Final Project

Delaney Norman

Data

The data is on sleep quality and comes from Kaggle. It contains variables on gender, age, occupation, sleep duration and quality, stress level, physical activity and daily steps, heart rate, BMI, blood pressure, and whether the person has a sleep disorder.

```
data = read_csv("Sleep_health_and_lifestyle_dataset.csv")
dim(data)
```

```
[1] 374  13
```

```
names(data)
```

```
 [1] "Person ID"               "Gender"
 [3] "Age"                     "Occupation"
 [5] "Sleep Duration"          "Quality of Sleep"
 [7] "Physical Activity Level" "Stress Level"
 [9] "BMI Category"            "Blood Pressure"
[11] "Heart Rate"              "Daily Steps"
[13] "Sleep Disorder"
```

```
head(data)
```

```
# A tibble: 6 x 13
  `Person ID` Gender   Age Occupation        `Sleep Duration` `Quality of Sleep`
        <dbl> <chr> <dbl> <chr>                        <dbl>              <dbl>
1           1 Male     27 Software Engineer             6.1                  6
```

```
2            2 Male       28 Doctor                              6.2                   6
3            3 Male       28 Doctor                              6.2                   6
4            4 Male       28 Sales Representa~                   5.9                   4
5            5 Male       28 Sales Representa~                   5.9                   4
6            6 Male       28 Software Engineer                   5.9                   4
# i 7 more variables: `Physical Activity Level` <dbl>, `Stress Level` <dbl>,
#   `BMI Category` <chr>, `Blood Pressure` <chr>, `Heart Rate` <dbl>,
#   `Daily Steps` <dbl>, `Sleep Disorder` <chr>
```

Goal

The goal is to determine what factors affect sleep quality the most, and what the average sleep quality and length of sleep are, both across the sample and when breaking the sample down into different groups based on health, occupation, age, etc.

Variables I plan on analyzing

- Sleep duration (quantitative) and sleep quality (ordinal)

- Compare differences in groups (compare across occupations, sleep disorders, gender, age, etc)

Hypotheses

- More people have poor quality of sleep (between 1-5) than good quality (between 6-10)

- Those with higher stress levels will get less sleep and have lower sleep quality than those with lower stress levels

- There is a difference in sleep quality based on the number of steps a person takes per day and the amount of physical activity they do