2023 | By: Shailesh Patel

Food-Delivery (Pandas)

[Exploratory Data Analysis by Pandas]

Food Delivery

Data of food Delivery company

Dataset

The Dataset is used in this project is taken from online site. The dataset in (.CSV) format, this dataset contains some details of food delivery.

Used Languages

In this analysis the language used widely is PYTHON and Python's libraries (Pandas, Matplotlib, Seaborn, Warnings).

Used Software

Jupyter notebook

Problem statement

The company's growth is going down. In this situation company is not getting much profit.

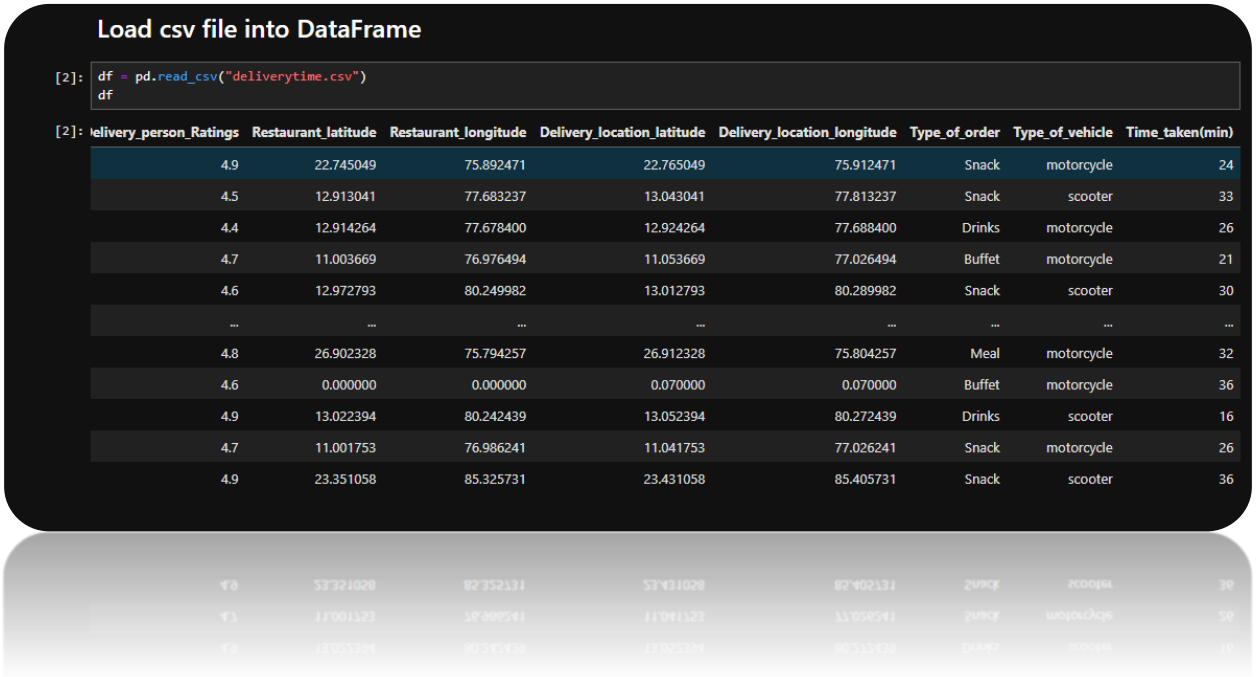
To solve this problem, we have to analyze the data and find some insights in the dataset.

Insights to find

- **Most ordered type of order**
- **Delivery Person Age**
- **Time Consumed in Delivery**
- **Type of Vehicle Used in Delivery**
- **Rating of Vehicles**

Loading Dataset

By using pandas functions the Dataset is loaded in a Data-Structure named DataFrame.



