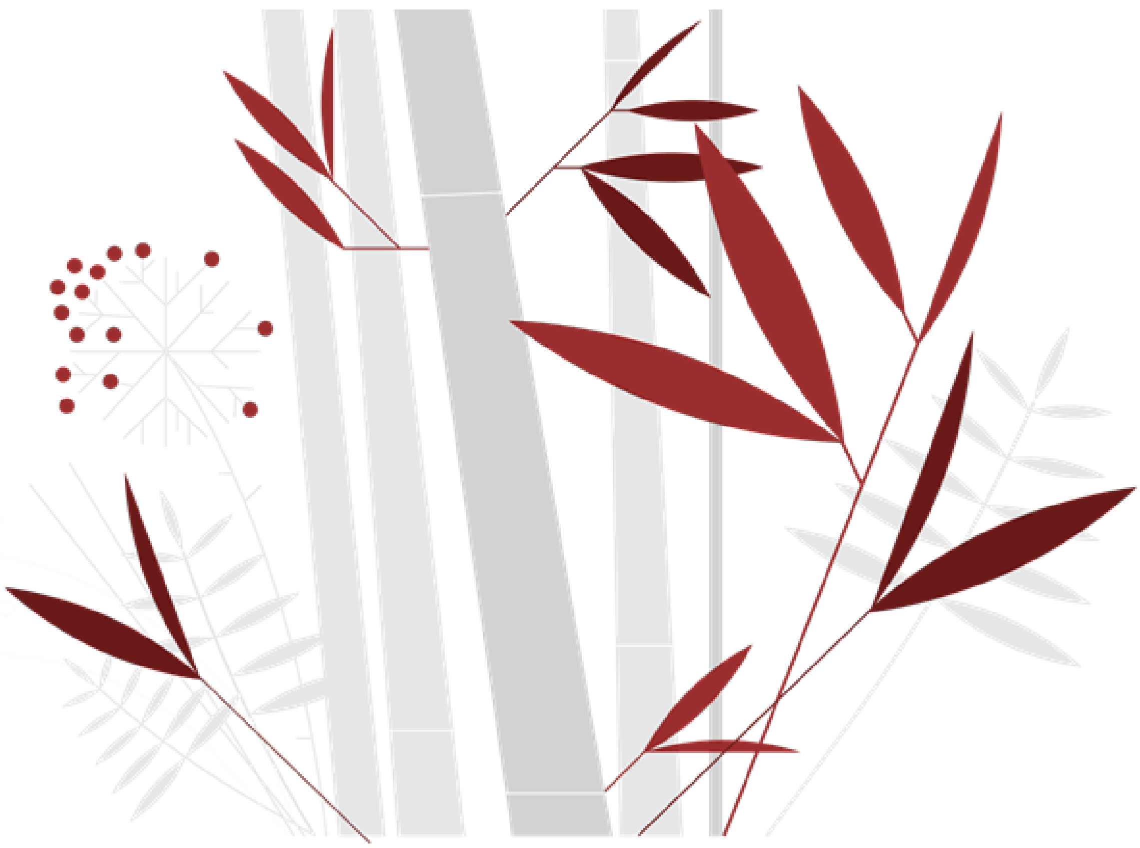
INFERENTIAL STATISTICS (IS) PROJECT-CODED

BY

Harsh Patel

7th April 2024



Sr No.	Contents	Page No.

Problem-1: Football Player Injury

A physiotherapist of a male football team is interested in studying the relationship between foot injuries and position at which the players are playing. The data collected can be seen in below table.

	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

Table-1 Football Dataset

Answer the following questions one by one to find out the probability for the given table.

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

Ans-1.1) The Probability can be calculated as follows,

$$P(Injured\ Player) = \frac{Total\ Number\ of\ Injured\ Players}{Total\ Number\ of\ Players}$$

$$= \frac{145}{235} \quad = 0.61702$$

Thus, The Probability for a randomly chosen player would suffer an injury is 61.702%.

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

Ans-1.2) We can calculate the Probability as follows,

$$P(Player\ Forword\ or\ Winger) = \frac{Total\ No.\ Forward\ Player}{Total\ No.\ Players} + \frac{Total\ No.\ Wingers\ Player}{Total\ No.\ Players}$$

= (94/235) + (29/235) = 0.523999

Thus, the probability that the player is a forward or a winger is 52.24%

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

Ans 1.3) The probability is calculated as follows,

$$P(Stri\ker\ and\ Injured) = \frac{Total\ no.\ of\ Striker}{Total\ no.\ of\ Players}$$

= (45/235) = 0.19148

Thus, the probability for a randomly chosen player is a striker and has foot injury is 19.14%

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

Ans-1.4) We calculated the probability as follows,

$$P(Injured\ and\ Stri\ker) = \frac{Total\ no.\ of\ Injured\ Striker}{Total\ no.\ of\ Stri\ker s}$$

= (45/145) = 0.31034

Thus, the probability that the randomly chosen injured player is a striker is 31.03%

Problem-2: Gunny Bags Breaking Strength

The breaking strength of gunny bags used for packing cement is normally distributed with the mean(μ) of 5 kg/sq.cm and has standard deviation(σ) of 1.5 kg/sq.cm. The quality team of the cement company wants to about the packaging material for better understand wastage or pilferage within the supply chain; Answer the questions below based on the given information

Given data,

Mu(μ) = 5 kg/sq.cm

Sigma(σ) = 1.5 kg/sq.cm

Before, we start we will load all the necessary libraries in the python like pandas, matplot.lib, seaborn, SciPy. Stats etc. As per our requirements.

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

Ans 2.1) We will calculate the z value as we know all the three variables,

Z = (x- μ)/ σ, Here X = 3.17 kg/sq.cm

= (3.17-5)/1.5 = -1.22

Now, we put this Z value in stats.norm.cdf function we the proportion of gunny bags having breaking strength of less than 3.17 kg/sq.cm is 0.1112.

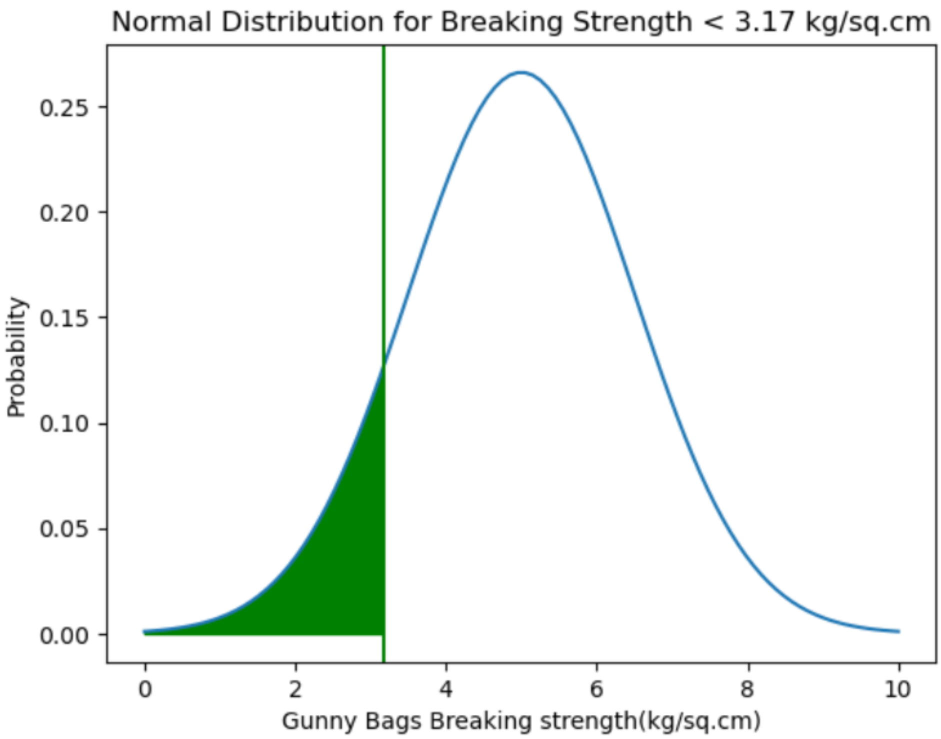


Fig-1 Normal Distribution for Breaking Strength < 3.17 kg/sq.cm

2.2 What proportion of the gunny bags have a breaking strength of at least 3.6 kg/sq.cm.?

Ans-2.2) We will calculate the z value as we know all the three variables,

Z = (x- μ)/ σ, Here X = 3.6 kg/sq.cm

= (3.6-5)/1.5 = -0.9333

Now, we put this Z value in stats.norm.cdf function we the proportion of gunny bags having breaking strength at least 3.6 kg/sq.cm is 0.8247.

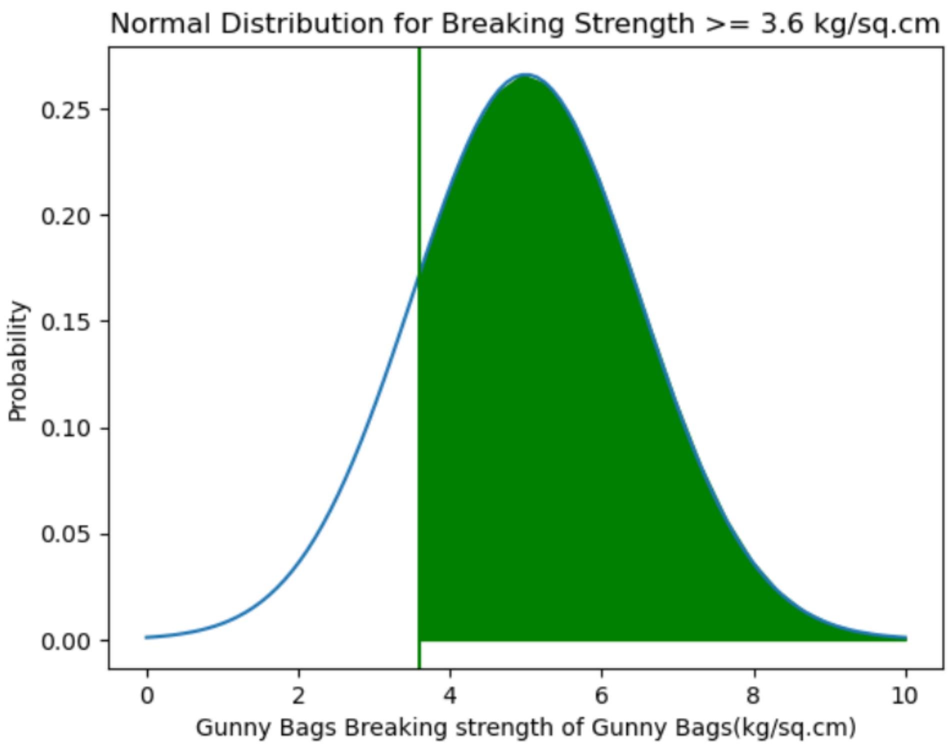


Fig-2 Normal Distribution for Breaking Strength >=3.6 kg/sq.cm

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

Ans-2.3) We will calculate the z value as we know all the three variables,

$Z1 = (x- \mu)/ \sigma$, Here X = 5 kg/sq.cm

$= (5-5)/1.5 = 0$

$Z2 = (x- \mu)/ \sigma$, Here X = 5.5 kg/sq.cm

$= (5.5-5)/1.5 = 0.3333$

Now, we put this Z1 and Z2 value in stats.norm.cdf function we the proportion of gunny bags having breaking strength between 5 and 5.5 kg/sq.cm is 0.1306.

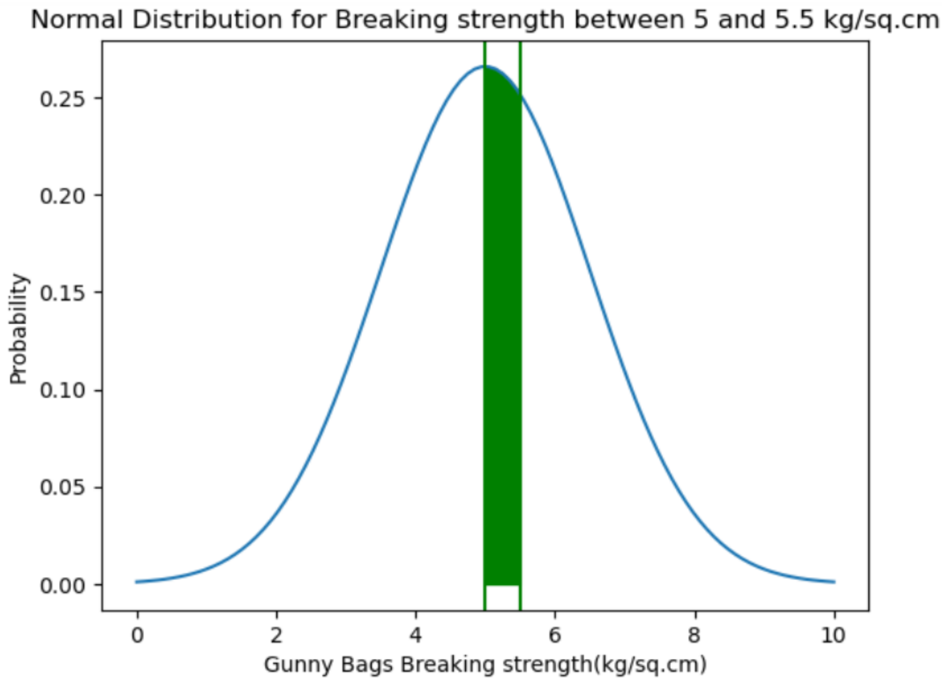


Fig-3 Normal Distribution for Breaking Strength between 5 and 5.5 kg/sq.cm

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

Ans-2.4) We will calculate the z value as we know all the three variables,

$Z1 = (x- \mu)/ \sigma$, Here $X = 3$ kg/sq.cm

$= (3-5)/1.5 = -1.333$

$Z2 = (x- \mu)/ \sigma$, Here $X = 7.5$ kg/sq.cm

$= (7.5-5)/1.5 = 1.6667$

Now, we put this Z1 and Z2 value in stats.norm.cdf function we the proportion of gunny bags having breaking strength Not between 3 and 7.5 kg/sq.cm is 0.139.

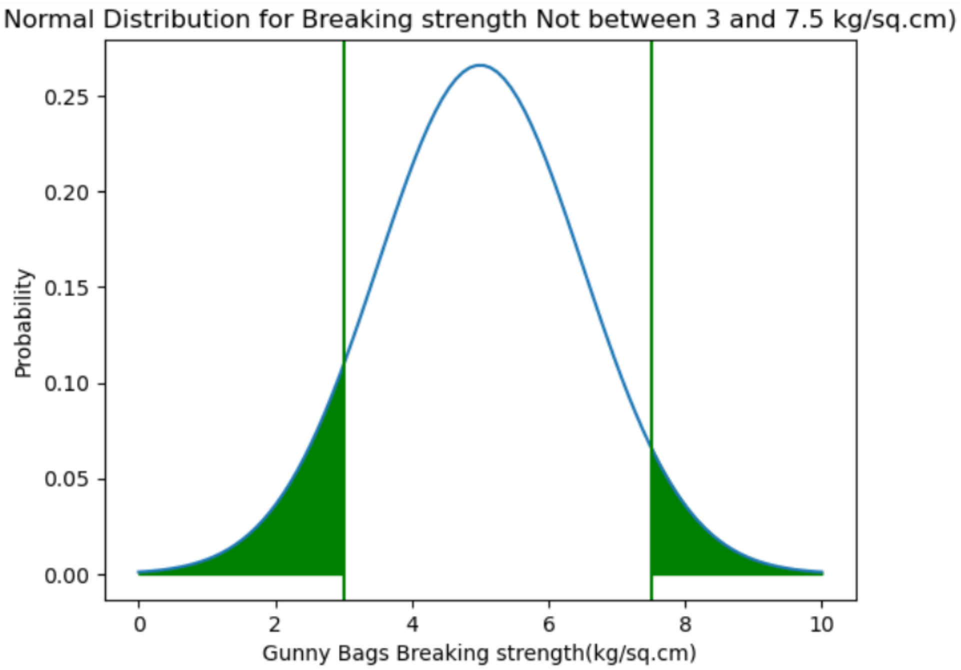


Fig4 Normal Distribution for Breaking Strength Not between 3 and 7.5 kg/sq.cm

Problem-3: Zingaro Company

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 must 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 i.e., α = 0.05)

Before, we start the calculation and visualisations we will load the Zingaro_Company csv file using read_csv function and use head and tail function to see get the visual representation of the file. After, we use shape function and get there are 75 rows and 2 columns. Were, columns names are 'Unpolished' and 'Treated and Polished'.

We checked for any duplicated rows, and we found there is none. We also, checked for any empty value using is null function and found no empty values. We used info function and got that both the columns are float data types with sample size of n=75. Finally, using the describe function we got the min, max, counts, mean, std, q1, q2, q3 values respectively for both columns.

Now, for both the given questions below we will define Null and Alternative Hypothesis.

Since, both the sample size (n)> 30 it is normally distributed, and we will take level of significance as 0.05. We do not have the standard deviation for the population from where the samples were randomly selected from.

So, we will perform t-distribution and tSTAT test statistic.

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?

Ans-3.1) For this question we will use one-tail t test.

NULL HYPOTHESIS: The mean Brinell's Hardness for unpolished stone is greater than equal to 150.

ALTERNATIVE HYPTHESIS: The mean Brinell's Hardness for unpolished stone is less than 150.

Ho: $\mu \geq 150$

Ha: $\mu < 150$

We will use ttest_1samp function to get the t-statistic and p-value. And as we are performing it for one tail, we will also divide the p-value obtained by 2 as by default it will give us the p-value for two tails.

After, performing the above function we get t-statistic = -4.1646 and p-value after dividing it by 2 = $4.1712e^{-5}$. The p-value < 0.05, Hence we have enough evidence for rejecting the Null hypothesis and consider the Alternative hypothesis, saying that the mean Brinell's Hardness for unpolished stone is less than 150 and the Zingaro company is right to think that the unpolished stone are not suitable for printing.

Now, we will find out the critical value using norm.ppf function and got –1.6448 value and plot the graph as shown below. From, the graph we can say that the t-statistic value lies within the rejection zone, and we can conclude that we will reject the null hypothesis as t-stats < critical value.

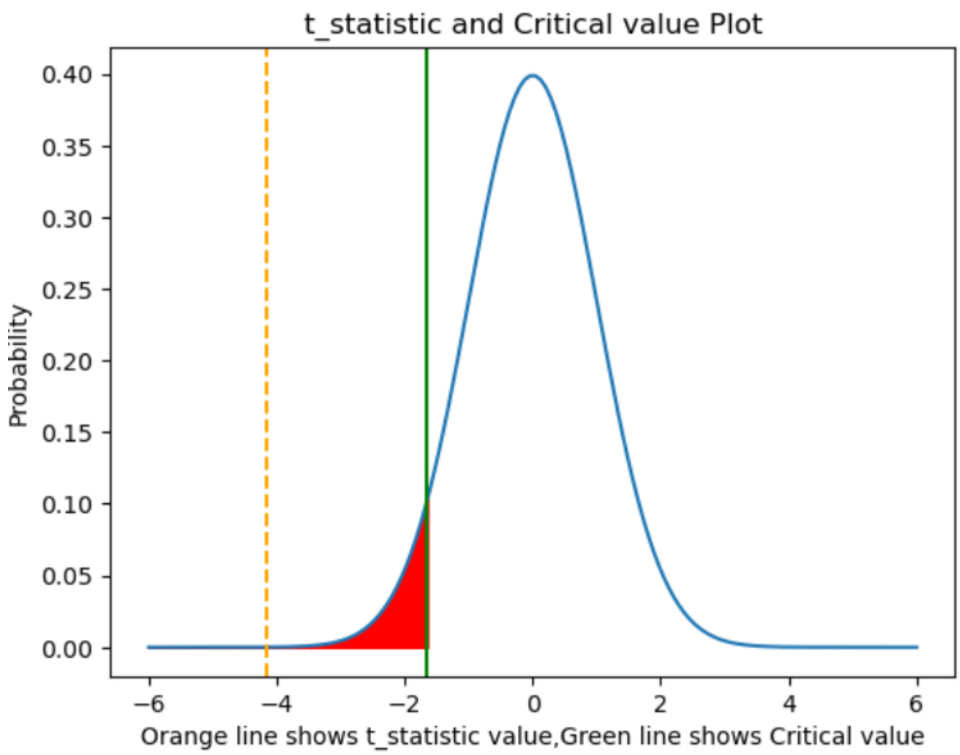


Fig-5 t_Statistic and Critical Value Plot

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

Ans-3.2) For this question we will perform two-tail t-test.

NULL HYPOTHESIS: The mean Brinell's Hardness for unpolished stone is same as mean Brinell's Hardness for treated and polished stone.

ALTERNATIVE HYPOTHESIS: The mean Brinell's Hardness for unpolished stone is not equal to mean Brinell's Hardness for treated and polished stone.

Ho: μ(unpolished) = μ(treated and polished)

Ha: μ(unpolished) ≠ μ(treated and polished)

We will use ttest_ind function as the two samples are independent. Using this we will get the t-statistic value and p-value for unpolished stones and for treated and polished stones.

After, performing the tow-tail for independent samples we got the t-statistic value = -3.2422 and p-value = 0.0014655. Here, the p-value is < 0.05, we will reject the Null Hypothesis and proceed with Alternative Hypothesis stating that mean Brinell's Hardness of unpolished stone is not equal to mean Brinell's Hardness of the treated and polished stones.

Now we will get two critical values for both samples using norm.ppf function and we got cv-1 = 1.9599 and cv-2 = -1.9599. We plot the graph as shown below and conclude that that the t-stats value lies in the rejection zone and so we reject the null hypothesis.

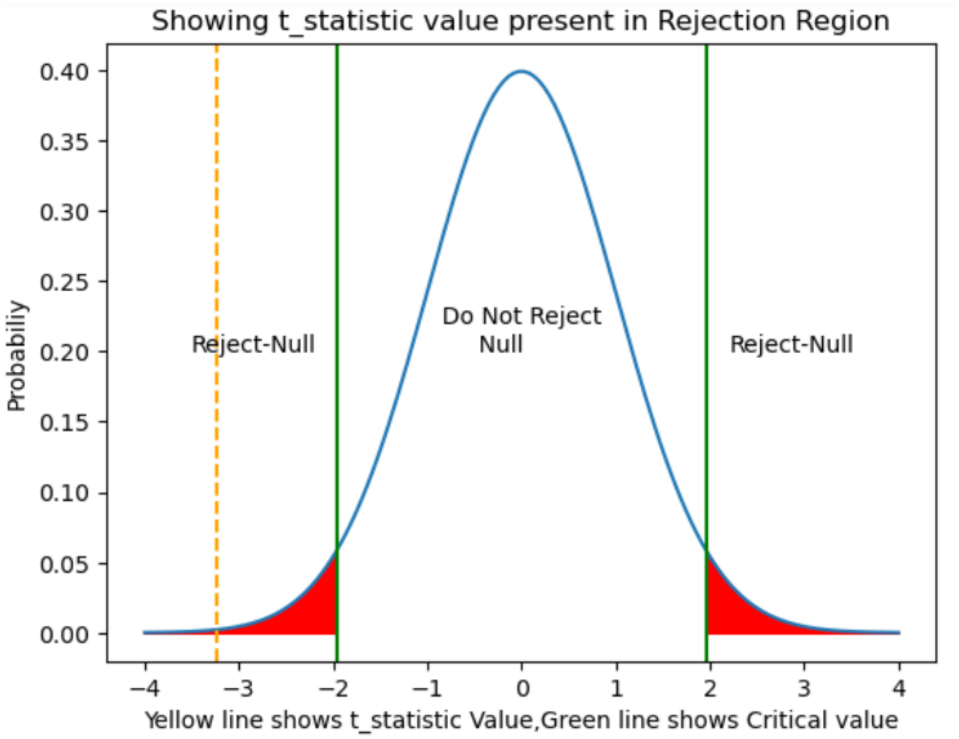


Fig-6 Showing t_statistic value in Rejection Region

From using describe function we can also see that the mean values are indeed different for both the samples as seen below.

	count	mean	std
Unpolished	75.0	134.110527	33.041804
Treated and Polished	75.0	147.788117	15.587355

Table-2 Describe function table

Problem-4: Dental Implants

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.

Before, we start anything we will load the Dental_Hardness_Data excel file using read_excel function. Read its data using head and tail function, we get the name of the variables, and they are Dentist, Method, Alloy, Temp and Response.

We will use shape function and got there are 90 rows and 5 columns. We checked for duplicate rows and there are none as well as checked for null or empty values and got there are no empty values. Then, we use info function and obtained that all the columns are numerical variables and of integer data type. The sample size n = 90 for the dataset.

Before we move further from the given questions, we know that we must perform One-way and Two-way ANOVA. We will check whether the given dataset satisfy the assumptions for both the test. The assumptions for both the test as follows:

One-way ANOVA:

1) The samples drawn from different populations are independent and random.

2) The response variables should be normally distributed for all populations.

3) For all population variances should be equal

Two-way ANOVA:

1) The dependent variable should be measured at continuous level.

2) Two independent variables should consist of 2 or more categorical, independent groups each.

3) No significant outliers should be present.

4) The dependent variable for each combination of groups of two independent variables should be approximately normally distributed.

First, we will use Histogram on dataset, and we see the variables are normally distributed or not and we obtain the graph as follows. And we can say that they are independent and randomly. The dependent variable is continuous.

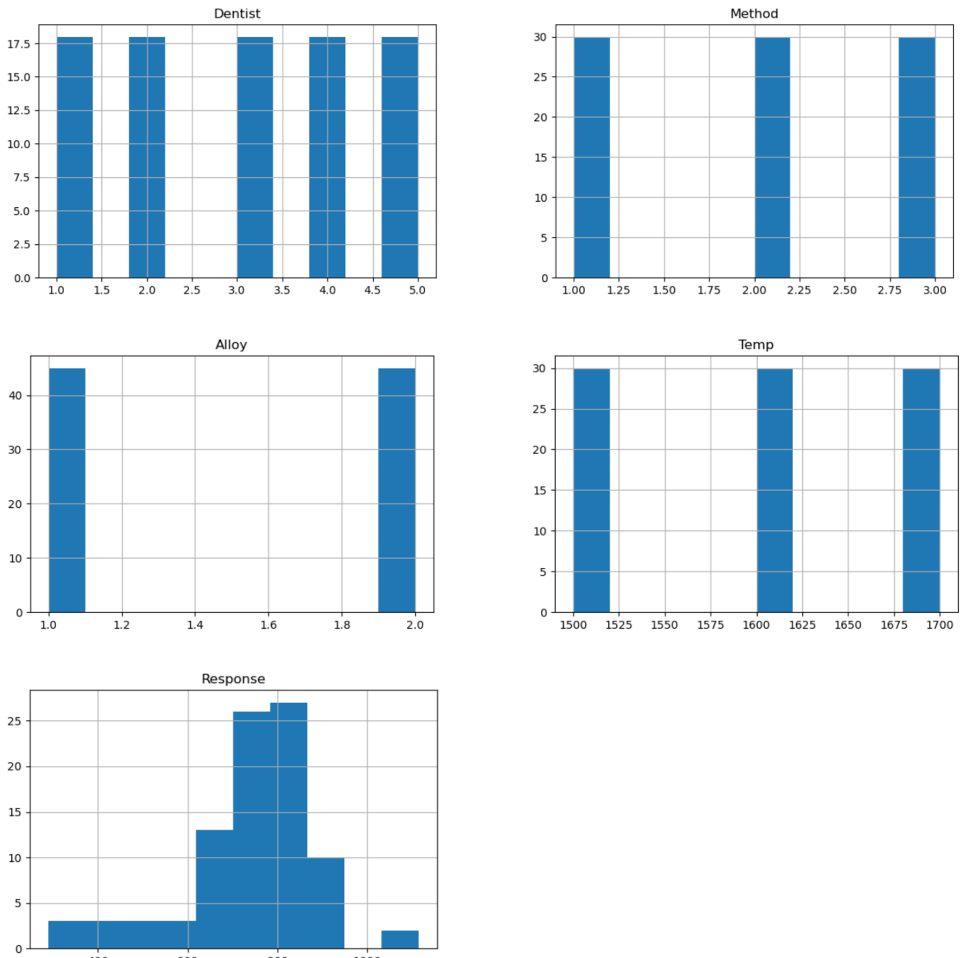


FIg-7 Histogram Plot for Dataset.

Now we will plot box plot for all the variables and treat if any outliers are present

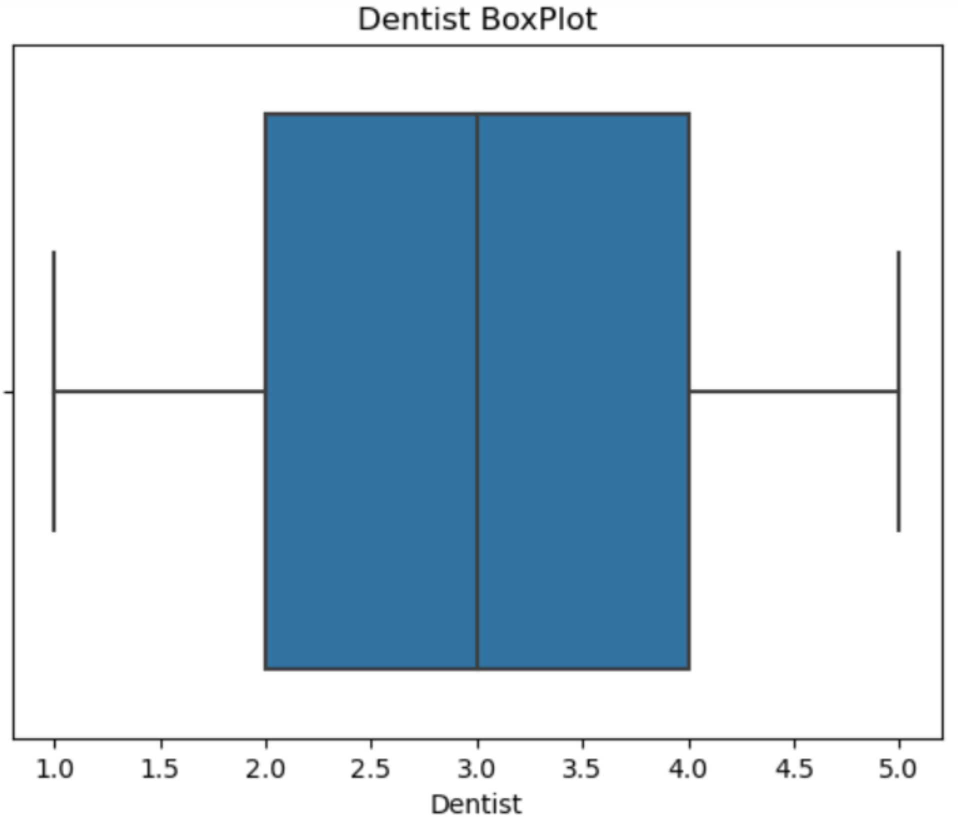


Fig-8 Dentist Boxplot

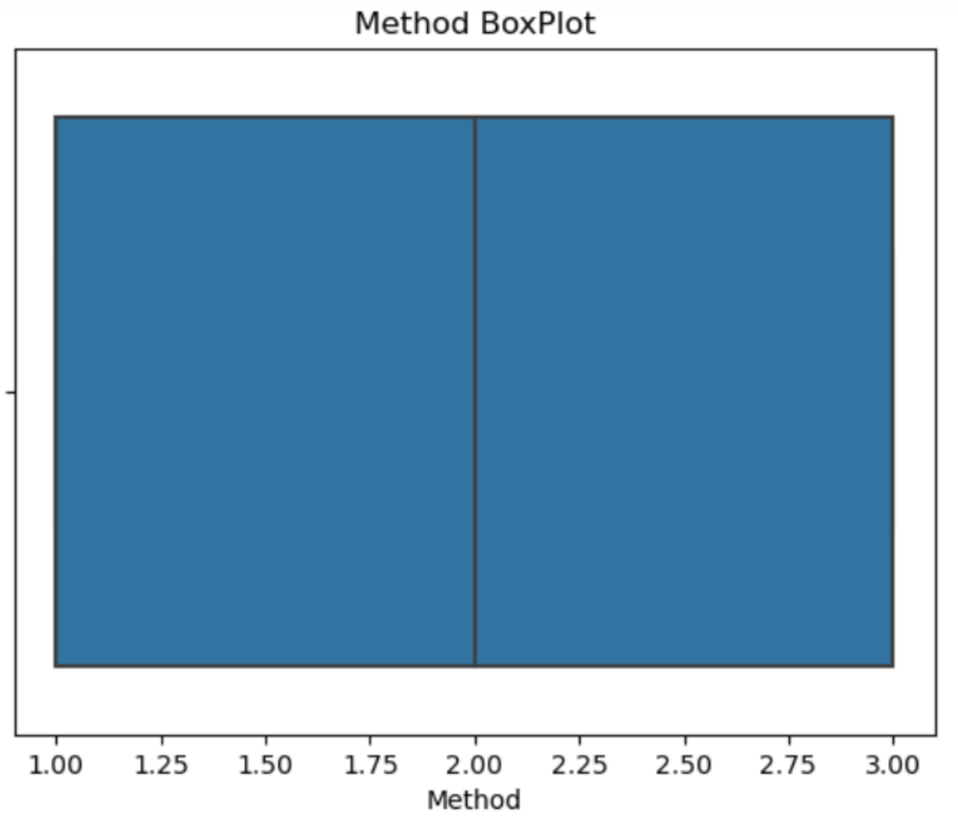


Fig-9 Method Boxplot

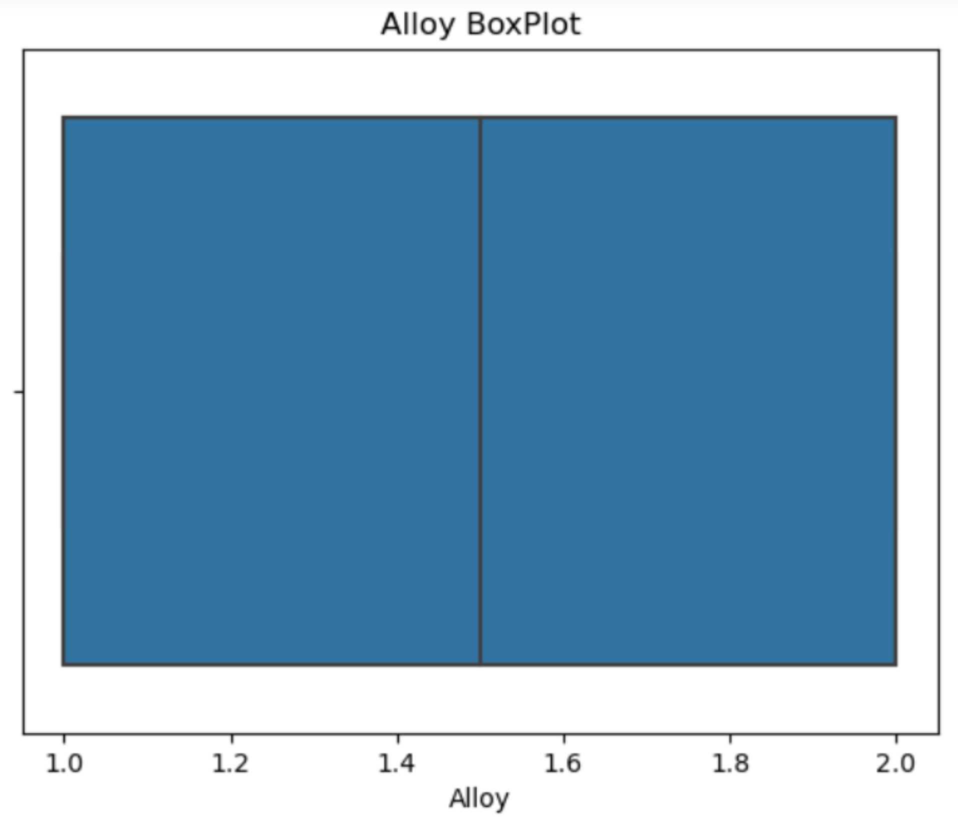


Fig-10 Alloy Boxplot

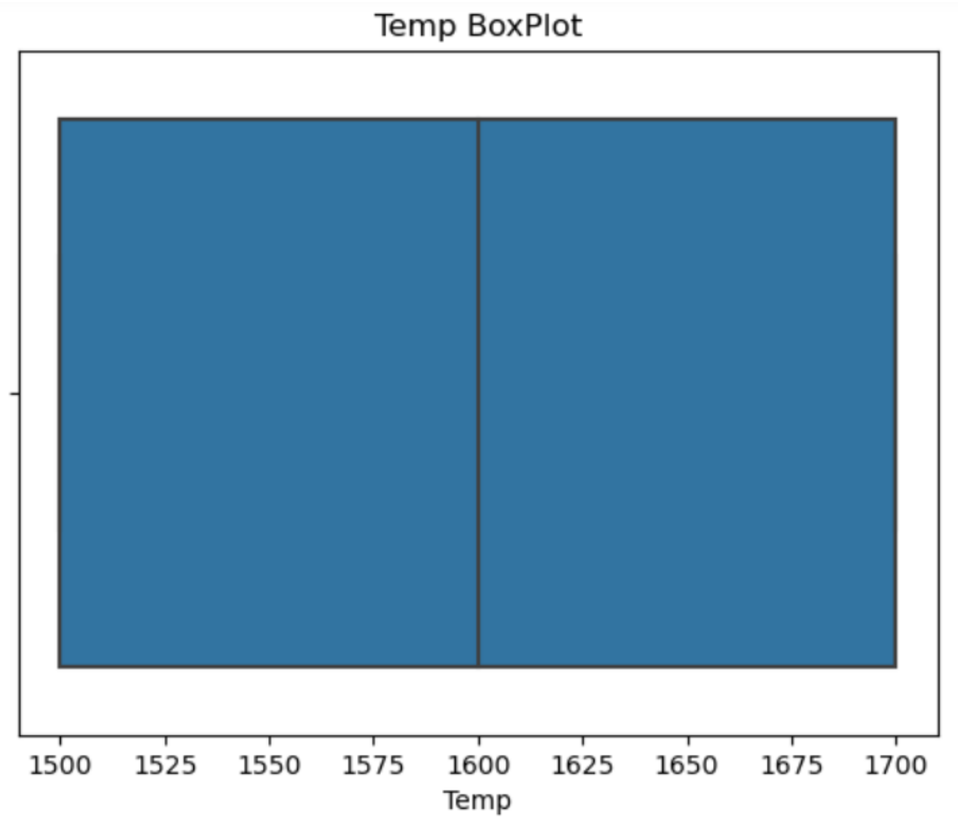


Fig-11 Temp Boxplot

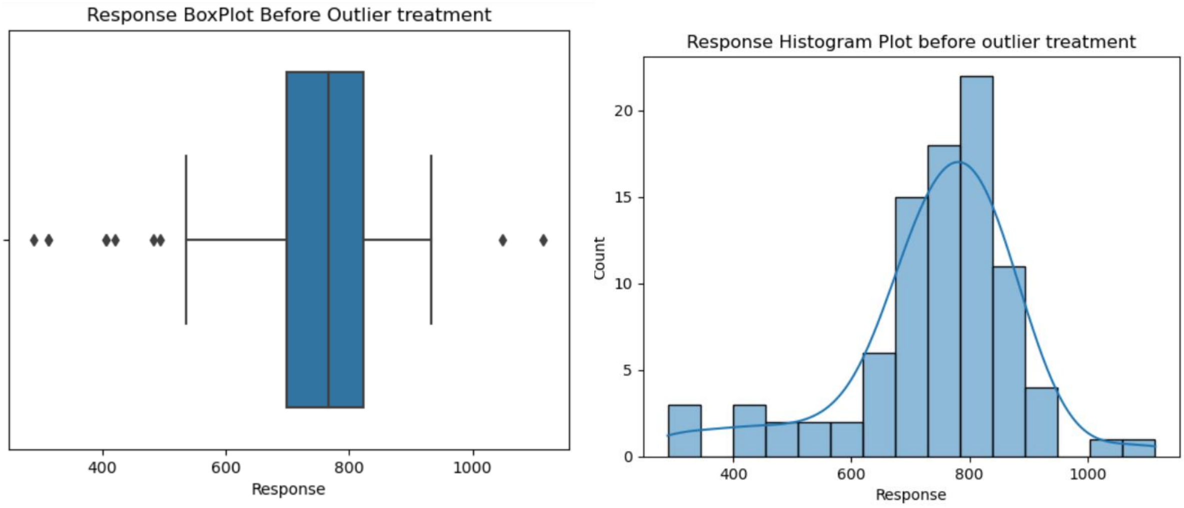


Fig-12 Response Boxplot and Histogram Plot Before Outlier Treatment

We will first get the q1 and q3 value using quantile function and get q1 = 698 and q3 = 824.

Now, we find the Inter-quantile range (IQR) = q3-q1=126.

We obtain the upper limit and lower limit of wishers using q3+(1.5*IQR) and q3-(1.5*IQR) obtain 1013 and 509 values respectively. And we replace any outliers to upper and lower limit of whiskers. And obtain the following graph.

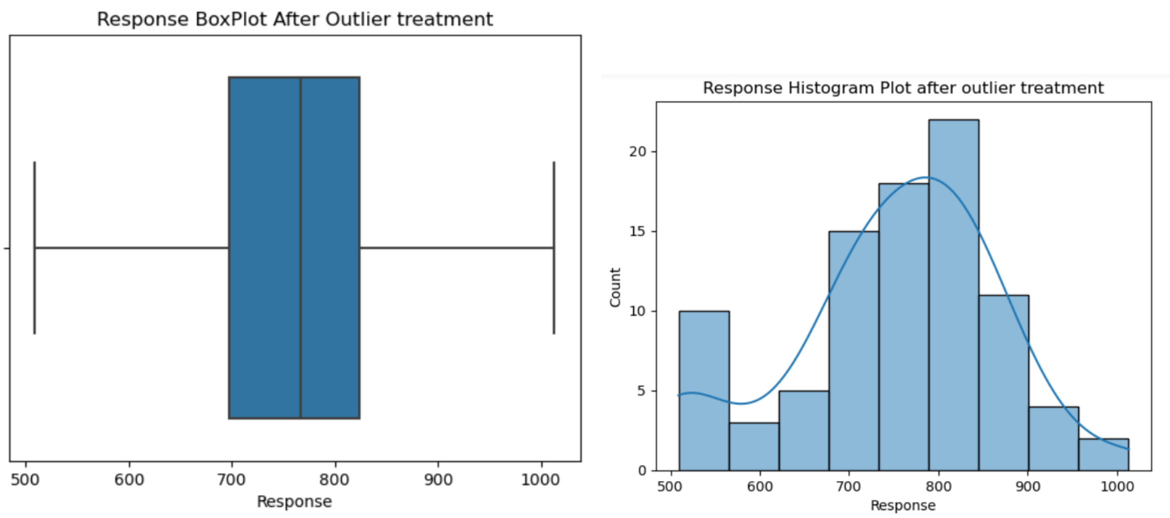


Fig-13 Response Boxplot and Histogram Plot After Outlier Treatment

For all, the question we cannot consider the alloys as same and therefore, we will create subsets for these alloys namely Alloy-1 and Alloy-2. We read the datasets using head for both subsets. As well as use info function to obtain the sample size and we got that both the dataset has n = 45. And all of them are numerical variables.

Now, we use Shapiro_wilk test to check the normality of the subsets as shown below.

```
ShapiroResult(statistic=0.8888034820556641, pvalue=0.00043243501568213105)
ShapiroResult(statistic=0.7937804460525513, pvalue=1.7407411405656603e-06)
ShapiroResult(statistic=0.7937804460525513, pvalue=1.7407411405656603e-06)
ShapiroResult(statistic=0.8888034820556641, pvalue=0.00043243501568213105)
ShapiroResult(statistic=0.7937804460525513, pvalue=1.7407411405656603e-06)
ShapiroResult(statistic=0.7937804460525513, pvalue=1.7407411405656603e-06)
```

Table-3 Shapiro_wilk Test Output

Since it failed the above test. We will perform Anderson Darling test and it passed all its test as shown below.

```
AndersonResult(statistic=1.564437360287279, critical_values=array([0.535, 0.609, 0.731, 0.853, 1.014]), significance_level=arra
y([15. , 10. ,  5. ,  2.5,  1. ]), fit_result=  params: FitParams(loc=3.0, scale=1.4301938838683885)
 success: True
 message: '`anderson` successfully fit the distribution to the data.')
AndersonResult(statistic=3.595656912040134, critical_values=array([0.535, 0.609, 0.731, 0.853, 1.014]), significance_level=arra
y([15. , 10. ,  5. ,  2.5,  1. ]), fit_result=  params: FitParams(loc=2.0, scale=0.8257228238447705)
 success: True
 message: '`anderson` successfully fit the distribution to the data.')
AndersonResult(statistic=3.595656912040134, critical_values=array([0.535, 0.609, 0.731, 0.853, 1.014]), significance_level=arra
y([15. , 10. ,  5. ,  2.5,  1. ]), fit_result=  params: FitParams(loc=1600.0, scale=82.57228238447705)
 success: True
 message: '`anderson` successfully fit the distribution to the data.')
AndersonResult(statistic=1.564437360287279, critical_values=array([0.535, 0.609, 0.731, 0.853, 1.014]), significance_level=arra
y([15. , 10. ,  5. ,  2.5,  1. ]), fit_result=  params: FitParams(loc=3.0, scale=1.4301938838683885)
 success: True
 message: '`anderson` successfully fit the distribution to the data.')
AndersonResult(statistic=3.595656912040134, critical_values=array([0.535, 0.609, 0.731, 0.853, 1.014]), significance_level=arra
y([15. , 10. ,  5. ,  2.5,  1. ]), fit_result=  params: FitParams(loc=2.0, scale=0.8257228238447705)
 success: True
 message: '`anderson` successfully fit the distribution to the data.')
AndersonResult(statistic=3.595656912040134, critical_values=array([0.535, 0.609, 0.731, 0.853, 1.014]), significance_level=arra
y([15. , 10. ,  5. ,  2.5,  1. ]), fit_result=  params: FitParams(loc=1600.0, scale=82.57228238447705)
 success: True
 message: '`anderson` successfully fit the distribution to the data.')
```

Table-4 Anderson Darling Test output

Now we will check the subsets homogeneity using Levene test and it passed its test as shown in below table.

```
LeveneResult(statistic=30.08802166864238, pvalue=9.417442084364688e-16)
LeveneResult(statistic=30.47328114592907, pvalue=6.458959800684149e-16)
LeveneResult(statistic=1.7299456023326063, pvalue=0.1625737109730439)
```

Table-5 Levene Test Output

Now, that it has satisfied all the assumptions for the one-way and two-way ANOVA test. We change all numerical variable like Dentist, Method and Temp to categorical variables except response variable as its dependent variable, using pd.Categorical function.

After that we use info function and seen that they have been converted to Categorical variables as seen in Table-6 below.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 45 entries, 0 to 44
Data columns (total 5 columns):
 #   Column    Non-Null Count  Dtype   
---  ------    --------------  -----   
 0   Dentist   45 non-null     category
 1   Method    45 non-null     category
 2   Alloy     45 non-null     int64   
 3   Temp      45 non-null     category
 4   Response  45 non-null     float64 
dtypes: category(3), float64(1), int64(1)
memory usage: 1.4 KB
None
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 45 entries, 0 to 44
Data columns (total 5 columns):
 #   Column    Non-Null Count  Dtype   
---  ------    --------------  -----   
 0   Dentist   45 non-null     category
 1   Method    45 non-null     category
 2   Alloy     45 non-null     int64   
 3   Temp      45 non-null     category
 4   Response  45 non-null     float64 
dtypes: category(3), float64(1), int64(1)
```

Table-6 Datatype of variables for Alloys.

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

Ans-4.1) We will make Null and Alternative Hypothesis for both the Alloys:

1.Null and Alternative Hypothesis for Alloy-1:

H01: Mean Hardness is same across all dentists for alloy 1.

Ha1: Mean Hardness is not same for at least one pair of dentists for alloy 1.

2.Null and Alternative Hypothesis for Alloy-2:

H02: Mean Hardness is same across all dentists for alloy 2.

Ha2: Mean Hardness is not same for at least one pair of dentists for alloy 2.

We will some functions for Statsmodel like ols, anova_lm, pair_wise_tukeyhsd etc.

We make a formula for Response as a function of Dentist variable for alloy1 and using ols function on them to create an Anova table as seen below.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4.0	66319.422222	16579.855556	2.026521	0.109066
Residual	40.0	327257.555556	8181.438889	NaN	NaN

Table-7 ANOVA table of Dentist for Alloy-1

From the table-7 we can see that the p-value > 0.05 and thus we fail to reject the null hypothesis and we can say that for alloy 1 the mean of implant hardness is same for all the dentists.

We plot the below graph using point plot obtain a Dentist vs Response graph. Since we fail to reject the null hypothesis, we cannot obtain which pairs of Dentist is different.

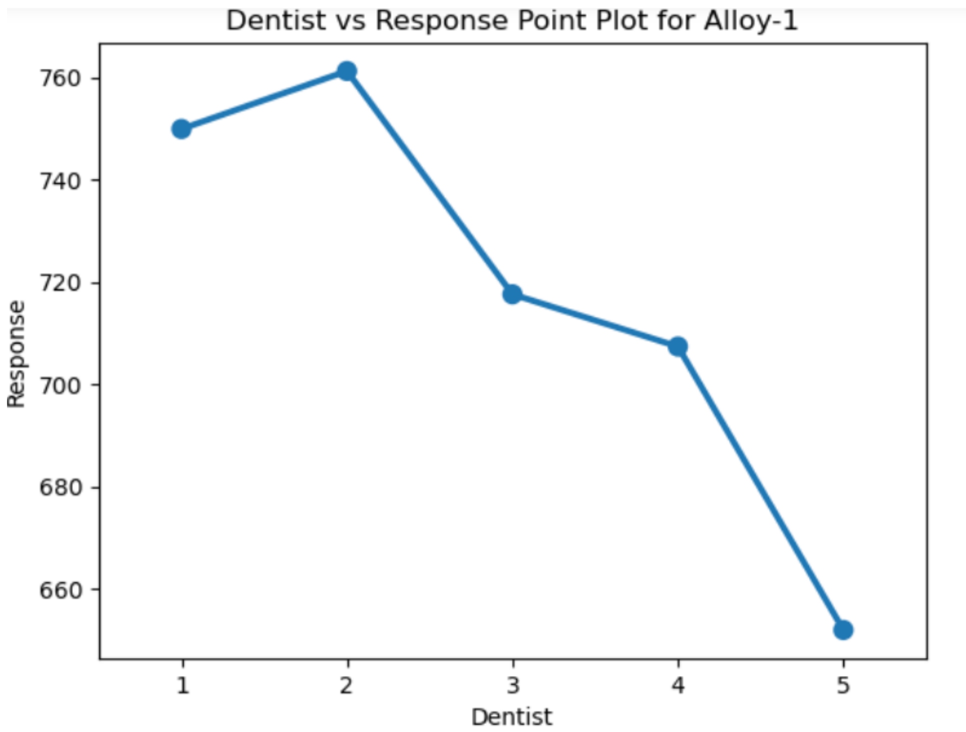


Fig-14 Dentist vs Response Point Plot for Alloy-1

Even though we failed to reject null hypothesis, we will still use Tukey Hsd test to find out the pairs of dentists for alloy 1 as shown in below table. We can see that all the pair have failed to be reject.

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1	2	11.3333	0.9989	-110.448	133.1147	False
1	3	-32.3333	0.9409	-154.1147	89.448	False
1	4	-42.5556	0.8547	-164.3369	79.2258	False
1	5	-98.0	0.1666	-219.7813	23.7813	False
2	3	-43.6667	0.8427	-165.448	78.1147	False
2	4	-53.8889	0.7143	-175.6702	67.8924	False
2	5	-109.3333	0.0968	-231.1147	12.448	False
3	4	-10.2222	0.9992	-132.0035	111.5591	False
3	5	-65.6667	0.5434	-187.448	56.1147	False
4	5	-55.4444	0.6925	-177.2258	66.3369	False

Table-8 Pairwise Comparison of Dentist for Alloy-1

Now, for Alloy-2 we repeat the same process we did for alloy 1 except in place of alloy1 we replace it with alloy 2 and obtain the below Anova table.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4.0	20704.977778	5176.244444	0.302518	0.874547
Residual	40.0	684420.666667	17110.516667	NaN	NaN

Table-9 ANOVA table of Dentist for Alloy-2

Here, p-value is also> 0.05 and we failed to reject the null hypothesis. Hence for alloy-2 also the mean implant hardness is same for all dentists.

We plot a line plot to see the visualisation of this and see that we cannot tell which pairs of dentists is different as shown in below graph.

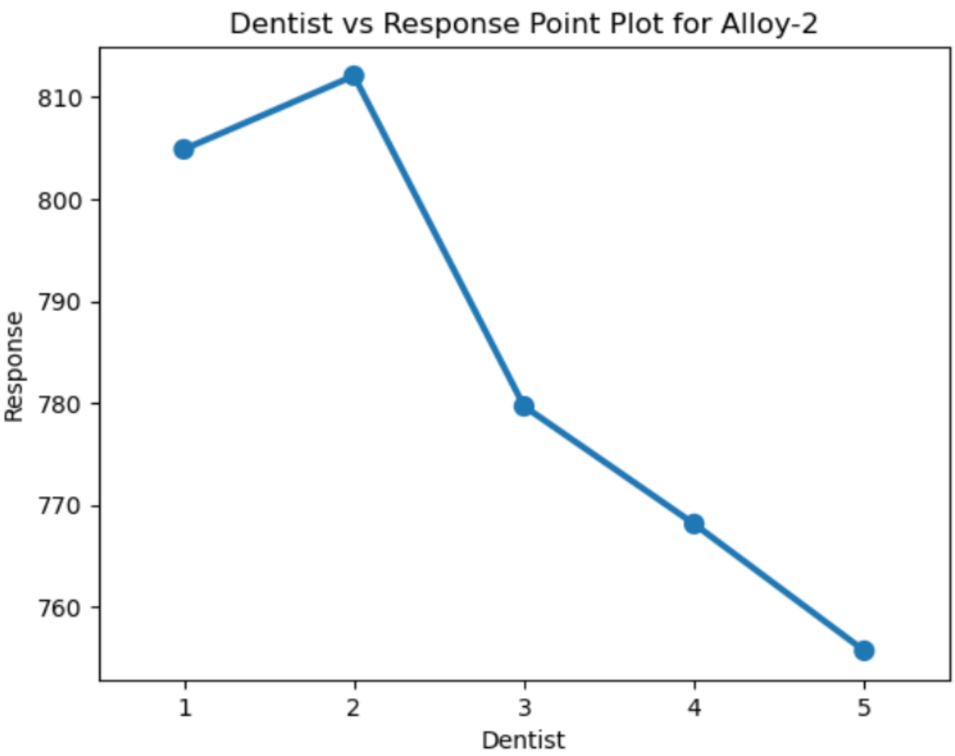


Fig-15 Dentist vs Response Point Plot for Alloy-2

We will still perform Tukey Hsd test and obtain the below table-10 and see that each pair has been failed to reject.

```
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
========================================================
group1 group2  meandiff p-adj    lower     upper   reject
--------------------------------------------------------
     1      2   11.3333 0.9989  -110.448 133.1147  False
     1      3  -32.3333 0.9409 -154.1147   89.448  False
     1      4  -42.5556 0.8547 -164.3369  79.2258  False
     1      5     -98.0 0.1666 -219.7813  23.7813  False
     2      3  -43.6667 0.8427  -165.448  78.1147  False
     2      4  -53.8889 0.7143 -175.6702  67.8924  False
     2      5 -109.3333 0.0968 -231.1147   12.448  False
     3      4  -10.2222 0.9992 -132.0035 111.5591  False
     3      5  -65.6667 0.5434  -187.448  56.1147  False
     4      5  -55.4444 0.6925 -177.2258  66.3369  False
--------------------------------------------------------
```

Table-10 Pairwise Comparison of Dentist for Alloy-2

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

Ans-4.2) We make Null and Alternative Hypothesis for Methods for both alloys.:

1. Null and Alternative Hypothesis for Alloy-1:

   H01: Mean Hardness is same for all Methods for alloy-1.

   Ha1: Mean Hardness is not same for at least one pair of Methods for alloy-1.

2. Null and Alternative Hypothesis for Alloy-2:

   H02: Mean Hardness is same for all Methods for alloy-2.

   Ha2: Mean Hardness is not same for at least one pair of Methods for alloy-2.

We create the table of Anova for response as a function of method variable and obtain the table-11 below.

	df	sum_sq	mean_sq	F	PR(>F)
C(Method)	2.0	83943.644444	41971.822222	5.693239	0.006489
Residual	42.0	309633.333333	7372.222222	NaN	NaN

Table-11 ANOVA table of Method for Alloy-1

We can see that the p-value<0.05.so, we reject the null and consider alternative hypothesis saying that for at least one pair of method the mean implant hardness is different for alloy1.

Now, we will plot the point plot for the above as shown below and see that Method 3 is way different from Method1 and 2.

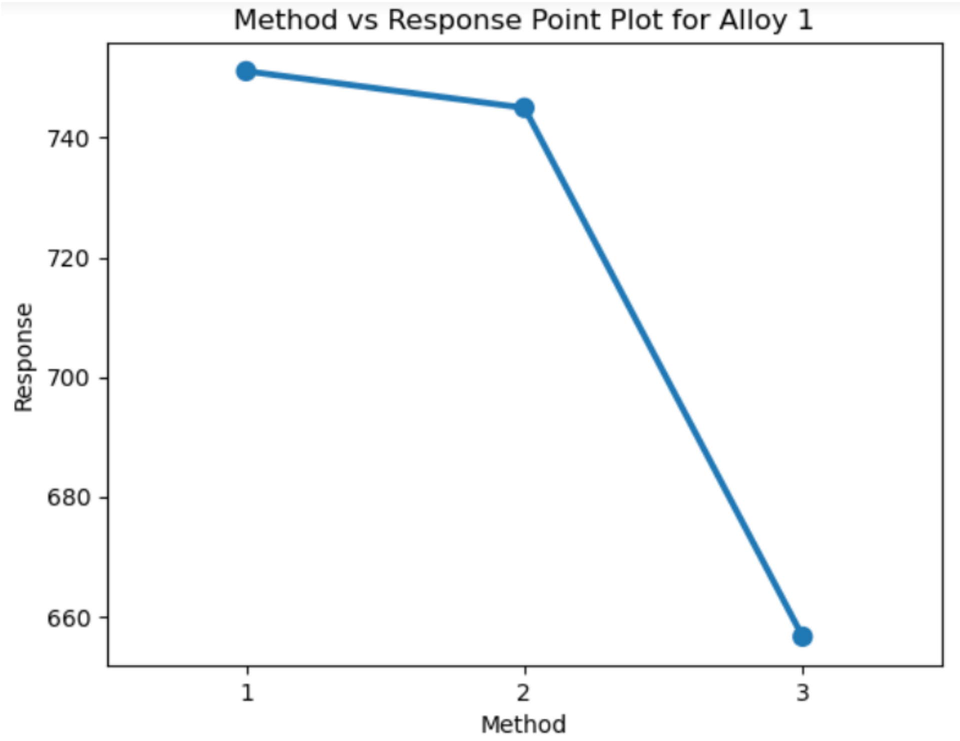


Fig-16 Method vs Response Point Plot for Alloy-1

Now, to find out which pairs are different we perform Tukey Hsd test and obtain the below table. From it we can say that the pairs 1-3 and 2-3 are different. And this may be due the Method 3 being significantly different compared to other two implant hardness means.

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1	2	-6.1333	0.9791	-82.3034	70.0367	False
1	3	-94.5333	0.0118	-170.7034	-18.3633	True
2	3	-88.4	0.0196	-164.5701	-12.2299	True

Table-12 Pairwise Comparison of Method for Alloy-1

Now we will repeat the same process for alloy-2 and obtain the Anova table as shown below.

	df	sum_sq	mean_sq	F	PR(>F)
C(Method)	2.0	344333.644444	172166.822222	20.042037	7.741020e-07
Residual	42.0	360792.000000	8590.285714	NaN	NaN

Table-13 ANOVA table of Method for Alloy-2

From the table-13 we conclude that p-value is < 0.05 and we reject the null and consider the alternative saying at least one pair of method is different.

We can see the visual representation using point plot and obtain the graph below. We can see that Method 3 is different from method 1and 2.

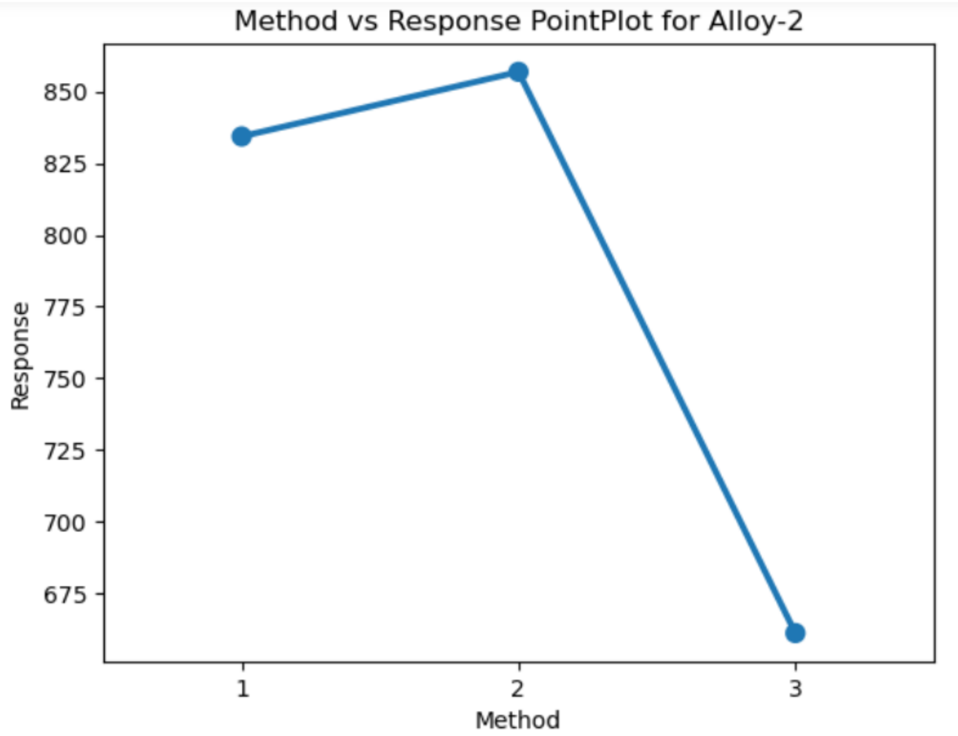


Fig-17 Method vs Response Point Plot for Alloy-2

Now we use Tukey Hsd test to find put which pairs of method are different for alloy 2. From the table we can see that pair 1-3 and 2-3 is different due to Method 3 being way different from Method1 and 2 implant hardness means.

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1	2	22.5333	0.7845	-59.6888	104.7555	False
1	3	-173.2667	0.0	-255.4888	-91.0445	True
2	3	-195.8	0.0	-278.0222	-113.5778	True

Table-14 Pairwise Comparison of Method for Alloy-2

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

Ans-4.3) The Interaction effect can be seen when we use response as a function of dentist, method and dentist: method. And obtain the table as seen below for alloy1.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4.0	66319.422222	16579.855556	3.539108	0.017620
C(Method)	2.0	83943.644444	41971.822222	8.959234	0.000890
C(Dentist):C(Method)	8.0	102771.244444	12846.405556	2.742172	0.021263
Residual	30.0	140542.666667	4684.755556	NaN	NaN

Table-15 ANOVA table for Alloy-1 for Dentist and Method

For better understanding we plot the interaction plot as shown below.

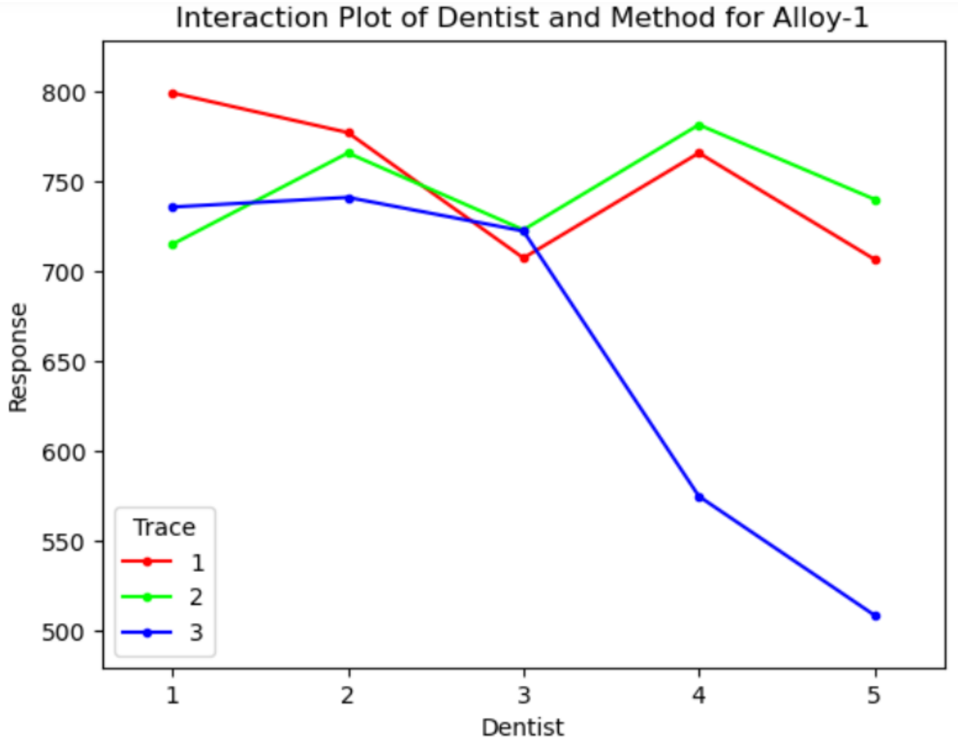


Fig-18 Interaction Plot for Alloy-1 between Method and Dentist

We can conclude that for alloy 1 different dentist using different method the mean implant hardness is not the same.

The Interaction effect can be seen when we use response as a function of dentist, method and dentist: method. And obtain the table as seen below for alloy2.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4.0	20704.977778	5176.244444	0.678500	0.612236
C(Method)	2.0	344333.644444	172166.822222	22.567548	0.000001
C(Dentist):C(Method)	8.0	111218.355556	13902.294444	1.822306	0.111831
Residual	30.0	228868.666667	7628.955556	NaN	NaN

Table-16 ANOVA table for Alloy-2 for Dentist and Method

For better understanding we plot the interaction plot as shown below.

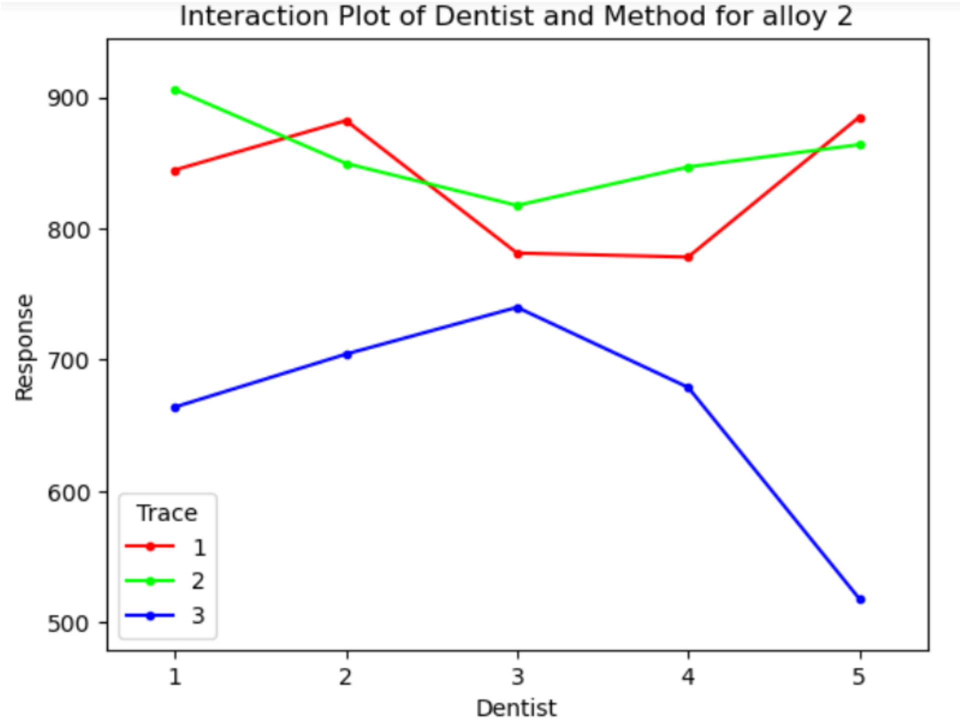


Fig-19 Interaction Plot for Alloy-2 between Method and Dentist

We can conclude that for alloy 2 different dentist using different method the mean implant hardness is not the same.

4.4 How does the hardness of implants vary depending on dentists and methods together?

Ans-4.4) For alloy-1 we will consider the following Null and Alternative Hypothesis:

Ho1: mean implant hardness for Dentist and Method is same for all for alloy-1. And there is no interaction effect between dentist and method for alloy-1.

Ha1: mean implant hardness for Dentist and Method at least one pair of dentists and one pair of method is different for alloy-1. And there is interaction effect between dentist and method for alloy-1.

For alloy-2 we will consider the following Null and Alternative Hypothesis:

Ho2: mean implant hardness for Dentist and Method is same for all for alloy-2. And there is no interaction effect between dentist and method for alloy-2.

Ha2: mean implant hardness for Dentist and Method at least one pair of dentists and one pair of method is different for alloy-2. And there is interaction effect between dentist and method for alloy-2.

We make the Anova table using response as function of Dentist and Method and obtain the below table for alloy 1.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4.0	66319.422222	16579.855556	2.589390	0.052004
C(Method)	2.0	83943.644444	41971.822222	6.555027	0.003583
Residual	38.0	243313.911111	6402.997661	NaN	NaN

Table-17 ANOVA table for Dentist and Method for Alloy-1

Thus, the p-value for dentist>0.05 and p-value for method<0.05. We can say that for Alloy-1 the means implant hardness for dentist is same across the dataset but there is at least one pair of method that has different mean implant hardness.

We will use Tukey Hsd test to see which pair is different and we can see from table-18 that pairs for Method 1-3 and 2-3 is different for alloy 1.

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1	2	11.3333	0.9989	-110.448	133.1147	False
1	3	-32.3333	0.9409	-154.1147	89.448	False
1	4	-42.5556	0.8547	-164.3369	79.2258	False
1	5	-98.0	0.1666	-219.7813	23.7813	False
2	3	-43.6667	0.8427	-165.448	78.1147	False
2	4	-53.8889	0.7143	-175.6702	67.8924	False
2	5	-109.3333	0.0968	-231.1147	12.448	False
3	4	-10.2222	0.9992	-132.0035	111.5591	False
3	5	-65.6667	0.5434	-187.448	56.1147	False
4	5	-55.4444	0.6925	-177.2258	66.3369	False

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1	2	-6.1333	0.9791	-82.3034	70.0367	False
1	3	-94.5333	0.0118	-170.7034	-18.3633	True
2	3	-88.4	0.0196	-164.5701	-12.2299	True

Table-18 Pairwise Comparison for Alloy-1

The Interaction effect can be seen when we use response as a function of dentist, method and dentist: method. And obtain the table as seen below for alloy 1. We can see the F- statistics value for dentist: method and can say there is some interaction between method and dentist variables.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4.0	66319.422222	16579.855556	3.539108	0.017620
C(Method)	2.0	83943.644444	41971.822222	8.959234	0.000890
C(Dentist):C(Method)	8.0	102771.244444	12846.405556	2.742172	0.021263
Residual	30.0	140542.666667	4684.755556	NaN	NaN

Table-19 Interaction Between Dentist and Method for Alloy-1

We make the Anova table using response as function of Dentist and Method and obtain the below table for alloy 2.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4.0	20704.977778	5176.244444	0.578373	0.680067
C(Method)	2.0	344333.644444	172166.822222	19.237251	0.000002
Residual	38.0	340087.022222	8949.658480	NaN	NaN

Table-20 ANOVA table for Dentist and Method for Alloy-2

Thus, the p-value for dentist>0.05 and p-value for method<0.05. We can say that for Alloy-2 the means implant hardness for dentist is same across the dataset but there is at least one pair of method that has different mean implant hardness.

We will use Tukey Hsd test to see which pair is different and we can see from table-21 that pairs for Method 1-3 and 2-3 is different for alloy 2.

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1	2	7.2222	1.0	-168.8932	183.3376	False
1	3	-25.2222	0.9939	-201.3376	150.8932	False
1	4	-36.7778	0.9748	-212.8932	139.3376	False
1	5	-49.2222	0.9296	-225.3376	126.8932	False
2	3	-32.4444	0.9842	-208.5599	143.671	False
2	4	-44.0	0.9522	-220.1154	132.1154	False
2	5	-56.4444	0.8893	-232.5599	119.671	False
3	4	-11.5556	0.9997	-187.671	164.5599	False
3	5	-24.0	0.9949	-200.1154	152.1154	False
4	5	-12.4444	0.9996	-188.5599	163.671	False

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1	2	22.5333	0.7845	-59.6888	104.7555	False
1	3	-173.2667	0.0	-255.4888	-91.0445	True
2	3	-195.8	0.0	-278.0222	-113.5778	True

Table-21 Pairwise Comparison for Alloy-2

The Interaction effect can be seen when we use response as a function of dentist, method and dentist: method. And obtain the table as seen below for alloy 2. We can see the F- statistics value for dentist: method and can say there is some interaction between method and dentist variables.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4.0	20704.977778	5176.244444	0.678500	0.612236
C(Method)	2.0	344333.644444	172166.822222	22.567548	0.000001
C(Dentist):C(Method)	8.0	111218.355556	13902.294444	1.822306	0.111831
Residual	30.0	228868.666667	7628.955556	NaN	NaN

Table-22 Interaction Between Dentist and Method for Alloy-2