

Proposal_11/11

Yimeng Cai

2023-11-08

Group member:

Yimeng Cai (yc3577), Zheshu Jiang (zj2379), Ze Li (zl2746), Qianying Wu (qw2418)

Project title:

Multiple Linear Regression Analysis on Socio-economic factors and Bioindicators of Type II Diabetes by using R

Project Motivation

Type II Diabetes is a global health concern that affects more than 37 million Americans and places a significant burden on healthcare systems. As a multifaceted disease, Type II diabetes was influenced by both socio-economic factors and biological indicators. By analyzing a comprehensive dataset from the Behavioral Risk Factor Surveillance System, we aim to uncover the correlations between diabetes prevalence and factors such as sex, education, income, lifestyle habits, and bioindicators like blood glucose and cholesterol levels. Through visualizations and statistical analysis in R, we aspire to highlight patterns that could inform better prevention and management strategies, thereby contributing to the broader dialogue on public health and socio-economic policies related to diabetes.

Data resources:

<https://www.kaggle.com/datasets/alexteboul/diabetes-health-indicators-dataset/data>

- Data Description:

The dataset was collected by the Behavioral Risk Factor Surveillance System (BRFSS) in 2015. This original dataset contains responses from 441,455 individuals and has 330 features. For this project, we use the one with 253,680 responses and a three-class diabetes target variable indicating no diabetes, prediabetes, or diabetes. The other potential factors for diabetes are blood pressure, cholesterol, cholesterol check, BMI, stroke, heart disease, and habits of smoking, physical exercise, eating fruit.

The planned timeline

| Timeline | Requirement | Predicted Timeline |
|----------|------------------------|--------------------|
| 11/13-17 | Project review meeting | Nov 15th or 16th |
| 12/9 | Report | Nov 24 |
| 12/9 | webpage and screencast | Dec 1-7 |
| 12/9 | peer assessment | Dec 7-8 |
| 12/14 | In class discussion | |