# Analyzing the Weight Lifting Performance on the 2000-2016 Olympics Games

Elyse Renshaw
Advisor: Jebessa Mijena

December 1, 2016

# 1 Abstract

In the first part of this research, we fit a model that will predict total as a function of Weight level, height, and actual bodyweight. We build potential best models using 70 percent of our data and test these models using Root Mean Squared Errors. We choose the model with the most minimal RMSE as well as one that satisfies our constant variance and normality assumptions.

This article also studies Weightlifting results from the 2000-2016 Olympic Games for both males and females to determine the differences between gender performance. The data showed that there are considerable differences in the competitive level between male and female lifters. This article compares the results of winners in a weightclass with an average lifter in the weightclass from year to year. This article shows that males in every competition studied had a lower percent difference from first than the same comparison for women. This article also compares the winner from each weightclass to the weightclass above them and observes the place that they would have received. We show that there is a negative correlation between weight class and percentage difference from 1st place for males, but no correlation for females. This shows that the upper weightclasses for males are more competitive than the lower weightclasses. We show that China is by far, outperforming all other countries. Over 40 percent of Gold Medals awarded to females in the last 5 Olympic Games have been given to Chinese women, and around $\frac{1}{3}$ of Gold Medals awarded to males have been given to Chinese male lifters

We can see that as place increases, the difference between male and female results increases. This shows that females are less competitive than males, as the place that the lifter attained increases.

Overall we see that for males, a winner in a weight class would be around 6th place in the weightclass above them, while for females, a winner in a weight class would be around 4th in a weightclass above them.

## 2 Introduction

Over the last 15 years, participation in Olympic Weightlifting has been growing rapidly worldwide and has seen especially large growth in the United States. The creation of a new popular fitness craze, Crossfit, which combines the movements of Olympic Weightlifting as well as gymnastics and cardio has brought more attention worldwide to the sport of weightlifting which was once relatively unknown. Crossfit also has a very large growing population of women participants which leads to more women learning the movements of Olympic weightlifting and has increased competition for both American female weightlifting, as well as increasing competitiveness in international competition.

Weightlifting measures overall strength, flexibility, and athleticism and includes two lifts for which each lifter gets three attempts. The highest successful attempt for each lift is recorded and makes up the athlete's total. First, each lifter gets three attempts at the snatch. In a properly executed snatch, the lifter moves the bar from the ground to overhead in one movement. The second lift is the clean and jerk. In this lift, the lifter first lifts the bar to his or her shoulders, then he or she lifts the bar overhead. In both of these movements, the lift is complete when the lifter has the bar overhead and is motionless with complete control of the bar. Men have a much longer history in the sport of Weightlifting. Men began weightlifting in the first Olympic Games in 1896 while women first appeared in the Olympics in 2000. "Women have been in weightlifting for a very short time, when you consider the long history of men dominating the sport. So on the competition side of things, weightlifting is at a level for women where participants could actually achieve a lot more and go further[1]."

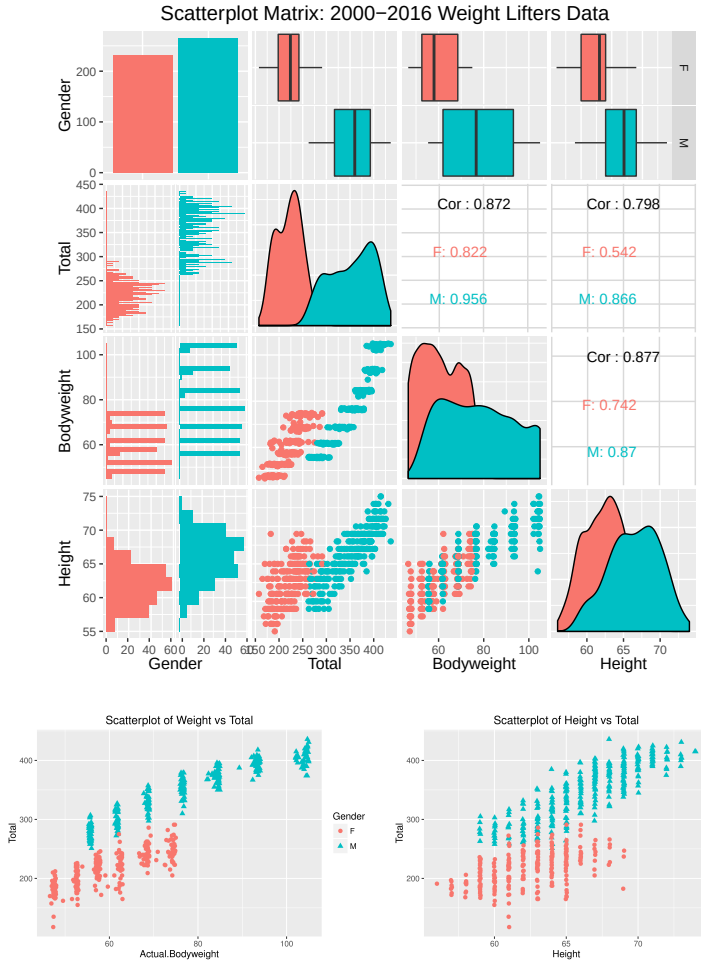
# 3 Chapter 1

## 3.1 Data

In this study, weightlifting data from 2000-2016 was collected for each lifter in each Olympic Games. Models were fit using the top 10 lifters in each weightclass, each year, and for each gender. For some weightclasses, there were less than 10 lifters in a particular year, so the model was developed using the number of lifters in the competition that year.Using only the top 10 lifters in most cases removed outliers and particularly low performing athletes, however we removed Irene Ajambo because although she was 9th in her weightclass and year, she was in last place and lifted over 80kg lower than the winner in her weightclass. For each lifter, their height was recorded as well as their actual bodyweight. There is official data on actual bodyweight of a lifter at every weightlifting competition, however, currently, there is not official height data available. The height data used in this study was obtained from "SportsReference.com". For each lifter that the data was not on this site, height data was obtained from Wikipedia. The superheavy weight categories for men and women, 105+ kg and 75+ kg respectively, were removed because of the high variation between the lifter's actual bodyweight in these categories. Lifters for which height data was not availble were removed. There were 36 lifters for which height was not available. The data was split into a training set whcih included 70 percent of the lifters and a testing set which included the remaining 30 percent. The training set of the data included 234 women and 191 men. The testing data set contained 99 men and 86 women.

## 3.2 Analysis

Following is a scatterplot matrix of the data set:




From these scatterplots, we can see that we may have an issue with muticollinearity since height and weight have a similar affect on total. Observing the variance inflation factor for male and female data separately, we obtain a value of 3.42 for males and 2.20 for females. Thus, we can conclude since neither of these values in greater than 5, that multicollinearity will not be an issue in our model. We want to determine for males and females separately. a model that will best predict total. we think that weightclass may be a good additional predictor, so we classify the weightclasses as follows:

| **Gender** | **Lower** | **Middle** | **Upper** |
|---|---|---|---|
| Male | 56, 62 | 69, 77, 85 | 94, 105 |
| Female | 48, 53 | 58, 63 | 69, 75 |

Next we used a box cox transformation to determine the optimal transformation for each of the variables in our model. We see that powers that we need

to raise total, bodyweight, and height to, respectively, are 2, .2, and 1.25. For interpretably, we use 1, $\frac{1}{4}$, and 1. Using our box cox transformations as well as data on what is currently expected to be the best model for predicting total, we obtain the following four models where $T$ is total, $W$ is actual bodyweight $H$ is height, $WL$ is weightlevel, and $\beta_0$,...$\beta_4$ are constants.

Table 1: Models for Male Performance

| | |
|---|---|
| Model 1 | $\log(T) = \beta_0 + \beta_1 \log(W) + \epsilon$ |
| Model 2 | $T = \beta_0 + \beta_1 W + \beta_2 (W)^2 + \epsilon$ |
| Model 3 | $T = \beta_0 + \beta_1 (W) + \beta_2 (W)^2 + \beta_3 (H) + \beta_4 (WL) + \epsilon$ |
| Model 4 | $T^2 = \beta_0 + \beta_1 (W)^{\frac{1}{4}} + \beta_2 (H) + \epsilon$ |

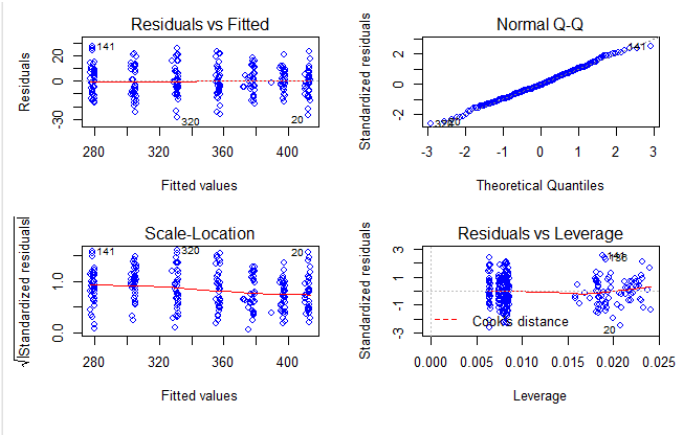
Test the root mean squared errors for these models, we obtain the following values:

Table 2: RMSE Values for Male Models

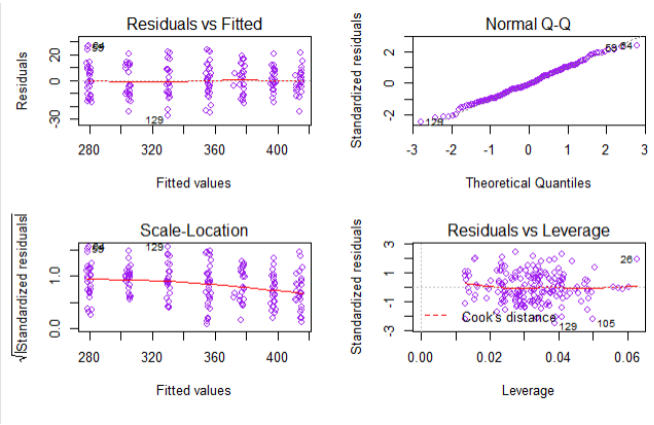
| | |
|---|---|
| Model 1 | 13.62 |
| **Model 2** | 10.55 |
| **Model 3** | 10.50 |
| Model 4 | 10.73 |

From this, we can see that the two best possible models are model 2 and model 3. Following the diagnostic plots for model 2 and model 3. Since they perform similarly in predicting, we want to see if one performs better in terms of constant variance assumption and normality.

$$T = \beta_0 + \beta_1 W + \beta_2 (W)^2 + \epsilon$$




$$T = \beta_0 + \beta_1 (W) + \beta_2 (W)^2 + \beta_3 (H) + \beta_4 (WL) + \epsilon$$




From these diagnostic plots, we can see that model 3 has slightly improved normality. For this reason, and since it's RMSE value is slightly lower, we choose th following model as the best predictor of male weightlifting performance:

$$T = \beta_0 + \beta_1(W) + \beta_2(W)^2 + \beta_3(H) + \beta_4(WL) + \epsilon$$

.

Similarly, we want to build a model to predict total for female weightlifters. Using box cox transformation, we see that we should transform total, bodyweight, and height by .8, .2, and .4 respectively. For interpretability, we will transform total, bodyweight, and height by 1, $\frac{1}{4}$, and $\sqrt{}$. Using box cox transformations as well as using previous knowledge about accepted best predictive models, we obtain the following 4 models.

Table 3: Models for Female Performance

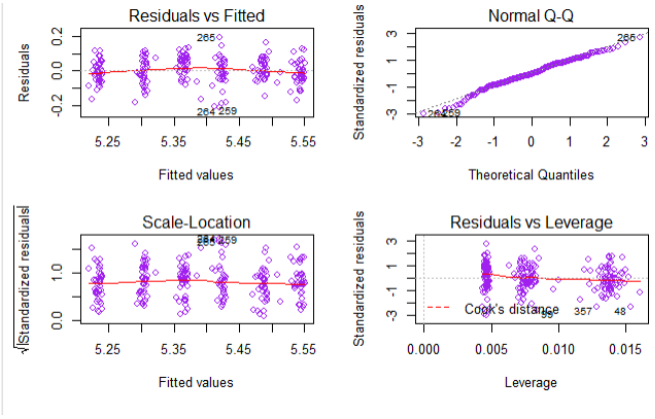
| | |
|---|---|
| Model 1 | $\log(T) = \beta_0 + \beta_1 \log(W) + \epsilon$ |
| Model 2 | $T = \beta_0 + \beta_1 W + \beta_2(W)^2 + \epsilon$ |
| Model 3 | $T = \beta_0 + \beta_1(W) + \beta_2(W)^2 + \beta_3(H) + \beta_4(WL) + \epsilon$ |
| Model 4 | $T = \beta_0 + \beta_1(W)^{\frac{1}{4}} + \beta_2\sqrt{H} + WL + \epsilon$ |

Table 4: RMSE Values for Female Models

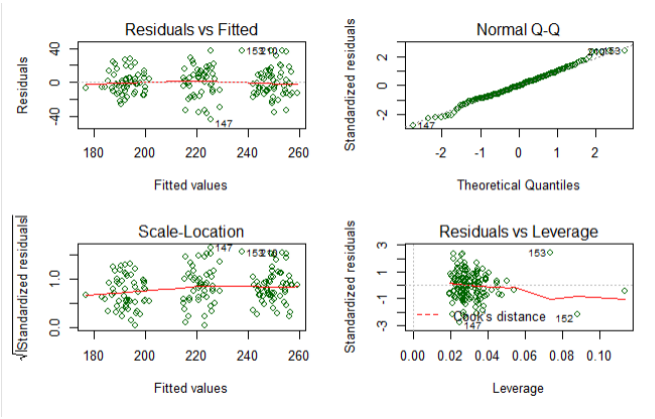
| | |
|---|---|
| **Model 1** | 16.86 |
| Model 2 | 17.35 |
| Model 3 | 17.01 |
| **Model 4** | 16.86 |

From these RMSE values, we see that the best model is possibly model 1 or model 4. Following are the diagnostic plots for these two models.

$$\log(T) = \beta_0 + \beta_1 \log(W) + \epsilon$$




$$T = \beta_0 + \beta_1 (W)^{\frac{1}{4}} + \beta_2 \sqrt{H} + WL + \epsilon$$



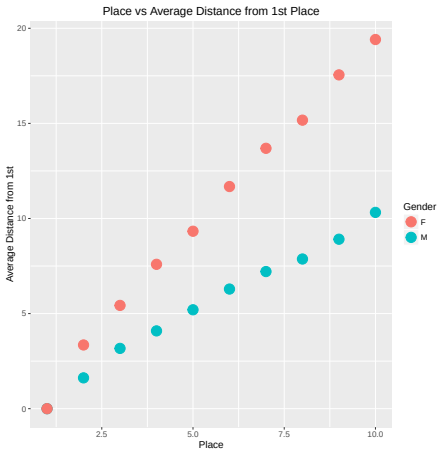

From this, we see that the two models perform very similarly. Since model 1 is much more simple, we choose model 1 which follows:

$$\log(T) = \beta_0 + \beta_1 \log(W) + \epsilon$$

# 4 Chapter 2

## 4.1 Analysis

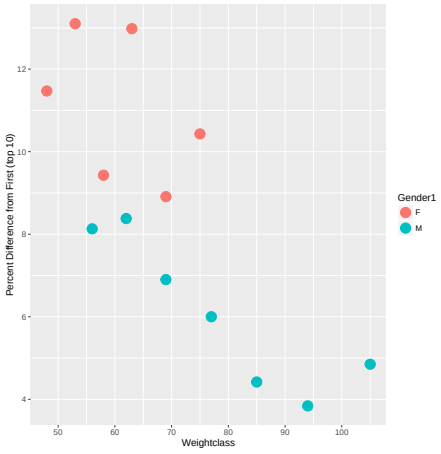
First, we will observe the affect of place on percent difference.




For males, Percent Difference=1.08(Place)-.46. For females, Percent Difference=2.08(Place)-1.13. We can see by comparing the slopes, that for each increase in place, the percent difference that the lifter is on average from first place grows at a rate around 2 times as fast for females as for males. For males, each increase in place corresponts to about a one percent increase in percent difference from first. We can conclude that as place increases, women are increasingly less competitive than males.

Next, we ask the question, does percent difference depend on weightclass? We want to see if certain weightclasses are more competitive or have more homogeneity in their data.

The Affect of Weightclass on Percent Difference




We can see that for men, there appears to be negative correlation between weightclass and percentage difference from 1st place. We can conclude that for males, the heavier weightclasses are more competitive than the lower weight-

classes. For women, there seems to be no correlation between weightclass and percentage difference from 1st.

Next, we will observe how competitive a winner can be in a weight class above them. Since we have small sample sizes, for the following tests we will use the Mann-Whitney U-Statistic as our test to compare samples. This test is nonparametric and does not assume constant variance.

$$\boxed{U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - W}$$

where $\bar{x_1}$ , $s_1^2$ and $N_1$ are the 1st sample mean, sample variance and sample size, respectively.

First, for males and for females, we want to test the winner of a weightclass with the winner of the weightclass above them to see if there is a difference. We want to know if a winner in one weightclass could also be a winner in the weightclass above them. For these test we obtain the following p-values:

| Gender | U | P-Value |
|---|---|---|
| M | 312 | .00382 |
| F | 384 | .0001372 |

Since both of these p-values are low compared to .05 we reject the null hypothesis and we can conclude that a winner in a weightclass cannot also be a winner in the weightclass above them.

| Weightclasses Compared | Average Place |
|---|---|
| 56-62 | 5 |
| 62-69 | 7 |
| 69-77 | 6 |
| 77-85 | 6 |
| 85-94 | 8 |
| 94-105 | 6 |
| Average | **6.33** |

Using the performance of winners from 2000-2016, we determine that for males, on average a winner of a weightclass would be 6.33 place in the weightclass above them.

| U | P-Value |
|---|---|
| 447 | .9705 |

Comparing male winners to 6th place in the weightclass above them, we can confidently conclude that being a winner of a male weightclass is the same of being 6th place in the weightclass above.

| Weightclasses Compared | Average Place |
|---|---|
| 48-53 | 4 |
| 53-58 | 5 |
| 58-63 | 3 |
| 63-69 | 4 |
| 69-75 | 3 |
| Average | **3.8** |

For women, on average, Gold medalists obtain 3.8 place.

| U | P-Value |
|---|---|
| 214.5 | .7048 |

Comparing winners to 4th place in the weightclass above them, we can conclude that a winner of a female weightclass would be expected to place 4th in the weightclass above them. We can see that men are much more competitive between weightclasses. While on average male winner are 3 spots away from being a medalist in the weightclass above them, we see that women are only one place away from medaling even when competing with other heavier women.

It is of interest to observe how similar weightclasses between genders compare. Due to structural differences in the anatomy of males and females, we know that males and females of the same weight will lift differently. We want to see how differently the genders perform. There is only one weightclass that is the same for both males and females (69kg), there are however many similar weightclasses. Following are the the comparable weightclasses as well as the average weight lifted by the winners.

| Male | Weight lifted | Female | Weight lifted |
|---|---|---|---|
| 56 | 298.4 | 58 | 238.6 |
| 62 | 322.8 | 63 | 246.6 |
| 69 | 349.8 | 69 | 265.1 |
| 77 | 373.3 | 75 | 272.9 |

For each of these comparisons, we compare winners for each year and for each individual weightclass. First we want to see if there is a difference between men and women of similar weights.

| U | p-value |
|---|---|
| 400 | 6.72e-8 |

From this, we can definitively conclude that men and women of similar weights do not lift the same. Next we want to compare how different the genders of similar sizes perform.

| **Year** | **Average Weight Diff** | $\sigma$ | **Average % Diff** |
|---|---|---|---|
| 2000-2016 | 80.28 | 37.83 | 23.25 |
| Rio 2016 | 79.75 | 44.84 | 22.77 |

On average, over the time period observed, male winners lift 80.28 kilograms more than female winners of similar size. Observing only the most recent Olympic competition, Rio 2016, we see that one average males lifted 79.75 kilograms more than when and the average weight difference is possibly decreasing.

Next, we will observe the gender difference of similar size lifters over the years 2000-2016.

| Event | Avg Weight Difference | Avg% Difference | Avg%Diff Previous Year |
|---|---|---|---|
| Athens 2000 | 99.9 | 29.236 | n/a |
| Sydney 2004 | 78.75 | 23.26 | -5.97 |
| Beijing 2008 | 66.5 | 20.42 | -2.84 |
| London 2012 | 75 | 22.12 | 1.7 |
| Rio 2016 | 79.75 | 23.24 | 1.12 |

As we can from Athens to Sydney there is a very large drop in the percent difference in the amount of weight lifted by the strongest men and women in similar weight classes. And again from Sydney to Beijing there is a decrease in the percent difference. But, from Beijing to London and London to Rio, we can see that the difference increases again. We can see though that the percent changes in these two year combined are less than the drop from Sydney to Beijing. We should consider the possibility that we are seeing the average percent difference leveling out as we get farther from 2000, when women first began competing. "The results in female weightlifting today have much higher level that 15-20 years ago...in the case of female competitors, there are no well balanced results, the non homogeneity is high; there are huge differences between the results of the winners and the others" We can measure the homogeneity of each of the genders performances by observing the relative difference between the winner of a weight class and an average lifter in that weight class.

Observing the 2012 Beijing Olympic Games, we noticed that performance of the winner in the 62kg category for males was very similar to performance in the 75+ kilogram category for females. Comparing these 2 weightclasses from 2000-2016, we obtain the following:

| U | p-value |
|---|---|
| 417 | .2594 |

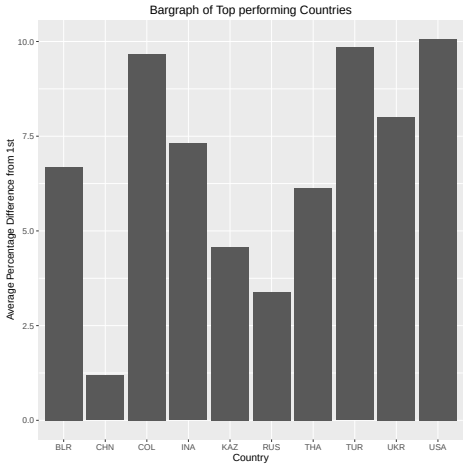
Therefore, we can conclude that a winner of the 62kg for males is the same as a winner of the 75+ kilogram category for females.

We also observed that winning the 56kg category for males may be comparable to winning the 75kg for women.

| U | p-value |
|---|---|
| 295.5 | .026 |

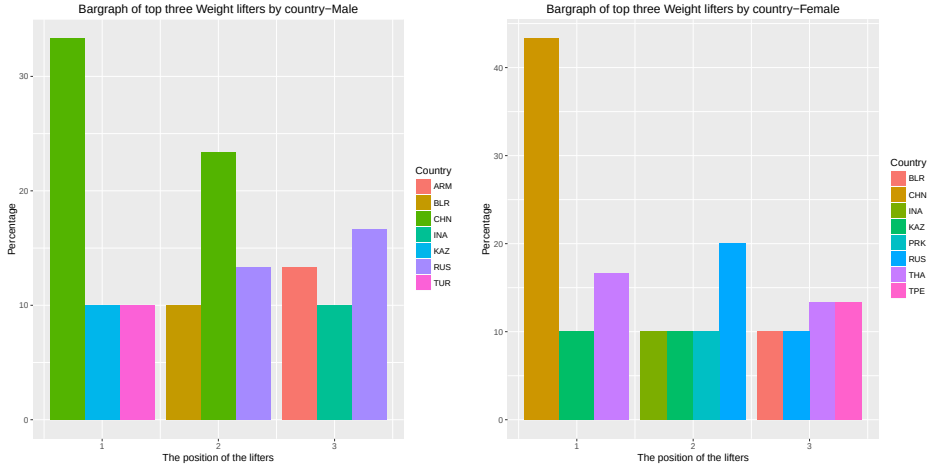
from this we see that our P-value is less than .05. Therefore, we reject the null hypothesis. We cannot conclude that a male winner of the 56kg weightclass is the same as a female winner of the 75kg female weightclass.

Next, we will observe how the percent difference differs between the countries.




How do the top performing countries compare?

From this, we can see the percentage difference from first place on average that a lifter from a particular country attains. As we can see, China has a less than 1 percent difference from 1st place on average. Further observation of the data shows us that of al 39 Chinese lifters in our data set, one obtained 5th place, 3 obtained 4th place, and the rest of the lifters were medalists.




From this graphic, we can see the percentage of medals given to each country. we can see that for males, over 30 percent of gold medals are given to Chinese men and for Female competitors, over 40 percent were given to Chinese Women.

### 4.2 Conclusion

Men have been competing in weightlifting in the Olympic games since the first event in 1986, while Women first appeared in the Olympics to weightlift in 2000. In the 1896 Olympics, men competed at the one hand lift and the two hand lift. Over the years, the sport has evolved and now includes the more complex movements, the Snatch and the Clean and Jerk, as well as changing to include women as competitors. This difference in years of development has led to many differences in performance between the genders. In every Olympic competition between 2000-2016, the average difference between the first place lifter in a weightclass and the 6th place lifter in the same weightclass was higher for females, showing a greater discrepancy between the best lifters and the average lifter for women than for men. Also, in each of these events, the average place of a a Gold Medalist in the weightclass just above them was higher for men than women showing that men have a more competitive field than women. When measuring average percent difference between men and women in the same year in comparable weightclasses we can see that the percent difference between the genders is lowering and extrapolating the data, we can speculate that the percent change from year to year will be getting closer to 0.

## 5 Further Research

As more data becomes available we should continue analyze what best predicts total. Including bodyfat as an additional prediction would possibly be a very strong predictor of total. In September, the International Weightlifting Federation created a new weight class for women. There are now 8 weightclasses for both men and women. The new weight category adopted is 90kg and the superheavy weight category will move from 75+ to 90+ kilograms. The new weight class will begin in December 2016 at the American Open. This change to the previous structure of weightlifting should be observed in future research. In future research, it could be researched what makes Chinese weightlifters different from the rest of the world. Training and body composition among other things should be observed.

## References


[1] Markovic, G., Sekulic, D. Modeling the Influence of Body Size on Weightlifting and Powerlifting Performance *Coll. Antropol.* , 30: 607-613, 2006.

[2] Olympic Sports *SportsReference.com.*

[3] Results by Event *International Weightlifting Federation*, iwf.net.

[4] Szabo, A. S., Nemith-Mora, A., Adamfi, A., Tolnay, P. Comparative Analysis of Balanced State of Weightlifting Results of Male and Female Competitors at the London 2012 Olympic Games. *Sport SPA*, 10(1):63-66