GROUP_C_5

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

Food is an essential need for man's survival. Everyone of us at certain points has either had a best food or cravings for certain kinds of foods and would most likely go to any length to get it either by cooking or eating out.

We also have ideal situations where the factors like the weather influences what we eat. For example, certain people wouldn't order for a certain meal when the weather is bad, because they feel delivery guys wouldn't work, or movement might be impeded, or maybe such meals might affect their health. There are also days a restaurant gets to be patronized the most.

This reality helped use define our objectives for the project.

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

Objectives

1. What food customers like the most?

2. Does weather influence what people eat?

3. What is the preferred mode of transportation used by the restaurant?

4. What time do people pick their orders most?

5. Which days are the busiest in the restaurant [which days do customers visit the restaurant the most]?

6. Which days of the week does the restaurant delivers most patronize?

7. What is the relationship between the time taken to deliver food and the distance travelled?

8. What is the correlation between days of the week with respect to the type of vehicle used to deliver food and the distance covered]?

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

Machine learning scenario:

Mrs Ndidi ordered meals at the amazon restaurant on Thursday morning. Predict how long before she received her order if the delivery man used motorcycle bearing in mind that the customer stays 9.760928 km away from the restaurant?

Models used:

Decision tree

Linear Regression

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

Importing libraries and loading datasets

```
In [48]:  import pandas as pd
          import numpy as np
          import matplotlib.pyplot as plt
          import seaborn as sn
```

```
In [2]:  food=pd.read_csv('cleaned_test.csv')
         food.head()
```

Out[2]:

	ID	Delivery_person_ID	Delivery_person_Age	Delivery_person_Ratings	Restaurant_latitude	Restaurant_longitude	Delivery_location_latitude	Delivery_loca
0	0x2318	COIMBRES13DEL01	NaN	NaN	11.003669	76.976494	11.043669	
1	0x3474	BANGRES15DEL01	28.0	4.6	12.975377	77.696664	13.085377	
2	0x9420	JAPRES09DEL03	23.0	4.5	26.911378	75.789034	27.001378	
3	0x72ee	JAPRES07DEL03	21.0	4.8	26.766536	75.837333	26.856536	
4	0xa759	CHENRES19DEL01	31.0	4.6	12.986047	80.218114	13.096047	

```
In [91]:  food.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 11399 entries, 0 to 11398
Data columns (total 23 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   ID                    11399 non-null  object
```

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

Data Pre-processing: Data Cleaning

converting the Order_Date to category datatype

```
In [3]: food['Order_Date']=pd.to_datetime(food['Order_Date'])
        food['Days']=food['Order_Date'].dt.strftime('%A')
        food['Days']=food['Days'].astype('category')
```

```
C:\Users\juliy\anaconda3\lib\site-packages\pandas\core\tools\datetimes.py:1047: UserWarning: Parsing '30-03-2022' in DD/MM/YYYY
format. Provide format or specify infer_datetime_format=True for consistent parsing.
  cache_array = _maybe_cache(arg, format, cache, convert_listlike)
C:\Users\juliy\anaconda3\lib\site-packages\pandas\core\tools\datetimes.py:1047: UserWarning: Parsing '29-03-2022' in DD/MM/YYYY
format. Provide format or specify infer_datetime_format=True for consistent parsing.
  cache_array = _maybe_cache(arg, format, cache, convert_listlike)
C:\Users\juliy\anaconda3\lib\site-packages\pandas\core\tools\datetimes.py:1047: UserWarning: Parsing '27-03-2022' in DD/MM/YYYY
format. Provide format or specify infer_datetime_format=True for consistent parsing.
  cache_array = _maybe_cache(arg, format, cache, convert_listlike)
C:\Users\juliy\anaconda3\lib\site-packages\pandas\core\tools\datetimes.py:1047: UserWarning: Parsing '15-02-2022' in DD/MM/YYYY
format. Provide format or specify infer_datetime_format=True for consistent parsing.
  cache_array = _maybe_cache(arg, format, cache, convert_listlike)
C:\Users\juliy\anaconda3\lib\site-packages\pandas\core\tools\datetimes.py:1047: UserWarning: Parsing '13-02-2022' in DD/MM/YYYY
format. Provide format or specify infer_datetime_format=True for consistent parsing.
  cache_array = _maybe_cache(arg, format, cache, convert_listlike)
C:\Users\juliy\anaconda3\lib\site-packages\pandas\core\tools\datetimes.py:1047: UserWarning: Parsing '17-02-2022' in DD/MM/YYYY
format. Provide format or specify infer_datetime_format=True for consistent parsing.
  cache_array = _maybe_cache(arg, format, cache, convert_listlike)
C:\Users\juliy\anaconda3\lib\site-packages\pandas\core\tools\datetimes.py:1047: UserWarning: Parsing '16-03-2022' in DD/MM/YYYY
format. Provide format or specify infer_datetime_format=True for consistent parsing.
  cache_array = _maybe_cache(arg, format, cache, convert_listlike)
```

