Classification Methods:

1. Decision Tree
   * Supervised learning (may work well if we know what station responded)
   * Decision trees can be visualized
   * Possibility of overfitting so trees may need to be pruned
   * Method included in DecisionTreeClassifier package in R
2. Logistic Regression
   * Covered in linear regression course
   * Used when response is categorical (and typically binary). This means we would need to utilize multinomial logistic regression for this problem.
   * Assumptions of multinomial logistic regression:
     + Independence of observations (If a station is already on a run, they will not respond when they normally would have responded.)
     + Categories of the outcome variable must be mutually exclusive or exhaustive
     + No multicollinearity between independent variables
     + Linear relationship between the continuous variables and the logit transformation of the outcome variable
     + No outliers or highly influential points
   * Method can be done in Minitab
3. Kernel Estimation
   * Unsupervised learning method
   * Method performs well in the case of bimodal or heavily skewed distributions (may not be applicable)
   * May not be able to be done in R
4. Random Forest
   * Covered in data mining course
   * Multiple decision trees are created using different random subsets of the data and features
   * Method included RandomForestClassifier package in R
   * Typical random forest method routine:
     + Split the data (into training and testing set)
     + Train the model
     + Tune the model
     + Assess model performance
5. Kernel SVM
   * Supervised learning method
   * “The main idea is that based on the labeled data (training data) the algorithm tries to find the optimal hyperplane which can be used to classify new data points. In two dimensions the hyperplane is a simple line.”
   * <https://towardsdatascience.com/svm-and-kernel-svm-fed02bef1200>
6. Naive Bayes
   * Assumption of independence between every pair of features
   * Work well in many real-world situations such as document classification and spam filtering (may not be applicable to this problem)
   * Small amount of training data necessary to train model
   * Method is very fast compared to classification methods
   * Known for being a bad estimator
7. MLP multi-layer perceptron (neural network) Classifier
   * Supervised learning
   * Method included in MLPClassifier package in R
   * Capability to learn nonlinear models
   * Capability to learn nonlinear models
   * Cons: I do not understand most of the terminology
8. Binary Classification
   * Classifies elements into two groups
   * Would not work for this problem (More than two stations)
9. Neural Networks
10. K-Nearest Neighbors
    * Simple algorithm
    * Class assigned is the most common amongst its K nearest neighbors measured by a distance function
    * Method can be done with a package in R