

Out [2]: [Click here to toggle on/off the raw code.](#)



by **Paulo Mario Medina, Gian Paolo Servañez,
Gregory Uy, and Ana Dominique Vasquez**

Applied Computational Statistics | MSDS 2024 Learning Team 4 | Asian Institute of
Management

Submitted 6 December 2023

Executive Summary

Background

This paper aimed to explore the performance of Hybrid Rice and Inbred Rice based on the study conducted by Ellanie Cabrera, and hosted on the Harvard Dataverse. In this study, the rice harvest data that was taken in 2005 was explored, with complete data for the farm situation, house hold, rice breed, and farm factors. The relationship of the breed (i.e. whether it is hybrid or inbred) to its yield was examined with exploratory data analysis and statistical methods to get a clear idea of the effect of various conditions on the productivity of both breeds of rice.

Data and Methodology

The data contains 324 rows and 126 columns of data on rice yields in the provinces of Bukidnon and Isabela. Each row is a separate plot of land which grew a different variety of rice while the columns include all relevant data on the land, irrigation, variety, input costs, and yield for that plot. We processed this data to show summary

values for yield, cost, and labor for each plot then performed EDA to show the impact of each feature on rice yield. Afterwards, Z-test, one-way ANOVA, and Mann-Whitney U Test is performed to test each hypothesis listed on the table. Below is a summary of the research questions and the corresponding hypothesis tests.

Table 1.

Research Question, Hypothesis Test, and Brief Rationale

Research Question No	Research Question	Hypothesis Test	Brief Rationale
1	Is there a significant difference between the mean yield per hectare of hybrid rice and inbred rice?	Z-test	To test whether there is a significant difference between the mean yield per hectare of the hybrid variety and the mean yield per hectare of the inbred variety.
2	Is there a significant difference between the mean yield per hectare of hybrid and inbred rice, given similar gravity irrigation system?	Z-test	To test whether there is a significant difference between the mean yield per hectare of the hybrid variety and the mean yield per hectare of the inbred variety, given a gravity irrigation system.
3	Is there at least one significant difference among the mean yield per hectare of the two breeds, given different land elevations? If there is, which breed-land elevation pair/s has/have a significant difference?	One-way ANOVA	To test whether at least one pair of land elevation-breed combination is significantly different.
4	Is there a significant difference between the mean yield per hectare per labor days (total days spent in planting, harvesting, and threshing the rice) of hybrid and inbred rice?	Mann-Whitney U Test	To test whether there is a significant difference between the mean yield per hectare per labor day of the hybrid variety and the mean yield per hectare per labor day of the inbred variety.
5.1	Lastly, is there a significant difference between the cost (of the seeds) per yield of each breed?	Mann-Whitney U Test	To test whether there is a significant difference between the mean total cost incurred from the hybrid variety and the mean total cost incurred from the inbred variety.

5.2 Is there also a significant difference between the gross profit per hectare of hybrid and inbred rice?	Mann-Whitney U Test	To test whether there is a significant difference between the mean net profit derived from the hybrid variety and the net profit derived from the inbred variety.
--	---------------------	---

Key Insight

All tests show that there is no significant difference between the rice breeds, given the metrics that were used such yield per hectare, total cost, and net profit.

Conclusion and Recommendations

The findings suggest that the decision to cultivate hybrid or inbred rice should be a strategic one, tailored to individual circumstances rather than a one-size-fits-all approach. Policymakers and agricultural advisors should focus on providing context-specific guidance and support to farmers, such as risk management training, market trend insights, and crop selection advice, helping them make informed decisions that align with their unique operational contexts and goals.

For future studies, we recommend a wider dataset encompassing different geographic locations and with different features of farming to further compare the profitability and productivity of farming.

Introduction

Rice is the number one food crop in the Philippines and is the staple food for millions of Filipinos that gets them through each day. Therefore, it is of no surprise that the cultivation of rice is one of the biggest in the country, covering about 4.7 million hectares [2], with an average of 4.1 metric tons per hectare yield [2]. The Philippines produces about 19 million tons of rice every year [2] and consumes most of it locally while also importing about 3 to 4 million [3] tons from regional trade partners, mostly Vietnam. One of the focuses of the Philippines in recent years has been food sustainability and protecting their local agricultural industry [4]. This has been a problem recently since even when the rest of the economy grew a total of 7.6% in 2022 [4], the agricultural sector trailed with only 0.5% growth in 2022 [4]. This was partly due to the increasing impact of natural disasters on Philippine Agriculture which was 22.4 billion pesos (in total cost) in 2022 and the rising cost of inputs such as fertilizer, irrigation, seed and fuel [4].

This makes things difficult for the local Filipino farmer. They cannot succeed through sheer diligence and putting in hours of labor. Farming is now a complex industry where skilled labor must manage each part of the process to use the best practices,

optimal amounts of fertilizer and pesticides, and even knowing just the right variety of rice to plant based on the conditions of the farm and the land. These are difficult challenges for many farmers who lack access to resources or are working with limited funds and with many costs. In many locations, land is underdeveloped without access to reliable irrigation. Aside from that, machinery or animal labor is still not fully utilized as investment into agriculture is not scaled up. This can be attributed to a lack of capital to invest in infrastructure as even farmers need to take loans for inputs every year. The DA alone had lent small and medium farm enterprises over 2.2 billion pesos [4] to sustain their livelihoods, or many might not even be able to afford seeds and fertilizer. Furthermore, many farmers do not even own the land they work, with only about 60% of farms self-owned [4].

To understand and to address some of the problems in this situation, our study focuses on the question of what kind of rice breed a farmer should grow. We focus on Hybrid vs. Inbred rice strains since these are the biggest choice for many farmers, and it is often at the start of the season when farmers have to make the choice about what variety of rice that they need to plant. The study will examine the factors that affect rice yields for hybrid and inbred rice, see if there is a significant difference in yields between the two, and examine the factors that may affect yields for these breeds differently. Finally, certain conditions for farmers that would impact what kind of rice strain that they should grow would also be proposed.

Research Questions

Given that introducing the hybrid rice variant to the Filipino farmers may prove as a game-changer to their livelihoods, the team proposed answering these research questions which are primarily focused on comparing the yield and profitability of both breeds:

- I. Is there a significant difference between the mean yield per hectare of hybrid rice and inbred rice?
- II. Is there a significant difference between the mean yield per hectare of hybrid and inbred rice, given similar gravity irrigation system?
- III. Is there at least one significant difference among the mean yield per hectare of the two breeds, given different land elevations? If there is, which breed-land elevation pair/s has/have a significant difference?
- IV. Is there a significant difference between the mean yield per hectare per labor days (total days spent in planting, harvesting, and threshing the rice) of hybrid and inbred rice?
- V. Lastly, is there a significant difference between the cost (of the seeds) per yield of each breed? Is there also a significant difference between the gross profit per hectare of hybrid and inbred rice?

By answering these questions, the appropriate conditions such as land, water availability, and labor days for growing hybrid rice (and inbred rice) may be ascertained, and most importantly, the important question of knowing whether hybrid rice provides more financial benefit to the Filipino farmers may be answered.

Methodology

Data Description (Scope and Limitations)

The dataset for the study is from a study done by Ellanie Cabrera in 2014 called "Socio-Economic Evaluation of Hybrid Rice Cultivation in the Philippines". Here, they studied the yield and profitability of hybrid and inbred grains.

