Business Report

Inferential Statistics for Business Insights: Probability, Distribution & Hypothesis Testing

Prepared by

Arkaprava Mazumder

Jun 15, 2025

Table of Contents

PROBLEM 1: PROBABILITY USING A CONTINGENCY TABLE 4

Business Context **4**

Data Summary **4**

Probability Analysis **4**

1.1 What is the probability that a randomly chosen player would suffer an injury? 4
1.2 What is the probability that a player is a Forward or a Winger? 5
1.3 What is the probability that a randomly chosen player plays in a Striker position *and* has a foot injury? 5
1.4 What is the probability that a randomly chosen injured player is a Striker? 5

PROBLEM 2: BREAKING STRENGTH OF GUNNY BAGS – NORMAL DISTRIBUTION ANALYSIS 6

Business Context **6**

Statistical Analysis **6**

2.1: What proportion of the gunny bags have a breaking strength of less than 3.17 kg per sq. cm? 6
2.2: What proportion of the gunny bags have a breaking strength of at least 3.6 kg per sq cm.? 7
2.3: What proportion of the gunny bags have a breaking strength between 5 and 5.5 kg per sq. cm.? 8
2.4: What proportion of the gunny bags have a breaking strength NOT between 3 and 7.5 kg per sq. cm.? 9

PROBLEM 3: COMPARATIVE ANALYSIS OF STONE HARDNESS – SUITABILITY AND TREATMENT EFFECT 11

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? **11**

Business Context 11
Hypotheses 11
Statistical Analysis 11
Test Results: 12
Conclusion 12

3.2 Is the mean hardness of the polished and unpolished stones the same? **13**

Business Context 13
Hypotheses 13
Statistical Analysis 13
Test Results 13
Conclusion 14

PROBLEM 4: MULTI-FACTOR ANALYSIS OF DENTAL IMPLANT HARDNESS 15

4.1 How does the hardness of implants vary depending on dentists? **15**

Problem Statement 15
Hypotheses 15
Assumptions Checks 15

One-Way ANOVA Results 17

4.2 How does the hardness of implants vary depending on methods? 19
Problem Statement 19
Hypotheses 19
Assumption Checks 20
One Way ANOVA Result 21
Post Hoc Testing Interpretation (Tukey HSD Results) 22

4.3 What is the interaction effect between the dentist and method on the hardness of dental implants for each type of alloy? 24
Objective 24
Interaction Plot Results 24
Conclusion 26

4.4 How does the hardness of implants vary depending on dentists and methods together? 27
Problem Statement 27
Hypotheses 27
Assumption Checks 28
ANOVA Results 29
Post Hoc Analysis (Tukey HSD) Results 31

Problem 1: Probability Using a Contingency Table

Business Context

A physiotherapist working with a male football team wants to explore if there's any link between foot injuries and the positions players play. The insights from this analysis can help with injury prevention strategies tailored by position.

Data Summary

The contingency table below shows the injury status of players across different positions:

	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

Probability Analysis

We compute key probabilities based on the data to understand injury trends:

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

Ans: This is a **marginal probability**.

- **Formula**:

$$P(Injury) = \frac{\text{Injured Players}}{\text{Total Players}} = \frac{145}{235}$$

- **Answer**:

$$P(Injury) = 0.617 \text{ } or \text{ } 61.7\%$$

1.2 What is the probability that a player is a Forward or a Winger?

Ans: This is a **marginal probability**.

- **Formula**:

$$P(Forward\ or\ Winger) = \frac{\text{Number of Forwards + Number of Wingers}}{\text{Total players}} = \frac{94 + 29}{235}$$

- **Answer**:

$$P(Forward\ or\ Winger) = \frac{123}{235} = 0.523\ or\ 52.3\%$$

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

Ans: This is a **joint probability**.

- **Formula**:

$$P(Striker\ and\ Injured) = \frac{\text{Number of injured strikers}}{\text{Total players}} = \frac{45}{235}$$

- **Answer**:

$$P(Striker\ and\ Injured) = 0.191\ or\ 19.1\%$$

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

Ans: This is a **conditional probability**.

- **Formula**:

$$P(\,Striker \mid Injured\,) = \frac{\text{Injured strikers}}{\text{Total injured players}} = \frac{45}{145}$$

- **Answer**:

$$P(\,Striker \mid Injured\,) = 0.310\ or\ 31.0\%$$

Problem 2: Breaking Strength of Gunny Bags – Normal Distribution Analysis

Business Context

In the cement industry, **packaging strength** plays a crucial role in reducing wastage and pilferage during transportation and handling. The quality control team has collected data on the **breaking strength of gunny bags** used to package cement.

- It has been determined that the **breaking strength is normally distributed**.
- Mean (μ) = **5 kg/cm²**
- Standard Deviation (σ) = **1.5 kg/cm²**

Statistical Analysis

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

Ans: We can find out the Z score using,

$$Z = \frac{X - \mu}{\sigma} = \frac{3.17 - 5}{1.5} = -1.22$$

Z-score of -1.22 tells us that 3.17 kg/cm² is 1.22 standard deviations below the average breaking strength of the gunny bags.

Using Python's ***scipy.stats.norm.cdf()*** to calculate the tail probability:

Using the standard normal distribution, the probability of a bag having a breaking strength **less than 3.17 kg per sq. cm** is:

$$P(X < 3.17) = 0.0112\ or\ 11.12\%$$

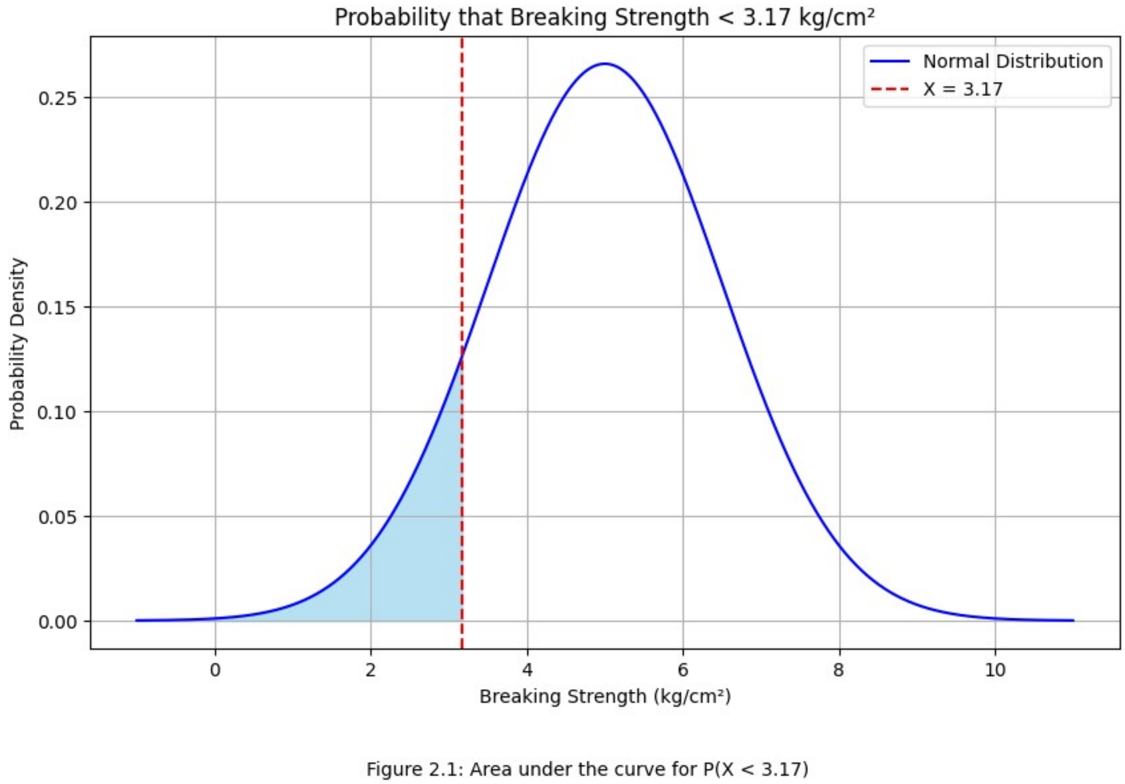


Figure 2.1: The shaded region under the normal distribution curve to the left of X = 3.17 represents the probability P(X < 3.17)

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

Ans: We can find out the Z score using,

$$Z = \frac{X - \mu}{\sigma} = \frac{3.6 - 5}{1.5} = -0.933$$

A Z-score of -0.93 indicates that 3.6 kg/cm² is 0.93 standard deviations below the mean.

Using Python's ***scipy.stats.norm.cdf()*** to calculate the tail probability:

Using the standard normal distribution, the probability of a bag having a breaking strength **at least 3.6 kg per sq. cm** is:

$$P(X \geq 3.6) = 1 - P(X < 3.6) = 0.8243 \; or \; 82.4\%$$

(This result was computed using Python and can be verified in the attached '.ipynb' file.)

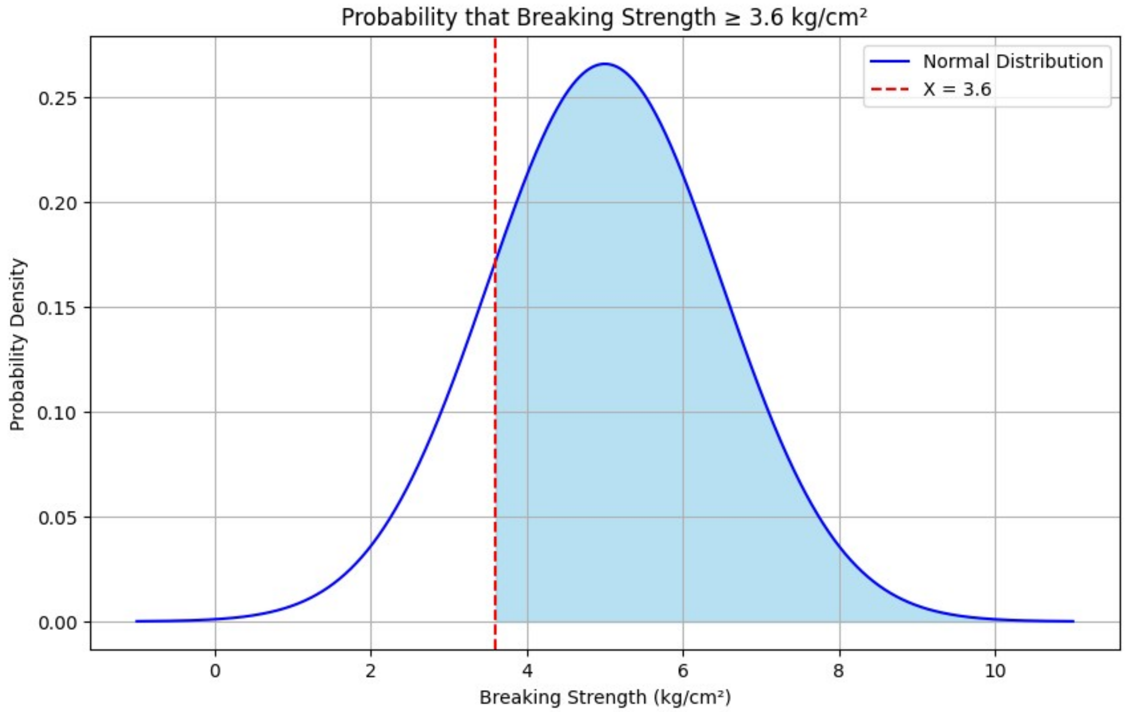


Figure 2.2: Area under the curve for $P(X \geq 3.6)$

The shaded area under the normal distribution curve to the right of X = 3.6 represents the probability $P(X \geq 3.6)$

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

Ans: We want to calculate:

$$P(5 \leq X \leq 5.5) = P(X \leq 5.5) - P(X \leq 5)$$

We can find out the Z score for upper and lower bound,

$$Z1 = \frac{5-5}{1.5} = 0.0$$

$$Z2 = \frac{5.5-5}{1.5} = 0.333$$

Using Python's ***scipy.stats.norm.cdf()*** to calculate the probability:

Using the standard normal distribution, the probability of a bag having a breaking strength **between 5 and 5.5 hg per sq. cm** is:

$$P(5 \leq X \leq 5.5) = 0.1306 \; or \; 13.06\%$$

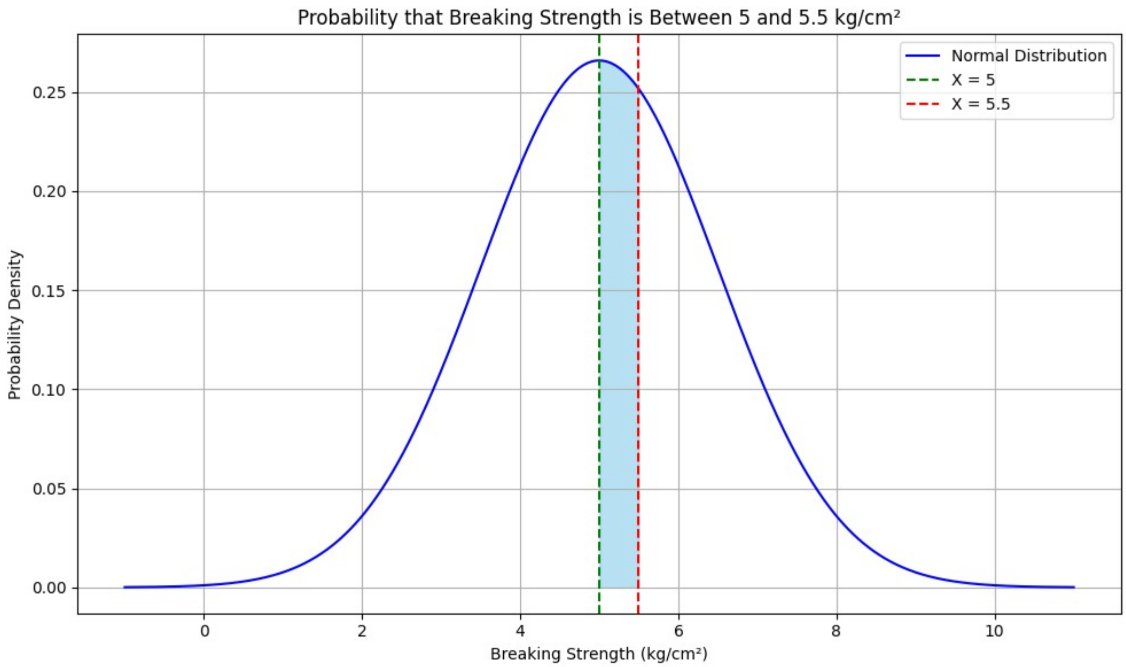


Figure 2.3: Area under the curve for P(5 ≤ X ≤ 5.5)

The shaded area between X = 5 and X = 5.5 under the normal distribution curve represents the probability P(5 ≤ X ≤ 5.5), indicating that approximately **13.08%** of the gunny bags fall within this strength range.

