Inferential Statistics – Coded Project

Business Report

DSBA – Course

Created by – Rishabh Gupta

Foreword

There are 4 Problem statements in this business report.

Hence, I have worked them out serial wise as mentioned in contents table.

I. **Problem 1**
II. **Problem 2**
III. **Problem 3**
IV. **Problem 4**

- Problem 1 and 2 doesn't require jupyter notebook, I have included mathematical calculations used to solve the questions.
- Appropriate visual repression sis used and can be referenced in List of Figures table.

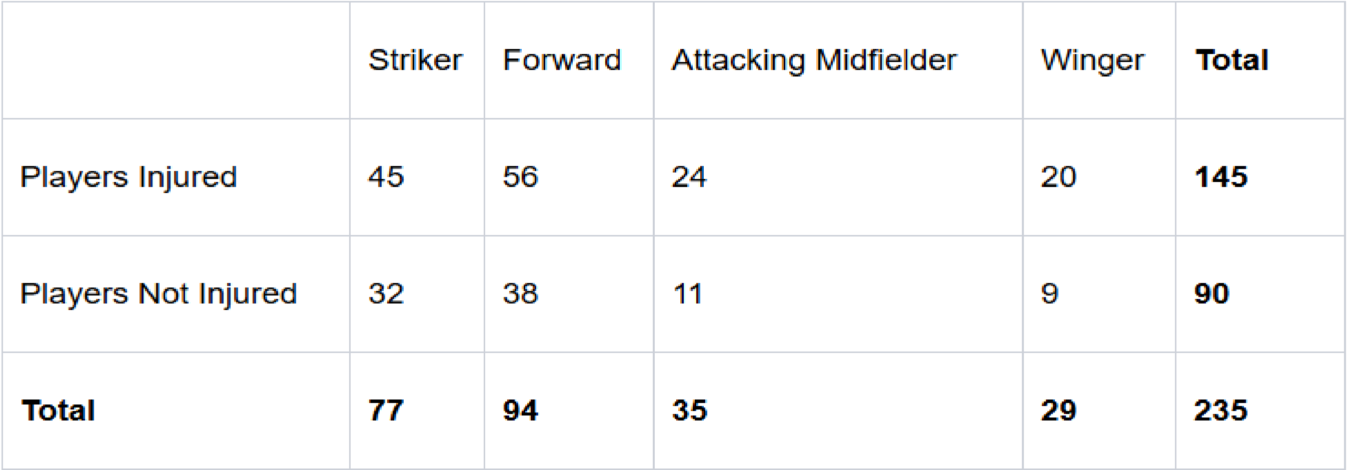
Contents

List of Tables

List of Figures

Problem Statement 1

A physiotherapist with a male football team is interested in studying the relationship between foot injuries and the positions at which the players play from the data collected.



	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 - Problem stmt - 1 Data

Based on above data, we need to answer following questions –

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

1.2 What is the probability that a player 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?

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

Data Analysis of problem statement –

On observing data as per table, let’s use below formula to calculate proportion of injured players for mentioned positions –

Injury Rate=Total Players / Injured Players

1. Striker: - 45 / 77 = 0.583 ~ 58.3 %

2. Forward: - 56 / 94 = 0.596 ~ 59.6 %

3. Attacking Midfielder: - 24 / 35 = 0.686 ~ 68.6 %

4. Winger: - 20 / 29 = 0.690 ~ 69 %

Observations

- The Winger position has the highest injury rate (69.0%).
- The Attacking Midfielder position has the second-highest injury rate (68.6%).
- The Forward position has a slightly lower injury rate (59.6%).
- The Striker position has the lowest injury rate (58.3%).

The provided data indicates a correlation between playing position and injury risk in football. Players occupying attacking roles, particularly Wingers and Attacking Midfielders, exhibit a higher incidence of injuries compared to those in the Striker and Forward positions. These findings underscore the need for tailored injury prevention programs for different player profiles. By identifying specific positions at greater risk, medical staff can implement targeted interventions to mitigate injuries.

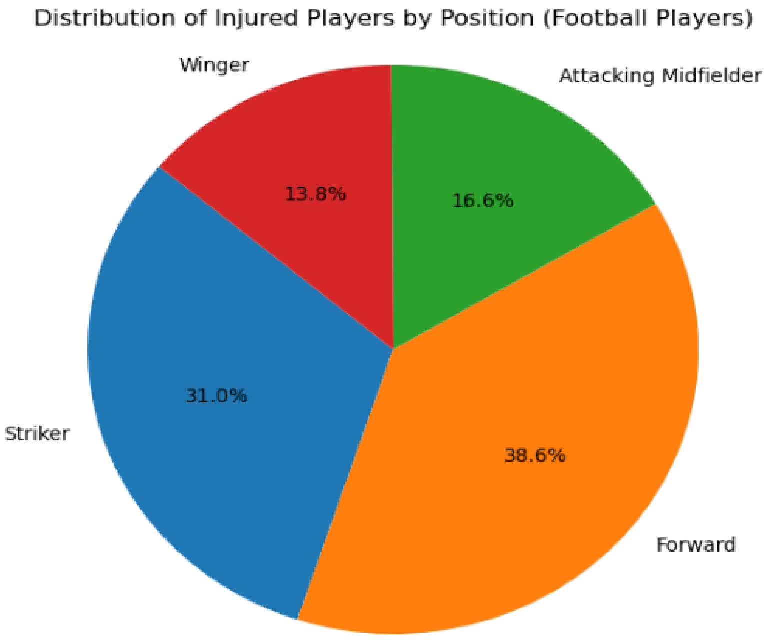


Figure 1 - Pie chart for data table

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

Answer – Here the probability that a randomly chosen player would suffer an injury, we need to use the total number of injured players and the total number of players.

As per Table 1 -

- Total Number of Injured Players: 145
- Total Number of Players: 235

The probability **P** of a randomly chosen player being injured can be calculated by –

P(injury) = Number of Injured Players **/** Total number of players

P(injury) = 145 / 235 = 29 / 47 = 0.617

Hence, the probability is approx. – **61.7 %**

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

Answer – Here the probability of a player being a forward or a winger, we need to first add the number of forward and winger players and then divide by the total number of players.

As per Table 1 -

- Number of Forwards: 94
- Number of Wingers: 29

Total number of Forward or Winger players: **-** 94+29=123

Total Number of Players: 235

Now, the probability **P** of a randomly chosen player being either a Forward or a Winger is –

P (Forwards or Winger) = Number of Forward or Winger / Total number of Players

P (Forwards or Winger) = 123 / 235 = 0.523

Hence, the probability that a randomly chosen player is a forward or a winger is **52.3%**

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

Answer – Here the probability of a randomly chosen player being a striker with a foot injury, we need to divide the number of players who are both strikers and injured by the total number of players.

As per Table 1 –

- Number of players who are strikers and injured = 45
- Total number of players = 235

The probability **P** of a player being both an injured Striker -

P (Injured Striker) = Number of players who are strikers and injured / Total number of Players

P (Injured Striker) = 45 / 235 = 0.191

Hence, there is a **19.1%** chance that a randomly selected player is a striker and has a foot injury.

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

Answer – Here, the probability that a randomly chosen injured player is a Striker can be calculated as –

As per Table 1 –

- Number of injured strikers = 45
- Total number of injured players = 145
-

The probability **P** of an injured player being a Striker is –

P (Striker **|** Injured) = Number of injured strikers / Total number of injured players

P (Striker **|** Injured) = 45 / 145 = 0.310

Hence the probability that a randomly chosen injured player is a striker is **31.0%**

Problem Statement 2

The breaking strength of gunny bags used for packaging cement is normally distributed with a mean of 5 kg per sq. centimeter and a standard deviation of 1.5 kg per sq. centimeter. 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

Based on above data, we need to answer following questions –

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

2.2 What proportion of the gunny bags have a breaking strength of 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.?

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

Data Analysis of problem statement –

We're given that the breaking strength of gunny bags follows a normal distribution with -

- Mean (μ) = 5 kg/cm²
- Standard deviation (σ) = 1.5 kg/cm²

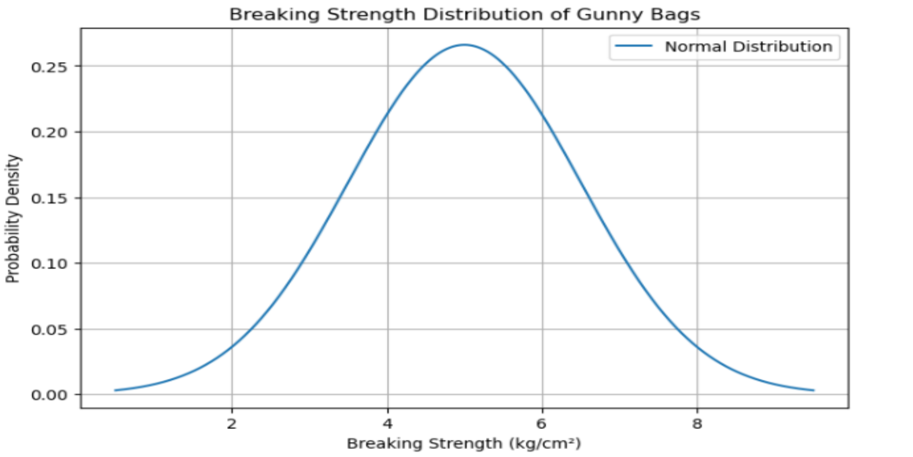


FIGURE 2 - BREAKING STRENGTH DISTRIBUTION GRAPH

Objective - The quality team wants to understand the packaging material better to identify potential wastage or pilferage.

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

Answer –

We have –

- Mean (μ): 5 kg per sq. cm
- Standard Deviation (σ): 1.5 kg per sq. cm
- Breaking Strength (X): 3.17 kg per sq. cm

Here, we are using Z-score formula –

$Z = (X - \mu) / \sigma$

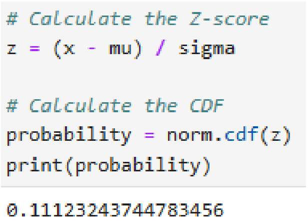
$Z = (3.17 - 5) / 1.5 = -1.22$

Now we can find Probability using the Z-score through -

From Z-table,

P $(Z \leq -1.22) \approx$ **0.1112**

Alternatively, we can use norm CDF function as below –



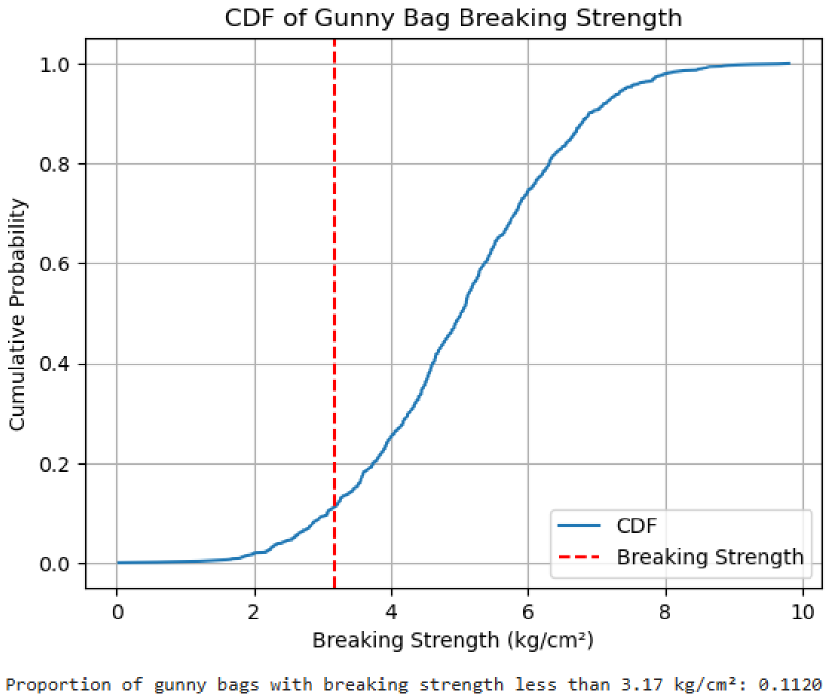
```
# Calculate the Z-score
z = (x - mu) / sigma

# Calculate the CDF
probability = norm.cdf(z)
print(probability)
```

```
0.11123243744783456
```

Figure 3 - Norm CDF function

Hence, the proportion of gunny bags with a breaking strength of less than 3.17 kg per sq. cm is **11.12 %**

Below CDF plot for reference (~ 1000 random samples) –



```
Proportion of gunny bags with breaking strength less than 3.17 kg/cm²: 0.1120
```

FIGURE 4 - CDF PLOT

- Follows normal distribution
- The red vertical line at 3.17 kg/cm² helps to visualize the proportion of bags with a breaking strength below this value, which was previously calculated to be approximately ~11 %.

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

Answer – Here, to find the proportion of gunny bags with a breaking strength of at least 3.6 kg/sq. cm, we need to calculate the probability that the breaking strength is greater than or equal to 3.6 kg/sq. cm.

As we know-

- Mean (μ): 5 kg per sq. cm
- Standard Deviation (σ): 1.5 kg per sq. cm

Calculation of Z-score:

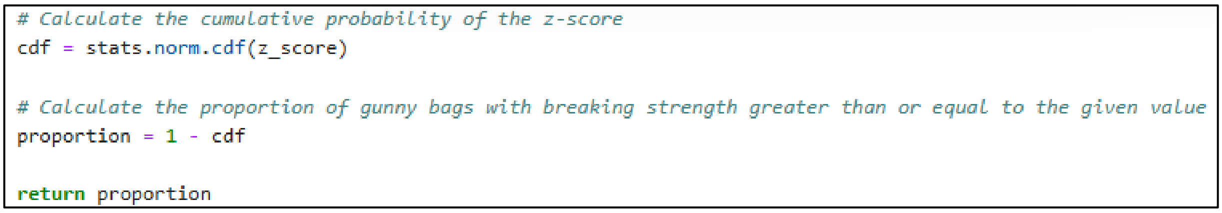
- Z = (X - μ) / σ
- Z = (3.6 - 5) / 1.5 = -0.9333

Calculating area left of Z –

We can use norm CDF function to find area left to Z = -0.9333 by subtracting as per below code-

Proportion = 1− Cumulative Probability

Proportion = 1−0.1757 = **0.8243**



```
# Calculate the cumulative probability of the z-score
cdf = stats.norm.cdf(z_score)

# Calculate the proportion of gunny bags with breaking strength greater than or equal to the given value
proportion = 1 - cdf

return proportion
```

Figure 5 - CDF for Problem 2.2

Hence, approx. **82.38 %** of the gunny bags have a breaking strength of at least 3.6 kg/cm².

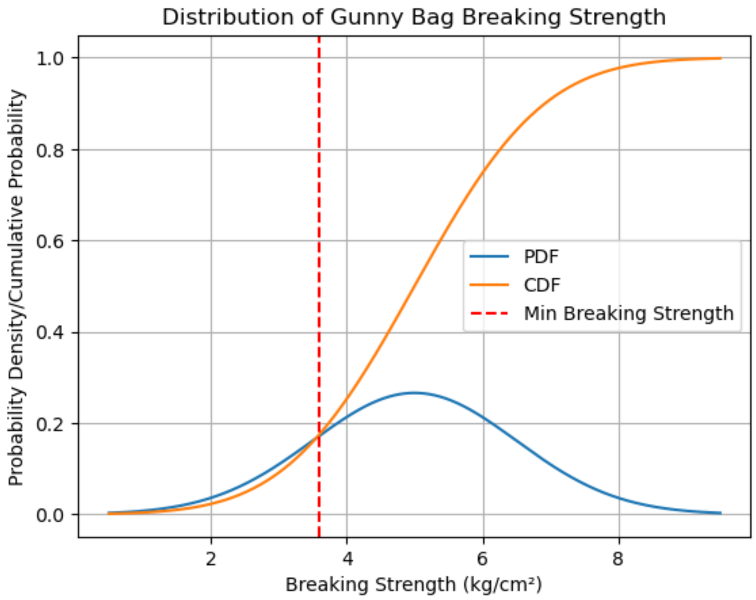


Figure 6 - Plot for Problem 2.2

- **Lines:**
  - Blue line: Probability Density Function (PDF)
  - Orange line: Cumulative Distribution Function (CDF)
  - Red dashed line: Minimum Breaking Strength

- The shape of the PDF curve indicates a normal distribution.
- The peak of the PDF curve represents the most likely breaking strength, which is around 5 kg/cm². This aligns with the mean value.
- The CDF curve shows the cumulative probability of a bag having a breaking strength less than or equal to a mentioned value. As the breaking strength increases, the CDF approaches 1, indicating that all bags have a breaking strength less than or equal to the maximum value.
- The red dashed line represents a critical breaking strength value. Bags with breaking strengths below this line might be considered subpar or defective.

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

Answer – Here to find the proportion of gunny bags with a breaking strength between 5 and 5.5 kg/sq. cm,

As we know –

- Mean (μ): 5 kg per sq. cm
- Standard Deviation (σ): 1.5 kg per sq. cm

Converting the Breaking Strength Values to Z-Scores:

For the lower bound (5 kg/sq. cm): z1= (5 – 5) / 1.5 = 0

For the upper bound (5.5 kg/sq. cm): z2 = (5.5 – 5) / 1.5 = 0.333

Now, we can calculate Cumulative Probability for Z- scores –

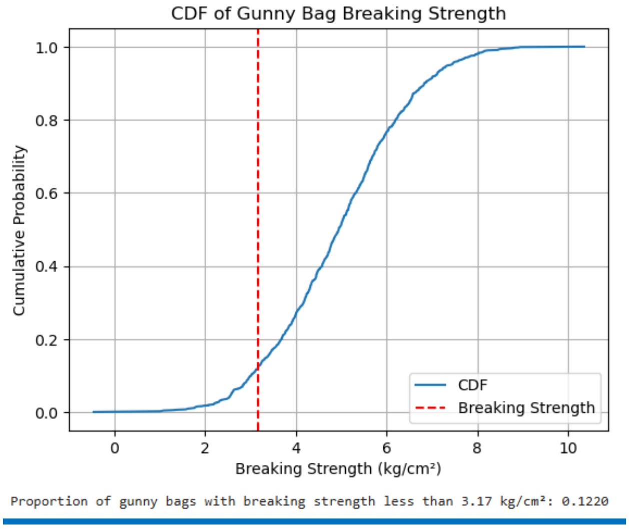
- For z1 = 0, the cumulative probability is 0.5000 (since a Z-score of 0 corresponds to the 50th percentile).
- For z2=0.333, we used norm CDF function (refer code) to calculate the cumulative probability which is ~ 0.6304

Finally, we can calculate the Proportion Between the Two Values –

Proportion = Cumulative Probability at z2 − Cumulative Probability at z1

Proportion = 0.6293 − 0.5000 = 0.1293

Hence, approximately **12.93% ~ 13%** of the gunny bags have a breaking strength between 5 and 5.5 kg/sq. cm.



```
Proportion of gunny bags with breaking strength less than 3.17 kg/cm²: 0.1220
```

Figure 7 - Question 2.3 CDF Plot

From random samples with n=1000, as per graph we can observe that –

- The distribution of gunny bag breaking strengths is concentrated between 4 and 6 kg/cm², as indicated by the steep incline of the cumulative distribution function (CDF) in this range.
- A smaller percentage of bags exhibit significantly lower or higher breaking strengths, reflected in the CDF's flatter sections at the extremes.
- The calculated proportion of 12.20% of bags with a breaking strength below 3.17 kg/cm² corroborates the visual analysis of the CDF.

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

Answer –

Here, to find the proportion of gunny bags that have a breaking strength outside the range of 3 to 7.5 kg/cm² is the objective.

Convert the Breaking Strength Values to Z-Scores:

For the lower bound (3 kg/sq. cm) : z1 = (3 – 5) / 1.5 = **-1.333**

For the upper bound(7.5 kg/sq. cm) : z2 = (7.5 – 5) / 1.5 = **1.667**

Now, we can calculate cumulative probabilities for z-scores as per code(norm CDF function) -

For z1=−1.333, the cumulative probability is approximately 0.0918

For z2=1.667, the cumulative probability is approximately 0.9525

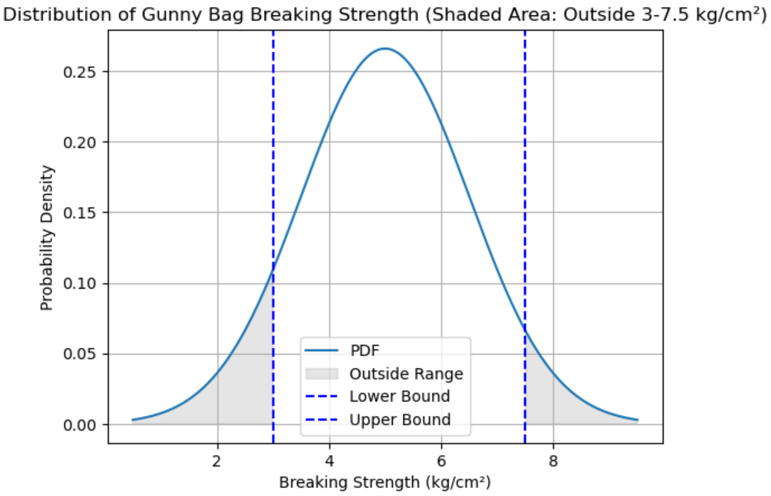


FIGURE 8 - QUESTION 2.4 PDF PLOT

Now to find the proportion between of gunny bags with a breaking strength between 3 and 7.5 kg/sq. cm –

P (between 3 and 7.5) = Cumulative P(z2) – Cumulative P(z1) = 0.9525 – 0.0918 =0.8607

Finally, to find proportion of gunny bags NOT in between the range –

P (NOT between 3 and 7.5) = 1 - P (between 3 and 7.5) = 1- 0.8607 = **0.1393**

Hence, approximately **13.93%** of the gunny bags have a breaking strength outside the range of 3 to 7.5 kg/sq. cm.

Observations -

- The distribution of gunny bag breaking strengths primarily falls within the acceptable range of 3 to 7.5 kg/cm², evident in the graph's central, unshaded area.
- Instances of bags with extremely low or high breaking strengths are infrequent, represented by the smaller shaded regions.
- This normal distribution, indicated by the bell-shaped curve, allows for the establishment of quality control standards to detect manufacturing anomalies.
- The curve's width signifies the degree of variation in bag strength, with narrower curves indicating greater consistency.

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);

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?

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

Data Analysis of problem statement –

We're given that the stone surface has to have a Brinell's hardness index of at least 150 and assuming 5% significance level.

The dataset contains data corresponding to polished and unpolished stones.

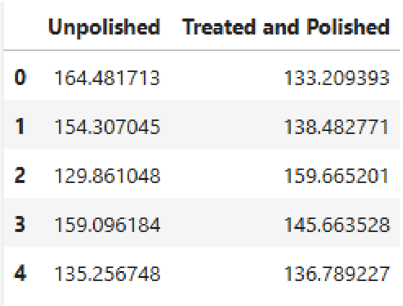
Sheet name – *Zingaro_Company.csv*

Data Dictionary –

- **Unpolished** – Brinell's hardness index for unpolished stones
- **Treated and Polished** - Brinell's hardness index for unpolished stones

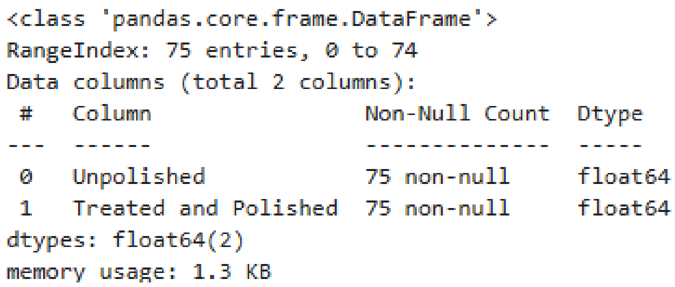
Data Overview –

To start the data analysis, we need to import the necessary libraries, specify the working directory, and load the dataset. Then, we will view the first five rows using head () function to get a preliminary understanding of the data. The Dataset has 75 number of rows with 2 columns.



	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

TABLE 2 - TOP 5 ROWS OF DATASET



```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 75 entries, 0 to 74
Data columns (total 2 columns):
 #   Column                Non-Null Count  Dtype  
---  ------                --------------  -----  
 0   Unpolished            75 non-null     float64
 1   Treated and Polished  75 non-null     float64
dtypes: float64(2)
memory usage: 1.3 KB
```

TABLE 3 - BASIC INFO. OF DATASET

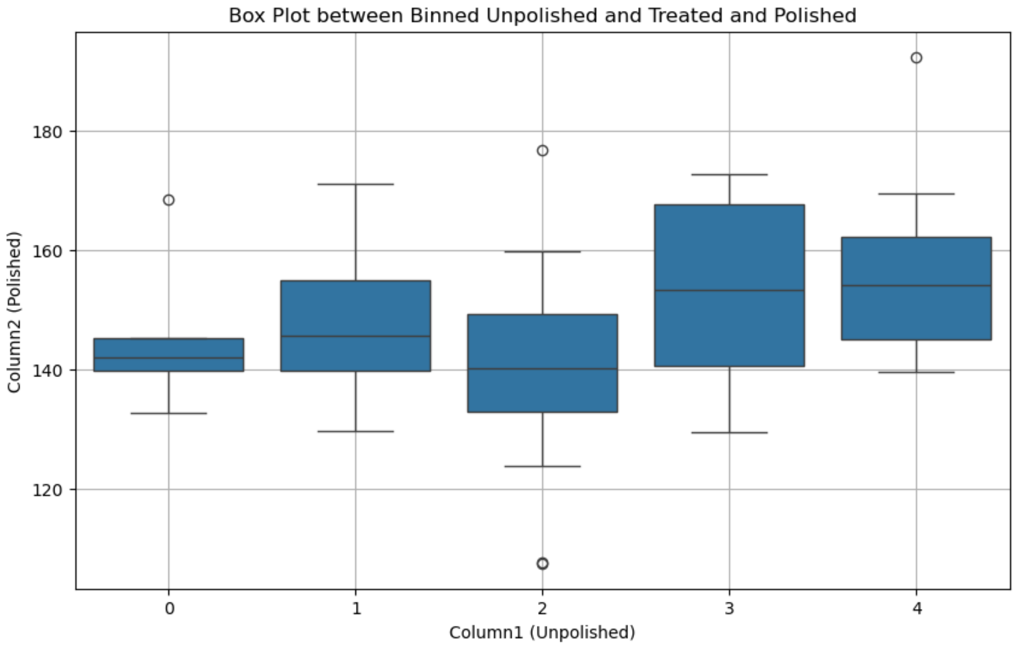


FIGURE 9 - BOXPLOT OF DATASET

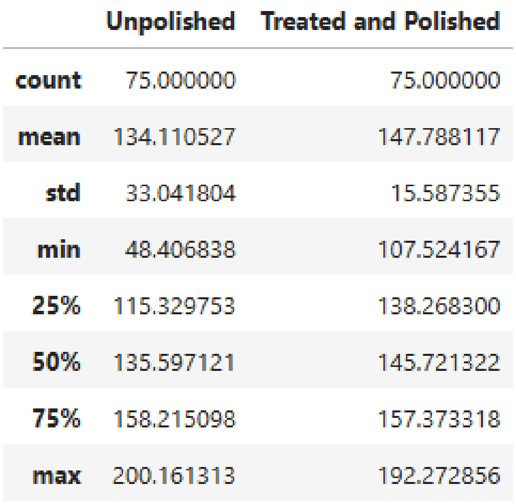
- Both groups exhibit a rightward skew, indicated by longer tails on the right side of the box plots.
- Outliers are present in both groups, as shown by data points beyond the box plot's whiskers.
- The median value is higher for the Polished group compared to the Unpolished group.
- The Polished group also shows slightly more variability than the Unpolished group, as indicated by its larger interquartile range.

Missing value treatment -

On analysis, we can observe there are no null values in below two columns in dataset.

Statistical Summary –

Using Describe () function, we can analyses the summary statistics of the dataset –



	Unpolished	Treated and Polished
count	75.000000	75.000000
mean	134.110527	147.788117
std	33.041804	15.587355
min	48.406838	107.524167
25%	115.329753	138.268300
50%	135.597121	145.721322
75%	158.215098	157.373318
max	200.161313	192.272856

Table 4 - Statistical summary of dataset

Observations-

- The average value for the "Treated and Polished" group (147.79) is higher than for the "Unpolished" group (134.11). This suggests that the treatment and polishing process has a beneficial effect, leading to higher average values.

- The "Treated and Polished" group has a lower standard deviation (15.59) compared to the "Unpolished" group (33.04). This indicates that the values in the "Treated and Polished" group are more consistent and less dispersed around the mean, while the "Unpolished" values show greater variability.

- The range of values (from minimum to maximum) for the "Treated and Polished" group (107.52 to 192.27) is somewhat wider than that of the "Unpolished" group (48.41 to 200.16). However, the "Treated and Polished" values are more concentrated around the mean, as evidenced by the lower standard deviation and closer percentiles.

- The difference in the range and the standard deviations suggest that the "Unpolished" data may be more skewed compared to the "Treated and Polished" data. The "Unpolished" data's larger variability and wider range could indicate a more spread-out distribution, possibly with more extreme values affecting the mean.

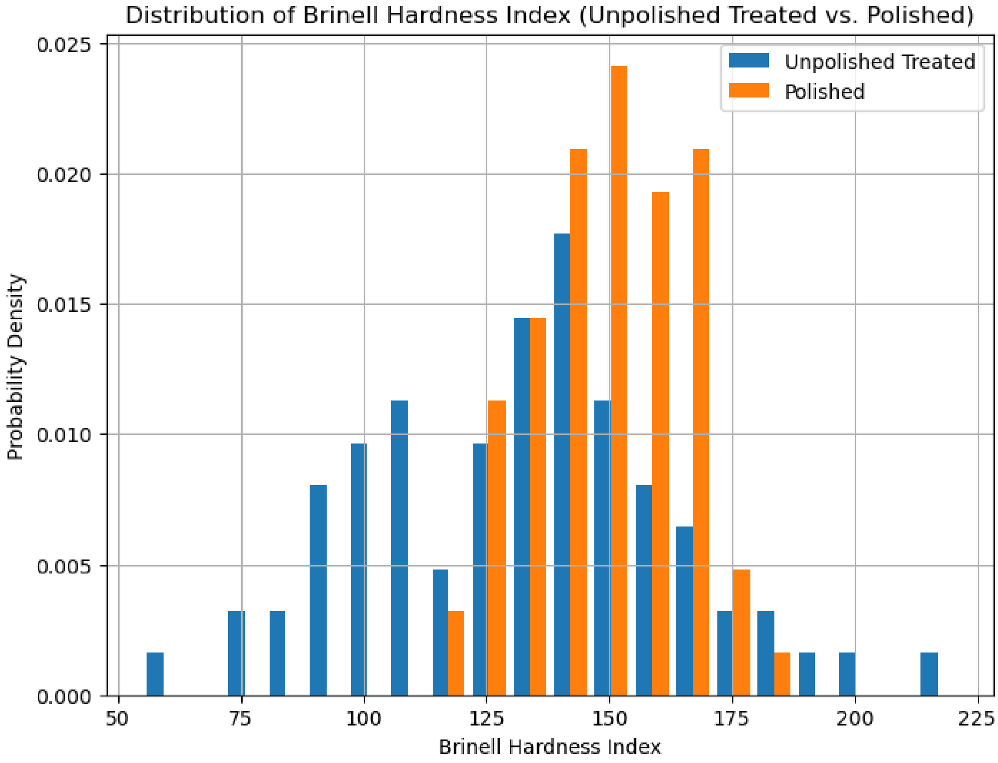


FIGURE 10 - BRINELL HARDNESS INDEX PLOT

Hypothesis Testing –

1: Define null and alternative hypotheses

The null and alternative hypotheses can be formulated as:

H0 (Null hypothesis): The mean hardness index of Unpolished stones is equal to or greater than 150

Ha (Alternate Hypothesis) : The mean hardness index of Unpolished stones is less than 150

2: Select appropriate test

We will do **1 sample z-test** as the dataset is a one-sided sample with sample size is 75>30. We know standard deviation as well.

Hence, n = 75 and μ = 150

3: Decide Significance level

As provided in problem statement, let have significance level (α) = 5 % or 0.05

It means there is a 5% risk of rejecting the null hypothesis when it is actually true.

4: Calculate P value

We can use two-sample z-test comparing the means of the "Unpolished" and "Treated and Polished" stone groups, we'll use the following information:

Unpolished Stone:

- Count (Sample Size, n1): 75
- Mean (X̄1): 134.11
- Standard Deviation (σ1): 33.04

Treated and Polished Stone:

- Count (Sample Size, n2): 75
- Mean (X̄2): 147.79
- Standard Deviation (σ2): 15.59

➢ The **z-test statistic** for comparing two sample means is given by:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Figure 11 - Two sample z test formula

➢ Calculate the **Standard Error (SE):**

$$SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Figure 12 - Standar Error Formula

As per calculation in code, **SE = 4.22**

➢ **Now,** Calculate the Z-Statistic:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2)}{SE}$$

Figure 13 - Z statistic formula

As per calculation in code, **Z = -3.24**

Find the P-Value:

Using a standard normal CDF statistical calculation as per formula: 2×(1−CDF(|Z|))

we find the p-value associated with Z = −3.24 for a two-tailed test.

The p-value is approximately **0.0012**

5: Compare the p-value with α

Significance Level (α): As alpha is specified at 0.05, then since the p-value (0.0012) is less than 0.05, **we can reject the null hypothesis**

6: Draw Interference - So, unpolished stones do not have a Brinell's hardness index of at least 150

On other side, **the p-value for treated and polished stones is 0.21(refer notebook)**, which is greater than the significance level of 0.05. Therefore, **we fail to reject the null hypothesis**. This suggests that both treated and polished stones have a Brinell hardness index of at least 150

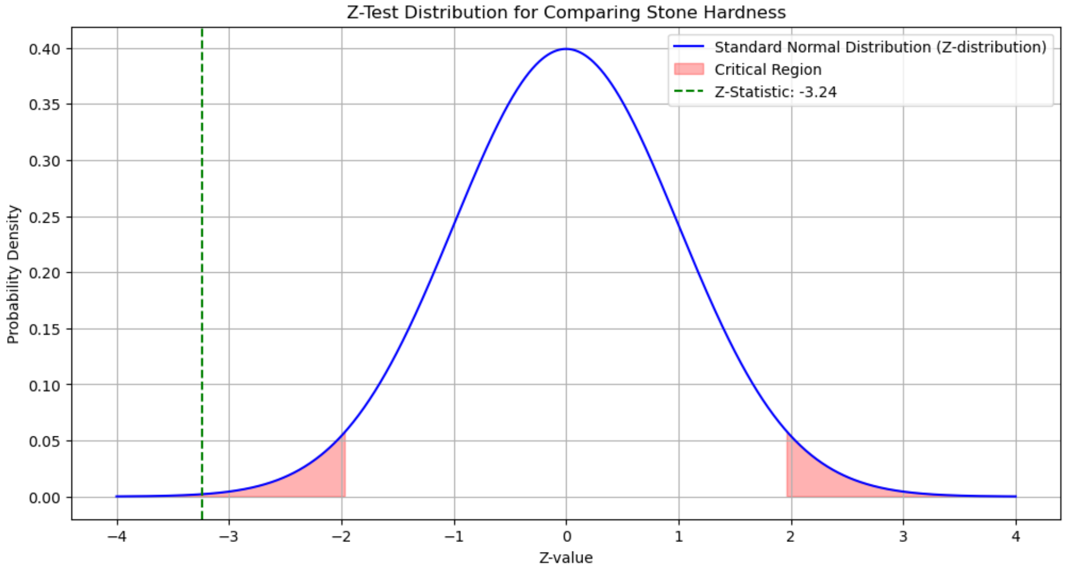


Figure 14 - Z- Test Distribution

Observations:

- The presence of critical regions in both tails confirms a two-tailed test.
- The total area of the critical regions is 5%, indicating an alpha level of 0.05.
- The Z-statistic of -3.24 falls within the left critical region.
- The Z-distribution is symmetrical around zero, as expected for a standard normal distribution.
- Since the Z-statistic falls within the left critical region, we reject the null hypothesis at the 5% significance level. The Z-statistic of -3.24 is quite far from the mean (zero), indicating strong evidence against the null hypothesis.

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?

Answer – Yes.

Zingaro's concern that the unpolished stones might not be suitable for printing is well-established. The p-value shows a statistically significant difference, indicating that the mean Brinell hardness index of the unpolished stones falls notably below the required threshold of 150. This suggests that the unpolished stones likely do not meet the hardness criteria necessary for optimal printing and may therefore be unsuitable for the intended printing process

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

Answer – No, the hardness of polishes and unpolished stones are significantly different.

Here,

H0 (Null hypothesis) - The mean hardness index of Unpolished stones is equal to Polished stones

Ha (Alternate Hypothesis) - The mean hardness index of Unpolished stones is NOT equal to Polished stones

Here, as we do not know which group may have higher or lower values, so to determine if there is statistical difference between two groups…**we will use Two tailed T Test.**

Unpolished Stone:

- Count (n1): 75

- Mean (X¯1): 134.11
- Standard Deviation (σ1): 33.04

Treated and Polished Stone:

- Count (n2): 75
- Mean (X¯2): 147.79
- Standard Deviation (σ2): 15.59

Significance Level (α) - 5% or (0.05)

Steps for the Two-Tailed T-Test – It is solved in jupyter notebook file. Plz refer.

1. Calculate the Pooled Standard Deviation (if variances are assumed equal) – 25.77
2. Calculate the Standard Error (SE) of the difference between the means – 4.21
3. Calculate the T-Statistic :- -3.24

$$T = \frac{\bar{X}_1 - \bar{X}_2}{SE}$$

Figure 15 - T-Statistic formula

4. Calculate the Degrees of Freedom (df) – 148

$$df = n_1 + n_2 - 2$$

Figure 16 - Difference of Freedom formula

5. Calculate the P-Value using the t-distribution with the calculated degrees of freedom. – p-value=2×(1−CDF(|T|))

p-value – 0.0014

Hence, with a p-value of 0.001, the results of the two-tailed test are statistically significant at the 0.05 level. The very low p-value strongly suggests that the difference observed is real and not due to chance. We can accept the alternative hypothesis.

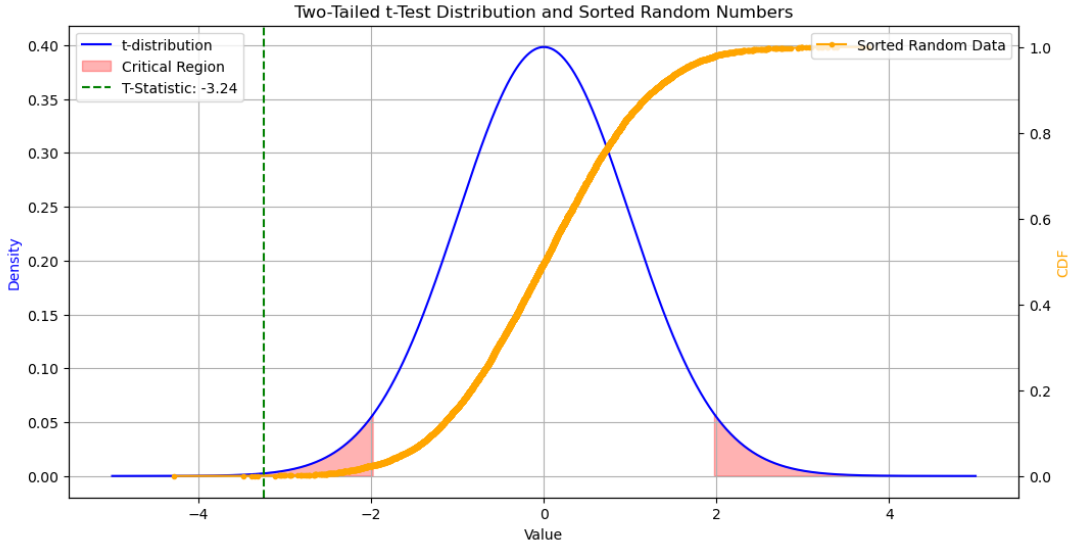


FIGURE 17 - TWO TAILED DISTRIBUTION

Observations:

- The critical regions are symmetrical, indicating a two-tailed test.
- The t-statistic of -3.24 falls within the left critical region, suggesting a significant result in our favour that we can accept Ha.
- The sorted random data follows a generally increasing pattern, as expected for a CDF.
- The t-distribution is bell-shaped and symmetric, typical of a t-distribution

Based on the plot, it appears that the calculated t-statistic is significantly different from zero, leading to the rejection of the null hypothesis.

Problem Statement 4

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.

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

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

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

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

Data Analysis of problem statement –

We're given that the dataset for dental implants on basis of alloys used, methods and temperature.

The dataset contains data corresponding to polished and unpolished stones.

Sheet name – *Dental+Hardness+data.csv*

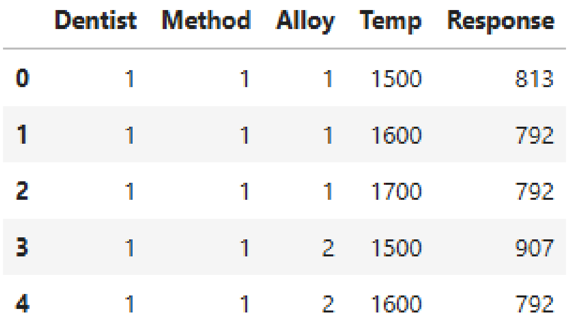
Data Dictionary –

- **Dentist:** Categorical variable representing different dentists.
- **Method:** Categorical variable representing different implant methods.
- **Alloy:** Categorical variable representing different alloys used.
- **Temp:** Numerical variable representing the temperature at which the metal is treated.
- **Response:** Numerical variable representing the hardness of the implant (the variable of interest)

Data Overview –

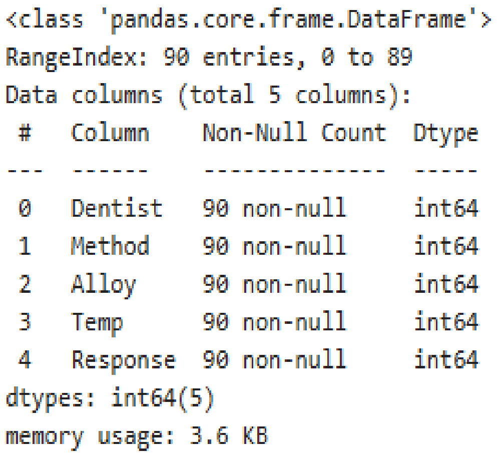
To start the data analysis, we need to import the necessary libraries, specify the working directory, and load the dataset.

Then, we will view the first five rows using head () function to get a preliminary understanding of the data. The Dataset has 90 number of rows with 5 columns.



	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 5 - TOP 5 ROWS OF Q4 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 6 - Basic Info. of Dental dataset

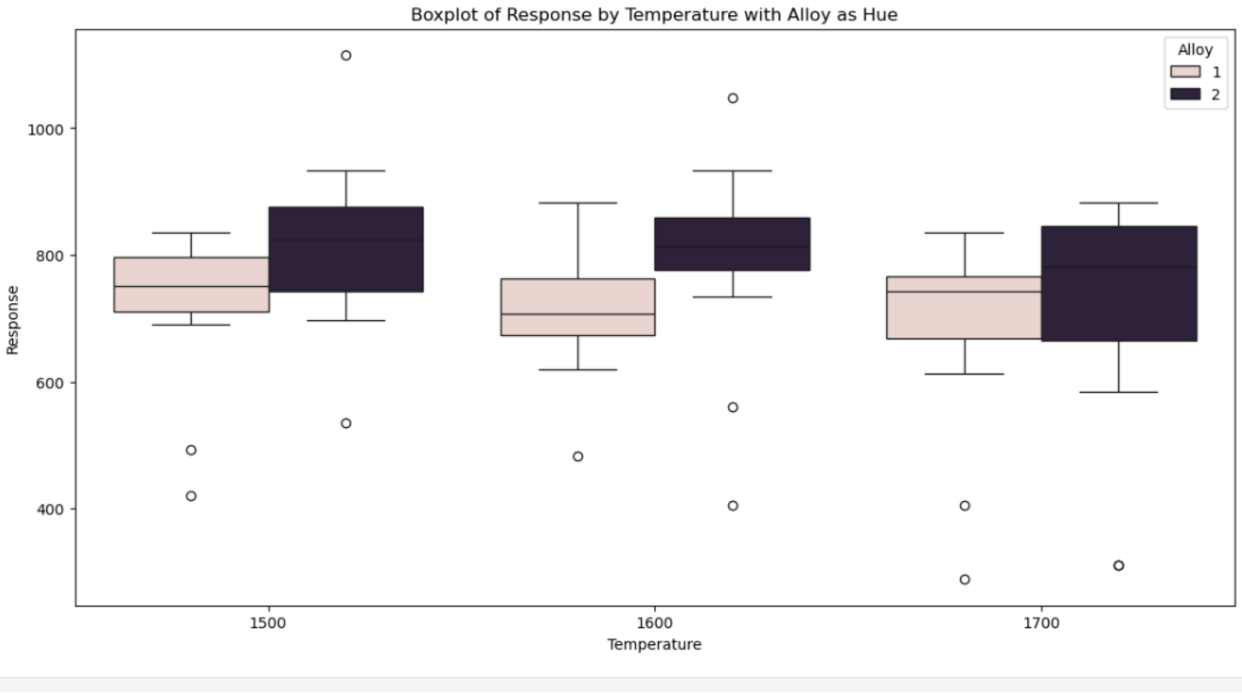


FIGURE 18 - BOXPLOT OF DENTAL DATASET WITH HUE AS ALLOY

The boxplot illustrates the distribution of the "Response" variable—potentially indicating implant hardness or a similar property—varies with different "Temperature" levels, while distinguishing between "Alloy" types.

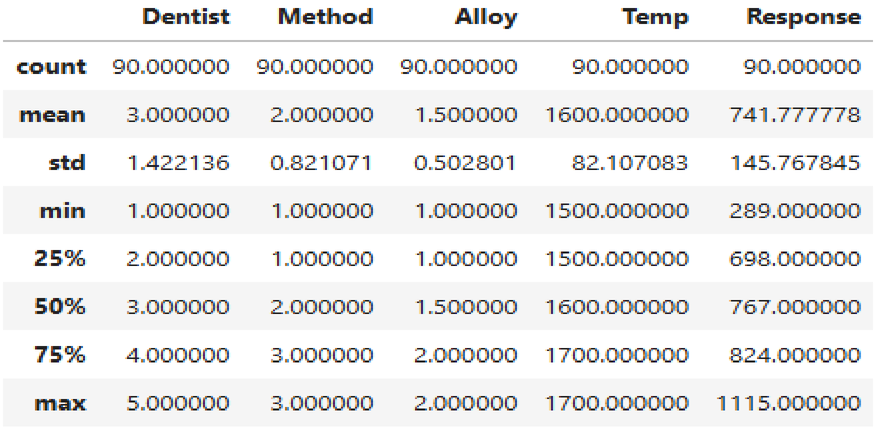
- There appears to be a slight downward trend in the median response as temperature rises from 1500 to 1700. However, this trend is not very strong and may not be statistically significant without additional analysis.
- The data's spread, as shown by the boxes and whiskers, seems relatively consistent across various temperatures.
- Alloy 2 consistently exhibits higher median response values compared to Alloy 1 at all temperature levels, suggesting that Alloy 2 generally yields higher response values.
- The data dispersion for both alloys appear to be similar at each temperature level.
- There are several outliers in the data, particularly at lower response values. These outliers could indicate unusual cases or experimental errors and warrant further investigation.

Missing value treatment -

On analysis, we can observe there are no null values in below two columns in dataset.

Statistical Summary –

Using Describe () function, we can analyses the summary statistics of the dataset –



	Dentist	Method	Alloy	Temp	Response
count	90.000000	90.000000	90.000000	90.000000	90.000000
mean	3.000000	2.000000	1.500000	1600.000000	741.777778
std	1.422136	0.821071	0.502801	82.107083	145.767845
min	1.000000	1.000000	1.000000	1500.000000	289.000000
25%	2.000000	1.000000	1.000000	1500.000000	698.000000
50%	3.000000	2.000000	1.500000	1600.000000	767.000000
75%	4.000000	3.000000	2.000000	1700.000000	824.000000
max	5.000000	3.000000	2.000000	1700.000000	1115.000000

Table 7 - Statistical summary of Dental dataset

Observations-

- The dataset includes 90 observations.
- The data is evenly distributed across dentists (mean = 3, std = 1.42) and methods (mean = 2, std = 0.82).
- There are two alloys (mean = 1.5, std = 0.5), suggesting a balanced representation.
- The temperature varies between 1500 and 1700 degrees (mean = 1600, std = 82.1).
- The response (hardness) has a significant range (min = 289, max = 1115), indicating considerable variation in implant hardness. The standard deviation (145.77) suggests a relatively high level of dispersion around the mean (741.78).

To get more understanding, lets plot histogram for all columns with response as variable.

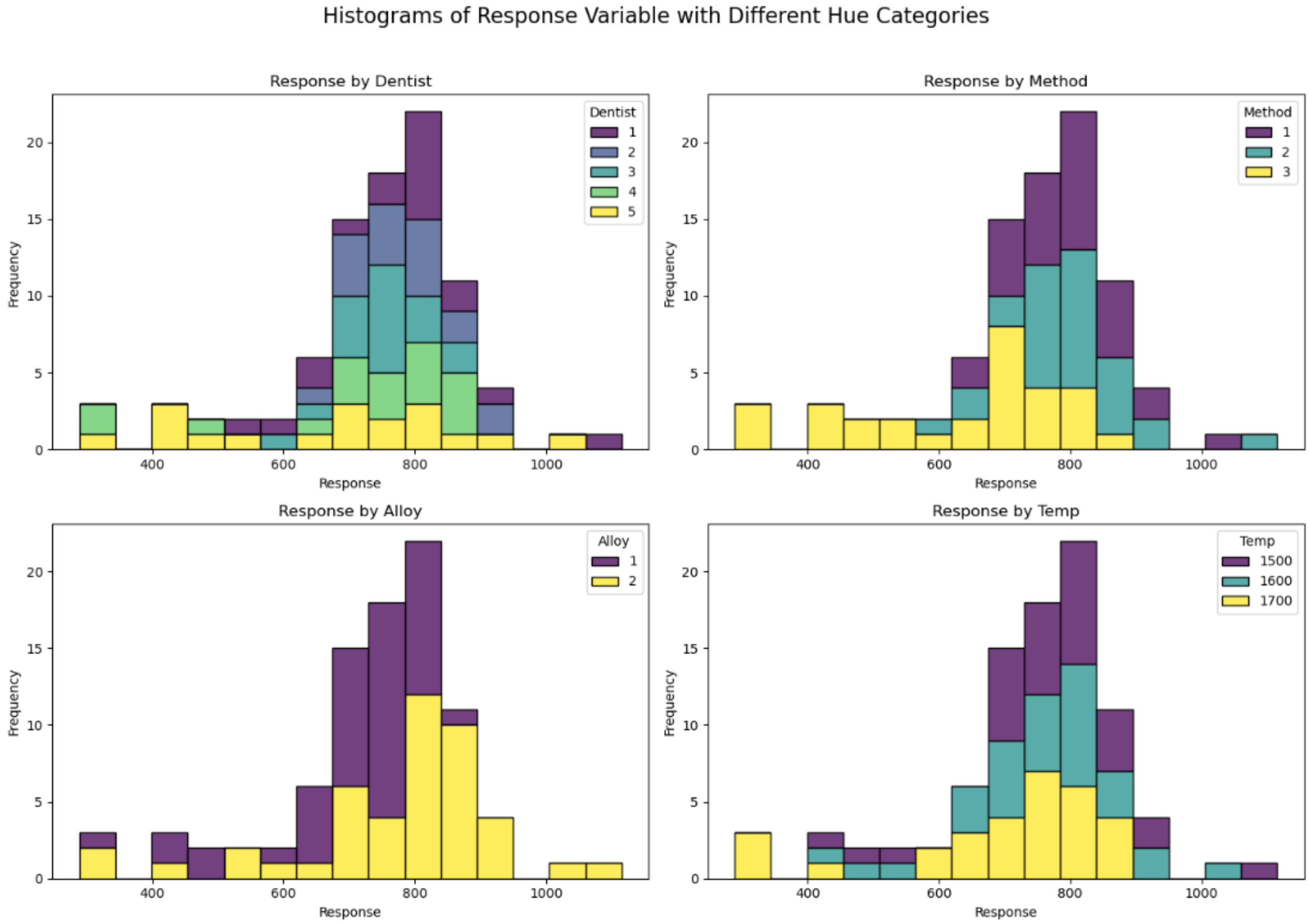


FIGURE 19 - HISTOGRAM FOR DENTAL DATASET WITH VARIABLE AS RESPONSE

Overall, the response data is skewed to the right, with most values clustering between 600 and 800.

- **Dentist:** Different dentists produced varying response distributions. Some dentists had lower-valued results, while others had a wider range.
- **Method:** The implant method seemed to have little impact on the response distribution, although Method 1 might slightly favour higher values.
- **Alloy:** Alloy 1 showed a wider range of responses, while Alloy 2 tended to produce more middle-of-the-road results.
- **Temperature:** Higher temperatures were associated with lower response values.

Hypothesis Testing –

Alloy 1 and Alloy 2-

1. Define null and alternative hypotheses State the null and alternate hypotheses

Ho(alloy1): The mean Hardness of Dental implant is same for all 5 dentists provided the Alloy 1 is used

Ha(alloy1): for at least 1 dentist the mean Hardness of Dental implant is different when using Alloy 1 used

Similarly,

Ho(alloy2): The mean Hardness of Dental implant is same for all 5 dentists provided the Alloy 2 is used

Ha(alloy2): for at least 1 dentist the mean Hardness of Dental implant is different when using Alloy 2 used

2. Select Appropriate test

Since each alloy is need to be analyzed separately, it involved conducting **one-way ANOVA** tests for each alloy type independently. This approach isolates the effect of dentists on hardness for each alloy type separately.

Assumptions for One-Way ANOVA:

1. Observations are independent of each other.
2. The residuals (errors) are normally distributed within each group.
3. The variance among the groups should be approximately equal.

Given the dataset,

- We will analyse using histograms.
- Check variances using Levene’s test

Steps are as follows (plz refer jupyter notebook for actual calculations) -

a) **Calculate Means:**

- Calculate the overall mean of the Response variable for Alloy 1 and alloy 2
- Calculate the mean Response for each Dentist group within Alloy 1 and alloy 2

b) **Calculate Sum of Squares:**

- Calculate the Total Sum of Squares (SST), Sum of Squares Between Groups (SSB), and Sum of Squares Within Groups (SSW).

$$SSB = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2 \qquad SSW = \sum_{i=1}^{k} (n_i - 1) \cdot s_i^2$$

c) **Calculate Degrees of Freedom:**

- Calculate the degrees of freedom for Between Groups (dfB) and Within Groups (dfW).

d) **Calculate Mean Squares:**

- Calculate the Mean Square Between Groups (MSB) and Mean Square Within Groups (MSW).

$$MSB = \frac{SSB}{dfB} \quad MSW = \frac{SSW}{dfW}$$

e) **Calculate F-statistic:**

- Calculate the F-statistic: F = MSB / MSW.

3) Determine p-value:

We can use a F-distribution table or statistical software as used in code to find the p-value corresponding to the calculated F-statistic, dfB, and dfW.

Hence, we used the f_oneway() function from the scipy.stats library to perform a one-way ANOVA test. The f_oneway() function takes the sample observations from the different groups and returns the test statistic and the p-value for the test. **Plz refer notebook.**

Alloy 1 –

ANOVA F-value: 1.98
ANOVA p-value: 0.1166
There is no statistically significant difference in hardness (Response) between dentists.

Alloy 2 -

ANOVA F-value: 0.52
ANOVA p-value: 0.7180
There is no statistically significant difference in hardness (Response) between dentists.

Also, on performing **Levene’s test**, we got below values-

Levene's Test p-value for alloy 1: 0.2566

Levene's Test p-value for alloy 2: 0.2369

The p-value is greater than 0.05 for both alloys. **There is not enough evidence to reject the null hypothesis.** This means there is no significant difference in the mean hardness across the methods and temperatures for Alloy 1 & Alloy 2

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

Answer – As per above calculations and inference, as well as p-value is greater than the alpha…we can conclude that there are statistically significant differences in hardness depending on the dentists.

Dentists with higher mean hardness tend to produce implants that are harder on average. However, those with higher standard deviations show greater inconsistency in their results.

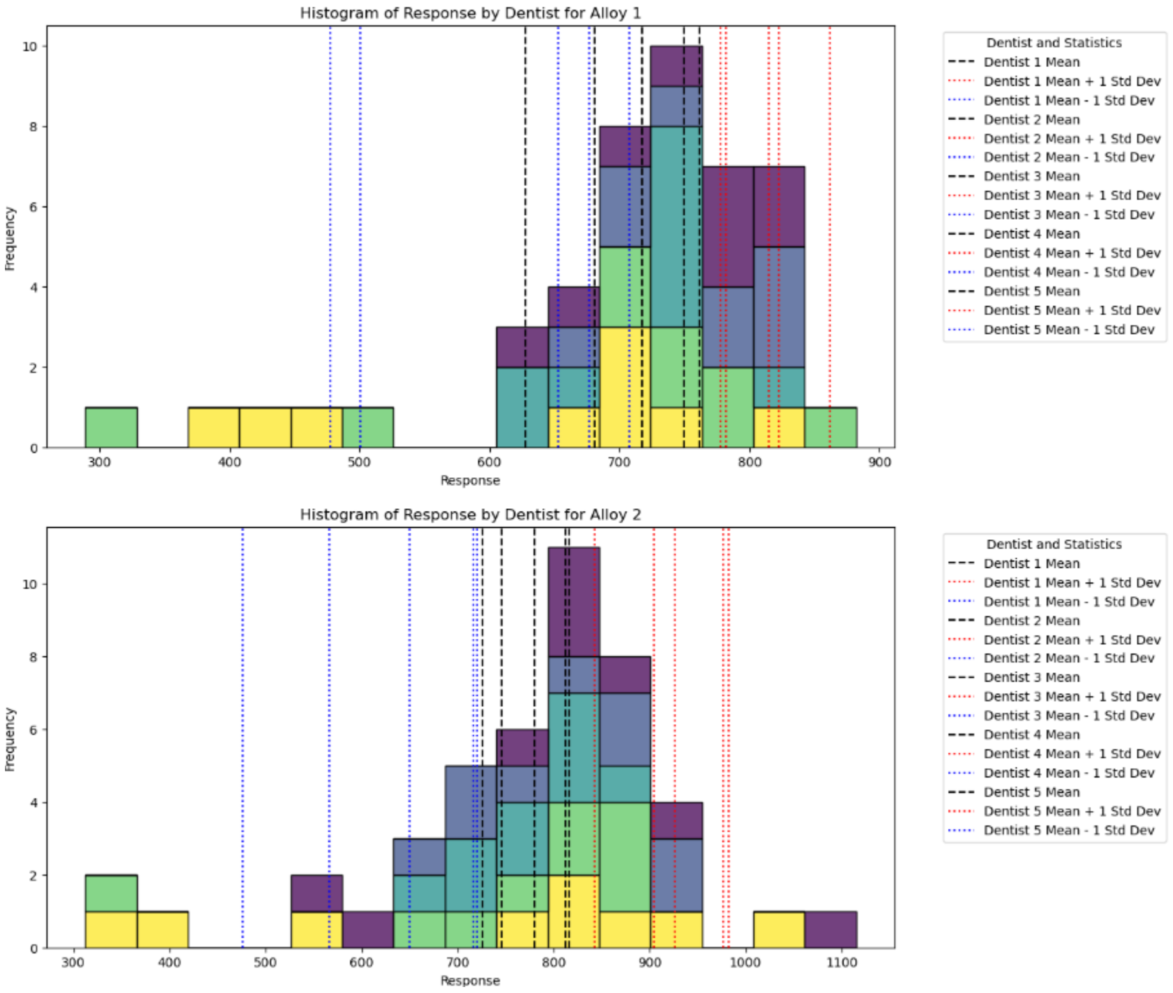


FIGURE 20 - HISTOGRAM PLOT FOR ALLOY 1 AND 2 FOR DENTISTS

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

Answer - Yes, hardness depends on methods. The alternate hypothesis that one or more methods are significantly different from the other can be accepted. Refer below solution.

1. Define null and alternative hypotheses State the null and alternate hypotheses

Ho(alloy1): The mean Hardness of Dental implant is same for all 3 methods provided the Alloy 1 is used

Ha(alloy1): for at least 1 method the mean Hardness of Dental implant is different when using Alloy 1 used

Similarly,

Ho(alloy2): The mean Hardness of Dental implant is same for all 3 methods provided the Alloy 2 is used

Ha(alloy2): for at least 1 method the mean Hardness of Dental implant is different when using Alloy 2 used

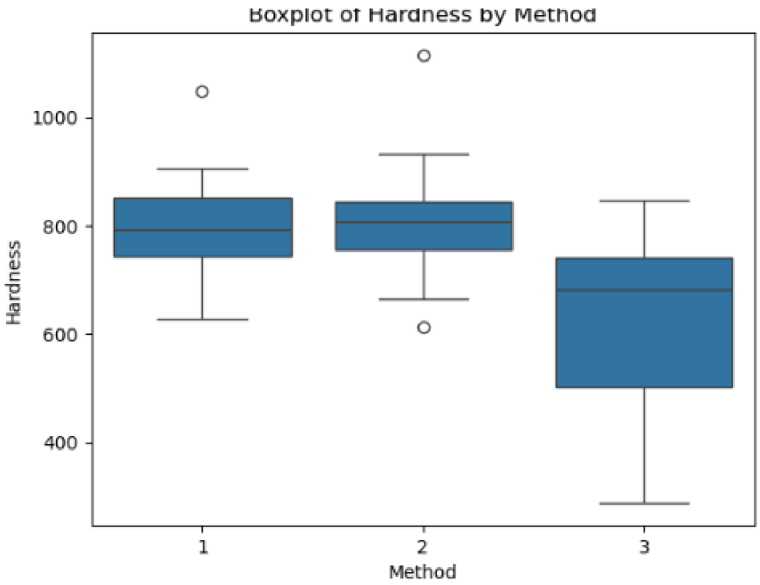


FIGURE 21 - BOXPLOT OF HARDNESS BY METHOD

- Method 1 has the highest median hardness, followed by Method 2 and then Method 3.
- Method 1 has the widest spread in hardness values, while Method 3 has the narrowest spread. Method 2's spread falls in between.
- There are two outliers present in the data, one for Method 1 and one for Method 2. Both of these outliers exhibit exceptionally high hardness values compared to the rest of the data points within their respective methods.

<u>Steps for ANOVA (similar to answer 1) - Plz refer jupyter notebook for calculations</u>

1. Calculate Group Means and Overall Mean
2. Calculate Sum of Squares (SS)
3. Calculate Degrees of Freedom (df)
4. Calculate Mean Squares (MS)
5. Calculate the F-statistic
6. Calculate p-value
7. Make inference

So as per calculations using python , we get below results-

Alloy 1 –

Levene's Test p-value: 0.003

ANOVA F-value: 6.26

ANOVA p-value: 0.0042

The differences in hardness (Response) between Methods are statistically significant.

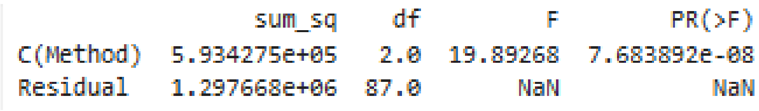
Alloy 2 –

Levene's Test p-value: 0.0447

ANOVA F-value: 16.41

ANOVA p-value: 0.0000

The differences in hardness (Response) between Methods are statistically significant.



```
                  sum_sq    df         F        PR(>F)
C(Method)   5.934275e+05   2.0  19.89268  7.683892e-08
Residual    1.297668e+06  87.0       NaN           NaN
```

Table 8 - C method and residual Anova results

- **sum_sq**: Sum of squares due to each factor.
- **df**: Degrees of freedom.
- **F**: F-statistic.
- **PR(>F)**: p-value.

As p-value (PR(>F)) is below 0.05, it suggests that there is a significant difference in hardness between the methods for both alloys. There is sufficient evidence against the null hypothesis that all methods have equal impact on hardness Hence, we can confirm that hardness of implants depends on methods.

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

Answer – Yes, there is interaction effect between the dentist and method on hardness level.

Here, we need to analyze the interaction effect between two categorical factors (Dentist and Method) on a response variable for different types of alloys in a dataset.

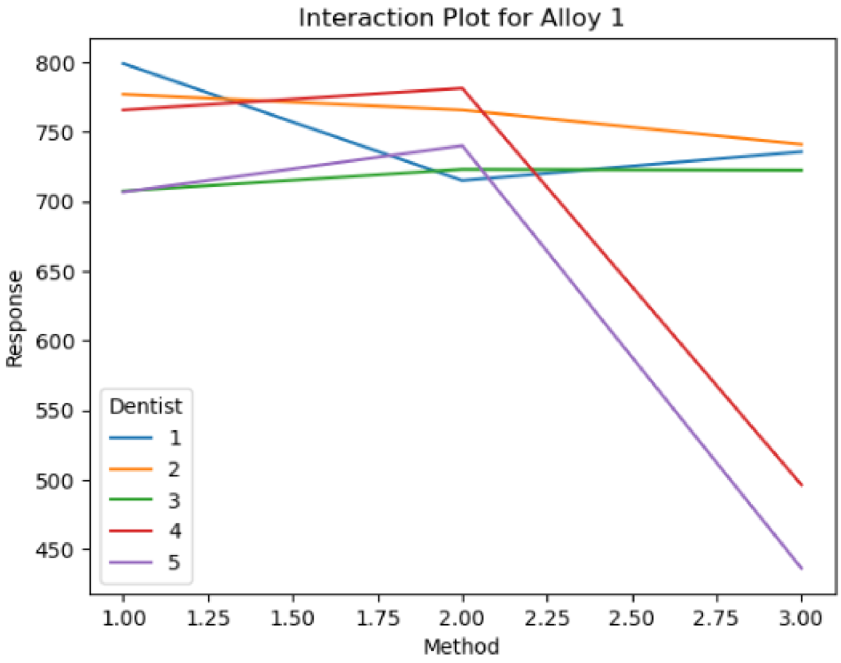


FIGURE 22 - PLOT BETW DENTIST AND METHOD FOR ALLOY 1

Response Trends:

- Dentists 1 and 2: Show a consistent decrease in Response as Method increases.
- Dentists 3 and 4: Exhibit a relatively stable Response across different Methods with slight variations.
- Dentist 5: Shows a slight increase in Response as Method increases.

- The lines representing different dentists (1 to 5) are not parallel. This suggests a notable interaction effect between the Dentist and Method on the Response for Alloy 1.

- There is some variation in the Response across different Dentists and Methods, but it is less pronounced than with Alloy 2.
- The interaction effect indicates that the combination of Dentist and Method has a significant impact on the Response for Alloy 1.
- The differing Response trends among Dentists imply potential variations in their expertise or techniques.
- The relatively lower variability in Response compared to Alloy 2 suggests that the factors influencing the Response for Alloy 1 might be simpler.

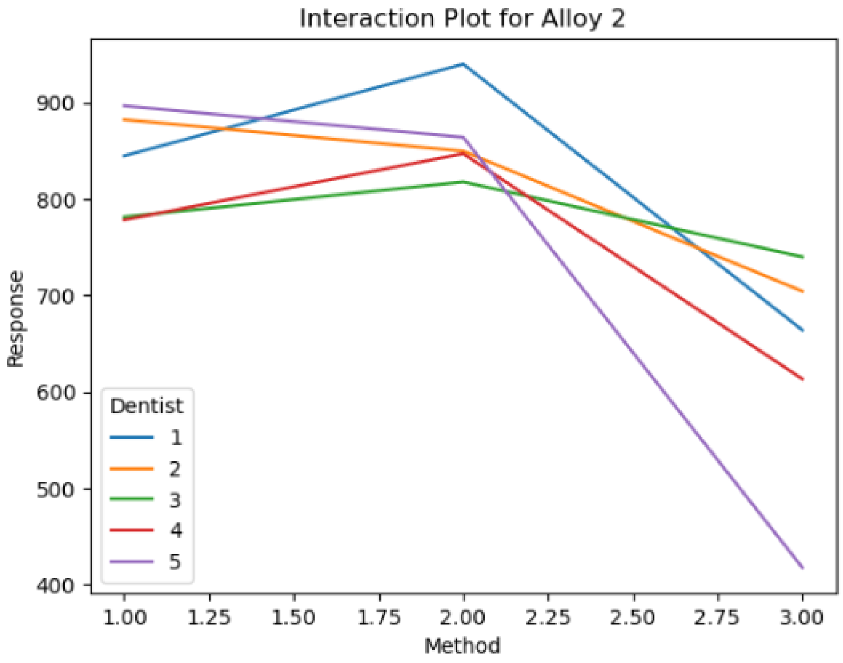


FIGURE 23 - PLOT BETW DENTIST AND METHOD FOR ALLOY 2

<u>Response Trends:</u>

- ➢ Dentist 1: Shows a steady increase in Response as Method increases.
- ➢ Dentist 2: Has a similar trend to Dentist 1 but with slightly lower Response values.
- ➢ Dentist 3: Starts with a lower Response, increases rapidly, and then plateaus.
- ➢ Dentist 4: Shows an initial increase in Response followed by a sharp decline.
- ➢ Dentist 5: Exhibits a consistent decrease in Response as Method increases.

- The interaction plot reveals that the lines representing different dentists (1 to 5) are not parallel, indicating a significant interaction effect between Dentist and Method on the

Response for Alloy 2. This suggests that the impact of the Method on the Response changes depending on which Dentist is performing the procedure.
- There is substantial variation in Response across different Dentists and Methods. Some combinations lead to much higher or lower Response values than others.
- This interaction effect emphasizes the importance of carefully selecting the right combination of Dentist and Method to achieve optimal Response for Alloy 2. The differing response patterns among Dentists imply that each may influence the outcome in a unique way.
- The considerable variability in Response indicates that further investigation is needed to understand the underlying factors driving these differences.

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

1. Define null and alternative hypotheses State the null and alternate hypotheses

Null Hypotheses-

H(0a): There is no main effect of Dentist on the hardness of implants.

H(0b): There is no main effect of Method on the hardness of implants.

H (0): There is no interaction effect between Dentist and Method on the hardness of implants.

Alternative Hypotheses-

H(1a): There is a main effect of Dentist on the hardness of implants.

H(1b): There is a main effect of Method on the hardness of implants.

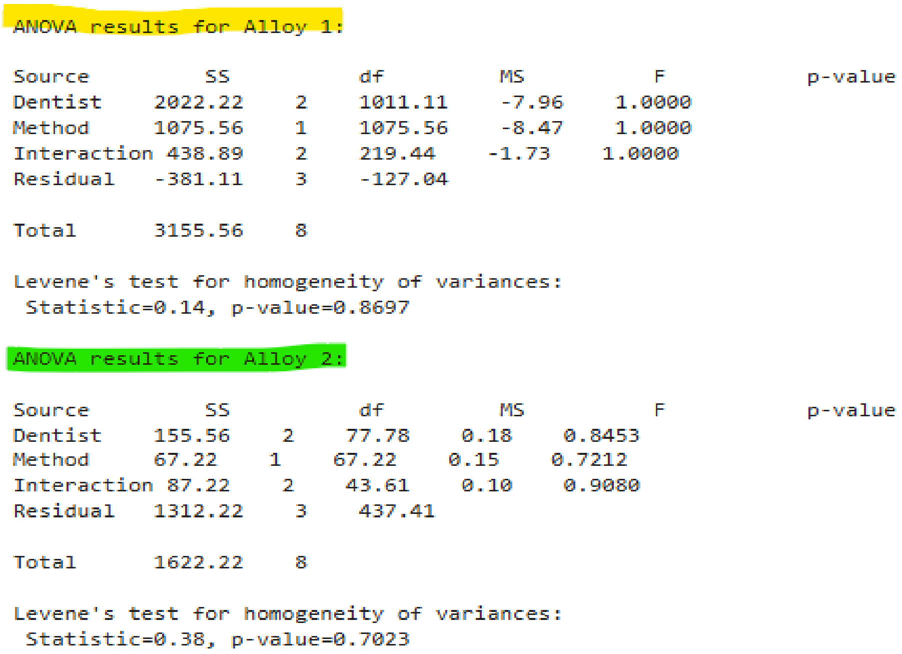
H (1): There is an interaction effect between Dentist and Method on the hardness of implants.

Please refer notebook for calculations. We are following below steps similar to previous solutions-

- Group the data by Dentist and Method to prepare for 2-way ANOVA.
- Calculate Means and Sums of Squares:

- **SST**: Total variability.
- **SSB**: Variability between groups.
- **SSW**: Variability within groups.
- **Degrees of Freedom**: For between groups and within groups.

- **Mean Squares**: Compute for between groups and within groups.
- **F-Statistic**: Ratio of mean squares.
- **P-Value**: Calculate based on the F-distribution



```
ANOVA results for Alloy 1:

Source          SS          df          MS          F           p-value
Dentist     2022.22    2    1011.11     -7.96    1.0000
Method      1075.56    1    1075.56     -8.47    1.0000
Interaction 438.89     2    219.44     -1.73    1.0000
Residual    -381.11    3    -127.04

Total       3155.56    8

Levene's test for homogeneity of variances:
 Statistic=0.14, p-value=0.8697

ANOVA results for Alloy 2:

Source          SS          df          MS          F           p-value
Dentist     155.56    2    77.78     0.18    0.8453
Method      67.22    1    67.22    0.15    0.7212
Interaction 87.22     2    43.61     0.10    0.9080
Residual    1312.22    3    437.41

Total       1622.22    8

Levene's test for homogeneity of variances:
 Statistic=0.38, p-value=0.7023
```

TABLE 9 - P-VALUE FOR 2-WAY ANOVA

Observations and Insights as per table 5 -

- As per calculation, When the p-value is greater than 0.05 in a two-way ANOVA, it means that there is insufficient evidence to reject the null hypothesis. In essence, a p-value greater than 0.05 in a two-way ANOVA indicates that the data does not provide enough evidence to conclude that the factors or their interaction have a significant impact on the response variable.
- For finding, the hardness of implants varies depending on dentists and methods together, we'll review the results of the ANOVA for both Alloy 1 and Alloy 2.
    - Alloy 1:
      Dentist: F=−7., p=1.0000
      Method: F=−8.47 , p=1.0000
      Interaction: F=−1.73 , p=1.0000
      Levene's Test: p=0.8697 (indicates that variances are equal across groups)

    - Alloy 2:
      Dentist: F=0.18 , p=0.8453
      Method: F=0.15 , p=0.7212
      Interaction: F=0.10 , p=0.908
      Levene's Test: p=0.7023 (indicates that variances are equal across groups)

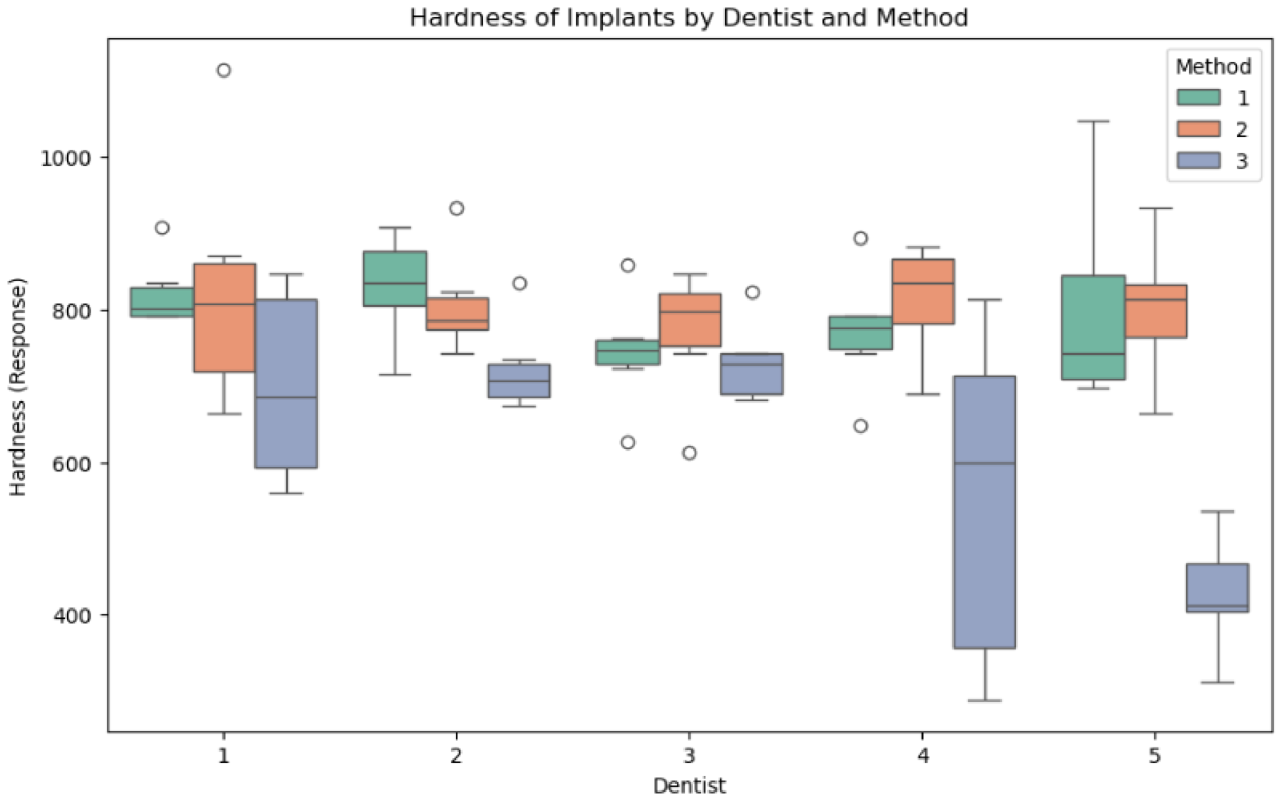


FIGURE 24 – BOXPLOT – HARDNESS W.R.T TO DENTIST AND METHOD

Effect of Dentists:

- **Alloy 1:** The F-value for the dentist is negative and the p-value is 1.0000, which is atypical for an F-test. Normally, F-values should be positive and p-values range between 0 and 1. The negative F-value might indicate a problem with the data or a computational error.
- **Alloy 2:** The F-value is 0.18 with a p-value of 0.8453, suggesting that there is no statistically significant difference in hardness due to the dentist.

Effect of Methods:

- **Alloy 1:** The F-value for the method is negative and the p-value is 1.0000, which is also unusual. This result implies no significant effect of the method on hardness, though the negative F-value raises concerns.
- **Alloy 2:** The F-value is 0.15 with a p-value of 0.7212, indicating that the method does not significantly affect hardness.

Interaction between Dentists and Methods:

- **Alloy 1:** The interaction F-value is negative with a p-value of 1.0000, suggesting no significant interaction effect between dentists and methods, although the negative F-value is unusual.

- **Alloy 2:** The interaction F-value is 0.10 with a p-value of 0.9080, indicating that there is no significant interaction between dentists and methods.

Summary:

For both alloys, the analysis shows no significant effects of dentists, methods, or their interaction on the hardness of the implants. The high p-values confirm that there are no statistically significant differences. However, the unusual negative F-values observed raises concern.