# *UBER DATA ANALYSIS*

BY

- SAKSHI MOGHA
- ANJALI NIMJE
- RIYA SONI
- RISHABH SHARMA

In [1]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import os
import seaborn as sns
import time,datetime
from matplotlib.pyplot import figure
```

In [2]:

```
df = pd.read_csv("C:\\Users\\Sakshi\\Downloads\\My Uber Drives.csv")
```

```
df
```

```
In [3]:
Out[3]:
```

| | START_DATE* | END_DATE* | CATEGORY* | START* | STOP* | MILES* | PURPOSE* |
|---|---|---|---|---|---|---|---|
| **0** | 01-01-2016 21:11 | 01-01-2016 21:17 | Business | Fort Pierce | Fort Pierce | 5.1 | Meal/Entertain |
| **1** | 01-02-2016 01:25 | 01-02-2016 01:37 | Business | Fort Pierce | Fort Pierce | 5.0 | NaN |
| **2** | 01-02-2016 20:25 | 01-02-2016 20:38 | Business | Fort Pierce | Fort Pierce | 4.8 | Errand/Supplies |
| **3** | 01-05-2016 17:31 | 01-05-2016 17:45 | Business | Fort Pierce | Fort Pierce | 4.7 | Meeting |
| **4** | 01-06-2016 14:42 | 01-06-2016 15:49 | Business | Fort Pierce | West Palm Beach | 63.7 | Customer Visit |
| **...** | ... | ... | ... | ... | ... | ... | ... |
| **1151** | 12/31/2016 13:24 | 12/31/2016 13:42 | Business | Kar?chi | Unknown Location | 3.9 | Temporary Site |
| **1152** | 12/31/2016 15:03 | 12/31/2016 15:38 | Business | Unknown Location | Unknown Location | 16.2 | Meeting |
| **1153** | 12/31/2016 21:32 | 12/31/2016 21:50 | Business | Katunayake | Gampaha | 6.4 | Temporary Site |
| **1154** | 12/31/2016 22:08 | 12/31/2016 23:51 | Business | Gampaha | Ilukwatta | 48.2 | Temporary Site |

```
In [4]:
```

|  |  |  |  |  |  |  |  |
|---|---|---|---|---|---|---|---|
| **1155** | Totals | NaN | NaN | NaN | NaN | 12204.7 | NaN |

**1156** rows × 7 columns

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1156 entries, 0 to 1155
Data columns (total 7 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
0   START_DATE*  1156 non-null   object
1   END_DATE*    1155 non-null   object
2   CATEGORY*    1155 non-null   object
3   START*       1155 non-null   object
4   STOP*        1155 non-null   object
5   MILES*       1156 non-null   float64
6   PURPOSE*     653 non-null    object
dtypes: float64(1), object(6)
memory usage: 63.3+ KB
```

```
In [5]:  df.isnull().sum()
```

```
Out[5]: START_DATE*      0
        END_DATE*        1
        CATEGORY*        1
        START*           1
        STOP*            1
        MILES*           0

        PURPOSE*       503

        dtype: int64 df =
In [6]:
        df.drop(1155)
```

```
In [7]:
         df
```

Out[7]:

|  | START_DATE* | END_DATE* | CATEGORY* | START* | STOP* | MILES* | PURPOSE* |
|---|---|---|---|---|---|---|---|
| **0** | 01-01-2016 21:11 | 01-01-2016 21:17 | Business | Fort Pierce | Fort Pierce | 5.1 | Meal/Entertain |
| **1** | 01-02-2016 01:25 | 01-02-2016 01:37 | Business | Fort Pierce | Fort Pierce | 5.0 | NaN |
| **2** | 01-02-2016 20:25 | 01-02-2016 20:38 | Business | Fort Pierce | Fort Pierce | 4.8 | Errand/Supplies |
| **3** | 01-05-2016 17:31 | 01-05-2016 17:45 | Business | Fort Pierce | Fort Pierce | 4.7 | Meeting |
| **4** | 01-06-2016 14:42 | 01-06-2016 15:49 | Business | Fort Pierce | West Palm Beach | 63.7 | Customer Visit |
| **...** | ... | ... | ... | ... | ... | ... | ... |
| **1150** | 12/31/2016 1:07 | 12/31/2016 1:14 | Business | Kar?chi | Kar?chi | 0.7 | Meeting |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **1151** | 12/31/2016 13:24 | 12/31/2016 13:42 | Business | Kar?chi | Unknown Location | 3.9 | Temporary Site |
| **1152** | 12/31/2016 15:03 | 12/31/2016 15:38 | Business | Unknown Location | Unknown Location | 16.2 | Meeting |
| **1153** | 12/31/2016 21:32 | 12/31/2016 21:50 | Business | Katunayake | Gampaha | 6.4 | Temporary Site |
| **1154** | 12/31/2016 22:08 | 12/31/2016 23:51 | Business | Gampaha | Ilukwatta | 48.2 | Temporary Site |

1155 rows × 7 columns

In [8]:

```
#Now lets rename the columns first. To make my code easy to understand i am
renaming df = df.rename(columns = {df.columns[0]:'startdate',
df.columns[1]:'enddate', df.col
```

In [9]: df

Out[9]:

| | startdate | enddate | category | start | stop | miles | purpose |
|---|---|---|---|---|---|---|---|
| **0** | 01-01-2016 21:11 | 01-01-2016 21:17 | Business | Fort Pierce | Fort Pierce | 5.1 | Meal/Entertain |
| **1** | 01-02-2016 01:25 | 01-02-2016 01:37 | Business | Fort Pierce | Fort Pierce | 5.0 | NaN |
| **2** | 01-02-2016 20:25 | 01-02-2016 20:38 | Business | Fort Pierce | Fort Pierce | 4.8 | Errand/Supplies |
| **3** | 01-05-2016 17:31 | 01-05-2016 17:45 | Business | Fort Pierce | Fort Pierce | 4.7 | Meeting |
| **4** | 01-06-2016 14:42 | 01-06-2016 15:49 | Business | Fort Pierce | West Palm Beach | 63.7 | Customer Visit |
| **...** | ... | ... | ... | ... | ... | ... | ... |
| **1150** | 12/31/2016 1:07 | 12/31/2016 1:14 | Business | Kar?chi | Kar?chi | 0.7 | Meeting |
| **1151** | 12/31/2016 13:24 | 12/31/2016 13:42 | Business | Kar?chi | Unknown Location | 3.9 | Temporary Site |
| **1152** | 12/31/2016 15:03 | 12/31/2016 15:38 | Business | Unknown Location | Unknown Location | 16.2 | Meeting |
| **1153** | 12/31/2016 21:32 | 12/31/2016 21:50 | Business | Katunayake | Gampaha | 6.4 | Temporary Site |
| **1154** | 12/31/2016 22:08 | 12/31/2016 23:51 | Business | Gampaha | Ilukwatta | 48.2 | Temporary Site |

1155 rows × 7 columns

**startdate**

```
df
```

**enddate category start stop miles purpose**

In [10]:

Out[10]:

| | startdate | enddate | category | start | stop | miles | purpose |
|---|---|---|---|---|---|---|---|
| **0** | 01-01-2016 21:11 | 01-01-2016 21:17 | Business | Fort Pierce | Fort Pierce | 5.1 | Meal/Entertain |
| **1** | 01-02-2016 01:25 | 01-02-2016 01:37 | Business | Fort Pierce | Fort Pierce | 5.0 | NaN |
| **2** | 01-02-2016 20:25 | 01-02-2016 20:38 | Business | Fort Pierce | Fort Pierce | 4.8 | Errand/Supplies |
| **3** | 01-05-2016 17:31 | 01-05-2016 17:45 | Business | Fort Pierce | Fort Pierce | 4.7 | Meeting |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **4** | 01-06-2016 14:42 | 01-06-2016 15:49 | Business | Fort Pierce | West Palm Beach | 63.7 | Customer Visit |
| **...** | ... | ... | ... | ... | ... | ... | ... |
| **1150** | 12/31/2016 1:07 | 12/31/2016 1:14 | Business | Kar?chi | Kar?chi | 0.7 | Meeting |
| **1151** | 12/31/2016 13:24 | 12/31/2016 13:42 | Business | Kar?chi | Unknown Location | 3.9 | Temporary Site |
| **1152** | 12/31/2016 15:03 | 12/31/2016 15:38 | Business | Unknown Location | Unknown Location | 16.2 | Meeting |
| | **startdate** | **enddate** | **category** | **start** | **stop** | **miles** | **purpose** |
| **1153** | 12/31/2016 21:32 | 12/31/2016 21:50 | Business | Katunayake | Gampaha | 6.4 | Temporary Site |
| **1154** | 12/31/2016 22:08 | 12/31/2016 23:51 | Business | Gampaha | Ilukwatta | 48.2 | Temporary Site |

1155 rows × 7 columns

```
df.isnull().sum()
```

```
startdate      0
enddate        0
category       0
start          0
stop           0
miles          0
purpose      502
dtype: int64
```

```
df.category.value_counts()
```

In [11]:

Out[11]:

```
In [12]:


Out[12]: Business    1078
In [13]: Personal      77
         Name: category, dtype: int64

          print("\n..........Average Length of the Trip.............\n")
          print('Business:', round(df[df['category'] == 'Business'].miles.mean(), 3))
          print('Personal:', round(df[df['category'] == 'Personal'].miles.mean(), 3))
          print('Meal/Entertain:', round(df[df['purpose'] ==
          'Meal/Entertain'].miles.mean(), 3

         ..........Average Length of the Trip.............

         Business: 10.656
         Personal: 9.321
         Meal/Entertain: 5.698
          #Now to change the data types of start date and end
          date df.startdate = pd.to_datetime(df.startdate)
          df.enddate = pd.to_datetime(df.enddate)

In [14]:
In [15]:
          df.info()


         <class 'pandas.core.frame.DataFrame'>
         Int64Index: 1155 entries, 0 to 1154
         Data columns (total 7 columns):
          #   Column     Non-Null Count  Dtype
         ---  ------     --------------  -----
         0   startdate  1155 non-null   datetime64[ns]
         1   enddate    1155 non-null   datetime64[ns]
         2   category   1155 non-null   object
         3   start      1155 non-null   object
         4   stop       1155 non-null   object
         5   miles      1155 non-null   float64         6   purpose     653 non-null
             object         dtypes: datetime64[ns](2), float64(1), object(4) memory usage:
             72.2+ KB
In [16]:
```

```
df[df.start.str.contains('\?') == True]
```

Out[16]:

| | startdate | enddate | category | start | stop | miles | purpose |
|---|---|---|---|---|---|---|---|
| **140** | 2016-02-20 14:50:00 | 2016-02-20 15:54:00 | Business | R? walpindi | R?walpindi | 23.1 | Meeting |
| **141** | 2016-02-20 16:59:00 | 2016-02-20 17:54:00 | Personal | R? walpindi | Unknown Location | 16.5 | NaN |
| **656** | 2016-08-15 09:05:00 | 2016-08-15 09:52:00 | Business | R? walpindi | Unknown Location | 15.6 | NaN |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **670** | 2016-08-17 15:32:00 | 2016-08-17 15:47:00 | Business | R?walpindi | Islamabad | 6.4 | NaN |
| **679** | 2016-08-19 10:57:00 | 2016-08-19 11:06:00 | Business | R?walpindi | Unknown Location | 2.0 | NaN |
| **696** | 2016-08-22 20:53:00 | 2016-08-22 21:31:00 | Business | R?walpindi | R?walpindi | 4.1 | NaN |
| **697** | 2016-08-22 22:31:00 | 2016-08-22 23:00:00 | Business | R?walpindi | Unknown Location | 18.7 | NaN |
| **763** | 2016-09-19 06:18:00 | 2016-09-19 06:49:00 | Business | R?walpindi | Unknown Location | 18.2 | NaN |
| **788** | 2016-10-06 17:23:00 | 2016-10-06 17:40:00 | Business | R?walpindi | Unknown Location | 112.6 | NaN |
| **804** | 2016-10-12 19:18:00 | 2016-10-12 19:21:00 | Business | R?walpindi | Unknown Location | 18.4 | NaN |
| **811** | 2016-10-14 10:16:00 | 2016-10-14 10:52:00 | Business | R?walpindi | Unknown Location | 12.4 | NaN |
| **1119** | 2016-12-27 07:02:00 | 2016-12-27 07:14:00 | Business | Kar?chi | Kar?chi | 4.9 | Temporary Site |
| **1120** | 2016-12-27 08:37:00 | 2016-12-27 08:59:00 | Business | Kar?chi | Kar?chi | 5.0 | Meal/Entertain |
| **1121** | 2016-12-27 12:53:00 | 2016-12-27 12:57:00 | Business | Kar?chi | Kar?chi | 0.6 | Meal/Entertain |
| **1122** | 2016-12-27 14:49:00 | 2016-12-27 15:03:00 | Business | Kar?chi | Unknown Location | 3.1 | Customer Visit |
| **1124** | 2016-12-27 19:19:00 | 2016-12-27 19:50:00 | Business | Kar?chi | Kar?chi | 5.5 | Customer Visit |
| **1125** | 2016-12-28 08:34:00 | 2016-12-28 09:06:00 | Business | Kar?chi | Unknown Location | 10.3 | Meal/Entertain |
| **1127** | 2016-12-28 13:53:00 | 2016-12-28 14:01:00 | Business | Kar?chi | Kar?chi | 2.0 | Errand/Supplies |
| **1128** | 2016-12-28 15:04:00 | 2016-12-28 15:39:00 | Business | Kar?chi | Unknown Location | 8.5 | Meal/Entertain |
| **1130** | 2016-12-28 18:33:00 | 2016-12-28 18:56:00 | Business | Kar?chi | Kar?chi | 3.8 | Errand/Supplies |
| **1131** | 2016-12-28 22:44:00 | 2016-12-28 23:18:00 | Business | Kar?chi | Kar?chi | 5.1 | Errand/Supplies |
| **1132** | 2016-12-29 00:49:00 | 2016-12-29 01:06:00 | Business | Kar?chi | Kar?chi | 3.8 | Errand/Supplies |
| **1133** | 2016-12-29 09:44:00 | 2016-12-29 10:07:00 | Business | Kar?chi | Unknown Location | 11.6 | Meal/Entertain |
| **1135** | 2016-12-29 12:25:00 | 2016-12-29 12:33:00 | Business | Kar?chi | Kar?chi | 1.4 | Errand/Supplies |

