Business Report: Inferential Statistics

Prepared by: Indusudha

Table of Contents

List of Graphs

List of Tables

Context

This report presents a comprehensive statistical analysis of four diverse real-world scenarios using foundational and advanced concepts from data science. Each problem explores unique domains—sports medicine, industrial packaging, materials science, and dental healthcare—providing insight into how data-driven decision-making can improve outcomes.

- **Problem 1** investigates foot injury patterns in football players by analyzing positional data to determine injury probabilities and identify high-risk roles.
- **Problem 2** evaluates the quality and consistency of gunny bags used for cement packaging by modeling their breaking strength with normal distribution techniques.
- **Problem 3** examines the suitability of polished versus unpolished stones for industrial printing based on their hardness, using hypothesis testing and comparison of means.
- **Problem 4** explores the factors affecting the hardness of dental implants, including the method used, the dentist's role, and their interaction, utilizing ANOVA and Tukey HSD testing.

Each section is supported by clearly interpreted statistical outputs, visualizations, and actionable insights to bridge theory with practical business relevance.

Problem 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.

Based on the above data, answer the 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?

	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**

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

Output:

0.6170212765957447

Formula:

P(Injury) = Number of Players with Injury / Total Number of Players

Interpretation:

Based on the dataset, the probability that a randomly selected player suffers an injury is approximately 61.7%. This indicates that injuries are relatively common among players.

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

Output:

0.5234042553191489

Formula:

The combined probability of two categories — Forward or Winger — using:

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

Interpretation:

There is a 52.3% chance that a randomly selected player is either a Forward or a Winger. These two positions together make up the majority of the player base in this dataset, highlighting their key role in the team's offensive lineup.

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

Output:

0.19148936170212766

Formula:

P(Striker and Foot Injury) = Number of Players who are Strikers with Foot Injury / Total Number of Players

Interpretation:

There is a 19.1% chance that a randomly selected player is a striker and has suffered a foot injury. This may highlight a higher injury risk associated with the striker position, potentially due to frequent contact, sudden movements, or high match involvement.

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

Output:

0.3103448275862069

Formula:

To calculate the probability that a randomly chosen injured player is a striker, we use the following formula:

P(Striker | Injured) = (Number of Injured Strikers) / (Total Number of Injured Players)

Interpretation:

There is a 31.0% chance that a randomly selected injured player is a striker. This means that about 31.0% of the injured players in the dataset are strikers. This suggests that strikers, while an important part of the team, make up a relatively smaller proportion of the injured player pool compared to other positions.

Problem 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; (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 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.?

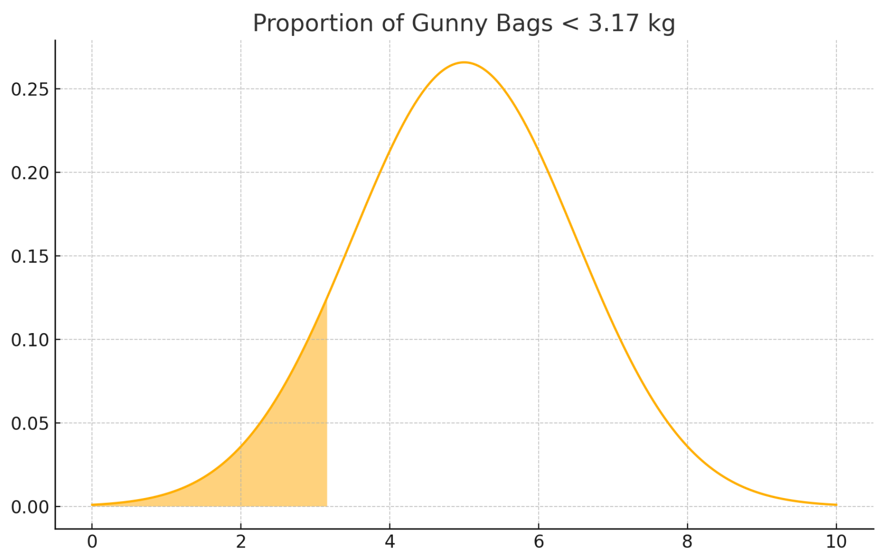
2.4 What proportion of the gunny bags have a breaking strength NOT between 3 and 7.5 kg per sq cm.?

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

Output:

Proportion of bags with breaking strength less than 3.17: 0.11123243744783456

Graph 1:



To answer this question, we need to calculate the proportion of gunny bags that have a breaking strength of less than 3.17 kg per square centimeter. Since the breaking strength follows a normal distribution with a mean of 5 kg per square centimeter and a standard deviation of 1.5 kg per square centimeter, we will use the Z-score formula to standardize the value of 3.17 kg:

Z-score Formula:

$Z = X - \mu / \sigma$

Where:

$X = 3.17$ (the value we want to compare)

$\mu = 5$ kg (mean)

σ=1.5 kg (standard deviation)

Z-score Calculation:

Z=3.17−5/1.5

Once we have the Z-score, we can use the standard normal distribution table or a statistical function to find the cumulative probability (the proportion) of gunny bags with a breaking strength less than 3.17 kg per square centimeter.

The proportion of gunny bags with a breaking strength of less than 3.17 kg per square centimeter is approximately 0.1112 or 11.12%.This means that about 11.12% of the gunny bags have a breaking strength less than 3.17 kg per square centimeter.

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

Output:

Proportion of bags with breaking strength at least 3.6: 0.8247

Graph 2:

We are asked to find the proportion of gunny bags with breaking strength ≥ 3.6 kg/sq. cm

Step 1: Use Z-Score Formula

To find the proportion under a normal distribution curve, we first convert the given value (3.6) to a z-score using the formula:

$z = x - \mu / \sigma$

Where:

x = value of interest (3.6),μ = mean (5),σ = standard deviation (1.5)

$z = 3.6 - 5 / 1.5 = -0.9333$

Step 2: Use standard normal distribution table

$P(Z \geq -0.9333) = 1 - P(Z < -0.9333) = 1 - 0.1753 = 0.8247$

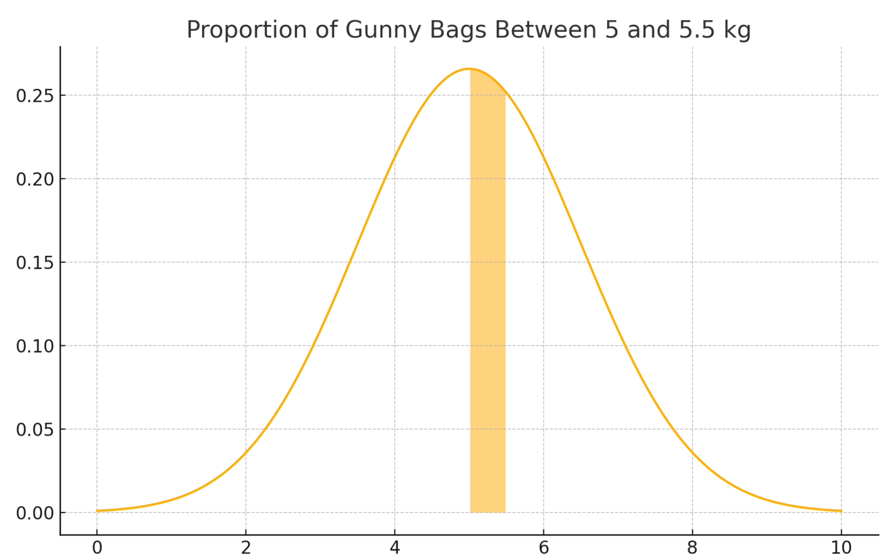
✅ So, 82.47% of the gunny bags have a breaking strength of at least 3.6 kg/cm^2.

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

Output:

Proportion of bags with breaking strength between 5 and 5.5: 0.1306

Graph 3:



Step 1: Convert to Z-scores

For X = 5:

$Z1 = 5 - 5/1.5 = 0$

For X = 5.5:

$Z2 = 5.5 - 5/1.5 = 0.5/1.5 \approx 0.3333$

Step 2:

From the Z-table:

$P(Z < 0.3333) \approx 0.6293$

$P(Z<0) = 0.5$

$P(5 \leq X \leq 5.5) = 0.6293 - 0.5 = 0.1306$

Interpretation:

Approximately 13.06% of the gunny bags have a breaking strength between 5 and 5.5 kg per square centimeter.

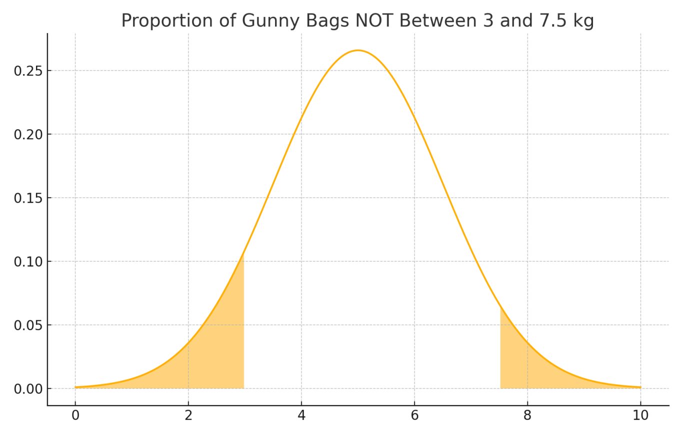
This indicates that a small but notable portion of the bags falls within this strength range, which is slightly above the average (mean = 5). While this strength is not considered weak, it’s also not exceptionally strong.

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

Output:

Proportion of bags with breaking strength NOT between 3 and 7.5: 0.1390

Graph 4:



Formula:

$P(X<3 \text{ or } X>7.5)=1-P(3 \leq X \leq 7.5)$

Interpretation:

Approximately 13.90% of the gunny bags have a breaking strength that falls outside the acceptable range of 3 to 7.5 kg per sq cm, indicating potential risks of underperforming or over engineered packaging.

Problem 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?

Read the Dataset given and display the first 5 rows in data using head():

	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

Describe the Data frame using describe() Function and shape is (75, 2)

	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

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%."

State the null and alternate hypotheses

Null Hypothesis

H_0 (Null): Mean hardness of unpolished stones ≥ 150

H_1 (Alternative): Mean hardness of unpolished stones < 150

Columns are:

['Unpolished ', 'Treated and Polished']

Find Mean and Standard Deviation

Mean: 134.11052653373332
Standard Deviation: 33.0418044136061

Conduct the hypothesis test and compute the p-value

Output:

=== 3.1: One-Sample T-Test (Unpolished vs 150) ===
Mean (Unpolished): 134.11
T-statistic: -4.1646
One-tailed p-value: 0.00004

Write down conclusions from the test results

Output:

Conclusion: Reject the null hypothesis. The unpolished stones likely have hardness < 150.

Test Result

Sample Mean (Unpolished): 134.11

One-tailed p-value: 0.00004

Since the p-value is much less than 0.05, we reject the null hypothesis.

Yes, Zingaro is justified in believing that the unpolished stones may not be suitable for printing, as their average hardness is significantly below 150.

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%.

State the null and alternate hypotheses

Null Hypothesis

H_0 (Null): Mean hardness of polished = unpolished

H_1 (Alternative): Mean hardness of polished ≠ unpolished

Conduct the hypothesis test

Output:
=== 3.2: Two-Sample T-Test (Unpolished vs Polished) ===
Mean (Polished): 147.79
Mean (Unpolished): 134.11
T-statistic: -3.2422
Two-tailed p-value: 0.00159

Write down conclusions from the test results.

Output:

Conclusion: Reject the null hypothesis. The means are significantly different.

Mean (Unpolished): 134.11

Mean (Polished): 147.79

Two-tailed p-value: 0.00159

Since the p-value is less than 0.05, we reject the null hypothesis.The mean hardness of polished and unpolished stones is significantly different.

Problem 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 favor one method above another and may work better in his/her favorite 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?

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 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."

Reading the Dataset and displays the dataframe head, describe, shape,data type,Unique values.

Head of Data:

	Dentist	Method	Alloy	Temp	Response
cell output actions		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

Describe of Data:

	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

Data info:

```
<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
```

Checking for Missing Values :

	0
Dentist	0
Method	0
Alloy	0
Temp	0
Response	0

Observation- No Missing Values, 5 Columns are there-'Dentist',
'Method','Alloy','Temp','Response'.

Change integers to Categorical Variable Conversion.

Unique Values:

Unique values in column 'Dentist': [1 2 3 4 5]
Unique values in column 'Method': [1 2 3]
Unique values in column 'Alloy': [2]
Unique values in column 'Temp': [1500 1600 1700]
Unique values in column 'Response': [813 792 907 835 782 698 665 1115 870 752 620
847 560 585
715 803 858 882 772 743 933 824 673 734 681 627 762 724
613 894 649 690 493 707 289 312 1048 421 483 405 536]

State the null and alternate

Null Hypothesis (H_0):

The mean implant hardness is the same for all dentists (no difference across dentists).

Alternate Hypothesis (H_1):

At least one dentist has a significantly different mean implant hardness.

μ_1 = Mean hardness for Dentist 1

μ_2 = Mean hardness for Dentist 2

μ_3 = Mean hardness for Dentist 3… and so on

Null Hypothesis (H_0): $H_0 : \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$ $H_0 : \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$

(The mean hardness is the same for all dentists)

Alternate Hypothesis (H_1):

Alternate Hypothesis (H_1): H_1 : At least one $\mu_i \neq \mu_j$

(At least one dentist has a different mean hardness)

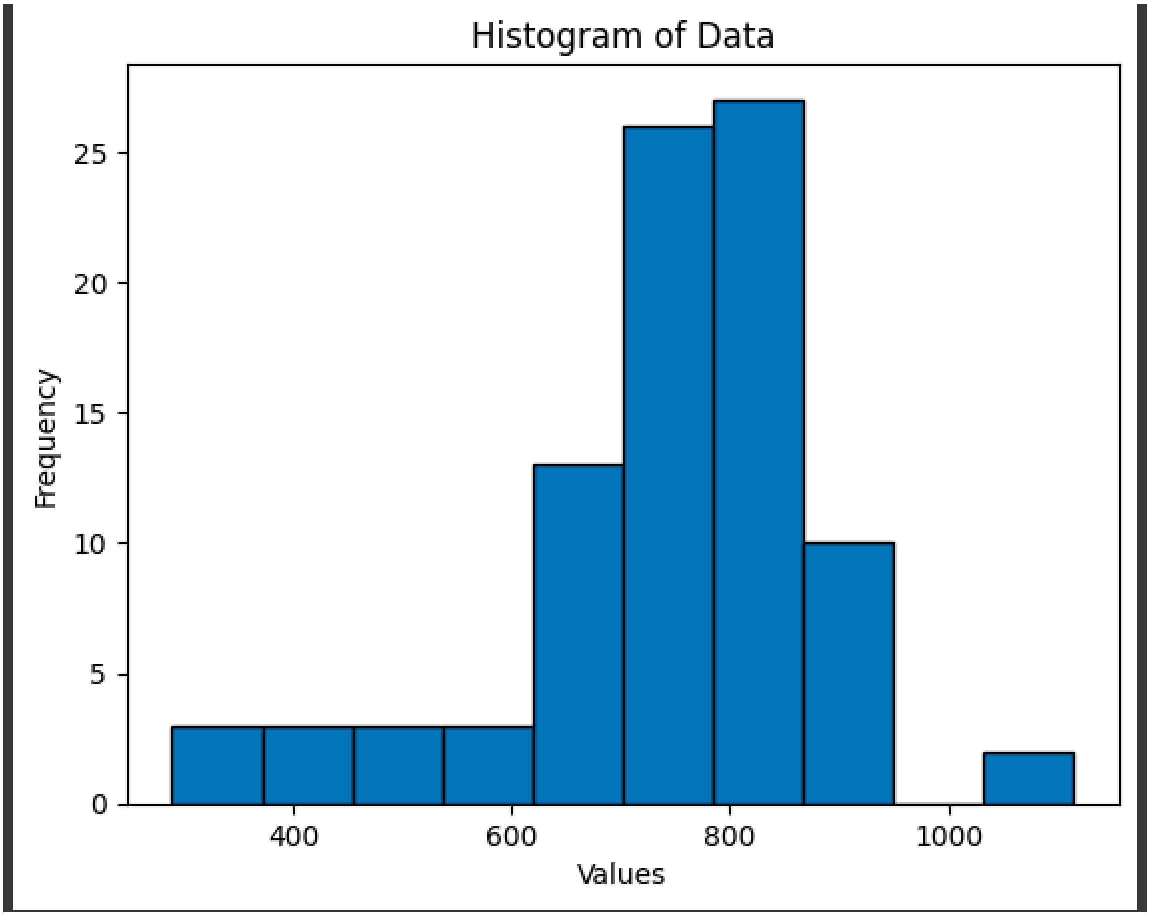
Check the assumptions of the hypothesis test

*Independence: Assumed, as different dentists work independently.

*Normality: Check visually using box plot or histogram.

*Homogeneity of variances: Can be tested using Levene's test

Graph 5-Normality: Check visually using histogram.



Output:

Shapiro-Wilk Test: Statistic = 0.9065922596770956, p-value = 8.079982979541578e-06

Interpretation:

If p-value > 0.05 → normality assumption is satisfied

If p-value < 0.05 → normality assumption is not satisfied (but as per instructions, continue the test)

Since the p-value < 0.05, we reject the null hypothesis of the Shapiro-Wilk test.
This means that the data does not follow a normal distribution — the normality assumption is not satisfied.

However, as per instructions, we can proceed with the next test anyway, even though the data is not normally distributed.

Homogeneity of variances: Levene's Test

Output:

Alloy-1: The p-value is 0.2566
Alloy-2: The p-value is 0.2369

Conclusion:

Homogeneity of variances is likely met the data. This means there is no significant difference in the variability of the 'Response' variable between the different dentist groups for both Alloy-1 and Alloy-2.

We can proceed with statistical tests that assume equal variances, such as ANOVA, without major concerns about violating this assumption.

Conduct the hypothesis test and compute the p-value

Output:

Alloy-1 ANOVA results:
F-value: 1.9771119908770842
P-value: 0.11656712140267628

Alloy-2 ANOVA results:
F-value: 0.5248351000282961
P-value: 0.7180309510793431

Write down conclusions from the test results

For both Alloy-1 and Alloy-2, the p-values are greater than the significance level of 0.05. This means we fail to reject the null hypothesis for both alloys.

Therefore, we conclude that there is no statistically significant difference in the mean hardness values between the different dentist groups for either Alloy-1 or Alloy-2.

In simpler terms, the hardness of implants does not appear to be significantly influenced by the dentist performing the assessment for either type of alloy.

Further Interpretation:

Alloy-1: While the p-value for Alloy-1 is relatively closer to the significance level (0.05) compared to Alloy-2, it is still not considered statistically significant. There might be a slight trend towards differences in hardness for Alloy-1 between dentists, but the evidence is not strong enough to conclude a definite effect. It might be worth investigating further with a larger sample size.

Alloy-2: The high p-value for Alloy-2 strongly suggests that the hardness measurements are consistent across different dentists for this type of alloy.

Overall:

The results indicate that the choice of dentist does not appear to have a statistically significant impact on the measured hardness of either Alloy-1 or Alloy-2. This suggests a level of consistency and agreement among the dentists in their hardness assessment techniques.

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

Technically, we don't need to run post-hoc tests, because we already concluded that there's no significant difference.

However, we can still run a Tukey's HSD test to see pairwise comparisons.

Alloy 1

Output:

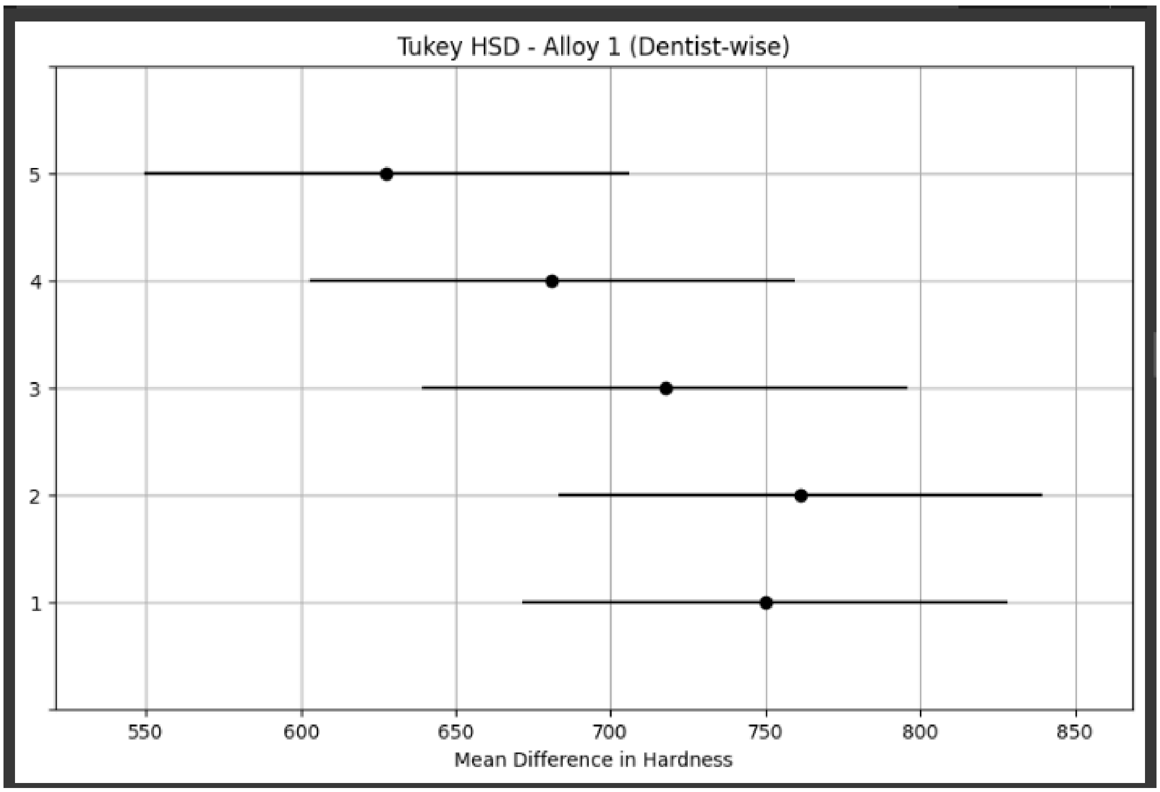
```
Multiple Comparison of Means - Tukey HSD, FWER=0.05
=============================================================
group1 group2  meandiff p-adj    lower    upper   reject
-------------------------------------------------------------
     1      2   11.3333 0.9996 -145.0423  167.709  False
     1      3  -32.3333 0.9757  -188.709 124.0423  False
     1      4  -68.7778 0.7189 -225.1535  87.5979  False
     1      5 -122.2222 0.1889 -278.5979  34.1535  False
     2      3  -43.6667 0.9298 -200.0423  112.709  False
     2      4  -80.1111 0.5916 -236.4868  76.2646  False
     2      5 -133.5556 0.1258 -289.9312  22.8201  False
     3      4  -36.4444 0.9626 -192.8201 119.9312  False
     3      5  -89.8889 0.4805 -246.2646  66.4868  False
     4      5  -53.4444 0.8643 -209.8201 102.9312  False
-------------------------------------------------------------
```

reject = True ➜ means significant difference between those two groups

reject = False ➜ no significant difference

But Reject= All False Means no Significant difference

Graph 6:



Alloy 2

Output:

Multiple Comparison of Means - Tukey HSD, FWER=0.05

```
=====================================================
group1 group2 meandiff p-adj    lower     upper   reject
-----------------------------------------------------
     1      2  -4.1111    1.0 -225.5687 217.3465  False
     1      3 -36.5556 0.9895 -258.0131  184.902  False
```

1	4	-70.0	0.8941	-291.4576	151.4576	False
1	5	-90.1111	0.7724	-311.5687	131.3465	False
2	3	-32.4444	0.9933	-253.902	189.0131	False
2	4	-65.8889	0.9132	-287.3465	155.5687	False
2	5	-86.0	0.8008	-307.4576	135.4576	False
3	4	-33.4444	0.9925	-254.902	188.0131	False
3	5	-53.5556	0.9574	-275.0131	167.902	False
4	5	-20.1111	0.999	-241.5687	201.3465	False

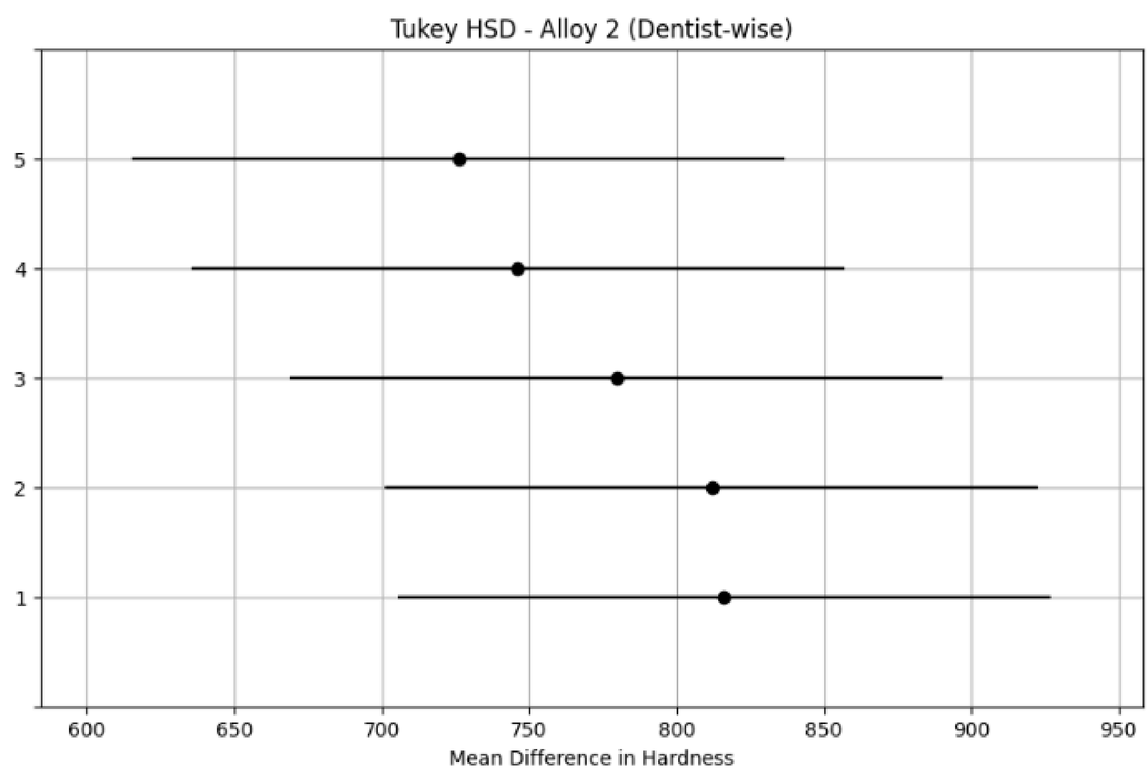
----

reject = True ➜ means significant difference between those two groups

reject = False ➜ no significant difference

But Reject= All False Means no Significant difference

Graph 7:



Based on the Tukey's HSD test, none of the dentist pairs show statistically significant differences in implant hardness for either Alloy-1 or Alloy-2.This supports the ANOVA result, suggesting consistent assessment across dentists.

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."

State the null and alternate hypotheses

Null Hypothesis (H_0): Mean hardness is the same across all methods.

Alternative Hypothesis (H_1): Mean hardness is different for at least one method.

Method 1, Method 2, Method 3 and comparing the mean implant hardness (denoted as μ) for each method.

Null Hypothesis (H_0): There is no difference in mean implant hardness across all methods:

$H0:\mu 1=\mu 2=\mu 3$ H 0 :μ 1 =μ 2 =μ 3

Alternative Hypothesis (H_1): There is a difference in mean implant hardness for at least one method:

$H1$: At least one $\mu i \neq \mu j$ for $i \neq j$

Check the assumptions of the hypothesis test

Output:

Shapiro-Wilk test for Method 1: p-value = 0.758239031460697
Shapiro-Wilk test for Method 2: p-value = 0.001051113443000027
Shapiro-Wilk test for Method 3: p-value = 0.10259048044955665

Conclusion (Shapiro-Wilk Normality Test – Alloy-wise Methods):

The Shapiro-Wilk test is used to assess whether the data is normally distributed within each method group.

