Sleep Health and Lifestyle Analysis

Project Overview

The project aims to analyze the relationship between sleep patterns, health metrics, and the prevalence of sleep disorders among individuals from diverse occupations and age groups. Key attributes such as sleep duration, quality of sleep, physical activity level, stress level, BMI category, blood pressure, heart rate, daily steps, and presence of sleep disorders will be examined.

Importing Libraries

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import datetime as dt
from numpy import nan as NA
```

```
sl_data = pd.read_csv('/content/13_Sleep Health and Lifestyle Analysis.csv')
sl_data
```

	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

```
sl_data.shape
```

```
(374, 13)
```

```
sl_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
```

```
sl_data.head(10)
```

	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
5	6	Male	28	Software Engineer	5.9	4	30	8	Obese	140/90	85	3000	Insomnia
6	7	Male	29	Teacher	6.3	6	40	7	Obese	140/90	82	3500	Insomnia

```
sl_data.columns
```

```
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')
```

Data Cleaning

```
sl_data.duplicated()
```

```
0      False
1      False
2      False
3      False
4      False
       ...
369    False
370    False
371    False
372    False
373    False
Length: 374, dtype: bool
```

Preparation and Processing Data

```
sl_data.isnull().sum()
```

```
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
```

```
for col in sl_data.describe(include='object').columns:
    print(col)
    print(sl_data[col].unique())
```

```
Gender
['Male' 'Female']
Occupation
['Software Engineer' 'Doctor' 'Sales Representative' 'Teacher' 'Nurse'
 'Engineer' 'Accountant' 'Scientist' 'Lawyer' 'Salesperson' 'Manager']
BMI Category
['Overweight' 'Normal' 'Obese' 'Normal Weight']
Blood Pressure
['126/83' '125/80' '140/90' '120/80' '132/87' '130/86' '117/76' '118/76'
 '128/85' '131/86' '128/84' '115/75' '135/88' '129/84' '130/85' '115/78'
 '119/77' '121/79' '125/82' '135/90' '122/80' '142/92' '140/95' '139/91'
 '118/75']
Sleep Disorder
[nan 'Sleep Apnea' 'Insomnia']
```

```
sl_data_encoded = pd.get_dummies(sl_data)
print (sl_data_encoded.head())
```

```
   Person ID  Age  Sleep Duration  Quality of Sleep  Physical Activity Level  \
0          1   27             6.1                 6                       42
1          2   28             6.2                 6                       60
2          3   28             6.2                 6                       60
3          4   28             5.9                 4                       30
4          5   28             5.9                 4                       30

   Stress Level  Heart Rate  Daily Steps  Gender_Female  Gender_Male  ...  \
0             6          77         4200          False         True  ...
1             8          75        10000          False         True  ...
2             8          75        10000          False         True  ...
3             8          85         3000          False         True  ...
4             8          85         3000          False         True  ...

   Blood Pressure_131/86  Blood Pressure_132/87  Blood Pressure_135/88  \
0                  False                  False                  False
1                  False                  False                  False
2                  False                  False                  False
3                  False                  False                  False
4                  False                  False                  False

   Blood Pressure_135/90  Blood Pressure_139/91  Blood Pressure_140/90  \
0                  False                  False                  False
1                  False                  False                  False
2                  False                  False                  False
3                  False                  False                   True
4                  False                  False                   True

   Blood Pressure_140/95  Blood Pressure_142/92  Sleep Disorder_Insomnia  \
0                  False                  False                    False
1                  False                  False                    False
2                  False                  False                    False
3                  False                  False                    False
4                  False                  False                    False

   Sleep Disorder_Sleep Apnea
0                       False
1                       False
2                       False
3                        True
4                        True

[5 rows x 52 columns]
```

```
sl_data.select_dtypes(include = 'object').nunique()
```

```
Gender             2
Occupation        11
BMI Category       4
Blood Pressure    25
Sleep Disorder     2
dtype: int64
```

```
sl_data.drop(columns= 'Person ID' , inplace=True)
sl_data.head()
```

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

