Final Project

CHECKLIST

* Dataset
* Visuals x3
* Hypothesis Testing 1: Simulate a Chance Model
* Hypothesis Testing 2: Compare 2 samples using permutation test.
* Hypothesis Testing 3: Create a bootstrapped confidence interval.
* Presentation
* Clarity and Creativity

Dataset:

Sleep Health and Lifestyle Dataset

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

Google Collab:

<https://colab.research.google.com/drive/1WQmFPha8gXBwhZ848vz91OhenuY0zchk?authuser=0#scrollTo=qsQIGd-EUSKd>

Presentation:

<https://www.canva.com/design/DAF2Em5fqWg/M7EW27boqawL-gYMX-ALZA/edit?utm_content=DAF2Em5fqWg&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton>

|  |  |  |
| --- | --- | --- |
| Column Name | Categorical/Quantitative | Description |
| Person ID | Categorical | An identifier for each individual |
| Gender | Categorical | The gender of the person (M/F) |
| Age | Quantitative | The age of the person in years |
| Occupation | Categorical | The occupation or profession of the person |
| Sleep Duration | Quantitative | The number of hours the person sleeps per day |
| Quality of Sleep (scale: 1-10) | Categorical | A subjective rating of the quality of sleep, ranging from 1 to 10. |
| Physical Activity Level | Quantitative | The number of minutes the person engages in physical activity daily |
| Stress Level (scale 1-10) | Categorical | A subjective rating of the stress level experience by the person, ranging from 1 to 10. |
| BMI Category | Categorical | The BMI category of the person (e.g., Underweight, Normal, Overweight). |
| Blood Pressure | Quantitative | The blood pressure measurement of the person, indicated as stolic pressure over diastolic pressure |
| Heart Rate | Quantitative | The resting heart rate of the person in beats per minute. |
| Daily Steps | Quantitative | The number of steps the person takes per day |
| Sleep Disorder | Categorical | The presence or absence of a sleep disorder in person (None, Insomnia, Sleep Apnea). |

Hypothesis Testing 1: Simulate a Chance Model

We want to know if **doctors** statistically experience more stress than all the other occupations.

NULL: The stress level for doctors is equal to the mean stress level of all the other occupations

ALT: The stress level for doctors is greater than the mean stress level of all countries

Alpha: 0.05

Conclusion: We conclude that the mean stress level of doctors is greater than the other occupations thus we fail to reject the null hypothesis

A graph of a function

Description automatically generated

Hypothesis Testing 2: Compare 2 samples using permutation test.

We want to know if the average quality of sleep with Software Engineers is significantly different than the average quality of sleep with Engineers.

Null: Software Engineers and Engineers have similar quality of sleep

Alt: The quality of sleep of SWE and Engineers is different

Alpha: 0.05

Conclusion: We conclude that the quality of sleep between Software Engineers and Engineers is significantly different thus we reject the null hypothesis.

A graph of a blue line

Description automatically generated with medium confidence

Hypothesis Testing 3: Create a bootstrapped confidence interval.

We want to estimate the mean BMI of Doctors. To do this I bootstrapped a 95% confidence interval.

Does the mean BMi of nurses fall inside the confidence region and is therefore similar to the BMI of Doctors or not.

Null: The BMI of Nurses and Doctors are similar

Alt: The BMI of Nurses and Doctors is different

Alpha: 0.05

Conclusion: We can conclude that the BMI of Doctors and Nurses is different as the BMI of Nurses is significantly greater. Therefore, we reject the null hypothesis.

A graph of a number of doctors and nurses

Description automatically generated

General Questions:

What is the average sleep in the dataset?

A graph of blue bars

Description automatically generated

Is there a specific group or gender that sleeps better based on the sleep duration?

Conclusion: Females on average get more sleep.

A blue rectangles with white text

Description automatically generated

Are there any Occupations’ who tend to have more sleep disorders?

Conclusion: The occupation with the highest is Doctor, the occupation with the lowest is Manager.

A graph of sleep disorders

Description automatically generated