


```

dataset['END_DATE'] = pd.to_datetime(dataset['END_DATE'],
                                     errors='coerce')

from datetime import datetime

dataset['date'] = pd.DatetimeIndex(dataset['START_DATE']).date
dataset['time'] = pd.DatetimeIndex(dataset['START_DATE']).hour

#changing into categories of day and night
dataset['day-night'] = pd.cut(x=dataset['time'],
                             bins = [0,10,15,19,24],
                             labels =
                             ['Morning','Afternoon','Evening','Night'])

dataset.dropna(inplace=True)

dataset.drop_duplicates(inplace=True)

obj = (dataset.dtypes == 'object')
object_cols = list(obj[obj].index)

unique_values = {}
for col in object_cols:
    unique_values[col] = dataset[col].unique().size
unique_values

{'CATEGORY': 1, 'START': 69, 'STOP': 74, 'PURPOSE': 6, 'date': 85}

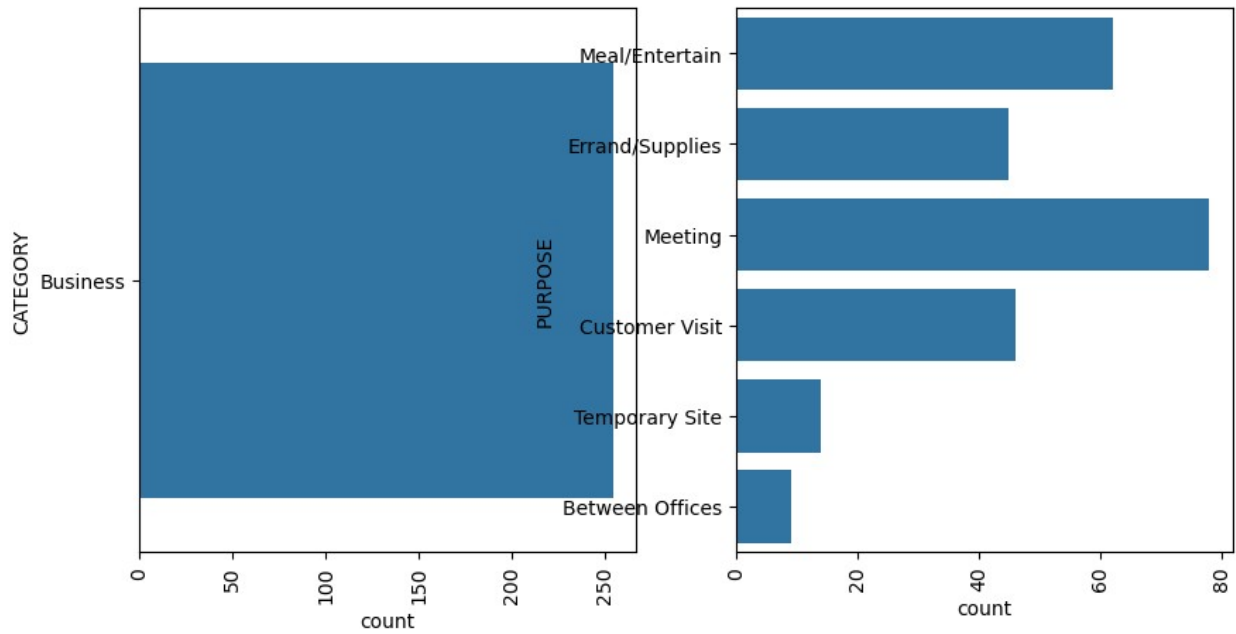
plt.figure(figsize=(10,5))

plt.subplot(1,2,1)
sns.countplot(dataset['CATEGORY'])
plt.xticks(rotation=90)

plt.subplot(1,2,2)
sns.countplot(dataset['PURPOSE'])
plt.xticks(rotation=90)

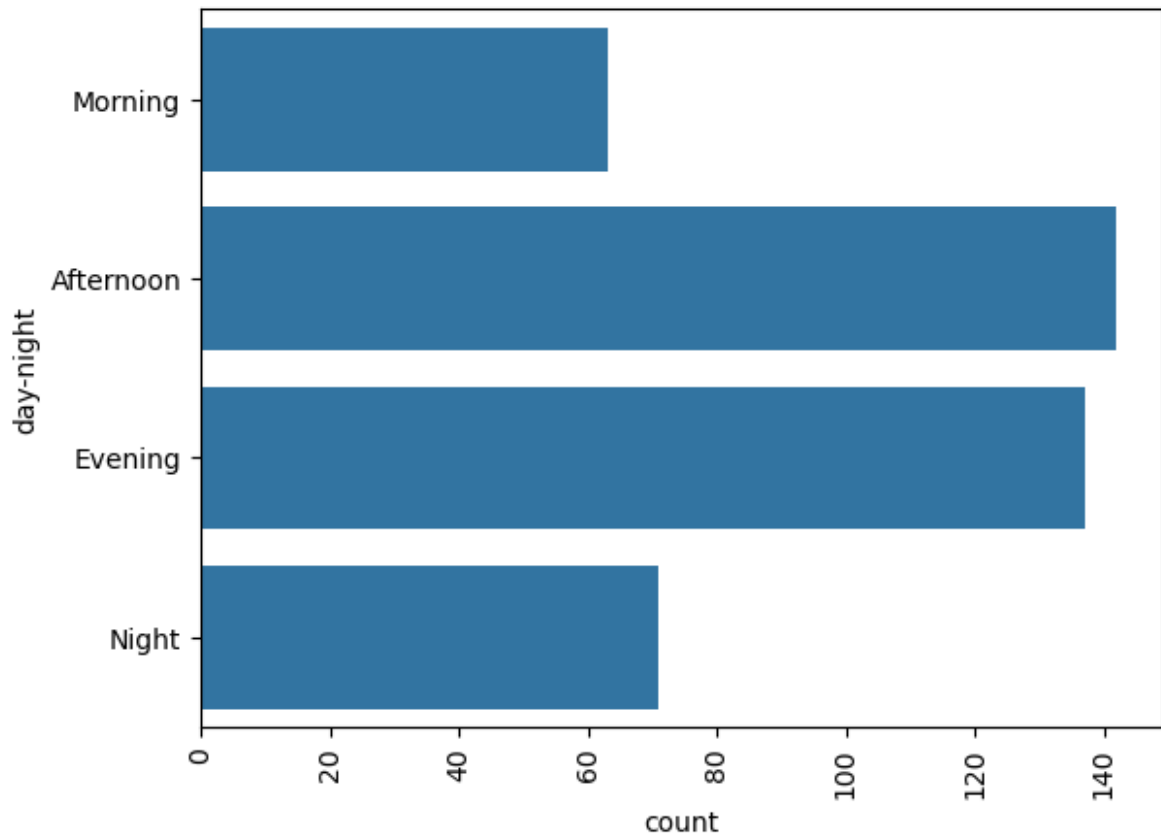