| startdate | | | | | | | |
|---|---|---|---|---|---|---|---|
| **1136** | 2016-12-29 13:17:00 | 2016-12-29 13:24:00 | Business | Kar?chi | Kar?chi | 1.1 | Errand/Supplies |
| **1137** | 2016-12-29 13:56:00 | 2016-12-29 14:11:00 | Business | Kar?chi | Kar?chi | 4.1 | Airport/Travel |
| **1138** | 2016-12-29 14:42:00 | 2016-12-29 14:58:00 | Business | Kar?chi | Kar?chi | 6.1 | Between Offices |
| **1139** | 2016-12-29 15:05:00 | 2016-12-29 15:16:00 | Business | Kar?chi | Kar?chi | 1.3 | Errand/Supplies |
| **1140** | 2016-12-29 18:59:00 | 2016-12-29 19:14:00 | Business | Kar?chi | Unknown Location | 3.0 | Meal/Entertain |
| **1142** | 2016-12-29 20:15:00 | 2016-12-29 20:45:00 | Business | Kar?chi | Kar?chi | 7.2 | Meeting |
| **1143** | 2016-12-29 20:53:00 | 2016-12-29 21:42:00 | Business | Kar?chi | Unknown Location | 6.4 | NaN |
| **1145** | 2016-12-30 10:15:00 | 2016-12-30 10:33:00 | Business | Kar?chi | Kar?chi | 2.8 | Errand/Supplies |
| **1146** | 2016-12-30 11:31:00 | 2016-12-30 11:56:00 | Business | Kar?chi | Kar?chi | 2.9 | Errand/Supplies |
| **1147** | 2016-12-30 15:41:00 | 2016-12-30 16:03:00 | Business | Kar?chi | Kar?chi | 4.6 | Errand/Supplies |
| **1148** | 2016-12-30 16:45:00 | 2016-12-30 17:08:00 | Business | Kar?chi | Kar?chi | 4.6 | Meeting |
| **1149** | 2016-12-30 23:06:00 | 2016-12-30 23:10:00 | Business | Kar?chi | Kar?chi | 0.8 | Customer Visit |
| **1150** | 2016-12-31 01:07:00 | 2016-12-31 01:14:00 | Business | Kar?chi | Kar?chi | 0.7 | Meeting |
| **1151** | 2016-12-31 13:24:00 | 2016-12-31 13:42:00 | Business | Kar?chi | Unknown Location | 3.9 | Temporary Site |

```
df['start']   =   df['start'].replace({"\?":"a"}, = True)
regex
df["stop"] = df["stop"].replace({"\?":"a"}, regex True)
=
```

```
df
```

| | enddate | category | start | stop | miles | purpose |
|---|---|---|---|---|---|---|

```
In [17]:
```

```
In [18]:
```

Out[18]:

| | startdate | enddate | category | start | stop | miles | purpose |
|---|---|---|---|---|---|---|---|
| **0** | 2016-01-01 21:11:00 | 2016-01-01 21:17:00 | Business | Fort Pierce | Fort Pierce | 5.1 | Meal/Entertain |
| **1** | 2016-01-02 01:25:00 | 2016-01-02 01:37:00 | Business | Fort Pierce | Fort Pierce | 5.0 | NaN |
| **2** | 2016-01-02 20:25:00 | 2016-01-02 20:38:00 | Business | Fort Pierce | Fort Pierce | 4.8 | Errand/Supplies |
| **3** | 2016-01-05 17:31:00 | 2016-01-05 17:45:00 | Business | Fort Pierce | Fort Pierce | 4.7 | Meeting |
| **4** | 2016-01-06 14:42:00 | 2016-01-06 15:49:00 | Business | Fort Pierce | West Palm Beach | 63.7 | Customer Visit |
| **...** | ... | ... | ... | ... | ... | ... | ... |
| **1150** | 2016-12-31 01:07:00 | 2016-12-31 01:14:00 | Business | Karachi | Karachi | 0.7 | Meeting |
| **1151** | 2016-12-31 13:24:00 | 2016-12-31 13:42:00 | Business | Karachi | Unknown Location | 3.9 | Temporary Site |
| **1152** | 2016-12-31 15:03:00 | 2016-12-31 15:38:00 | Business | Unknown Location | Unknown Location | 16.2 | Meeting |
| **1153** | 2016-12-31 21:32:00 | 2016-12-31 21:50:00 | Business | Katunayake | Gampaha | 6.4 | Temporary Site |
| **1154** | 2016-12-31 22:08:00 | 2016-12-31 23:51:00 | Business | Gampaha | Ilukwatta | 48.2 | Temporary Site |

1155 rows × 7 columns

```
In [19]:
```

```
print("Popular                          Starting                           Points
\n..................\n",df.start.sort_values().value_
```

```
Popular Starting Points
..................
 Cary                          201
Unknown Location              148
Morrisville                    85
Whitebridge                    68
Islamabad                      57
...
Tribeca                         1
Mcvan                           1
NOMA                            1
Eastgate                        1
Townes at Everett Crossing      1
Name: start, Length: 175, dtype: int64
```

```
 df.stop.sort_values().value_counts()
```

```
         Cary                   203
         Unknown Location       149
         Morrisville             84
         Whitebridge             65
         Islamabad               58
         ...
         Nugegoda                 1
         Isles of Buena Vista     1
         Seattle                  1
         St Thomas                1
In [20]: University District      1
Out[20]: Name: stop, Length: 186, dtype: int64
In [21]: df["year"] = df.startdate.dt.year
          df["month"] = df.startdate.dt.month
          df["date"] = df.startdate.dt.date
          df["week"] = df.startdate.dt.isocalendar().week
          df["time"] = df.startdate.dt.time
          df["minutes"] = (df["enddate"]-df["startdate"]).dt.total_seconds()/60
```

```
In [22]:  df
```

Out[22]:

| | startdate | enddate | category | start | stop | miles | purpose | year | month | da |
|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 2016-01-01 21:11:00 | 2016-01-01 21:17:00 | Business | Fort Pierce | Fort Pierce | 5.1 | Meal/Entertain | 2016 | 1 | 201 01- |
| **1** | 2016-01-02 01:25:00 | 2016-01-02 01:37:00 | Business | Fort Pierce | Fort Pierce | 5.0 | NaN | 2016 | 1 | 201 01- |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **2** | 2016-01-02 20:25:00 | 2016-01-02 20:38:00 | Business | Fort Pierce | Fort Pierce | 4.8 | Errand/Supplies | 2016 | 1 | 201 01- |
| **3** | 2016-01-05 17:31:00 | 2016-01-05 17:45:00 | Business | Fort Pierce | Fort Pierce | 4.7 | Meeting | 2016 | 1 | 201 01- |
| **4** | 2016-01-06 14:42:00 | 2016-01-06 15:49:00 | Business | Fort Pierce | West Palm Beach | 63.7 | Customer Visit | 2016 | 1 | 201 01- |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| **1150** | 2016-12-31 01:07:00 | 2016-12-31 01:14:00 | Business | Karachi | Karachi | 0.7 | Meeting | 2016 | 12 | 201 12- |
| **1151** | 2016-12-31 13:24:00 | 2016-12-31 13:42:00 | Business | Karachi | Unknown Location | 3.9 | Temporary Site | 2016 | 12 | 201 12- |
| **1152** | 2016-12-31 15:03:00 | 2016-12-31 15:38:00 | Business | Unknown Location | Unknown Location | 16.2 | Meeting | 2016 | 12 | 201 12- |
| **1153** | 2016-12-31 21:32:00 | 2016-12-31 21:50:00 | Business | Katunayake | Gampaha | 6.4 | Temporary Site | 2016 | 12 | 201 12- |
| **1154** | 2016-12-31 22:08:00 | 2016-12-31 23:51:00 | Business | Gampaha | Ilukwatta | 48.2 | Temporary Site | 2016 | 12 | 201 12- |

