ADVANCED STATISTICS PROJECT

Submitted by,
VIDYA V
PGPDSBA.O.2023.B
03.06.2023

CONTENTS

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)** 8

3.1 What proportion of the gunny bags have a breaking strength less than 3.17 kg per sq cm? 8

3.2 What proportion of the gunny bags have a breaking strength at least 3.6 kg per sq cm.? 9

3.3 What proportion of the gunny bags have a breaking strength between 5 and 5.5 kg per sq cm.? 9

3.4 What proportion of the gunny bags have a breaking strength NOT between 3 and 7.5 kg per sq cm.? 10

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. 10

4.1 What is the probability that a randomly chosen student gets a grade below 85 on this exam? 10

4.2 What is the probability that a randomly selected student scores between 65 and 87? 11

4.3 What should be the passing cut-off so that 75% of the students clear the exam? 11

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); 12

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 12 believe now that the unpolished stones may not be suitable for printing. Do you think Zingaro is justified in thinking so?

5.2. Is the mean hardness of the polished and unpolished stones the same? 12

12

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.

13

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.

1. Test whether there is any difference among the dentists on the implant hardness. 13 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 14 assumptions fulfilled? Comment separately on both alloy types.?

3. Irrespective of your conclusion in 2, we will continue with the testing procedure. 22 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?

4. Now test whether there is any difference among the methods on the hardness of 23 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?

5. Now test whether there is any difference among the temperature levels on the 24 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?

6. Consider the interaction effect of dentist and method and comment on the 25
interaction plot, separately for the two types of alloys?

7. Now consider the effect of both factors, dentist, and method, separately on each 26
alloy. What do you conclude? Is it possible to identify which dentists are different, which methods are different, and which interaction levels are different?

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?

P(Injury)=Players Injured/Total Players
= 145/235
= 0.62

The probability that a randomly chosen player would suffer an injury is 0.62.

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

P(Forward or winger)= (Winger + Forward)/ Total players
= (56+20)/235
= 0.32

The probability that a player is a forward or a winger is 0.32.

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

P(striker and injured)= Injured striker/ Total Players
=45/235
= 0.19

The probability that a randomly chosen player plays in a striker position and has a foot injury is 0.19.

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

P(Injured striker) = Striker/ Total Injured Players
= 45/145
=0.31

The probability that a randomly chosen injured player is a striker is 0.31.

1.5. What is the probability that a randomly chosen injured player is either a forward or an attacking midfielder?

P(Injured Forward/attacking midfielder)= (Forward + attacking midfielder)/Total injured players
= (56+24) / 145
=0.55

The probability that a randomly chosen injured player is either a forward or an attacking midfielder is 0.55.

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?

This is an example of conditional probability. The formula is:

- P (Fire)= P (Radiation and fire)/ P (Radiation given fire)
- P (Mechanical failure) = P (Radiation and mechanical failure)/ P (Radiation given mechanical failure)
- P (Human error) = P (Radiation and human error)/ P (Radiation given human error)

Substituting the values, the probabilities of a fire, a mechanical failure, and a human error are 0.005, 0.003, 0.012 respectively.

2.2. What is the probability of a radiation leak?

P(Radiation)= P(Radiation given fire)* P(Fire) + P(Radiation given mechanical failure) * P(Mechanical failure) + P(Radiation given human error) * P(Human error)

Substituting the values, the probability of a radiation leak is 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.**
- **A Mechanical Failure.**
- **A Human Error.**

Applying Bayes' theorem to the current problem,

P(Fire given radiation) = P(radiation given fire) * P(fire) / P(Radiation)
P(Mechanical failure given radiation) = P(radiation given mechanical failure) * P(Mechanical failure) / P(Radiation)
P(Human error given radiation) = P(radiation given human error) * P(human error) / P(Radiation)

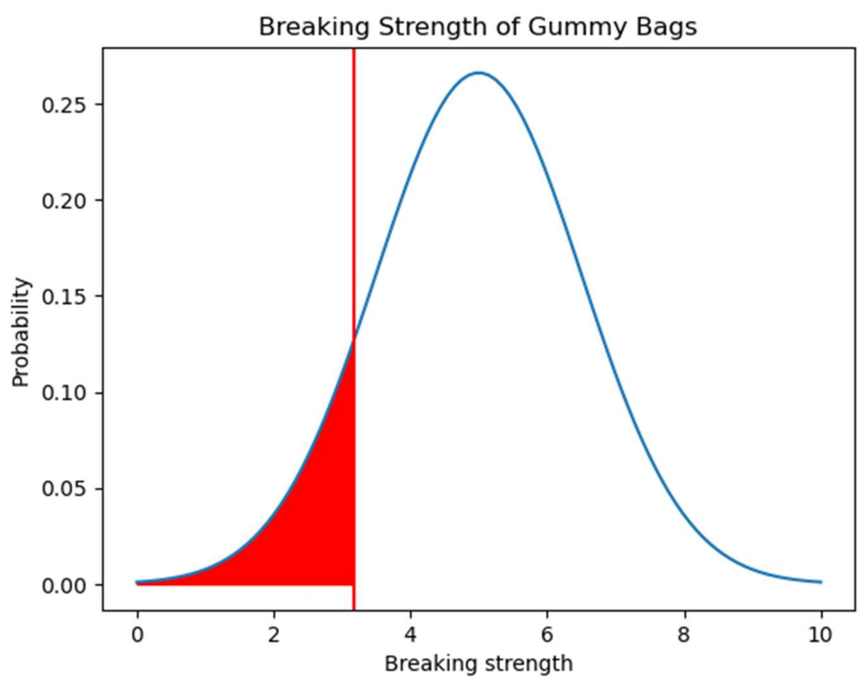
Substituting the values,

The probability of the radiation leak caused by fire is 0.27
The probability of the radiation leak caused by mechanical failure is 0.405
The probability of the radiation leak caused by human error is 0.324

Problem 3:

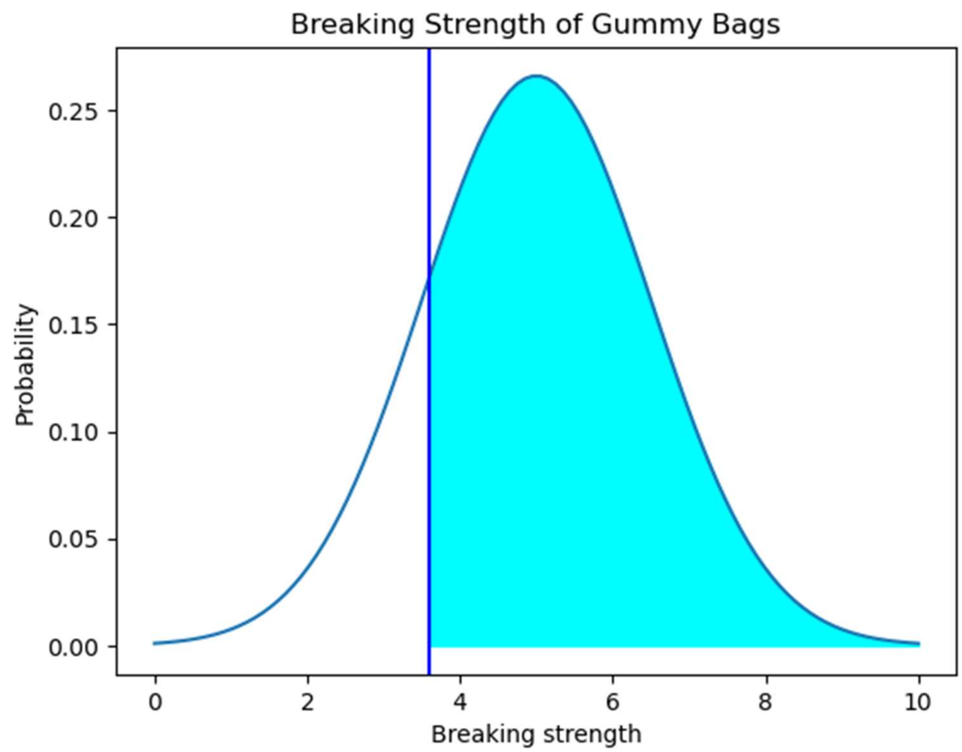
The breaking strength of gunny bags used for packaging cement is normally distributed with a mean of 5 kg per sq. centimeter and a standard deviation of 1.5 kg per sq. centimeter. The quality team of the cement company wants to know the following about the packaging material to better understand wastage 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?