(array([ 0., 20., 40., 60., 80., 100.]),
 [Text(0.0, 0, '0'),
  Text(20.0, 0, '20'),
  Text(40.0, 0, '40'),
  Text(60.0, 0, '60'),
  Text(80.0, 0, '80'),
  Text(100.0, 0, '100')])

```

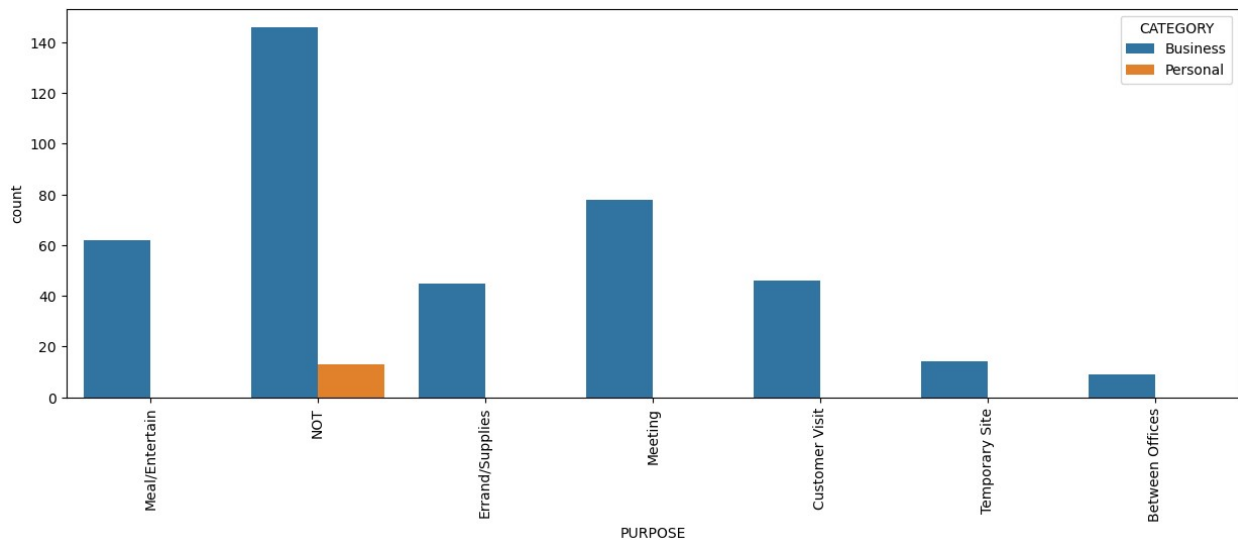


```
sns.countplot(dataset['day-night'])
plt.xticks(rotation=90)

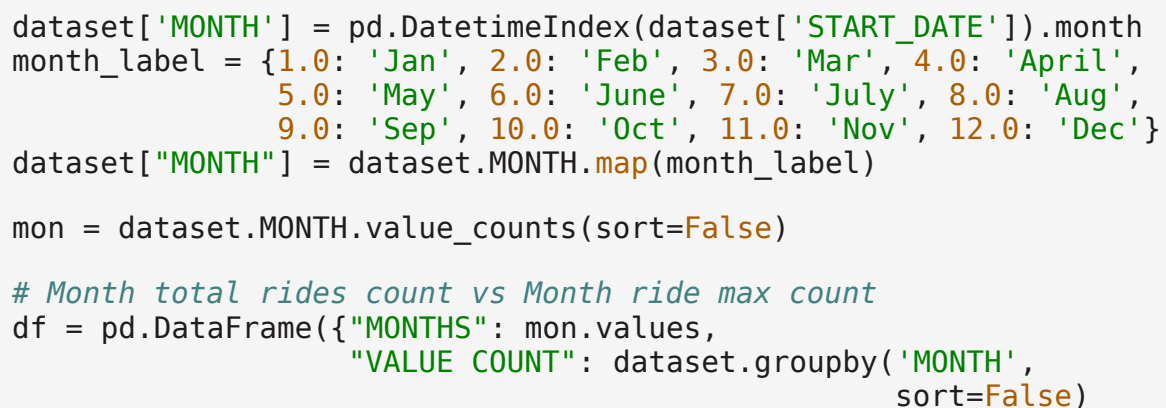
(array([ 0., 20., 40., 60., 80., 100., 120., 140., 160.]),
 [Text(0.0, 0, '0'),
  Text(20.0, 0, '20'),
  Text(40.0, 0, '40'),
  Text(60.0, 0, '60'),
  Text(80.0, 0, '80'),
  Text(100.0, 0, '100'),
  Text(120.0, 0, '120'),
