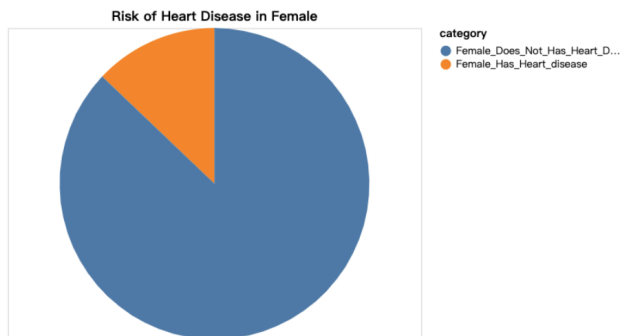
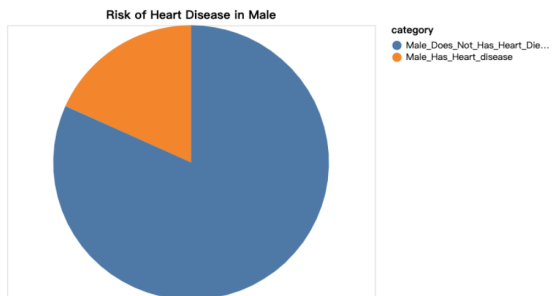


Visualizations:

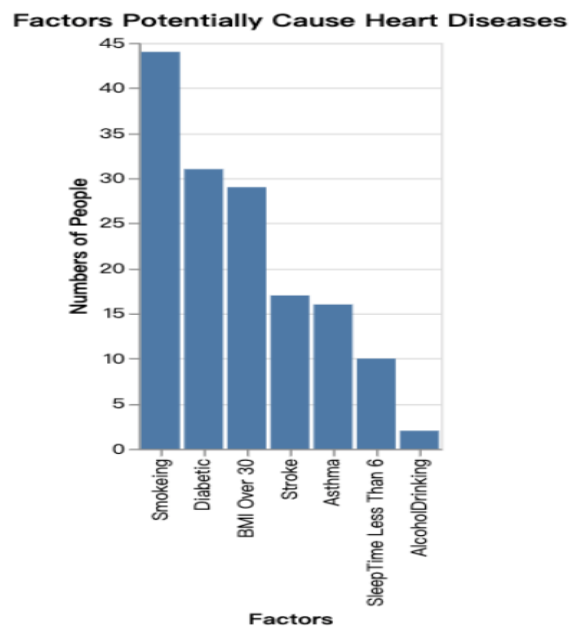
1. Risk of Heart Disease in Male and Female

```
      a  b
0 Male_Has_Heart_disease  29
1   Male_Total          159

      a  b
0 Female_Has_Heart_disease  44
1   Female_Total          341
```



2. Factors Potentially Cause Heart Diseases



Data Source:

<https://www.kaggle.com/kamilpytlak/personal-key-indicators-of-heart-disease>

To Grader:

Because the csv file contains too many rows of data (using all the data will cause “read_csv” to fail compiling), so only the first 500 rows of data are used in all these data visualizations.

Descriptions:

1. The first data visualization uses pie charts to show the incidence of heart disease in males and females. The reason to use pie charts here rather than all the other data visualization methods is because the dataset only contains the ratio of the number of men/women with heart disease. Therefore, using a pie chart can make people more intuitive to see the incidence of heart disease in men and women. According to the data on the chart, among the 500 people, there are 159 men and 341 women, of which 29 men have heart disease and 44 women have heart disease. In the pie chart, orange represents the number of men and women with heart disease, while blue represents the number of men and women without heart disease. According to the pie chart, we can intuitively see that the probability of males suffering from heart disease is much higher than that of females (male: $29/159=18.2\%$, female: $44/341=12.9\%$)
2. The second data visualization uses a bar chart to show factors that could potentially cause heart disease. Unlike pie charts, for all the “factors”, there is an union relationship between each factor. For example, a person with heart disease may have both asthma and sleep deprivation, so using a pie chart to show the factors that cause heart disease may double count among these 73 people who have heart disease. Thus, a bar chart is the best data visualization method. By using a bar chart and sorting according to the number of people for each factor, people can intuitively see the main factors that lead to heart disease.

According to the chart, the main cause of heart disease is smoking. There are 44 people, 44 of 73 people with heart disease are smokers. At the same time, drinking alcohol is the least direct cause of heart disease, only 2 people with heart disease have the habit of drinking too much alcohol. According to the bar graph, the main factors of heart disease from small to large are: Smoking, Diabetic, BMI Over 30, Stroke, Asthma, SleepTime Less Than 6 Hours and Alcohol Drinking.