```
sl_data['Sleep Duration'].mean()
```

```
7.132085561497325
```

```
sl_data.describe()
```

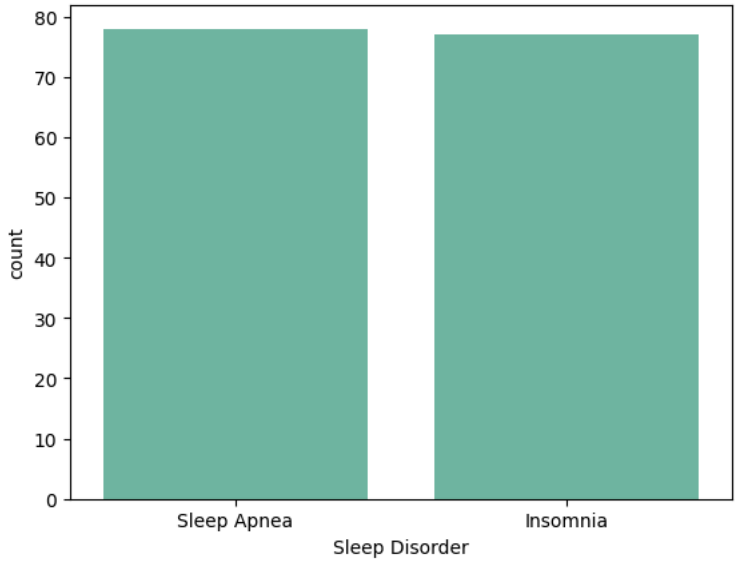
	Age	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	Heart Rate	Daily Steps
count	374.000000	374.000000	374.000000	374.000000	374.000000	374.000000	374.000000
mean	42.184492	7.132086	7.312834	59.171123	5.385027	70.165775	6816.844920
std	8.673133	0.795657	1.196956	20.830804	1.774526	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

Data Visualization

EDA.Measures of Sleep Disorder

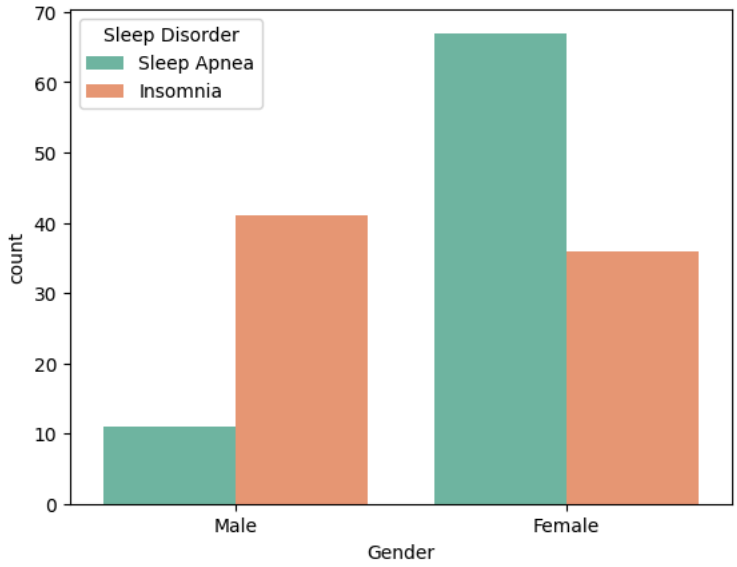
```
sns.set_palette(palette="Set2")

sns.countplot(x='Sleep Disorder', data=sl_data)
plt.show()
```

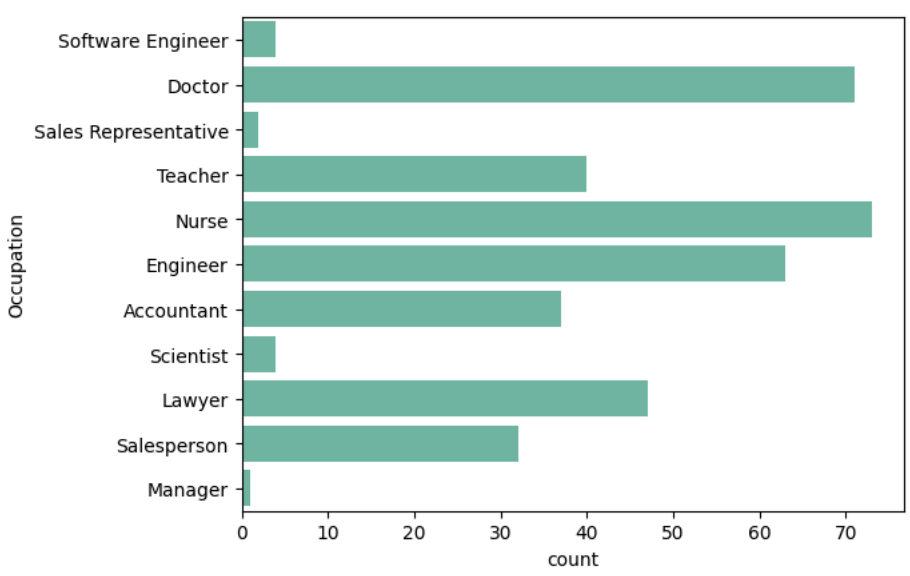


The chart above illustrates sleep disorder occurrences across genders, with males primarily showing no sleep disorder, whereas females predominantly suffer from sleep apnea.

