

What is the distribution of total revenue in year 2015 for
companies in Information Technology sector?
Is it symmetric?

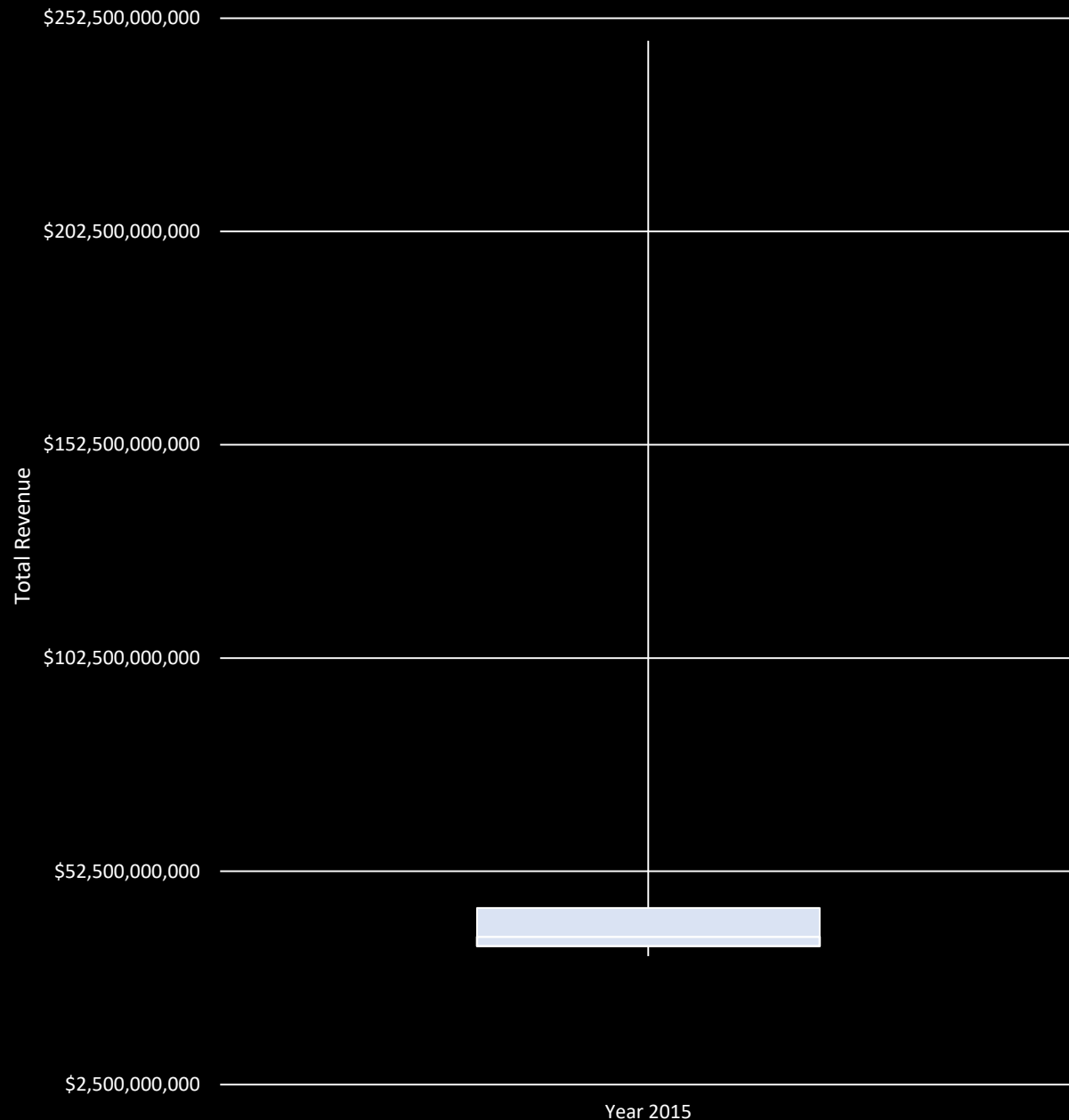
Project #2
Udacity Business Analytics Nanodegree
Javad Ebadi

Summary Statistics for Total Revenue of Companies in Information Technology Sector in year 2015

| Measure | Value |
|--------------------------|-------------------|
| Mean | \$16,035,034,618 |
| Standard deviation | \$32,615,202,621 |
| | |
| Minimum | \$1,059,366,000 |
| 1 st quartile | \$3,353,632,500 |
| Median | \$5,568,700,000 |
| 3 rd quartile | \$12,318,500,000 |
| Maximum | \$215,639,000,000 |
| Range | \$214,579,634,000 |

- Since the mean and median are not equal the distribution is not symmetric
- The standard deviation is rather big, it means that the total revenue of different companies varies a lot
- Since data is skewed the range is a better measure of spread than standard deviation
- The range is about \$200B which is a very wide range, which indicates a large spread
- Because the median is smaller than mean, the distribution is right skewed
- A box plot is a good choice to see these results in a visualization

Box plot for distribution of total revenue in Information Technology sector in year 2015



- Box plot for distribution of total revenue in Information Technology sector in year 2015
- The box plot show that there are huge outlier in total revenue and therefore the mean is not a good measure for center
- The median is a good measure of the center
- The maximum value is about 50 times the median value