**Question 1 – Perform 3 Analysis on Enron Dataset**

**Analysis 1:**

Created Folder structure based on month-year combination of Enron dataset.

Read the data from the sent email list instead of reading data from inbox to avoid data from spam emails.

Defined a list which contains meeting related words like: ['meeting', 'attach', 'email', 'thank', 'inform', 'receive', 'file']

It is well known fact that CEO Jeﬀrey Skilling and Chairman Kenneth Lay of Enron at the time of the scandal, were holding regular meeting with their top executives to pressure them into ﬁnding new ways to hide Enron’s debt.

Graph shows the same results, count of email containing those meeting related words are highest between the period Jan 2001 to Oct 2001. Enron has declared bankruptcy in Dec 2001 after that the count of meeting related words decreased significantly.

**Analysis 2:**

Created a text file and stored it to the corresponding years folder structure.

Text file contains the list of employers those are involved in discussion of any kind of ‘gifts’ and ‘rewards’ related emails.

Also, text file contains a keyword which gives the clear idea of discussion.

**Analysis 3:**

Keywords like ‘Gift’ and ‘Rewards’ were significantly increased by 76% in the email content in year 2001 as compared to 1999.

Created a bar chart using matplotlib to show this difference.

**Question 2 – Perform 3 Analysis on NYT API**

**Analysis 1:**

Downloaded and collected data from the Book API

Created a CSV file which stores the Rank, Author’s name, Title of the book and list published date using CVS module.

**Analysis 2:**

Downloaded and collected data for the year 2000 from Archive API

Stored them into required folder structure

Found top 10 buzz word for that year.