Data Visualization-Final Report

1. **Section 1. Introduction** - Write a brief introduction of your project. Include clickable link to your published project if available.

   https://public.tableau.com/app/profile/cole.schmidt1799/viz/Schmidt_FinalProject/Story1?publish=yes

   A fighter's effectiveness in the UFC is often boiled down to just a fighter's record which is a record of a fighter's highest and lowest moments in their sport. The UFCs welterweight division is arguably the most competitive division in the sport currently with many fighters using their own unique and diverse styles to defeat their opponents and climb the rankings. In this final report I will use all relevant stats such as, finish rate, striking efficiency, method of victory, etc. to determine how fighters win at the highest levels of the sport. My report will evaluate a fighter's efficiency, specialization, tolerance to risk, and predictability to determine if certain styles are better for success in the UFC's most competitive division.

2. **Section 2. Data Description**

3. Describe the domain of the data set

   This data set was made by me using publicly available stats from the UFC's official rankings and fighter's official records. Each row was manually made by me to represent one fighter while each column describes performance data or a calculated field, I made to make analysis easier. I'm able to group the columns into one of three groups.

   The first group is used to identify the fighter, their fighting style, and their rank in the welterweight division. The second group is the fighters record information such as, wins/losses, wins split by method of victory, total fights, and total finishes (separated by fighter of course). Using data in the second group I expanded it by creating calculated fields such as, win rate, number of losses, KO rate, sub rate, and decision rate. The third group is raw data from the fighters performance in the cage, it is usually averaged and changes very often; these data types are, significant strikes landed per minute, significant strikes absorbed per minute, takedowns per 15 minutes, takedown accuracy, sub attempts per 15 minutes, and striking efficiency (this last one is a calculated field).

4. Describe the data file (be specific!)

   The data file is a tableau desktop .twbx file that contains my project which I have separated into 6 worksheets that each solve a goal I made, one dashboard that combines data from all 6 goals to give a good overview/final look at the project, and finally a story that contains my analysis of my goals and charts made from the raw excel data set.

5. Clearly state the number of rows and columns you are using in this project.

I'm using 16 rows which show the top 15 ranked fighters in the welterweight division and the champion. I also have 21 columns that show a fighter's relevant stats which have already been listed in the report. In total the rows and columns together make 336 individual data points.

6. Describe the data source (be specific!)

The data source I made is a .xlsx file that I created by combining raw data the UFC collects and data from calculated fields I made to go a bit more into depth on the fighter's skills and stats.

7. **Section 3. Data Cleaning Strategies** - what did you do to clean your data? If no cleaning was needed, state this.

I only had to do a small amount of cleaning for this project as I made the data sheet myself. The cleaning I did was create calculated fields, switching data's numeric format to and from percentage values.

8. **Section 4. Clean Dataset** - Describe the clean dataset that you will visualize. Show an excerpt of the data (you don't have to show all the rows if your data set is huge).

Fighter name	Rank	Wins	losses	KO Wins	Sub Wins	Decision Wins	Significant Strikes Landed per Min	Significant Strikes Absorbed per Min	Takedowns per 15 Minutes	Takedown Accuracy	Sub Attempts per 15 Minutes	Total Fights	Finish Wins	Finish Rate	Win Rate	KO Rate	Sub Rate	Decision Rate	Striking Efficiency	Style
Islam M	C	28	1	5	13	10	2.45	1.45	3.1	56%	0.98	29	18	64.29%	96.55%	17.86%	46.43%	35.71%	1	Grappler
Jack D M	1	18	3	12	2	4	5.57	3.84	0.13	10%	0.1	21	14	77.78%	85.71%	66.67%	11.11%	22.22%	1.73	Striker
Ian M G	2	17	1	7	1	9	4.78	2.86	0.92	31%	0.37	18	8	47.06%	94.44%	41.18%	5.88%	52.94%	1.92	Striker
Michael M	3	19	0	14	1	4	5.68	3.26	1.04	42%	0	19	15	78.95%	100.00%	73.68%	5.26%	21.05%	2.42	Striker
Belal M	4	24	5	5	1	18	4.43	3.82	2.14	37%	0.14	29	6	25.00%	82.76%	20.83%	4.17%	75.00%	0.61	Grappler
Carlos P	5	23	7	18	3	2	3.77	4.53	0.21	100%	0	30	21	91.30%	76.67%	78.26%	13.04%	8.70%	-0.76	Striker
Sean B	6	18	2	3	6	9	3.96	3.05	3.53	53%	0.91	20	9	50.00%	90.00%	16.67%	33.33%	50.00%	0.91	Grappler
Kamaru U	7	21	4	9	1	11	4.16	2.67	2.79	44%	0.09	25	10	47.62%	84.00%	42.86%	4.76%	52.38%	1.49	Grappler
Leon E	8	22	6	7	3	12	2.6	2.44	1.25	37%	0.36	28	10	45.45%	78.57%	31.82%	13.64%	54.55%	0.16	Striker
Joquin B	9	21	7	15	0	6	3.88	2.93	1.54	45%	0.08	28	15	71.43%	75.00%	71.43%	0.00%	28.57%	0.95	Striker
Gabriel B	10	19	1	4	13	2	4.61	3.68	3.6	56%	1.44	20	17	89.47%	95.00%	21.05%	68.42%	10.53%	0.93	Grappler
Gilbert B	11	22	9	6	11	5	3.15	3.64	2.12	38%	0.46	31	17	77.27%	70.97%	27.27%	50.00%	22.73%	-0.49	MMA
Geoff N	12	16	7	10	2	4	5.05	5.48	0.54	45%	0.11	23	12	75.00%	69.57%	62.50%	12.50%	25.00%	-0.43	Striker
Daniel R	13	20	5	9	4	7	7.2	5.4	0.51	50%	0.08	25	13	65.00%	80.00%	45.00%	20.00%	35.00%	1.8	Striker
Michael P	14	24	3	13	3	8	2.39	1.68	0.23	17%	0	27	16	66.67%	88.89%	54.17%	12.50%	33.33%	0.71	Striker
Colby C	15	17	5	4	4	9	3.81	3.09	3.64	43%	0.16	22	8	47.06%	77.27%	23.53%	23.53%	52.94%	0.72	Grappler

I don't think my data set is huge but it is very descriptive.

