Contents

Advanced Statistics Project

Import some libraries like Numpy, Pandas, Seaborn, Matplotlib, Scipy.stats for statistical functions, for n-way ANOVA etc.

Problem 1

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

	Striker	Forward	Attacking Midfielder	Winger	Total
Players Injured	45	56	24	20	145
Players Not Injured	32	38	11	9	90
Total	77	94	35	29	235

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

Total Players Injured = 145
Total Players=235
The probability that a randomly chosen player would suffer an injury =
= Total Players Injured ÷ Total Players = 145/235
The probability that a randomly chosen player would suffer an injury = 0.6170212765957447

- 61.70% of players are suffer.

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

Forward Players = 94
Winger Player = 29
Forward and winger = 94+29=123
Total Players=235
The probability that a player is a forward or a winger = Forward and winger ÷ Total Players
The probability that a player is a forward or a winger = 123/235
The probability that a player is a forward or a winger 0.5234042553191489

- 52.34% of players are forward and winger.

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

Players Injured Striker = 45
Total Striker = 77
The probability that a randomly chosen player plays in a striker position and has a foot injury = 45/77
The probability that a randomly chosen player plays in a striker position and has a foot injury 0.5844155844155844

- 58.44% of player plays in a striker position and has a foot injury.

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

Players Injured Striker = 45
Total Players Injured = 145
The probability that a randomly chosen injured player is a striker = 45/145
The probability that a randomly chosen injured player is a striker 0.3103448275862069

- 31.03% of randomly chosen injured player is a striker.

1.5 What is the probability that a randomly chosen injured player is either a forward or an attacking midfielder?
Players Injured Forward = 56
Players Injured Attacking = 24
Forward and Attacking = 56+24 = 80
Total Players Injured = 145
The probability that a randomly chosen injured player is either a forward or an attacking midfielder = 80/145
The probability that a randomly chosen injured player is either a forward or an attacking midfielder 0.5517241379310345

- 55.17% of randomly chosen injured player is either a forward or an attacking midfielder.

Problem 2

An independent research organization is trying to estimate the probability that an accident at a nuclear power plant will result in radiation leakage. The types of accidents possible at the plant are, fire hazards, mechanical failure, or human error. The research organization also knows that two or more types of accidents cannot occur simultaneously.

According to the studies carried out by the organization, the probability of a radiation leak in case of a fire is 20%, the probability of a radiation leak in case of a mechanical 50%, and the probability of a radiation leak in case of a human error is 10%. The studies also showed the following;

The probability of a radiation leak occurring simultaneously with a fire is 0.1%.
The probability of a radiation leak occurring simultaneously with a mechanical failure is 0.15%.

The probability of a radiation leak occurring simultaneously with a human error is 0.12%.
On the basis of the information available, answer the questions below:

2.1 What are the probabilities of a fire, a mechanical failure, and a human error respectively?

According to question

- P(R|F)=20%=0.20
- P(R|M)=50%=0.50
- P(R|H)=10%=0.10
- P($R \cap F$)=0.1%=0.001
- P($R \cap M$)=0.15%=0.0015
- P($R \cap H$)=0.12%=0.0012

We know that : P($R \cap F$)=P(F).P(R/F)

- The probabilities of a fire P(F)=P($R \cap F$) / P(R/F)
- The probabilities of a mechanical failure P(M)=P($R \cap M$) / P(R/M)
- The probabilities of a human error P(H)=P($R \cap H$) / P(R/H)

➢ The probabilities of a fire = 0.001/0.20

The probabilities of a fire 0.005

- The probabilities of a mechanical failure = 0.0015/0.50

The probabilities of a mechanical failure 0.003

- The probabilities of a human error = 0.0012/0.10

The probabilities of a human error 0.011999999999999999

Insights
The probabilities of a fire, a mechanical failure, and a human error respectively 0.005, 0.003, 0.012