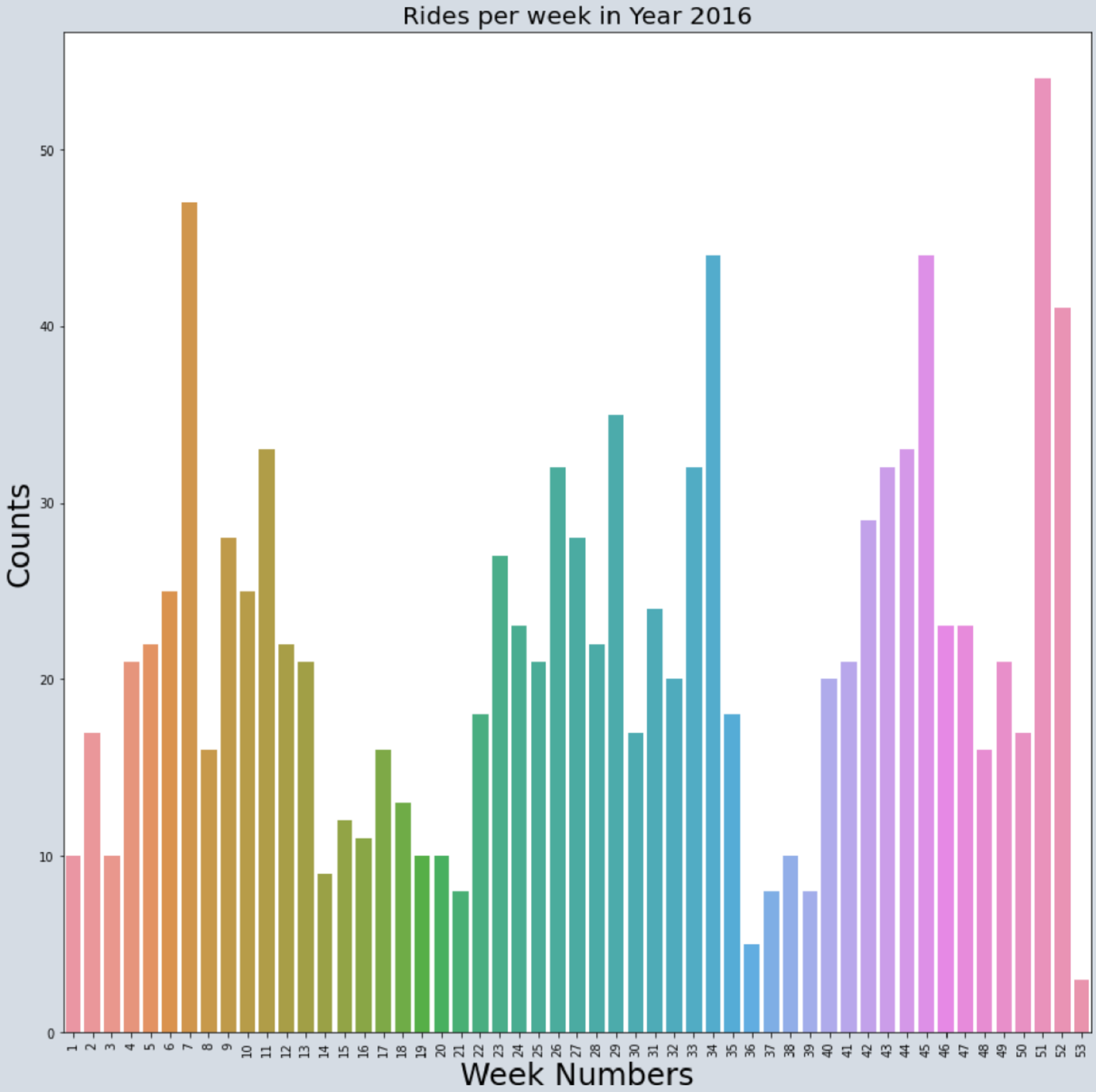
1155 rows × 13 columns

In [23]:

```
fig_dims = (15, 15)
fig, ax = plt.subplots(figsize=fig_dims) gg =
sns.countplot(x = "week", ax = ax, data = df) p1 =
plt.setp(gg.get_xticklabels(), rotation = 90)
plt.title("Rides per week in Year 2016", fontsize =
20) plt.xlabel("Week Numbers", fontsize = 25)
plt.ylabel("Counts", fontsize = 25)
```

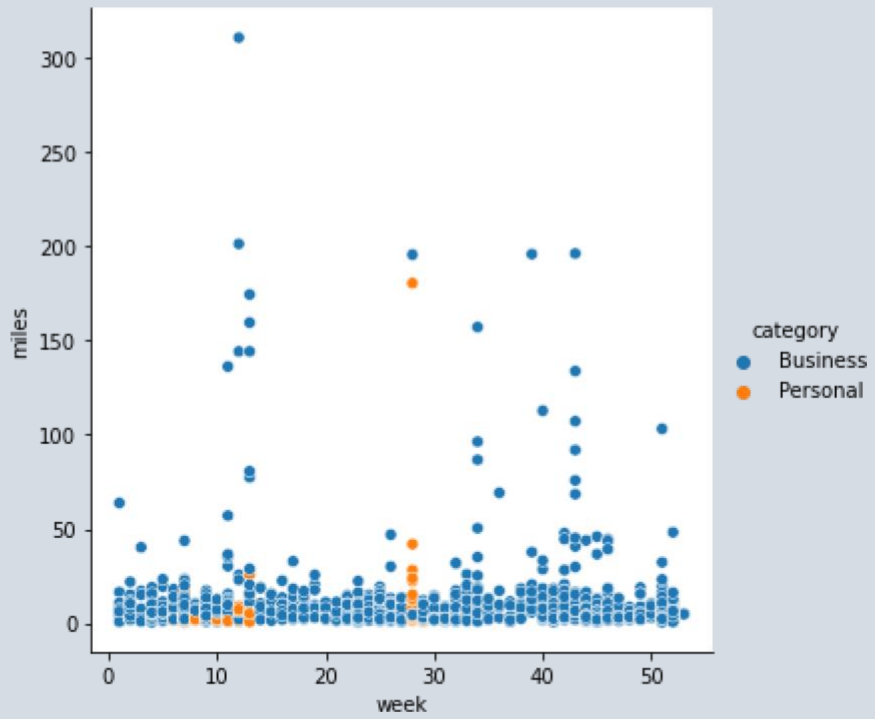
Out[23]:

```
Text(0, 0.5, 'Counts')
```



```
In [24]: sns.relplot(x = "week", y = "miles", hue = "category", data = df)
```

