EDA Report

Personal Key Indicators of Heart Disease

EDA Report

Lina Li (jinghua@usc.edu)
Indira Lakshmi Gundavarapu (igundava@usc.edu)
Alyssa Pan (dingwenp@usc.edu)
Xinyao Tang (xinyaota@usc.edu)

Project Overview

Business Objectives

Our audience is some health institutions, or some health research centers. They can use our EDA report to select the attribute in their experiments. As for our purpose of the dashboard. We will find the relationship between heart disease and other attributes. Which attribute will have a greater impact, and the impact of different attribute combinations on heart disease.

Data Source

Personal Key Indicators of Heart Disease: 2020 annual CDC survey data of 400k adults related to their health status

Link: https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease

Attribute Grouping

Output:
- Heart Disease

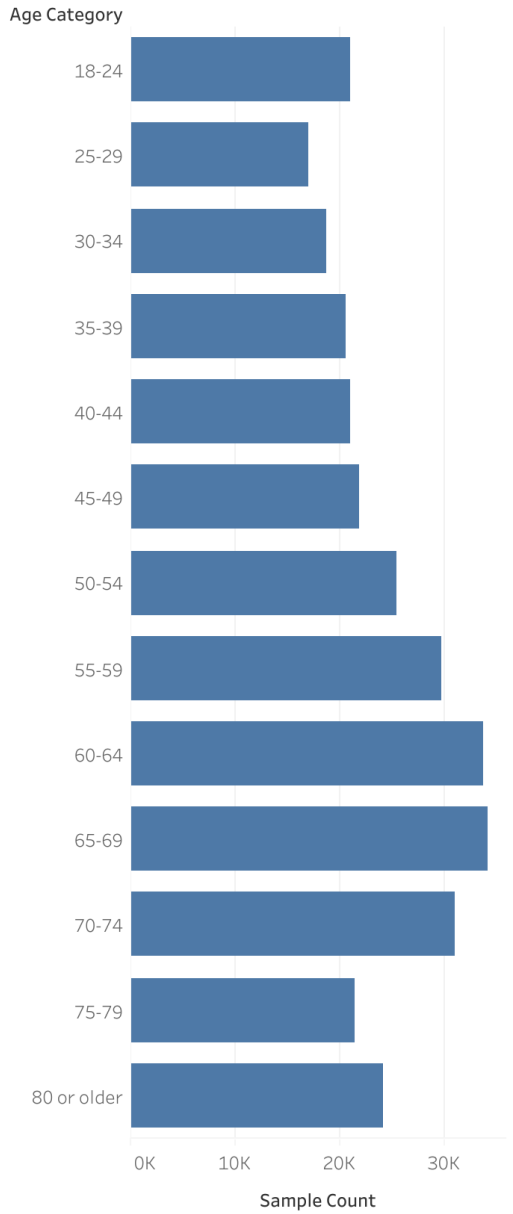
Demographic:
- Age
- Sex
- Race

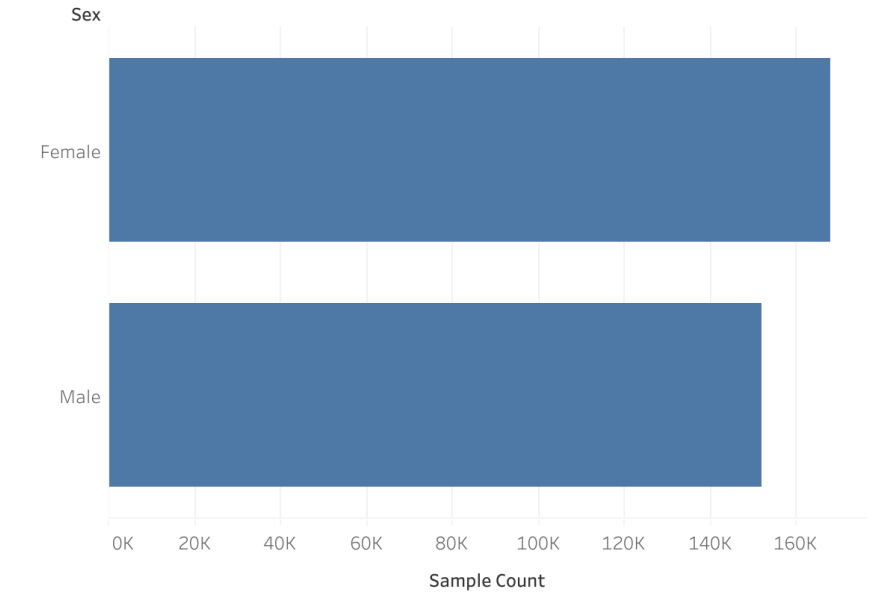
Habits:
- Alcohol Drinking
- Physical Activity
- Smoking

