EXPLORING THE IMPACT OF SLEEP HEALTH ON LIFESTYLE

By: SAMROOTH RIFAS (COADDS222F-005)

BACHELOR OF SCIENCE (HONS) IN DATA SCIENCE

NATIONAL INSTITUTE OF BUSINESS MANAGEMENT (NIBM)- NIC

Colombo, Sri Lanka

Date of submission- 14th September 2023

"A Project Proposal submitted for the partial fulfillment of the requirements of Advanced Diploma in Data Science (Full time) Programme."

DECLARATION

I hereby declare that the work presented in this project report was carried out independently by myself and have cited the work of others and given due reference diligently.

……………….

Samrooth Rifas

……………….

Date

I certify that the above student carried out his/her project under my supervision and guidance.

………………..

W.M.S.G.D.C Wanigasekara

………………..

Date

ACKNOWLEDGMENT

I would like to express my heartfelt gratitude to all those who have contributed to the successful completion of this project. First and foremost, I would like to extend my sincere gratitude to my Course Director and my research supervisor, Mrs. W.M.S.G.D.C Wanigasekara for her valuable support, guidance and encouragement throughout this project. Additionally, I would like to thank other Senior Lecturer Mr.Thurairasa Balakumar, for guiding me in this project work.

Finally, I want to thank my Family and Friends for their valuable support, motivation, help and their encouragement throughout this journey.

Executive Summary

This research study investigates the relationship between sleep health and various lifestyle factors such as age, gender, occupation, stress level and physical activity level. The objectives have been proved using some statistical tests including Multiple Linear Regression, Chi-squared test, ANOVA test. The findings reveal that sleep duration influenced by as age, gender, occupation, stress level and physical activity level. There is a significant association between gender and presence of sleep disorders. Moreover, this study recognizes a meaningful difference between the quality of sleep and physical activity level which emphasizes the importance of in maintaining an active life. The results of this research shows the role of sleep in overall well-being and it emphasizes the need of improving sleep habits and preventing sleep disorders. Additionally this research study provides necessary recommendations to prevent the disorders related to sleep. Overall this study contributes valuable insights into optimizing sleep health for a healthier and more fulfilling life.

Contents

List of Table

List of Figure

Chapter 1: Introduction

1.1 Background

Sufficient and restful sleep is essential for proper physical, mental and emotional functioning. Sleep health has become an important issue due to prevalence of sleep disorders, increased work-related stress, and lack of physical activity. Early philosophers and medical practitioners recognized importance of sleep in maintaining a balanced and healthy life. However, the systematic scientific investigation of sleep health began in the late 19th century by some scientists. Conditions such as insomnia and sleep apnea are increasing nowadays resulting in a variety of health concerns and a lower quality of life. Numerous studies have shown that sleep health directly impacts on physical health. Poor sleep leads to some disorders such as obesity and cardiovascular diseases. At the same time poor sleep leads to some mental health issues like anxiety, depression, and mood disorders. Sleep disorder will blunt decision making skills, concentration and overall productivity. Having that on mind this research will contribute to the development of targeted interventions to improve healthy sleep habits and overall well-being for individuals in modern society.

1.2 Research Problems

The research problem for this project is focused on understanding the association between sleep disorders and lifestyle factors in a diverse population. Sleep disorders are prevalent health conditions that can significantly impact on an individual's overall well-being and quality of life. Sleep disorders have been connected to various negative health issues including increased risk of cardiovascular diseases, obesity and mental health disorders. Understanding the relationship between sleep disorders and lifestyle

factors which surrounds a wide range of variables including physical activity level, stress levels, occupation, age and gender is crucial for developing effective prevention strategies and interventions in a diverse population.

1.3 Objectives of the Project

The objectives of this project are to:

I.   Predict the association between Sleep duration and Gender, Age, Occupation, Stress level and Physical activity level.

II.   Determine the association between Gender and Sleep disorder (presence or absence of sleep disorders)

III. Investigate the mean difference between Physical activity level and Quality of sleep

1.4 Research Questions

I.   How do Gender, Age, Occupation, Stress level and Physical activity level relate to Sleep duration?

II.   Is there a significant difference in the prevalence of Sleep disorders between males and females?

III. How does the Quality of sleep affect the duration of Physical activity?

1.5 Scope of the Research

The main scope of the research is to investigate the multidimensional interaction between sleep health and lifestyle through a comprehensive examination of factors that influence sleep patterns and lifestyle aspects in a comprehensive information gathered from a diverse sample of participants. By exploring the impact of sleep health on

lifestyle factors using statistical methods and other analytical methods the research can shed insight on the relationship between them. This research will help to better understand the interconnections which means how sleep habits influence Age, Gender, Occupation, stress level and physical activity level.

1.6 Justification of the Research

This research is justified due to the important influence of sleep on various lifestyle aspects. Understanding this association can lead to improve physical and mental health and overall quality of life. By exploring these connections, the research aims to develop the personalized approaches to the awareness and prevention of several disorders caused by sleep disorders.

1.7 Expected Limitations

The main limitation of this research is not being able to collect data on my own since the instruction is to use secondary dataset. The study has been done at the people of a particular age group and occupations. Another limitation would be not being able to focus on a region in which we live. Based on the data obtained to conduct this research study, Statistical analysis is the only tool that can be used and the data obtained is subjected to these limitations.

1.8 Proposed work schedule

24th May 2023 – Started to find a dataset for the project.

11th June 2023 – Finalized the dataset.

10th July 2023 – Started working on proposal presentation.

17th July 2023 – Proposal presentation.

20th July 2023 – Introduction and literature review.

23rd July 2023 – Project Proposal Submission

9th September 2023 - Finalizing data analysis.

11th September 2023 – Discussion and Recommendation

15th September 2023 - Final Project Report Submission

Chapter 2: Literature Review

2.1 Introduction to the research theme

Sleep health is an important pillar of overall being, influencing various factors such as physical activity level, stress level, occupation, age, and gender of an individual. A balanced and restful sleep pattern is an important aspect in maintaining physical and mental functioning. The theme of the research aims to explore the impact of sleep duration, sleep disorders and quality of sleep on lifestyle factors such as physical activity level, stress level, occupation, age and gender. Through a deep exploration of the impact of sleep health on lifestyle factors this research aims to give a complete understanding of essential role sleep plays in promoting overall well being of an individual and this study seeks to identify the patterns that can inform targeted interventions to improve sleep and well-being of different population groups.

2.2 Theoretical explanation about the Key Words in the Topic

Key words	Theoretical Definitions
Impact	It refers the effect or influence of a certain factor on an individual. In the context of my research topic, I am interested in exploring the impact of sleep health on lifestyle factors.
Sleep Health	It refers to the quality of an individual's sleep patterns and habits. In this research sleep health includes sleep duration, quality of sleep, sleep disorder.

Lifestyle factors	It refers to the habits, choices and behaviors of an individual in their daily life. In this research lifestyle factors includes physical activity level, stress level, age, gender and occupation.
Insomnia	The individual experiences difficulty falling asleep or staying asleep, leading to inadequate or poor- quality sleep.
Obstructive Sleep Apnea	The individual suffers from pauses in breathing during sleep, resulting in disrupted sleep patterns and potential health risks.
Quality of sleep	It means how restful and uninterrupted of an individual's sleep during the night. Sleep duration and sleep disturbances are the factors affect quality of sleep.

Table 1.Theoretical explanation of the Key Words

2.3 Findings by other researchers

The following part tries to learn more about past researchers' investigations and findings on themes related to this research study.

