Predicting Incidence of Increasing Mental Health Disease in Relationship to the COVID-19 Pandemic

Team 109: Delta 6

Andrew Piscioneri, Eric Nagel, Jeffery Zhang, Jerrod Pelley, Keven Ayers, and Vipul Kotti

Motivation/Introduction

Information about the COVID-19 pandemic has been made readily available. While people have done different analyses on the impact of COVID-19 and the workplace, there is a lack of understand in how the COVID-19 Pandemic will continue to impact the remote workforce. In addition, with the ongoing pandemic, mental health has been put into focus, especially the lack of resources and equity. **Using CDC COVID-19 data and various sources from the US Census Bureau, we questioned the impact COVID-19 had on the mental health of the US, especially the remote workforce**. This information has the potential to enable the distribution of mental health resources to areas of need before the system is overwhelmed

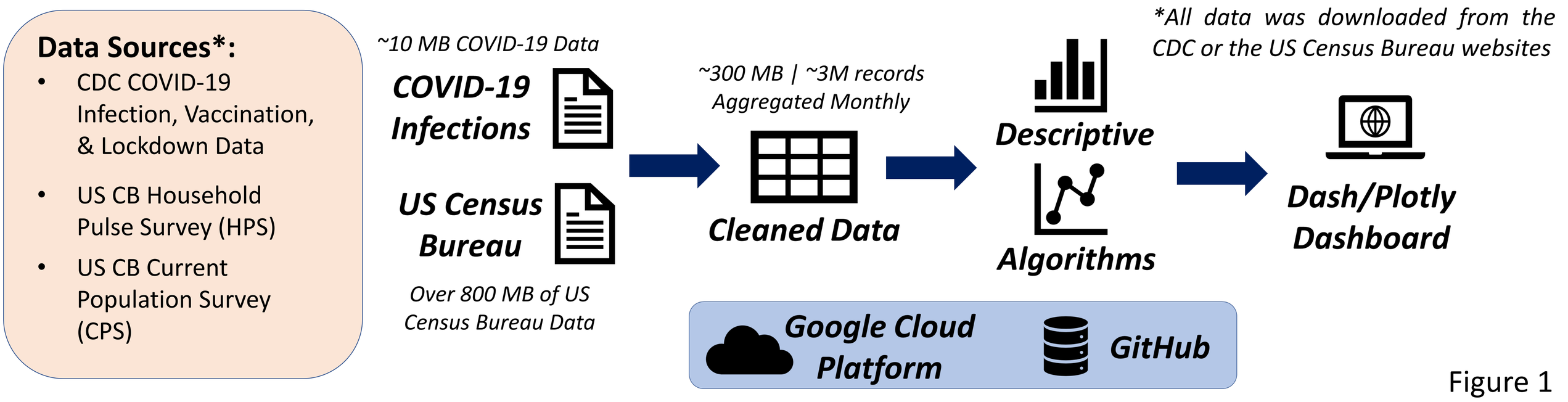

Figure 1

*In certain cases

Approach

Methods: All data was downloaded in CSV format, stored in **Google Cloud Platform (GCP)**. The data was joined and cleaned via **DB Browser for SQLite** and **Python**. Variable selection via PCA was used to eliminate correlated and unnecessary predictor*. Multiple ML models were tested with **XGBOOST Classification** enabling the best individual and **PCA Linear Reg.** enabling the best aggregate prediction of our mental stress index (MSI), which determines low/high risk. Therefore, using COVID-19 and US Census Bureau parameters, we were able to predict mental health status using our MSI. **Innovation:** To the team's knowledge, no one has previously attempted to predict mental health status based on COVID-19 infections and remote workplace status.

Experiments & Results

Individual MSI Results			
Model	MAE	RMSE	Acc.
XGBoost	**.20777**	**.45582**	**79.22%**
Random Forest	.20782	**.45582**	**79.22%**
Naïve Bayes	.21237	.46084	78.76%
MLP NN	.20967	.45790	79.03%
Log. Regression	.20967	.45790	79.03%
KNN	.20960	.45790	79.04%

Table: 1: Individual Model Predictive Accuracy

Aggregated MSI Results			
Model	MAE	RMSE	Acc.
XGBoost	.24457	.08756	30.50%
MLP NN	.21419	.07209	42.78%
PCA Lin. Reg.	**.18180**	**.05165**	**59.00%**

Table 2: Aggregate Model Predictive Accuracy

The impact remote work has on the mental health of individuals is mixed. Regarding prediction of mental health risks of individuals based on COVID-19 information, as well as quantitative aggregation of mental health stress levels throughout the United States, the answer is much more promising. Using any of the classifying models, the prediction accuracy of determining if an individual is at high-risk of mental stress was impressive at 79.22% (Table 1). While the quantitative model only scored 59.0% accuracy (Table 2), it too is likely something that can be built upon or used as an initial guide for federal agencies and organizations to see which states have a higher overall Mental Stress Index aggregate value.