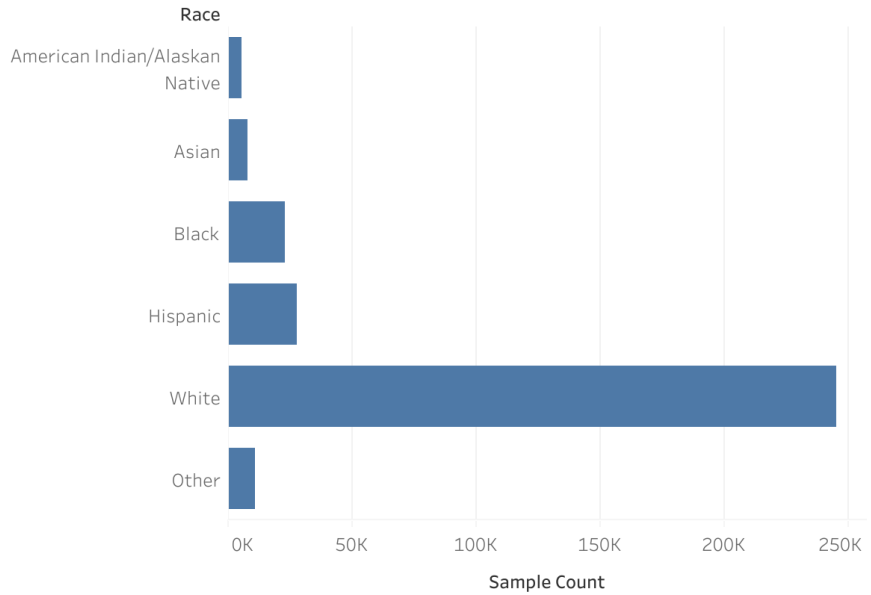
Health History:
- Asthma
- Diabetic
- Diff Walking
- Kidney Disease
- Gen Health
- Skin Cancer
- Stroke

