Work Hard, Sleep Hard?

By: The Standard Deviants





01

Introduction

Background, Motivation, and Research Question

Introduction

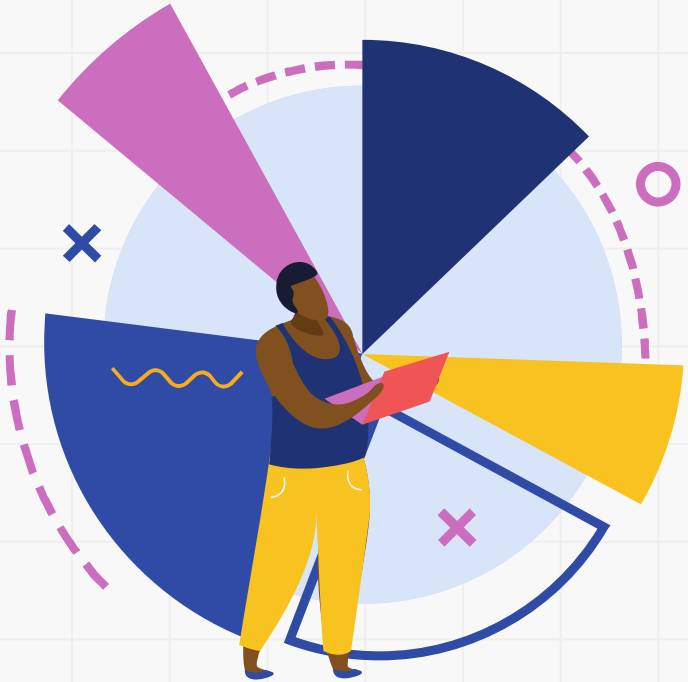
Background and Motivation: Our group chose to explore this question to better understand the factors that influence sleep quality and duration. Because sleep plays a significant role in overall health and quality of life, we believe it is an important topic to research.

Research Question: How does the level of physical activity associated with an individual's occupation influence their sleep quality and duration?

02

Data

Data Description and Data Cleaning



Data Description

Source: The data used in our analysis comes from the Lifestyle and Sleep Patterns dataset, published on Kaggle by Minahil Fatima.

Description: The used data set contains synthetic data generated based on specific algorithms, distributions, and assumptions designed to simulate individuals' sleep patterns, screen time habits, and related lifestyle factors.

