Inferential Statistics

Coded Project

SRINIVASAN T

CONTENT

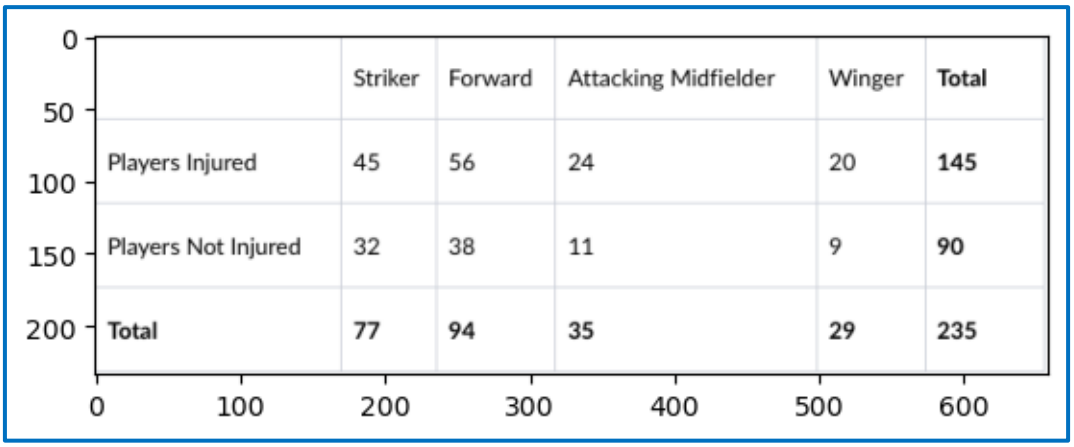
LIST OF FIGURES

LIST OF TABLES

Scoring guide (Rubric) - AS Graded Project (Coded) Rubric Criteria	Points
1.1 What is the probability that a randomly chosen player would suffer an injury?	1
1.2 What is the probability that a player is a forward or a winger?	1
1.3 What is the probability that a randomly chosen player plays in a striker position and has a foot injury?	2
1.4 What is the probability that a randomly chosen injured player is a striker?	2
2.1 What proportion of the gunny bags have a breaking strength less than 3.17 kg per sq cm?	1
2.2 What proportion of the gunny bags have a breaking strength at least 3.6 kg per sq cm.?	1
2.3 What proportion of the gunny bags have a breaking strength between 5 and 5.5 kg per sq cm.?	2
2.4 What proportion of the gunny bags have a breaking strength NOT between 3 and 7.5 kg per sq cm.?	2
3.1 Zingaro has reason to believe that the unpolished stones may not be suitable for printing. Do you think Zingaro is justified in thinking so? "- State the null and alternate hypotheses - Conduct the hypothesis test and compute the p-value - Write down conclusions from the test results Note: Consider the level of significance as 5%."	4
3.2 Is the mean hardness of the polished and unpolished stones the same? - State the null and alternate hypotheses. - Conduct the hypothesis test. - Write down conclusions from the test results. Note: Consider the level of significance as 5%.	4
4.1 How does the hardness of implants vary depending on dentists? "- State the null and alternate hypotheses - Check the assumptions of the hypothesis test. - Conduct the hypothesis test and compute the p-value - Write down conclusions from the test results - In case the implant hardness differs, identify for which pairs it differs Note: 1. Both types of alloys cannot be considered	10

Scoring guide (Rubric) - AS Graded Project (Coded) Rubric Criteria	Points
together. You must conduct the analysis separately for the two types of alloys. 2. Even if the assumptions of the test fail, kindly proceed with the test."	
4.2 How does the hardness of implants vary depending on methods? "- State the null and alternate hypotheses - Check the assumptions of the hypothesis test. - Conduct the hypothesis test and compute the p-value - Write down conclusions from the test results - In case the implant hardness differs, identify for which pairs it differs Note: 1. Both types of alloys cannot be considered together. You must conduct the analysis separately for the two types of alloys. 2. Even if the assumptions of the test fail, kindly proceed with the test."	10
4.3 What is the interaction effect between the dentist and method on the hardness of dental implants for each type of alloy? "- Create Interaction Plot - Inferences from the plot Note: Both types of alloys cannot be considered together. You must conduct the analysis separately for the two types of alloys."	4
4.4 How does the hardness of implants vary depending on dentists and methods together? "- State the null and alternate hypotheses - Check the assumptions of the hypothesis test. - Conduct the hypothesis test and compute the p-value - Write down conclusions from the test results - Identify which dentists and methods combinations are different, and which interaction levels are different. Note: 1. Both types of alloys cannot be considered together. You must conduct the analysis separately for the two types of alloys. 2. Even if the assumptions of the test fail, kindly proceed with the test."	10
Quality of Business Report	6
Points	60

1. Problem Statement 1:



	Striker	Forward	Attacking Midfielder	Winger	Total
Players Injured	45	56	24	20	**145**
Players Not Injured	32	38	11	9	**90**
Total	**77**	**94**	**35**	**29**	**235**

Fig: 1.1 - Male football team dataset

1.1 What is the probability that a randomly chosen player would suffer an injury?

There is 61.7 % chance of a randomly chosen player would suffer an injury

1.2 What is the probability that a player is a forward or a winger?

There is 52.34 % chance of a player would is a forward or a winger

1.3 What is the probability that a randomly chosen player plays in a striker position and has a foot injury?

There is 19.15 % chance of a randomly chosen player plays in a striker position and has an injury

1.4 What is the probability that a randomly chosen injured player is a striker?

There is 31.03 % that a randomly chosen injured player is a striker

2. Problem Statement 2:

The breaking strength of gunny bags used for packaging cement is normally distributed with a mean of 5 kg per sq. centimetre and a standard deviation of 1.5 kg per sq. centimetre. The quality team of the cement company wants to know the following about the packaging material to better understand wastage or pilferage within the supply chain; Answer the questions below based on the given information; (Provide an appropriate visual representation of your answers, without which marks will be deducted)

2.1 What proportion of the gunny bags have a breaking strength less than 3.17 kg per sq.. cm?

The Z score is -1.22
There is 11.12 % chance that the gunny bags have a breaking strength less than 3.17 kg per sq. cm

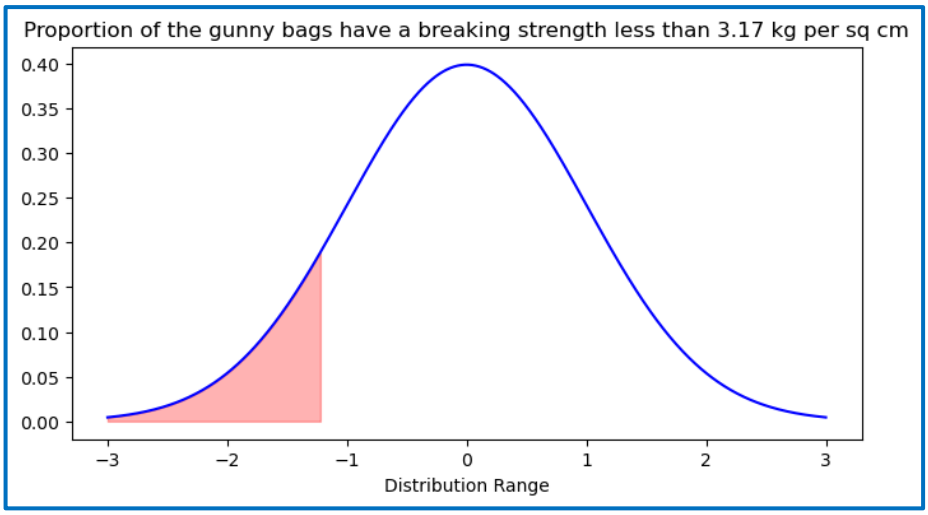


Fig: 1.2 - Proportion of the gunny bags have a breaking strength less than 3.17 kg per sq. cm

2.2 What proportion of the gunny bags have a breaking strength at least 3.6 kg per sq. cm.?
The Z score is -1.0
There is 84.13 % chance that the gunny bags have a breaking strength at least 3.6 kg per sq. cm

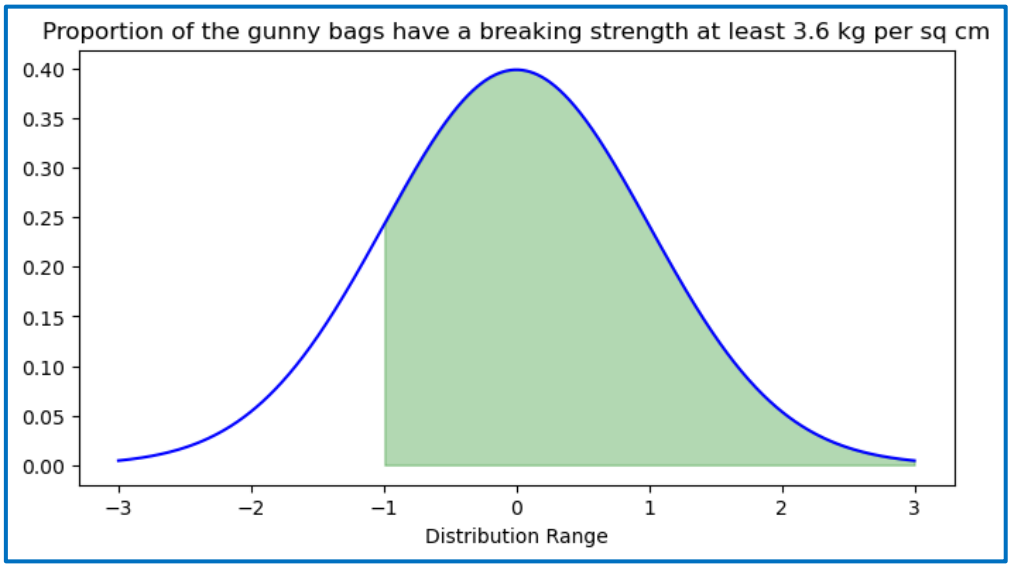


Fig: 1.3 - Proportion of the gunny bags have a breaking strength at least 3.6 kg per sq. cm

2.3 What proportion of the gunny bags have a breaking strength between 5 and 5.5 kg per sq. cm.?

The Z score of 5kg per sq. cm is 0.0
The Z score is 5.5kg per sq. cm 0.3333333333333333
There is 13.06 % chance that the gunny bags have a breaking strength between 5 and 5.5 kg per sq. cm.

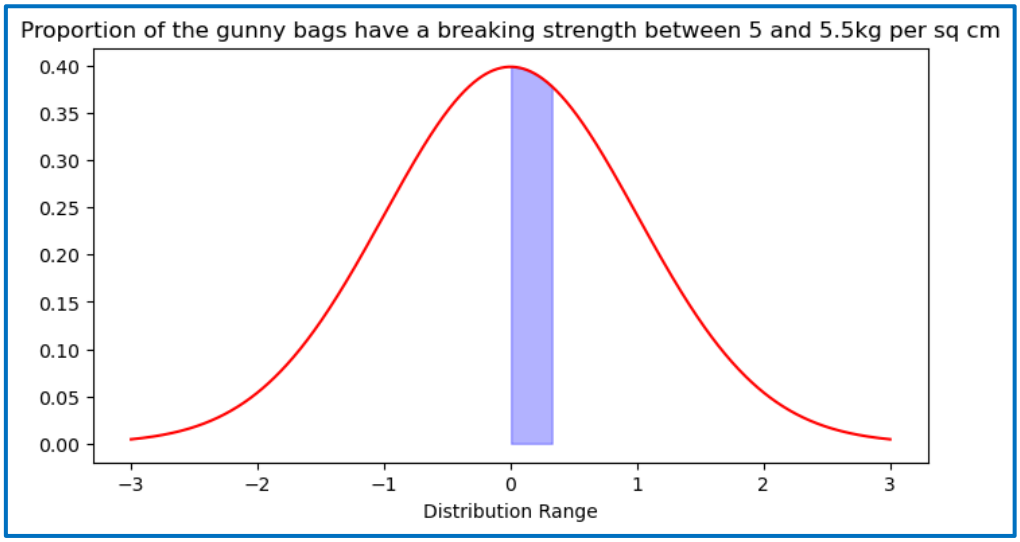


Fig: 1.4 - Proportion of the gunny bags have a breaking strength between 5 and 5.5kg per sq. cm

2.4 What proportion of the gunny bags have a breaking strength NOT between 3 and 7.5 kg per sq. cm.?

The Z score of less than 3kg per sq. cm is -1.3333333333333333
The Z score is greater than 7.5kg per sq. cm 1.6666666666666667
There is 13.9 % chance that the gunny bags have a breaking strength NOT between 3 and 7.5kg per sq. cm.

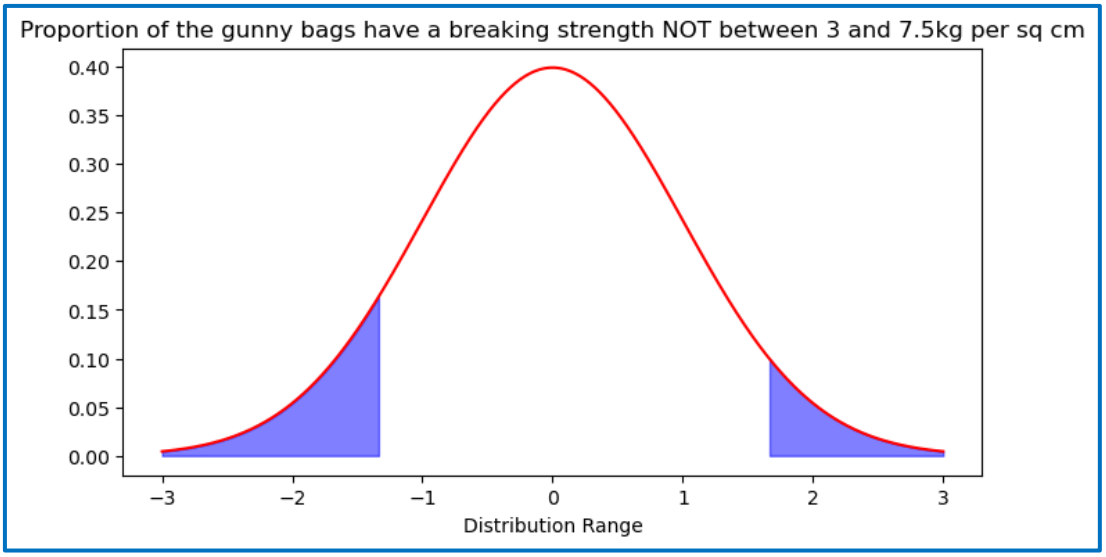


Fig: 1.5 - Proportion of the gunny bags have a breaking strength NOT between 3 and 7.5kg per sq. cm

3. Problem Statement 3:

Zingaro stone printing is a company that specializes in printing images or patterns on polished or unpolished stones. However, for the optimum level of printing of the image, the stone surface has to have a Brinell's hardness index of at least 150. Recently, Zingaro has received a batch of polished and unpolished stones from its clients. Use the data provided to answer the following (assuming a 5% significance level);

	Unpolished	Treated and Polished
0	164.481713	133.209393
1	154.307045	138.482771
2	129.861048	159.665201
3	159.096184	145.663528
4	135.256748	136.789227
...	...	...
70	123.067611	142.293544
71	171.822218	140.124092
72	88.135994	141.393091
73	145.150397	131.370530
74	170.854823	144.502647

75 rows × 2 columns

Table 1.1 - Zingaro dataset for polished and unpolished stones

3.1 Zingaro has reason to believe that the unpolished stones may not be suitable for printing. Do you think Zingaro is justified in thinking so?

3.1.1 State the null and alternate hypothesis

Ho: The unpolished stones may be suitable for printing (where the stone surface has Brinell's hardness index >=150)

Ha: The unpolished stones may not be suitable for printing (where the stone surface has Brinell's hardness index <150)

3.1.2 Conduct the hypothesis test

t-stat value = -4.16
P Value is = 0.0001

3.1.3 Write down conclusions from the test results

Since the p value is less than 0.05, we reject the null hypothesis and conclude that the mean hardness of polished and unpolished stones are not the same.

3.2 Is the mean hardness of the polished and unpolished stones the same?

3.2.1 State null and alternate hypothesis

Ho: Mean hardness of polishes and unpolished stones are the same

Ha: Mean hardness of polishes and unpolished stones are not the same

3.2.2 Hypothesis test

```
t-stat value = -3.24
P Value      =  0.0015
```

3.2.3 Write down the conclusions from the test results

Since the p value is less than 0.05, we reject the null hypothesis and conclude that the mean hardness of polished and unpolished stones are not the same

4. Problem Statement 4

4.1 How does the hardness of implants vary depending on dentists?

Dental implant data: The hardness of metal implants in dental cavities depends on multiple factors, such as the method of implant, the temperature at which the metal is treated, the alloy used as well as the dentists who may favour one method above another and may work better in his/her favourite method. The response is the variable of interest.

Checking whether the data is properly loaded for dataset.

	Dentist	Method	Alloy	Temp	Response
0	1	1	1	1500	813
1	1	1	1	1600	792
2	1	1	1	1700	792
3	1	1	2	1500	907
4	1	1	2	1600	792
...	...	...	...	...	...
85	5	3	1	1600	483
86	5	3	1	1700	405
87	5	3	2	1500	536
88	5	3	2	1600	405
89	5	3	2	1700	312

90 rows × 5 columns

Table 1.2 - Dataset of Dental Implant

	count	mean	std	min	25%	50%	75%	max
Dentist	90.0	3.000000	1.422136	1.0	2.0	3.0	4.0	5.0
Method	90.0	2.000000	0.821071	1.0	1.0	2.0	3.0	3.0
Alloy	90.0	1.500000	0.502801	1.0	1.0	1.5	2.0	2.0
Temp	90.0	1600.000000	82.107083	1500.0	1500.0	1600.0	1700.0	1700.0
Response	90.0	741.777778	145.767845	289.0	698.0	767.0	824.0	1115.0

Table 1.3 - Statistical summary of Dental implant dataset

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 90 entries, 0 to 89
Data columns (total 5 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   Dentist   90 non-null     int64
 1   Method    90 non-null     int64
 2   Alloy     90 non-null     int64
 3   Temp      90 non-null     int64
 4   Response  90 non-null     int64
dtypes: int64(5)
memory usage: 3.6 KB
```

Table 1.4 - Data types of the Dental implant dataset

All the independent variable datatypes need to be changed to object type since python consider that types a int64 for ANOVA

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 90 entries, 0 to 89
Data columns (total 5 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   Dentist   90 non-null     category
 1   Method    90 non-null     category
 2   Alloy     90 non-null     category
 3   Temp      90 non-null     int64
 4   Response  90 non-null     int64
dtypes: category(3), int64(2)
memory usage: 2.3 KB
```

Table 1.5 - Data types of the Dental implant dataset as categorical

Since the analysis need to be conducted for different types of alloys, we'll be separating the into 2 based on the two types of alloys. The Alloy Type 1 will be store in a variable named "Alloy Type1" and The Alloy Type 2 will be store in a variable named "Alloy Type2"

4.1.1 Alloy Type1

	Dentist	Method	Alloy	Temp	Response
0	1	1	1	1500	813
1	1	1	1	1600	792
2	1	1	1	1700	792
6	1	2	1	1500	782
7	1	2	1	1600	698

Table 1.6 - Alloy Type1 dataset of Dentist implant

4.1.1.1 State the null and alternate hypothesis

Ho: Means of all type of Dentists are equal for the Response of Dentist types for alloy type 1
Ha: At least one of the mean of a Dentist type differs for the Response of Dentist types for alloy type 1

Visually we can see the hardness of implant of Response and Dentist types for Alloy Type1 using a boxplot.

