Stress Level By Sleep Duration

May 27, 2024

1 Raza Abbas 2k22/CSE/83

2 Assignment Stress Level By Sleep Duration

2.0.1 https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset

[1]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from pandas.core.common import random_state
from sklearn.linear_model import LinearRegression
```

[2]:

```
df=pd.read_csv('Sleep_health_and_lifestyle_dataset.csv')
df.head()
```

[2]:

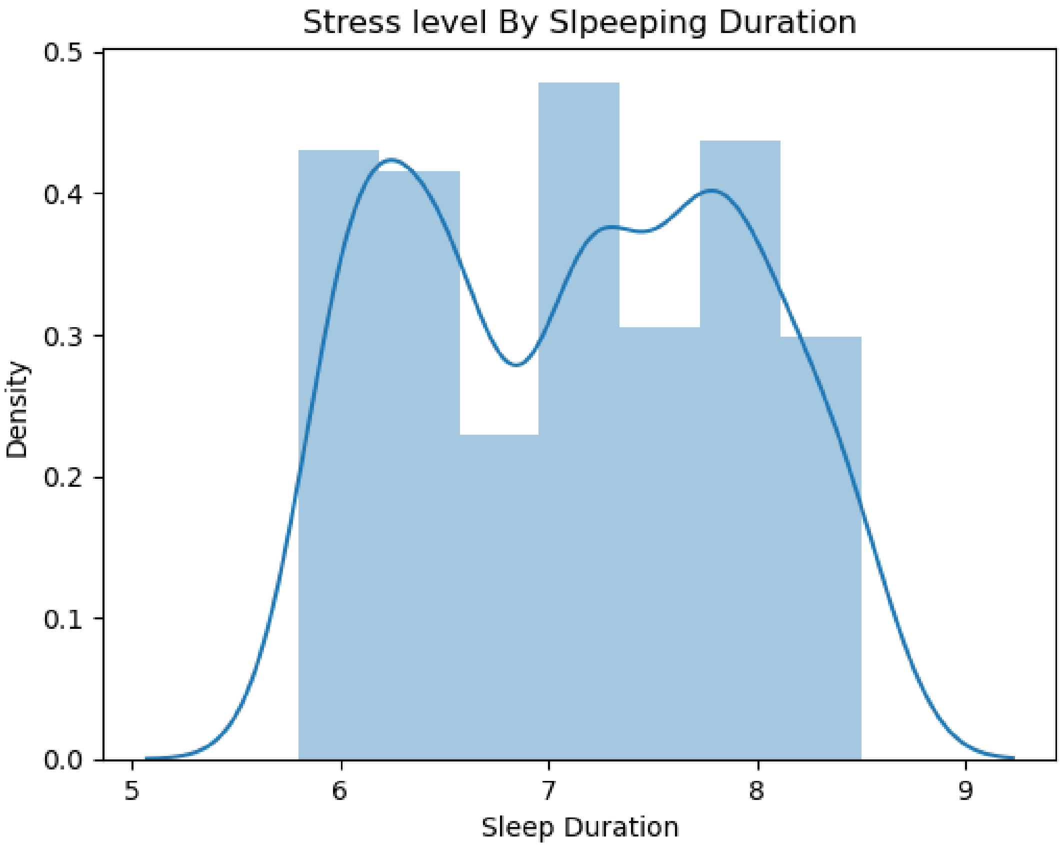
```
   Person ID Gender  Age            Occupation  Sleep Duration  \
0          1   Male   27     Software Engineer             6.1
1          2   Male   28                Doctor             6.2
2          3   Male   28                Doctor             6.2
3          4   Male   28  Sales Representative             5.9
4          5   Male   28  Sales Representative             5.9

   Quality of Sleep  Physical Activity Level  Stress Level BMI Category  \
0                 6                       42             6   Overweight
1                 6                       60             8       Normal
2                 6                       60             8       Normal
3                 4                       30             8        Obese
4                 4                       30             8        Obese