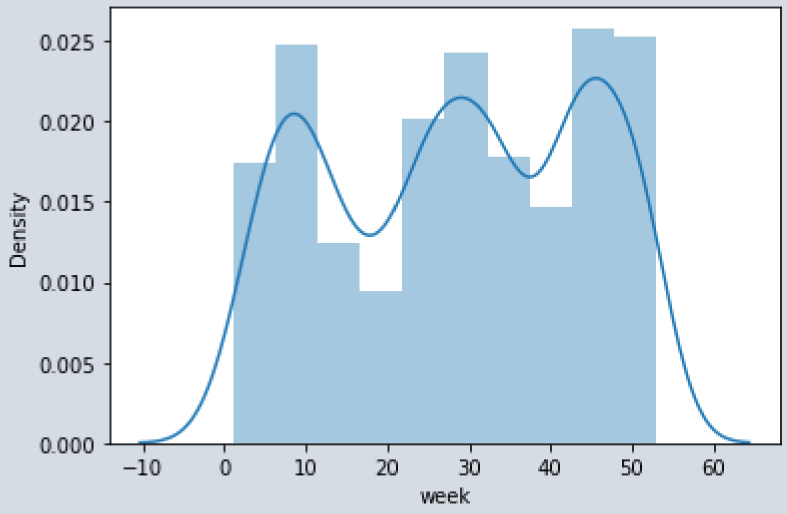
```
Out[24]: < seaborn.axisgrid.FacetGrid at 0x2a4ce92c0d 0>
```



```
In [25]: sns.distplot(df["week"])
```

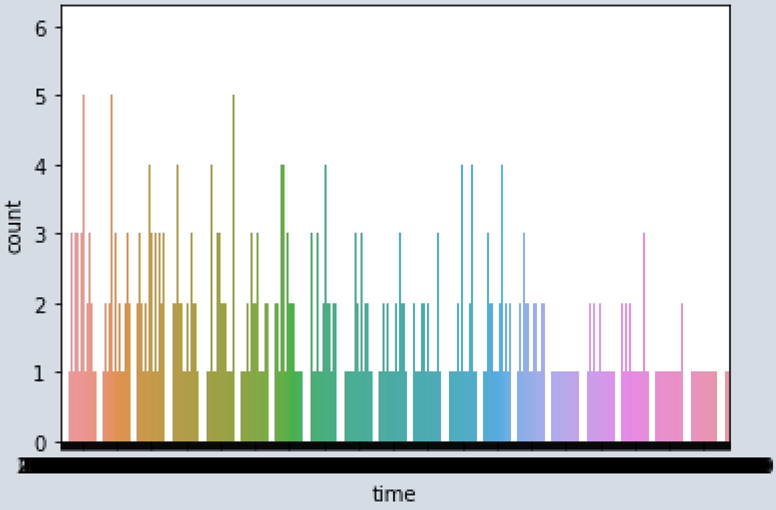
```
C:\Users\Sakshi\anaconda3\lib\site-packages\seaborn\distributions.py:2557:
FutureWar ning: `distplot` is a deprecated function and will be removed in a future
version. P lease adapt your code to use either `displot` (a figure-level function
with similar  flexibility) or `histplot` (an axes-level function for histograms).
warnings.warn(msg, FutureWarning)
```

```
Out[25]: < AxesSubplot:xlabel='week', ylabel='Density' >
```



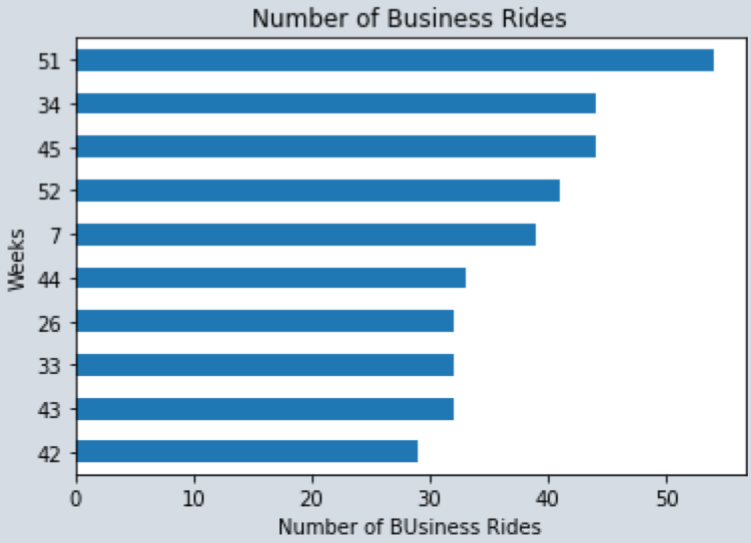
```
In [26]:  sns.countplot(x = "time", data = df)
```

```
Out[26]: < AxesSubplot:xlabel='time', ylabel='count' >
```



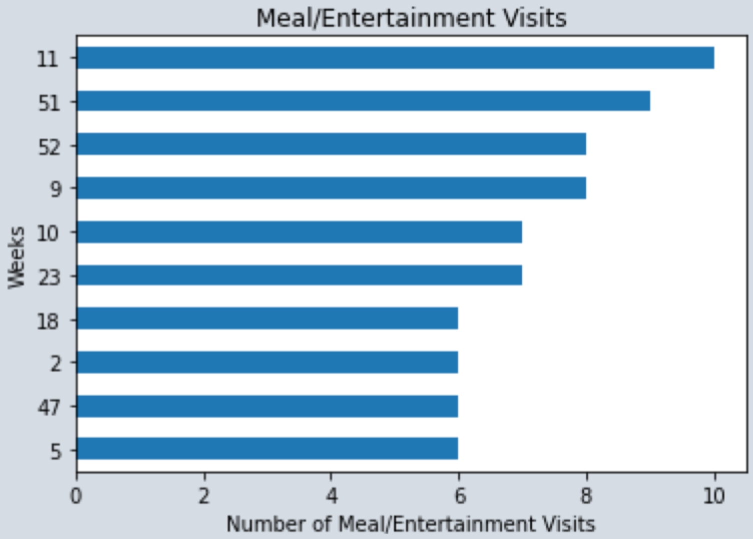
```
In [27]: df[df['category'] ==
         'Business']['week'].value_counts()[:10].sort_values().plot.barh plt.title("Number
         of Business Rides") plt.xlabel("Number of BUsiness Rides") plt.ylabel("Weeks")
```

```
Out[27]:   Text(0, 0.5, 'Weeks')
```



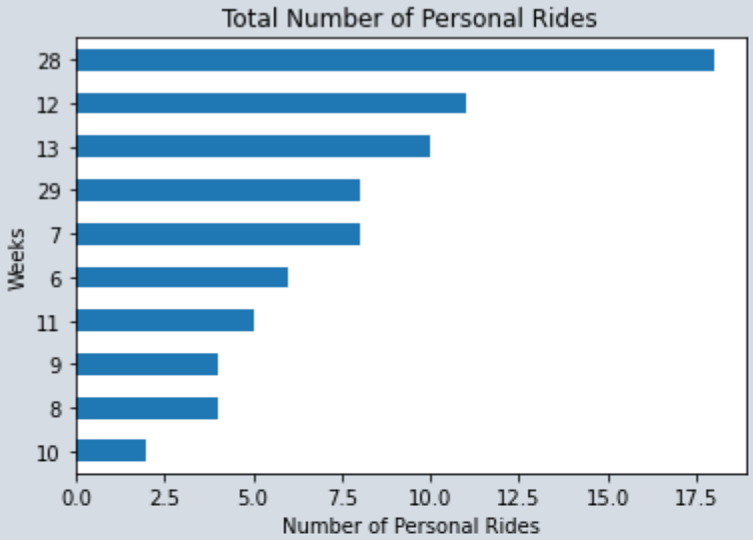
```
In [28]:  df[df["purpose"] ==
          'Meal/Entertain']["week"].value_counts()[:10].sort_values().plot
          plt.title("Meal/Entertainment Visits")
          plt.xlabel("Number of Meal/Entertainment Visits")
          plt.ylabel("Weeks")
```

```
Out[28]:   Text(0, 0.5, 'Weeks')
```



