* Firstly, I import library such as pandas, numpy, matplotlib, sklearn. Pandas library is the most popular that makes data manipulation easy. In the next line I read it using the pd.read\_csv('Tweets.csv'), since dataset is a csv file and then we print the dataset.
* The dataset contains a lot more parameters about the tweet like tweet\_id, airline\_Sentiment, airline, name, text, tweet\_created, user\_timezone. But for the sake of building a model we are only required about 2 parameters: the actual ***text*** of the tweet and the ***airline sentiment***.
* For data preprocessing, we need to remove unnecessary attributes from the dataset and unwanted text from the ***text*** attribute.
* To remove unnecessary words, I am going to use the following techniques:

1. Removing Stop Words: Basically words like ***this, an, a, the, etc*** that do not affect the meaning of the tweet
2. Removing Punctuation: ***',.\*!'*** and other punctuation marks that are not really needed by the model
3. Stemming: Basically reducing words like ***'running, run, ran'*** to its root word, which is ***run*** in this case. Since all variations of the root word define the same meaning.

* I converted the data into the numeric form. As the last step before we train our algorithms, we need to divide our data into training and testing sets. The training set will be used to train the algorithm while the test set will be used to evaluate the performance of the model.
* In the last step, I am going to build the NLP model using the input and the output. Here, I used ***DecisionTreeClassifier*** and ***LogisticRegression*** to build the model. I used sklearn library to build model and find classification report, accuracy score.