We'll be using the same data as used in the logistic regression. You dont have to re-do the data prep parts . Use that data as is.

1. For running a random forest model , you dont have to consider VIF etc. This step is to tell you to NOT do all that .
2. Convert your response variable to factor type [you need to do this to build a classification model with randomforest package in R]
3. Add a random column to your data [rnorm to generate random values]
4. Build a randomforest model with 300 trees , keep mtry value close to number of features in the data
5. Extract variable importance and take note of the columns which have more importance than the random column added earlier
6. Keep only those columns which were selected in the earlier step . decide on range of values to tune on for these parameters : `ntree`, `mtry`, `maxnodes`, `nodesize`.
7. Tune parameters using the method discussed in the class. what is the cv performance of the model?