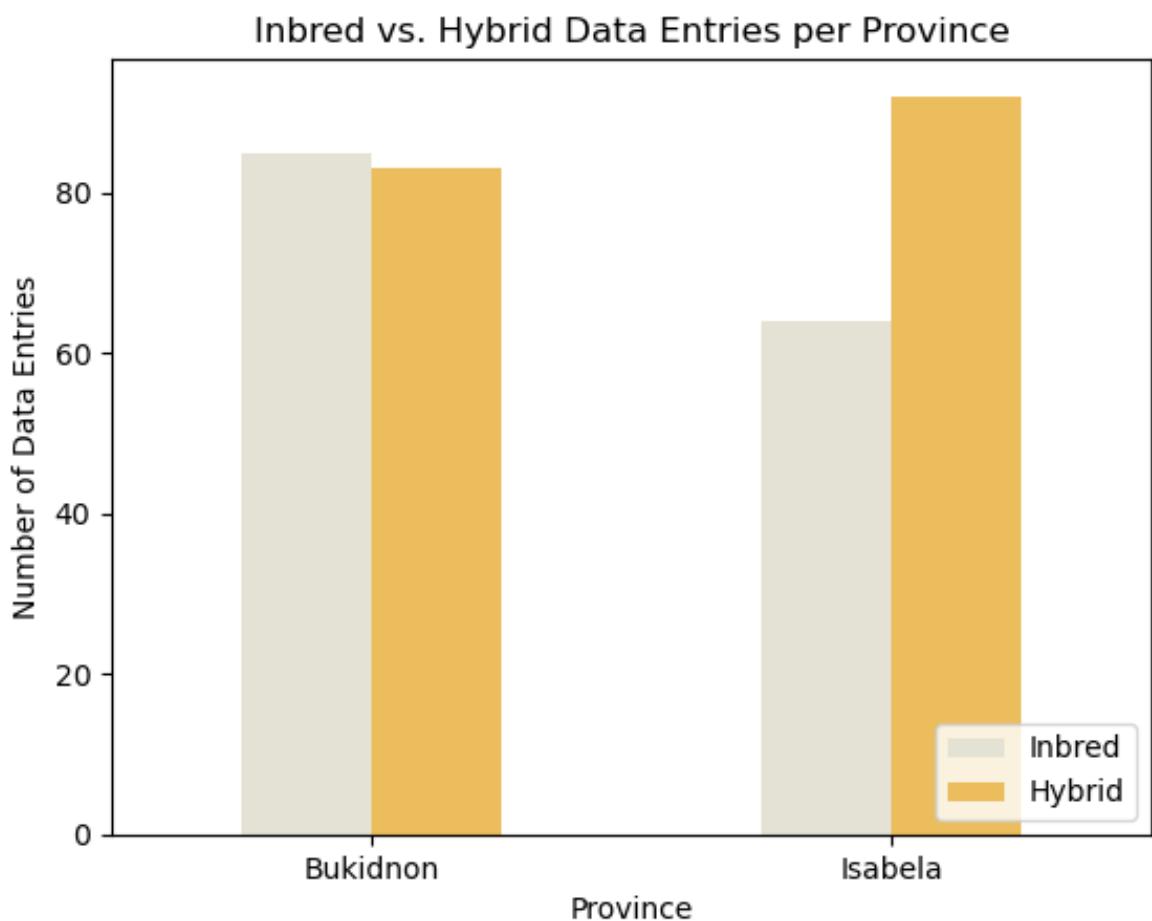


Figure 1. Inbred vs. Hybrid Data Entries per Province

The dataset originally contains 325 rows and 126 columns. It is only constrained to the provinces of Bukidnon and Isabela, with 6 villages and a total of 673 hectares of farm data. There were 44 varieties of rice grown in these farms, but these varieties were simplified to both their primary groups, hybrid and inbred rice. Lastly, each data point in the dataset represents a plot of land which was planted with either hybrid or inbred rice, which proves that each data point is independently sampled.

Due to the limitations of the dataset, it only covers the wet season of 2005 and the wet season of 2006. There were a total of 128 unique farms, but a total of 324 plots of land over the two seasons. This also only covered a limited area for two provinces, so there may be some information lacking in the dataset's features.

With this, the methodology or how the study was undertaken is divided into three parts:

1. Data Preprocessing
2. Exploratory Data Analysis
3. Hypothesis Testing



Data Preprocessing

The dataset is first loaded using the pandas module, where it is converted into a dataframe.

After checking the dataset for null or empty values, it was found that only one row needed to be dropped since there was no yield data, and there were otherwise no missing values in the dataset.

Feature engineering was conducted on the dataset to get aggregate values for the **yield per hectare, total cost, yield in KG, and total labor days** for more points to consider and explore in the exploratory data analysis and to better represent the data.

Finally, columns were renamed and then rows were split into two working datasets for inbred and hybrid rice for easier parsing and data analysis.

Features related to rice growth conditions, inputs, and total yields were more focused on so there would be a clearer picture of the rice yield situation.

Table 2.

Cleaned Rice Dataset DataFrame

Out[153...]

Fno	RECNO	Wet Season	Dry Season	Area	Hybrid	Inbred	Variety	Direct Seed	Transplan
0	1	1	1.0	1	8.0	1.0	0.0	M3	0.0
1	2	2	1.0	1	0.6	1.0	0.0	72H	0.0
2	3	2	1.0	1	0.9	1.0	0.0	M3	0.0

Exploratory Data Analysis

Overview Descriptive Statistics

Table 3.

Descriptive Statistics of the DataFrame

Out[48] :

	Fno	RECNO	Wet Season	Dry Season	Area	Hybrid
count	324.000000	324.000000	324.000000	324.0	324.000000	324.000000
mean	161.787037	66.509259	0.506173	1.0	2.076667	0.540123
std	93.786701	37.035209	0.500735	0.0	2.195546	0.499158
min	1.000000	1.000000	0.000000	1.0	0.200000	0.000000
25%	80.750000	35.000000	0.000000	1.0	1.000000	0.000000
50%	161.500000	69.500000	1.000000	1.0	1.500000	1.000000
75%	242.250000	99.000000	1.000000	1.0	2.027500	1.000000
max	324.000000	128.000000	1.000000	1.0	16.000000	1.000000

Mean Yield per Hectare of Hybrid Rice vs. Inbred Rice

Count of Data Points for Mean Yield per Hectare of Hybrid vs. Inbred Rice

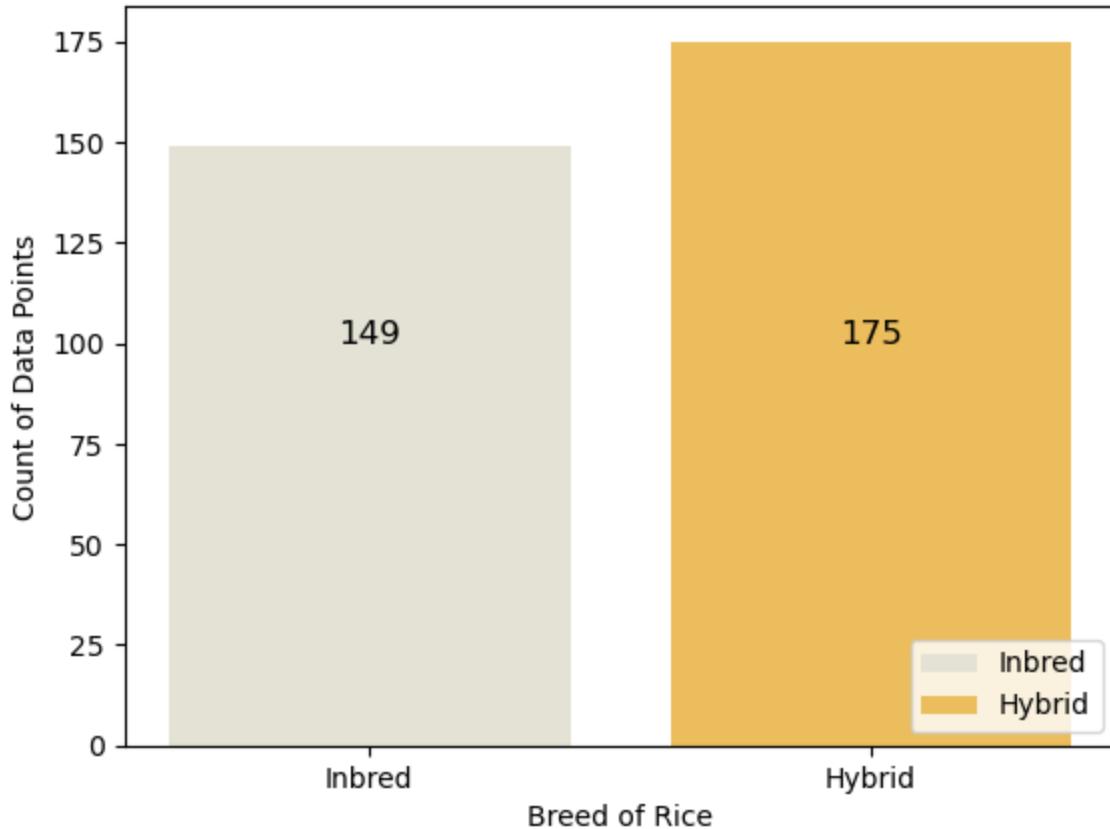


Figure 2. Count of Data Points for Mean Yield per Hectare Hybrid vs. Inbred Rice

This figure can be used as a support for how despite not knowing the population standard deviation, Z-test can still be used for comparing the mean yields of both groups, given that the sample sizes of both groups are greater than 30.

