- **Document your approach to solve the problem, discussing the difficulties and how your proposed solution tackles them.**

  Consequently, many sophisticated and high performing algorithms have been invented to analyze text data and predict its sentiments. But application of more advanced algorithm doesn't necessarily mean that our prediction is of high accuracy.

  We still need to go back to the basics and understand the nature of data, its challenges for any further processing.

  The main purpose of this project is to find whether the tweets polarity as positive and negative or neutral and word frequency counts.

- **Discuss the technique used and the reason why you have chosen it.**

  The dataset consisted of 2 rows × 43347 columns Where I aggregated the data set and created a data frame changed the dataset for 43347 rows × 2 columns to have a better understanding over the dataset.

  I decided to go ahead to prepare the data for quick and dirty classification task .Of course, one has to do at least some basic text pre-processing that is acceptable by the classifiers . Some of the basic pre-processing techniques expand contracted words, lower case all the text, remove non letter strings, @, hyperlinks, stop words etc, expand contracted words, e.g, 'isn't' : 'is not', 'won't 've': 'will not have' etc.

- **Discuss the shortcomings or mistakes of your proposed solution with a few examples.**

  The mistakes I've done in the project was I haven't modelled the data the problem statement seems only to find the entity frequency counts and the polarity as positive, negative and neutral where I only pre-processed the Data and done some visualisations for statements functionality

- **If there are any shortcomings or mistakes, discuss how you would go about tackling them given more resources and time.**

  Where the mistakes in this project maybe modelling the data and the not having a better visualisations of the entities. if the there are more resources and time available where I can produce the project even accurate by modelling and by giving visualisations to understand the data for a better outcome.