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

Markdown ▾

```python
In [9]: import math
        from math import radians, sin, cos, acos

        def distance(restaurant, destination):

            lat1, long1 = restaurant
            lat2, long2 = destination
            radius = 6371
            # distance in kilometers

            dlat = math.radians(lat2 - lat1)
            dlong = math.radians(long2 - long1)

            a = (math.sin(dlat / 2) * math.sin(dlat / 2) + math.cos(math.radians(lat1)) * math.cos(math.radians(lat2)) * math.sin(dlong /
            c = 2 * math.atan2(math.sqrt(a), math.sqrt(1 - a))
            d = radius * c

            return d
```

```python
In [10]: food['distance_km'] = food.apply(lambda x: distance((x['Restaurant_latitude'], x['Restaurant_longitude']), (x['Delivery_location_
```

```python
In [11]: food.head()
```

Out[11]:

	ID	Delivery_person_ID	Delivery_person_Age	Delivery_person_Ratings	Restaurant_latitude	Restaurant_longitude	Delivery_location_latitude	Delivery_loca
0	0x2318	COIMBRES13DEL01	NaN	NaN	11.003669	76.976494	11.043669	

File   Edit   View   Insert   Cell   Kernel   Widgets   Help   Trusted | Python 3 (ipykernel) O
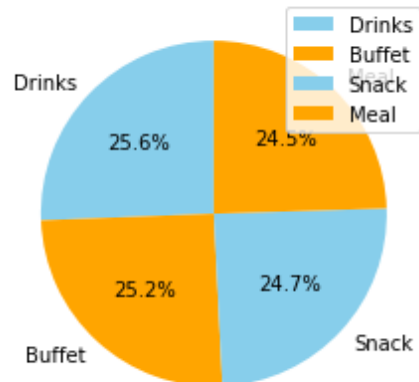
```
Meal        2794
Name: Type_of_order, dtype: int64
```

In [20]:
```python
Order_type=food['Type_of_order'].isnull().sum()
Order_type
```

