Mental Health in Workplace

Cohort B Team 6 - Alvaro Chinchayan, Leighton Li, Andrey Lifar, Yoki Liu, Yue Ping, Sherry Zuo

Objectives

The following project explores the idea of mental health, with a particular focus on its relation to the workplace environment from a country-level perspective.

We utilize survey data and perform various machine learning models in order to: predict whether an employee sought treatment for his/her mental health condition based on other factors obtained, whether there is correlation between mental health and whether or not the workplace provides access to mental health services.

Individuals that would benefit from our findings include human resource departments that can begin to think about the potential benefits/harms of the existence (or lack thereof) mental health services.

Introduction

- Perform supervised and unsupervised *machine learning* models

- Identify which supervised machine learning model obtained the highest accuracy for prediction

- Analyze characteristics in clusters where people have highest chance or lowest chance of seeking mental health treatment after segmentations

- Merge suicide rates and world happiness score back to survey data in order to find correlation with mental health issues by using *linear regression*

Exploratory Data Analysis
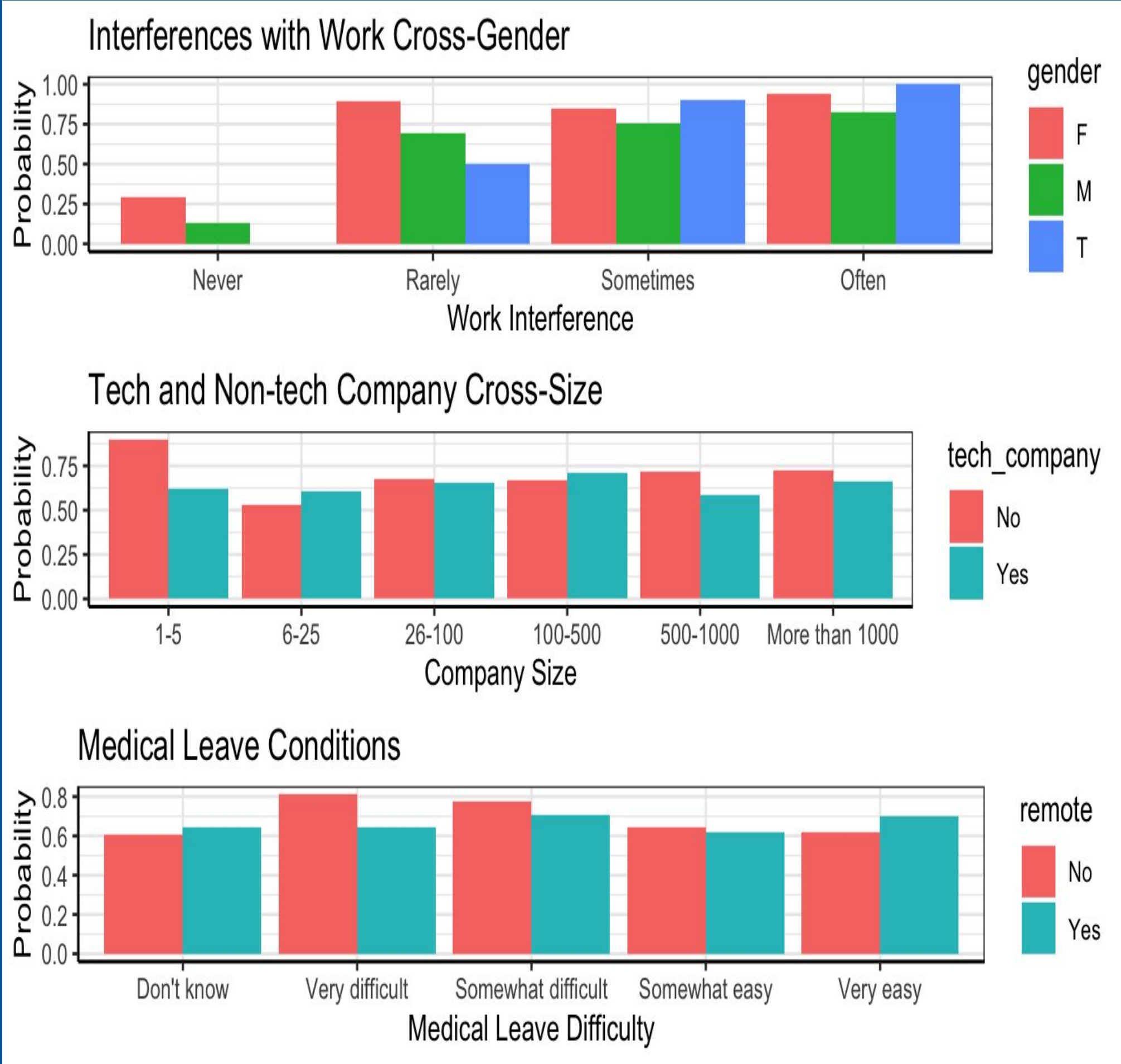
First plot shows respondents who sought treatment and either felt mental health often/sometimes/rarely/never interferes with work, across different genders.