- A study published in 2021on "Lifestyle factors and Sleep Health across the Lifespan" focused on sleep health across different age groups impacts on lifestyle factors. A cross sectional survey conducted between 3284 participants on sleep health, lifestyle habits and demographics shows that older adults reported the highest sleep health followed by young and middle aged adults. It

states that some negative lifestyle habits such as fast food consumption and social media are influencing sleep health across different age groups. (Joseph M, Sahar M, Sarah M, Elliottnell Perez ,2021)

- A study published in 2022 on " Sleep and Health – A Lifestyle Medicine Approach" found that sleep is an important aspect for overall mental and physical well being. It is related with some important health issues such as obesity, heart disease diabetes and Insomnia. Sleep disorder leads accidents and errors. Cognitive Therapy for Insomnia (CBT-I) is an effective treatment for Insomnia. Stress management, engaging in physical activities are the lifestyle factors which promotes a good sleep health. (Param Dedhia, Robert Maurer, 2022)

- A cross sectional study titled as "Health, Lifestyle, and Psycho- Social Determinants of Poor Sleep Quality During the Early Phase of the COVID-19 Pandemic" shows that the influence of COVID-19 related lockdowns on sleep quality in older adults aged more than 50 years whom 523 were grouped as COVID-19CEV out of 5558 were surveyed. The survey results shows that 37% of participants were reported with poor sleep quality during lockdown. Age, female sex, higher BMI, mental health disorders, loneliness, higher anxiety and depressive symptoms are the factors that are related to poor sleep. This study highlights the poor sleep in older adults who were affected by COVID-19 during the lockdown. (Chinedu T, Tamlyn, Shireen,2021)

- Another study was done on "Association of healthy lifestyle with the risk of obstructive sleep apnea" among Chinese adults. The results showed that a higher HLS (Healthy Lifestyle Score) which developed by the researchers based on seven lifestyle factors such as physical activity, mental status, diet, smoking, alcohol consumption and waist-hip ratio was associated with lower obstructive sleep apnea (OSA) and this study highlights the significance of preventing OSA to follow a healthy life. (Xueru Duan, Jun Huang, Murui Zheng, 2022)

- A research on "Anxiety and depressive symptoms are associated with poor sleep health during a period of COVID-19 induced nationwide lockdown" found that majority of the participants out of 1240 were experienced anxiety and depression during lockdown which is associated with over 60% reported sleep problems and shorter sleep duration. (Yazan , Aziz, Wafaa, 2020)

- Another study on "Stress and sleep" investigates the impacts of specific stressors on sleep among Hispanics/Latinos in the United States. The study has been done over 5000 adults and found that stress was associated with poor sleep outcomes such as Insomnia, daytime sleepiness and less sleep duration. (Carmela, Sanjay, Sheila F,2017)

- A study published in 2020 on "A composite measure of sleep health is associated with symptoms of depression among Japanese female hospital nurses" states the association between sleep health and symptoms of depression which includes 2482 female nurses with an average age of 31.2 years and they reported their sleep and at least one depressive symptom. (Ryuji, Kaori, 2020)

- A study on "The impact of sleep, physical activity and sedentary behavior on symptoms of depression and anxiety before and during the COVID-19 pandemic in a sample of South African participants" found that Insomnia symptoms and excessive sedentary screen time were highly linked to more depression and anxiety during lockdown through a study of 1048 South African adults. These findings suggest that engaging in physical activity helped to improve to sleep quality and reduced symptoms of depression and anxiety during lockdown. (R.Lewis, L.C.Roden, 2021)

- A research on "Changes in lifestyle during the COVID-19 lockdown in Brazil: Impact on sleep quality" was done states that among 589 participants in a Brazilian population, 41.93% reported poor sleep quality during lockdown. Engaging in physical activity and being male were protective measure for the poor sleep quality. This study highlighted that sleep patterns were associated with daily routines disorder, increased stress and changes in health related habits during the pandemic. (Gabriala, Diogo, Andrea, 2022)

- Another research on "The influence of sleep and movement on mental health and life satisfaction during the transition to parenthood" involved 157 couples showing various parental statuses. This study revealed that getting sufficient amount of sleep played a significant role in improving mental health specially among new parents. However this study did not find a strong association between physical activity and mental health outcomes during the transition to parenthood. (Alison, Chris, Ryan, 2022)

2.4 The research gap

The reviewed studies have explored mostly about the association between stress, anxiety, depression and the sleep habits and also the relationship between sleep health and age but the studies explored about the association between sleep disorders such as insomnia, sleep apnea and physical activities are less. They did not mainly focus on the sleep health and physical activities as there is still a need of further researches to investigate the complex relationship between them like how sleep health and physical activity contribute each other. Furthermore, there is still a gap in understanding the extent to which the sleep health associates with different types of occupations and across males and female. Overall, the study needs to fill the research gap by exploring the relationship between more lifestyle factors and sleep health.

2.5 Table for Variables, their definitions and sources

Variables	Definition	Sources
Person ID	An identifier for each individual	Kaggle Sleep Health and Lifestyle Dataset
Gender	The gender of the person (Male/Female)	Kaggle Sleep Health and Lifestyle Dataset
Age	The age of the person in years	Kaggle Sleep Health and Lifestyle Dataset
Occupation	The occupation or profession of the person	Kaggle Sleep Health and Lifestyle Dataset

Sleep Duration	The number of hours the person sleeps per day	Kaggle Sleep Health and Lifestyle Dataset
Quality of sleep (scale: 1-10)	A subjective rating of the quality of sleep, ranging from 1-10	Kaggle Sleep Health and Lifestyle Dataset
Physical activity level (minutes/day)	The number of minutes the person engages in physical activity daily	Kaggle Sleep Health and Lifestyle Dataset
Stress level (scale: 1-10)	A subjective rating of the stress level experienced by the person, ranging from 1-10	Kaggle Sleep Health and Lifestyle Dataset
Sleep Disorder	The presence or absence of a sleep disorder in the person (None, Insomnia, Apnea)	Kaggle Sleep Health and Lifestyle Dataset
Blood Pressure (Systolic/Diastolic)	The blood pressure measurement of the person, indicated as systolic pressure over diastolic pressure	Kaggle Sleep Health and Lifestyle Dataset
Heart Rate (bpm)	The resting heart rate of the person in beats per minute	Kaggle Sleep Health and Lifestyle Dataset
Daily Steps	The number of steps the person takes per day	Kaggle Sleep Health and Lifestyle Dataset

BMI category	The BMI category of the person (Underweight, Normal, Overweight)	Kaggle Sleep Health and Lifestyle Dataset

Table 2.Variables, their definitions and sources

2.6 Chapter conclusion

The literature review given in this chapter focused to provide a comprehensive understanding between sleep health and lifestyle factors as sleep health is a complex matter which is influenced by various lifestyle aspects such as age, gender, occupation, stress level and physical activity. While this chapter reviewed and analyzed the findings from various range of studies with different populations, still this study seeks to fill the research gap regarding the understanding between sleep health and different occupation, gender and physical activity, This research study aims to fill this research gap by exploring the impact of sleep health on more lifestyle factors.

Chapter 3: Methodology

3.1 Introduction

In this chapter we will be focusing on the population, sample for this study, the types of data to be collected and the data sources, data collection tools which means from where the dataset was obtained, the conceptual framework to be done for this research, hypothesis, operationalization table which contains variable, definitions, references and at the last the methods of data analysis which is being used to this research study. This chapter will give a clear understanding regarding this study which includes statistical tests such as Multiple linear regression model, Chi-squared test and ANOVA test used to determine the relationship between variables for each objectives and the hypothesis to be tested respectively as the research aims to explore the connection between sleep health and lifestyle factors.

3.2 Population, sample and Sampling technique

As it is statistical research it is important to test a population which means a collection of individuals with the features that has to be investigated. The population for this study includes adults both males and females of different occupations. Th sample size which was tested under different experiments is 374 participants aged from 27 to 59 years consisting of 51% of males and 49% of females. This sample is consisting of individuals who has been examined and analyzed in order to arrive at a conclusion for this study. This ensures that the sample reflects the diversity of population which allows more accurate findings on how sleep health effects different lifestyle groups. The sampling technique used here is not specified in the dataset documentation.

3.3 Type of Data to be collected and data sources

The type of data to be collected for this research would involve both qualitative and quantitative data. The categorical variables such as gender, occupation, stress level, sleep disorder and quality of sleep could be included under qualitative data and the numerical variables such as age, physical activity level and sleep duration could be included under quantitative data. These quantitative and qualitative data types can be used to get a better understanding about this research. The data source for this project is the Sleep Health and Lifestyle Dataset I got from "Kaggle".

3.4 Data collections tools and plan

The data to be used for this research study is from a dataset obtained from "Kaggle" which is an online community platform for data scientists and machine learning enthusiasts. So that we can call it is a secondary data source. 374 participants especially from the ages of 27 to 59 were tested to get more accurate data. This sample size gives more appropriate and securely stored data about sleep health and lifestyle factors which is being explored in this research study. The data will be integrated to provide a clear understanding on impact of sleep health on lifestyle aspects.

3.5 Conceptual framework

i.   Predict the association between sleep duration and gender, age, occupation, stress level and physical activity level.

- Independent variable - Gender, Age, Occupation, Stress level and physical activity level

- Dependent variable – Sleep duration

ii. Determine the association between Gender and Sleep disorder

- Gender

- Sleep disorder

iii. Investigate the mean difference between Physical activity level and Quality of sleep

- Independent variable – Quality of sleep

- Dependent variable – Physical activity level

3.6 Hypothesis

I. Predict the association between sleep duration and gender, age, occupation, stress level and physical activity level.

- Null hypothesis – There is no significant linear relationship between Gender, Age, Occupation, Stress level and physical activity level and Sleep duration.

- Alternative hypothesis – There is a significant relationship between Gender, Age, Occupation, Stress level and physical activity level and Sleep duration.

II. Determine the association between Gender and Sleep disorder.

- Null hypothesis – There is no significant association between Gender and Sleep disorder.