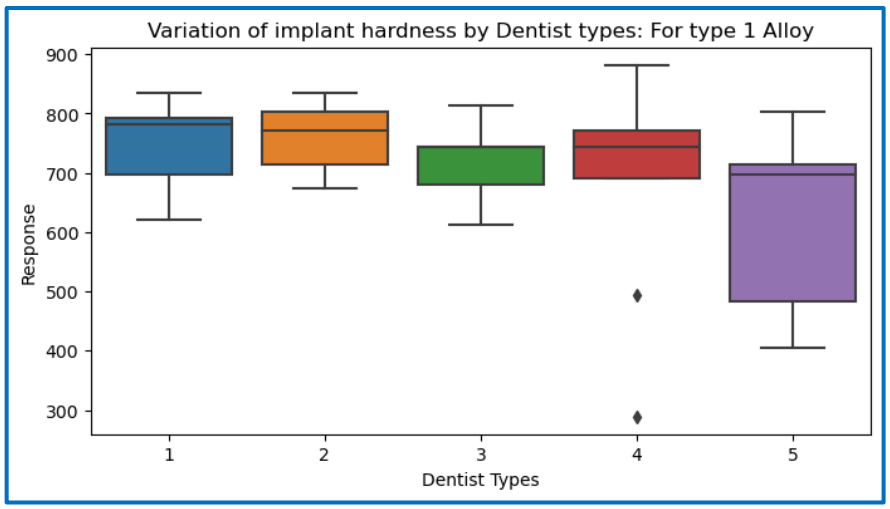


Fig 1.7 - Boxplot for hardness of implants of dentist types of alloy type 1

4.1.1.2 Check the assumptions of the hypothesis test

Checking the normality of the hardness of the implant on Dentist types for Alloy Type1. This can be done using Shapiro's Test

Shapiro's Test:

Stating the Hypothesis:

Ho: The Response follows a normal distribution of all dentist types for alloy type 1
Ha: The Response does not follow a normal distribution for all dentist types for alloy type 1

```
ShapiroResult(statistic=0.9113541841506958, pvalue=0.3254688084125519)
ShapiroResult(statistic=0.9642462134361267, pvalue=0.8415456414222717)
ShapiroResult(statistic=0.8721169233322144, pvalue=0.12953516840934753)
ShapiroResult(statistic=0.8368974328041077, pvalue=0.05333680287003517)
ShapiroResult(statistic=0.8534296751022339, pvalue=0.08127813786268234)
```

All 5 p_value are greater than 0.05 which states that the response for all type of dentists for alloy type1 is normal. Hence the data follow a normal distribution. Ho is True

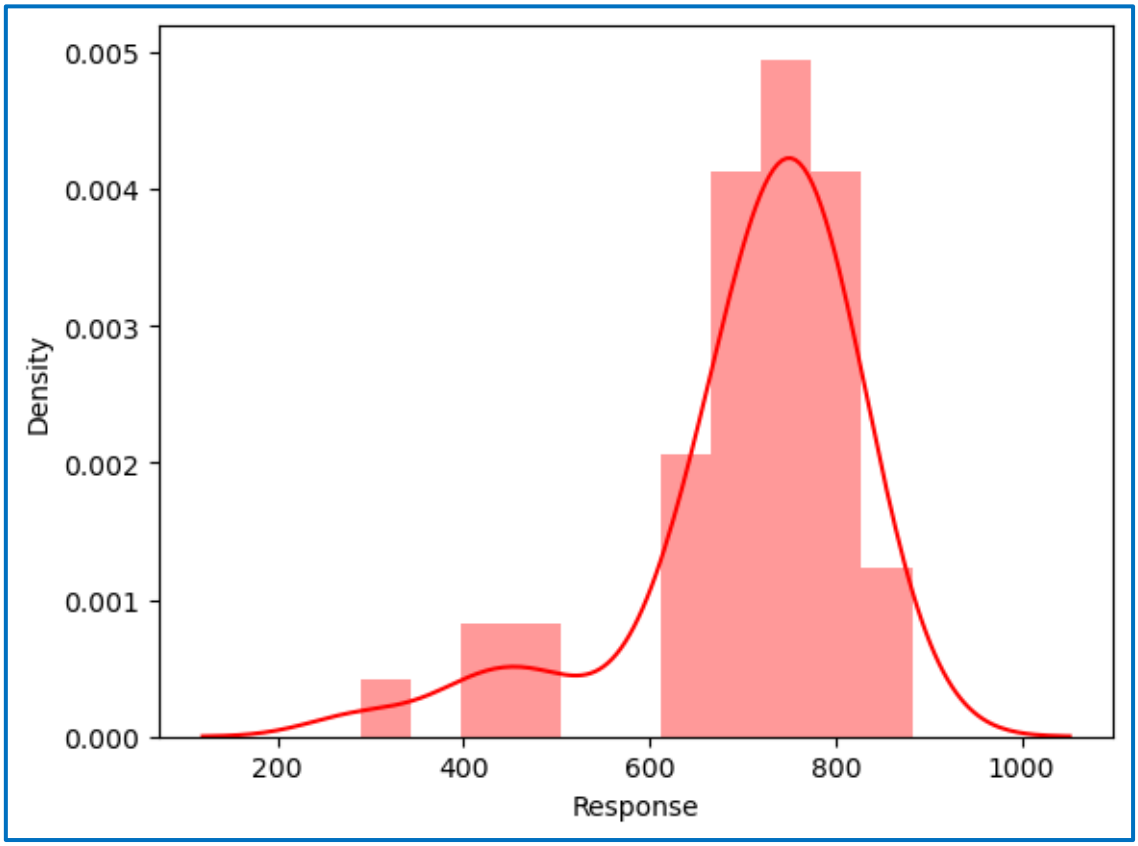


Fig 1.7 - Distribution of Response for Alloy Type 1

Checking the Homogeneity of Variance of the hardness of the implant on Dentist types for Alloy Type 1 using Levene's test

Levene's Test:

Stating the Hypothesis:

Ho: All the variance for Response are equal of all dentist types for alloy type 1
Ha: At least one variance for Response is not equal of all dentist types for alloy type 1

```
Statistics =  1.385
P_value    =  0.257
```

The p value 0.257 is greater than 0.05 and we conclude that the Ho is true as there is no difference between the variance for Response are equal of all dentist types for alloy type 1

4.1.1.3 Conduct the hypothesis test and compute the p-value

The one way ANOVA testing is done using the ols() & anova.lm() functions on the independent (Dentist) and dependent variable (Response)

	df	sum_sq	mean_sq	F	PR(>F)
Dentist	4.0	106683.688889	26670.922222	1.977112	0.116567
Residual	40.0	539593.555556	13489.838889	NaN	NaN

Table 1.7 - ONE WAY ANOVA table for hardness of implant of Dentist types for alloy type 1

4.1.1.4 Write down conclusions from the test results.

p_value is 0.1165 which is greater than 0.05 where we fail to reject null hypothesis (Ho is True). Means of all type of Dentists are equal of the Response for alloy type 1.

4.1.1.5 In case the implant hardness differs, identify for which pairs it differs

NO pairs differs on the implant hardness of all Dentist types for alloy type 1

4.1.2 Alloy Type2

Checking whether the data is properly loaded for Alloy Type2

	Dentist	Method	Alloy	Temp	Response
3	1	1	2	1500	907
4	1	1	2	1600	792
5	1	1	2	1700	835
9	1	2	2	1500	1115
10	1	2	2	1600	835

Table 1.8 - Alloy Type2 dataset of Dentist implant

4.1.2.1 State the null and alternate hypothesis

Ho: Means of all type of Dentists are equal for the Response of Dentist types for alloy type 2
Ha: At least one of the mean of a Dentist type differs for the Response of Dentist types for alloy type 2

Visually we can see the hardness of implant of Response and Dentist types for Alloy Type1 using a boxplot.

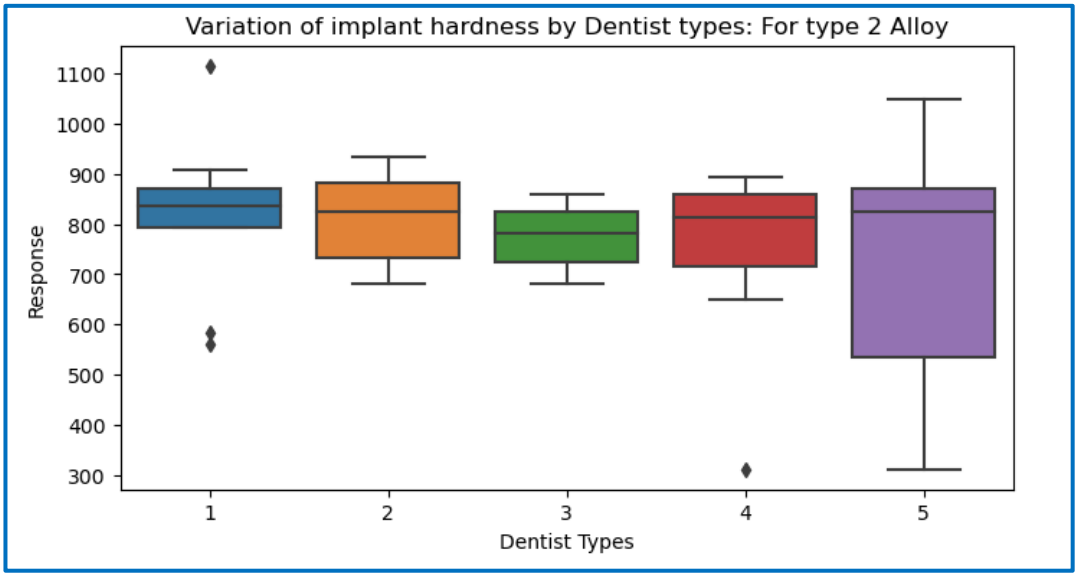


Fig 1.8 - Boxplot for hardness of implants of dentist types of alloy type 2

4.1.2.2 Check the assumptions of the hypothesis test

Checking the normality of the hardness of the implant on Dentist types for Alloy Type2. This can be done using Shapiro's Test

Shapiro's Test:

Stating the Hypothesis:

Ho: The Response follows a normal distribution of all dentist types for alloy type 2
Ha: The Response does not follow a normal distribution for all dentist types for alloy type 2

```
ShapiroResult(statistic=0.9039731621742249, pvalue=0.27593979239463806)
ShapiroResult(statistic=0.9392004013061523, pvalue=0.5735077857971191)
ShapiroResult(statistic=0.9340971112251282, pvalue=0.5213080644607544)
ShapiroResult(statistic=0.7613219022750854, pvalue=0.007332688197493553)
ShapiroResult(statistic=0.9131584167480469, pvalue=0.33861100673675537)
```

80% of p values are greater than 0.05 which states that the response for the different type of dentists for alloy type2 is normal. Hence the data follow a normal distribution. Ho is True.

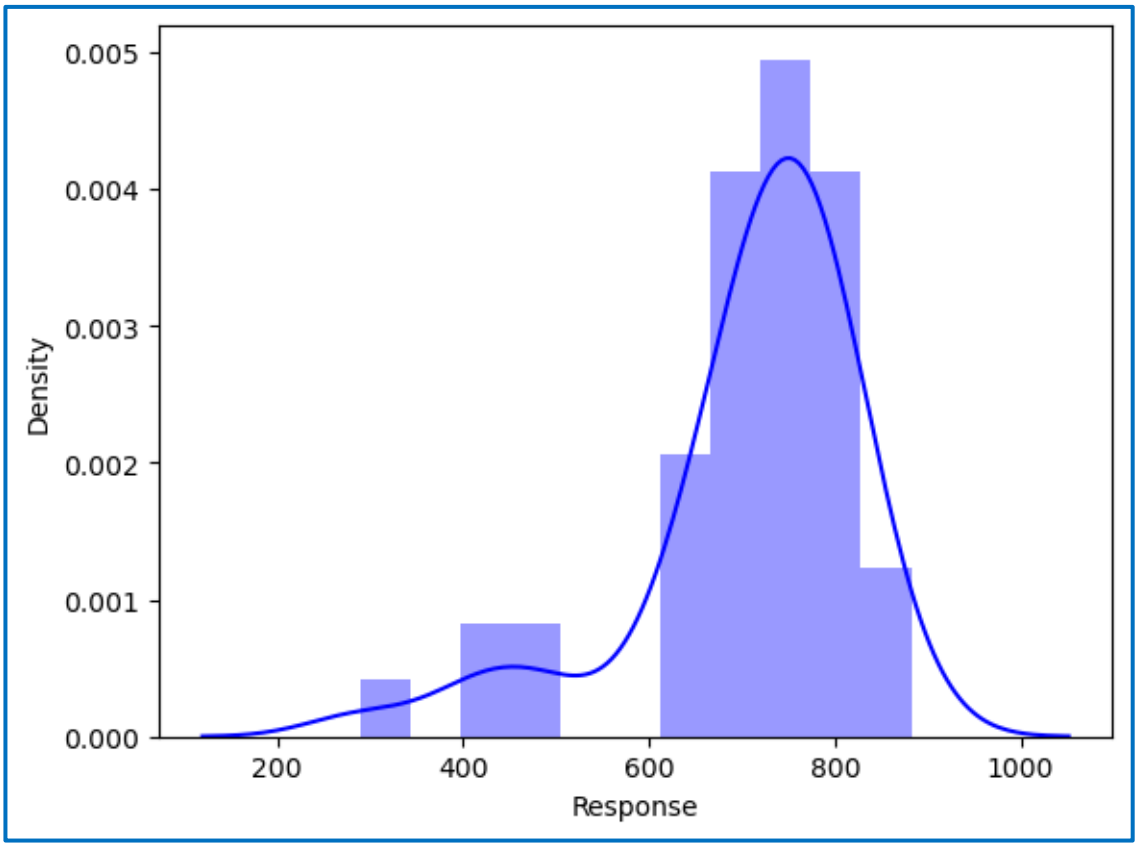


Fig 1.9 - Distribution of Response for Alloy Type 2

Checking the Homogeneity of Variance of the hardness of the implant on Dentist types for Alloy Type 2 using Levene's test.

Levene's Test:

Stating the Hypothesis:

Ho: All the variance for Response are equal of all dentist types for alloy type 2
Ha: At least one variance for Response is not equal of all dentist types for alloy type 2

```
Statistics =  1.446
P_value    =  0.237
```

The p value 0.237 is greater than 0.05 and we conclude that the Ho is true as there is no difference between the variance for Response are equal of all dentist types for alloy type 2.

4.1.2.3 Write down conclusions from the test results

The one way ANOVA testing is done using the ols() & anova.lm() functions on the independent (Dentist) and dependent variable (Response)

	df	sum_sq	mean_sq	F	PR(>F)
Dentist	4.0	5.679791e+04	14199.477778	0.524835	0.718031
Residual	40.0	1.082205e+06	27055.122222	NaN	NaN

Table 1.9 - ONE WAY ANOVA table for hardness of implant of Dentist types for alloy type 2

4.1.2.4 In case the implant hardness differs, identify for which pairs it differs

p_value is 0.7180 which is greater than 0.05 where we fail to reject null hypothesis (Ho is True). Means of all type of Dentists are equal of the Response for alloy type 2

4.2 How does the hardness of implants vary depending on methods?

Dental implant data: The hardness of metal implants in dental cavities depends on multiple factors, such as the method of implant, the temperature at which the metal is treated, the alloy used as well as the dentists who may favour one method above another and may work better in his/her favourite method. The response is the variable of interest.

4.2.1 Alloy Type1

Checking whether the data is properly loaded for Alloy Type1

	Dentist	Method	Alloy	Temp	Response
0	1	1	1	1500	813
1	1	1	1	1600	792
2	1	1	1	1700	792
6	1	2	1	1500	782
7	1	2	1	1600	698

Table 1.10 - Alloy Type1 dataset of hardness implant

4.2.1.1 State the null and alternate hypotheses

Ho: Means of all type of Dentists are equal for the Response of Dentist types for alloy type 1
Ha: At least one of the mean of a Dentist type differs for the Response of Dentist types for alloy type 1

Visually we can see the hardness of implant of Response and Dentist types for Alloy Type1 using a boxplot.

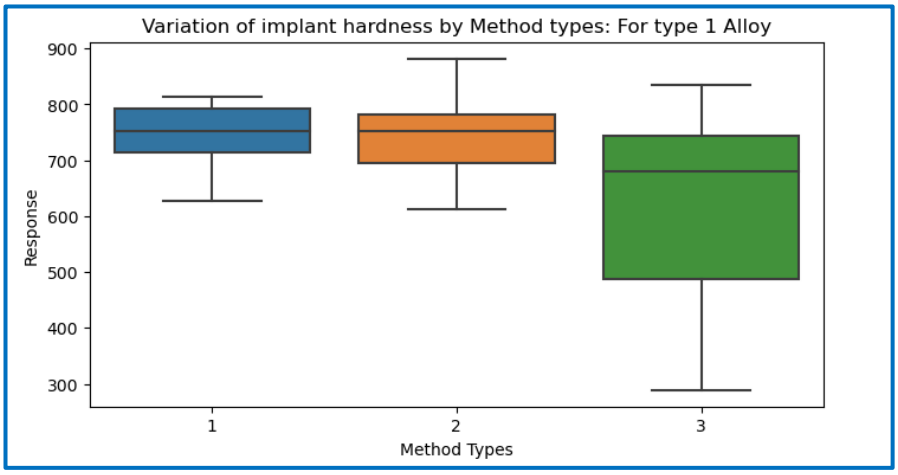


Fig 1.10 - Boxplot for hardness of implants of dentist types of alloy type 2

4.2.1.2 Check the assumptions of the hypothesis test

Checking the normality of the hardness of the implant on Methods types for Alloy Type1. This can be done using Shapiro's Test

Shapiro's Test:

Stating the Hypothesis:

Ho: The Response follows a normal distribution of all method types for alloy type 1
Ha: The Response does not follow a normal distribution for all method types for alloy type 1

```
ShapiroResult(statistic=0.9183822870254517, pvalue=0.18198540806770325)
ShapiroResult(statistic=0.9732585549354553, pvalue=0.9030335545539856)
ShapiroResult(statistic=0.9114548563957214, pvalue=0.14254699647426605)
```

All 3 p_value are greater than 0.05 which states that the response for all type of method for alloy type1 is normal. Hence the data follow a normal distribution. Ho is True

Levene's Test:

Stating the Hypothesis:

Ho: All the variance for Response are equal of all method types for alloy type 1
Ha: At least one variance for Response is not equal of all method types for alloy type 1

```
Statistics =  6.521
P_value    =  0.003
```

The p_value 0.003 is lesser than 0.05 and we conclude that the H1 is true. Therefore, there is difference between the variance for Response which are not equal of all method types for alloy type 1.

4.2.1.3 Conduct the hypothesis test and compute the p-value

The one way ANOVA testing is done using the ols() & anova.lm() functions on the independent (Method) and dependent variable (Response)

	df	sum_sq	mean_sq	F	PR(>F)
Method	2.0	148472.177778	74236.088889	6.263327	0.004163
Residual	42.0	497805.066667	11852.501587	NaN	NaN

Table 1.11 - ONE WAY ANOVA table for hardness of implant of Method types for alloy type 1

4.2.1.4 Write down conclusions from the test results

P value is 0.004163 which is lesser than 0.05 where we reject the null hypothesis (Ha is True). At least one type of Method is not equal of the Response for alloy type 1

4.2.1.5 In case the implant hardness differs, identify for which pairs it differs

To identify the pairs which differ from each other we conduct the Tukey's Pair Test using MultiComparsion() and MultiComp.tukeyhsd() functions

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1	2	-6.1333	0.987	-102.714	90.4473	False
1	3	-124.8	0.0085	-221.3807	-28.2193	True
2	3	-118.6667	0.0128	-215.2473	-22.086	True

Table 1.12 - Tukey's HSD comparison table for response and methods on Alloy Type 1

The pairs (1,3) & (2,3) methods differs as the p value is lesser than 0.05

4.2.2 Alloy Type2

Checking whether the data is properly loaded for Alloy Type2

	Dentist	Method	Alloy	Temp	Response
3	1	1	2	1500	907
4	1	1	2	1600	792
5	1	1	2	1700	835
9	1	2	2	1500	1115
10	1	2	2	1600	835

Table 1.14 - Alloy Type 2 dataset of hardness implant

4.2.2.1 State the null and alternate hypothesis

Ho: Means of all type of Dentists are equal for the Response of Dentist types for alloy type 1
Ha: At least one of the mean of a Dentist type differs for the Response of Dentist types for alloy type 1

Visually we can see the hardness of implant of Response and Dentist types for Alloy Type1 using a boxplot.

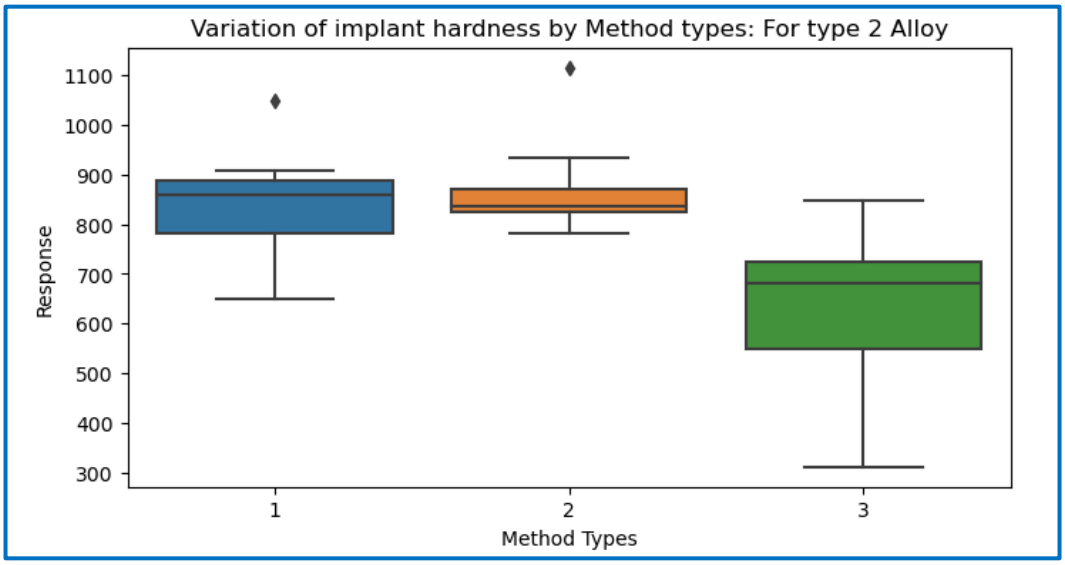


Fig 1.11 - Boxplot for hardness of implants of dentist types of alloy type 2

4.2.2.2 Check the assumptions of the hypothesis test

Checking the normality of the hardness of the implant on Methods types for Alloy Type2. This can be done using Shapiro's Test

Shapiro's Test:

Stating the Hypothesis:

Ho: The Response follows a normal distribution of all method types for alloy type 2
Ha: The Response does not follow a normal distribution for all method types for alloy type 2

```
ShapiroResult(statistic=0.963810384273529, pvalue=0.7582374811172485)
ShapiroResult(statistic=0.755793035030365, pvalue=0.001051110913977027)
ShapiroResult(statistic=0.9021322131156921, pvalue=0.1025901660323143)
```

66.6% p values are greater than 0.05 which states that the response for all type of method for alloy type2 is normal. Hence the data follow a normal distribution. Ho is True

Levene's Test:

Stating the Hypothesis:

Ho: All the variance for Response are equal of all method types for alloy type 2
Ha: At least one variance for Response is not equal of all method types for alloy type 2

```
Statistics =  3.35
P_value    =  0.045
```

The p value 0.045 is lesser than 0.05 and we conclude that the H1 is true. Therefore, there is difference between the variance for Response which are not equal of all method types for alloy type 2.

4.2.2.3 Conduct the hypothesis test and compute the p-value

The one way ANOVA testing is done using the ols() & anova.lm() functions on the independent (Method) and dependent variable (Response)

	df	sum_sq	mean_sq	F	PR(>F)
Method	2.0	499640.4	249820.200000	16.4108	0.000005
Residual	42.0	639362.4	15222.914286	NaN	NaN

Table 1.15 - ONE WAY ANOVA table for hardness of implant of Method types for alloy type 2

4.2.2.4 Write down conclusions from the test results

p_value is 0.000005 which is lesser than 0.05 where we reject the null hypothesis (Ha is True). At least one type of Method is not equal of the Response for alloy type 2

4.2.2.5 In case the implant hardness differs, identify for which pairs it differs.

To identify the pairs which differ from each other we conduct the Tukey's Pair Test using MultiComparsion() and MultiComp.tukeyhsd() functions

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1	2	27.0	0.8212	-82.4546	136.4546	False
1	3	-208.8	0.0001	-318.2546	-99.3454	True
2	3	-235.8	0.0	-345.2546	-126.3454	True

Table 1.16 - Tukey's HSD comparison table for response and methods on Alloy Type 2

The pairs (1,3) & (2,3) methods differs as the p value is lesser than 0.05.

4.3 What is the interaction between the dentist and method on the hardness of dental implants for each type of alloy?

	Dentist	Method	Alloy	Temp	Response
0	1	1	1	1500	813
1	1	1	1	1600	792
2	1	1	1	1700	792
3	1	1	2	1500	907
4	1	1	2	1600	792

Table 1.17 - Implant hardness dataset

For Interaction plot for two type of alloys we will create two different dataset based on the alloy type and storing the same in new variables named Alloy_type3 & Alloy_type4

Checking whether the data is properly loaded for Alloy Type1 with resetting the index

	index	Dentist	Method	Alloy	Temp	Response
0	0	1	1	1	1500	813
1	1	1	1	1	1600	792
2	2	1	1	1	1700	792
3	6	1	2	1	1500	782
4	7	1	2	1	1600	698

Table 1.18 - Alloy Type 1 dataset of hardness implant

Checking whether the data is properly loaded for Alloy Type1 with resetting the index

	index	Dentist	Method	Alloy	Temp	Response
0	3	1	1	2	1500	907
1	4	1	1	2	1600	792
2	5	1	1	2	1700	835
3	9	1	2	2	1500	1115
4	10	1	2	2	1600	835

Table 1.19 - Alloy Type 2 dataset of hardness implant

4.3.1 Create Interaction Plot – Alloy Type1

The interaction plot is done using the interaction_plot() function with Dentist as x, Method as trace and Response as response for Alloy Type 1

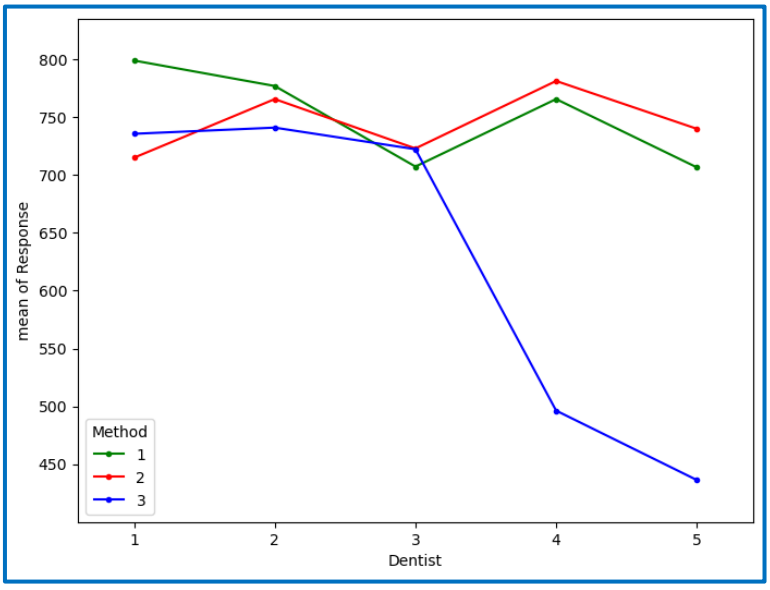


Fig 1.12 - Interaction plot between Dentist, Methods & Response for alloy type 1

4.3.1.1 Inferences

From the above figure as the lines intersect with each other the Dentist types interacts at all the Method type.

4.3.2 Create Interaction Plot – Alloy Type2

The interaction plot is done using the interaction_plot() function with Dentist as x, Method as trace and Response as response for Alloy Type 1

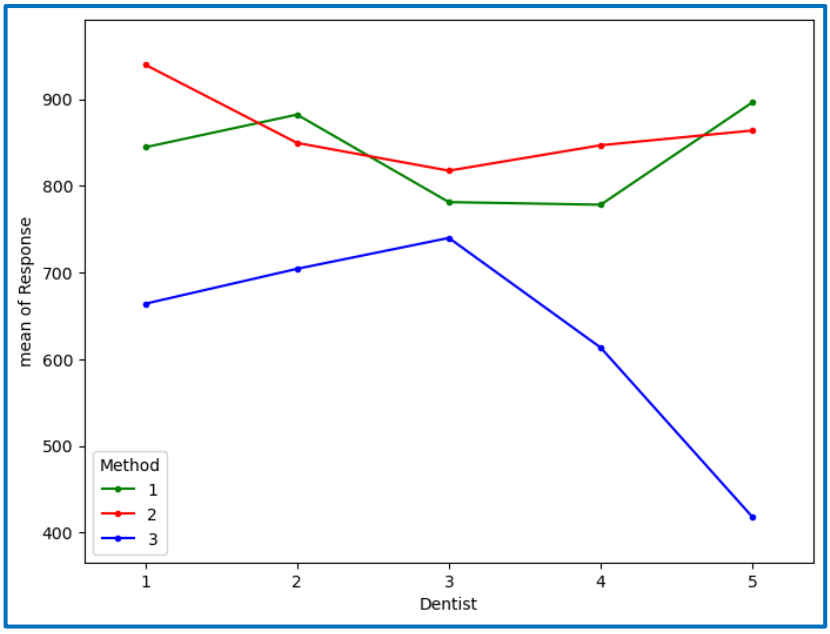


Fig 1.13 - Interaction plot between Dentist, Methods & Response for alloy type 2

4.3.2.1 Inferences

Method 1&2 on all dentist types interacts but the Method 3 for all dentist type does not interact.

4.4 What is the interaction effect between the dentist and method on the hardness of dental implants for each type of alloy?

4.4.1 Alloy Type1

4.4.1.1 State the null and alternate hypothesis

Ho: Types of Dentists and Methods together have equal effect or mean on the response (hardness of the implant) for Alloy Type 1

Ha: Types of Dentists and Methods together does not have a effect or mean not equal on the response (hardness of the implant) for Alloy Type 1

Visually we can see the hardness of implant of Dentist and Method types for Alloy Type1 using a boxplot

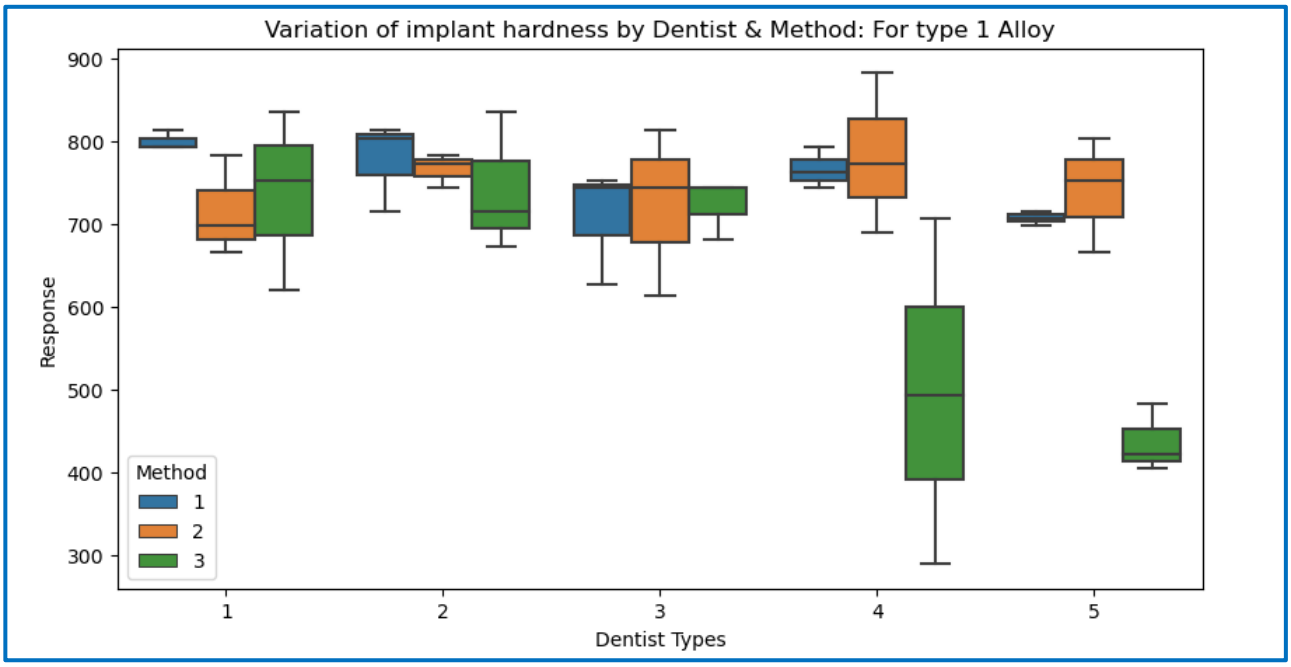


Fig 1.14 - Hardness of implant on Dentist, Methods & Response for alloy type 1

4.4.1.2 Check the assumptions of the hypothesis test

Checking the normality of the hardness of the implant for Alloy Type1. This can be done using Shapiro's Test

Shapiro's Test:

Stating the Hypothesis:

Ho: The Response follows a normal distribution for all the combinations of Dentists and Methods for Alloy Type 1
Ha: The Response does not follow a normal distribution for all the combinations of Dentists and Methods for Alloy Type 1

Checking the Homogeneity of Variance of the hardness of the implant for type 1 Alloy

Levene's Test:

Stating the Hypothesis:

Ho: All the combinations of Dentists and Methods have equal variance for Alloy Type 1
Ha: At least one combinations of Dentists and Methods does not have equal variance for Alloy Type 1

```
Statistics = 1.22
P_value    = 0.31
```

The p value is 0.31 which is greater than 0.05, hence we fail to reject the null hypothesis. Ho is True. All the combinations of Dentists and Methods have equal variance for Alloy Type 1

4.4.1.3 Conduct the hypothesis test and compute the p-value

Two Way ANOVA without interaction of Dentist & Method types using the functions ols() & anova_lm()

	df	sum_sq	mean_sq	F	PR(>F)
Dentist	4.0	106683.688889	26670.922222	2.591255	0.051875
Method	2.0	148472.177778	74236.088889	7.212522	0.002211
Residual	38.0	391121.377778	10292.667836	NaN	NaN

Table 1.20 - Two Way ANOVA for the types Dentist and Methods w.r.t Response for Alloy Type 1

Two Way ANOVA with interaction of Dentist & Method types using the functions ols() & anova_lm()

	df	sum_sq	mean_sq	F	PR(>F)
Dentist	4.0	106683.688889	26670.922222	3.899638	0.011484
Method	2.0	148472.177778	74236.088889	10.854287	0.000284
Dentist:Method	8.0	185941.377778	23242.672222	3.398383	0.006793
Residual	30.0	205180.000000	6839.333333	NaN	NaN

Table 1.21 - Two Way ANOVA for the types Dentist and Methods and Interaction of Dentist and Methods w.r.t Response for Alloy Type 1

4.4.1.4 Write down conclusions from the test results

The p value for Dentist and Method together is 0.006793 which is less than 0.05, hence Ha is True. We reject the null hypothesis. And we can conclude that for the Types of Dentists and Methods together does not have an effect or mean not equal on the response (hardness of the implant) for Alloy Type 1

4.4.1.5 In case the implant hardness differs, identify for which pairs it differs

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1:1	1:2	-84.0	0.9933	-332.8283	164.8283	False
1:1	1:3	-63.3333	0.9996	-312.1617	185.495	False
1:1	2:1	-22.0	1.0	-270.8283	226.8283	False
1:1	2:2	-33.3333	1.0	-282.1617	215.495	False
1:1	2:3	-58.0	0.9999	-306.8283	190.8283	False
1:1	3:1	-91.6667	0.9853	-340.495	157.1617	False
1:1	3:2	-76.0	0.9975	-324.8283	172.8283	False
1:1	3:3	-76.6667	0.9972	-325.495	172.1617	False
1:1	4:1	-33.3333	1.0	-282.1617	215.495	False
1:1	4:2	-17.6667	1.0	-266.495	231.1617	False
1:1	4:3	-302.6667	0.007	-551.495	-53.8383	True
1:1	5:1	-92.3333	0.9844	-341.1617	156.495	False
1:1	5:2	-59.0	0.9998	-307.8283	189.8283	False
1:1	5:3	-362.6667	0.0007	-611.495	-113.8383	True
1:2	1:3	20.6667	1.0	-228.1617	269.495	False
1:2	2:1	62.0	0.9997	-186.8283	310.8283	False
1:2	2:2	50.6667	1.0	-198.1617	299.495	False
1:2	2:3	26.0	1.0	-222.8283	274.8283	False
1:2	3:1	-7.6667	1.0	-256.495	241.1617	False
1:2	3:2	8.0	1.0	-240.8283	256.8283	False
1:2	3:3	7.3333	1.0	-241.495	256.1617	False
1:2	4:1	50.6667	1.0	-198.1617	299.495	False
1:2	4:2	66.3333	0.9994	-182.495	315.1617	False
1:2	4:3	-218.6667	0.1324	-467.495	30.1617	False
1:2	5:1	-8.3333	1.0	-257.1617	240.495	False
1:2	5:2	25.0	1.0	-223.8283	273.8283	False
1:2	5:3	-278.6667	0.0173	-527.495	-29.8383	True
1:3	2:1	41.3333	1.0	-207.495	290.1617	False

1:3	2:2	30.0	1.0	-218.8283	278.8283	False
1:3	2:3	5.3333	1.0	-243.495	254.1617	False
1:3	3:1	-28.3333	1.0	-277.1617	220.495	False
1:3	3:2	-12.6667	1.0	-261.495	236.1617	False
1:3	3:3	-13.3333	1.0	-262.1617	235.495	False
1:3	4:1	30.0	1.0	-218.8283	278.8283	False
1:3	4:2	45.6667	1.0	-203.1617	294.495	False
1:3	4:3	-239.3333	0.0688	-488.1617	9.495	False
1:3	5:1	-29.0	1.0	-277.8283	219.8283	False
1:3	5:2	4.3333	1.0	-244.495	253.1617	False
1:3	5:3	-299.3333	0.0079	-548.1617	-50.505	True
2:1	2:2	-11.3333	1.0	-260.1617	237.495	False
2:1	2:3	-36.0	1.0	-284.8283	212.8283	False
2:1	3:1	-69.6667	0.999	-318.495	179.1617	False
2:1	3:2	-54.0	0.9999	-302.8283	194.8283	False
2:1	3:3	-54.6667	0.9999	-303.495	194.1617	False
2:1	4:1	-11.3333	1.0	-260.1617	237.495	False
2:1	4:2	4.3333	1.0	-244.495	253.1617	False
2:1	4:3	-280.6667	0.016	-529.495	-31.8383	True
2:1	5:1	-70.3333	0.9989	-319.1617	178.495	False
2:1	5:2	-37.0	1.0	-285.8283	211.8283	False
2:1	5:3	-340.6667	0.0016	-589.495	-91.8383	True
2:2	2:3	-24.6667	1.0	-273.495	224.1617	False
2:2	3:1	-58.3333	0.9999	-307.1617	190.495	False
2:2	3:2	-42.6667	1.0	-291.495	206.1617	False
2:2	3:3	-43.3333	1.0	-292.1617	205.495	False
2:2	4:1	0.0	1.0	-248.8283	248.8283	False
2:2	4:2	15.6667	1.0	-233.1617	264.495	False
2:2	4:3	-269.3333	0.0243	-518.1617	-20.505	True
2:2	5:1	-59.0	0.9998	-307.8283	189.8283	False
2:2	5:2	-25.6667	1.0	-274.495	223.1617	False
2:2	5:3	-329.3333	0.0025	-578.1617	-80.505	True
2:3	3:1	-33.6667	1.0	-282.495	215.1617	False
2:3	3:2	-18.0	1.0	-266.8283	230.8283	False
2:3	3:3	-18.6667	1.0	-267.495	230.1617	False
2:3	4:1	24.6667	1.0	-224.1617	273.495	False
2:3	4:2	40.3333	1.0	-208.495	289.1617	False
2:3	4:3	-244.6667	0.0576	-493.495	4.1617	False
2:3	5:1	-34.3333	1.0	-283.1617	214.495	False
2:3	5:2	-1.0	1.0	-249.8283	247.8283	False
2:3	5:3	-304.6667	0.0065	-553.495	-55.8383	True
3:1	3:2	15.6667	1.0	-233.1617	264.495	False
3:1	3:3	15.0	1.0	-233.8283	263.8283	False
3:1	4:1	58.3333	0.9999	-190.495	307.1617	False
3:1	4:2	74.0	0.9981	-174.8283	322.8283	False
3:1	4:3	-211.0	0.166	-459.8283	37.8283	False
3:1	5:1	-0.6667	1.0	-249.495	248.1617	False
3:1	5:2	32.6667	1.0	-216.1617	281.495	False
3:1	5:3	-271.0	0.0229	-519.8283	-22.1717	True
3:2	3:3	-0.6667	1.0	-249.495	248.1617	False
3:2	4:1	42.6667	1.0	-206.1617	291.495	False
3:2	4:2	58.3333	0.9999	-190.495	307.1617	False
3:2	4:3	-226.6667	0.1035	-475.495	22.1617	False
3:2	5:1	-16.3333	1.0	-265.1617	232.495	False
3:2	5:2	17.0	1.0	-231.8283	265.8283	False
3:2	5:3	-286.6667	0.0128	-535.495	-37.8383	True
3:3	4:1	43.3333	1.0	-205.495	292.1617	False
3:3	4:2	59.0	0.9998	-189.8283	307.8283	False
3:3	4:3	-226.0	0.1057	-474.8283	22.8283	False
3:3	5:1	-15.6667	1.0	-264.495	233.1617	False
3:3	5:2	17.6667	1.0	-231.1617	266.495	False
3:3	5:3	-286.0	0.0131	-534.8283	-37.1717	True
4:1	4:2	15.6667	1.0	-233.1617	264.495	False
4:1	4:3	-269.3333	0.0243	-518.1617	-20.505	True
4:1	5:1	-59.0	0.9998	-307.8283	189.8283	False
4:1	5:2	-25.6667	1.0	-274.495	223.1617	False

4:1	5:3	-329.3333	0.0025	-578.1617	-80.505	True
4:2	4:3	-285.0	0.0137	-533.8283	-36.1717	True
4:2	5:1	-74.6667	0.9979	-323.495	174.1617	False
4:2	5:2	-41.3333	1.0	-290.1617	207.495	False
4:2	5:3	-345.0	0.0013	-593.8283	-96.1717	True
4:3	5:1	210.3333	0.1692	-38.495	459.1617	False
4:3	5:2	243.6667	0.0596	-5.1617	492.495	False
4:3	5:3	-60.0	0.9998	-308.8283	188.8283	False
5:1	5:2	33.3333	1.0	-215.495	282.1617	False
5:1	5:3	-270.3333	0.0234	-519.1617	-21.505	True
5:2	5:3	-303.6667	0.0067	-552.495	-54.8383	True

Table 1.22 - Interaction table of dentist and methods together for alloy type 1

The pairs which has the reject status as True or the pairs with p-adj value less than 0.05 are the pairs which differs in implant hardness from rest of the pairs. There are totally 18 pairs which differs.

4.4.2 Alloy Type2

4.4.2.1 State the null and alternate hypothesis

Ho: Types of Dentists and Methods together have equal effect or mean on the response (hardness of the implant) for Alloy Type 1

Ha: Types of Dentists and Methods together does not have an effect or mean not equal on the response (hardness of the implant) for Alloy Type 1

Visually we can see the hardness of implant of Dentist and Method types for Alloy Type1 using a boxplot

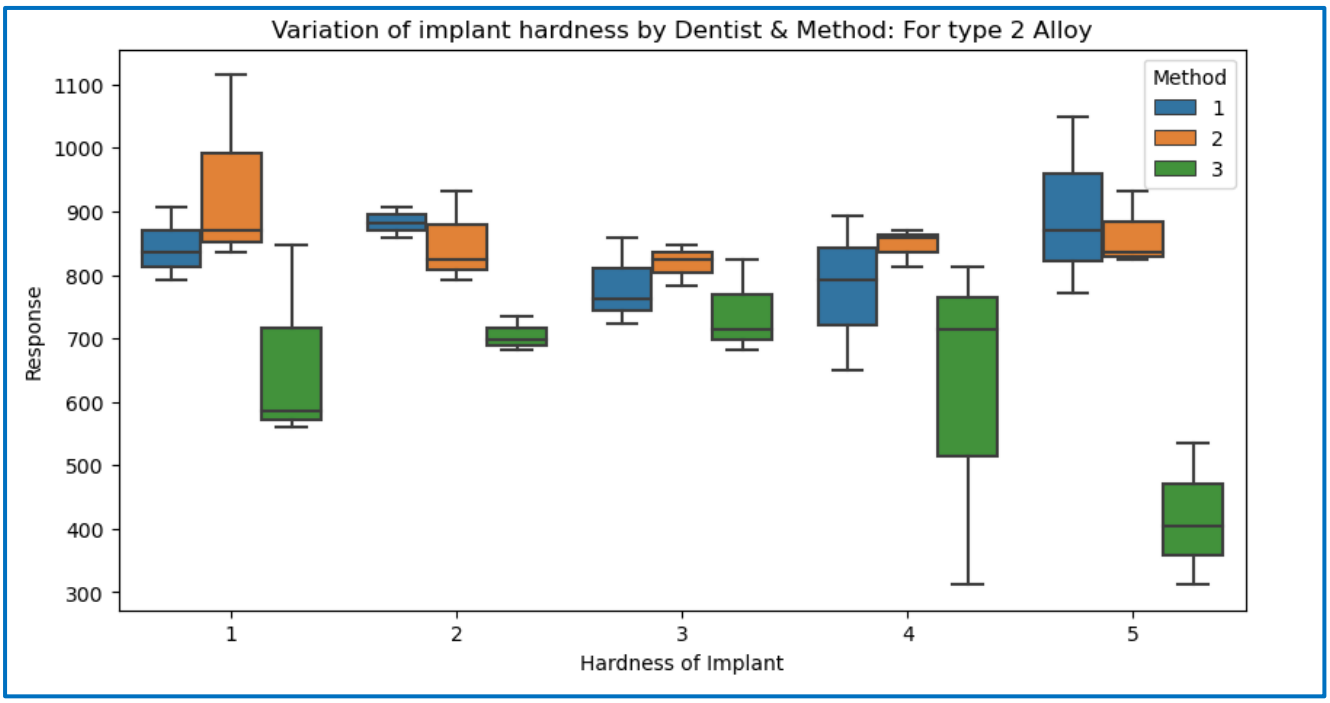


Fig 1.15 - Hardness of implant on Dentist, Methods & Response for alloy type 2

4.4.2.2 Check the assumptions of the hypothesis test

Checking the normality of the hardness of the implant for Alloy Type2. This can be done using Shapiro's Test

Shapiro's Test:

Stating the Hypothesis:

Ho: The Response follows a normal distribution for all the combinations of Dentists and Methods for Alloy Type 2
Ha: The Response does not follow a normal distribution for all the combinations of Dentists and Methods for Alloy Type 2

Checking the Homogeneity of Variance of the hardness of the implant for type 2 Alloy

Levene's Test:

Stating the Hypothesis:

Ho: All the combinations of Dentists and Methods have equal variance for Alloy Type 2
Ha: At least one combinations of Dentists and Methods does not have equal variance for Alloy Type 2

```
Statistics = 0.67
P_value    = 0.78
```

The p value is 0.78 which is greater than 0.05, hence we fail to reject the null hypothesis. Ho is True. All the combinations of Dentists and Methods have equal variance for Alloy Type 2

4.4.2.3 Conduct the hypothesis test and compute the p-value

Two Way ANOVA without interaction of Dentist & Method types using the functions ols() & anova_lm()

	df	sum_sq	mean_sq	F	PR(>F)
Dentist	4.0	56797.911111	14199.477778	0.926215	0.458933
Method	2.0	499640.400000	249820.200000	16.295479	0.000008
Residual	38.0	582564.488889	15330.644444	NaN	NaN

Table 1.23 - Two Way ANOVA for the types Dentist and Methods w.r.t Response for Alloy Type 2

Two Way ANOVA with interaction of Dentist & Method types using the functions ols() & anova_lm()

	df	sum_sq	mean_sq	F	PR(>F)
Dentist	4.0	56797.911111	14199.477778	1.106152	0.371833
Method	2.0	499640.400000	249820.200000	19.461218	0.000004
Dentist:Method	8.0	197459.822222	24682.477778	1.922787	0.093234
Residual	30.0	385104.666667	12836.822222	NaN	NaN

Table 1.24 - Two Way ANOVA for the types Dentist and Methods and Interaction of Dentist and Methods w.r.t Response for Alloy Type 2

4.4.2.4 Write down conclusions from the test results

The p value for Dentist and Method together is 0.093234 which is greater than 0.05, hence Ho is True. We fail reject the null hypothesis. And we can conclude that for the Types of Dentists and Methods together have equal effect or mean on the response (hardness of the implant) for Alloy Type 2

4.4.2.5 In case the implant hardness differs, identify for which pairs it differs

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1:1	1:2	95.3333	0.999	-245.5625	436.2292	False
1:1	1:3	-180.6667	0.8085	-521.5625	160.2292	False
1:1	2:1	37.6667	1.0	-303.2292	378.5625	False
1:1	2:2	5.0	1.0	-335.8958	345.8958	False
1:1	2:3	-140.3333	0.9635	-481.2292	200.5625	False
1:1	3:1	-63.3333	1.0	-404.2292	277.5625	False
1:1	3:2	-27.0	1.0	-367.8958	313.8958	False
1:1	3:3	-104.6667	0.9973	-445.5625	236.2292	False
1:1	4:1	-66.3333	1.0	-407.2292	274.5625	False
1:1	4:2	2.3333	1.0	-338.5625	343.2292	False
1:1	4:3	-231.3333	0.4686	-572.2292	109.5625	False
1:1	5:1	52.0	1.0	-288.8958	392.8958	False
1:1	5:2	19.3333	1.0	-321.5625	360.2292	False
1:1	5:3	-427.0	0.0049	-767.8958	-86.1042	True
1:2	1:3	-276.0	0.2169	-616.8958	64.8958	False
1:2	2:1	-57.6667	1.0	-398.5625	283.2292	False
1:2	2:2	-90.3333	0.9994	-431.2292	250.5625	False
1:2	2:3	-235.6667	0.4396	-576.5625	105.2292	False
1:2	3:1	-158.6667	0.912	-499.5625	182.2292	False
1:2	3:2	-122.3333	0.9884	-463.2292	218.5625	False
1:2	3:3	-200.0	0.6868	-540.8958	140.8958	False
1:2	4:1	-161.6667	0.9005	-502.5625	179.2292	False
1:2	4:2	-93.0	0.9992	-433.8958	247.8958	False

1:2	4:3	-326.6667	0.0709	-667.5625	14.2292	False
1:2	5:1	-43.3333	1.0	-384.2292	297.5625	False
1:2	5:2	-76.0	0.9999	-416.8958	264.8958	False
1:2	5:3	-522.3333	0.0003	-863.2292	-181.4375	True
1:3	2:1	218.3333	0.5587	-122.5625	559.2292	False
1:3	2:2	185.6667	0.7793	-155.2292	526.5625	False
1:3	2:3	40.3333	1.0	-300.5625	381.2292	False
1:3	3:1	117.3333	0.992	-223.5625	458.2292	False
1:3	3:2	153.6667	0.9291	-187.2292	494.5625	False
1:3	3:3	76.0	0.9999	-264.8958	416.8958	False
1:3	4:1	114.3333	0.9937	-226.5625	455.2292	False
1:3	4:2	183.0	0.7951	-157.8958	523.8958	False
1:3	4:3	-50.6667	1.0	-391.5625	290.2292	False
1:3	5:1	232.6667	0.4596	-108.2292	573.5625	False
1:3	5:2	200.0	0.6868	-140.8958	540.8958	False
1:3	5:3	-246.3333	0.3717	-587.2292	94.5625	False
2:1	2:2	-32.6667	1.0	-373.5625	308.2292	False
2:1	2:3	-178.0	0.8234	-518.8958	162.8958	False
2:1	3:1	-101.0	0.9981	-441.8958	239.8958	False
2:1	3:2	-64.6667	1.0	-405.5625	276.2292	False
2:1	3:3	-142.3333	0.9594	-483.2292	198.5625	False
2:1	4:1	-104.0	0.9975	-444.8958	236.8958	False
2:1	4:2	-35.3333	1.0	-376.2292	305.5625	False
2:1	4:3	-269.0	0.2485	-609.8958	71.8958	False
2:1	5:1	14.3333	1.0	-326.5625	355.2292	False
2:1	5:2	-18.3333	1.0	-359.2292	322.5625	False
2:1	5:3	-464.6667	0.0017	-805.5625	-123.7708	True
2:2	2:3	-145.3333	0.9525	-486.2292	195.5625	False
2:2	3:1	-68.3333	1.0	-409.2292	272.5625	False
2:2	3:2	-32.0	1.0	-372.8958	308.8958	False
2:2	3:3	-109.6667	0.9958	-450.5625	231.2292	False
2:2	4:1	-71.3333	1.0	-412.2292	269.5625	False
2:2	4:2	-2.6667	1.0	-343.5625	338.2292	False
2:2	4:3	-236.3333	0.4352	-577.2292	104.5625	False
2:2	5:1	47.0	1.0	-293.8958	387.8958	False
2:2	5:2	14.3333	1.0	-326.5625	355.2292	False
2:2	5:3	-432.0	0.0043	-772.8958	-91.1042	True
2:3	3:1	77.0	0.9999	-263.8958	417.8958	False
2:3	3:2	113.3333	0.9942	-227.5625	454.2292	False
2:3	3:3	35.6667	1.0	-305.2292	376.5625	False
2:3	4:1	74.0	0.9999	-266.8958	414.8958	False
2:3	4:2	142.6667	0.9586	-198.2292	483.5625	False
2:3	4:3	-91.0	0.9994	-431.8958	249.8958	False
2:3	5:1	192.3333	0.7376	-148.5625	533.2292	False
2:3	5:2	159.6667	0.9083	-181.2292	500.5625	False
2:3	5:3	-286.6667	0.1746	-627.5625	54.2292	False
3:1	3:2	36.3333	1.0	-304.5625	377.2292	False
3:1	3:3	-41.3333	1.0	-382.2292	299.5625	False
3:1	4:1	-3.0	1.0	-343.8958	337.8958	False
3:1	4:2	65.6667	1.0	-275.2292	406.5625	False
3:1	4:3	-168.0	0.8735	-508.8958	172.8958	False
3:1	5:1	115.3333	0.9932	-225.5625	456.2292	False
3:1	5:2	82.6667	0.9998	-258.2292	423.5625	False
3:1	5:3	-363.6667	0.0279	-704.5625	-22.7708	True
3:2	3:3	-77.6667	0.9999	-418.5625	263.2292	False
3:2	4:1	-39.3333	1.0	-380.2292	301.5625	False
3:2	4:2	29.3333	1.0	-311.5625	370.2292	False
3:2	4:3	-204.3333	0.657	-545.2292	136.5625	False
3:2	5:1	79.0	0.9999	-261.8958	419.8958	False
3:2	5:2	46.3333	1.0	-294.5625	387.2292	False
3:2	5:3	-400.0	0.0105	-740.8958	-59.1042	True
3:3	4:1	38.3333	1.0	-302.5625	379.2292	False
3:3	4:2	107.0	0.9967	-233.8958	447.8958	False
3:3	4:3	-126.6667	0.9842	-467.5625	214.2292	False
3:3	5:1	156.6667	0.9191	-184.2292	497.5625	False
3:3	5:2	124.0	0.9869	-216.8958	464.8958	False

3:3	5:3	-322.3333	0.0786	-663.2292	18.5625	False
4:1	4:2	68.6667	1.0	-272.2292	409.5625	False
4:1	4:3	-165.0	0.8868	-505.8958	175.8958	False
4:1	5:1	118.3333	0.9914	-222.5625	459.2292	False
4:1	5:2	85.6667	0.9997	-255.2292	426.5625	False
4:1	5:3	-360.6667	0.0302	-701.5625	-19.7708	True
4:2	4:3	-233.6667	0.4529	-574.5625	107.2292	False
4:2	5:1	49.6667	1.0	-291.2292	390.5625	False
4:2	5:2	17.0	1.0	-323.8958	357.8958	False
4:2	5:3	-429.3333	0.0046	-770.2292	-88.4375	True
4:3	5:1	283.3333	0.1871	-57.5625	624.2292	False
4:3	5:2	250.6667	0.3458	-90.2292	591.5625	False
4:3	5:3	-195.6667	0.7158	-536.5625	145.2292	False
5:1	5:2	-32.6667	1.0	-373.5625	308.2292	False
5:1	5:3	-479.0	0.0011	-819.8958	-138.1042	True
5:2	5:3	-446.3333	0.0028	-787.2292	-105.4375	True

Table 1.25 - Interaction table of dentist and methods together for alloy type 2

The pairs which has the reject status as True or the pairs with p-adj value less than 0.05 are the pairs which differs in implant hardness from rest of the pairs. There are totally 18 pairs which differs.