```
sns.countplot(x='Gender', data = sl_data, hue='Sleep Disorder')
plt.show()
```

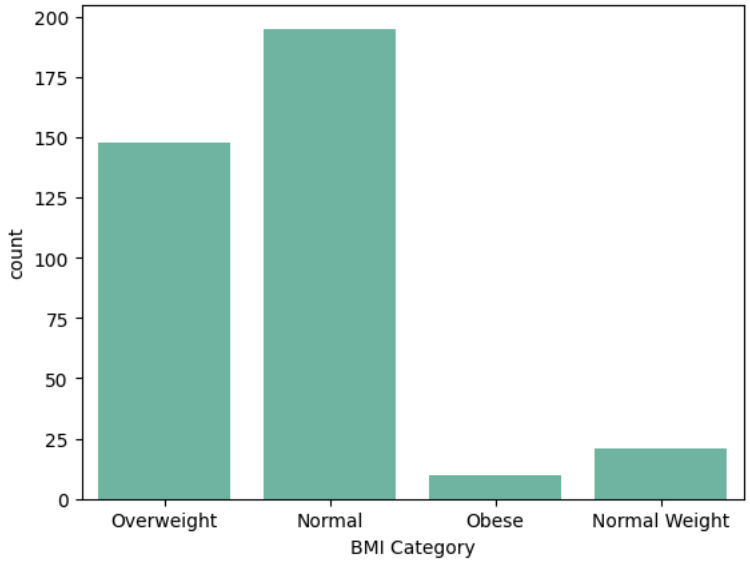


```
sns.countplot(y='Occupation' ,data=sl_data)
plt.show()
```



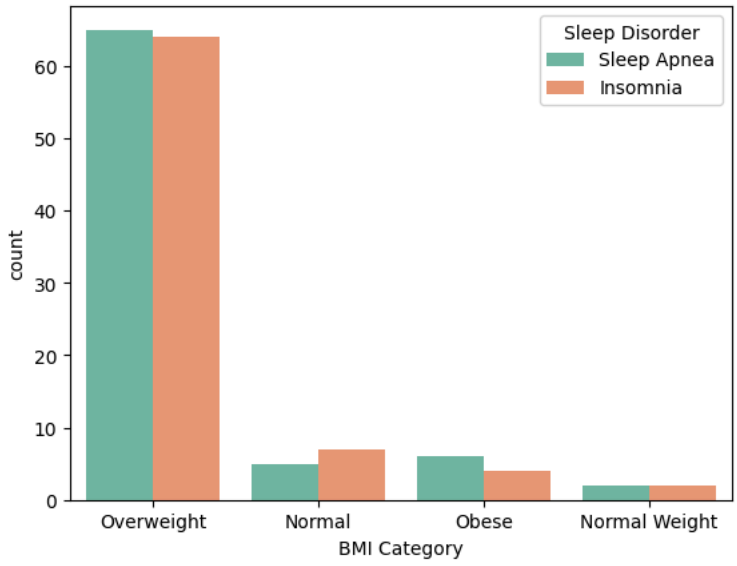

