Introduction

Evaluating competitive performance in powerlifting is important for a competitor to improve his overall strength. This report plans to investigate what elements might be associated with the overall strength development for a powerlifter. The data used in this report contains 4900 independent records of 17 variables related to competitive powerlifters. The features of the dataset used in this report are detailed in the table below.

Variable Name	Data Type	Description
ID		Participant ID
Sex	Categorical (2 levels)	Sex: M=Male, F=Female
AverageTime	Assume continuous	Average time per training session
Equipment	Categorical (4 levels)	Hand covering type
Age	Assume continuous	Age (years)
Schedule	Categorical (2 levels)	Whether some trains in the morning or night
LiquidConsumed	Assume continuous	Average liquid consumed per training session
GymCost	Assume continuous	Monthly amount paid for all gym costs
BodyweightKg	Assume continuous	Bodyweight in kilograms
BestSquatKg	Assume continuous	A competitor's best squat in kilograms
BestBenchKg	Assume continuous	A competitor's best bench press in kilograms
BestDeadliftKg	Assume continuous	A competitor's best deadlift in kilograms
TotalKg	Assume continuous	The total amount for a competitor's combined squat, bench and deadlift in kilograms
Wilks	Assume continuous	Wilks score (strength adjusted for body mass)
Winner	Categorical (2 levels)	Whether a competitor won their last event
Displacement	Assume continuous	Distance during an additional challenge that a competitor was able to push a weighted object from its starting location
Group	Categorical (10 levels)	Which group of judges a competitor was assigned (ranges from A-J)

In this paper, questions like whether the total amount a competitor lifts is associated with the competitor's sex or their choice of equipment will be researched and gave interpretations. Additionally, several models will be built to select a subset of predictors which are better at predicting the Wilks score or explaining whether a competitor won their last event.

Exploratory Data Analysis

For this report, we work with 4900 independent records of 17 variables related to competitive powerlifters. Of these variables, 5 out of 17 are categorical and the rest are continuous except for the variable 'ID'. The below **Table 1.1** shows the basic descriptive statistics, including number of observations with non-missing values, mean, minimum and maximum which helps us get a basic description for the continuous variables in the dataset.

Descriptive Statistics for Continuous Variables

Variable	N	Mean	Std Dev	Minimum	Maximum
AverageTime	4900	54.5002041	20.7597431	18.0000000	90.0000000
Age	1832	31.8084061	12.9877321	9.5000000	84.5000000
LiquidConsumed	4900	2106.09	1088.37	200.0000000	4000.00
GymCost	4900	33.1867367	15.6885170	6.0100000	60.0000000
BodyweightKg	4875	86.9895881	23.1940769	25.4000000	197.6300000
BestSquatKg	3778	175.3020487	67.9675116	-375.0000000	444.5200000
BestBenchKg	4543	118.2484768	53.9556623	-327.5000000	381.0200000
BestDeadliftKg	4037	194.2284989	60.3751090	-327.5000000	410.0000000
TotalKg	4604	423.1623132	194.1788582	27.2200000	1077.50
Displacement	4594	1.9028938	327.9368216	-1005.73	923.6086849
Wilks	4594	300.6032490	114.6128358	21.5200000	629.3500000

Table 1.1 Descriptive statistics for continuous variables

In general, competitor tries to lift the maximum weight as possible, therefore it is of interest to investigate what potential predictors might be correlated with the total amount a competitor lift. In other words, we will generate an initial impression for the association between the continuous variable 'TotalKg' and some categorical variables such as the competitor's sex or their choice of equipment.

The summary statistic for the variable 'Sex' and 'Equipment' are derived below in the **Table 1.2** and **Table 1.3**.

Summary Statistic for Equipment Variable

Number of Variable Levels	
Variable	Levels
Equipment	4

Equipment	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Multi-ply	110	2.24	110	2.24
Raw	2331	47.57	2441	49.82
Single-ply	1894	38.65	4335	88.47
Wraps	565	11.53	4900	100.00

Table 1.2 Summary Statistics for 'Equipment' variable

Summary Statistic for Sex Variable

Number of Variable Levels	
Variable	Levels
Sex	2

Sex	Frequency	Percent	Cumulative Frequency	Cumulative Percent
F	1127	23.00	1127	23.00
M	3773	77.00	4900	100.00

Table 1.3 Summary Statistics for 'Sex' variable

A Scatter Plot Matrix is generated for the dataset as follows in **Table 1.4**. The goal here is to investigate any possible associations between variables in the dataset.

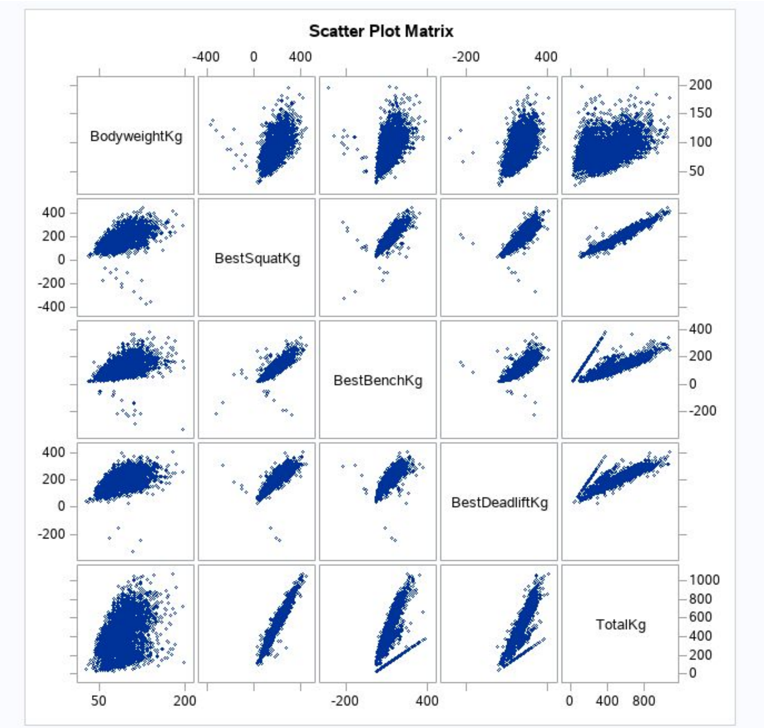

Table 1.4 Scatter Plot Matrix for continuous variables

Based on the visual output in **Table 1.4**, we can discover a high correlation between some of the continuous variables in the dataset.

In the following section, the association between the total amount a competitor lift with their sex or choice of equipment will be researched more deeply and clearly explained. In addition, considering 'Wilks' and 'Winner' as the response variable respectively, several models will be built and procedures like data preparation, data splitting as well as model selection would be conducted later on for the enhancement of model performance.

Formal Analysis

Questions of Interest

1. Association between the mean total amount a competitor lift and sex

In the exploratory analysis section, we have already got an initial impression between variables in the dataset. A two sample t-test will be performed here to investigate whether the average total amount a competitor lift is different based on their sex.

To ensure our conclusions are valid, we will proceed to check the modeling assumptions first. By looking at our study design, the dataset contains 4900 independent records of variables, therefore the assumption of independence between observations appears reasonable.
Next, we can check the Normality assumption using **Table 2.1** and **Table 2.2**.

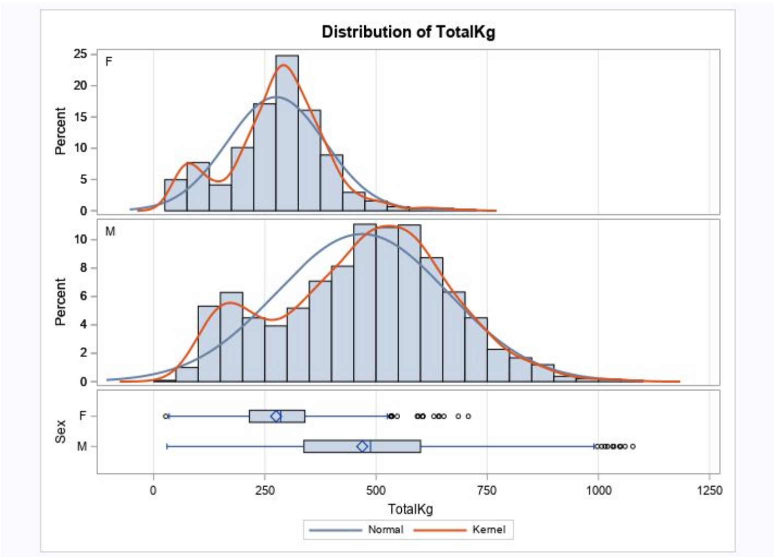

Table 2.1 Histogram of TotalKg for Sex group

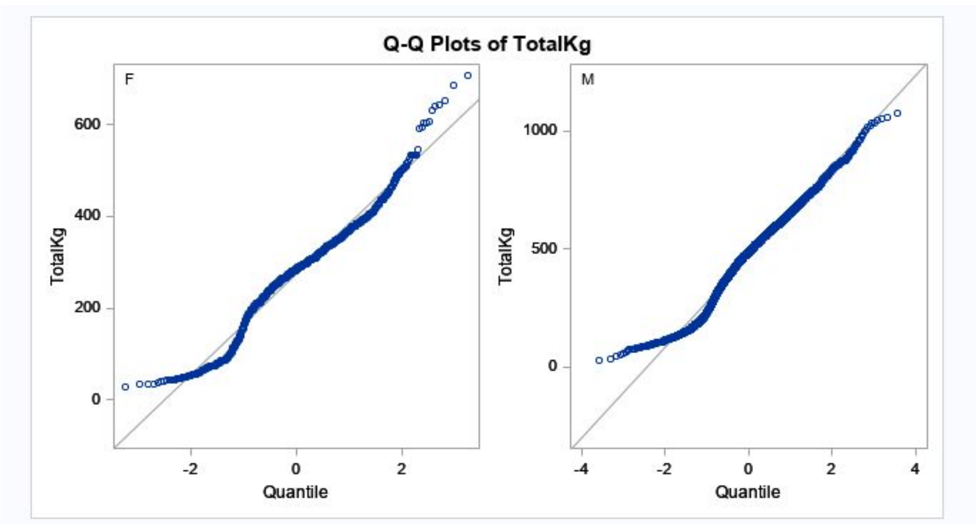

Table 2.2 Normal QQ plots of TotalKg for different Sex group

By looking at the **Table 2.1** and **Table 2.2**, for each sex group, only half of the QQ plot points lie on the straight diagonal line and there are some fanning out points at the initials, so the assumption of normality seems doubtable.
To resolve the non-normality problem, a log transformation was applied here of the 'TotalKg' variable in **Table 2.3**.

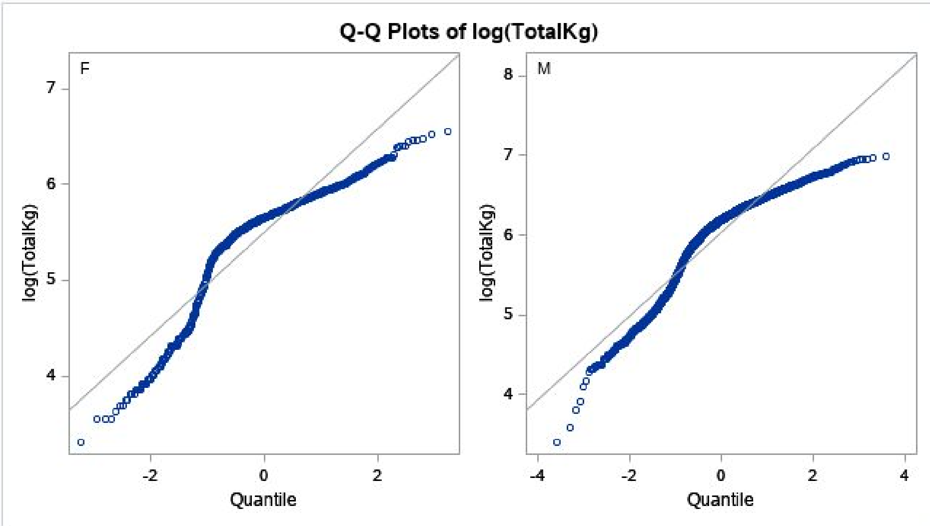

Table 2.3 Normal QQ Plot for log of TotalKg for different Sex group

By **Table 2.3**, it seems the log transformation does not help too much. Back to the problem, the assumption of normality is ambiguous and requires further exploration. Some resolution could be Box-Cox Transformation.

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	3515	1087	3.06	<.0001

Table 2.4 Equality of Variances Table

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	4602	-31.74	<.0001
Satterthwaite	Unequal	3224	-41.74	<.0001

Table 2.5 Test Statistic for Pooled and Satterthwaite Method

In order to see whether the equality of variance assumption is valid, we look at **Table 2.4**. With a significance level of 5%, we reject the null hypothesis that the variance is the same for both sex groups. In other words, we proceed under the assumption that they are not equal and will refer to the outcome of the Satterthwaite method (**Table 2.5**).

Two Sample t-test comparing Sex(F/M)

Variable: TotalKg

Sex	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
F		1088	275.3	109.7	3.3258	27.2200	707.5
M		3516	468.9	191.8	3.2347	30.0000	1077.5
Diff (1-2)	Pooled		-193.7	175.9	6.1024		
Diff (1-2)	Satterthwaite		-193.7		4.6394		

Sex	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
F		275.3	268.7	281.8	109.7	105.3	114.5
M		468.9	462.6	475.3	191.8	187.4	196.4
Diff (1-2)	Pooled	-193.7	-205.6	-181.7	175.9	172.4	179.6
Diff (1-2)	Satterthwaite	-193.7	-202.8	-184.6			

Table 2.6 Summary statistics for Two Sample t-test

Results by two different methods (Pooled and Satterthwaite) were generated in the above **Table 2.6**. The Pooled method compares the means for populations with the same variance for each group, whereas the Satterthwaite method compares the means for populations with unequal variances.

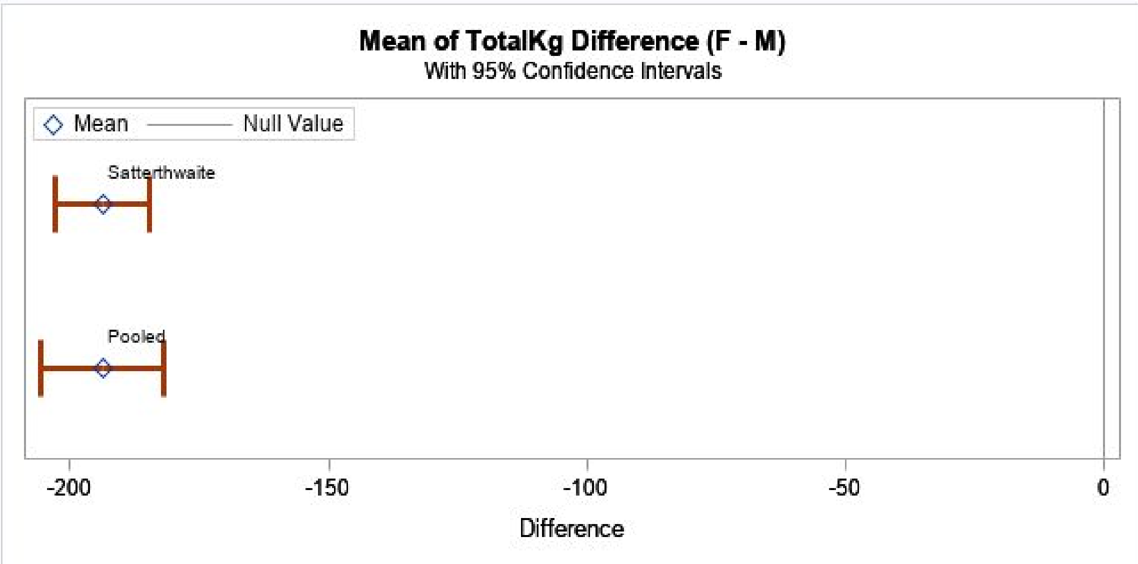

Table 2.7 Mean of TotalKg Difference and Confidence Interval Plot

Recall from **Table 2.5**, both methods reject the null hypothesis, so we can conclude that there is a significant difference between the mean total amount a competitor lift and their sex.

By looking at **Table 2.7** the difference and confidence Interval plot, we can discover that the total amount a competitor lifts is higher for males than for females. Since we rejected the assumption of equal variances, we would use the confidence interval for the Satterthwaite method to make the conclusion that the mean difference of the total amount a competitor lift for females and males is highly likely to lie between -180 and -205 kilograms.

2. Association between the average total amount a competitor lift and choice of Equipment

Aside from sex, it is quite possible that the total amount a competitor lift is associated with their choice of equipment. A one-way analysis of variance will be performed here and the goal is to check whether the average total amount a competitor lift is different based on their choice of equipment.

We proceed to check the modelling assumptions first. By checking the study design, each observation corresponds to a different competitor with different types of Equipment, hence the assumption of independence seems to hold.

Table 3.1 displays different diagnostics plots for the variable 'TotalKg'.

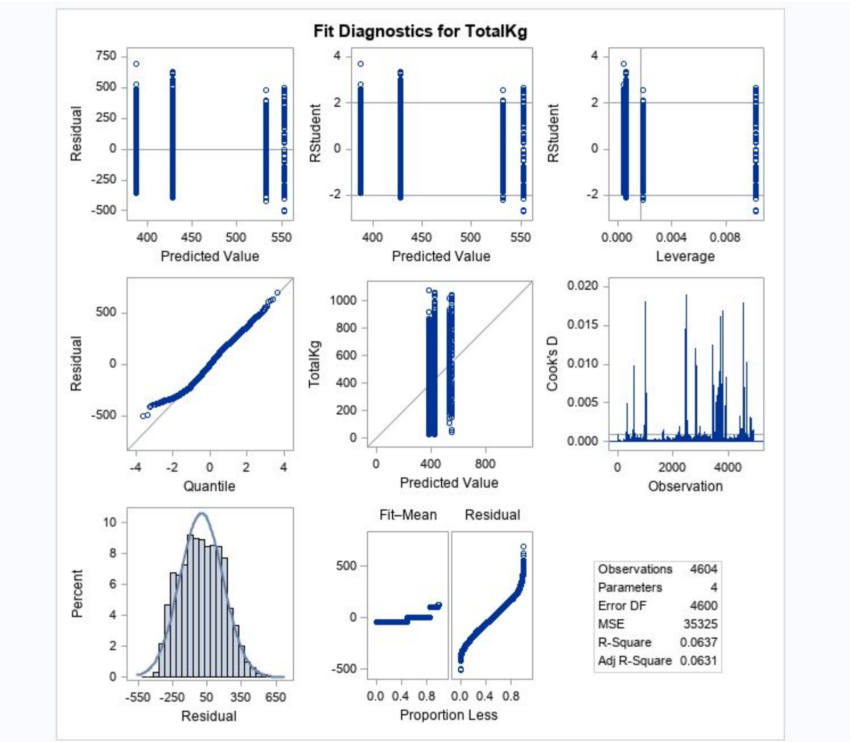

Table 3.1 Diagnostics plots for the 'TotalKg' variable

From the above visual output, we can see that the histogram shows a bell-shaped distribution and most of the points on the QQ plot lie on the diagonal line except for a small deviation at the initials, therefore the assumption of Normality seems to hold.
In order to see whether the equality of variance assumption is valid, we look at **Table 3.2**.

Levene's Test for Homogeneity of TotalKg Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Equipment	3	1.616E11	5.388E10	31.24	<.0001
Error	4600	7.933E12	1.7245E9		

Table 3.2 Levene's Test for Homogeneity of TotalKg Variance

With a significance level of 5%, we reject the null hypothesis of equal variance. Instead, we will use the Welch's variance weighted one-way ANOVA (**Table 3.3**). By **Table 3.3**, we proceed under the Welch's correction and shall conclude that the difference is considered to be statistically significant.

Welch's ANOVA for TotalKg			
Source	DF	F Value	Pr > F
Equipment	3.0000	106.28	<.0001
Error	422.0		

Table 3.3 Welch's variance-weighted one-way ANOVA

Table 3.4 displays the test statistic for one-way ANOVA with Equipment as explanatory variable, and the interpretation is carried on as follows.

One-way Anova with Equipment as Explanatory

The GLM Procedure

Dependent Variable: TotalKg

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	11062986.4	3687662.1	104.39	<.0001
Error	4600	162495103.1	35325.0		
Corrected Total	4603	173558089.5			

R-Square	Coeff Var	Root MSE	TotalKg Mean
0.063742	44.41547	187.9495	423.1623

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Equipment	3	11062986.36	3687662.12	104.39	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Equipment	3	11062986.36	3687662.12	104.39	<.0001

Table 3.4 Test Statistic One-way ANOVA with Equipment as Explanatory

With a significance level of 5%, we reject the null hypothesis that the average total amount a competitor lifts is the same across different equipment categories ($F=104.39$, $p < 0.0001$).

By the statistically significant ANOVA, we will perform a follow-up analysis (Post-Hoc Analysis) to explore which specific group differs. To avoid the Type I error, Tukey's Honestly Significant Difference (HSD) was applied here to adjust the p-values for multiple comparisons.

Post-Hoc Analysis of ANOVA-Equipment as Explanatory

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

Equipment	TotalKg LSMEAN	LSMEAN Number
Multi-ply	552.892143	1
Raw	387.178689	2
Single-ply	428.480315	3
Wraps	532.266697	4

Least Squares Means for effect Equipment
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: TotalKg

i/j	1	2	3	4
1		<.0001	<.0001	0.7495
2	<.0001		<.0001	<.0001
3	<.0001	<.0001		<.0001
4	0.7495	<.0001	<.0001	

Table 3.5 Post-Hoc Analysis of ANOVA - Equipment as Explanatory

Average total amount a competitor lift across different equipment categories were displayed in the above visual output. For the first picture in **Table 3.5**, "LSMEAN Number" gives the index for SAS output (e.g. Equipment "Multi-ply" is index 1). For the second picture in **Table 3.5**, p-values for each pair of indexed levels is produced.

By **Table 3.5**, we can see that the mean total amount a competitor lift is statistically different between all equipment categories, with the exception of "Multi-ply" and "Wraps".

The difference plot in below **Table 3.6** also shows us these pairwise differences.
In the **Table 3.6**, the solid line refers to the confidence interval's width for the difference between each mean equipment group. If the solid line touches the 45 degree dashed line, it indicates an interval that contains 0 which means the difference between the group means is not significant. By looking at **Table 3.6**, we can discover that the solid line for difference of "Multi-ply" and "Wraps" group touches the 45 degree dashed line, therefore the difference between the group means of "Multi-ply" and "Wraps" is not significant.

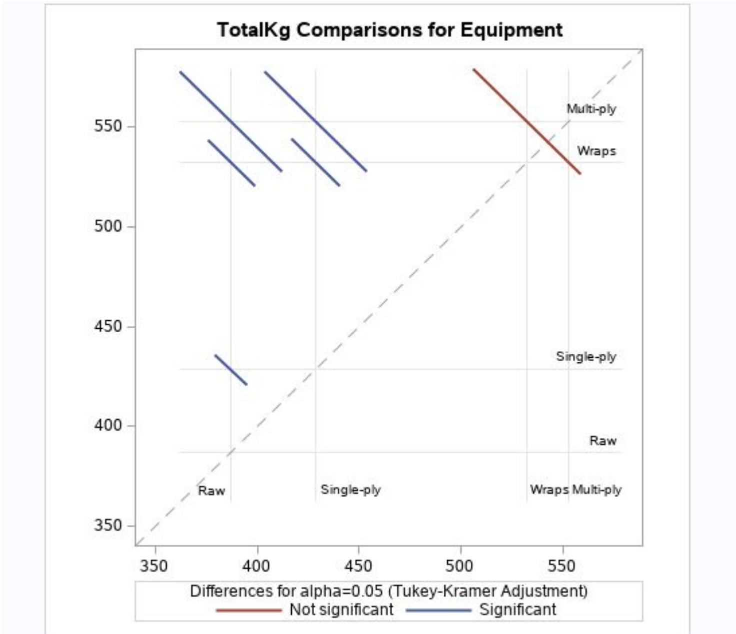

Table 3.6 Difference plot of TotalKg comparisons for Equipment

Based on our investigations on post-hoc analysis, we successfully navigate where the difference is and conclude that the overall test is statistically significant. The mean total amount a competitor lifts is statistically different between all equipment categories, with the exception of "Multi-ply" and "Wraps" group.

3. Fitting a Model for variable Wilks

To select a subset of predictors that better explain the variable wilks score, a forward selection will be applied here to decide which potential predictors may have a major influence in our model. We will use Wilks as the response variable and all other variables as potential predictors except for the variable 'ID' and 'Winner'.

Forward Selection Summary				
Step	Effect Entered	Number Effects In	Number Parms In	AIC
0	Intercept	1	1	12063.7684
1	BestSquatKg	2	2	10997.9570
2	BodyweightKg	3	3	10233.1023
3	Sex	4	4	9746.4479
4	TotalKg	5	5	8966.3068
5	Equipment	6	8	8960.1445
6	AverageTime	7	9	8958.0486
7	GymCost	8	10	8956.5999*
* Optimal Value of Criterion				

Table 4.1 Forward Selection Summary

Table 4.1 displays the AIC value for each step that we include another variable. In our example, the selection process stopped at the eighth step.

Forward Selection Model (7 Independent variables) : Intercept BestSquatKg BodyweightKg Sex TotalKg Equipment Averagetime GymCost

To assess how the model selection process performed, a number of plots was obtained.

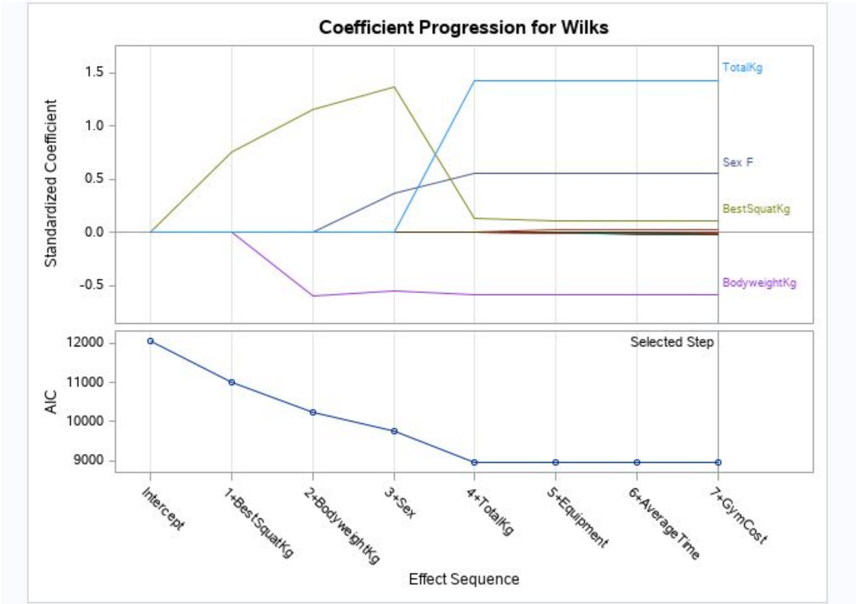

Table 4.2 Coefficients for Wilks at each iteration of the model selection process

By **Table 4.2**, we can discover that the coefficients do not behave erratically when each variable was included in the model.

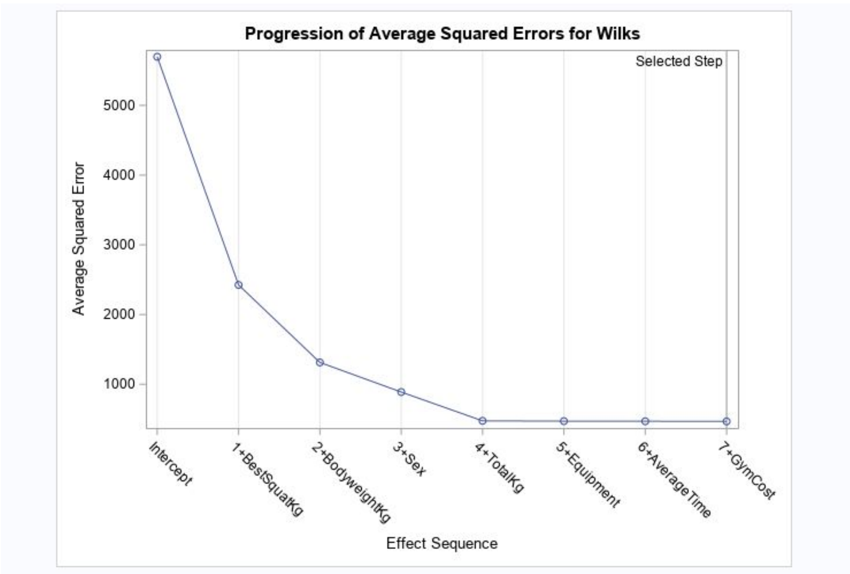

Table 4.4 Progression of Average Squared Errors for Wilks

By **Table 4.4**, we can discover that the averaged squared error is lowest at the final iteration step. In general, models with lower error will be preferred. Therefore, the forward selection includes: Intercept, BestSquatKg, BodyweightKg, Sex, TotalKg, Equipment, AverageTime and GymCost in the model.

Forward Selection guides us towards better modelling. Next, we would use the predictors obtained in the forward selection and performed a two-way ANOVA model to measure the association between Wilks and those selected predictors in the dataset.

Source	DF	Type III SS	Mean Square	F Value	Pr > F
BestSquatKg	1	105215.262	105215.262	90.32	<.0001
BodyweightKg	1	4208909.650	4208909.650	3612.98	<.0001
TotalKg	1	1012585.938	1012585.938	869.22	<.0001
AverageTime	1	6.708	6.708	0.01	0.9395
GymCost	1	193.175	193.175	0.17	0.6839

Table 4.5 Type III Sum of Squares

By **Table 4.5**, the sum of squares output shows the addition to the sum of squares when adding an effect into the model. For a significance level of 5%, we can see that the Wilks score is significantly different depending on the competitor's best squat, the competitor's bodyweight and the total amount the competitor lifts.

Before making any conclusions, we would go on checking the modelling assumptions first.

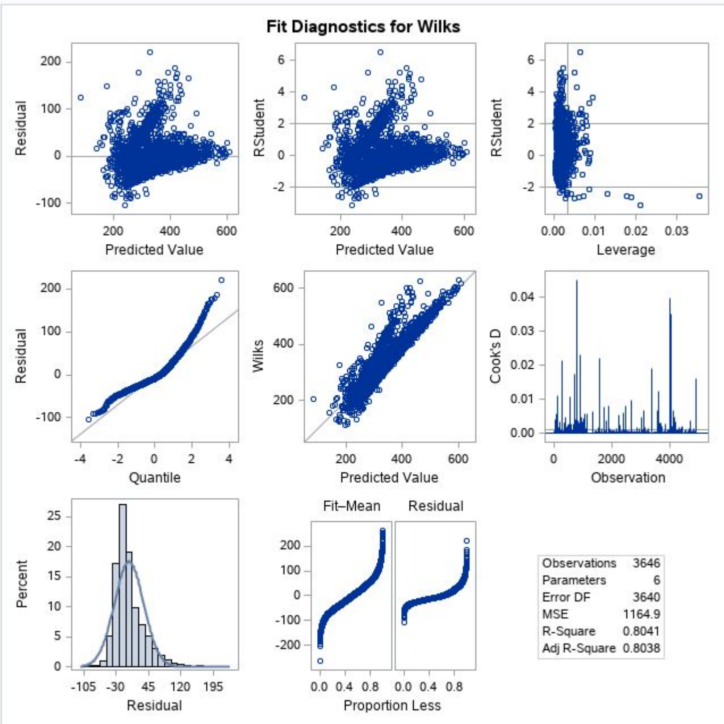

Table 4.6 Diagnostics Plots for the Wilks variable

By looking at **Table 4.6**, points in the Normal QQ plot does not lie on the diagonal line and there are points fanning out greatly, therefore the assumption of normality does not hold. From the (Residual vs Fitted Values Plot), some pattern appears which means the assumption of constant variance also violates.

Dealing with non-normality and non-constant variance, some resolutions could be:

1. Try a transformation
2. Try a non-parametric test
3. Detect the existence of any outlier

Here, we will transform the corresponding variables into log form to resolve the problem.

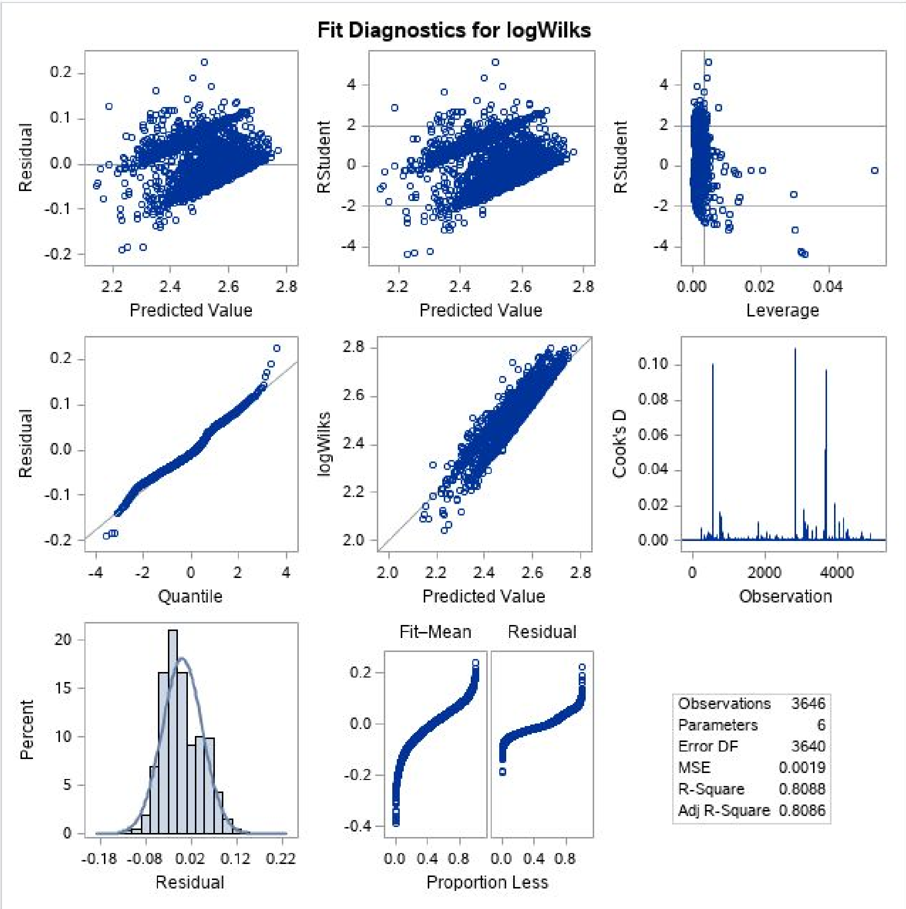

Table 4.7 Diagnostics Plots for log of Wilks variable

By **Table 4.7**, problems of violation of assumptions seem to improve. The histogram shows a bell-shaped distribution and most of the Normal QQ plot points lie on the straight diagonal line, therefore the assumption of normality seems to hold. By looking at the (Residual vs Fitted Values Plot), the error variance is not constant over time where the fitted value increases, the residual decreases, so the assumption of constant variance still does not hold. Some resolutions could be add a squared term to the model or try the Box-Cox transformation instead of log transformation. Clearly, more work should be done here to resolve the problem of non-constant variance.

4. Fitting a Model for Categorical Variable 'Winner'

In addition to Wilks Score, we are also interested in the subset of predictors which are better explaining at whether a competitor won their last event.
Before we start fitting the model for the variable Winner, it is important to be aware of some issues present in our datasets.

Clustering Categorical Input Levels

For this example, quasi-complete separation problem arises where the categorical variables having too many levels. One resolution for dealing with categorical variables with too many levels is to combine the different levels.

Notice that the 'Group' variable has up to 10 levels, therefore we will perform Greenacre's method to evaluate which levels could be clustered together for the 'Group' variable.

Cluster History						
Number of Clusters	Clusters Joined		Freq	Semipartial R-Square	R-Square	Tie
8	G	I	178	0.0000	1.00	
7	H	J	2344	0.0006	.999	
6	C	CL8	1060	0.0007	.999	
5	E	CL7	2414	0.0011	.998	
4	CL5	F	2508	0.0106	.987	
3	A	B	1142	0.0276	.959	
2	CL6	CL4	3568	0.0936	.866	
1	CL3	CL2	4710	0.8659	.000	

Table 5.1 Cluster History Table for 'Group' variable

Based on the **Table 5.1**, we can see the reduction in R-squared value for each of the clusters. A total of eight different cluster levels were presented as well as the dendrogram in **Table 5.2** which helps for better visual interpretation.

Based on the below **Table 5.2**, we will cluster level "A" and "B" together which lead to an R-squared reduction of only 2.7%. Meanwhile, the rest levels could be clustered together because we still retain 86.6% of the original chi-squared test statistic value.

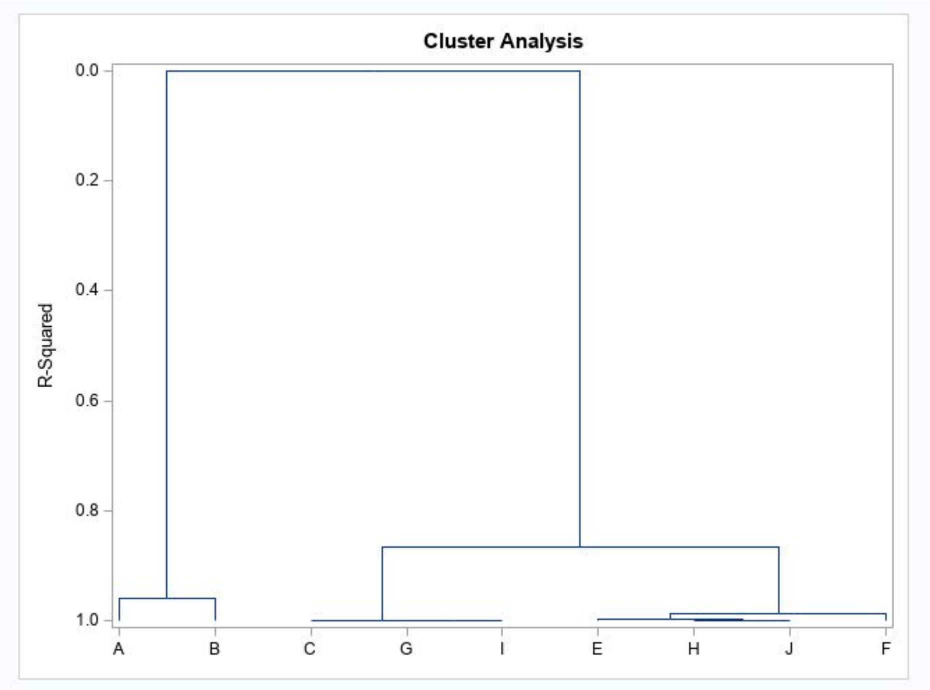

Table 5.2 Dendrogram of Clusters vs R-squared values for 'Group' variable

Model Selection

After we prepared the dataset, a logistic regression is modelled to explore the association between 'Winner' and other potential predictors. We will use 'Winner' as the response variable and all other variables as potential predictors except for variable 'ID' and 'Wilks'. Model selection in proc logistic have been performed here for convergence. Otherwise, the model convergence will not be satisfied due to the singular information matrix. Based on the forward selection in proc logistic, we get a final model with 3 predictors.

Forward Selection Model (3 independent variables) :
Intercept Age Group_Combined Equipment
* (Group_Combined is newly created based on the assigned cluster of levels for 'Group')

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Table 6.1 Model Convergence Status

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1873.106	1786.439
SC	1878.315	1817.690
-2 Log L	1871.106	1774.439

Table 6.2 Model Fit Statistics

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	96.6675	5	<.0001
Score	92.7240	5	<.0001
Wald	84.2130	5	<.0001

Table 6.3 Likelihood Ratio, Score and Wald Testing results

By looking at **Table 6.1**, we reached the convergence status, so we will carry on the interpretation of results. By **Table 6.2** and **Table 6.3**, we discover that the model has an AIC=1873.106 and is valid for the Likelihood Ratio Test, Score Test based on the corresponding p-value.

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.1391	0.2278	25.0030	<.0001
Age		1	0.0251	0.00493	25.9851	<.0001
Equipment	Multi-ply	1	0.2768	0.3841	0.5195	0.4711
Equipment	Raw	1	-0.0993	0.1478	0.4515	0.5016
Equipment	Single-ply	1	-0.4499	0.1773	6.4427	0.0111
Group_Combined	level1	1	-0.6488	0.1052	38.0362	<.0001

Table 6.4 Analysis of Maximum Likelihood Estimates

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	63.2	Somers' D	0.271
Percent Discordant	36.1	Gamma	0.273
Percent Tied	0.7	Tau-a	0.135
Pairs	455700	c	0.636

Table 6.5 Predicted Probabilities and Observed Responses

Now we will look at **Table 6.4** and **Table 6.5**, with a significance level of 5%, we can conclude that there is a significant effect of the Age, Equipment (Single-ply) and Group (level1) on the odds of whether a competitor won their last event. We also get some predicted probabilities metrics as well as odds ratio estimates and confidence intervals.

Odds Ratio Estimates and Profile-Likelihood Confidence Intervals				
Effect	Unit	Estimate	95% Confidence Limits	
Age	1.0000	1.025	1.016	1.036
Equipment Multi-ply vs Wraps	1.0000	1.004	0.372	2.990
Equipment Raw vs Wraps	1.0000	0.690	0.520	0.912
Equipment Single-ply vs Wraps	1.0000	0.486	0.327	0.718
Group_Combined level1 vs level2	1.0000	0.273	0.179	0.408

Table 6.6 Odds Ratio Estimates and Profile-Likelihood Confidence Intervals

Table 6.6 displays how the predictors of our model are relevant. If the interval cross 1.00, it means that the corresponding levels are not relevant. In our case, for the 'Equipment' variable, odds of Multi-ply vs Wraps indicates that there is not significant difference for that ratio. The odds ratio plot graphically shows in **Table 6.7** helps for visual illustration.

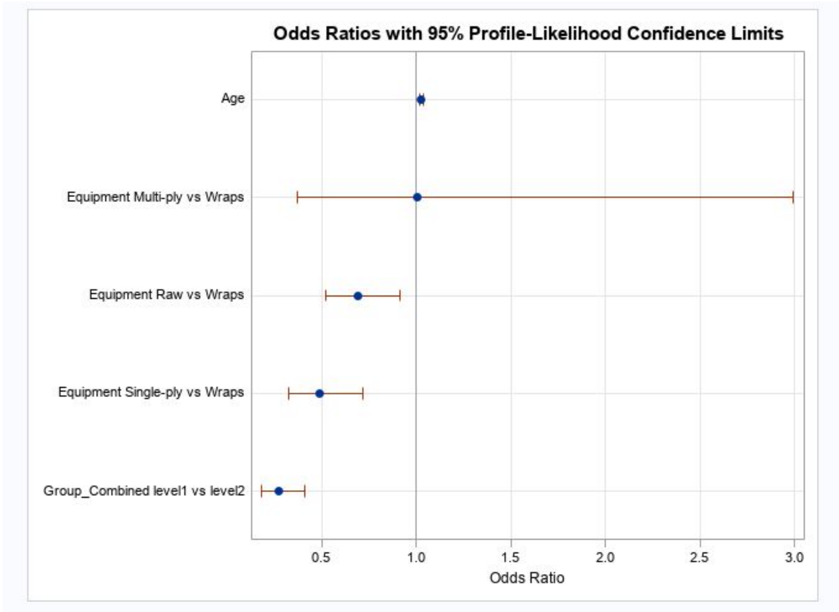

Table6.7 Odds Ratios with 95% Confidence Limits

Finally, we can also compare the predicted probabilities for won the last event based on the different levels of our explanatory variables.

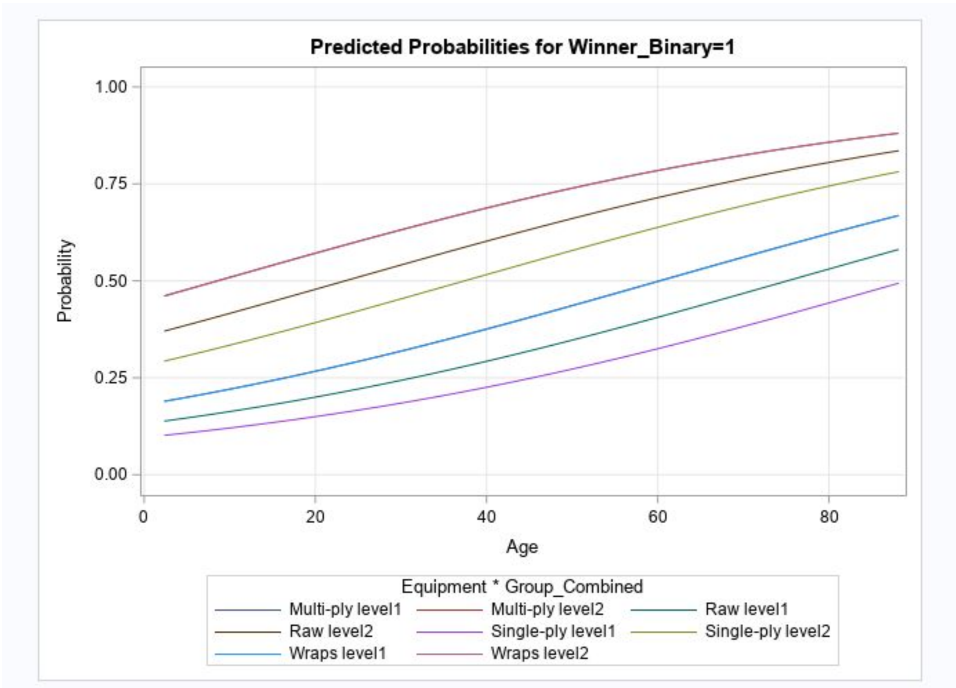

Table 6.8 Predicted Probabilities for variable Winner_Binary
(*Winner_Binary is a newly created binary version of variable 'Winner')

5. Model Comparison (Out-of-Sample Performance)

It is of interest to see whether a single variable can be used to make better predictions about whether a competitor won their last event. In this section, we will consider a model that only uses the variable 'Wilks' as predictor. We will compare this simpler model with the fitted logistic model we obtained in previous section to measure the out-of-sample predictive performance.

Data Splitting

When assessing the model performance, data splitting enables us to evaluate the model and have an idea of how the model would perform on unseen data. Since our model has many variables, to avoid the problem of overfitting, we will split the data into training data (70% of the original dataset) and test data (30% of the original dataset) by variable 'Winner'.

Frequency in Training Dataset

Winner	Frequency	Percent	Cumulative Frequency	Cumulative Percent
N	1484	46.22	1484	46.22
Y	1727	53.78	3211	100.00
Frequency Missing = 220				

Table 7.1 Frequencies in Training Dataset

Frequency in Testing Dataset

Winner	Frequency	Percent	Cumulative Frequency	Cumulative Percent
N	635	46.18	635	46.18
Y	740	53.82	1375	100.00
Frequency Missing = 94				

Table 7.2 Frequencies in Testing Dataset

The above **Table 7.1** and **Table 7.2** helps us make sure that there are roughly the same proportion of Winner events in both training and testing dataset.

After successfully splitting the original data into training and testing dataset, the simpler model is fitted which is a logistic regression model using 'Winner' as response variable and 'Wilks' as the only explanatory variable in the training dataset.

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Table 7.3 Model Convergence Table

By looking at **Table 7.3**, the simpler model have reached the convergence status.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	6320.737	6265.734
SC	6327.165	6278.591
-2 Log L	6318.737	6261.734

Table 7.4 Model Fit Statistics

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	57.0029	1	<.0001
Score	56.5711	1	<.0001
Wald	55.9425	1	<.0001

Table 7.5 Likelihood Ratio, Score and Wald Testing results

By **Table 7.4** and **Table 7.5**, we discover that the simpler model has an AIC=6265.734 and is valid for the Likelihood Ratio Test, Score Test based on the corresponding p-value.

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.7479	0.0854	76.6857	<.0001
Wilks	1	-0.00197	0.000264	55.9425	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Wilks	0.998	0.998	0.999

Table 7.6 Analysis of Maximum Likelihood Estimates and Odds Ratio Estimates

Now we will look at **Table 7.6**, with a significance level of 5%, we can conclude that there is a significant effect of the Wilks Score on the odds of whether a competitor won their last event. The Odds Ratio Estimates table also indicates the significance information.

For both complex and simpler model, we have derived their statistics information and give interpretations. Next, we will compare the model predictive performance using ROC curves.

ROC Association Statistics							
	Mann-Whitney						
ROC Model	Area	Standard Error	95% Wald Confidence Limits		Somers' D	Gamma	Tau-a
Complex Model	0.6528	0.0237	0.6064	0.6992	0.3057	0.3078	0.1510
Simple Model	0.5682	0.0250	0.5191	0.6172	0.1363	0.1363	0.0673

Table 7.7 ROC Association Statistics

ROC Contrast Test Results			
Contrast	DF	Chi-Square	Pr > ChiSq
Comparing Models	1	0.1877	0.6648

Table 7.8 ROC Contrast Test Results

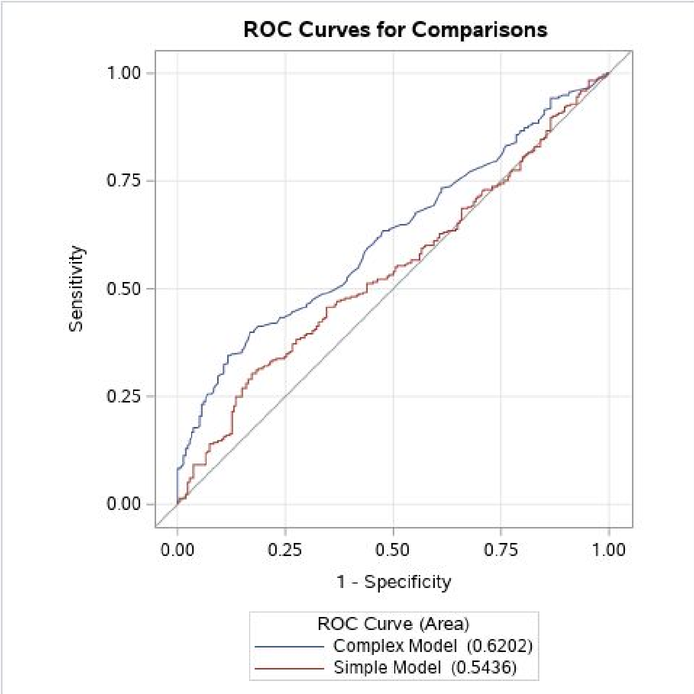

Table 7.9 ROC Curves Comparison

By examining the **Table 7.8** and **Table 7.9**, the complex model returns an AUROC value of 0.62, and the simpler model returns an AUROC value of 0.54. The chi-square test reject the null hypothesis that the two ROC curves are the same. Therefore, we would conclude that the complex model has a better performance than the simpler one.

Conclusion

By implementing several analyses, we figured out that there is a significant association between the average total amount a competitor lift and their sex or choice of equipment. Moreover, we established different models for better prediction of Wilks Score and whether the competitor won their last event. For the prediction of Wilks Score, further analysis such as data transformation should be applied to overcome the problem of non-constant variance. For the prediction of 'Winner' as response variable, we obtained a relatively good model to make predictions about whether a competitor won their last event. However, more improvements such as dealing with missing values could be done to enhance its predictive performance.