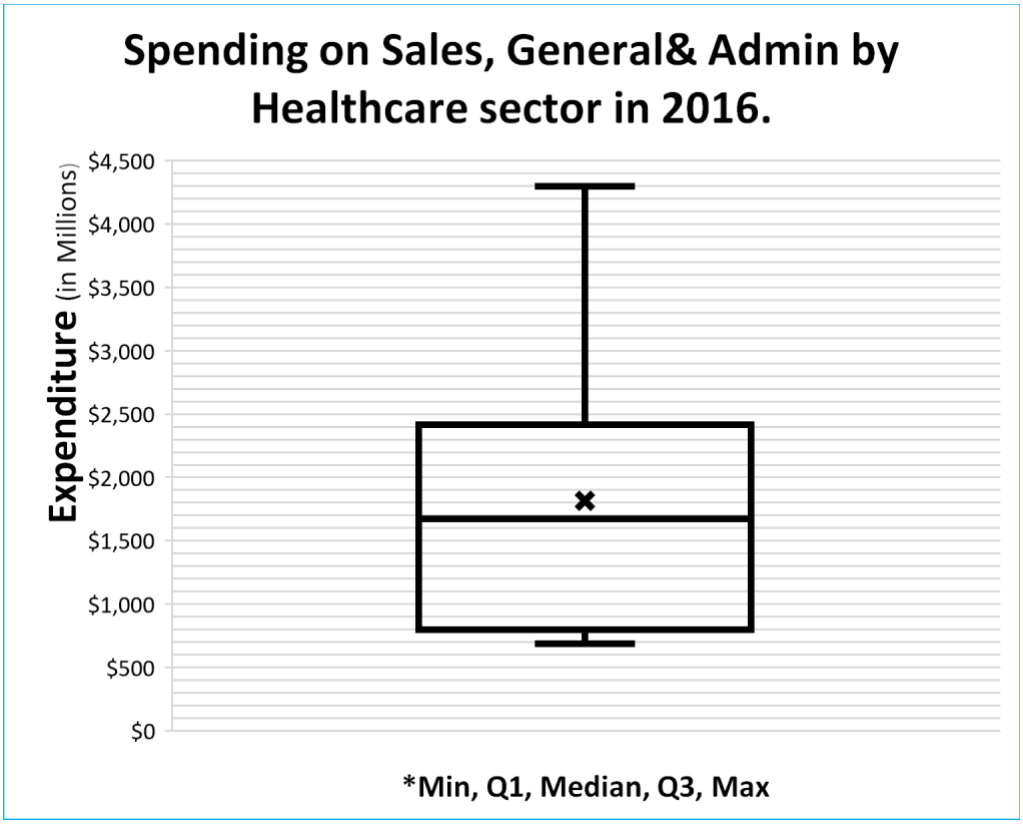
Project N#2: Analyze NYSE Data.

QUESTION: *Did Healthcare sector spend more on **Sales, General & Admin** than Information Technology in 2016?*

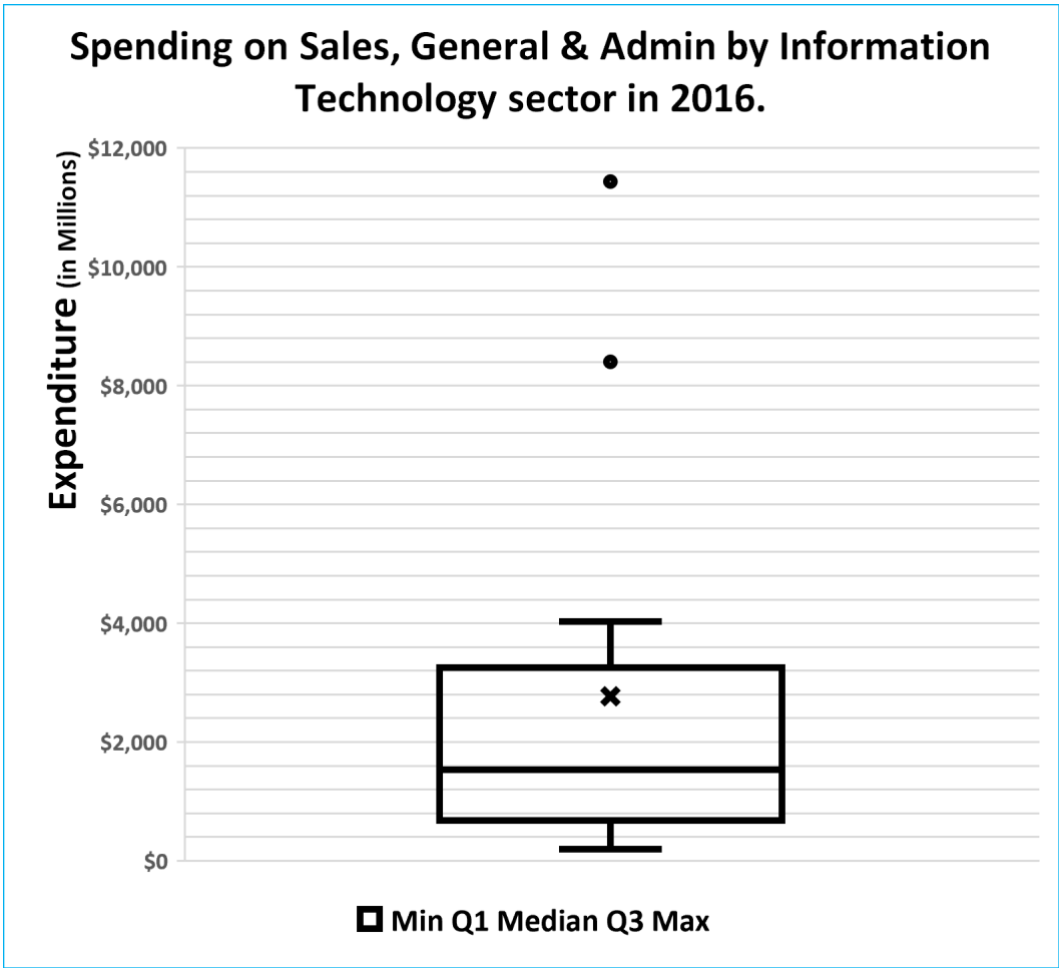
In the following paragraphs, 3 charts will be represented to answer the question above: in the first 2 charts, spending of Heathcare and Information Technology sectors on Sales, General & Admin is shown in visual manner, along with a concise statistical analysis. In the last chart, the average expenditure of these two sectors is illustrated (mean and standard deviation graphed) and summary statistics about the data is presented.