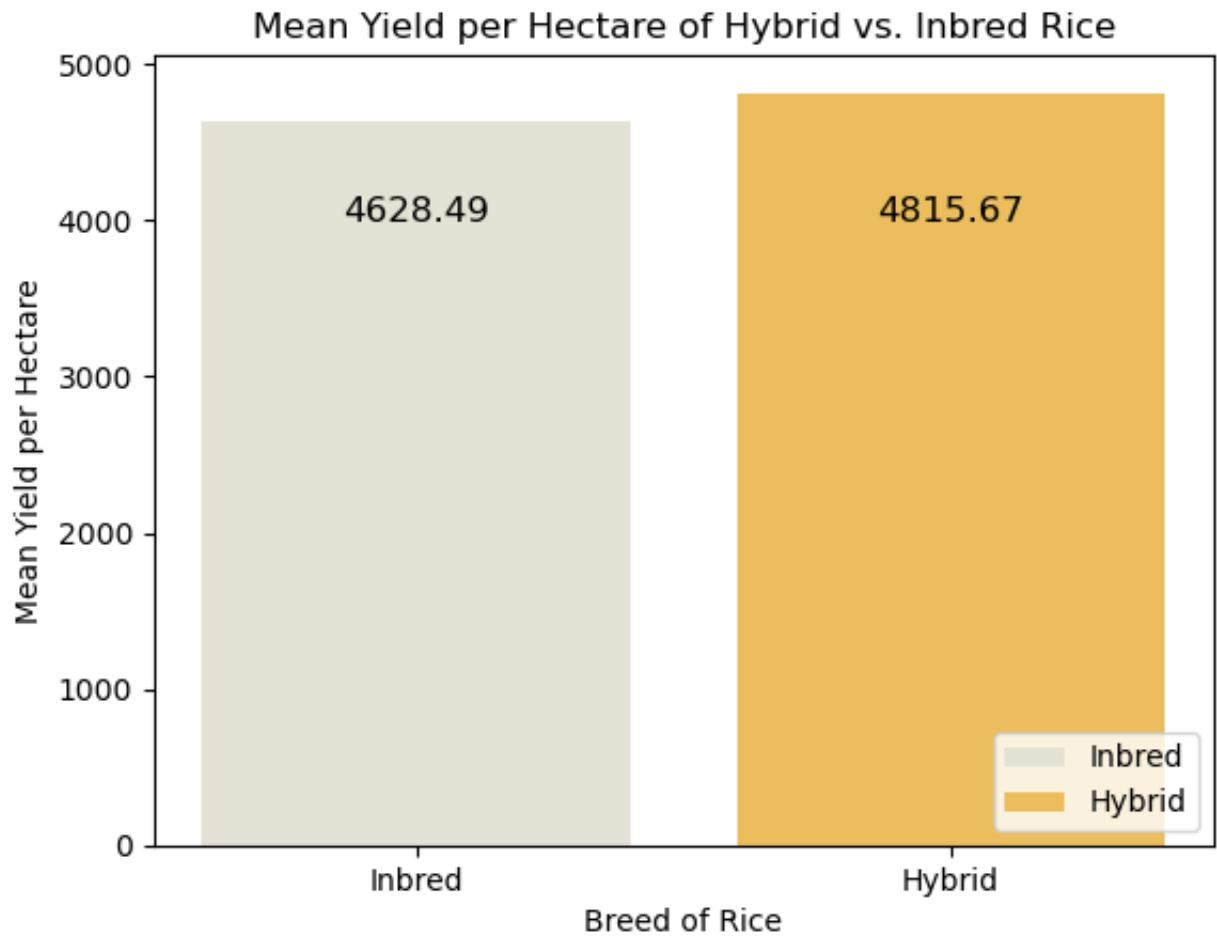


Figure 3. Mean Yield per Hectare of Hybrid vs. Inbred Rice

Given that the comparison of the mean yield per hectare between the two breeds would be delved into in the hypothesis testing phase, it can be seen from this figure that there is an apparent difference in mean yield per hectare between the two breeds. This further necessitates the need for proving whether there is indeed a significant difference between their mean yields per hectare.

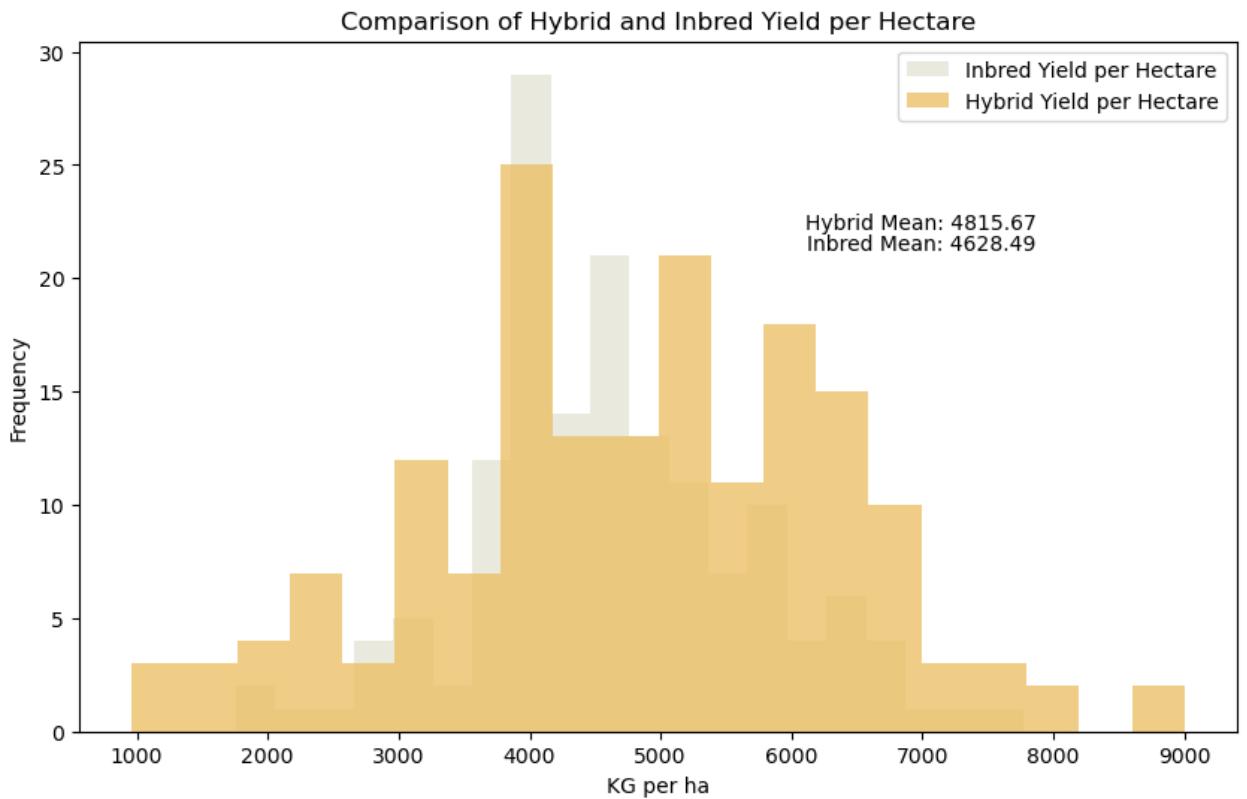


Figure 4. Yield per Hectare of Hybrid vs. Inbred Rice Histogram

As seen here in this figure, both groups approximately follow a normal distribution, which further supports that Z-test, which requires that the data points are normally distributed, can be used in comparing the means of both groups. Moreover, given the wider distribution of the hybrid yield, it can be said that it varies more than the inbred yield, which means that it is not as consistent as the inbred yield in producing mean yields per hectare. Lastly, there are no extreme outliers in both groups, which is further supported by the box-plots below.

