In [1]:
```
import pandas as pd
```

In [2]:
```
data = pd.read_csv('.//sleephealth.csv', index_col='Person ID')
data.head(10)
```

Out[2]:

Person ID	Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	BMI Category	Blood Pressure	Heart Rate	Daily Steps	Dis
1	Male	27	Software Engineer	6.1	6	42	6	Overweight	126/83	77	4200	
2	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	
3	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	
4	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	
5	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	
6	Male	28	Software Engineer	5.9	4	30	8	Obese	140/90	85	3000	Ins
7	Male	29	Teacher	6.3	6	40	7	Obese	140/90	82	3500	Ins
8	Male	29	Doctor	7.8	7	75	6	Normal	120/80	70	8000	
9	Male	29	Doctor	7.8	7	75	6	Normal	120/80	70	8000	
10	Male	29	Doctor	7.8	7	75	6	Normal	120/80	70	8000	

Las preguntas que buscaremos responder con los datos de este conjunto son las siguientes: ¿Quiénes tienen peor calidad del sueño, hombres o mujeres? ¿Existe una relación entre la calidad del sueño de las personas y su profesión? ¿La actividad física afecta al sueño? ¿Qué profesión presenta la peor calidad de sueño? ¿Cuál profesión tienen a las personas con el mayor nivel de estrés y el mayor índice de masa corporal? ¿En qué rango de edades se encuentran la mayor cantidad de trastornos del sueño? La cantidad de pasos al día, ¿afecta la calidad del sueño? ¿al índice de masa corporal?

In [3]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 374 entries, 1 to 374
Data columns (total 12 columns):
 #   Column                   Non-Null Count  Dtype  
---  ------                   --------------  -----  
 0   Gender                   374 non-null    object 
 1   Age                      374 non-null    int64  
 2   Occupation               374 non-null    object 
 3   Sleep Duration           374 non-null    float64
 4   Quality of Sleep         374 non-null    int64  
 5   Physical Activity Level  374 non-null    int64  
 6   Stress Level             374 non-null    int64  
 7   BMI Category             374 non-null    object 
 8   Blood Pressure           374 non-null    object 
 9   Heart Rate               374 non-null    int64  
 10  Daily Steps              374 non-null    int64  
 11  Sleep Disorder           155 non-null    object 
```

```
dtypes: float64(1), int64(6), object(5)
memory usage: 38.0+ KB
```

In [4]:
```
data.describe()
```

Out[4]:

	Age	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	Heart Rate	Daily Steps
count	374.000000	374.000000	374.000000	374.000000	374.000000	374.000000	374.000000
mean	42.184492	7.132086	7.312834	59.171123	5.385027	70.165775	6816.844920
std	8.673133	0.795657	1.196956	20.830804	1.774526	4.135676	1617.915679
min	27.000000	5.800000	4.000000	30.000000	3.000000	65.000000	3000.000000
25%	35.250000	6.400000	6.000000	45.000000	4.000000	68.000000	5600.000000
50%	43.000000	7.200000	7.000000	60.000000	5.000000	70.000000	7000.000000
75%	50.000000	7.800000	8.000000	75.000000	7.000000	72.000000	8000.000000
max	59.000000	8.500000	9.000000	90.000000	8.000000	86.000000	10000.000000

In [5]:
```
#Age Analysis
data['Gender'].value_counts()
```

Out[5]:
```
Gender
Male      189
Female    185
Name: count, dtype: int64
```

In [6]:
```
data.Age.mean()
```

Out[6]:
```
42.18449197860963
```

In [7]:
```
data.Age.std()
```

Out[7]:
```
8.673133465547243
```

In [8]:
```
data.loc[data['Age']<=40, 'Age_group']= 'Young_Adult'
data.loc[data['Age']>=41, 'Age_group']= 'Older_Adult'
```

In [9]:
```
data
```

Out[9]:

Person ID	Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	BMI Category	Blood Pressure	Heart Rate	Daily Steps	Dis
1	Male	27	Software Engineer	6.1	6	42	6	Overweight	126/83	77	4200	
2	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	
3	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	
4	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	
5	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	
...	...	...	...	...	...	...	...	...	...	...	...	...

370	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000
371	Female	59	Nurse	8.0	9	75	3	Overweight	140/95	68	7000
372	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000
373	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000
374	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000

374 rows × 13 columns

In [10]:
```
data.Age_group.value_counts()
```

Out[10]:
```
Age_group
Older_Adult    209
Young_Adult    165
Name: count, dtype: int64
```

In [11]:
```
gms= data.groupby('Gender')['Age'].agg(['mean','std'])
gms
```

Out[11]:

Gender	mean	std
Female	47.405405	8.093407
Male	37.074074	5.662006

In [12]:
```
#Analyzing quality of sleep by occupation
```

In [13]:
```
data.groupby('Occupation')['Gender'].value_counts()
```

Out[13]:
```
Occupation            Gender
Accountant            Female    36
                      Male       1
Doctor                Male      69
                      Female     2
Engineer              Female    32
                      Male      31
Lawyer                Male      45
                      Female     2
Manager               Female     1
Nurse                 Female    73
Sales Representative  Male       2
Salesperson           Male      32
Scientist             Female     4
Software Engineer     Male       4
Teacher               Female    35
                      Male       5
Name: count, dtype: int64
```

In [14]:
```
a=data['Sleep Duration'].min()
b=data['Sleep Duration'].max()
print("Minimum sleep hours-",a,"-","Maximum sleep hours-",b)
```

