Name: Jackton Mumo

Course: Advanced Data Science & AI

**First Project: Sleep Health and Lifestyle(Graph analysis)**

**Type of Analysis:** Graph analysis

**Application Area:** Health and Wellness, Medical Research

**Industry Application:** This dataset can be used in healthcare for diagnosing sleep disorders, assessing lifestyle impacts on sleep quality, and tailoring health recommendations.

**Dataset Description:**

The first project linked below is a Sleep Health and Lifestyle Dataset from Kaggle which comprises 400 rows and 13 columns, covering a wide range of variables related to sleep and daily habits. It includes details such as gender, age, occupation, sleep duration, quality of sleep, physical activity level, stress levels, BMI category, blood pressure, heart rate, daily steps, and the presence or absence of sleep disorders. It contains 20 various graphs which help evaluate the dataset.

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

****Details about Sleep Disorder Column:****

* None: The individual does not exhibit any specific sleep disorder.
* Insomnia: The individual experiences difficulty falling asleep or staying asleep, leading to inadequate or poor-quality sleep.
* Sleep Apnea: The individual suffers from pauses in breathing during sleep, resulting in disrupted sleep patterns and potential health risks.

Here is the link to the file:

<https://colab.research.google.com/drive/12eyyXinHM-P2IP3TVzdxgMjGD1bycSa4>

**Second Project: Obesity Classification**

**Type of Analysis:** Classification

**Application Area:** Health and Medicine, Public Health

**Industry Application:** This dataset is crucial for public health initiatives aimed at understanding and combating obesity. It can aid in developing targeted health interventions and policies based on demographic and lifestyle factors.

**Dataset Description:**

This dataset include data for the estimation of obesity levels in individuals from the countries of Mexico, Peru and Colombia, based on their eating habits and physical condition. The data contains 17 attributes and 2111 records, the records are labeled with the class variable Obesity Level, that allows classification of the data using the values of Insufficient Weight, Normal Weight, Overweight Level I, Overweight Level II, Obesity Type I, Obesity Type II and Obesity Type III. It focuses on the health and medicine field.

Here is the link to the notebook:

[https://colab.research.google.com/drive/1VUquNiZDYGs\_r7huQIUp3WDEJWnEQ9Cm#scrollTo=NwbJChme-iUE](https://colab.research.google.com/drive/1VUquNiZDYGs_r7huQIUp3WDEJWnEQ9Cm" \l "scrollTo=NwbJChme-iUE)

**Blood pressure analysis(Clustering)**

**Type of Analysis:** Clustering

**Application Area:** Healthcare, Predictive Medicine

**Industry Application:** This dataset can be used in healthcare for predictive modeling to foresee future health events and enable proactive measures. It also aids in personalized medicine by identifying specific patient groups with unique hypertension characteristics.

**Dataset Description:**  
This project involves patient segmentation using blood pressure data to tailor treatments and interventions for specific patient groups. Clustering techniques are applied to identify different types of hypertension, facilitating disease diagnosis and management.

The dataset contains the following columns:

* patient: Patient identifier
* sex: Gender of the patient
* agegrp: Age group of the patient
* bp\_before Blood pressure measurement before a certain intervention
* bp\_after :Blood pressure measurement after the intervention

Here is the link to the file: <https://colab.research.google.com/drive/138EIgpRsWIOSZ8sEzZgW375YjO81B7k6>