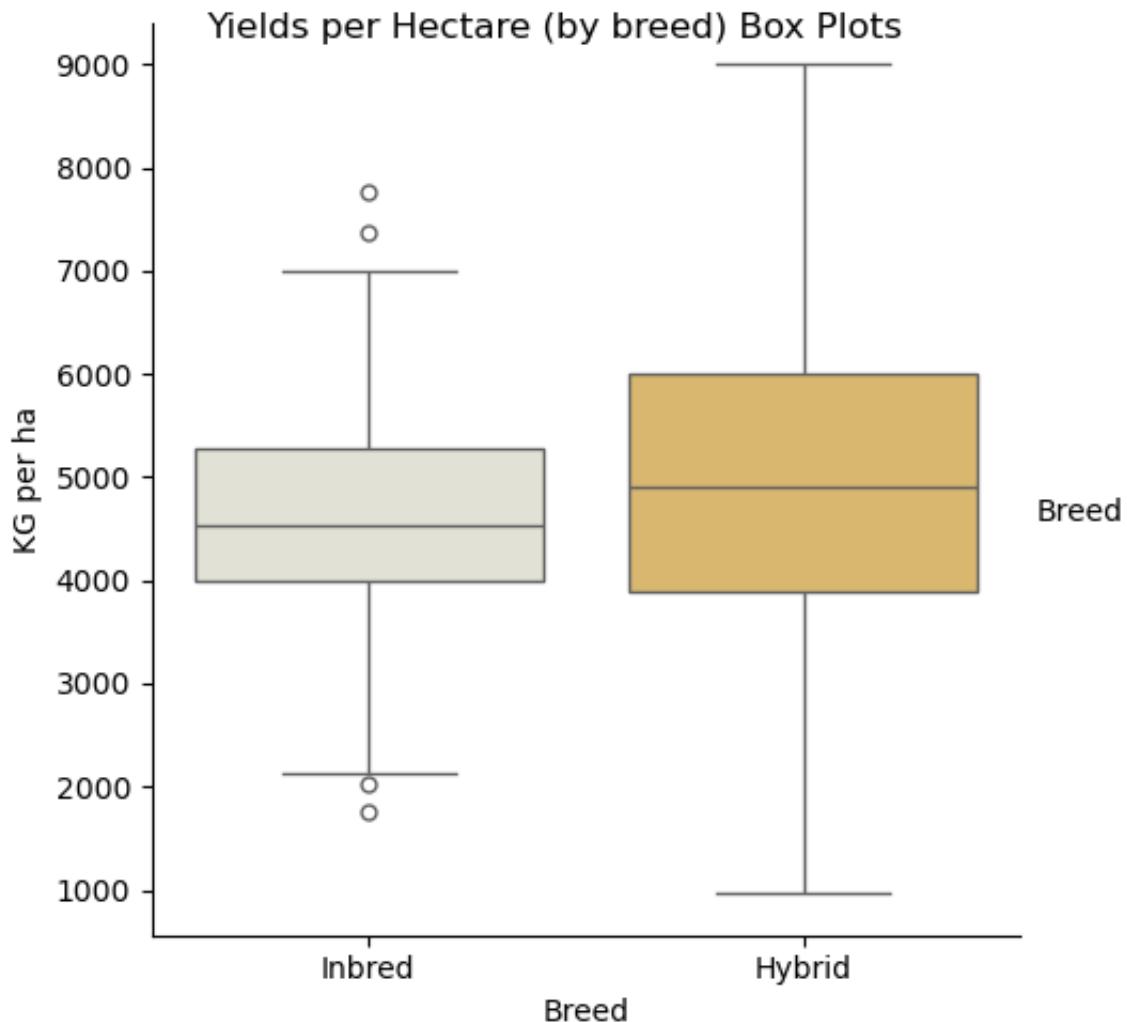


Figure 5. Yield per Hectare of Hybrid vs. Inbred Rice Box Plots

This figure further supports how the yield per hectare of the hybrid rice has a wider distribution than the inbred rice. Moreover, although there are outliers in the inbred rice group, they are not extreme and would skew the mean of the inbred rice very much. And so, they would not be dropped from the dataset.

Mean Yield per Hectare of Hybrid Rice vs. Inbred Rice Given A Gravity-based Irrigation System

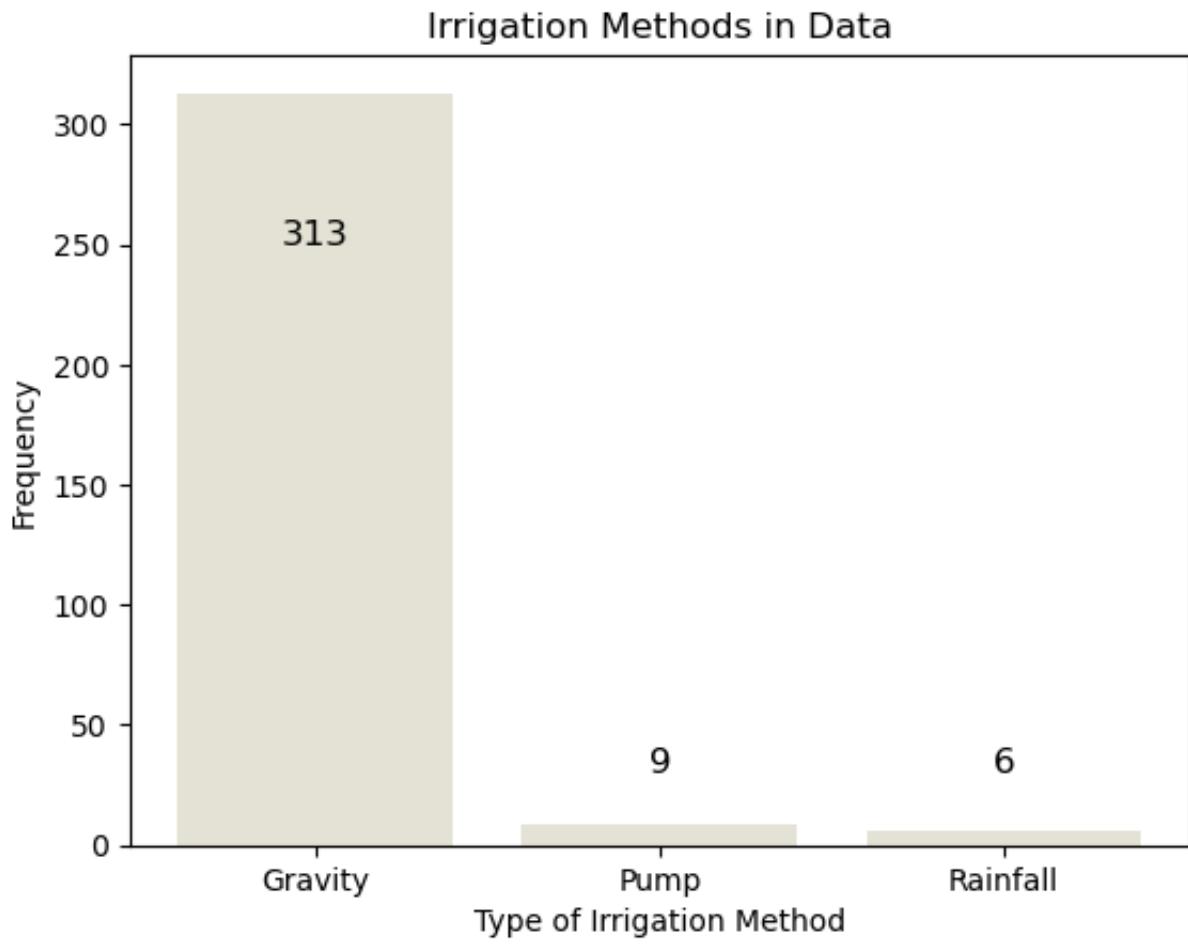


Figure 6. Frequency of Data Points for Each Irrigation Method

As seen in the figure above, other irrigation methods such as pump and rainfall do not have high frequencies, and so, it would be difficult to compare their mean yield per hectare, given that they would not be able to approximate a normal distribution and would not reach the given rule-of-thumb of a sample size of at least 15 data points for each group (given 2-9 groups) [5], which is a prerequisite for Analysis of Variance (ANOVA), which would allow for comparison of means between groups. Given this, it would be better to focus instead on whether there is a significant difference between the yield per hectare of each breed given a gravity-based irrigation method.

