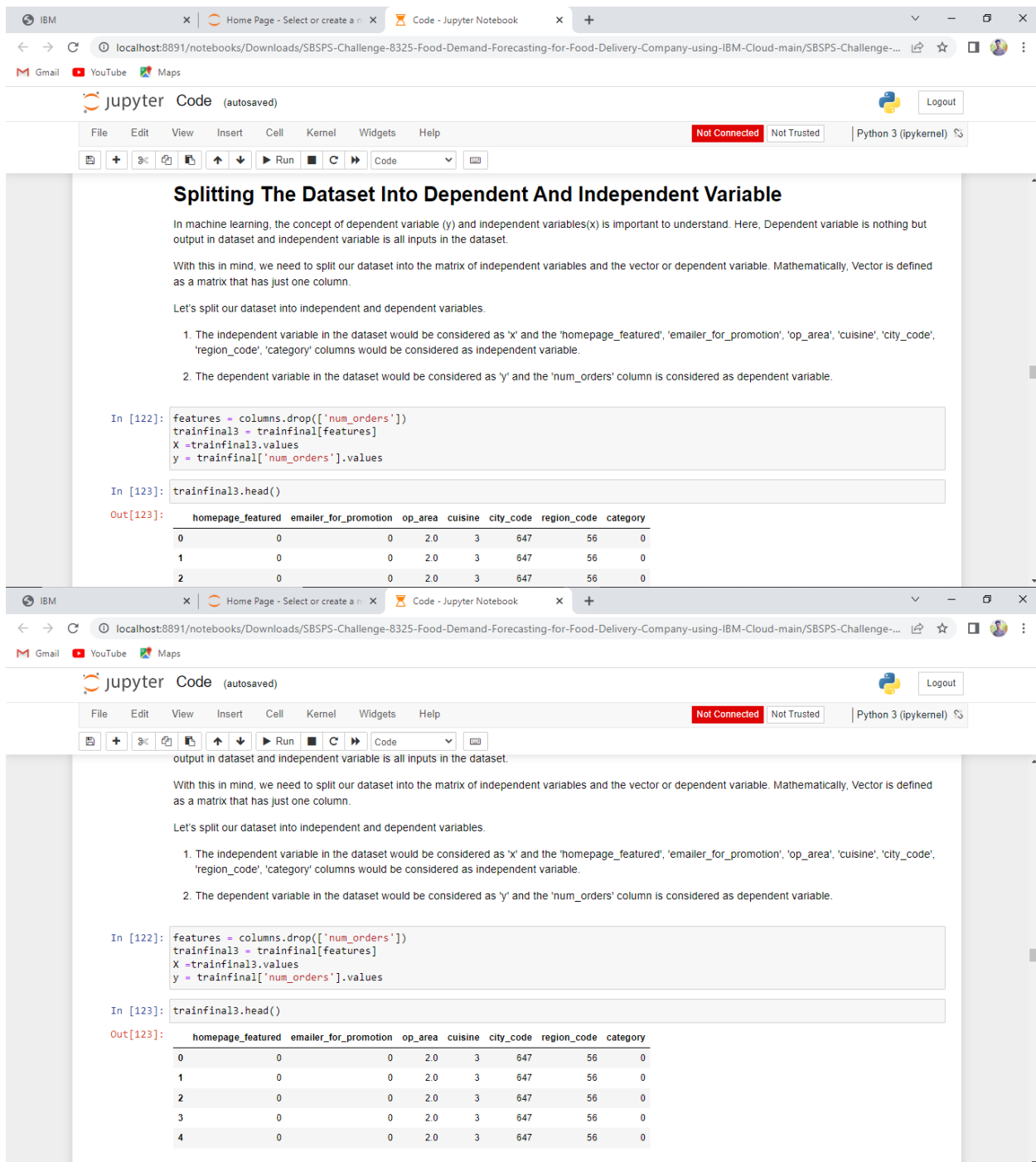


TEAM ID: PNT2022TMID17645

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

### Team Leader



**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

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 1

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 Notebook interface has a top bar with 'jupyter Code (autosaved)' and a 'Logout' button. Below the top bar is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Cell', 'Kernel', 'Widgets', and 'Help'. The 'Kernel' status is 'Not Connected' and 'Not Trusted'. The language is 'Python 3 (ipykernel)'. The notebook content includes a title 'Splitting The Dataset Into Dependent And Independent Variable', an introduction to dependent and independent variables, and two numbered points explaining the dataset split. The first point lists independent variables: 'homepage\_featured', 'emailer\_for\_promotion', 'op\_area', 'cuisine', 'city\_code', 'region\_code', and 'category'. The second point identifies 'num\_orders' as the dependent variable. Below the text are two code cells. The first cell (In [122]) contains Python code to drop the 'num\_orders' column and create 'X' and 'y' variables. The second cell (In [123]) contains the command 'trainfinal3.head()'. The output (Out[123]) shows a table with 7 columns and 3 rows of data.

## 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 same Jupyter Notebook interface as the first screenshot, but with the second cell of the notebook. The first cell's content is visible at the top. The second cell (In [122]) contains the same Python code as the first cell. The third cell (In [123]) contains the command 'trainfinal3.head()'. The output (Out[123]) shows a table with 7 columns and 5 rows of data, including the first three rows from the first screenshot and two additional rows.

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.