- Alternative hypothesis – There is a significant association between Gender and Sleep disorder.

III. Investigate the mean difference between Physical activity level and Quality of sleep.

- Null hypothesis – There is no significant differences between the means of Quality of sleep and Physical activity level.
- Alternative hypothesis – There is at least one significant difference between the means of Quality of sleep and Physical activity level.

3.7 Operationalization Table

Variable	Indicators	Measures
Gender	Male or Female	Categorical Male Female
Age	Age	Numerical 27 - 59
Occupation	Occupation	Nurse Doctor Sales Representative Teacher Software Engineer Engineer Accountant Scientist Lawyer Salesperson

Sleep Duration	Hours	Numerical 5.8 – 8.5
Physical activity level	Minutes/day	Numerical 30 - 90
Stress level	Scale: 1- 10	Categorical 3 - 8
Sleep Disorder	Presence or absence of a sleep disorder	Categorical None Insomnia Sleep Apnea

Table 3.Operationalization Table

3.8 Methods of data analysis

The data obtained will be analyzed using Python and R Studio with the use of Inferential statistics (Multiple Linear Regression model, Chi- squared test, ANOVA test), Descriptive statistics and Graphical Representation.

All the objectives will be achieved for this research study by the following statistical tests,

i.   Predict the association between sleep duration and gender, age, occupation, stress level and physical activity level.

- Multiple Linear Regression Model

ii.   Determine the association between Gender and Sleep disorder

- Chi-squared test

iii.   Investigate the mean difference between Physical activity level and Quality of sleep
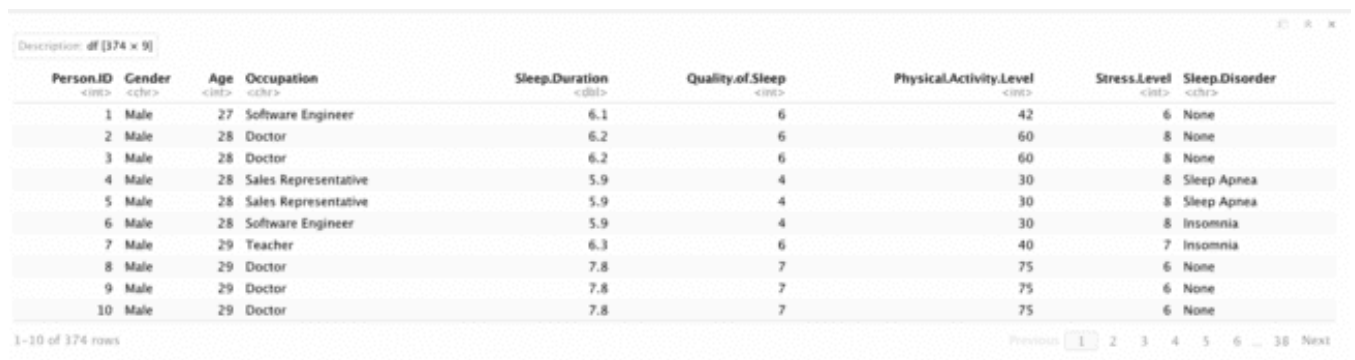
- ANOVA test

After performing the tests in R Studio, the hypothesis will be tested, accordingly the null or alternative hypothesis will be accepted or declined. The relationships between the variables to be tested in this research will be proved and we can conclude that sleep health impacts on lifestyle factors.

Chapter 4: Data Analysis

4.1 Data Analysis

4.1.1 Data Preprocessing

All Data Preprocessing, Data Visualization, Descriptive Statistics and Statistical tests are done using R-Studio. In Data Preprocessing the first step is the variables that are not needed in the above research objectives are removed from the original dataset which is obtained from Kaggle. Initially the original dataset had 13 variables and from there 4 variables have been removed to visualize the data related to the objectives using 'select' function in R-Studio.

Figure 1.Dataset after removing unwanted variables

The next step of the data preprocessing is out of 9 variables which are obtained from the original dataset, the variables that are having the structure as 'integer' have been converted into 'numeric' and the variables that are having the structure as 'character' have been converted into 'factor'.

```
'data.frame':   374 obs. of  9 variables:
 $ Person.ID              : int  1 2 3 4 5 6 7 8 9 10 ...
 $ Gender                 : Factor w/ 2 levels "Female","Male": 2 2 2 2 2 2 2 2 2 2 ...
 $ Age                    : num  27 28 28 28 28 28 29 29 29 29 ...
 $ Occupation             : Factor w/ 11 levels "Accountant","Doctor",..: 10 2 2 7 7 10 11 2 2 2 ...
 $ Sleep.Duration         : num  6.1 6.2 6.2 5.9 5.9 5.9 6.3 7.8 7.8 7.8 ...
 $ Quality.of.Sleep       : Factor w/ 6 levels "4","5","6","7",..: 3 3 3 1 1 1 3 4 4 4 ...
 $ Physical.Activity.Level: num  42 60 60 30 30 30 40 75 75 75 ...
 $ Stress.Level           : Factor w/ 6 levels "3","4","5","6",..: 4 6 6 6 6 6 5 4 4 4 ...
 $ Sleep.Disorder         : Factor w/ 3 levels "Insomnia","None",..: 2 2 2 3 3 1 1 2 2 2 ...
```

Figure 2.Structure of the variables after changing as 'numeric' and 'factor'

The final step of data preprocessing is the dataset have been tested whether the data under each and every variables have null values or not and here the above dataset has no any null values.

4.1.2 Data Visualization for Categorical Variables

1.Gender

gender <chr>	count.Var1 <fctr>	count.Freq <int>
Female	Female	185
Male	Male	189

Figure 3.Count table of Gender

Figure 4.Bar graph for Gender

Gender Distribution

Female
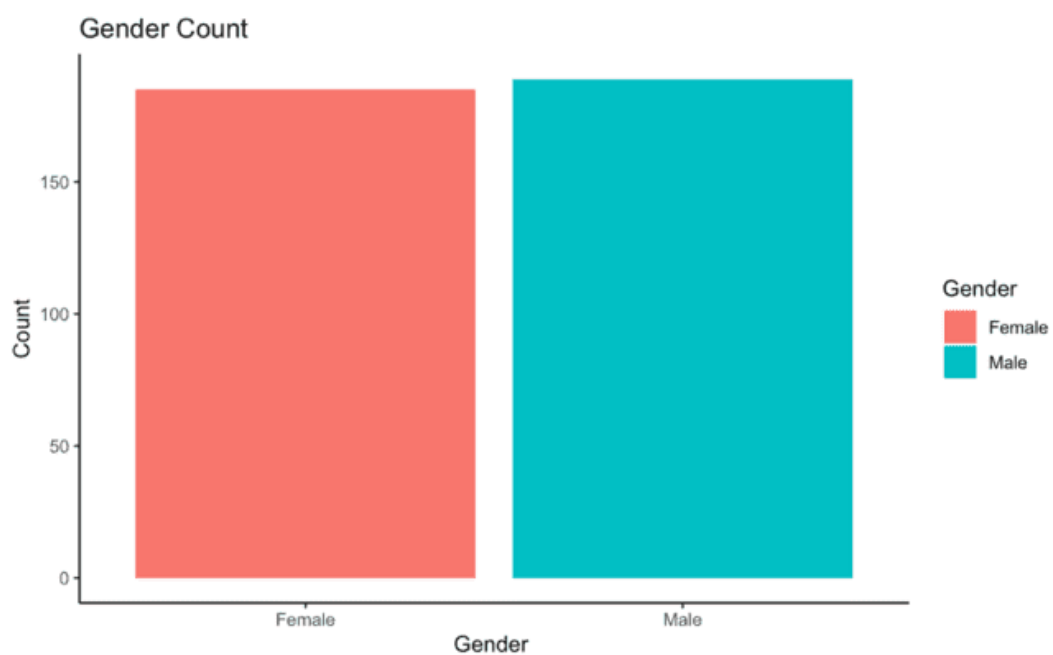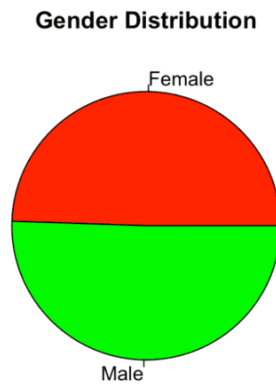
Male

Figure 5.Pie chart for Gender

According to the above bar graph and pie chart there are 185 females and males and 189 males under gender variable in this dataset.