2.2 What is the probability of a radiation leak?
The probability of Radiation leak due to Fire, Mechanical Failure, Human Error
Total Probability =1
P(N)= Probability of no accident occur
P(N)=1-(P(F)+P(M)+P(H))
P(N)=0.98 So, P(R|N)=0
P(R)=P($R \cap F$) + P($R \cap M$) + P($R \cap H$) + P(R|N)
The probability of a radiation leak = P(R) = 0.001+0.0015+0.0012+0
The probability of a radiation leak 0.0037

2.3 Suppose there has been a radiation leak in the reactor for which the definite cause is not known. What is the probability that it has been caused by:

A Fire.
The probability of radiation leak caused by Fire
P(F|R)=P(F∩R)/P(R)
The probability of radiation leak caused by Fire = 0.001/0.0037
The probability of a radiation leak 0.2702702702702703

A Mechanical Failure.
The probability of radiation leak caused by Mechanical Failure
P(F|R)=PP(M∩R)/P(R)
The probability of radiation leak caused by Mechanical Failure = 0.0015/0.0037
The probability of radiation leak caused by Mechanical Failure 0.4054054054054054

A Human Error.
The probability of radiation leak caused by Human Error
P(F|R)=PP(H∩R)/P(R)
The probability of radiation leak caused by Human Error = 0.0012/0.0037
The probability of radiation leak caused by Human Error 0.3243243243243243

Problem 3
The breaking strength of gunny bags used for packaging cement is normally distributed with a mean of 5 kg per sq. centimeter and a standard deviation of 1.5 kg per sq. centimeter. The quality team of the cement company wants to know the following about the packaging material to better understand wastage or pilferage within the supply chain; Answer the questions below

based on the given information; (Provide an appropriate visual representation of your answers, without which marks will be deducted)

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

Formula of z-score
z=(x-mean)/std

mean=5
std=1.5
x=3.17
z=(3.17-5)/1.5

Proportion of the gunny bags have a breaking strength less than 3.17 kg per sq cm is %1.5f = 0.11123

In a normal distribution, based on the empirical rule, sometimes known as the 68-95-99.7 rule, the number 4 is selected as an acceptable estimate. Approximately 68% of the values, 95%, and 99.7% of the values, respectively, lie within one standard deviation of the mean, two standard deviations, and three standard deviations of the mean, respectively.

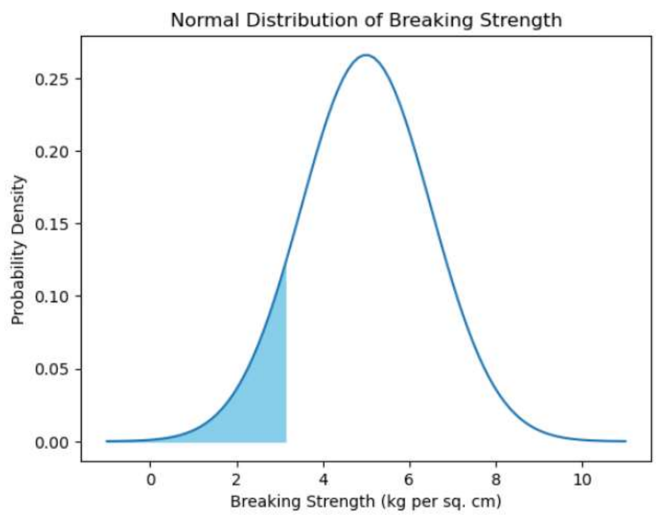

Fig 1 : Normal Distribution of Breaking Strength less than 3.17 kg per sq cm

- The percentage of gunny bags having a breaking strength less than amount is indicated by the shaded region beneath the curve beginning at the value of 3.17 kg per sq. cm.