9. **Section 5. Visualization Tools** - Tell us which tools you selected for this project. Tell us why you chose them. Include clickable links to your working files (hosting your project in GitHub is great for your portfolio but is not required and will not be checked for this assignment). From the document - and from your repo if you make one - provide clickable links to your work in Tableau Public.

The two tools I used for my data visualization were excel for the bones and structure of the project, and Tableau desktop, for the final visualizations, charts, and summaries. Excel is where I input my raw data, it's also where I organized and cleaned my data. Tableau desktop is where I made the interactive charts, dashboard, and the story to make sense of the raw data. I also used Tableau public to share the project if anyone is interested.
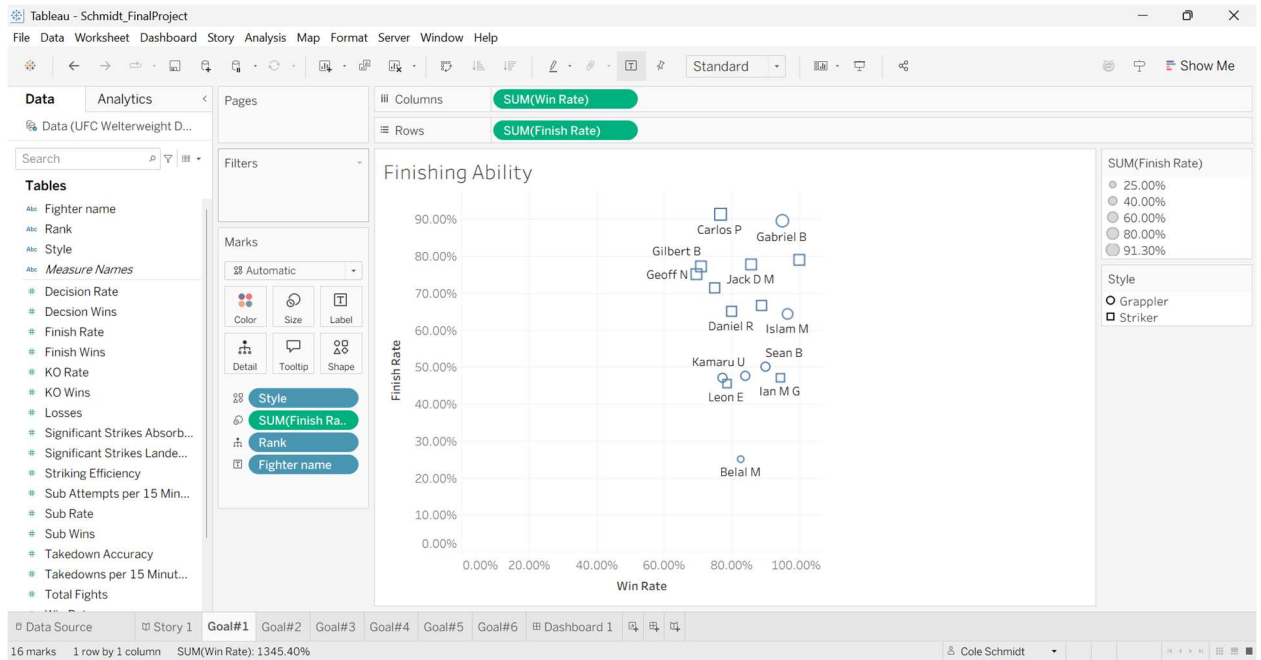
Excel link: Dataset

Tableau desktop link: Project
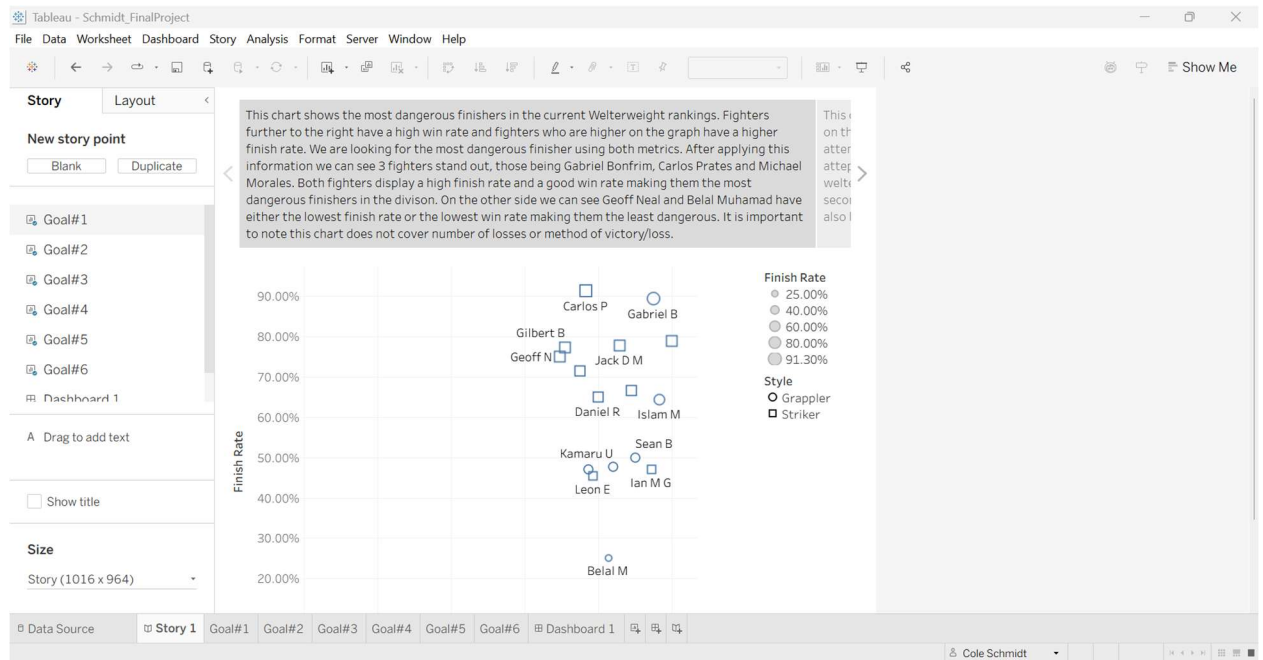
Tableau public link: Public

10. **Section 6. Visualizations and Stories**

11. List the goal: Goal #1: Finishing ability

12. Display the chart(s) for that goal
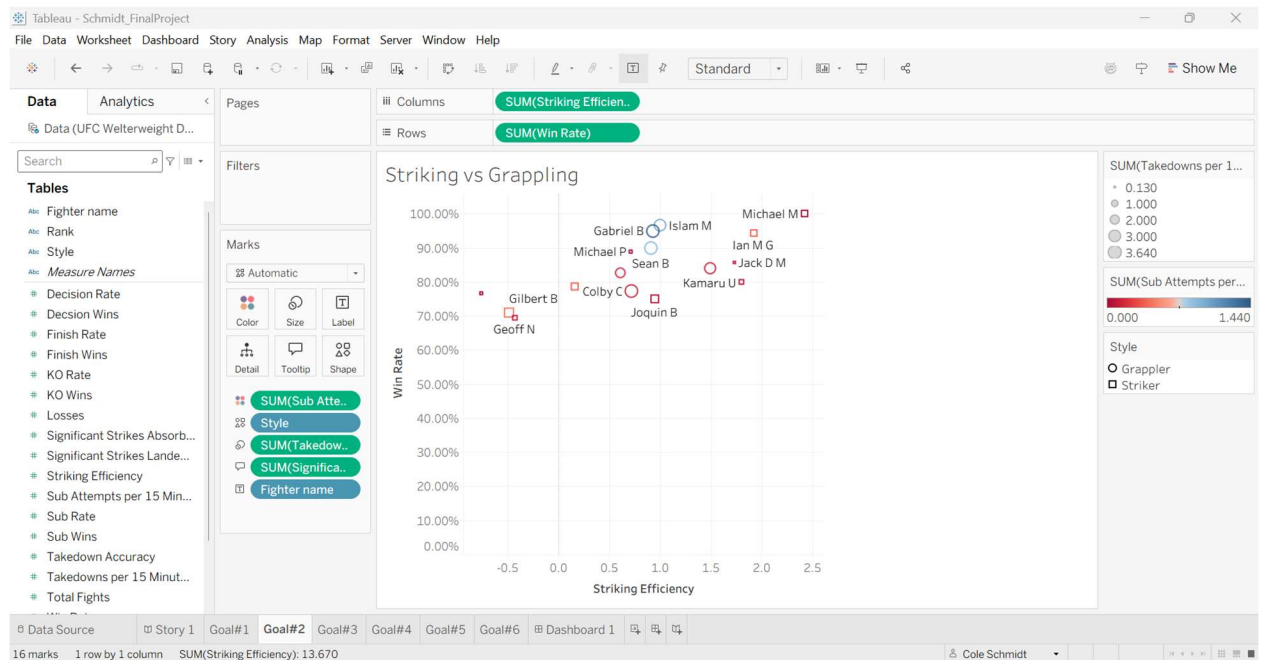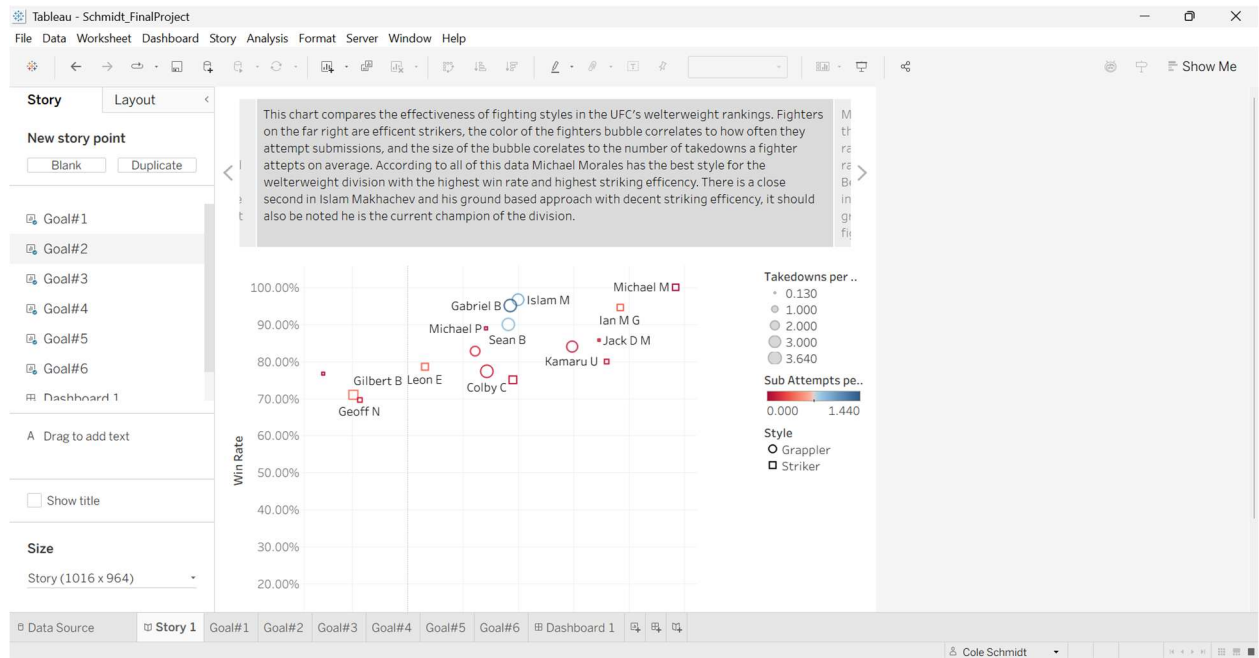
13. Display the story for that goal.