  Text(140.0, 0, '140'),
  Text(160.0, 0, '160')])
```



```
plt.figure(figsize=(15, 5))
sns.countplot(data=dataset, x='PURPOSE', hue='CATEGORY')
plt.xticks(rotation=90)
plt.show()
```



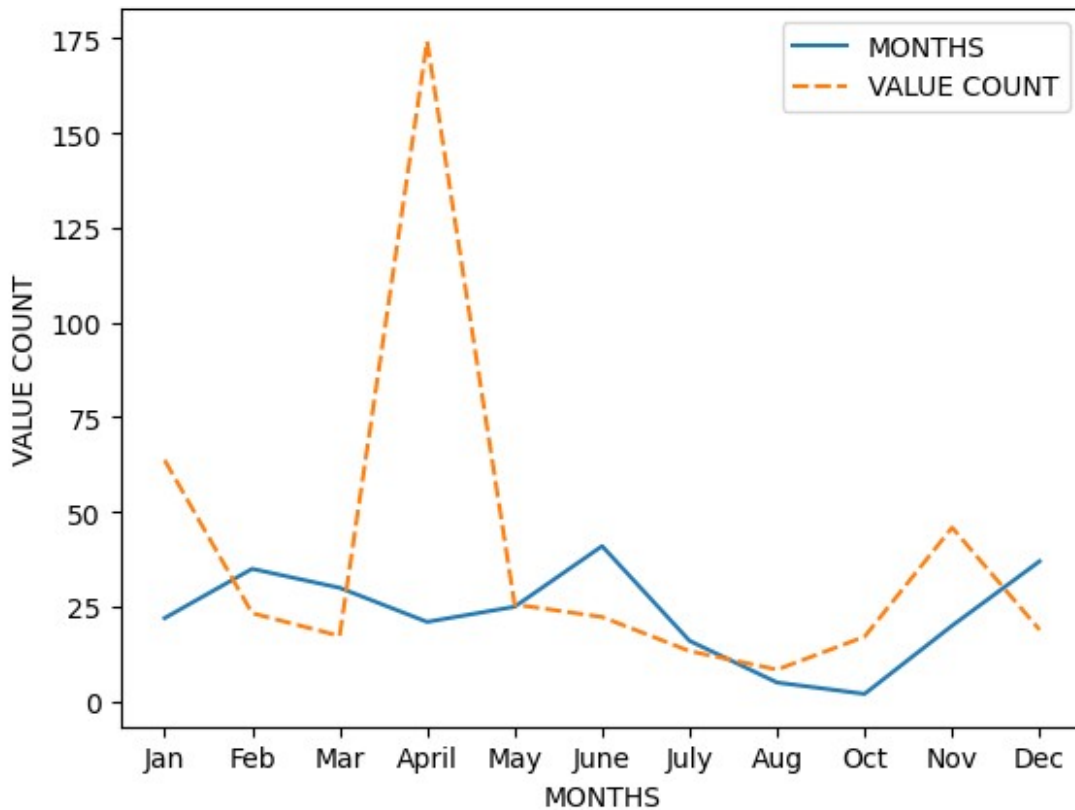
```
<Axes:  >
```



```
[ 'MILES' ].max()})

p = sns.lineplot(data=df)
p.set(xlabel="MONTHS", ylabel="VALUE COUNT")

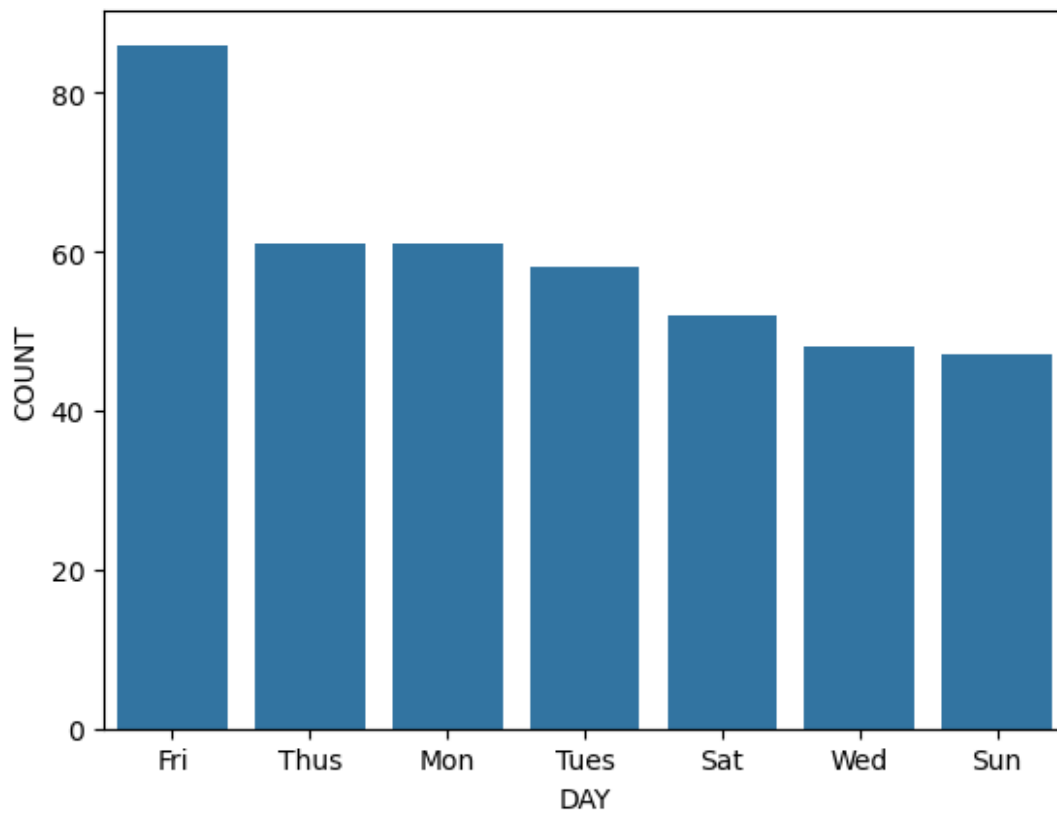
[Text(0.5, 0, 'MONTHS'), Text(0, 0.5, 'VALUE COUNT')]
```



```
dataset['DAY'] = dataset.START_DATE.dt.weekday
day_label = {
    0: 'Mon', 1: 'Tues', 2: 'Wed', 3: 'Thus', 4: 'Fri', 5: 'Sat', 6:
    'Sun'
}
dataset['DAY'] = dataset['DAY'].map(day_label)

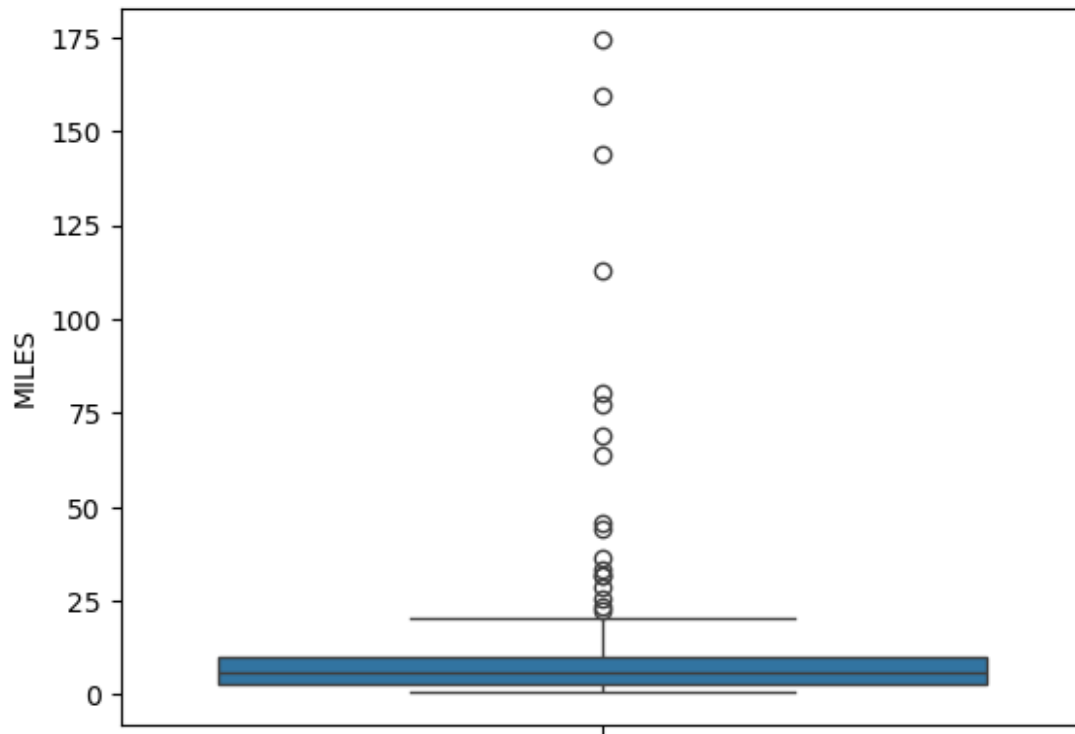
day_label = dataset.DAY.value_counts()
sns.barplot(x=day_label.index, y=day_label);
plt.xlabel('DAY')
plt.ylabel('COUNT')

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

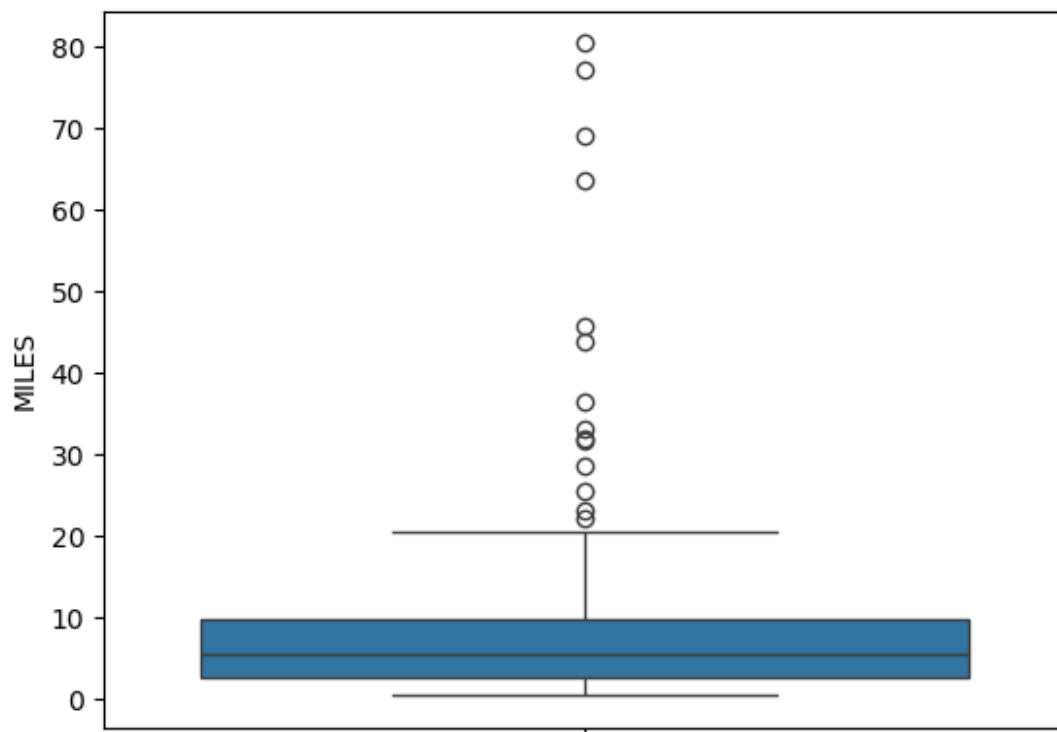


```
sns.boxplot(dataset['MILES'])
```

```
<Axes: ylabel='MILES'>
```



```
sns.boxplot(dataset[dataset['MILES']<100]['MILES'])  
<Axes: ylabel='MILES'>
```




```
sns.distplot(dataset[dataset['MILES']<40]['MILES'])
```

C:\Users\LENOVO\AppData\Local\Temp\ipykernel_10632\1678554178.py:1:
UserWarning:

`distplot` is a deprecated function and will be removed in seaborn
v0.14.0.

Please adapt your code to use either `displot` (a figure-level
function with
similar flexibility) or `histplot` (an axes-level function for
histograms).

For a guide to updating your code to use the new functions, please see
<https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751>

```
sns.distplot(dataset[dataset['MILES']<40]['MILES'])
```

```
<Axes: xlabel='MILES', ylabel='Density'>
```