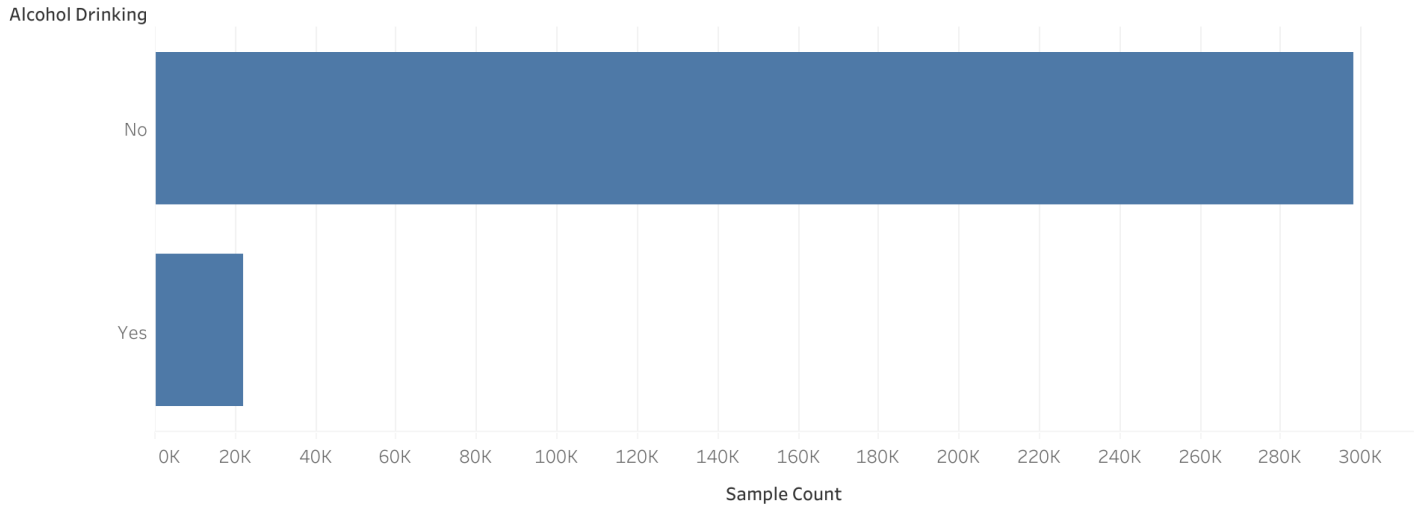
Habits:
- Sleep Time

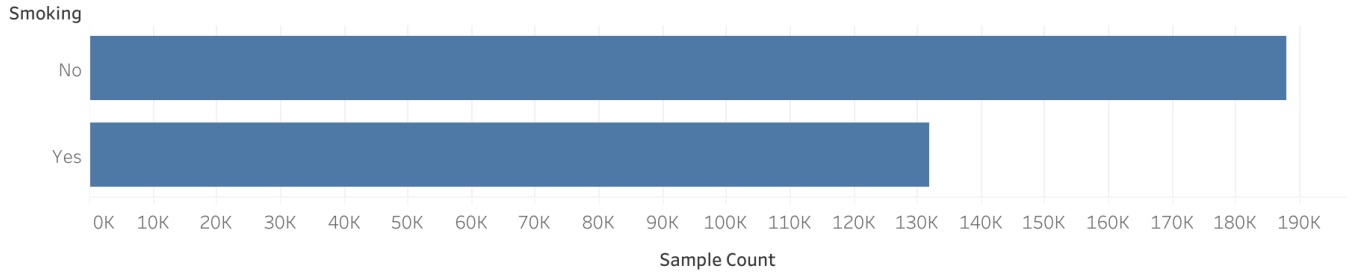
Health History:
- BMI
- Mental Health
- Physical Health

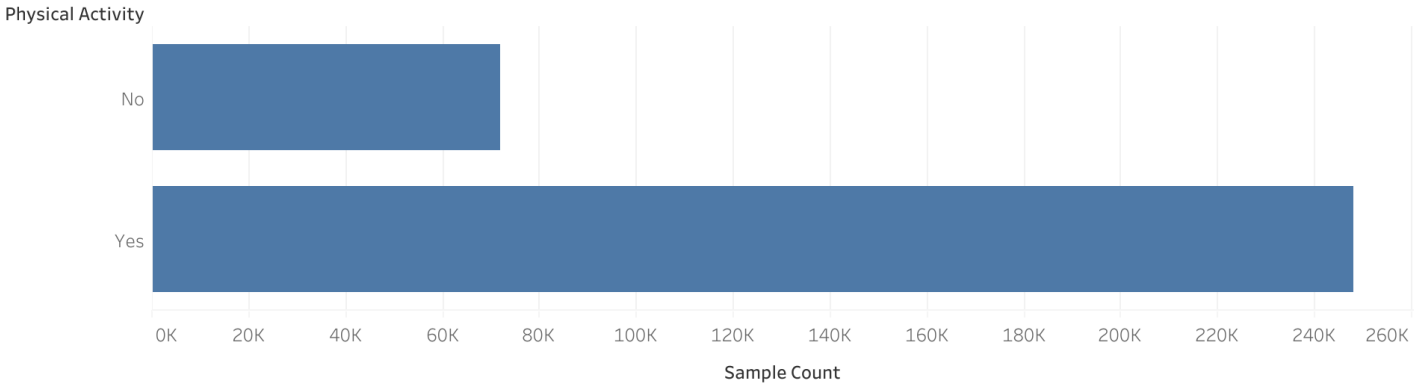

Age Category
18-24
25-29
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-69
70-74
75-79
80 or older
0K
10K
20K
30K
Sample Count