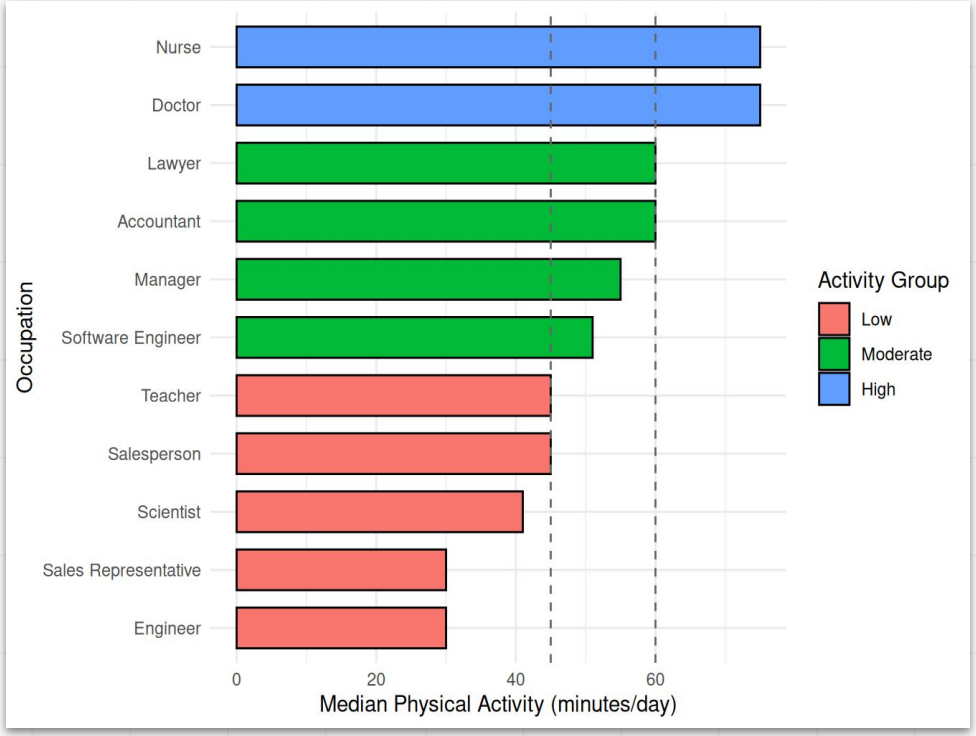
Variables of Interest

- **Occupation**: Self-reported profession (11 categories).
- **Sleep Duration**: Average nightly sleep (hours).
- **Quality of Sleep**: Self-reported sleep quality (1–10 scale; higher = better).
- **Physical Activity Level**: Average daily physical activity (minutes/day).
- **Occupation Activity Group**: Low, Moderate, or High, based on each occupation's median physical activity level.

Data Cleaning

1. Selected four variables of interest: occupation, sleep duration, sleep quality, and physical activity level.
2. Calculated the median daily physical activity level for each of the 11 occupations.
3. Classified occupations into three activity groups based on their median physical activity levels:
    - Low: 45 minutes/day or less (n = 141)
    - Moderate: More than 45, up to 60 minutes/day (n = 89)
    - High: More than 60 minutes/day (n = 144)
4. Assigned each participant to an activity group based on their occupation.
5. Checked for missing values and verified group assignments.

Activity Group Categorization





03

EDA

Exploratory Data Analysis

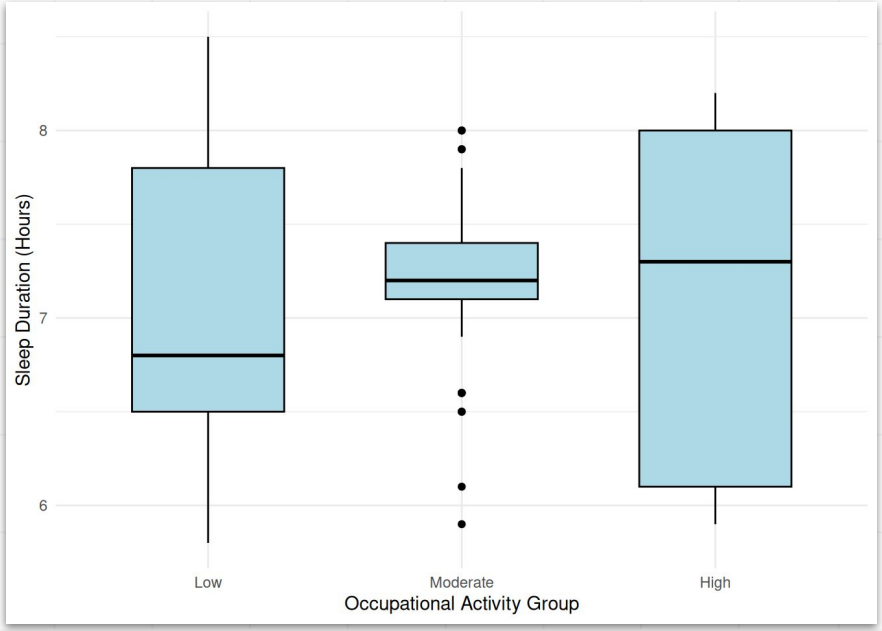
Summary Statistics

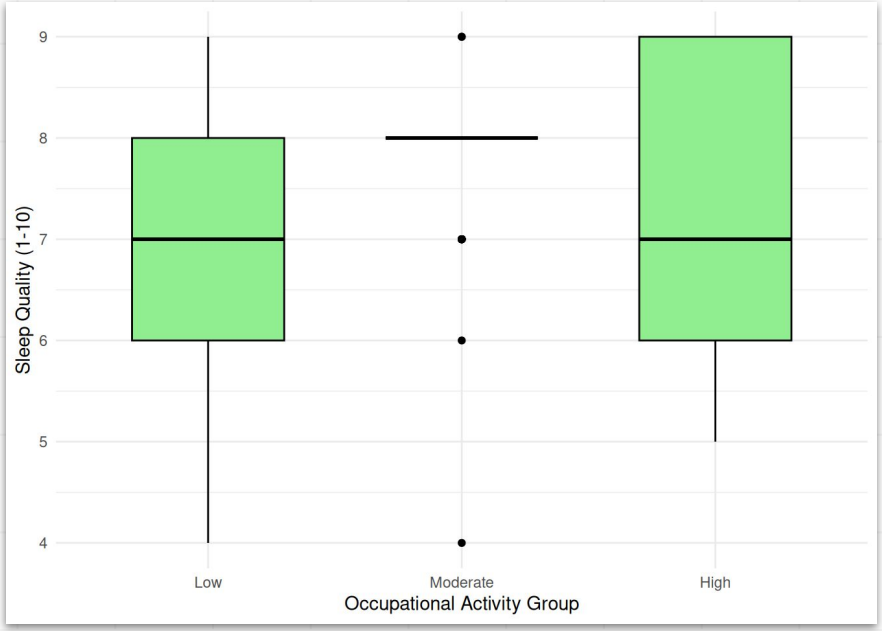
Activity Group	Sample Size	Mean (Hours)	Median (Hours)	SD	IQR
<fct>	<int>	<dbl>	<dbl>	<dbl>	<dbl>
1 Low	141	7.17	6.8	0.83	1.3
2 Moderate	89	7.25	7.2	0.37	0.3
3 High	144	7.02	7.3	0.93	1.9