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

Ans:

Using the Complement Rule we can get,

$$P(X \; not \; between \; 3 \; and \; 7.5) = 1 - P(3 \leq X \leq 7.5)$$

For finding the Z score values,

$$Z1 = \frac{3 - 5}{1.5} = -1.33$$

$$Z2 = \frac{7.5 - 5}{1.5} = 1.66$$

Using Python's ***scipy.stats.norm.cdf()*** to calculate the probability:

$$P(X\ not\ between\ 3\ and\ 7.5) = 0.1390\ or\ 13.9\%$$

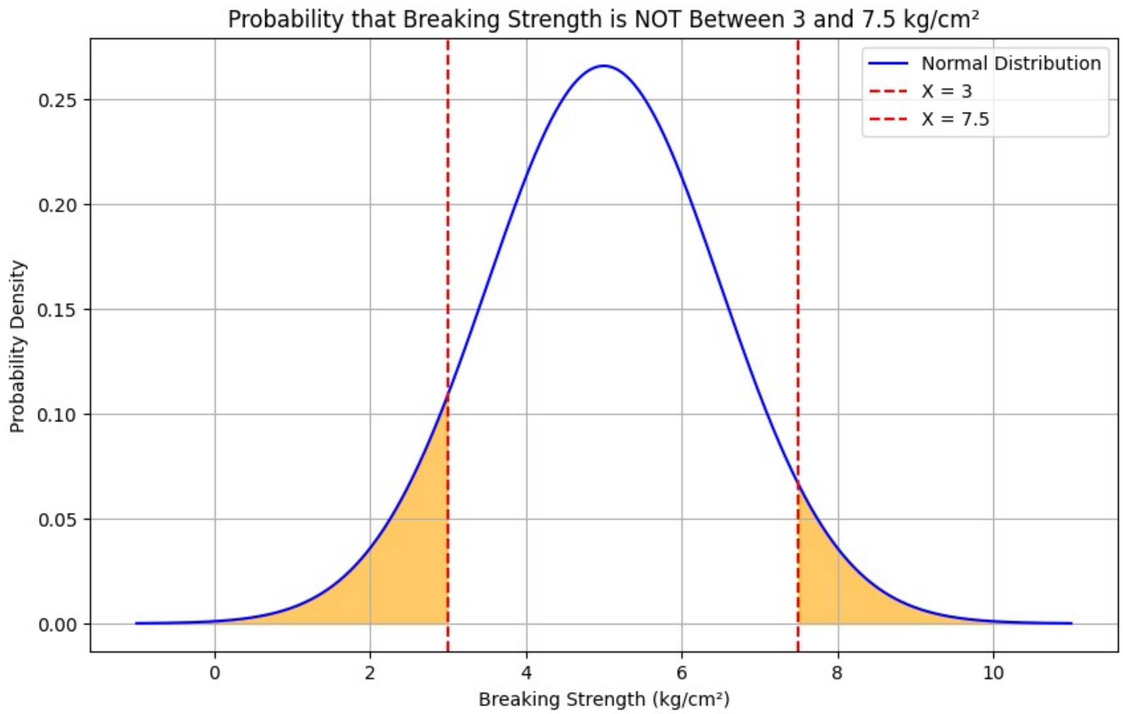


Figure: Area under the curve for $P(X < 3 \text{ or } X > 7.5)$

The shaded area under the normal distribution curve to the left of X = 3 and to the right of X = 7.5 represents the probability $P(X < 3 \text{ or } X > 7.5)$, which is the probability that a bag's breaking strength does not lie between 3 and 7.5 kg/cm².

Problem 3: Comparative Analysis of Stone Hardness – Suitability and Treatment Effect

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:

Business Context

Zingaro Stone Printing is a company that specializes in printing high-quality images onto both polished and unpolished stone surfaces. For the printing process to be effective, it is essential that the stones meet a minimum surface hardness level. Based on industry standards and internal testing, the company requires stones to have a **Brinell Hardness Index (BHI) of at least 150** for optimal print quality.

Recently, Zingaro received a batch of **unpolished stones** from a supplier. There is growing concern within the quality assurance team that these unpolished stones may not meet the minimum hardness threshold and could potentially lead to poor printing results or increased rework costs.

To validate this concern, Zingaro wishes to perform a statistical test to determine whether the **mean hardness of unpolished stones** is **significantly less than 150**, and thereby assess if the stones are indeed **unsuitable for printing**.

Hypotheses

- **Null Hypothesis (H_0):** $\mu \geq 150$
  (The average hardness of unpolished stones is at least 150 — they are suitable.)
- **Alternative Hypothesis (H_1):** $\mu < 150$
  (The average hardness of unpolished stones is less than 150 — they are unsuitable.)

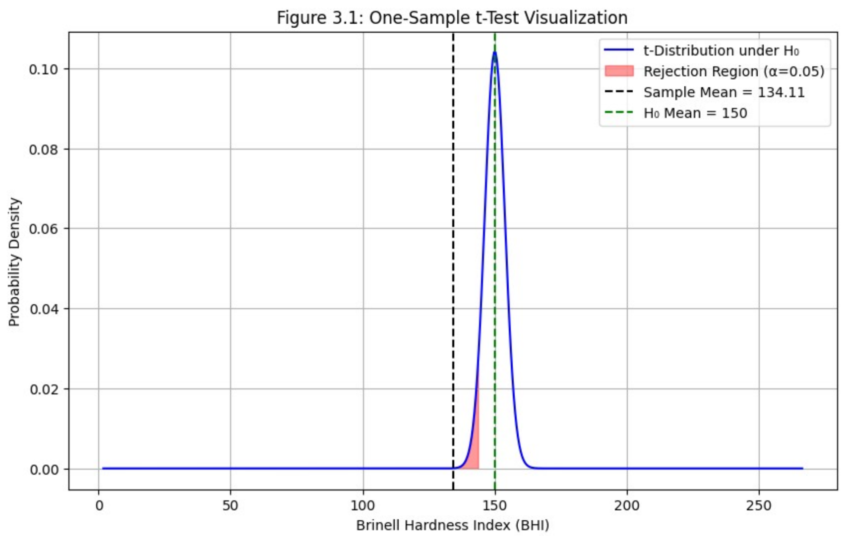
This is a **one-sample, one-tailed t-test**.

Statistical Analysis

To assess whether the unpolished stones are suitable for printing, a **one-sample t-test** was conducted using Python's ***ttest_1samp()*** function from the ***scipy.stats*** library. This test checks whether the **mean Brinell Hardness Index (BHI)** of unpolished stones is statistically lower than the required threshold of **150**.

Test Results:

- **Sample Mean**: 134.11
- **Test Statistic (t)**: -4.165
- **p-value (one-tailed)**: 0.0000



- **Blue Curve:** The t-distribution assuming the null hypothesis (H_0: $\mu = 150$).
- **Green Dashed Line:** The hypothesized mean **μ = 150**.
- **Red Shaded Area: Rejection region (α = 0.05)** for the **one-tailed** test.
- **Black Dashed Line:** The **sample mean = 134.11**, well into the rejection region.

Since the **sample mean falls in the rejection region**, and the **p-value is effectively 0**, we **reject H_0** and conclude that **Zingaro is justified** in believing the unpolished stones are unsuitable for printing.

Conclusion

At a **5% level of significance**, the **one-sample t-test** reveals that the **mean Brinell Hardness Index** of unpolished stones (**mean = 134.11**) is **significantly lower than 150**.

Since the **one-tailed p-value is 0.0000**, which is **less than 0.05**, we **reject the null hypothesis**. This statistically supports Zingaro's concern:

Zingaro is justified in believing that the unpolished stones may not be suitable for printing, as their average hardness falls below the minimum required threshold.

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

Business Context

Zingaro Stone Printing also uses **treated and polished stones** alongside unpolished ones. While polishing is expected to enhance surface quality, the company wants to validate whether it actually leads to a significant increase in **Brinell Hardness Index (BHI)**. This comparison helps determine whether the treatment process adds measurable value and justifies the additional processing costs.

To investigate this, Zingaro intends to statistically test whether the **mean hardness of polished stones** differs from that of **unpolished stones**.

Hypotheses

- **Null Hypothesis (H_0):** $\mu_1 = \mu_2$

(The average hardness of polished stones is equal to that of unpolished stones.)

- **Alternative Hypothesis (H_1):** $\mu_1 \neq \mu_2$

(The average hardness of polished stones is different from that of unpolished stones.)

This is a **two-sample, two-tailed t-test** for comparing independent means.

Statistical Analysis

To determine if there is a statistically significant difference in mean hardness between polished and unpolished stones, a **two-sample t-test** was conducted using Python's ***ttest_ind()*** function from the ***scipy.stats*** library. The test compares the means of two independent samples:

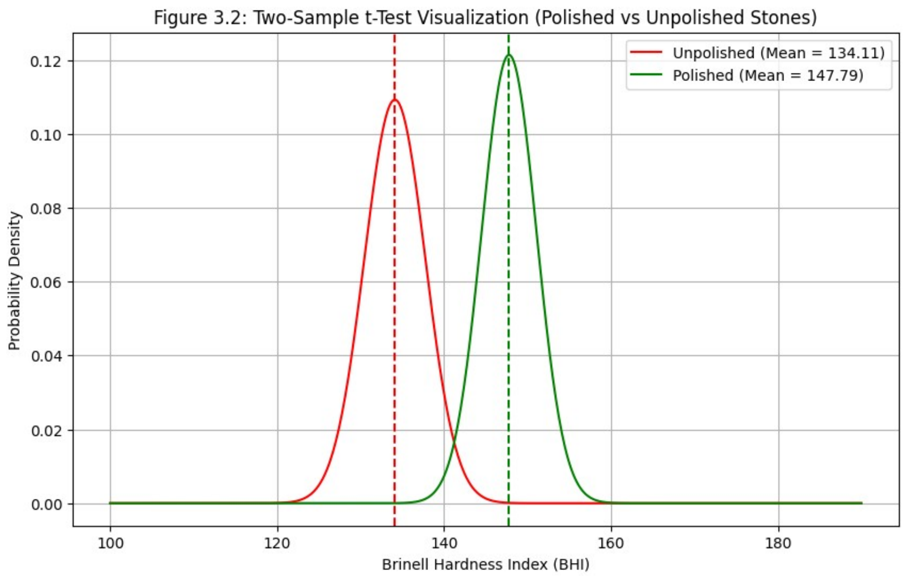
- ***unpolished_data:*** Brinell Hardness Index of unpolished stones
- ***polished_data:*** Brinell Hardness Index of treated and polished stones

The ***equal_var=False*** parameter is used to account for unequal variances (Welch's t-test).

Test Results

- Mean of unpolished stones: **134.11**
- Mean of polished stones: **147.79**

• Test Statistic (t): **-3.242**
• p-value (two-tailed): **0.0016**



- The **mean of the polished stones (green)** is **visibly shifted to the right** compared to unpolished stones (red), showing **higher average hardness**.
- The **two-sample t-test result (p = 0.0016)** confirms this **visual difference is statistically significant** at the 5% significance level.
- The **confidence in rejecting the null hypothesis is strong.**

Zingaro is justified in continuing the polishing process as it **measurably improves surface hardness**, making the extra processing cost worthwhile.

Conclusion

At a **5% level of significance**, the p-value is **0.0016**, which is **less than 0.05**.
Therefore, we **reject the null hypothesis**.

This provides **strong evidence** that the mean hardness of polished stones is **significantly different** from that of unpolished stones.

The visual and statistical evidence together confirm **a significant difference in hardness** between polished and unpolished stones.

Problem 4: Multi-Factor Analysis of Dental Implant Hardness

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

Problem Statement

The hardness of dental implants may vary based on the dentist performing the procedure. Dentists might prefer specific techniques or have varying skill levels that influence implant quality. To ensure consistency and optimize patient outcomes, Zingaro Dental Solutions seeks to determine if the **mean hardness of implants** significantly differs across dentists. The analysis is conducted **separately for Alloy 1 and Alloy 2**, as the material properties may interact differently with each dentist's techniques.

Hypotheses

For each alloy, we perform a **one-way ANOVA** with the factor being Dentist.

Both For Alloy 1 and Alloy 2 we can assume:

- **H_0 (Null Hypothesis):** The mean hardness is the same across all dentists.
- **Ha (Alternative Hypothesis):** At least one dentist has a different mean hardness.

Assumptions Checks

Before performing a One-Way ANOVA, two key assumptions must be checked:

1 Normality of Residuals (Shapiro-Wilk Test)

- The **Shapiro-Wilk test** tests the **null hypothesis that the data is normally distributed**.
- **We applied this test to the implant hardness ("Response") values for each dentist.**

Using Python ***shapiro()*** function under ***scipy.stats*** we can get,

Alloy	**Shapiro-Wilk p-value**	**Conclusion**
Alloy 1	0.0000	Reject H_0 → Not normal
Alloy 2	0.0004	Reject H_0 → Not normal

Interpretation:

- Both alloys **fail the normality test**, indicating that the implant hardness data is **not normally distributed**.
- However, **ANOVA is robust to mild violations of normality**, especially with approximately equal group sizes (as is the case here).

2 Homogeneity of Variances (Levene's Test)

The **Levene's test** checks whether **the variance (spread) of data is equal across groups (dentists)**.

Using Python's ***levene()*** function in the ***scipy.stats*** library we can get,

Alloy	**Levene's p-value**	**Conclusion**
Alloy 1	0.2566	Fail to reject → Equal variances
Alloy 2	0.2369	Fail to reject → Equal variances

Interpretation:

- The **variance of implant hardness is consistent across dentists** for both alloys.
- **This satisfies the equal variance assumption required for One-Way ANOVA.**

Despite non-normality, **we proceed with ANOVA due to robustness to this violation and homogeneity of variances**.

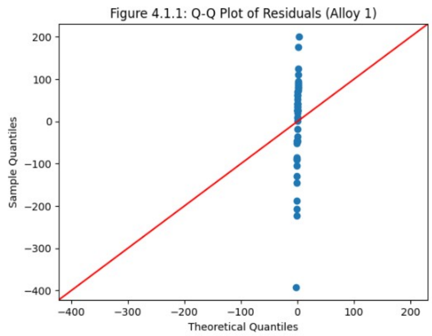


Figure 4.1.1: Q-Q Plot of Residuals (Alloy 1)

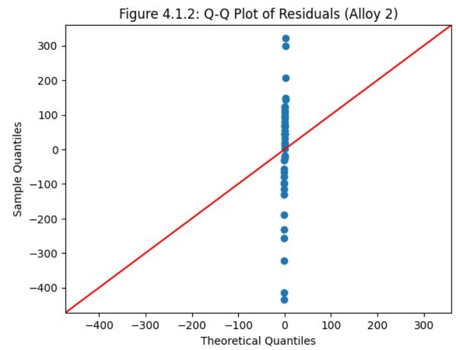


Figure 4.1.2: Q-Q Plot of Residuals (Alloy 2)

Q-Q Plots (Alloy 1 & Alloy 2)

Points deviate from the red line, especially at the ends (tails).

This confirms the non-normality we saw with Shapiro-Wilk test.

Some extreme outliers → particularly for Alloy 2.

One-Way ANOVA Results

```
Alloy 1 ANOVA Table:
                   sum_sq    df         F    PR(>F)
C(Dentist)  106683.688889   4.0  1.977112  0.116567
Residual    539593.555556  40.0       NaN       NaN

Alloy 2 ANOVA Table:
                  sum_sq    df         F    PR(>F)
C(Dentist)  5.679791e+04   4.0  0.524835  0.718031
Residual    1.082205e+06  40.0       NaN       NaN
```

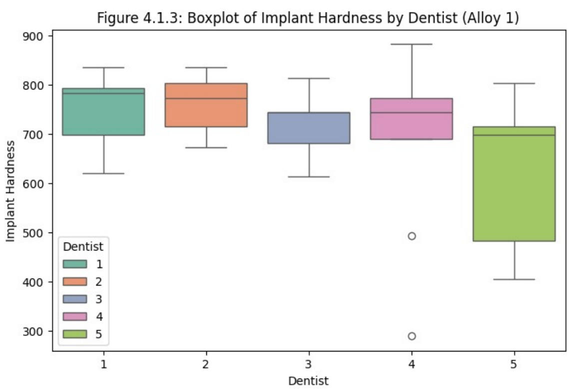
Alloy 1:

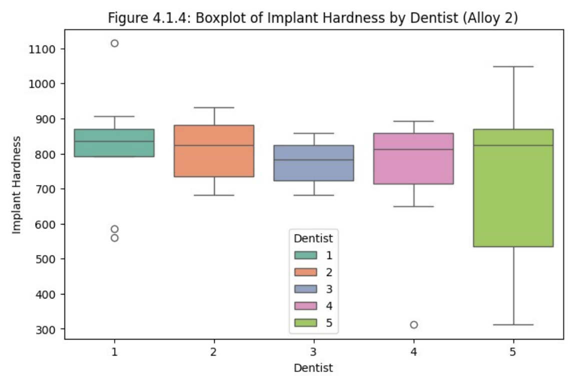
- **F-statistic = 1.955**, **p-value = 0.1166**
- The probability of observing such variation in implant hardness between dentists *by chance alone* is approximately 11.66%. Since **p > 0.05**, this result is **not statistically significant**.
- Fail to reject the null hypothesis.
  → **We do not have sufficient evidence to conclude that dentist affects implant hardness for Alloy 1.**

Alloy 2:

- **F-statistic = 0.5248**, **p-value = 0.7180**
- The observed variability between dentists is **very likely due to chance** (p = 71.80%). Again, this result is **not statistically significant**.
- Fail to reject the null hypothesis.
→ **We do not have sufficient evidence to conclude that dentist affects implant hardness for Alloy 2.**

Post Hoc Testing:
Post hoc analysis (e.g., Tukey HSD) **was not performed** because the p-values for both alloys were greater than 0.05. Since there was **no statistically significant difference** detected by ANOVA, **further pairwise comparisons were unnecessary**.





Boxplots (Alloy 1 & Alloy 2)

Variation between dentists is visible.

Dentist 5 has greater spread (variability) in both alloys.

Alloy 1: Dentists 1 & 2 seem to have higher median hardness than others, Dentist 5 appears lower and more variable.

Alloy 2: Same trend; but the spread is even greater for Dentist 5.

Outliers: Visible across both alloys → contributes to non-normality.

Based on the results of this analysis, **there is no statistically significant evidence to suggest that the hardness of dental implants differs depending on the dentist performing the procedure for either Alloy 1 or Alloy 2**. Although **visualizations revealed variability**, particularly in the case of Dentist 5, the statistical tests confirmed that **any observed differences are likely due to random variation**.

Post hoc analysis **was not conducted** as the overall ANOVA results were **not statistically significant ($p > 0.05$)**.

➞ **Conclusion:** The dentist does **not** have a significant influence on implant hardness in this dataset.

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

Problem Statement

Different implant methods may influence the hardness of dental implants, affecting product reliability and patient outcomes. To ensure consistency and optimal implant performance, **Zingaro Dental Solutions** seeks to determine if the **mean hardness of implants differs based on implant method**. The analysis is conducted **separately for Alloy 1 and Alloy 2**, as different alloys may respond differently to implant techniques.

Hypotheses

For each alloy, a **One-Way ANOVA** is performed with the factor being **Method**:

Hypothesis	**Statement**
H_0 (Null)	The mean hardness is the same across all methods.
Ha (Alternative)	At least one implant method has a different mean hardness.

Assumption Checks

1. Normality of Residuals (Shapiro-Wilk Test)

```
===== Analysis for Alloy 1 =====
Shapiro-Wilk Test: W = 0.8305, p = 0.0000
Levene's Test:    F = 6.5214, p = 0.0034

===== Analysis for Alloy 2 =====
Shapiro-Wilk Test: W = 0.8878, p = 0.0004
Levene's Test:    F = 3.3497, p = 0.0447
```

Alloy	**Shapiro-Wilk p-value**	**Conclusion**
Alloy 1	0.0000	Reject H_0 → Not Normal
Alloy 2	0.0004	Reject H_0 → Not Normal

Interpretation:

- Both alloys **fail the normality test**, indicating that the implant hardness data is **not normally distributed**.
- **ANOVA is robust to mild normality violations**, especially with balanced group sizes.

2. Homogeneity of Variances (Levene's Test)

Alloy	**Levene's p-value**	**Conclusion**
Alloy 1	0.0034	Reject H_0 → Variances not equal
Alloy 2	0.0447	Reject H_0 → Variances not equal

Interpretation:

- The variance of implant hardness **differs across methods** for both alloys.
- **Despite assumption violations, we proceed with ANOVA** as instructed, with caution in interpreting the results. Based on the results, we **fail to reject the null hypothesis**

One Way ANOVA Result

```
Alloy 1 ANOVA Table:
                   sum_sq    df         F    PR(>F)
C(Method)  148472.177778   2.0  6.263327  0.004163
Residual   497805.066667  42.0       NaN       NaN

Alloy 2 ANOVA Table:
              sum_sq    df        F    PR(>F)
C(Method)  499640.4   2.0  16.4108  0.000005
Residual   639362.4  42.0      NaN       NaN
```

Alloy 1:

- **F-statistic = 6.263**, **p-value = 0.0042**
- The probability of observing such variation in implant hardness between implant methods by chance alone is approximately **0.42%**. Since **$p < 0.05$**, this result **is statistically significant**.
- **Reject the null hypothesis. → We have sufficient evidence to conclude that implant method affects implant hardness for Alloy 1.**

Alloy 2:

- **F-statistic = 16.411**, **p-value = 0.000005**
- The observed variability between implant methods is **extremely unlikely** to be due to chance (**$p = 0.000005$**). This result **is statistically significant**.
- **Reject the null hypothesis. → We have sufficient evidence to conclude that implant method affects implant hardness for Alloy 2.**

Post Hoc Testing: Required and Conducted

- **Post hoc testing (Tukey HSD)** was conducted to identify **which specific implant methods differ significantly**.

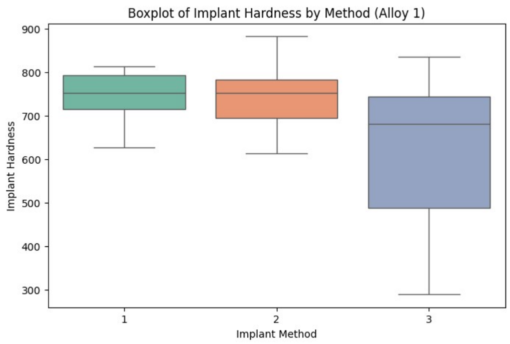


Figure 4.2.1

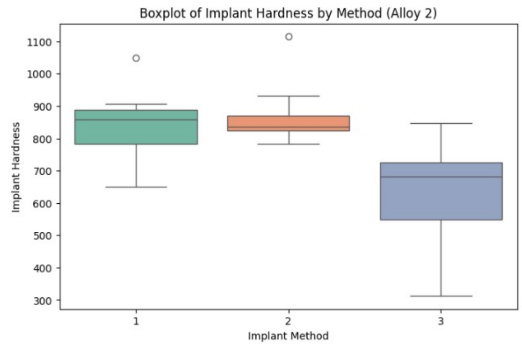


Figure 4.2.2

- **Method 3** shows a **wider spread and lower median hardness** compared to Methods 1 and 2.
- **Method 3** exhibits **lower implant hardness values** with greater variability.
- **Methods 1 and 2 produce consistently higher hardness.**

Post Hoc Testing Interpretation (Tukey HSD Results)

The overall ANOVA tests revealed significant effects of implant method on hardness for both alloys (p = 0.0042 for Alloy 1, p = 0.000005 for Alloy 2), prompting Tukey HSD to identify specific method differences.

```
===== Tukey HSD Post Hoc Test for Alloy 1 =====

ANOVA Table:
                  sum_sq    df         F    PR(>F)
C(Method)  148472.177778   2.0  6.263327  0.004163
Residual   497805.066667  42.0       NaN       NaN

Tukey HSD Results:
  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
---------------------------------------------------------
```

```
===== Tukey HSD Post Hoc Test for Alloy 2 =====

ANOVA Table:
              sum_sq    df        F    PR(>F)
C(Method)   499640.4   2.0  16.4108  0.000005
Residual    639362.4  42.0      NaN       NaN

Tukey HSD Results:
  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
--------------------------------------------------------
```

Alloy 1:

- No significant difference between Method 1 and Method 2.
- Significant difference between:
  - Method 1 vs Method 3
  - Method 2 vs Method 3
- **Conclusion: Implant Method 3** produces **significantly lower hardness** than Methods 1 and 2.

Alloy 2:

- No significant difference between Method 1 and Method 2.
- Significant difference between:
    - Method 1 vs Method 3
    - Method 2 vs Method 3
- **Conclusion: Implant Method 3 consistently produces significantly lower implant hardness**.

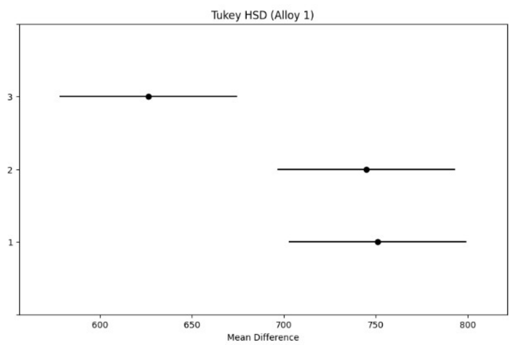


Figure 4.2.3

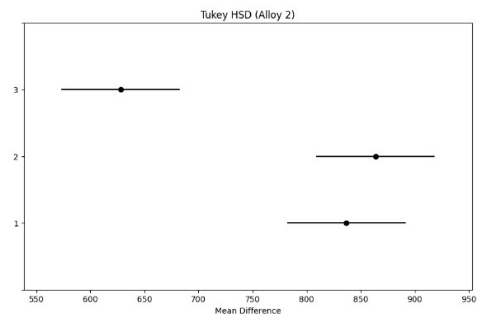


Figure 4.2.4

The **Tukey HSD plots** visually reinforce the statistical findings from the post hoc analysis. In **both Alloy 1 and Alloy 2**, **Method 3** shows **confidence intervals that do not overlap** with those of **Methods 1 and 2**, clearly indicating **statistically significant differences**. This confirms that **Method 3 consistently results in lower implant hardness** compared to the other two methods. Conversely, **Methods 1 and 2** have **overlapping confidence intervals**, suggesting **no significant difference** between them. These plots provide a **clear and intuitive visual representation** of the differences in implant hardness across methods, **complementing** the numerical results from the **Tukey HSD tables**.

Summary Interpretation:

- **In both Alloy 1 and Alloy 2, Implant Method 3 consistently results in significantly lower implant hardness compared to Methods 1 and 2.**
- **Methods 1 and 2 are statistically similar.**

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

Objective

The objective of this analysis is to determine whether there is an **interaction effect** between **dentist** and **implant method** on **implant hardness**. Understanding these interactions helps ensure consistency in implant quality and guides process improvements.

Alloy 1 and Alloy 2 were analyzed independently.

Interaction Plot Results

Alloy 1:

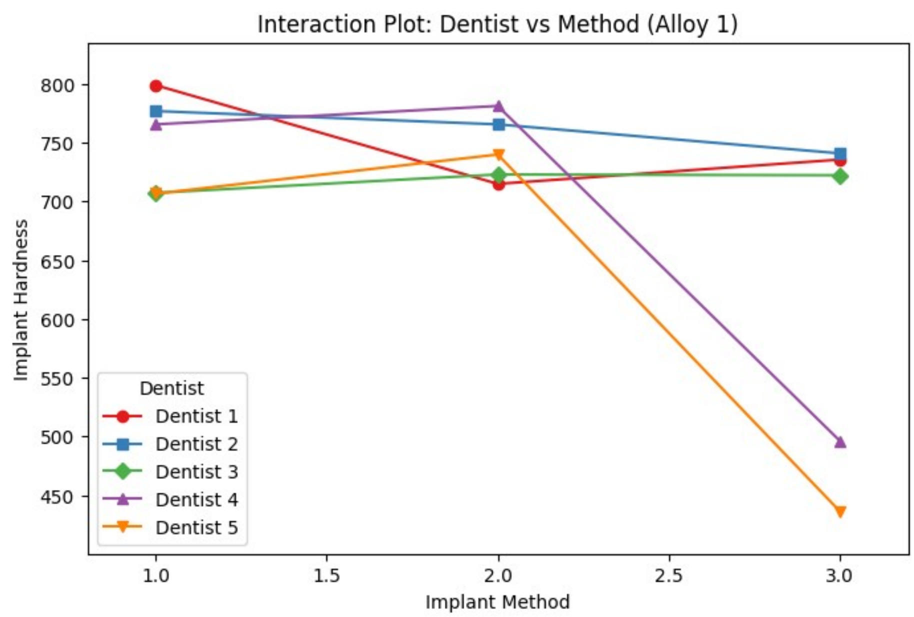


Figure 4.3.1

Inference for Alloy 1:

- The interaction plot for **Alloy 1** displays **non-parallel lines** and **line crossings**, which visually suggest a possible interaction between **dentist** and **implant method**.
- **Dentist 4 and Dentist 5** show the most pronounced changes — both exhibit sharp decreases in implant hardness from Method 2 to Method 3.
- **Dentist 3** shows minimal change across all methods, while others (Dentist 1 and 2) have moderate shifts.

- → **Conclusion**: These patterns provide **visual evidence** of an interaction between **dentist** and **method** for Alloy 1. The impact of method on implant hardness **varies across dentists**, indicating a potential interaction effect.

Alloy 2:

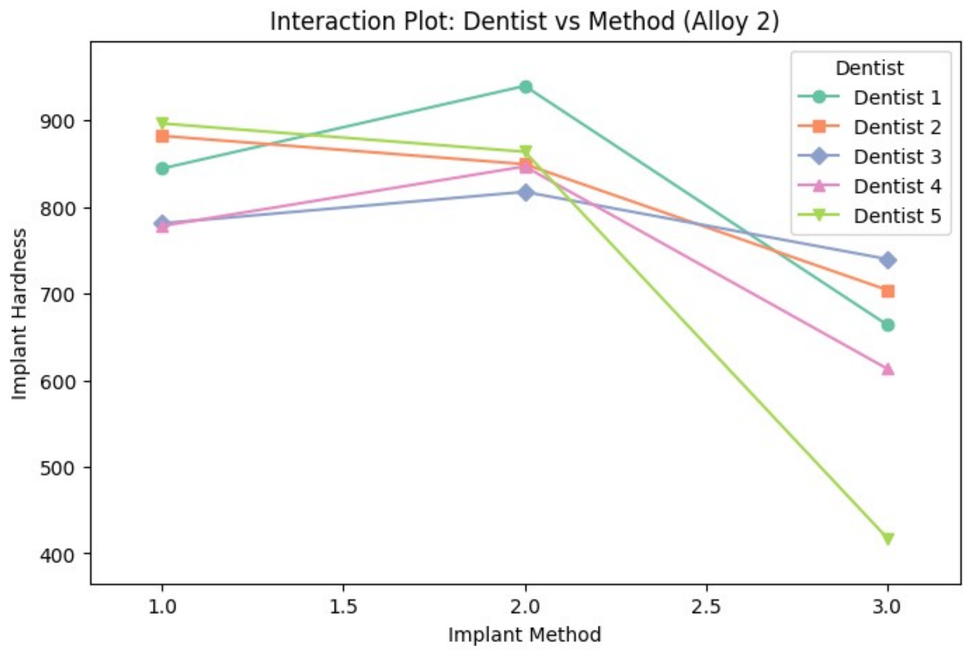


Figure 4.3.2

Inference for Alloy 2:

- The lines for different dentists are **not parallel** — this is the first clue that interaction **may** be present.
- Some lines **cross each other** — another classic indicator of potential interaction.
- Most dentists show moderate trends across methods (slight increase or decrease).
- **Dentist 4** and **Dentist 5** display a **steep drop** in hardness from Method 2 to Method 3.
- Dentists 2 and 3 show fairly stable or slightly rising/falling trends.
- **Dentist 3's line is nearly flat**, suggesting their results are unaffected by method.

Conclusion

- **Both Alloy 1 and Alloy 2 demonstrate a clear interaction effect between dentist and implant method.**
- **➔ The choice of implant method influences implant hardness differently for different dentists.**
- **➔ Dentist 5 consistently shows greater variability in implant hardness across methods, suggesting that certain combinations of dentist and method may result in significantly lower hardness values.**
- **➔ These interaction effects highlight the importance of training consistency and potentially standardizing implant techniques across dentists to ensure optimal implant hardness.**

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

Problem Statement

The hardness of dental implants may be influenced by multiple factors, including **the dentist performing the procedure** and **the method of implantation used**. While Zingaro Dental Solutions has already examined the individual effects of dentists and methods separately, it is critical to determine whether **these factors have a combined effect** on implant hardness.

Understanding this combined effect helps answer two key business questions:

1. **Does the implant method influence hardness differently for different dentists?**
2. **Are certain combinations of dentists and methods producing better or worse results?**

The objective of this analysis is to conduct a **Two-Way ANOVA with interaction** to test whether there are **statistically significant main effects of dentist and method**, and whether there is a **significant interaction effect between them** on the **hardness of implants**.

Hypotheses

For each alloy, we perform a **Two-Way ANOVA with Interaction** to assess how implant hardness varies **depending on dentists, methods, and their combined effect**.

Null Hypotheses (H_0):

Main Effect of Dentist (H_{01}):
→ The **mean hardness of implants is the same for all dentists**.

Main Effect of Method (H_{02}):
→ The **mean hardness of implants is the same for all implant methods**.

Interaction Effect (H_{03}):
→ There is **no interaction** between **dentist** and **method** on implant hardness.
(This means that the effect of implant method on hardness is consistent across all dentists.)

Alternative Hypotheses (H_1):

Main Effect of Dentist (H_{11}):
→ The **mean hardness of implants differs for at least one dentist**.

Main Effect of Method (H_{12}):
→ The **mean hardness of implants differs for at least one implant method**.

Interaction Effect (H_{13}):
→ There **is an interaction** between **dentist** and **method**, meaning **the effect of implant method on hardness varies depending on which dentist performs the procedure**.

Assumption Checks

Before conducting the Two-Way ANOVA with interaction, we tested two important assumptions for each alloy:

```
===== Assumption Checks for Alloy 1 =====
Shapiro-Wilk Test (Normality of residuals): p = 0.0902
Levene's Test (Homogeneity of variances): p = 0.3128

===== Assumption Checks for Alloy 2 =====
Shapiro-Wilk Test (Normality of residuals): p = 0.0946
Levene's Test (Homogeneity of variances): p = 0.7832
```

Alloy 1:

Assumption	Test	p-value	Conclusion
Normality of Residuals	Shapiro-Wilk Test	0.0902	Fail to reject H_0 → Normal
Equal Variances	Levene's Test	0.3128	Fail to reject H_0 → Equal

➞ **Conclusion (Alloy 1):** Both normality and equal variances assumptions are satisfied.

Alloy 2:

Assumption	Test	p-value	Conclusion
Normality of Residuals	Shapiro-Wilk Test	0.0946	Fail to reject H_0 → Normal
Equal Variances	Levene's Test	0.7832	Fail to reject H_0 → Equal

➞ **Conclusion (Alloy 2):** Both assumptions of normality and equal variances are satisfied.

- For **both Alloy 1 and Alloy 2**, the **p-values are greater than 0.05** for both tests.
- **Therefore, we fail to reject the null hypotheses for both tests,** indicating:
    - The residuals **do not deviate significantly from normality**.
    - The variances of implant hardness **are consistent across all combinations of dentists and methods**.
- The assumptions for performing Two-Way ANOVA with interaction are **satisfactorily met** for both alloys.

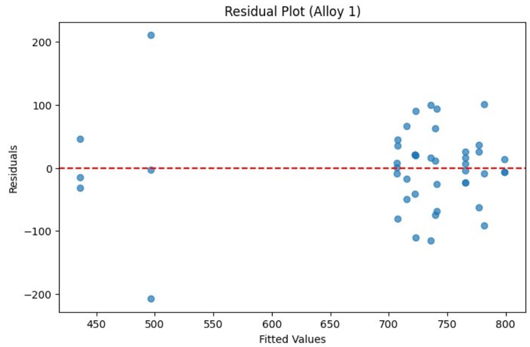


Figure no. 4.4.1

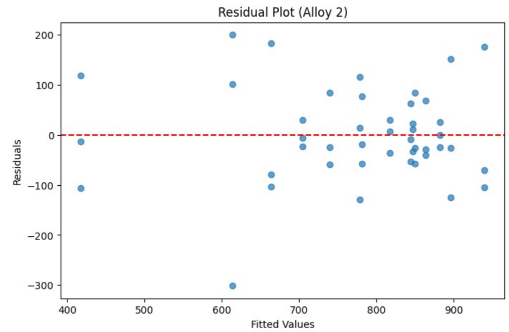


Figure no. 4.4.2

Alloy 1:

- The residuals exhibit variability at lower fitted values, with some extreme outliers observed.
- At higher fitted values, the residuals are more tightly clustered around zero, indicating a better fit in this range.

Alloy 2:

- Residuals are more spread out across the range of fitted values.
- There is evidence of outliers and potential non-constant variance (heteroscedasticity).

ANOVA Results

Alloy 1 (Two-Way ANOVA with Interaction):

Effect	F	p-value	Conclusion
C(Dentist)	3.900	0.0115	Significant difference → Dentist affects hardness
C(Method)	10.854	0.0003	Significant difference → Method affects hardness
C(Dentist):C(Method)	3.398	0.0068	Significant interaction → Effect of Method depends on Dentist

Interpretation for Alloy 1:

- **Both Dentist and Method have significant effects on implant hardness.**
- **There is a significant interaction → the effect of implant method on hardness depends on which dentist performs the procedure.**
- **→ Further analysis recommended on specific dentist-method pairs (Tukey HSD conducted).**

Alloy 2 (Two-Way ANOVA with Interaction):

Effect	F	p-value	Conclusion
C(Dentist)	1.106	0.3718	Not significant
C(Method)	19.461	0.0000	Highly significant difference → Method affects hardness
C(Dentist):C(Method)	1.923	0.0932	Not significant (Borderline, p ≈ 0.093)

Interpretation for Alloy 2:

- **Method has a highly significant effect on implant hardness.**
- **No statistically significant interaction → effect of method on hardness is similar across dentists.**
- **→ No further post hoc required for interaction. Tukey HSD was conducted for Method.**

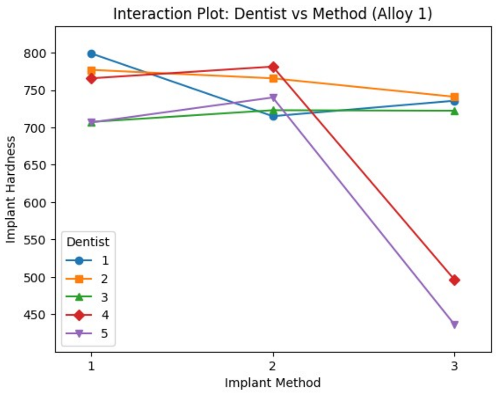


Figure no. 4.4.3

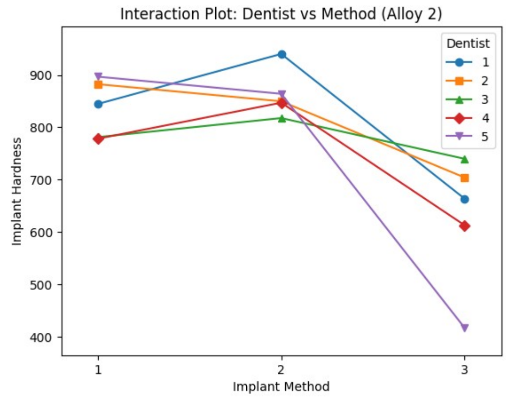


Figure no. 4.4.4

The interaction plot for Alloy 1 reveals that the lines representing different dentists are not parallel, indicating a significant interaction effect between dentist and method on implant hardness. This suggests that the effect of the method on hardness depends on the dentist performing the procedure. Specifically, Method 3 shows a drastic decrease in implant hardness for certain dentists, such as Dentist 5, emphasizing that Method 3 consistently underperforms, albeit with varying effects across dentists. In contrast, Methods 1 and 2 exhibit relatively stable hardness levels across dentists, with smaller variations.

For Alloy 2, the interaction plot shows more parallel lines compared to Alloy 1, supporting the non-significant interaction effect observed in the ANOVA results. This implies that the method's effect on implant hardness is consistent across different dentists. However, Method 3 still consistently results in lower implant hardness across all dentists, with some minor variations. Methods 1 and 2 display similar trends, with only minor differences in hardness levels across dentists.

Post Hoc Analysis (Tukey HSD) Results

```
Tukey HSD Results for Method (Alloy 1):
  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
----------------------------------------------------------

Tukey HSD Results for Method (Alloy 2):
  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
----------------------------------------------------------
```

Since the ANOVA indicated a significant effect of **Method** on implant hardness for both alloys, **Tukey HSD post hoc tests** were conducted to determine *which specific methods* differed from each other.

Alloy 1: Tukey HSD Results Summary

Comparison	Mean Difference	p-value	Conclusion
Method 1 vs 2	-6.13	0.987	Not significant
Method 1 vs 3	**-124.80**	**0.0085**	**Significant difference**
Method 2 vs 3	**-118.67**	**0.0128**	**Significant difference**

Interpretation for Alloy 1:

- **Methods 1 and 2 are similar.**
- **Method 3 produces significantly different (lower) implant hardness** than both **Method 1** and **Method 2**.

Alloy 2: Tukey HSD Results Summary

Comparison	Mean Difference	p-value	Conclusion
Method 1 vs 2	27.00	0.821	Not significant
Method 1 vs 3	**-208.80**	**0.0001**	**Highly significant difference**
Method 2 vs 3	**-235.80**	**0.0000**	**Highly significant difference**

Interpretation for Alloy 2:

- **Methods 1 and 2 are similar.**
- **Method 3 produces significantly lower implant hardness** than both **Method 1** and **Method 2**.

In **both Alloy 1 and Alloy 2**, **Method 3 consistently leads to significantly lower implant hardness.**

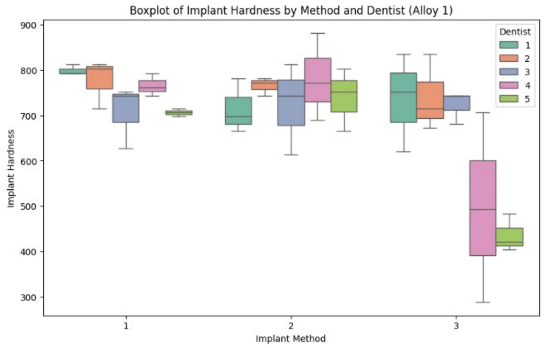


Figure no. 4.4.5

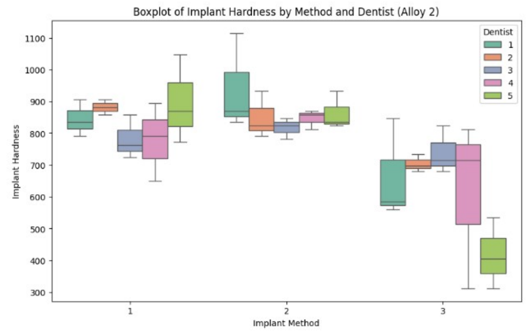
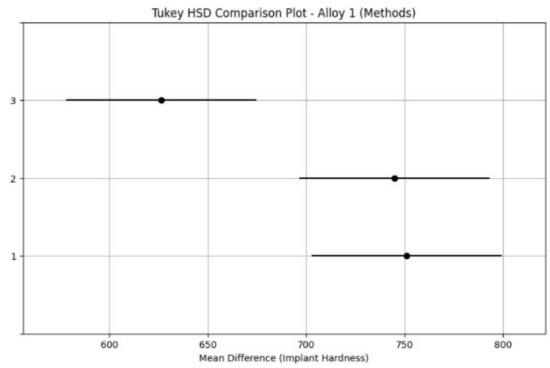


Figure no. 4.4.6

The boxplot for Alloy 1 highlights that Method 3 consistently results in the lowest implant hardness, with higher variability across dentists, further supporting the significant interaction effect identified in the analysis. In contrast, Methods 1 and 2 demonstrate more uniform implant hardness across dentists, characterized by smaller interquartile ranges and fewer outliers. Notably, Dentist 5 exhibits the most pronounced reduction in hardness when employing Method 3.

For Alloy 2, the boxplot similarly shows that Method 3 yields the lowest implant hardness, accompanied by considerable variability across dentists. However, the interaction effect is not statistically significant for this alloy. Methods 1 and 2 continue to display higher and more consistent implant hardness, with relatively smaller variations in performance across dentists. Although Dentist 5 still shows lower hardness values for Method 3, the disparity is less marked compared to Alloy 1.



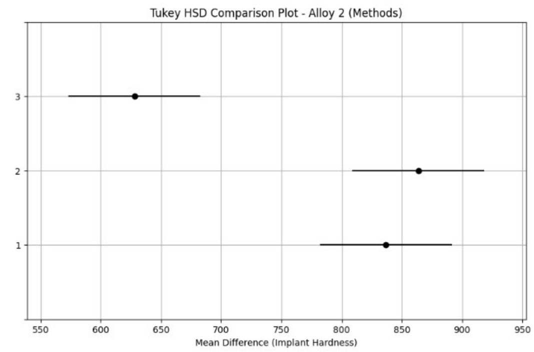


Figure no. 4.4.7

Figure no. 4.4.8

The purpose of the Tukey HSD plots is to compare the means of different groups, such as implant methods, while accounting for multiple comparisons. The confidence intervals in the plots indicate the range of mean differences between groups. For Alloy 1, the intervals for some methods overlap, suggesting no significant differences between them. However, Method 3 appears distinct from Methods 1 and 2 in terms of implant hardness, indicating a potential significant difference. In contrast, for Alloy 2, the intervals reveal that Method 1 is clearly distinct from Method 3, as their confidence intervals do not overlap. Meanwhile, Methods 2 and 3 show considerable overlap, suggesting their means may not significantly differ. These findings indicate that, for Alloy 1, some methods may perform similarly, with one method standing out, whereas for Alloy 2, at least one method demonstrates a significant difference in performance compared to the others.

Summary of Findings:

The analysis revealed that both the choice of dentist and implant method influence implant hardness for Alloy 1, with a significant interaction effect indicating that the effect of method varies by dentist. For Alloy 2, implant method is the primary influencing factor, while dentist choice does not significantly alter hardness outcomes. Method 3 consistently results in significantly lower implant hardness for both alloys. These findings suggest that Zingaro Dental Solutions should reconsider the use of Method 3 or ensure standardized procedures among dentists to maintain implant quality.

➔ **Business Recommendation:** Review and optimize implant techniques, particularly focusing on Method 3, and consider targeted training for dentists to reduce variability.