# Remote Work On Mental Health

By

Robert Solomon

Supervised By

Dr. Vinayak Deshpande

# Introduction

Research Topic: Investigating the Impact of Remote Work on Mental Health using Data Science Approaches.

The shift to remote work has been accelerated due to global events like COVID-19.

This project aims to explore both qualitative and quantitative data to understand the mental health implications.
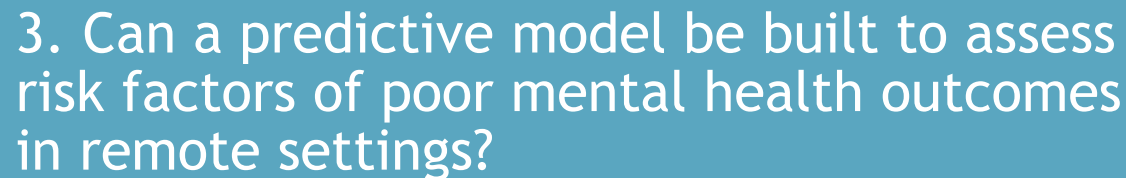
# Research Questions

1. How does remote work influence mental health indicators such as stress, isolation, and burnout?

2. Is there a statistically significant correlation between remote work and reported anxiety levels?

3. Can a predictive model be built to assess risk factors of poor mental health outcomes in remote settings?

# Methodology

Primary data collection using online surveys distributed to remote/hybrid and onsite workers.

Secondary dataset sourced from publicly available mental health research databases. (Kaggle)

Statistical tools and techniques used: correlation analysis, regression modeling, and ANOVA testing.

# Data Preprocessing & Analysis

Data Cleaning: Removed missing/null entries, standardized responses.

Used Python libraries (Pandas, NumPy) for preprocessing.

Correlation, Heatmaps, Histograms and Boxplots were used for EDA.

Linear Regression & ANOVA applied to test hypothesis.

# Data Preprocessing & Analysis for Primary Data

- Snapshot of Primary Dataset (Survey) highlighting incomplete survey responses (missing data shown in green).

# Data Preprocessing & Analysis for Secondary Data

- Secondary dataset sourced from publicly available mental health records

- Outlier detection applied across numeric variables
  - Notable outliers in:
  - Age
  - Years of experience
  - Virtual meetings per week
  - Hours worked per week

- Preprocessed for ANOVA & regression modeling
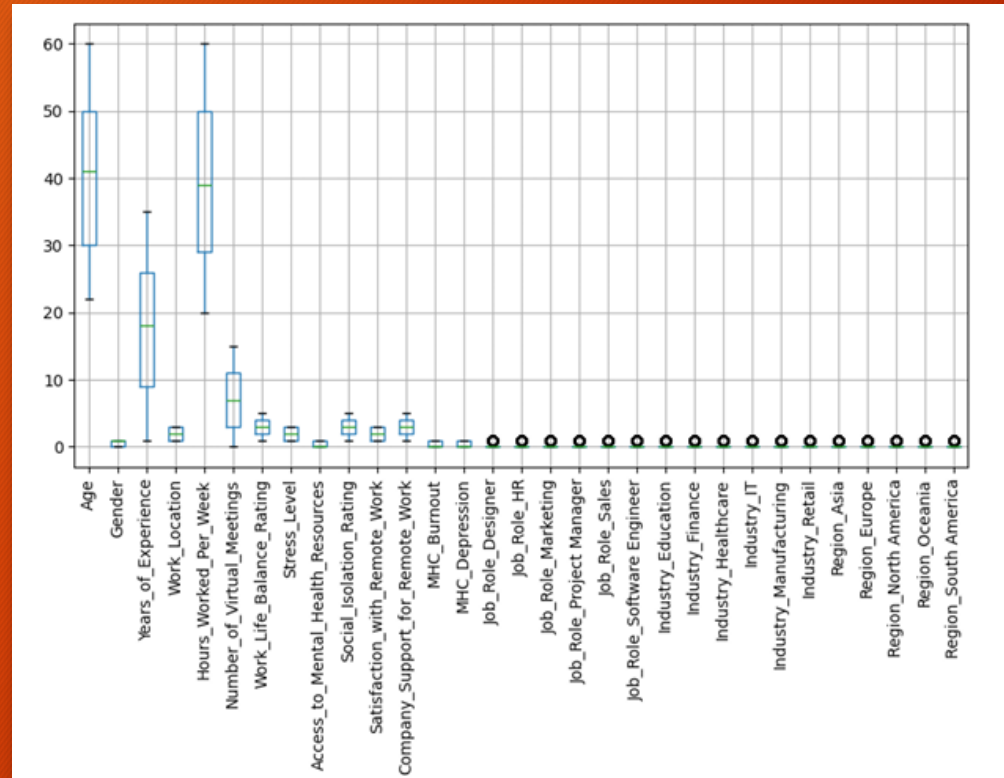
# Analysis & Findings - Secondary Data