A tibble: 3 × 6

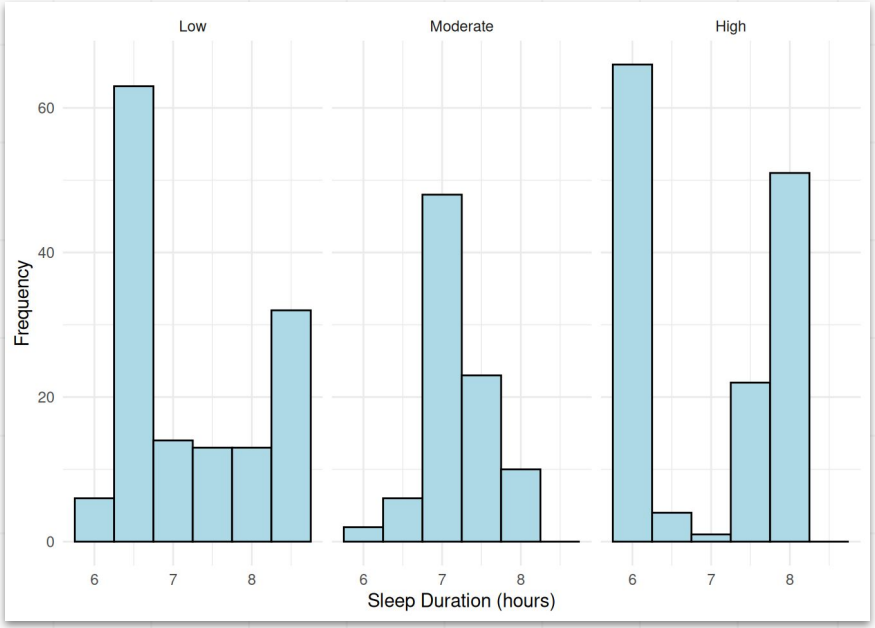
Activity Group	Sample Size	Mean Quality	Median Quality	SD	IQR
<fct>	<int>	<dbl>	<dbl>	<dbl>	<dbl>
1 Low	141	7.3	7	1.3	2
2 Moderate	89	7.82	8	0.59	0
3 High	144	7.01	7	1.27	3

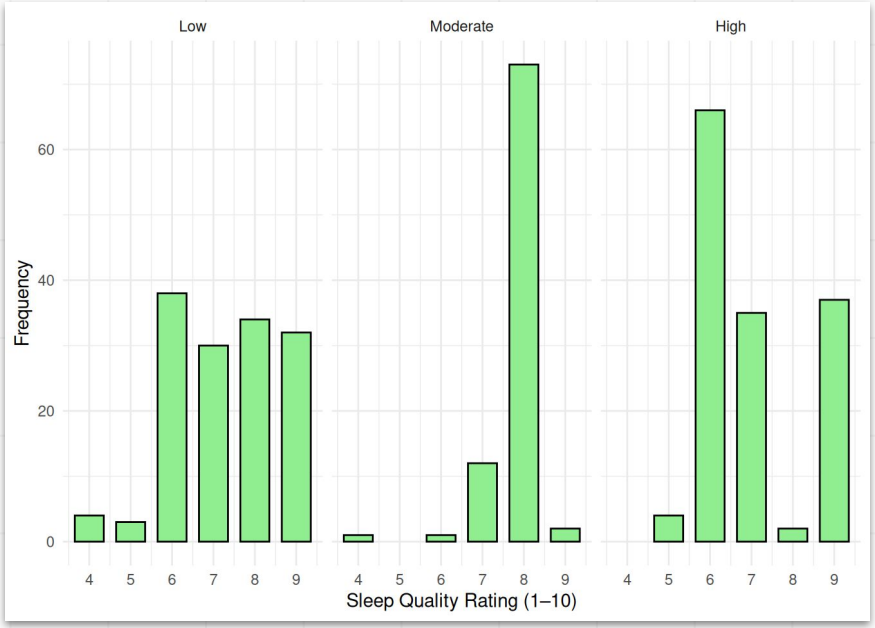
Distribution of Sleep Duration and Quality



