```
In [1]: import pandas as pd
In [2]: data = pd.read_csv('.//sleephealth.csv', index_col='Person ID')
    data.head(10)
```

Out[2]:

Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	BMI Category	Blood Pressure	Heart Rate	Daily Steps	Dis
Male	27	Software Engineer	6.1	6	42	6	Overweight	126/83	77	4200	
Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	
Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	
Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	ı
Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	,
Male	28	Software Engineer	5.9	4	30	8	Obese	140/90	85	3000	Ins
Male	29	Teacher	6.3	6	40	7	Obese	140/90	82	3500	Ins
Male	29	Doctor	7.8	7	75	6	Normal	120/80	70	8000	
Male	29	Doctor	7.8	7	75	6	Normal	120/80	70	8000	
Male	29	Doctor	7.8	7	75	6	Normal	120/80	70	8000	
	Male Male Male Male Male Male Male Male	Male 28 Male 28 Male 28 Male 28 Male 28 Male 29 Male 29 Male 29 Male 29	Male 27 Software Engineer Male 28 Doctor Male 28 Doctor Male 28 Sales Representative Male 28 Software Engineer Male 28 Teacher Male 29 Doctor Male 29 Doctor	Male28Doctor Engineer6.1Male28Doctor6.2Male28Doctor6.2Male28Doctor6.2Male28Sales Representative5.9Male28Software Engineer5.9Male28Software Engineer5.9Male29Teacher6.3Male29Doctor7.8Male29Doctor7.8	Gender Age Occupation Sleep Duration of Sleep Puration Male 27 Software Engineer 6.1 6 Male 28 Doctor 6.2 6 Male 28 Doctor 6.2 6 Male 28 Sales Representative 5.9 4 Male 28 Software Engineer 5.9 4 Male 29 Teacher 6.3 6 Male 29 Doctor 7.8 7 Male 29 Doctor 7.8 7 Male 29 Doctor 7.8 7	Gender Age Occupation Sleep Duration of Sleep Sleep Activity Level Male 27 Software Engineer 6.1 6 42 Male 28 Doctor 6.2 6 60 Male 28 Doctor 6.2 6 60 Male 28 Sales Sales Representative 5.9 4 30 Male 28 Software Engineer 5.9 4 30 Male 29 Teacher 6.3 6 40 Male 29 Doctor 7.8 7 75 Male 29 Doctor 7.8 7 75	Gender Age Occupation Sleep Duration of Sleep Sleep Activity Level Stress Level Male 27 Software Engineer 6.1 6 42 6 Male 28 Doctor 6.2 6 60 8 Male 28 Doctor 6.2 6 60 8 Male 28 Sales Sales Representative 5.9 4 30 8 Male 28 Software Engineer 5.9 4 30 8 Male 29 Teacher 6.3 6 40 7 Male 29 Doctor 7.8 7 75 6 Male 29 Doctor 7.8 7 75 6	Gender Male Age Level Occupation Duration Duration Sleep Sleep Sleep Sleep Sleep Level Activity Level Level Sleep Level Stress Shift Shift Sleep Sleep Level Male 27 Software Engineer 6.1 6 42 6 Overweight Male 28 Doctor 6.2 6 60 8 Normal Male 28 Doctor 6.2 6 60 8 Normal Male 28 Sales Sales Representative 5.9 4 30 8 Obese Male 28 Software Engineer 5.9 4 30 8 Obese Male 29 Teacher 6.3 6 40 7 Obese Male 29 Doctor 7.8 7 75 6 Normal Male 29 Doctor 7.8 7 75 6 Normal	Gender Male Age Image: Age Male Image: Age Male Male Male Male Male Male Male Mal	Gender Age Occupation Sileep Duration Activity Level Level Sites Level Level Bill Level Level Level Bill Bill Bill Bill Bill Bill Bill Bill	Gender Age Occupation Distriction of Sleep Activity Level Trees Bill Indicates Activity Indicates Bill Indicates Bill Indicates Activity Indicates Bill Indicates Activity Ind

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()
```

```
Data columns (total 12 columns):