Out[20]: 0

In [21]:
```python
plt.pie(food['Type_of_order'].value_counts().head(10),
        labels=food['Type_of_order'].value_counts().head().index,autopct='%1.1f%%',
        colors=['skyblue', 'orange'],startangle=90


       );
plt.legend();
```
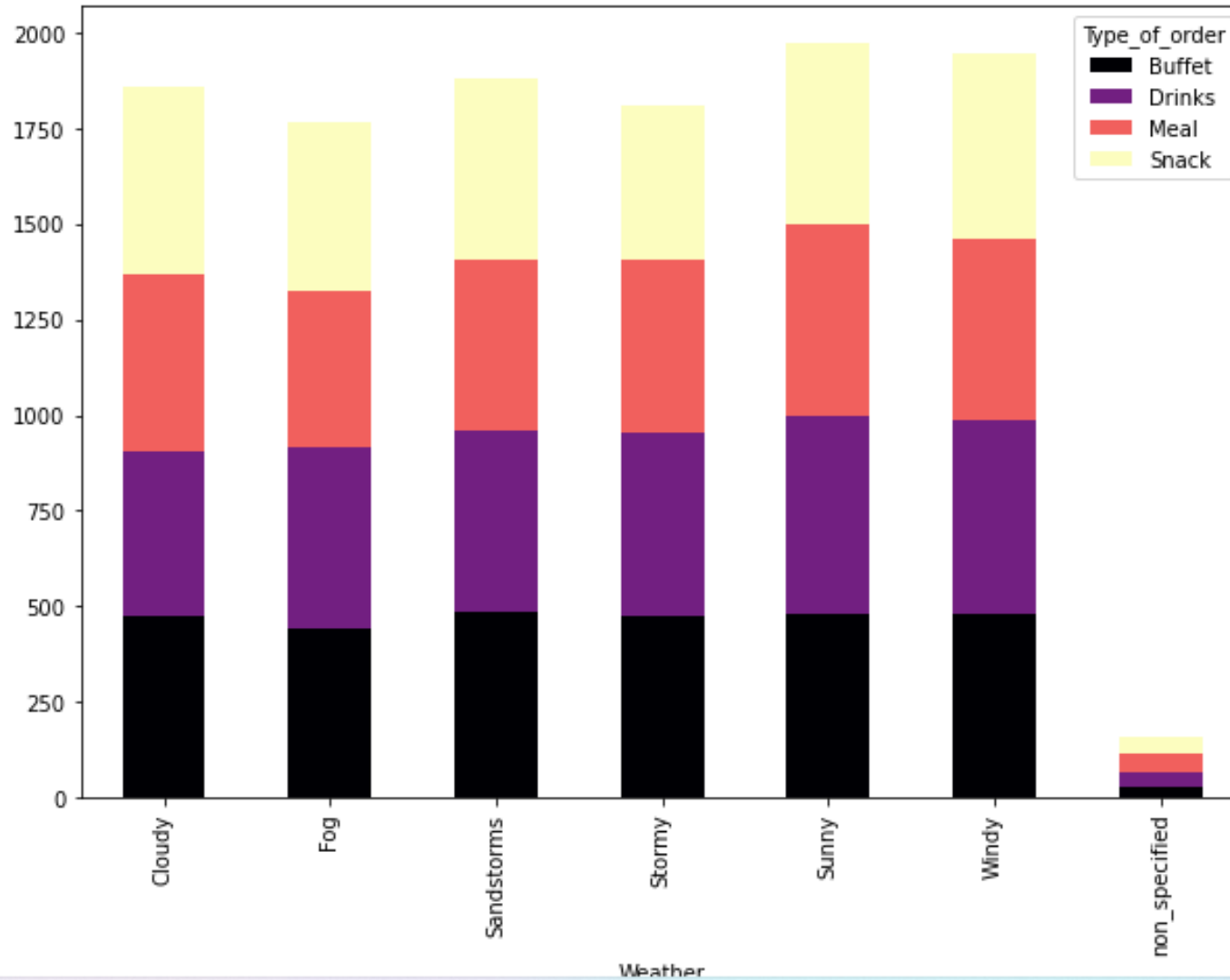
AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

```
In [93]: Order_Weather.plot(kind='bar',stacked=True,figsize=(10,7),cmap=cmap,edgecolor='None');
```
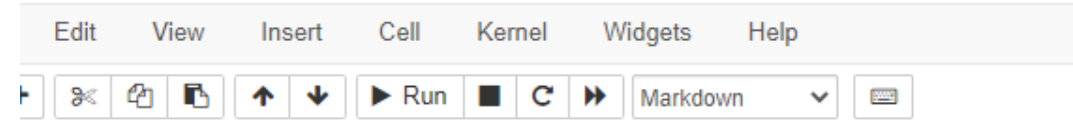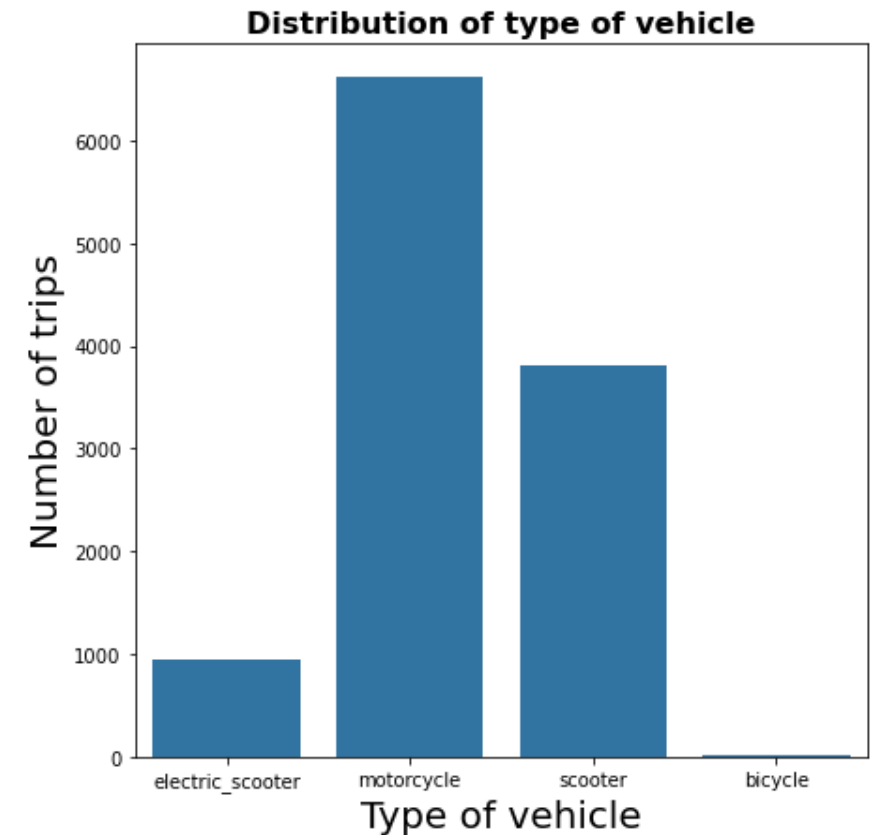
AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

