**REPORT**

I changed the name of my dataset to make it easier to understand and code.

When choosing my data set, I chose one that would be suitable for the logit model and the probit model. The main goal in the data I was interested in was to estimate whether the patient had a 10-year future risk of coronary heart disease (CHD).

In the descriptive data analysis section, I obtained more diverse descriptive statistical data using the describe function from the psych library.

I used boxplots and histograms while visualizing the data, and for the sake of convenience, I combined the variables into a single code using a function that uses them under a single roof.

In the boxplots, I identified outliers for variables other than the gender variable and made brief comments on them.

I used multiple linear model, logit model and probit model as statistical models. I decided that the most suitable model among these models was the logit model. I saw that multiple linear regression was not suitable for the data at all, because the R-squared value explained the model by 9.8%, which shows us how inappropriate the model is.

As a result of the logit model, I interpreted that the independent variables always positively affected the dependent variable. As an interpretation of the data, the variables show that the patient's risk of having a coronary heart attack within 10 years increases. At the same time, we can comment that all the variables are significant.