2.Occupation

occupation <chr>	count.Var1 <fctr>	count.Freq <int>
Accountant	Accountant	37
Doctor	Doctor	71
Engineer	Engineer	63
Lawyer	Lawyer	47
Manager	Manager	1
Nurse	Nurse	73
Sales Representative	Sales Representative	2
Salesperson	Salesperson	32
Scientist	Scientist	4
Software Engineer	Software Engineer	4

Figure 6.Count table for Occupation

Job position frequency

Figure 7.Bar graph for Occupation

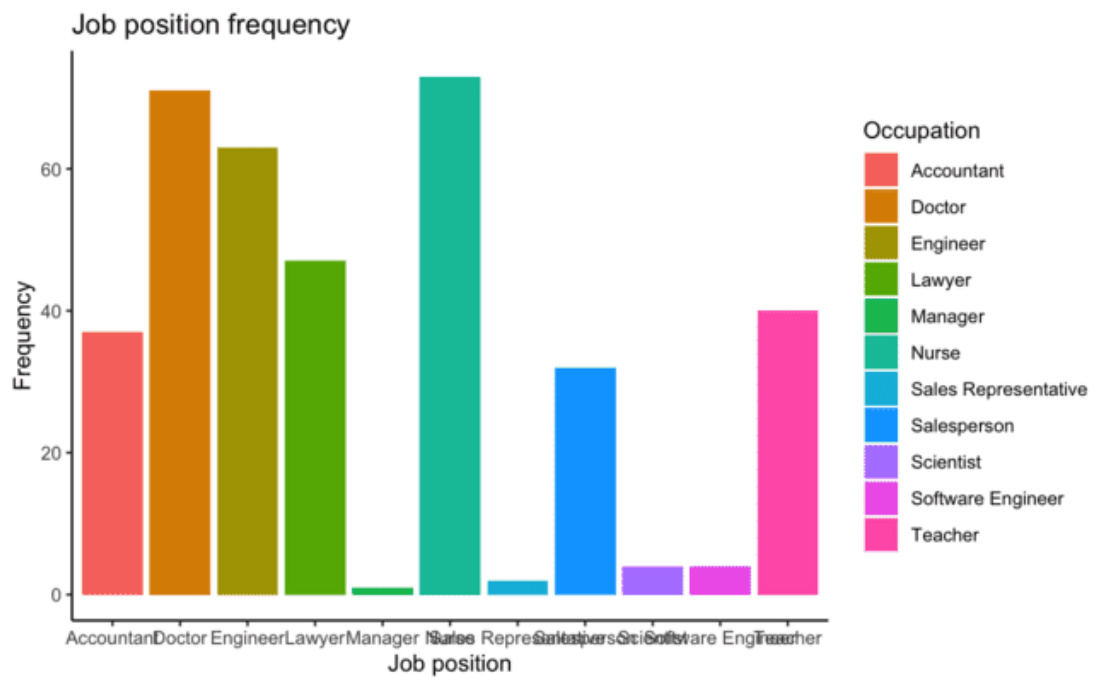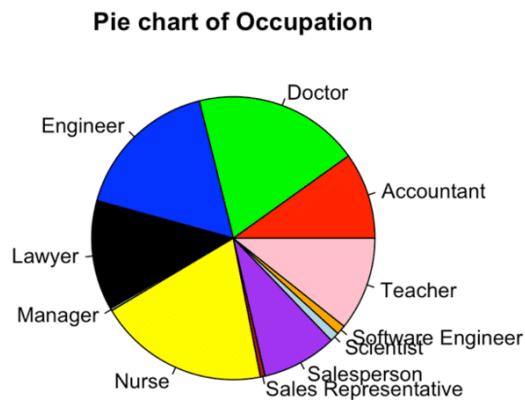## Pie chart of Occupation

Figure 8.Pie chart for Occupation

According to the above bar graph and pie chart there are 37 Accountants, 71 Doctors, 63 Engineers, 47 Lawyers, 73 Nurses, 32 Salespersons, 4 Scientists and Software Engineers, 2 Sales Representatives and only one Manager.

3.Quality of Sleep

Quality.of.Sleep <chr>	count.Var1 <fctr>	count.Freq <int>
4	4	5
5	5	7
6	6	105
7	7	77
8	8	109
9	9	71

Figure 9.Count table for Quality of sleep
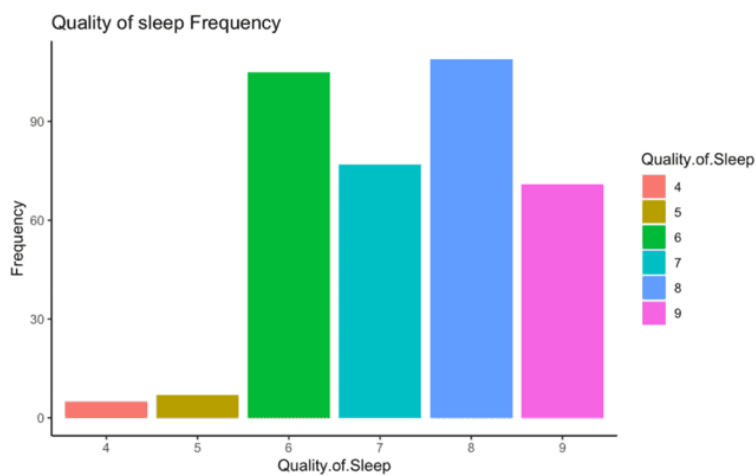
Figure 10.Bar graph for Quality of sleep

It is a subjective rating of Quality of Sleep ranging from 1 to 10. According to this bar graph in this dataset data is available for only ranging from 4 to 9. There are only 5 number of rating of 4, 7 number of rating of 5, 105 number of rating of 6, 77 number of rating of 7, 109 number of rating of 8 and 71 number of rating of 9. Most of the peoples Quality of Sleep is good as it shows the rating of 8.

4.Stress Level

Description: df [6 × 3]

Stress.Level <chr>	count.Var1 <fctr>	count.Freq <int>
3	3	71
4	4	70
5	5	67
6	6	46
7	7	50
8	8	70

Figure 11.Count table for Stress level
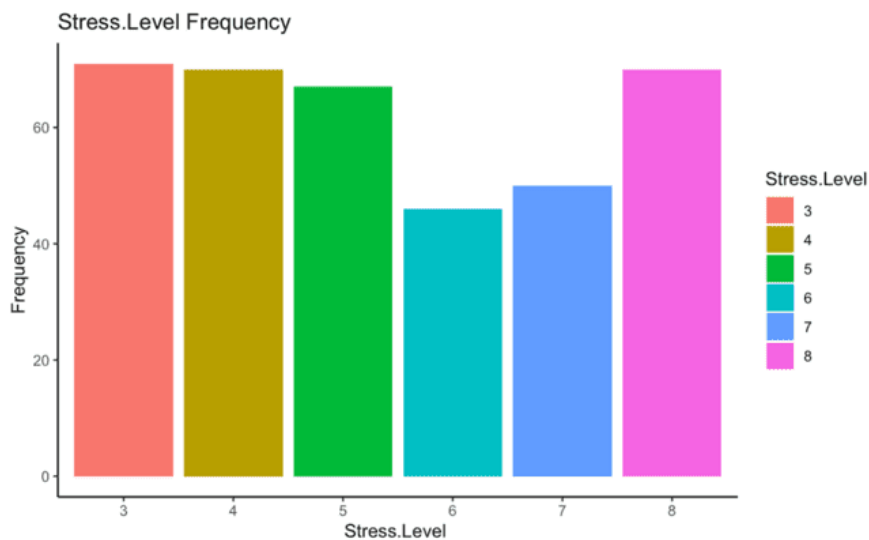
Figure 12.Bar graph for Stress level

It is a subjective rating of the stress level experienced by the person, ranging from 1-10. According to this bar graph in this dataset data is available for only ranging from 3 to 8. There are 71 number of rating of 3, 70 number of rating of 4, 67 number of rating of 5, 46 number of rating of 6, 50 number of rating of 7 and 70 number of rating of 8. Highest number of people are experiencing the stress level at rating of 3.

