ISyE 6740 – Spring 2021
Project Proposal Final

Team Member Names: Simon Yee - 903563821

Project Title: Are Men That Much Stronger Than Women? An Analysis of Biological Relative Strength and How Different Compound Exercises Relate to Gender.

Problem Statement:

Media has a history of portraying men's natural greater biological physical strength through depictication of men assisting their partners in opening jars, how women are unfit for physical labor, and cultures based around women staying at home, being provided too and needing protection[2]. In fact, some of these differences have led to controversial topics like male and female athletic pay and treatment[5,6]. Is this all true? Do men's greater amount of muscle mass and bone density through naturally higher levels of testosterone reflect in their ability to gain greater strength over women[3]? If physical appearances were ignored, could we differentiate sex through physical performances?

Do men or women's natural distribution in muscle or fat mass contribute to greater physical strength in certain activities (i.e. does a woman's greater distribution of body mass the lower body to a stronger squat?)[4]. The objective is to observe and classify 1. Is it possible to determine an individual's sex through attributes (i.e. weight), and performance metrics (squat, deadlift, and bench press) provided through powerlifting data and 2. Are there any observable relationships between the squat, deadlift and bench press between the genders and exercises.

Data Source:

Data source is provided by Kaggle[8] through OpenPowerLifting[9] as it is. The data set contains 37 columns and over 1.4 million rows, ranging from sex to age, location of competition, and amount of weights lifted in the competition. Since rows are based on competitions, there are possibilities that athletes have competed multiple times, and depending on the structure of the data, it may be treated as an individual data point, or become aggregated. As an example, averaging the individuals best successive attempts lifts in each competition and taking their average weight as well could be a potential method.

There is a "Tested" column that indicates whether competitors were tested for doping in that event. Naturally, doping or any use of drugs plays a big role, as it changes several factors and disrupts the playing field for all competitors involved, and should be a significant consideration in attributes when measuring performance. The difficulty revolves around the controversy of doping in sports, specifically in regards to uses (in non-tested competitions in this case), and avoidance (of being caught) in tested environments[1]. If such issues are addressed in a consistent manner, then it could be considered, otherwise it is beyond the scope of this research. Effectively, during the cleaning process, only those who were disqualified or no show were removed. As tested and non-tested individuals in events could not be completely confidently assessed equally for competitors on this dataset.

Height is another considerable variable that is not available for assessment. Using an oversimplified example, the effort required to move a specific exercise at any weight is considered Work. Work in physics is defined as Work = Force x Distance. Force = Mass x Acceleration, an individual's weight could hypothetically be mass, and their strength will determine the acceleration. Under this simplified example, an individual's height in relation to their limb size will determine Distance. Using the squat as an example, the official International Powerlifting Federation (IPF) rulebook states: "...the lifter must bend the knees and lower the body until the top surface of the legs at the hip joint is lower than the top of the knees... The attempt is deemed to have commenced when the lifters' knees have unlocked."[7] Given that two competitors are equal in all but Distance, in a squatting competition, the individual with longer legs will produce a greater Work to meet the requirement of a successful squat attempt at the same weight.

Methodology:

The first step in the analysis is to observe of the 37 available variables, which are potentially practical and realistically in determining an individuals sex. After reviewing the data documentation and analyzing the data, there are 99,758 rows that are disqualified individuals, 803 doping disqualifications and 9609 no shows. Only disqualified and no shows are dropped due to the amount of missing data, doping individuals are not dropped because there is an inability to assess the uses of illegal substances and is outside the scope of this analysis.

The next step is to determine the usability of the data, to minimize the amount of missing data, and as a last resort, truncate any rows with missing data. There were some trade offs in precision vs. usability that were made in this analysis. As an example, attributes Age and AgeClass provide the age of a competitor, Age provides an exact number and AgeClass provides a range. The issue is that instances of where Age is available and AgeClass is not, is relatively significant, with a difference of 161 and 28,027 (see Fig. 1). This later becomes a nonexisting issue as it is discovered that Age and AgeClass is missing from 49% and 45% of the data, respectively.

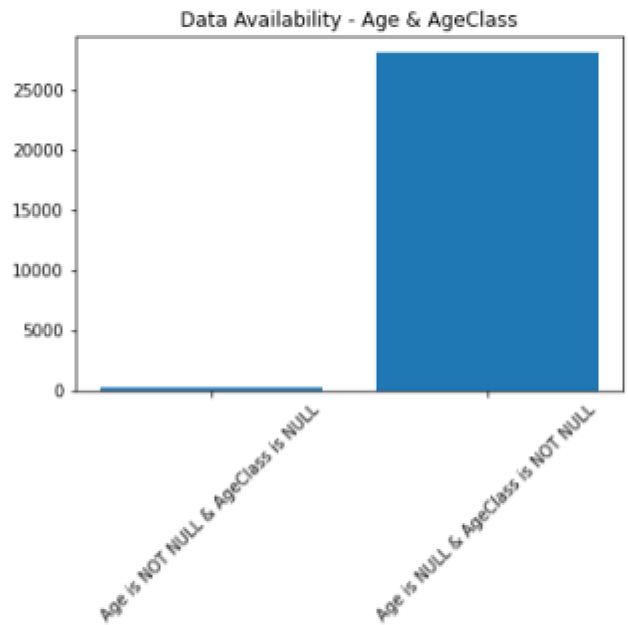

Fig. 1

Similar issues arise for attributes BodyweightKg and WeightClassKg, where BodyweightKg is the bodyweight of the competitors in kilograms, and WeightClassKg specifies a minimum or maximum. Per documentation, 90 would mean "up to (and including) 90kg", and 90+ would indicate ("90kg and above, excluding"). BodyweightKg is the superior choice due to high availability, where only 8000 rows out of the 1.3 million rows are missing, and because WeightClassKg has an unaccountable low precision.

Due to excessive missing data in other predictive columns, the data is aggregated, grouped by names and taking averages, as an attempt to reduce the amount of missing data. The attempt is unsuccessful and rows with missing predictors are ultimately imputed. To avoid introducing too much bias, missing values are not replaced with predicted values, and missing values have no observable patterns that determine why or how they are missing. The final cleaned and aggregated data is scaled, split into a 80% training set and 20% test set, and are used in scikit-learn's random forest classifier and logistic regression model.

Evaluation and Final Results:

Conclusively, both the logistic regression (LR) and random forest (RF) model both perform really well, with an accuracy score of 92% and 94% predicted on the tested data. The LR model's coefficient provides an interesting insight in explaining some of these predictors, where the bench has an coefficient of -52.75 and squat has 24.34 (see Fig. 2). The coefficient by itself does not mean a lot, except that in general, keeping all other predictors constant, as the bench press increases, it increases the likelihood that the gender will be male (0), while squat is the opposite, indicating that it'll likely be female(1). We can further derive more information by taking the exponential of the coefficient. For example, the exponential of the bench coefficient (-52.75) gives an odds ratio of 1.23e-23 or .23. This explains that for every increase in one unit of the bench press in kilograms, there is a 23% increase in odds of the individual being a male. Of all measured predictors, a strong bench press appears to be a strong predictor of the individual being male.

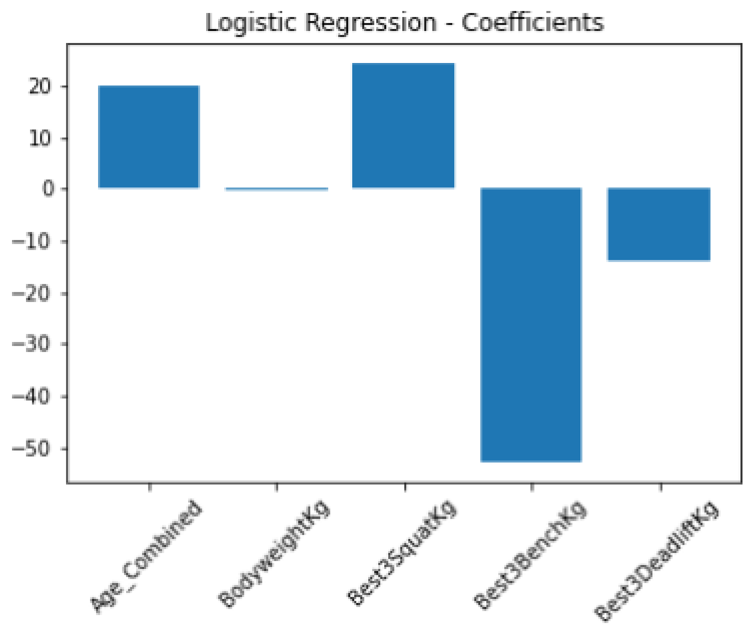

Fig. 2

The RF model also produces an exceptional accuracy score of 94%. Using the Gini coefficient a measurement in variable importance is produced, and the higher the score is, the better the predictor is (see Fig. 3). The results are similar to LR coefficients, where the bench press is a strong predictor.

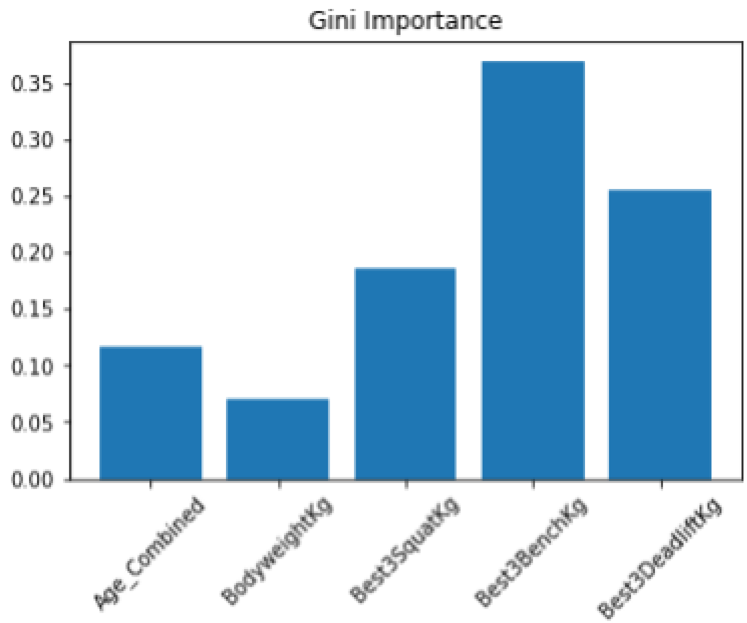

Fig. 3

Before reaching a conclusion, it is important to take at what the data that was cleaned and aggregated looks like:

1. The data contains about 2/3rds male and 1/3rds female. While there are twice the amount of males to females, I do not believe it had any impediment to the accuracy of the model.
2. The average age for females is 28.3 and for males is 27.9. There is not a huge discrepancy between the ages, which leads to differences in strength between ages not being an issue.
3. The average male competitor is ~89.7kg, while the average female is ~68kg.
4. By dividing the average lift by the average lifter weight, and dividing the difference between sex's to determine difference in relativity strength, is it found that on average that the men are stronger than females by about 36% on the squat, 65% on the bench press, and 28%.

If physical appearances are ignored, it is very likely that an individual's sex could determined, and further more men's natural high levels of testosterone are reflective in their ability to train and acquire greater strength. What this difference in strength fails to mention is, despite natural differences, women are still able to acquire tremendous strength in their own ways. It is important for media and culture to forgo the traditional roles, that women are unfit for physical labor, that traditional families should not be the expectation, and that differences in athletic performance should not be a justification to be treated poorly. Not everyone is born equally, everyone is trying their best, and if their best does not exceed others, it does not invalidate their efforts or the individual.

References:

[1] Robert Alexandru Vlad, Gabriel Hancu, Gabriel Cosmin Popescu, Ioana Andreea Lungu
Doping in Sports, a Never-Ending Story?
Adv Pharm Bull. 2018 Nov; 8(4): 529–534. Published online 2018 Nov 29.
doi: 10.15171/apb.2018.062
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6311632/

[2] J R Morrow Jr, W W Hosler
Strength comparisons in untrained men and trained women athletes
Med Sci Sports Exerc. 1981;13(3):194-7.
https://pubmed.ncbi.nlm.nih.gov/7253873/

[3] A E Miller 1, J D MacDougall, M A Tarnopolsky, D G Sale
Gender differences in strength and muscle fiber characteristics
Eur J Appl Physiol Occup Physiol. 1993;66(3):254-62.
https://pubmed.ncbi.nlm.nih.gov/8477683/

[4] Melanie Schorr, Laura E. Dichtel, Anu V. Gerweck, Ruben D. Valera, Martin Torriani, Karen K. Miller & Miriam A. Bredella
Sex differences in body composition and association with cardiometabolic risk
https://bsd.biomedcentral.com/articles/10.1186/s13293-018-0189-3

[5] 2021 Women's NCAA Tournament: NCAA says women got smaller gym than men because there wasn't enough space
Gabriel Fernandez Mar 19
https://www.cbssports.com/college-basketball/news/2021-womens-ncaa-tournament-ncaa-says-women-got-smaller-gym-than-men-because-there-wasnt-enough-space/

[6] Sue Bird chats economics of women's hoops with economist Steven Levitt
Ariail
https://www.swishappeal.com/wnba/2021/1/26/22248717/wnba-sue-bird-podcast-people-mostly-admire-steven-levitt-economics-womens-basketball-pay-salary-cba

[7] IPF. “IPF Technical Rules Book 2021.” International Powerlifting Federation, Jan. 2019,
www.powerlifting.sport/fileadmin/ipf/data/rules/technical-rules/english/IPF_Technical_Rules_Book_2021docx.pdf.

[8] https://www.kaggle.com/open-powerlifting/powerlifting-database

[9] https://www.openpowerlifting.org/