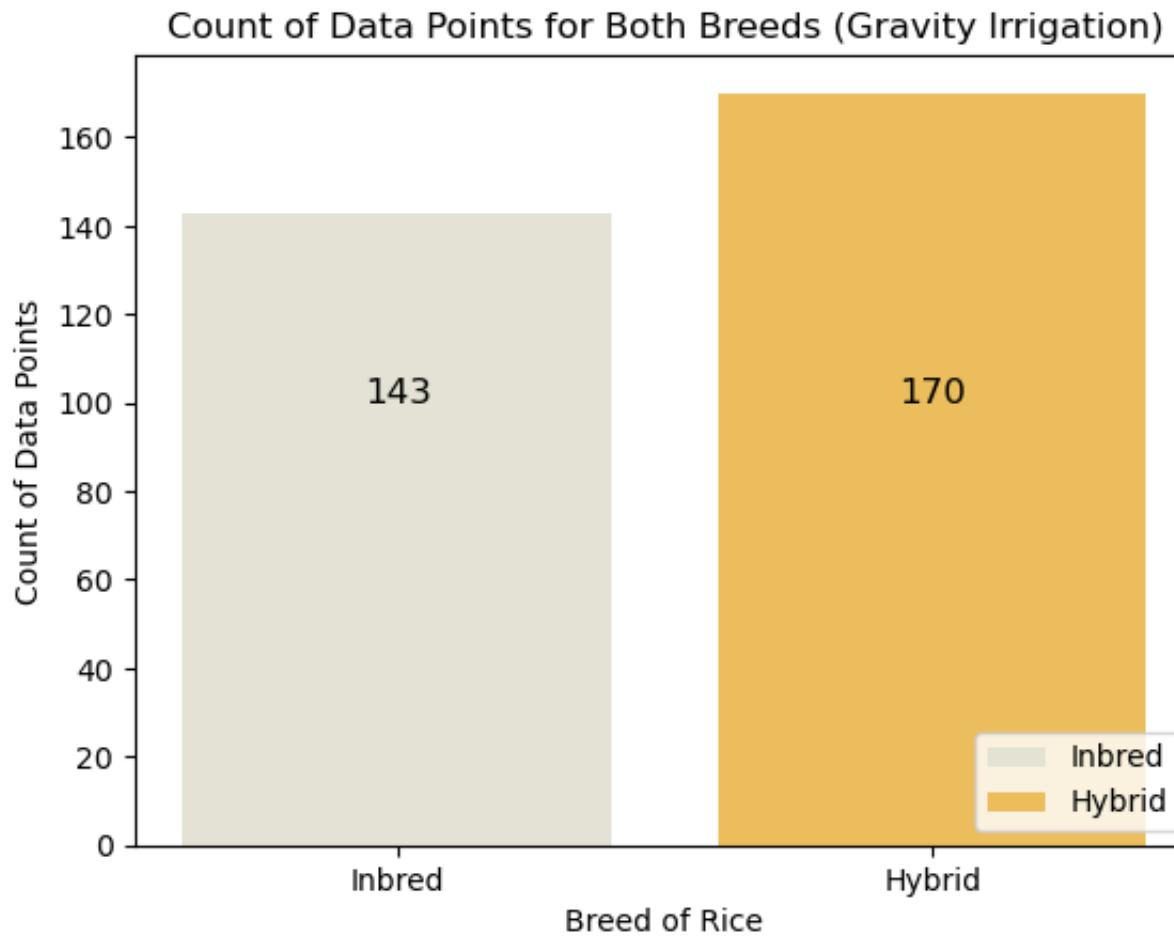


Figure 7. Frequency of Data Points (Gravity Irrigation)

Focusing on both breeds given a gravity-based irrigation system, both are greater than 30, and so, similarly, Z-test can also be used to compare both groups' mean yields per hectare.

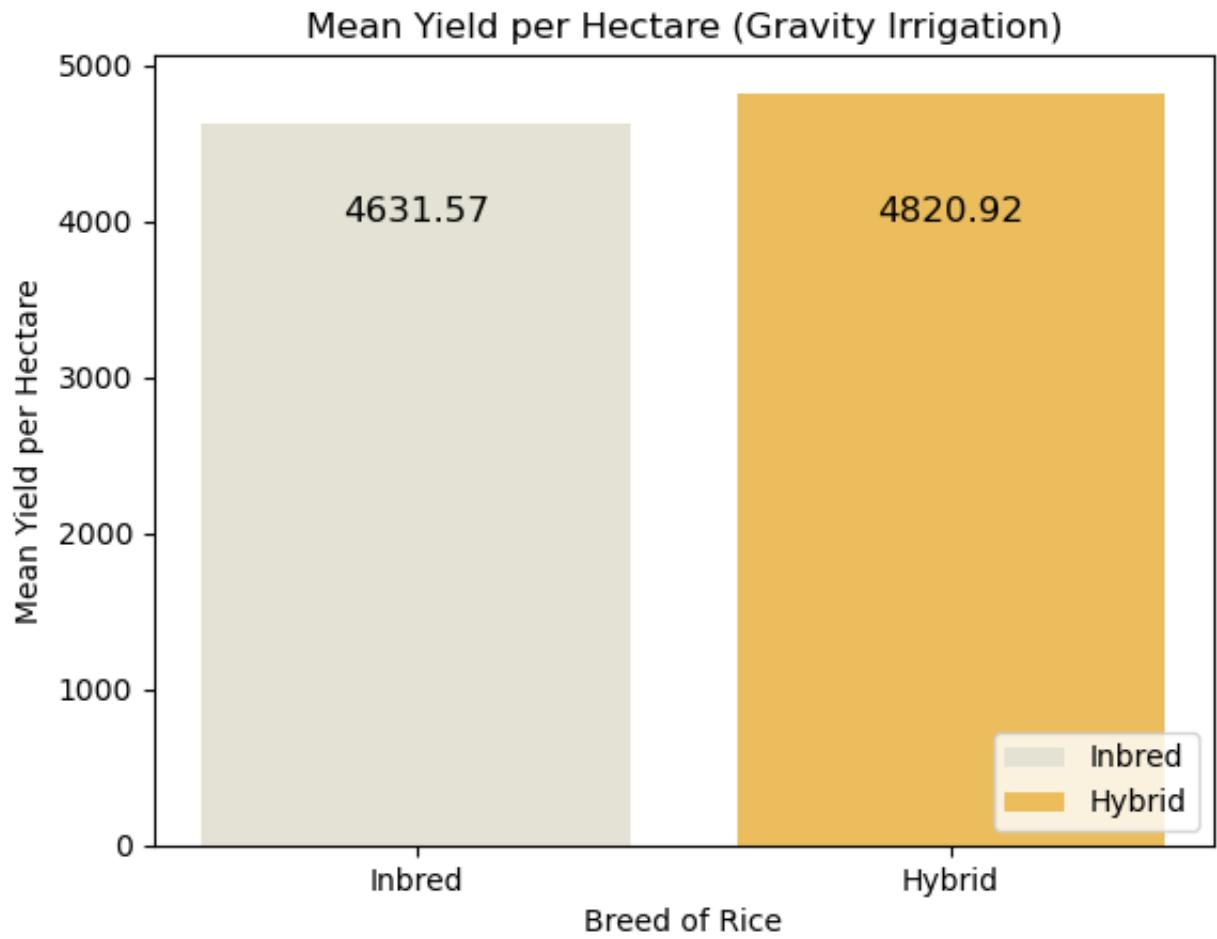


Figure 8. Mean Yield per Hectare of Hybrid vs. Inbred Rice (Gravity Irrigation)

Similarly, from this figure, it can be said that there is a possible significant difference between the mean yields of both groups, although this would be confirmed through a hypothesis test.

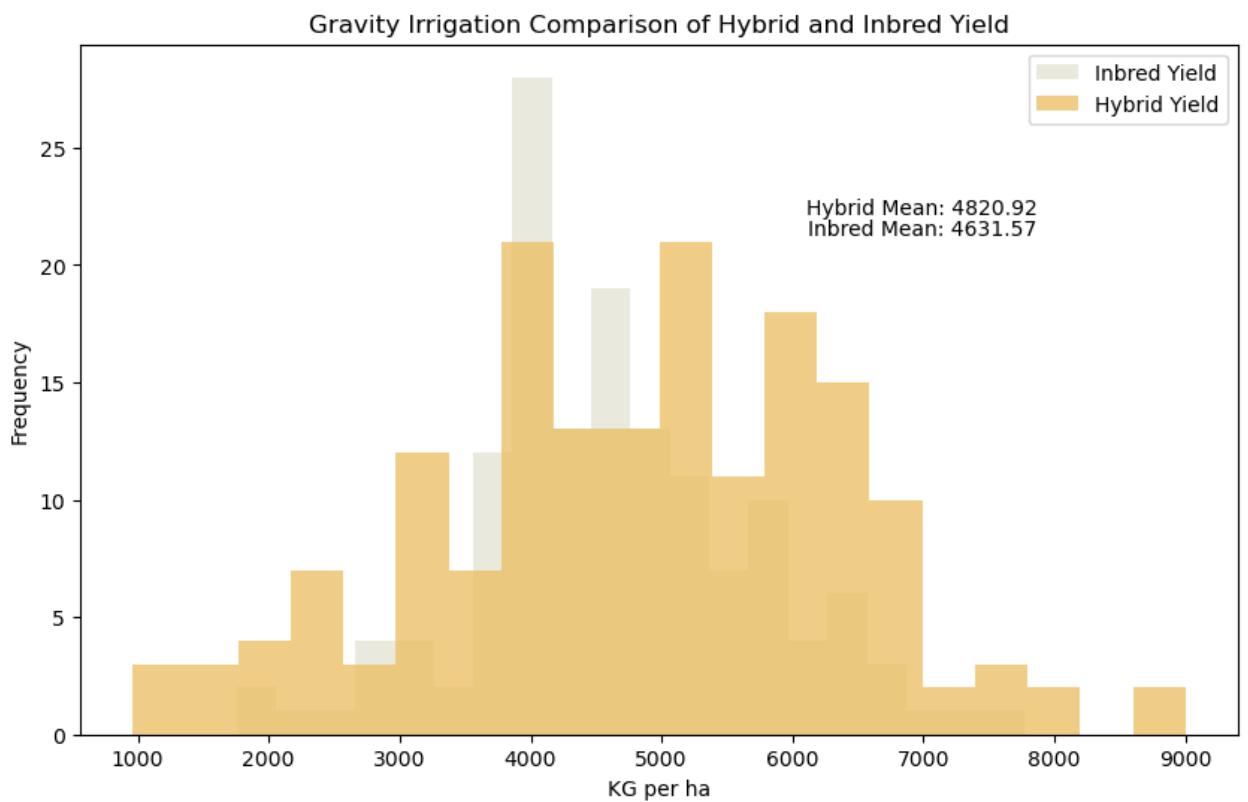


Figure 9. Gravity Irrigation Comparison of Hybrid and Inbred Yield

Similarly, from this figure, it can be said that the hybrid yield has a wider distribution again than the inbred yield and that they both follow approximately a normal distribution. This further supports how the Z-test can be used for the hypothesis associated with both groups.

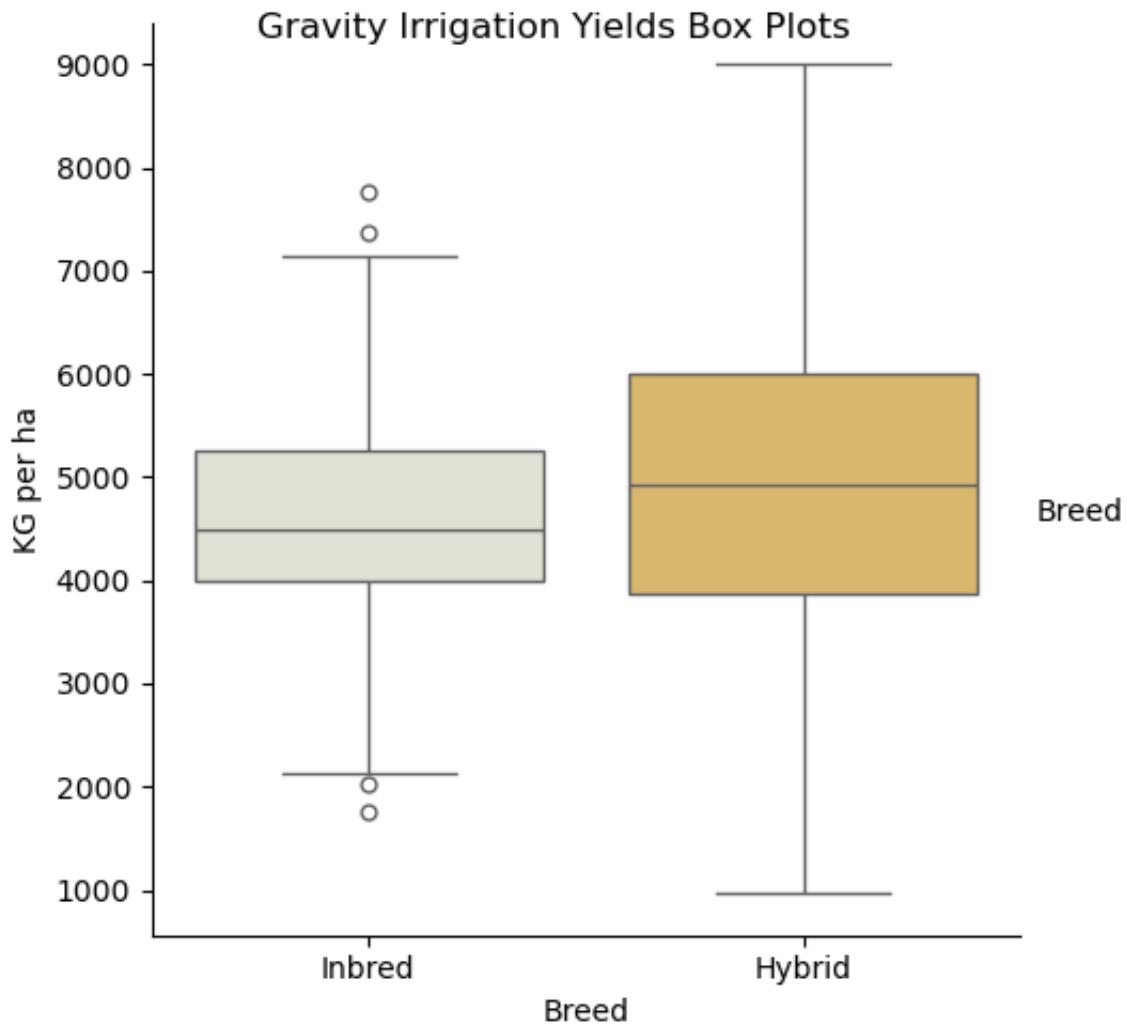


Figure 10. Yield per Hectare of Hybrid vs. Inbred Rice (Gravity Irrigation) Box Plots

Similarly, this figure is very similar to the box plot, and so, the same interpretation that there are no **extreme** outliers in both groups in this section still holds.

