**Week 4 – The plan**

**Goal**:

Find the features that are important

We need to experiment with different techniques to test the importance of features, in order to cut out ¾ of the variables

**Techniques for feature selection:**

Logistic regression

Random Forest Classifier – accuracy, precision, recall

Correlation analysis

Univariate Feature Selection

Recursive Feature Elimination (RFE)

~~~~~~~~

Random Forest Classifier:

* Joshiah
* Wahson

Logistic regression:

* Weihao
* Fares

Recursive Feature Elimination (RFE)

* Madison
* Vu

Qualitative:

* See min, max appropriate ranges for features – refer to the excel sent in Moodle.
* Madison
* Wahson
* Weihao
* Fares
* Joshiah
* Vu

Due Date: Saturday 23rd August

1. Rank your feature importance
2. For each of the top 10 variables, be able to chose either min, mean or max
3. Upload your code to GitHub