

Team ID: PNT2022TMID23576

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

### TEAM LEADER:

```
+ Code + Text
train.head()
id week center_id meal_id checkout_price base_price emailer_for_promotion homepage_featured num_orders
0 1379560 1 55 1885 136.83 152.29 0 0 177
1 1466964 1 55 1993 136.83 135.83 0 0 270
2 1346989 1 55 2539 134.86 135.86 0 0 189
3 1338232 1 55 2139 339.50 437.53 0 0 54
4 1448490 1 55 2631 243.50 242.50 0 0 40

[ ] test.head()

id week center_id meal_id checkout_price base_price emailer_for_promotion homepage_featured
0 1028232 146 55 1885 158.11 159.11 0 0
1 1127204 146 55 1993 160.11 159.11 0 0
2 1212707 146 55 2539 157.14 159.14 0 0
3 1082698 146 55 2631 162.02 162.02 0 0
4 1400926 146 55 1248 163.93 163.93 0 0

[ ] train.info()
<class 'pandas.core.frame.DataFrame'
RangeIndex: 456548 entries, 0 to 456547
Data columns (total 9 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   id          456548 non-null  int64  
 1   week        456548 non-null  int64  
 2   center_id   456548 non-null  int64  
 3   meal_id     456548 non-null  int64  
 4   checkout_price  456548 non-null  float64
 5   base_price   456548 non-null  float64
 6   emailer_for_promotion 456548 non-null  int64  
 7   homepage_featured 456548 non-null  int64  
 8   num_orders   456548 non-null  int64  
dtypes: float64(2), int64(7)
memory usage: 31.3 MB
```

```
+ Code + Text
[ ] train.info()
<class 'pandas.core.frame.DataFrame'
RangeIndex: 456548 entries, 0 to 456547
Data columns (total 9 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   id          456548 non-null  int64  
 1   week        456548 non-null  int64  
 2   center_id   456548 non-null  int64  
 3   meal_id     456548 non-null  int64  
 4   checkout_price  456548 non-null  float64
 5   base_price   456548 non-null  float64
 6   emailer_for_promotion 456548 non-null  int64  
 7   homepage_featured 456548 non-null  int64  
 8   num_orders   456548 non-null  int64  
dtypes: float64(2), int64(7)
memory usage: 31.3 MB

[ ] train['num_orders'].describe()
count    456548.000000
mean     261.872760
std      395.922798
min     13.000000
25%     54.000000
50%     136.000000
75%     324.000000
max    24299.000000
Name: num_orders, dtype: float64
```

## TEAM MEMBER 1:

The screenshot shows two sessions of code execution in Google Colab. The top session displays the head of the training and testing datasets, while the bottom session shows the full dataset and its descriptive statistics.

**Session 1 (Top):**

```
+ Code + Text
[ ] train.head()

   id week center_id meal_id checkout_price base_price emailer_for_promotion homepage_featured num_orders
0 1379560 1 55 1885 136.83 152.29 0 0 177
1 1466964 1 55 1993 136.83 135.83 0 0 270
2 1346989 1 55 2539 134.86 135.86 0 0 189
3 1338232 1 55 2139 339.50 437.53 0 0 54
4 1448490 1 55 2631 243.50 242.50 0 0 40

[ ] test.head()

   id week center_id meal_id checkout_price base_price emailer_for_promotion homepage_featured
0 1028232 146 55 1885 158.11 159.11 0 0
1 1127204 146 55 1993 160.11 159.11 0 0
2 1212707 146 55 2539 157.14 159.14 0 0
3 1082698 146 55 2631 162.02 162.02 0 0
4 1400926 146 55 1248 163.93 163.93 0 0

[ ] train.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 456548 entries, 0 to 456547
Data columns (total 9 columns):
 # Column Non-Null Count Dtype 
---  --- 
 0 id      456548 non-null int64 
 1 week    456548 non-null int64 
 2 center_id 456548 non-null int64 
 3 meal_id  456548 non-null int64 
 4 checkout_price 456548 non-null float64
 5 base_price 456548 non-null float64
 6 emailer_for_promotion 456548 non-null int64 
 7 homepage_featured 456548 non-null int64 
 8 num_orders 456548 non-null int64 
dtypes: float64(2), int64(7)
memory usage: 31.3 MB
```

**Session 2 (Bottom):**

```
+ Code + Text
[ ] train['num_orders'].describe()

count    456548.000000
mean     261.872760
std      395.922798
min      13.000000
25%      54.000000
50%     136.000000
75%     324.000000
max     24299.000000
Name: num_orders, dtype: float64


```

## TEAM MEMBER 2:

The screenshot shows a Jupyter Notebook interface with three tabs at the top: 'IBM', 'WhatsApp', and 'Data\_preprocessing\_and\_Model'. The main area contains the following code and output:

```
+ Code + Text
[ ] train.head()
<ipython.core.debugger> In [1]: train = pd.read_csv('train.csv')
Out[1]: DataFrame[9 columns]
   id week center_id meal_id checkout_price base_price emailer_for_promotion homepage_featured num_orders
0 1379560 1 55 1885 136.83 152.29 0 0 177
1 1466964 1 55 1993 136.83 135.83 0 0 270
2 1346989 1 55 2539 134.86 135.66 0 0 189
3 1338232 1 55 2139 339.50 437.53 0 0 54
4 1448490 1 55 2631 243.50 242.50 0 0 40

[ ] test.head()
<ipython.core.debugger> In [2]: test = pd.read_csv('test.csv')
Out[2]: DataFrame[9 columns]
   id week center_id meal_id checkout_price base_price emailer_for_promotion homepage_featured num_orders
0 1028232 146 55 1885 158.11 159.11 0 0 0
1 1127204 146 55 1993 160.11 159.11 0 0 0
2 1212707 146 55 2539 157.14 159.14 0 0 0
3 1082698 146 55 2631 162.02 162.02 0 0 0
4 1400926 146 55 1248 163.93 163.93 0 0 0

[ ] train.info()
<ipython.core.debugger> In [3]: train.info()
Out[3]: <class 'pandas.core.frame.DataFrame'>
RangeIndex: 456548 entries, 0 to 456547
Data columns (total 9 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   id              456548 non-null  int64  
 1   week            456548 non-null  int64  
 2   center_id       456548 non-null  int64  
 3   meal_id         456548 non-null  int64  
 4   checkout_price  456548 non-null  float64
 5   base_price      456548 non-null  float64
 6   emailer_for_promotion 456548 non-null  int64  
 7   homepage_featured 456548 non-null  int64  
 8   num_orders      456548 non-null  int64  
dtypes: float64(2), int64(7)
memory usage: 31.3 MB
```

The notebook also includes a search bar at the bottom and a taskbar with various icons.

## TEAM MEMBER 3:

```
+ IBM WhatsApp Model evaluation last 3 pdfs - sn Data_preprocessing_and_Model +  
colab.research.google.com/drive/10liQx3kt5FxailbDB7MRxkgXxBcpAOX  
+ Code + Text  
train.head()  
[x] id week center_id meal_id checkout_price base_price emailer_for_promotion homepage_featured num_orders  
0 1379560 1 55 1885 136.83 152.29 0 0 177  
1 1466964 1 55 1993 136.83 135.83 0 0 270  
2 1346989 1 55 2539 134.86 135.86 0 0 189  
3 1338232 1 55 2139 339.50 437.53 0 0 54  
4 1448490 1 55 2631 243.50 242.50 0 0 40  
[ ] test.head()  
id week center_id meal_id checkout_price base_price emailer_for_promotion homepage_featured  
0 1028232 146 55 1885 158.11 159.11 0 0  
1 1127204 146 55 1993 160.11 159.11 0 0  
2 1212707 146 55 2539 157.14 159.14 0 0  
3 1082696 146 55 2631 162.02 162.02 0 0  
4 1400926 146 55 1248 163.93 163.93 0 0  
[ ] train.info()  
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 456548 entries, 0 to 456547  
Data columns (total 9 columns):  
 # Column Non-Null Count Dtype  
--- ---  
 0 id 456548 non-null int64  
 1 week 456548 non-null int64  
 2 center_id 456548 non-null int64  
 3 meal_id 456548 non-null int64  
 4 checkout_price 456548 non-null float64  
 5 base_price 456548 non-null float64  
 6 emailer_for_promotion 456548 non-null int64  
 7 homepage_featured 456548 non-null int64  
 8 num_orders 456548 non-null int64  
dtypes: float64(2), int64(7)  
memory usage: 31.3 MB  
Type here to search
```

```
+ IBM WhatsApp Model evaluation last 3 pdfs - sn Data_preprocessing_and_Model +  
colab.research.google.com/drive/10liQx3kt5FxailbDB7MRxkgXxBcpAOX  
+ Code + Text  
[ ] 2 1212707 146 55 2539 157.14 159.14 0 0  
3 1082696 146 55 2631 162.02 162.02 0 0  
4 1400926 146 55 1248 163.93 163.93 0 0  
[ ] train.info()  
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 456548 entries, 0 to 456547  
Data columns (total 9 columns):  
 # Column Non-Null Count Dtype  
--- ---  
 0 id 456548 non-null int64  
 1 week 456548 non-null int64  
 2 center_id 456548 non-null int64  
 3 meal_id 456548 non-null int64  
 4 checkout_price 456548 non-null float64  
 5 base_price 456548 non-null float64  
 6 emailer_for_promotion 456548 non-null int64  
 7 homepage_featured 456548 non-null int64  
 8 num_orders 456548 non-null int64  
dtypes: float64(2), int64(7)  
memory usage: 31.3 MB  
[ ] train['num_orders'].describe()  
count 456548.000000  
mean 261.872760  
std 395.922798  
min 13.000000  
25% 54.000000  
50% 136.000000  
75% 324.000000  
max 24299.000000  
Name: num_orders, dtype: float64  
Type here to search
```

## TEAM MEMBER 4:

The screenshot shows a Jupyter Notebook interface with two tabs: 'Data\_preprocessing\_and\_Model' and 'Model\_evaluation last 3 pdfs - sri'. The current tab is 'Data\_preprocessing\_and\_Model'.

Code cells are visible, showing the execution of various pandas DataFrame operations:

- `train.head()`: Displays the first 5 rows of the training dataset. The columns include id, week, center\_id, meal\_id, checkout\_price, base\_price, emailer\_for\_promotion, homepage\_featured, and num\_orders.
- `test.head()`: Displays the first 5 rows of the test dataset, which has the same structure as the training data.
- `train.info()`: Provides detailed information about the training DataFrame, including the number of entries (456548), column names, non-null counts, and data types.
- `train.describe()`: Provides descriptive statistics for the 'num\_orders' column, including count, mean, std, min, 25%, 50%, 75%, and max values.
- `train['num_orders'].describe()`: Another call to describe the 'num\_orders' column, showing similar results to the previous command.

The interface includes a toolbar at the top with various icons for file operations, and a taskbar at the bottom with icons for different applications like WhatsApp, Google Chrome, and Microsoft Word.