Project #2

Evette Awa

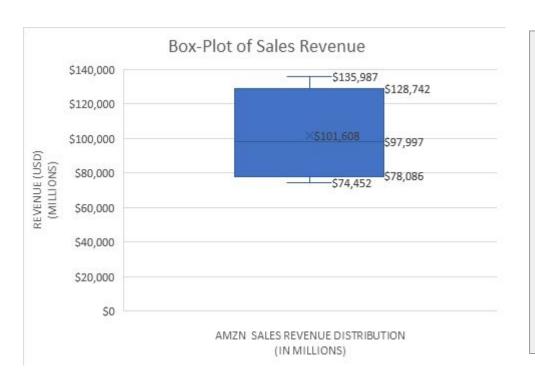
NYSE Data Project Submission

5/27/19



This is a Sunburst diagram of all of the industries under the GICS categorization. All featured sectors from our data set are broken down into these main groups..

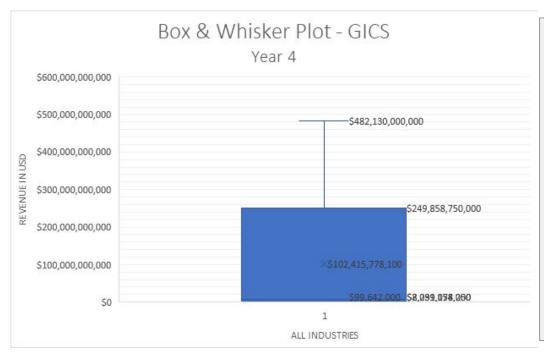
I used this chart to show the breakdown of my categorical data, as opposed to the bar chart that I had previously used. Using this graph, I can gather insights on the size of the market where AMZN plays (Consumer Discretionary). Here we can say that the company is a part of the biggest slice, which we can argue, where they contribute the most of all companies represented. You can see why in the next slides.



Here is a box-whisker plot for AMZN that includes all revenue data for Years 1 through Year 4.

I created this chart to show the measures of spread in the data. Here we can see not only the median but also the min, max, first quartile, and third. This is partially left-skewed but data is pretty consistent across the data set. No potential outliers. The standard deviation here is \$26,506,394,076, which you'll see in the next slide, is slightly more concentrated around the mean than spread out.

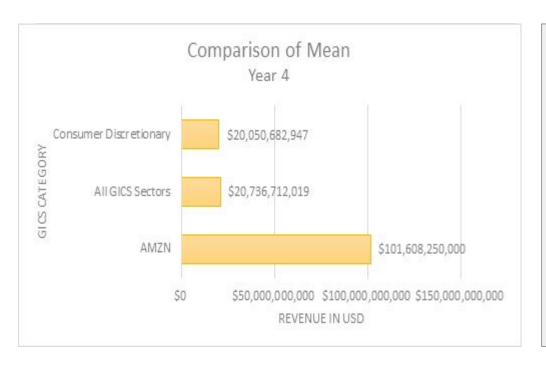
This distribution of data shows that the average of all AMZN sales revenue was \$101,608,250,000 and the actual median is \$97,997,000,000. This median amount is actually lower than the mean, but still above the overall revenues amount in the previous slide of \$20,736,712,000 for all GICS sectors. Amazon earned their highest annual sales in the amount of \$135,987,000,000 and the lowest was measured in its first year at the revenue amount of \$74,452,000,000. All of these figures are more than 5x the amount (on average) of all other sectors featured in this data set. For more details and graphs of findings, visit my workbook attached.



Here is a box-whisker plot for all GICS sectors and categories..

I created this chart to show the measures of spread in the data. Here we can see not only the median but also the min, max, first quartile, and third. This chart is left-skewed and not symmetrical in that most of the data sits around the first thirds of the data. It could be that the max (\$482,130,000,000) is an outlier, but that's what the data tells us.

This distribution of data shows the average of all sectors' sales revenue for Year 4. The standard deviation is \$39,921,574,021, which is calculated on my spreadsheet in my workbook. This measure of spread was necessary in visualizing it's proximity to the mean for this chart. The mean is slightly higher here at \$102,415,778,100 across all industries, as opposed to \$101,608,250,000 for AMZN for all four years combined. Insights gathered here is that this box plot proves that all in all, AMZN does NOT (although close) earn more annual revenue than all other categories, even though the previous plot shows a substantial difference between the categories for one year alone.



Here is a bar chart composed of data sourced from all GICS Sectors annual revenues for Year 4.

This chart shows us the calculated means for each category, and how it relates to the other. I chose to measure the data this way because we can use this to help interpret that AMZN sold more on average than not only the Consumer Discretionary Market, but in all of the GICS Sectors as well.

There were (76) companies in the given dataset with the GICS Sector 'Consumer Discretionary' and its mean was \$20,050,682,947. This figure was just about as much as the mean for all GICS Sectors, which was \$20,736,712,019. We can see that from this graph, that Amazon (AMZN) had nearly 5x as much in sales revenue in Year 4 as any other company, \$101,608,250,000. However, after gathering insight from the previous graphs, we understand that on average, AMZN is roughly the same in sales revenue as the overall GICS companies year over year. It looks skewed when comparing numbers on an annual basis, but when looking at measures of spread and center, we learn that over the 4 years, the gap isn't so big.