Second plot shows the portion of respondents who sought treatment and worked for either tech or non-tech companies across different company sizes.
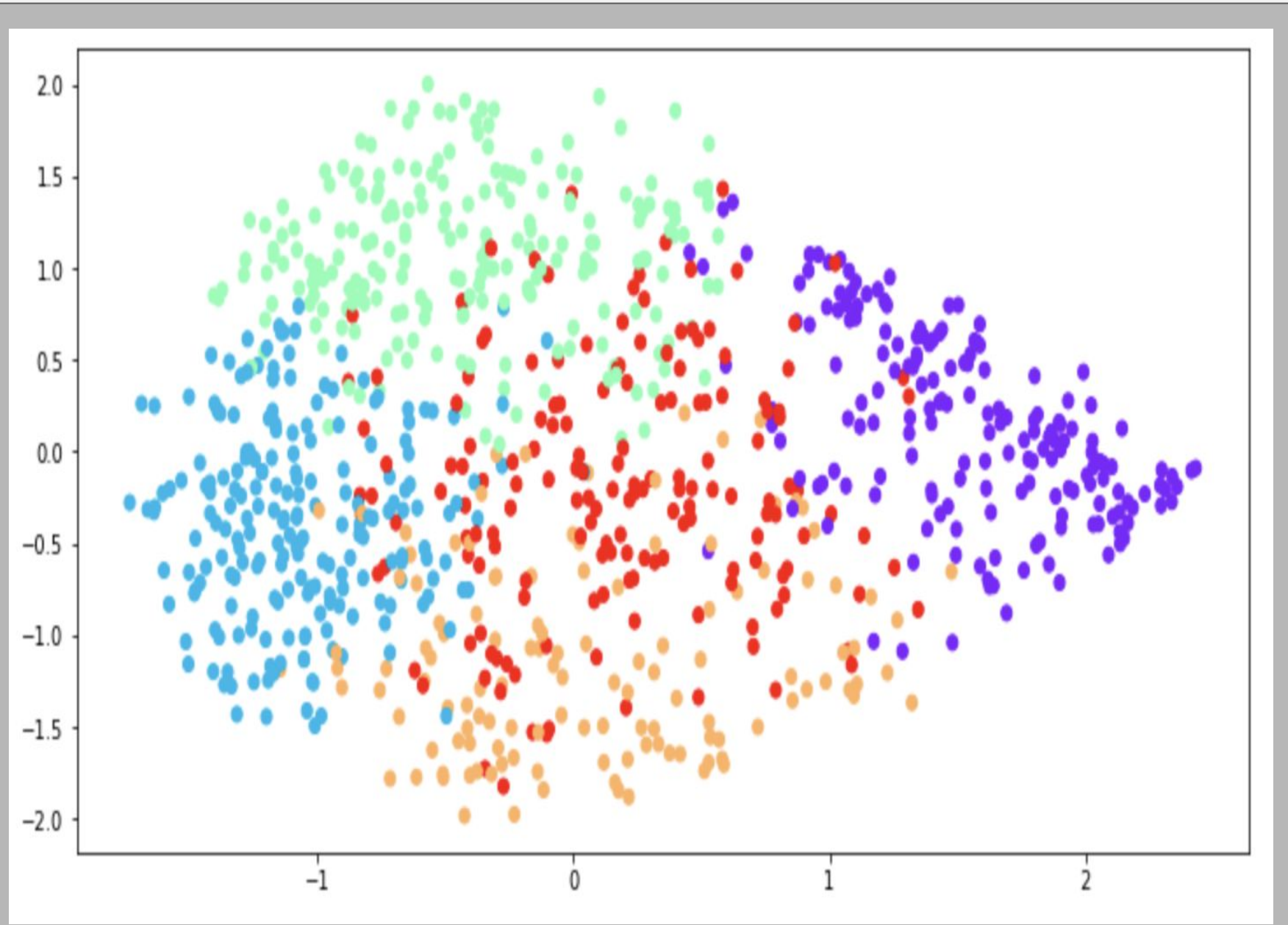
On average, despite medical leave difficulty, non-remote