Dataset is divided into various folders which include-

* **FamousNewsChannel**- Contains the tweets of famous business news accounts that have mentioned name of company. (Separate file for each company in Big4).
* **Happy Or Sad**- Tweets are captured based on the basic of positive and negative words in the tweet with company name. (Separate file for each Big4 company and each sentiment i.e. good or bad).
* **Services**- Tweets are captured on the basis of various services provided by these BIG4. (Separate file for each company and each service i.e. consultancy, tax, assurance and Strategy).
* **Geographical** Tweets captured on the basis of twitter account tagged in the tweet for a particular company of a particular location. (Separate file for each company and countries which include US, UK, India, Australia, Canada and Africa).
* **Event** Tweets containing company’s name and event(like COVID19) or technology(like Data Science, AI, Machine Learning) in it.
* **BIG4 file**- Contains all tweets with hashtag BIG4.
* **Sample Data**- Data provided by EY organization.