In [29]:

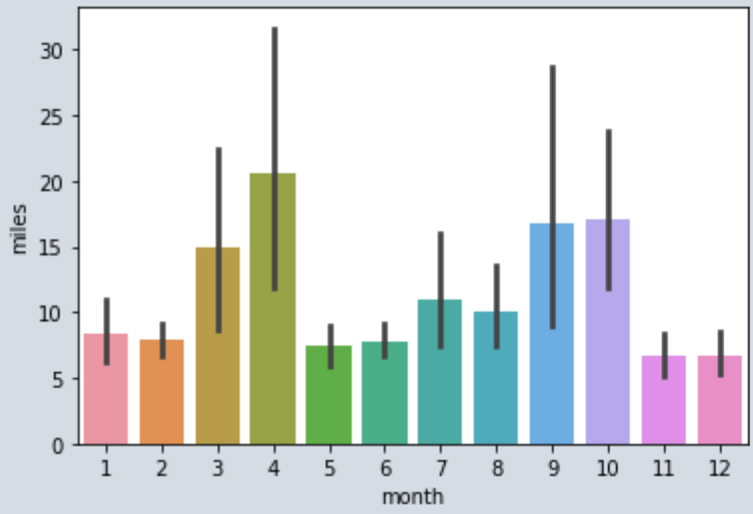
```
df[df["category"]=="Personal"]["week"].value_counts()[:10].sort_values().plot.barh
() plt.title("Total Number of Personal Rides") plt.xlabel("Number of Personal
Rides") plt.ylabel("Weeks") plt.show()
```



In [30]:

```
# To calculate the Number of Miles Each Month the Traveler has Travelled
sns.barplot(x = "month", y = "miles", data = df)
```

```
Out[30]: < AxesSubplot:xlabel='month', ylabel='miles' >
```

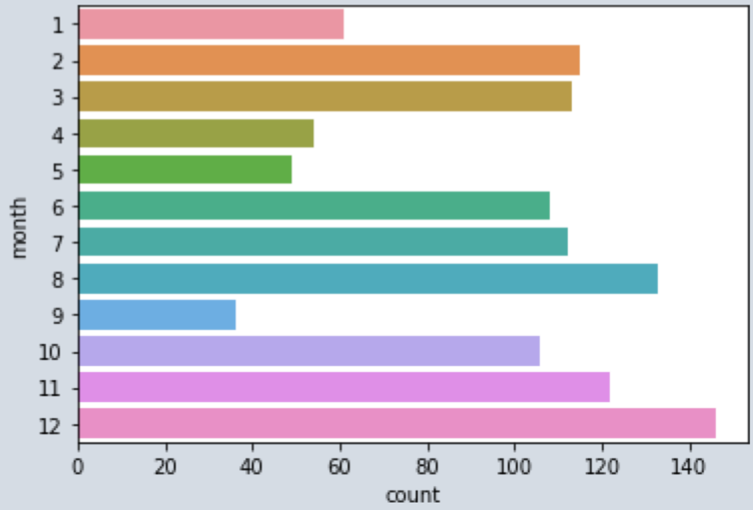


In [31]:

```
sns.countplot(y = "month", data = df)
```

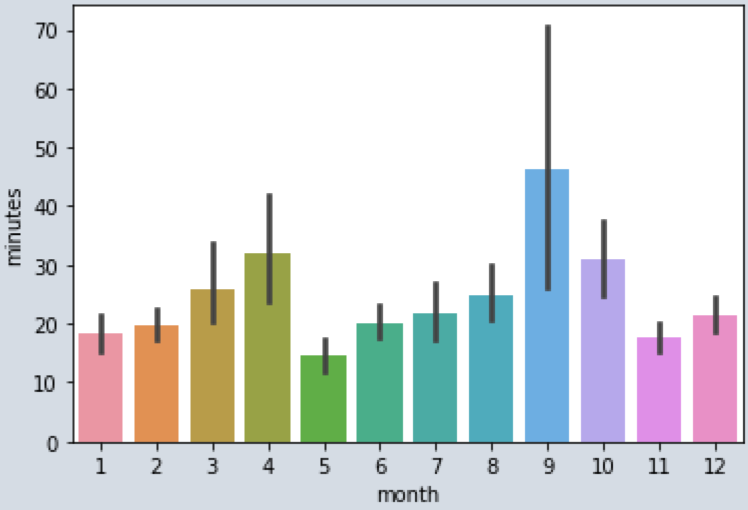
Out[31]:

```
< AxesSubplot:xlabel='count', ylabel='month' >
```



```
In [32]: sns.barplot(x = "month", y = "minutes", data = df)
```

```
<AxesSubplot:xlabel='month', ylabel='minutes'>
```

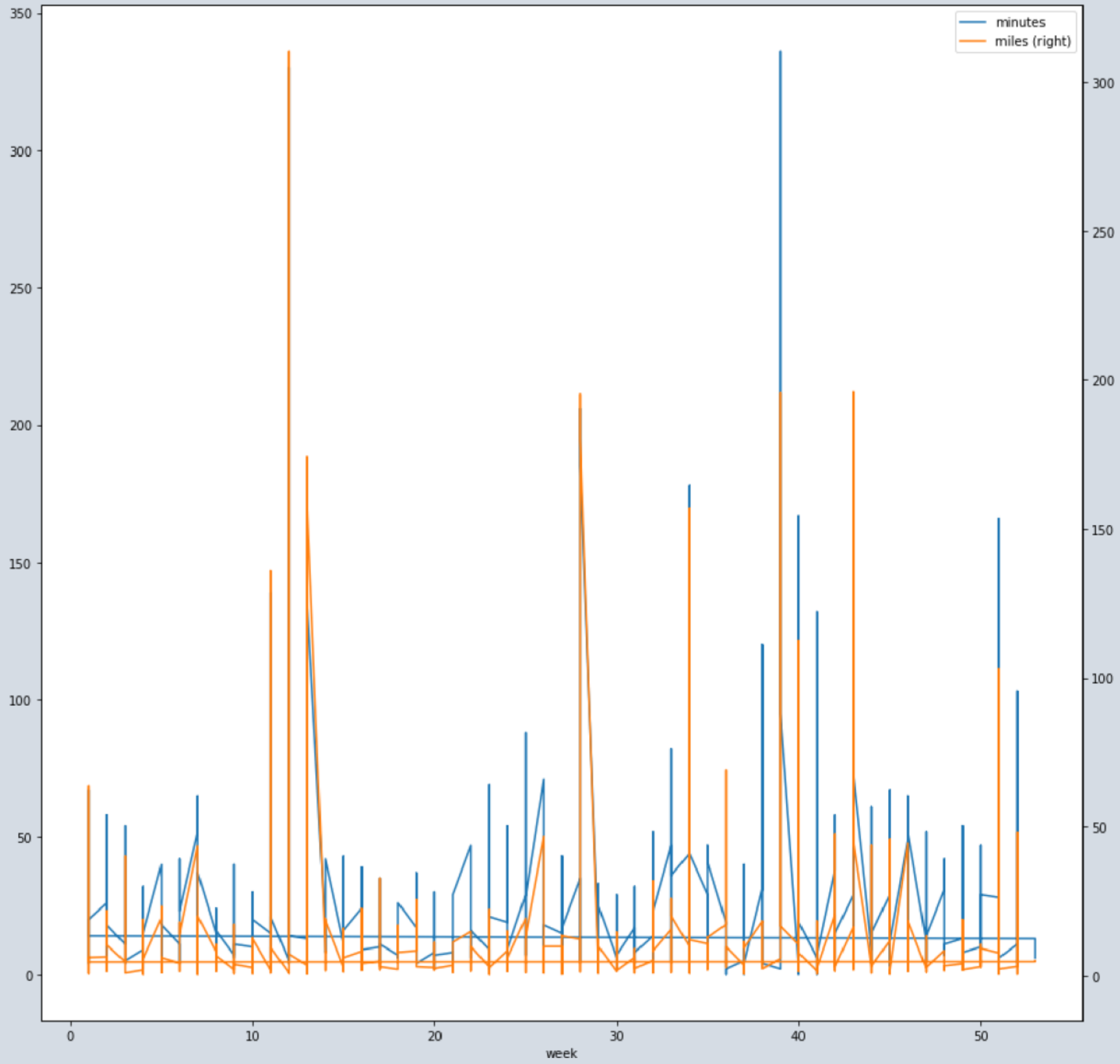


```
fig_dims = (15, 15)
fig, ax = plt.subplots(figsize=fig_dims)
```

```
Out[32]:
```