Load csv file into DataFrame

```
[2]: df = pd.read_csv("deliverytime.csv")
     df
```

[2]:

elivery_person_Ratings	Restaurant_latitude	Restaurant_longitude	Delivery_location_latitude	Delivery_location_longitude	Type_of_order	Type_of_vehicle	Time_taken(min)
4.9	22.745049	75.892471	22.765049	75.912471	Snack	motorcycle	24
4.5	12.913041	77.683237	13.043041	77.813237	Snack	scooter	33
4.4	12.914264	77.678400	12.924264	77.688400	Drinks	motorcycle	26
4.7	11.003669	76.976494	11.053669	77.026494	Buffet	motorcycle	21
4.6	12.972793	80.249982	13.012793	80.289982	Snack	scooter	30
...	...	...	...	...	...	...	...
4.8	26.902328	75.794257	26.912328	75.804257	Meal	motorcycle	32
4.6	0.000000	0.000000	0.070000	0.070000	Buffet	motorcycle	36
4.9	13.022394	80.242439	13.052394	80.272439	Drinks	scooter	16
4.7	11.001753	76.986241	11.041753	77.026241	Snack	motorcycle	26
4.9	23.351058	85.325731	23.431058	85.405731	Snack	scooter	36

Data Cleaning

Before performing any analytical task, we must clean the data. In this scenario we have used some python functions to clean the current dataset.

Data Cleaning ¶

```
[3]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 45593 entries, 0 to 45592
Data columns (total 11 columns):
 #   Column                       Non-Null Count  Dtype
---  ------                       --------------  -----
 0   ID                           45593 non-null  object
 1   Delivery_person_ID           45593 non-null  object
 2   Delivery_person_Age          45593 non-null  int64
 3   Delivery_person_Ratings      45593 non-null  float64
 4   Restaurant_latitude          45593 non-null  float64
 5   Restaurant_longitude         45593 non-null  float64
 6   Delivery_location_latitude   45593 non-null  float64
 7   Delivery_location_longitude  45593 non-null  float64
 8   Type_of_order                45593 non-null  object
 9   Type_of_vehicle              45593 non-null  object
 10  Time_taken(min)              45593 non-null  int64
dtypes: float64(5), int64(2), object(4)
memory usage: 3.8+ MB
```

```
[4]: df.isnull().count()
```

```
[4]: ID                           45593
     Delivery_person_ID           45593
     Delivery_person_Age          45593
     Delivery_person_Ratings      45593
     Restaurant_latitude          45593
     Restaurant_longitude         45593
```

Cleaned Data

```
[5]: df.describe()
```

[5]:

	Delivery_person_Age	Delivery_person_Ratings	Restaurant_latitude	Restaurant_longitude	Delivery_location_latitude	Delivery_location_longitude	Time_taken(min)
count	45593.000000	45593.000000	45593.000000	45593.000000	45593.000000	45593.000000	45593.000000
mean	29.544075	4.632367	17.017729	70.231332	17.465186	70.845702	26.294607
std	5.696793	0.327708	8.185109	22.883647	7.335122	21.118812	9.383806
min	15.000000	1.000000	-30.905562	-88.366217	0.010000	0.010000	10.000000
25%	25.000000	4.600000	12.933284	73.170000	12.988453	73.280000	19.000000
50%	29.000000	4.700000	18.546947	75.898497	18.633934	76.002574	26.000000
75%	34.000000	4.800000	22.728163	78.044095	22.785049	78.107044	32.000000
max	50.000000	6.000000	30.914057	88.433452	31.054057	88.563452	54.000000