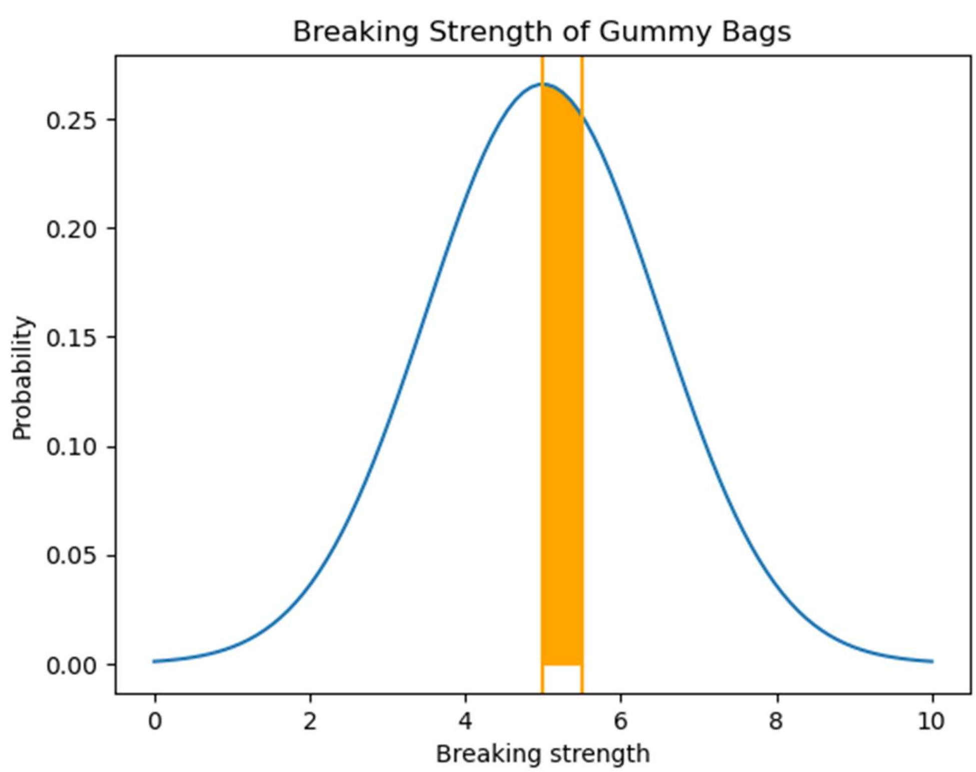
The proportion of gunny bags having breaking strength less than 3.17 kg per sq cm is 11.12 %.

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



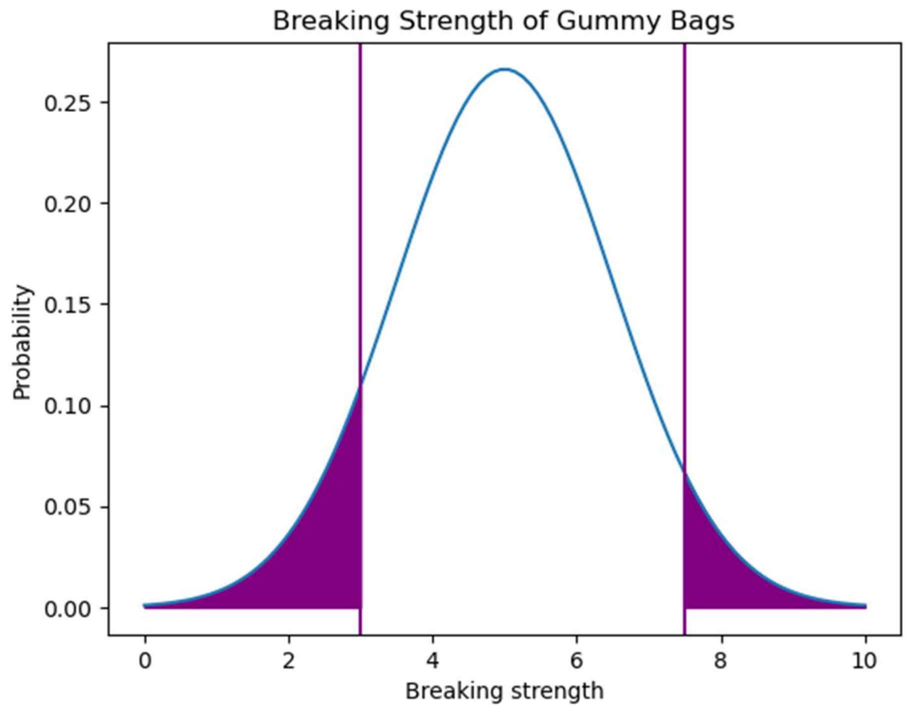
The proportion of gunny bags having breaking strength at least 3.6 kg per sq cm is 82.47%.

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



The proportion of gunny bags having breaking strength between 5 to 5.5 kg per sq cm is 13.06 %.

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

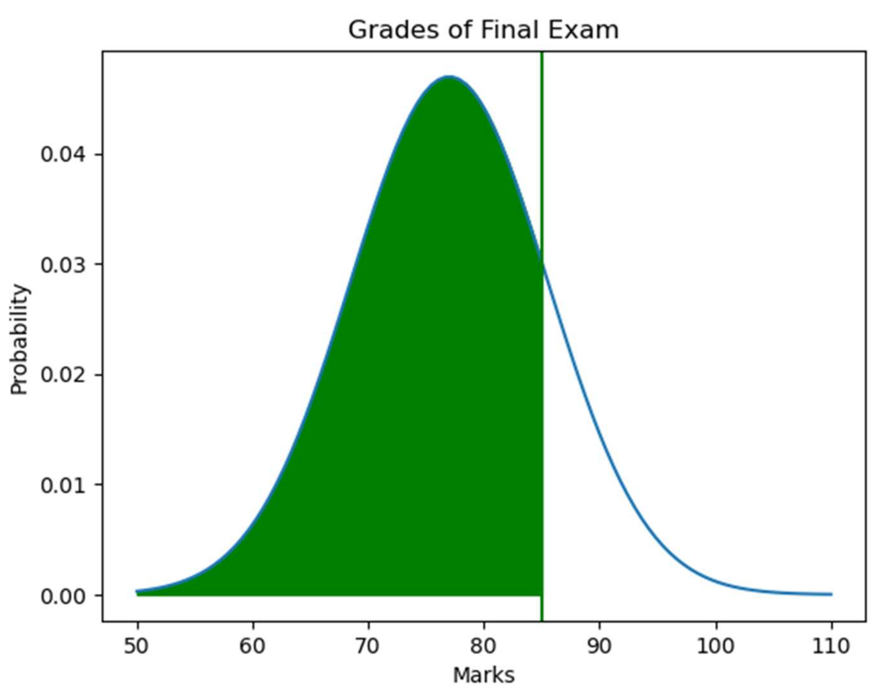


The proportion of gunny bags having breaking strength not between 3 and 7.5 kg per sq cm is 13.9%.