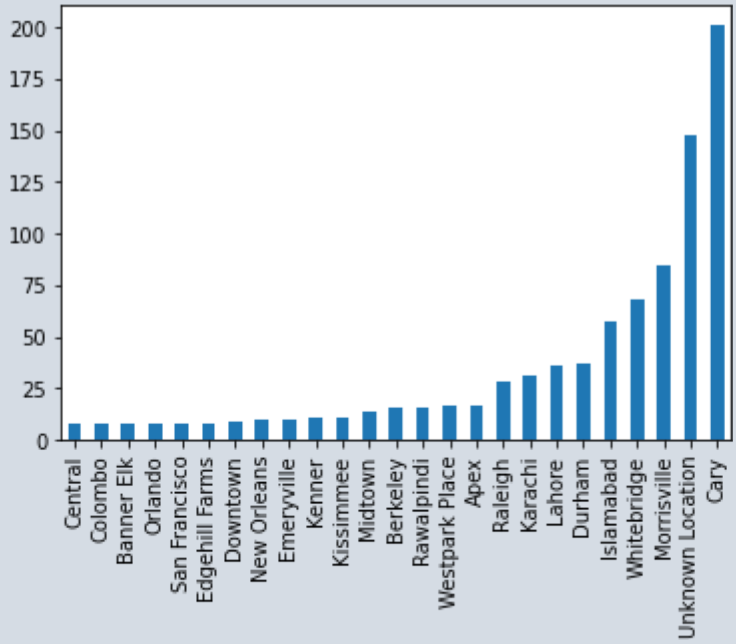
```
In [33]:
df.plot(x = 'week', y = 'minutes', ax = ax) df.plot(x =
'week', y = 'miles', ax = ax, secondary_y = True)
```

```
Out[33]: < AxesSubplot:label='3955fef6-04a2-4b21-8af2-a4f15e3b6555' >
```



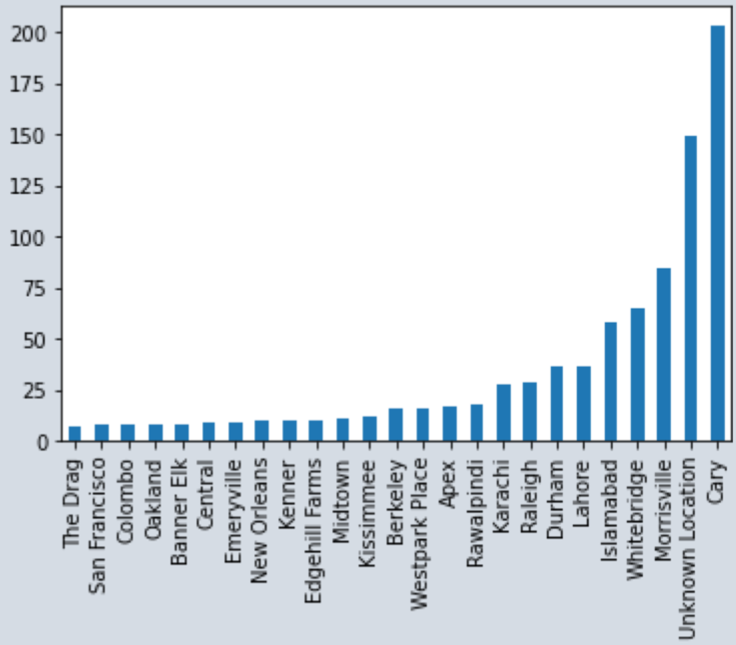
```
In [34]:  df["start"].value_counts()[:25].sort_values().plot.bar()
```

```
Out[34]: < AxesSubplot :>
```



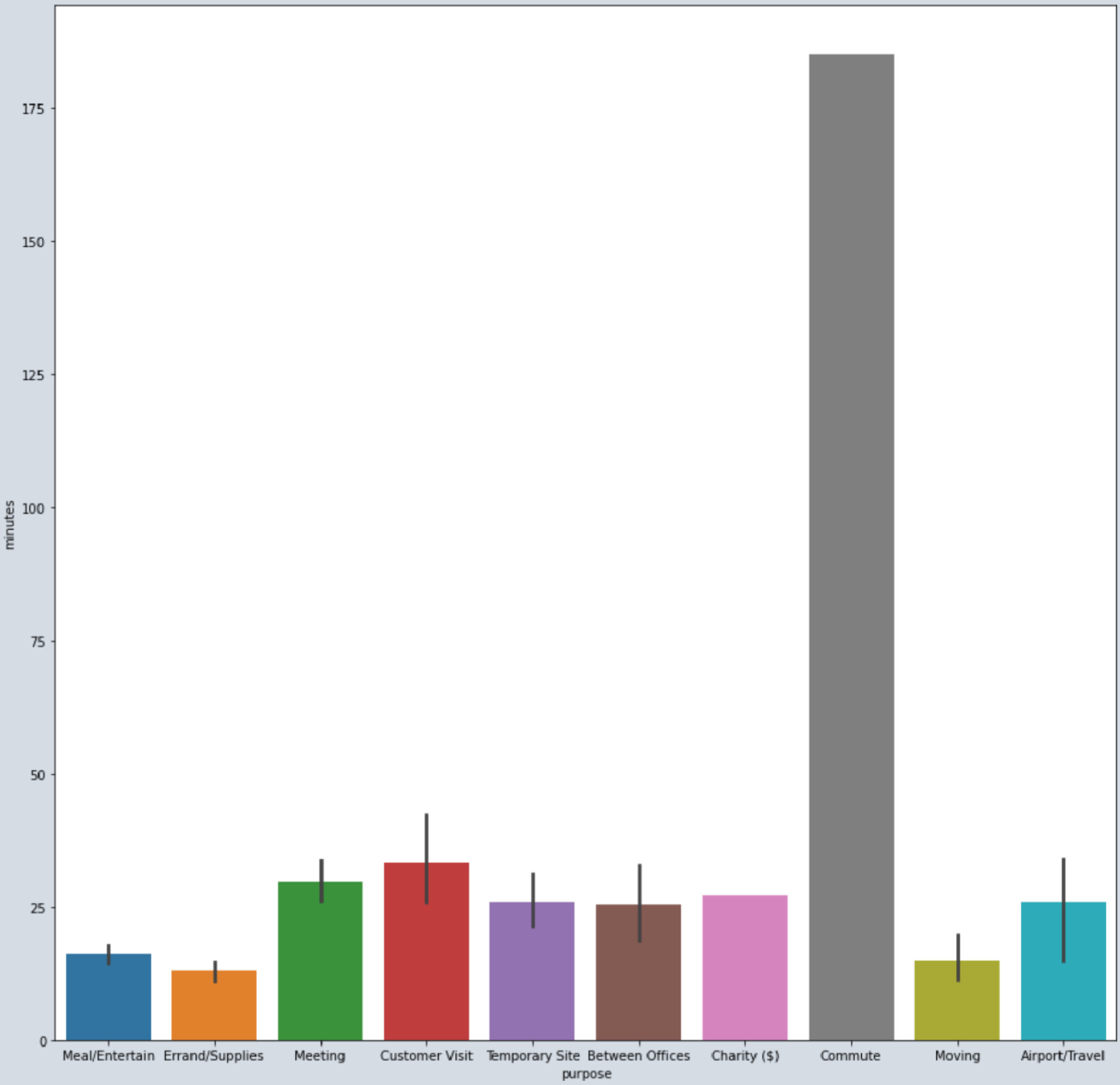
```
In [35]: df["stop"].value_counts()[:25].sort_values().plot.bar()
```

```
Out[35]: < AxesSubplot :>
```



```
In [36]: fig_dims =
           (15, 15)
fig, ax =
plt.subplots(figsiz
e=fig_dims)
sns.barplot(x =
'purpose', y =
'minutes', data =
df, ax = ax)
```