Mean Yield per Hectare Across Land Elevations

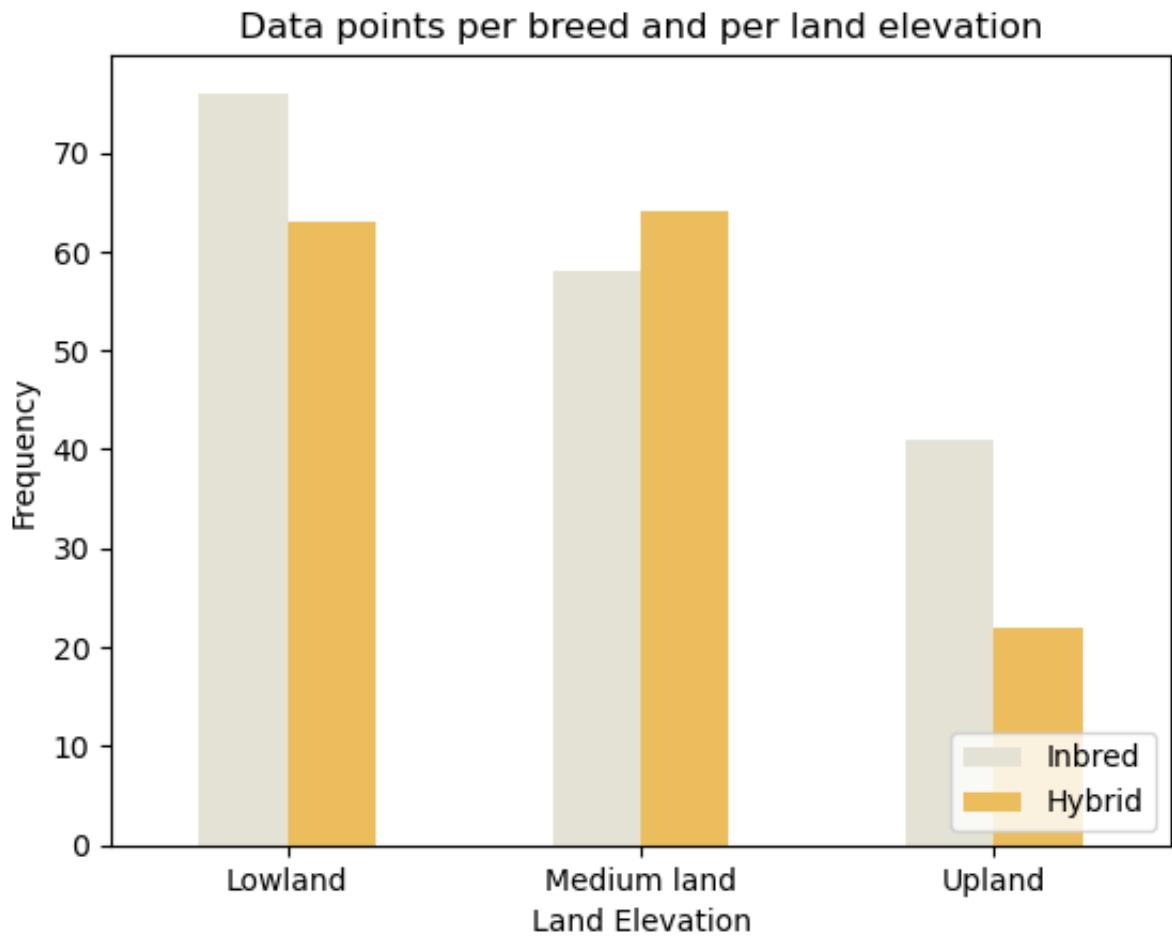


Figure 11. Frequency of Data Points Per Breed and Per Land Elevation

As seen here in this figure, most of the plot of lands are in the lowland and in the medium land, as compared to the upland. Moreover, the plots of land where hybrid is planted and inbred is planter differs across land elevations. Since one-way analysis of variance (ANOVA) would be the best hypothesis test for this, in order to maintain robust statistical power, each pair would be truncated to the number of data points in the pair with lowest sample size, i.e., upland-hybrid pair where 22, the number of the hybrid-upland pairs are enough for the sample size especially for ANOVA involving six groups (where the rule-of-thumb is at least 15). [5]

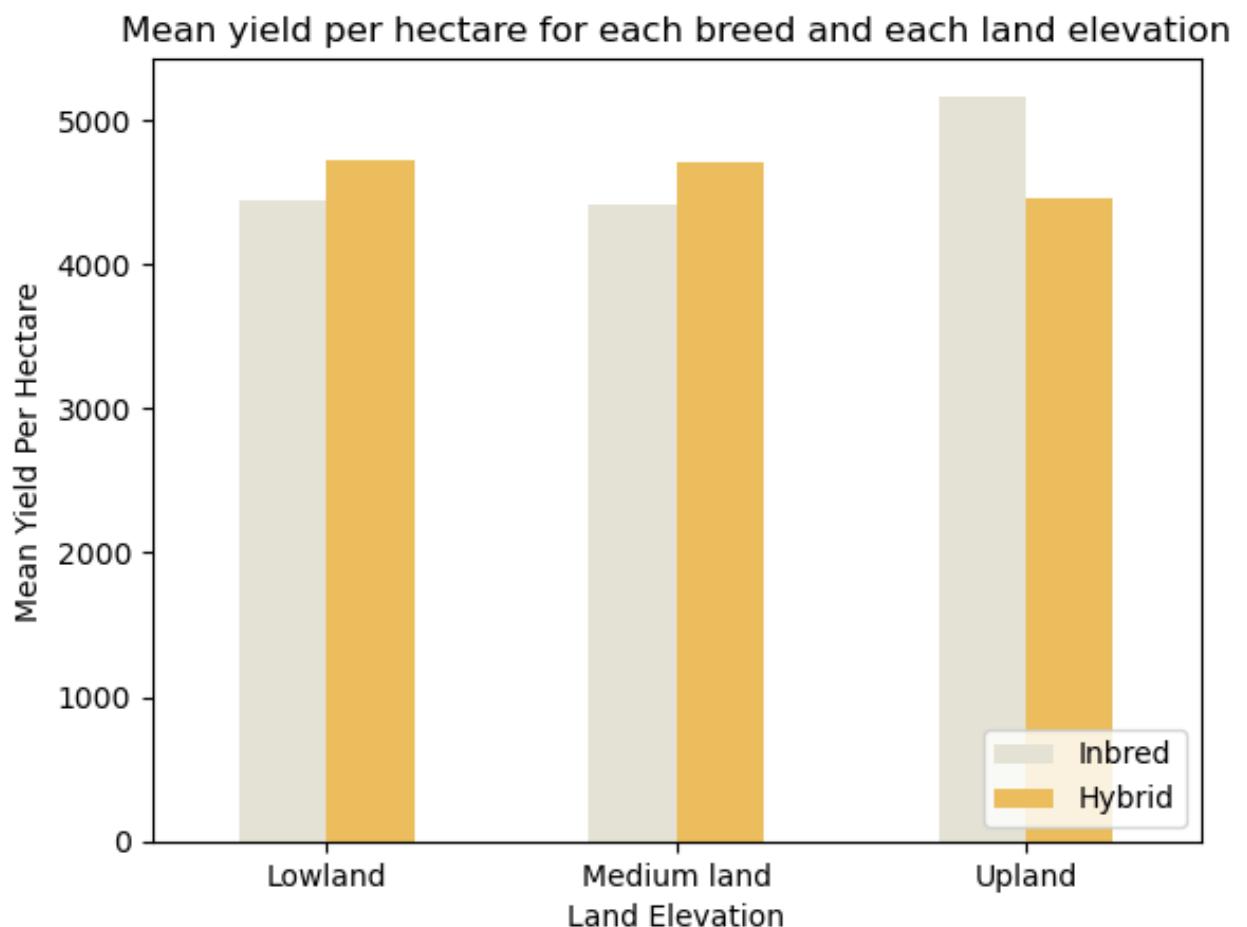


Figure 12. Mean Yield per Hectare Per Breed and Per Land Elevation

As seen here, there is not much difference for the mean yield per hectare for all given pairs. However, there seems to be significant differences in the mean yield per hectare for the upland pairs where inbred rice has higher yield per hectare for both.