Findings showed a moderate negative correlation between number of hours worked remotely and social connectedness.

Regression indicated that age and workload were significant predictors of mental health score.

# Analysis & Findings - Secondary Data
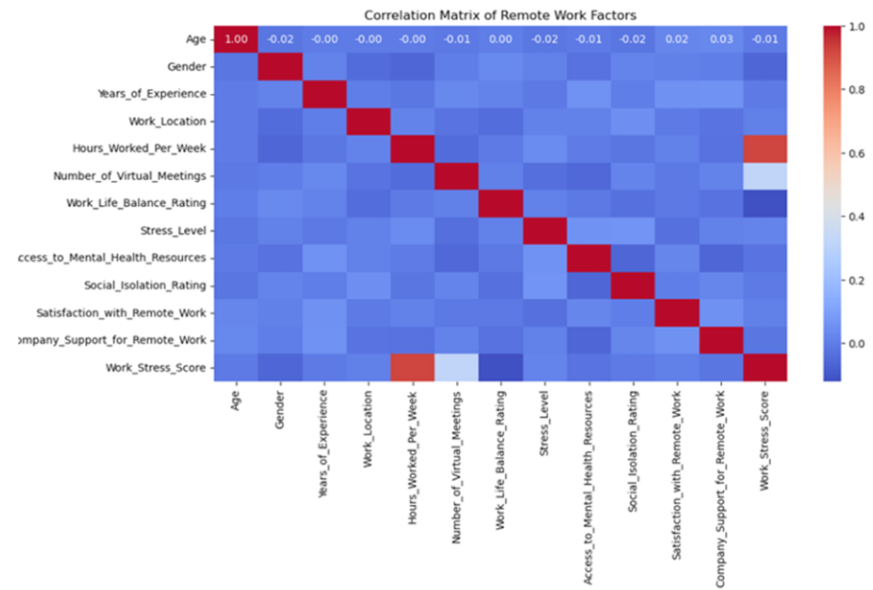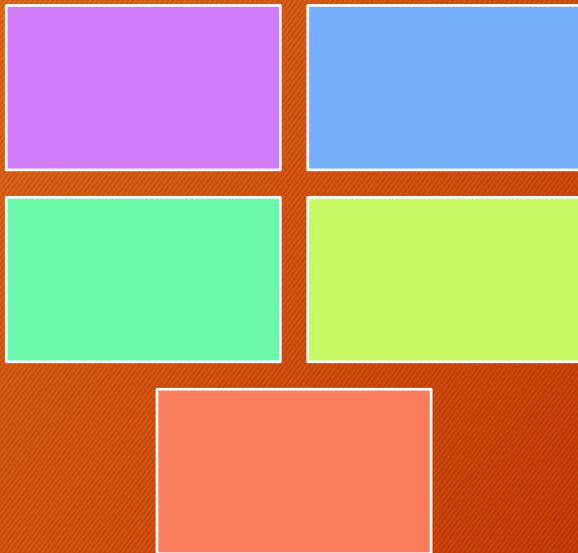


Figure 3 - Heatmap for correlation matrix of Remote Work factors (Secondary Data)

# Primary & Secondary Analysis & Findings

Key findings from primary data

Key findings from secondary data

# Model Comparisons for Secondary Dataset

| | Model | Score_Type | Score_Value |
|---|---|---|---|
| 0 | OLS Regression | R-squared | 0.0060 |
| 1 | Binary Logistic Regression | Pseudo R-squared | 0.0001 |
| 2 | Decision Tree | Accuracy | 0.5860 |
| 3 | Random Forest | Accuracy | 0.6250 |