5.Sleep Disorder

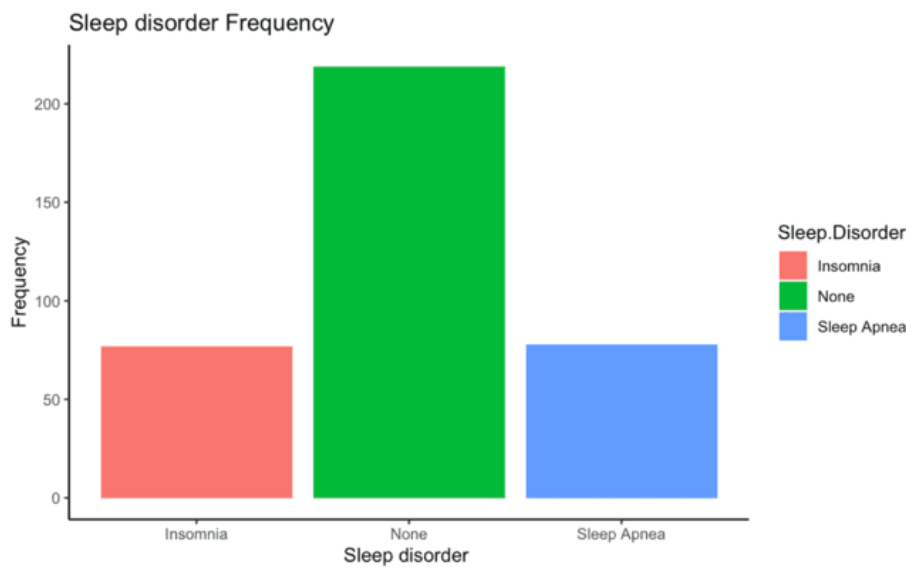| Sleep.Disorder <chr>	count.Var1 <fctr>	count.Freq <int>
Insomnia	Insomnia	77
None	None	219
Sleep Apnea	Sleep Apnea	78

Figure 13.Count table for Sleep disorder

Figure 14.Bar graph for Sleep disorder
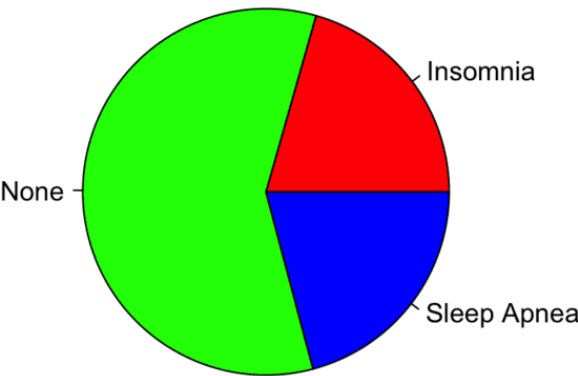
Figure 15.Pie chart for Sleep disorder

According to this bar graph and pie chart there are 77 people are having Insomnia, 78 people are having Sleep Apnea and highest number of 219 people are not having any sleep disorder.
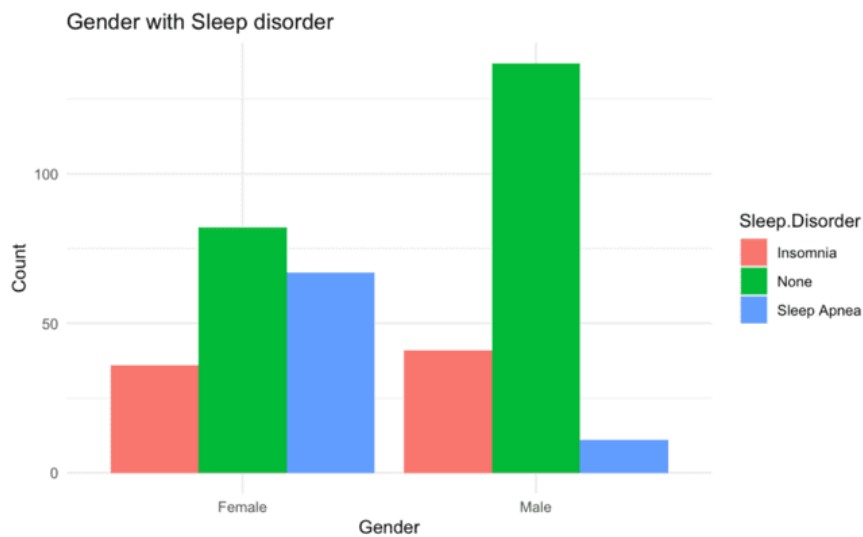
Figure 16.Bar graph for Gender and Sleep disorder

Sleep.Disorder <fctr>	Gender <fctr>	Count <int>
Insomnia	Female	36
None	Female	82
Sleep Apnea	Female	67
Insomnia	Male	41
None	Male	137
Sleep Apnea	Male	11

Figure 17.Count table for Gender and Sleep disorder

This above bar graph shows the count of female and male having different types of sleep disorder. Accordingly the highest number of 82 females and 137 males are not having any sleep disorder. Only 36 females having Insomnia and 11 having Sleep Apnea.
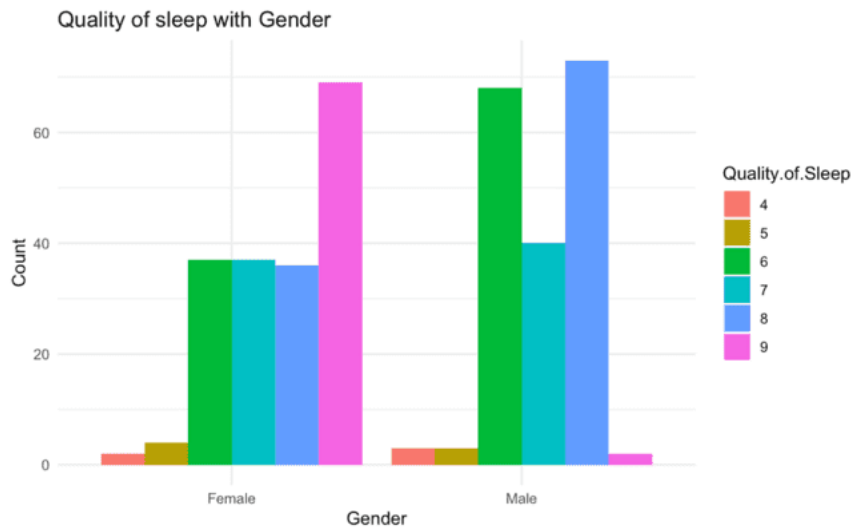
Figure 18.Bar graph for Gender and Quality of sleep

Quality.of.Sleep <fctr>	Gender <fctr>	Count <int>
4	Female	2
5	Female	4
6	Female	37
7	Female	37
8	Female	36
9	Female	69
4	Male	3
5	Male	3
6	Male	68
7	Male	40

Figure 19.Count table for Gender and Quality of sleep

The above bar chart shows the count of males and females with different rating in quality of sleep. The highest number of 69 females having the quality of sleep at the rating of 9 while 68 males having the quality of sleep at the rating of 6. The lowest number of 2 females having the quality of sleep at the rating of 4 while each 3 males having the quality of sleep at the ratings of 4 and 5.

4.1.3 Data Visualization for Numerical Variables

Descriptive statistics for Numerical Variables

Mean, Median, 1st Quartile, 3rd Quartile, Minimum and Maximum value of the Numerical Variables such as Age, Physical Activity level and Sleep Duration are respectively as follows,

```
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   27.00   35.25   43.00   42.18   50.00   59.00
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   30.00   45.00   60.00   59.17   75.00   90.00
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   5.800   6.400   7.200   7.132   7.800   8.500
```

Figure 20.Descriptive analysis for Numerical variables

1.Age
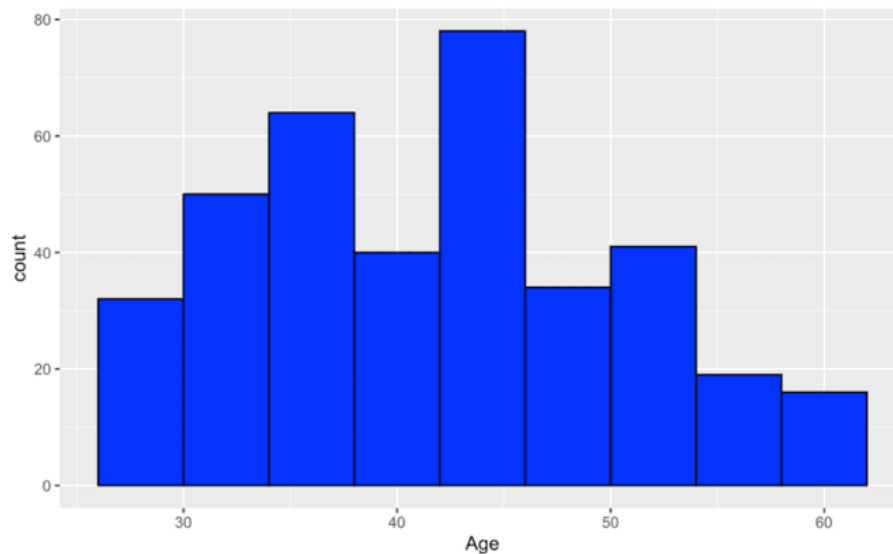
Figure 21.Histogram for Age

The tallest bar for the age range 30-40 and the tallest bar for the age range 40-50 have more people than the other. Overall, the histogram shows that the most people in this dataset are in their 30s and 40s which fewer peoples are in their 20s, 50s and 60s.