Problem 4:

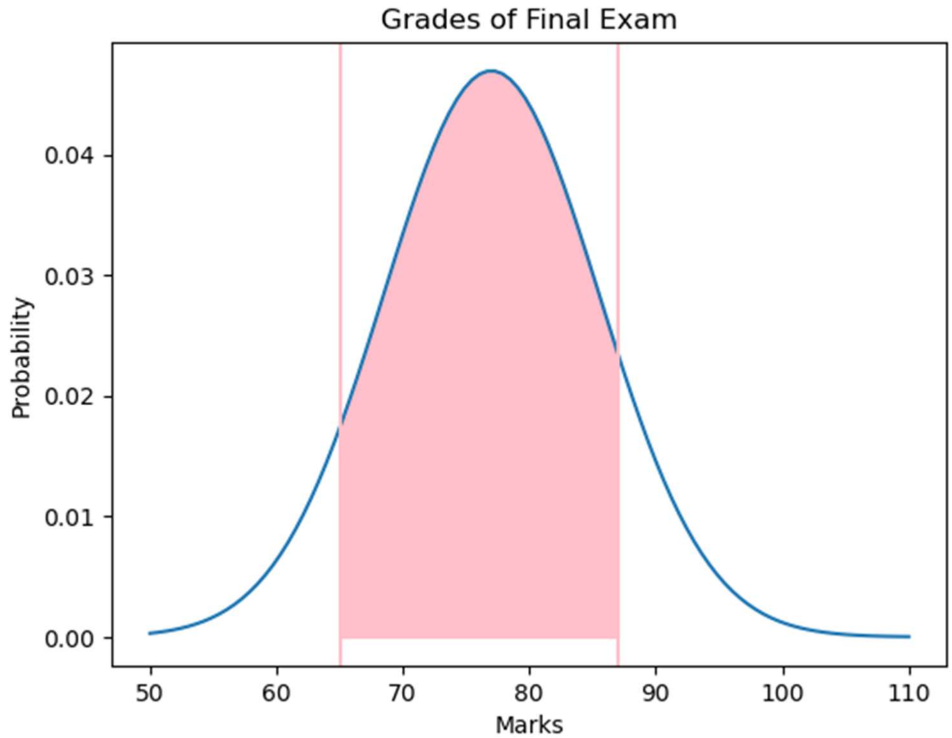
Grades of the final examination in a training course are found to be normally distributed, with a mean of 77 and a standard deviation of 8.5. Based on the given information answer the questions below.

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



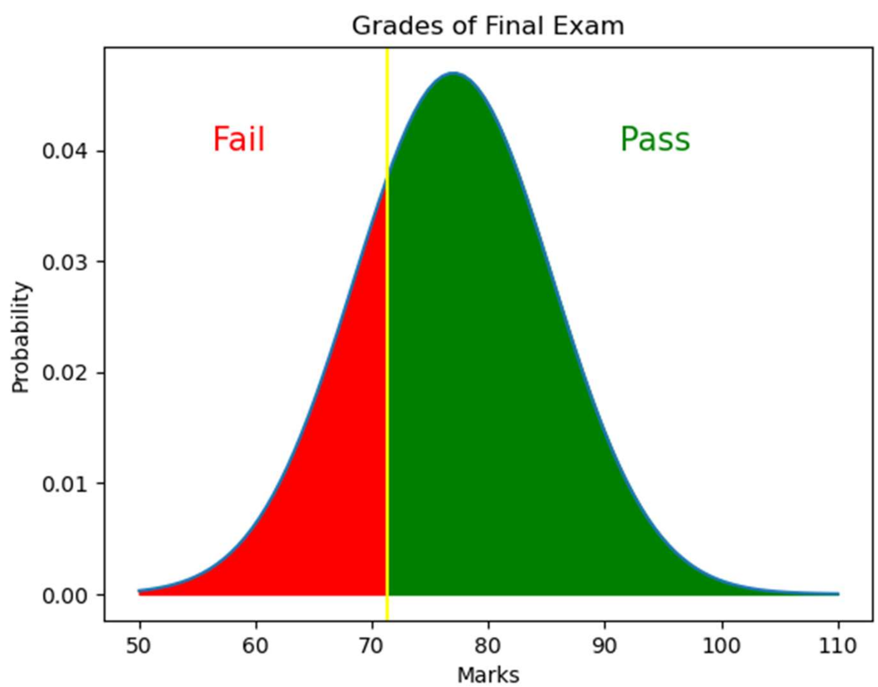
The probability that a randomly chosen student gets a grade below 85 on this exam is 0.83.

4.2. What is the probability that a randomly selected student scores between 65 and 87?



The probability that a randomly selected student scores between 65 and 87 is 0.8.

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



The passing cut-off so that 75% of the students clear the exam is 71.27.

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);

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?**

This is an example of one tail, one sample t test.
Stating Hypothesis:
Null hypothesis: The unpolished stones are suitable for printing
H0: Hardness>=150
Alternate Hypothesis: The unpolished stones are not suitable for printing
H1: Hardness<150
Alpha: 0.05.

After performing the test, p value for the one tail test is found to be 4.17e-05.

Conclusion: Since, p value is less than alpha, null hypothesis is rejected. Considering a confidence of 95%, the unpolished stones are not suitable for printing and the statement of the company is right.

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

This is an example of independent two sample t test.

Stating the hypothesis:
Null Hypothesis: The mean hardness of the polished and unpolished stones are the same
H0: Polished mean = Unpolished mean
Alternate Hypothesis: The mean hardness of the polished and unpolished stones are not the same
H1: Polished mean not = Unpolished mean
Alpha = 0.05.

After performing the test, the p value is 0.0014.

Conclusion: Since p value is less than alpha, null hypothesis is rejected. Considering a confidence of 95%, the mean hardness of polished and unpolished stones is not equal.

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.

This is an example of one tail, Paired t test

Establishing hypothesis:

Null Hypothesis: The difference betwen the mean number of sit-ups done by the candidates after and before the training is less than 5

H0: mu_after-mu_before < 5

Alternative hypothesis: The difference betwen the mean number of sit-ups done by the candidates after and before the training is atleast 5

H1: H0: mu_after-mu_before >= 5

Alpha=0.05.

After performing the test, the p value was found to be 1.14e-35

Conclusion: Since the p value is less than alphs, null hypothesis is rejected. At 95% confidence, the difference between the mean number of sit-ups done by the candidates after and before the training is atleast 5. Hence, the program is successful.

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.

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.?**

   Alloy 1:

   Null Hypothesis
   H0: The mean responses for different dentists is the same for alloy 1
   Alternate Hypothesis
   H1: The mean response is different from the rest for atleast one dentist for alloy 1

   Alloy 2:

Null Hypothesis

H0: The mean responses for different dentists is the same for alloy 2

Alternate Hypothesis

H1: The mean response is different from the rest for atleast one dentist for alloy 2

