

Excel: Beyond Analytics

**Group 2:**

Angel Cuevas

Chaminda Ariyaratne

Dylan Lam

Joyce Lam

California State University, Los Angeles

**Group Information**

|  |  |  |
| --- | --- | --- |
| **Member Name** | **Percent Contribution** | **Activities Completed by Member** |
| Angel Cuevas | 25% | Visuals 1, 2, 3 |
| Chaminda Ariyaratne | 25% | Visuals 4, 5, 6 |
| Dylan Lam | 25% | Visuals 7, 8, 9 |
| Joyce Lam | 25% | Visuals 10, 11, 12 |
| **Total** | **100%** |  |

**Table of Contents**

[Abstract](#_heading=h.gjdgxs) 4

[Sample](#_heading=h.30j0zll) 5

[Visual 1](#_heading=h.1fob9te) 6

[Visual 2](#_heading=h.3znysh7) 7

[Visual 3](#_heading=h.2et92p0) 8

[Visual 4](#_heading=h.tyjcwt) 9

[Visual 5](#_heading=h.3dy6vkm) 10

[Visual 6 1](#_heading=h.1t3h5sf)1

[Visual 7 1](#_heading=h.4d34og8)2

[Visual 8 1](#_heading=h.2s8eyo1)3

[Visual 9 1](#_heading=h.17dp8vu)4

[Visual 10 1](#_heading=h.3rdcrjn)5

[Visual 11 1](#_heading=h.26in1rg)6

[Visual 12 1](#_heading=h.lnxbz9)7

[References](#_heading=h.2jxsxqh) 18

**Abstract**

The objective of this project aims to show that all participating members of this group fully understand the basic concept of data visuals created in Excel with the use of Pivot Tables, Pivot Charts, and graphs. Our group will utilize data within Excel files provided by Professor Pantaleon. Our group will also create visuals based on videos from the ‘Additional Student Resource’ section of *Business Analytics: Communicating with Numbers, 1st Edition*. Moving forward, every member within our group will provide a brief description that describes what the data visualization is showing and what video they used for that specific visual. Furthermore, each data visualization within this project is unique and different from one another.

# Sample

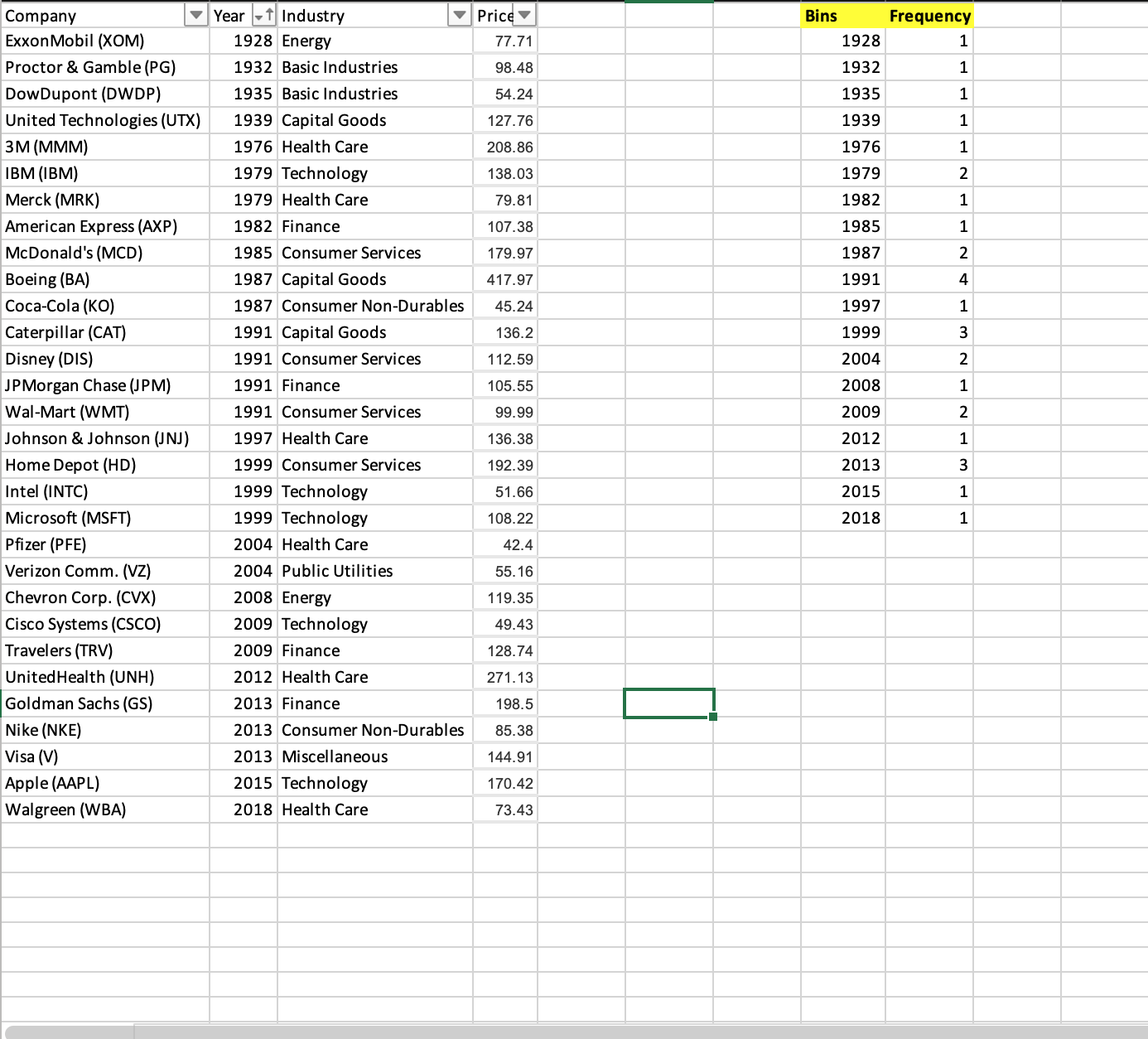
This is an example of regression analysis with beds, bath, sqft, and rent as variables. It includes regression statistics, ANOVA, and Coefficients.

Table

Description automatically generated

# Visual 1

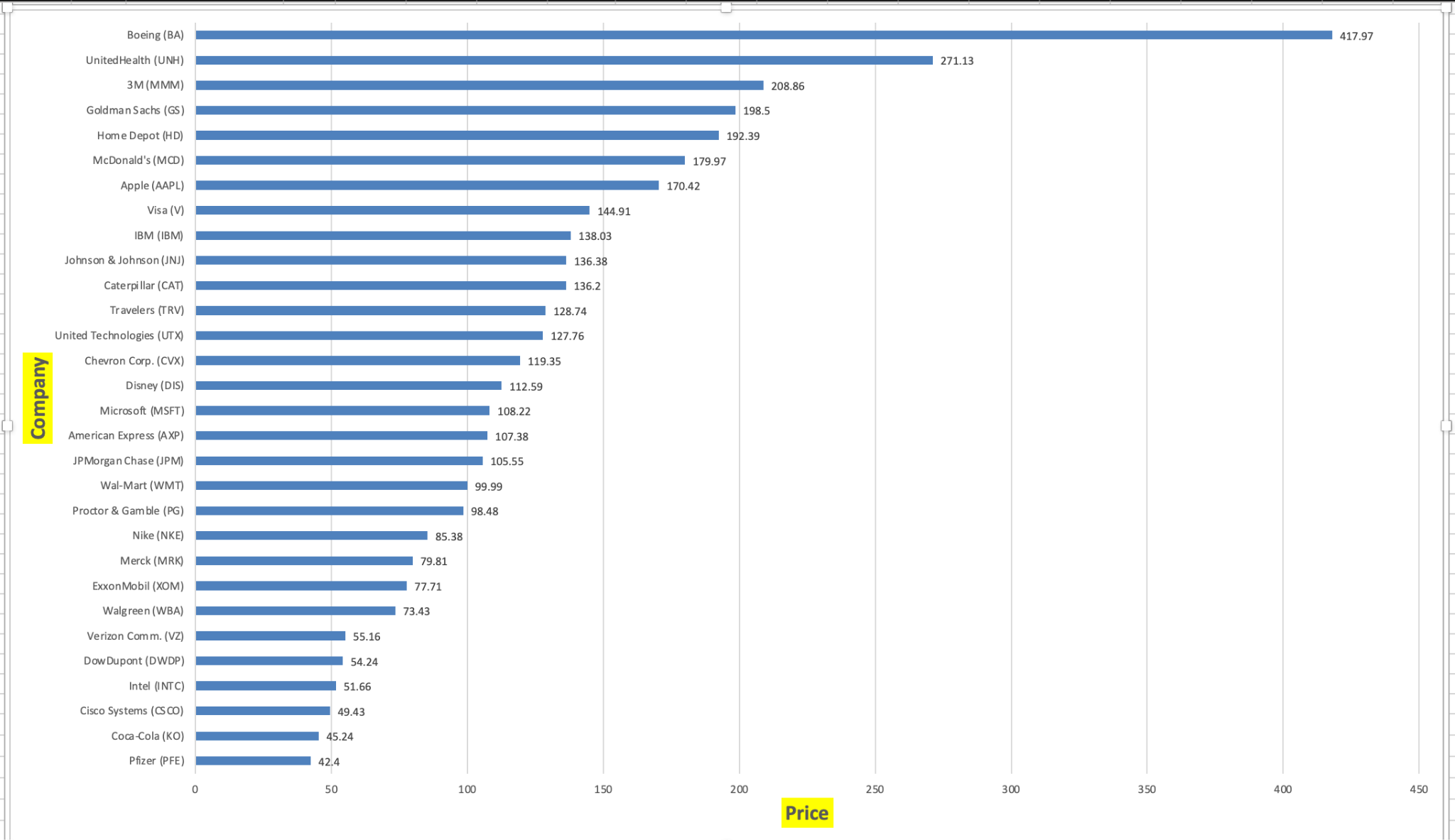
This is an example of a frequency table using the frequency array in Excel. The highlighted tables (on the right side) represent the Bins and the Frequency. The Bins display the Year and the Frequency displays the number of times each specific year appears in the dataset. (Video #01)



# 

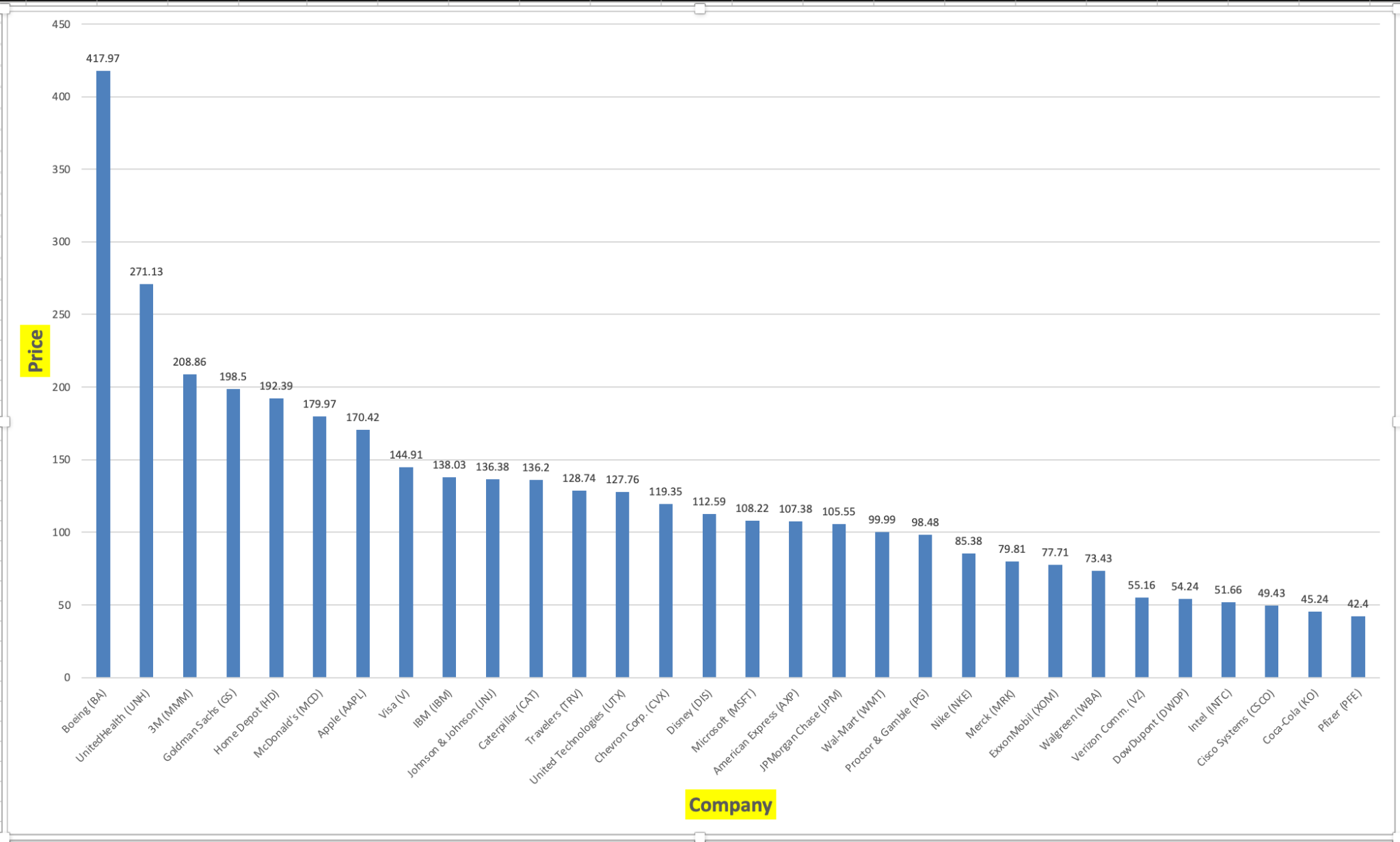
# Visual 2

This second visual is an example of a horizontal bar chart. The X-axis represents the DOW Jones stock price and the Y-axis represents a company. The overall representation is the price of the stock price of each company in ascending order. The most expensive stock price is the Boeing company at $417.97 at the top of the chart and the least expensive stock price is Pfizer at under $42.40 at the bottom of the chart. (Video #02)



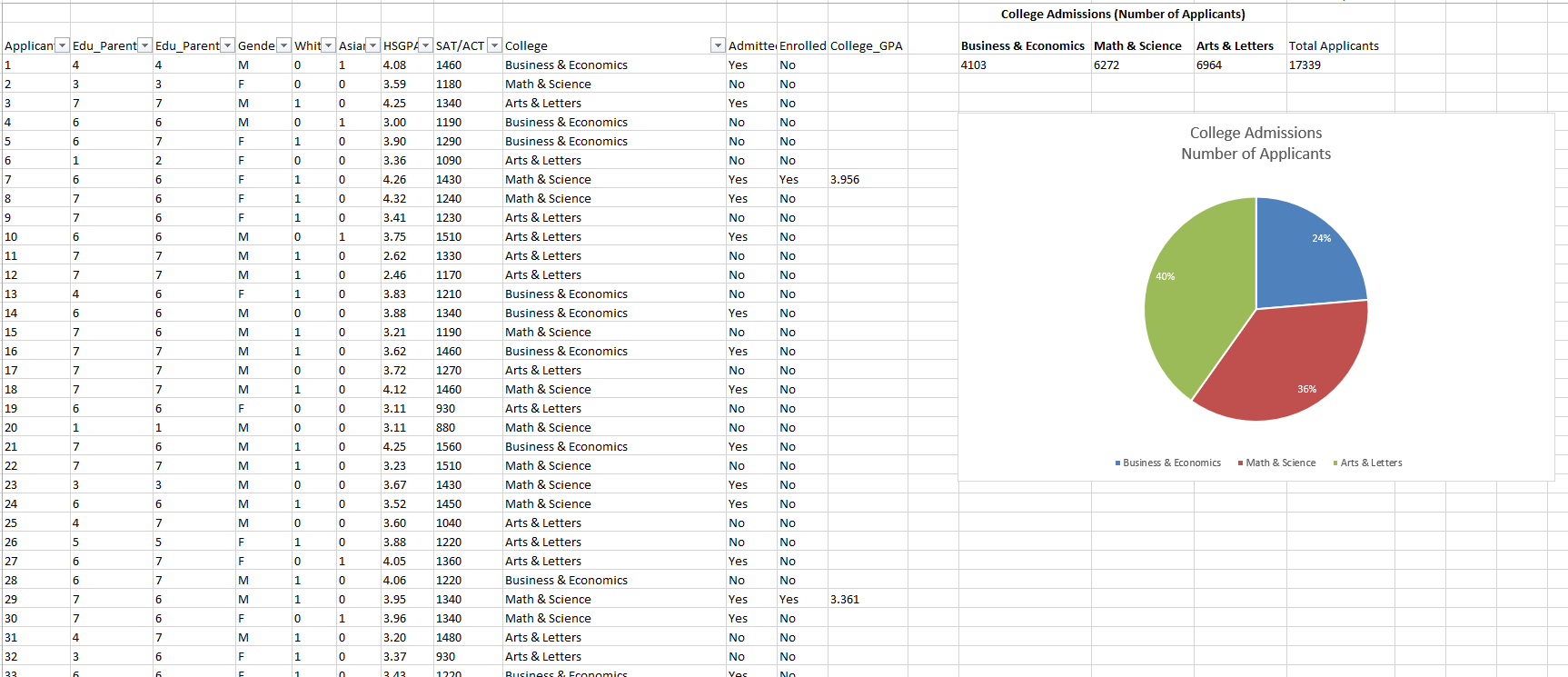
# Visual 3

This third visual is an alternate format of the previous visual. This is an example of a vertical bar chart. The information displayed is the same, it is just the visual representation that has changed. On the Y-axis we have the price of the DOW Jones stock and on the X-axis we have the name of the company that belongs to each price. Listed in descending order, on the far left we have the Boeing company with the highest stock price of $417.97 and at the far right we have the Pfizer company with the lowest stock price at $42.40. (Video #03)



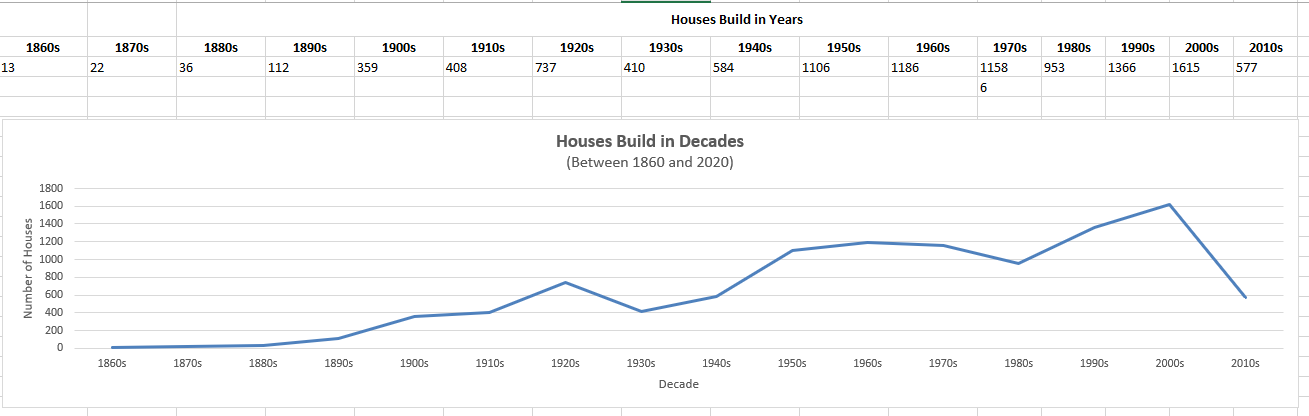
# Visual 4

The fourth visual is an example of a bar chart view of the college applicant’s area of study. The collected data represent three different types of colleges: Business & Economics, Math & Science, and Arts & Letters. Seventeen thousand three hundred thirty-nine applicant data were processed in the example. The bar chart shows the total number of applicants for each college as a percentage of the total. (Video #04)



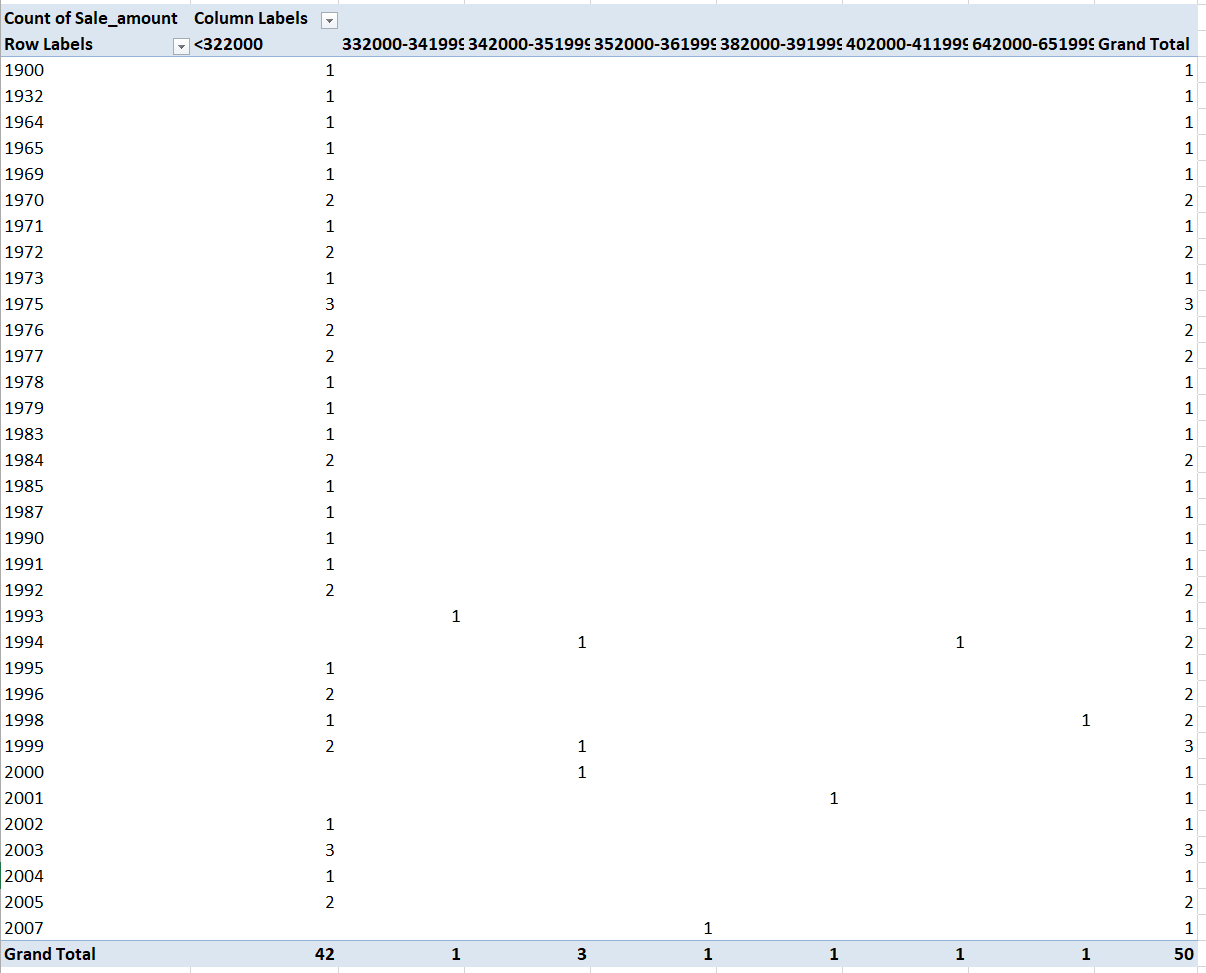
# Visual 5

The fifth visual is an example of a line chart. The data sample contained 10659 collected house data with their prices, built year, town, state, sale price, etc. This visual represents the decades' houses were built based on the data. Homes on the market before 1860 were low, so the data was not included in this visual. The line chart shows that the construction of the new house went up in 1950 and the trend declined slightly in 1980 but went back up in the 1990s and continued to the 2000s. (Video #05)



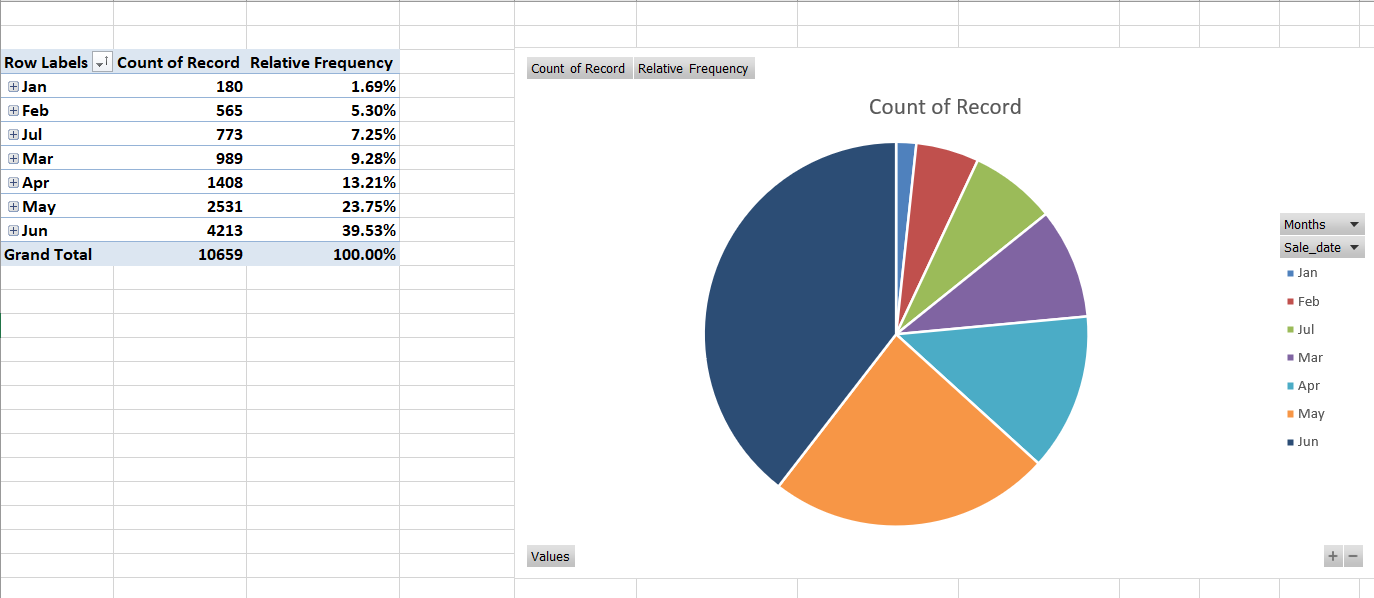
# Visual 6

The sixth visual is an example of joint frequency distribution using a pivot table based on the previous housing sales data. The first 50 sales were selected to keep the pivot table more manageable and visually simple. The pivot table sales data were grouped to reduce too many data points. The pivot table shows how the prices fell into different price classes. (Video #06)



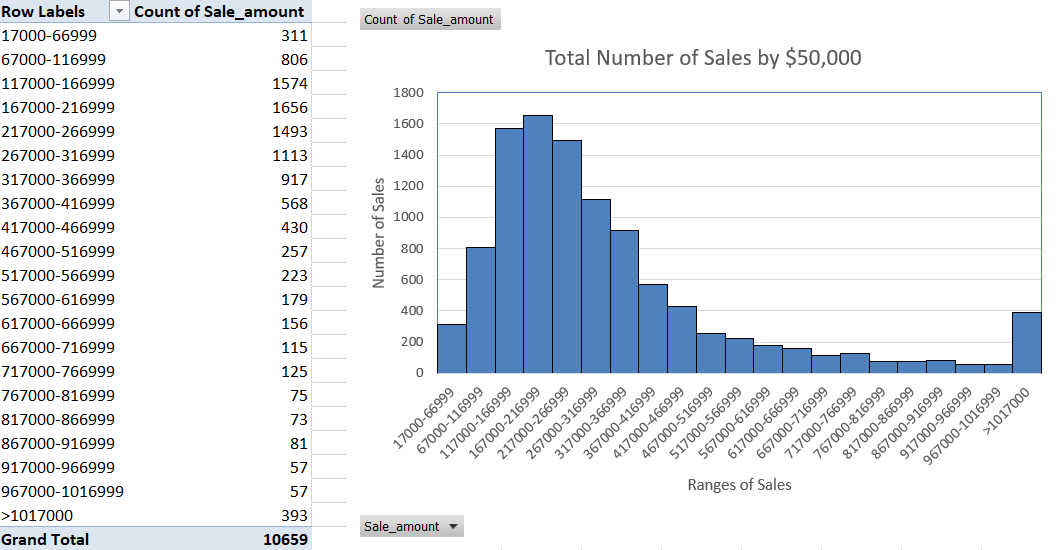
# Visual 7

The seventh visual contains a frequency table built by using a pivot table. The Count of Record is the number of records (or occurrences) in a specific month if not a day in that month. On the right-hand side of the Count of Record, there is a column that calculates the relative frequency of the number of records. The relative frequency is calculated by finding the number of occurrences within a certain month divided by all occurrences in all months in the table. Furthermore, there is a pie chart to help the audience visualize the data better. Here, we can see that June is taking up the majority of the pie chart by 39.53%. (Video #7)



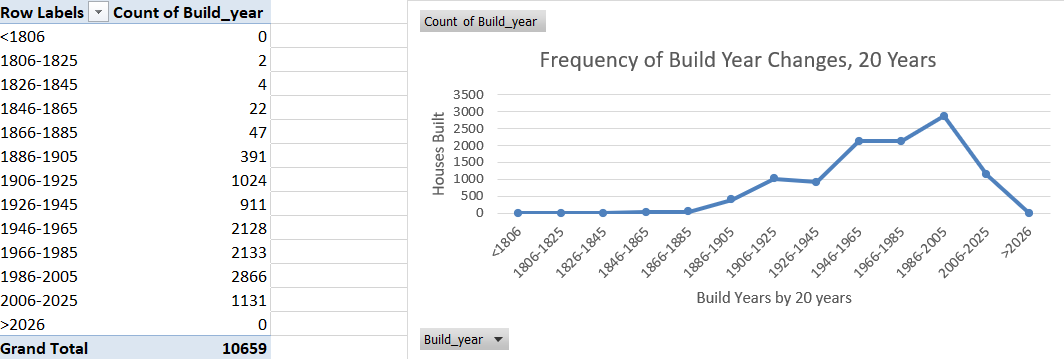
# Visual 8

The eight visual contains a frequency distribution built by using a pivot table. The Count of Sale Amount is the number of sales (or occurrences) in a specific range. Going into this further, there is a clustered column chart that helps the audience visualize the data better. Looking at the chart, the dominant column would be the number of sales in the range of 167,000 to 216,999. Plus, if we look at the overall view of the chart, we can see that the chart is skewed to the right as it has a tail on the right side. Furthermore, the range of 1,017,000 and beyond is excluded in this case. The reason is that if we group the values under Row Labels by default, then we would see all of the data in 5 rows. However, there would only be one row being 98% as other rows would be less than 1%. Therefore, we choose the range of 17,000 to 1,016,699 to group our data. (Video #8)



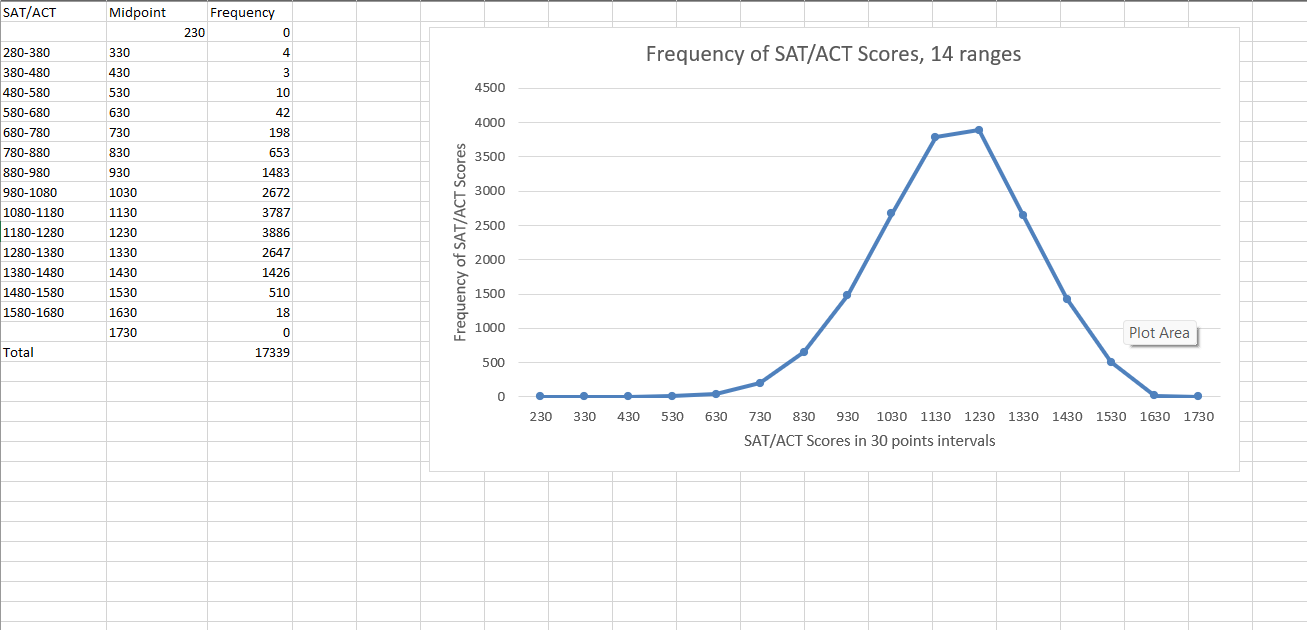
# Visual 9

The ninth visual contains a frequency polygon chart built by using a pivot table. The Count of Build Year is the number of houses built in a specific year in a specific range of 20 years. Going into this further, there is a line chart with markers that helps the audience visualize the data better. Looking at the chart, the dominant interval would be the number of houses built from 1986 to 2005. Furthermore, we can see that the line chart is skewed more towards the left as the majority of houses are built after the 1900s. (Video #10)



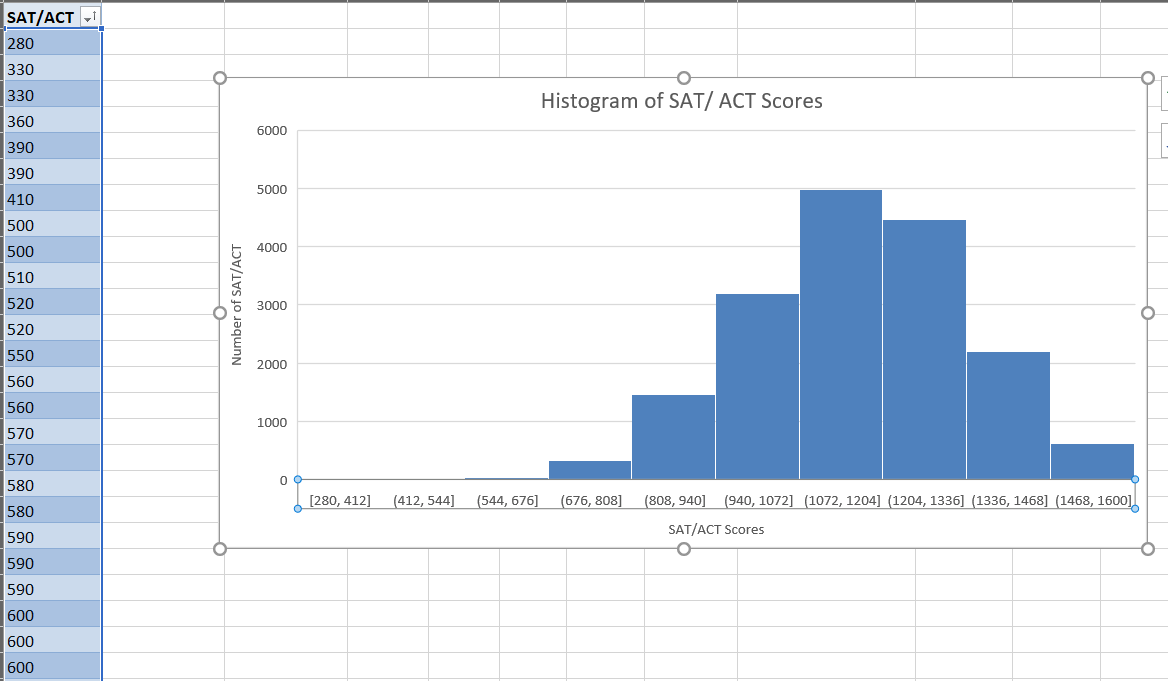
# Visual 10

The tenth visual contains a frequency polygon chart built by using frequency distribution. The Frequency is the number of occurrences of SAT/ACT scores in 14 specific ranges. This is a line chart with markers that will allow the audience to visualize the data. Looking at the line chart, you can see most students scored around 1180-1280. (Video #11)



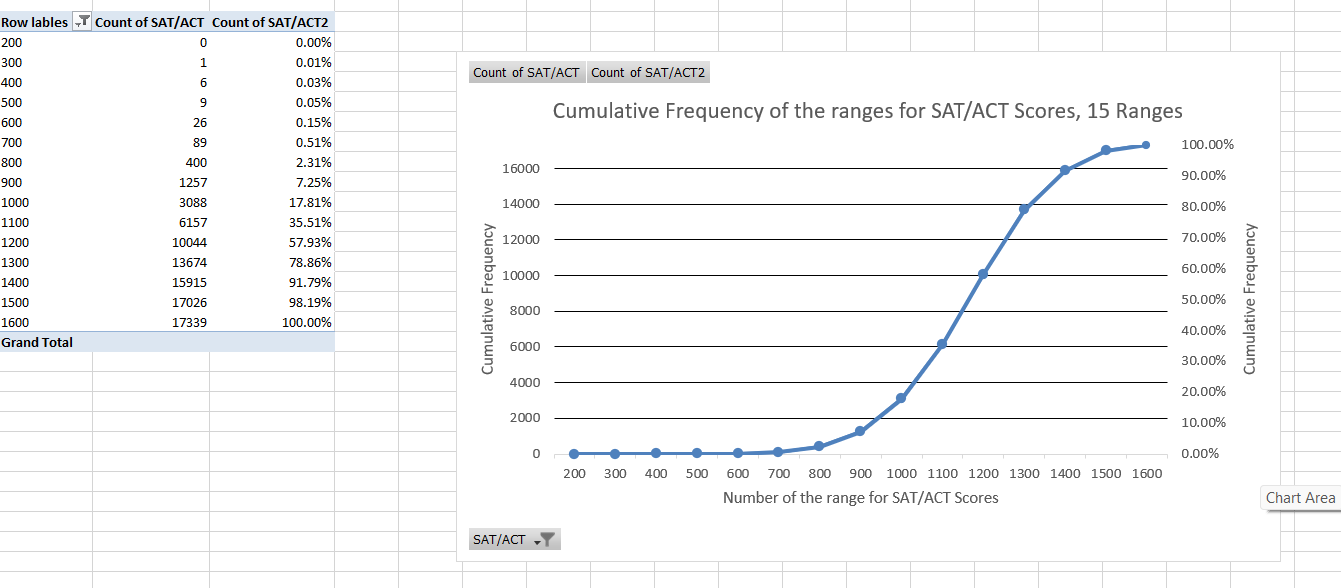
# Visual 11

The eleventh visual contains a histogram built by using the histogram feature. This histogram condenses the SAT/ACT scores by grouping them into bins, so it would allow the audience to easily interpret the chart. (Video # 12)



# Visual 12

The twelfth visual contains a cumulative frequency polygon chart built by using a pivot table. The Count of SAT/ACT is the number of occurrences in a specific range of scores. On the right-hand side of the Count of SAT/ACT, there is a column that calculates the percent of the running total of the number of occurrences. This is a line chart with markers that will allow the audience to visualize the data better. (Video #13)



**References**

Jaggia, S., Kelly, A., Lertwachara, K., & Chen, L. (2021). *Business analytics: Communicating with numbers*. McGraw Hill LLC.