```
sns.countplot(x='BMI Category', data=sl_data, )
```

```
<Axes: xlabel='BMI Category', ylabel='count'>
```




```
sns.countplot(x='BMI Category', data=sl_data, hue='Sleep Disorder' )
```

```
<Axes: xlabel='BMI Category', ylabel='count'>
```

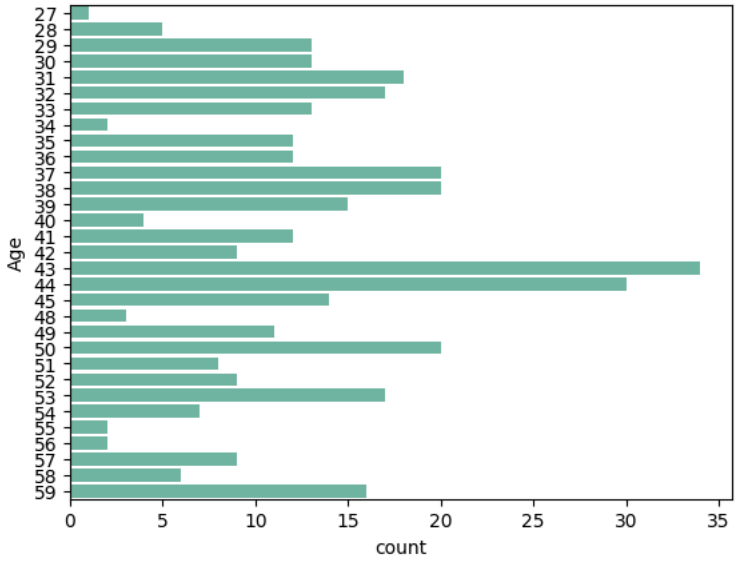


The majority of individuals classified as obese experience either sleep apnea or insomnia, while a significant portion of overweight individuals also suffer from some form of sleep disorder. Conversely, there is a notably lower prevalence of sleep disorders among individuals with a

normal BMI.

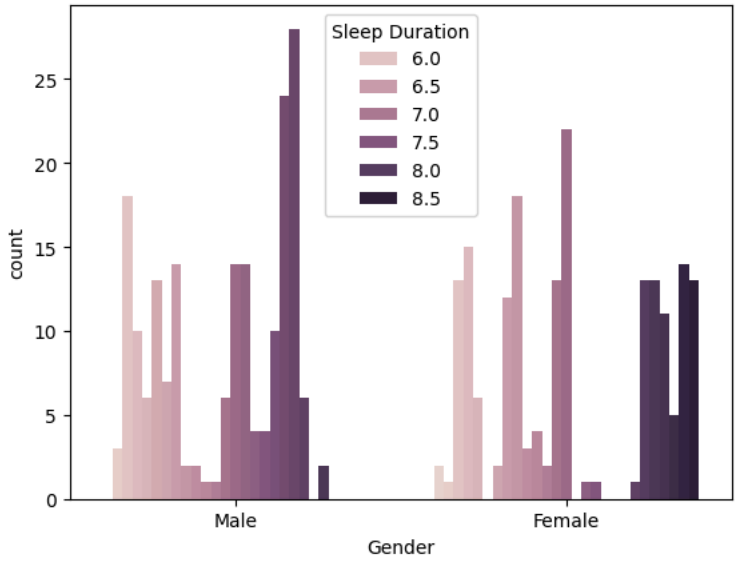
```
sns.countplot(y='Age', data=sl_data )
```

```
<Axes: xlabel='count', ylabel='Age'>
```




```
sns.countplot(x='Gender', data=sl_data, hue='Sleep Duration' )
```

```
<Axes: xlabel='Gender', ylabel='count'>
```

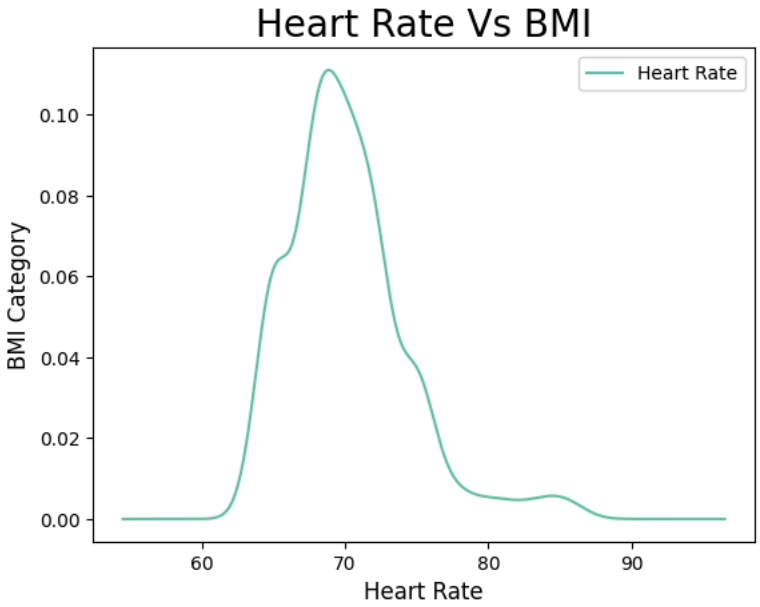



```
sl_data.plot(kind='scatter',x="Stress Level", y="Quality of Sleep",figsize=(12,4),fontsize=10)
plt.title("Relation between stress level and quality of sleep",fontsize=17)
plt.ylabel("Stress Level",fontsize=15)
plt.xlabel("Quality of Sleep",fontsize=15)
```

```
Text(0.5, 0, 'Quality of Sleep')
```

```
sl_data['Heart Rate'].plot(kind='kde')
plt.title('Heart Rate Vs BMI',fontsize=20)
plt.xlabel('Heart Rate',fontsize=12)
plt.ylabel('BMI Category',fontsize=12)
plt.legend()
```

```
<matplotlib.legend.Legend at 0x7a083c09df60>
```




```
sl_data['Daily Steps'].plot(kind='kde')
plt.title('The Relationship between Daily Steps and Blood pressure',fontsize=20)
plt.xlabel('Daily Steps',fontsize=12)
plt.ylabel('Blood Pressure',fontsize=12)
plt.legend()
```

```
<matplotlib.legend.Legend at 0x7a0836b97b80>
```