What tells you that you have a heart disease?*

An analysis of Heart Disease indicators

Puyu Liu (1005092971)

06 April 2022

Abstract

This paper analyzed the data of personal key indicators of heart disease. The result indicates that age, sex, BMI, etc are correlated to having a heart disease. Based on the analysis, this paper consists of a linear regression analysis, with the use of the programming language R. Keywords: Heart Diseases, Physical health, Mental Health, kaggle, age & sex, stroke, Smoking, Alcohol Drinking

1 Introduction

As heart diseases have caused most deaths for people in the US, this dataset include about 20 variables that involve several risk factors of heart disease. Around half of all americans have at least one heart disease indicator: blood pressure, high cholesterol, smoking, alcohol drinking, diabete and obesity (high BMI), physical and mental health. Hopefully we can derive a regression analysis to use the logistic regression model to make future predictions for population with heart disease risk factors. Computational developments have a great effect on detecting the early stage of some typical heart diseases.

The paper starts with an overview of the data with 300 thousands observations. And also, we will look into the correlation of every variable then try to conduct a regression analysis. Before that, the dataset should be properly cleaned and outliers and leverage should be well taken care of. After processing the data, we will be using the ANOVA test to determine our variables in our model.

2 Data

3 Model

4 Results

5 Discussion

5.1 First discussion point

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

*Code and data are available at: https://github.com/liupuyu2/Puyu-Liu.

Appendix

A Additional details

B References