# Column Non-Null Count Dtype
--- --- O 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 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
```

<class 'pandas.core.frame.DataFrame'>

Index: 374 entries, 1 to 374

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
374.000000	374.000000	374.000000	374.000000	374.000000	374.000000	374.000000
42.184492	7.132086	7.312834	59.171123	5.385027	70.165775	6816.844920
8.673133	0.795657	1.196956	20.830804	1.774526	4.135676	1617.915679
27.000000	5.800000	4.000000	30.000000	3.000000	65.000000	3000.000000
35.250000	6.400000	6.000000	45.000000	4.000000	68.000000	5600.000000
43.000000	7.200000	7.000000	60.000000	5.000000	70.000000	7000.000000
50.000000	7.800000	8.000000	75.000000	7.000000	72.000000	8000.000000
59.000000	8.500000	9.000000	90.000000	8.000000	86.000000	10000.000000
	374.000000 42.184492 8.673133 27.000000 35.250000 43.000000 50.0000000	Age Duration 374.000000 374.000000 42.184492 7.132086 8.673133 0.795657 27.000000 5.800000 35.250000 6.400000 43.000000 7.200000 50.000000 7.800000	Age Duration Sleep 374.000000 374.000000 374.000000 42.184492 7.132086 7.312834 8.673133 0.795657 1.196956 27.000000 5.800000 4.000000 35.250000 6.400000 6.000000 43.000000 7.200000 7.000000 50.000000 7.800000 8.000000	Age Duration Sleep Level 374.000000 374.000000 374.000000 42.184492 7.132086 7.312834 59.171123 8.673133 0.795657 1.196956 20.830804 27.000000 5.800000 4.000000 30.000000 35.250000 6.400000 6.000000 45.000000 43.000000 7.200000 7.000000 75.000000	Age Duration Sleep Level Level 374.000000 374.000000 374.000000 374.000000 42.184492 7.132086 7.312834 59.171123 5.385027 8.673133 0.795657 1.196956 20.830804 1.774526 27.000000 5.800000 4.000000 30.000000 3.000000 35.250000 6.400000 6.000000 45.000000 4.000000 43.000000 7.200000 7.000000 75.000000 7.000000	Age Duration Sleep Level Level Heart Rate 374.000000 374.000000 374.000000 374.000000 374.000000 374.000000 42.184492 7.132086 7.312834 59.171123 5.385027 70.165775 8.673133 0.795657 1.196956 20.830804 1.774526 4.135676 27.000000 5.800000 4.000000 30.000000 3.000000 65.000000 35.250000 6.400000 6.000000 45.000000 4.000000 68.000000 43.000000 7.200000 7.000000 75.000000 7.000000 72.000000

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]:

•		Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	BMI Category	Blood Pressure	Heart Rate	Daily Steps	Dis
	Person ID												
	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	J
	5	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	1
	•••												

```
371
                Female
                         59
                                  Nurse
                                             8.0
                                                             75
                                                                    3 Overweight
                                                                                  140/95
                                                                                           68
                                                                                                7000
                Female
                         59
                                             8.1
                                                      9
                                                             75
                                                                    3 Overweight
                                                                                  140/95
                                                                                                7000
           372
                                  Nurse
                                                                                           68
           373
                Female
                         59
                                  Nurse
                                             8.1
                                                      9
                                                             75
                                                                    3 Overweight
                                                                                  140/95
                                                                                           68
                                                                                                7000
                                                      9
                         59
                                             8.1
                                                             75
                                                                    3 Overweight
                                                                                                7000
           374
                Female
                                  Nurse
                                                                                  140/95
                                                                                           68
        374 rows × 13 columns
         data.Age group.value counts()
In [10]:
         Age_group
Out[10]:
         Older Adult
                         209
         Young_Adult
                         165
         Name: count, dtype: int64
In [11]:
         gms= data.groupby('Gender')['Age'].agg(['mean','std'])
Out[11]:
                   mean
                             std
         Gender
         Female 47.405405 8.093407
           Male 37.074074 5.662006
         #Analyzing quality of sleep by occupation
In [12]:
         data.groupby('Occupation')['Gender'].value counts()
In [13]:
                                Gender
         Occupation
Out[13]:
         Accountant
                                Female
                                           36
                                Male
                                           1
         Doctor
                                Male
                                           69
                                Female
                                          2
         Engineer
                                Female
                                           32
                                Male
                                           31
                                          45
         Lawyer
                                Male
                                Female
                                           2
         Manager
                                Female
                                           1
                                Female
                                          73
         Nurse
         Sales Representative Male
                                           2
                                           32
         Salesperson
                                Male
                                          4
         Scientist
                                Female
         Software Engineer
                                Male
                                           4
         Teacher
                                Female
                                           35
                                Male
                                            5
         Name: count, dtype: int64
         a=data['Sleep Duration'].min()
In [14]:
         b=data['Sleep Duration'].max()
         print("Minimum sleep hours-",a,"-","Maximum sleep hours-",b)
         Minimum sleep hours- 5.8 - Maximum sleep hours- 8.5
         data.groupby('Gender')[['Sleep Duration', 'Quality of Sleep','Heart Rate']].mean()
In [15]:
```

370 Female

59

Nurse

8.1

9

75

3 Overweight

140/95

68

7000

Out[15]: Sleep Duration Quality of Sleep Heart Rate Gender **Female** 7.229730 7.664865 69.259459 Male 7.036508 6.968254 71.052910 quabyoccu=data.groupby('Occupation')[['Sleep Duration', 'Quality of Sleep',]].mean() In [16]: quabyoccu.sort values(by='Quality of Sleep') Out[16]: Sleep Duration Quality of Sleep Occupation **Sales Representative** 5.900000 4.000000 Scientist 6.000000 5.000000 6.403125 6.000000 Salesperson **Software Engineer** 6.750000 6.500000

6.970423

6.690000

6.900000

7.063014

7.113514

7.410638

7.987302

Doctor

Teacher

Manager

Accountant

Nurse

Lawyer

Engineer

In []: #BODY MASS INDEX BY OCCUPATION

56

6.647887

6.975000

7.000000

7.369863

7.891892

7.893617

8.412698

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

Normal

Engineer

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]: Daily Steps Physical Activity Level Sleep Duration Quality of Sleep

Occupation				
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	[]:	