Topic: Enhance Customer service by studying Customer Support of top companies on Twitter.

#Question/need:
-What is the framing question of your analysis, or the purpose of the model/system you plan to build?
-Who benefits from exploring this question or building this model/system?

The goal of this project was to aid innovation in natural language understanding and conversational models by using classification model. In this project, the study will focus on modern customer support practices on one of the most popular social media platforms 'Twitter' for Over 3 million tweets and replies from the biggest brands and companies its impact. This will give a wide view to enhance my job as customer service representative in a mall which customer support on twitter is new filed to my company.

The size and breadth of this dataset inspires many interesting questions:

- Can we predict company responses? Given the bounded set of subjects handled by each company?
- Do requests get stale? How quickly do the best companies respond, compared to the worst?
- Can we learn high quality dense embeddings or representations of similarity for topical clustering?
- How does tone affect the customer support conversation? Does saying sorry help?
- Can we help companies identify new problems, or ones most affecting their customers?

#Data Description:
-What dataset(s) do you plan to use, and how will you obtain the data?
-What is an individual sample/unit of analysis in this project? What characteristics/features do you expect to work with?
-If modeling, what will you predict as your target?

This dataset "https://www.kaggle.com/thoughtvector/customer-support-on-twitter" contains 2811774 row, where each row is a tweet and there is 7 columns. Every conversation included has at least one request from a consumer and at least one response from a company. Which user IDs are company user IDs can be calculated using the inbound field which is the target.

#Tools:
-How do you intend to meet the tools requirement of the project?
-Are you planning in advance to need or use additional tools beyond those required?

- Numpy and Pandas for data manipulation.
- Scikit-learn for modeling.
- Matplotlib and Seaborn for plotting.
- Tableau for interactive visualizations.

#MVP Goal:
-What would a minimum viable product (MVP) look like for this project?

* Basic data cleaning:
    - Removing punctuation, numbers, extra whitespace, stop-words, duplicates and emojis.
    - Handling missing data.

* Basic Data Exploration such as:

    1) Create new columns such as the companies column.
    2) The shape, head, info, describe, and Summary of the dataset.
    3) Convert some data types if needed.
    4) Statistical Insight mean, max, min.
    5) Plotting the distribution of the data to handle imbalanced distribution.

* Some visualizations to the data for more understanding.