3.2 What proportion of the gunny bags have a breaking strength at least 3.6 kg per sq cm.?
mean=5
std=1.5
x=3.6
z=(3.6-5)/1.5

Proportion of the gunny bags have a breaking strength less than 3.6 kg per sq cm is %1.5f = 0.82468

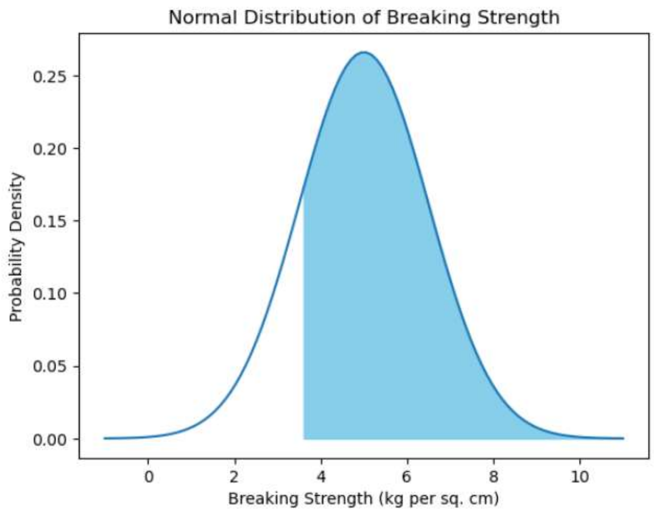

Fig 2 : Normal Distribution of Breaking Strength less than 3.6 kg per sq cm

- The percentage of gunny bags having a breaking strength at least amount is indicated by the shaded region beneath the curve beginning at the value of 3.6 kg per sq. cm.

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

Z1=(5.5-5)/1.5
Z2=(5-5)/1.5

Proportion of the gunny bags have a breaking strength between 5 and 5.5 kg per sq cm. %3.4f = 0.1306

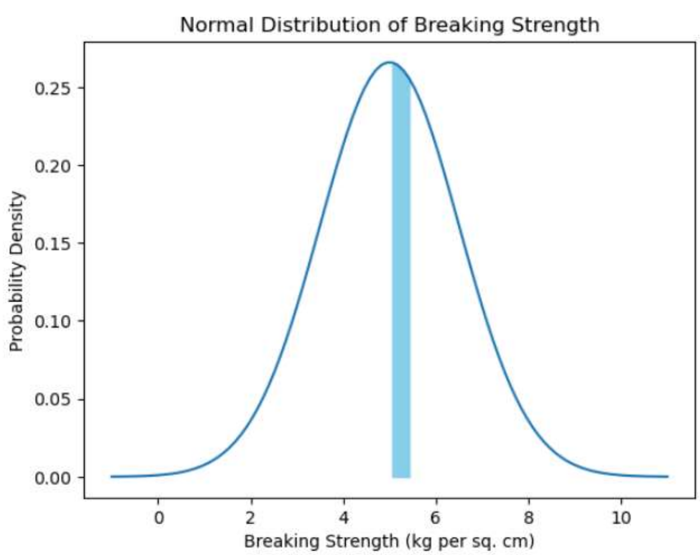

Fig 3 : Normal Distribution of Breaking Strength between 5 and 5.5 kg per sq cm

- The percentage of gunny bags having a breaking strength between 5 and 5.5 kg per sq. cm is shown by the shaded region under the curve.

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

Z1=(7.5-5)/1.5
Z2=(3-5)/1.5

Proportion of the gunny bags have a breaking strength between 3 and 7.5 kg per sq cm.0.1390