What meaans of transportation delievers food faster?

```
In [28]: plt.figure(figsize = [7, 7])
         color = sn.color_palette()[0]
         sn.countplot(data = food, x = 'Type_of_vehicle', color = color)
         plt.title('Distribution of type of vehicle',y=1.0, fontsize=16, fontweight='bold')
         plt.xlabel('Type of vehicle',fontsize=20)
         plt.ylabel('Number of trips',fontsize=20);
```

| Edit | View | Insert | Cell | Kernel | Widgets | Help |

```
plt.ylabel('Number of trips',fontsize=20);
```
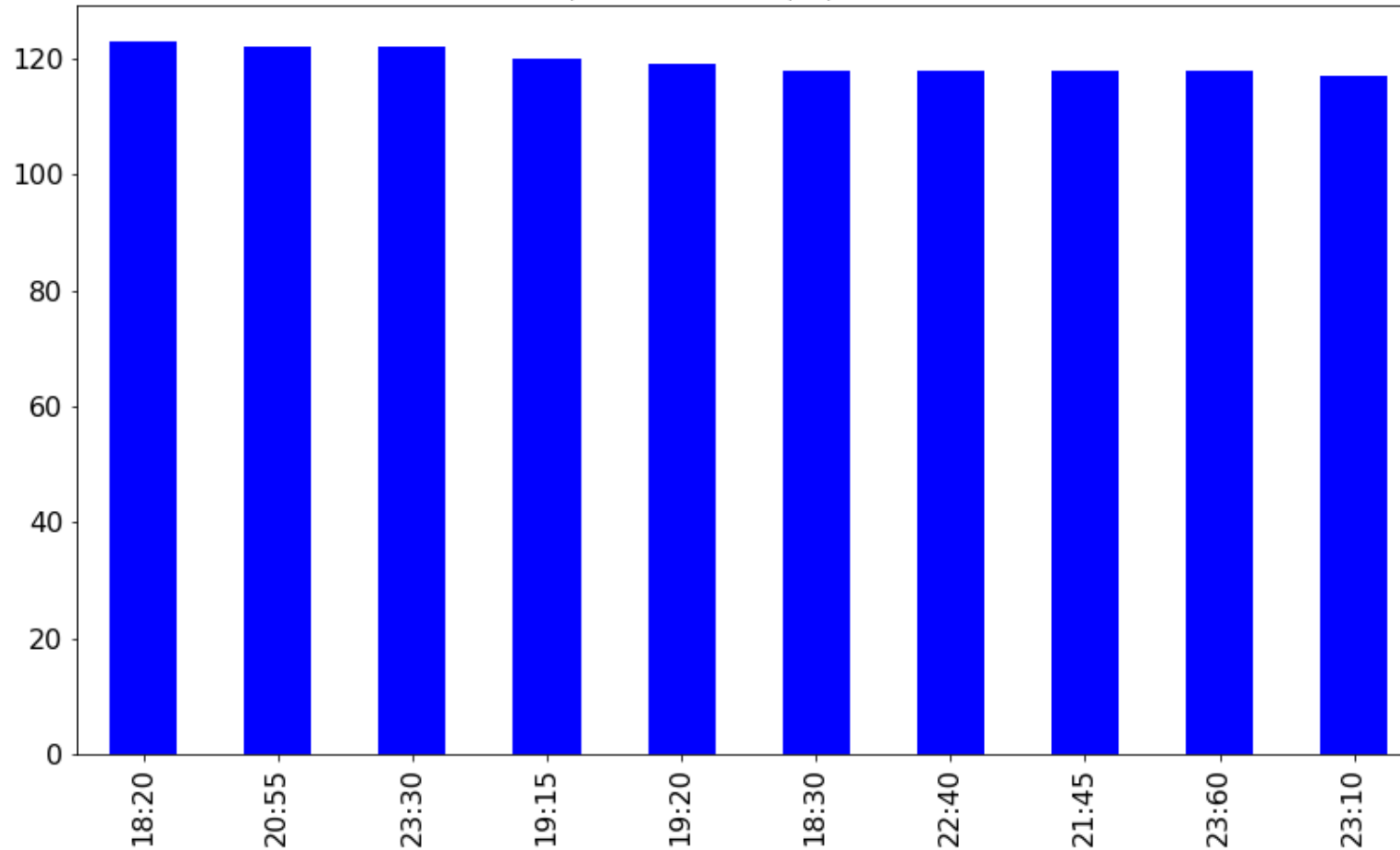

Distribution of type of vehicle

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

At what time do people make most orders?