```
Minimum sleep hours- 5.8 - Maximum sleep hours- 8.5
```

In [15]:
```
data.groupby('Gender')[['Sleep Duration', 'Quality of Sleep','Heart Rate']].mean()
```

Out[15]:

Gender	Sleep Duration	Quality of Sleep	Heart Rate
Female	7.229730	7.664865	69.259459
Male	7.036508	6.968254	71.052910

In [16]:
```
quabyoccu=data.groupby('Occupation')[['Sleep Duration', 'Quality of Sleep',]].mean()
quabyoccu.sort_values(by='Quality of Sleep')
```

Out[16]:

Occupation	Sleep Duration	Quality of Sleep
Sales Representative	5.900000	4.000000
Scientist	6.000000	5.000000
Salesperson	6.403125	6.000000
Software Engineer	6.750000	6.500000
Doctor	6.970423	6.647887
Teacher	6.690000	6.975000
Manager	6.900000	7.000000
Nurse	7.063014	7.369863
Accountant	7.113514	7.891892
Lawyer	7.410638	7.893617
Engineer	7.987302	8.412698

In [ ]:
```
#BODY MASS INDEX BY OCCUPATION
```

In [35]:
```
occ_bmi=data.groupby(['Occupation','BMI Category'])['BMI Category'].value_counts()
occ_bmi.sort_values()
```

Out[35]:
```
Occupation            BMI Category
Manager               Overweight        1
Teacher               Obese             1
Software Engineer     Overweight        1
Lawyer                Normal Weight     1
Software Engineer     Obese             1
Lawyer                Overweight        2
Software Engineer     Normal Weight     2
Doctor                Normal Weight     2
Lawyer                Obese             2
Sales Representative  Obese             2
Engineer              Overweight        3
Doctor                Obese             4
Engineer              Normal Weight     4
Scientist             Overweight        4
Accountant            Normal Weight     5
Teacher               Normal            6
Accountant            Overweight        6
Nurse                 Normal Weight     7
Accountant            Normal           26
Salesperson           Overweight       32
Teacher               Overweight       33
Lawyer                Normal           42
Engineer              Normal           56
```

```
Doctor                  Normal          65
Nurse                   Overweight      66
Name: count, dtype: int64
```

In [ ]:
```
#worse quality of sleep by occupation
```

In [41]:
```
occ_steps=data.groupby('Occupation')[['Daily Steps','Physical Activity Level','Sleep Dur
occ_steps.sort_values(by='Daily Steps', ascending = False)
```

Out[41]:

Occupation	Daily Steps	Physical Activity Level	Sleep Duration	Quality of Sleep
Nurse	8057.534247	78.589041	7.063014	7.369863
Lawyer	7661.702128	70.425532	7.410638	7.893617
Accountant	6881.081081	58.108108	7.113514	7.891892
Doctor	6808.450704	55.352113	6.970423	6.647887
Salesperson	6000.000000	45.000000	6.403125	6.000000
Engineer	5980.952381	51.857143	7.987302	8.412698
Teacher	5957.500000	45.625000	6.690000	6.975000
Software Engineer	5800.000000	48.000000	6.750000	6.500000
Manager	5500.000000	55.000000	6.900000	7.000000
Scientist	5350.000000	41.000000	6.000000	5.000000
Sales Representative	3000.000000	30.000000	5.900000	4.000000

In [ ]:

In [ ]:

In [ ]:

In [ ]: