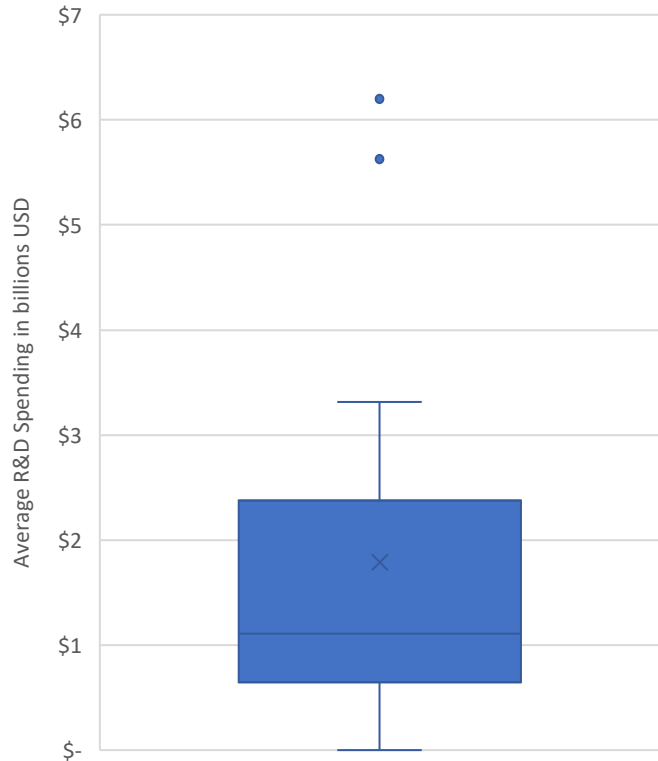


## **NYSE Data Analysis Project Question**

Which Information Technology sub-industry had the most average spending on R&D in 2016?

# Which Information Technology sub-industry had the most average spending on R&D in 2016?

Average R&D Spending by Sub-Industry of IT

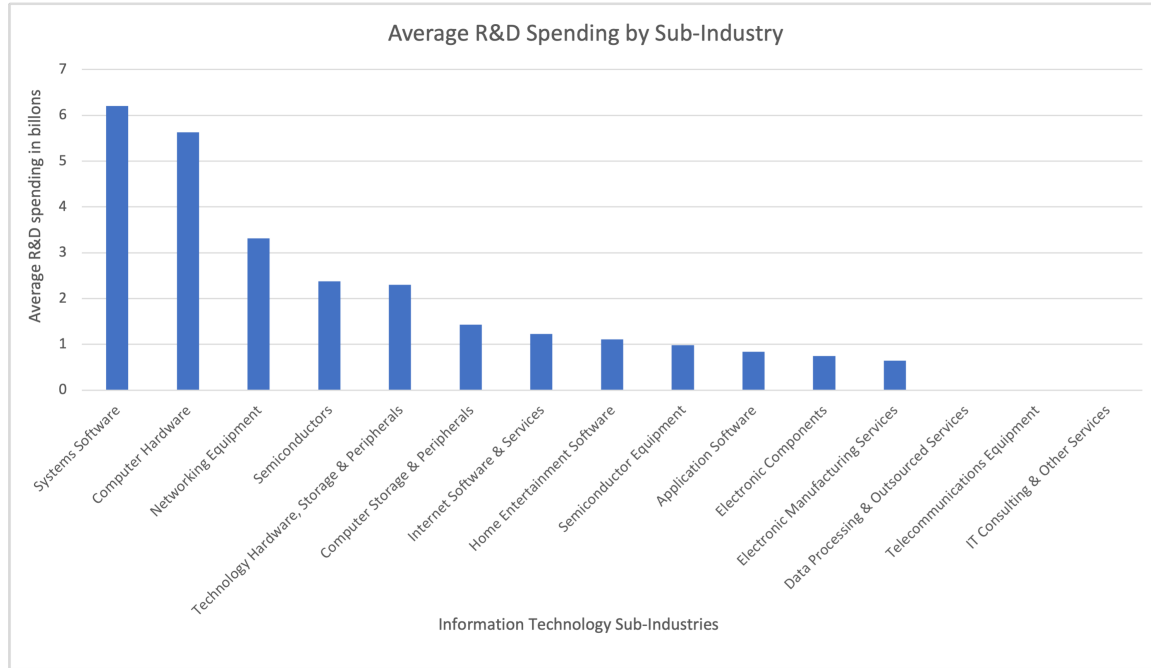


Looking at the box plot of average spending of IT subsectors on R&D, clearly there are outliers. The means that the spending on R&D is not evenly distributed on all subsectors of IT. There are two subsectors with the highest expenditures (6.2 and 5.6 billions) while the remaining sub-sectors are from range of 0 USD to 3.3 billions USD.

<b>MEDIAN</b>	\$861,000,000
<b>MEAN</b>	\$1,965,002,366
<b>STDEV</b>	\$3,107,139,263
<b>RANGE</b>	\$1,147,457,500

It is clearly defined that standard deviation is much higher. What does this mean? It means that there are sub-industries existing which their expenditure on R&D is almost more than the half of other sub-industries' average expenditure. It means outliers or extreme values are existing.

# Which Information Technology sub-industry had the most average spending on R&D in 2016?



Here is a bar chart of the average spending on R&D in 2016 by sub-industry. The data is right-skewed, resulting in a mean larger than the median.

The mean spending was \$1.96B, with a standard deviation of \$3.1B. This indicates most spending was by a small group of industries with systems software at the top with over 6 billion in their spending.

Of the top two industries, the range for Systems software at 11B, higher than the range of Computer Hardware at 8B. It appears the Systems software industry has more variability in R&D because their range is more spread out.