Barchart Representation of time people make their orders

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

What day of the week take longer distance to delivery the food

```
In [16]: plt.figure(figsize = [10, 8])
         orderby = ['Monday','Tuesday','Wednesday','Thursday','Friday','Saturday','Sunday'];
         g = sn.FacetGrid(data = food, col = 'Days',col_wrap = 4,col_order =orderby);
         plt.title("Restaurant days distribution", y=1.0, fontsize=20, fontweight='bold')
         g.map(plt.hist, "distance_km");
```
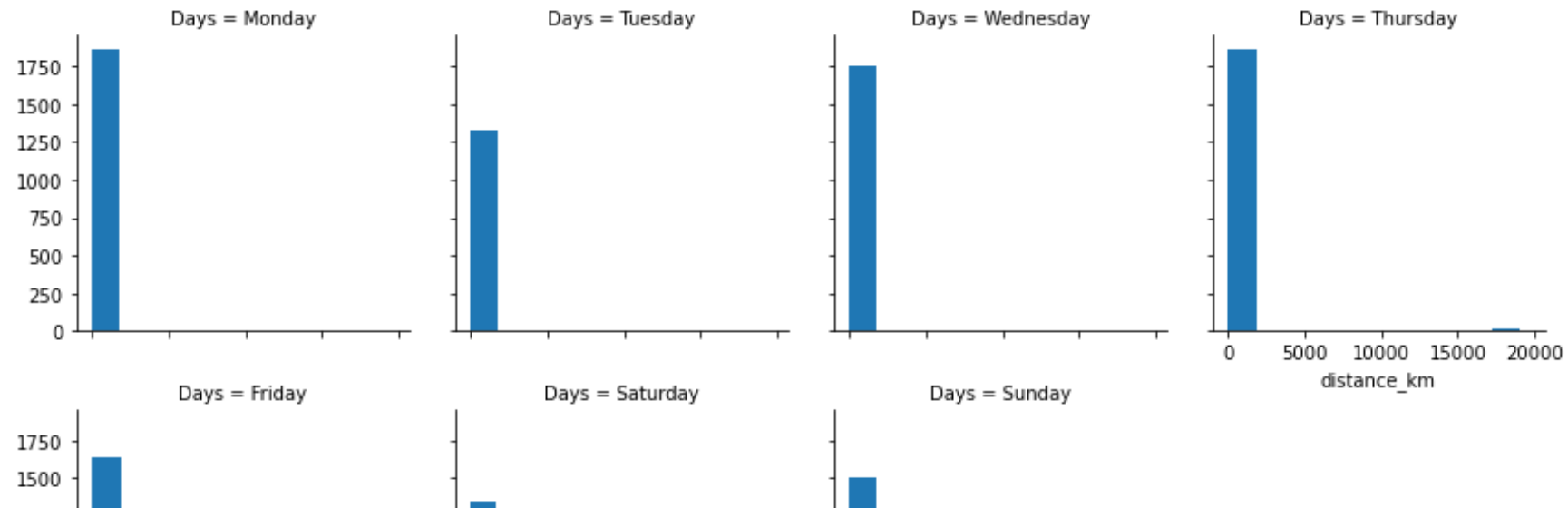
<Figure size 720x576 with 0 Axes>