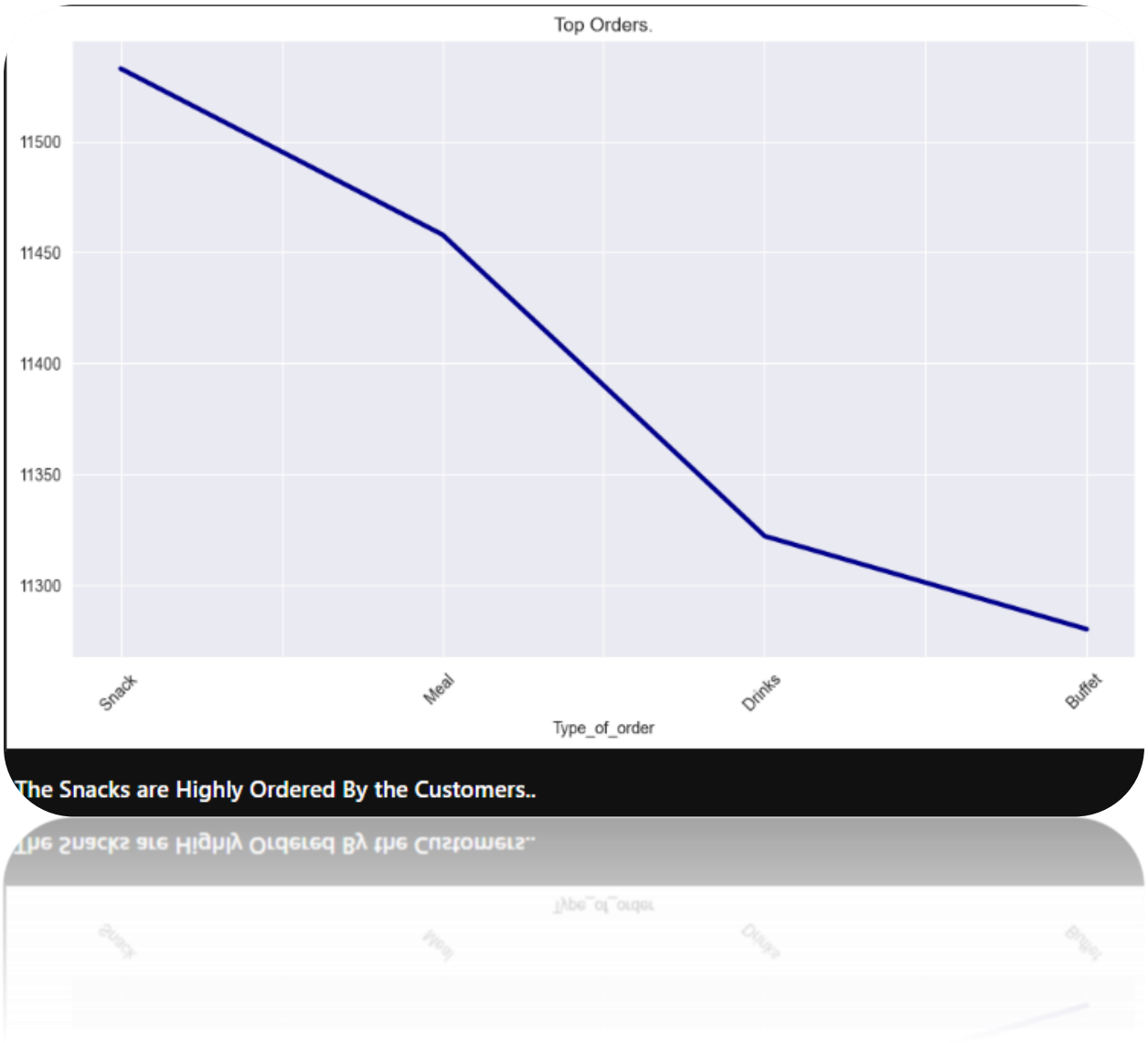
Data is clean !

Now the Data is cleaned and ready to perform analytical task.