Chart N#1.



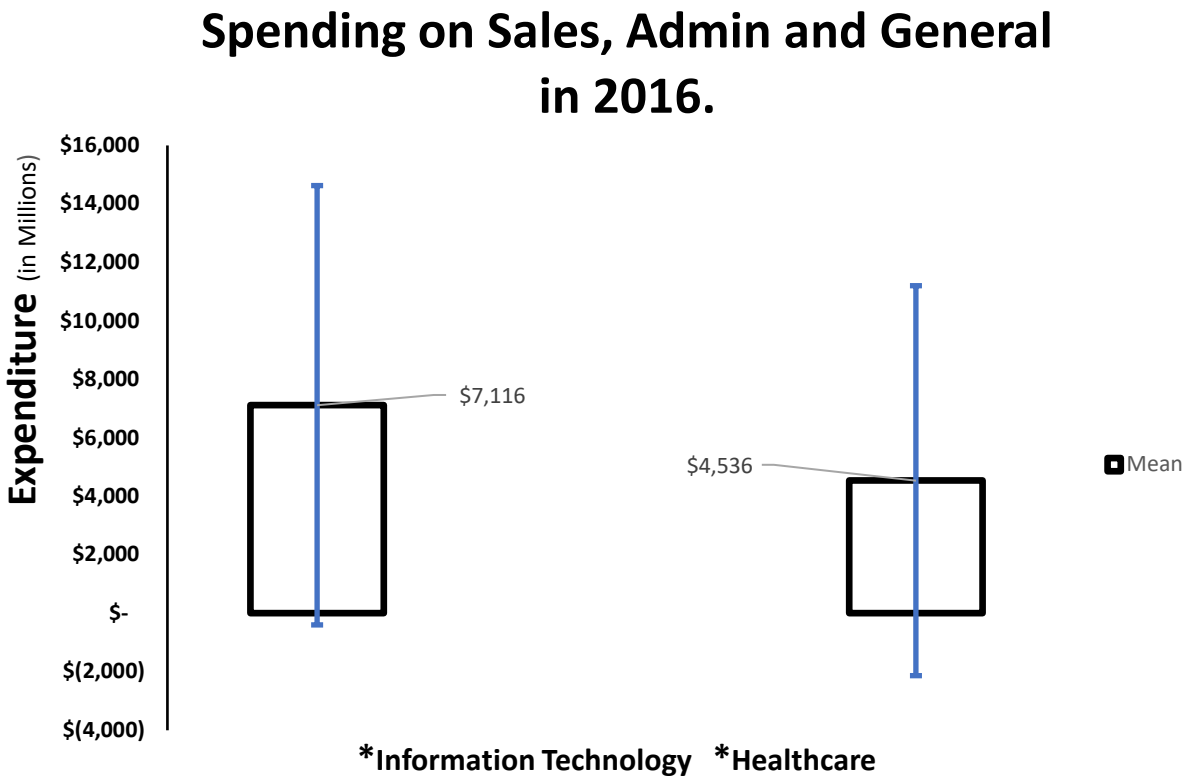
Description: To analyze the data, 5 Number Summary Technique is used. Min value is 682,400,000 and max is 4,295,000,000. It can be clearly seen that data distribution is right skewed (positive), the average amount of data (1,814,382,400) is higher than the median (1,668,450,000).

Chart N#2.



Description: It is better off using 5 Number Summary Method than mean and standard deviation when it comes to deal with the data that has outliers (two dots above). If there is a data set that greater than the amount (IQR multiplied by 1.5 and Q3 added), it is an outlier (in excel spreadsheet). It is shown that there are 2 outliers: 8,397,000,000 and 11,433,000,000 and the rest of the data is skewed positively.

Chart N#3.



Summary: On average, Information technology sector has spent more on "Sales, Admin and General" than Healthcare in 2016 ($7,116 and $4,536 millions respectively). It can be clearly seen that in both cases, standard deviation (blue line) is higher than the mean because of high variation between values, abnormal distribution (right skewed) for data and existing outliers (explained above). If we compare two standard deviations here, Information Technology has slightly higher indication than Healthcare (7,511 and 6,668 respectively) as to there is a wide range of variation amongst in IT sector data sets.