2.Physical Activity Level
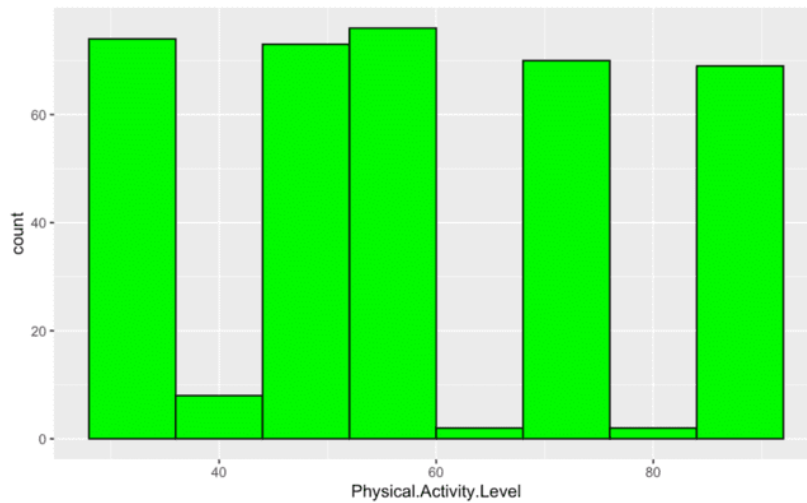
Figure 22.Histogram for Physical Activity level

The horizontal axis shows the physical activity level in minutes and the vertical axis shows the number of people in each physical activity level range. There are two tallest bar for physical activity level range 40-60 minutes which means most of the people in this dataset engage in physical activity for 40-60 minutes per day while the other physical activity level ranges have fewer people in them.

3.Sleep Duration
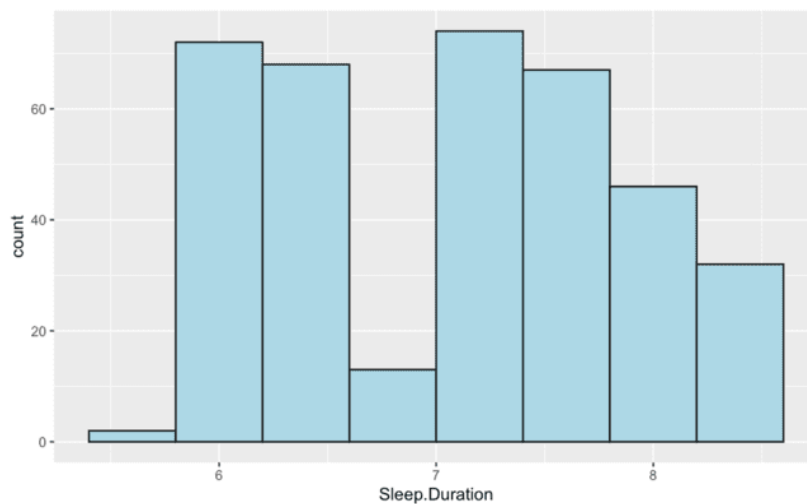
Figure 23.Histogram for Sleep duration

There are two tallest bar for sleep duration range of 7-8 hours which means most of the people in this dataset sleep for 7-8 hours while there are fewer people who sleep for less than 6 hours and more than 8 hours per day.

4.1.4 Data Visualization for Both Categorical and Numerical Variables
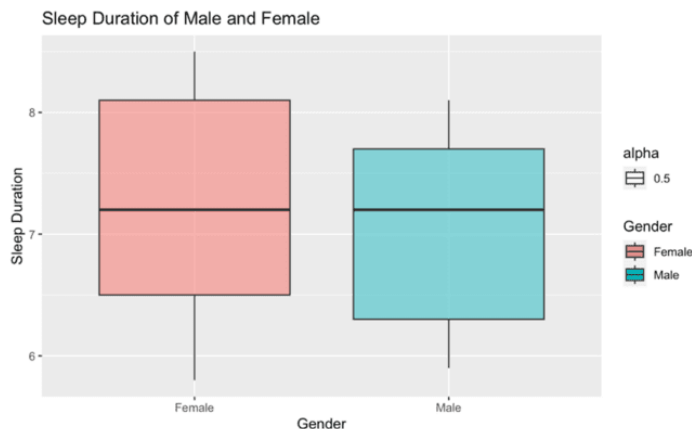
Figure 24.Boxplot for Gender and Sleep duration

There are no outliers in this data. This boxplot compares the sleep duration of males and females and this shows that females tend to sleep more than males on average. Both genders have similar sleep duration with most values falls between 6-8 hours. The median sleep duration for both genders is same which is between 7-7.5 hours.
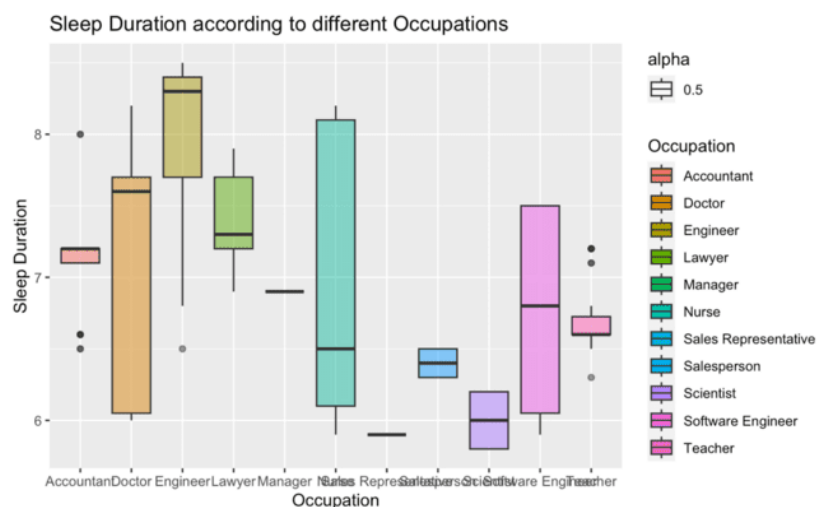
Figure 25.Boxplot for Occupation and Sleep duration

According to this boxplot the occupation with higher median sleep duration is Engineer between 8- 8.5 hours and the occupation with lower median sleep duration is Sales Representative around 6 hours. This shows that Engineers sleep more hours and Sales Representatives sleeps less than the other occupations. There are some outliers for Accountant, Engineer and Teacher.
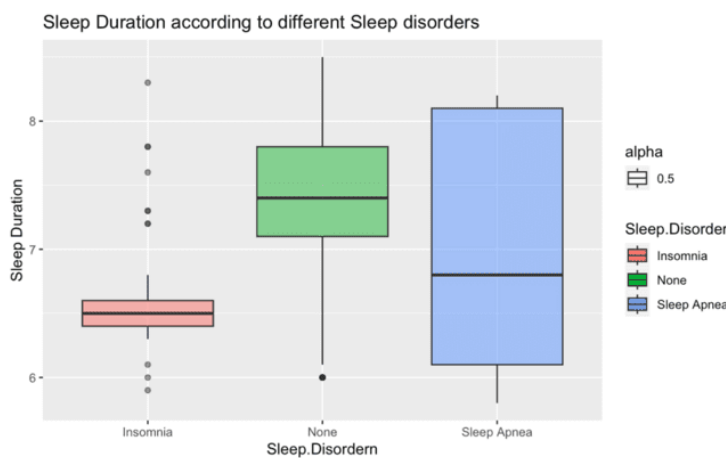
Figure 26.Boxplot for Sleep disorder and Sleep duration

There is no outlier for only Sleep Apnea while Insomnia and no sleep disorder have outliers. The median sleep duration for Insomnia is 6.5 hours, for no sleep disorder is around 7.4 hours and for Sleep Apnea is between 6.5-7 hours. This shows that the people with Insomnia and Sleep Apnea sleep less than the people with no sleep disorder.

4.2 Findings and Interpretation

All the objectives of this research study had been achieved after performing the following statistical tests in R-Studio and finally arrived at a conclusion.

