print("Exploring Sleep Health and Lifestyle Patterns Through Python Data Analysis")

Afonso Duarte Do Fundo Ruela Branco Carreira (107988)

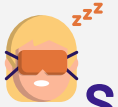
Beatriz Nunes Lavado (102495)

Tomás Vasconcelos Branco Serras Geraldes (107508)

universidade de aveiro

print("Index")

01 print("Introduction")



Sleep Health

- Physical health
- Cognitive function
- Emotional well-being

Sleep Disorders

- Sleep Apnea
- Insomnia

Data Analysis in Python

- Physiological metrics
- Sleep patterns
- Environmental conditions

02 print("Data Processing and Explanation")

```
df['Sleep Disorder'] = df['Sleep Disorder'].fillna('None')
df['Sleep Disorder'] = pd.Categorical(df['Sleep Disorder'],
                                      categories=['None', 'Sleep Apnea', 'Insomnia'],
                                      ordered=True)
```

```
# Função para categorizar a pressão sistólica
def categorizar_pressao(leitura):
    sistolica = int(leitura.split('/')[0])

    if 110 <= sistolica < 120:
        return "110-119"
    elif 120 <= sistolica < 130:
        return "120-129"
    elif 130 <= sistolica < 140:
        return "130-139"
    else:
        return "140-149"

# Aplicando a função e criando a nova coluna 'BloodPressureGroups'
df['BloodPressureGroups'] = df['Blood Pressure'].apply(categorizar_pressao)
```

```
shape = df.shape
print('The dataset has {} observations and {} variables'.format(shape[0],shape[1]))
```

```
The dataset has 374 observations and 12 variables
```

Gender	Stress Level
Age	BMI Category
Occupation	Blood Pressure
Sleep Duration	Heart Rate
Quality of Sleep	Daily Steps
Physical Act. Level	**Sleep Disorder**

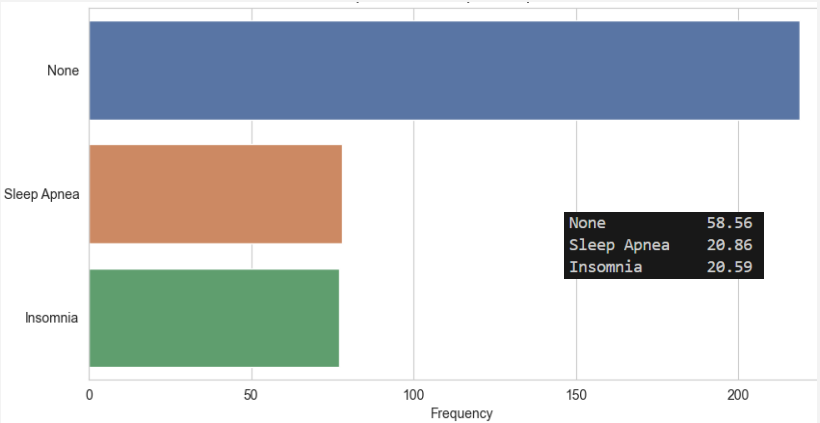
02 print("Data Processing and Explanation")

```
#vars numericas
estatistica1=df.describe().round(2).drop('count')
print(tabu(estatistica1, headers='keys', tablefmt='fancy_grid'))

#vars categoricas
estatistica2=df.describe(include='O').drop('count').drop('Blood Pressure',axis=1)
estatistica2['Sleep Disorder']=df['Sleep Disorder'].describe()
print(tabu(estatistica2,headers='keys',tablefmt='fancy_grid'))
```

	Age	Sleep Duration	Quality of Sleep
mean	42.18	7.13	7.31
std	8.67	0.8	1.2
min	27	5.8	4
25%	35.25	6.4	6
50%	43	7.2	7
75%	50	7.8	8
max	59	8.5	9

	Gender	Sleep Disorder
unique	2	3
top	Male	None
freq	189	219

```
#percentagem de doentes
percent = df['Sleep Disorder'].value_counts(normalize=True) * 100
percent=percent.round(2)
print(percent)

#Distribuição da amostra por sleep disorder
plt.figure(figsize = (10, 5))
sleepD_order = df['Sleep Disorder'].value_counts().index
sns.set_style("whitegrid")
plt.gca().set_title('Sample Distribution per Sleep Disorder')
sns.countplot(y = 'Sleep Disorder', orient='h',
              palette = 'deep', data = df, order=sleepD_order)
plt.xlabel('Frequency')
plt.show()
```



04 print("Code Structure")

Libraries

Pandas- Reading data;

