In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
data=pd.read_csv("Sleep_health_and_lifestyle_dataset[1].csv")
data
```

Out[1]:

	Person ID	Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	BMI Category	Blood Pressure	Heart Rate	Daily Steps	Sleep Disorder
0	1	Male	27	Software Engineer	6.1	6	42	6	Overweight	126/83	77	4200	NaN
1	2	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	NaN
2	3	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	NaN
3	4	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
4	5	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
...	...	...	...	...	...	...	...	...	...	...	...	...	...
369	370	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
370	371	Female	59	Nurse	8.0	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
371	372	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
372	373	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
373	374	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea

374 rows × 13 columns

In [3]:
```
data.shape
```

Out[3]:
```
(374, 13)
```

In [5]:
```
data.columns
```

Out[5]:
```
Index(['Person ID', 'Gender', 'Age', 'Occupation', 'Sleep Duration',
       'Quality of Sleep', 'Physical Activity Level', 'Stress Level',
       'BMI Category', 'Blood Pressure', 'Heart Rate', 'Daily Steps',
       'Sleep Disorder'],
      dtype='object')
```

In [7]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 374 entries, 0 to 373
Data columns (total 13 columns):
 #   Column                   Non-Null Count  Dtype  
---  ------                   --------------  -----  
 0   Person ID                374 non-null    int64  
 1   Gender                   374 non-null    object 
 2   Age                      374 non-null    int64  
 3   Occupation               374 non-null    object 
 4   Sleep Duration           374 non-null    float64
 5   Quality of Sleep         374 non-null    int64  
 6   Physical Activity Level  374 non-null    int64  
 7   Stress Level             374 non-null    int64  
 8   BMI Category             374 non-null    object 
 9   Blood Pressure           374 non-null    object 
 10  Heart Rate               374 non-null    int64  
 11  Daily Steps              374 non-null    int64  
 12  Sleep Disorder           155 non-null    object 
dtypes: float64(1), int64(7), object(5)
memory usage: 38.1+ KB
```

In [9]:
```
data.isnull().sum()
```

Out[9]:
```
Person ID                    0
Gender                       0
Age                          0
Occupation                   0
Sleep Duration               0
Quality of Sleep             0
Physical Activity Level      0
Stress Level                 0
BMI Category                 0
Blood Pressure               0
Heart Rate                   0
Daily Steps                  0
Sleep Disorder             219
dtype: int64
```

In [11]:
```
data.fillna("Normal",inplace=True)
data
```

Out[11]:

	Person ID	Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	BMI Category	Blood Pressure	Heart Rate	Daily Steps	Sleep Disorder
0	1	Male	27	Software Engineer	6.1	6	42	6	Overweight	126/83	77	4200	Normal
1	2	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	Normal
2	3	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	Normal
3	4	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
4	5	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
...	...	...	...	...	...	...	...	...	...	...	...	...	...
369	370	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
370	371	Female	59	Nurse	8.0	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
371	372	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
372	373	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
373	374	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea

374 rows × 13 columns

In [13]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 374 entries, 0 to 373
Data columns (total 13 columns):
 #   Column                   Non-Null Count  Dtype  
---  ------                   --------------  -----  
 0   Person ID                374 non-null    int64  
 1   Gender                   374 non-null    object 
 2   Age                      374 non-null    int64  
 3   Occupation               374 non-null    object 
 4   Sleep Duration           374 non-null    float64
 5   Quality of Sleep         374 non-null    int64  
 6   Physical Activity Level  374 non-null    int64  
 7   Stress Level             374 non-null    int64  
 8   BMI Category             374 non-null    object 
 9   Blood Pressure           374 non-null    object 
 10  Heart Rate               374 non-null    int64  
 11  Daily Steps              374 non-null    int64  
 12  Sleep Disorder           374 non-null    object 
dtypes: float64(1), int64(7), object(5)
memory usage: 38.1+ KB
```

In [15]:
```
data.head()
```

Out[15]:

	Person ID	Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	BMI Category	Blood Pressure	Heart Rate	Daily Steps	Sleep Disorder
0	1	Male	27	Software Engineer	6.1	6	42	6	Overweight	126/83	77	4200	Normal
1	2	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	Normal
2	3	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	Normal
3	4	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
4	5	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea

In [17]:
```
data.tail()
```

Out[17]:

	Person ID	Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	BMI Category	Blood Pressure	Heart Rate	Daily Steps	Sleep Disorder
369	370	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
370	371	Female	59	Nurse	8.0	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
371	372	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
372	373	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
373	374	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea

In [19]:
```
data.rename(columns={"Person ID":"ID","Physical Activity Level":"Physical Activity in min","BMI Category":"BMI"
data
```

Out[19]:

	ID	Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity in min	Stress Level	BMI	BP	Heart Rate	Daily Steps	Sleep Disorder
0	1	Male	27	Software Engineer	6.1	6	42	6	Overweight	126/83	77	4200	Normal
1	2	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	Normal
2	3	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	Normal
3	4	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
4	5	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
...	...	...	...	...	...	...	...	...	...	...	...	...	...
369	370	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
370	371	Female	59	Nurse	8.0	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
371	372	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
372	373	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
373	374	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea

374 rows × 13 columns

In [21]:
```
data.set_index("ID",inplace=True)
data
```

Out[21]:

ID	Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity in min	Stress Level	BMI	BP	Heart Rate	Daily Steps	Sleep Disorder
1	Male	27	Software Engineer	6.1	6	42	6	Overweight	126/83	77	4200	Normal
2	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	Normal
3	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	Normal
4	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
5	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
...	...	...	...	...	...	...	...	...	...	...	...	...
370	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
371	Female	59	Nurse	8.0	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
372	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
373	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
374	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea

374 rows × 12 columns

In [23]:
```
data[["Gender","Age","Occupation","Sleep Duration","Quality of Sleep","Physical Activity in min","Stress Level"
```

Out[23]:
```
Gender                     2
Age                       31
Occupation                11
Sleep Duration            27
Quality of Sleep           6
Physical Activity in min  16
Stress Level               6
BMI                        4
BP                        25
Heart Rate                19
Daily Steps               20
Sleep Disorder             3
dtype: int64
```

In [25]:
```
#gender
data["Gender"].value_counts()
```

Out[25]:
```
Gender
Male      189
Female    185
Name: count, dtype: int64
```

In [27]:
```
# age categary
data["Age"].value_counts()
```

Out[27]:
```
Age
43    34
44    30
37    20
38    20
50    20
31    18
32    17
53    17
59    16
39    15
45    14
33    13
30    13
29    13
35    12
36    12
41    12
49    11
57     9