Method 1:p-value = 0.7582

Since the p-value is greater than 0.05, we fail to reject the null hypothesis.

Conclusion: The data for **Method 1 is normally distributed.**

Method 2:p-value = 0.0011

Since the p-value is less than 0.05, we reject the null hypothesis.

Conclusion: The data for **Method 2 is not normally distributed.**

Method 3: p-value = 0.1026

Since the p-value is greater than 0.05, we fail to reject the null hypothesis.

Conclusion: The data for **Method 3 is normally distributed**

Interpretation:

Normality assumption is violated for Method 2.

Since not all groups satisfy the normality assumption, caution is advised when using traditional ANOVA.

Output:

🔍 Levene's Test for Equality of Variances (Method-wise):

➡️ Alloy1: Statistic = 6.5214, p-value = 0.003416
➡️ Alloy2: Statistic = 3.3497, p-value = 0.044693

Conclusion

Conclusion (Levene's Test for Methods – Alloy-wise)

Alloy1

Statistic = 6.5214

p-value = 0.0034

Since the p-value is less than 0.05, we reject the null hypothesis.This indicates that the variances in implant hardness differ significantly across methods for Alloy1.

Alloy2

Statistic = 3.3497

p-value = 0.0447

Since the p-value is also less than 0.05, we reject the null hypothesis for Alloy2 as well.This means the variances in implant hardness also differ significantly across methods for Alloy2.

Conduct the hypothesis test and compute the p-value

One-Way ANOVA for Method-Wise Comparison (for Both Alloys)

Output:

Alloy-1 ANOVA results (Method-wise):
F-value: 6.263326635486233
P-value: 0.004163412167505543

Alloy-2 ANOVA results (Method-wise):
F-value: 16.41079988438482
P-value: 5.415871051443187e-06

Write down conclusions from the test results

Based on the one-way ANOVA results, there is strong evidence to suggest that the mean implant hardness significantly differs across the different methods for both Alloy-1 and Alloy-2.

For Alloy-1, the p-value is 0.0042, which is less than the significance level of 0.05, indicating a statistically significant difference in hardness across methods.

Similarly, for Alloy-2, the p-value is extremely small (5.42×10^{-6}), reinforcing a highly significant difference in hardness based on the method used.

Therefore, we reject the null hypothesis for both alloys and conclude that the method of implantation has a significant impact on the hardness of dental implants.**

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

Alloy-1 method-wise

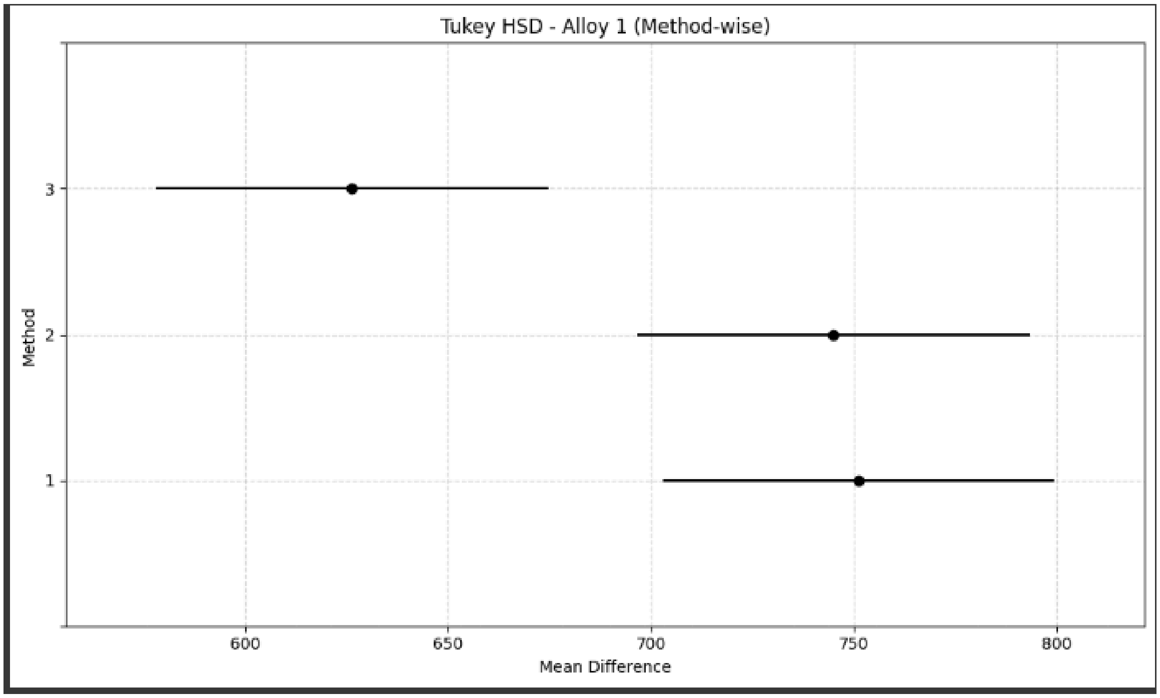
Run Tukey HSD test

Output:

```
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

Graph 8:



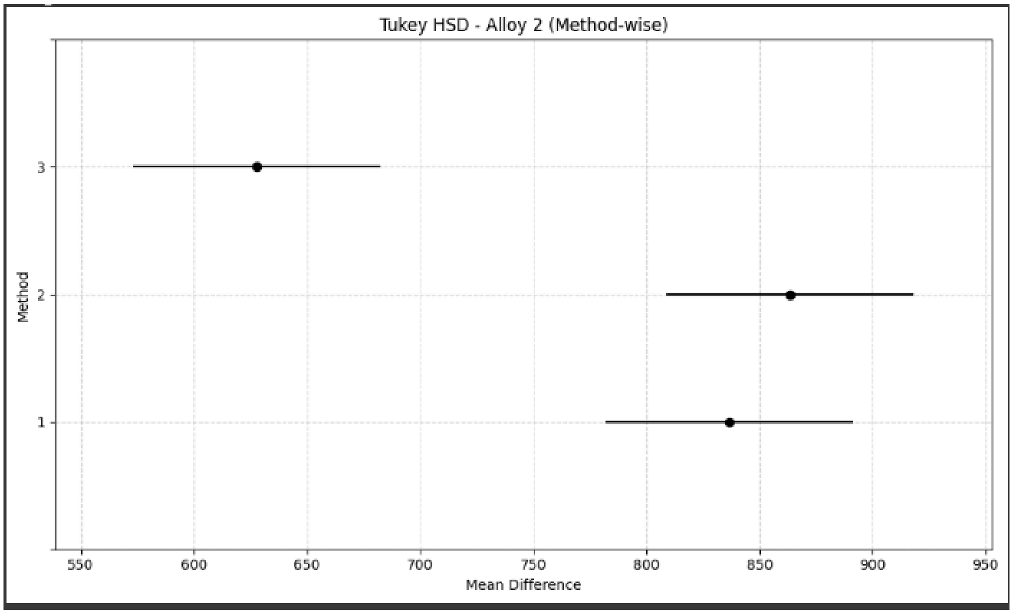
Alloy-2 method-wise

Run Tukey HSD test

Output:

```
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
------------------------------------------------------
```

Graph 9:



Conclusion:

Alloy-1

The Tukey HSD test reveals statistically significant differences in implant hardness between the following method pairs:

Method 1 vs Method 3

Method 2 vs Method 3

This suggests that Method 3 produces significantly different hardness results compared to Methods 1 and 2. However, no significant difference was found between Method 1 and Method 2.

Alloy-2

The Tukey HSD test for Alloy-2 also shows significant differences between multiple method pairs, indicating that method of assessment has a substantial impact on implant hardness for Alloy-2.

The results highlight inconsistencies in measurement across methods, especially involving Method 3.

Interpretation:

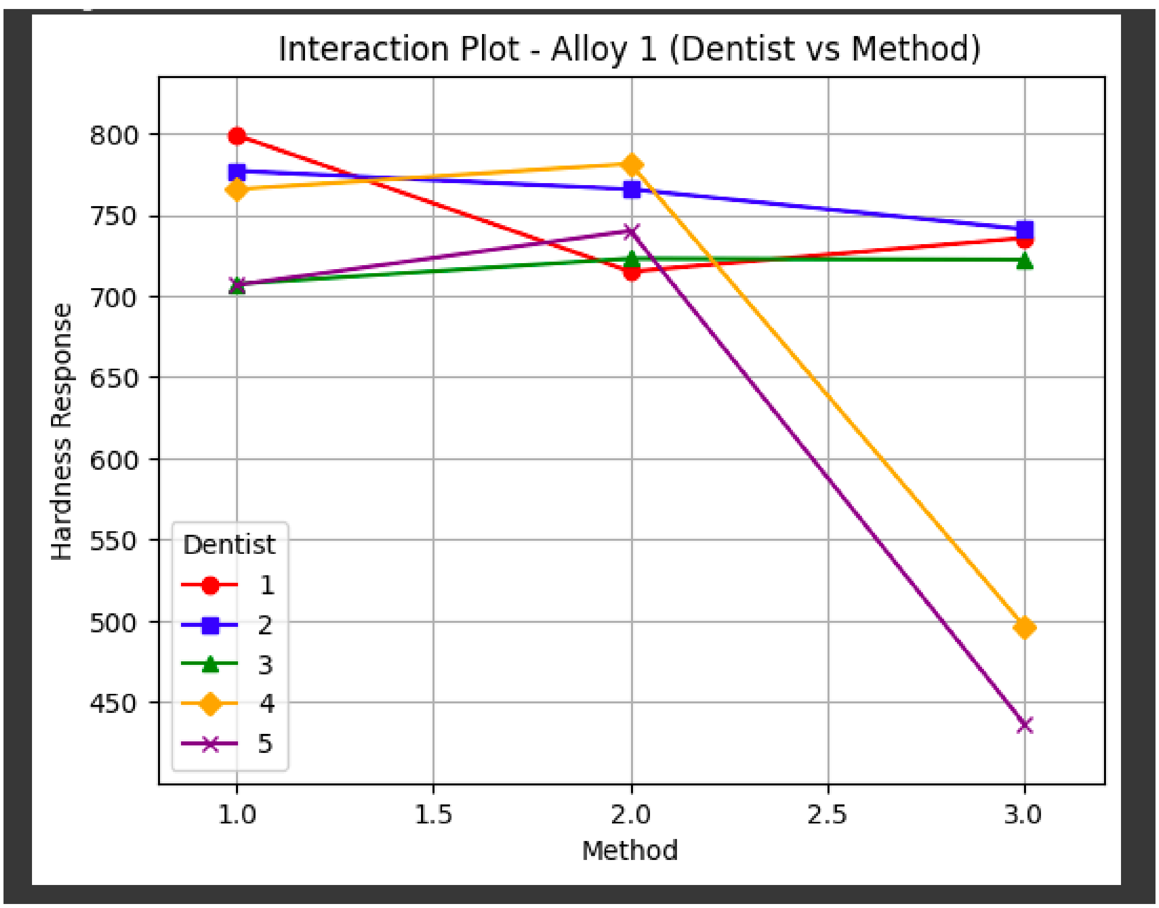
In both alloys, Method 3 appears to be the outlier, showing significantly different results compared to other methods. This could indicate differences in procedure, equipment, or evaluator subjectivity associated with Method 3. Further investigation may be necessary to understand why Method 3 differs consistently.

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."

Create Interaction Plot

Alloy 1 Interaction Plot: Graph 10



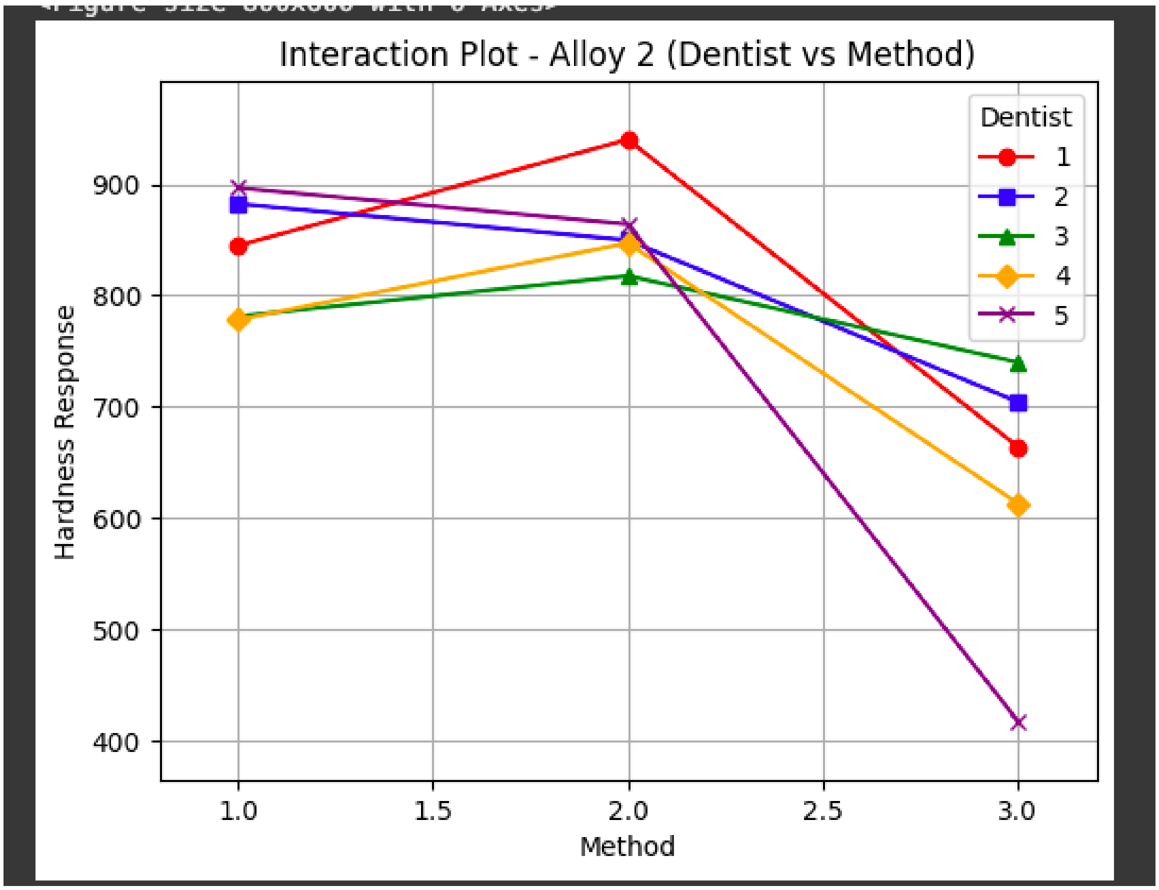
Inferences from the plot

Interaction Effect Alloy 1

The interaction plot displays non-parallel lines among the different dentists across methods, particularly a sharp divergence for Dentist 4 and Dentist 5 at Method 3. This suggests the presence of a potential interaction effect between the Dentist and Method factors.

In other words, the effect of the method on the hardness response is not consistent across all dentists. Such patterns indicate that the dentists may differ in their evaluation of implant hardness depending on the method used, pointing to a moderating role of the dentist in the relationship between method and measured hardness.

Alloy 2 Interaction Plot: Graph 11



Interaction Effect – Alloy 2

The interaction plot for Alloy-2 shows pronounced non-parallel trends, especially at Method 3, where dentists show a significant divergence in hardness responses. For instance, Dentist

5 reports a much lower hardness than others at Method 3, while Dentist 1 exhibits a strong peak at Method 2 followed by a steep drop.

These variations indicate a strong interaction effect between Dentist and Method, suggesting that the impact of the method on implant hardness is highly dependent on the dentist evaluating it.

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."

State the null and alternate hypotheses

Hypotheses (Two-Way ANOVA with Interaction)

Null Hypotheses (H_0):

H_{01}: There is no difference in implant hardness across different methods.

H_{02}: There is no difference in implant hardness across different dentists.

H_{03}: There is no interaction effect between dentist and method on implant hardness.

Alternative Hypotheses (H_1):

H_{11}: At least one method differs in mean hardness.

H_{12}: At least one dentist differs in mean hardness.

H_{13}: There is an interaction effect between dentist and method.

Fit the Two-Way ANOVA Model

Check the assumptions of the hypothesis test

Normality: Use Shapiro-Wilk test for residuals.

Homogeneity of Variance: Use Levene’s test. (Note: Even if assumptions fail, proceed as instructed.)

1. Shapiro-Wilk Test (Normality of Residuals)

Output:

Alloy-1 Shapiro-Wilk Test (Normality):
Statistic = 0.9565, p-value = 0.0902
Residuals are approximately normal.

Normality Assumption Check (Shapiro-Wilk Test) – Alloy-1

The Shapiro-Wilk test for the residuals of Alloy-1 yielded a test statistic of 0.9565 and a p-value of 0.0902.

Since the **p-value is greater than the significance level of 0.05**, we fail to reject the null hypothesis of the test. This indicates that the residuals are approximately **normally distributed**, satisfying the assumption of normality for ANOVA.

Output:

Alloy-2 Shapiro-Wilk Test (Normality):
Statistic = 0.9571, p-value = 0.0946
Residuals are approximately normal.

Normality Assumption Check (Shapiro-Wilk Test) – Alloy-2

The Shapiro-Wilk test for the residuals of Alloy-2 produced a test statistic of 0.9571 and a p-value of 0.0946.

As the **p-value exceeds the standard significance level of 0.05**, we fail to reject the null hypothesis of normality.

This suggests that the residuals are approximately **normally distributed**, thereby satisfying the assumption of normality required for conducting ANOVA.

2.Levene’s Test (Homogeneity of Variance)

Alloy 1-Levene's Test

Output: Alloy-1 Levene’s Test (Equal Variance):
Statistic = 1.2189, p-value = 0.3128
Variances are equal.

Alloy 2-Levene's Test

Output:

Alloy-2 Levene’s Test (Equal Variance):
Statistic = 0.6710, p-value = 0.7832
Variances are equal.

Levene’s Test Conclusion and Interpretation (Equal Variance Assumption):

Alloy-1:

The Levene’s Test produced a statistic of 1.2189 with a p-value of 0.3128, which is greater than the common significance level of 0.05.

✅ Conclusion: We fail to reject the null hypothesis of equal variances.

➡️ Interpretation: This indicates that the assumption of homogeneity of variances holds true for Alloy-1, meaning that the variance in implant hardness is consistent across the different dentist-method combinations.

Alloy-2:

The Levene’s Test resulted in a statistic of 0.6710 and a p-value of 0.7832, again greater than 0.05.

✅ Conclusion: We fail to reject the null hypothesis of equal variances.

➡️ Interpretation: For Alloy-2 as well, the equal variance assumption is met, suggesting uniform variability in hardness measurements across groups.

Conduct the hypothesis test and compute the p-value

Two-Way ANOVA with Interaction

Output:

Two-way ANOVA with Interaction - Alloy 1

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

Output:

Two-way ANOVA with Interaction - Alloy 2

```
                     sum_sq    df         F    PR(>F)
C(Method)           499640.400000   2.0  19.461218  0.000004
C(Dentist)          56797.911111   4.0   1.106152  0.371833
C(Method):C(Dentist)  197459.822222   8.0   1.922787  0.093234
Residual            385104.666667  30.0       NaN       NaN
```

Write down conclusions from the test results

Conclusion (Alloy 1):

For Alloy 1, both the method used, the dentist, and their interaction significantly affect the hardness of dental implants. The significant interaction implies that the effect of method on hardness depends on which dentist is using it — there is a combined influence.

Conclusion (Alloy 2):

For Alloy 2, only the method significantly affects implant hardness. The dentist and the interaction between method and dentist do not have a significant impact. This suggests that no matter who performs the procedure, the method alone influences the hardness for Alloy 2.

Interaction Effects

In Alloy 1, the significant interaction indicates:

The effect of a method changes depending on the dentist using it.

In Alloy 2, since the interaction is not significant, it means:

The effect of the method is consistent regardless of the dentist.

Identify which dentists and methods combinations are different, and which interaction levels are different.

Run Tukey's HSD Post-hoc Test

After finding a significant interaction in Alloy-1-Tukey's HSD for the interaction of Dentist × Method.

Output:

Significant differences:

	group1	group2	meandiff	p-adj	lower	upper	reject
0	1_1	4_3	-302.6667	0.0070	-551.4950	-53.8383	True
1	1_1	5_3	-362.6667	0.0007	-611.4950	-113.8383	True
2	1_2	5_3	-278.6667	0.0173	-527.4950	-29.8383	True
3	1_3	5_3	-299.3333	0.0079	-548.1617	-50.5050	True
4	2_1	4_3	-280.6667	0.0160	-529.4950	-31.8383	True
5	2_1	5_3	-340.6667	0.0016	-589.4950	-91.8383	True
6	2_2	4_3	-269.3333	0.0243	-518.1617	-20.5050	True
7	2_2	5_3	-329.3333	0.0025	-578.1617	-80.5050	True
8	2_3	5_3	-304.6667	0.0065	-553.4950	-55.8383	True
9	3_1	5_3	-271.0000	0.0229	-519.8283	-22.1717	True
10	3_2	5_3	-286.6667	0.0128	-535.4950	-37.8383	True
11	3_3	5_3	-286.0000	0.0131	-534.8283	-37.1717	True
12	4_1	4_3	-269.3333	0.0243	-518.1617	-20.5050	True
13	4_1	5_3	-329.3333	0.0025	-578.1617	-80.5050	True
14	4_2	4_3	-285.0000	0.0137	-533.8283	-36.1717	True
15	4_2	5_3	-345.0000	0.0013	-593.8283	-96.1717	True
16	5_1	5_3	-270.3333	0.0234	-519.1617	-21.5050	True
17	5_2	5_3	-303.6667	0.0067	-552.4950	-54.8383	True

After finding a significant interaction in Alloy-2-Tukey's HSD for the interaction of Dentist × Method.

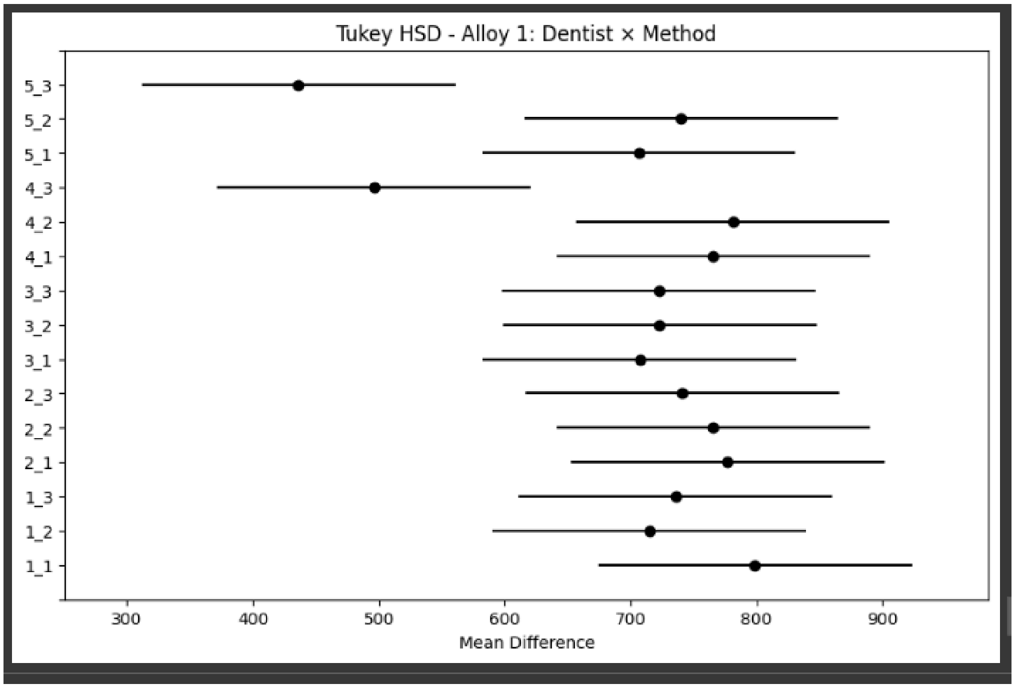
Output:

Significant differences:

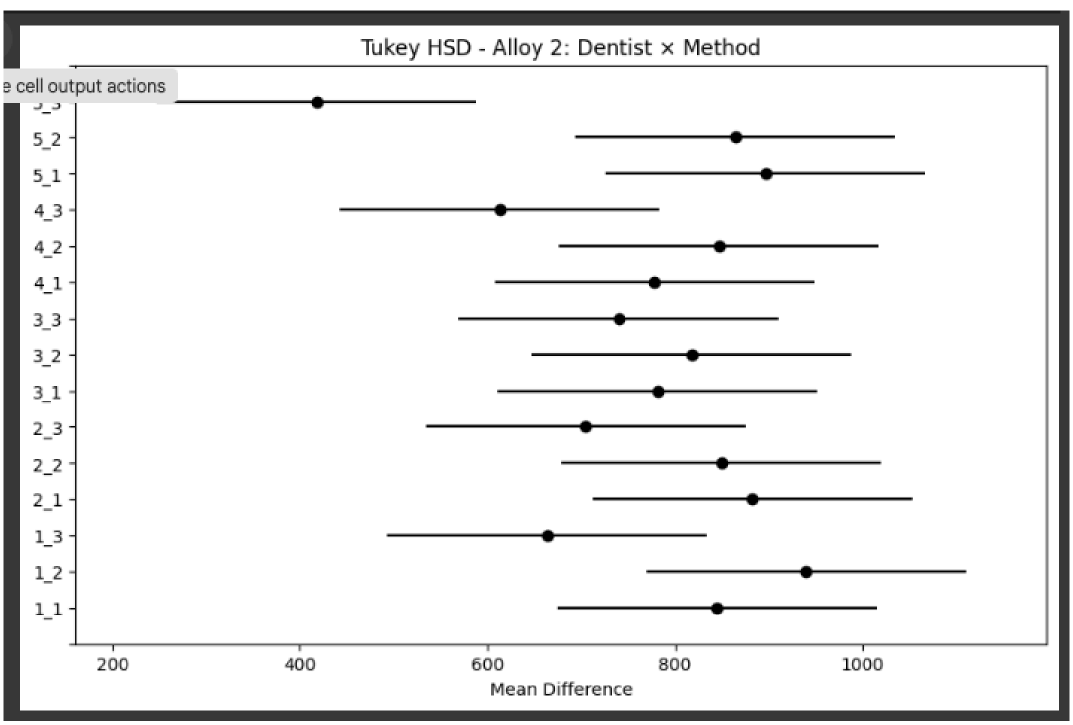
	group1	group2	meandiff	p-adj	lower	upper	reject
0	1_1	5_3	-427.0000	0.0049	-767.8958	-86.1042	True
1	1_2	5_3	-522.3333	0.0003	-863.2292	-181.4375	True
2	2_1	5_3	-464.6667	0.0017	-805.5625	-123.7708	True
3	2_2	5_3	-432.0000	0.0043	-772.8958	-91.1042	True
4	3_1	5_3	-363.6667	0.0279	-704.5625	-22.7708	True
5	3_2	5_3	-400.0000	0.0105	-740.8958	-59.1042	True
6	4_1	5_3	-360.6667	0.0302	-701.5625	-19.7708	True
7	4_2	5_3	-429.3333	0.0046	-770.2292	-88.4375	True
8	5_1	5_3	-479.0000	0.0011	-819.8958	-138.1042	True
9	5_2	5_3	-446.3333	0.0028	-787.2292	-105.4375	True

Visual Interpretation

Graph 12:



Graph 13:



Alloy 1: Result of Tukey Test

There are significant differences between certain group pairs — meaning the results they produced are not just due to chance. Specifically, for Alloy 1:

Dentist 1 using Method 1 had significantly lower results compared to Dentist 4 using Method 3 and Dentist 5 using Method 3.

Similar patterns are seen across many other group comparisons where Method 3 (especially from Dentist 5) consistently shows higher values.

Dentist 5 with Method 3 appears in almost every significant comparison, meaning their results are consistently higher than others.

Alloy 2: Result of Tukey Test

There are 10 pairs of Dentist_Method combinations that show statistically significant differences.

In every single one, the group "5_3" (Dentist 5 using Method 3) has significantly higher results than the others.

Final Conclusion:

Dentist 5 using Method 3 is the best-performing combination for both Alloy 1 and Alloy 2.

Recommendation:

If you're looking for the most effective and reliable choice, choose

👉 **Method 3 with Dentist 5 — it's clearly the top performer for both alloys.**