In [22]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

In [90]:
```
import math
```

In [94]:
```
import numpy as np
```

In [60]:
```
#Imports DF
df = pd.read_csv("C:\\Users\\taylo\\Desktop\\GA\\DSI_Assignments_New\\Projector\\ProjectorTask2.csv")
```

In [42]:
```
#Checks out DF
df
```

Out[42]:

	Month of	Hours	Percent of Total	Revenue
0	Jun-19	10.0	NaN	NaN
1	Jul-19	120.0	NaN	NaN
2	Aug-19	100.0	NaN	NaN
3	Sep-19	70.0	NaN	NaN
4	Oct-19	100.0	NaN	NaN
5	Nov-19	50.0	NaN	NaN
6	Dec-19	100.0	NaN	NaN
7	Jan-20	10.0	NaN	NaN
8	TOTAL	NaN	NaN	$0.32

In [36]:
```
#Checks out columns
df.columns
```

Out[36]:
```
Index(['Month of', 'Hours', 'Percent of Total', 'Revenue'], dtype='object')
```

In [35]:
```
#Checks out data types
df.dtypes
```

Out[35]:
```
Month of            object
Hours              float64
Percent of Total   float64
Revenue             object
dtype: object
```

In [39]:
```
#Info of DF
df.info
```

Out[39]:
```
<bound method DataFrame.info of       Month of  Hours  Percent of Total  Reve
nue
0       Jun-19   10.0               NaN      NaN
1       Jul-19  120.0               NaN      NaN
2       Aug-19  100.0               NaN      NaN
3       Sep-19   70.0               NaN      NaN
4       Oct-19  100.0               NaN      NaN
5       Nov-19   50.0               NaN      NaN
6       Dec-19  100.0               NaN      NaN
7       Jan-20   10.0               NaN      NaN
8        TOTAL    NaN               NaN   $0.32
Total      560  560.0             560.0      560>
```

In [76]:
```
#Gets rid of Row 8
df = df[:8]
```

In [77]:
```
#Confirms Row 8 not in DF
df
```

Out[77]:

	Month of	Hours	Percent of Total	Revenue
0	Jun-19	10.0	0.017857	1.0
1	Jul-19	120.0	0.214286	7.0
2	Aug-19	100.0	0.178571	6.0
3	Sep-19	70.0	0.125000	4.0
4	Oct-19	100.0	0.178571	6.0
5	Nov-19	50.0	0.089286	3.0
6	Dec-19	100.0	0.178571	6.0
7	Jan-20	10.0	0.017857	1.0

In [78]:
```
#Creates function for calculating monthly percentage
def calculate_percent(row):
    return row['Hours'] / 560

df.apply(calculate_percent, axis=1)
```

Out[78]:
```
0    0.017857
1    0.214286
2    0.178571
3    0.125000
4    0.178571
5    0.089286
6    0.178571
7    0.017857
dtype: float64
```

In [79]:
```
#Applies monthly percentage function to new column
df['Percent of Total'] = df.apply(calculate_percent, axis=1)
df
```

```
C:\Users\taylo\Anaconda3\envs\dsi\lib\site-packages\ipykernel_launcher.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/st
able/indexing.html#indexing-view-versus-copy
  """Entry point for launching an IPython kernel.
```

Out[79]:

	Month of	Hours	Percent of Total	Revenue
0	Jun-19	10.0	0.017857	1.0
1	Jul-19	120.0	0.214286	7.0
2	Aug-19	100.0	0.178571	6.0
3	Sep-19	70.0	0.125000	4.0
4	Oct-19	100.0	0.178571	6.0
5	Nov-19	50.0	0.089286	3.0
6	Dec-19	100.0	0.178571	6.0
7	Jan-20	10.0	0.017857	1.0

In [80]:
```
#Creates calculate revenue function
def calculate_revenue(row):
    return row['Percent of Total'] * 32

df.apply(calculate_revenue, axis=1)
```

Out[80]:
```
0    0.571429
1    6.857143
2    5.714286
3    4.000000
4    5.714286
5    2.857143
6    5.714286
7    0.571429
dtype: float64
```

In [81]:
```
#Creates new column applying revenue function
df['Revenue'] = df.apply(calculate_revenue, axis=1)
df
```

```
C:\Users\taylo\Anaconda3\envs\dsi\lib\site-packages\ipykernel_launcher.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/st
able/indexing.html#indexing-view-versus-copy
  """Entry point for launching an IPython kernel.
```

Out[81]:

	Month of	Hours	Percent of Total	Revenue
0	Jun-19	10.0	0.017857	0.571429
1	Jul-19	120.0	0.214286	6.857143
2	Aug-19	100.0	0.178571	5.714286
3	Sep-19	70.0	0.125000	4.000000
4	Oct-19	100.0	0.178571	5.714286
5	Nov-19	50.0	0.089286	2.857143
6	Dec-19	100.0	0.178571	5.714286
7	Jan-20	10.0	0.017857	0.571429

In [82]:
```
#Checks out data types
df.dtypes
```

Out[82]:
```
Month of            object
Hours              float64
Percent of Total   float64
Revenue            float64
dtype: object
```

In [83]:
```
#Rounds down Revenue column to whole integer
df['Revenue'] = df.Revenue.astype(float).round()
```

```
C:\Users\taylo\Anaconda3\envs\dsi\lib\site-packages\ipykernel_launcher.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/st
able/indexing.html#indexing-view-versus-copy
  """Entry point for launching an IPython kernel.
```

In [ ]:
```
#Creates total Row
df.loc['Total',:]= df.sum(axis=0)
```

In [91]:
```
#Creates Second Revenue Column, because first was greater than 32
df['Revenue_RD'] = df.apply(calculate_revenue, axis=1)
df
```

Out[91]:

	Month of	Hours	Percent of Total	Revenue	Revenue_RD
0	Jun-19	10.0	0.017857	1.0	0.571429
1	Jul-19	120.0	0.214286	7.0	6.857143
2	Aug-19	100.0	0.178571	6.0	5.714286
3	Sep-19	70.0	0.125000	4.0	4.000000
4	Oct-19	100.0	0.178571	6.0	5.714286
5	Nov-19	50.0	0.089286	3.0	2.857143
6	Dec-19	100.0	0.178571	6.0	5.714286
7	Jan-20	10.0	0.017857	1.0	0.571429
Total	Jun-19Jul-19Aug-19Sep-19Oct-19Nov-19Dec-19Jan-20	560.0	1.000000	34.0	32.000000

In [96]:
```
#Rounds down new Revenue column
df['Revenue_RD'] = df['Revenue_RD'].apply(np.floor)
```

In [97]:
```
#Confirms new Revenue Column rounds down and Total is equal to 32
df
```

Out[97]:

	Month of	Hours	Percent of Total	Revenue	Revenue_RD
0	Jun-19	10.0	0.017857	1.0	0.0
1	Jul-19	120.0	0.214286	7.0	6.0
2	Aug-19	100.0	0.178571	6.0	5.0
3	Sep-19	70.0	0.125000	4.0	4.0
4	Oct-19	100.0	0.178571	6.0	5.0
5	Nov-19	50.0	0.089286	3.0	2.0
6	Dec-19	100.0	0.178571	6.0	5.0
7	Jan-20	10.0	0.017857	1.0	0.0
Total	Jun-19Jul-19Aug-19Sep-19Oct-19Nov-19Dec-19Jan-20	560.0	1.000000	34.0	32.0