52     9
42     9
51     8
54     7
58     6
28     5
40     4
48     3
55     2
56     2
34     2
27     1
Name: count, dtype: int64
```

In [29]:
```
# occapation categary
data["Occupation"].value_counts()
```

Out[29]:
```
Occupation
Nurse                   73
Doctor                  71
Engineer                63
Lawyer                  47
Teacher                 40
Accountant              37
Salesperson             32
Software Engineer        4
Scientist                4
Sales Representative     2
Manager                  1
Name: count, dtype: int64
```

In [31]:
```
data["Sleep Duration"].value_counts()
```

Out[31]:
```
Sleep Duration
7.2    36
6.0    31
7.8    28
6.5    26
6.1    25
7.7    24
6.6    20
7.1    19
8.1    15
7.3    14
8.4    14
6.3    13
8.5    13
8.0    13
6.2    12
8.2    11
7.6    10
6.4     9
7.9     7
7.4     5
6.7     5
7.5     5
6.8     5
8.3     5
5.9     4
6.9     3
5.8     2
Name: count, dtype: int64
```

In [33]:
```
data["Quality of Sleep"].value_counts()
```

Out[33]:
```
Quality of Sleep
8    109
6    105
7     77
9     71
5      7
4      5
Name: count, dtype: int64
```

In [35]:
```
data.groupby("Quality of Sleep")["Sleep Duration"].value_counts()
```

Out[35]:
```
Quality of Sleep  Sleep Duration
4                 5.9               3
                  5.8               2
5                 6.5               3
                  6.4               2
                  6.6               2
6                 6.0              31
                  6.1              25
                  6.3              13
                  6.5              13
                  6.2              12
                  6.4               7
                  6.8               2
                  5.9               1
                  6.9               1
7                 6.6              18
                  7.7              14
                  6.5              10
                  7.8               9
                  6.7               5
                  7.6               5
                  7.9               4
                  6.8               3
                  7.1               3
                  6.9               2
                  7.3               2
                  7.4               2
8                 7.2              36
                  7.8              19
                  7.1              16
                  7.3              12
                  7.7              10
                  7.5               5
                  7.6               5
                  7.4               3
                  7.9               3
9                 8.1              15
                  8.4              14
                  8.0              13
                  8.5              13
                  8.2              11
                  8.3               5
Name: count, dtype: int64
```

In [39]:
```
data["Physical Activity in min"].value_counts()
```

Out[39]:
```
Physical Activity in min
60    70
30    68
45    68
75    67
90    67
40     6
55     6
35     4
50     4
70     3
42     2
32     2
80     2
65     2
85     2
47     1
Name: count, dtype: int64
```

In [41]:
```
data["Stress Level"].value_counts()
```

Out[41]:
```
Stress Level
3    71
8    70
4    70
5    67
7    50
6    46
Name: count, dtype: int64
```

In [43]:
```
data["BMI"].value_counts()
```

Out[43]:
```
BMI
Normal           195
Overweight       148
Normal Weight     21
Obese             10
Name: count, dtype: int64
```

In [45]:
```
data["BP"].value_counts()
```

Out[45]:
```
BP
130/85    99
140/95    65
125/80    65
120/80    45
115/75    32
135/90    27
140/90     4
125/82     4
132/87     3
128/85     3
126/83     2
115/78     2
139/91     2
142/92     2
119/77     2
135/88     2
129/84     2
128/84     2
131/86     2
117/76     2
130/86     2
118/75     2
121/79     1
122/80     1
118/76     1
Name: count, dtype: int64
```

In [47]:
```
data["Heart Rate"].value_counts()
```

Out[47]:
```
Heart Rate
68    94
70    76
72    69
65    67
75    36
78     5
85     3
80     3
84     2
83     2
73     2
67     2
74     2
77     2
81     2
76     2
69     2
86     2
82     1
Name: count, dtype: int64
```

In [49]:
```
data["Daily Steps"].value_counts()
```

Out[49]:
```
Daily Steps
8000     101
6000      68
5000      68
7000      66
10000     36
5500       4
3000       3
3500       3
4000       3
6800       3
4800       2
7300       2
7500       2
4200       2
3300       2
5600       2
5200       2
4100       2
3700       2
6200       1
Name: count, dtype: int64
```

In [51]:
```
data["Sleep Disorder"].value_counts()
```

Out[51]:
```
Sleep Disorder
Normal         219
Sleep Apnea     78
Insomnia        77
Name: count, dtype: int64
```

in this data there are

Gender 2 (Male -189 Female -185, total-374)

Age 31

Occupation 11 (Nurse 73

