Estimate Medical Equipment required for Heart patients in Covid19 environment.

- By Simi Sudhakaran

04/25/2020

VERSION 1.0 | 04/25/2019

Saint Peter's UNIVERSITY
The Jesuit University of New Jersey

1. Introduction 3
Purpose 3
Goal 3
2. Assumptions and Constraint 4
3. Implementation steps 5
4. Visualizations: 7
State wise % Prevalence of Heart Disease from 2011 – 2017 7
Predicting state wise 2020 % Prevalence of Heart disease. 7
Plot a state-wise comparison graph with the maximum number of vulnerable population 8
Plot a heatmap to show the states that are most vulnerable 8
Elbow plot to show the number of clusters that can be created between the vulnerable population and number of staffed beds in each state 9
K-Mean cluster analysis of the vulnerable population against the staffed beds in each state 9
5. References 10
DISCLAIMER 11

1. Introduction

Purpose

The purpose of this project is to find the number of medical equipments required by each state in Covid19 environment. For this project we will specifically focus on heart disease and calculating the number of heart disease equipments required by health care providers. This estimation will be useful for healthcare providers and Government officials to estimate and prioritize the purchase of cardiac equipment in Covid-19 situation. Source of this dataset is from the National Cardiovascular disease surveillance data. For this project we will be concentrating specifically on Cardiovascular diseases. Although, this model can be replicated for any other diseases during a pandemic.

This particular dataset consist of data from all the states in USA from the timeframe 2011 to 2017. The 2020 data has not been published yet but we will predict it using the Linear regression model. We will also be clustering the state wise data using K-Mean clustering, so that we can prioritize the states which are at a higher risk and therefore require more medical equipment relative to other states.

Goal

The main purpose of this project is to filter and analyze the estimated percentage of prevalence of heart disease which will in turn help in estimating the total cardiac devices required in a given region.

2. Assumptions and Constraint

We will be using the current estimated mortality rate of 51% in patients with pre-existing heart conditions as published by the Medical association of China & South Korea.

Following dataset will be used for analysis:

- BRFSS - CDC Survey data from 2011 to 2017 disease report.
    - https://catalog.data.gov/dataset?tags=coronary-heart-disease
      File : Behavioral_Risk_Factor_Surveillance_System__BRFSS__-__National_Cardiovascular_Disease_Surveillance_Data.csv
- Census data : https://www2.census.gov/programs-surveys/popest/datasets/2010-2019/state/detail/SCPRC-EST2019-18+POP-RES.csv
- American Hospital Directory - Hospital Statistics by State

  https://www.ahd.com/state_statistics.html

3. Implementation steps

This analysis will be performed in the following steps:

1. We imported the BRFSS – National Cardiovascular Disease Data from the CVS file.
2. We then removed null and invalid data. Removed unnecessary columns.
3. We then iterated through all the rows and created a data-frame of all the states with its respective Percentage prevalence of heart disease for each year.
    - Here we added the disease related to heart disease each year. The resulting dataframe consisted of '**state'**, '**year'** and estimated **'%Prevalence of Disease'**.
4. Since BRFSS is based on survey, we did not have data for 2020. So now the goal was to predict the estimated value of '%Prevalence of Disease' in each state in 2020.
5. Using the prior year data and linear regression, we calculated the predicted value for 2020 of percentage prevalence of heart disease for each state.
6. We then Plot linear regression for each state from 2011-2017 and predicted percentage prevalence for 2020 for each state.

At this point, we have '%Prevalence of Heart Disease' in each state for 2020 stored in a data frame.

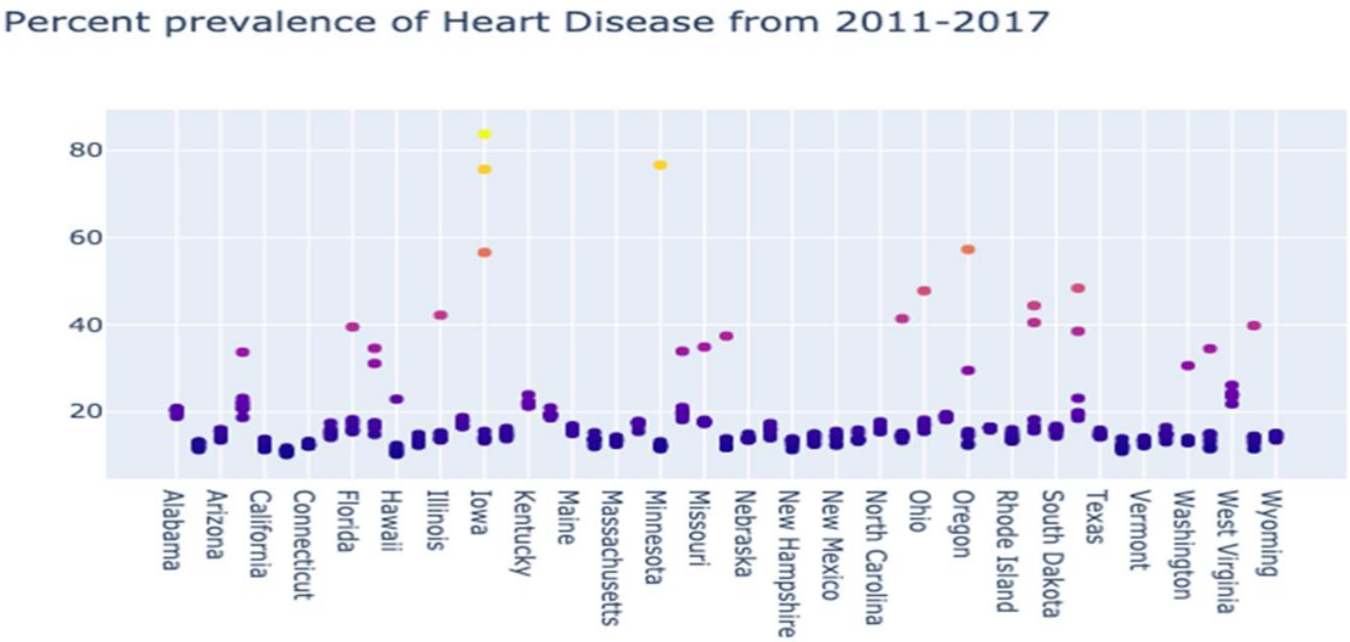
7. Next step is to get the current total US population and the no. of staffed hospital beds in each state.
8. We fetched the total US state wise population from the official US Cencus.gov website.
9. We fetched the total hospital bed data from the American Hospital Directory Ahd.com.
10. Since we predicted percentage prevalence for 2020 for each state, we can now multiply it with the actual state census data to get the current vulnerable population.
11. The current estimated mortality rate of 51% in patients with pre-existing heart conditions as published by the Medical association of China & South Korea. We will apply this to the vulnerable population to get the most vulnerable population per state.
12. We then created a dataframe with all the states, respective staffed beds and Covid-19 vulnerable population.
13. We then plot a state-wise comparison graph with the maximum number of vulnerable population
14. We then plot a heatmap to show the states that are most vulnerable.
15. Export the final dataset created to clean it
16. We then used the Elbow method of K-Mean Clustering to determine the number of cluster of vulnerable population

17. We lastly plotted the K-Mean clusters based on the k=3 value determined by the elbow method.

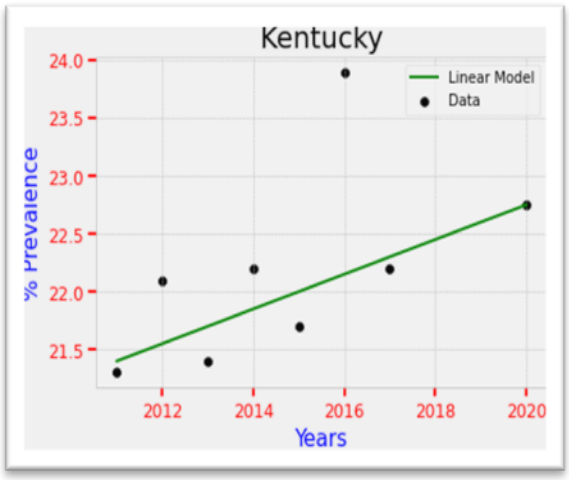
These estimations will be useful for healthcare providers and Government officials to estimate and prioritize the purchase of cardiac equipment in Covid-19 situation.

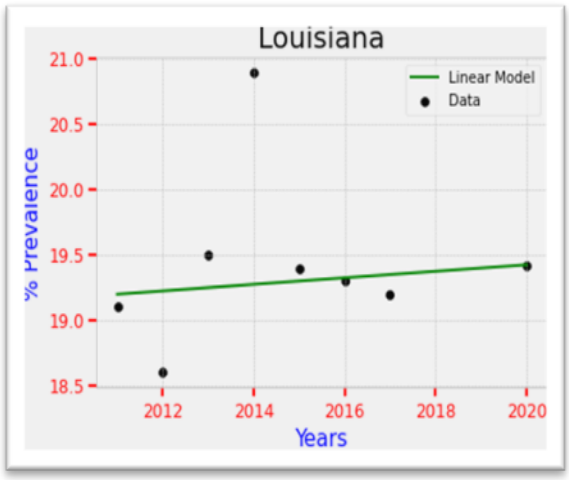
4. Visualizations:

State wise % Prevalence of Heart Disease from 2011 – 2017

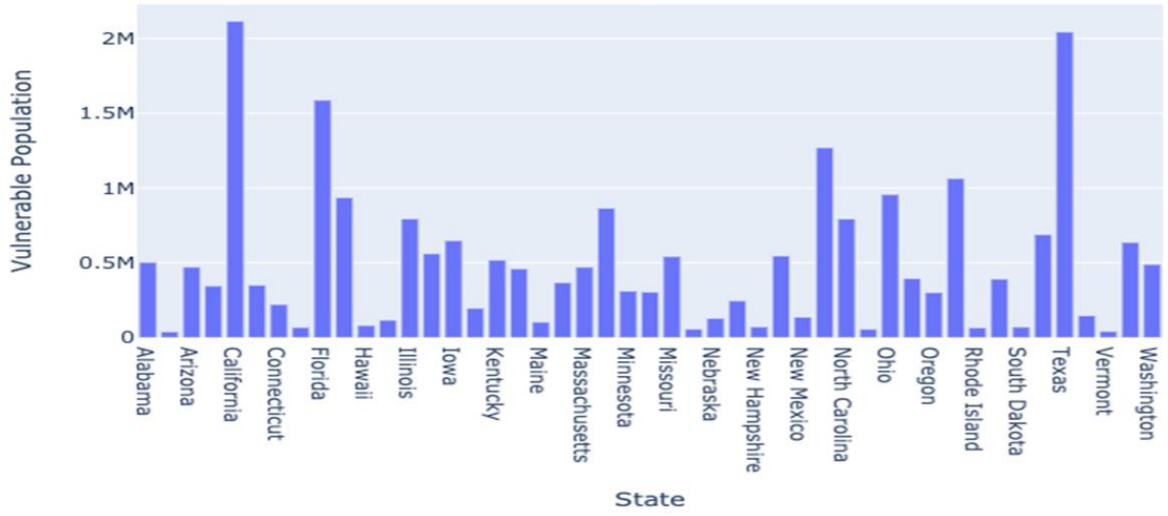


Predicting state wise 2020 % Prevalence of Heart disease.

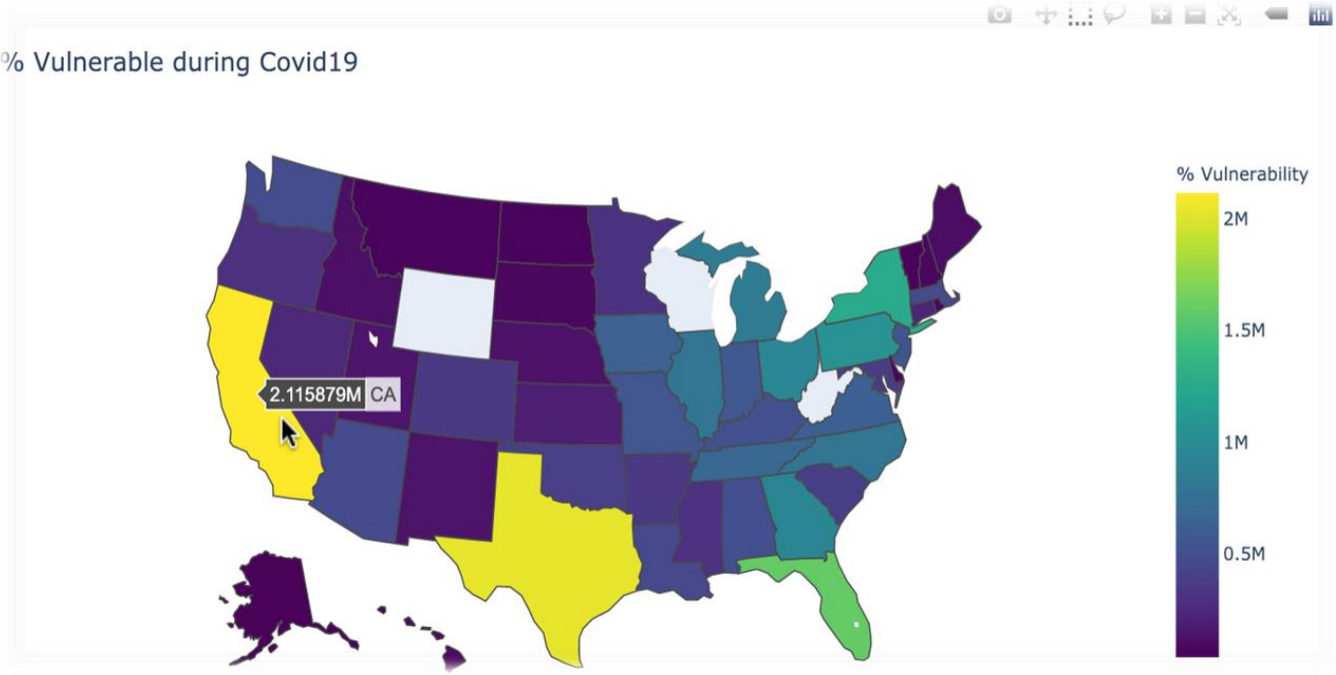




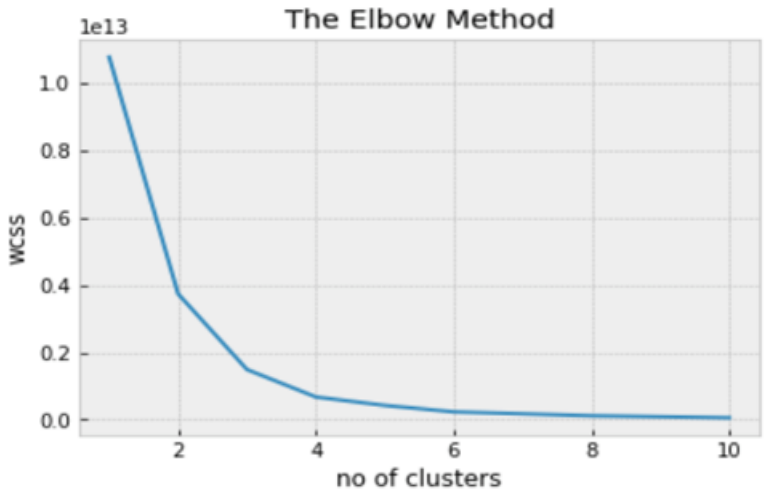
Plot a state-wise comparison graph with the maximum number of vulnerable population



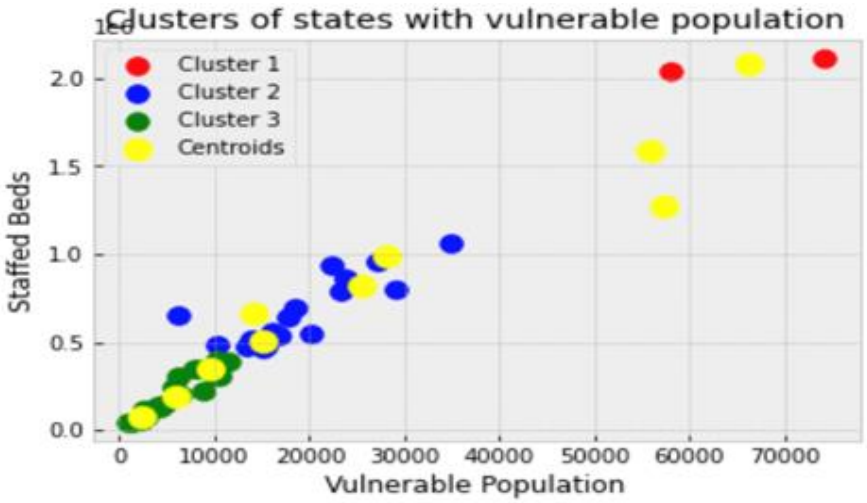
Plot a heatmap to show the states that are most vulnerable



Elbow plot to show the number of clusters that can be created between the vulnerable population and number of staffed beds in each state



K-Mean cluster analysis of the vulnerable population against the staffed beds in each state



5. References

https://www.cdc.gov/brfss/data_documentation/index.htm

https://catalog.data.gov/dataset?tags=coronary-heart-disease

https://www2.census.gov/programs-surveys/popest/datasets/2010-2019/state/detail/SCPRC-EST2019-18+POP-RES.csv

https://www.ahd.com/state_statistics.html

DISCLAIMER

Any articles, templates, or information provided is this document is purely for academic purpose. Actual results may vary. We strive to keep the information up to date and correct, and make no representations any kind, express or implied, about the accuracy, completeness, suitability, reliability, or availability with respect to the information and analysis articles, Any reliance you place on such information is therefore strictly at your own risk.