  Blood Pressure  Heart Rate  Daily Steps Sleep Disorder
0         126/83          77         4200            NaN
1         125/80          75        10000            NaN
2         125/80          75        10000            NaN
3         140/90          85         3000    Sleep Apnea
4         140/90          85         3000    Sleep Apnea
```

[3]:
```
df.info()
```

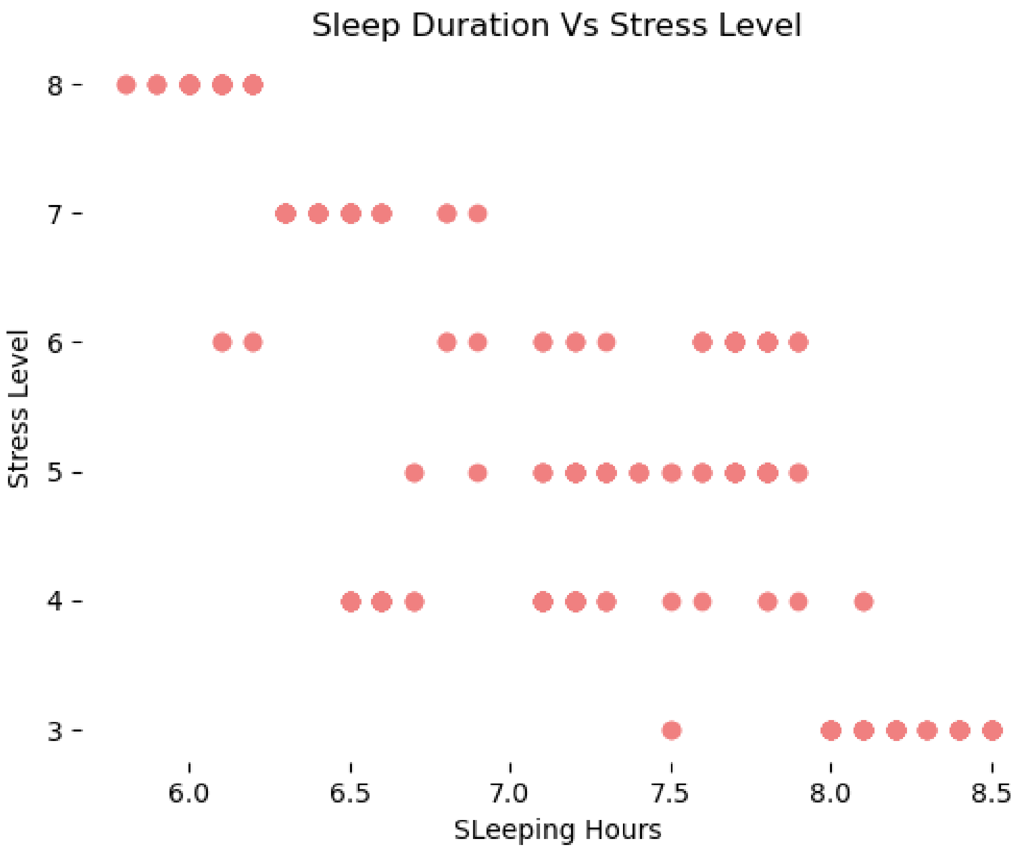
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 374 entries, 0 to 373
Data columns (total 13 columns):
 #   Column                   Non-Null Count  Dtype
---  ------                   --------------  -----
 0   Person ID                374 non-null    int64
 1   Gender                   374 non-null    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
```

[5]:
```
plt.title('Stress level By Slpeeping Duration')
sns.distplot(df['Sleep Duration'])
plt.show()
```

```
C:\Users\beydaah\AppData\Local\Temp\ipykernel_16024\3056138834.py:2:
UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df['Sleep Duration'])
C:\ProgramData\anaconda3\Lib\site-packages\seaborn\_oldcore.py:1119:
FutureWarning: use_inf_as_na option is deprecated and will be removed in a
future version. Convert inf values to NaN before operating instead.
  with pd.option_context('mode.use_inf_as_na', True):
```