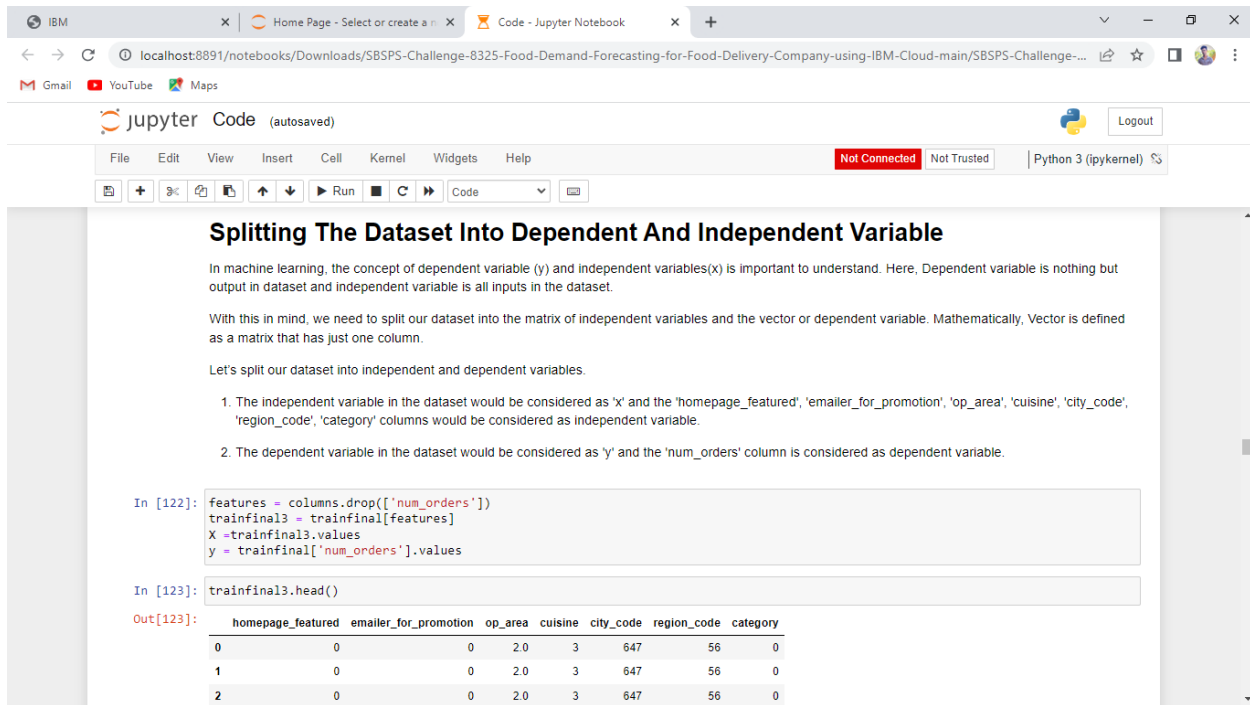
```
In [122]: features = columns.drop(['num_orders'])
trainfinal3 = trainfinal[features]
X = trainfinal3.values
y = trainfinal['num_orders'].values
```

```
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 2



The screenshot shows a Jupyter Notebook titled "Code (autosaved)" with a "Python 3 (ipykernel)" kernel. The notebook content includes a title, an introductory paragraph, a list of two points, and two code cells. The first code cell (In [122]) defines features, trainfinal3, X, and y. The second code cell (In [123]) displays the first three rows of the dataset using trainfinal3.head().

### 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.

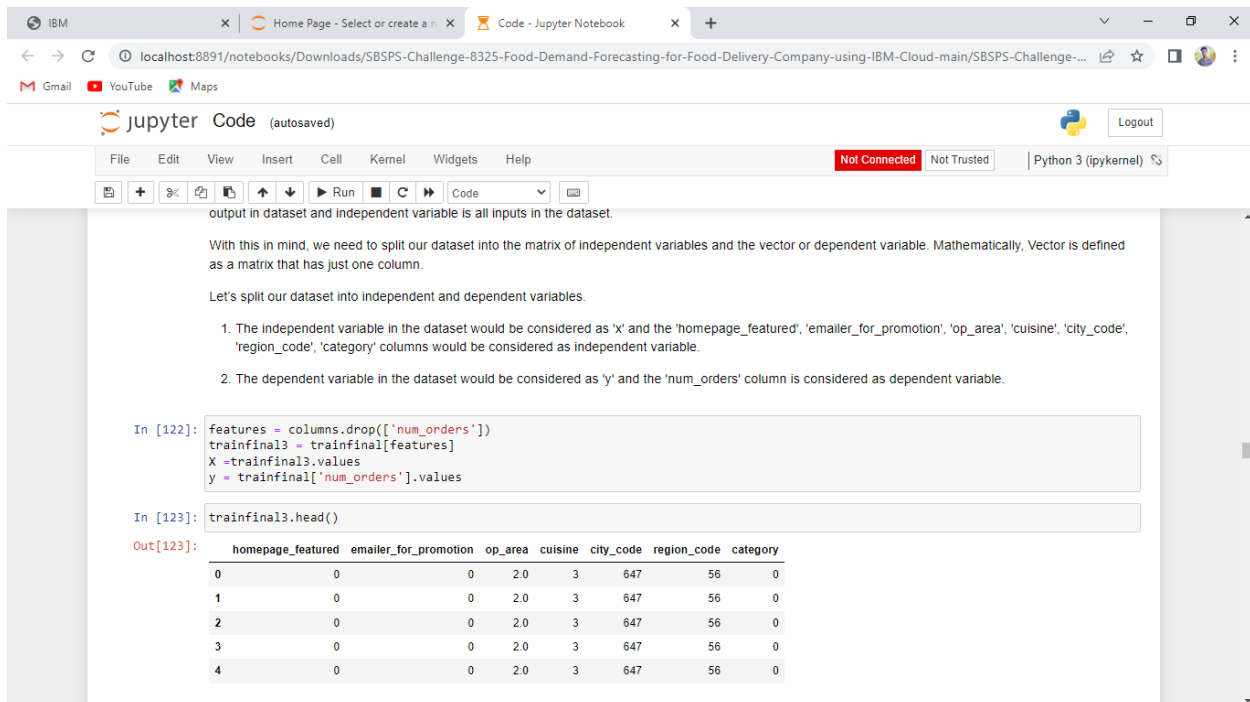
1. The independent variable in the dataset would be considered as 'x' and the 'homepage\_featured', 'emailer\_for\_promotion', 'op\_area', 'cuisine', 'city\_code', 'region\_code', 'category' columns would be considered as independent variable.
2. The dependent variable in the dataset would be considered 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 a continuation of the Jupyter Notebook interface, showing the same content as the previous block. The output table for Out[123] now includes an additional row (index 4).

```
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 titled "Code (autosaved)" with a toolbar at the top. The notebook content is as follows:

### 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 a Jupyter Notebook titled "Code (autosaved)" with a toolbar at the top. The notebook content is as follows:

### 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
3	0	0	2.0	3	647	56	0
4	0	0	2.0	3	647	56	0