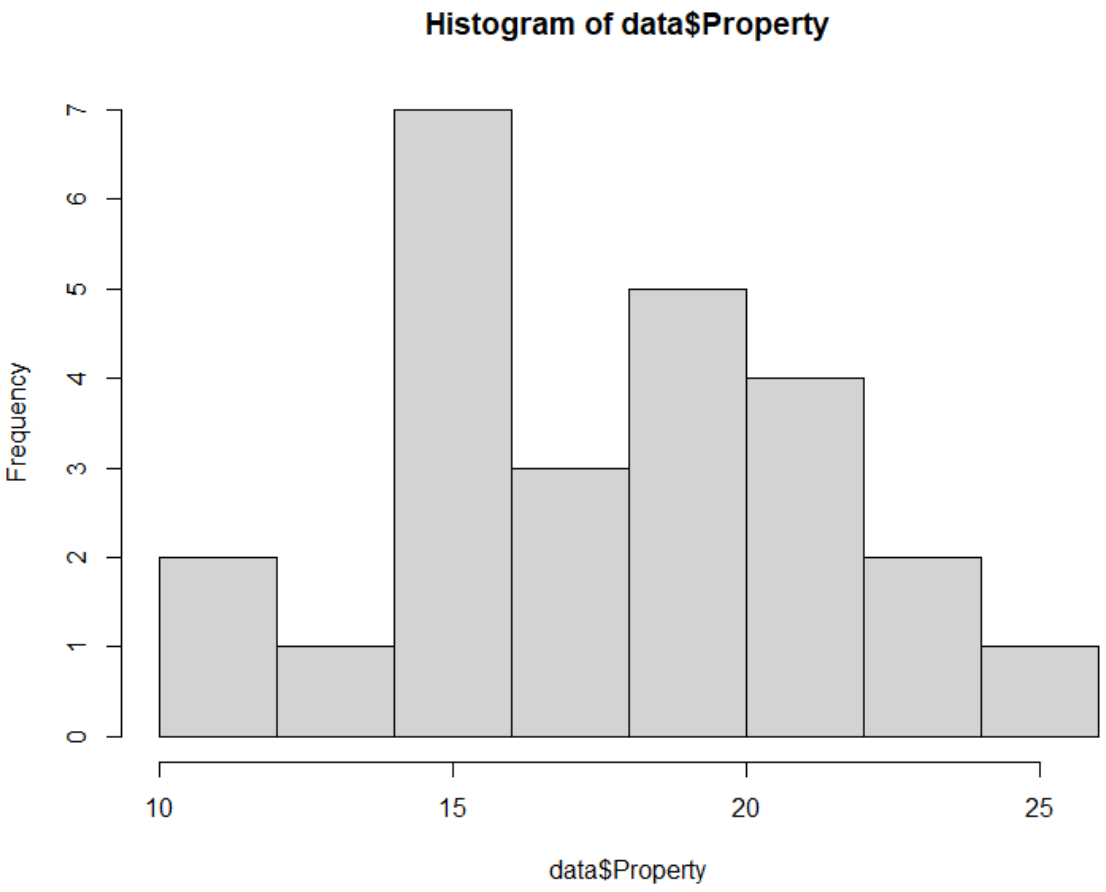
Homework 2
Nikhil Ambati

Histogram:



1. SVM

The SVM model's findings after being trained and tested using the provided dataset are shown below:

```
                truth
svm.pred_train -1 1
            -1  8 4
             1  2 6
```

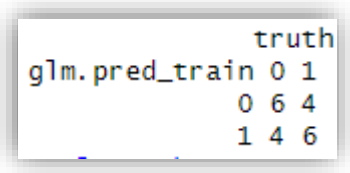
```
                truth
svm.pred_test -1 1
           -1  3 0
            1  0 2
```

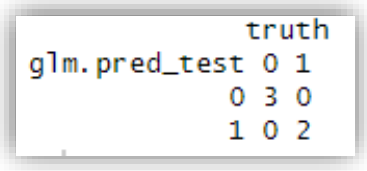
SVM_Train_R2 = 0.7	**SVM_Test_R2 = 1**
Sensitivity=0.8	**Sensitivity=1**
Specificity=0.6	**Specificity=1**

Homework 2
Nikhil Ambati

2. Logistic Regression

The logistic regression model's findings after being trained and tested using the provided dataset are shown below:





logistic_Train_R2 = 0.6	**logistic_Test_R2 = 1**
Sensitivity 0.6	**Sensitivity 1**
Specificity 0.6	**Specificity 1**

On the test data, the models for logistic and SVM are producing good results, although 100% accuracy is not the desired outcome. The modest size of the data collection may have contributed to the high accuracy.

Given that both models' accuracy is high, I have a moderately high confidence in them. The approaches listed below are less dependent on the distribution of test and training data. As a result, altering the distribution of the data will not reveal any appreciable variations.

One of the key features for defining the model is feature 4, which has a high correlation with the property field. In a similar vein, feature 3, which is not correlated with any other fields, also has a respectable impact on the model constructed.