1. How you proceeded with the data. Your thought process before you started working!
I have gone through the text and thought of processing the data and then obtaining the high frequency words so that the company could be found as it will be used more commonly in the speech.
2. Tools, languages, software's used
NLTK library ,python ,Google Collab
3. Answers to these 2 simple questions: (you may need to read the text a little bit)
a. The company you think this is related to?
Twitter
b. Financial year CEO is discussing?
2016
4. The visualizations, insights you extracted from data – images, thoughts, anything!
- Separate the text into sentences
- Separate the text into words
- Delete the stopping words from the tokenized words.
- Plot the frequency of most common words.
- Get the parts of speech of the words.
- Trying to get the company name using the pos_tags

5. What more you think can be done in future on this data!
    As a future enhancement below can be done :
- Sentiment Analysis :we can do sentiment analysis of the speech and get to know the tone of the text/call.If it ends on a positive note then we can conclude that the goals were achieved for the Financial year. Else then the company didn't do well during the Financial year.
- Text Summarization : Summary of the text can be generated which gives details about the growth and developments during the Financial year.

6. Any code (if you have written) should be uploaded in a public GitHub repository. The link should be included in the report.
https://github.com/harichandanamone/DATA_CHALLENGE