Matplotlib.pyplot and **Seaborn-** Plot visualization.

```
import pandas as pd
ficheiro = "Sleep_health_and_lifestyle_dataset.csv"
df = pd.read_csv(ficheiro)
```

```
import matplotlib.pyplot as plt
plt.subplot(1, 2, 1)
plt.figure(figsize=(12, 6))
plt.xlabel()
```

```
import seaborn as sns
sns.heatmap()
sns.boxplot()
```

Functions

- Perform the descriptive statistics, generating plots, and saving output files;
- Called by the main function as needed.

```
def datasetStr():
    # ESTATISTICA DESCRITIVA
    print("Showing the first 5 observations...")
    print(df.head())
    # print(df.columns)
```

```
def saveOnFile():
    BMI_order = df['BMI Category'].value_counts().index

    # Abrir o arquivo em modo de escrita (irá sobrescrever ou criar um novo ar
    with open('Dados_e_gráficos_Projeto.txt', 'w', encoding="utf-8") as file:
        # Escrever título e primeiras observações (head)
        file.write('### First 5 Observations of the DataFrame ###\n')
        file.write(str(df.head()))  # Salva as primeiras 5 linhas
        file.write('\n\n')  # Espaço entre seções
```

Main function

- Defines the user menu;
- Starts the program.

```
def main():
    print("\n\nExploring Sleep Health and Lifestyle Patterns Through Pyt
    while True:
        print("-------------------||-------------------")
        print("Sleep is a physiological need present in animals, but it
        print(
            "To learn more about the work, the following options are ava
        # fazer contxtualização sobre o sono
        print("1 - Contextualization about sleep.")
        # Ver os nomes, ver medias, medianas, sd, etc; ver gráficos de f
        print("2 - Study variables. (see graphics and dispersion measure
        print("3 - Most relevant variables' associations")
```

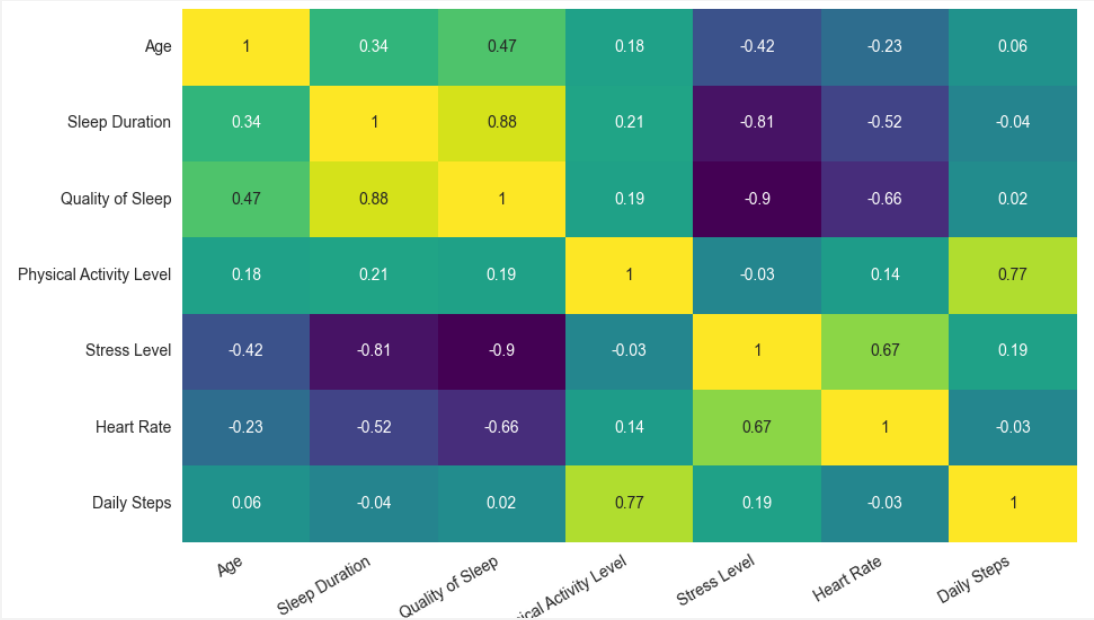
(...)

```
        if resposta == "1":
            print("Sleep is an essential biological process that allo
        elif resposta == "2":
            mostrarVariaveis()
        elif resposta == "3":
            mostraCorrs()
        elif resposta == "4":
            datasetStr()
        elif resposta == "5":
            saveOnFile()
```