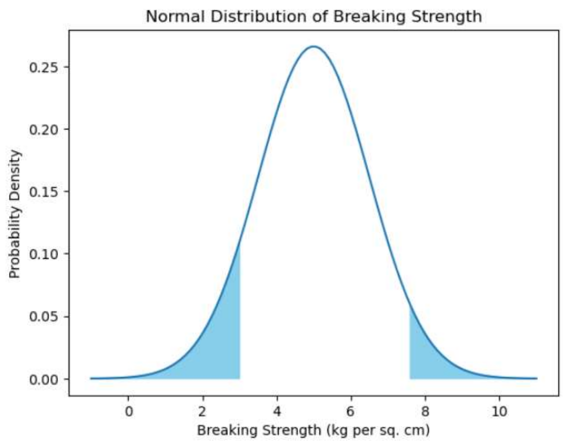

Fig 4 : Normal Distribution of Breaking Strength between 3 and 7.5 kg per sq cm

- The percentage of gunny bags having a breaking strength outside of that range is represented by the shaded region beneath the curve outside of the range between 3 and 7.5 kg per sq. cm.

Problem 4:

Grades of the final examination in a training course are found to be normally distributed, with a mean of 77 and a standard deviation of 8.5. Based on the given information answer the questions below.

4.1 What is the probability that a randomly chosen student gets a grade below 85 on this exam?

Formula of z-score

z=(x-mean)/std

mu=77
sigma=8.5
x=85

The probability that a randomly chosen student gets a grade below 85 on this exam 0.8266927837484748

- The probability that a randomly chosen student gets a grade below 85 on this exam is 82.66%.

4.2 What is the probability that a randomly selected student scores between 65 and 87?
mu=77
sigma=8.5
x1=65
x2=87

Probability that a randomly selected student scores between 65 and 87 0.8012869336779058

- The probability that a randomly selected student scores between 65 and 87 is 80.13%

4.3 What should be the passing cut-off so that 75% of the students clear the exam?

mu=77
sigma=8.5
x=0.75

The passing cut-off so that 75% of the students clear the exam 82.7331628766667

- The passing cut-off so that 75% of the students clear the exam should be 82.73%.

Problem 5:

Zingaro stone printing is a company that specializes in printing images or patterns on polished or unpolished stones. However, for the optimum level of printing of the image the stone surface has to have a Brinell's hardness index of at least 150. Recently, Zingaro has received a batch of polished and unpolished stones from its clients. Use the data provided to answer the following (assuming a 5% significance level)

Load our data set, Zingaro_Company.csv, and use the head() function to view the data. Find out the characteristics of the columns using the info() method and we can determine that there are 75 rows and 2 columns. The datatypes for the float64(2) columns are present.

- **head()** it given by default top five data

	Unpolished	Treated and Polished
0	164.481713	133.209393
1	154.307045	138.482771
2	129.861048	159.665201
3	159.096184	145.663528
4	135.256748	136.789227

Table 1 : Top five data of dataset Zingaro_Company

- **info()** it tells a concise summary of a DataFrame

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

Table 2: Information about the Zingaro_Company structure and content.

- **Describe()** it tells summary of the central tendency, dispersion, and shape of the distribution of the data.

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

Table 3: Descriptive statistics of the Zingaro_Company

5.1 Earlier experience of Zingaro with this particular client is favorable as the stone surface was found to be of adequate hardness. However, Zingaro has reason to believe now that the unpolished stones may not be suitable for printing. Do you think Zingaro is justified in thinking so?

Step 1: Define the Null Hypothesis(H_0) and Alternative hypothesis(H_A)
Null hypothesis states that,The mean hardness of unpolished stone, μ_o is at least 150.
Alternative hypothesis states that, The mean hardness of unpolished stone, μ_a is less than 150.

H_0 : μ_o >= 150
H_A : μ_a< 150

Step 2: Decide the Significance Level
Significance Level is given in the question
Here we select $\alpha = 0.05$.

Step 3: Identify the Test Statistics
Here n = 75

Step 4: Calculate the p - value and test statistic

scipy.stats.ttest_1samp calculates the t test for the mean of one sample given the sample observations and the expected value in the null hypothesis. This function returns t statistic and the two-tailed p value.

1. Calculate sample statistics

- Sample Mean of Zig dataest is:
  Unpolished 134.110527
  Treated and Polished 147.788117

- Standard Deviation of Zig dataest is:
  Unpolished 33.041804
  Treated and Polished 15.587355

2. Calculate the test statistic and p-value
   Unpolished -4.164630

Treated and Polished   -1.228911
array([4.17128700e-05, 1.11499484e-01])

Step 5: Decide to Reject or Fail to reject Null Hypothesis

- Perform a t-test
    - Level of significance: 0.05
    - We have evidence to reject the null hypothesis since p_value < Level of signifi cance
    - Our one-sample t-test p-value= 0.0014655150194628353

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

Calculate sample statistics

- Sample Mean of Unpolished is: 134.11052653373332
- Sample Mean of Treated and Polished is: 147.78811718133335
- The mean hardness of the polished and unpolished stones is not the same

Problem 6:

Aquarius health club, one of the largest and most popular cross-fit gyms in the country has been advertising a rigorous program for body conditioning. The program is considered successful if the candidate is able to do more than 5 push-ups, as compared to when he/she enrolled in the program. Using the sample data provided can you conclude whether the program is successful? (Consider the level of Significance as 5%)

Note that this is a problem of the paired-t-test. Since the claim is that the training will make a difference of more than 5, the null and alternative hypotheses must be formed accordingly.

Load our data set, Aquarius_gym.csv, and use the head() function to view the data. Find out the characteristics of the columns using the info() method and we can determine that there are 100 rows and 3 columns. The datatypes for the int64(3) columns are present.

- **head()** it given by default top five data

	Sr no.	Before	After
0	1	39	44
1	2	25	25
2	3	39	39
3	4	6	13
4	5	40	44

Table 4 : Top five data of dataset Aquarius_gym

- **info()** it tells a concise summary of a DataFrame

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 3 columns):
 #   Column   Non-Null Count  Dtype
---  ------   --------------  -----
 0   Sr no.   100 non-null    int64
 1   Before   100 non-null    int64
 2   After    100 non-null    int64
dtypes: int64(3)
memory usage: 2.5 KB
```

Table 5: Information about the Aquarius_gym structure and content.

Step 1: Define null and alternative hypotheses
In testing Since the claim is that the training will make a difference of more than 5,
The null hypothesis states that, The average difference in push-up counts before and after the program is less than or equal to 5. $\mu_0 \leq 5$.
The alternative hypthesis states that, The average difference in push-up counts before and after the program is greater than 5. $\mu_a > 5$
$H_0 : \mu_0 \leq 55$

$H_A : \mu_a > 55$

Step 2: Decide the significance level
Here we select $\alpha = 0.05$ as given in the question.

Step 3: Calculate the p - value and test statistic
We use the scipy.stats.ttest_rel to calculate the T-test on TWO RELATED samples of data. This is a two-sided test for the null hypothesis that 2 related or repeated samples have identical average (expected) values. Here we give the two sample observations as input. This function returns t statistic and two-tailed p value.

Sample Mean: 5.55
t-statistic: 1.9148542155126753
p-value: 0.058397744282022435
Critical value: 1.6603911559963895

Step 4: Decide to reject or accept null hypothesis
Conclusion: The null hypothesis cannot be rejected.
There is not enough evidence to conclude that the program is successful in terms of body conditioning.

Problem 7:

Dental implant data: The hardness of metal implant 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 on 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.

Load our data set, Dental Hardness data1.xlsx, and use the head() function to view the data. Find out the characteristics of the columns using the info() method and we can determine that there are 100 rows and 3 columns. The datatypes for the int64(3) columns are present.

- **head()** it given by default top five data

	Dentist	Method	Alloy	Temp	Response
0	1	1	1	1500	813
1	1	1	1	1600	792
2	1	1	1	1700	792
3	1	1	2	1500	907
4	1	1	2	1600	792

Table 6 : Top five data of dataset Dental Hardness

- **info()** it tells a concise summary of a DataFrame

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 90 entries, 0 to 89
Data columns (total 5 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   Dentist   90 non-null     int64
 1   Method    90 non-null     int64
 2   Alloy     90 non-null     int64
 3   Temp      90 non-null     int64
 4   Response  90 non-null     int64
dtypes: int64(5)
memory usage: 3.6 KB
```

Table 7: Information about the Dental Hardness structure and content.

- **Describe()** it tells summary of the central tendency, dispersion, and shape of the distribution of the data.

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

Table 8: Descriptive statistics of the Dental Hardness

1. Test whether there is any difference among the dentists on the implant hardness. State the null and alternative hypotheses. Note that both types of alloys cannot be considered together. You must state the null and alternative hypotheses separately for the two types of alloys.?

2. Before the hypotheses may be tested, state the required assumptions. Are the assumptions fulfilled? Comment separately on both alloy types.?

- Shapiro_Walk Test or Anderson Darling Test
    - The Shapiro_Walk test, Tests the Null Hypothesis that the data was drawn from a Normal Distribution.
    - Use level of significance $\alpha = 0.05$

Alloy 1 - Normality Test
Shapiro-Wilk test (p-value): 1.1945070582441986e-05
Anderson-Darling test (statistic, critical values): 2.561066309273201 [0.535 0.609 0.731 0.853 1.014]

Alloy 2 - Normality Test
Shapiro-Wilk test (p-value): 0.00040293222991749644
Anderson-Darling test (statistic, critical values): 1.8931311726356412 [0.535 0.609 0.731 0.853 1.014]

- Homogeniety and The Levene tests the Null Hypothesis that all input sample are from populations with equal variance

Levene's Test - Alloy 1
Levene's test (statistic, p-value): 1.4194717470917784 0.23669380462584474

Levene's Test - Alloy 2
Levene's test (statistic, p-value): 1.4194717470917784 0.23669380462584474

- The Levene's test will provide statistic and p-value.Here p-value is greater then α

1. **Define null and alternative hypotheses**

Test whether there is any difference among the dentists on the implant hardness. Let take two Alloys as A1 and A2

For A1:
Null Hypotheses states that, There is no difference among the dentists on the implant hardness for Alloy A1.
Alternative hypothesis state that, There is a difference among the dentists on the implant hardness for Alloy A1.
$H_0 : \mu_0 == A1$
$H_A : \mu_a \neq\neq A1$

For A2:
Null Hypotheses states that, There is no difference among the dentists on the implant hardness for Alloy A2.

Alternative hypothesis state that, There is a difference among the dentists on the implant hardness for Alloy A2.

$H_0 : \mu_0 == A2$

$H_A : \mu_a \neq\neq A2$

ANOVA Result for Alloy A1:

	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

let significance level α=0.05

p-value = 0.116567, which is greater than α, so we fail to reject the Null hypothesis. We conclude that There is no difference among the dentists on the implant hardness for Alloy A1.

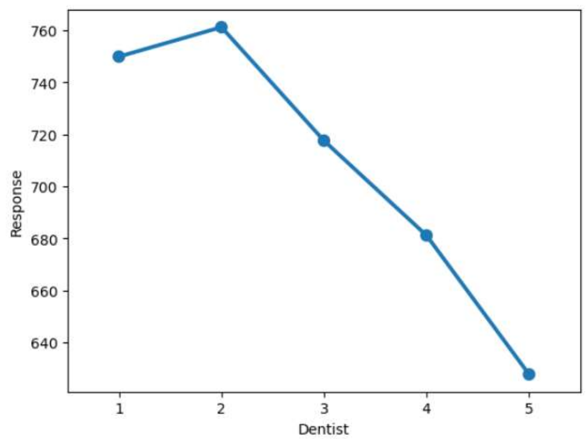

Fig 5 : Pointplot for implant hardness for Alloy A1

ANOVA Result for Alloy A2:

	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

let significance level α=0.05

p-value = 0.718031, which is greater than α, so we fail to reject the Null hypothesis. We conclude that There is no difference among the dentists on the implant hardness for Alloy A2.

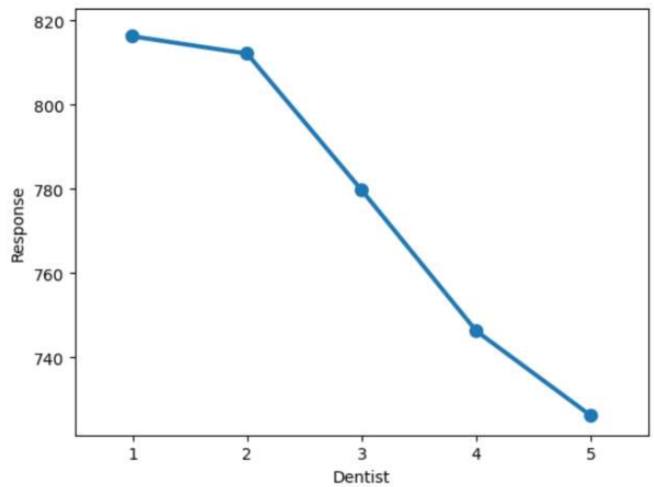

Fig 6 : Pointplot for implant hardness for Alloy A2

Insights: Here, the p-value is greater than the significance level (usually 0.05), so it fails to reject the null hypothesis and concludes that there is no significant difference in variance among alloys 1 and 2.

3. Irrespective of your conclusion in 2, we will continue with the testing procedure. What do you conclude regarding whether implant hardness depends on dentists? Clearly state your conclusion. If the null hypothesis is rejected, is it possible to identify which pairs of dentists differ?

Here, the p-value is greater than the significance level (α = 0.05), so it fails to reject the null hypotheses and concludes that there is no significant difference in variance among alloys 1 and 2. That means implant hardness is the same across all dentists for both alloys 1 and 2.

- If the null hypothesis is rejected, which means there is a significant difference, we can apply Tukey's test to identify which specific pairs of dentists differ significantly in terms of implant hardness. This test will help us compare the means of different dentists and determine which ones are significantly different from each other.
- **In above dataset null hypothesis is fail to rejected.**

4. Now test whether there is any difference among the methods on the hardness of dental implant, separately for the two types of alloys. What are your conclusions? If the null hypothesis is rejected, is it possible to identify which pairs of methods differ?

Here, the p-value is greater than the significance level (α = 0.05), so it fails to reject the null hypotheses

5. Now test whether there is any difference among the temperature levels on the hardness of dental implant, separately for the two types of alloys. What are your conclusions? If the null hypothesis is rejected, is it possible to identify which levels of temperatures differ?

Define null and alternative hypotheses

Test whether there is any difference among the temperature levels on the implant hardness. Let take two Alloys as A1 and A2

For A1:

Null Hypotheses states that, There is no difference among the temperature levels on the implant hardness for Alloy A1.

Alternative hypothesis state that, There is a difference among the temperature levels on the implant hardness for Alloy A1.

$H_0 : \mu == TA1$

$H_a : \mu \# \# TA1$

For A2:

Null Hypotheses states that, There is no difference among the temperature levels on the implant hardness for Alloy A2.

Alternative hypothesis state that, There is a difference among the temperature levels on the implant hardness for Alloy A2.

$H_0 : \mu == TA2$

$H_a : \mu \# \# TA2$

ANOVA Test - Alloy 1

ANOVA test (F-statistic, p-value): 0.3352235344077172 0.7170741113686678

ANOVA Test - Alloy 2

ANOVA test (F-statistic, p-value): 1.883492290995591 0.16467846603141556

Insights:

- The p-value of ANOVA Test - Alloy 1 is 0.71707, Which is greater than the significance level (α= 0.05), so it fails to reject the null hypotheses.
- The p-value of ANOVA Test - Alloy 2 is 0.16467, Which is greater than the significance level (α = 0.05), so it fails to reject the null hypotheses

Conclude that there is no significant difference in variance among alloys 1 and 2. That means the temperature levels of dental implants hardness for alloys 1 and 2 are the same.

6. Consider the interaction effect of dentist and method and comment on the interaction plot, separately for the two types of alloys?

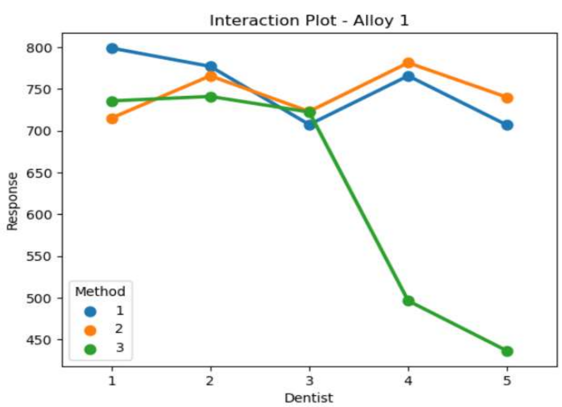

Fig 7 :Interaction Plot – Alloy1

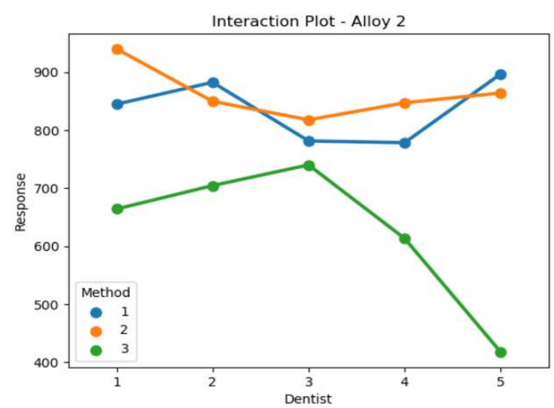

Fig 8 :Interaction Plot – Alloy1

There is a significant interaction effect between dentist and method on implant hardness of Alloy1 for Method 1 to 3 but in Alloy 2 for Method 1 and 2. Hence, it indicates an interaction effect.

7. Now consider the effect of both factors, dentist, and method, separately on each alloy. What do you conclude? Is it possible to identify which dentists are different, which methods are different, and which interaction levels are different?

Two-way ANOVA for Alloy 1:

	sum_sq	df	F	PR(>F)
Dentist	94802.677778	1.0	9.160485	0.004212
Method	116812.800000	1.0	11.287254	0.001670
Residual	434661.766667	42.0	NaN	NaN

The null hypothesis rejected for both factors Dentist and Method because the p-values repectivily 0.0042 and 0.0016, both are less than the significance level (α = 0.05) that means there are significant differences among Dentists and Method regarding the implant hardness for Alloy 1.

Two-way ANOVA for Alloy 2:

	sum_sq	df	F	PR(>F)
Dentist	54513.611111	1.0	3.022503	0.089443
Method	326980.800000	1.0	18.129428	0.000113
Residual	757508.388889	42.0	NaN	NaN

Dentist do not have a significant impact on the implant hardness because p-value 0.0894 is greater than signification level 0.05, and the method used for the implant has a significant effect on the implant hardness because p-value 0.0003 is less than signification level 0.05 for Alloy 2. So, we perform the interaction between dentist and method.

Two-way ANOVA for Alloy 2:

	sum_sq	df	F	PR(>F)
Dentist	54513.611111	1.0	3.163974	0.082695
Method	326980.800000	1.0	18.977993	0.000086
Dentist:Method	51100.016667	1.0	2.965849	0.092577
Residual	706408.372222	41.0	NaN	NaN

Here, the Interaction of Dentist:Method's p-value of 0.0925 is greater than the signification level of 0.05, so it does not significantly influence the implant hardness for Alloy 2