1. Predict the association between sleep duration and gender, age, occupation, stress level and physical activity level.

- **Multiple Linear Regression Model**

  - Independent variable - Gender, Age, Occupation, Stress level and physical activity level

  - Dependent variable – Sleep duration

```
Call:
lm(formula = data2$Sleep.Duration ~ data2$Age + data2$Physical.Activity.Level +
    data2$Stress.Level + data2$Gender + data2$Occupation)

Residuals:
    Min      1Q   Median      3Q     Max
-0.68295 -0.19992 -0.02051  0.16971  1.04444

Coefficients:
                                       Estimate Std. Error t value Pr(>|t|)
(Intercept)                           8.1228071  0.1731698  46.907  < 2e-16 ***
data2$Age                             0.0063841  0.0033649   1.897  0.05860 .
data2$Physical.Activity.Level         0.0077840  0.0009903   7.861 4.51e-14 ***
data2$Stress.Level                   -0.3725408  0.0121390 -30.690  < 2e-16 ***
data2$GenderMale                     -0.1064605  0.0767339  -1.387  0.16618
data2$OccupationDoctor                0.8197045  0.0867291   9.451  < 2e-16 ***
data2$OccupationEngineer              0.6645799  0.0812353   8.181 4.94e-15 ***
data2$OccupationLawyer                0.4763605  0.0921991   5.167 3.96e-07 ***
data2$OccupationManager              -0.0755032  0.2967576  -0.254  0.79931
data2$OccupationNurse                 0.0646497  0.0762183   0.848  0.39688
data2$OccupationSales Representative  0.4517078  0.2257759   2.001  0.04618 *
data2$OccupationSalesperson           0.3663802  0.1132915   3.234  0.00133 **
data2$OccupationScientist            -0.0480295  0.1570844  -0.306  0.75997
data2$OccupationSoftware Engineer     0.3957666  0.1680243   2.355  0.01904 *
data2$OccupationTeacher              -0.3552701  0.0696542  -5.100 5.50e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2922 on 359 degrees of freedom
Multiple R-squared:  0.8702,    Adjusted R-squared:  0.8652
F-statistic: 171.9 on 14 and 359 DF,  p-value: < 2.2e-16
```

Figure 27.P-value for Multiple Linear Regression

Null hypothesis: There is no significant linear relationship between Gender, Age, Occupation, Stress level and physical activity level and Sleep duration.

Alternative hypothesis: There is a significant relationship between Gender, Age, Occupation, Stress level and physical activity level and Sleep duration.

As the P value=2.2e-16 is less than 0.05 we reject Null hypothesis and accept Alternative hypothesis and we can conclude that there is a significant relationship between Gender, Age, Occupation, Stress level and physical activity level and Sleep duration.

2. Determine the association between Gender and Sleep disorder

▪ **Chi-squared test**

- Gender

- Sleep disorder

```
        Pearson's Chi-squared test

data:  data2$Gender and data2$Sleep.Disorder
X-squared = 54.306, df = 2, p-value = 1.613e-12
```

Figure 28.P-value for Chi-squared test

Null hypothesis: There is no significant association between Gender and Sleep disorder.

Alternative hypothesis: There is a significant association between Gender and Sleep disorder.

As p-value =1.613e-12 less than 0.05 so we reject Null hypothesis and accept Alternative hypothesis and we can conclude that there is a significant association between Gender and Sleep disorder.

3. Investigate the mean difference between Physical activity level and Quality of sleep

▪ **ANOVA test**

- Independent variable – Quality of sleep

- Dependent variable – Physical activity level

```
                  Df  Sum Sq  Mean Sq  F value   Pr(>F)
Quality.of.Sleep   1    6022     6022    14.38  0.000175 ***
Residuals        372  155831      419
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 29.P-value for ANOVA test

Null hypothesis: There is no significant differences between the means of Quality of sleep and Physical activity level.

Alternative hypothesis: There is at least one significant difference between the means of Quality of sleep and Physical activity level.

As p value=0.000175 is less than 0.05 we reject Null hypothesis and accept Alternative hypothesis and we can conclude that there is at least one significant difference between the means of Quality of sleep and Physical activity level.

Chapter 5: Discussion and Recommendations

5.1 Discussion

According to the statistical test which is performed to achieve all three objectives of this research study having the topic Exploring the Impact of Sleep health on Lifestyle Factors, It has been proved that sleep habits influence Age, Gender, Occupation, stress level and physical activity level from the dataset taken here. Here Multiple Linear Regression Model had been used to Predict the association between sleep duration and gender, age, occupation, stress level and physical activity level and according to the test we can decide that there is a relationship between them. In accordance to the Chi-squared test, it has been proved that the presence of sleep disorder varies with both males and females. By ANOVA test we can decide that there is at least one significant difference between the means of Quality of sleep and Physical activity level. While the past research studies have proved that the sleep health impacts on the lifestyle factors this research study also has confirmed this using the statistical tests. Finally this research will help to better understand the interconnections which means how sleep habits influence Age, Gender, Occupation, stress level and physical activity level. By exploring these connections, the research aims to develop the personalized approaches to the awareness and prevention of several disorders caused by sleep disorders.

5.2 Recommendation

The research problem that has been recognized in this research study is that the sleep disorders have been connected to various negative health issues among different age groups and different occupations. By exploring this research study we can understand the effect of sleep disorders on individual's overall well-being and quality of life and

we can develop effective prevention strategies and interventions. Some suggestions to avoid these problems are,

- Promote sleep education on the importance of sleep health by implementing educational programs among the individuals of all ages and genders to improve sleep duration and quality.

- Workplaces should conduct provide flexible work hours particularly to support the well-being of high-stress occupations.

- Promote sleep education in schools and workplaces on how high-quality sleep contributes to better physical activity level which means the ability to engage physical activity for a longer duration.

- Implementing awareness programs on Stress Management among all ages by informing them how poor sleep quality and insufficient sleep duration lead to high stress level.

5.3 Conclusions

The aim of this research study is to explore the connections between sleep health and lifestyle factors such as age, gender, occupation, stress level and physical activity level. According to the research questions the sleep duration may vary with females and males and adults and youngsters. Also sleep durations differ among high stress jobs and low stress jobs. Physical activity level and stress level also vary with the sleep duration. Gender was found to be associated with the presence and absence of sleep disorder. Quality of sleep was significantly influenced by physical activity level. Furthermore, there is still a gap in understanding the extent to which the sleep health associates with

different types of occupations and across males and female and complex relationship

between sleep disorders such as insomnia, sleep apnea and physical activities.

APPENDICES

1.

```{r}
library(tidyverse)
```

```
Registered S3 methods overwritten by 'dbplyr':
  method          from
  print.tbl_lazy
  print.tbl_sql
── Attaching packages ──────────────────────────────────── tidyverse 1.3.2 ──✓
✓ ggplot2 3.4.0      ✓ purrr   0.3.5
✓ tibble  3.1.8      ✓ dplyr   1.0.10
✓ tidyr   1.2.1      ✓ stringr 1.4.1
✓ readr   2.1.3      ✓ forcats 0.5.2 ── Conflicts
──────────────────────────────────── tidyverse_conflicts() ──
✗ dplyr::filter() masks stats::filter()
✗ dplyr::lag()    masks stats::lag()
```

```{r}
library(ggplot2)
```

2.

```{r}
s1 <- select(data,-(Daily.Steps),-(BMI.Category),-(Blood.Pressure),-(Heart.Rate))
s1
```

3.

```{r}
data1= write.csv(s1,file="sleep_health_new.csv",row.names = F)
```

aftr prprcng - with only 9 variables
```{r}
data2=read.csv("sleep_health_new.csv")
data2
```

4.

```{r}
data2$Gender <- as.factor(data2$Gender)
data2$Age <- as.numeric(data2$Age)
data2$Occupation  <- as.factor(data2$Occupation)
data2$Quality.of.Sleep <- as.factor(data2$Quality.of.Sleep)
data2$Physical.Activity.Level <- as.numeric(data2$Physical.Activity.Level)
data2$Stress.Level  <- as.factor(data2$Stress.Level)
data2$Sleep.Disorder   <- as.factor(data2$Sleep.Disorder)

```

```{r}
str(data2)
```

5.

```{r}
sum(is.na(data2))
```

6.

```{r}
summary(data2)
```

```{r}
summary(data2$Age)
summary(data2$Physical.Activity.Level)
summary(data2$Sleep.Duration)
```

7.

```{r}
ggplot(data=data2)+geom_bar(mapping=aes(x=Gender ,fill=Gender))+
labs(title="Gender Count",x="Gender ",y="Count")+theme_classic()
```

```{r}
chart_data <- data.frame(gender = names(gender_counts), count = gender_counts)
chart_data
```

```{r}
gender.count = table(data2$Gender)
pie(gender.count, col = c("Red","Green"), main = "Gender Distribution")
```

8.

```{r}
ggplot(data=data2)+geom_bar(mapping=aes(x=Occupation ,fill=Occupation))+
labs(title="Job position frequency",x="Job position",y="Frequency")+theme_classic()
```

```{r}
chart_data <- data.frame(occupation = names(occupation_counts), count = occupation_counts)
chart_data
```

```{r}
occupation.count = table(data2$Occupation)
pie(occupation.count, col = c("Red","Green","Blue","Black","light green","Yellow","Brown","Purple","light
blue","Orange","Pink"), main = "Pie chart of Occupation")
```

9.

