_code	region_code	category
0	0	0	2.0	3	647	56	0
1	0	0	2.0	3	647	56	0
2	0	0	2.0	3	647	56	0



This screenshot is identical to the one above, showing the same Jupyter Notebook interface and content. The only difference is in the output of the second cell (Out[123]), which now displays the first four rows of data from the table.

	homepage_featured	emailer_for_promotion	op_area	cuisine	city_code	region_code	category
0	0	0	2.0	3	647	56	0
1	0	0	2.0	3	647	56	0
2	0	0	2.0	3	647	56	0
3	0	0	2.0	3	647	56	0

# Team Member 1

The screenshot shows a Jupyter Notebook interface with the following content:

## Splitting The Dataset Into Dependent And Independent Variable

In machine learning, the concept of dependent variable (y) and independent variables(x) is important to understand. Here, Dependent variable is nothing but output in dataset and independent variable is all inputs in the dataset.

With this in mind, we need to split our dataset into the matrix of independent variables and the vector or dependent variable. Mathematically, Vector is defined as a matrix that has just one column.

Let's split our dataset into independent and dependent variables.

1. The independent variable in the dataset would be considered as 'x' and the 'homepage\_featured', 'emailer\_for\_promotion', 'op\_area', 'cuisine', 'city\_code', 'region\_code', 'category' columns would be considered as independent variable.
2. The dependent variable in the dataset would be considered as 'y' and the 'num\_orders' column is considered as dependent variable.

```
In [122]: features = columns.drop(['num_orders'])
trainfinal3 = trainfinal[features]
X = trainfinal3.values
y = trainfinal['num_orders'].values
```

```
In [123]: trainfinal3.head()
```

	homepage_featured	emailer_for_promotion	op_area	cuisine	city_code	region_code	category
0	0	0	2.0	3	647	56	0
1	0	0	2.0	3	647	56	0
2	0	0	2.0	3	647	56	0

The screenshot shows the continuation of the Jupyter Notebook with the following content:

output in dataset and independent variable is all inputs in the dataset.

With this in mind, we need to split our dataset into the matrix of independent variables and the vector or dependent variable. Mathematically, Vector is defined as a matrix that has just one column.

Let's split our dataset into independent and dependent variables.

1. The independent variable in the dataset would be considered as 'x' and the 'homepage\_featured', 'emailer\_for\_promotion', 'op\_area', 'cuisine', 'city\_code', 'region\_code', 'category' columns would be considered as independent variable.
2. The dependent variable in the dataset would be considered as 'y' and the 'num\_orders' column is considered as dependent variable.

```
In [122]: features = columns.drop(['num_orders'])
trainfinal3 = trainfinal[features]
X = trainfinal3.values
y = trainfinal['num_orders'].values
```

```
In [123]: trainfinal3.head()
```

	homepage_featured	emailer_for_promotion	op_area	cuisine	city_code	region_code	category
0	0	0	2.0	3	647	56	0
1	0	0	2.0	3	647	56	0
2	0	0	2.0	3	647	56	0
3	0	0	2.0	3	647	56	0
4	0	0	2.0	3	647	56	0

## Team Member 2

The screenshot shows a Jupyter Notebook interface with the following content:

### Splitting The Dataset Into Dependent And Independent Variable

In machine learning, the concept of dependent variable (y) and independent variables(x) is important to understand. Here, Dependent variable is nothing but output in dataset and independent variable is all inputs in the dataset.

With this in mind, we need to split our dataset into the matrix of independent variables and the vector or dependent variable. Mathematically, Vector is defined as a matrix that has just one column.

Let's split our dataset into independent and dependent variables.

1. The independent variable in the dataset would be considered as 'x' and the 'homepage\_featured', 'emailer\_for\_promotion', 'op\_area', 'cuisine', 'city\_code', 'region\_code', 'category' columns would be considered as independent variable.
2. The dependent variable in the dataset would be considered as 'y' and the 'num\_orders' column is considered as dependent variable.

```
In [122]: features = columns.drop(['num_orders'])
trainfinal3 = trainfinal[features]
X = trainfinal3.values
y = trainfinal['num_orders'].values
```

```
In [123]: trainfinal3.head()
```

```
Out[123]:
```

	homepage_featured	emailer_for_promotion	op_area	cuisine	city_code	region_code	category
0	0	0	2.0	3	647	56	0
1	0	0	2.0	3	647	56	0
2	0	0	2.0	3	647	56	0

The screenshot shows the continuation of the Jupyter Notebook with the following content:

output in dataset and independent variable is all inputs in the dataset.

With this in mind, we need to split our dataset into the matrix of independent variables and the vector or dependent variable. Mathematically, Vector is defined as a matrix that has just one column.