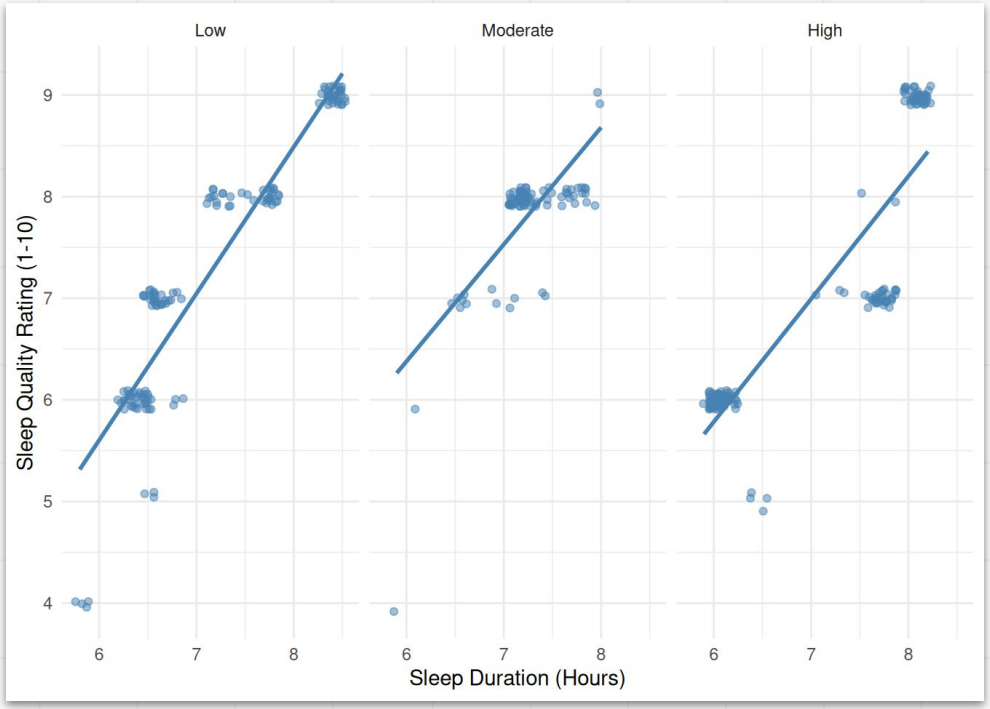


Distribution of Sleep Duration and Quality

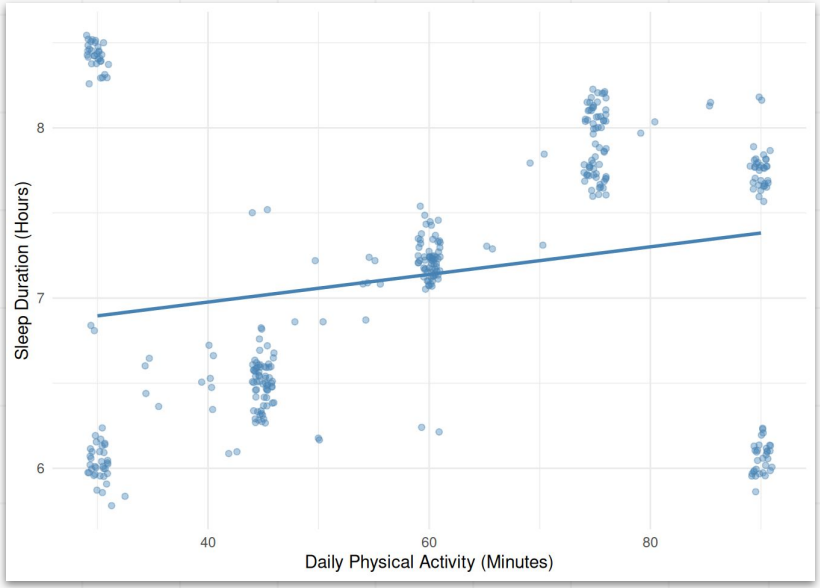


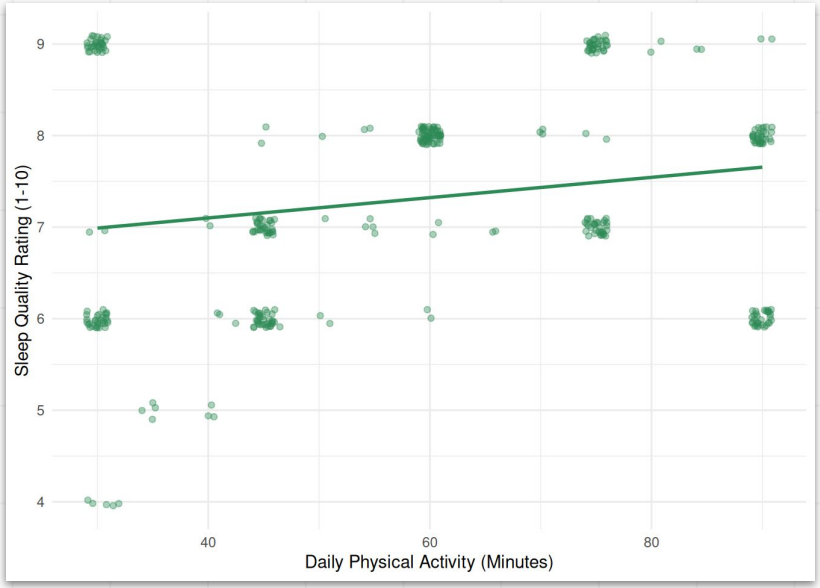


Relationship between Sleep Duration and Sleep Quality



Overall Trends in Physical Activity and Sleep Outcomes

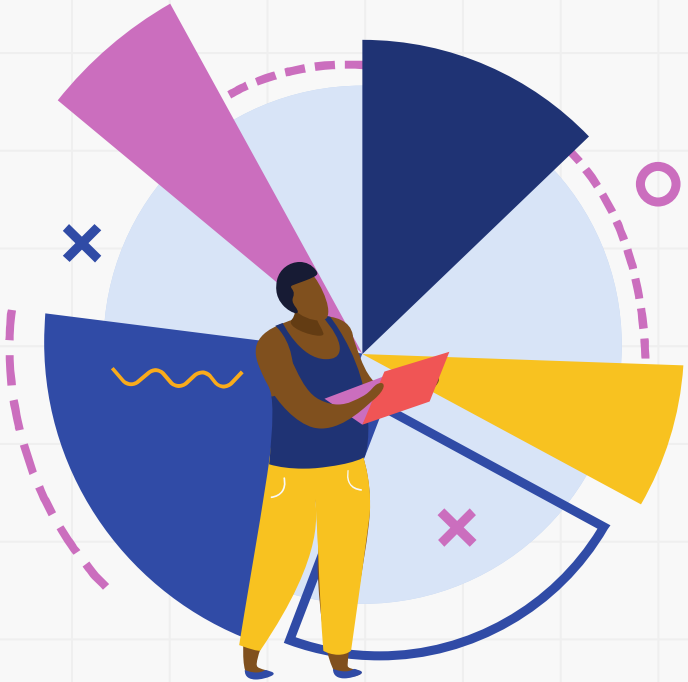




04

Discussion

Data Interpretation, Conclusion, and Critique



Highlights from EDA

- Occupations in the “High” physical activity group have:
  - Lowest mean quality of sleep
  - Lowest mean duration of sleep
- “Moderate” physical activity group has the **greatest average quality and duration of sleep**
- **High variability** within “Low” and “High” groups, **low variability** in “Moderate” group

Discussion and Conclusion

- **No clear trend** as physical activity level increases
- Generally, “Moderate” physical activity jobs associated with greatest sleep duration and quality
- Individuals from **any physical activity level** can have relatively high sleep duration and quality due to high variability

Critique and Limitation

- Having a “Moderate” job might be optimal for sleep but not necessarily
  - Our analyses **do not establish causal relationships**
- One **limitation**: analysis is based on a synthetic dataset; if the data deviates largely from real world information, our conclusions are not applicable to real people and jobs