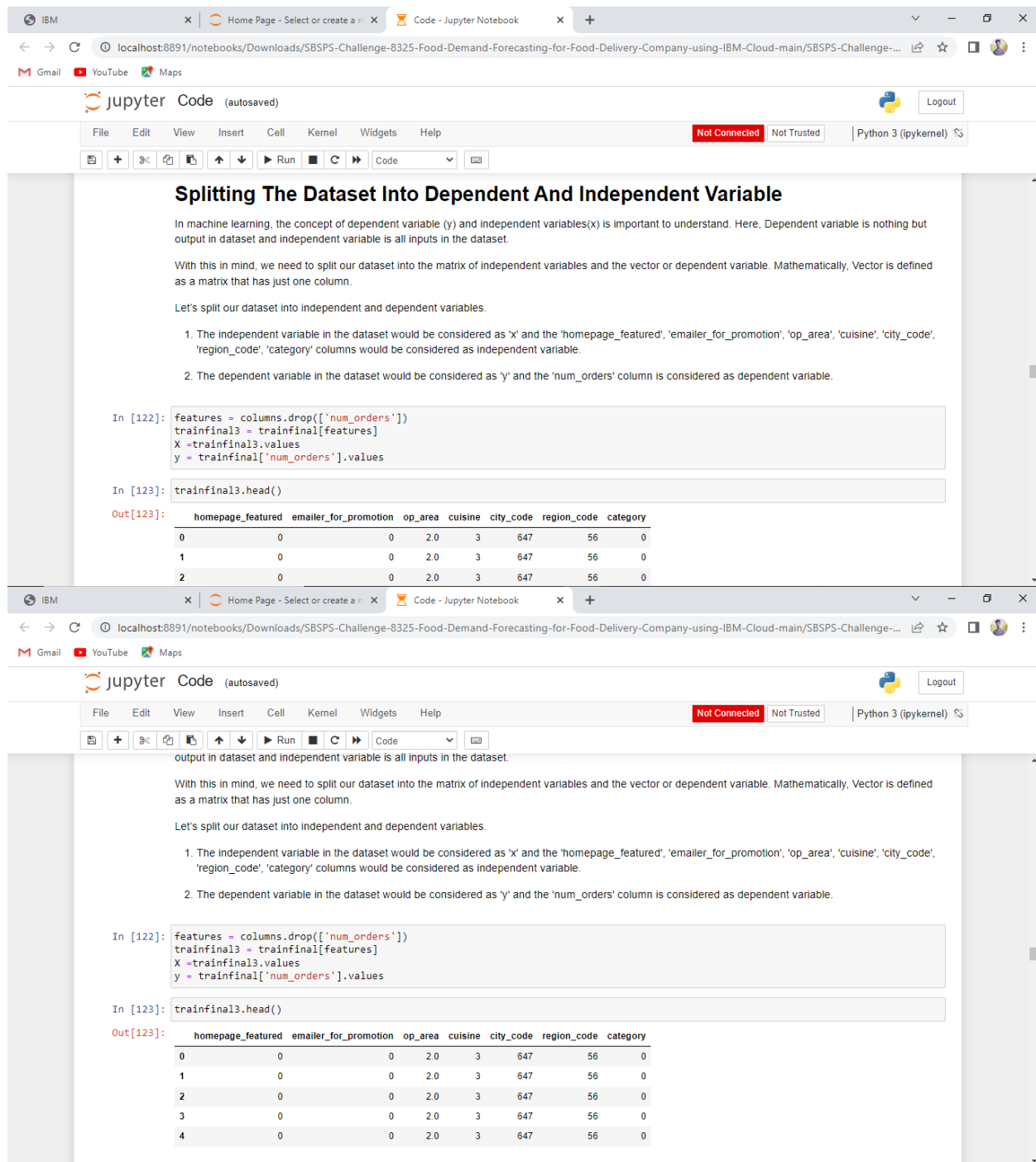


TEAM ID: PNT2022TMID17554

PROJECT NAME: DemandEst - AI powered Food Demand Forecaster Team Leader



The screenshot displays a Jupyter Notebook interface within 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 top bar with the 'jupyter Code' logo, a 'Logout' button, and a status bar indicating 'Not Connected' and 'Not Trusted' for the 'Python 3 (ipykernel)' environment. The notebook content is titled 'Splitting The Dataset Into Dependent And Independent Variable'. It contains a text block explaining the importance of splitting the dataset into independent variables (x) and dependent variables (y). Below this, there are two numbered 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. The code cell (In [122]) shows the following Python code:

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

 The output cell (Out[123]) shows the result of `trainfinal3.head()`, which is a table with 7 columns: `homepage_featured`, `emailer_for_promotion`, `op_area`, `cuisine`, `city_code`, `region_code`, and `category`. The table contains 3 rows of data, with the first row being `[0, 0, 0, 2.0, 3, 647, 56, 0]`.

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

Team Member 1

The screenshot shows a Jupyter Notebook titled "Code (autosaved)" running on a local host. The notebook content includes a title, an introductory paragraph about dependent and independent variables, a list of two points explaining the dataset split, and two code cells. The first code cell (In [122]) defines 'features' as all columns except 'num_orders', and 'y' as the 'num_orders' column. The second code cell (In [123]) displays the first three rows of the 'trainfinal3' dataset using the .head() method.

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 continues the Jupyter Notebook from the previous one. It shows the same introductory text and list of points. The code cell (In [122]) is identical. The code cell (In [123]) now displays the first five rows of the 'trainfinal3' dataset using the .head() method.

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 trainfinal3 dataset as a table.

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

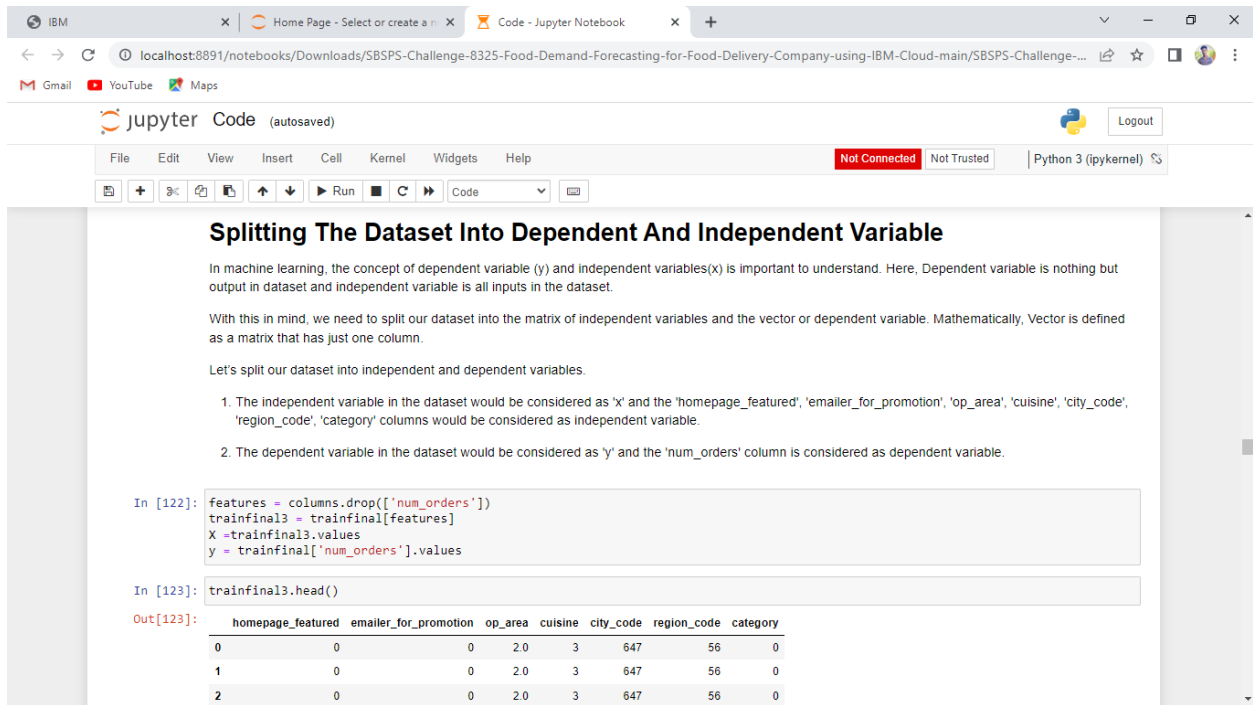
