BUSINESS REPORT

List of Content

List of Figures

LIST OF TABLES

Objective:

Problem 1:

To investigate the correlation between foot injuries and player positions within a male football team, aiming to provide insights into injury patterns that may inform targeted preventive measures and optimize player performance and well-being.

Problem 2:

To analyse the breaking strength distribution of gunny bags used for packaging cement, characterized by a normal distribution with a mean of 5 kg per sq. cm and a standard deviation of 1.5 kg per sq. cm. The aim is to address inquiries related to the quality and durability of the packaging material, thereby enhancing insights into potential wastage or pilferage within the supply chain.

Problem 3:

To assess the Brinell's hardness index of a batch of stones received by Zingaro Stone Printing, aiming to determine their suitability for achieving the desired printing quality. The objective is to ensure that the stone surfaces meet the minimum requirement of a Brinell's hardness index of at least 150, thereby enabling optimal image printing quality and maintaining client satisfaction standards.

Problem 4:

To analyse dental implant data, investigating the influence of various factors—including implant method, treatment temperature, alloy composition, and dentist preference—on the hardness of metal implants within dental cavities. The aim is to discern patterns and dependencies that contribute to variations in implant hardness, thereby facilitating informed decision-making for optimizing dental implant procedures and outcomes.

Problem 1:

1.1 probability that a randomly chosen player would suffer an injury:

The probability that a randomly chosen player would suffer an injury is 0.6170212765957447

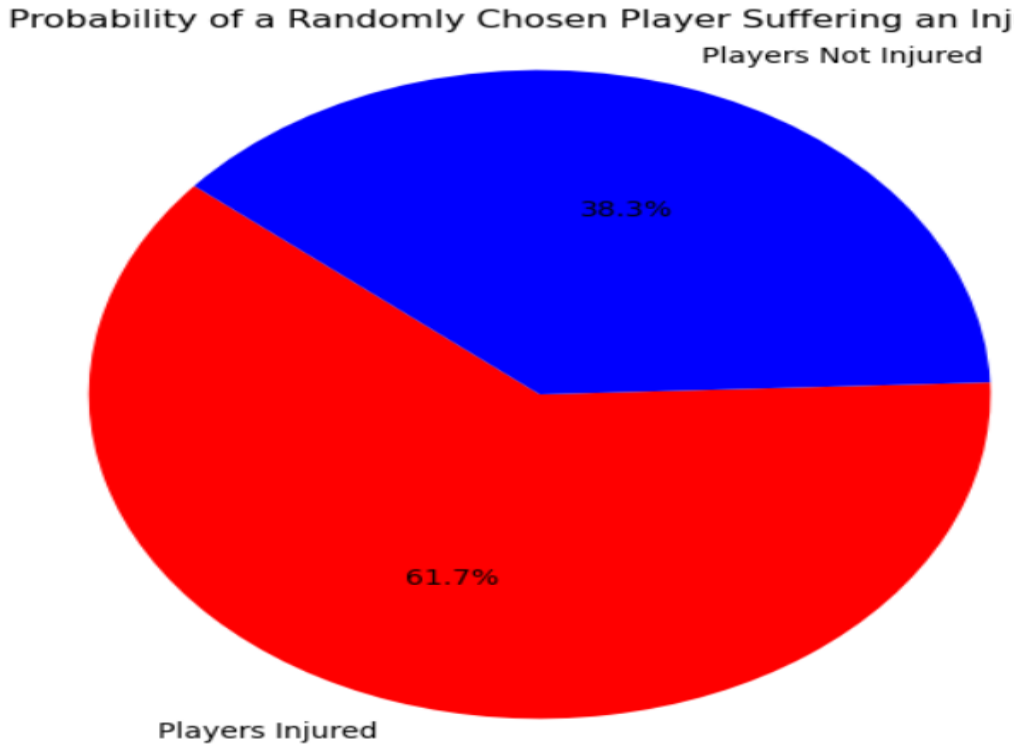


Fig 1.1

1.2 probability that a player is a forward or a winger:

The probability that a player is a forward or a winger 0.5234042553191489.

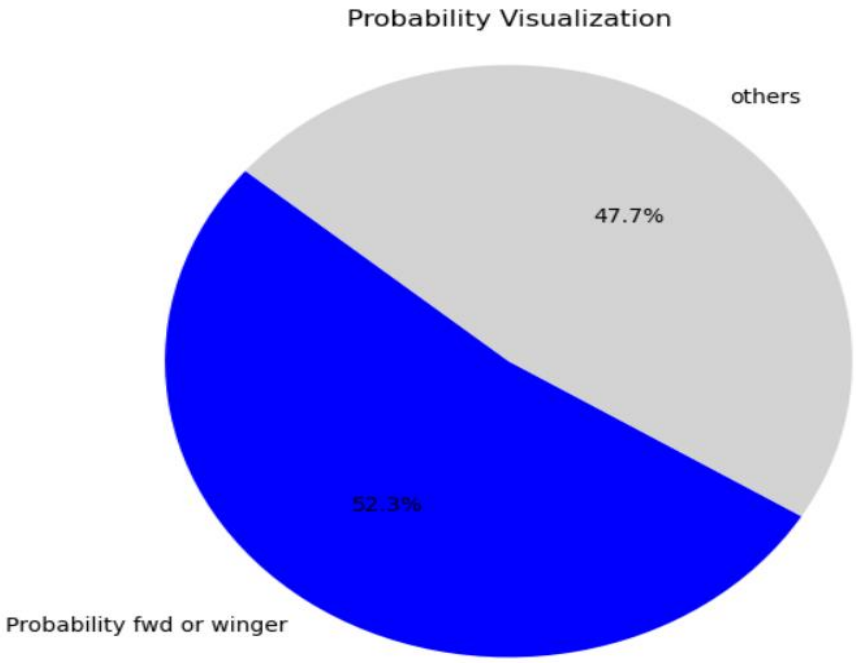


Fig 1.2

1.3 probability that a randomly chosen player plays in a striker position and has a foot injury:

Probability of a randomly chosen player being a striker and having a foot injury: 0.19148936 170212766.