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

▶ Run   ■   C   ▶▶   Markdown ⌄   ⌨

od and the distance covered?

```
plt.ylabel(' distance_km',fontsize=20);
```

In [54
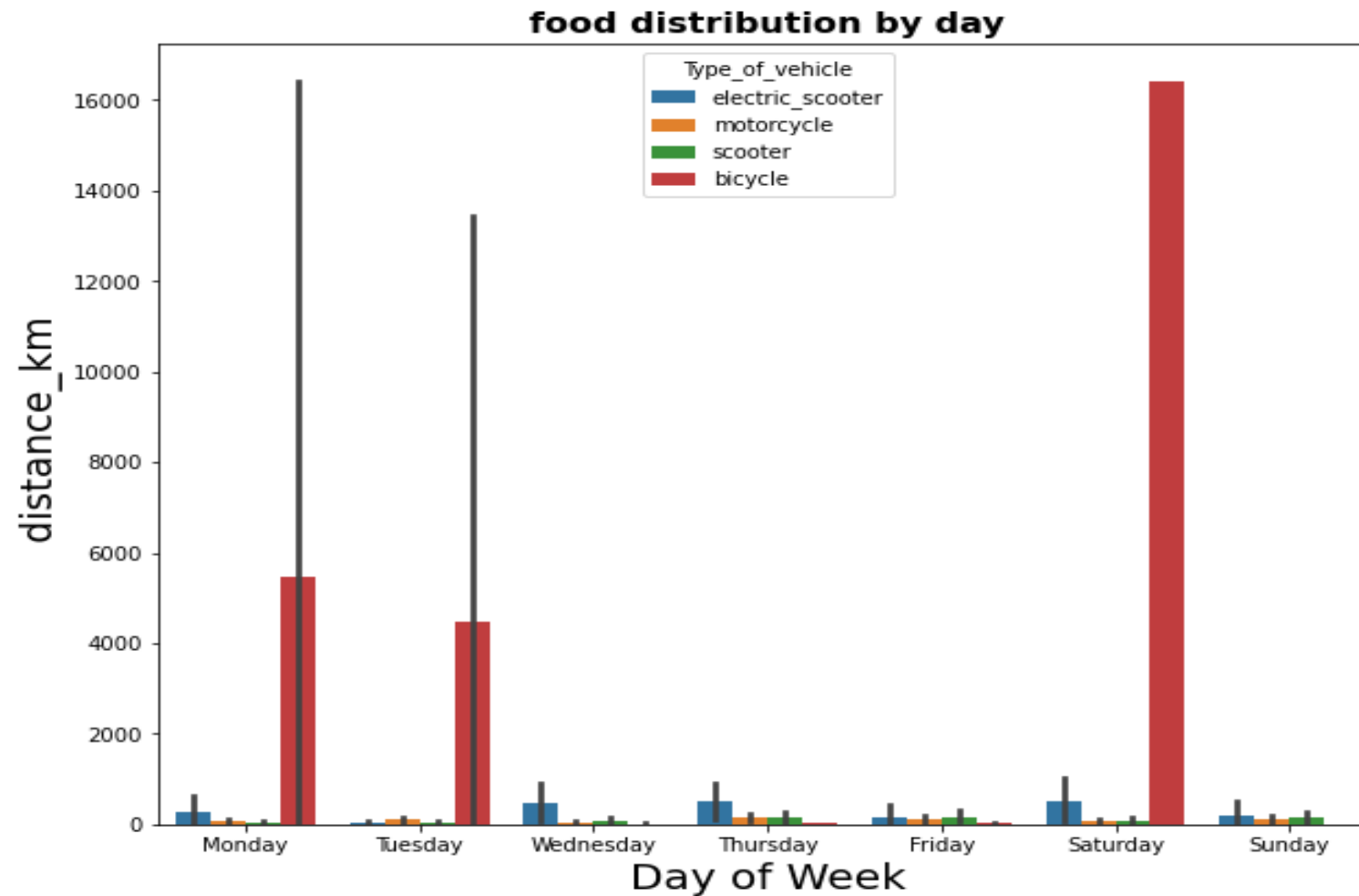
= orderby);
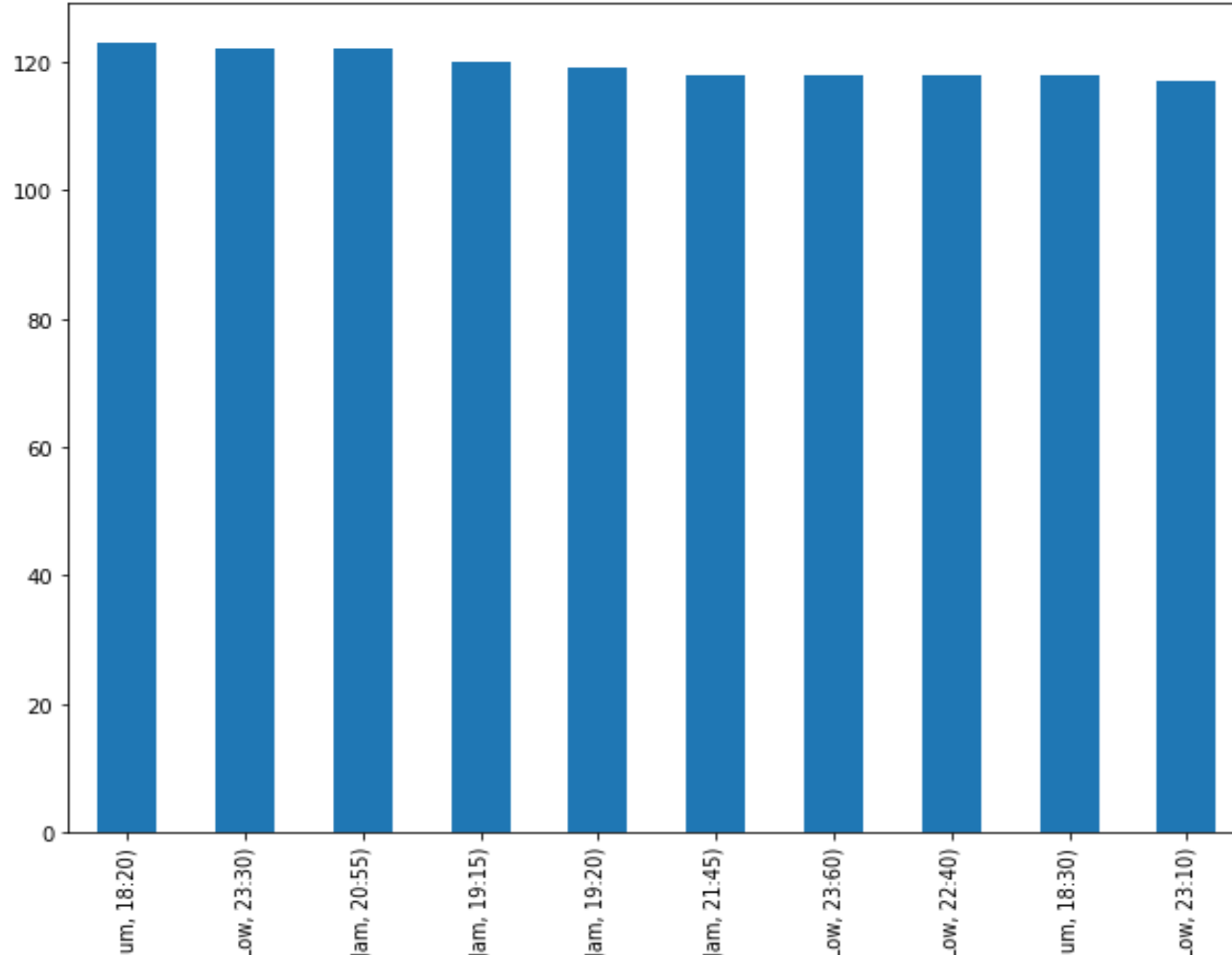