[6]:
```
plt.scatter(df['Sleep Duration'], df['Stress Level'], color = 'lightcoral')
plt.title('Sleep Duration Vs Stress Level')
plt.xlabel('SLeeping Hours')
plt.ylabel('Stress Level')
plt.box(False)
plt.show()
```



[7]:
```
X = df[['Sleep Duration']]
y = df['Stress Level']
```

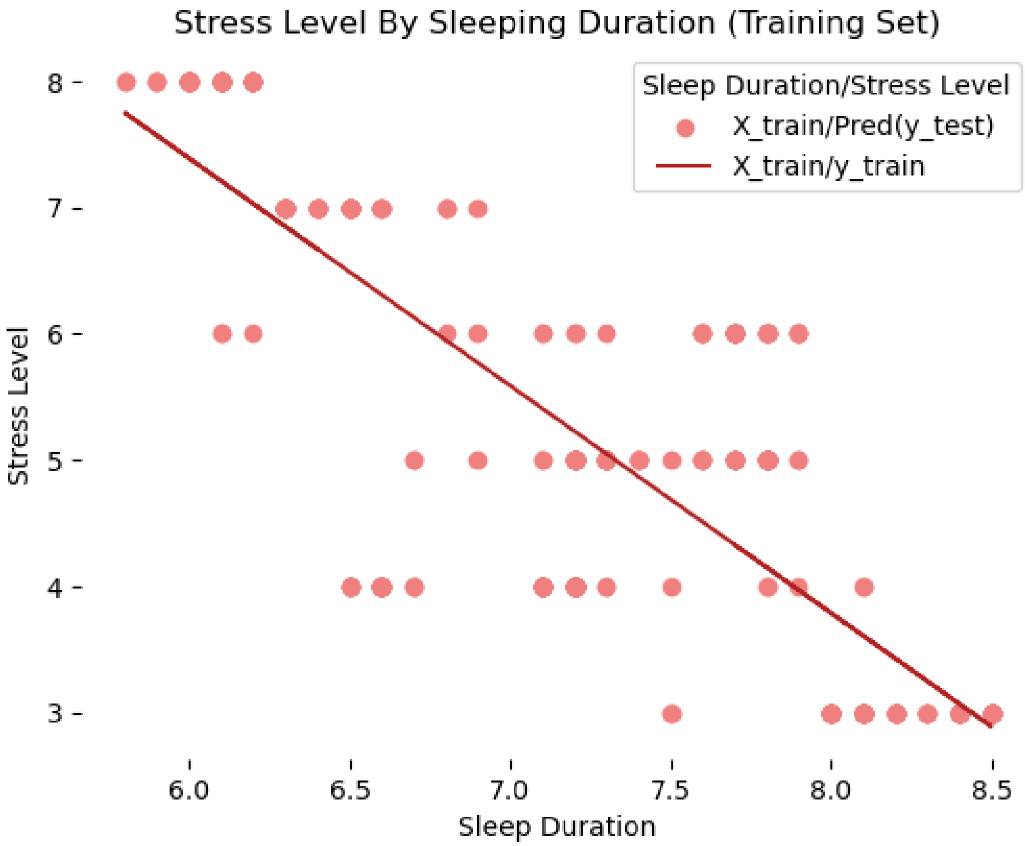
[8]:
```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,␣
 ↪random_state=0)
```

