

## Exploratory Data Analysis

In [96]: train.head()

Out[96]:

	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

In [97]: test.head()

Out[97]:

	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

```
In [98]: 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
```

```
In [99]: test.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 32573 entries, 0 to 32572  
Data columns (total 8 columns):  
#   Column                Non-Null Count  Dtype  
---  ---  
0   id                    32573 non-null  int64  
1   week                 32573 non-null  int64  
2   center_id            32573 non-null  int64  
3   meal_id              32573 non-null  int64
```