CPSC 583 Initial Data Analysis

Charlie Cheung 10079808

February 26th, 2016

Introduction

For my CPSC 583 Project 2 data set, I chose the results from the 13th World Wushu Championships held in 2015 in Jakarta, Indonesia. The World Wushu Championship is hosted by the International Wushu Federation (IWUF) and takes place biannually, featuring the highest competitive level within the sport. I chose this data set because I had previously competed in Wushu at the International level and since I am now an instructor in Calgary, I am personally very close to the sport. Also, it gives me a chance to connect the two biggest passions of my life: computer science and martial arts.

Initial Questions to be Answered

The initial questions I had about the data set included:

- How many medals were won by each country?
- What was the average score of each category?
- Was there a significant difference between the scores of males and females who competed in the same form or weapons category?

The first question is something very obvious to consider when gathering data on an international sports competition. I asked the question about the average score of each category because hand forms, short weapon forms, and long weapon forms all have some variation in judging criteria. Also, wushu is a male dominated sport, so the number of men competing in a category would generally be greater than the number of females. I would like to know if the larger number of males would raise or lower the average score or if they were about the same as the female scores, which leads into the third initial question I had for the data set. The judging criteria does not change between the male and female categories in wushu so I was curious to see if there was a difference in the highest and lowest scores of both genders in the same categories.

Organizing the Data

The data was originally in PDF format so first I wanted to change it by organizing it in Excel. I had to copy and paste the athlete ID, country, first name, last name, and score into each

cell for every category and add a category code to each athlete. Wushu is separated into two categories: sanda (which is Chinese kickboxing that incorporates grappling and throws) and taolu (the forms component of wushu). For this project, I decided to focus only on the taolu portion of the 13th WWC. From taolu, wushu is again separated into hand forms, short weapon, and long weapon.

The three most popular hand forms are: Changquan ("Northern"), Nanquan ("Southern"), and Taijiquan ("Tai Chi"). The hand forms then diverge to have their own weapon styles. In Changquan, short weapons include: Daoshu ("Broadsword") and Jianshu ("Sword") and long weapons include: Gunshu ("Staff") and Qiangshu ("Spear"). In Nanquan, the short weapon is Nandao ("Southern Sabre") and the long weapon is Nangun ("Southern Staff"). For Taijiquan, the only weapon is Taijijian ("Tai Chi Sword"). Below is a tree diagram showing the different categories of Wushu.

Athletes specialize in a hand form style based on their body type. Lean, athletic body types are given Changquan, Jianshu, and Qiangshu because they require more athleticism, flexibility, speed, and jumping ability. Nanquan, Nandao, and Nangun are given to athletes built with wide shoulders, shorter height, and stronger torsos which are required to perform "Southern" movements better. Finally, Taijiquan requires athletes to build up their basics from Changquan. Taijiquan requires athletes that are tall. Long legs, long arms, and long fingers are ideal for Taijiquan athletes. It is extremely rare to find athletes that specialize in a mixture of these three hand forms. For example, it is common to find an athlete that does Changquan and Jianshu. It is very uncommon for an athlete to do Taijiquan and Nandao.

I have excluded other forms of taolu because they will not be used in this data set. The boxes in green below are going to be used in my project. There are ten categories multiplied by two for men and women so in total, I have twenty events in my data set.

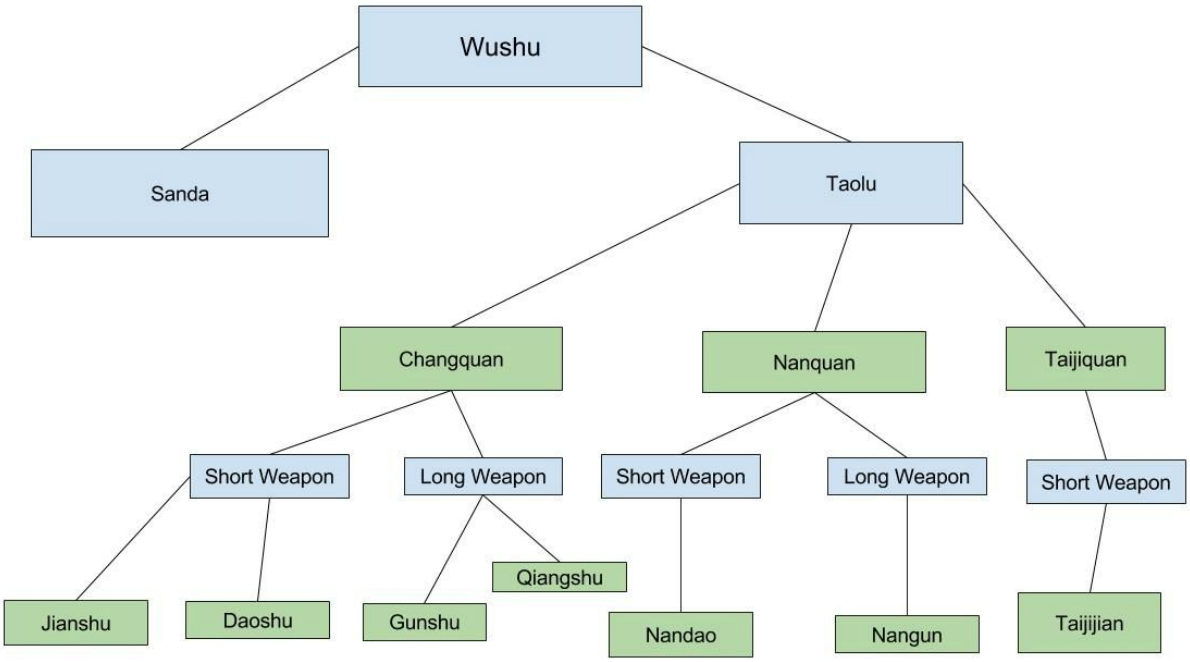

Figure I. Tree diagram of the organization of wushu.

After coming up with a unique category key for each of the events, I organized 515 entries into an Excel spreadsheet. Below is an example taken from the spreadsheet for the Women's Taijiquan category.

Category	Place	AthleteID	Country	First Name	Last Name	Score
WTJQ	1	1227	INA	Lindswell	LINDSWELL	9.7
WTJQ	2	1262	HKG	Suijin	CHEN	9.68
WTJQ	3	1105	MAX	Shin	YII NG	9.66
WTJQ	4	1235	VIE	Khanh Ly	TRAN THI	9.55
WTJQ	5	1142	JPN	Shiho	SAITO	9.39
WTJQ	6	1218	IND	Sanatombi	CHANU LEIMAPOKPAM	9.32
WTJQ	7	1141	JPN	Naoko	ICHIKIZAKI	9.09
WTJQ	8	1084	CAN	Wei Jen	LEE	8.35
WTJQ	Default	1126	NEP	Sabita	RAI	0

Figure II. Screenshot of the data organized in Microsoft Excel.

All scores are out of 10 for all categories. From above, there is an athlete who has not been given a place number but instead, have "Default" next to their names. This is given to athletes who were supposed to compete in the category but did not end up doing so. The

athletes are then given a score of 0 because I did not want to leave the cell blank. The athlete with the default rank from the above table is highlighted below.

Category	Place	AthleteID	Country	First Name	Last Name	Score
WTJQ	1	1227	INA	Lindswell	LINDSWELL	9.7
WTJQ	2	1262	HKG	Suijin	CHEN	9.68
WTJQ	3	1105	MAX	Shin	YII NG	9.66
WTJQ	4	1235	VIE	Khanh Ly	TRAN THI	9.55
WTJQ	5	1142	JPN	Shiho	SAITO	9.39
WTJQ	6	1218	IND	Sanatombi	CHANU LEIMAPOKPAM	9.32
WTJQ	7	1141	JPN	Naoko	ICHIKIZAKI	9.09
WTJQ	8	1084	CAN	Wei Jen	LEE	8.35
WTJQ	Default	1126	NEP	Sabita	RAI	0

Figure III. Excel spreadsheet screenshot highlighting an athlete that placed with a "Default".

However, one athlete in the Men's Daoshu category competed in the category but still received a score of 0. This was probably due to a disqualification most likely caused by the weapon breaking in the middle of competing. The athlete with both a number ranking and an official score of 0 is highlighted below.

MDS	33	1037	GER	Mathieu	SYCHLA	8.35
MDS	34	1092	LIB	Avo Mrad	MANDJIAN	7.95
MDS	35	1147	SRI	Hasika Chameera Bandara	DISANAYAKA MUDIYANSEL	7.22
MDS	36	1001	ALG	Boudaoud	YACINE	6.92
MDS	37	1036	GER	David	TOEROEK	0
MDS	Default	1171	UKR	Roman	REVA	0
MDS	Default	1128	NGR	Friday	NNADI	0
MDS	Default	1245	CHN	Pei Yuan	SUN	0
MDS	Default	1089	GAB	Nzengue Boukanguo	GLENN-RIENK	0
MDS	Default	1240	CHI	Ariel	MANCILLA BARRIENTOS	0

Figure IV. Excel spreadsheet screenshot highlighting an athlete that received a score of O.

The next thing I did was put my data into Tableau. While I was copying the data from the PDF it occurred to me how many more male competitors there were in the competition than females. I created the bar chart shown on the following page to show the total number of athletes in each category.

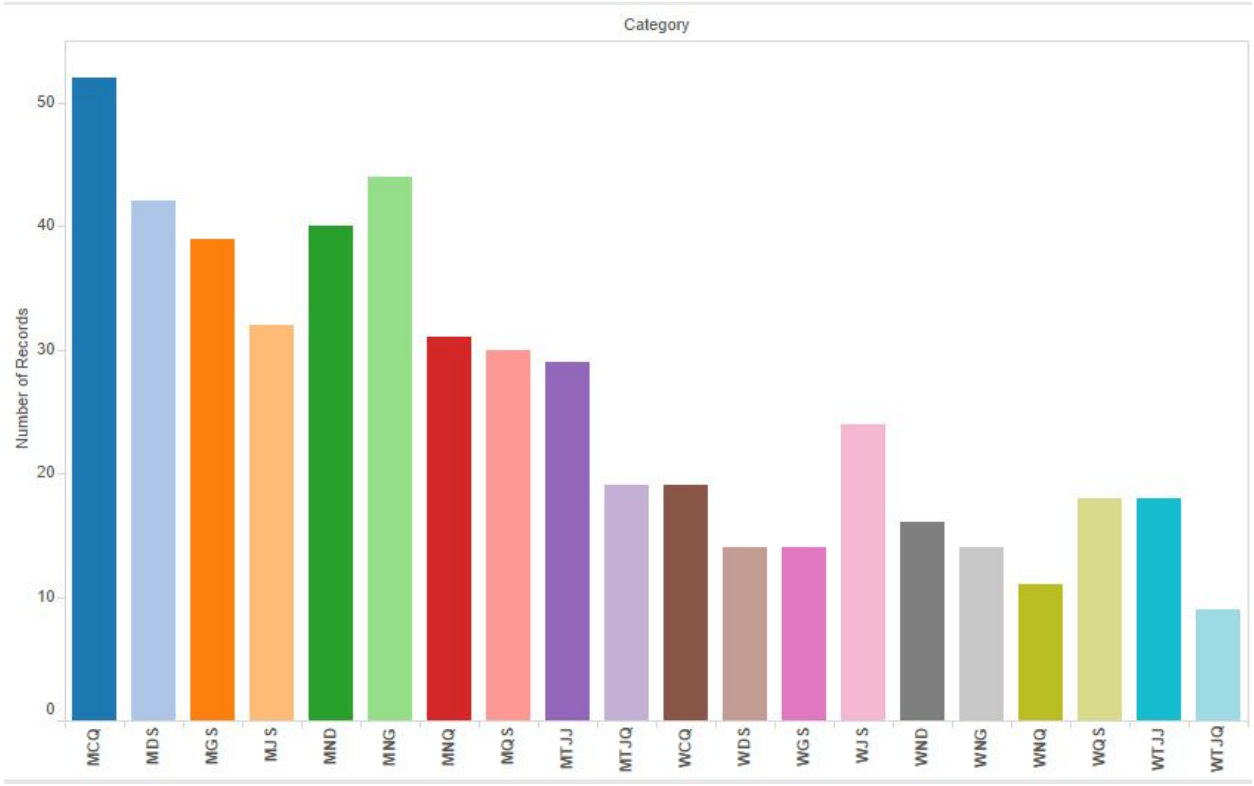

Figure V. Bar chart showing the total number of competitors in each event.

The bar chart shows that each male category has more participants than their female category counterpart. I sorted the data again to show the descending order of the sum of records. Then, I highlighted where a women's category had a higher number of participants than a male category and the women's category that had the same number of participants as the men's category. The result is the bar chart on the following page.

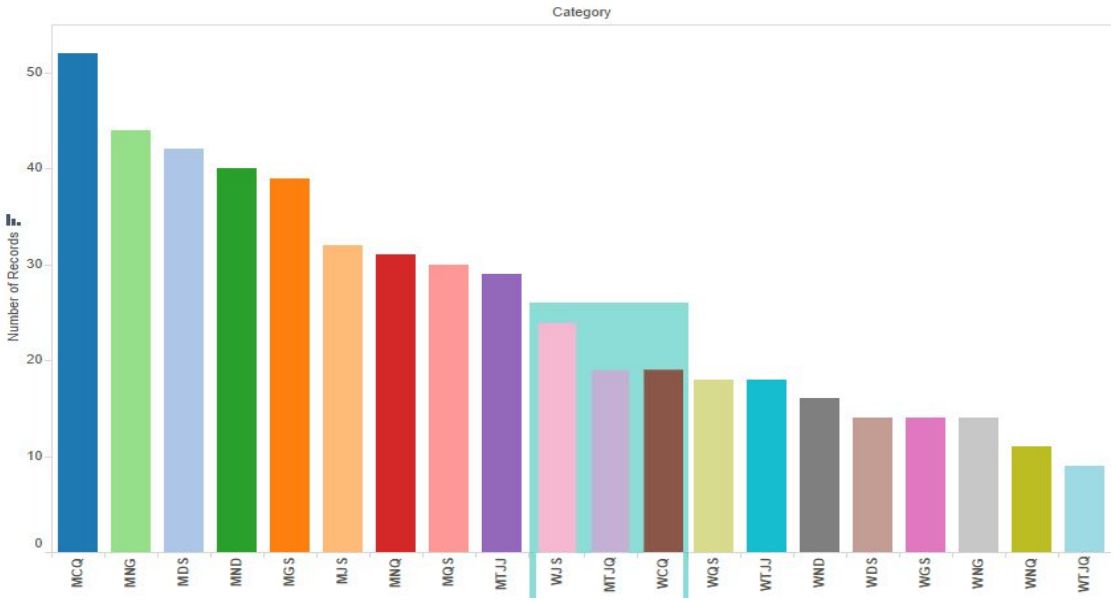

Figure VI. Bar chart showing the number of athletes competing in each event ordered from largest to smallest. WJS, MTJQ, and WCQ highlighted.

After filtering out all the default values (athletes that did not end up competing) from the chart, some of the categories did end up changing places. One of the more noticeable changes was the difference between the largest participant category and the runner up (men's changquan and men's nandao). Also, even though in the above bar charts, men's nangun had the second highest number of participants, after removing the default values, it dropped down to become the fourth largest category. The resulting bar chart is shown on the following page.

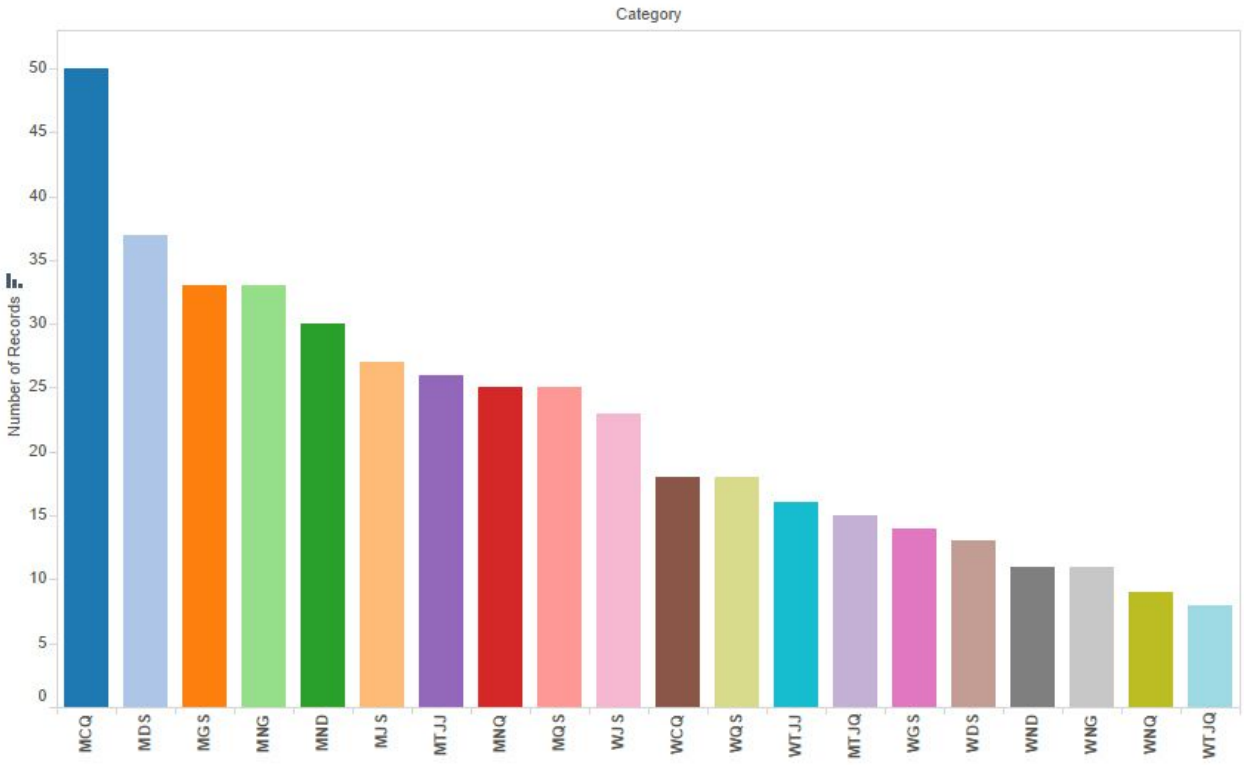

Figure VII. Bar chart showing the number of athletes competing in each event ordered from largest to smallest. "Defaults" have been removed.

The next thing I wanted to do was look at the number of athletes sent by each country. Each athlete is allowed to compete in a maximum of four events and a team can only consist of a total of 8 athletes. First, I wanted to see which country registered their athletes in the most events. To do this, I created a tree graph of the sum of records and the countries they were from. The tree graph is shown on the following page.

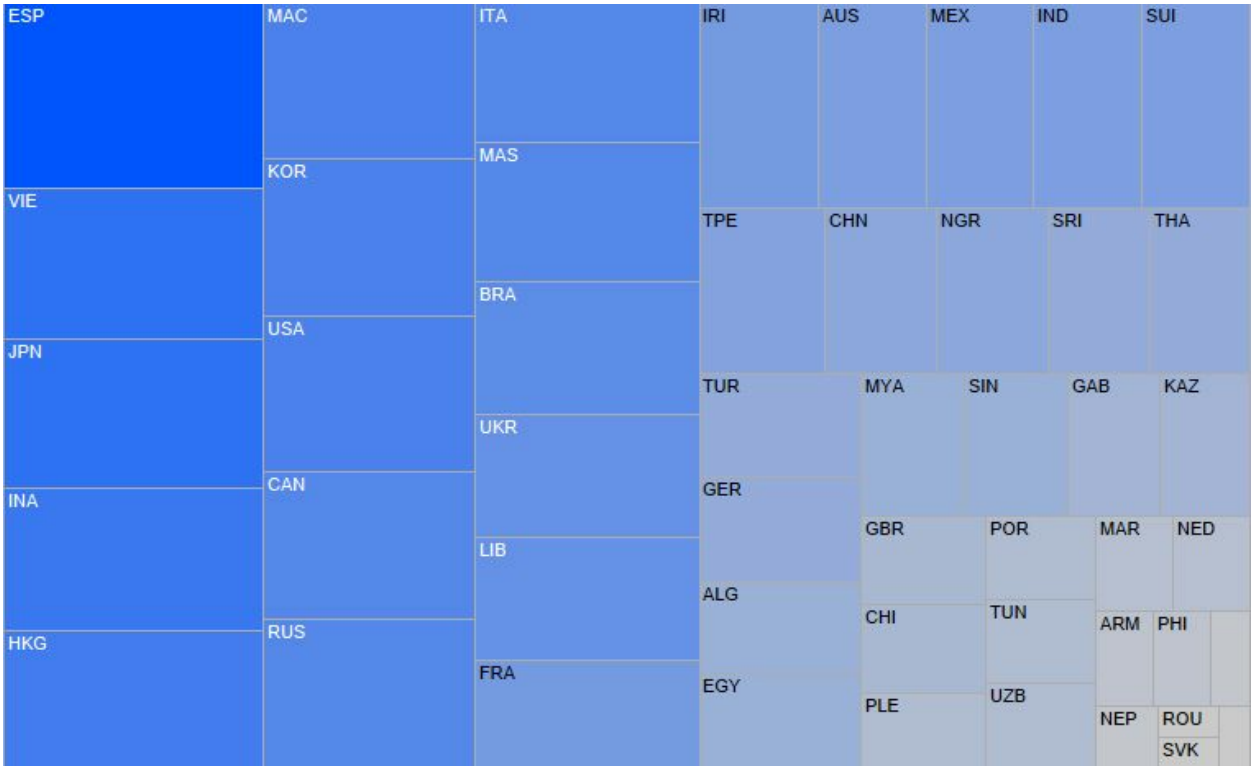

Figure VIII. Tree graph of countries and how many events they registered their athletes in.

From the tree graph above, Spain seems to have competed in the most number of categories at a total of 26. The host country, Indonesia competed in 20 categories and China in 10. At the bottom right are countries that only sent one athlete to represent them which includes: Hungary, Slovakia, and Romania. Next, I filtered out all the athletes that did not actually compete in their registered category. The new tree graph is shown on the following page.

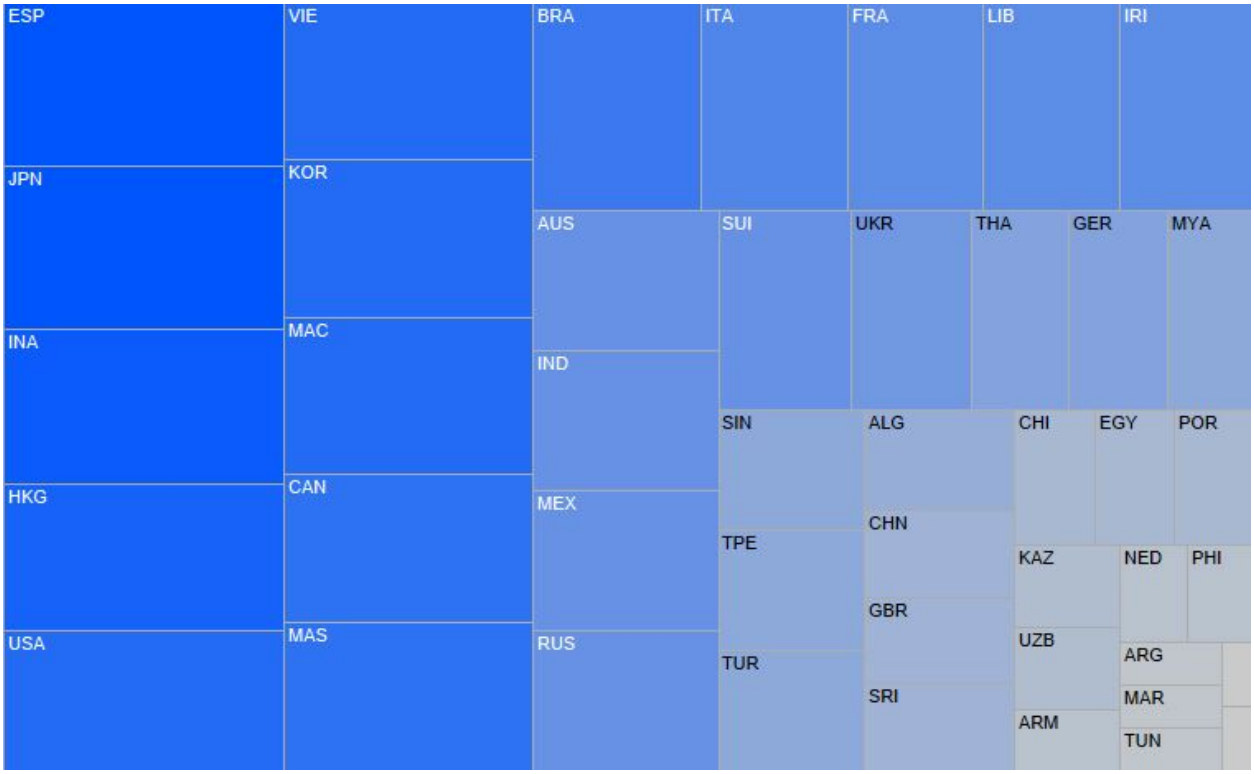

Figure IX. Tree graph of countries and how many events they registered their athletes in. "Defaults" have been removed.

The tree graph became more evenly distributed with Spain's count dropping to 21. This put them at the same number as Japan and only one above Indonesia. China had also dropped to 6 meaning that 40% of their athletes did not compete in their registered event.

Question One: How many medals were won by each country?

To start answering my initial questions, I needed to reorganize my data to show how many medals were won by each country. On the following page is a partial chart that shows the countries and the number of times they placed in the top ten. The colours represent different categories in the competition.

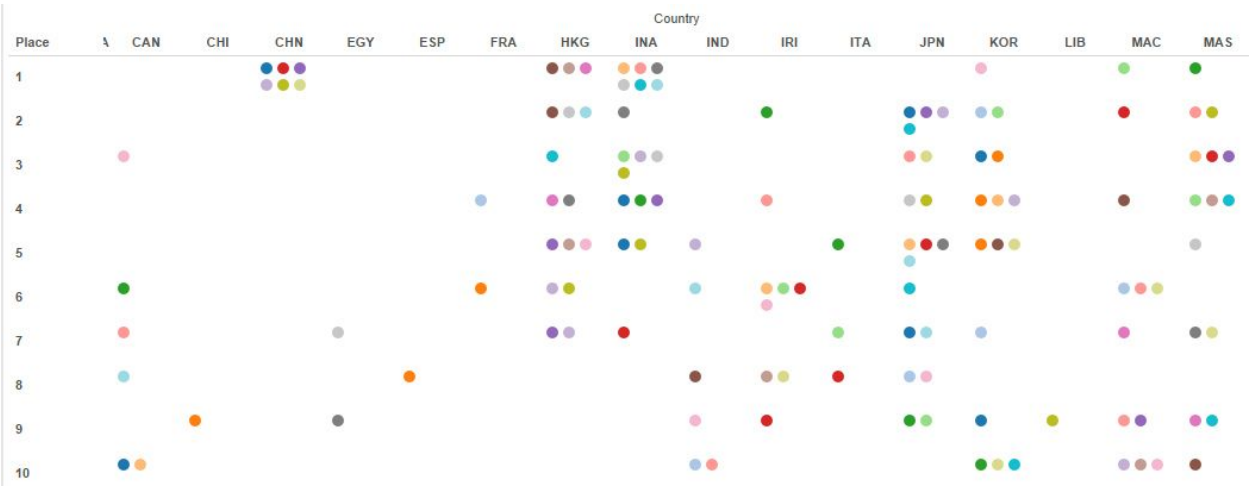

Figure X. Partial table showing countries that have placed in the top ten of different categories.

From the chart above, it is interesting to note that for all categories competed in by team China, they ended up winning first place. Other countries that seemed to dominate the competition based on medal count include Hong Kong, Indonesia, Korea, and Japan. After recognizing this, another question was raised for the data set and that was:

Does a country's geographic location have any influence on the calibre of wushu athletes?

To answer this question, I sorted my data to show in descending order, the countries with the most medals awarded. I also filtered my data to show only the top eight in each category. I did this because it is only the top eight that are awarded with certificates in the competition. Finally, I filtered the countries to show the top twelve countries with the highest medal count. The final chart to answer my initial question is shown on the following page.

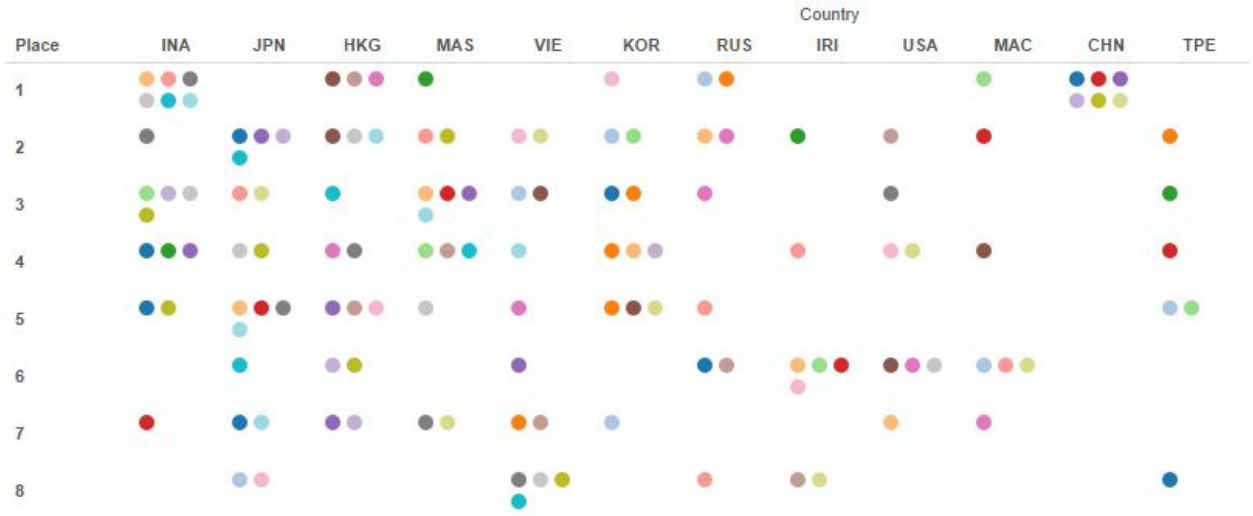

Figure XI. Partial table showing countries that have placed in the top eight of different categories in descending order.

From the above chart, Indonesia, the host country, seemed to have acquired the most medals in the entire competition. They are followed by Japan, Hong Kong, Malaysia, Vietnam, Korea, Russia, Iran, United States, Macau, China, and Taiwan (Chinese Taipei). There seems to be an obvious pattern regarding the geographic location of the top twelve countries with almost all of them located relatively close to China, the origin of Wushu. This final chart answers my initial question of which countries earned the most medals by showing the top twelve countries with the highest medal count. The countries are organized in a table that shows what place the country received and the colours represent the event in which the country placed. When highlighted in Tableau, the specific category is also given to the viewer.

Question Two. What was the average score of each category?

The next question I had asked of the average score of each category. I started by simply taking each category and finding the average. The table below shows the averages of all categories in the competition. The darker the colour of the background, the higher the average of the scores.

Category	
MCQ	8.438
MDS	7.758
MGS	7.739
MJS	7.599
MND	6.833
MNG	6.831
MNQ	7.306
MQS	7.543
MTJJ	8.077
MTJQ	7.376
WCQ	8.658
WDS	8.582
WGS	9.118
WJS	8.752
WND	6.133
WNG	7.011
WNQ	7.555
WQS	9.036
WTJJ	7.929
WTJQ	8.304

Figure XII. Table showing the average score of each category.

Since the averages here still included all of the Defaults that acquired a score of zero, I had to include the "Place" of each category and filter out the "Defaults". Below is the filtered table that calculates the average of only the athletes that actually competed.

Category	
MCQ	8.7758
MDS	8.8068
MGS	9.1467
MJS	9.0063
MND	9.1110
MNG	9.1085
MNQ	9.0592
MQS	9.0516
MTJJ	9.0088
MTJQ	9.3427
WCQ	9.1389
WDS	9.2423
WGS	9.1179
WJS	9.1326
WND	8.9200
WNG	8.9236
WNQ	9.2333
WQS	9.0361
WTJJ	8.9206
WTJQ	9.3425

Figure XII. Table showing the average score of each category excluding "Defaults".

The averages for all categories have increased because of the removal of "Defaults". I believe this visualization is the best for simply answering my initial question. However, having only a chart with numbers and colours does not aid in comparing the averages between different categories. So a new question I wanted to ask was:

How do the averages between categories compare with one another?

In order to answer this more specific question, I created a scatterplot in Tableau. After seeing how the the values range so close to each other, I zoomed in on my y-axis to show the values between 8.6 and 9.5. The resulting chart is depicted below.

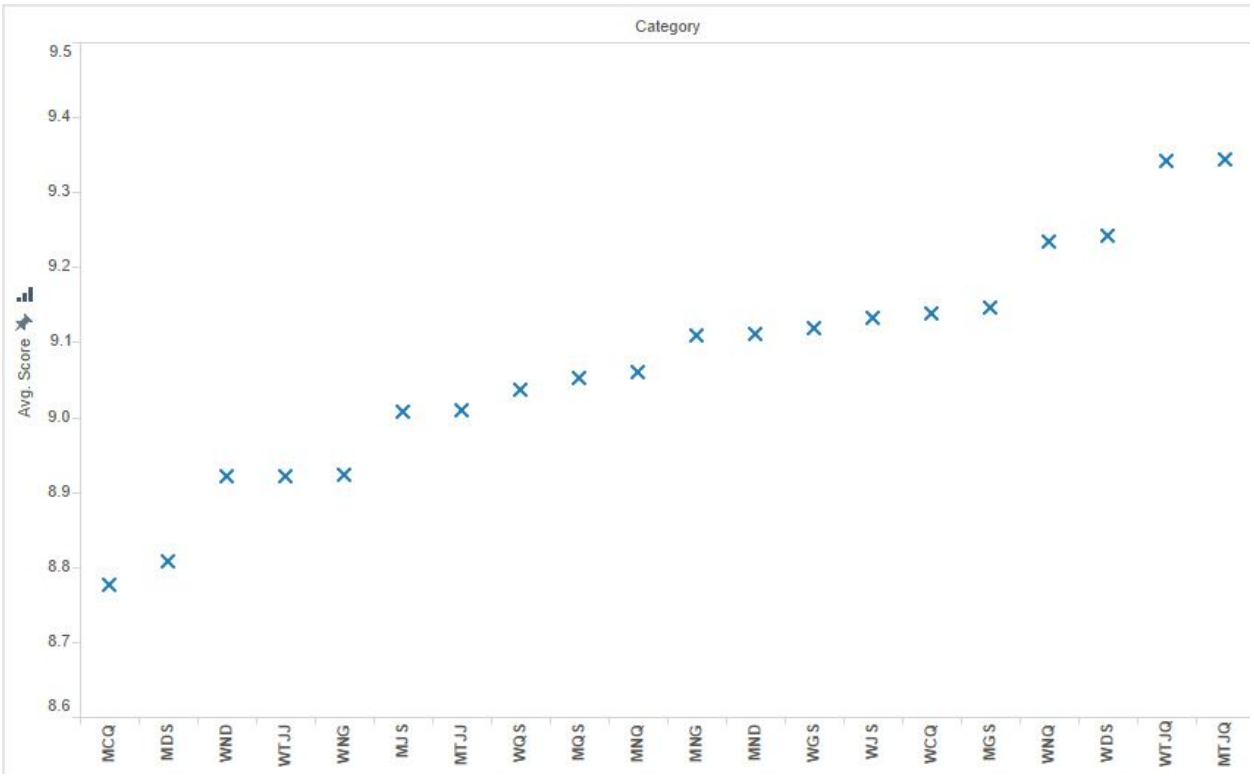

Figure XIII. Scatterplot of the average scores in different categories in ascending order.

From this chart it is easier to compare the averages of different categories with each other. There was also over half a point difference between the events with the highest averages and the lowest average. Naturally, a question that came to my mind was:

Why was there over a 0.5 point difference between Men's & Women's Taijiquan averages and Men's Changquan average?

The two highest averages come from Women's Taijiquan and Men's Taijiquan which differ by 0.0002 from each other. I decided to isolate both these categories to see why the average score was so high.

For men's Taijiquan I removed the "Default" athletes and showed each athlete's score as a coloured circle. The athletes are identified by their athlete ID. The scores range from 7.5 to 10. The chart is shown below.

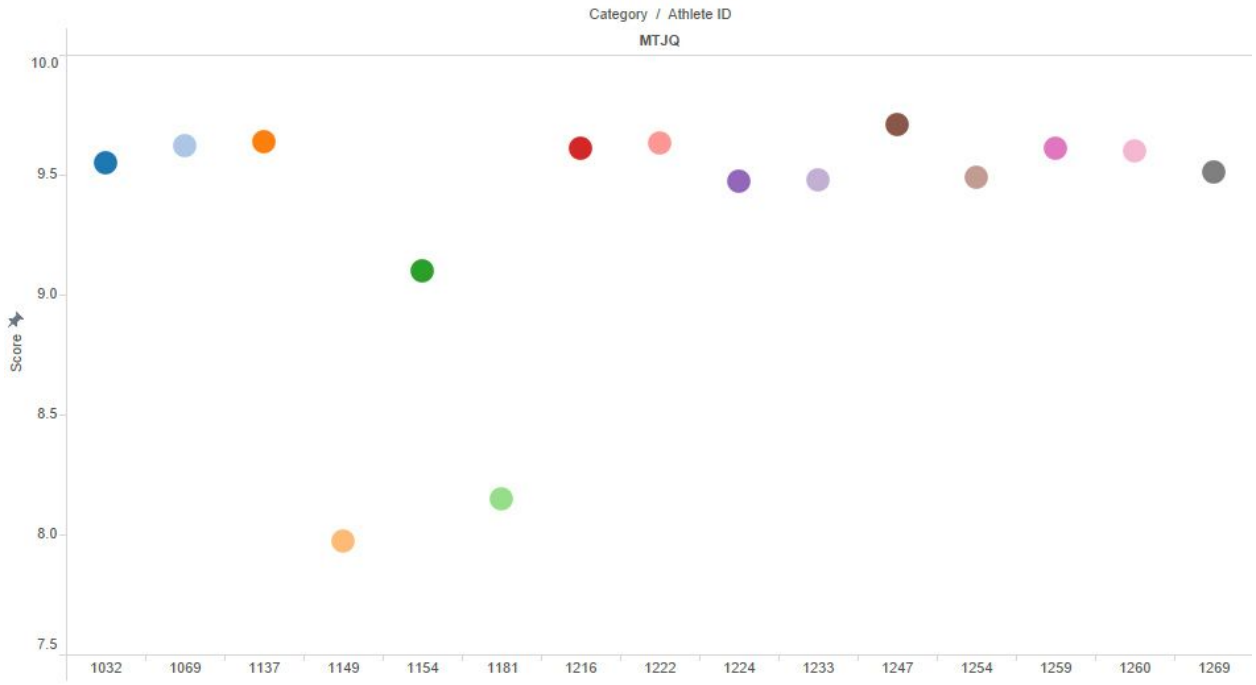

Figure XIII. Scores of Men's Taijiquan.

The highest score here is shown to belong to China's Chen Zhou Li at 9.71 and the lowest is Sri Lanka's Melina Prema Kumara at 7.97. The men's Taijiquan category is also the smallest category in any men's events in the entire competition. This could explain the high average of the category.

For Women's Taijiquan, I applied the same dimensions, measures, and filters with the women's event data. The result is the chart on the following page.

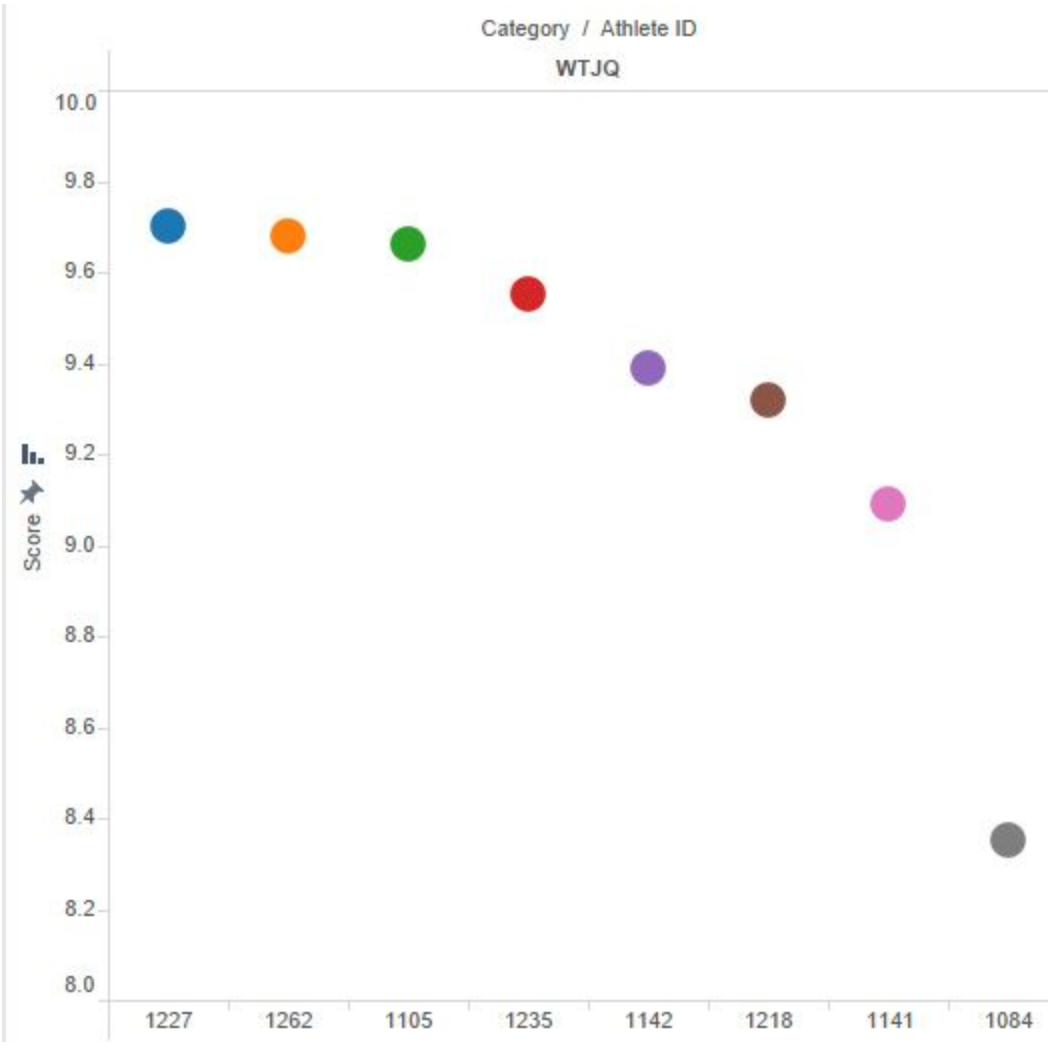

Figure XIV. Scores of Women's Taijiquan.

The highest score here is shown to belong to Indonesia's Lindswell at 9.70 and the lowest is Canada's Wei Jen Lee at 8.35. The women's Taijiquan category is one of the smallest categories in the entire competition with only eight competitors in total. This could also explain the high average of the category.

The event with the lowest average is Men's Changquan. This is also the largest category in the entire competition. Using the same dimensions and measures from Figure XIII and Figure XIV, the chart below shows the scores of all athletes who competed in Men's Changquan. Again, the y-axis ranges from 7.5 to 10.

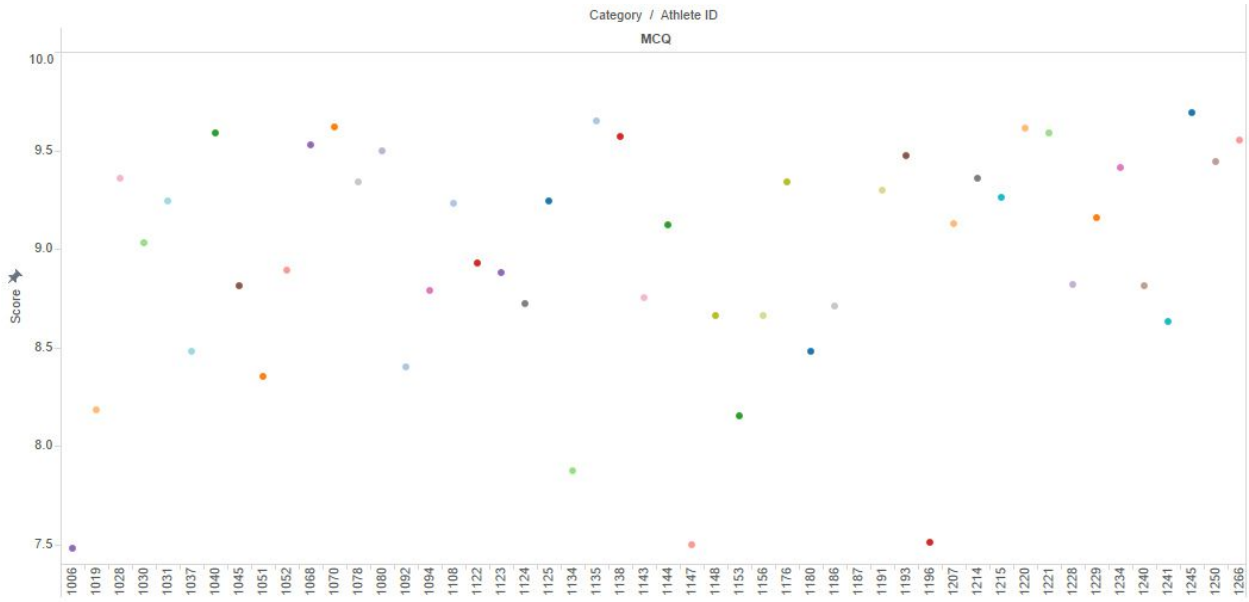

Figure XV. Men's Changquan scores by athlete ID.

The above chart shows the scores of 50 athletes that range from 7.48 to 9.69. The large number of athletes and the range between their scores is what gives Men's Changquan the lowest average score in the entire competition. Since the events with the highest averages also have the lowest number of athletes and the event with the lowest average has the highest number of athletes, it can be assumed that the larger the number of athletes competing in an event, the lower the average score will be. This leads into my final initial question which asks whether there is a significant difference between the scores of males and females who competed in the same form or weapons category.

Question Three: Is there a significant difference between the scores of males and females who competed in the same forms or weapons categories?

To answer this question I started with the Changquan category first, filtering out all of the "Defaults" and excluding all of the other categories. Then, I took a measure of all the athletes that remained in each category and compared the number of athletes that competed in Men's Changquan and Women's Changquan in horizontal bars.

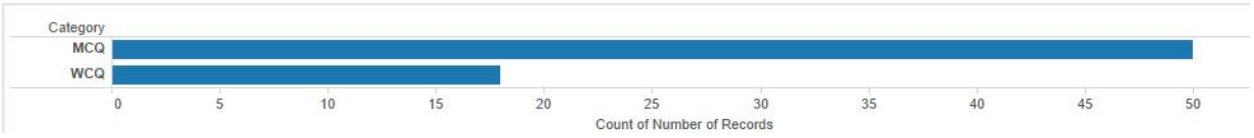

Figure XVI. Comparison of athletes competing in Men's Changquan and Women's Changquan.

I proceeded to do this with the 9 other categories in the data set. The results are shown below.

Nanquan

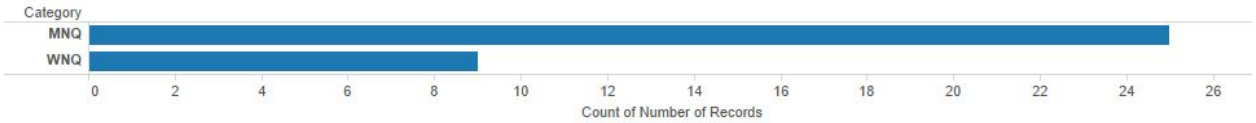

Figure XVII. Comparison of athletes competing in Men's Nanquan and Women's Nanquan.

Taijiquan

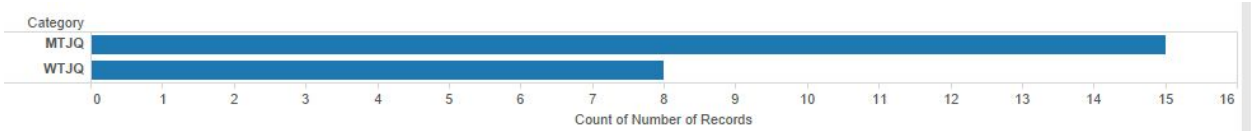

Figure XVIII. Comparison of athletes competing in Men's Taijiquan and Women's Taijiquan.

Jianshu

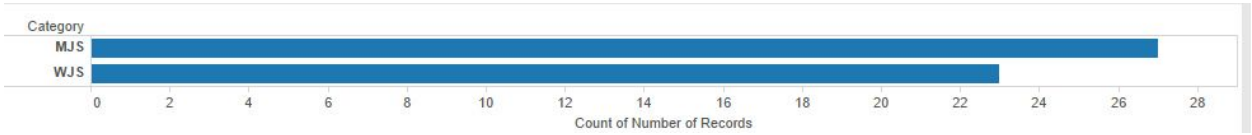

Figure XIX. Comparison of athletes competing in Men's Jianshu and Women's Jianshu.

Daoshu

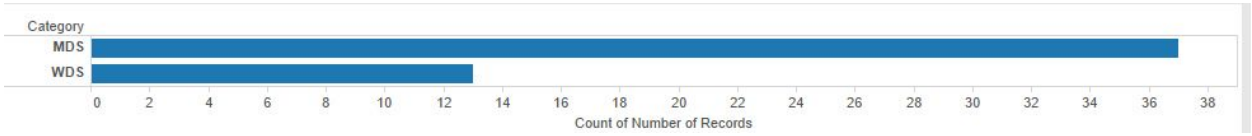

Figure XX. Comparison of athletes competing in Men's Daoshu and Women's Daoshu.

Gunshu

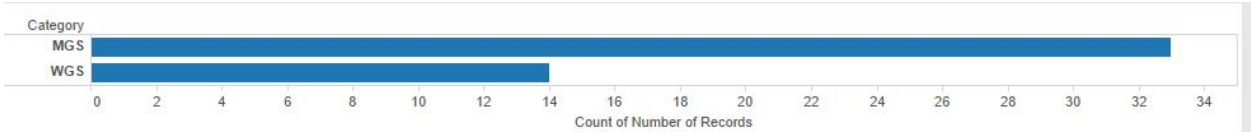

Figure XXI. Comparison of athletes competing in Men's Gunshu and Women's Gunshu.

Qiangshu

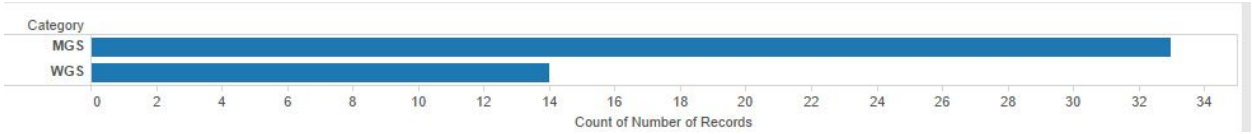

Figure XXII. Comparison of athletes competing in Men's Qiangshu and Women's Qiangshu.

Nandao

Figure XXIII. Comparison of athletes competing in Men's Nandao and Women's Nandao.

Nangun

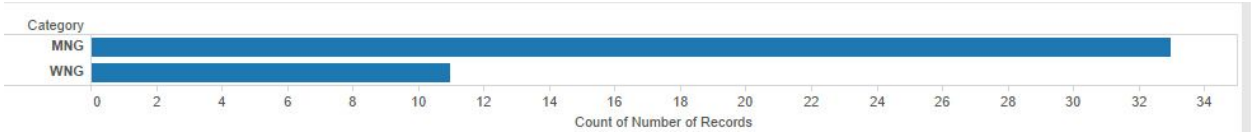

Figure XXIV. Comparison of athletes competing in Men's Nangun and Women's Nangun.

Taijijian

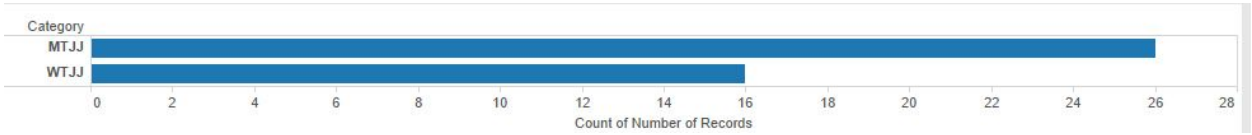

Figure XXV. Comparison of athletes competing in Men's Taijijian and Women's Taijijian.