workers still have a high chance of having sought mental health treatment.
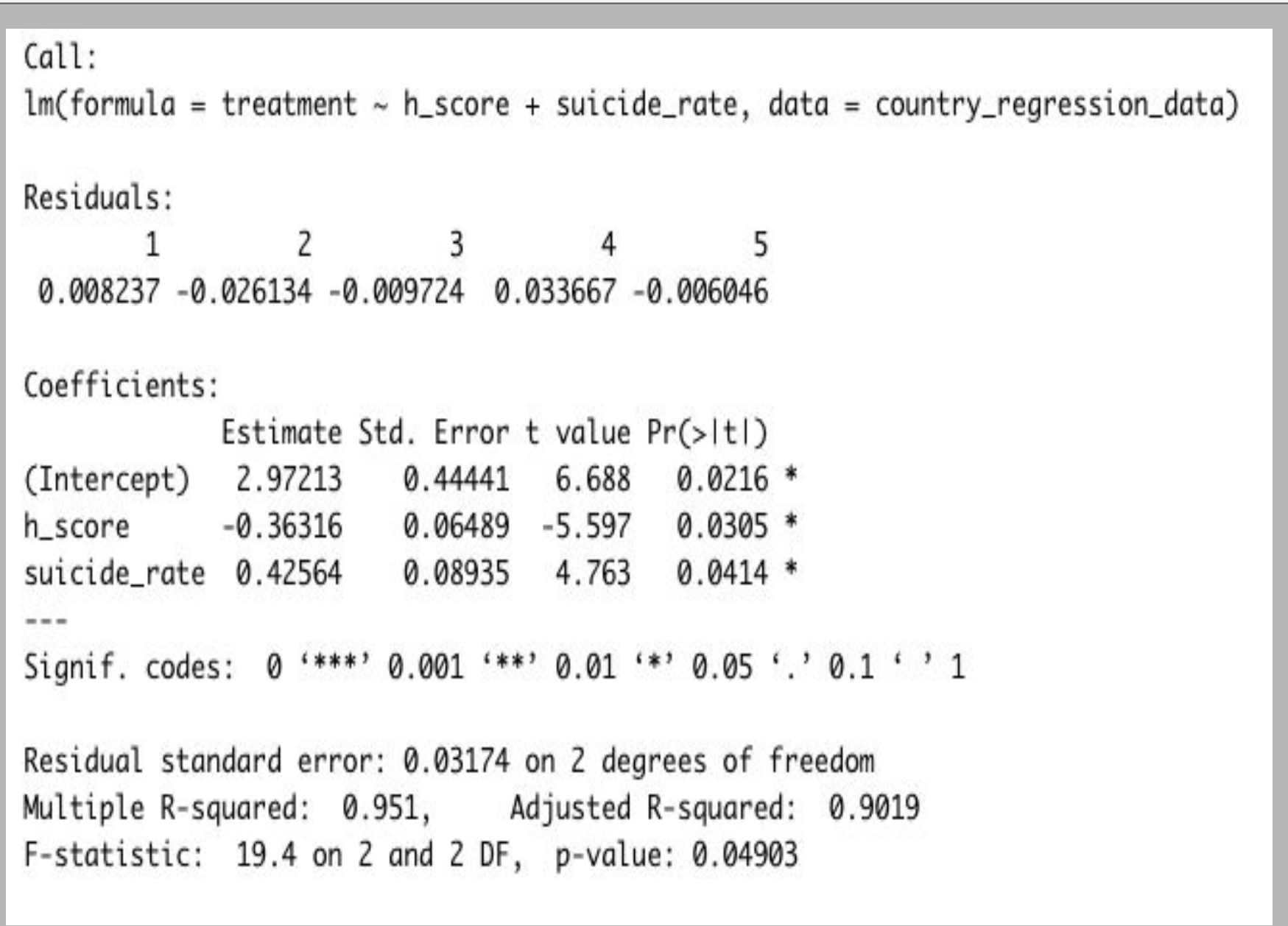
Results
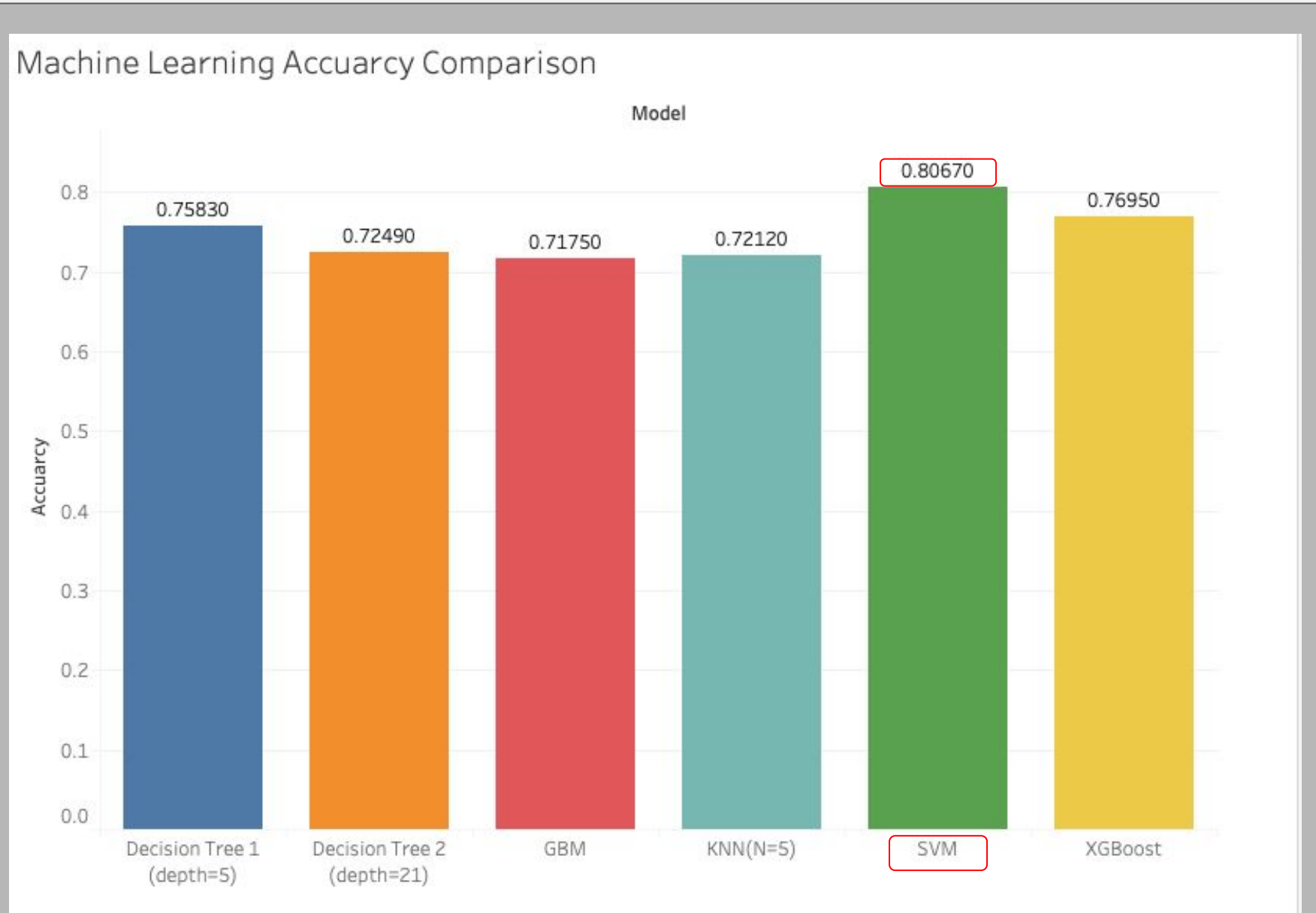
The bar chart above compares different supervised machine learning models' accuracies.

