Comprehensive Physical Health and Occupational Statistics Dataset

About

The data aims to explore the nuanced relationship between occupation, salary, physical health, and lifestyle while using reputable data from the U.S. Bureau of Labor Statistics to help discover potential patterns and correlations. The dataset comprises comprehensive information focusing on various job sectors including, healthcare, finance, legal professions, and more. Additionally, it encompasses diverse income brackets, offering a clear perspective on its correlation to health. This approach will allow us to observe health trends across different jobs and the average salaries. Our goal is to provide a quantified pattern of how the nature and remuneration of work influence sleep health, offering insights for both individuals and policymakers to foster healthy work environments and lifestyles, and encouraging work-life balance among specific occupations.

Data Creation Range: 2019-2023

Created By:
Ivette Ivanov (imivanov@uw.edu),
Lele Zhang (lelez@uw.edu),
Sera Lee (hayunl2@uw.edu),
Sohyun Han (sohyuh@uw.edu)

Content:
Statistical computation and combined data using R
Data frame CSV(Comma-Separated Values) files

Dataframe Sources:
sleep_df (Link inserted)
job_df (Link inserted)
conversion_df (Manually created by the authors)

Use Cases

Potential real-world applications of the dataset

1. Predicting quantifiable patterns of how the nature and compensation of work affects physical health.
2. Policymakers can utilize this data to formulate policies aimed at improving public health by addressing the work-related stress.
3. Promoting a healthy work environment and work-life balance across different industries.

Data Nutrition Label

Alert Count	**3**
Completeness Data contains the sleep duration, quality, physical activity, stress, BMI, average hourly and annual wage.	**0**
Provenance/Collection	**2**
Dataset Sources Bias The occupation dataset is from the US Bureau of Labor Statistics, while Sleep Health and Lifestyle Dataset from Kaggle may cause bias as a student research.	1
Unmatched Timeframe The May 2022 Occupation Dataset spans three years, while Sleep Health and Lifestyle Dataset is projected to be collected in 2023, last modified on April 25, 2023.	1
Composition	**1**
Socioeconomic Bias The Sleep Health and Lifestyle Dataset includes occupations limited to 11 professionals bias centered healthcare area. Meanwhile, Wage statistics obtained from Occupation Dataset may be negatively impacted by the COVID-19 pandemic.	1
Objective The dataset analyzes the overall health quality of each occupation including sleep health to provide information for practical implementation, fostering a healthier work environment, and improving individual's well-being.	**0**

INFO 201 Final Project Data Nutrition Label

About

The data aims to explore the nuanced relationship between occupation, salary, physical health, and lifestyle while using reputable data from the U.S. Bureau of Labor Statistics to help discover potential patterns and correlations. The dataset comprises comprehensive information focusing on various job sectors including, healthcare, finance, legal professions, and more. Additionally, it encompasses diverse income brackets, offering a clear perspective on its correlation to health. This approach will allow us to observe health trends across different jobs and the average salaries. Our goal is to provide a quantified pattern of how the nature and remuneration of work influence sleep health, offering insights for both individuals and policymakers to foster healthy work environments and lifestyles, and encouraging work-life balance among specific occupations.

Data Creation Range: 2019-2023

Created By:
Sera Lee - hayunl2@uw.edu
Sohyun Han - sohyuh@uw.edu
Lele Zhang - lelez@uw.edu
Ivette Ivanov - imivanov@uw.edu

Content:
Statistical computation and combined data using R
Dataframe CSV(Comma-Separated Values) files

Dataframe Sources:

Sleep_df (Sleep Health and Lifestyle Dataset)
*job_dfl (*Occupational Employment and Wage Statistics)
Conversion_df (Manually created by the authors)

Data Nutrition Label

Alert Count: 3

Completeness: 0

Data contains the sleep duration, quality, physical activity, stress, BMI, and average hourly wage, and average annual wage.

Composition: 1

Socioeconomic bias

The *Sleep, Health, and Lifestyle Dataset* includes occupations limited to 11 major groups with a notable bias towards healthcare-related professions. Moreover, it is important to acknowledge that the wage statistics obtained from the *May 2022 National Occupational Employment and Wage Estimates* may be negatively impacted by the COVID-19 pandemic.

Provenance/Collection: 2

Dataset sources

The occupation dataset was acquired from the US Bureau of Labor Statistics. The sleep health and lifestyle dataset was acquired from Kaggle for student research purposes and, therefore may contain bias.

Timeframe of data collection

The *May 2022 National Occupational Employment and Wage Estimates* dataset contains responses collected over three years: May 2022, November 2021, May 2021, November 2020, May 2020, and November 2019. The *Sleep, Health, and Lifestyle Dataset* is estimated to be collected in 2023, based on the data that was last modified on April 25, 2023.

Objective: 0

The dataset analyzes the overall health quality of each occupation to provide information for practical implementation, fostering a healthier work environment, and improving individual's well-being.

Use Cases

Potential real-world applications:

1. Predicting quantifiable patterns of how the nature and compensation of work affects physical health
2. Policymakers can utilize this data to formulate policies aimed at improving public health by addressing the work-related stress
3. Promoting a healthy work environment and work-life balance across different industries