From the diagram above, it shows that the restaurant utilized bicycles the most to cover longer distance on Saturdays

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant


```
[31]: State_traffic_density.plot(x='Road_traffic_density',y='Time_Orderd',kind='bar',figsize=(10,8));
```
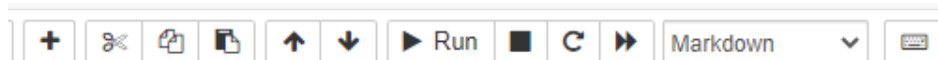
Most of the orders are made after 17h where the traffic density is not high

`.head(10)`

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

| + | ✂ | ⎘ | 📋 | ↑ | ↓ | ▶ Run | ■ | C | ⏭ | Markdown ⌄ | ⌨ |

Machine learning scenario

In [37]: `food.head()`

Out[37]:

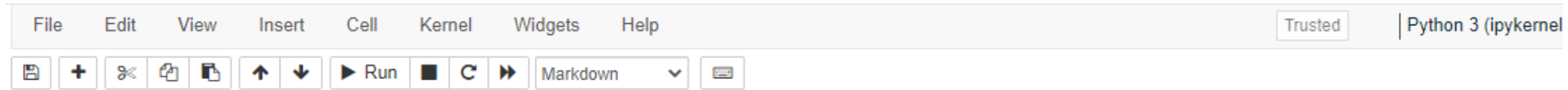|	ID	Delivery_person_ID	Delivery_person_Age	Delivery_person_Ratings	Restaurant_latitude	Restaurant_longitude	Delivery_location_latitude	Delivery_loca
0	0x2318	COIMBRES13DEL01	NaN	NaN	11.003669	76.976494	11.043669	
1	0x3474	BANGRES15DEL01	28.0	4.6	12.975377	77.696664	13.085377	
2	0x9420	JAPRES09DEL03	23.0	4.5	26.911378	75.789034	27.001378	
3	0x72ee	JAPRES07DEL03	21.0	4.8	26.766536	75.837333	26.856536	
4	0xa759	CHENRES19DEL01	31.0	4.6	12.986047	80.218114	13.096047	

5 rows × 23 columns

In [38]: `food_delivery=food.drop(['ID','Delivery_person_ID','Delivery_person_Age','Delivery_person_Ratings','Weather','Road_traffic_densit`
`food_delivery.head(10)`

Out[38]:

	Type_of_order	Type_of_vehicle	Time_taken	Days	distance_km
0	Drinks	electric_scooter	24.0	Wednesday	6.232373
1	Snack	motorcycle	33.0	Tuesday	17.076617
2	Drinks	motorcycle	42.0	Monday	13.406029
3	Meal	scooter	19.0	Friday	13.413644

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant

| File | Edit | View | Insert | Cell | Kernel | Widgets | Help | | Trusted | | Python 3 (ipykernel |

Markdown ▾

Converting some columns from words to digit using labelEncoder and fit transform

```
In [39]:  from sklearn.preprocessing import LabelEncoder

          le_Type_of_order=LabelEncoder()
          le_Type_of_vehicle=LabelEncoder()
          le_Days=LabelEncoder()
```

```
In [40]:  food_delivery['Type_of_order']=le_Type_of_order.fit_transform(food_delivery['Type_of_order'])
          food_delivery['Type_of_vehicle']=le_Type_of_vehicle.fit_transform(food_delivery['Type_of_vehicle'])
          food_delivery['Days']=le_Days.fit_transform(food_delivery['Days'])
```

```
In [41]:  food_delivery
```

Out[41]:

	Type_of_order	Type_of_vehicle	Time_taken	Days	distance_km
0	1	1	24.0	6	6.232373
1	3	2	33.0	5	17.076617
2	1	2	42.0	1	13.406029
3	2	3	19.0	0	13.413644
4	1	3	25.0	3	17.076259
...	...	...	...	...	...
11394	3	3	NaN	3	16.384901
11395	1	3	NaN	4	1.489842

AMAZON BUSINESS RESEARCH ANALYST DATASET
Case in Study; Restaurant


```
In [48]: from sklearn import tree
         model=tree.DecisionTreeClassifier()
```

```
In [49]: model.fit(inputs,target)
```

Out[49]: DecisionTreeClassifier()

Mrs Ndidi ordered meals at the amazon restaurant on Thursday morning. Predict how long before she received her order if the delivery man used motorcycle bearing in mind that the customer stays 9.760928 km away from the restaurant.

```
In [50]: model.predict([[2,4,5,9.760928 ]])
```

C:\Users\juliy\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning: X does not have valid feature names, but DecisionTreeClassifier was fitted with feature names
    warnings.warn(

Out[50]: array([26.])

