

DAB 304 Project Proposal

Project Title: Analyzing Mental Well-being in a Remote Work Era

Course: DAB 304 (HealthCare Analytics)

Instructor: Abiodun Sodiq Shofoluwe

Group: 13

Group Members:

Bassey Esin Elliot – W0859047

Jensen Paul – W0875856

Papa Kofi Arhin – W0865495

Kayode Babalola – W08575164

Harini Gunasekaran – W0873771

Date: January 27, 2026

1. Introduction

Project Overview

This project analyses the impact of remote work on mental well-being using healthcare data analytics techniques. Since the COVID-19 pandemic, remote and hybrid work models have become widespread. While remote work offers flexibility and reduced commuting time, it also presents challenges such as isolation, burnout, blurred work-life boundaries, and reduced social interaction (WHO, 2022).

We will examine mental health indicators such as stress levels, job satisfaction, work-life balance, productivity, and burnout among remote workers. Using exploratory data analysis (EDA), statistical techniques, and data visualization, we aim to identify patterns and relationships that influence mental well-being in remote environments.

Related Work

Research suggests remote work can improve autonomy and work-life balance, but may also increase loneliness and emotional exhaustion (OECD, 2023). Reports from Gallup (2023) and WHO (2022) highlight rising levels of workplace stress and burnout in hybrid and remote settings.

Our project builds on existing literature by applying practical data analytics methods to identify measurable trends and correlations in mental health data.

Connection to the Course

This project aligns directly with DAB 304 objectives by applying:

- Data cleaning and preprocessing
- Exploratory data analysis
- Statistical measures and correlation analysis
- Data visualization and dashboard development
- Interpretation of healthcare-related datasets

It demonstrates real-world application of healthcare analytics concepts to a societal issue.

2. Motivation

Mental health in the workplace is a growing public health concern. Workplace stress and burnout significantly affect productivity, absenteeism, employee turnover, and long-term health outcomes (WHO, 2022; Gallup, 2023). These impacts extend beyond individuals to organizations and society.

Through this project, we aim to identify key stressors and protective factors within remote work environments. By generating data-driven insights, we hope to contribute recommendations that support healthier remote work policies.

Personal Learning Motivation

From a data analytics perspective, this project allows us to:

- Work with real-world, human-centered datasets
- Strengthen statistical analysis and visualization skills
- Practice interpreting behavioral and survey-based data
- Apply healthcare analytics techniques to a relevant issue

The topic is also relatable to students and professionals experiencing remote or hybrid work environments.

3. Evaluation

Successful Outcome

- The project will be considered successful if it identifies statistically significant relationships between remote work conditions and mental health outcomes using the Kaggle Remote Work and Mental Health dataset. The primary objective is to examine whether factors such as working hours, work-life balance, job satisfaction, and work arrangement (remote, hybrid, in-office) are significantly associated with stress, burnout, and loneliness levels.
- Descriptive statistics (mean, median, and standard deviation) will be used to summarize the dataset and identify overall trends. Pearson correlation analysis will assess relationships between continuous variables, while Spearman correlation will be applied when variables are ordinal or not normally distributed.
- To evaluate differences in mental health outcomes across work arrangements, ANOVA will be conducted when stress or burnout is measured on a numerical scale. If burnout or stress is categorized (for example, high versus low), a Chi-square test of independence will be used to examine associations between categorical variables.
- Multiple linear regression will be applied to determine the impact of workplace factors such as working hours, job satisfaction, and work-life balance on stress or burnout levels while controlling for other variables.

- In addition, a Logistic Regression model will be implemented to predict the likelihood of high burnout or stress levels. Model performance will be evaluated using accuracy, F1-score, and ROC-AUC for classification models. If a continuous outcome is modeled, R² and RMSE will be used for evaluation.
- Success will be measured by statistically significant findings (p < 0.05), strong model performance, clear visualizations, and logical, data-driven conclusions aligned with the research objectives.

4. Resources Dataset

Primary dataset:

Waqi786. (2023). *Remote work and mental health dataset*. Kaggle.

<https://www.kaggle.com/datasets/waqi786/remote-work-and-mental-health>

Additional sources:

- Gallup workplace reports
- World Health Organization reports
- OECD employment outlook reports
- Academic literature

Tools and Software

- Microsoft Excel (data cleaning and pivot analysis)
- Python (pandas, NumPy, matplotlib, seaborn)
- Jupyter Notebook
- Tableau (dashboards)
- GitHub and Jira for project management

Computational Requirements

A standard personal computer with a Python environment installed.

5. Contributions

To ensure balanced workload distribution:

- **Kayode & Bassey:** Data collection, preprocessing, EDA, correlation analysis, GitHub management
- **Harini & Jensen:** Visualization development, dashboards, interpretation, drafting findings, Jira and project charter
- **Papa:** Literature review, final editing, presentation preparation

All members will contribute to report writing and proofreading to maintain consistency.

6. References

- Gallup. (2023). *State of the global workplace report 2023*. Gallup Press.
- OECD. (2023). *OECD employment outlook 2023*. OECD Publishing.
- Waqi786. (2023). *Remote work and mental health dataset*. Kaggle.
- World Health Organization. (2022). *Mental health at work*.