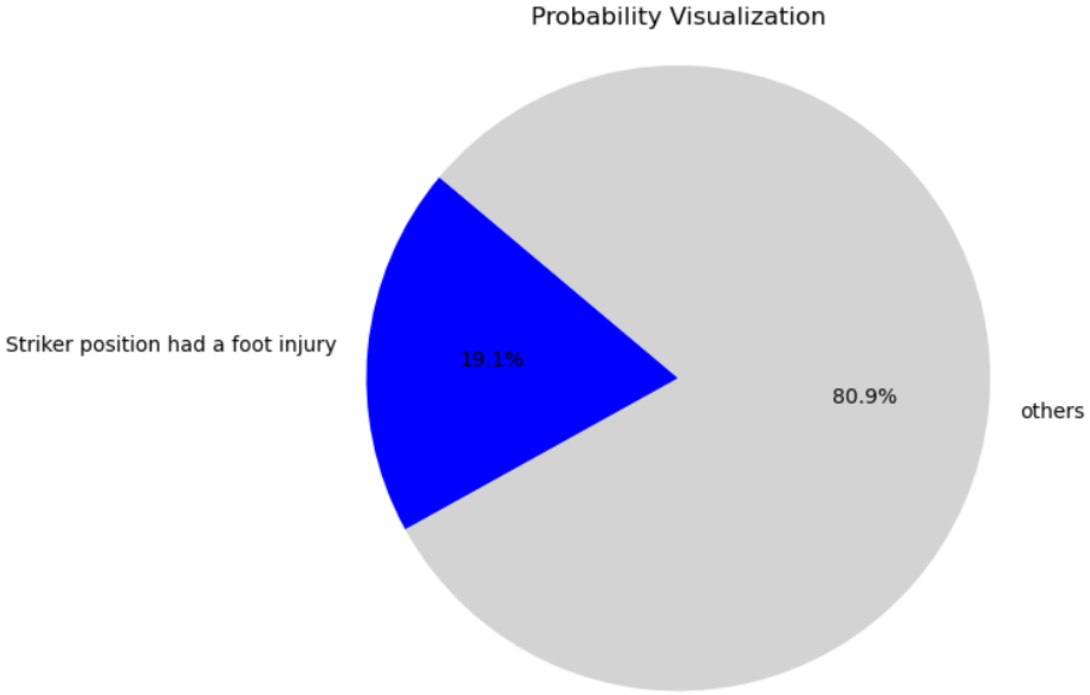


Fig 1.3

1.4 probability that a randomly chosen injured player is a striker:

The probability that a randomly chosen injured player is a striker 0.3103448275862069

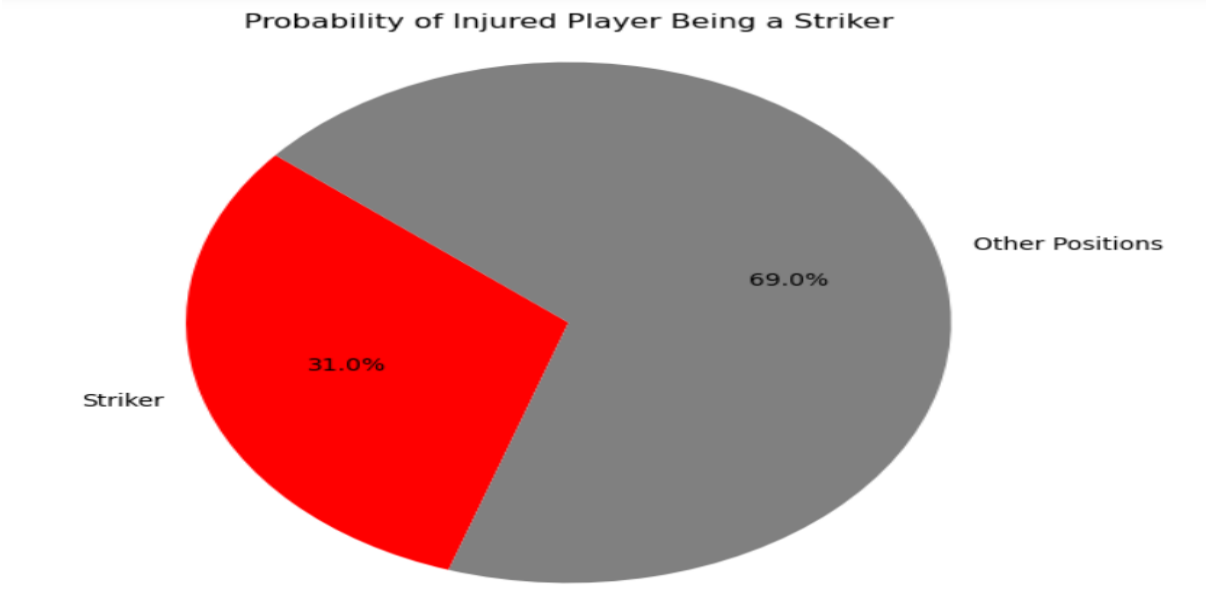
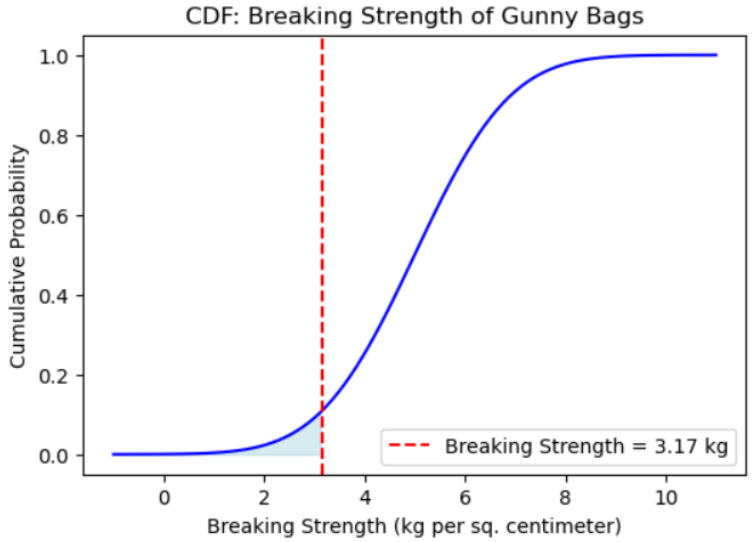


Fig 1.4

Problem 2:

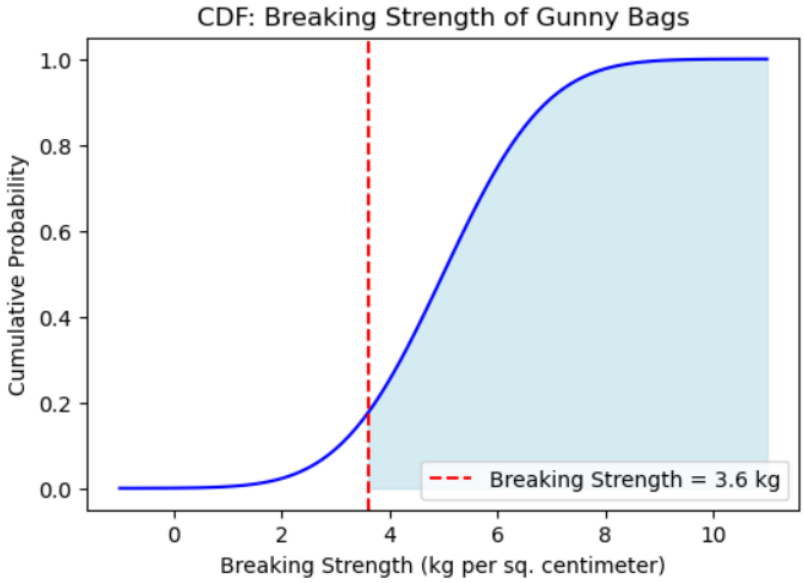
2.1 Proportion of the gunny bags have a breaking strength of less than 3.17 kg per square cm:




```
Proportion of gunny bags with a breaking strength of less than 3.17 kg per sq cm: 0.11123243744783456
```

Fig 2.1

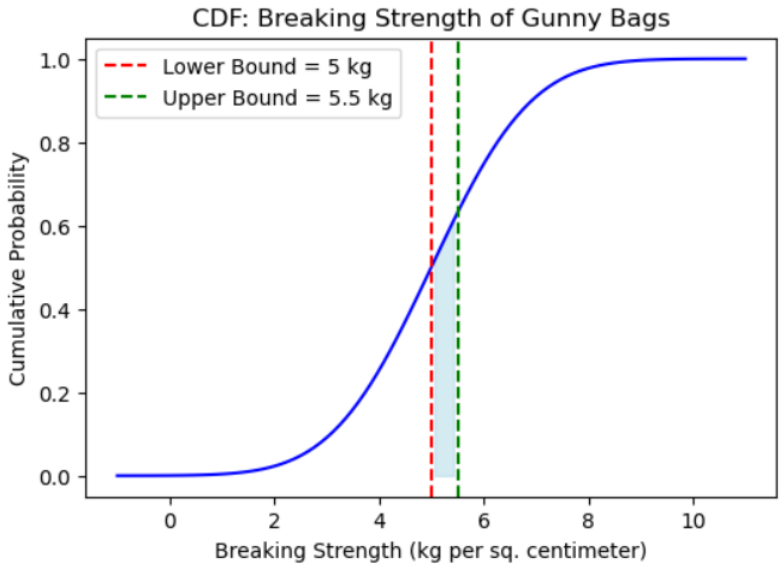
2.2 Proportion of the gunny bags have a breaking strength of at least 3.6 kg per square cm:




```
Proportion of gunny bags with a breaking strength of at least 3.6 kg per sq cm: 0.8246760551477705
```

