Business Report

Inferential Statistics

Coded

April 7, 2024

Authored by: Kartik Trivedi

Table of Contents

List of Figures

List of Tables

Data Dictionary

Problem 3

Column Name	Description	Data Type
Unpolished	Brinell's hardness index value of unpolished stones	Float 64
Treated and Polished	Brinell's hardness index value of polished stones	Float 64

Problem 4

Column Name	Description	Data Type
Dentist	Dentist who has done the implant	
Method	Type of method used by dentist	
Alloy	Type of allow used	
Temp	Temperature at which the metal is treated	
Response	Hardness of metal implant	

Executive Summary

Problem 1

Problem Statement

The objective is to help a physiotherapist of a male football team understand the relationship between foot injuries and the positions at which the players play.

Key Takeaways

Probability of a randomly chosen player suffering an injury

- There is 61.7% chance that a randomly chosen player will suffer an injury.

Probability of a randomly chosen a player being a forward or a winger

- There is 52.34% chance that a randomly chosen a player is a forward or a winger.

Probability of a randomly chosen player playing in a striker position and has a foot injury

- there is 19.15% chance that a randomly chosen player plays in a striker position and has a foot injury.

Probability of a randomly chosen injured player being a striker

- There is 31.03% chance that a randomly chosen injured player is a striker.

Problem 2

Problem Statement

The objective of this analysis is to help the quality team of a cement company better understand wastage or pilferage within the supply chain by studying the breaking strength of gunny bags used by them for packaging cement.

Key Takeaways

- Proportion of the gunny bags having a breaking strength of less than 3.17 kg per sq cm are 11.12%.
- Proportion of the gunny bags having a breaking strength of at least 3.6 kg per sq cm are 82.47%.
- Proportion of the gunny bags with breaking strength between 5 and 5.5 kg per sq cm are 13.06%.
- Proportion of the gunny bags not having breaking strength between 3 and 7.5 kg per sq cm are 13.9%.

Problem 3

Problem Statement

The objective of this analysis is to study the Brinell's hardness index value of a new batch of polished and unpolished stones received by Zingaro stone printing company.

Key Takeaways

- Brinell's hardness index value of unpolished stone was 4.16 units lower on an average when compared to the minimum value of 150 required for optimum level of printing of the image.
- Probability of Mean hardness of the polished and unpolished stones being same is 0.00008342, which is significantly lower than the level of significance (0.05).
- Based on these results, with a 95% confidence level, we can conclude that the assertion that unpolished stones may not be suitable for printing is justified.
- Brinell's hardness index value for polished and unpolished stones showed on an average a variance of 3.56 units.
- Probability of Mean hardness of the polished and unpolished stones being same is 0.00065, which is significantly lower than the level of significance (0.05).
- Based on these results, with a 95% confidence level, we can conclude that mean hardness of the polished and unpolished stones is not same.

Problem 4

Business Objective

The objective is to study the effect of different factors over the hardness of metal implants in dental cavities.

Key Takeaways

For Alloy 1

- Dentist as a factor can explain only 16.5% variability in hardness of metal implant.
- Method as a factor can explain only 22.97% variability in hardness of metal implant.
- Dentist and Method together explains 68.25% variability in hardness of metal implant.
- There is notable interaction effect between Dentists and Methods, particularly indicating significant differences in the mean values observed for Method 3 utilized by Dentist 5 compared to almost all methods employed by any dentist.

For Alloy 2

- Dentist as a factor can explain only 4.98% variability in hardness of metal implant.
- Method as a factor can explain 43.87% variability in hardness of metal implant.
- Dentist and Method together explains 66.19% variability in hardness of metal implant.
- Method as a predictor is the most important factor influencing the Hardness of dental implant, based on the provided results.

Problem 1

1.1 Context

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

1.2 Problem Statement

The objective is to help a physiotherapist of a male football team understand the relationship between foot injuries and the positions at which the players play.

1.3 Methodology

1. **Data Collection:** A table was shared which contained information regarding number of players at various positions and their injury status.
2. **Tools and Software:** We have carried out the analysis using programming language python on Jupyter notebook.

1.4 Data Overview

1. **Data Overview:** Following crosstab was shared where each column represented a playing position and rows represented injury status i.e. number of players injured or not injured.

	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 1 crosstab

1.5 Analysis and Findings

1.5.1 Probability of a randomly chosen player suffering an injury.

Based on the above provided table probability of a randomly chosen player suffering an injury is 0.617 or there is 61.7% chance that a randomly chosen player will suffer an injury.

1.5.2 Probability of a randomly chosen a player being a forward or a winger.

Probability of a randomly chosen a player being a forward or a winger is 0.5234 or there is 52.34% chance that a randomly chosen a player is a forward or a winger.

1.5.3 Probability of a randomly chosen player playing in a striker position and has a foot injury.

Probability of a randomly chosen player playing in a striker position and has a foot injury is 0.1915 or there is 19.15% chance that a randomly chosen player plays in a striker position and has a foot injury.

1.5.4 Probability of a randomly chosen injured player being a striker.

Probability of a randomly chosen injured player being a striker is 0.3103 or there is 31.03% chance that a randomly chosen injured player is a striker.

Problem 2

2.1 Context

The quality team of a cement company wants some insights on the breaking strength of gunny bags used for packaging cement so that they can better understand wastage or pilferage within the supply chain.

2.2 Problem Statement

The objective of this analysis is to help the quality team of a cement company better understand wastage or pilferage within the supply chain by studying the breaking strength of gunny bags used by them for packaging cement.

2.3 Methodology

1. **Data Collection:** For this problem a dataset was not provided, however, some key data like mean and standard deviation for the breaking strength of gunny bags and type of data distribution were provided.
2. **Visualization Technique:** We have utilized line plots to generate distribution curves, aiming to facilitate the quality team's comprehension of the explanations provided for the raised queries.

3. **Tools and Software:** We have carried out the analysis using programming language python on Jupyter notebook. For this analysis Python libraries Numpy, Pandas, Matplotlib, Seaborn and Scipy were used.
4. **Assumptions and Limitations:** For this problem dataset was not provided, only figures for mean and standard deviation were provided along with which it was stated that the data is normally distributed. The entire analysis is done based on the assumption that the information shared is correct.

2.4 Analysis and Findings

2.4.1 Proportion of the gunny bags with breaking strength of less than 3.17 kg per sq cm

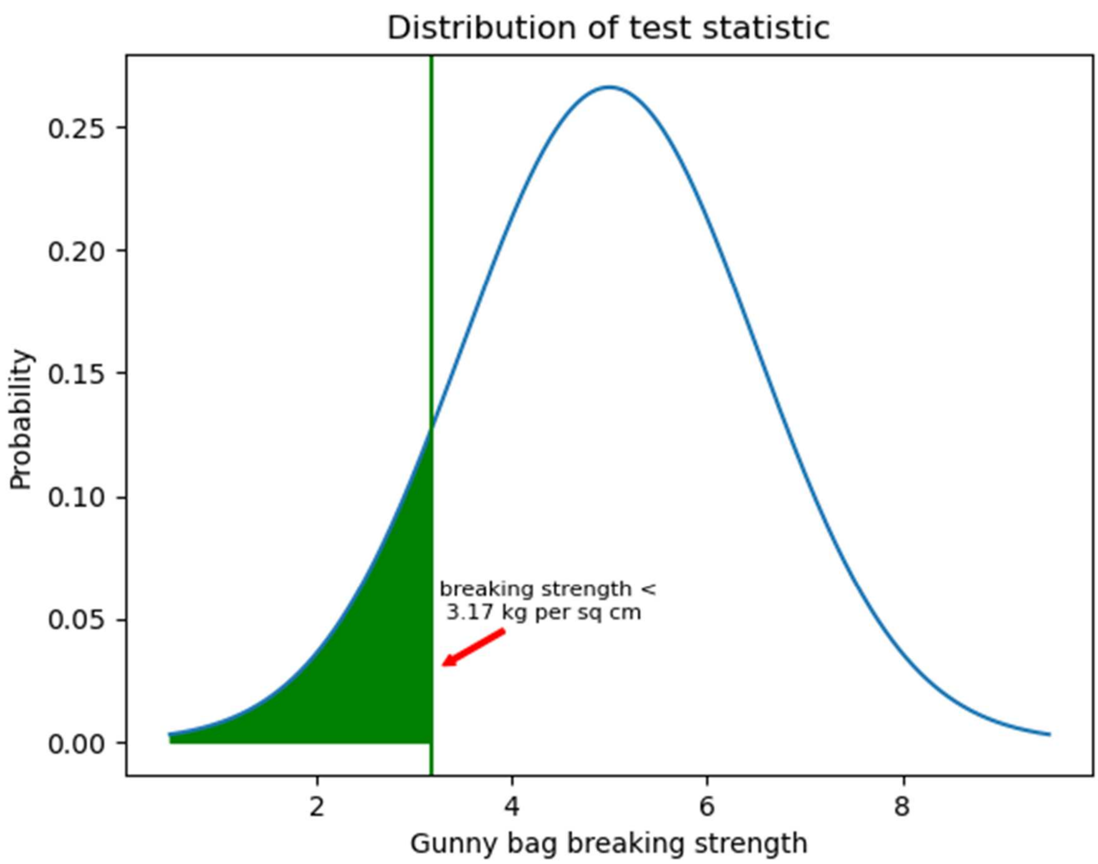

Figure1: Distribution plot highlighting proportion of the gunny bags with breaking strength of less than 3.17 kg per sq cm

In the above figure green coloured area in the curve highlights the region where breaking strength of gunny bags is less than 3.17 kg per sq cm. By proportion this area accounts for 11.12%.

Based on the above data we can conclude that almost 11.12% gunny bags have breaking strength of less than 3.17 kg per sq cm.

2.4.2 Proportion of the gunny bags with a breaking strength of at least 3.6 kg per sq cm.

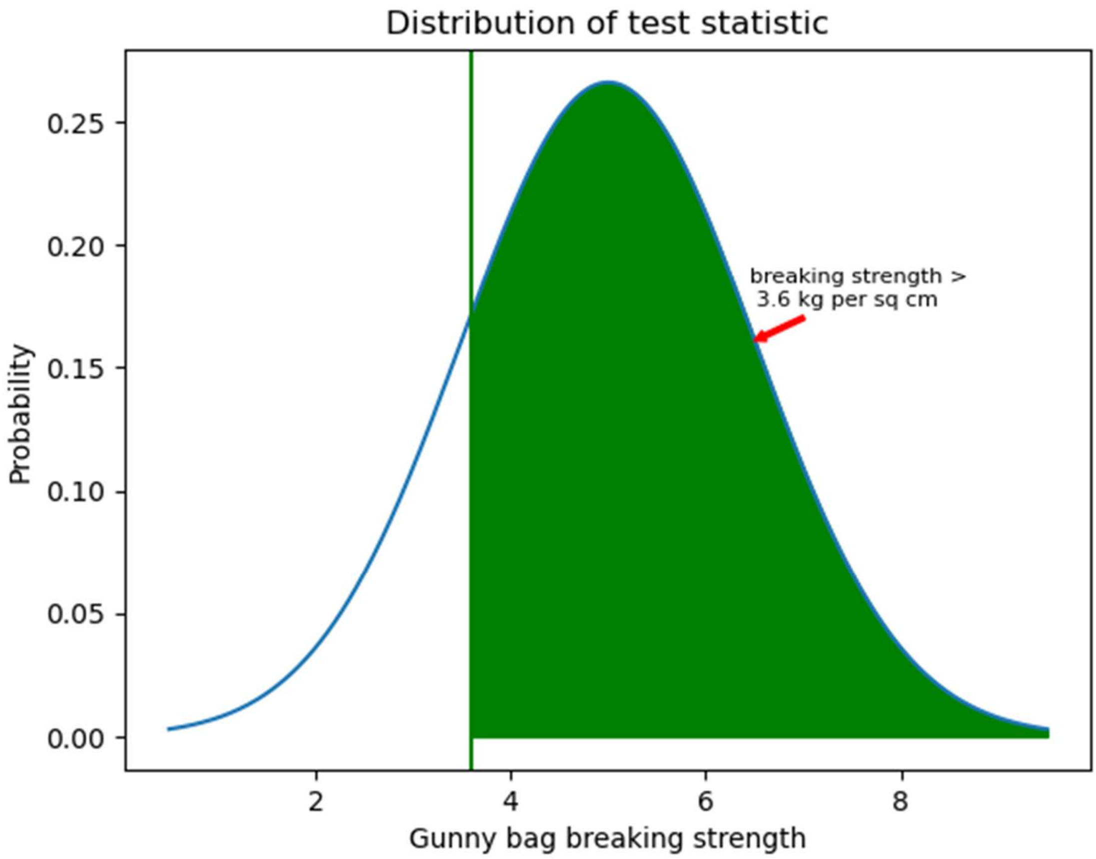

Figure2: Distribution plot highlighting proportion of the gunny bags with breaking strength of at least 3.6 kg per sq cm

In the above figure green coloured area in the curve highlights the region where gunny bags have a minimum breaking strength of 3.6 kg per sq cm. By proportion this area accounts for 82.47%.

Based on the above data we can conclude that almost 82.47% gunny bags have minimum breaking strength of 3.6 kg per sq cm.

2.4.3 Proportion of gunny bags with breaking strength between 5 and 5.5 kg per sq cm

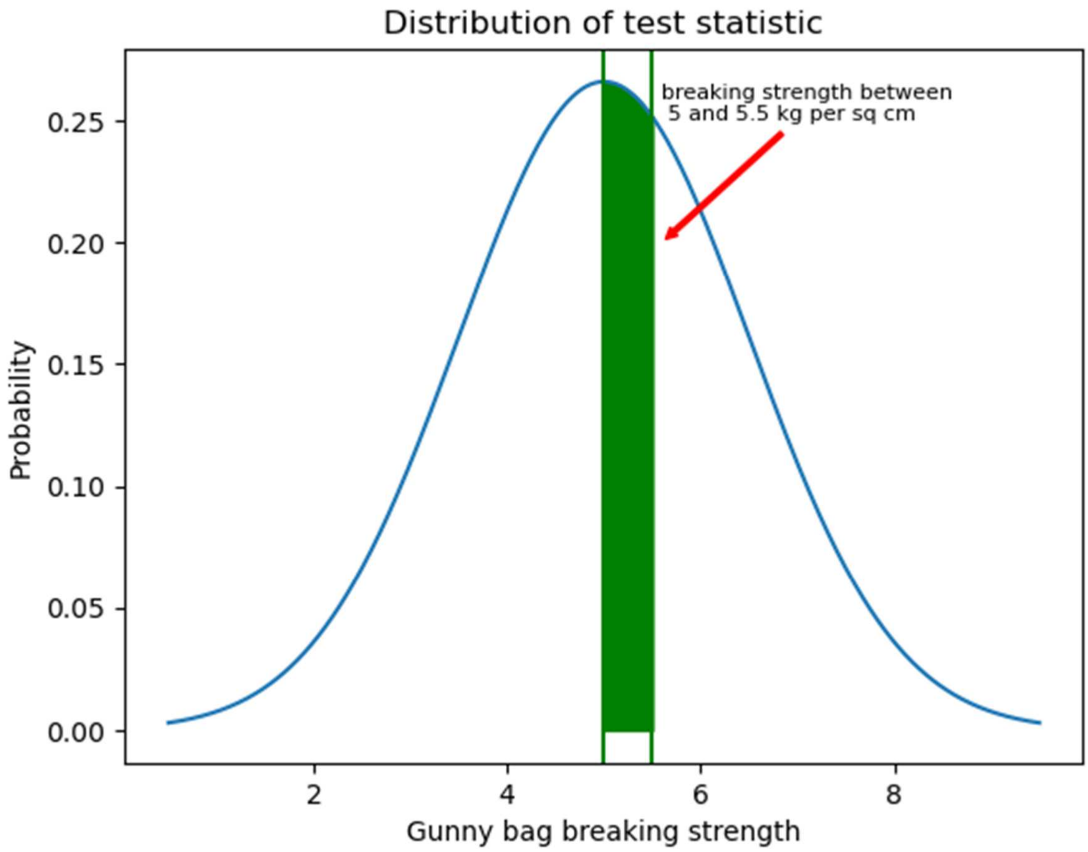

Figure3: Distribution plot highlighting proportion of the gunny bags with breaking strength between 5 and 5.5 kg per sq cm

In the above figure green coloured area in the curve highlights the region where gunny bags have breaking strength between 5 and 5.5 kg per sq cm. By proportion this area accounts for 13.06%.

Based on the above data we can conclude that almost 13.06% gunny bags have minimum breaking strength between 5 and 5.5 kg per sq cm.

2.4.4 Proportion of the gunny bags not having breaking strength between 3 and 7.5 kg per sq cm.

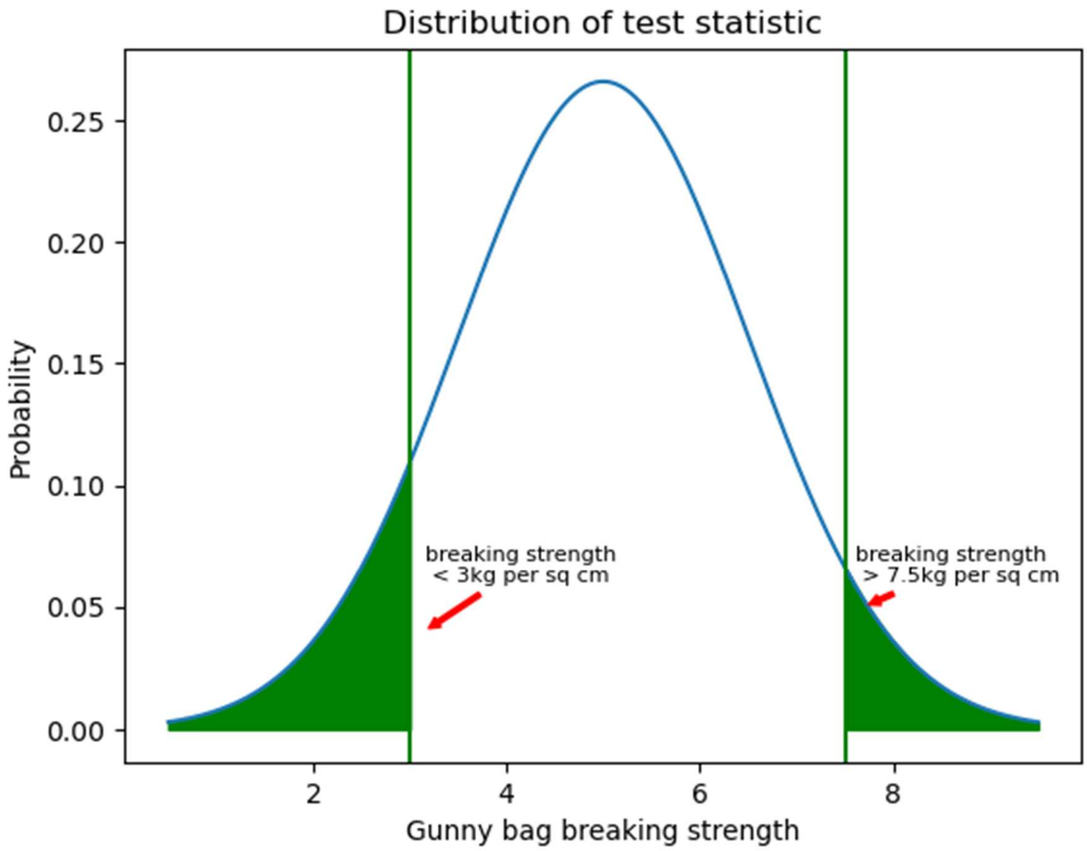

Figure4: Distribution plot highlighting proportion of the gunny bags with breaking strength below 3 or above 7.5 kg per sq cm

In the above figure green coloured area in the curve highlights the region where breaking strength of gunny bags is not between3 and 7.5 kg per sq cm or in other words breaking strength is below 3 or above 7.5 kg per sq cm. By proportion this area accounts for 13.9%.

Based on the above data we can conclude that almost 13.9% gunny bags have breaking strength below 3 or above 7.5 kg per sq cm.

Problem 3

3.1 Context

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.

3.2 Problem Statement

The objective of this analysis is to study the Brinell's hardness index value of a new batch of polished and unpolished stones received by Zingaro stone printing company.

3.3 Methodology

Import the libraries - Load the data - Check the structure of the data - Check the types of the data – Check for and treat (if needed) missing values - Check the statistical summary - Check for and treat (if needed) data irregularities – Univariate Analysis – Draft null and alternate hypothesis – Choose correct test type – Test the hypothesis – Interpret the result and provide appropriate conclusion.

Key Points

1. **Data Collection:** A dataset was provided by the data engineering team of Zingaro stone printing company that contains Brinell's hardness index value of polished and unpolished stones of the new batch.
2. **Data cleaning and Pre-processing:** Dataset was checked for missing values, bad data and datatype of columns, no anomalies were found.
3. **Univariate Analysis:** Individual variables were analysed using boxplot and histogram to understand distribution, central tendency and variability of variables.
4. **Visualization Techniques:** We have used histogram and boxplot to understand the data distribution.
5. **Statistical Test:** For this analysis we have used different types of t test.
6. **Tools and Software:** We have carried out the analysis using programming language python on Jupyter notebook. For this analysis Python libraries Numpy, Pandas, Matplotlib, Seaborn and Scipy were used.
7. **Assumptions:** For this analysis we have considered level of significance at 5% (0.05).

3.4 Data Overview

1. **Data Description:** Dataset has 75 rows and 2 columns, both the columns namely Unpolished, Treated and polished are of float type.

```
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)
```

Table 2: Problem 3 Data description

2. **Statistical Summary:**

	count	mean	std	min	25%	50%	75%	max
Unpolished	75.0	134.110527	33.041804	48.406838	115.329753	135.597121	158.215098	200.161313
Treated and Polished	75.0	147.788117	15.587355	107.524167	138.268300	145.721322	157.373318	192.272856

Table3: Problem 3 Statistical Summary

Observations and Insights

Based on the data cleaning and exploration we've conducted, here are some key observations and insights.

- In the current batch number of polished and unpolished stones are equal at 75 each.
- For Treated and Polished stones mean Brinell's hardness index is 147.79 which is significantly higher than that of Unpolished stones at 134.11.
- Mean Brinell's hardness index for both polished and unpolished stones is below the critical level of 150 required for the optimum level of printing of the image.

3.5 Analysis and Findings

3.5.1 Univariate Analysis

Explore all the variables – Observations and Insights

Distribution of Unpolished Stones

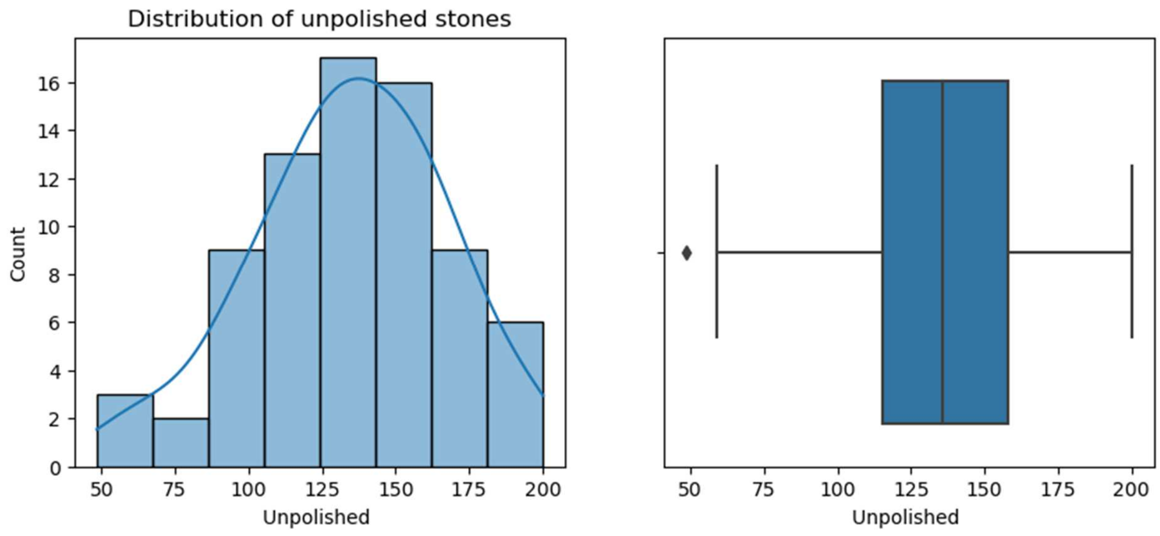

Figure 5: Distribution of Unpolished Stones

Distribution of Polished and Treated Stones

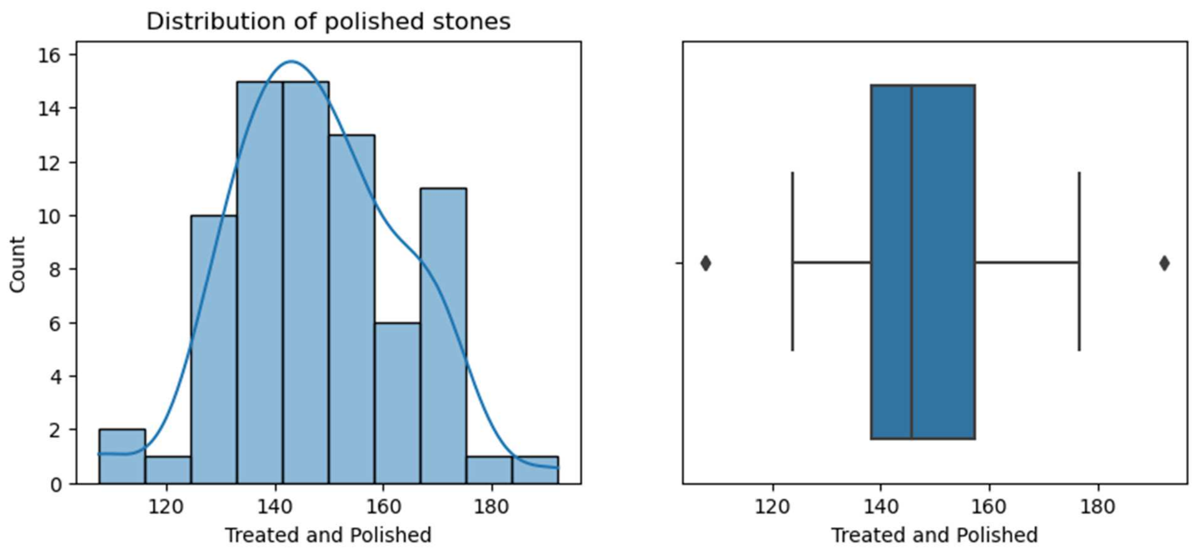

Figure 6: Distribution of Treated and Polished Stones

Observations and Insights

From the above plots of Unpolished stones and Polished stones and Treated stones we can conclude that the data is normally distributed for both.

3.5.2 Zingaro has reason to believe that the unpolished stones may not be suitable for printing.

Hypothesis

Null Hypothesis (H0): Unpolished stones are suitable for printing.

Alternate Hypothesis (Ha): Unpolished stones are not suitable for printing.

Key Values

Level of significance (Alpha): 5% or 0.05

Population Mean: Minimum Brinell's hardness index value required for optimum level of printing = 150

Test Used

One sample t-test

Calculation of t statistics

We calculated t-statistic for One sample t-test using the following equation

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Equation 1: Calculation of t-statistic for one sample t-test

- x bar is sample mean
- Mu is population mean as per null hypothesis
- s is standard deviation
- n is sample size

Using the above-mentioned equation, we calculated t-statistic whose value came as -4.16.

This implies that the mean Brinell's hardness index for unpolished stones is 4.16 units lower than the desired Brinell's hardness index of 150.

Calculation of p-value

The p-value is calculated by referring to the t-distribution table on the basis of

- t-statistics
- Degree of freedom (n-1)

Using the t-distribution for t-statistics of -4.16 and degree of freedom of 74 (75-1) we found the p-value at 0.00008342

Comparison of p-value with level of significance

For any hypothesis test if level of significance (0.05) is less than or equal to p-value then we fail to reject the null hypothesis otherwise we reject the null hypothesis

Here, p-value is 0.00008342 while level of significance is 0.05 from this it is evident that

Level of significance > p-value

Since p-value is significantly smaller than level of significance we can reject the null hypothesis.

Conclusion

Based on the test results above, since the p-value is below the 5% significance level, we reject the null hypothesis. This suggests, with 95% confidence, that Zingaro's consideration that unpolished stones may not be suitable for printing is justified.

3.5.3 Mean hardness of the polished and unpolished stones are same

Hypothesis

Null Hypothesis (H0): Mean hardness of the polished and unpolished stones is same.

Alternate Hypothesis (Ha): Mean hardness of the polished and unpolished stones is not same.

Key Values

Level of significance (Alpha): 5% or 0.05

Test Used

Two-Sample t-Test

Calculation of t statistic

We calculated t-statistic for Independent Two-Sample t-Test using the following equation

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

Equation 2: Calculation of t-statistic for two sample t-test

- x1 bar and x2 bar are means for group1 and group2
- s1 and s2 are standard deviation for group1 and group2
- n1 and n2 are sample size for group1 and group2

Using the above-mentioned equation, we calculated t-statistic whose value came as 3.56.

This indicates a difference of 3.56 units in the Brinell hardness index between polished and treated stones compared to unpolished stones.

Calculation of p-value

The p-value is calculated by referring to the t-distribution table on the basis of

- t-statistic
- Degree of freedom (n1+n2-2)

Using the t-distribution for t-statistic of 3.56 and degree of freedom of 148 (75+75-2) we found the p-value at 0.00065

Comparison of p-value with level of significance

For any hypothesis test if level of significance (0.05) is less than or equal to p-value then we fail to reject the null hypothesis otherwise we reject the null hypothesis

Here, p-value is 0.00065 while level of significance is 0.05 from this it is evident that

Level of significance > p-value

Since p-value is significantly smaller than level of significance we can reject the null hypothesis.

Conclusion

Based on the test results above, since the p-value is below the 5% significance level, we reject the null hypothesis. This suggests, with 95% confidence, that mean hardness of the polished and unpolished stones is not same.

Problem 4

4.1 Context

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

4.2 Problem Statement

The objective is to study the effect of different factors over the hardness of metal implants in dental cavities.

4.3 Methodology

Import the libraries - Load the data - Check the structure of the data - Check the types of the data and convert is (if needed) - Check for and treat (if needed) missing values - Check the statistical summary - Check for and treat (if needed) data irregularities – Create data subsets – Hypothesis testing

Key Points

1. **Data Collection:** A dental implant dataset was provided by data engineering team which consist of attributes pertaining to hardness of metal implants and factors affecting it.
2. **Data Cleaning and Pre-processing:** Dataset was checked for missing values, bad data and data type of columns. All the columns had numeric data, however, except for response column all the others were categorical type so data was converted accordingly.
3. **Data Subset:** The data was segregated based on alloys, resulting in the creation of two new data frames from the original dataset for Alloy 1 and Alloy 2 using the .copy command.
4. **Visualization Technique:** In the report, we utilized box plots to understand the variation in data based on different factors, and interaction plots were employed to examine whether various factors interact with each other.
5. **Tools and Software:** We have carried out the analysis using programming language python on Jupyter notebook. For this analysis Python libraries Numpy, Pandas, Matplotlib, Seaborn, Scipy, Statsmodel and Warnings – to ignore the warning messages were used.
6. **Assumptions:** For this analysis it is assumed that samples are independently drawn.

4.4 Data Overview

1. **Data Description:** Dataset has 90 rows and 5 columns with all having numeric data and int64 as datatype. There were no null or missing values in the data.

```
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)
```

Table 4: Problem 4 Data Overview

2. **Statistical Summary:**

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

Table 5: Problem 4 Statistical Summary

3. **Data Pre-processing:** Four of the five columns have categorical data which has to be converted
   a. **Converting Datatypes:** Data for columns Dentist, Method, Alloy and Temp are categorical in nature, They were converted to category datatype using **.astype**.

```
RangeIndex: 90 entries, 0 to 89
Data columns (total 5 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   Dentist   90 non-null     category
 1   Method    90 non-null     category
 2   Alloy     90 non-null     category
 3   Temp      90 non-null     category
 4   Response  90 non-null     int64
dtypes: category(4), int64(1)
```

Table 6: Problem 4 Updated Data Overview

4. **Data Segregation:** Data has to divided into 2 subsets on the basis alloys. New datasets namely alloy1 and alloy2 were created using groupby and copy.

	count	unique	top	freq	mean	std	min	25%	50%	75%	max
Dentist	45.0	5.0	1.0	9.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Method	45.0	3.0	1.0	15.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Alloy	45.0	1.0	1.0	45.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Temp	45.0	3.0	1500.0	15.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Response	45.0	NaN	NaN	NaN	707.488889	121.194551	289.0	681.0	743.0	782.0	882.0
Dentist_Method	45	15	1:1	3	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Table 7: alloy1 Statistical Summary

	count	unique	top	freq	mean	std	min	25%	50%	75%	max
Dentist	45.0	5.0	1.0	9.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Method	45.0	3.0	1.0	15.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Alloy	45.0	1.0	2.0	45.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Temp	45.0	3.0	1500.0	15.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Response	45.0	NaN	NaN	NaN	776.066667	160.892595	312.0	715.0	824.0	858.0	1115.0

Table 8: alloy2 Statistical Summary

4.5 Hypothesis Testing

Draft null and alternate hypothesis - Choose correct test type - Test the hypothesis - Interpret the result and provide appropriate conclusion.

How does the hardness of implants vary depending on dentists?

Hypothesis

Null hypothesis (H0): Mean hardness of implant remains same across all dentists.

Alternate hypothesis (Ha): For at least one dentist mean hardness of implant varies.

Key Values

Level of significance (Alpha): 5% or 0.05

Test Used

One Way ANOVA Test

Analysis for Alloy 1

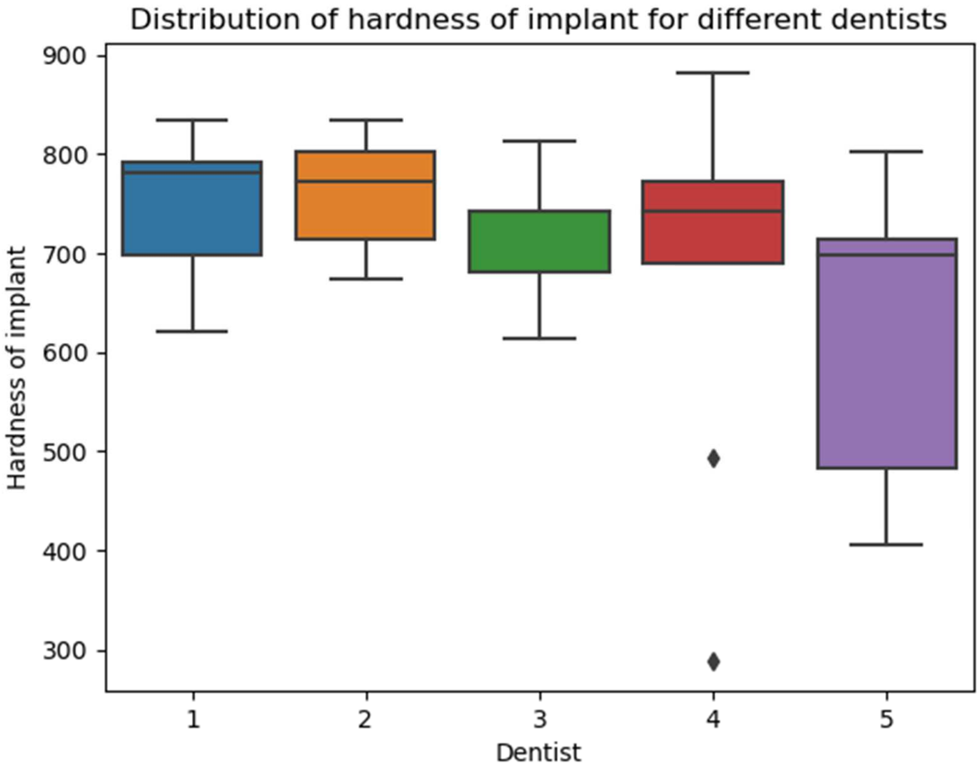

Figure 7: Distribution of hardness of implant for different dentist

Based on the plot above, it's evident that there is some variation in hardness of implant across different dentists. To explore the causes of variation we will conduct ANOVA test.

The ANOVA test relies on certain assumptions that must be met by our data before applying the test.

1. Data has to be normally distributed. We will test using Shapiro-Wilk's test.
2. Population variance is equal for which we will do Levene's test.
3. Samples are randomly and independently drawn.

Shapiro-Wilk's Test

Hypothesis

Null Hypothesis (H0): Data is normally distributed.

Alternate Hypothesis (Ha): Data is not normally distributed.

Shapiro-Wilk's test provide W and p-value where W represents how much data is close to being perfectly normally distributed with 1 representing data is normally distributed and 0 meaning data is not normally distributed.

For alloy1 dataset, W is 0.8304 which means data is 83% normally distributed.

However, p-value is 0.0000119 and if p-value is less than level of significance (0.05), we reject the null hypothesis.

Since in this case p-value is less than level of significance we reject the null hypothesis that data is normally distributed.

Levene's Test

Hypothesis

Null Hypothesis (H0): Variance for all dentists is equal

Alternate Hypothesis (Ha): Variance for at least one dentist is not equal

Levene's test compares variance for all the groups and provides p-value, where if p-value is smaller than level of significance (0.05) then we reject the null hypothesis.

For alloy1, p-value for Levene's test is 0.2566 which is significantly higher than level of significance.

Since in this case p-value is higher than level of significance we fail to reject the null hypothesis that Variance for all dentists is equal.

Based on the above tests alloy1 dataset does not fulfil all the assumptions as it fails Shapiro-Wilks test meaning data is not normally distributed, but still, we will continue with ANOVA test.

One Way ANOVA Test

For one way ANOVA test we need to calculate degree of freedom (dof) and sum of squares

Degree of freedom

Dof = n-1

Where n is the number of observations in a dataset.

This degree of freedom is divided into 2 parts

1. Between-Group Degrees of Freedom (DFbetween):
   DFbetween = k-1
   k = number of groups
2. Within-Group Degrees of Freedom (DFwithin):
   DFwithin = n-k

Sum of squares

Sum of squares can also be called as sum of variances.

Sum of squares are of two types:

1. Between-Group Sum of Squares (SSB):
   It calculates the sum of square of the mean of dataset from each element, depicted by this equation

$$SSB = \sum n_i(\bar{Y_i} - \bar{Y})^2$$

Equation 3: Sum of squares between

   Here
   ni represent number of elements in a group
   Yi bar represent mean of a group
   Y bar represent mean of a dataset

2. Within-Group Sum of Squares (SSW):
   It calculates the sum of squares of the mean of each group from each element in that group, depicted by this equation

$$SSW = \sum\sum(Y_{ij} - \bar{Y_i})^2$$

Equation 4: Sum of squares within

   Here
   Yij represent each element in a group
   Yi bar represents mean of that group

The total of SSB and SSW is called Total sum of squares (SST)

$$SST = SSB + SSW$$

Equation 5: Total sum of squares

Using the above-mentioned equations, we created an ANOVA Table

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

Table 9: ANOVA Table

In this table
C(Dentist) represents between the table
Residual represents within the table
df = degree of freedom
sum_sq = sum of squares (SSB/SSW)
mean_sq = sum_sq / df
F = mean_sq between / mean_sq within
PR(>F) = p-value of getting the f-statistic

Using these values, we can find if variable of interest gets affected by the factors and to what levels and whether the observed result is by chance

Conclusion

From the above table total sum of squares (SST) is 646277.25 and with Dentist as predictor, (106683.69/646277.25) only 16.5% variability is explained by it. Also, p-value of 0.116567 is significantly greater than alpha (0.05).

Thus, to conclude we fail to reject the Null Hypothesis(H0) that mean hardness of implant remains same across all dentists.

Analysis for Alloy 2

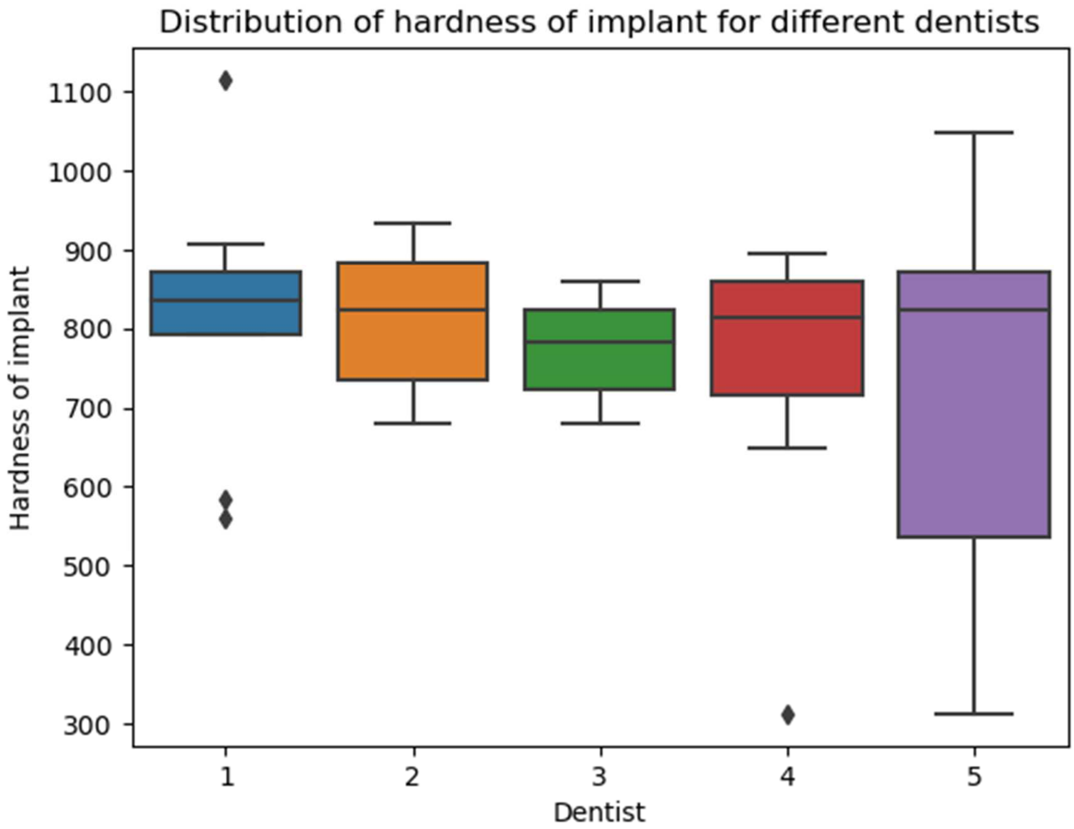

Figure 8: Distribution of hardness of implant for different dentist

Based on the plot above, it's evident that there is some variation in hardness of implant across different dentists. To explore the causes of variation we will conduct ANOVA test.

The ANOVA test relies on certain assumptions that must be met by our data before applying the test.

1. Data has to be normally distributed. We will test using Shapiro-Wilk's test.
2. Population variance is equal for which we will do Levene's test.
3. Samples are randomly and independently drawn.

Shapiro-Wilk's Test

Hypothesis

Null Hypothesis (H0): Data is normally distributed.

Alternate Hypothesis (Ha): Data is not normally distributed.

Shapiro-Wilk's test provide W and p-value where W represents how much data is close to being perfectly normally distributed with 1 representing data is normally distributed and 0 meaning data is not normally distributed.

For alloy2 dataset, W is 0.8877 which means data is 88.77% normally distributed.

However, p-value is 0.000402 and if p-value is less than level of significance (0.05), we reject the null hypothesis.

Since in this case p-value is less than level of significance we reject the null hypothesis that data is normally distributed.

Levene's Test

Hypothesis

Null Hypothesis (H0): Variance for all dentists is equal

Alternate Hypothesis (Ha): Variance for at least one dentist is not equal

Levene's test compares variance for all the groups and provides p-value, where if p-value is smaller than level of significance (0.05) then we reject the null hypothesis.

For alloy2, p-value for Levene's test is 0.2369 which is significantly higher than level of significance.

Since in this case p-value is higher than level of significance we fail to reject the null hypothesis that Variance for all dentists is equal.

Based on the above tests, alloy2 dataset does not fulfil all the assumptions as it fails Shapiro-Wilks test meaning data is not normally distributed, but still, we will continue with ANOVA test.

One Way ANOVA Test

For one way ANOVA test we need to calculate degree of freedom (dof) and sum of squares

Degree of freedom

Dof = n-1

Where n is the number of observations in a dataset.

This degree of freedom is divided into 2 parts

1. Between-Group Degrees of Freedom (DFbetween):
   DFbetween = k-1
   k = number of groups
2. Within-Group Degrees of Freedom (DFwithin):
   DFwithin = n-k

Sum of squares

Sum of squares can also be called as sum of variances.

Sum of squares are of two types:

1. Between-Group Sum of Squares (SSB):
   It calculates the sum of square of the mean of dataset from each element, depicted by this equation

$$SSB = \sum n_i(\bar{Y_i} - \bar{Y})^2$$

Equation 6: Sum of squares between

Here
ni represent number of elements in a group
Yi bar represent mean of a group
Y bar represent mean of a dataset

2. Within-Group Sum of Squares (SSW):
   It calculates the sum of squares of the mean of each group from each element in that group, depicted by this equation

$$SSW = \sum\sum(Y_{ij} - \bar{Y_i})^2$$

Equation 7: Sum of squares within

Here
Yij represent each element in a group
Yi bar represents mean of that group

The total of SSB and SSW is called Total sum of squares (SST)

$$SST = SSB + SSW$$

Equation 8: Total sum of squares

Using the above-mentioned equations, we created an ANOVA Table

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

Table 10: ANOVA Table

In this table
C(Dentist) represents between the table
Residual represents within the table
df = degree of freedom
sum_sq = sum of squares (SSB/SSW)

mean_sq = sum_sq / df
F = mean_sq between / mean_sq within
PR(>F) = p-value of getting the f-statistic

Using these values, we can find if variable of interest gets affected by the factors and to what levels and whether the observed result is by chance

Conclusion

From the above table total sum of squares (SST) is 1139002.91 and with Dentist as predictor, (56797.91/1139002.91) only 4.98% variability is explained by it. Also, p-value of 0.718031 is significantly greater than alpha (0.05).

Thus, to conclude we fail to reject the Null Hypothesis(H0) that mean hardness of implant remains same across all dentists.

How does the hardness of implants vary depending on methods?

Hypothesis

Null hypothesis (H0): Mean hardness of implant remains same across all methods.

Alternate hypothesis (Ha): For at least one method mean hardness of implant varies.

Key Values

Level of significance (Alpha): 5% or 0.05

Test Used

One Way ANOVA Test

Analysis for Alloy 1

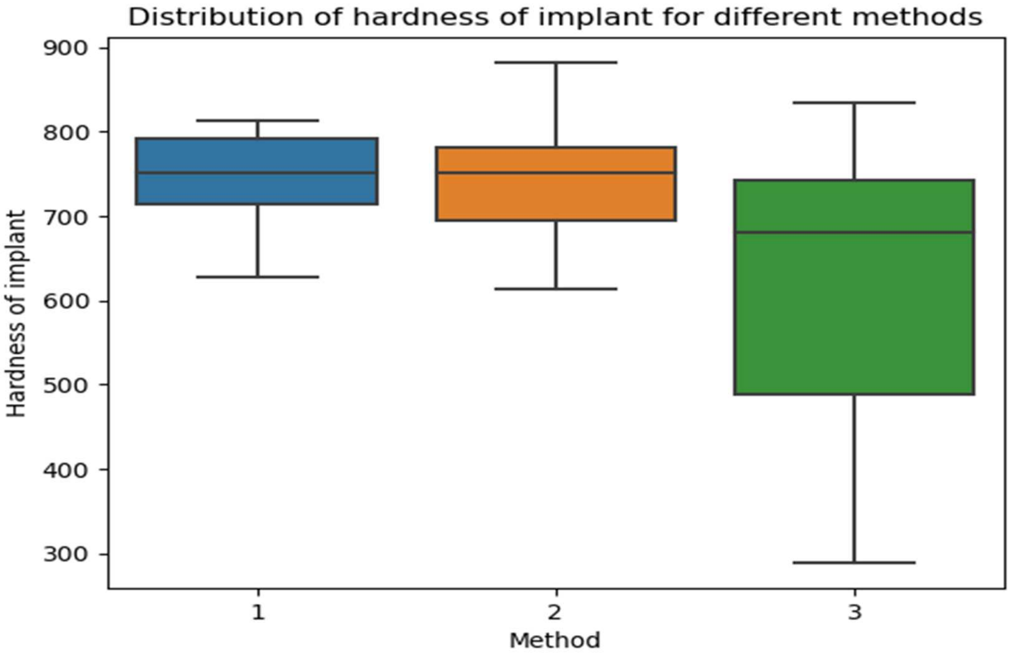

Figure 9: Distribution of hardness of implant for different Methods

Based on the plot above, it's evident that there is some variation in hardness of implant across different dentists. To explore the causes of variation we will conduct ANOVA test.

The ANOVA test relies on certain assumptions that must be met by our data before applying the test.

1. Data has to be normally distributed. We will test using Shapiro-Wilk's test.
2. Population variance is equal for which we will do Levene's test.
3. Samples are randomly and independently drawn.

Shapiro-Wilk's Test

Hypothesis

Null Hypothesis (H0): Data is normally distributed.

Alternate Hypothesis (Ha): Data is not normally distributed.

Shapiro-Wilk's test provide W and p-value where W represents how much data is close to being perfectly normally distributed with 1 representing data is normally distributed and 0 meaning data is not normally distributed.

For alloy1 dataset, W is 0.8304 which means data is 83% normally distributed.

However, p-value is 0.0000119 and if p-value is less than level of significance (0.05), we reject the null hypothesis.

Since in this case p-value is less than level of significance we reject the null hypothesis that data is normally distributed.

Levene's Test

Hypothesis

Null Hypothesis (H0): Variance for all methods is equal

Alternate Hypothesis (Ha): Variance for at least one method is not equal

Levene's test compares variance for all the groups and provides p-value, where if p-value is smaller than level of significance (0.05) then we reject the null hypothesis.

For alloy1, p-value for Levene's test is 0.0034 which is significantly smaller than level of significance.

Since in this case p-value is smaller than level of significance we reject the null hypothesis that Variance for all methods is equal.

Based on the above tests alloy1 dataset does not fulfil all the assumptions as it fails Shapiro-Wilks test and Levene's test meaning data is not normally distributed and variance for at least one method is not equal, but still, we will continue with ANOVA test.

One Way ANOVA Test

For one way ANOVA test we need to calculate degree of freedom (dof) and sum of squares

Degree of freedom

Dof = n-1

Where n is the number of observations in a dataset.

This degree of freedom is divided into 2 parts

2. Between-Group Degrees of Freedom (DFbetween):
   DFbetween = k-1
   k = number of groups
3. Within-Group Degrees of Freedom (DFwithin):
   DFwithin = n-k

Sum of squares

Sum of squares can also be called as sum of variances.

Sum of squares are of two types:

3. Between-Group Sum of Squares (SSB):
   It calculates the sum of square of the mean of dataset from each element, depicted by this equation

$$SSB = \sum n_i(\bar{Y_i} - \bar{Y})^2$$

Equation 9: Sum of squares between

Here
ni represent number of elements in a group
Yi bar represent mean of a group
Y bar represent mean of a dataset

4. Within-Group Sum of Squares (SSW):
It calculates the sum of squares of the mean of each group from each element in that group, depicted by this equation

$$SSW = \sum\sum(Y_{ij} - \bar{Y_i})^2$$

Equation 10: Sum of squares within

Here
Yij represent each element in a group
Yi bar represents mean of that group

The total of SSB and SSW is called Total sum of squares (SST)

$$SST = SSB + SSW$$

Equation 11: Total sum of squares

Using the above-mentioned equations, we created an ANOVA Table

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

Table 11: ANOVA Table

In this table
C(Method) represents between the table
Residual represents within the table
df = degree of freedom
sum_sq = sum of squares (SSB/SSW)
mean_sq = sum_sq / df
F = mean_sq between / mean_sq within
PR(>F) = p-value of getting the f-statistic

Using these values, we can find if variable of interest gets affected by the factors and to what levels and whether the observed result is by chance

Inference

- From the above table total sum of squares (SST) is 646277.25 and with Method as predictor, (148472.18/646277.25) 22.97% variability is explained by it.
- p-value of 0.004163 is significantly smaller than alpha (0.05).

Thus, to conclude we reject the Null Hypothesis(H0) and with 95% confidence can say that for at least one method mean hardness of implant varies.

Based on the ANOVA test results, we have determined that the means of at least one group is not equal. To recognize which specific groups differ from one another, we will employ a multiple comparison test.

Tukey HSD Test

Tukey HSD test identifies for which pair the mean is significantly different It provides confidence intervals for the differences in means between pairs of groups. If the confidence interval for a pair of groups does not include zero, it indicates that the mean difference is statistically significant, meaning that the means of those groups are significantly different from each other.

Using the above-mentioned explanation, we created a Table

Multiple Comparison of Means - Tukey HSD, FWER=0.05

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

Table 12: Tukey HSD Table

Here
FWER = level of significance
Group1, group2 = group pairs
Meandiff = average difference between pair of groups
p-adj = p-value
lower = lower point of confidence interval
upper = upper point of confidence interval
reject = reject null hypothesis

In the above table for pair of Method1 and Method2, zero lies in confidence interval thus reject is False. However, for pairs Method1 and Method3, Method2 and Method3 zero does not lie in confidence interval thus reject is True.

Based on the results of Tukey test:

- There is no statistically significant difference between means of response for Method1 and Method2.
- There are statistically significant differences between means of response for Method1 and Method3, as well as between Method2 and Method3.

Conclusion

From the above tests we can conclude that:

- With 95% confidence, mean of at least one method is significantly different from others.
- Method as a predictor explains 22.97% of variability.
- On comparing the means, we realised that Method3 shows statistically significant differences from Method1 and Method2.This suggests that the variation observed in the hardness of implant can be attributed to this difference between the Methods.

Analysis for Alloy 2

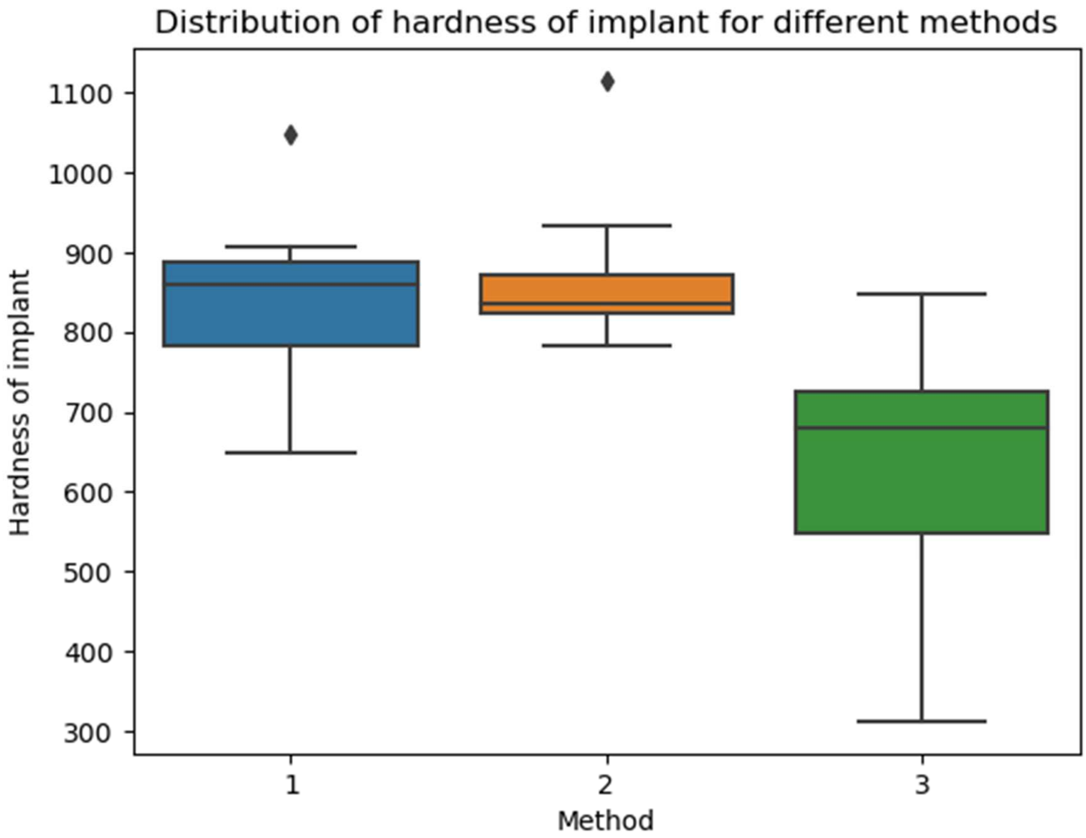

Figure 10: Distribution of hardness of implant for different Methods

Based on the plot above, it's evident that there is some variation in hardness of implant across different dentists. To explore the causes of variation we will conduct ANOVA test.

The ANOVA test relies on certain assumptions that must be met by our data before applying the test.

1. Data has to be normally distributed. We will test using Shapiro-Wilk's test.
2. Population variance is equal for which we will do Levene's test.
3. Samples are randomly and independently drawn.

Shapiro-Wilk's Test

Hypothesis

Null Hypothesis (H0): Data is normally distributed.

Alternate Hypothesis (Ha): Data is not normally distributed.

Shapiro-Wilk's test provide W and p-value where W represents how much data is close to being perfectly normally distributed with 1 representing data is normally distributed and 0 meaning data is not normally distributed.

For alloy2 dataset, W is 0.8877 which means data is 88.77% normally distributed.

However, p-value is 0.000402 and if p-value is less than level of significance (0.05), we reject the null hypothesis.

Since in this case p-value is less than level of significance we reject the null hypothesis that data is normally distributed.

Levene's Test

Hypothesis

Null Hypothesis (H0): Variance for all methods is equal

Alternate Hypothesis (Ha): Variance for at least one method is not equal

Levene's test compares variance for all the groups and provides p-value, where if p-value is smaller than level of significance (0.05) then we reject the null hypothesis.

For alloy2, p-value for Levene's test is 0.0447 which is less than level of significance.

Since in this case p-value is smaller than level of significance we reject the null hypothesis that Variance for all methods is equal.

Based on the above tests, alloy2 dataset does not fulfil all the assumptions as it fails Shapiro-Wilks test and Levene's test meaning data is not normally distributed and variance for at least one method is not equal, but still, we will continue with ANOVA test.

One Way ANOVA Test

For one way ANOVA test we need to calculate degree of freedom (dof) and sum of squares

Degree of freedom

Dof = n-1

Where n is the number of observations in a dataset.

This degree of freedom is divided into 2 parts

4. **4.** Between-Group Degrees of Freedom (DFbetween):
   DFbetween = k-1
   k = number of groups
5. **5.** Within-Group Degrees of Freedom (DFwithin):
   DFwithin = n-k

Sum of squares

Sum of squares can also be called as sum of variances.

Sum of squares are of two types:

4. **4.** Between-Group Sum of Squares (SSB):
   It calculates the sum of square of the mean of dataset from each element, depicted by this equation

$$SSB = \sum n_i(\bar{Y_i} - \bar{Y})^2$$

Equation 12: Sum of squares between

Here
ni represent number of elements in a group
Yi bar represent mean of a group
Y bar represent mean of a dataset

5. **5.** Within-Group Sum of Squares (SSW):
   It calculates the sum of squares of the mean of each group from each element in that group, depicted by this equation

$$SSW = \sum\sum(Y_{ij} - \bar{Y_i})^2$$

Equation 13: Sum of squares within

Here
Yij represent each element in a group
Yi bar represents mean of that group

The total of SSB and SSW is called Total sum of squares (SST)

$$SST = SSB + SSW$$

Equation 14: Total sum of squares

Using the above-mentioned equations, we created an ANOVA Table

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

Table 13: ANOVA Table

In this table
C(Method) represents between the table
Residual represents within the table
df = degree of freedom
sum_sq = sum of squares (SSB/SSW)
mean_sq = sum_sq / df
F = mean_sq between / mean_sq within
PR(>F) = p-value of getting the f-statistic

Using these values, we can find if variable of interest gets affected by the factors and to what levels and whether the observed result is by chance

Inference

- From the above table total sum of squares (SST) is 1139002.8 and with Method as predictor, (499640.4/1139002.8) 43.87% variability is explained by it.
- p-value of 0.000005 is significantly smaller than alpha (0.05).

Thus, to conclude we reject the Null Hypothesis(H0) and with 95% confidence can say that for at least one method mean hardness of implant varies.

Based on the ANOVA test results, we have determined that the means of at least one group is not equal. To recognize which specific groups differ from one another, we will employ a multiple comparison test.

Tukey HSD Test

Tukey HSD test identifies for which pair the mean is significantly different It provides confidence intervals for the differences in means between pairs of groups. If the confidence interval for a pair of groups does not include zero, it indicates that the mean difference is statistically significant, meaning that the means of those groups are significantly different from each other.

Using the above-mentioned explanation, we created a Table

Multiple Comparison of Means - Tukey HSD, FWER=0.05

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

Table 14: Tukey HSD Table

Here
FWER = level of significance
group1, group2 = group pairs
meandiff = average difference between pair of groups
p-adj = p-value
lower = lower point of confidence interval
upper = upper point of confidence interval
reject = reject null hypothesis

In the above table for pair of Method1 and Method2, zero lies in confidence interval thus reject is False. However, for pairs Method1 and Method3, Method2 and Method3 zero does not lie in confidence interval thus reject is True.

Based on the results of Tukey test:

- There is no statistically significant difference between means of response for Method1 and Method2.
- There are statistically significant differences between means of response for Method1 and Method3, as well as between Method2 and Method3.

Conclusion

From the above tests we can conclude that:

- With 95% confidence, mean of at least one method is significantly different from others.
- Method as a predictor explains 43.87% of variability.
- On comparing the means, we realised that Method3 shows statistically significant differences from Method1 and Method2.This suggests that the variation observed in the hardness of implant can be attributed to this difference between the Methods.

Interaction effect between the dentist and method on the hardness of dental implants for each type of alloy

When multiple factors influence a response, it's highly probable that the effect of one factor can influence the effect of another factor on the response. This phenomenon, where one factor affects the influence of another factor, is referred to as interaction.

To study this interaction between dentist and method we will create an interaction plot for both Alloy1 and Alloy2

Interaction plot for Alloy 1

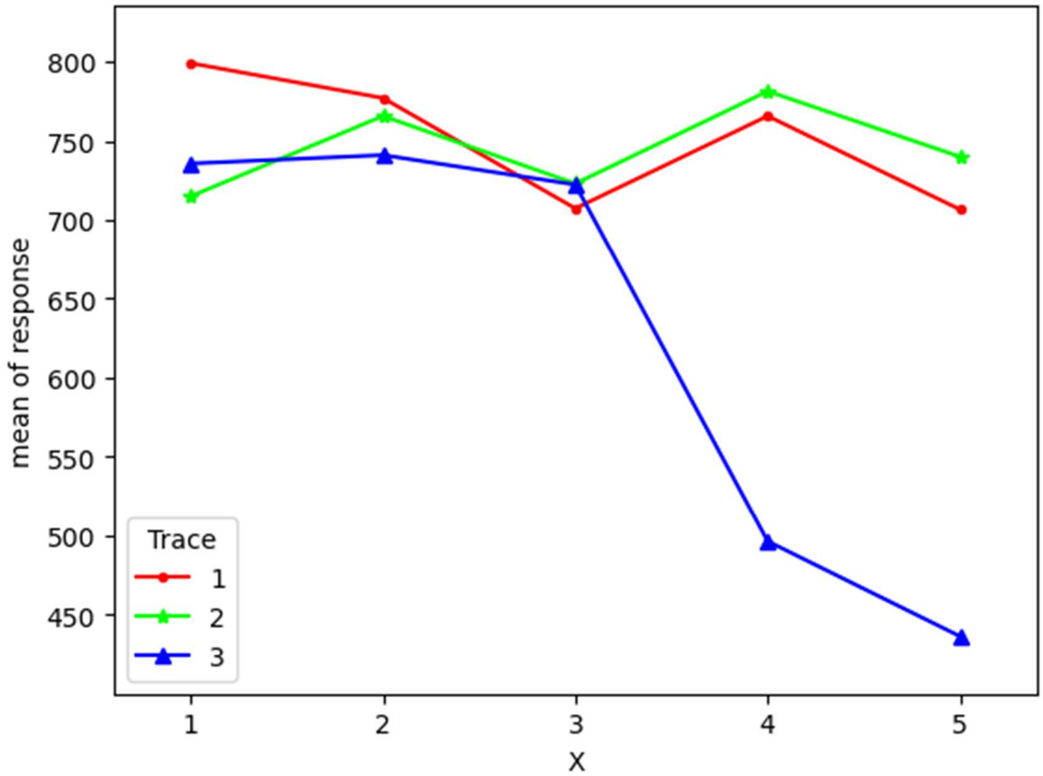

Figure 11: Interaction plot

Inference

In the above plot all three lines representing different methods interact with each other meaning that dentist and method together influence the hardness of implant.

Mean response values remain within 700 and 800 range for method1 and method2, however, in case of method3 values remain in 700 and 800 range for first 3 dentists but for dentist 4 and 5 there is a steep decline to 450.

Interaction plot for Alloy 2

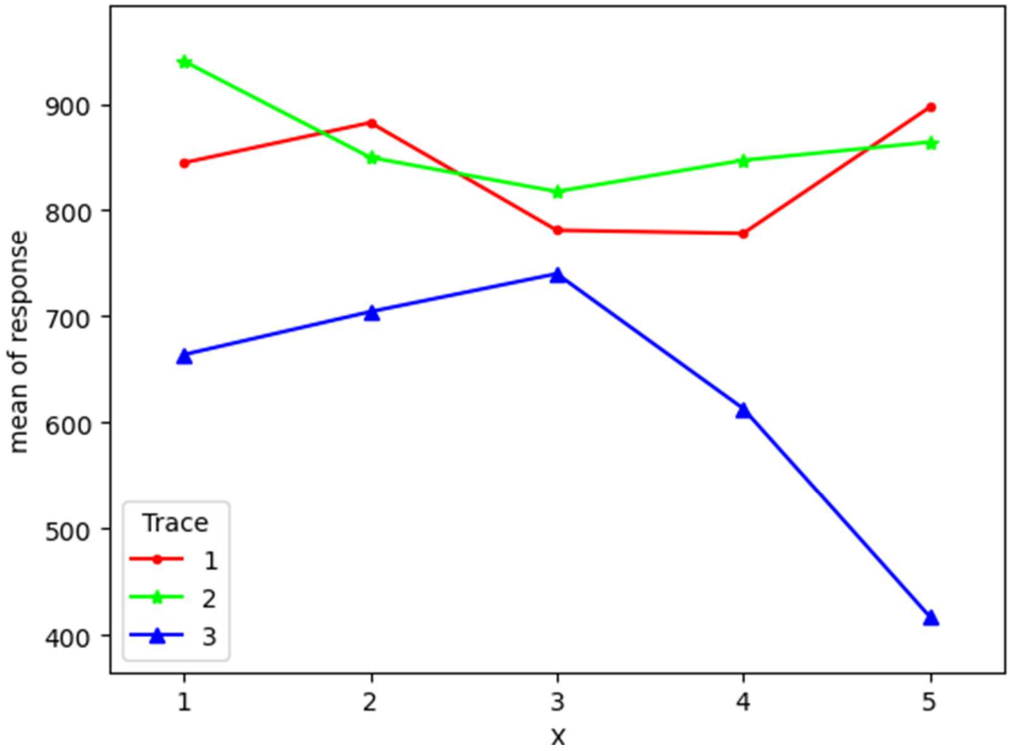

Figure 12: Interaction plot

Inference

There is some interaction for hardness of dental implant for different dentists based on different methods for Alloy2 meaning that dentist and method could together influence the hardness of implant.

Mean response values remain within 750 and 950 range for method1 and method2, however, in case of method3 values are lower hovering between 650 and 700 first 3 dentists and then there is a steep decline for dentist 4 and 5.

From the above plots we can see that there is some sort of interaction for hardness of dental implant for different dentists based on methods to study this influence we will require to carry out Two-way ANOVA test.

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

Here we will study how dentists and methods interact with each other.

Hypothesis

Null hypothesis (H0): Mean hardness of implant remains same for a dentist irrespective of the method involved. Similarly, mean hardness of implant for a method remains same irrespective of the dentist who performs the implant.

Alternate hypothesis (Ha): Mean hardness of implant varies for a dentist depending upon the method involved. Similarly, mean hardness of implant varies for a method depending upon the dentist who performs the implant.

Key Values

Level of significance (Alpha): 5% or 0.05

Test Used

Two Way ANOVA Test

Analysis for Alloy 1

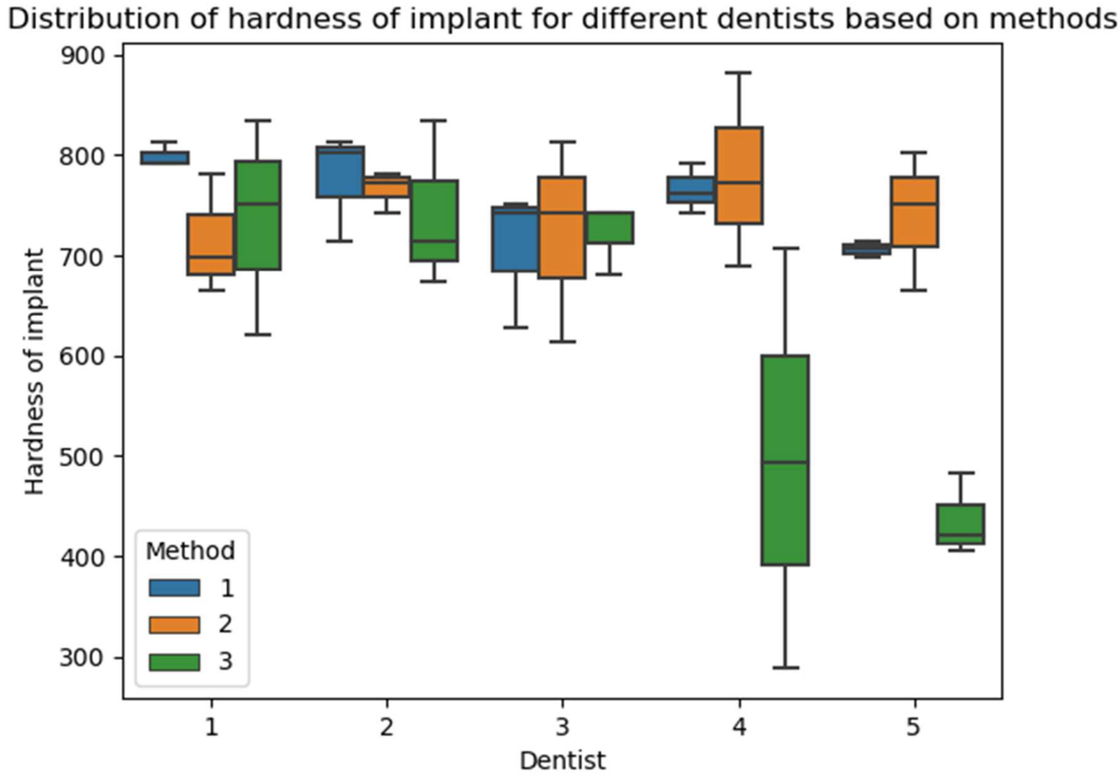

Figure 13: Distribution of hardness of implant for different dentists based on methods

The boxplot shows hardness of implant for different groups of Dentist and Method, range for boxes clearly vary for different groups which means there is some interaction between Dentist and Method which we will test using Two Way ANOVA test.

The ANOVA test relies on certain assumptions that must be met by our data before applying the test.

1. Data has to be normally distributed. We will test using Shapiro-Wilk's test.
2. Population variance is equal for which we will do Levene's test.
3. Samples are randomly and independently drawn.

Shapiro-Wilk's Test

We have already done this test twice earlier when doing One Way ANOVA and found that p-value remained same both times. p-value is significantly smaller than level of significance (0.05), thus we have rejected the Null Hypothesis(H0) and with 95% confidence stated that data is not normally distributed.

Levene's Test

Hypothesis

Null Hypothesis (H0): Variance for all groups is equal

Alternate Hypothesis (Ha): Variance for at least one group is not equal

Levene's test compares variance for all the groups and provides p-value, where if p-value is smaller than level of significance (0.05) then we reject the null hypothesis.

For alloy1, p-value for Levene's test is 0.3128 which is significantly greater than level of significance.

Since in this case p-value is greater than level of significance we fail to reject the null hypothesis that Variance for all groups is equal.

Based on the above tests alloy1 dataset does not fulfil all the assumptions as it fails Shapiro-Wilks test meaning data is not normally distributed, but still, we will continue with ANOVA test.

Two Way ANOVA Test

For two way ANOVA test we need to calculate degree of freedom (dof) and sum of squares

Degree of freedom

Dof = n-1

Where n is the number of observations in a dataset.

This degree of freedom is divided into 3 parts

1. Between-Factor1 Degrees of Freedom (DFFactor1):
   DFFactor1 = k-1
   k = number of groups in factor1
2. Between-Factor2 Degrees of Freedom (DFFactor2)
   DFFactor2 = l-1
   l = number of groups in factor2

3. Interaction Degrees of Freedom (DFInteraction)
   DFInteraction = (k-1)*(l-1)
4. Within-Group Degrees of Freedom (DFwithin):
   DFwithin = n-(k*l)

Sum of squares

Sum of squares can also be called as sum of variances.

Sum of squares are of two types:

1. Between-Factor 1 Sum of Squares (SSFactor1):
   It calculates the sum of squares of difference between mean of dataset from each element of factor1, depicted by this equation

$$SSFactor1 = \sum n_{i..}(\bar{Y}_{i..} - \bar{Y})^2$$

Equation 15: Sum of squares between factor1

Here
ni represent number of elements in a group of factor1
Yi bar represent mean of a group of factor1
Y bar represent mean of a dataset

2. Between-Factor 2 Sum of Squares (SSFactor2):
   It calculates the sum of squares of difference between mean of dataset from each element of factor2, depicted by this equation

$$SSFactor2 = \sum n_{.j.}(\bar{Y}_{.j.} - \bar{Y})^2$$

Equation 16: Sum of squares between factor2

Here
nj represent number of elements in a group of factor2
Yj bar represent mean of a group of factor2
Y bar represent mean of a dataset

3. Interaction Sum of Squares (SSInteraction)

$$SSInteraction = \sum\sum(Y_{ijk} - \bar{Y}_{ij})^2$$

Equation 17: Sum of squares Interaction

Here
Yijk = elements in group based on ith group of factor1 and jth group of factor2
Yij bar = mean of group based on ith group of factor1 and jth group of factor2

4. Within-Group Sum of Squares (SSW):

It calculates the sum of squares of the mean of each group from each element in that group, depicted by this equation

$$SSWithin = \sum\sum\sum(Y_{ijk} - \bar{Y_{ij}})^2$$

Equation 18: Sum of squares within

Here
Yijk = elements in group based on ith group of factor1 and jth group of factor2
Yij bar = mean of group based on ith group of factor1 and jth group of factor2

The total of all the above elements is called Total sum of squares (SST)

$$SST = SS_{Factor1} + SS_{Factor2} + SS_{Interaction} + SS_{Within}$$

Equation 19: Total sum of squares

Using the above-mentioned equations, we created an ANOVA Table

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

Table 15: ANOVA Table

In this table
C(Dentist) represents factor1
C(Method) represents factor2
C(Dentist): C(Method) represents interaction
Residual represents within the table
df = degree of freedom
sum_sq = sum of squares
mean_sq = sum_sq / df
F = mean_sq between / mean_sq within
PR(>F) = p-value of getting the f-statistic

Using these values, we can find if variable of interest gets affected by the new factor introduced in the form of interaction between factor1 and factor2 and to what level this factor influences the variable of interest.

Inference

- From One Way ANOVA Test, the total sum of squares (SST) is same at 646277.25. With only Dentist as predictor 16.5% variability is explained by it. With only Method as predictor 22.97% variability is explained by it. However, when both the predictors are in the model over two third, (441097.25/646277.25) 68.25% variability is explained by both main effects and their interaction effects.

- the p-value for all three cases is significantly below the level of significance (0.05).

Thus, to conclude we reject the Null Hypothesis(H0) which states that mean hardness of implant remains same for a dentist irrespective of the method involved. Similarly, mean hardness of implant for a method remains same irrespective of the dentist who performs the implant.

Based on the ANOVA test results, we have determined that the means of at least one group is not equal. To recognize which specific groups differ from one another, we will employ a multiple comparison test.

Tukey HSD Test

Tukey HSD test identifies for which pair the mean is significantly different It provides confidence intervals for the differences in means between pairs of groups. If the confidence interval for a pair of groups does not include zero, it indicates that the mean difference is statistically significant, meaning that the means of those groups are significantly different from each other.

To perform the Tukey HSD test based on Two Way ANOVA test a new column has to added to the data frame which will be based on concatenation of factor1 and factor2 columns separated by separator. This column will act as a group in the Tukey HSD test

We created a new column in alloy1 data frame with Dentist_Method name, based on which we will group our data for Tukey HSD test.

Using the above-mentioned explanation, we created a Table

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
1:1	1:2	-84.0	0.9933	-332.8283	164.8283	False
1:1	1:3	-63.3333	0.9996	-312.1617	185.495	False
1:1	2:1	-22.0	1.0	-270.8283	226.8283	False
1:1	2:2	-33.3333	1.0	-282.1617	215.495	False
1:1	2:3	-58.0	0.9999	-306.8283	190.8283	False
1:1	3:1	-91.6667	0.9853	-340.495	157.1617	False
1:1	3:2	-76.0	0.9975	-324.8283	172.8283	False
1:1	3:3	-76.6667	0.9972	-325.495	172.1617	False
1:1	4:1	-33.3333	1.0	-282.1617	215.495	False
1:1	4:2	-17.6667	1.0	-266.495	231.1617	False
1:1	4:3	-302.6667	0.007	-551.495	-53.8383	True
1:1	5:1	-92.3333	0.9844	-341.1617	156.495	False
1:1	5:2	-59.0	0.9998	-307.8283	189.8283	False
1:1	5:3	-362.6667	0.0007	-611.495	-113.8383	True
1:2	1:3	20.6667	1.0	-228.1617	269.495	False
1:2	2:1	62.0	0.9997	-186.8283	310.8283	False
1:2	2:2	50.6667	1.0	-198.1617	299.495	False
1:2	2:3	26.0	1.0	-222.8283	274.8283	False
1:2	3:1	-7.6667	1.0	-256.495	241.1617	False

1:2	3:2	8.0	1.0	-240.8283	256.8283	False
1:2	3:3	7.3333	1.0	-241.495	256.1617	False
1:2	4:1	50.6667	1.0	-198.1617	299.495	False
1:2	4:2	66.3333	0.9994	-182.495	315.1617	False
1:2	4:3	-218.6667	0.1324	-467.495	30.1617	False
1:2	5:1	-8.3333	1.0	-257.1617	240.495	False
1:2	5:2	25.0	1.0	-223.8283	273.8283	False
1:2	5:3	-278.6667	0.0173	-527.495	-29.8383	True
1:3	2:1	41.3333	1.0	-207.495	290.1617	False
1:3	2:2	30.0	1.0	-218.8283	278.8283	False
1:3	2:3	5.3333	1.0	-243.495	254.1617	False
1:3	3:1	-28.3333	1.0	-277.1617	220.495	False
1:3	3:2	-12.6667	1.0	-261.495	236.1617	False
1:3	3:3	-13.3333	1.0	-262.1617	235.495	False
1:3	4:1	30.0	1.0	-218.8283	278.8283	False
1:3	4:2	45.6667	1.0	-203.1617	294.495	False
1:3	4:3	-239.3333	0.0688	-488.1617	9.495	False
1:3	5:1	-29.0	1.0	-277.8283	219.8283	False
1:3	5:2	4.3333	1.0	-244.495	253.1617	False
1:3	5:3	-299.3333	0.0079	-548.1617	-50.505	True
2:1	2:2	-11.3333	1.0	-260.1617	237.495	False
2:1	2:3	-36.0	1.0	-284.8283	212.8283	False
2:1	3:1	-69.6667	0.999	-318.495	179.1617	False
2:1	3:2	-54.0	0.9999	-302.8283	194.8283	False
2:1	3:3	-54.6667	0.9999	-303.495	194.1617	False
2:1	4:1	-11.3333	1.0	-260.1617	237.495	False
2:1	4:2	4.3333	1.0	-244.495	253.1617	False
2:1	4:3	-280.6667	0.016	-529.495	-31.8383	True
2:1	5:1	-70.3333	0.9989	-319.1617	178.495	False
2:1	5:2	-37.0	1.0	-285.8283	211.8283	False
2:1	5:3	-340.6667	0.0016	-589.495	-91.8383	True
2:2	2:3	-24.6667	1.0	-273.495	224.1617	False
2:2	3:1	-58.3333	0.9999	-307.1617	190.495	False
2:2	3:2	-42.6667	1.0	-291.495	206.1617	False
2:2	3:3	-43.3333	1.0	-292.1617	205.495	False
2:2	4:1	0.0	1.0	-248.8283	248.8283	False
2:2	4:2	15.6667	1.0	-233.1617	264.495	False
2:2	4:3	-269.3333	0.0243	-518.1617	-20.505	True
2:2	5:1	-59.0	0.9998	-307.8283	189.8283	False
2:2	5:2	-25.6667	1.0	-274.495	223.1617	False
2:2	5:3	-329.3333	0.0025	-578.1617	-80.505	True
2:3	3:1	-33.6667	1.0	-282.495	215.1617	False
2:3	3:2	-18.0	1.0	-266.8283	230.8283	False
2:3	3:3	-18.6667	1.0	-267.495	230.1617	False
2:3	4:1	24.6667	1.0	-224.1617	273.495	False
2:3	4:2	40.3333	1.0	-208.495	289.1617	False
2:3	4:3	-244.6667	0.0576	-493.495	4.1617	False
2:3	5:1	-34.3333	1.0	-283.1617	214.495	False
2:3	5:2	-1.0	1.0	-249.8283	247.8283	False
2:3	5:3	-304.6667	0.0065	-553.495	-55.8383	True
3:1	3:2	15.6667	1.0	-233.1617	264.495	False
3:1	3:3	15.0	1.0	-233.8283	263.8283	False
3:1	4:1	58.3333	0.9999	-190.495	307.1617	False
3:1	4:2	74.0	0.9981	-174.8283	322.8283	False

3:1	4:3	-211.0	0.166	-459.8283	37.8283	False
3:1	5:1	-0.6667	1.0	-249.495	248.1617	False
3:1	5:2	32.6667	1.0	-216.1617	281.495	False
3:1	5:3	-271.0	0.0229	-519.8283	-22.1717	True
3:2	3:3	-0.6667	1.0	-249.495	248.1617	False
3:2	4:1	42.6667	1.0	-206.1617	291.495	False
3:2	4:2	58.3333	0.9999	-190.495	307.1617	False
3:2	4:3	-226.6667	0.1035	-475.495	22.1617	False
3:2	5:1	-16.3333	1.0	-265.1617	232.495	False
3:2	5:2	17.0	1.0	-231.8283	265.8283	False
3:2	5:3	-286.6667	0.0128	-535.495	-37.8383	True
3:3	4:1	43.3333	1.0	-205.495	292.1617	False
3:3	4:2	59.0	0.9998	-189.8283	307.8283	False
3:3	4:3	-226.0	0.1057	-474.8283	22.8283	False
3:3	5:1	-15.6667	1.0	-264.495	233.1617	False
3:3	5:2	17.6667	1.0	-231.1617	266.495	False
3:3	5:3	-286.0	0.0131	-534.8283	-37.1717	True
4:1	4:2	15.6667	1.0	-233.1617	264.495	False
4:1	4:3	-269.3333	0.0243	-518.1617	-20.505	True
4:1	5:1	-59.0	0.9998	-307.8283	189.8283	False
4:1	5:2	-25.6667	1.0	-274.495	223.1617	False
4:1	5:3	-329.3333	0.0025	-578.1617	-80.505	True
4:2	4:3	-285.0	0.0137	-533.8283	-36.1717	True
4:2	5:1	-74.6667	0.9979	-323.495	174.1617	False
4:2	5:2	-41.3333	1.0	-290.1617	207.495	False
4:2	5:3	-345.0	0.0013	-593.8283	-96.1717	True
4:3	5:1	210.3333	0.1692	-38.495	459.1617	False
4:3	5:2	243.6667	0.0596	-5.1617	492.495	False
4:3	5:3	-60.0	0.9998	-308.8283	188.8283	False
5:1	5:2	33.3333	1.0	-215.495	282.1617	False
5:1	5:3	-270.3333	0.0234	-519.1617	-21.505	True
5:2	5:3	-303.6667	0.0067	-552.495	-54.8383	True

Table 16: Tukey HSD Table

Here
FWER = level of significance
Group1, group2 = group pairs
Meandiff = average difference between pair of groups
p-adj = p-value
lower = lower point of confidence interval
upper = upper point of confidence interval
reject = reject null hypothesis

Method 3 used by Dentist 5 has significantly different mean from almost all the methods used by all the dentists. Also, Method 3 used by Dentist 4 is significantly different from Method 1 used by Dentists 1,2 and 4 respectively as well as Method 2 used by Dentist 2 and 4.

Conclusion

From the above tests we can conclude that:

- It is observed that variance in hardness of implant is significantly impacted by both Dentist and Method along with their interaction effect for Alloy1.

- And thus, to conclude with 95% confidence we can say that mean hardness of implant varies for a dentist depending upon the method involved. Similarly, mean hardness of implant varies for a method depending upon the dentist who performs the implant meaning that there is interaction effect between Dentist and Method.

Analysis for Alloy 2

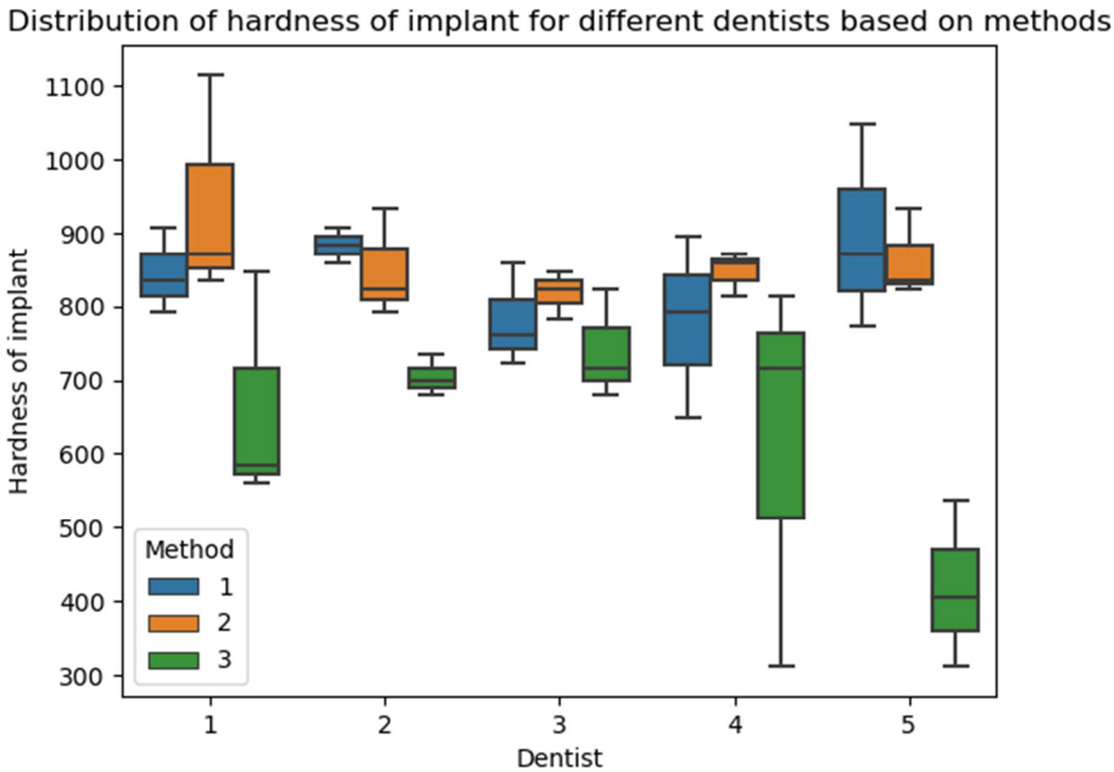

Figure 14: Distribution of hardness of implant for different dentists based on methods

The boxplot shows hardness of implant for different groups of Dentist and Method, range for boxes clearly vary for different groups which means there is some interaction between Dentist and Method which we will test using Two Way ANOVA test.

The ANOVA test relies on certain assumptions that must be met by our data before applying the test.

1. Data has to be normally distributed. We will test using Shapiro-Wilk's test.
2. Population variance is equal for which we will do Levene's test.
3. Samples are randomly and independently drawn.

Shapiro-Wilk's Test

We have already done this test twice earlier when doing One Way ANOVA and found that p-value remained same both times. p-value is significantly smaller than level of significance (0.05), thus we have rejected the Null Hypothesis(H0) and with 95% confidence stated that data is not normally distributed.

Levene's Test

Hypothesis

Null Hypothesis (H0): Variance for all groups is equal

Alternate Hypothesis (Ha): Variance for at least one group is not equal

Levene's test compares variance for all the groups and provides p-value, where if p-value is smaller than level of significance (0.05) then we reject the null hypothesis.

For alloy2, p-value for Levene's test is 0.7832 which is significantly greater than level of significance.

Since in this case p-value is greater than level of significance we fail to reject the null hypothesis that Variance for all groups is equal.

Based on the above tests alloy2 dataset does not fulfil all the assumptions as it fails Shapiro-Wilks test meaning data is not normally distributed, but still, we will continue with ANOVA test.

Two Way ANOVA Test

For two-way ANOVA test we need to calculate degree of freedom (dof) and sum of squares

Degree of freedom

Dof = n-1

Where n is the number of observations in a dataset.

This degree of freedom is divided into 3 parts

1. Between-Factor1 Degrees of Freedom (DFFactor1):
   DFFactor1 = k-1
   k = number of groups in factor1
2. Between-Factor2 Degrees of Freedom (DFFactor2)
   DFFactor2 = l-1
   l = number of groups in factor2
3. Interaction Degrees of Freedom (DFInteraction)
   DFInteraction = (k-1)*(l-1)
4. Within-Group Degrees of Freedom (DFwithin):
   DFwithin = n-(k*l)

Sum of squares

Sum of squares can also be called as sum of variances.

Sum of squares are of two types:

1. Between-Factor 1 Sum of Squares (SSFactor1):
   It calculates the sum of squares of difference between mean of dataset from each element of factor1, depicted by this equation

$$SSFactor1 = \sum n_{i..}(\bar{Y}_{i..} - \bar{Y})^2$$

Equation 20: Sum of squares between factor1

Here
ni represent number of elements in a group of factor1
Yi bar represent mean of a group of factor1
Y bar represent mean of a dataset

2. Between-Factor 2 Sum of Squares (SSFactor2):
   It calculates the sum of squares of difference between mean of dataset from each element of factor2, depicted by this equation

$$SSFactor2 = \sum n_{.j.}(\bar{Y}_{.j.} - \bar{Y})^2$$

Equation 21: Sum of squares between factor2

Here
nj represent number of elements in a group of factor2
Yj bar represent mean of a group of factor2
Y bar represent mean of a dataset

3. Interaction Sum of Squares (SSInteraction)

$$SSInteraction = \sum\sum(Y_{ijk} - \bar{Y}_{ij})^2$$

Equation 22: Sum of squares interaction

Here
Yijk = elements in group based on ith group of factor1 and jth group of factor2
Yij bar = mean of group based on ith group of factor1 and jth group of factor2

4. Within-Group Sum of Squares (SSW):
   It calculates the sum of squares of the mean of each group from each element in that group, depicted by this equation

$$SSWithin = \sum\sum\sum(Y_{ijk} - \bar{Y}_{ij})^2$$

Equation 23: Sum of squares within

Here
Yijk = elements in group based on ith group of factor1 and jth group of factor2
Yij bar = mean of group based on ith group of factor1 and jth group of factor2

The total of all the above elements is called Total sum of squares (SST)

$$SST = SS_{Factor1} + SS_{Factor2} + SS_{Interaction} + SS_{Within}$$

Equation 24: Total sum of squares

Using the above-mentioned equations, we created an ANOVA Table

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

Table 17: ANOVA Table

In this table
C(Dentist) represents factor1
C(Method) represents factor2
C(Dentist): C(Method) represents interaction
Residual represents within the table
df = degree of freedom
sum_sq = sum of squares
mean_sq = sum_sq / df
F = mean_sq between / mean_sq within
PR(>F) = p-value of getting the f-statistic

Using these values, we can find if variable of interest gets affected by the new factor introduced in the form of interaction between factor1 and factor2 and to what level this factor influences the variable of interest.

Conclusion

As noted from One Way ANOVA Test, the total sum of squares (SST) is same at 1139002.8. With only Dentist as predictor 4.98% variability is explained by it. With only Method as predictor 43.87% variability is explained by it. However, when both the predictors are in the model around two third, (753898.13/1139002.8) 66.19% variability is explained by both main effects and their interaction effects.

The p-value for only Method as predictor is significantly lower than level of significance (0.05) while for Dentist as predictor p-value (0.372) is significantly higher than level of significance (0.05)

Also, there is no significant interaction effect between Dentist and Method as the p-value for interaction effect (0.093) is significantly above the level of significance (0.05)

Thus, to conclude we fail to reject the Null Hypothesis(H0) of mean hardness of implant remains same for a dentist irrespective of the method involved. Similarly, mean hardness of implant for a method remains same irrespective of the dentist who performs the implant.

In case of alloy 2, Method as a predictor is the most important factor influencing the Hardness of dental implant, based on the provided results.