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

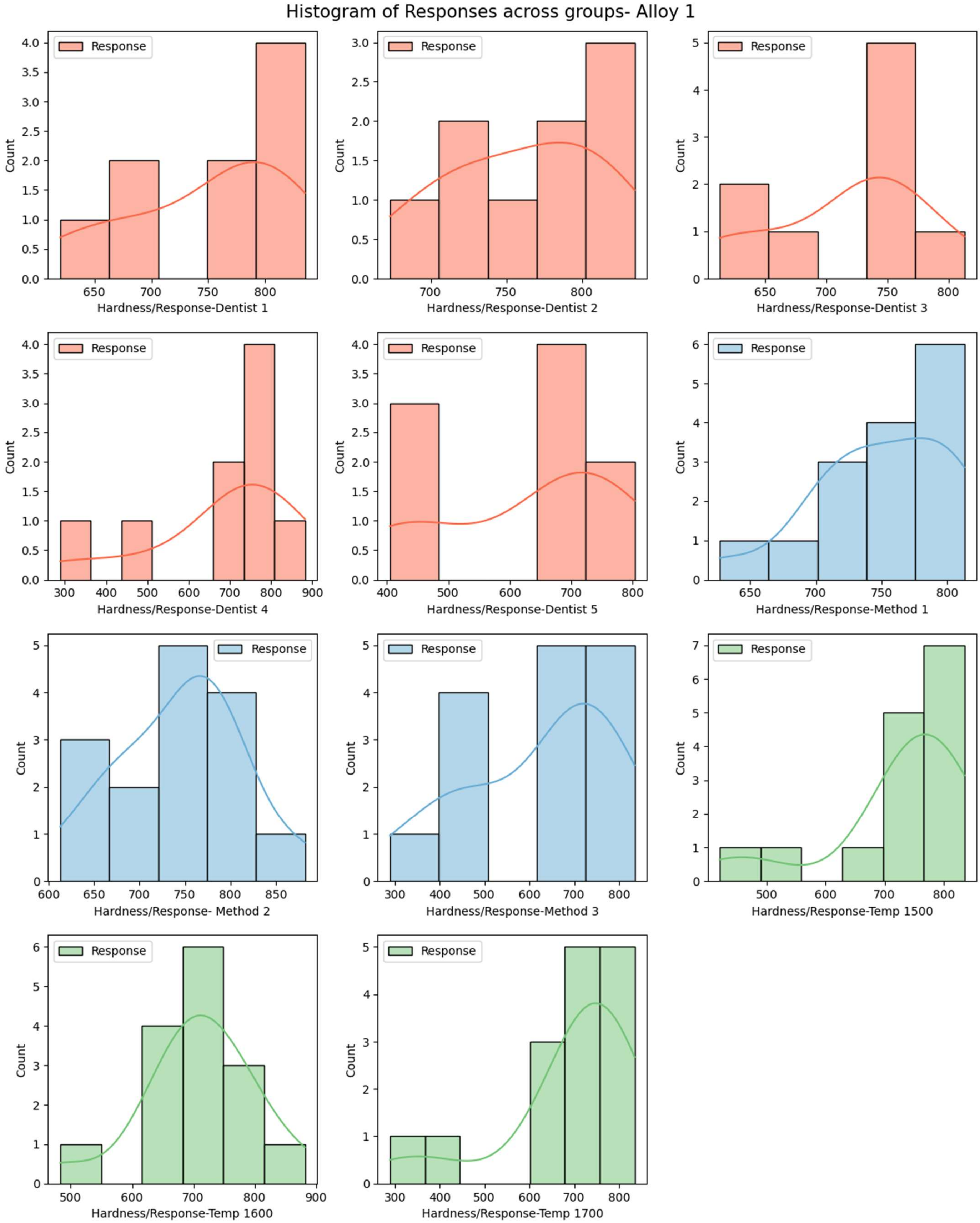
Assumptions of ANOVA:

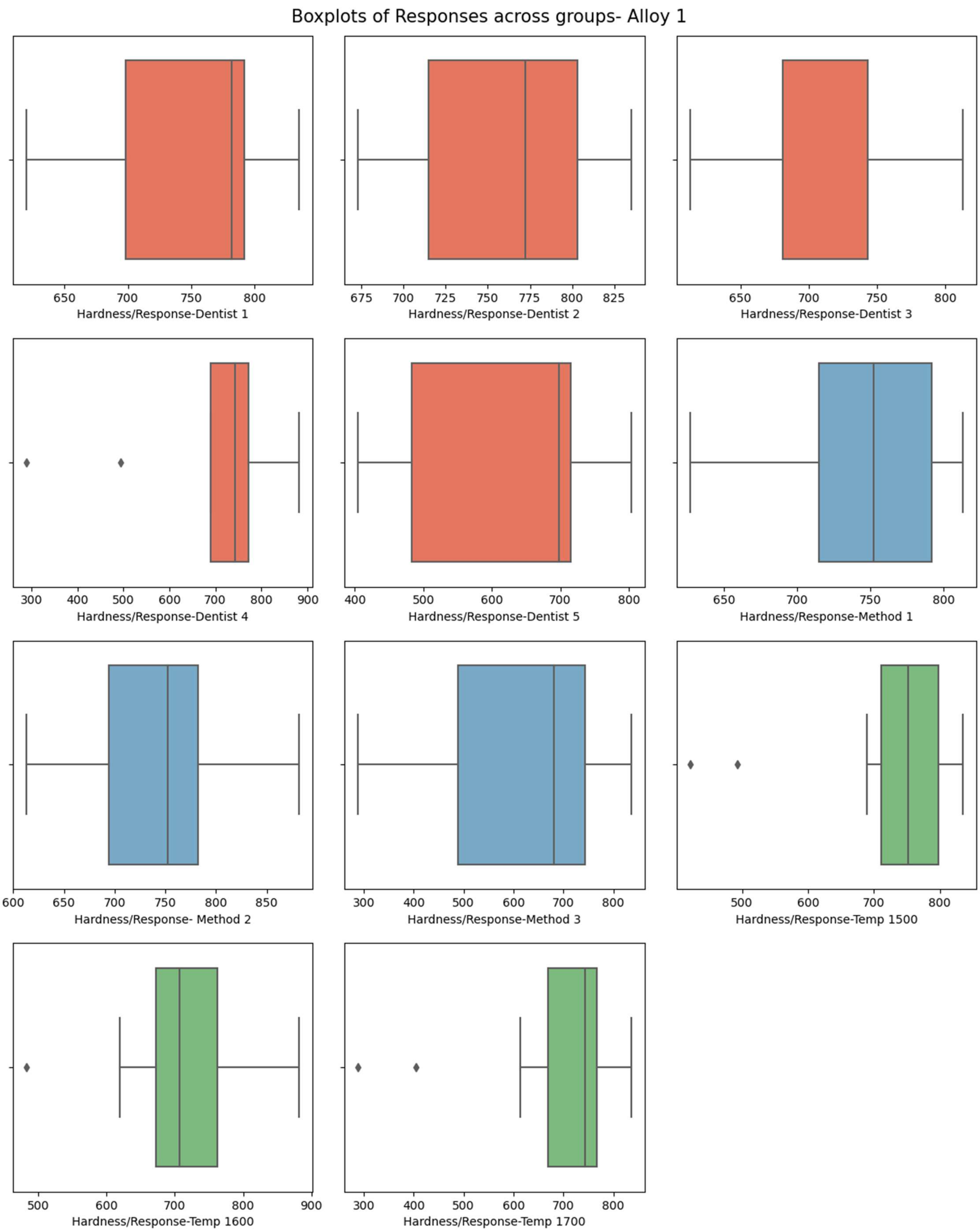
- Independent Sample - Sample should be selected randomly (Equally likely events). There should not be any pattern in the selection of sample
- Normal Distribution - Distribution of each group should be normal
- Homogenous Group - Variance between the group should be the same.

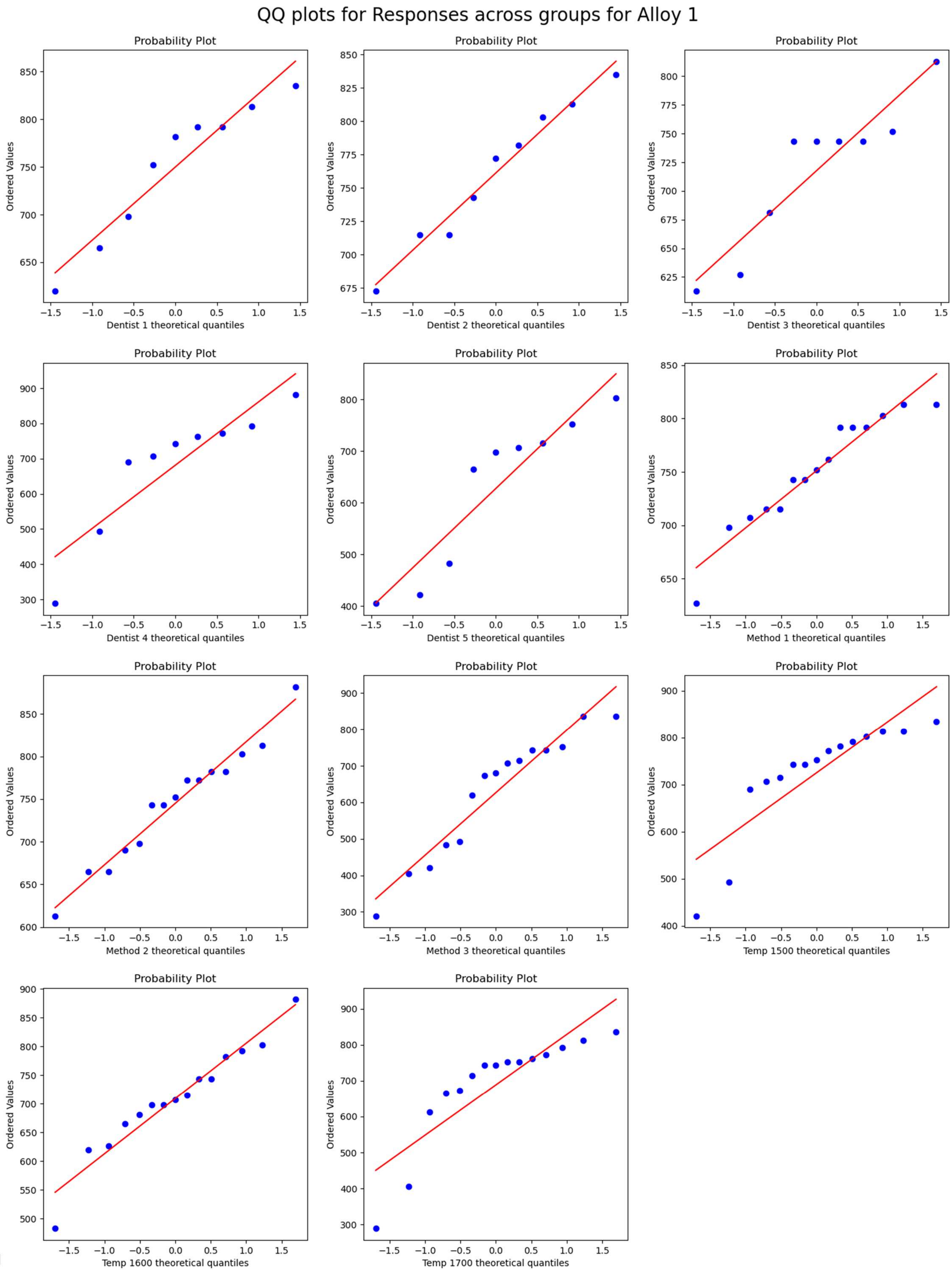
On conducting various tests, the following were the results obtained:

Alloy 1:

Normality test:




Boxplots of Responses across groups- Alloy 1
Hardness/Response-Dentist 1
Hardness/Response-Dentist 2
Hardness/Response-Dentist 3
Hardness/Response-Dentist 4
Hardness/Response-Dentist 5
Hardness/Response-Method 1
Hardness/Response- Method 2
Hardness/Response-Method 3
Hardness/Response-Temp 1500
Hardness/Response-Temp 1600
Hardness/Response-Temp 1700


QQ plots for Responses across groups for Alloy 1
Probability Plot
Ordered Values
Dentist 1 theoretical quantiles
Probability Plot
Ordered Values
Dentist 2 theoretical quantiles
Probability Plot
Ordered Values
Dentist 3 theoretical quantiles
Probability Plot
Ordered Values
Dentist 4 theoretical quantiles
Probability Plot
Ordered Values
Dentist 5 theoretical quantiles
Probability Plot
Ordered Values
Method 1 theoretical quantiles
Probability Plot
Ordered Values
Method 2 theoretical quantiles
Probability Plot
Ordered Values
Method 3 theoretical quantiles
Probability Plot
Ordered Values
Temp 1500 theoretical quantiles
Probability Plot
Ordered Values
Temp 1600 theoretical quantiles
Probability Plot
Ordered Values
Temp 1700 theoretical quantiles

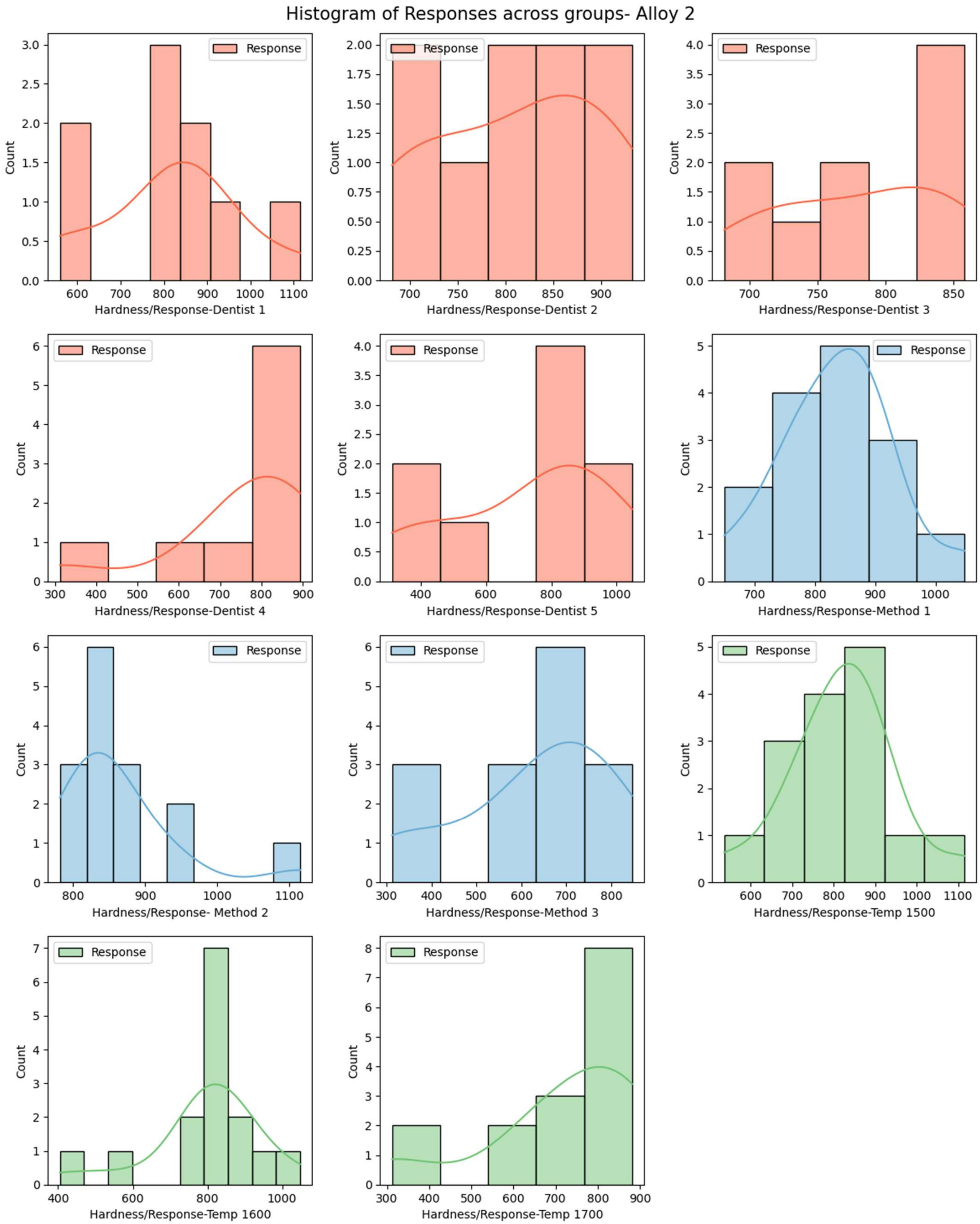
- Post the plotting of graphs, Shapiro and Anderson Darling tests for alloy 1 gave the following results
  - Shapiro test for normality had a p value less than alpha, and hence failed to prove normality
  - From the Anderson- Darling test result, if the test statistic is less than the critical value for a significance of 5%, we can fail to reject normal, i.e., the distribution is normal. Thus, for the given groups using Alloy 1, the following are the results
    - dentist1- normal
    - - dentist2- normal
    - - dentist3- normal
    - - dentist4- normal
    - - dentist5- normal
    - - method1- normal
    - - method2- normal
    - - method3- normal
    - - temp1500- not normal
    - - temp1600- normal
    - - temp1700- not normal
  - Hence, the ANOVA assumptions of normality fails for the above groups.