[9]:
```
regressor = LinearRegression()
regressor.fit(X_train, y_train)
```

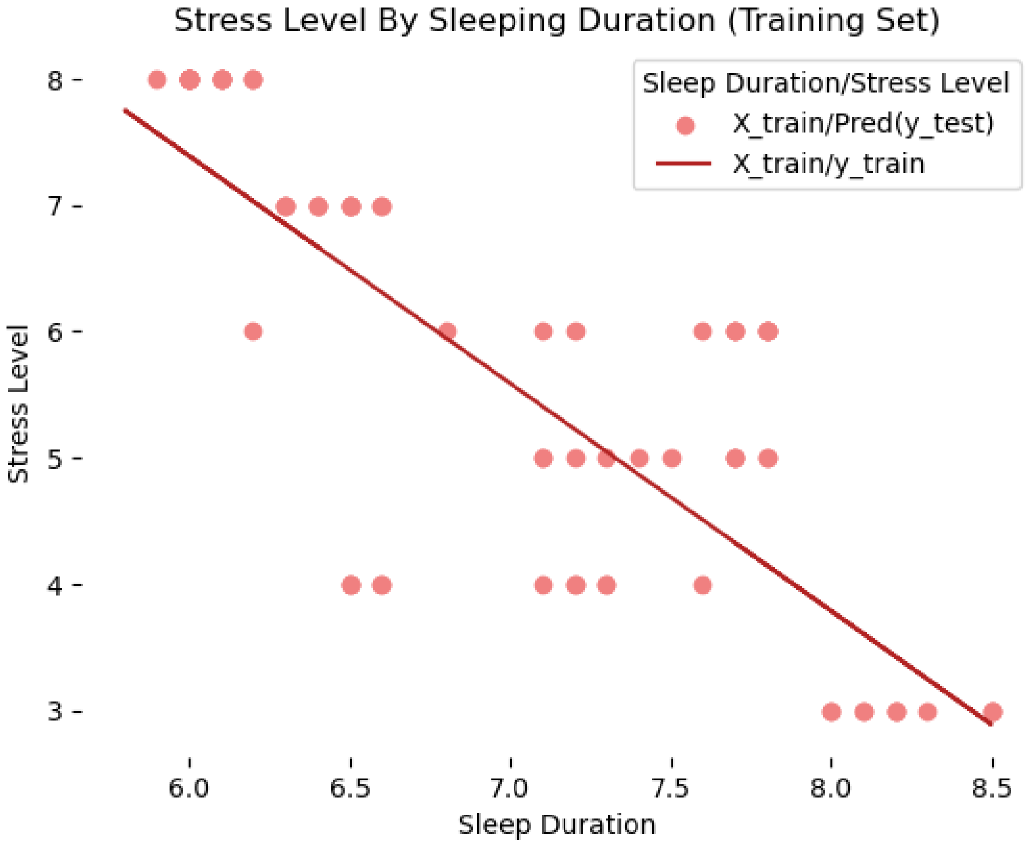
[9]:
```
LinearRegression()
```

[10]:
```
y_pred_test = regressor.predict(X_test)
y_pred_train = regressor.predict(X_train)
```

[11]:
```
plt.scatter(X_train, y_train, color = 'lightcoral')
plt.plot(X_train, y_pred_train, color = 'firebrick')
plt.title('Stress Level By Sleeping Duration (Training Set)')
plt.xlabel('Sleep Duration')
plt.ylabel('Stress Level')
plt.legend(['X_train/Pred(y_test)', 'X_train/y_train'], title = 'Sleep Duration/
 ↪Stress Level', loc='best', facecolor='white')
plt.box(False)
plt.show()
```



[12]:
```
plt.scatter(X_test, y_test, color = 'lightcoral')
plt.plot(X_train, y_pred_train, color = 'firebrick')
plt.title('Stress Level By Sleeping Duration (Training Set)')
plt.xlabel('Sleep Duration')
plt.ylabel('Stress Level')
plt.legend(['X_train/Pred(y_test)', 'X_train/y_train'], title = 'Sleep Duration/
 ↪Stress Level', loc='best', facecolor='white')
plt.box(False)
plt.show()
```


Stress Level By Sleeping Duration (Training Set)
Sleep Duration/Stress Level
● X_train/Pred(y_test)
— X_train/y_train
Stress Level
8
7
6
5
4
3
6.0
6.5
7.0
7.5
8.0
8.5
Sleep Duration

[ ]: