A3: Tableau Visualizations

## CSE412: Data Visualization

Drake Watson, drakerw@uw.edu, Section AB

Video Game Sales Dataset

For my analysis I chose to work with the Video Game Sales Dataset from the website Kaggle that contains data on regional and worldwide game sales from 2013-2018. I chose this dataset because I play lots of video games myself and I was also intrigued to see the way I could visualize this data because over the last decade video games have exploded in popularity/profitability. Particularly, Grand Theft Auto V (the most profitable video game of all time) was released during this period of time, so I was intrigued to see the way that visualized within the data set.
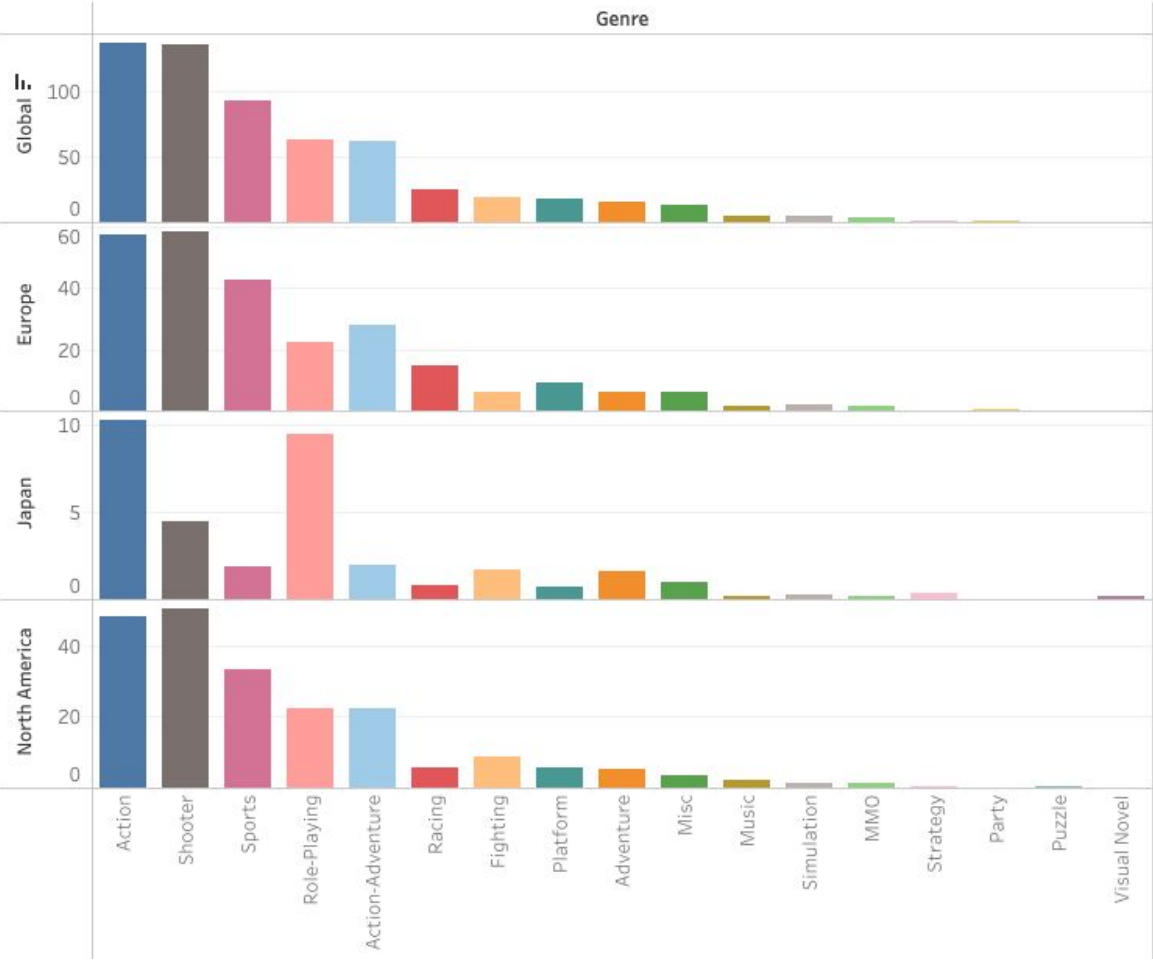
Link: https://www.kaggle.com/datasets/sidtwr/videogames-sales-dataset

Total Global Sales By Year
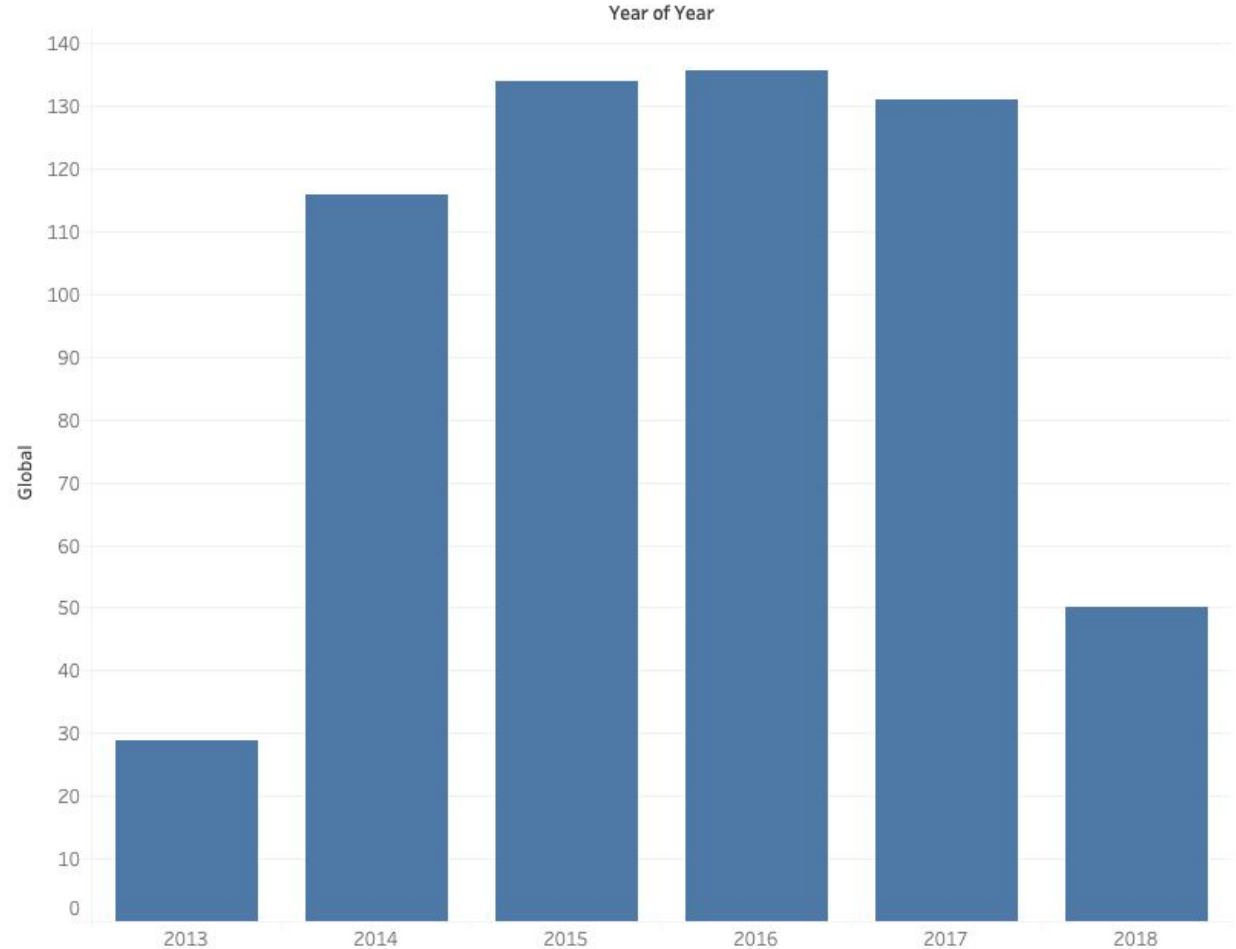
This visualization was used to both see the way that the data was distributed and also to see the way that the middle years compared to one another in overall sales numbers

Genre/Country Sales

The goal of this comparative bar chart is to see the distribution of of global sales across genres and also to see the way that the three largest regions differ in genre preference
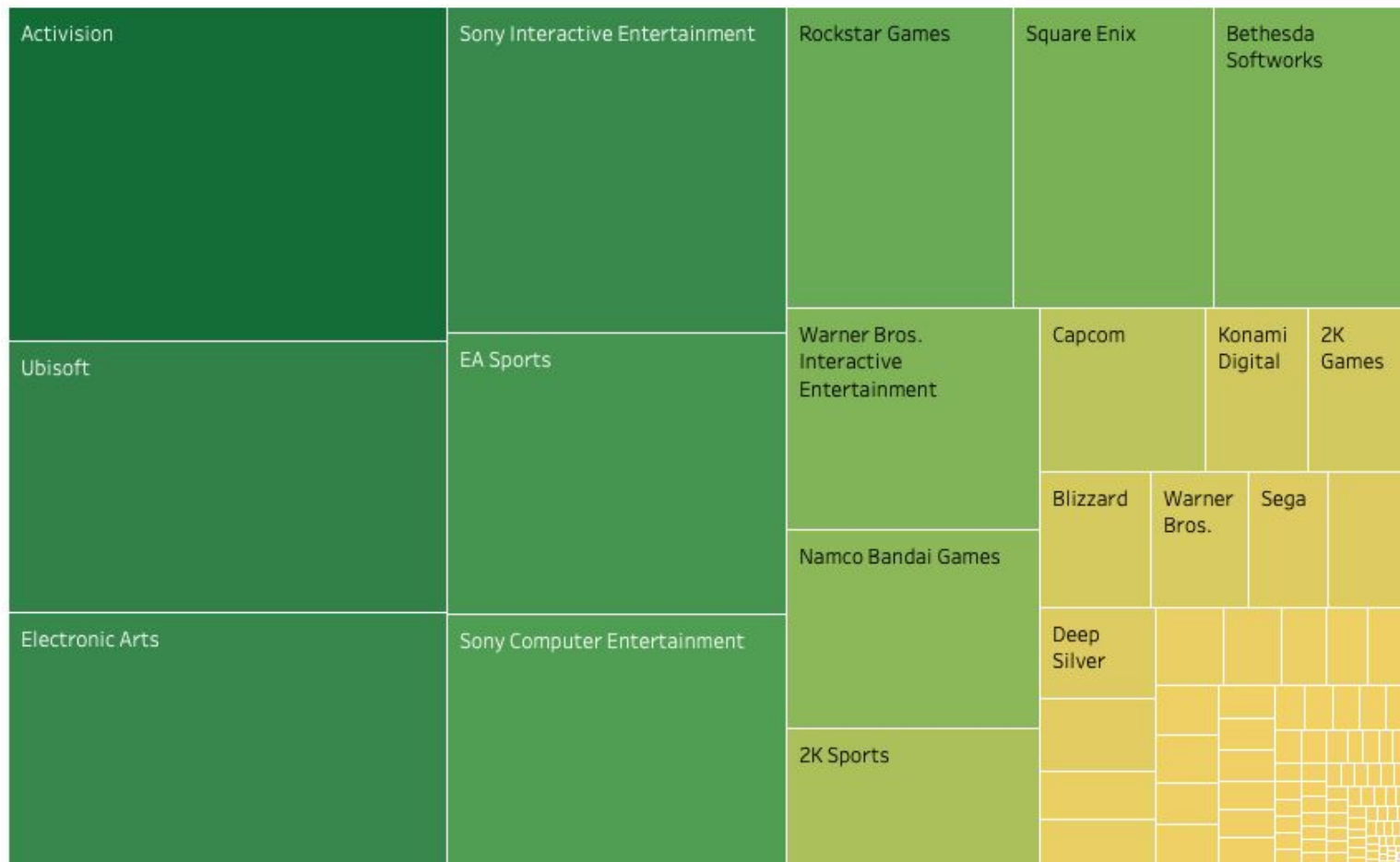
Genre Sales By Year
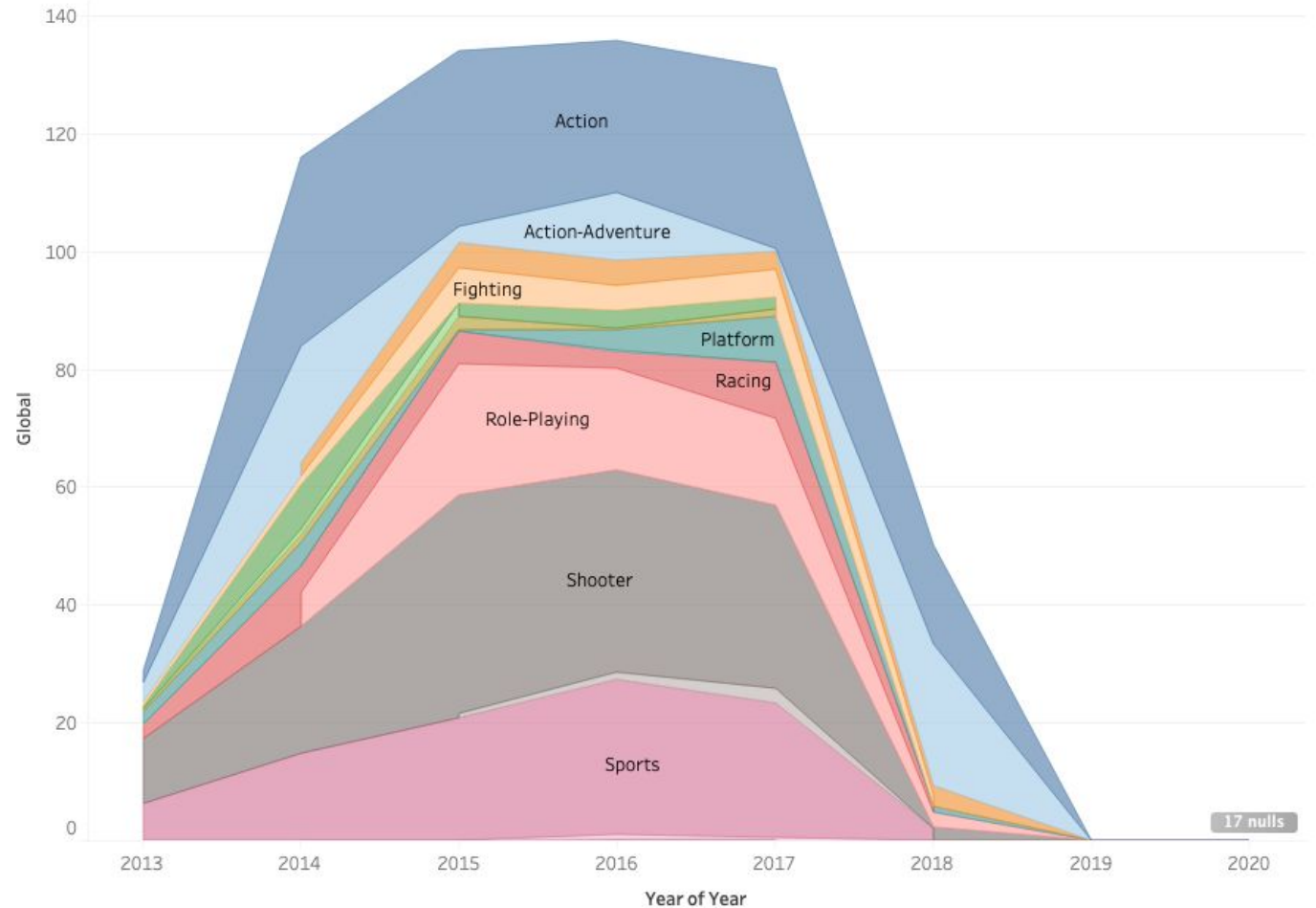
For this area/line chart I wanted to observe the way that the that genres grew/shrunk in popularity over this period in time

Publisher Sales

The goal of this visualization is to show what publishers dominated sales and to see the disparity between them by using a color scale that highlights the vast gaps in sales

Activision

Sony Interactive Entertainment

Rockstar Games

Square Enix

Bethesda Softworks

Ubisoft

EA Sports

Warner Bros. Interactive Entertainment

Capcom

Konami Digital

2K Games

Blizzard

Warner Bros.

Sega

Electronic Arts

Sony Computer Entertainment

Namco Bandai Games

Deep Silver

2K Sports

Publisher Genres

| Genre	2K Games	2K Sports	Activision	Bethesda Softworks	Capcom	EA Sports	Electronic Arts	Namco Bandai G..	Rockstar Games	Sony Com puter..	Sony Inter active En..	Square Enix	Ubisoft	Warner Bros. Int..
Action			7	5	16		1	16	1	6	13	7	12	14
Action-Adventure	1		2		4			1	1	3	6	3	4	3
Adventure			1					1	1	1	4	2	1	3
Fighting					3		1	10				1		3
Misc			2		1			4		1	1	3	7	2
MMO				2								4		
Music			1					1		1	3		7	
Party												2		
Platform			4		2		1	1		3	3		1	1
Puzzle			1										1	
Racing							3	2		2	3		4	1
Role-Playing				3	2		3	17		2		20	2	
Shooter	4		11	7	2		8			3	6		5	
Simulation							1	2			1			
Sports	1	7	1			19	3			3	4		1	
Strategy			1					1			1			

By isolating the top publishers from the previous visualization, I was able to uncover the different genres that they focus on and analyze the way that the genres are distributed between them

Sales By Game

FIFA 18
Call of Duty: Black Ops 3
FIFA 17
Fallout 4
Red Dead
Call of Duty: WWII
Grand Theft Auto V
FIFA 16
FIFA 15
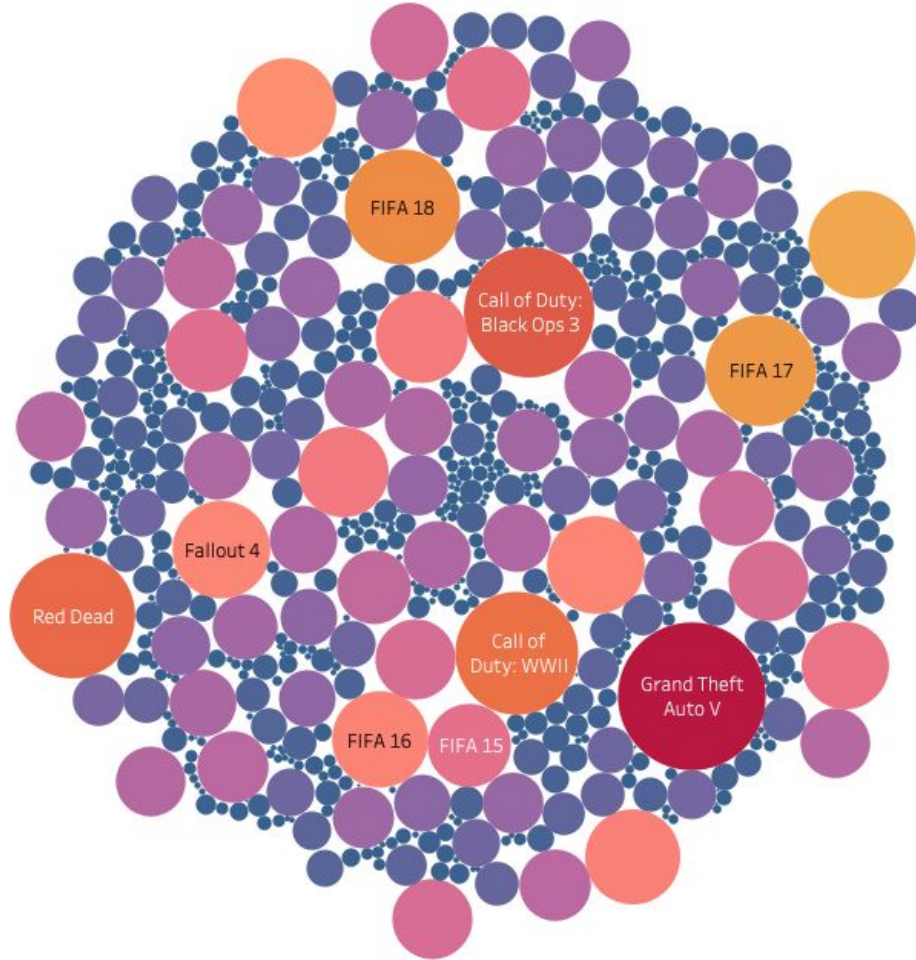
In order to see the distribution of game sales by title I made this visualization with a diverging color scale in order to show how the top sellers compare to the vast majority of the games in the dataset

Game B&W
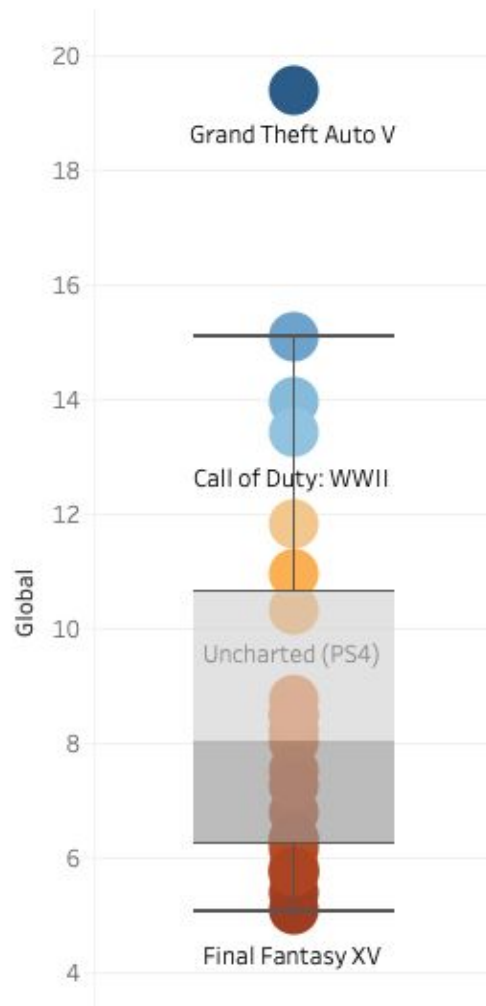
The focus of this visualization was to show the way that GTA V compared to even the highest selling games of the dataset, so by filtering for games that sold $5 million or more and placing them in a box and whisker plot it's easy to see that it's still a significant outlier

2014 Sales

Year of Year

To further display how great of an outlier GTA V is I filtered for only 2014 sales and then placed the games into a pie chart, and after doing so it became even clearer out great of a disparity there was between GTA V and every other game released that year

2014
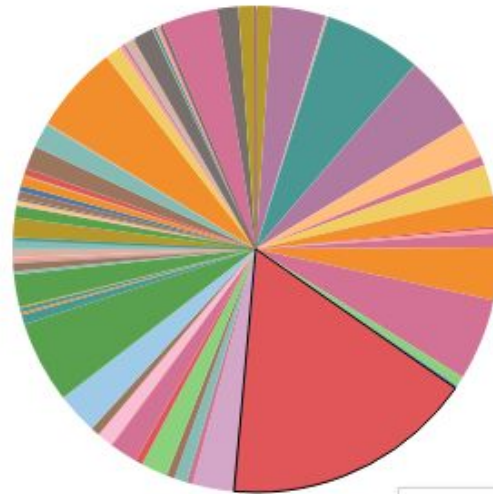
Game:	**Grand Theft Auto V**
Year of Year:	**2014**
Global:	**19.39**

Reflection

<u>Struggles</u>: There were several times that I struggled with finding the ways to visualize the answers to the questions that I had, and it wasn't after going back and reformatting the year attribute to a date value and finding ways to filter and label certain aspects of the data that I was able to visualize most of the ideas that I had.

<u>Enjoyed</u>: Finding different color styles and scales were both very enjoyable parts of the process as they played a major role in answering the questions that I had about the data. Learning how using two colors on a divergent scale show the way that certain data points are significant outliers in the context of the dataset.

<u>Learned</u>: This is an exercise that I've done many times in different programming languages and contexts, but it further certified how important it is to analyze the dataset before visualization (string vs date values, finding null values, etc.) and how much of effective visualization comes from a thorough understanding of the dataset and it's formatting.