Finding insights

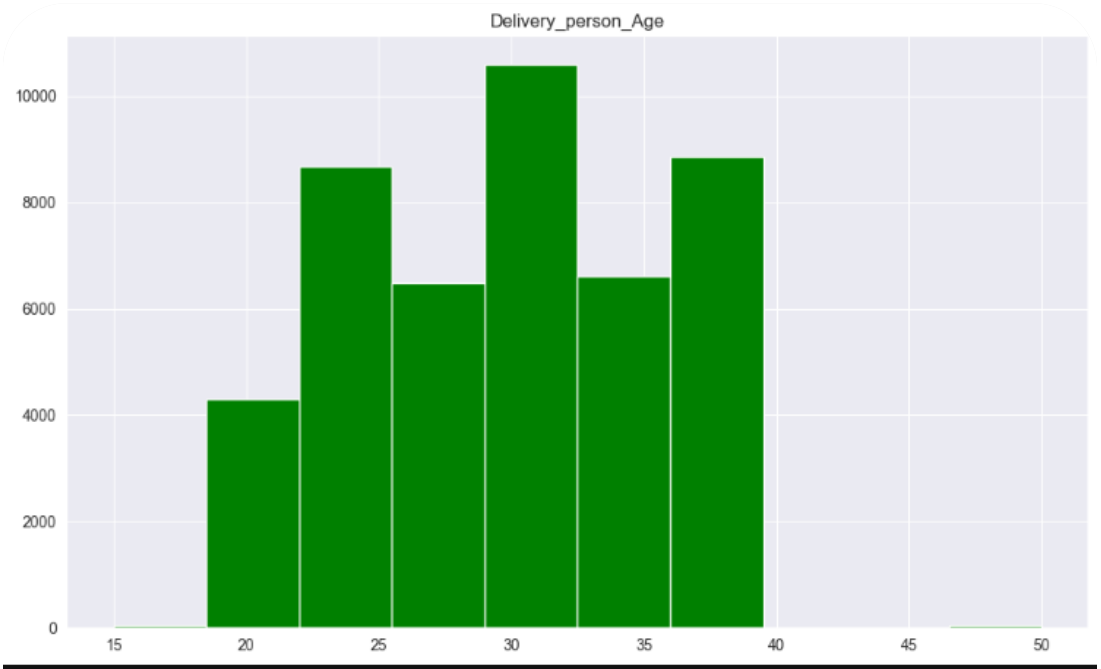
Most ordered type of Order

```
Type_of_order
Snack      11533
Meal       11458
Drinks     11322
Buffet     11280
Name: count, dtype: int64
```



In Graph we can observe that the trending Order is **Snacks**.

Delivery Person Age



Most of People are Age 30..

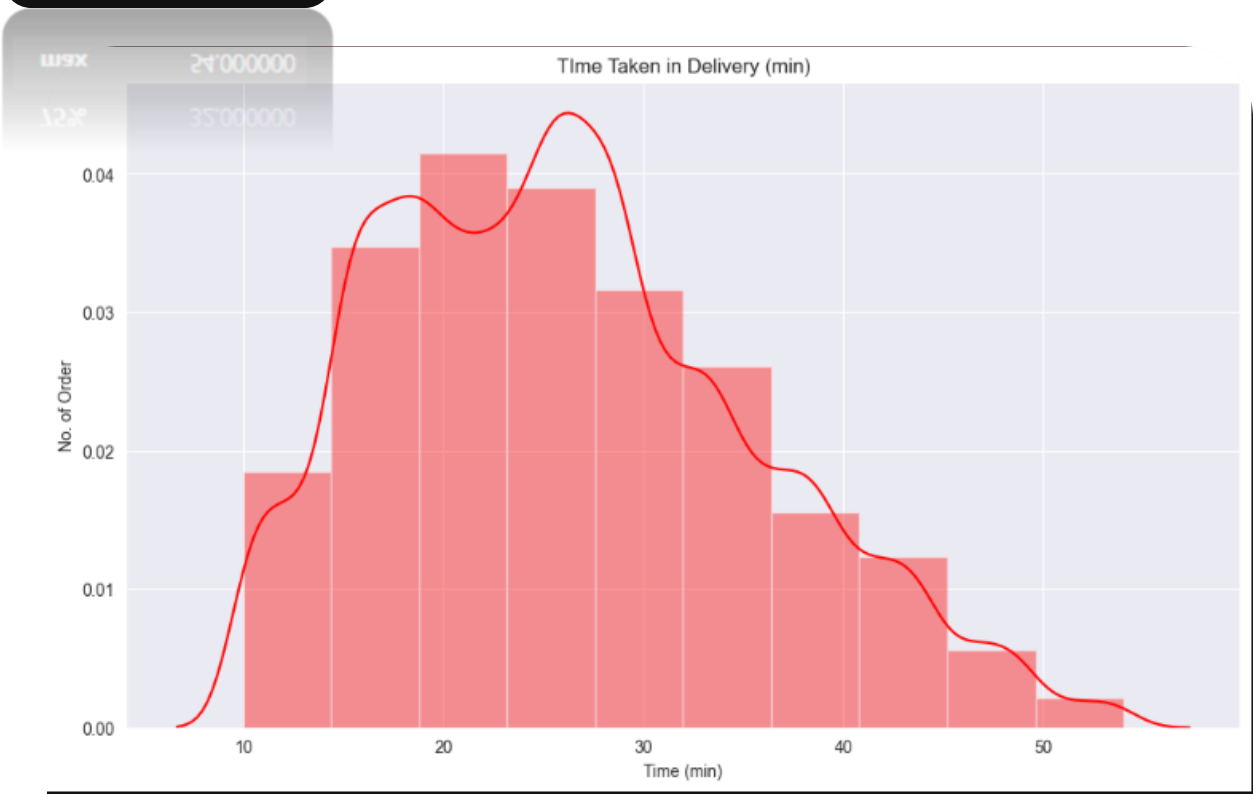
	Delivery_person_Age
count	45593.000000
mean	29.544075
std	5.696793
min	15.000000
25%	25.000000
50%	29.000000
75%	34.000000
max	50.000000