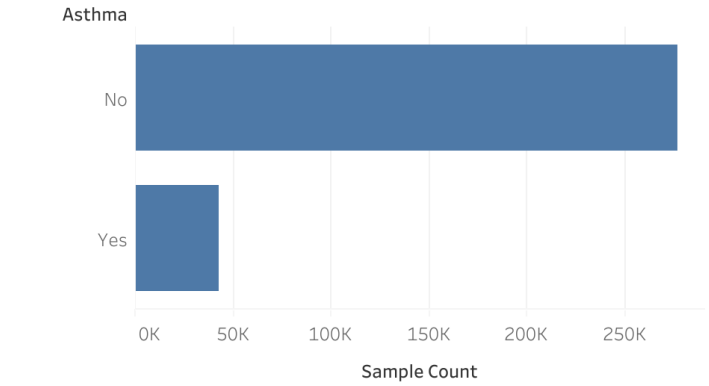

Sex
Female
Male
0K
20K
40K
60K
80K
100K
120K
140K
160K
Sample Count

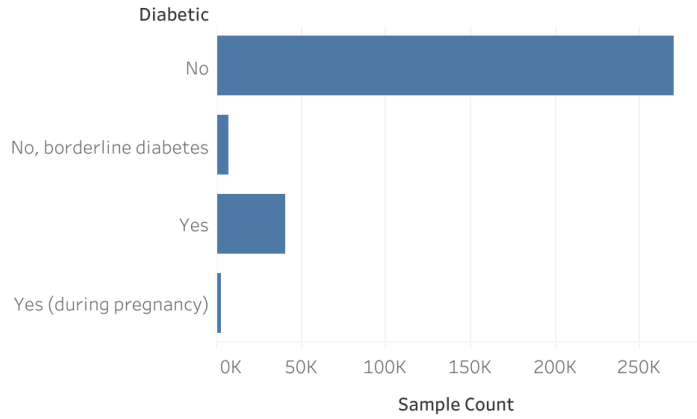

Race
American Indian/Alaskan Native
Asian
Black
Hispanic
White
Other
0K
50K
100K
150K
200K
250K
Sample Count

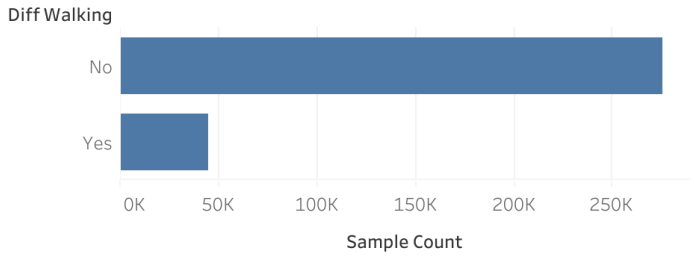

Alcohol Drinking
No
Yes
0K 20K 40K 60K 80K 100K 120K 140K 160K 180K 200K 220K 240K 260K 280K 300K
Sample Count

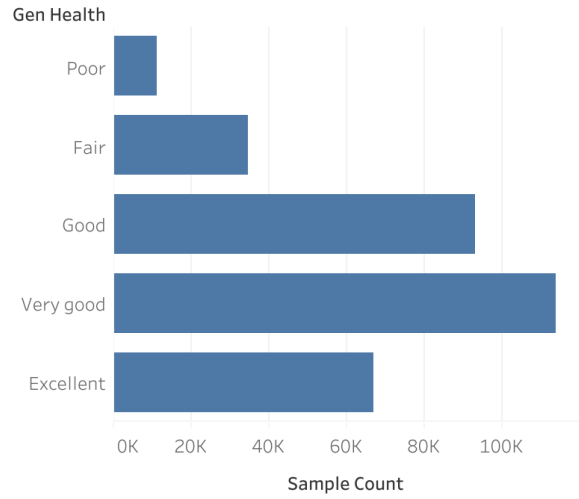

Smoking
No
Yes
0K 10K 20K 30K 40K 50K 60K 70K 80K 90K 100K 110K 120K 130K 140K 150K 160K 170K 180K 190K
Sample Count