This screenshot is a continuation of the Jupyter Notebook from the previous image. It shows the same text and code cells. The output of the second code cell (In [123]) now includes an additional row (index 4) in the table, showing the first four rows of the dataset.

```
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 3



The screenshot shows a Jupyter Notebook titled "Code (autosaved)" with a toolbar at the top. The notebook content is titled "Splitting The Dataset Into Dependent And Independent Variable". It explains the concept of dependent and independent variables in machine learning and provides 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 is a table with 8 columns: homepage_featured, emailer_for_promotion, op_area, cuisine, city_code, region_code, category, and an unlabeled column. The table shows the first three 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.

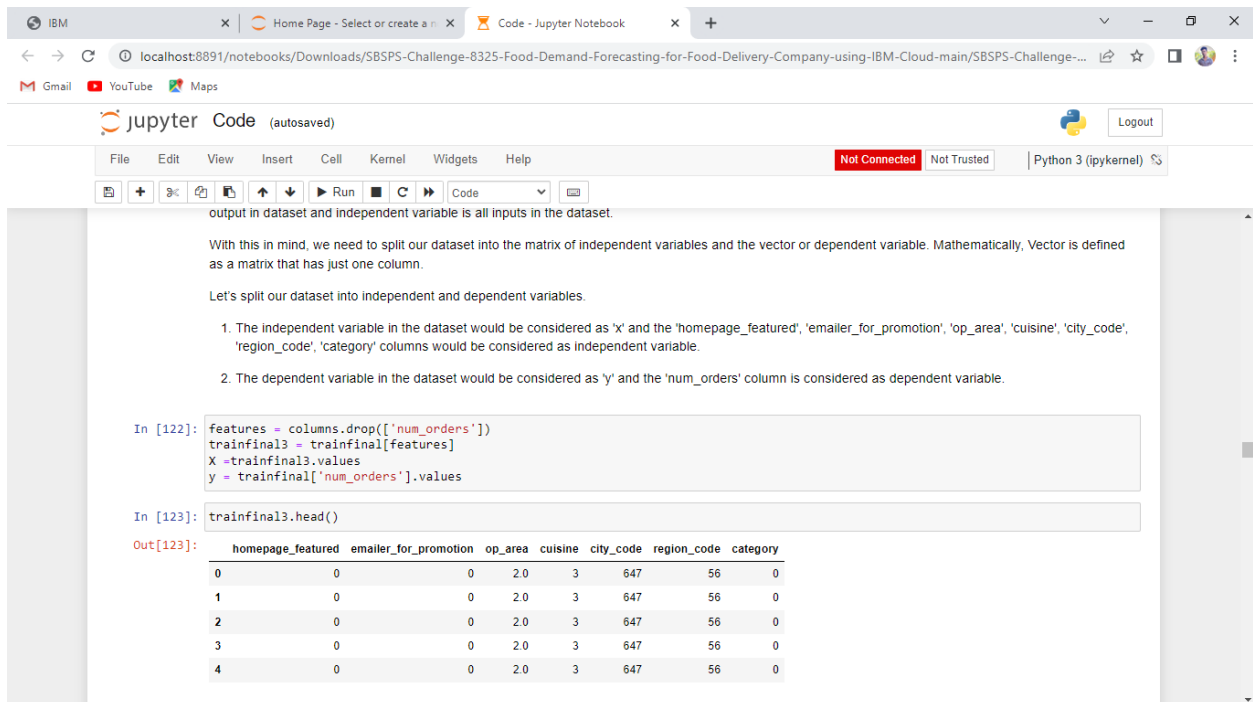
Let's split our dataset into independent and dependent variables.

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()
```

	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 titled "Splitting The Dataset Into Dependent And Independent Variable". It explains the concept of dependent and independent variables in machine learning and provides 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 is a table with 8 columns: homepage_featured, emailer_for_promotion, op_area, cuisine, city_code, region_code, category, and an unlabeled column. The table shows the first five rows of data.

Splitting The Dataset Into Dependent And Independent Variable

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