Table of Contents:

Proposal:

A. Subject of Course Project:

_____COVID-19: Using data mining applications to select clinical trial participants.

B. Business Problem Statement:

_____After months of research and development, a U.S. pharmaceutical company is ready to begin clinical trials on a new coronavirus vaccine. Before the vaccine can be approved, the company needs to determine how it performs on individuals who are at low risk of infection and those who are at a higher risk of infection. They need a fast, efficient means of selecting 2,000 participants for the trial. The participants should be selected from a range of significant demographics and characteristics. That way, the company is able to test the effectiveness of the vaccine on individuals from different age groups, ethnicities, geographic locations, and prior medical histories.

C. Company Name:

Group Four Technology LLC

D. Proposed Solution:

_____In order to find participants for the clinical trial, the company creates an application that integrates a number of different databases into a single data warehouse. These databases can include census data, data on COVID-19 infections and deaths, data from the company's prior clinical trials, and medical data from insurance companies and hospitals. With this integrated data warehouse, researchers should be able to draft an analysis of the effects of COVID-19 on the U.S. population. Such an analysis will identify those demographics most vulnerable to infection as well as those least vulnerable. From there, researchers can select their target