Here the maximum people age belongs to **28-33**

Time Consumed in Delivery

	Time_taken(min)
count	45593.000000
mean	26.294607
std	9.383806
min	10.000000
25%	19.000000
50%	26.000000
75%	32.000000
max	54.000000

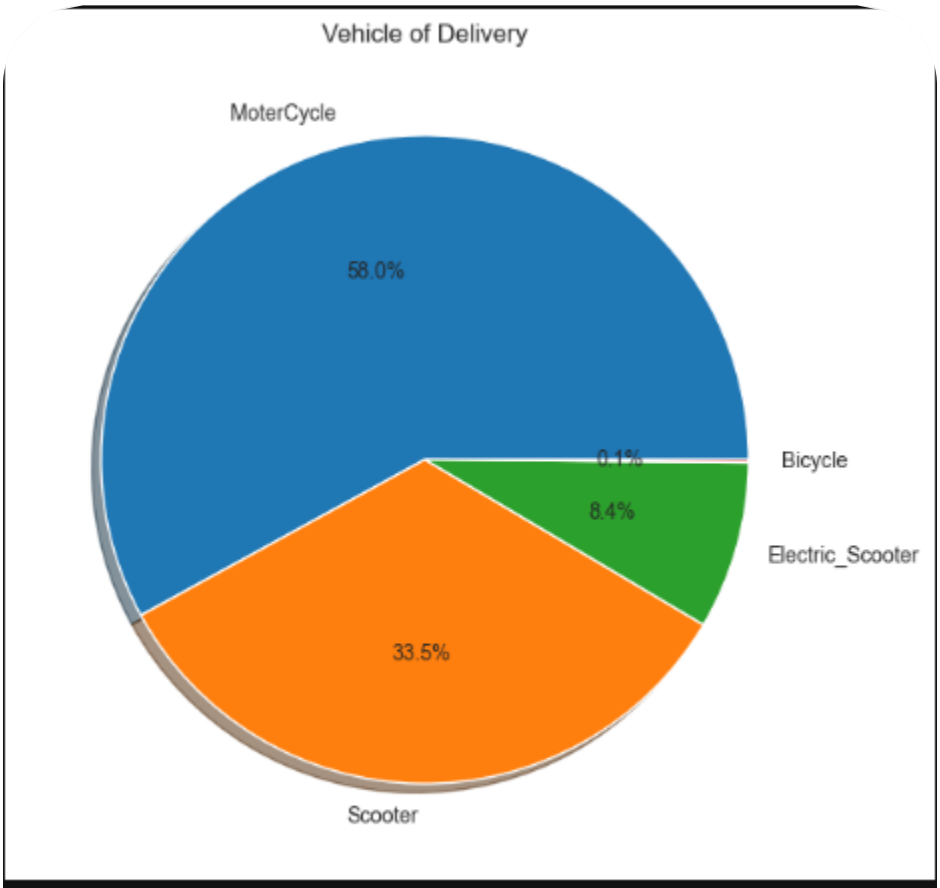
The maximum Delivery is Delivered in **20-26 minutes**