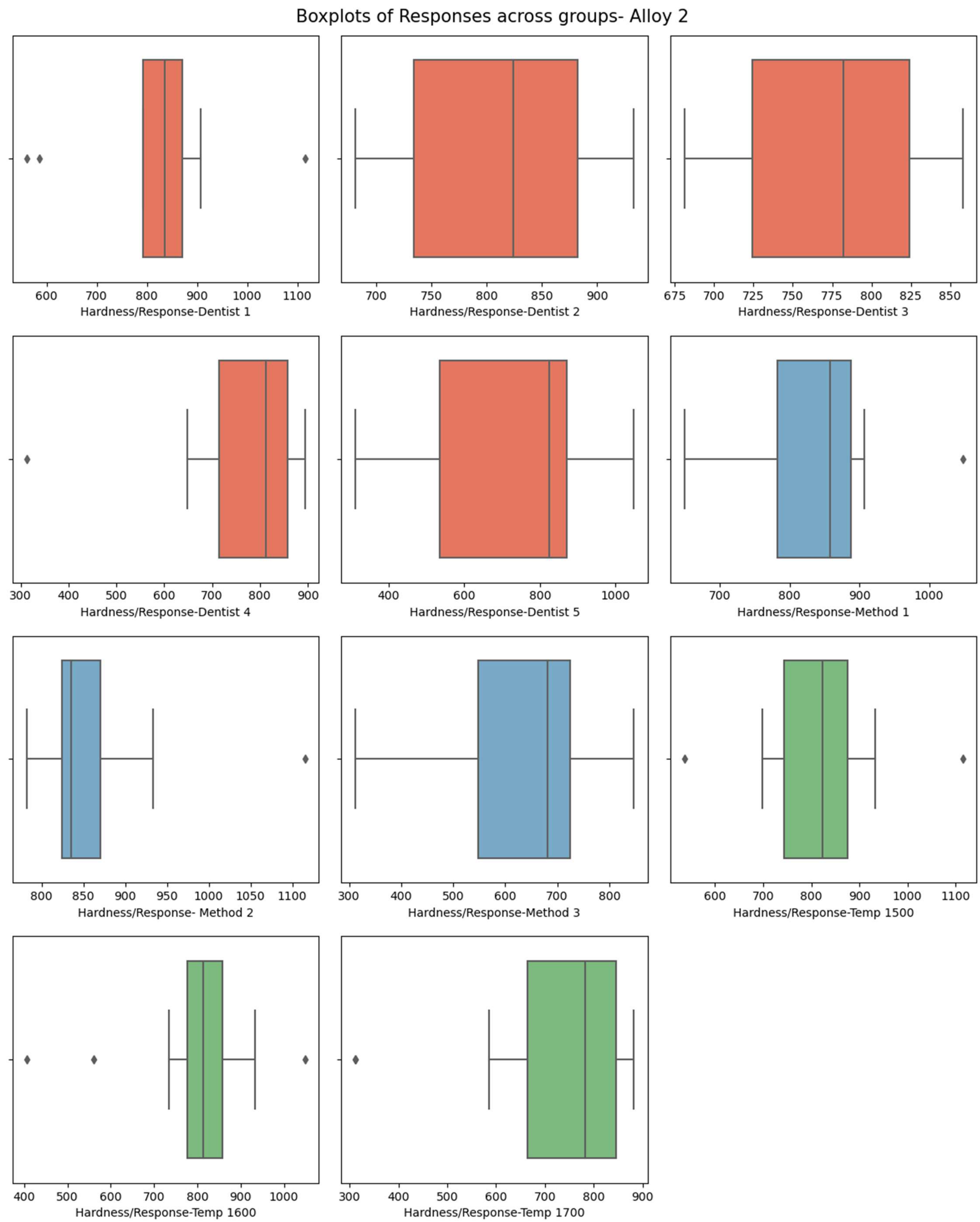
Homogeneity:

- **Null hypothesis H0:** All input samples are from population with different variances
- **Alternate Hypothesis H1:** The input samples are from populations with equal variances
- **Output:**
  - LeveneResult(statistic=1.3847146992797106, pvalue=0.2565537418543795)
  - LeveneResult(statistic=6.52140454403598, pvalue=0.0034160381460233975)
  - LeveneResult(statistic=0.26470963952464094, pvalue=0.7686994896007937)
- For a confidence of 95% if p value is > 0.05, we fail to reject null.
- Hence, based on the above output of the Levene's test, there is homogeneity across dentist groups and temperature groups. However, the homogeneity cannot be proven across the method groups. hence, the ANOVA assumption of homogeneity does not hold good.

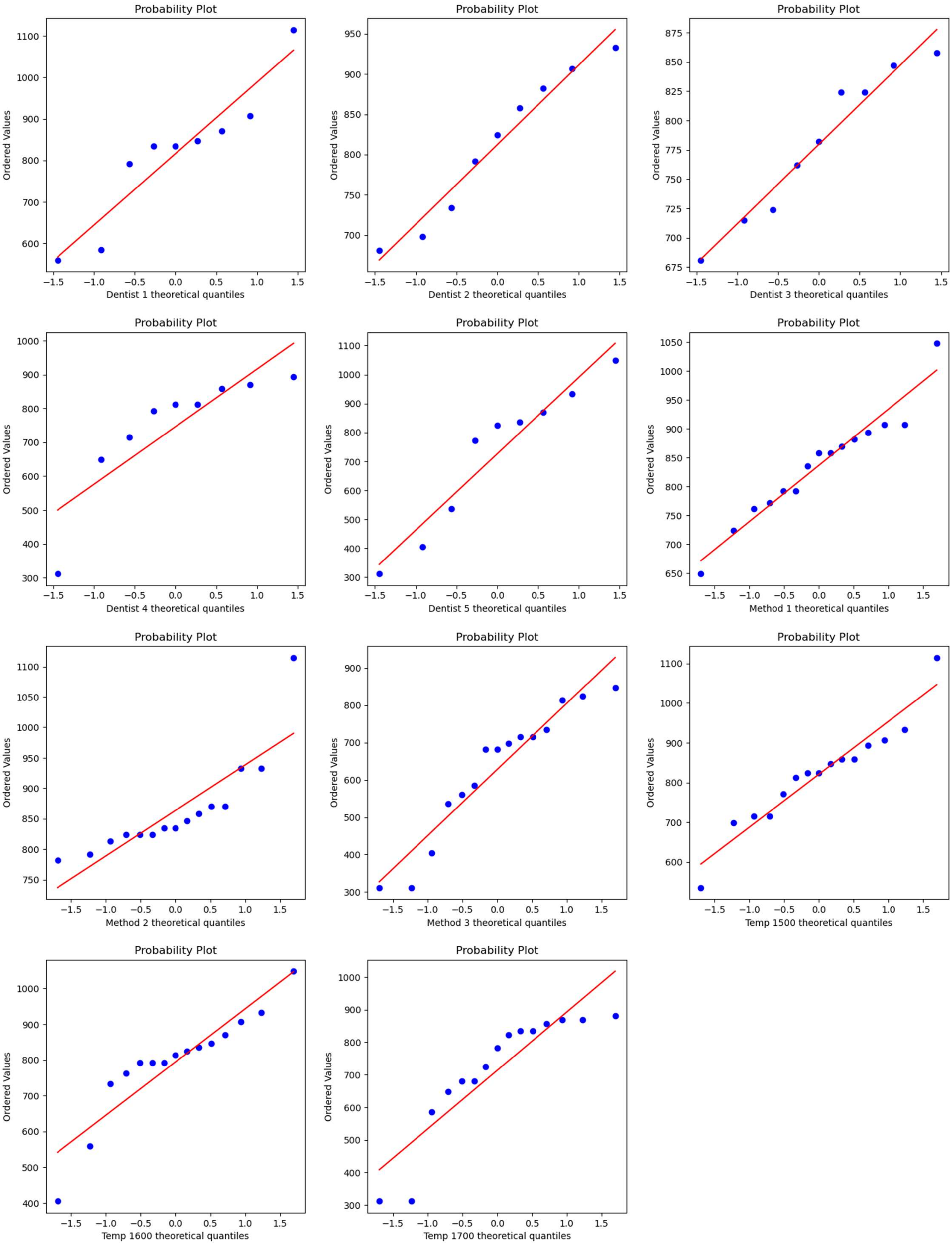
Alloy 2:

Normality:




Boxplots of Responses across groups- Alloy 2
Hardness/Response-Dentist 1
Hardness/Response-Dentist 2
Hardness/Response-Dentist 3
Hardness/Response-Dentist 4
Hardness/Response-Dentist 5
Hardness/Response-Method 1
Hardness/Response- Method 2
Hardness/Response-Method 3
Hardness/Response-Temp 1500
Hardness/Response-Temp 1600
Hardness/Response-Temp 1700

QQ plots for Responses across groups for Alloy 2



- Post the plotting of graphs, Shapiro and Anderson Darling tests for alloy 1 gave the following results
  - Shapiro test for normality had a p value less than alpha, and hence failed to prove normality
  - From the Anderson- Darling test result, if the test statistic is less than the critical value for a significance of 5%, we can fail to reject normal, i.e., the distribution is normal. Thus, for the given groups using Alloy 2, the following are the results
    - dentist1- normal
    - dentist2- normal
    - dentist3- normal
    - dentist4- not normal
    - dentist5- normal
    - method1- normal
    - method2- normal
    - method3- normal
    - temp1500- normal
    - temp1600- not normal
    - temp1700- normal
  - Hence, the ANOVA assumption of normality fails for the said groups.

Homogeneity:

- **Null hypothesis H0:** All input samples are from population with different variances
- **Alternate Hypothesis H1:** The input samples are from populations with equal variances
- **Output:**
  - LeveneResult(statistic=1.4456166464566966, pvalue=0.23686777576324952)
  - LeveneResult(statistic=3.349707184158617, pvalue=0.04469269939158668)
  - LeveneResult(statistic=0.6697956965974987, pvalue=0.5171946653062957)
- For a confidence of 95% if p value is > 0.05, we fail to reject null.
- Hence, based on the above output of the Levene's test, there is homogeneity across dentist groups and temperature groups. However, the homogeneity cannot be proven across the method groups. Hence, the ANOVA assumptions do not hold good.

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?**
   - **Null Hypothesis:**
     - H0: The mean response,i.e., hardness of implants is the same for different type of dentists
   - **Alternate Hypothesis:**
     - H1: The mean response,i.e., hardness of implants is different for atleast 1 dentist
   - **Alloy 1:**
     - **Output:**

	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

- **Conclusion:** Since the p value is greater than the significance level (0.05), we fail to reject the null hypothesis and thus conclude that the mean response, that is, the hardness of metal implant is the same for all 5 categories of dentists when using alloy 1.
- If the null hypothesis was rejected, we might be able to find the different pair by Tukey's significance test.

- **Alloy 2:**
  - **Output:**

	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

  - **Conclusion:** Since the p value is greater than the significance level (0.05), we fail to reject the null hypothesis and thus conclude that the mean response, that is, the hardness of metal implant is the same for all 5 categories of dentists when using alloy 2.
  - If the null hypothesis was rejected, we might be able to find the different pair by Tukey's significance test.

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?**

Null Hypothesis:
H0: The mean response, i.e. hardness of implants is the same for different type of Methods
Alternate Hypothesis:
H1: The mean response, i.e., hardness of implants is different for atleast 1 method.

- **Alloy 1:**
  - **Output**:

	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

  - **Conclusion:** Since the p value is less than the significance level (0.05), we reject the null hypothesis and thus conclude that the mean response, that is, the hardness of metal implant is the different for atleast one pair of methods when using alloy 1.
  - Since the null hypothesis is rejected, we can find the pair that causes the difference by Tukey Honest significant difference test
  - **Tukey test:**
    - **Null hypothesis: -** H0: The pairwise group means are equal
    - **Alternate Hypothesis: -** H1: The pairwise group means are not equal
    - **Output:**

```
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
=======================================================
group1 group2  meandiff p-adj    lower     upper  reject
-------------------------------------------------------
     1      2   -6.1333  0.987  -102.714  90.4473  False
     1      3    -124.8 0.0085 -221.3807 -28.2193   True
     2      3 -118.6667 0.0128 -215.2473  -22.086   True
-------------------------------------------------------
```

- **Conclusion:** For the pair 1&2, p>0.05, which means that we fail to reject null. Thus, while methods 1&3, and 2&3 show significant difference, methods 1&2 do not for alloy 1.

- **Alloy 2:**
  - **Output:**

```
             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
```

  - **Conclusion:** Since the p value is less than the significance level(0.05), we reject the null hypothesis and thus conclude that the mean response, that is, the hardness of metal implant is the different for atleast one pair of methods when using alloy 2.
  - Since the null hypothesis is rejected, we can find the pair that causes the difference by Tukey Honest significant difference test
  - **Tukey Test:**
    - **Null hypothesis: -** H0: The pairwise group means are equal
    - **Alternate Hypothesis: -** H1: The pairwise group means are not equal
    - **Output:**

```
 Multiple Comparison of Means - Tukey HSD, FWER=0.05
======================================================
group1 group2 meandiff p-adj    lower     upper   reject
------------------------------------------------------
     1      2     27.0 0.8212  -82.4546  136.4546  False
     1      3   -208.8 0.0001 -318.2546  -99.3454   True
     2      3   -235.8    0.0 -345.2546 -126.3454   True
------------------------------------------------------
```

    - **Conclusion:** For the pair 1&2, p>0.05, which means that we fail to reject null. Thus, while methods 1&3, and 2&3 show significant difference, methods 1&2 do not for alloy 2.