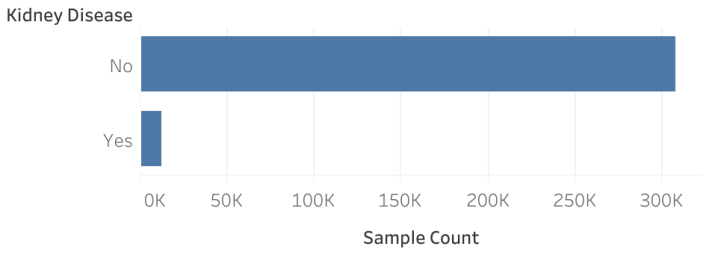

Physical Activity
No
Yes
0K 20K 40K 60K 80K 100K 120K 140K 160K 180K 200K 220K 240K 260K
Sample Count

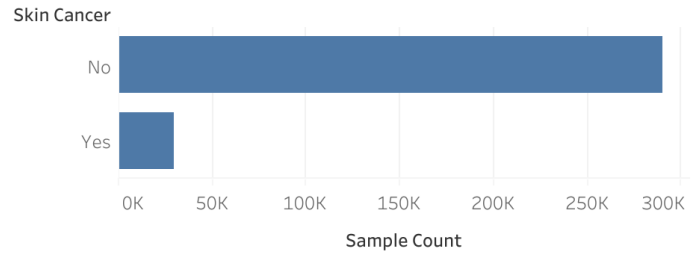

Asthma
No
Yes
0K
50K
100K
150K
200K
250K
Sample Count

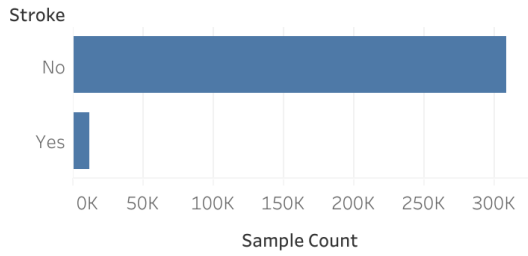

Diabetic
No
No, borderline diabetes
Yes
Yes (during pregnancy)
0K
50K
100K
150K
200K
250K
Sample Count

