Exercise 5: Linear or Logistic Regression – Task Outline

The data.txt file contained 100 individual’s Sex, Age, Height, BMI, Genotype and Disease data. A logistic regression was used to test the association between the genotype variable and the disease variable, with the other variables (Sex, Age, Height, BMI) added as covariates into the regression model. Where P values and Beta values were reported. A double loop structure was then designed to test the association between all possible pairs of columns, by selecting two columns and carrying out the association test.