According to our analysis, it indicates that the customer will receive her order in 26 minutes

```
In [51]: model.score(inputs,target)
```

Out[51]: 0.9840336871655408

Our model shows an accuracy score of 98%

using linear regression model

```
In [62]: from sklearn import linear_model
```

```
In [63]: reg = linear_model.LinearRegression()
         reg.fit(inputs, target)
```

```
Out[63]: LinearRegression()
```

```
In [64]: reg.predict([[2,4,5,9.760928]])
```

```
C:\Users\juliy\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning: X does not have valid feature names, but LinearReg
ression was fitted with feature names
  warnings.warn(
```

```
Out[64]: array([26.43485171])
```

Validating Prediction

```
In [65]: reg.coef_
```

```
Out[65]: array([-5.84737181e-02,  1.92261728e-01,  1.01874228e-02, -2.88044409e-05])
```

```
In [66]: reg.intercept_
```

```
Out[66]: 25.732096278581974
```

AMAZON BUSINESS RESEARCH ANALYST DATASET
## Case in Study; Restaurant

Recommendation;

Having observed that the peak periods where orders are made, were between the hours of 18:20- 0:00 , we recommend that the restaurant management reduce the number of employees who work during daytime and the amount of food prepared during the daytime so as to cut down on expenses, and capitalize on profit making.

GROUP_5_MEMBERS

"The only way to do great work is to love what you do. If you haven't found it, keep looking. Don't settle."

- Steve Jobs
Founder of Apple Computer
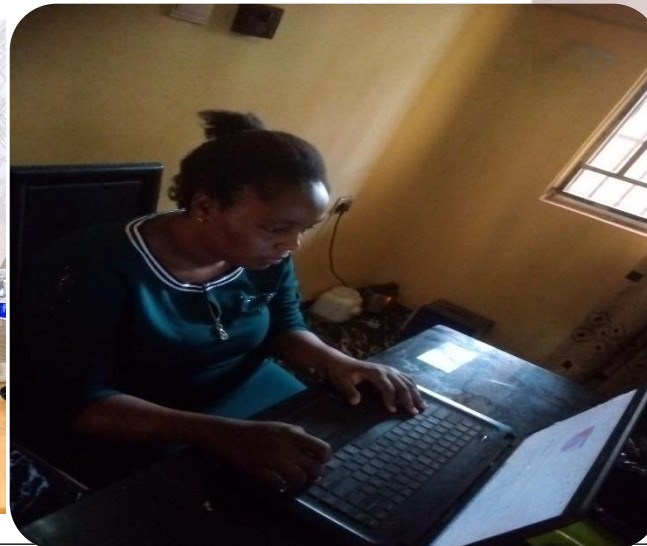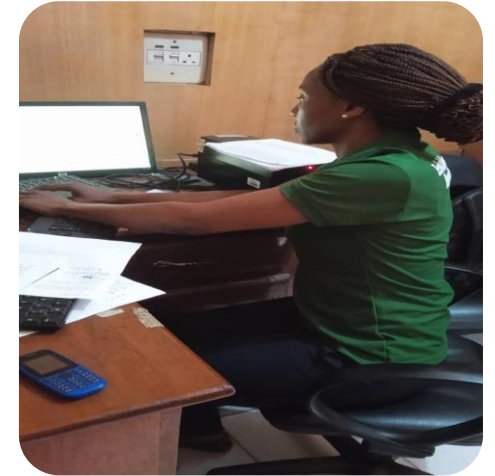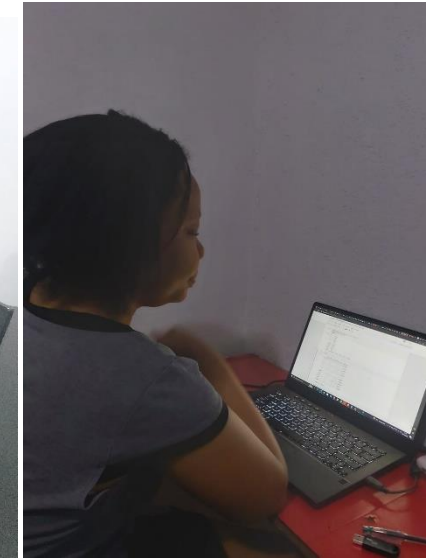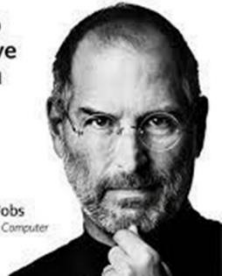
Contact Us

Instagram:Tech4Dev
Facebook:Tech4Dev
Twitter:Tech4Dev
Email:Info@tech4dev.com

Acknowledgement

KOSISOCHUKWU ASHARA

Thank You!