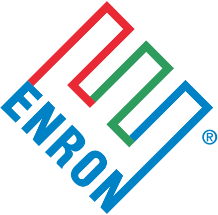
Analysis of Enron Data



**ABOUT ENRON:**

Enron Corporation was an American energy, commodities, and services company based in Houston, Texas. It was founded in 1985 as the result of a merger between Houston Natural Gas and InterNorth, both relatively small regional companies in the U.S. Enron employed approximately 20,000 staff and was one of the world's major electricity, natural gas, communications and pulp and paper companies, with claimed revenues of nearly $101 billion during 2000. Fortune named Enron "America's Most Innovative Company" for six consecutive years.

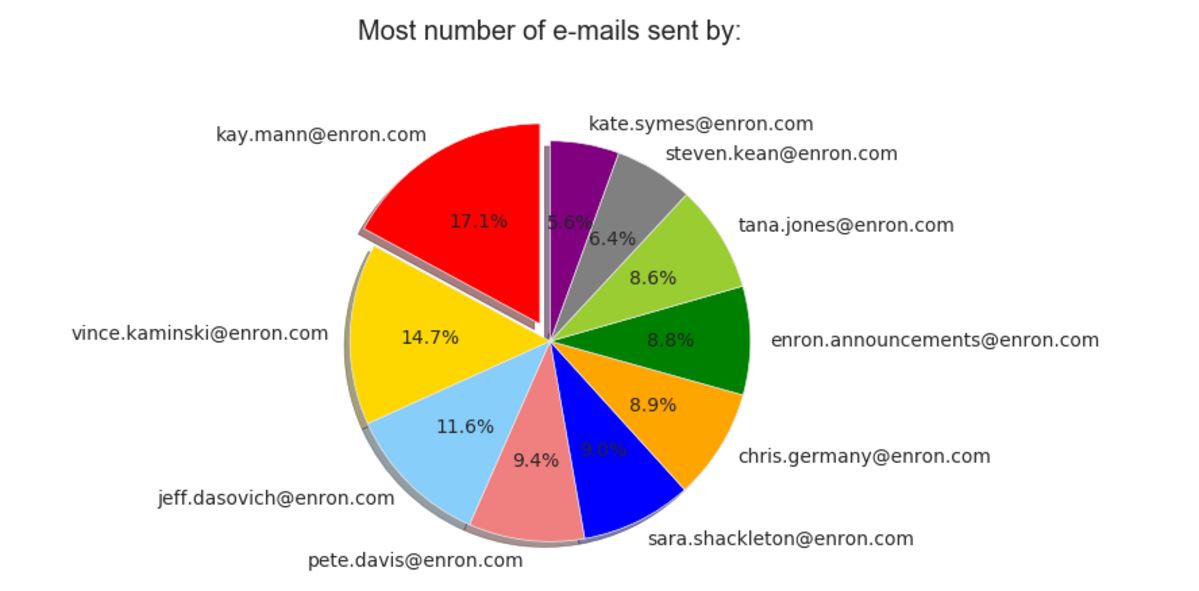
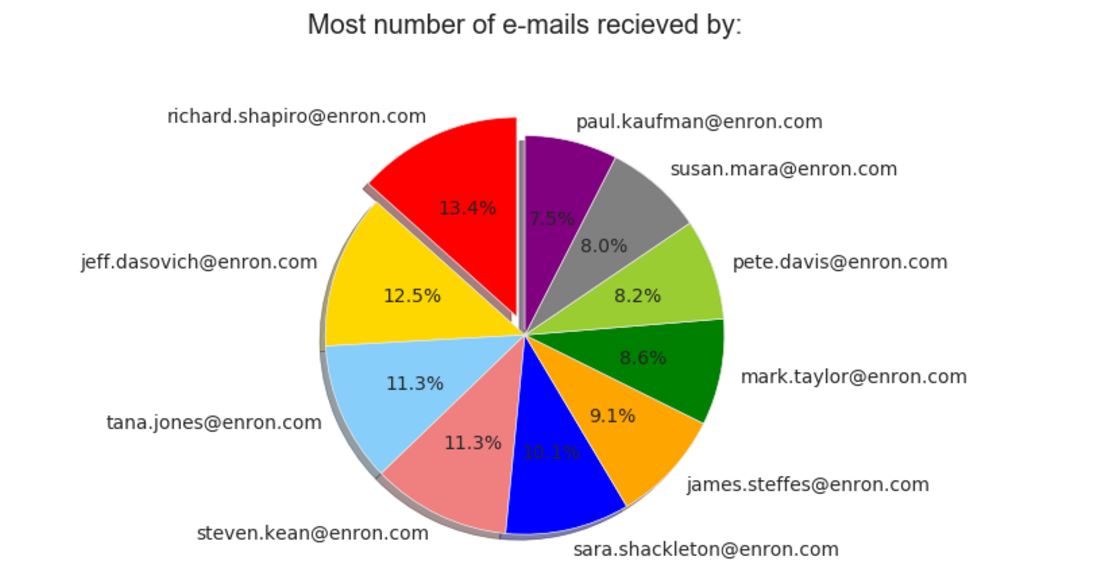
**DOWNFALL OF ENRON:**

By the fall of 2000, Enron was starting to crumble under its own weight. CEO Jeffrey Skilling had a way of hiding the financial losses of the trading business and other operations of the company; it was called mark-to-market accounting. This is a technique used when trading securities where you measure the value of a security based on its current market value, instead of its book value. In Enron's case, the company would build an asset, such as a power plant, and immediately claim the projected profit on its books, even though it hadn't made one dime from it. If the revenue from the power plant were less than the projected amount, instead of taking the loss, the company would then transfer these assets to an off-the-books corporation, where the loss would go unreported. This type of accounting enabled Enron to write off losses without hurting the company's bottom line.

**ANALYSIS OVER ENRON DATASET:**

Enron dataset contains data from about 150 users, mostly senior management of Enron, organized into folders. Kenneth lay the CEO and chairman of Enron Corporation. Upon analyzing his email directory and the content of the mails that he sent or received, contains words such as bankruptcy, lost, million etc. and they are the words from top 100 frequently used words by Lay. It clearly states there was something fishy going on in the organization. Later Lay was indicted by a grand jury and was found guilty of 10 counts of securities fraud. Similarly, top words from former CEO Jeffrey Skilling and Richard Shapiro who was the Vice President and lobbyist for Enron are attached in the skilling\_mail.csv and shapiro\_mail.csv files respectively.

The above report shows the number of mails that were exchanged by the Enron corporation in different year. It’s obvious that when the company was at its best performance and one of the biggest company in the world during year 2000 and 2001 the number of mails exchanged were more. The company got ceased on December 2, 2001 so we can see the impact and number of employees that were laid off which can be clearly seen in the number of mails that were exchanged in 2002 which clearly shows huge number of employees were laid off all of sudden.



**Pie chart-1: Employee with most email recepient Pie chart-2: Employee with most email sender**

From the pie chart-1 above you can clearly see that most numebr of mails were receieed by Rischar Shapiro who is the vice president of the Enron who was convicted to be the on of the lobbyist involved in the bribery and downfall of the Enron. And on the other hand from pie chart-2, most number of mails were sent by Kay Mann who handles legal department in corporate matters, energy, oil and gas, and construction. She is one of the good person that came out in the whole scandal and it can be clearly stated that though she must have sent many mails regarding the bad thing that is going on internally to higher officials, no one really cared or over looked the matter, which might have led the downfall of Enron.

**REFERENCE:**

[*Enron Scandal Summary*](http://www.investopedia.com/updates/enron-scandal-summary/)*- Link to Investopedia article to get a brief summary about the what the scandal was.*

[*https://en.wikipedia.org/wiki/Enron\_scandal*](https://en.wikipedia.org/wiki/Enron_scandal)

*http://www.investopedia.com/updates/enron-scandal-summary/*

**-*Sumeet Usturge***

***NUID-1676003***

NYT API to collect NYT data

**API USED:**

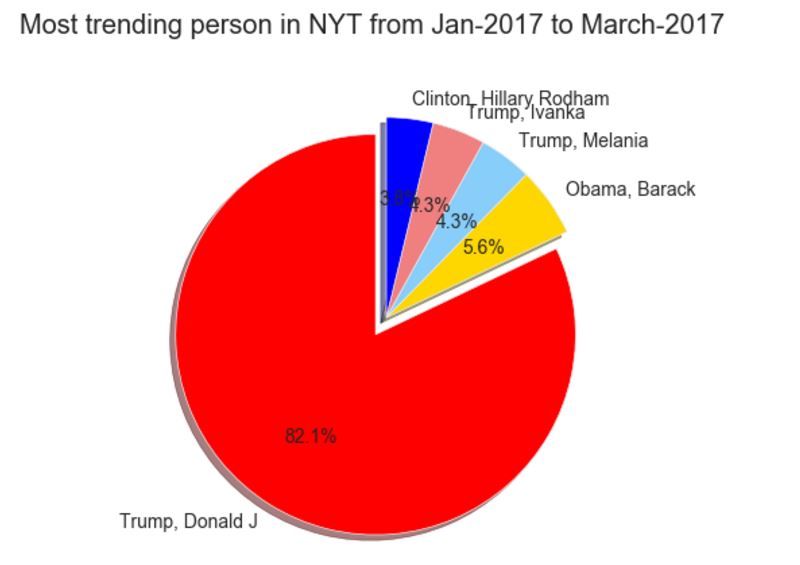
**1.Article Search API**: With the Article Search API, you can search New York Times articles from Sept. 18, 1851 to today, retrieving headlines, abstracts, lead paragraphs, links to associated multimedia and other article.

**2.Archive API:** The Archive API provides lists of NYT articles by month going back to 1851.

**Analysis Performed:**

1.Parsed the JSON response from the API and by drilling down the JSON(Dict) collected all the headlines for the specified time line of data and stored in the list.

2.From the data figured out the trending person from Jan-2017 to March-2017 and represented in pie-chart as below:



3.Calculated the frequently used words in the heading of NYT from Jan-2017 to March-2017.