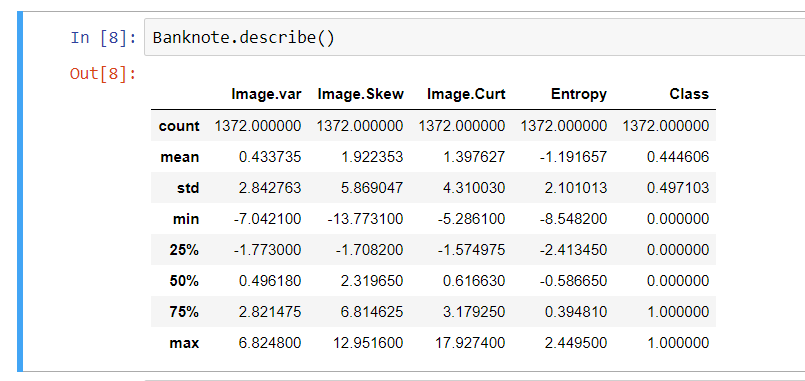
PART-A

* **Data cleaning**
  + Are there missing values? (10 %)
* No, this Dataset is fetched from Images of Notes that are complete.
  + Are there inappropriate values? (10 %)
* No, as they are from original values. That means from the images of Bank Notes. Hence, the data retrieved from those images must be authentic.
  + Remove or impute any bad data. (10 %)
* No need to remove any data as each row is important in decision making and cannot be ignored from accuracy perspective.
* **Answer the following questions for the data in each column:**
  + How is the data distributed? (10 %)
* This Dataset was simple as it was evenly distributed. It was not difficult for me to make fast decisions on algorithms. As it was an easily predictable data.
  + What are the summary statistics? (10 %)
* 
  + Are there anomalies/outliers? (10 %)
* No, there we no anomalies or outliers found in model or dataset during analyzing data
* **Plot each column as appropriate for the data type:**
  + Write a summary of what the plot tells you. (10 %)
* Plots give us a clear idea of how real and forged notes can be classified. Also, the pair plot gave the clear picture of distribution of both note features
* **Are any of the columns correlated? (10 %)**
* No, the columns are separated and do not create a problem of overfitting model as they are features of a note and do not impact each other’s correlation.
* **Write a clear summary of what the EDA tells you (20 %)**
* It was very easy to create a model on basis of EDA. EDA tells the clear picture of difference in features and how it effects the model.