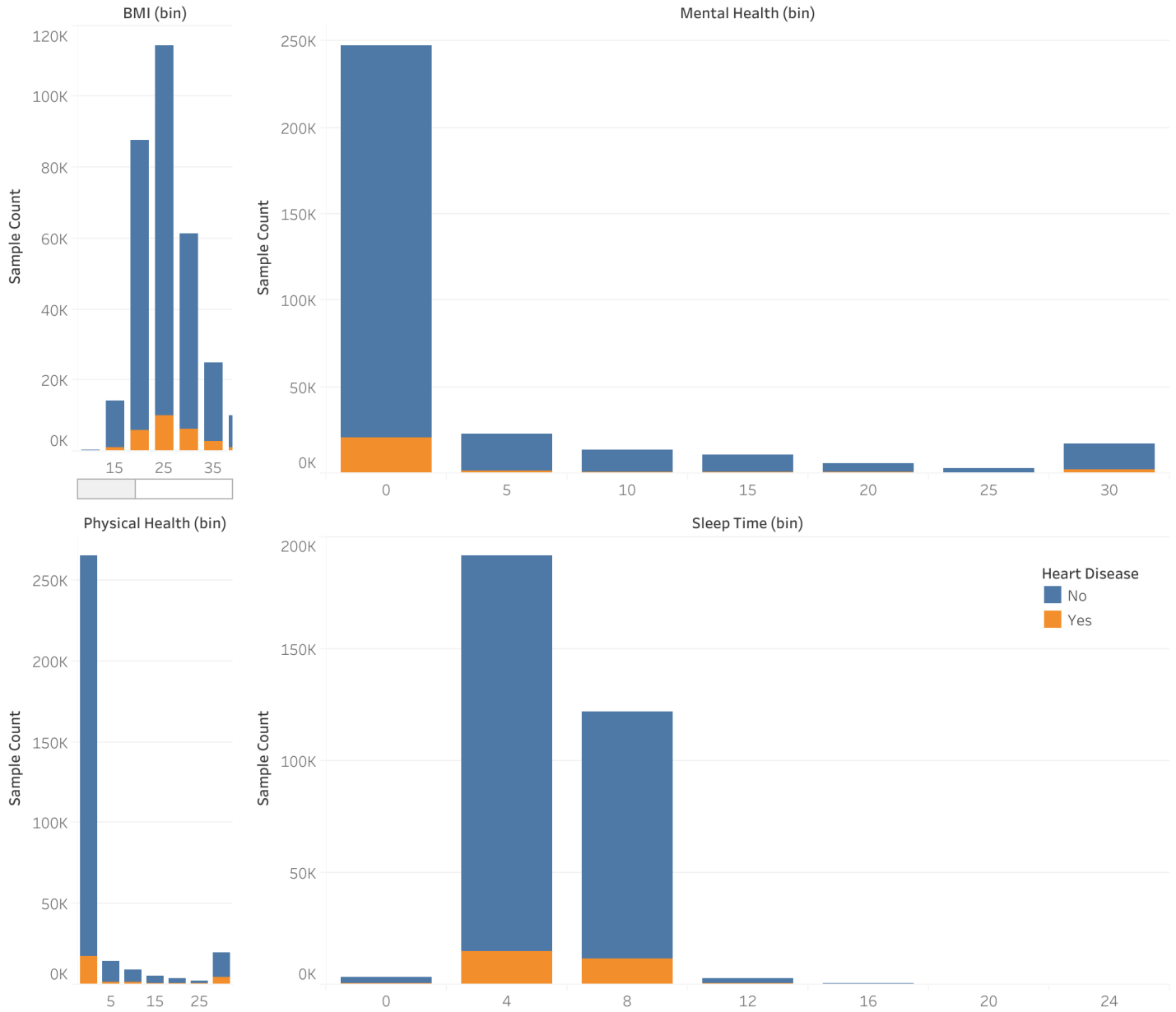

Diff Walking
No
Yes
0K
50K
100K
150K
200K
250K
Sample Count


Gen Health
Poor
Fair
Good
Very good
Excellent
0K
20K
40K
60K
80K
100K
Sample Count