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?**

   Null Hypothesis:
   H0: The mean response, i.e. hardness of implants is the same for different Temperatures
   Alternate Hypothesis:
   H1: The mean response, i.e., hardness of implants is different for ateast 1 particular temperature

- **Alloy 1:**
  - **Output:**

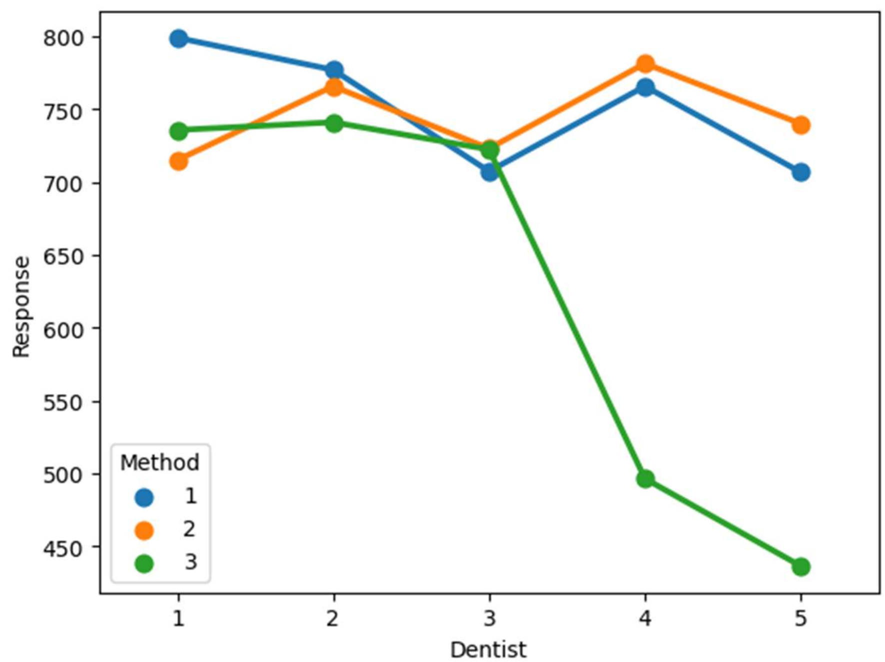
	df	sum_sq	mean_sq	F	PR(>F)
C(Temp)	2.0	10154.444444	5077.222222	0.335224	0.717074
Residual	42.0	636122.800000	15145.780952	NaN	NaN

  - **Conclusion:** Since the p value is greater than 0.05, we fail to reject the null hypothesis and thus conclude that the response, that is the mean hardness of dental implants is the same across different temperatures for alloy 1
- **Alloy 2:**
  - **Output:**

	df	sum_sq	mean_sq	F	PR(>F)
C(Temp)	2.0	9.374893e+04	46874.466667	1.883492	0.164678
Residual	42.0	1.045254e+06	24886.996825	NaN	NaN

  - **Conclusion:** Since the p value is greater than 0.05, we fail to reject the null hypothesis and thus conclude that the response, that is the mean hardness of dental implants is the same across different temperatures for alloy 2

6. **Consider the interaction effect of dentist and method and comment on the interaction plot, separately for the two types of alloys?**
- **Alloy 1:**
  - **Interaction Plot:**

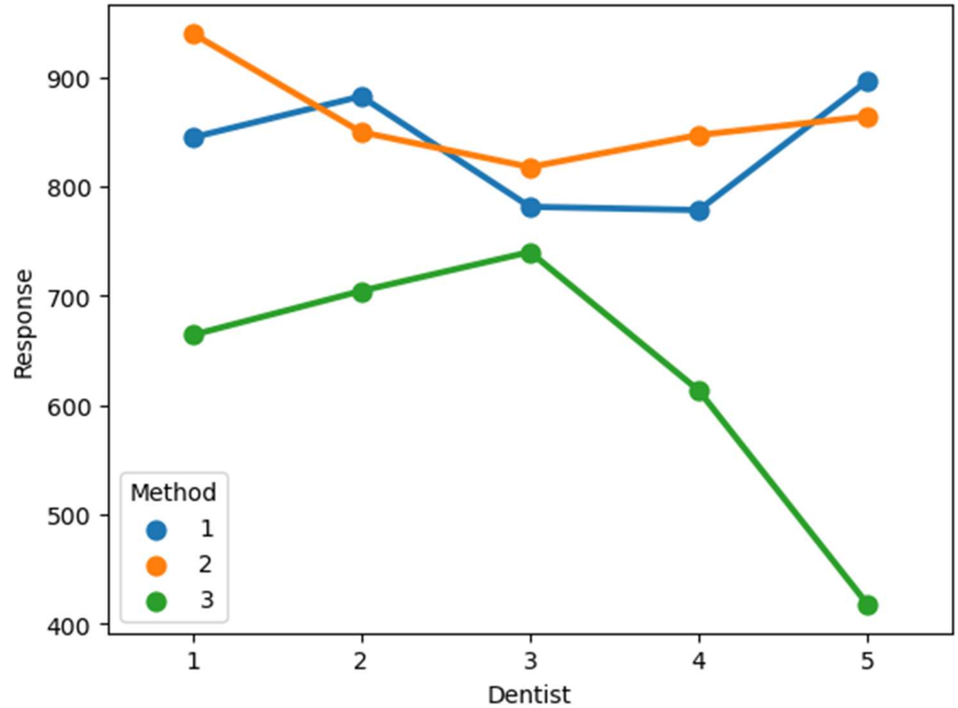


  - **ANOVA Output:**

	df	sum_sq	mean_sq	F	PR(>F)
C(Method):C(Dentist)	14.0	441097.244444	31506.946032	4.606728	0.000221
Residual	30.0	205180.000000	6839.333333	NaN	NaN

  - **Conclusion:** Now, from the result above, we observe that with the interaction of the two,for alloy 1, i.e., p(Dentist:Method) = 0.00022. Hence, we can reject the null hypothesis and conclude that the interaction between the dentists and the methods has an effect on the response variable, i.e., the dental implant hardness.

- **Alloy 2:**
  - **Interaction Plot:**



- From the plot above, we can see that there is an interaction effect of method 2 with dentist 1 and 2 and dentist 4 and 5.
- The same can be established by two way ANOVA with interaction as below

  - **ANOVA Output:**

	df	sum_sq	mean_sq	F	PR(>F)
C(Method):C(Dentist)	14.0	753898.133333	53849.866667	4.194953	0.000482
Residual	30.0	385104.666667	12836.822222	NaN	NaN

  - **Conclusion:** Now, from the result above, we observe that with the interaction of the two,for alloy 2, i.e., p(Dentist:Method) = 0.00048. Hence, we can reject the null hypothesis and conclude that the interaction between the dentists and the methods has an effect on the response variable, i.e., the dental implant hardness.

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?

- **Alloy 1:**
  - **Without interaction ANOVA Output:**

	df	sum_sq	mean_sq	F	PR(>F)
C(Method)	2.0	148472.177778	74236.088889	7.212522	0.002211
C(Dentist)	4.0	106683.688889	26670.922222	2.591255	0.051875
Residual	38.0	391121.377778	10292.667836	NaN	NaN

- **With Interaction ANOVA Output:**

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

- **Conclusion:** The results of the two-way ANOVA with and without interaction for alloy 1 is as below:
  - Without interaction, the p value for Dentist is slighlty greater than the significance level.
  - However, with interaction the p value becomes 0.01, and hence the null hypothesis can be rejected.
  - For the method category, with or without interaction the p value is less than 0.05.
  - Thus, the conclusion is that, for alloy 1, when there is an interaction between the types of dentists and methods, the dentist type has an effect on the hardness of the dental implant, and without interaction, it does not have an effect. And irrespective of the interaction, the method type has an impact on the dental implant hardness.

- **Alloy 2:**
  - **Without Interaction ANOVA Output:**

	df	sum_sq	mean_sq	F	PR(>F)
C(Method)	2.0	499640.400000	249820.200000	16.295479	0.000008
C(Dentist)	4.0	56797.911111	14199.477778	0.926215	0.458933
Residual	38.0	582564.488889	15330.644444	NaN	NaN

  - **With Interaction ANOVA Output:**

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

  - **Conclusion:** The results of the two-way ANOVA with and without interaction for alloy 2 is as below:
    - Without interaction, the p value for Dentist is 0.459, and is greater than the significance level of 0.05, and hence doesn't have an effect on hardness.
    - Even with interaction, the p value still remains 0.37, significantly higher than 0.05.
    - For the method category, with or without interaction the p value is less than 0.05.
    - Thus, the conclusion is that for alloy 2, even when there is an interaction between the types of dentists and methods, the dentist type does not have an effect on the hardness of the dental implant, and irrespective of the interaction, the method type has an impact on the same.