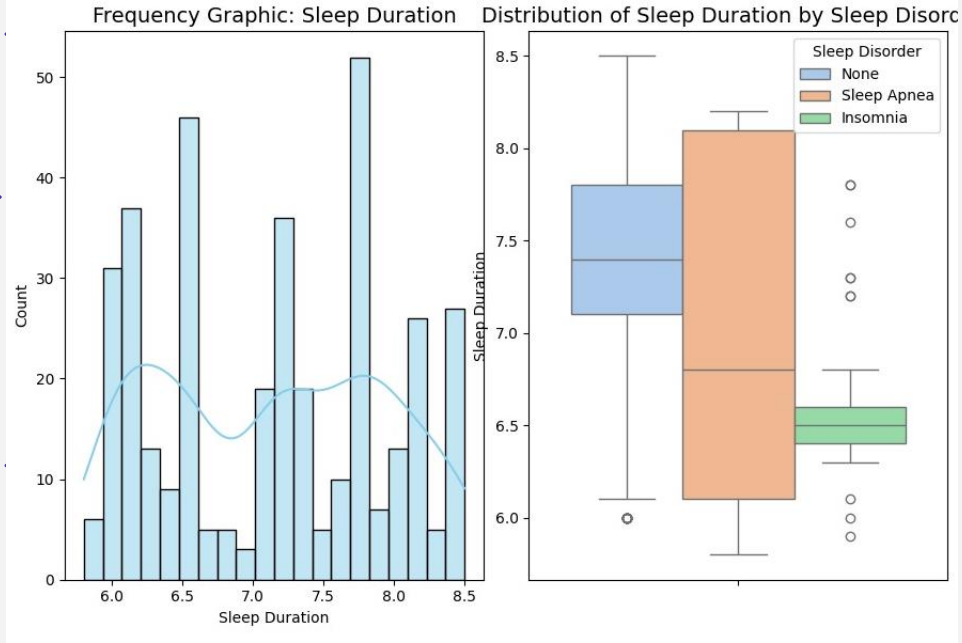
05 print("EDA")

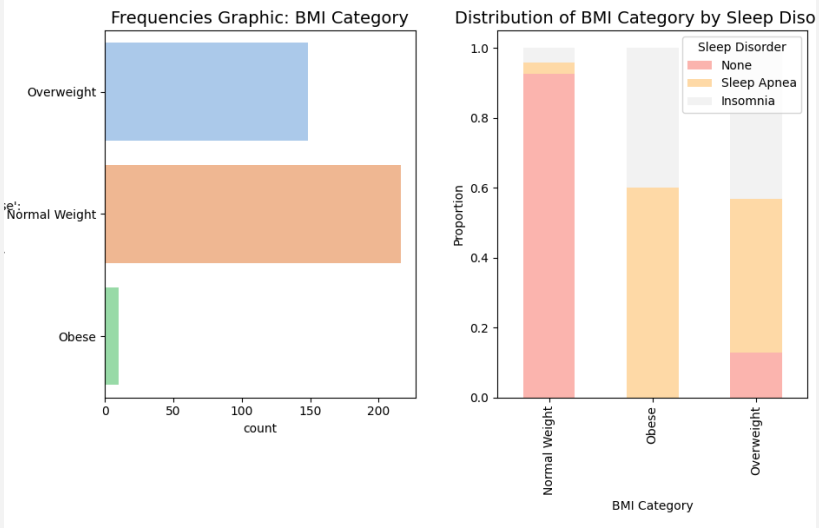
```
df_numeric = df.select_dtypes(include=['number'])
corr = df_numeric.corr().round(2)
plt.figure(figsize = (10,5))
sns.heatmap(corr, annot = True, cmap = 'viridis')
```

```
plt.xticks(rotation=30, ha='right')
plt.yticks(rotation=0)
plt.show()
```



05 print("EDA")





.....**Next step**: inferential statistics (for example, chi-squared independence test Showed significant dependence between sleep quality and sleep disorders)

06 print(“Conclusions”)

- The **data processing and EDA in Python** were successful, giving useful insights into the relationships between key variables like **Sleep Duration**, **Quality of Sleep**, and **Sleep Disorders**.

- The **user-friendly menu** provides a quick and easy way to explore relationships between variables, helping users gain initial insights without needing advanced technical skills.

- In the future, **inferential statistics** and **machine learning techniques** could be used to improve the analysis and get stronger conclusions.

Thank you...

Let’s go check the code! 🤓

Afonso Duarte Do Fundo Ruela Branco Carreira (107988)

Beatriz Nunes Lavado (102495)

Tomás Vasconcelos Branco Serras Geraldes (107508)