```
Doctor                 71
Engineer               63
Lawyer                 47
Teacher                40
Accountant             37
Salesperson            32
Software Engineer       4
Scientist               4
Sales Representative    2
Manager                 1)
```

Sleep Duration 27

Quality of Sleep 6 (normal 7-9 hours)

Physical Activity 16

in min

Stress Level 6 (normal >5)

BMI 4 (Normal 195

```
Overweight       148
Normal Weight     21
Obese             10)
```

BP 25 (normal 120/80)

Heart Rate 19 (normal 60-100 per min)

Daily Steps 20 (normal >5000)

Sleep Disorder 3 (Normal 219

```
Sleep Apnea      78
Insomnia         77)
```

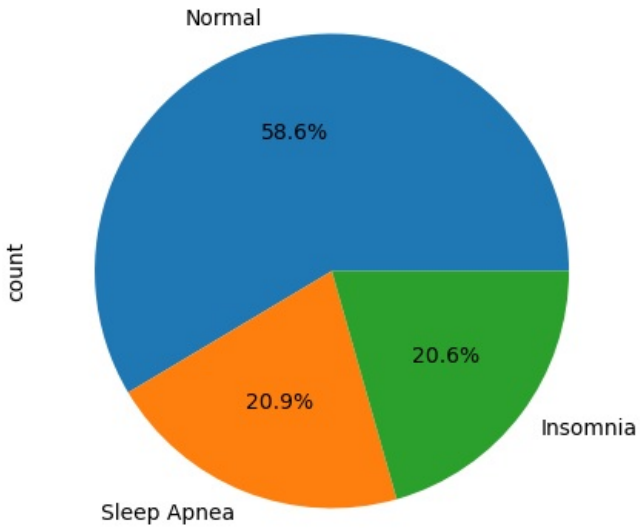
In [53]:
```
data.describe()
```

Out[53]:

	Age	Sleep Duration	Quality of Sleep	Physical Activity in min	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 [55]:
```
#percentage of sleep disorder and normal people with pie chart
plt.figure(figsize=(10,5))
data["Sleep Disorder"].value_counts().plot.pie(autopct="%.1f%%")
```

Out[55]:
```
<Axes: ylabel='count'>
```

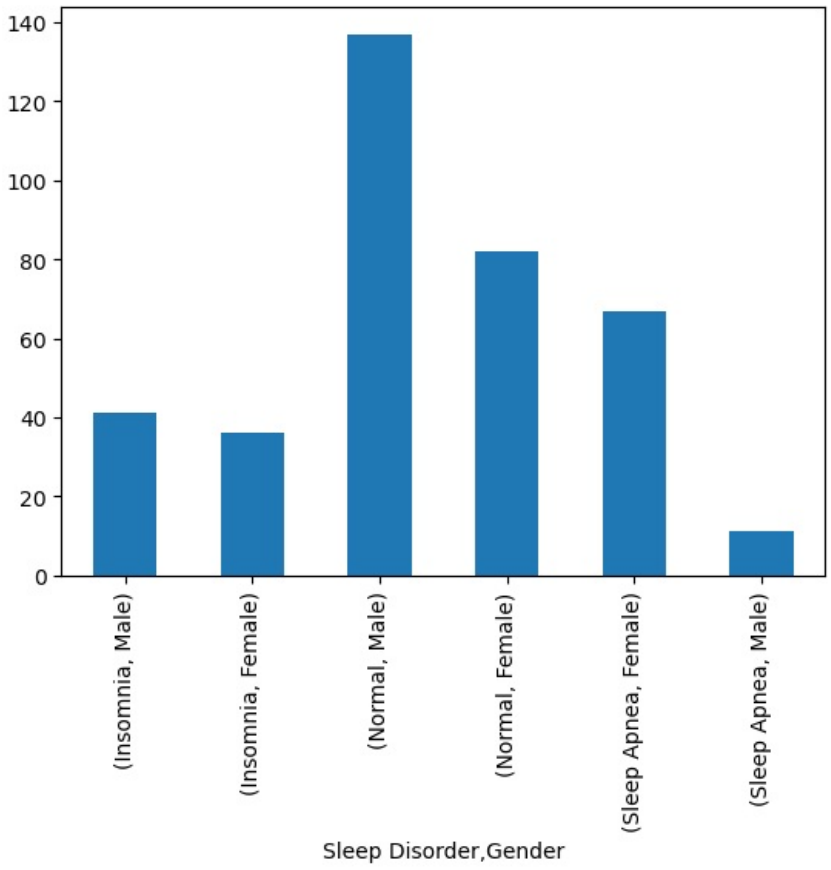


In [57]:
```
# gender wise affected
gender=data.groupby("Gender")["Sleep Disorder"].value_counts()
gender
```

Out[57]:
```
Gender  Sleep Disorder
Female  Normal             82
        Sleep Apnea        67
        Insomnia           36
Male    Normal            137
        Insomnia           41
        Sleep Apnea        11
Name: count, dtype: int64
```

In [59]:
```
#gender wise affected people and gender wise non affected people with bar
```

```
gender=data.groupby("Sleep Disorder")["Gender"].value_counts().plot(kind="bar")
```



In [65]:
```
# age wise affected people
age=data.groupby("Sleep Disorder")["Age"].agg(["max","min"])
age
```

Out[65]:

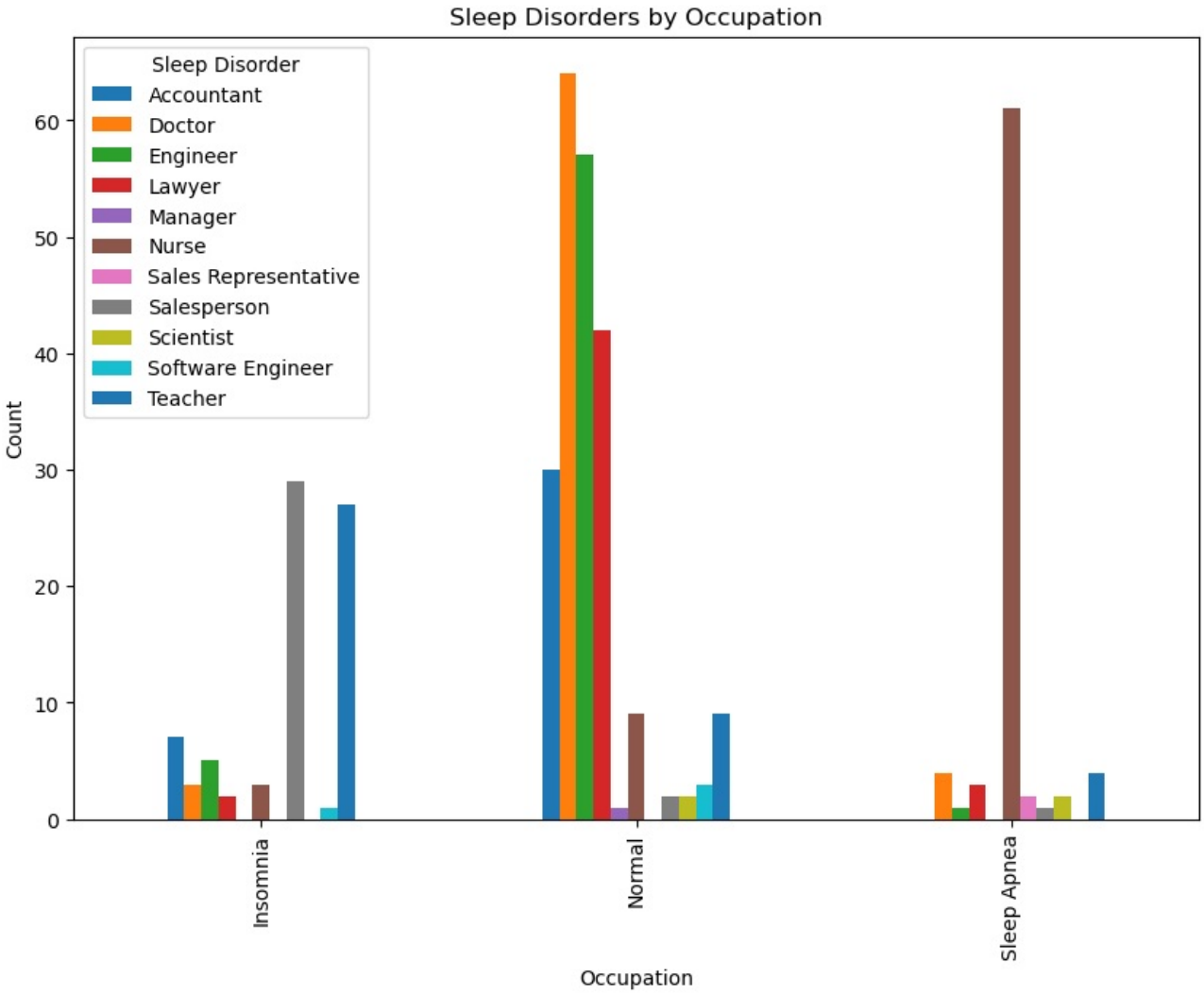
Sleep Disorder	max	min
Insomnia	53	28
Normal	59	27
Sleep Apnea	59	28

In [67]:
```
Occupation=data.groupby("Sleep Disorder")["Occupation"].value_counts().unstack(fill_value=0)
Occupation
```

Out[67]:

Occupation	Accountant	Doctor	Engineer	Lawyer	Manager	Nurse	Sales Representative	Salesperson	Scientist	Software Engineer	Teacher
Sleep Disorder											
Insomnia	7	3	5	2	0	3	0	29	0	1	27
Normal	30	64	57	42	1	9	0	2	2	3	9
Sleep Apnea	0	4	1	3	0	61	2	1	2	0	4

In [69]:
```
# with bar
Occupation.plot(kind='bar', figsize=(10,7))
plt.title("Sleep Disorders by Occupation")
plt.xlabel("Occupation")
plt.ylabel("Count")
plt.legend(title="Sleep Disorder")
plt.show()
```



In [73]:
```
# health analysis
data[["Quality of Sleep","Physical Activity in min","Stress Level","Heart Rate","Daily Steps"]].agg(["max","min
```

Out[73]:

	Quality of Sleep	Physical Activity in min	Stress Level	Heart Rate	Daily Steps
max	9.000000	90.000000	8.000000	86.000000	10000.00000
min	4.000000	30.000000	3.000000	65.000000	3000.00000
mean	7.312834	59.171123	5.385027	70.165775	6816.84492

In [77]:
```
#average bp
data["BP"].agg(["max","min"])
```

Out[77]:
```
max    142/92
min    115/75
Name: BP, dtype: object
```

In [83]:
```
Healthy=data.loc[(data["Quality of Sleep"]>=6) & (data["Physical Activity in min"]>=60) & (data["Stress Level"]
Healthy
```

Out[83]:

ID	Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity in min	Stress Level	BMI	BP	Heart Rate	Daily Steps	Sleep Disorder
86	Female	35	Accountant	7.2	8	60	4	Normal	115/75	68	7000	Normal
96	Female	36	Accountant	7.1	8	60	4	Normal	115/75	68	7000	Normal
97	Female	36	Accountant	7.2	8	60	4	Normal	115/75	68	7000	Normal
98	Female	36	Accountant	7.1	8	60	4	Normal	115/75	68	7000	Normal
99	Female	36	Teacher	7.1	8	60	4	Normal	115/75	68	7000	Normal
...	...	...	...	...	...	...	...	...	...	...	...	...
214	Male	43	Engineer	7.8	8	90	5	Normal	130/85	70	8000	Normal
215	Male	43	Engineer	7.8	8	90	5	Normal	130/85	70	8000	Normal
216	Male	43	Engineer	7.8	8	90	5	Normal	130/85	70	8000	Normal
217	Male	43	Engineer	7.8	8	90	5	Normal	130/85	70	8000	Normal
218	Male	43	Engineer	7.8	8	90	5	Normal	130/85	70	8000	Normal

82 rows × 12 columns

In [89]:
```
Healthy.describe()
```

Out[89]:

	Age	Sleep Duration	Quality of Sleep	Physical Activity in min	Stress Level	Heart Rate	Daily Steps
count	82.000000	82.000000	82.0	82.000000	82.000000	82.000000	82.000000
mean	39.121951	7.392683	8.0	70.609756	4.634146	68.658537	7634.146341
std	2.395419	0.278789	0.0	14.431410	0.484633	0.945679	484.632867
min	35.000000	7.100000	8.0	60.000000	4.000000	68.000000	7000.000000
25%	37.000000	7.200000	8.0	60.000000	4.000000	68.000000	7000.000000
50%	38.000000	7.200000	8.0	60.000000	5.000000	68.000000	8000.000000
75%	41.000000	7.700000	8.0	90.000000	5.000000	70.000000	8000.000000
max	43.000000	7.900000	8.0	90.000000	5.000000	70.000000	8000.000000

In [96]:
```
#healthy people in this people
Healthy.shape
```

Out[96]:
```
(82, 12)
```

suggession

here are the suggestion to be follow for the healthy life style

1. Sleep Duration should be 7-9 hours per day
2. Physical Activities atleast 60min per day
3. Stress Level should be less than 6
4. BMI should be Normal
5. Daily Step count should be above 6000 steps

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js