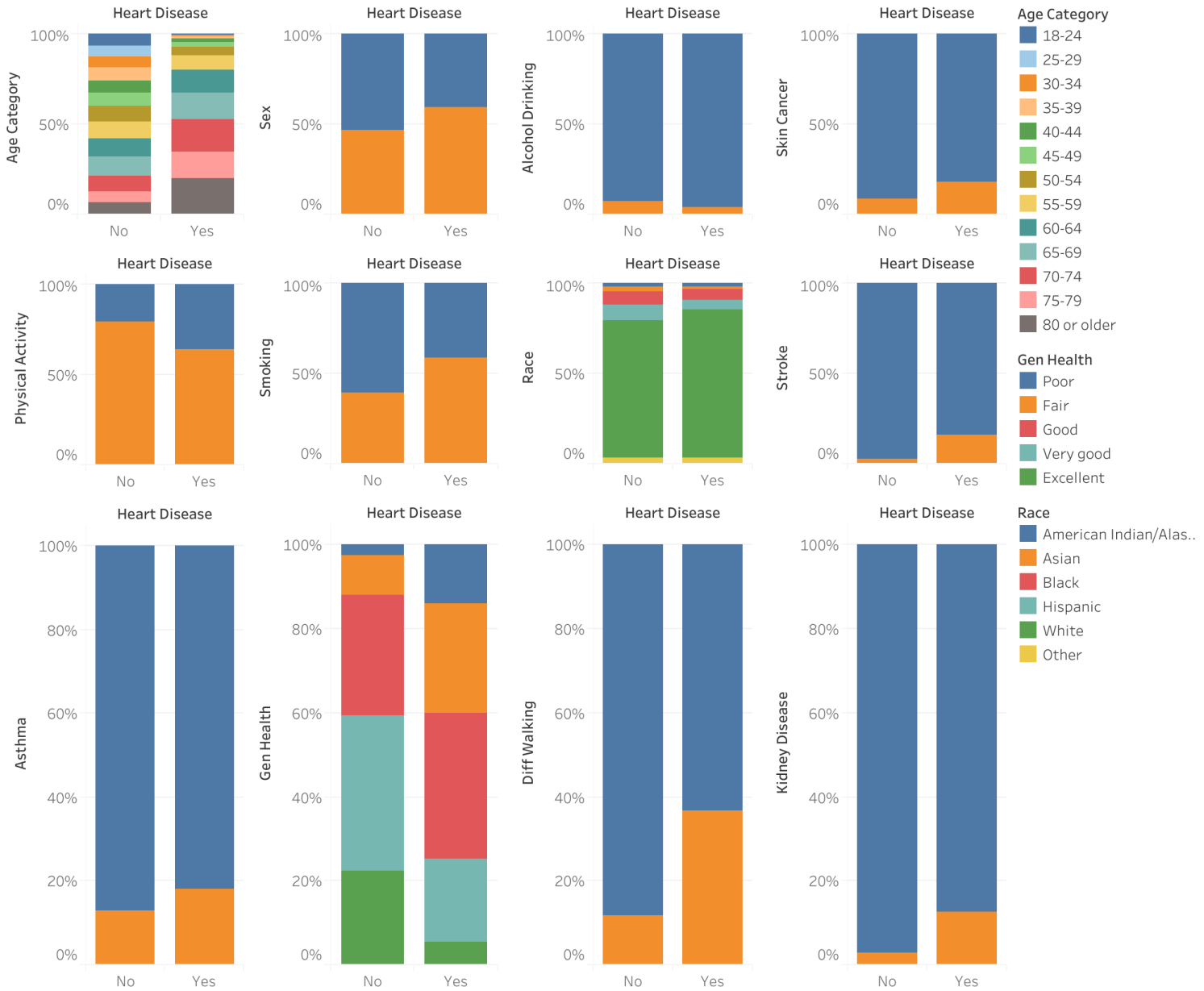

Kidney Disease
No
Yes
0K
50K
100K
150K
200K
250K
300K
Sample Count

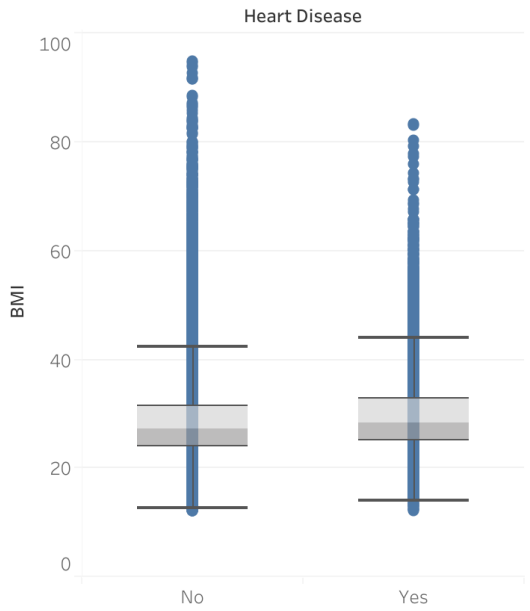

Skin Cancer
No
Yes
0K
50K
100K
150K
200K
250K
300K
Sample Count

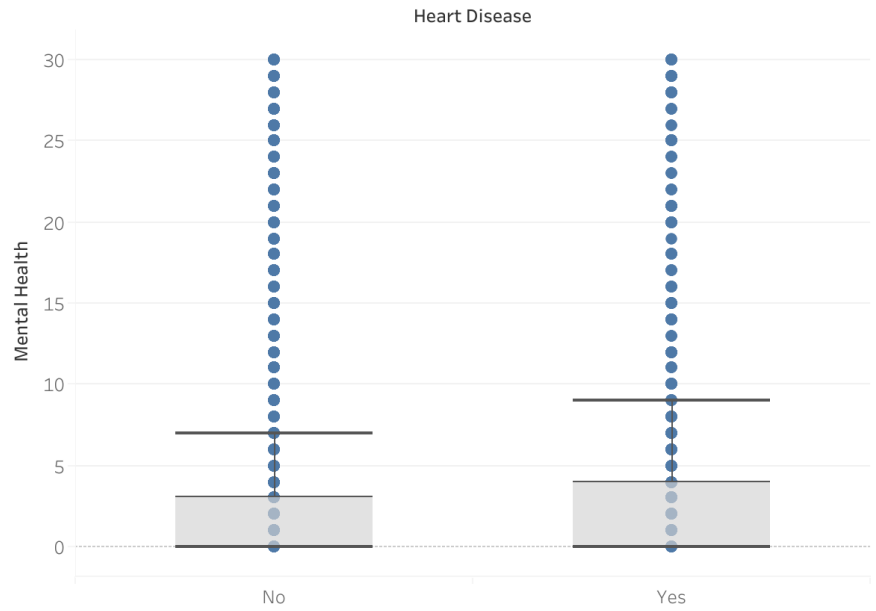

Stroke
No
Yes
0K
50K
100K
150K
200K
250K
300K
Sample Count

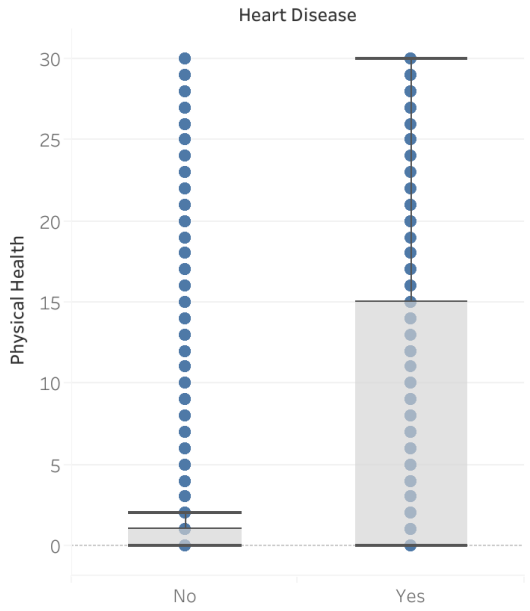

BMI (bin)
Sample Count
120K
100K
80K
60K
40K
20K
0K
15
25
35
Mental Health (bin)
Sample Count
250K
200K
150K
100K
50K
0K
0
5
10
15
20
25
30
Physical Health (bin)
