Cancer\_Data\_Using\_R

This file contains data from cancer patients whose tumor measurements are posted to determine if a patient has a Malignant or Benign tumor. It is written in R, using R Studio. It includes an in-depth Exploratory analysis and it creates a Word document with the code and the outputs of the code ‘knitted together.” Logistic regression was used to try to predict the outcome of the patient based on their tumor measurement. A cluster k-means algorithm is also used to try to see if any factors will cluster together into different groups. A set of different k values were used to try to determine which k value is the best one.