```r
Quality.of.Sleep_counts <- table(data2$Quality.of.Sleep)
chart_data <- data.frame(Quality.of.Sleep = names(Quality.of.Sleep_counts), count = Quality.of.Sleep_counts)
chart_data
```

```r
ggplot(data=data2)+geom_bar(mapping=aes(x=Quality.of.Sleep ,fill=Quality.of.Sleep))+
labs(title="Quality of sleep Frequency",x="Quality.of.Sleep",y="Frequency")+theme_classic()
```

10.

```r
Stress.Level_counts <- table(data2$Stress.Level)
chart_data <- data.frame(Stress.Level = names(Stress.Level_counts), count = Stress.Level_counts)
chart_data
```

```r
ggplot(data=data2)+geom_bar(mapping=aes(x=Stress.Level,fill=Stress.Level ))+
labs(title="Stress.Level Frequency",x="Stress.Level ",y="Frequency")+theme_classic()
```

11.

```r
ggplot(data=data2)+geom_bar(mapping=aes(x=Sleep.Disorder,fill=Sleep.Disorder ))+
labs(title="Sleep disorder Frequency",x="Sleep disorder ",y="Frequency")+theme_classic()
```

```r
Sleep.disorder_counts <- table(data2$Sleep.Disorder)
chart_data <- data.frame(Sleep.Disorder = names(Sleep.disorder_counts), count = Sleep.disorder_counts)
chart_data
```

```r
sleep_disorder.count = table(data2$Sleep.Disorder )
pie(sleep_disorder.count, col = c("Red","Green","Blue"), main = "Pie chart for Sleep disorder")
```

12.

```r
ggplot(data=data2)+geom_bar(mapping=aes(x=Gender ,fill=Sleep.Disorder),position="dodge")+
labs(title="Gender with Sleep disorder",x="Gender",y="Count")+theme_minimal()
```

```r
sleep.disorder_gender_table=table(data2$Sleep.Disorder,data2$Gender)
table2<- as.data.frame(sleep.disorder_gender_table)
colnames(table2)<- c("Sleep.Disorder","Gender","Count")
table2
```

```r
ggplot(data=data2)+geom_bar(mapping=aes(x=Gender ,fill=Quality.of.Sleep),position="dodge")+
labs(title="Quality of sleep with Gender ",x="Gender",y="Count")+theme_minimal()
```

```{r}
quality.of.sleep_gender_table=table(data2$Quality.of.Sleep,data2$Gender)
table2<- as.data.frame(quality.of.sleep_gender_table)
colnames(table2)<- c("Quality.of.Sleep","Gender","Count")
table2
```

13.

```{r}
gplot(data=data2)+geom_histogram(mapping=aes(x=Age ),binwidth = 4,color="black",fill="blue")
```

```{r}
ggplot(data=data2)+geom_histogram(mapping=aes(x=Physical.Activity.Level),binwidth = 8,color="black",fill="green")
```

```{r}
ggplot(data=data2)+geom_histogram(mapping=aes(x=Sleep.Duration),binwidth = 0.4,color="black",fill="light blue")
```

14.

```{r}
ggplot(data=data2)+geom_boxplot(mapping=aes(x=Gender,y=Sleep.Duration,fill=Gender,alpha=0.5))+
labs(title="Sleep Duration of Male and Female",x="Gender",y="Sleep Duration")+theme_grey()
```

```{r}
ggplot(data=data2)+geom_boxplot(mapping=aes(x=Occupation,y=Sleep.Duration,fill=Occupation,alpha=0.5))+
labs(title="Sleep Duration according to different Occupations",x="Occupation",y="Sleep Duration")+theme_grey()
```

```{r}
ggplot(data=data2)+geom_boxplot(mapping=aes(x=Sleep.Disorder,y=Sleep.Duration,fill=Sleep.Disorder,alpha=0.5))+
labs(title="Sleep Duration according to different Sleep disorders",x="Sleep.Disordern",y="Sleep Duration")+theme_grey()
```

```{r}
sleep_lm<-lm(data2$Sleep.Duration~data2$Age+data2$Physical.Activity.Level+data2$Stress.Level+data2$Gender+data2$Occupation)
sleep_lm
```

15.

```{r}
summary(sleep_lm)
```

16.

```{r}
test <- chisq.test(data2$Gender, data2$Sleep.Disorder)
test
```

17.

```{r}
one.way<-aov(Physical.Activity.Level~Quality.of.Sleep,data=data2)
summary(one.way)
```

REFERENCES

- Sleep Health and Lifestyle Dataset available from <https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset?resource=download> [Retrieved 11 June 2023]

- Dzierzewski, Joseph M., Sahar M. Sabet, Sarah M. Ghose, Elliottnell Perez, Pablo Soto, Scott G. Ravyts, and Natalie D. Dautovich. 2021. "Lifestyle Factors and Sleep Health across the Lifespan" *International Journal of Environmental Research and Public Health,*18(12) [online].available from< https://www.mdpi.com/1660-4601/18/12/6626 >[16 July 2023]

- Param Dedhia, MD, Robert Maurer.2022" Sleep and Health-A Lifestyle Medicine Approach (2022)"[online].available from <https://lifestylemedicine.org/wp-content/uploads/2022/08/ACLM-Article-LM-Sleep-and-Health.pdf >[16 July 2023]

- Chinedu T. Udeh-Momoh, Tamlyn Watermeyer, Shireen Sindi, Parthenia Giannakopoulou, Catherine E. Robb, Sara Ahmadi-Abhari, Bang Zheng, Amina Waheed, James McKeand, David Salman, Thomas Beaney,Celeste A. de Jager Loots, Geraint Price, Christina Atchison, Josip Car, Azeem Majeed, Alison. H. McGregor, Miia Kivipelto, Helen Ward and Lefkos T. Middleton.2021." Health ,Lifestyle ,and Psycho-Social Determinants of Poor Sleep Quality During the Early Phase of the Covid-19 Pandemic" *Front.PublicHealth*[online].availablefrom<

https://www.frontiersin.org/articles/10.3389/fpubh.2021.753964/full >[16 July 2023]

- Xueru Duan, Jun Huang, Murui Zheng, Wenjing Zhao, Lixian Lao, Haiyi Li, Zhiwei Wang, Jiahai Lu, Weiqing Chen, Hai Deng and Xudong Liu.2022." Association of healthy lifestyle with risk of obstructive sleep apnea" *BMC Pulmonary Medicine*[online].available from< https://bmcpulmmed.biomedcentral.com/articles/10.1186/s12890-021-01818-7 >[16 July 2023]

- Yazan A Al- Ajlouni, Su Hyun Park, Jude Alawa, Ghaith Shamaileh, Aziz Bawab, Wafaa M El- Sadr, Dustin T Duncan.2020." Anxiety and depressive symptoms are associated with poor sleep health during a period of COVID-19 induced nationwide lockdown"*BMJ Journals*[online].available from< https://bmjopen.bmj.com/content/10/12/e041995 >[16 July 2023]

- Carmela Alcantara, Sanjay R., Mercedes Carnethon, Sheila F., Carmen R., Sonia Davis, Alberto R., Elva Arredondo, Susan Redline, Phyllis C., Linda C.2017."Stress and sleep"*SSM-Population Health*[online].availablefrom< https://www.sciencedirect.com/science/article/pii/S2352827317300459 >[16 July 2023]

- Ryuji Furihata, Kaori Saitoh, Masahiro Suzuki, Maki Jike, Yashitaka Kaneita, Takashi Ohida, Daneil J., Makoto Uchiyama.2020. "A composite measure of sleep health is associated with symptoms of depression among Japanese female

hospital nurses"*Comprehensive Psychiatry*[online].available from<

https://www.sciencedirect.com/science/article/pii/S0010440X19300744

> [22 July 2023]

- R.Lewis, L.C.Roden, K.Scheuermaier, F.X.Gomez-Olive, D.E.Rae,
  S.Iacovides, A.Bentley, J.P.Davy, C.J.Christie, S.Zschernack, J,Rochd,
  G.Lipinska.2021."The impact of sleep, physical activity and sedentary
  behaviour on symptoms of depression and anxiety before and during the
  COVID-19 pandemic in a sample of South African participants"*Scientific
  reports*[online].available from< https://www.nature.com/articles/s41598-021-
  02021-8 > [12 September 2023]

- Gabriela Vieira Minetto, Diogo Thimoteo da Cunha, Andrea Maculano
  Esteves.2022."Changes in lifestyle during the COVID-19 lockdown in
  Brazil:Impact on sleep quality"*Sleep Epidemiolody*[online].available
  from<https://www.sciencedirect.com/science/article/pii/S2667343622000142
  > [12 September 2023]

- Alison Divine, Chris Blanchard, Cecilia Benoit, Danielle Symons Downs, Ryan
  E.Rhodes.2022."The influence of sleep health and life satisfaction during the
  transition to parenthood"*Sleep Health*[online].available
  from<https://www.sciencedirect.com/science/article/pii/S2352721822001139
  > [12 September 2023]