Fig 2.2

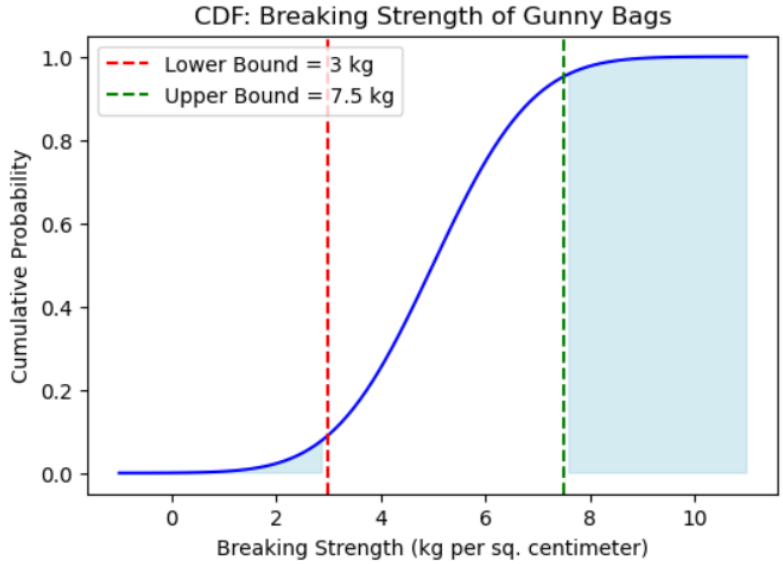
2.3 Proportion of the gunny bags have a breaking strength between 5 and 5.5 kg per square cm:




```
Proportion of gunny bags with breaking strength between 5 and 5.5 kg per sq cm: 0.13055865981823633
```

Fig 2.3

2.4 Proportion of the gunny bags have a breaking strength NOT between 3 and 7.5 kg per square cm.




```
Proportion of gunny bags with breaking strength NOT between 3 and 7.5 kg per sq cm: 0.13900157199868257
```

Fig 2.4

Problem 3:

3.1 Zingaro has reason to believe that the unpolished stones may not be suitable for printing:

Null Hypothesis :

- The mean Brinell's hardness index of unpolished stones is equal to or greater than 150 (suitable for printing).

Alternative Hypothesis:

- The mean Brinell's hardness index of unpolished stones is less than 150 (not suitable for printing).

Fig 3.1.1

Loaded the required file. The dataset consists of 75 rows and 2 columns.

	Unpolished	Treated and Polished
0	164.481713	133.209393
1	154.307045	138.482771
2	129.861048	159.665201
3	159.096184	145.663528
4	135.256748	136.789227
...	...	...
70	123.067611	142.293544
71	171.822218	140.124092
72	88.135994	141.393091
73	145.150397	131.370530
74	170.854823	144.502647

75 rows × 2 columns

Table 3.1.1

```
RangeIndex: 75 entries, 0 to 74
Data columns (total 2 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Unpolished            75 non-null     float64
 1   Treated and Polished  75 non-null     float64
dtypes: float64(2)
memory usage: 1.3 KB
```

Fig 3.1.2

Test Statistic: -4.164629601426757
P-value: 8.342573994839304e-05
Reject the null hypothesis. There is sufficient evidence to conclude that the mean Brinell's hardness index of unpolished stone s is less than 150.

Rejected the Null Hypothesis

Fig 3.1.3

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

Null Hypothesis :

- The mean hardness of polished stones is equal to the mean hardness of unpolished stones.

Alternative Hypothesis :

- The mean hardness of polished stones is not equal to the mean hardness of unpolished stones.

Fig 3.2.1

Test Statistic: 3.2422320501414053
P-value: 0.0014655150194628353
Reject the null hypothesis. There is sufficient evidence to conclude that the mean hardness of polished and unpolished stones i s not the same.

Fig 3.2.2

Problem 4:

4.1 Hardness of implants vary depending on dentists:

Loaded the required file, loaded the data. The dataset consists of 107 rows and 14 columns.

	Dentist	Method	Alloy	Temp	Response	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9	Unnamed: 10	Unnamed: 11	Unnamed: 12	U
0	1.0	1.0	1.0	1500.0	813.0	NaN	NaN	Anova: Two-Factor Without Replication	NaN	NaN	NaN	NaN	NaN	
1	1.0	1.0	1.0	1600.0	792.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
2	1.0	1.0	1.0	1700.0	792.0	NaN	NaN	SUMMARY	Count	Sum	Average	Variance	NaN	
3	1.0	1.0	2.0	1500.0	907.0	NaN	NaN	1	4	2315	578.75	523721.583333	NaN	
4	1.0	1.0	2.0	1600.0	792.0	NaN	NaN	1	4	2394	598.5	584819	NaN	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
102	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Rows	513114.013889	89	5765.325999	0.778206	0.917216	
103	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Columns	156164270.208333	3	52054756.736111	7026.375161	0.0	
104	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Error	1978064.041667	267	7408.479557	NaN	NaN	
105	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
106	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Total	158655448.263889	359	NaN	NaN	NaN	

107 rows × 14 columns

Table 4.1.1

```
RangeIndex: 107 entries, 0 to 106
Data columns (total 14 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   Dentist      90 non-null     float64
 1   Method       90 non-null     float64
 2   Alloy        90 non-null     float64
 3   Temp         90 non-null     float64
 4   Response     90 non-null     float64
 5   Unnamed: 5   0 non-null      float64
 6   Unnamed: 6   0 non-null      float64
 7   Unnamed: 7   102 non-null    object
 8   Unnamed: 8   100 non-null    object
 9   Unnamed: 9   100 non-null    object
 10  Unnamed: 10  99 non-null     object
 11  Unnamed: 11  98 non-null     object
 12  Unnamed: 12  3 non-null      object
 13  Unnamed: 13  3 non-null      object
dtypes: float64(7), object(7)
memory usage: 11.8+ KB
```

Fig 4.1.1

State the null and alternate hypotheses

Null and Alternative Hypotheses:

Null Hypothesis :

- There is no difference in the mean hardness of implants among different dentists.

Alternative Hypothesis:

- There is a difference in the mean hardness of implants among different dentists.

Fig 4.1.2

Check the assumptions of the hypothesis test.

```
Shapiro-Wilk Normality Test Results:
Dentist 1.0: p-value = 0.1856756955385208
Dentist 2.0: p-value = 0.7310513854026794
Dentist 3.0: p-value = 0.5520960092544556
Dentist 4.0: p-value = 0.002037237398326397
Dentist 5.0: p-value = 0.5022943019866943

Levene's Test for Homogeneity of Variances:
Test Statistic: 3.7082282441933954
P-value: 0.007858817382355401
```

Fig 4.1.3

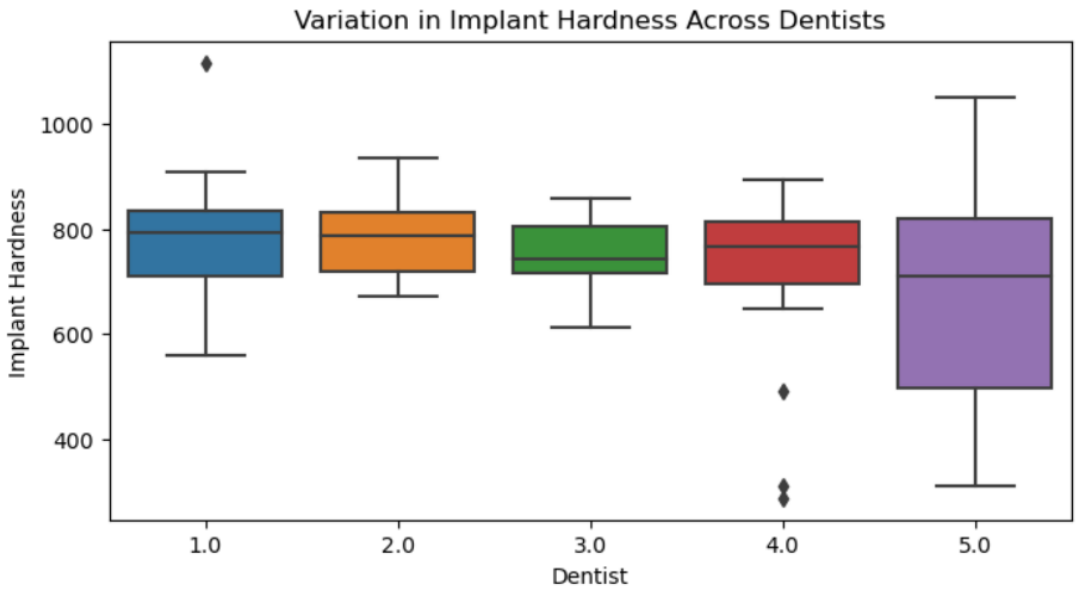


Fig 4.1.4

Conduct the hypothesis test and compute the p-value and write down conclusions from the test results.

```
One-way ANOVA Test Results:
F-statistic: 1.9345366474464367
P-value: 0.11206595023098852
The p-value is greater than or equal to the significance level, indicating no significant differences exist among dentists in t
erms of implant hardness.
```

Hypothesis Test:

One-way ANOVA Test: The p-value obtained is approximately 0.112, which is greater than the significance level of 0.05. Therefore, we fail to reject the null hypothesis. This indicates that there are no significant differences in implant hardness among the different dentists.

Conclusion:

Based on the results of the hypothesis test, we conclude that there are no significant differences in implant hardness among the different dentists.

Fig 4.1.5

There is no sufficient data for Alloy type 1&2.

4.2 Hardness of implants vary depending on methods:

State the null and alternate hypotheses.

Null Hypothesis :

- The mean hardness of implants is the same across all methods.

Alternative Hypothesis :

- The mean hardness of implants varies depending on the method used.

Fig 4.2.1

Check the assumptions of the hypothesis test.

```
Shapiro-Wilk Normality Test Results:
Method 1.0: Shapiro-Wilk Statistic = 0.9652422666549683, P-value = 0.41838711500167847
Method 2.0: Shapiro-Wilk Statistic = 0.9371030926704407, P-value = 0.07601878046989441
Method 3.0: Shapiro-Wilk Statistic = 0.9091948866844177, P-value = 0.014201466925442219

Levene's Test for Homogeneity of Variances:
Test Statistic: 5.848571977210466
P-value: 0.004138452940152019
```

Fig 4.2.2

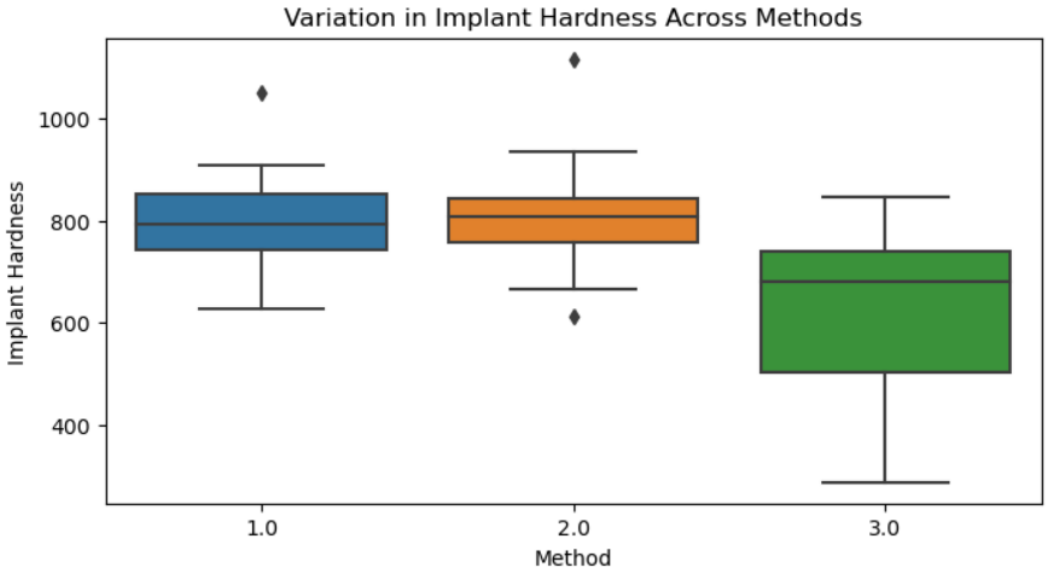


Fig 4.2.3

Conduct the hypothesis test and compute the p-value - Write down conclusions from the test results.

```
Welch's ANOVA Test Results:
F-statistic: 19.89268013119534
P-value: 7.683891892977992e-08
```

Conclusion:

P value : 7.683891892977992e-08

- ANOVA Test shows that there are statistically significant differences in hardness amonng the different methods of dental implantation.Atleast one pair of methods had significantly diffrent hardness.
- Since the p-value is negligible we can reject the Null hypothesis.

Fig 4.2.4

separate analysis for the two types of alloys.

```
Games-Howell Post-Hoc Test Results for Alloy Type 1:
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
=====================================================
group1 group2  meandiff p-adj    lower    upper   reject
-----------------------------------------------------
   1.0    2.0    -6.1333  0.987  -102.714  90.4473  False
   1.0    3.0     -124.8 0.0085 -221.3807 -28.2193   True
   2.0    3.0 -118.6667 0.0128 -215.2473  -22.086   True
-----------------------------------------------------

Games-Howell Post-Hoc Test Results for Alloy Type 2:
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
=====================================================
group1 group2 meandiff p-adj    lower      upper   reject
-----------------------------------------------------
   1.0    2.0      27.0 0.8212  -82.4546  136.4546  False
   1.0    3.0    -208.8 0.0001 -318.2546  -99.3454   True
   2.0    3.0    -235.8    0.0 -345.2546 -126.3454   True
-----------------------------------------------------
```

Fig 4.2.5

The Games-Howell post-hoc test results for each group of alloys are as follows:

Alloy Type 1:

- There is a statistically significant difference in implant hardness between Method 1 and Method 3 (reject null hypothesis).
- There is a statistically significant difference in implant hardness between Method 2 and Method 3 (reject null hypothesis).
- There is no statistically significant difference in implant hardness between Method 1 and Method 2 (fail to reject null hypothesis).

Alloy Type 2:

- There is a statistically significant difference in implant hardness between Method 1 and Method 3 (reject null hypothesis).
- There is a statistically significant difference in implant hardness between Method 2 and Method 3 (reject null hypothesis).
- There is no statistically significant difference in implant hardness between Method 1 and Method 2 (fail to reject null hypothesis).

Fig 4.2.6

4.3 Interaction effect between the dentist and method:

Create Interaction Plot - Inferences from the plot

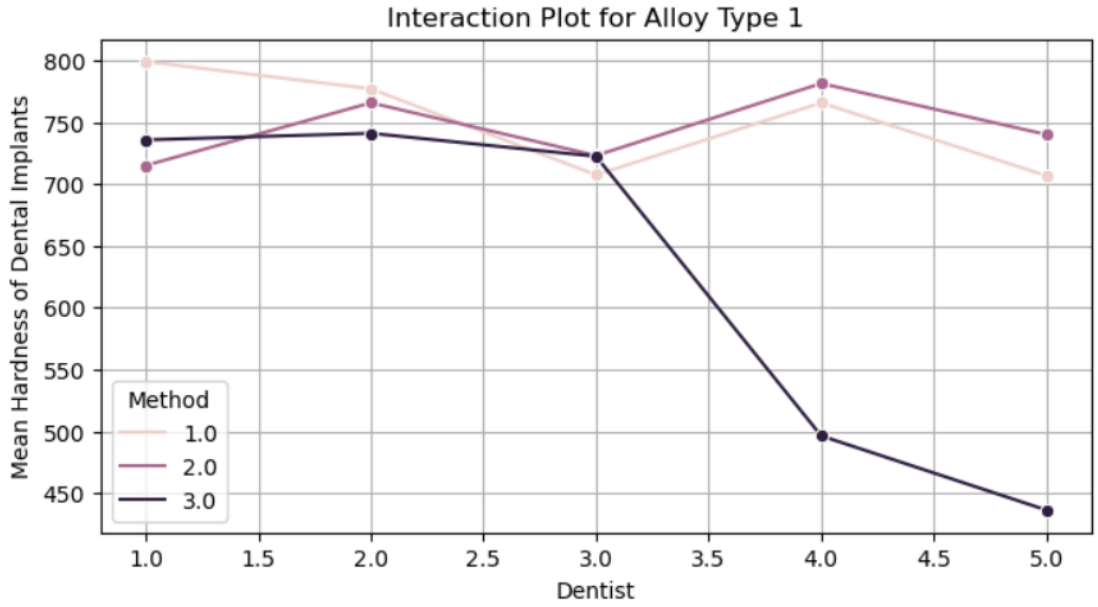
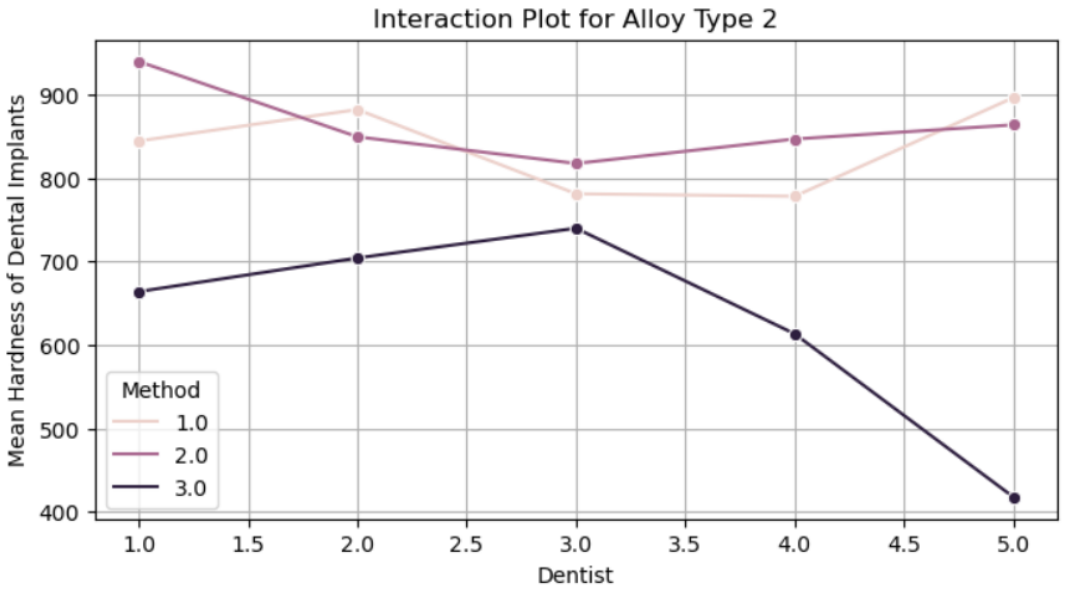


Fig 4.3.1



- Here we can see the mean hardness differs gradually among the dentists methods for both alloy types.

Fig 4.3.2

4.4 Hardness of implants vary depending on dentists and methods together:

State the null and alternate hypotheses.

Null Hypothesis :

- There is no interaction effect between dentists and methods on the hardness of implants.

Alternative Hypothesis :

- There is a significant interaction effect between dentists and methods on the hardness of implants.

Fig 4.4.1

Check the assumptions of the hypothesis test.

```
Dentist 1.0: Shapiro-Wilk p-value = 0.1856756955385208
Dentist 2.0: Shapiro-Wilk p-value = 0.7310513854026794
Dentist 3.0: Shapiro-Wilk p-value = 0.5520960092544556
Dentist 4.0: Shapiro-Wilk p-value = 0.002037237398326397
Dentist 5.0: Shapiro-Wilk p-value = 0.5022943019866943
Method 1.0: Shapiro-Wilk p-value = 0.41838711500167847
Method 2.0: Shapiro-Wilk p-value = 0.07601878046989441
Method 3.0: Shapiro-Wilk p-value = 0.014201466925442219

Levene's Test for Homogeneity of Variances:
Test Statistic: 3.9381047054545255
P-value: 0.0005175374842450017
```

Fig 4.4.2

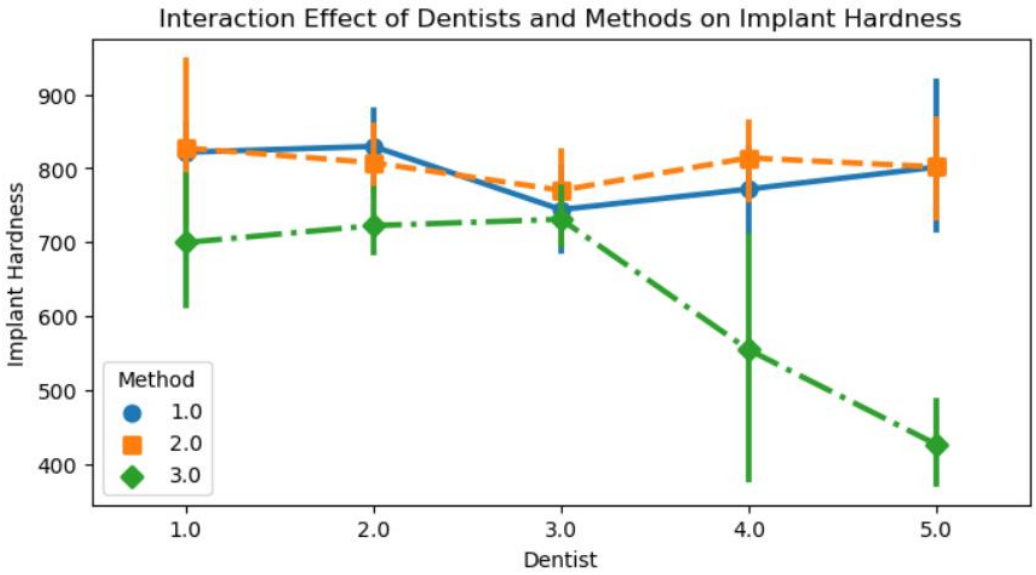


Fig 4.4.3

Based on the Shapiro-Wilk tests for normality and Levene's test for homogeneity of variances:

Normality of Residuals:

- Dentist 1.0: Shapiro-Wilk p-value = 0.186 (normal)
- Dentist 2.0: Shapiro-Wilk p-value = 0.731 (normal)
- Dentist 3.0: Shapiro-Wilk p-value = 0.552 (normal)
- Dentist 4.0: Shapiro-Wilk p-value = 0.002 (not normal)
- Dentist 5.0: Shapiro-Wilk p-value = 0.502 (normal)
- Method 1.0: Shapiro-Wilk p-value = 0.418 (normal)
- Method 2.0: Shapiro-Wilk p-value = 0.076 (normal)
- Method 3.0: Shapiro-Wilk p-value = 0.014 (not normal)

Homogeneity of Variances:

- Test Statistic: 3.938
- P-value: 0.001 Since the p-value is less than the chosen significance level, we reject the null hypothesis.This indicates that the variances of the residuals are not equal across all groups.

Fig 4.4.4

Conduct the hypothesis test and compute the p-value - Write down conclusions from the test results.

```
Two-way ANOVA Test Results:
P-value: 0.0019695151879771935
```

- The p-value obtained from the two-way ANOVA test is 0.00197. This p-value indicates that there is a statistically significant interaction effect between Dentist and Method on the hardness of dental implants.
- The significance of this result suggests that the effect of one factor (Dentist) on implant hardness depends on the level of the other factor (Method), and vice versa.

Fig 4.4.5

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

```
Post-hoc Tests for Alloy Type 1:

Dentist:
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
=============================================================
group1 group2  meandiff p-adj    lower    upper   reject
-------------------------------------------------------------
   1.0    2.0   11.3333 0.9996 -145.0423  167.709  False
   1.0    3.0  -32.3333 0.9757  -188.709 124.0423  False
   1.0    4.0  -68.7778 0.7189 -225.1535  87.5979  False
   1.0    5.0 -122.2222 0.1889 -278.5979  34.1535  False
   2.0    3.0  -43.6667 0.9298 -200.0423  112.709  False
   2.0    4.0  -80.1111 0.5916 -236.4868  76.2646  False
   2.0    5.0 -133.5556 0.1258 -289.9312  22.8201  False
   3.0    4.0  -36.4444 0.9626 -192.8201 119.9312  False
   3.0    5.0  -89.8889 0.4805 -246.2646  66.4868  False
   4.0    5.0  -53.4444 0.8643 -209.8201 102.9312  False
-------------------------------------------------------------

Method:
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
=============================================================
group1 group2  meandiff p-adj    lower    upper   reject
-------------------------------------------------------------
   1.0    2.0   -6.1333  0.987  -102.714  90.4473  False
   1.0    3.0    -124.8 0.0085 -221.3807 -28.2193   True
   2.0    3.0 -118.6667 0.0128 -215.2473  -22.086   True
-------------------------------------------------------------
```

Fig 4.4.6

```
Post-hoc Tests for Alloy Type 2:

Dentist:
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
=====================================================
group1 group2 meandiff p-adj    lower    upper   reject
-----------------------------------------------------
   1.0    2.0  -4.1111    1.0 -225.5687 217.3465  False
   1.0    3.0 -36.5556 0.9895 -258.0131  184.902  False
   1.0    4.0    -70.0 0.8941 -291.4576 151.4576  False
   1.0    5.0 -90.1111 0.7724 -311.5687 131.3465  False
   2.0    3.0 -32.4444 0.9933  -253.902 189.0131  False
   2.0    4.0 -65.8889 0.9132 -287.3465 155.5687  False
   2.0    5.0    -86.0 0.8008 -307.4576 135.4576  False
   3.0    4.0 -33.4444 0.9925  -254.902 188.0131  False
   3.0    5.0 -53.5556 0.9574 -275.0131  167.902  False
   4.0    5.0 -20.1111  0.999 -241.5687 201.3465  False
-----------------------------------------------------

Method:
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
=======================================================
group1 group2 meandiff p-adj    lower     upper   reject
-------------------------------------------------------
   1.0    2.0     27.0 0.8212  -82.4546  136.4546  False
   1.0    3.0   -208.8 0.0001 -318.2546  -99.3454   True
   2.0    3.0   -235.8    0.0 -345.2546 -126.3454   True
-------------------------------------------------------
```

Fig 4.4.7

Based on the post-hoc tests results:

AlloyType 1:

- Dentist: There are no significant differences observed among all pairs of dentists.
- Method: Significant differences are observed between Method 1 and Method 3, as well as between Method 2 and Method 3.

Alloy Type 2:

- Dentist: There are no significant differences observed among all pairs of dentists.
- Method: Significant differences are observed between Method 1 and Method 3, as well as between Method 2 and Method 3.
- These results suggest that for both types of alloys, the method used significantly affects the hardness of implants, with Method 3 showing differences compared to Method 1 and Method 2. However, there are no significant differences observed among dentists for either type of alloy.

Fig 4.4.8