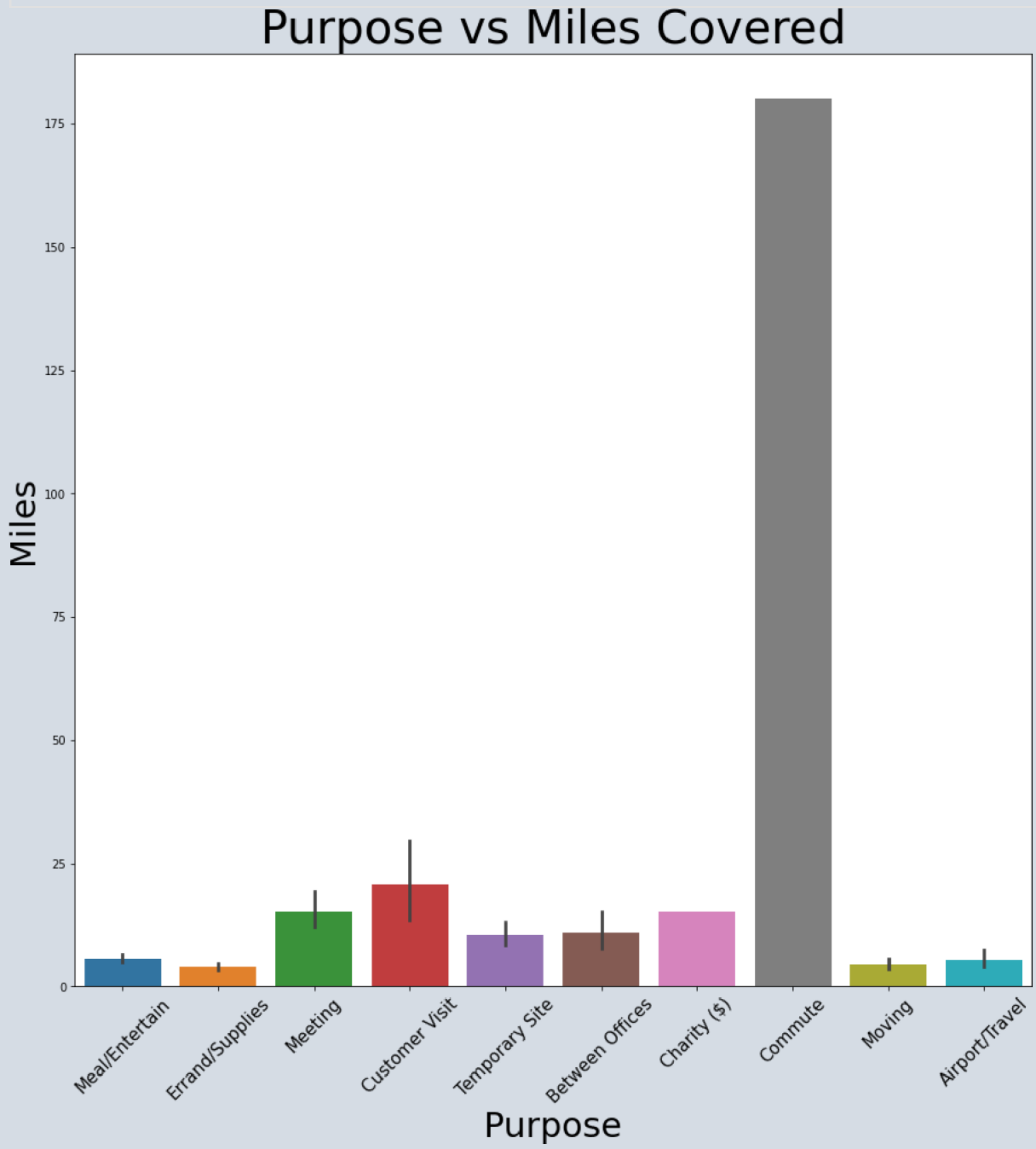
```
Out[36]:   < AxesSubplot:xlabel='purpose', ylabel='minutes' >
```



```
In [37]:  fig_dims = (15, 15)
          fig, ax = plt.subplots(figsize=fig_dims)
          sns.barplot(x = 'purpose', y = 'miles', data = df, ax =
          ax) plt.title("Purpose vs Miles Covered", fontsize =  40)
          plt.ylabel("Miles", fontsize = 30) plt.xlabel("Purpose",
          fontsize = 30) plt.xticks(rotation = 45, fontsize = 15)
```

```
Out[37]:   (array([0, 1, 2, 3, 4, 5, 6, 7, 8, 9]),
         [Text(0, 0, 'Meal/Entertain'),
           Text(1, 0, 'Errand/Supplies'),
           Text(2, 0, 'Meeting'),
           Text(3, 0, 'Customer Visit'),
           Text(4, 0, 'Temporary Site'),
           Text(5, 0, 'Between Offices'),
           Text(6, 0, 'Charity ($)'),
         Text(7, 0, 'Commute'),
           Text(8, 0, 'Moving'),
           Text(9, 0, 'Airport/Travel')])
```

```
In [ ]:
```


Purpose vs Miles Covered
Miles
175
150
125
100
75
50
25
0
Meal/Entertain
Errand/Supplies
Meeting
Customer Visit
Temporary Site
Between Offices
Charity ($)
Commute
Moving
Airport/Travel
Purpose