**DECISION TREE**

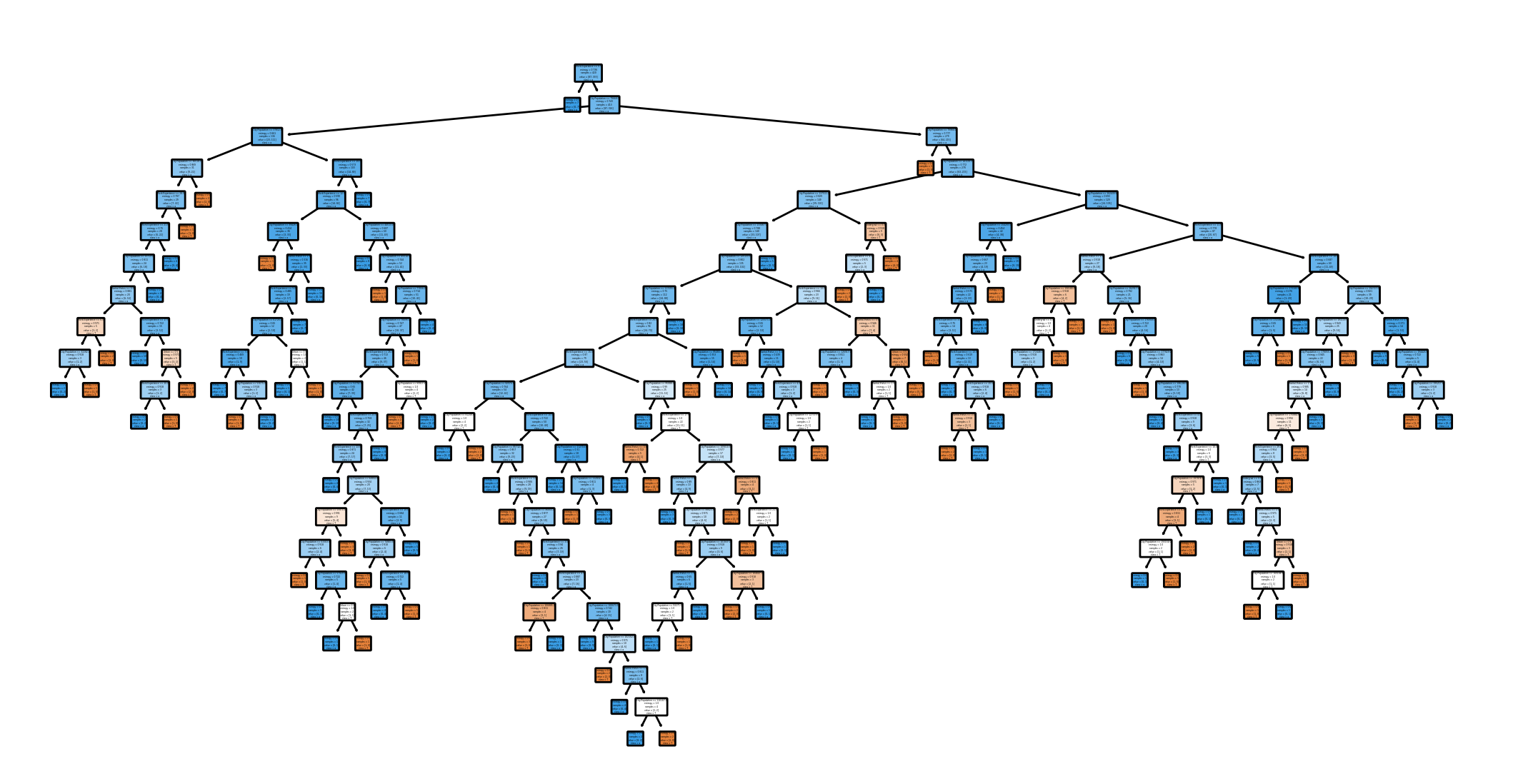
**Business Problem** = ﻿﻿﻿Use Decision Trees to prepare a model on fraud data to check customer Risky or Good.

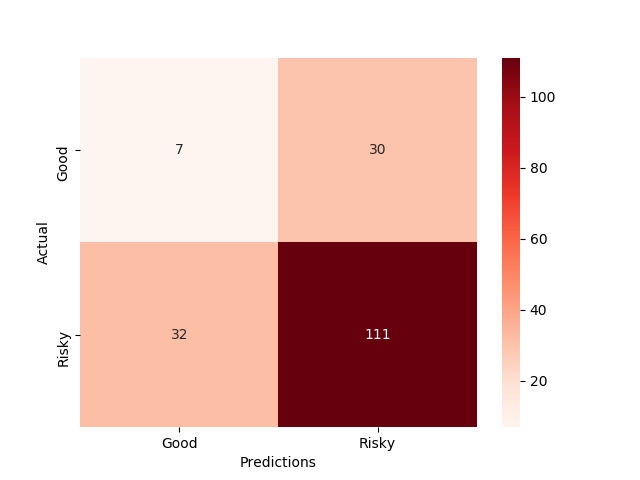
* **Name of the File: -** Fraud\_check.csv
* **Size of the File: -** 438 KB
* **Necessary Data : -** 600 Observations, 6 Features.

**Exploratory data Analysis** =

* **Outliers: -**  Outliers are not presents.
* **Missing Value: -** Data don’t have Missing Values
* **Output:** - Categorical
* **Sampling:**- Stratified Sampling (79% - 21%)

**Building Decision Tree =** Building decision tree by considering entropy criterion.

* **﻿Tree :-**

****

* **﻿Accuracy Score :-** 67 %
* **Confusion Matrix : -**

Accuracy of model on Train data is very high but on test data it is quite low so this is overfitting problem. We have to overcome this problem by building ML model using different techniques such as Bagging, Boosting or Stacking of Ensembles Methods.

**Python code file**: - [Fraud\_check Analysis.py](https://github.com/nilaydeshmukh0/Decision-Tree/blob/master/Fraud%20Checks%20Analysis/Fraud_check%20Analysis.py)