Figure 6 - Bar chart of Model Comparisons (Secondary Data)

# Model Comparisons for Secondary Dataset

# Discussion

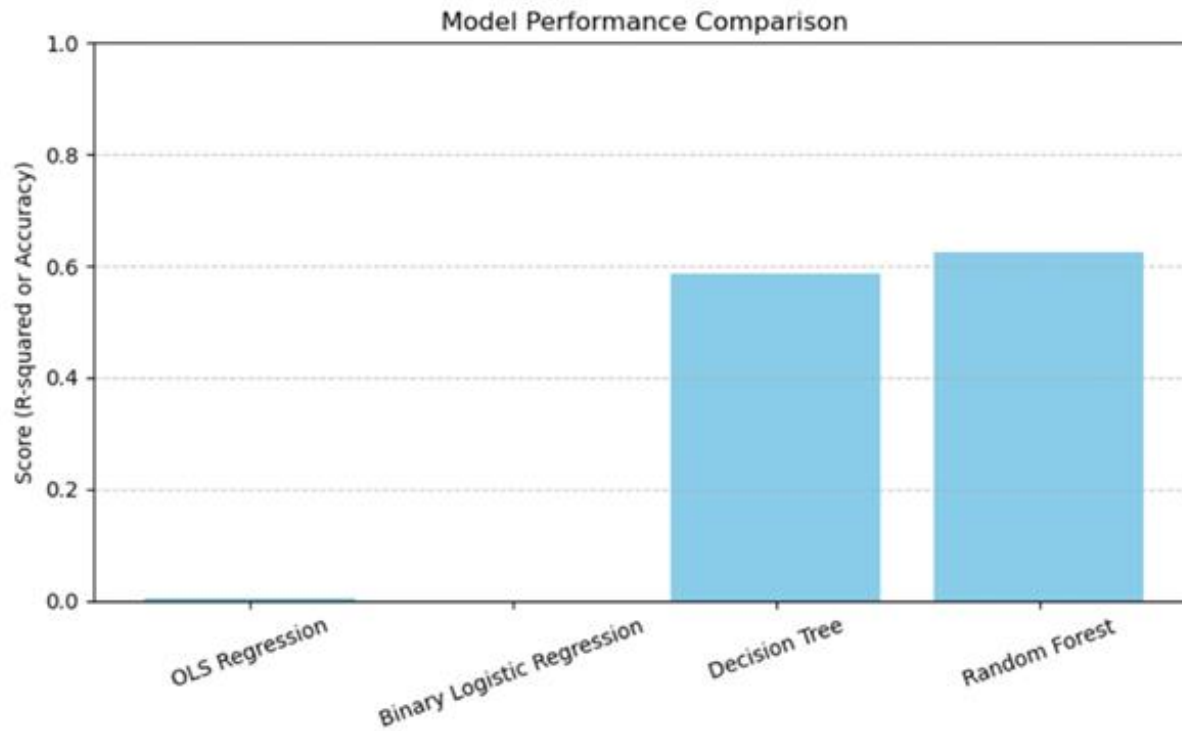- Interpretation of results

- Comparison with existing literature

- Implications

# Conclusion & Recommendations

Conclusion: Remote work presents both risks and benefits—flexibility vs. isolation.

Recommendations: Employers should foster community and encourage regular check-ins.

# Self-Reflection & Challenges

Challenge: Planning and executing primary data collection required persistence and adaptability.

Data preprocessing for secondary data was time-intensive due to inconsistencies and noise.

Lessons learned: Importance of version control (used GitHub privately), and documenting all preprocessing steps clearly.

My Github: Robert-Solomon12

# Future Work

- Recommendations for future research

- Improvements and extensions

# References

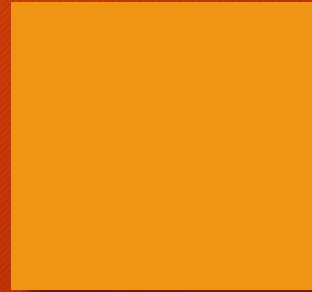APA / Harvard style references

Tools used for citation

# Thank You!

OPEN TO QUESTIONS?

# CONTACT INFORMATION

**Linkedin:** robertosolomon

**GITHUB:** ROBERT-SOLOMON12