The top 3 models by accuracy scores are SVM, XGBoost, and Decision Tree with depth=5.

Unsupervised Machine Learning:

The elbow method with distortion and inertia suggests K = 5. We asserted this by comparing segmentation with and without PCA to choose 5 clusters from k-means with PCA of Dimension 12.

Linear Regression with suicide rate & happiness score as IVs, treatment as DV.

R^2 is 0.9 which suggests the result is relatively reliable. P-values for both coefficients are statistically significant.

Discussion

Based on our findings with particularly focus on the target variable, *treatment sought for mental health,* we obtained the following findings:

- Most respondents (65%) have sought treatment for mental health conditions.
- Based on our clustering results, the group with the highest mean rate of seeking mental health treatment has the following attributes: more likely to have family history of mental health issues, more likely to work for larger companies and consider discussing a mental/physical health issue with an employer would have negative consequences.
- Based on our linear regression model, family history of mental health conditions as well as access to mental health benefits/resources have an impact on employees' decisions on whether they sought for treatment.
- Based on survey Results from 2016 - 2019, only about 1% of responders feel comfortable discussing their mental health conditions, while about 60% of them feel more comfortable discussing their physical health.

Takeaways

- Corporations should raise more awareness for providing access to mental health support;

- Corporations should help employees understand that both mental health and physical health are equally important;

- Corporations should encourage employees to utilize and take advantage of existing mental health support programs