

TEAM ID: PNT2022TMID32006

PROJECT NAME: DemandEst - AI powered Food DemandForecaster

Team Leader

The screenshot displays a Jupyter Notebook interface with a browser window at the top showing the URL: localhost:8891/notebooks/Downloads/SBSPS-Challenge-8325-Food-Demand-Forecasting-for-Food-Delivery-Company-using-IBM-Cloud-main/SBSPS-Challenge-... The notebook is titled "jupyter Code (autosaved)" and shows a menu bar with File, Edit, View, Insert, Cell, Kernel, Widgets, and Help. The status bar indicates "Not Connected", "Not Trusted", and "Python 3 (ipykernel)".

The notebook content includes a section titled "Splitting The Dataset Into Dependent And Independent Variable". The text explains that in machine learning, the concept of dependent variable (y) and independent variables (x) is important. It states that the dependent variable is the output in the dataset, and independent variables are all inputs. It then explains that the dataset needs to be split into a matrix of independent variables and a vector of dependent variables. Mathematically, a vector is defined as a matrix with just one column. The text concludes by stating that the dataset will be split into independent and dependent variables.

Two numbered points are listed:

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 shows the following Python code:

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

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

The output cell shows the following 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

The second screenshot shows the same Jupyter Notebook interface, but with the output cell displaying a larger table with 5 rows:

	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

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

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

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