A Statistical Analysis of Bowling Balls

Daniel Magro (484497M)

Bachelor of Science in Information Technology (Honours) (Artificial Intelligence)

SOR1231 - Hypothesis Testing and Modelling using SPSS

Lecturer: Mark A. Caruana

1. Introduction 3
2. Aims and Objectives 5
3. Descriptive Statistics & Illustrations 6
3.1 Descriptive Statistics 6
3.1.1 Measures of Location 6
3.1.2 Measures of Dispersion 6
3.2 Illustrations 8
4. Parametric/Non-Parametric Tests 12
4.1 Tests of Normality 12
4.2 Kruskal-Wallis Test 13
4.2.1 Does the Brand affect the PerfectScale of a Bowling Ball? 13
4.2.2 Does Performance affect the Price of a Bowling Ball? 13
4.3 Chi-Squared Test 15
4.3.1 Does The Finish affect the intended Lane Condition of a Bowling Ball? 15
4.4 Pearson Correlation 16
5. Regression 18
5.1 Simple Linear Regression 18
5.1.1 SLR of PerfectScale vs Year 18
5.2 Multiple Linear Regression 22
5.2.1 MLR of Price with all correlated covariates 22
6. General Linear Model 27
6.1 5-Way ANOVA 27
6.2 ANCOVA 34
7. Conclusion 46
8. Appendix – The Data Set 48
9. References 50

1. Introduction

1.1 Why this data set was chosen

This data set was chosen because of my interest in Ten Pin Bowling, I also believe it would be very interesting to show how the many features and specifications of different bowling balls affect the way it performs, i.e. its potential to hook.

The data set was compiled using the specifications page of 47 different bowling balls from http://www.bowlingball.com/shop/all/bowling-balls/ . This data set only includes data regarding 'strike balls', 'spare balls' were not included with the data since they are usually all identical, regardless of Brand, no interesting results would have been obtained and because they would have skewed the data.

1.2 Variables

The Fixed Factors used were the Brand, Finish, Lane Condition, Performance and Core Type.

The Covariates used were the Differential, Mass Bias Diff, RG and Year

The Dependent Variables were the Price and Perfectscale® (Perfectscale® measurement was used with permission from http://www.bowlingball.com/).

Brand – Fixed Factor – The Brand of the bowling ball – Ranges from 1-7, Selected brands were DV8, Columbia 300, Hammer, Brunswick, Ebonite, Storm and Track.

Finish – Fixed Factor – The finish of the bowling ball's surface/coverstock – Ranges from 1-4: High Polish, Polished, Matte and Sanded.

Lane Condition – Fixed Factor – The amount of oil on the lane the ball is designed for – Ranges from 1-5: Dry, MediumDry, Medium, MediumHeavy and Heavy.

Differential – Covariate – The difference in the Radius of Gyration or RG on the x-axis and the y-axis. RG differential indicates the amount of flare potential of a bowling ball – Range <0.06"

Mass Bias Differential – Covariate – A measure of the 'Asymmetry' of a bowling ball – Ranges from 0.008"-0.037"

Performance – Fixed Factor – A measure of how ‘reactive’ a bowling ball is – Ranges from 1-5: Entry, Performance, Advanced, High, Pro.

Radius of Gyration (RG) – Covariate – A measure of the distribution of mass inside the bowling ball – Ranges from 2.46"-2.8"

Core Type – Fixed Factor – Whether the core of a bowling ball is symmetric or asymmetric – Ranges from 1-2: Symmetric and Asymmetric.

Year – Covariate – The year in which the bowling ball was released.

Price – Covariate – Dependent Variable – The MSRP of the bowling ball in $.

PerfectScale® – Covariate – Dependent Variable – A measure of a bowling ball’s ‘Hook Potential’ – Ranges from 0-300

2. Aims and Objectives

In Chapter 3 – Descriptive Statistics & Illustrations, I will display Descriptive Statistics: Measures of Location and Measures of Dispersion of the data set, such as the Mean, Median and Mode, as well as Standard Deviation, Variance, Skewness and Kurtosis. I will also present various Illustrations - graphical properties of the data set through Bar Charts, Pie Charts, Clustered Bar Graphs, Histograms, Scatter Diagrams and Box Plots.

In Chapter 4 – Parametric/Non-Parametric Tests, I will come up with several Hypothesis and make use of several Tests to come up with conclusions about the Data Set. Examples of such tests may be Tests of Normality, Tests of effect of Fixed Factors on Dependent Variables (Covariates), Tests of association between two Fixed Factors and Tests of association between multiple Covariates, amongst others.

In Chapter 5 – Regression, I will try to relate the Dependent Variables to Covariates in the form of models (similar to 'equations') through Linear Regression.

In Chapter 6 – General Linear Model, I will relate the Dependent Variables to multiple Fixed Factors (n-way ANOVA) and the Dependent Variables to both Fixed Factors and Covariates (ANCOVA) to acquire models which describe the data set better (Dependent Variables in terms of Independent Variables).

3. Descriptive Statistics & Illustrations

3.1 Descriptive Statistics

3.1.1 Measures of Location

From Table 3.1.1.1, The Mean of the Price was found to be $181.62, whereas the Median Price was found to be 179.99. The difference between the mean and the median is very small, which suggests that there might be skewness in the data.

From Table 3.1.1.1, it was also found that the Mean Perfect Scale was 187, whereas the Median was 200.2. This difference between the mean and the median might again suggest some skewness in the data.

Table 3.1.1.1

		Price	PerfectScale
N	Valid	47	47
	Missing	0	0
Mean		181.6283	187.049
Median		179.9900	200.200

3.1.2 Measures of Dispersion

Table 3.1.2.1 shows a collection of measures of dispersion of the dependent variables in the data set. As suggested in Section 3.1.1, both dependent variables show a degree of negative skewness. It can also be seen that both dependent variables have a negative kurtosis (platykurtic).

Table 3.1.2.1

		Price	PerfectScale
N	Valid	47	47
	Missing	0	0
Std. Deviation		63.21682	46.6749
Variance		3996.366	2178.547
Skewness		-.132	-.940
Std. Error of Skewness		.347	.347
Kurtosis		-1.206	-.167
Std. Error of Kurtosis		.681	.681
Range		232.00	159.6

Table 3.1.3.1 shows a collection of descriptive statistics regarding all the covariates in the data set.

Table 3.1.3.1

		Differential	MassBiasDiff	RG	Year	Price	PerfectScale
N	Valid	47	47	47	47	47	47
	Missing	0	0	0	0	0	0
Mean		.04223	.00572	2.53009	2014.62	181.6283	187.049
Median		.04600	.00000	2.50000	2015.00	179.9900	200.200
Std. Deviation		.014424	.007362	.061736	2.280	63.21682	46.6749
Variance		.000	.000	.004	5.198	3996.366	2178.547
Skewness		-.935	.794	1.237	-2.021	-.132	-.940
Std. Error of Skewness		.347	.347	.347	.347	.347	.347
Kurtosis		-.117	-.849	.541	4.266	-1.206	-.167
Std. Error of Kurtosis		.681	.681	.681	.681	.681	.681
Range		.052	.024	.230	10	232.00	159.6
Percentiles	25	.03200	.00000	2.48100	2014.00	119.9900	150.500
	50	.04600	.00000	2.50000	2015.00	179.9900	200.200
	75	.05400	.01300	2.57000	2016.00	239.9900	230.200

Table 3.1.3.2 shows the mode of all the fixed factors in the data set.

Table 3.1.3.2

		Brand	Finish	LaneCondition	Performance	CoreType
N	Valid	47	47	47	47	47
	Missing	0	0	0	0	0
Mode		3.00	2.00	4.00	5.00	1.00

3.2 Illustrations

The Bar Graph in Figure 3.2.1 shows the amount of balls that fall under each performance category. It can be seen that in the sample chosen for the data set most balls fall under the 'Pro Performance' category. In recent years bowling balls are becoming 'stronger' which is why more 'Pro' balls are present (The Year variable is negatively skewed which means more balls from recent years are present in the data set than older balls). This claim will be further substantiated in Figure 3.2.5.

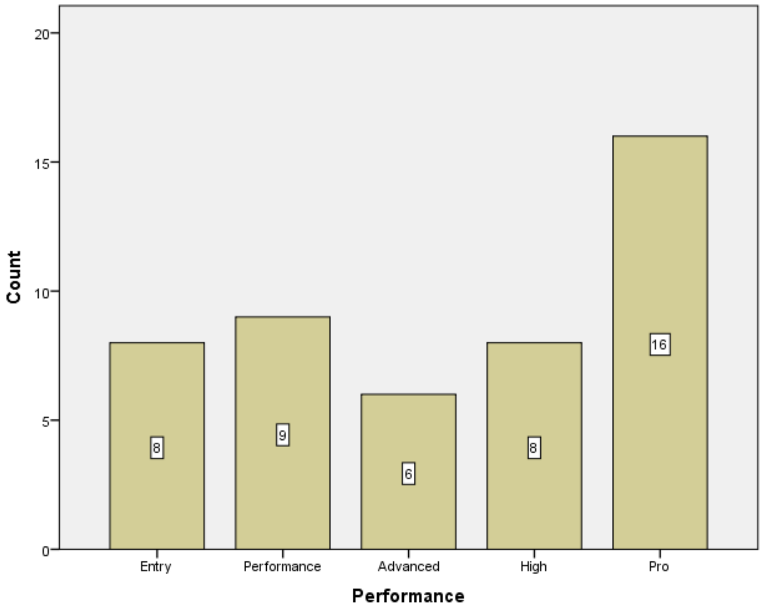


Figure 3.2.1: Bar Graph showing the amount of bowling balls per performance category

The Pie Chart in Figure 3.2.2 shows what percentage of bowling balls have a symmetric vs. asymmetric core.

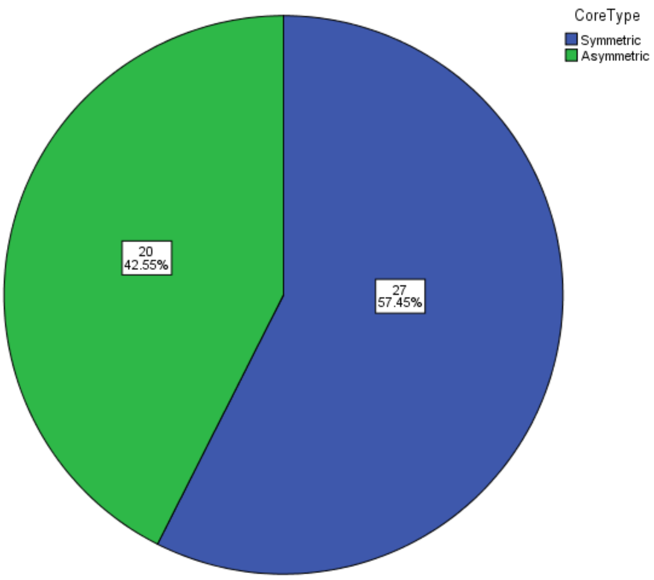


Figure 3.2.2: Pie Chart showing the percentage of bowling balls that have a symmetric vs asymmetric core

The Clustered Bar Graph in Figure 3.2.3 shows how bowling balls are distributed under performance categories, grouped by their core type. From this graph, one can observe that higher performing bowling balls are more likely to have an asymmetric core, i.e. bowling balls that have an asymmetric core usually perform better than bowling balls with symmetric cores.

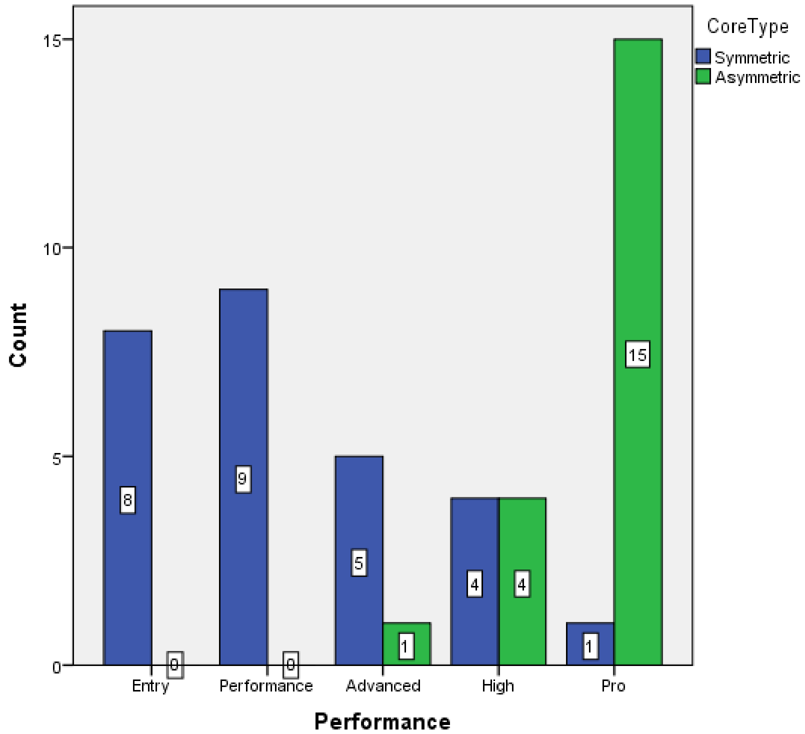
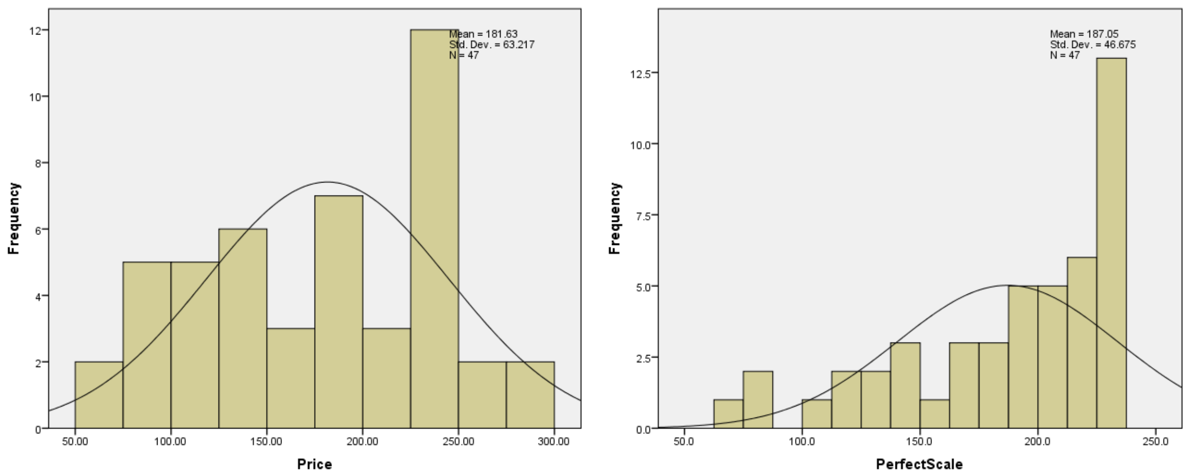


Figure 3.2.3: Clustered Bar Graph showing bowling ball Performance, clustered by Core Type

The Histograms in Figures 3.2.4a and 3.2.4b substantiate the claims made in section 3.1.2 regarding the negative skewness of both independent variables.



Figures 3.2.4a and 3.2.4b: Histogram for each independent variable (Price and PerfectScale respectively) fitted with the normal distribution

The Scatter Diagram in Figure 3.2.5 shows how the PerfectScale of a bowling ball relates to the Year in which it was released. It is instantly clear that bowling balls released in more recent years have a higher PerfectScale than those released in earlier years. The Linear Fit Line suggests that every year, the PerfectScale has increased by an average of 14.6.

This suggested Correlation will be further discussed in Section 4.4: Pearson Correlation and in Section 5.1: Simple Linear Regression.

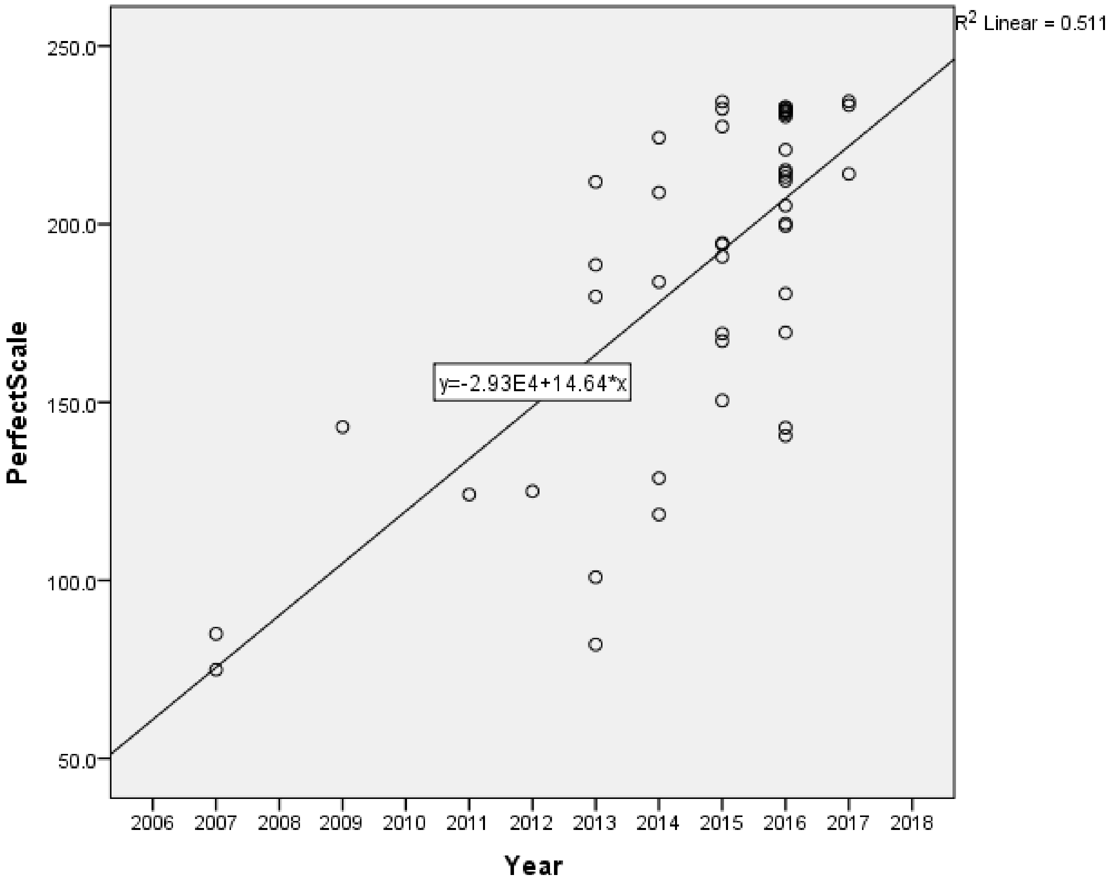


Figure 3.2.5: Scatter Diagram showing PerfectScale in relation to Year of ball production

The Box Plot in Figure 3.2.6 depicts the descriptives of the PerfectScale grouped by the ball's Finish graphically. It can be clearly seen that the Median PerfectScale of bowling balls with Matte or Sanded surfaces is noticeably higher than those with High Polish or Polished surfaces.

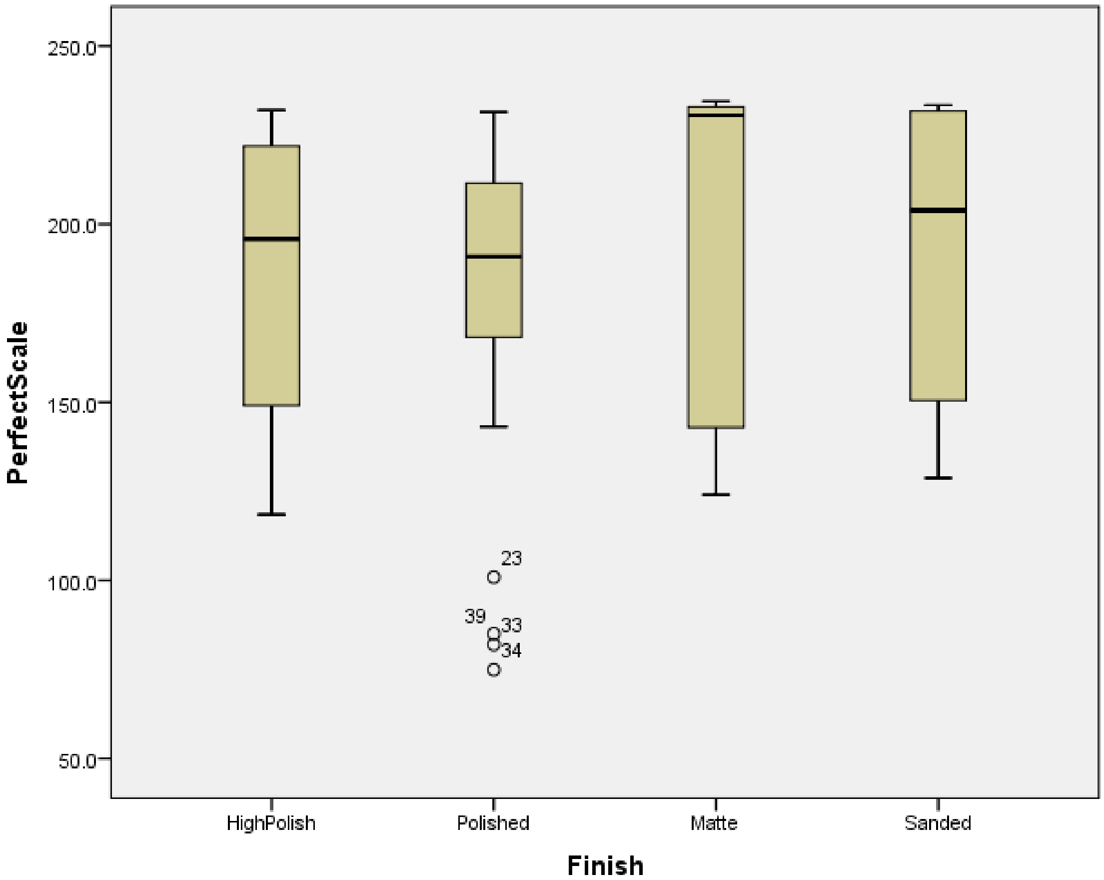


Figure 3.2.6: Box Plot showing the Perfect Scale of bowling balls grouped by their Finish

4. Parametric/Non-Parametric Tests

4.1 Tests of Normality

In order to determine whether we can use Parametric or Non-Parametric tests, we test the normality assumption by checking whether both dependent variables are Normally Distributed or not. This will be accomplished with the Kolmogorov-Smirnov and the Shapiro-Wilk tests.

H_0: Price is Normally Distributed

H_1: Price is not Normally Distributed

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Price	.141	47	.020	.945	47	.028

a. Lilliefors Significance Correction

Table 4.1.1: Test for Normality - Price

As can be seen in Table 4.1.1, the p-value was found to be 0.28, which is less than 0.05. Thus, with a 95% degree of confidence, H_0 can be rejected and H_1 is accepted. i.e. Price is not normally distributed.

H_0: PerfectScale is Normally Distributed

H_1: PerfectScale is not Normally Distributed

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
PerfectScale	.155	47	.007	.871	47	.000

a. Lilliefors Significance Correction

Table 4.1.2: Test for Normality - PerfectScale

As can be seen in Table 4.1.2, the p-value was found to be $9.5x10^{-5}$, which is less than 0.05. Thus, with a 95% degree of confidence, H_0 can be rejected and H_1 is accepted. i.e. PerfectScale is not normally distributed.

4.2 Kruskal-Wallis Test

Since the data was determined not to be normally distributed, the Kruskal-Wallis Test was deemed suitable to study further this data set. The Kruskal-Wallis Test will help determine whether certain Fixed Factors significantly affect the dependent variables.

4.2.1 Does the Brand affect the PerfectScale of a Bowling Ball?

H_0: All Brands produce the same PerfectScale bowling balls.

H_1: Some Brands produce higher PerfectScale bowling balls.

Test Statistics[a,b]

	PerfectScale
Chi-Square	4.643
df	6
Asymp. Sig.	.590

a. Kruskal Wallis Test

b. Grouping Variable: Brand

Table 4.2.1: Kruskall-Wallis Test on PerfectScale vs Brand

As can be seen in Table 4.2.1, since the p-value $0.590 > 0.05$, H_0 is accepted and thus, with a 95% degree of confidence it can be said that all Brands produce bowling balls with similar PerfectScale (i.e. Brand is not a significant factor in the PerfectScale of a bowling ball).

4.2.2 Does Performance affect the Price of a Bowling Ball?

H_0: Bowling Balls from All Performance classes have the same Price.

H_1: Bowling Balls from Some Performance classes have a higher Price.

Test Statistics[a,b]

	Price
Chi-Square	34.559
df	4
Asymp. Sig.	.000

a. Kruskal Wallis Test

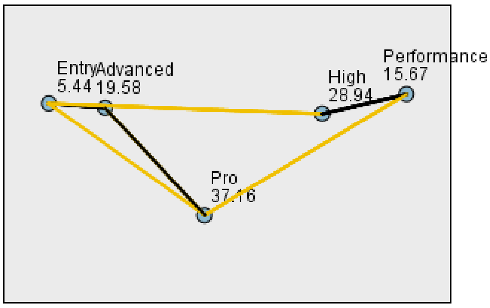
b. Grouping Variable: Performance

Table 4.2.2.1: Kruskall-Wallis Test on Price vs Performance

As can be seen in Table 4.2.2.1, since the p-value $5.72\times10^{-7} < 0.05$, H_0 is rejected and H_1 is accepted, and thus, with a 95% degree of confidence it can be said that Bowling Balls from some Performance classes have a higher Price.

With this result in mind, the Pairwise Comparisons were generated, as shown in Figure 4.2.2.2. The Pairwise Comparisons suggest that there is a significant change in Price when comparing Entry and High, Entry and Pro, and Performance and Pro categories of Performance bowling balls.

Pairwise Comparisons of Performance



Each node shows the sample average rank of Performance.

Sample1-Sample2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
Entry-Performance	-10.229	6.645	-1.539	.124	1.000
Entry-Advanced	-14.146	7.385	-1.915	.055	.554
Entry-High	-23.500	6.838	-3.437	.001	.006
Entry-Pro	-31.719	5.922	-5.357	.000	.000
Performance-Advanced	-3.917	7.207	-.543	.587	1.000
Performance-High	-13.271	6.645	-1.997	.046	.458
Performance-Pro	-21.490	5.698	-3.771	.000	.002
Advanced-High	-9.354	7.385	-1.267	.205	1.000
Advanced-Pro	-17.573	6.546	-2.684	.007	.073
High-Pro	-8.219	5.922	-1.388	.165	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

Figure 4.2.2.2: Pairwise Comparisons of Performance

4.3 Chi-Squared Test

4.3.1 Does The Finish affect the intended Lane Condition of a Bowling Ball?

The Chi-Squared Test was deemed to be yet another suitable test to further analyse this data set. The Chi-Squared Test helps in determining whether an association exists between two fixed factors.

H_0: There is no association between Finish and Lane Condition

H_1: There is an association between Finish and Lane Condition

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	21.803[a]	12	.040
Likelihood Ratio	26.458	12	.009
Linear-by-Linear Association	3.725	1	.054
N of Valid Cases	47		

a. 19 cells (95.0%) have expected count less than 5. The minimum expected count is .51.

Table 4.3.1: Chi-Squared Test on Lane Condition vs Finish

As can be seen in Table 4.3.1, since the p-value $0.04 < 0.05$, H_0 is rejected and H_1 is accepted, and thus, with a 95% degree of confidence it can be said that the Finish of a Bowling Ball does affect what Lane Condition a Bowling Ball is designed for (ie. There is an association between the two Fixed Factors).

4.4 Pearson Correlation

The Pearson Correlation Test was deemed to be a suitable test to further analyse this data set. The Pearson Correlation Matrix helps in determining whether any 2 covariates are related/correlated.

H_0: The Correlation between the 2 variables is 0

H_1: There is a Significant Correlation between the 2 variables

Correlations

		Differential	MassBiasDiff	RG	Year	Price	PerfectScale
Differential	Pearson Correlation	1	.614**	-.777**	.358*	.681**	.791**
	Sig. (2-tailed)		.000	.000	.014	.000	.000
	N	47	47	47	47	47	47
MassBiasDiff	Pearson Correlation	.614**	1	-.528**	.348*	.682**	.671**
	Sig. (2-tailed)	.000		.000	.016	.000	.000
	N	47	47	47	47	47	47
RG	Pearson Correlation	-.777**	-.528**	1	-.546**	-.594**	-.793**
	Sig. (2-tailed)	.000	.000		.000	.000	.000
	N	47	47	47	47	47	47
Year	Pearson Correlation	.358*	.348*	-.546**	1	.670**	.715**
	Sig. (2-tailed)	.014	.016	.000		.000	.000
	N	47	47	47	47	47	47
Price	Pearson Correlation	.681**	.682**	-.594**	.670**	1	.842**
	Sig. (2-tailed)	.000	.000	.000	.000		.000
	N	47	47	47	47	47	47
PerfectScale	Pearson Correlation	.791**	.671**	-.793**	.715**	.842**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	
	N	47	47	47	47	47	47

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 4.4: Pearson Correlation Matrix of all Covariate Variables

From Table 4.4, the following significant (non-zero) correlations can be extracted:

With a 95% degree of confidence:

$$Differential \overset{+}{\leftrightarrow} Year$$

$$MassBiasDiff \overset{+}{\leftrightarrow} Year$$

With a 99% degree of confidence:

$$Differential \overset{+}{\leftrightarrow} MassBiasDiff$$
$$Differential \overset{-}{\leftrightarrow} RG$$
$$Differential \overset{+}{\leftrightarrow} Price$$
$$Differential \overset{+}{\leftrightarrow} PerfectScale$$
$$MassBiasDiff \overset{-}{\leftrightarrow} RG$$
$$MassBiasDiff \overset{-}{\leftrightarrow} RG$$
$$MassBiasDiff \overset{+}{\leftrightarrow} Price$$
$$MassBiasDiff \overset{+}{\leftrightarrow} PerfectScale$$
$$RG \overset{-}{\leftrightarrow} Year$$
$$RG \overset{-}{\leftrightarrow} Price$$
$$RG \overset{-}{\leftrightarrow} PerfectScale$$
$$Year \overset{+}{\leftrightarrow} Price$$
$$Year \overset{+}{\leftrightarrow} PerfectScale$$
$$Price \overset{+}{\leftrightarrow} PerfectScale$$

Some correlations that stand out are the following:

The positive correlation between the Differential, MassBiasDiff and the Year, and the negative correlation between the RG and the Year; and the same correlation of each variable with the PerfectScale; further prove the strength of the positive correlation between the Year and PerfectScale, which was also suggested in an earlier Scatter Diagram (Figure 3.2.5). This means that over the years, further research and development by competitive companies in the industry has developed stronger and more reactive bowling balls.

The positive correlation between the Price and PerfectScale also suggests that the stronger a bowling ball (higher PerfectScale) is, the more expensive it is.

The above correlations shall be used in Chapter 5 for regression analysis.

5. Regression

5.1 Simple Linear Regression

5.1.1 SLR of PerfectScale vs Year

From Table 5.1.1.1 and the Scatter Diagram in Chapter 3 (Figure 3.2.5), the R^2 coefficient of determination is found to be 0.511, meaning the regression line explains 51.1% of the variability of the data set.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.715[a]	.511	.500	32.9910

a. Predictors: (Constant), Year

b. Dependent Variable: PerfectScale

Table 5.1.1.1: Model Summary of PerfectScale vs Year

H_0: The model $PerfectScale = b_0$ is appropriate for this data set

H_1: The model $PerfectScale = b_0 + b_1 Year$ fits the data set better than the constant model

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	51234.793	1	51234.793	47.073	.000[b]
	Residual	48978.365	45	1088.408		
	Total	100213.157	46			

a. Dependent Variable: PerfectScale

b. Predictors: (Constant), Year

Table 5.1.1.2: ANOVA of PerfectScale vs Year

As can be seen in Table 5.1.1.2, since the p-value $1.64 \times 10^{-8} < 0.05$, H_0 is rejected and H_1 is accepted, and thus, with a 95% degree of confidence it can be said that the model $PerfectScale = b_0 + b_1 Year$ fits the data set better than the constant model.

From Table 5.1.1.3, b_0 and b_1 can be found to be -29303.266 and 14.638 respectively. This makes the model:

$$Expected\ PerfectScale = -29303.266 + 14.368 * Year$$

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-29303.266	4298.264		-6.817	.000
	Year	14.638	2.134	.715	6.861	.000

a. Dependent Variable: PerfectScale

Table 5.1.1.3: Coefficients of PefecstScale vs Year

By looking at the Studentized Residuals for every entry in the data set, it is discovered that there are 2 cases which are outliers - Balls 24 and 33 ($\because Studentized\ Residuals \notin [-2,2]$).

Next, the Centered Leverage Values are examined, but first, the cut-off point must be calculated using the formula $\frac{2p}{n}$; where p is the number of parameters and n is the sample size. Using $p = 2$ and $n = 47$ the cut-off point is found to be $\frac{2*2}{47} = 0.085$. When looking through the data set for which cases had a Centered Leverage Value higher than the cut-off point (0.085), 3 outliers were found, these being cases 34, 36 and 39.

Finally, by examining the Cook's Distance for all the cases, 1 outlier was found to have an abnormally high distance, that being case 36.

The Normality of the Unstandardized Residuals was checked using the Kolmogorov-Smirnov and Shapiro-Wilk tests. As can be seen in Table 5.1.1.4, the p-value was found to be 0.007<0.05. Thus, with a 95% degree of confidence it is concluded that the Unstandardized Residuals are not normally distributed, and it can hence be stated that this model should not be applied to this data set.

H_0: Residuals are Normally Distributed

H_1: Residuals are not Normally Distributed

Tests of Normality

	Kolmogorov-Smirnov[a] Statistic	Kolmogorov-Smirnov[a] df	Kolmogorov-Smirnov[a] Sig.	Shapiro-Wilk Statistic	Shapiro-Wilk df	Shapiro-Wilk Sig.
RES_1	.134	47	.034	.929	47	.007

a. Lilliefors Significance Correction

Table 5.1.1.4: Test for Normality – Unstandardized Residuals

#	Unstandardized Predicted Value	Unstandardized Residual	Studentized Residual	Cook's Distance	Centered Leverage Value
1	207.2932	24.70678	0.7601	0.00871	0.008
2	207.2932	25.60678	0.78779	0.00936	0.008
3	178.0169	46.28313	1.41922	0.02357	0.00159
4	192.6551	-42.1551	-1.29199	0.01868	0.00061
5	192.6551	-25.4551	-0.78016	0.00681	0.00061
6	207.2932	-7.79322	-0.23976	0.00087	0.008
7	207.2932	7.10678	0.21864	0.00072	0.008
8	207.2932	22.90678	0.70473	0.00749	0.008
9	134.1024	-10.0024	-0.3154	0.00409	0.05472
10	207.2932	-26.7932	-0.82429	0.01025	0.008
11	192.6551	-23.3551	-0.7158	0.00573	0.00061
12	207.2932	6.00678	0.1848	0.00051	0.008
13	178.0169	-49.3169	-1.51225	0.02676	0.00159
14	207.2932	23.60678	0.72626	0.00795	0.008
15	207.2932	24.80678	0.76318	0.00878	0.008
16	178.0169	-59.5169	-1.82502	0.03898	0.00159
17	207.2932	4.90678	0.15096	0.00034	0.008
18	163.3787	25.2213	0.77711	0.01005	0.01094
19	221.9314	-7.83139	-0.24291	0.00139	0.02375
20	221.9314	11.46861	0.35573	0.00298	0.02375
21	207.2932	-7.09322	-0.21822	0.00072	0.008
22	192.6551	41.74495	1.27942	0.01832	0.00061
23	163.3787	-62.4787	-1.92507	0.06167	0.01094
24	207.2932	-66.6932	-2.05181	0.06348	0.008
25	192.6551	-1.75505	-0.05379	0.00003	0.00061
26	192.6551	39.74495	1.21813	0.0166	0.00061
27	207.2932	-2.09322	-0.0644	0.00006	0.008
28	221.9314	12.56861	0.38985	0.00358	0.02375
29	178.0169	5.78313	0.17733	0.00037	0.00159
30	178.0169	30.88313	0.947	0.01049	0.00159
31	192.6551	2.04495	0.06267	0.00004	0.00061
32	163.3787	48.5213	1.49502	0.0372	0.01094
33	163.3787	-81.3787	-2.50741	0.10463	0.01094
34	75.54965	-0.64965	-0.02295	0.00009	0.24265
35	192.6551	34.74495	1.06488	0.01269	0.00061
36	104.826	38.274	1.26074	0.14381	0.13195
37	207.2932	24.50678	0.75395	0.00857	0.008
38	148.7405	-23.7405	-0.73827	0.01432	0.02864
39	75.54965	9.45035	0.33388	0.01999	0.24265
40	207.2932	24.20678	0.74472	0.00836	0.008
41	163.3787	16.3213	0.50289	0.00421	0.01094
42	207.2932	-37.6932	-1.15963	0.02028	0.008
43	207.2932	13.60678	0.41861	0.00264	0.008

44	207.2932	-64.3932	-1.98105	0.05918	0.008
45	192.6551	1.64495	0.05042	0.00003	0.00061
46	207.2932	7.90678	0.24325	0.00089	0.008
47	207.2932	25.60678	0.78779	0.00936	0.008

Table 5.1.1.5: Residuals and Distances – Used for checking Normality and Finding Outliers

5.2 Multiple Linear Regression

5.2.1 MLR of Price with all correlated covariates

(It was already discovered in Section 4.1 that Price is not normally distributed, this technically invalidates the validity of this model)

It has already been shown that all the covariates are correlated with the dependent variable, Price, in Section 4.4 using the Pearson Correlation test.

Thus, the Linear Regression is calculated using SPSS with Price as the dependent and all the other covariates as independents.

H_0: The model $Expected\ Price = b_0$ is appropriate for this data set

H_1: The model $Expected\ Price = b_0 + b_1 Differential + b_2 MassBiasDiff + b_3 RG + b_4 Year$ fits the data set better than the constant model.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	138733.683	4	34683.421	32.300	.000[b]
	Residual	45099.169	42	1073.790		
	Total	183832.851	46			

a. Dependent Variable: Price

b. Predictors: (Constant), Year, MassBiasDiff, Differential, RG

Table 5.2.1.1: ANOVA of Price vs Covariates

As can be seen in Table 5.2.1.1, since the p-value $2.58 \times 10^{-12} < 0.05$, H_0 is rejected and H_1 is accepted, and thus, with a 95% degree of confidence it can be said that the model $Price = b_0 + b_1 Differential + b_2 MassBiasDiff + b_3 RG + b_4 Year$ fits the data set better than the constant model.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-28774.369	5360.405		-5.368	.000
	Differential	2034.789	583.534	.464	3.487	.001
	MassBiasDiff	2896.558	844.319	.337	3.431	.001
	RG	225.555	139.547	.220	1.616	.114
	Year	14.039	2.576	.506	5.451	.000

a. Dependent Variable: Price

Table 5.2.1.2: Coefficients of Price vs Covariates

As can be seen in Table 5.2.1.2, since the p-value is 0.114>0.05 H_0 is accepted and thus, with a 95% degree of confidence, it can be said that the RG of a Bowling Ball does not affect its Price (i.e. it is insignificant).

Thus, RG is removed from the list of Covariates, and the Coefficients are re-calculated.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-24433.321	4725.239		-5.171	.000
	Differential	1401.502	440.472	.320	3.182	.003
	MassBiasDiff	2860.377	859.702	.333	3.327	.002
	Year	12.181	2.348	.439	5.188	.000

a. Dependent Variable: Price

Table 5.2.1.3: Coefficients of Price vs Covariates

In Table 5.2.1.3, all listed Covariates are significant (all p-values < 0.05). Thus, the model becomes $Expected\ Price = b_0 + b_1 Differential + b_2 MassBiasDiff + b_3 Year$

From Table 5.2.1.3, the values of b_0, b_1, b_2, b_3 and b_4 can be found. This makes the model:

$$Expected\ Price = -24433.321 + 1401.502 * Differential + 2860.377 * MassBiasDiff + 12.181 * Year$$

From Table 5.2.1.4, the R^2 coefficient of determination is found to be 0.721, meaning the regression line explains 72.1% of the variability of the data set.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.860[a]	.739	.721	33.37751

a. Predictors: (Constant), Year, MassBiasDiff, Differential

Table 5.2.1.4: Model Summary of Price vs Covariates

By looking at the Studentized Residuals for every entry in the data set, it is discovered that there are 3 cases which are outliers - Balls 5, 37 and 39 ($\because Studentized\ Residuals \notin [-2,2]$).

Next, the Centered Leverage Values are examined, but first, the cut-off point must be calculated using the formula $\frac{2p}{n}$; where p is the number of parameters and n is the sample size. Using $p = 4$ and $n = 47$ the cut-off point is found to be $\frac{2*4}{47} = 0.17$. When looking through the data set for which cases had a Centered Leverage Value higher than the cut-off point (0.17), 2 outliers were found, these being cases 34 and 39. (Note that Case 39 was already pointed out to be an outlier by the Studentized Residuals)

Finally, by examining the Cook's Distance for all the cases, 2 outliers were found to have an abnormally high distance, these being cases 37 and 39. (Note that both cases were already pointed out to be outliers by previous tests)

The Normality of the Unstandardized Residuals was checked using the Kolmogorov-Smirnov and Shapiro-Wilk tests. As can be seen in Table 5.2.1.5, the p-value was found to be 0.165>0.05. Thus, H_0 is accepted, and with a 95% degree of confidence it is concluded that the Unstandardized Residuals are normally distributed, and can hence state that this model is correct for this data set.

H_0: Residuals are Normally Distributed

H_1: Residuals are not Normally Distributed

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.098	47	.200*	.965	47	.165

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 5.2.1.5: Test for Normality – Unstandardized Residuals

#	Unstandardized Predicted Value	Unstandardized Residual	Studentized Residual	Cook's Distance	Centered Leverage Value
1	241.4411	8.54889	0.26303	0.00094	0.03052
2	235.7777	24.21227	0.74238	0.00653	0.02394
3	203.1795	16.81051	0.53033	0.00765	0.07683
4	131.7536	28.23643	0.90123	0.02739	0.0976
5	185.0107	-65.0207	-2.03898	0.09931	0.06594
6	180.3733	-0.3833	-0.01182	0	0.03441
7	249.9649	-9.97487	-0.30918	0.00168	0.04442
8	227.3687	12.62128	0.38724	0.00183	0.02519
9	106.8564	23.13358	0.72152	0.0109	0.05597
10	164.9568	-44.9668	-1.39183	0.0326	0.0418
11	169.5941	-19.6041	-0.60092	0.00422	0.02338
12	198.5928	-48.6028	-1.53602	0.06647	0.08001
13	143.3984	-33.4084	-1.01964	0.00981	0.0151
14	232.86	-12.87	-0.39418	0.00175	0.02186
15	247.1619	2.82813	0.08759	0.00013	0.04294
16	111.1639	-1.17389	-0.03827	0.00007	0.1341
17	194.3883	25.60168	0.80211	0.01505	0.06426
18	150.8388	29.1512	0.90223	0.01367	0.04166
19	210.7735	29.2165	0.93274	0.02946	0.09802
20	269.3825	-19.3925	-0.61629	0.01188	0.08996
21	190.1838	-20.1938	-0.62836	0.00776	0.05165
22	237.8989	2.09106	0.06555	0.0001	0.06537
23	94.77871	-4.78871	-0.15832	0.00136	0.15755
24	153.7448	26.24524	0.82652	0.01791	0.07365
25	163.9881	16.00188	0.48921	0.00247	0.01835
26	233.5223	26.46768	0.81823	0.01083	0.03949
27	203.1989	26.79106	0.81501	0.00514	0.00877
28	239.492	40.49799	1.24573	0.021	0.03007
29	140.5954	-30.6054	-0.93557	0.00898	0.01813
30	219.8828	20.10723	0.62438	0.00723	0.04782
31	178.29	-28.3	-0.85871	0.00474	0.00379
32	183.475	-43.485	-1.34825	0.03224	0.04497
33	114.3997	-24.4097	-0.76034	0.0117	0.0536
34	70.74726	-2.75726	-0.0985	0.00102	0.27535
35	230.7767	-30.7867	-0.95454	0.01616	0.04498
36	96.5101	-5.5201	-0.18239	0.0018	0.15649
37	197.1913	102.7987	3.23873	0.27748	0.07441
38	130.2491	-37.2591	-1.15217	0.02168	0.04003
39	41.31571	58.67429	2.07032	0.41474	0.25776
40	241.4411	8.54889	0.26303	0.00094	0.03052
41	125.6118	-50.6218	-1.55288	0.02915	0.02484
42	163.5553	-13.5653	-0.4205	0.00312	0.04456
43	215.8125	34.17754	1.04383	0.01067	0.01642

44	143.9342	16.05576	0.51904	0.01106	0.1198
45	170.9956	28.99437	0.88974	0.00971	0.02551
46	231.4585	-31.4685	-0.96276	0.00991	0.01974
47	268.6434	-28.6534	-0.94214	0.04537	0.14847

Table 5.2.1.6: Residuals and Distances – Used for checking Normality and Finding Outliers

6. General Linear Model

6.1 5-Way ANOVA

The ANOVA test requires the normality assumption, however neither of the two dependent variables of this data set are normally distributed.

H_0: The Brand of a Bowling Ball does not affect its Price

H_1: The Brand of a Bowling Ball affects its Price

Tests of Between-Subjects Effects

Dependent Variable: Price

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	12440.351[a]	6	2073.392	.484	.816
Intercept	1500883.597	1	1500883.597	350.280	.000
Brand	12440.351	6	2073.392	.484	.816
Error	171392.500	40	4284.813		
Total	1734308.265	47			
Corrected Total	183832.851	46			

a. R Squared = .068 (Adjusted R Squared = -.072)

Table 6.1.1: Effects of Bowling Ball Brand on Price

As can be seen in Table 6.1.1, since the p-value is 0.816>0.05 H_0 is accepted and thus, with a 95% degree of confidence, it can be said that the Brand of a Bowling Ball does not affect its Price (i.e. it is insignificant).

Thus, Brand is removed from the list of Fixed Factors, and the significance of the Finish is checked.

H_0: The Finish of a Bowling Ball does not affect its Price

H_1: The Finish of a Bowling Ball affects its Price

Tests of Between-Subjects Effects

Dependent Variable: Price

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	25617.930[a]	3	8539.310	2.321	.089
Intercept	967406.471	1	967406.471	262.924	.000
Finish	25617.930	3	8539.310	2.321	.089
Error	158214.921	43	3679.417		
Total	1734308.265	47			
Corrected Total	183832.851	46			

a. R Squared = .139 (Adjusted R Squared = .079)

Table 6.1.2: Effects of Bowling Ball Finish on Price

As can be seen in Table 6.1.2, since the p-value is 0.089>0.05 H_0 is accepted and thus, with a 95% degree of confidence, it can be said that the Finish of a Bowling Ball does not affect its Price (i.e. it is insignificant).

Thus, Finish is removed from the list of Fixed Factors, and the significance of the LaneCondition is checked.

H_0: The LaneCondition of a Bowling Ball does not affect its Price

H_1: The LaneCondition of a Bowling Ball affects its Price

Tests of Between-Subjects Effects

Dependent Variable: Price

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	106304.314[a]	4	26576.079	14.397	.000
Intercept	1331159.326	1	1331159.326	721.137	.000
LaneCondition	106304.314	4	26576.079	14.397	.000
Error	77528.537	42	1845.918		
Total	1734308.265	47			
Corrected Total	183832.851	46			

a. R Squared = .578 (Adjusted R Squared = .538)

Table 6.1.3: Effects of Bowling Ball LaneCondition on Price

As can be seen in Table 6.1.3, since the p-value is $1.76 \times 10^{-7} < 0.05$ H_0 is rejected and H_1 is accepted and thus, with a 95% degree of confidence, it can be said that the LaneCondition of a Bowling Ball affects its Price (i.e. it is significant).

Thus, LaneCondition is kept in the list of Fixed Factors, and the significance of the Performance is checked.

H_0: The Performance of a Bowling Ball does not affect its Price

H_1: The Performance of a Bowling Ball affects its Price

Tests of Between-Subjects Effects

Dependent Variable: Price

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	144951.639[a]	8	18118.955	17.708	.000
Intercept	1109577.779	1	1109577.779	1084.430	.000
LaneCondition	4431.372	4	1107.843	1.083	.379
Performance	38647.325	4	9661.831	9.443	.000
Error	38881.212	38	1023.190		
Total	1734308.265	47			
Corrected Total	183832.851	46			

a. R Squared = .788 (Adjusted R Squared = .744)

Table 6.1.4: Effects of Bowling Ball LaneCondition and Performance on Price

As can be seen in Table 6.1.4, when the effects of Lane Condition and Performance together on Price are tested, the significance of LaneCondition becomes 0.379>0.05, which would imply that it is insignificant. This suggests that there is an interaction between LaneCondition and Performance. To choose whether to leave LaneCondition or Performance as a Fixed Factor, we first test the effects of Performance alone on Price.

Tests of Between-Subjects Effects

Dependent Variable: Price

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	140520.268[a]	4	35130.067	34.065	.000
Intercept	1222536.438	1	1222536.438	1185.488	.000
Performance	140520.268	4	35130.067	34.065	.000
Error	43312.583	42	1031.252		
Total	1734308.265	47			
Corrected Total	183832.851	46			

a. R Squared = .764 (Adjusted R Squared = .742)

Table 6.1.5: Effects of Bowling Ball Performance on Price

From Table 6.1.5, The adjusted R squared value of Performance on Price is 0.742 and from Table 6.1.3 the adjusted R squared value of LaneCondition on Price is 0.538. Therefore, Performance is chosen from the 2 Fixed Factors since it has a higher R Squared value, meaning it fits the model better.

Lastly, the significance of the CoreType is checked.

H_0: The CoreType of a Bowling Ball does not affect its Price

H_1: The CoreType of a Bowling Ball affects its Price

Tests of Between-Subjects Effects

Dependent Variable: Price

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	140522.213[a]	5	28104.443	26.605	.000
Intercept	897202.462	1	897202.462	849.336	.000
Performance	57403.029	4	14350.757	13.585	.000
CoreType	1.945	1	1.945	.002	.966
Error	43310.638	41	1056.357		
Total	1734308.265	47			
Corrected Total	183832.851	46			

a. R Squared = .764 (Adjusted R Squared = .736)

Table 6.1.6: Effects of Bowling Ball Performance and CoreType on Price

From Table 6.1.6, since the p-value of CoreType is 0.966>0.5, H_0 is accepted and thus, with a 95% degree of confidence, it can be said that the CoreType of a Bowling Ball does not affect its Price (i.e. it is insignificant).

Thus, as can be seen in Table 6.1.7, the final model will be based on the following Fixed Factors, i.e. Performance only.

Tests of Between-Subjects Effects

Dependent Variable: Price

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	140520.268[a]	4	35130.067	34.065	.000
Intercept	1222536.438	1	1222536.438	1185.488	.000
Performance	140520.268	4	35130.067	34.065	.000
Error	43312.583	42	1031.252		
Total	1734308.265	47			
Corrected Total	183832.851	46			

a. R Squared = .764 (Adjusted R Squared = .742)

Table 6.1.7: Effects of Bowling Ball Performance and CoreType on Price

From the Adjusted R Squared value, it is found that the model explains 74% of the variability of the data set.

From table 6.1.8, the model for Price becomes:

$$Expected\ Price = 240.6 - 147.4(Entry) - 98.3(Performance) - 77.292(Advanced) - 30.625(High)$$

Parameter Estimates

Dependent Variable: Price

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	240.615	8.028	29.971	.000	224.413	256.817
[Performance=1.00]	-147.375	13.905	-10.598	.000	-175.437	-119.313
[Performance=2.00]	-98.292	13.380	-7.346	.000	-125.295	-71.289
[Performance=3.00]	-77.292	15.373	-5.028	.000	-108.316	-46.268
[Performance=4.00]	-30.625	13.905	-2.202	.033	-58.687	-2.563
[Performance=5.00]	0[a]	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

Table 6.1.8: Parameter Estimates of Price vs Performance

By looking at the Studentized Residuals for every entry in the data set, it is discovered that there are 2 cases which are outliers - Balls 13 and 32 ($\because Studentized\ Residuals \notin [-2,2]$).

Next, the Centered Leverage Values are examined, but first, the cut-off point must be calculated using the formula $\frac{2p}{n}$; where p is the number of parameters and n is the sample size. Using $p = 5$ and $n = 47$ the cut-off point is found to be $\frac{2*5}{47} = 0.213$. When looking through the data set for which cases had a Centered Leverage Value higher than the cut-off point (0.213), no outliers were found, as all cases had a Centered Leverage Value lower than the cut-off point.

Finally, by examining the Cook's Distance for all the cases, 1 outlier was found to have an abnormally high distance, that being case 13. (Note that case 13 was already pointed out to be an outlier by the Studentized Residuals.)

The Normality of the Unstandardized Residuals was checked using the Kolmogorov-Smirnov and Shapiro-Wilk tests. As can be seen in Table 6.1.9, the p-value was found to be 0.0005<0.05 (from the Shapiro-Wilk test). Thus, with a 95% degree of confidence it is concluded that the Unstandardized Residuals are not normally distributed, and it can hence be stated that this model should not be applied to this data set.

H_0: Residuals are Normally Distributed

H_1: Residuals are not Normally Distributed

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Residual for Price	.119	47	.092	.894	47	.000

a. Lilliefors Significance Correction

Table 6.1.9: Test for Normality – Unstandardized Residuals

#	Unstandardized Predicted Value	Unstandardized Residual	Studentized Residual	Cook's Distance	Centered Leverage Value
1	240.61	9.38	0.3	0	0.06
2	240.61	19.38	0.62	0.01	0.06
3	240.61	-20.62	-0.66	0.01	0.06
4	142.32	17.67	0.58	0.01	0.11
5	142.32	-22.33	-0.74	0.01	0.11
6	163.32	16.67	0.57	0.01	0.17
7	209.99	30	1	0.03	0.13
8	240.61	-0.62	-0.02	0	0.06
9	142.32	-12.33	-0.41	0	0.11
10	93.24	26.75	0.89	0.02	0.13
11	163.32	-13.33	-0.45	0.01	0.17
12	163.32	-13.33	-0.45	0.01	0.17
13	209.99	-100	-3.33	0.32	0.13
14	240.61	-20.62	-0.66	0.01	0.06
15	240.61	9.38	0.3	0	0.06
16	93.24	16.75	0.56	0.01	0.13
17	209.99	10	0.33	0	0.13
18	163.32	16.67	0.57	0.01	0.17
19	209.99	30	1	0.03	0.13
20	240.61	9.38	0.3	0	0.06
21	163.32	6.67	0.23	0	0.17
22	240.61	-0.62	-0.02	0	0.06
23	93.24	-3.25	-0.11	0	0.13
24	142.32	37.67	1.24	0.04	0.11
25	142.32	37.67	1.24	0.04	0.11
26	240.61	19.38	0.62	0.01	0.06
27	209.99	20	0.67	0.01	0.13
28	240.61	39.38	1.27	0.02	0.06
29	142.32	-32.33	-1.07	0.03	0.11
30	209.99	30	1	0.03	0.13
31	163.32	-13.33	-0.45	0.01	0.17
32	240.61	-100.62	-3.24	0.14	0.06
33	93.24	-3.25	-0.11	0	0.13
34	93.24	-25.25	-0.84	0.02	0.13

35	240.61	-40.62	-1.31	0.02	0.06
36	142.32	-51.33	-1.7	0.07	0.11
37	240.61	59.38	1.91	0.05	0.06
38	93.24	-0.25	-0.01	0	0.13
39	93.24	6.75	0.22	0	0.13
40	240.61	9.38	0.3	0	0.06
41	93.24	-18.25	-0.61	0.01	0.13
42	142.32	7.67	0.25	0	0.11
43	240.61	9.38	0.3	0	0.06
44	142.32	17.67	0.58	0.01	0.11
45	209.99	-10	-0.33	0	0.13
46	209.99	-10	-0.33	0	0.13
47	240.61	-0.62	-0.02	0	0.06

Table 6.1.10: Residuals and Distances – Used for checking Normality and Finding Outliers

6.2 ANCOVA

PerfectScale vs All Significant Fixed Factors and Covariates

The ANCOVA test requires the normality assumption, however neither of the two dependent variables of this data set are normally distributed.

First, it is required to check the significance of all the Fixed Factors. This is done one Fixed Factor at a time, starting with the Brand.

H_0: The Brand of a Bowling Ball does not affect its PerfectScale

H_1: The Brand of a Bowling Ball affects its PerfectScale

Tests of Between-Subjects Effects

Dependent Variable: PerfectScale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	10544.107[a]	6	1757.351	.784	.588
Intercept	1623734.973	1	1623734.973	724.323	.000
Brand	10544.107	6	1757.351	.784	.588
Error	89669.051	40	2241.726		
Total	1744616.470	47			
Corrected Total	100213.157	46			

a. R Squared = .105 (Adjusted R Squared = -.029)

Table 6.2.1: Effects of Bowling Ball Brand on PerfectScale

As can be seen in Table 6.2.1, since the p-value is 0.588>0.05 H_0 is accepted and thus, with a 95% degree of confidence, it can be said that the Brand of a Bowling Ball does not affect its PerfectScale (i.e. it is insignificant).

Thus, Brand is removed from the list of Fixed Factors, and the significance of the Finish is checked.

H_0: The Finish of a Bowling Ball does not affect its PerfectScale

H_1: The Finish of a Bowling Ball affects its PerfectScale

Tests of Between-Subjects Effects

Dependent Variable: PerfectScale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	5265.338[a]	3	1755.113	.795	.504
Intercept	1075836.391	1	1075836.391	487.225	.000
Finish	5265.338	3	1755.113	.795	.504
Error	94947.819	43	2208.089		
Total	1744616.470	47			
Corrected Total	100213.157	46			

a. R Squared = .053 (Adjusted R Squared = -.014)

Table 6.2.2: Effects of Bowling Ball Finish on PerfectScale

As can be seen in Table 6.2.2, since the p-value is $0.504>0.05$ H_0 is accepted and thus, with a 95% degree of confidence, it can be said that the Finish of a Bowling Ball does not affect its PerfectScale (i.e. it is insignificant).

Thus, Finish is removed from the list of Fixed Factors, and the significance of the LaneCondition is checked.

H_0: The LaneCondition of a Bowling Ball does not affect its PerfectScale

H_1: The LaneCondition of a Bowling Ball affects its PerfectScale

Tests of Between-Subjects Effects

Dependent Variable: PerfectScale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	64148.213[a]	4	16037.053	18.676	.000
Intercept	1423443.412	1	1423443.412	1657.693	.000
LaneCondition	64148.213	4	16037.053	18.676	.000
Error	36064.945	42	858.689		
Total	1744616.470	47			
Corrected Total	100213.157	46			

a. R Squared = .640 (Adjusted R Squared = .606)

Table 6.2.3: Effects of Bowling Ball LaneCondition on PerfectScale

As can be seen in Table 6.2.3, since the p-value is $6.9\text{x}10^{-9}<0.05$ H_0 is rejected and H_1 is accepted and thus, with a 95% degree of confidence, it can be said that the LaneCondition of a Bowling Ball affects its PerfectScale (i.e. it is significant).

Thus, LaneCondition is kept in the list of Fixed Factors, and the significance of the Performance is checked.

H_0: The Performance of a Bowling Ball does not affect its PerfectScale

H_1: The Performance of a Bowling Ball affects its PerfectScale

Tests of Between-Subjects Effects

Dependent Variable: PerfectScale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	85369.455[a]	8	10671.182	27.318	.000
Intercept	1216775.256	1	1216775.256	3114.955	.000
LaneCondition	8987.184	4	2246.796	5.752	.001
Performance	21221.242	4	5305.310	13.582	.000
Error	14843.703	38	390.624		
Total	1744616.470	47			
Corrected Total	100213.157	46			

a. R Squared = .852 (Adjusted R Squared = .821)

Table 6.2.4: Effects of Bowling Ball LaneCondition and Performance on PerfectScale

As can be seen in Table 6.2.4, when the effects of Lane Condition and Performance together on PerfectScale are tested, the significance of LaneCondition changes, this suggests that there is an interaction between LaneCondition and Performance. However, since both Fixed Factors have a p-value<0.05, we can consider both to be significant.

Lastly, the significance of the CoreType is checked.

H_0: The CoreType of a Bowling Ball does not affect its PerfectScale

H_1: The CoreType of a Bowling Ball affects its PerfectScale

Tests of Between-Subjects Effects

Dependent Variable: PerfectScale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	85927.328[a]	9	9547.481	24.728	.000
Intercept	894218.575	1	894218.575	2316.007	.000
LaneCondition	8997.799	4	2249.450	5.826	.001
Performance	14973.734	4	3743.434	9.695	.000
CoreType	557.873	1	557.873	1.445	.237
Error	14285.830	37	386.104		
Total	1744616.470	47			
Corrected Total	100213.157	46			

a. R Squared = .857 (Adjusted R Squared = .823)

Table 6.2.5: Effects of Bowling Ball LaneCondition, Performance and CoreType on PerfectScale

From Table 6.2.5, since the p-value of CoreType is 0.237>0.05, H_0 is accepted and thus, with a 95% degree of confidence, it can be said that the CoreType of a Bowling Ball does not affect its PerfectScale (i.e. it is insignificant).

Therefore, the Fixed Factors which affect (are significant) the PerfectScale of a Bowling Ball are the LaneCondition and the Performance.

Next, the significance of the Covariates on the PerfectScale are checked. First, the significance of the Differential is checked.

H_0: The Differential of a Bowling Ball does not affect its PerfectScale

H_1: The Differential of a Bowling Ball affects its PerfectScale

Tests of Between-Subjects Effects

Dependent Variable: PerfectScale

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	89726.849[a]	9	9969.650	35.177	.000
Intercept	30002.131	1	30002.131	105.860	.000
LaneCondition	10341.869	4	2585.467	9.123	.000
Performance	15826.696	4	3956.674	13.961	.000
Differential	4357.395	1	4357.395	15.375	.000
Error	10486.308	37	283.414		
Total	1744616.470	47			
Corrected Total	100213.157	46			

a. R Squared = .895 (Adjusted R Squared = .870)

Table 6.2.6: Effects of Bowling Ball LaneCondition, Performance and Differential on PerfectScale

From Table 6.2.6, since the p-value of Differential is $3.68\text{x}10^{-4}$<0.05, H_1 is accepted and thus, with a 95% degree of confidence, it can be said that the Differential of a Bowling Ball affects its PerfectScale (i.e. it is significant).

Thus, Differential is kept in the list of significant Covariates, and the significance of the MassBiasDiff is checked.

H_0: The MassBiasDiff of a Bowling Ball does not affect its PerfectScale

H_1: The MassBiasDiff of a Bowling Ball affects its PerfectScale

Tests of Between-Subjects Effects

Dependent Variable: PerfectScale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	89952.596[a]	10	8995.260	31.561	.000
Intercept	29952.646	1	29952.646	105.091	.000
LaneCondition	10183.937	4	2545.984	8.933	.000
Performance	13305.370	4	3326.343	11.671	.000
Differential	3855.349	1	3855.349	13.527	.001
MassBiasDiff	225.747	1	225.747	.792	.379
Error	10260.561	36	285.016		
Total	1744616.470	47			
Corrected Total	100213.157	46			

a. R Squared = .898 (Adjusted R Squared = .869)

Table 6.2.7: Effects of Bowling Ball LaneCondition, Performance and Differential and MassBiasDiff on PerfectScale

From Table 6.2.7, since the p-value of MassBiasDiff is 0.379>0.05, H_0 is accepted and thus, with a 95% degree of confidence, it can be said that the MassBiasDiff of a Bowling Ball does not affect its PerfectScale (i.e. it is insignificant).

Thus, MassBiasDiff is removed from the list of significant Covariates, and the significance of the RG is checked.

H_0: The RG of a Bowling Ball does not affect its PerfectScale

H_1: The RG of a Bowling Ball affects its PerfectScale

Tests of Between-Subjects Effects

Dependent Variable: PerfectScale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	90413.300[a]	10	9041.330	33.214	.000
Intercept	1369.414	1	1369.414	5.031	.031
LaneCondition	6362.903	4	1590.726	5.844	.001
Performance	11321.067	4	2830.267	10.397	.000
Differential	1308.688	1	1308.688	4.807	.035
RG	686.451	1	686.451	2.522	.121
Error	9799.857	36	272.218		
Total	1744616.470	47			
Corrected Total	100213.157	46			

a. R Squared = .902 (Adjusted R Squared = .875)

Table 6.2.8: Effects of Bowling Ball LaneCondition, Performance and Differential and RG on PerfectScale

As can be seen in Table 6.2.8, when the effects of Differential and RG together on PerfectScale are tested, the significance of Differential becomes 0.035, and that of RG 0.121 which would imply that it is insignificant. This suggests that there is an interaction between Differential and RG. When tested individually, both Covariates are found to be significant. However, Differential has an adjusted R^2 value of 0.87, while RG has an adjusted R^2 value of 0.862. Since having Differential as a significant Covariate fits the model better, it is kept in the list of significant Covariates and RG is removed.

Next, the significance of Year is checked.

H_0: The RG of a Bowling Ball does not affect its PerfectScale

H_1: The RG of a Bowling Ball affects its PerfectScale

Tests of Between-Subjects Effects

Dependent Variable: PerfectScale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	91518.278[a]	10	9151.828	37.892	.000
Intercept	1744.118	1	1744.118	7.221	.011
LaneCondition	2411.254	4	602.813	2.496	.060
Performance	4864.496	4	1216.124	5.035	.003
Differential	4455.097	1	4455.097	18.446	.000
Year	1791.429	1	1791.429	7.417	.010
Error	8694.879	36	241.524		
Total	1744616.470	47			
Corrected Total	100213.157	46			

a. R Squared = .913 (Adjusted R Squared = .889)

Table 6.2.9: Effects of Bowling Ball LaneCondition, Performance and Differential and Year on PerfectScale

From Table 6.2.9, since the p-value of Year is $0.01<0.05$, H_0 is rejected and H_1 is accepted and thus, with a 95% degree of confidence, it can be said that the Year of a Bowling Ball affects its PerfectScale (i.e. it is significant).

Lastly, the significance of the Price is checked.

H_0: The Price of a Bowling Ball does not affect its PerfectScale

H_1: The Price of a Bowling Ball affects its PerfectScale

Tests of Between-Subjects Effects

Dependent Variable: PerfectScale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	91974.539[a]	11	8361.322	35.521	.000
Intercept	861.632	1	861.632	3.660	.064
LaneCondition	2631.759	4	657.940	2.795	.041
Performance	3270.916	4	817.729	3.474	.017
Differential	3960.111	1	3960.111	16.824	.000
Year	885.277	1	885.277	3.761	.061
Price	456.261	1	456.261	1.938	.173
Error	8238.619	35	235.389		
Total	1744616.470	47			
Corrected Total	100213.157	46			

a. R Squared = .918 (Adjusted R Squared = .892)

Table 6.2.10: Effects of Bowling Ball LaneCondition, Performance and Differential, Year and Price on PerfectScale

As can be seen in Table 6.2.10, when the effects of Year and Price together on PerfectScale are tested, the significance of Differential becomes 0.061, and that of RG 0.173 which would imply that it is insignificant. This suggests that there is an interaction between Year and Price. When tested individually, both Covariates are found to be significant. However, Year has an adjusted R^2 value of 0.889, while Price has an adjusted R^2 value of 0.884. Since having Year as a significant Covariate fits the model better, it is kept in the list of significant Covariates and Price is removed.

Finally, the significance of 2-Way Interactions between Fixed Factors is checked

H_0: The LaneCondition*Performance of a Bowling Ball does not affect its PerfectScale

H_1: The LaneCondition*Performance of a Bowling Ball affects its PerfectScale

Tests of Between-Subjects Effects

Dependent Variable: PerfectScale

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	93731.877[a]	14	6695.134	33.056	.000
Intercept	1329.376	1	1329.376	6.564	.015
LaneCondition	2994.270	4	748.568	3.696	.014
Performance	3814.225	4	953.556	4.708	.004
Differential	5307.479	1	5307.479	26.205	.000
Year	1367.546	1	1367.546	6.752	.014
LaneCondition * Performance	2213.599	4	553.400	2.732	.046
Error	6481.280	32	202.540		
Total	1744616.470	47			
Corrected Total	100213.157	46			

a. R Squared = .935 (Adjusted R Squared = .907)

*Table 6.2.11: Effects of Bowling Ball LaneCondition, Performance and Differential, Year and LaneCondition*Performance on PerfectScale*

From Table 6.2.11, since the p-value of Year is 0.046<0.05, H_0 is rejected and H_1 is accepted and thus, with a 95% degree of confidence, it can be said that the LaneCondition*Performance of a Bowling Ball affects its PerfectScale (i.e. it is significant).

Therefore, as can be seen in Table 6.2.11, the final model will be based on the following Fixed Factors and Interactions (i.e. Lane Condition, Performance and LaneCondition*Performance) and Covariates (i.e. Differential and Year).

From the Adjusted R Squared value, it is found that the model explains 90.7% of the variability of the data set.

From Table 6.2.12, the model for PerfectScale becomes:

$$\begin{aligned} &Expected\ PerfectScale \\ &\quad = -8579.969 - 8.295(Dry) + 7.458(MediumDry) - 4.822(Medium) \\ &\quad - 1.982(MediumHeavy) - 96.805(Entry) - 38.660(Performance) \\ &\quad - 18.149(Advanced) - 18.435(High) + 35.394(Dry)(Entry) \\ &\quad + 66.463(MediumDry)(Entry) + 10.147(Medium)(High) \\ &\quad - 10.671(MediumHeavy)(High) + 1385.536(Differential) \\ &\quad + 4.335(Year) \end{aligned}$$

Parameter Estimates

Dependent Variable: PerfectScale

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	-8579.969	3363.105	-2.551	.016	-15430.390	-1729.548
[LaneCondition=1.00]	-8.295	24.660	-.336	.739	-58.526	41.936
[LaneCondition=2.00]	7.458	24.031	.310	.758	-41.493	56.408
[LaneCondition=3.00]	-4.822	15.243	-.316	.754	-35.870	26.226
[LaneCondition=4.00]	-1.982	7.475	-.265	.793	-17.208	13.245
[LaneCondition=5.00]	0[a]	.	.	.	.	.
[Performance=1.00]	-96.805	25.442	-3.805	.001	-148.628	-44.982
[Performance=2.00]	-38.660	23.580	-1.640	.111	-86.691	9.371
[Performance=3.00]	-18.149	15.728	-1.154	.257	-50.187	13.889
[Performance=4.00]	-18.435	11.540	-1.597	.120	-41.942	5.072
[Performance=5.00]	0[a]	.	.	.	.	.
Differential	1385.536	270.663	5.119	.000	834.213	1936.859
Year	4.335	1.668	2.598	.014	.937	7.733
[LaneCondition=1.00] * [Performance=1.00]	35.394	23.606	1.499	.144	-12.689	83.478
[LaneCondition=1.00] * [Performance=2.00]	0[a]	.	.	.	.	.
[LaneCondition=2.00] * [Performance=1.00]	66.463	22.685	2.930	.006	20.255	112.671
[LaneCondition=2.00] * [Performance=2.00]	0[a]	.	.	.	.	.
[LaneCondition=3.00] * [Performance=1.00]	0[a]	.	.	.	.	.
[LaneCondition=3.00] * [Performance=2.00]	0[a]	.	.	.	.	.
[LaneCondition=3.00] * [Performance=3.00]	0[a]	.	.	.	.	.
[LaneCondition=3.00] * [Performance=4.00]	10.147	23.238	.437	.665	-37.188	57.481
[LaneCondition=3.00] * [Performance=5.00]	0[a]	.	.	.	.	.
[LaneCondition=4.00] * [Performance=4.00]	-10.671	14.075	-.758	.454	-39.340	17.998
[LaneCondition=4.00] * [Performance=5.00]	0[a]	.	.	.	.	.
[LaneCondition=5.00] * [Performance=4.00]	0[a]	.	.	.	.	.
[LaneCondition=5.00] * [Performance=5.00]	0[a]	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

Table 6.2.12: Parameter Estimates of PerfectScale

By looking at the Studentized Residuals for every entry in the data set, it is discovered that there are 2 cases which are outliers - Balls 5 and 13 ($\because$ $Studentized\ Residuals \notin [-2,2]$).

Next, the Centered Leverage Values are examined, but first, the cut-off point must be calculated using the formula $\frac{2p}{n}$; where p is the number of parameters and n is the sample size. Using $p = 15$ and $n = 47$ the cut-off point is found to be $\frac{2*15}{47} = 0.638$. When looking through the data set for which cases had a Centered Leverage Value higher than the cut-off point (0.638), 4 outliers were found, these being cases 27, 34, 36 and 40.

Finally, by examining the Cook's Distance for all the cases, 2 outliers were found to have an abnormally high distance, these being cases 5 and 13. (Note that both cases were already pointed out to be outliers by the Studentized Residuals.)

The Normality of the Unstandardized Residuals was checked using the Kolmogorov-Smirnov and Shapiro-Wilk tests. As can be seen in Table 6.2.13, the p-value was found to be 0.012<0.05 (from the Shapiro-Wilk test). Thus, with a 95% degree of confidence it is concluded that the Unstandardized Residuals are not normally distributed, and it can hence be stated that this model should not be applied to this data set.

H_0: Residuals are Normally Distributed

H_1: Residuals are not Normally Distributed

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Residual for PerfectScale	.167	47	.002	.936	47	.012

a. Lilliefors Significance Correction

Table 6.2.13: Test for Normality – Unstandardized Residuals

#	Unstandardized Predicted Value	Unstandardized Residual	Studentized Residual	Cook's Distance	Centered Leverage Value
1	234.34	-2.34	-0.18	0	0.14
2	233.55	-0.65	-0.05	0	0.14
3	208.25	16.05	1.29	0.03	0.24
4	143.98	6.52	0.53	0.01	0.26
5	196.63	-29.43	-2.54	0.22	0.34
6	192.57	6.93	0.54	0	0.18
7	208.01	6.39	0.51	0	0.22
8	223.26	6.94	0.53	0	0.14
9	134.44	-10.34	-1.17	0.14	0.61
10	177.41	3.09	0.28	0	0.39
11	189.62	-20.32	-1.57	0.03	0.17
12	210.58	2.72	0.21	0	0.21
13	163.31	-34.61	-3.05	0.35	0.36
14	236.32	-5.42	-0.41	0	0.15
15	236.32	-4.22	-0.32	0	0.15
16	139.65	-21.15	-1.92	0.16	0.4
17	210.96	1.24	0.12	0	0.51
18	186.49	2.11	0.17	0	0.23
19	206.8	7.3	0.59	0.01	0.24
20	239.27	-5.87	-0.45	0	0.16
21	202.27	-2.07	-0.16	0	0.18
22	227.23	7.17	0.54	0	0.13
23	84.33	16.57	1.48	0.09	0.38
24	158.01	-17.41	-1.37	0.03	0.21
25	175.85	15.05	1.17	0.02	0.18
26	231.39	1.01	0.08	0	0.13
27	205.2	0			1
28	232.34	2.16	0.17	0	0.16
29	160.43	23.37	1.82	0.05	0.19
30	199.34	9.56	0.77	0.01	0.24
31	184.08	10.62	0.83	0.01	0.18
32	219.95	-8.05	-0.63	0.01	0.2
33	103.73	-21.73	-1.82	0.09	0.29
34	74.9	0			1
35	225.85	1.55	0.12	0	0.13
36	143.1	0			1
37	230.18	1.62	0.12	0	0.13
38	127.11	-2.11	-0.19	0	0.38
39	77.72	7.28	0.72	0.03	0.5
40	231.5	0			1
41	161.64	18.06	1.6	0.1	0.37
42	167.71	1.89	0.15	0	0.18
43	228.8	-7.9	-0.6	0	0.13

44	132.56	10.34	1.17	0.14	0.61
45	195.54	-1.24	-0.12	0	0.51
46	203.85	11.35	0.9	0.01	0.21
47	234.94	-2.04	-0.15	0	0.14

Table 6.2.14: Residuals and Distances – Used for checking Normality and Finding Outliers

7. Conclusion

Throughout this assignment, for each test done on the Data Set some very interesting results and outcomes were discovered.

In Section 4.2.1 it was discovered that the Brand does not really influence the PerfectScale of a bowling ball, and in Section 4.2.2 it was shown that a higher Performance normally implies a higher Price. In Section 4.3 it was concluded that the Finish of a bowling ball affects what Lane Conditions it is meant to be played on. In Section 4.4 the relation between every pair of Covariates was displayed; one notable result was the positive correlation between Price and PerfectScale, which means that the more expensive a ball is, the better its overall Hook Potential. Section 4.4 was also needed for Regression in Chapter 5.

In Section 5.1, a model for PerfectScale in terms of a covariate (Year) was produced. The model suggested that as the Years have gone by, PerfectScale has always been increasing. This was also shown in Section 3.2 by a Scatter Diagram (Figure 3.2.5). This is due to the fact that over the years further research and product development has been done which allows for the production of better performing bowling balls.

Similarly, In Section 5.2, a model for Price in terms of several covariates was produced. This model suggested that Differential, MassBiasDiff and Year were all positively correlated with Price (already shown in Section 4.4). However, since Balls of CoreType 1 (Symmetric) always have a MassBiasDiff of 0 (zero), we can elicit that Bowling Balls with a symmetric Core are generally cheaper.

In Section 6.1, a model for Price in terms of Fixed Factors was produced. This model showed that with higher Performance categories, the Price increases.

In Section 6.2, a model for PerfectScale against both Covariates and Fixed Factors was produced. In this model, PerfectScale was modelled on LaneCondition, Performance, LaneCondition*Performance, Differential and Year. From the model it was evident that the most significant variable was the Differential,

Two limitations, or areas of improvement, of this study are the following:

Firstly, a bigger sample size can result in more accurate and insightful results.

Secondly, the data set includes specifications of undrilled, or ‘raw’ bowling balls. Using measurements taken from bowling balls which are drilled would result in more realistic results. That being said, it would be very costly and time consuming to take a multitude of measurements (Hand Span, PAP, Pitch, BowlingStyle, …) of different drilling layouts of multiple bowling balls from different bowlers, and the current measurements have proven to be descriptive enough to give a potential buyer a decent idea of what to expect.

8. Appendix – The Data Set

	Brand	Finish	Lane Condition	Differential	Mass Bias Diff	Performance	RG	Core Type	Year	Price	PerfectScale
1	1	1	4	0.056	0.014	5	2.49	2	2016	249.99	232
2	4	3	5	0.054	0.013	5	2.477	2	2016	259.99	232.9
3	2	2	5	0.042	0.016	5	2.5	2	2014	219.99	224.3
4	3	4	2	0.015	0	2	2.65	1	2015	159.99	150.5
5	5	2	2	0.053	0	2	2.53	1	2015	119.99	167.2
6	2	2	3	0.041	0	3	2.5	1	2016	179.99	199.5
7	3	2	4	0.058	0.016	4	2.5	2	2016	239.99	214.4
8	4	3	4	0.048	0.013	5	2.504	2	2016	239.99	230.2
9	3	3	1	0.032	0	2	2.57	1	2011	129.99	124.1
10	4	2	2	0.03	0	1	2.539	1	2016	119.99	180.5
11	3	2	3	0.042	0	3	2.51	1	2015	149.99	169.3
12	1	4	3	0.054	0	3	2.481	1	2016	149.99	213.3
13	3	4	4	0.032	0	4	2.57	1	2014	109.99	128.7
14	2	3	5	0.056	0.011	5	2.48	2	2016	219.99	230.9
15	5	3	5	0.056	0.016	5	2.48	2	2016	249.99	232.1
16	6	1	2	0.009	0	1	2.57	1	2014	109.99	118.5
17	6	3	5	0.051	0	4	2.48	1	2016	219.99	212.2
18	6	2	3	0.046	0	3	2.57	1	2013	179.99	188.6
19	3	2	4	0.054	0	4	2.49	1	2017	239.99	214.1
20	1	4	5	0.055	0.02	5	2.486	2	2017	249.99	233.4
21	5	2	3	0.048	0	3	2.48	1	2016	169.99	200.2
22	6	3	4	0.054	0.018	5	2.48	2	2015	239.99	234.4
23	6	2	1	0.006	0	1	2.69	1	2013	89.99	100.9
24	6	3	2	0.022	0	2	2.57	1	2016	179.99	140.6
25	7	2	2	0.038	0	2	2.5	1	2015	179.99	190.9
26	7	3	4	0.057	0.015	5	2.48	2	2015	259.99	232.4
27	7	2	3	0.043	0.007	4	2.48	2	2016	229.99	205.2
28	3	3	5	0.05	0.012	5	2.51	2	2017	279.99	234.5
29	1	2	2	0.03	0	2	2.562	1	2014	109.99	183.8
30	3	2	4	0.058	0.014	4	2.5	2	2014	239.99	208.9
31	7	2	3	0.038	0.005	3	2.49	2	2015	149.99	194.7
32	5	1	4	0.055	0.007	5	2.5	2	2013	139.99	211.9
33	2	2	1	0.02	0	1	2.65	1	2013	89.99	82
34	4	2	3	0.041	0	1	2.684	1	2007	67.99	74.9
35	7	2	4	0.053	0.016	5	2.49	2	2015	199.99	227.4
36	5	2	3	0.042	0	2	2.54	1	2009	90.99	143.1
37	4	4	4	0.053	0	5	2.584	1	2016	299.99	231.8
38	2	3	1	0.04	0	1	2.46	1	2012	92.99	125
39	5	2	1	0.02	0	1	2.65	1	2007	99.99	85

40	1	2	3	0.056	0.014	5	2.49	2	2016	249.99	231.5
41	4	1	2	0.028	0	1	2.578	1	2013	74.99	179.7
42	7	2	2	0.029	0	2	2.57	1	2016	149.99	169.6
43	2	2	4	0.052	0.007	5	2.47	2	2016	249.99	220.9
44	3	3	1	0.015	0	2	2.65	1	2016	159.99	142.9
45	6	4	5	0.043	0	4	2.5	1	2015	199.99	194.3
46	1	3	4	0.055	0.011	4	2.499	2	2016	199.99	215.2
47	6	3	5	0.055	0.024	5	2.48	2	2016	239.99	232.9

9. References

1. Data Set compiled from Product Pages of Bowling Balls from: http://www.bowlingball.com/shop/all/bowling-balls/
2. Explanation of some bowling ball specifications taken from (used mostly in Chapter 1): https://www.bowlingthismonth.com/bowling-balls/
3. More in depth explanation of what Differential, RG and Mass Bias Diff are: http://www.bowlingball.com/BowlVersity/bowlingball-com-differential-of-rg-specifications