14. List the goal: Goal #2: Striking vs Grappling
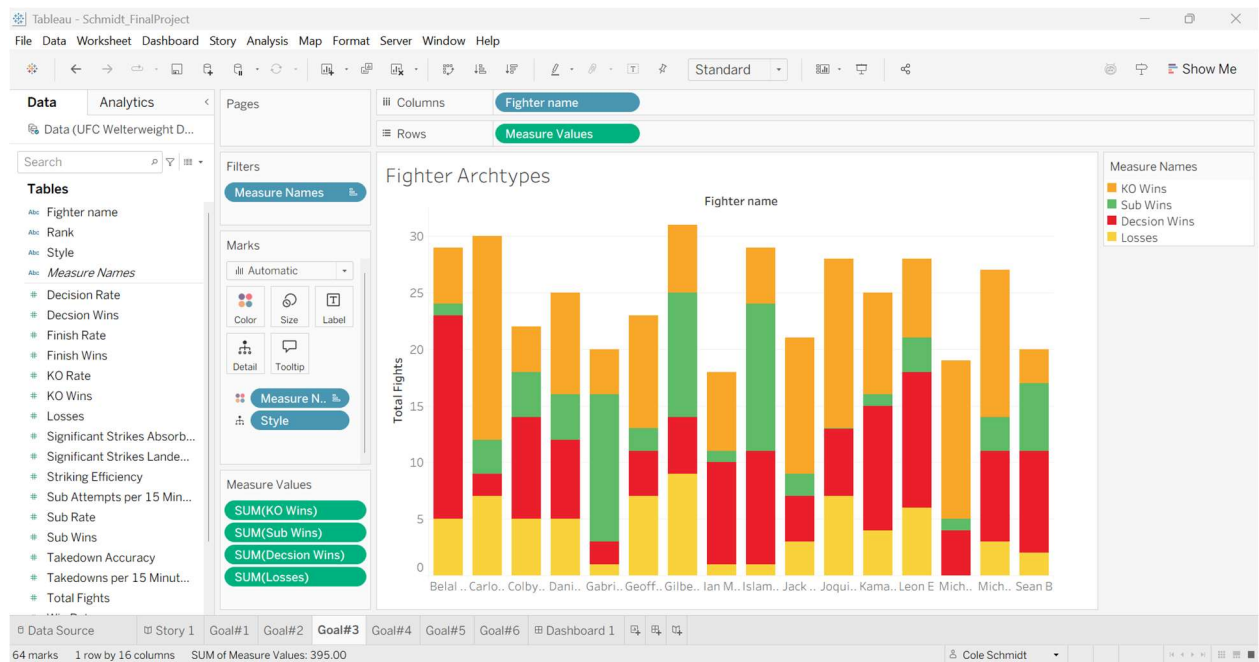
15. Display the chart(s) for that goal
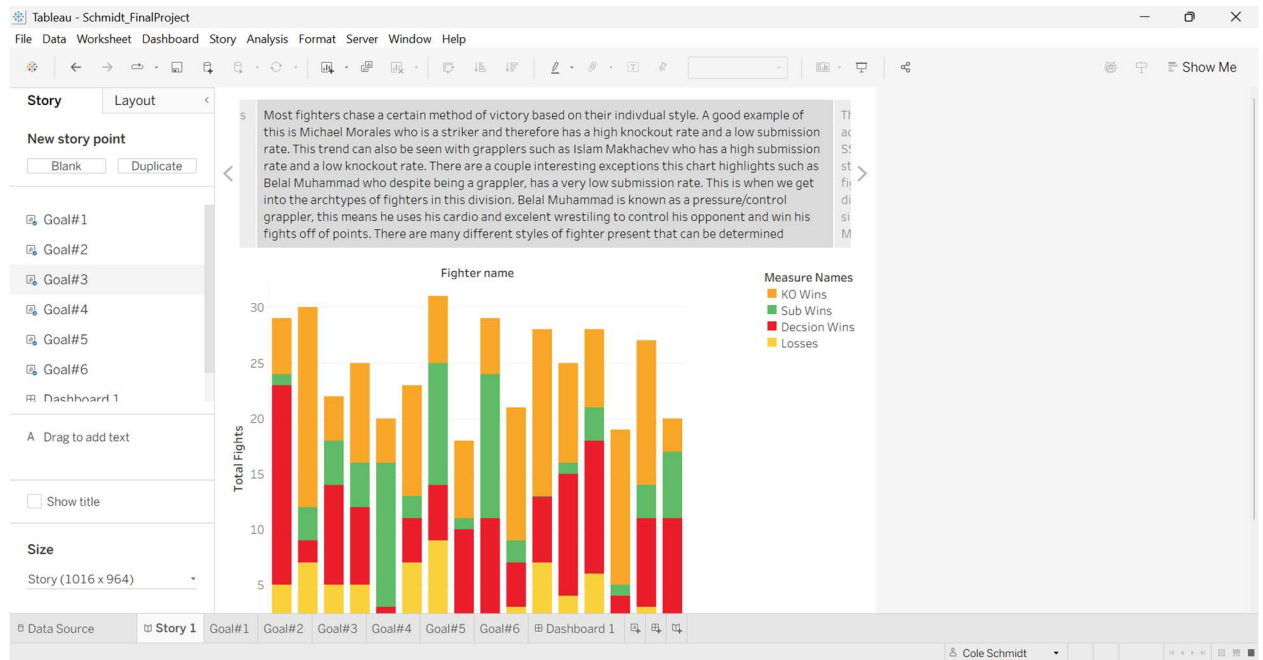
16. Display the story for that
    goal.

17. List the goal: Goal #3: Fighter Archetypes

18. Display the chart(s) for that goal

19. Display the story for that
    goal.

20. List the goal: Goal #4: Efficiency vs Activity
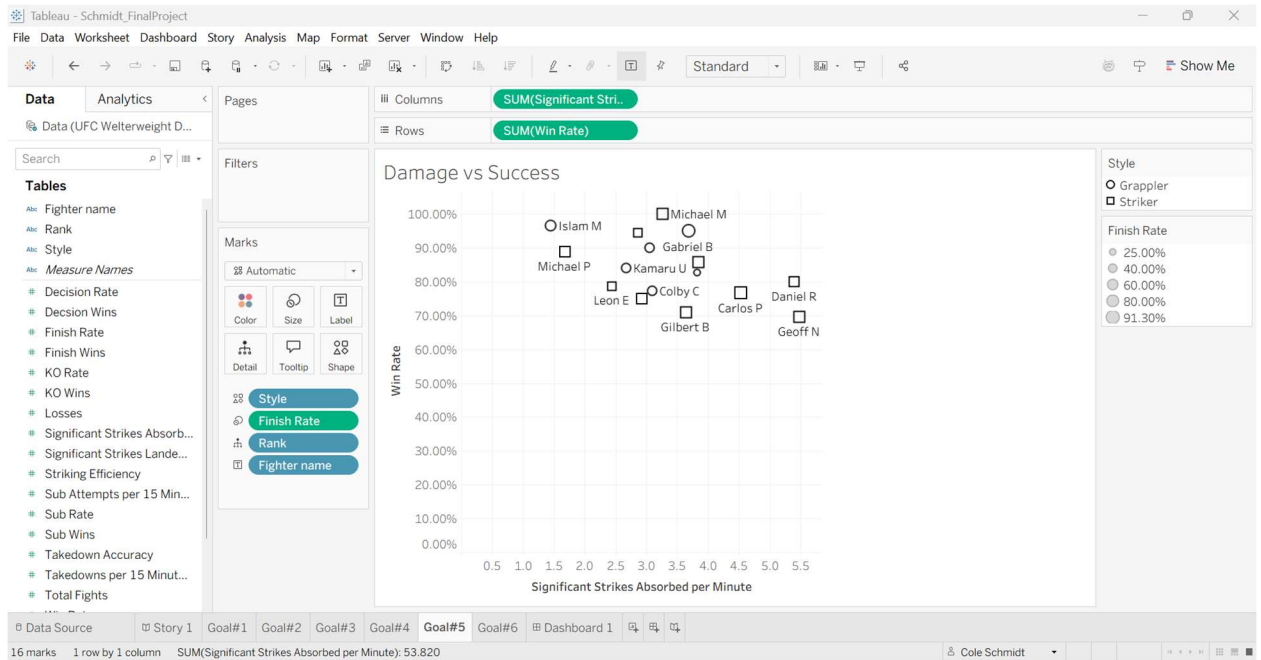
21. Display the chart(s) for that goal
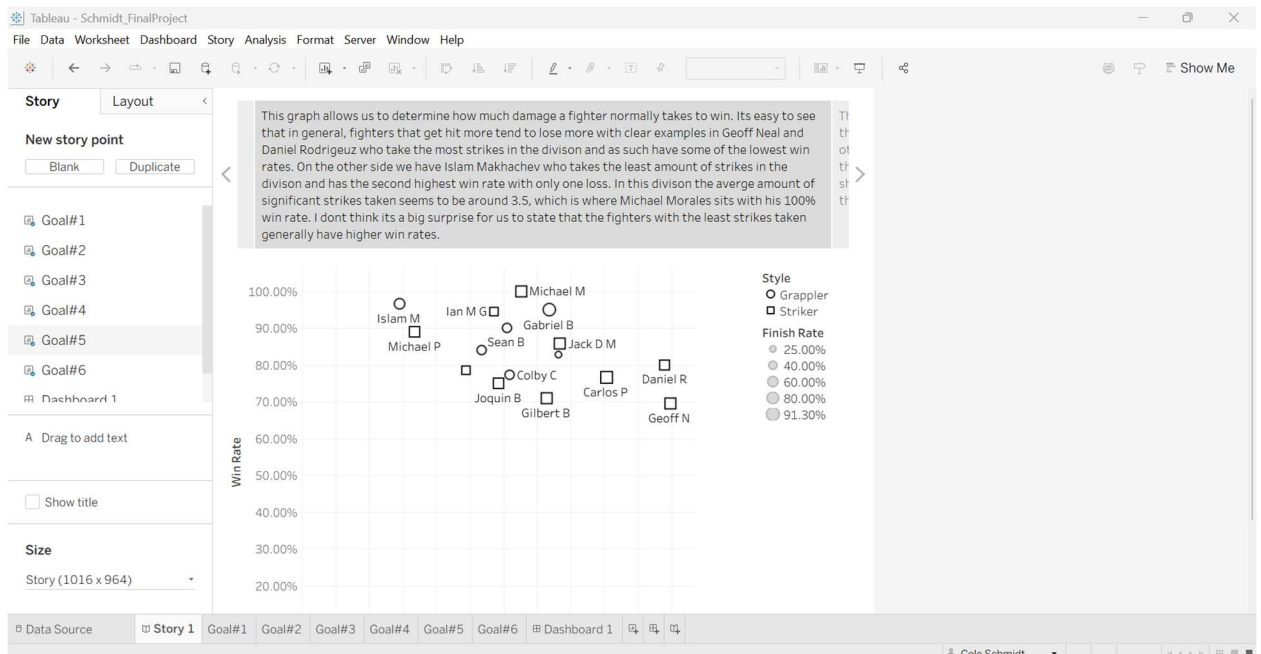
22. Display the story for that goal.
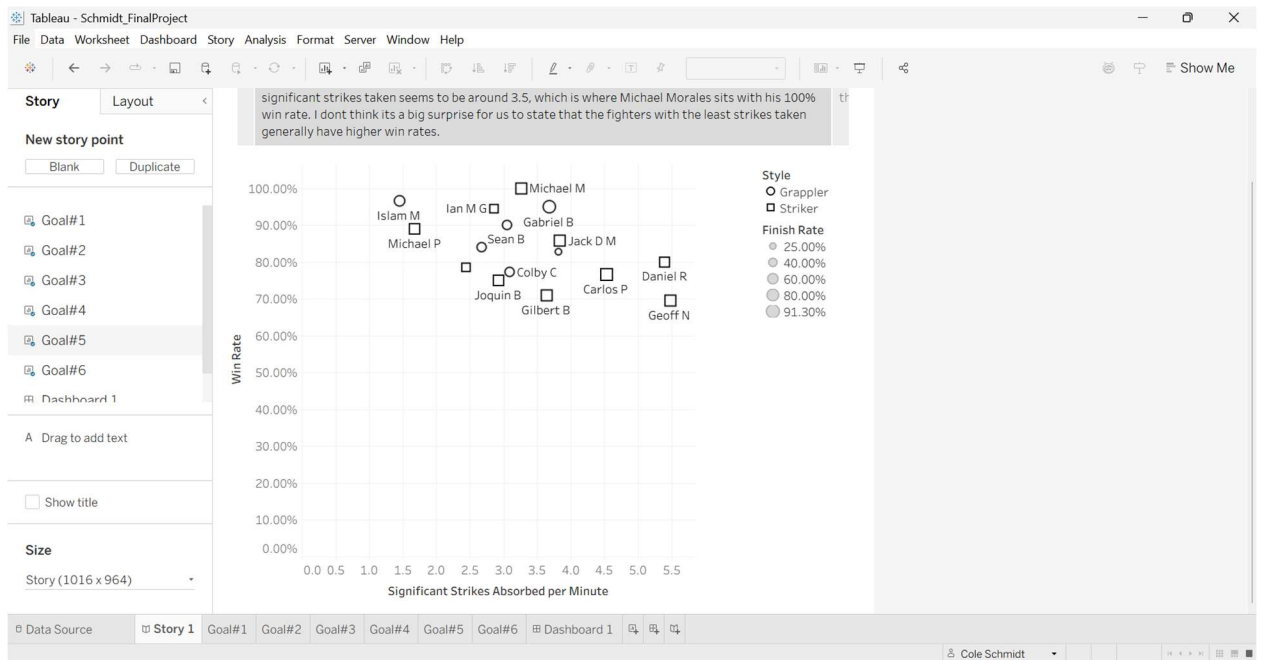

23. List the goal: Goal #5: Damage vs Success

24. Display the chart(s) for that goal

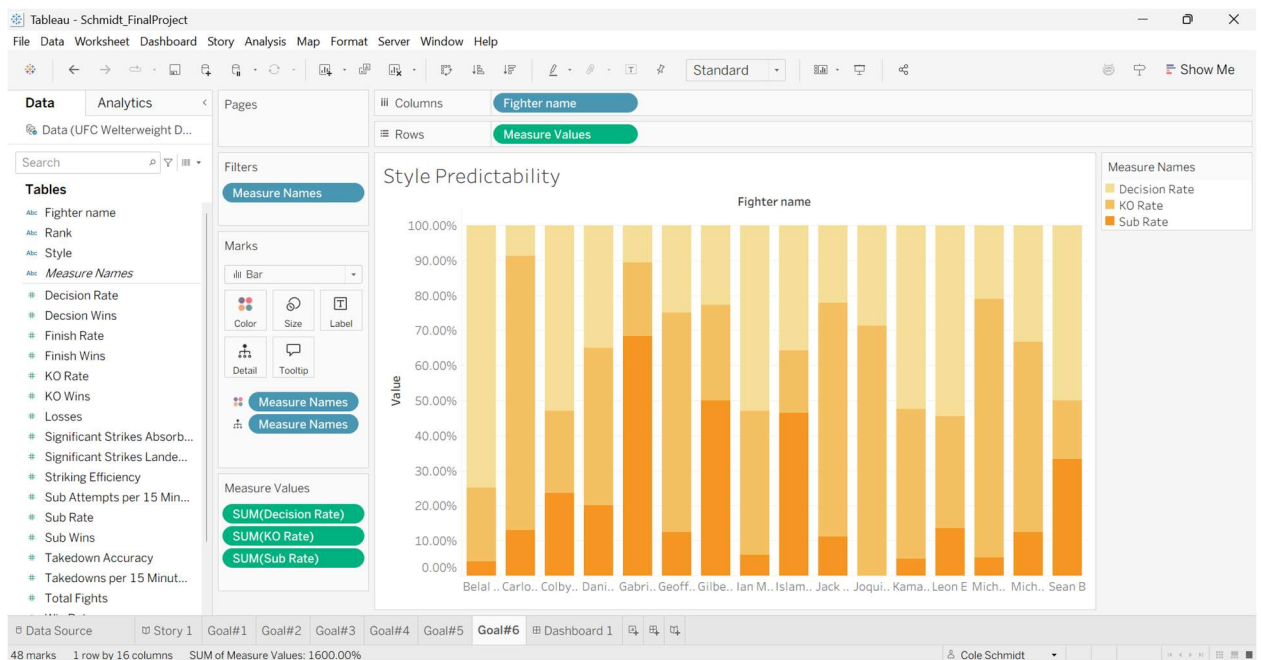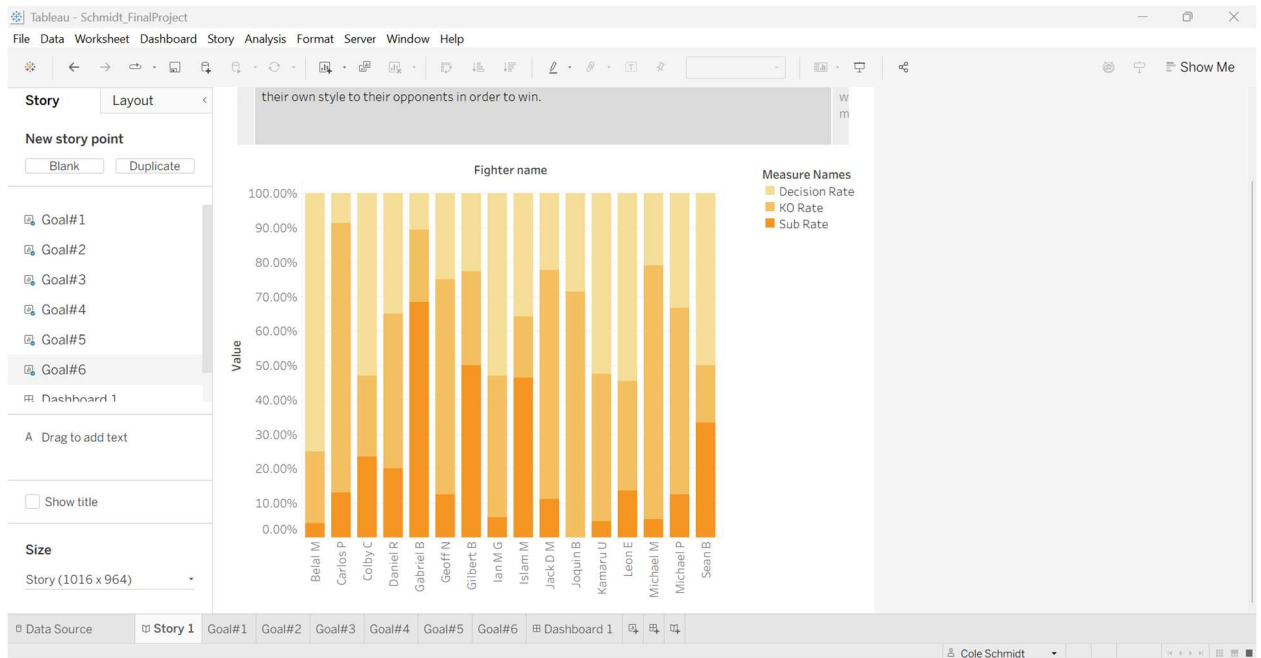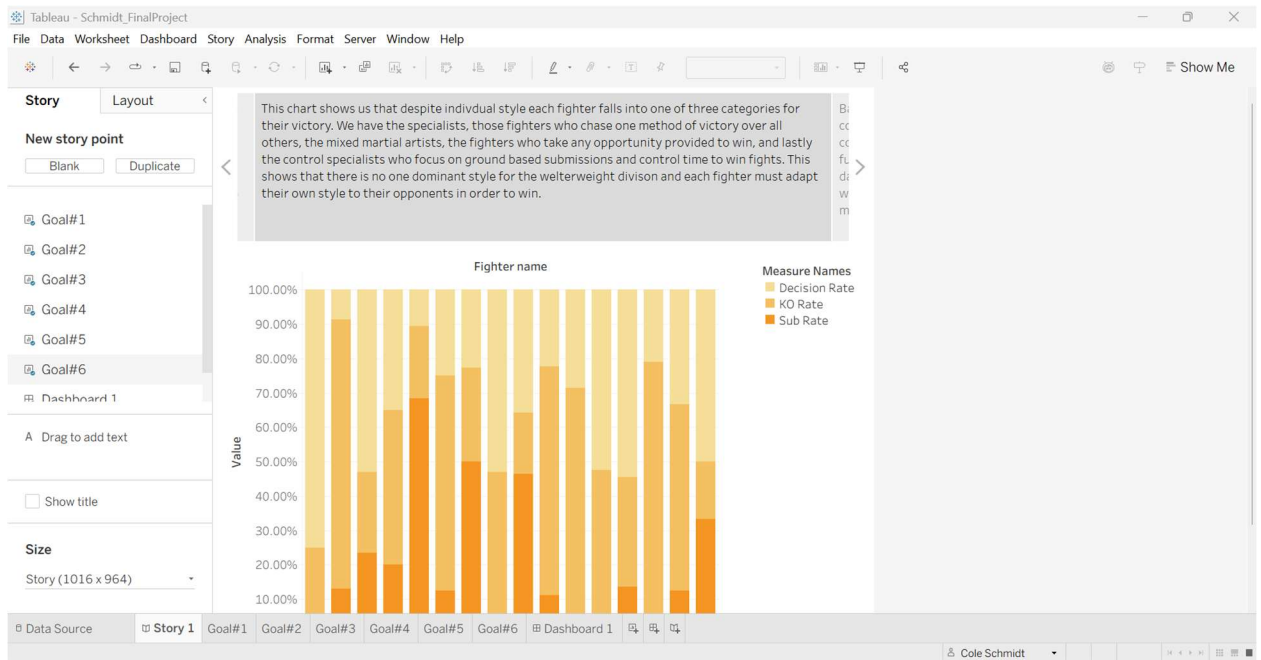## 25. Display the story for that goal.

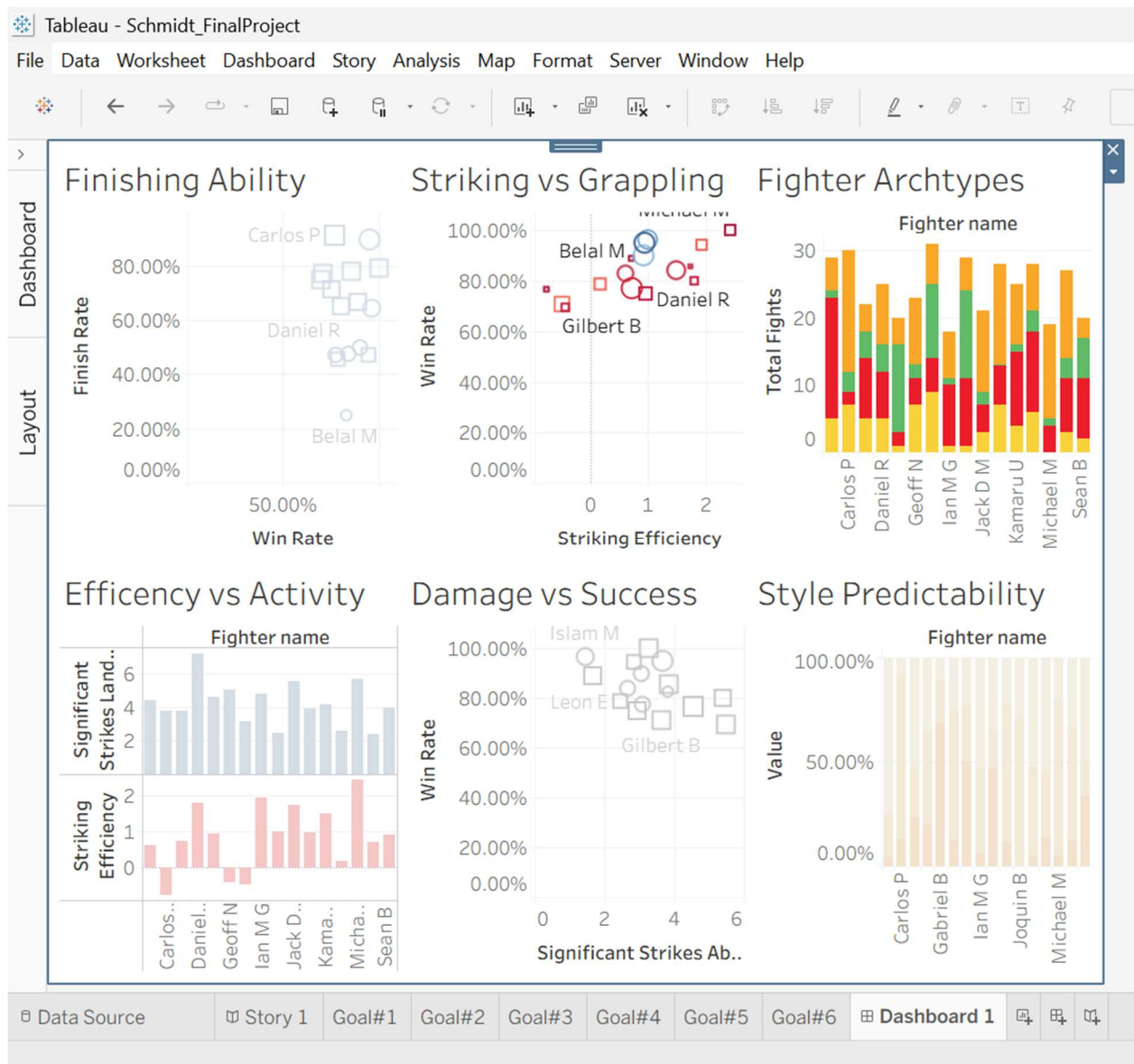## 26. List the goal: Goal #6: Style Predictability

27. Display the chart(s) for that goal

28. Display the story for that goal.

Dashboard:

29. **Section 7. Conclusions**

The top 4 fighters in the UFC's welterweight division are currently Islam M. (champion), Jack D. M. (No 1 contender), Ian M. G. (No 2 contender), Michael M. (No 3 contender). Based off their individual stats and records I believe these 4 fighters are the most likely to hold or capture the belt in the near future. Islam M. has the best striking defense at 1.45 strikes absorbed per minute with a 96.55% win rate, in addition to this he is arguably the most dominant grappler the sport has ever seen; due to this and his previous fights I think its fair to say he is most likely to retain the belt for the foreseeable future due to these stats, especially when we consider the stats and styles of the top 3 contenders.

The No 1 contender, Jack D. M. is without a doubt a striker with 5.57 significant strikes per minute and a 78% finish rate Jack is a big threat on the feet; however Jack has already lost to Islam in a very dominant display of grappling from Islam, so its safe to say Jack will likely not regain his belt anytime soon.

The No 2 contender Ian M. G. may have a better chance as he is more of an all-around fighter who is unpredictable. Ian only absorbs 2.86 significant strikes a minute while dishing out 4.78 a minute making it clear he's a safe fighter who stays active. Ian has 7 KO wins, 9 decision wins, and 1 sub victory which makes it seem like he isn't active on the ground; however, this isn't true, Ian has 0.37 submission attempts per minute which is the third highest in the current welterweight division. Out of the three title contenders we will talk about I believe Ian poses the biggest threat to the champion on the ground as he is active though his raw stats don't reflect that.

The No 3 contender Michael Morales is currently the scariest striker in the welterweight division and the contender I believe has the greatest chance of beating Islam. Michael has an impressive record of 19/0 with 14 KO victories. Michael has the second highest significant strikes per minute in the division at 5.68 and the highest striking efficiency in the division at 2.42. Michael Morales is a very frightening fighter for anyone in the division, and I believe he is the greatest threat to Islam in the division due to his heavy hands and great defense. Also, I did not include fighters' physical characteristics such as height or reach; though he is one of the tallest fighters and has the longest reach in the division; all these factors together make Michael Morales one of the scariest in the division.

Before completing this project, I thought I would find a specific style that is the most effective in the division; however, I didn't find any style or archetype that really dominates the rankings of the division. I believe this may be because welterweight is a moderately sized division where fighters are more diverse but I'll have to do the other weight classes in the future to make a determination.