Let's split our dataset into independent and dependent variables.

1. The independent variable in the dataset would be considered as 'x' and the 'homepage\_featured', 'emailer\_for\_promotion', 'op\_area', 'cuisine', 'city\_code', 'region\_code', 'category' columns would be considered as independent variable.
2. The dependent variable in the dataset would be considered as 'y' and the 'num\_orders' column is considered as dependent variable.

```
In [122]: features = columns.drop(['num_orders'])
trainfinal3 = trainfinal[features]
X = trainfinal3.values
y = trainfinal['num_orders'].values
```

```
In [123]: trainfinal3.head()
```

```
Out[123]:
```

	homepage_featured	emailer_for_promotion	op_area	cuisine	city_code	region_code	category
0	0	0	2.0	3	647	56	0
1	0	0	2.0	3	647	56	0
2	0	0	2.0	3	647	56	0
3	0	0	2.0	3	647	56	0
4	0	0	2.0	3	647	56	0

## Team Member 3

The screenshot shows a Jupyter Notebook interface with a browser window at the top. The notebook title is "Splitting The Dataset Into Dependent And Independent Variable". The text explains the importance of splitting a dataset into independent variables (X) and a dependent variable (y). It lists two points: 1. Independent variables include 'homepage\_featured', 'emailer\_for\_promotion', 'op\_area', 'cuisine', 'city\_code', 'region\_code', and 'category'. 2. The dependent variable is 'num\_orders'. Below the text, there are two code cells. The first cell (In [122]) shows the code to split the dataset: `features = columns.drop(['num_orders'])`, `trainfinal3 = trainfinal[features]`, `X = trainfinal3.values`, and `y = trainfinal['num_orders'].values`. The second cell (In [123]) shows `trainfinal3.head()`. The output (Out[123]) is a table with 7 columns: homepage\_featured, emailer\_for\_promotion, op\_area, cuisine, city\_code, region\_code, and category. The first three rows of data are shown.

### Splitting The Dataset Into Dependent And Independent Variable

In machine learning, the concept of dependent variable (y) and independent variables(x) is important to understand. Here, Dependent variable is nothing but output in dataset and independent variable is all inputs in the dataset.

With this in mind, we need to split our dataset into the matrix of independent variables and the vector or dependent variable. Mathematically, Vector is defined as a matrix that has just one column.

Let's split our dataset into independent and dependent variables.

1. The independent variable in the dataset would be considered as 'x' and the 'homepage\_featured', 'emailer\_for\_promotion', 'op\_area', 'cuisine', 'city\_code', 'region\_code', 'category' columns would be considered as independent variable.
2. The dependent variable in the dataset would be considered as 'y' and the 'num\_orders' column is considered as dependent variable.

```
In [122]: features = columns.drop(['num_orders'])
          trainfinal3 = trainfinal[features]
          X = trainfinal3.values
          y = trainfinal['num_orders'].values
```

```
In [123]: trainfinal3.head()
```

```
Out[123]:
```

	homepage_featured	emailer_for_promotion	op_area	cuisine	city_code	region_code	category
0	0	0	2.0	3	647	56	0
1	0	0	2.0	3	647	56	0
2	0	0	2.0	3	647	56	0

This screenshot is similar to the first one, but it includes an additional line of text above the code cells: "output in dataset and independent variable is all inputs in the dataset." The rest of the content, including the explanation, points, code cells, and the output table, is identical to the first screenshot.

output in dataset and independent variable is all inputs in the dataset.

With this in mind, we need to split our dataset into the matrix of independent variables and the vector or dependent variable. Mathematically, Vector is defined as a matrix that has just one column.

Let's split our dataset into independent and dependent variables.

1. The independent variable in the dataset would be considered as 'x' and the 'homepage\_featured', 'emailer\_for\_promotion', 'op\_area', 'cuisine', 'city\_code', 'region\_code', 'category' columns would be considered as independent variable.
2. The dependent variable in the dataset would be considered as 'y' and the 'num\_orders' column is considered as dependent variable.

```
In [122]: features = columns.drop(['num_orders'])
          trainfinal3 = trainfinal[features]
          X = trainfinal3.values
          y = trainfinal['num_orders'].values
```

```
In [123]: trainfinal3.head()
```

```
Out[123]:
```

	homepage_featured	emailer_for_promotion	op_area	cuisine	city_code	region_code	category
0	0	0	2.0	3	647	56	0
1	0	0	2.0	3	647	56	0
2	0	0	2.0	3	647	56	0
3	0	0	2.0	3	647	56	0
4	0	0	2.0	3	647	56	0