From the above images, all men's categories have a greater number of athletes than their female counterparts. Therefore, I became to think that my initial question was too simple and I presented more questions to my data set when it came to the differences between male and female athletes. The additional questions I wanted answered included:

What is the difference in number of athletes between male and females in the same event?

What was the highest score in each event? What was the lowest?

Was it a male or female that received the highest score? What about the lowest score?

What was the average in the general category considering both male and female scores?

To answer these questions, I created a table out of the information acquired in Tableau. The table can be seen below.

Event	Men	Women	Diff	Highest Score	M or W	Lowest Score	M or W	Average
CQ	50	18	32	9.69	M	7.78	W	8.96
NQ	25	9	16	9.73	M	7.67	M	9.12
TJQ	15	8	7	9.71	M	7.97	M	9.34
JS	27	23	4	9.64	Both	7.76	W	9.07
DS	37	13	24	9.66	M	0	M	9.0
GS	33	14	19	9.65	M	7.32	W	9.13
QS	25	18	7	9.71	W	7.46	M	9.04
ND	30	11	19	9.67	M	7.05	W	9.02
NG	33	11	22	9.66	M	7.35	W	9.02
TJJ	26	16	10	9.69	M	6.67	W	8.97

Figure XXVI. Table showing events, highest scores, lowest scores, averages, number of athletes competing.

In this table, the average is calculated from the average of both the men's and women's events. It can be seen that more men compete in each category than women with the difference varying between four people and thirty two people in a category. Men had the highest scores in 8 out of the 10 categories, women had the highest score in 1 category, and there was a tie in Jianshu. The lowest score is more evenly divided with women getting 6 of the 10 lowest scores and men getting the lowest score in four categories.

After analyzing all of the data, the final new questions I had were:

What was the highest score in the entire competition?
Who and what country won in the most categories?
Who and what country won first place in each category?
Which athlete had the highest average score in the competition while competing in two or more categories?

First, I organized a chart to show the gold medal winner in each of the categories. The dimensions I considered were the athlete's last name, country, and category. The resulting chart is shown below.

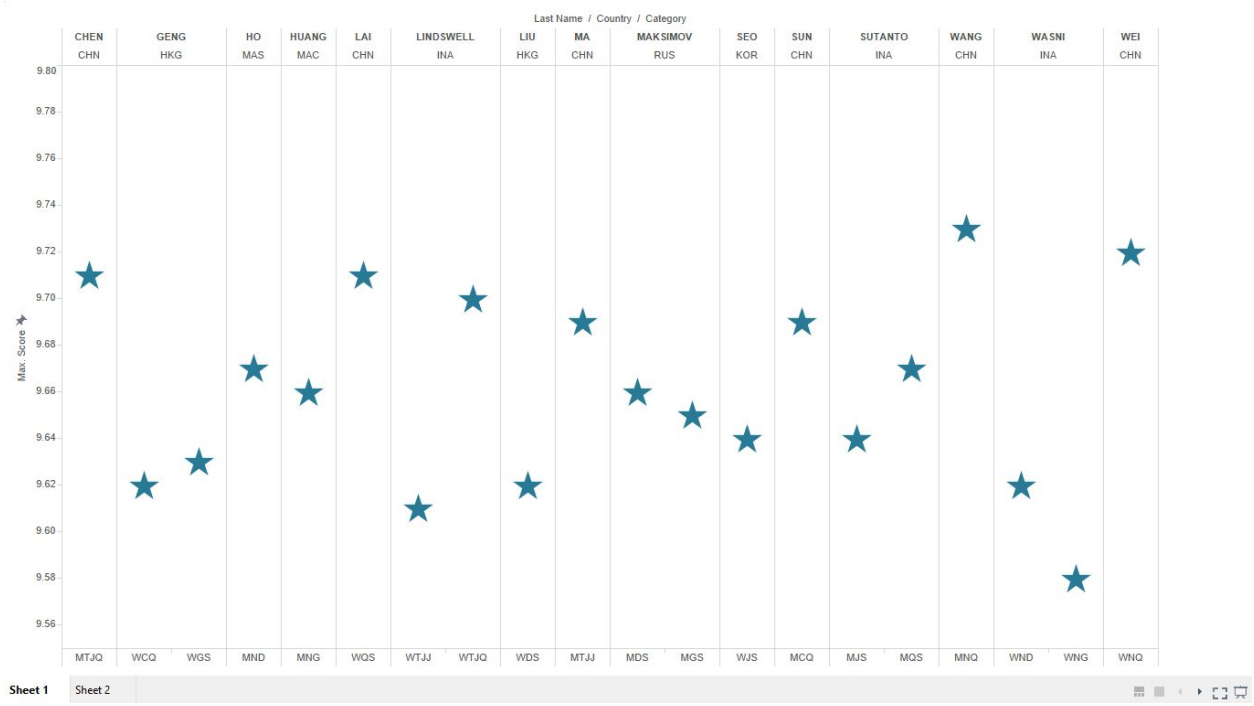

Figure XXVII. The gold medal winners in all twenty categories.

The highest score in the competition came from China's Wang Di in Men's Nanquan with a score of 9.73. China also won gold in six out of the twenty possible categories. In order to find the athlete with the highest average score, I took the measure of the average score and the dimensions of country, category, last name, and first name. I also excluded athletes that only competed in one category and filtered the list down to athletes who placed in the top eight to narrow down the search. The athlete with the highest average score was Macau's Huang Jun Hua with an average score of 9.67 after competing in Men's Nanquan and Men's Nangun.

Final Visualizations

To answer my first initial question about which country won the most medals, the chart below shows the final visualization that answers the question. The chart shows only the countries that placed in the top eight of each category. The colours represent the place the country received in a category with the lighter the colour, the higher they placed. The visualization also shows the number of times they placed in a particular rank and can be summed to find the total number of awards won by a country. As I continue to analyze this

data, I hope to delve deeper into the reasons why medals are won mostly by Eastern-Asian and Southeastern-Asian countries. I will do so by looking at the Chinese population in certain countries, the geographical location of countries, and the political relationship those top twelve countries have with China.

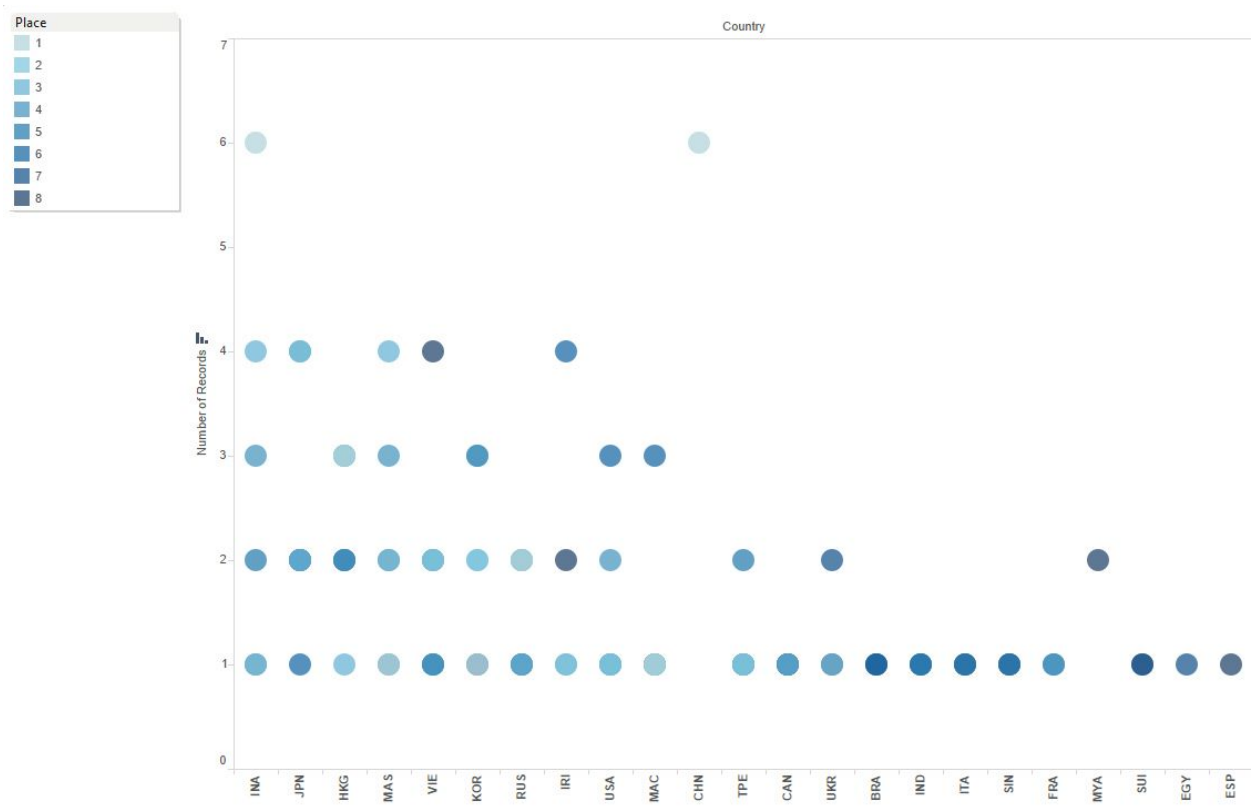

Figure XXVIII. Final visualization to show how many medals each country won.

To answer the second question about the average score of each category, I believe it is still best to show the data in a simple table because the question itself is so simple. The only difference between the final visualization below and Figure XII is that I removed the colours because I don't think it is necessary to represent the average values of the scores across all categories. However, I think my initial question was too simple and as I continue working with this data set I hope to explore this question further by comparing the averages between different categories as I have done above in Figure XIII and analyze the average scores according to countries and wushu styles. I could also analyze the minimums, maximums, standard deviations, medians, and variants of the scores to find new trends in the data set.

Category	
MCQ	8.7758
MDS	8.8068
MGS	9.1467
MJS	9.0063
MND	9.1110
MNG	9.1085
MNQ	9.0592
MQS	9.0516
MTJJ	9.0088
MTJQ	9.3427
WCQ	9.1389
WDS	9.2423
WGS	9.1179
WJS	9.1326
WND	8.9200
WNG	8.9236
WNQ	9.2333
WQS	9.0361
WTJJ	8.9206
WTJQ	9.3425

Figure XXIX. Final visualization for the average scores of each category.

The final question asked whether there was a difference between male and female scores in the same events. I believe Figure XXVI does a good job in comparing many different aspects of the competition between males and females in order to answer this question.

Event	**Men**	**Women**	**Diff Mvs.W**	**Highest Score**	**M or W**	**Lowest Score**	**M or W**	**Average**
CQ	50	18	+32	9.69	M	7.78	W	8.96
NQ	25	9	+16	9.73	M	7.67	M	9.12
TJQ	15	8	+7	9.71	M	7.97	M	9.34
JS	27	23	+4	9.64	Both	7.76	W	9.07
DS	37	13	+24	9.66	M	0	M	9.0
GS	33	14	+19	9.65	M	7.32	W	9.13
QS	25	18	+7	9.71	W	7.46	M	9.04

ND	30	11	+19	9.67	M	7.05	W	9.02
NG	33	11	+22	9.66	M	7.35	W	9.02
TJJ	26	16	+10	9.69	M	6.67	W	8.97

Figure XXX. Table showing events, highest scores, lowest scores, averages, number of athletes competing.

I only edited the table slightly by adding that the "Difference" is the number of men in a category compared to women. Each cell has a "+" symbol added to it to show that men have that amount of athletes more competing in the category over women. However, this table does not show the patterns in each category so my final visualization to answer this question would also need to include the following ten charts that show the patterns in each category.

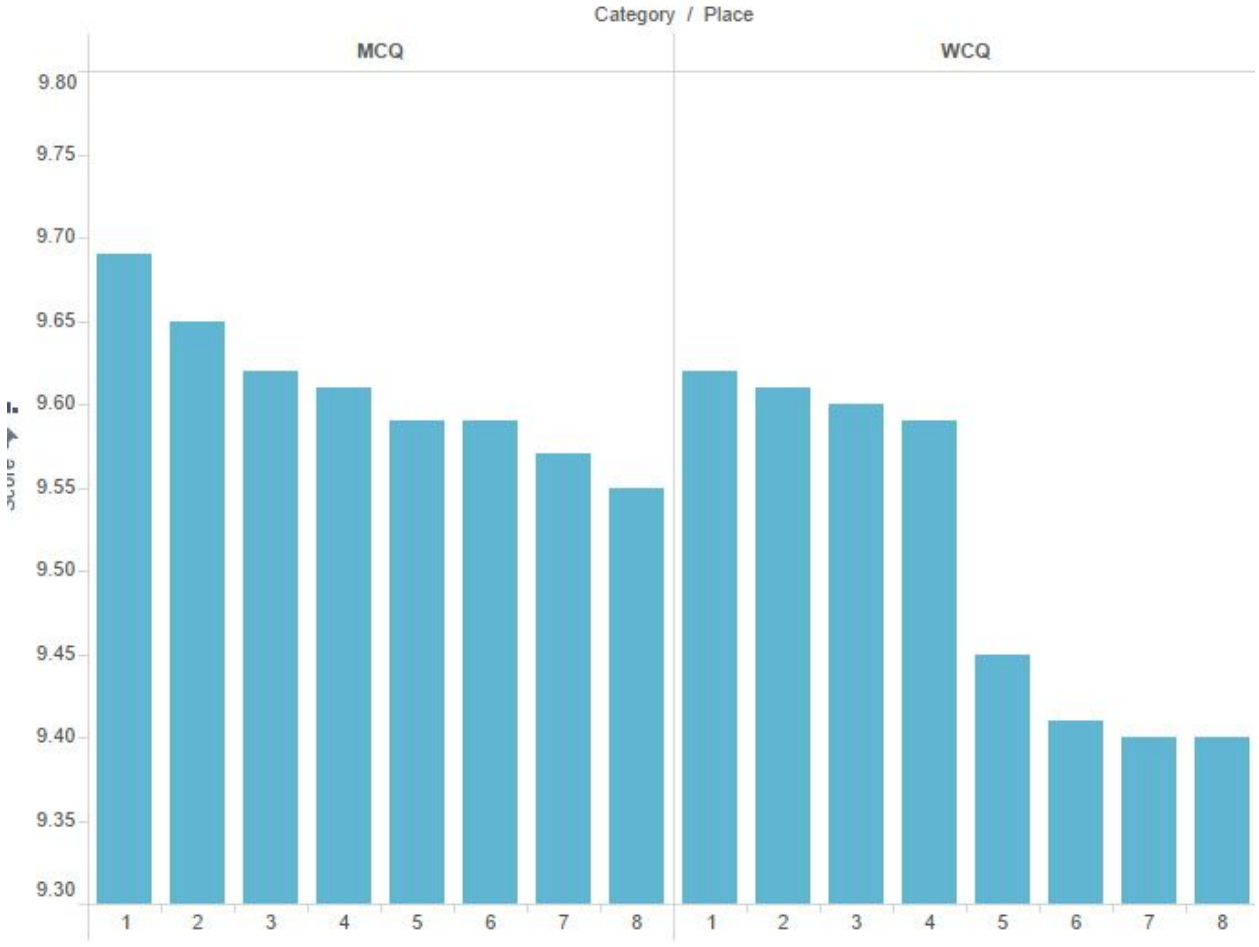

Figure XXXI. Changquan

Through this visualization we can see that the men's scores were much more uniform in the top eight competitors while there is a difference that can be easily seen in the women's

category between the top 1-4 and the top 5-8 athletes. The highest score here is also given to a male athlete in this category.

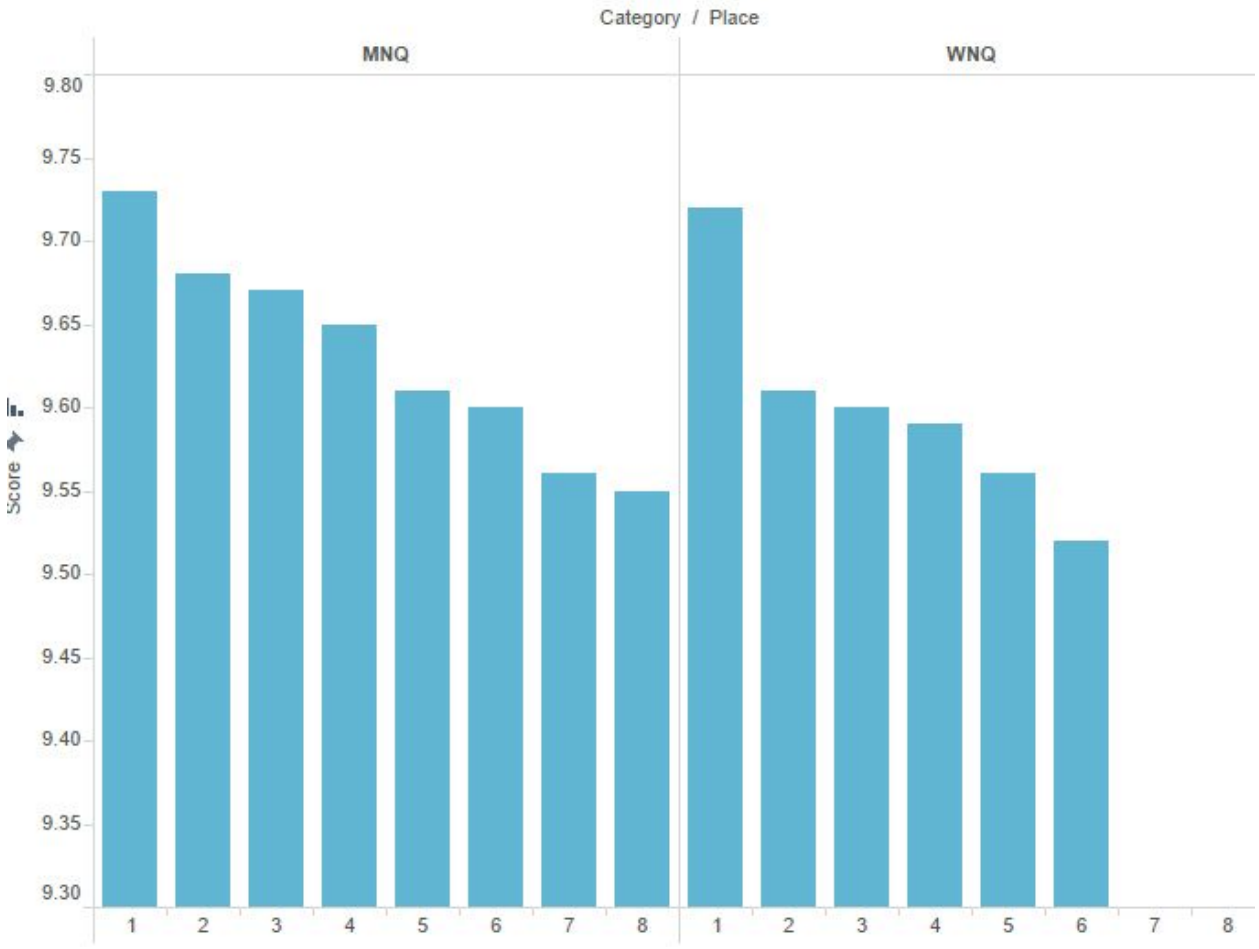

Figure XXXII. Nanquan

In the nanquan event, the highest score was given to a male athlete but there is only a very slight difference between first place in the men's and women's category. Also, the women's category was so small that only a top 6 could be taken and the seventh and eighth place are left blank. There is also another significant difference between first and second place in the women's category.

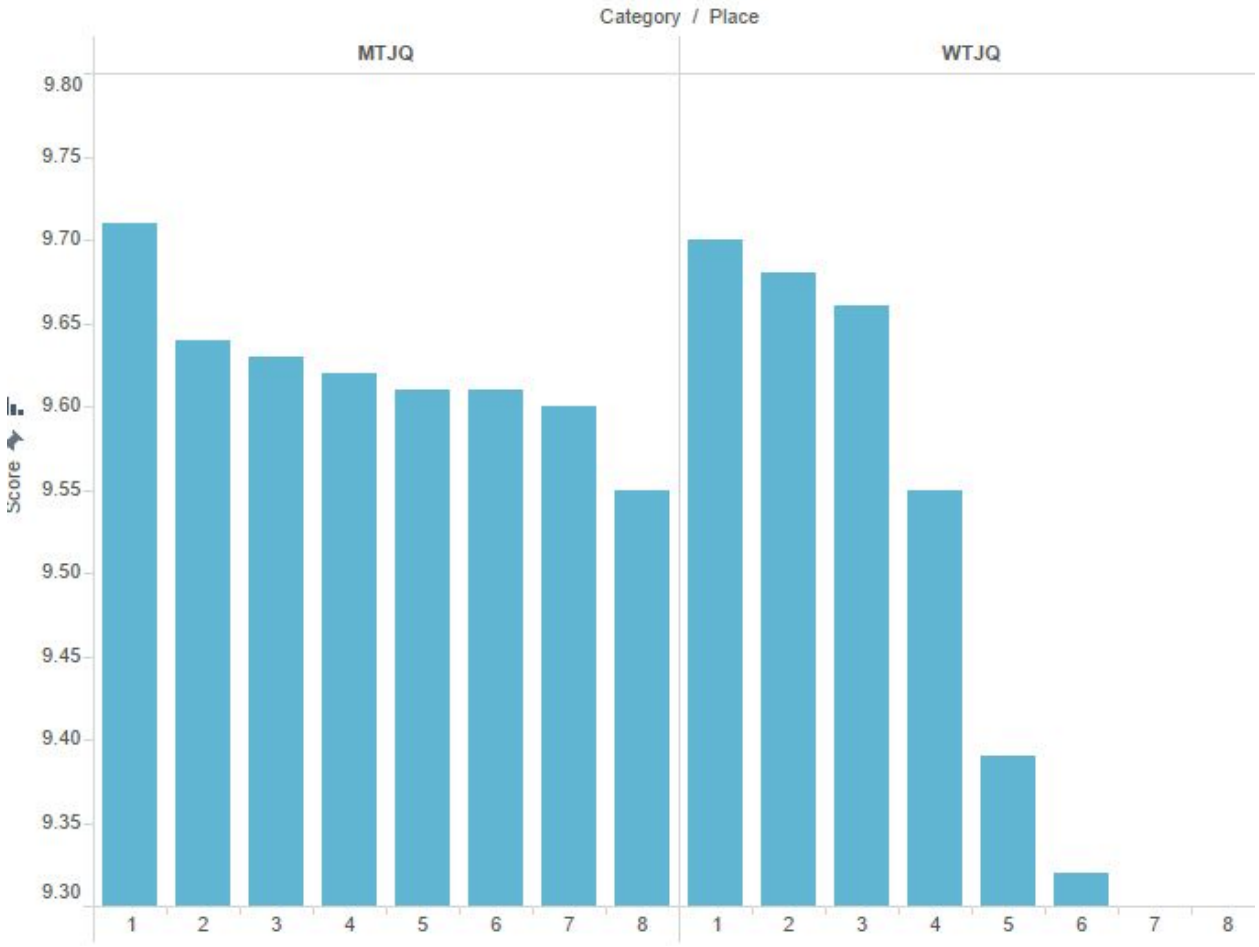

Figure XXXIII. Taijiquan

Like the Nanquan event above, the men's and women's first place scores are very similar. In the men's category, there is over a 0.05 point difference between first and second place then the scores seem fairly close for the rest of the top eight. In the women's category, there were no athletes for seventh and eighth place. There also seems to be very big differences from fourth, fifth, and sixth place. The sixth place athlete seems to have scored very low because the bar height is so short however the baseline is at 9.30 which is still a very good score to receive.

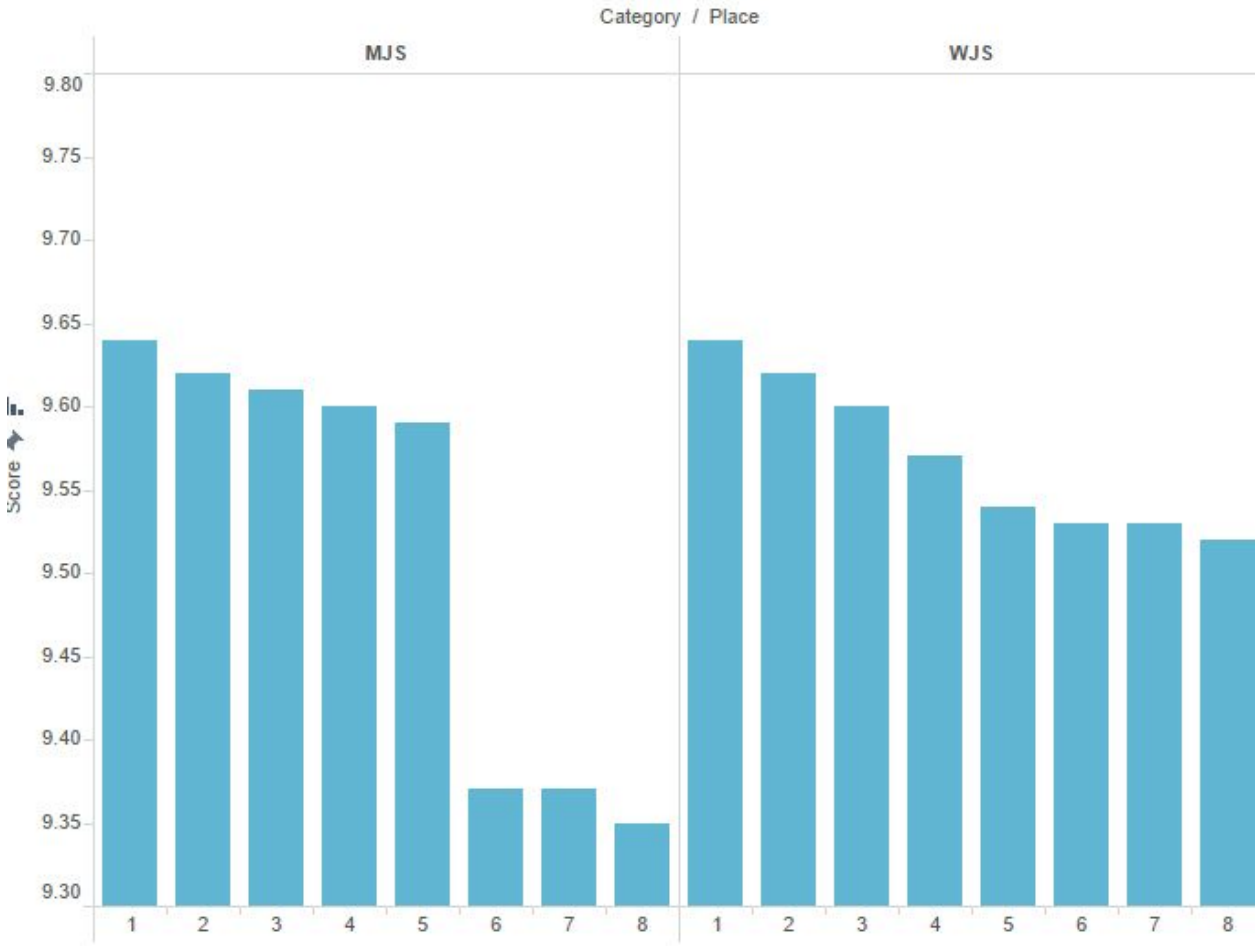

Figure XXXIV. Jianshu

In the Jianshu category, both men and women first place winners received the same score. Up until fifth place, the scores for men were fairly close then there was a o.2 drop to sixth place. The women's scores seem much closer than the men's while they still maintain a staircase kind of look. Given the data shown here, it can be assumed that the women's category was much more competitive than the men's; with the female athletes being much closer in skill level all around.

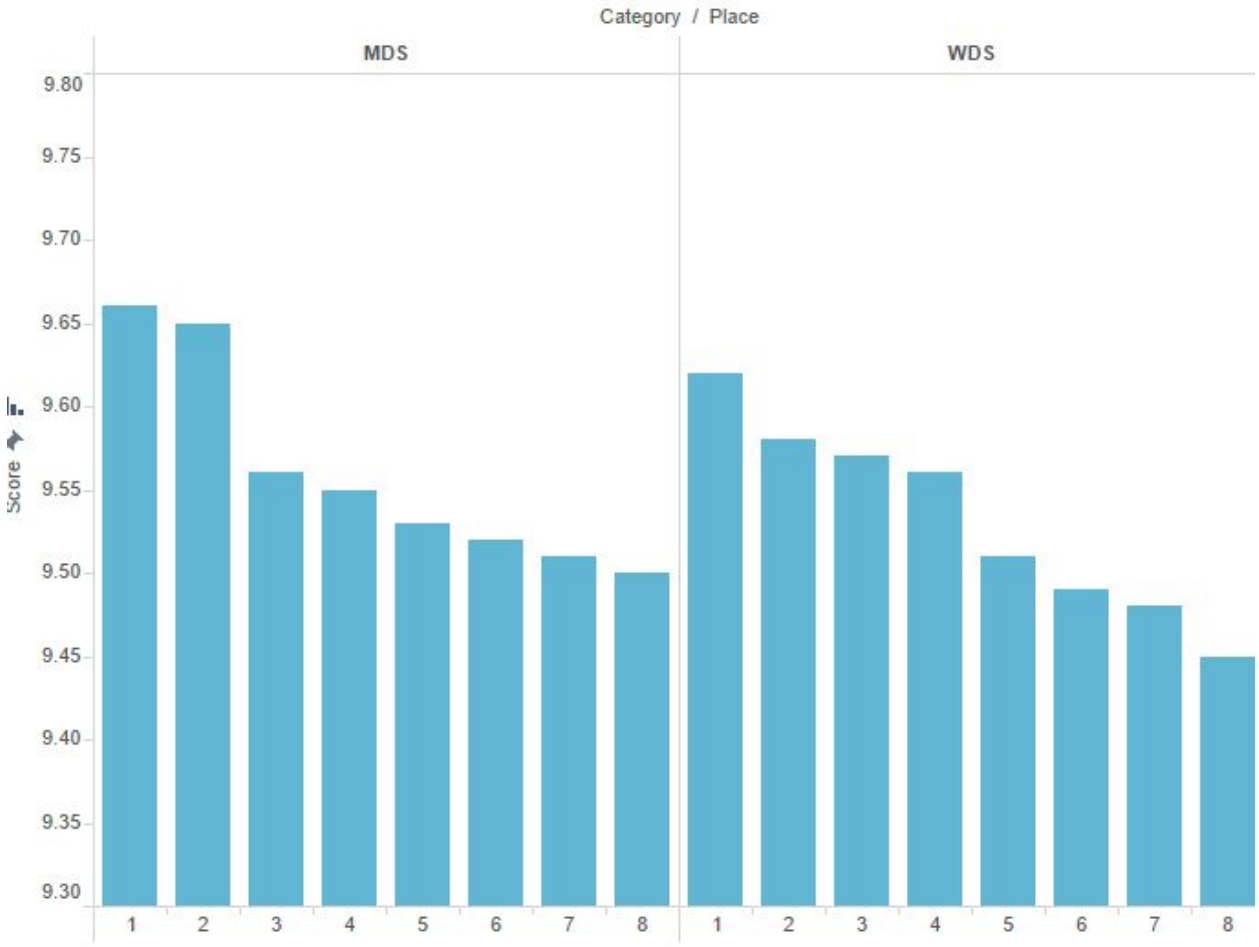

Figure XXXV. Daoshu

The men's category for Daoshu looks like first and second place were extremely close compared to the rest of the top eight. The women's scores were much more evenly distributed. It seems that there is a greater difference between the men's and women's scores between the top eight but it is only because there is such a big difference between first and second place compared to third place in the men's category. In reality, from first to eighth, the scores differ only slighter more than 0.1 points.

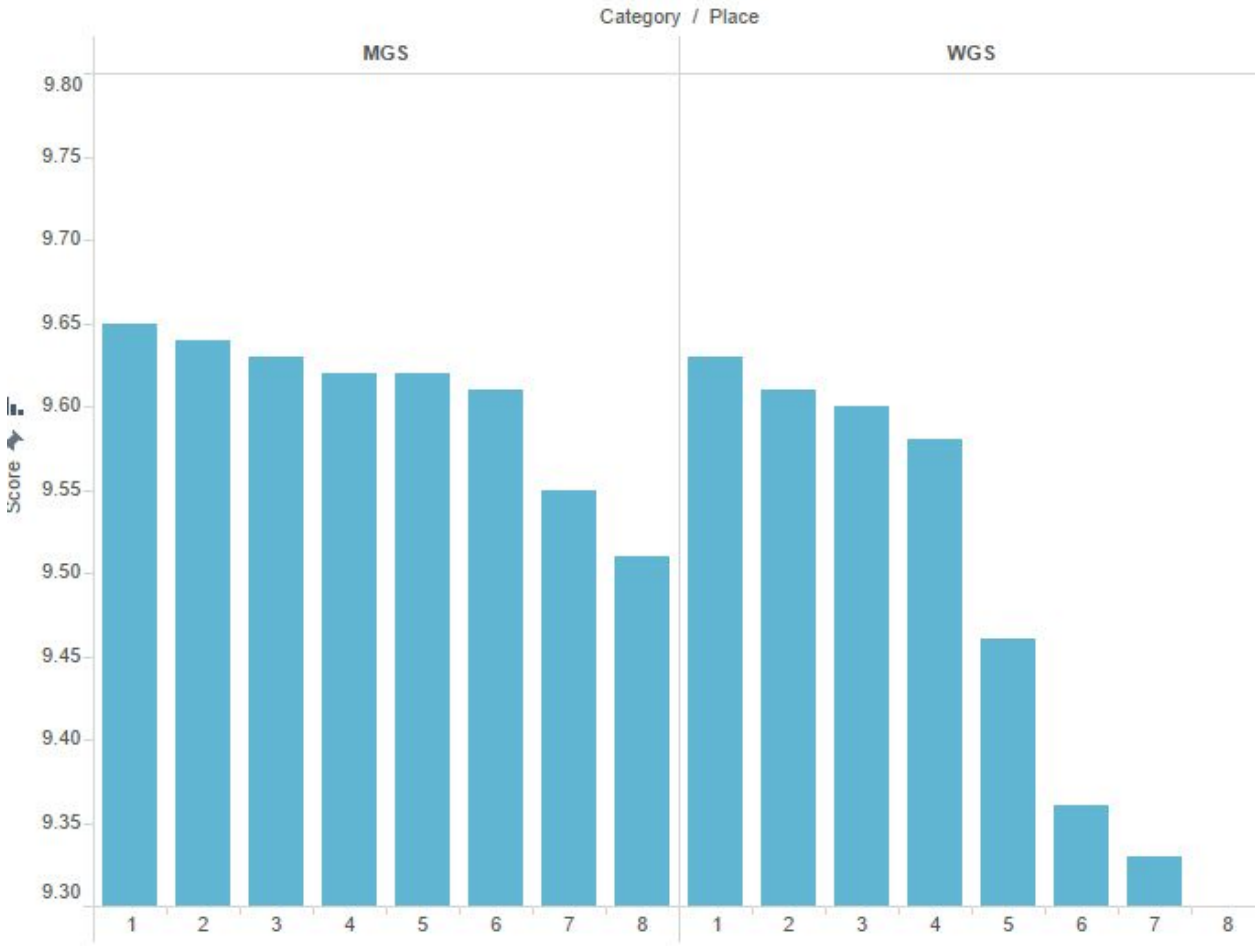

Figure XXXVI. Gunshu

In the female event, not enough athletes were present for there to be an eighth place. However, there is a drastic difference between seventh place for the men's event and seventh place for the women's event. There are definitely more dramatic differences between the female athletes in this category than in the men's category. Even so, the first place score for both men and women are very close.

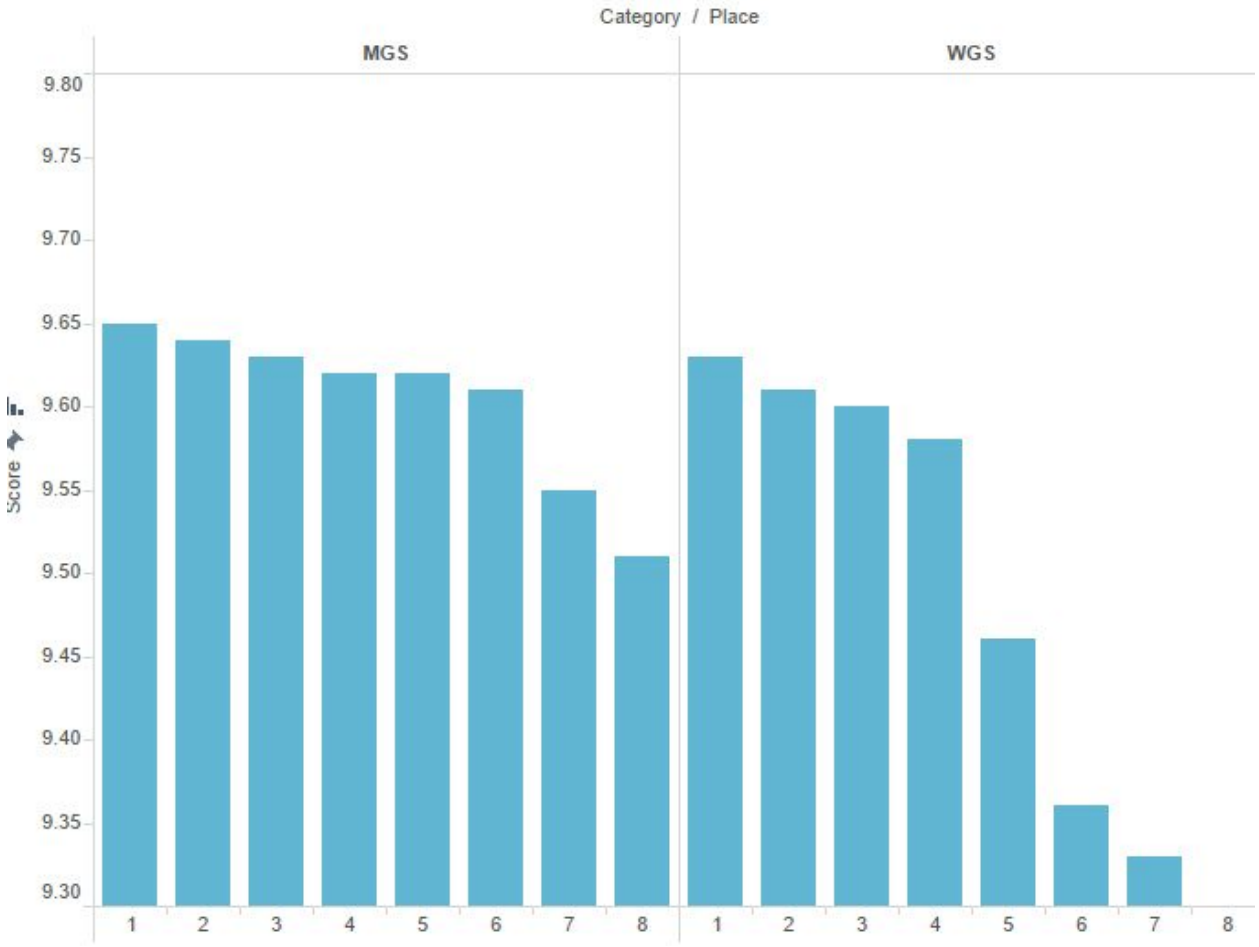

Figure XXXVII. Qiangshu

Again, Qiangshu provides not enough female athletes for there to be an eighth place. Because of the small number of athletes, there is a noticeable difference in the scores between first and seventh place. Though the first place scores only differs by 0.01 points, there is a drastic difference between the athletes that round out the top eight. As also seen in Gunshu and Taijiquan, there is a big difference between female scores, much more so than the male scores.

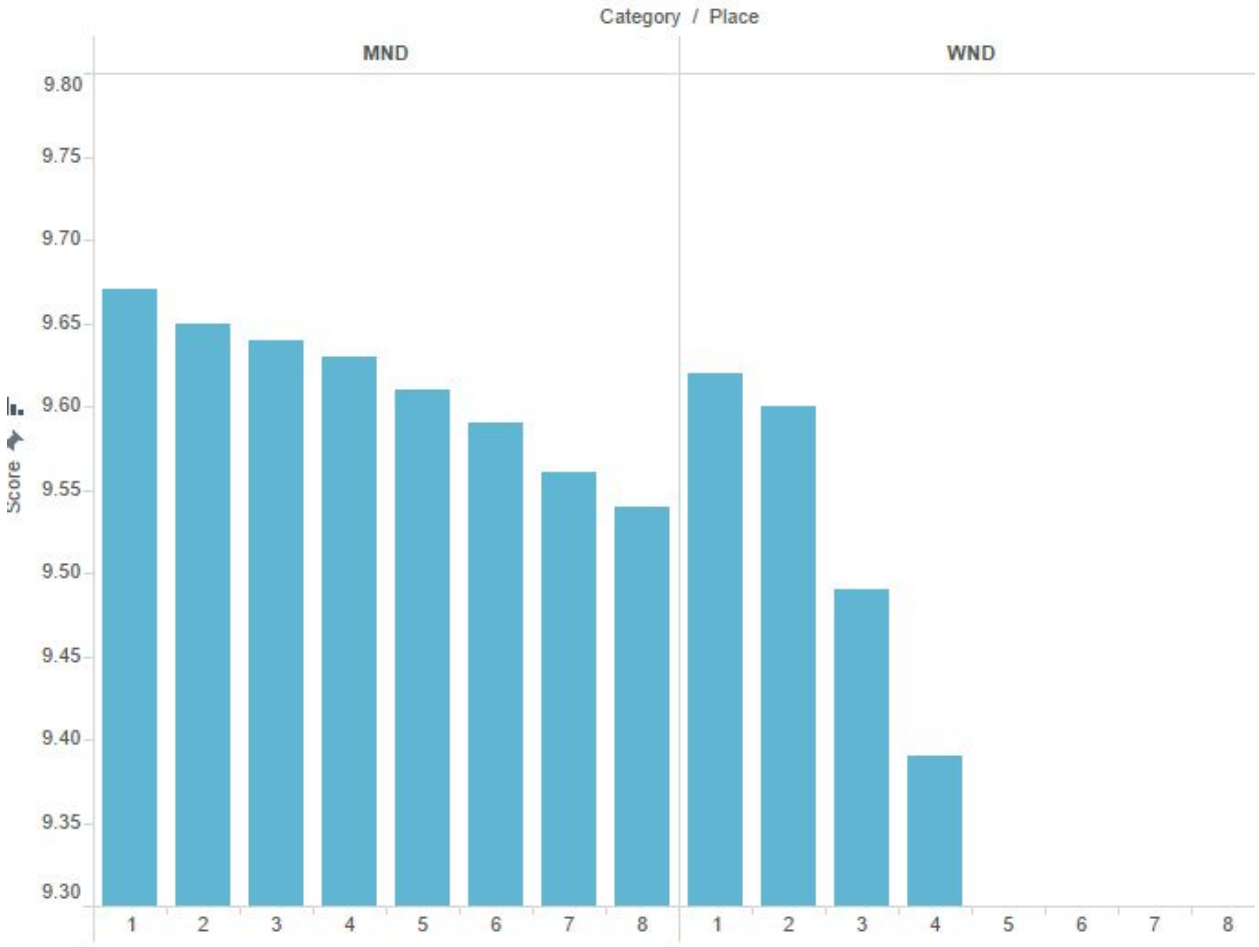

Figure XXXVIII. Nandao

In Nandao, the women's event only had four athletes competing in it so the bar chart looks very incomplete. The difference in the skill level is also more apparent in the women's category because of this reason. The men's category is much more evenly distributed compared to the women's and the difference between each score is very similar as well since there are no huge drops between them. The low number of women competing in this event can be due to the nature of Nandao in general. Nanquan, Nandao, and Nangun are all heavily male dominated because they require a specific body type in order to perform well. Since few women have this body type, not many female competitors competed in those three categories.

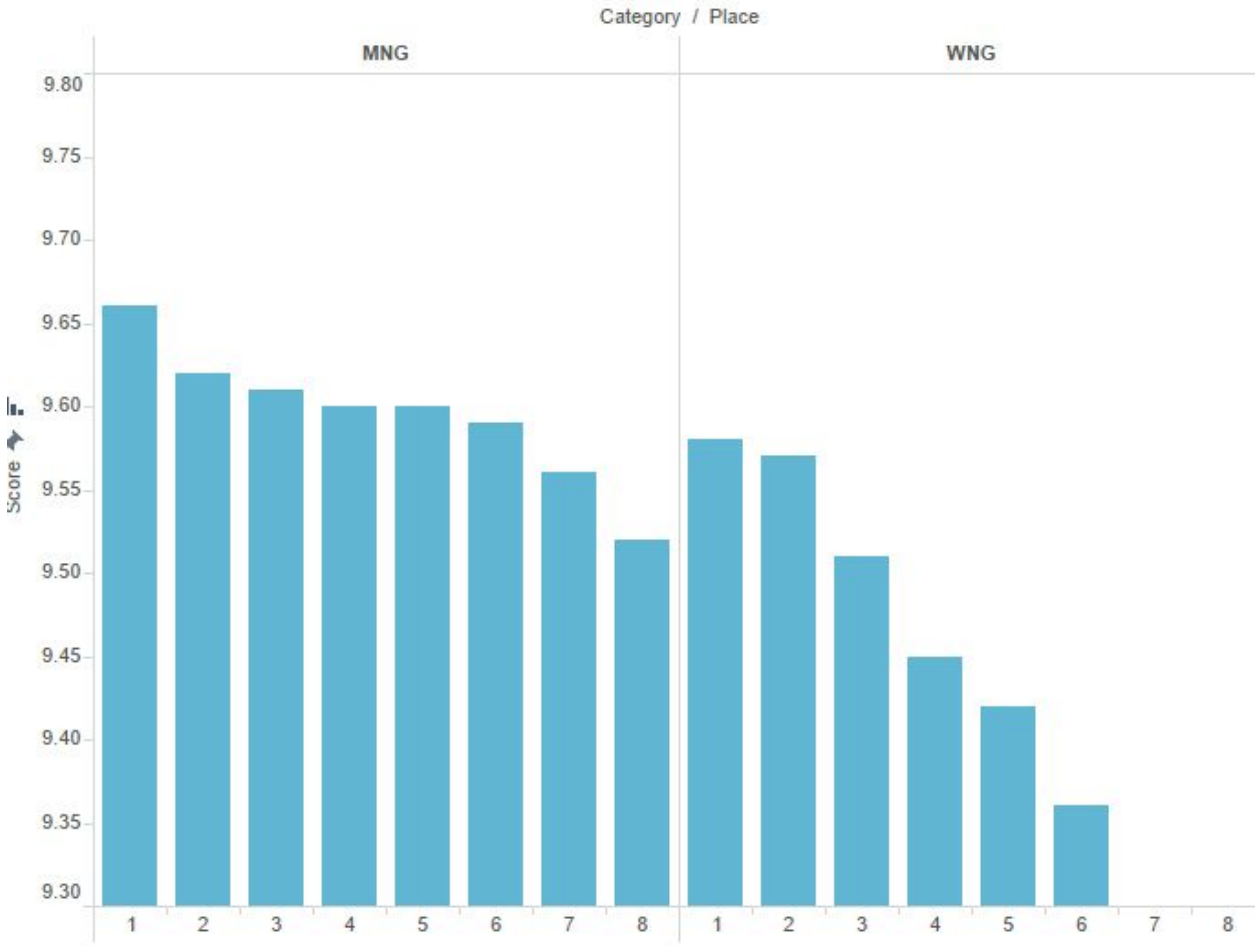

Figure XXXIX. Nangun

First thing to notice in the Nangun event is that there is no seventh and eighth place in the women's category. The difference between the highest scores in the categories also differ by more than 0.05 points. The women's category also falls a lot steeper than the men's. The scores are interesting in the Nangun category because it almost looks like the graph can be the graph of one category. The women's scores are almost all lower than the top eight scores in the men's category. If one were to ignore the men's seventh and eighth place, the scores look like they are all arranged in descending order.

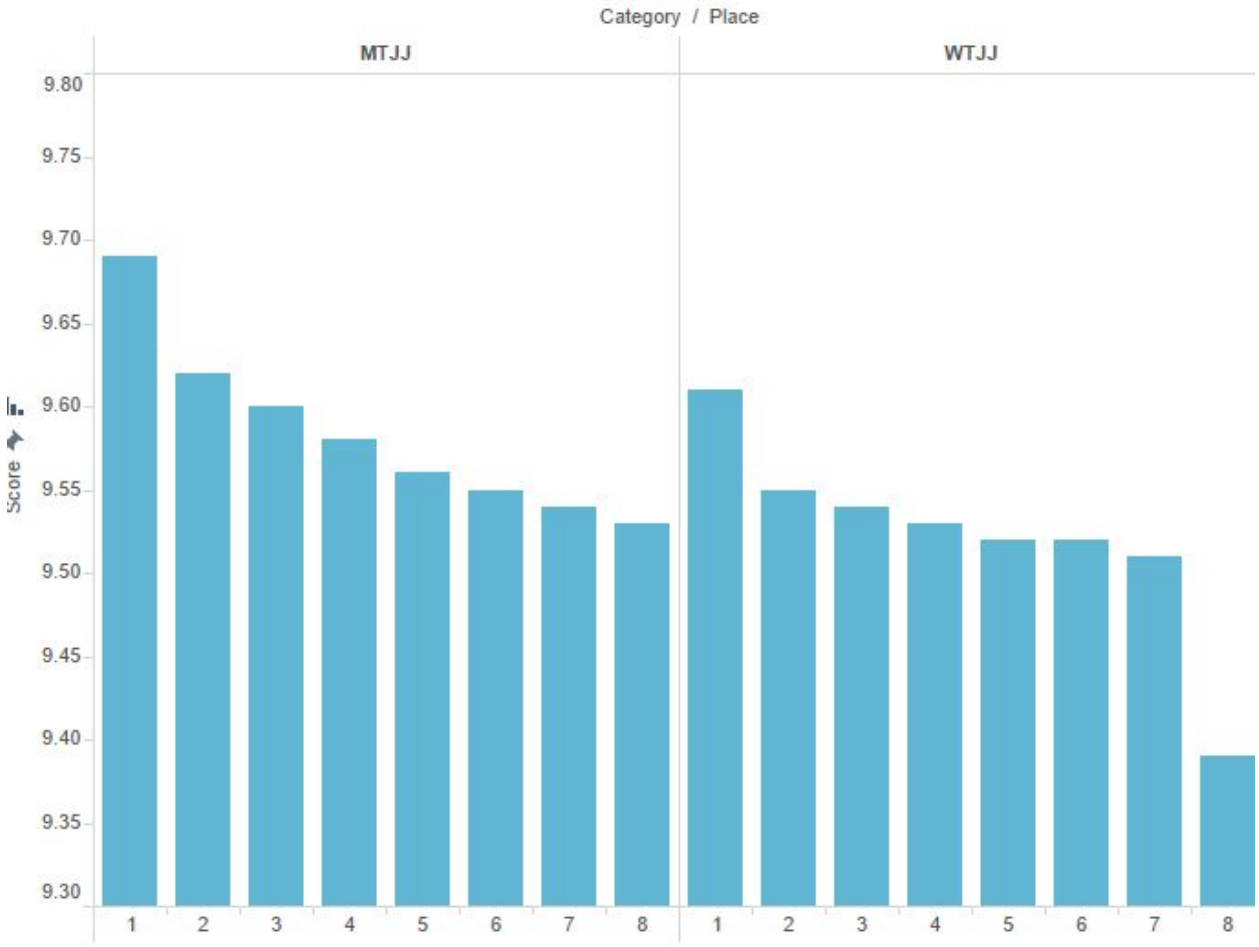

Figure XXXX. Taijijian

The Taijijian category looks fairly similar for both the men's and women's categories. The stand outs are the first place winners who scored at least 0.05 points higher than the runner up. The rest of the scores are all fairly even except for eighth place in the women's category. Between seventh and eighth place in the women's category there was over a 0.1 point difference.

After individually comparing the ten categories and the scores of the top eight men and women of each category I was able to see patterns in the data I would not have otherwise noticed. While making the charts, I did consider using a line chart briefly but since there is no time measured in scores and the bar charts used length to show an increase in value I decided to leave the scores as bars. Since my initial question that asks if there was a significance in the difference of scores between men and women needed so much data to prove anything, I had to use both a chart and multiple bar graphs to illustrate how body type influences the number of participants in a category. I also had to use the table and chart to prove that even though the number of athletes was greater in the men's categories, men still managed a higher average

score in a category. As I continue with this data set I hope to do more comparisons between the two genders and how each category is influenced because of this.

Conclusion

Finally, my data set taken from the 13th World Wushu Championship was used to find the number of medals won by countries, the average score of each category, and the difference in scores between males and females competing in the same category. My data had to be reorganized in many ways so that patterns can be found that might have initially gone unnoticed. Through my analysis, I came to find many more questions that became more specific as I dove deeper into the data. Though my questions started out very straightforward, they provided great insight into my data set and had me consider the different reasons why such conclusions were made. Along with my background knowledge and experience in the sport, I was able to explore the data in depth and have a good understanding of the results of the competition.