## Titanic Practice Exam

## Task 1 – Explore the data (5 points)

## Task 2 – Create a Title variable and simplify the factor levels (7 points)

## Task 3 – Create a variable Family Size (2 points)

## Task 4 – Use Kmeans clustering to look for outliers (10 points)

## Task 5 – Select which variables should be used in modeling (3 points)

## Task 6 – Fit a Random Forest (7 points)

## Task 7 – Select the Top 7 Features (7 points)

## Task 8 – Fit a Logistic Regression and interpret the signs of the coefficients (4 points)

## Task 9 – Calculate the probabilities of a person of age 10, 30, and 60 surviving (9 points)

## Task 10 –Validate the Logistic Regression and compare the AUC and Accuracy to the random forest in Task 5 (4 points)

## Task 11 – Fit an Elastic Net with all pairwise interactions (6 points)

## Task 12 – Choose the value of Lambda so that between 5-10 variables remain (6 points)

## Task 13 – Compare the AUC with the Logit’s AUC (6 points)

## Task 14 – Executive summary (24 points)