The Highest Time of Delivery is 50-55 min is there

Type of Vehicle Used in Delivery

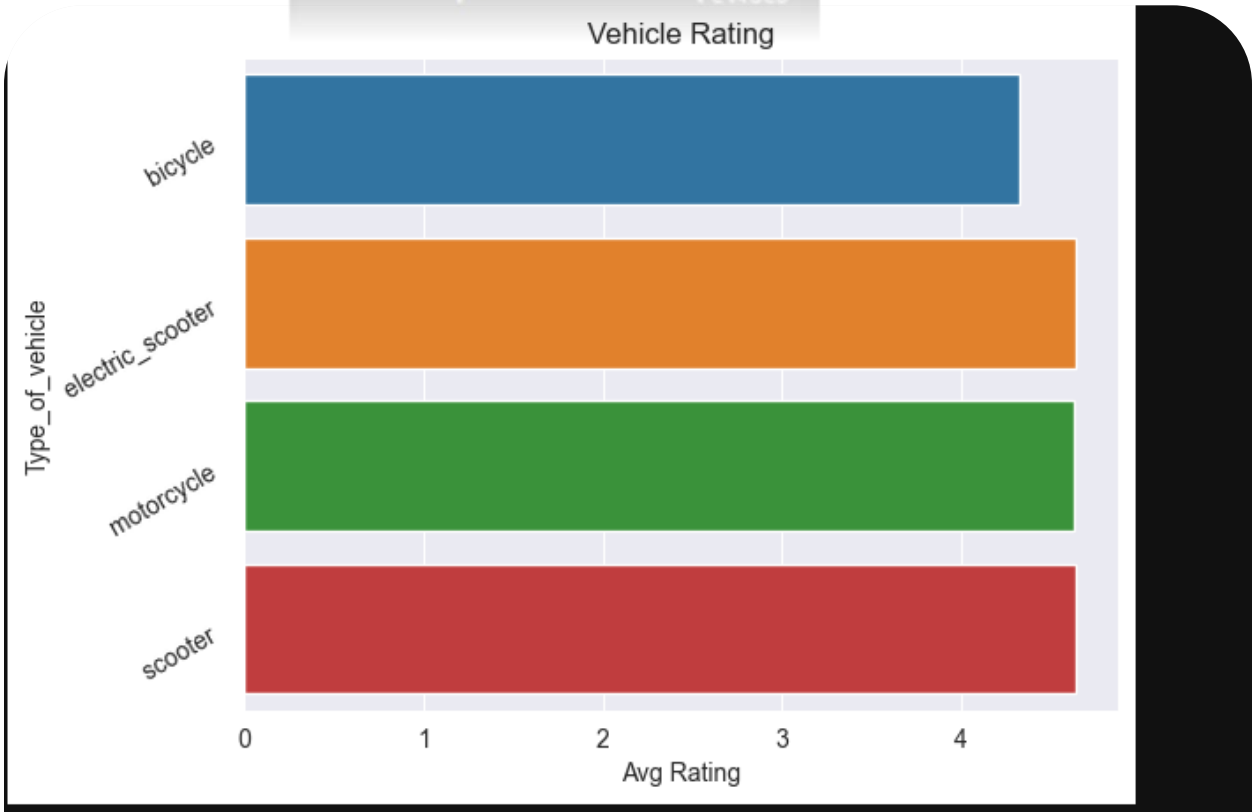
```
Type_of_vehicle
motorcycle           26435
scooter              15276
electric_scooter      3814
bicycle                 68
Name: count, dtype: int64
```



The Most Used Vehicle is "MoterCycle".

Ratings of Vehicles

Type_of_vehicle	Delivery_person_Ratings
bicycle	4.320588
electric_scooter	4.641846
motorcycle	4.626253
scooter	4.641968



Now here we can see the Highest avg rating goes to "Scooter".

Solution Suggestion

- Improve other foods quality and add new food type. Adding new food type would be attract more people to Order
- Try to decrease the timing of delivery. Some deliveries are taking 1 hour approx.
- Improve the vehicle used in delivery because the rating of scooter and motorcycle are high instead of others vehicles.

Thank you 😊