Random Forest Worksheet

Excel Model

1. What is the shape of the dataset? \_\_\_\_\_\_\_\_
2. What is the mean for each of the independent variables? \_\_\_\_\_

\_\_\_\_\_\_ \_\_\_\_\_\_ \_\_\_\_\_\_\_\_

1. What is the minimum? \_\_\_\_\_\_\_ Maximum ?\_\_\_\_\_\_\_\_\_
2. What are the predicted values from the model?
3. In the confusion matrix what are the following values for TP \_\_\_, TN \_\_\_\_, FN \_\_\_\_, FP
4. How accurate is your model? \_\_\_\_\_\_\_\_
5. List the features of importance.

Comma Separated Values

1. What are two ways can you use to create a csv file?
2. How to you separate the independent variables?
3. What is the shape of the dataset?
4. In the predicted values, how many passes? Failed?
5. In the confusion matrix, how many are TN?
6. List the criteria and their scores for most important reasons for passing or failing .

Internal data model

1. What is JSON? Explain.
2. What is the name of the array holding all the data?
3. What does the X variable contain?
4. What does the y variable hold?
5. How many true negatives are there? Name them.
6. Are there any false positives?
7. List the important criteria for passing and failing and their percentages.

Panda dataframes and file formats

1. What is a Pandas dataframe?
2. What is a Pandas dictionary?
3. What is a Pandas series?
4. What are the areas that each of the schools are judged on?
5. List three ways to get data into a Python program.