Sample Count
250K
200K
150K
100K
50K
0K
5
15
25
Sleep Time (bin)
Sample Count
200K
150K
100K
50K
0K
0
4
8
12
16
20
24
Heart Disease
No
Yes

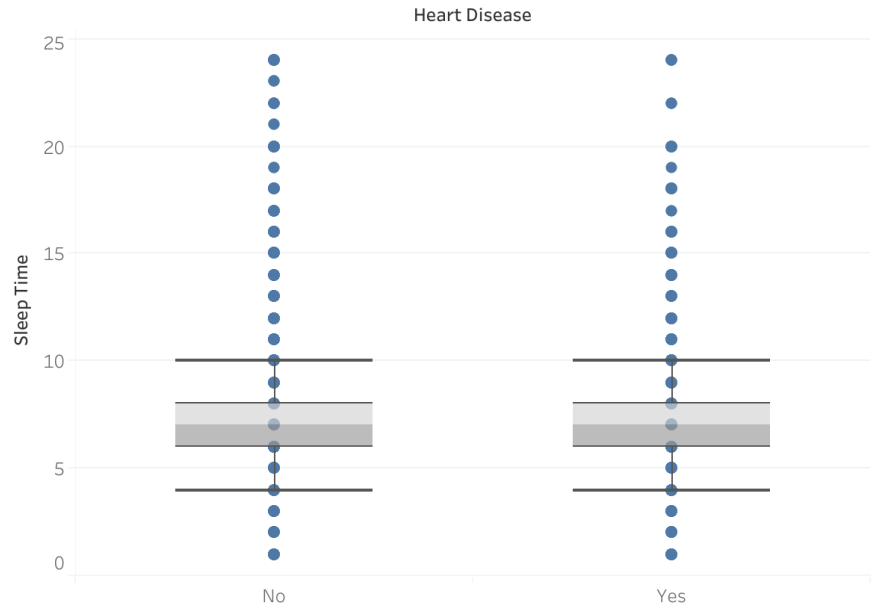
Attrition Category/Category Exploratory analysis indicates that the following attributes appear to have some relationship with heart desease




Heart Disease
BMI
100
80
60
40
20
0
No
Yes


Heart Disease
Mental Health
30
25
20
15
10
5
0
No
Yes


Heart Disease
Physical Health
30
25
20
15
10
5
0
No
Yes


Heart Disease
Sleep Time
25
20
15
10
5
0
No
Yes