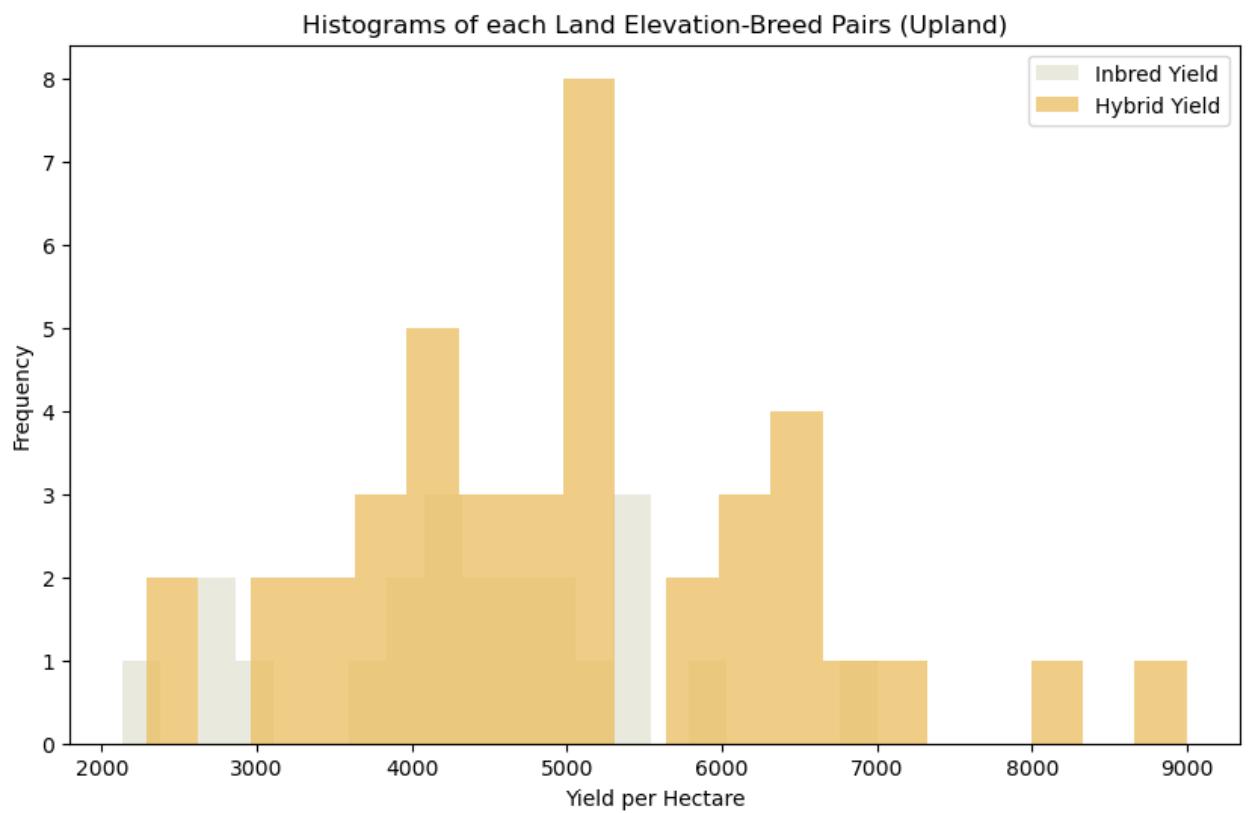


Figure 13. Mean Yield per Hectare Per Breed (Upland) Histograms

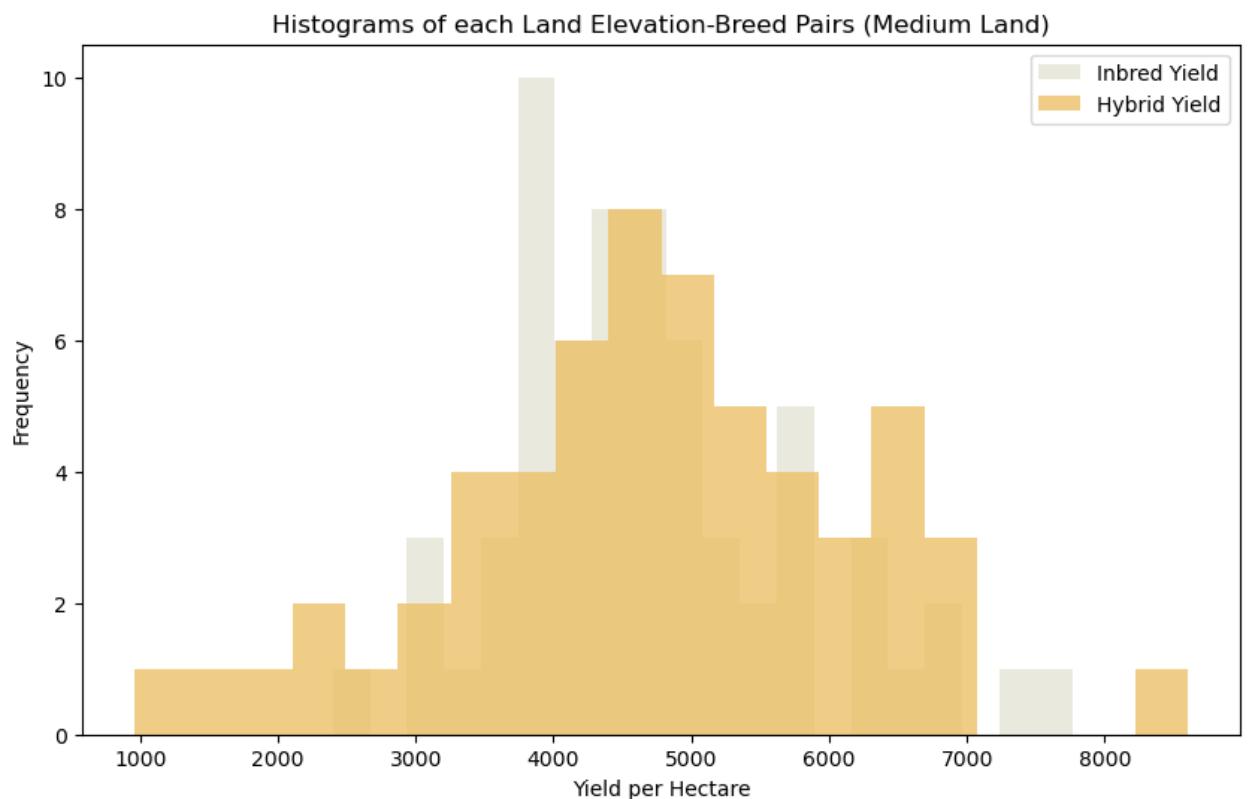


Figure 14. Mean Yield per Hectare Per Breed (Medium Land) Histograms

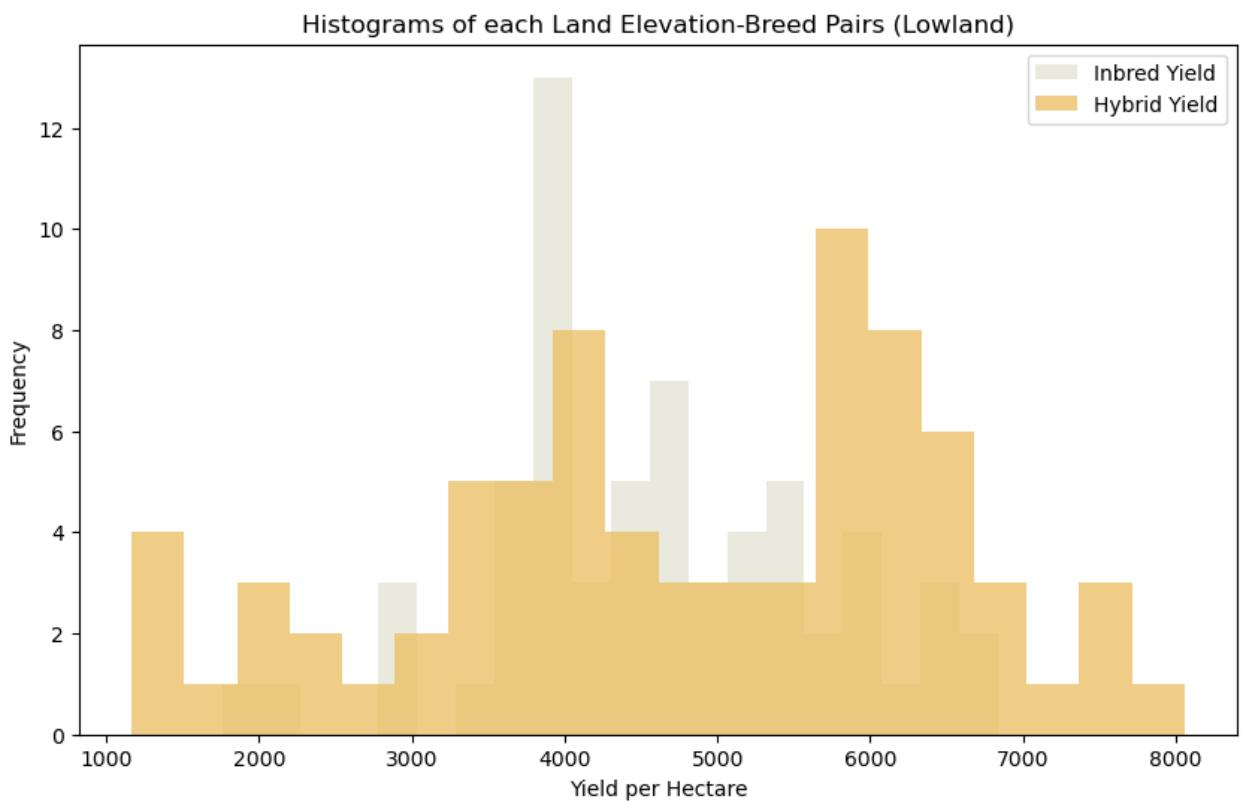


Figure 15. Mean Yield per Hectare Per Breed (Lowland) Histograms

It can be seen from this set of figures that while both lowland and medium land pairs nearly approximately follow a normal distribution, the upland pairs except for the hybrid-upland pair show little signs of normality. Although the last one does not show signs of normality, ANOVA is still said to be robust against the violation of its assumption of normality, allowing the use of ANOVA for the given hypothesis. [6]

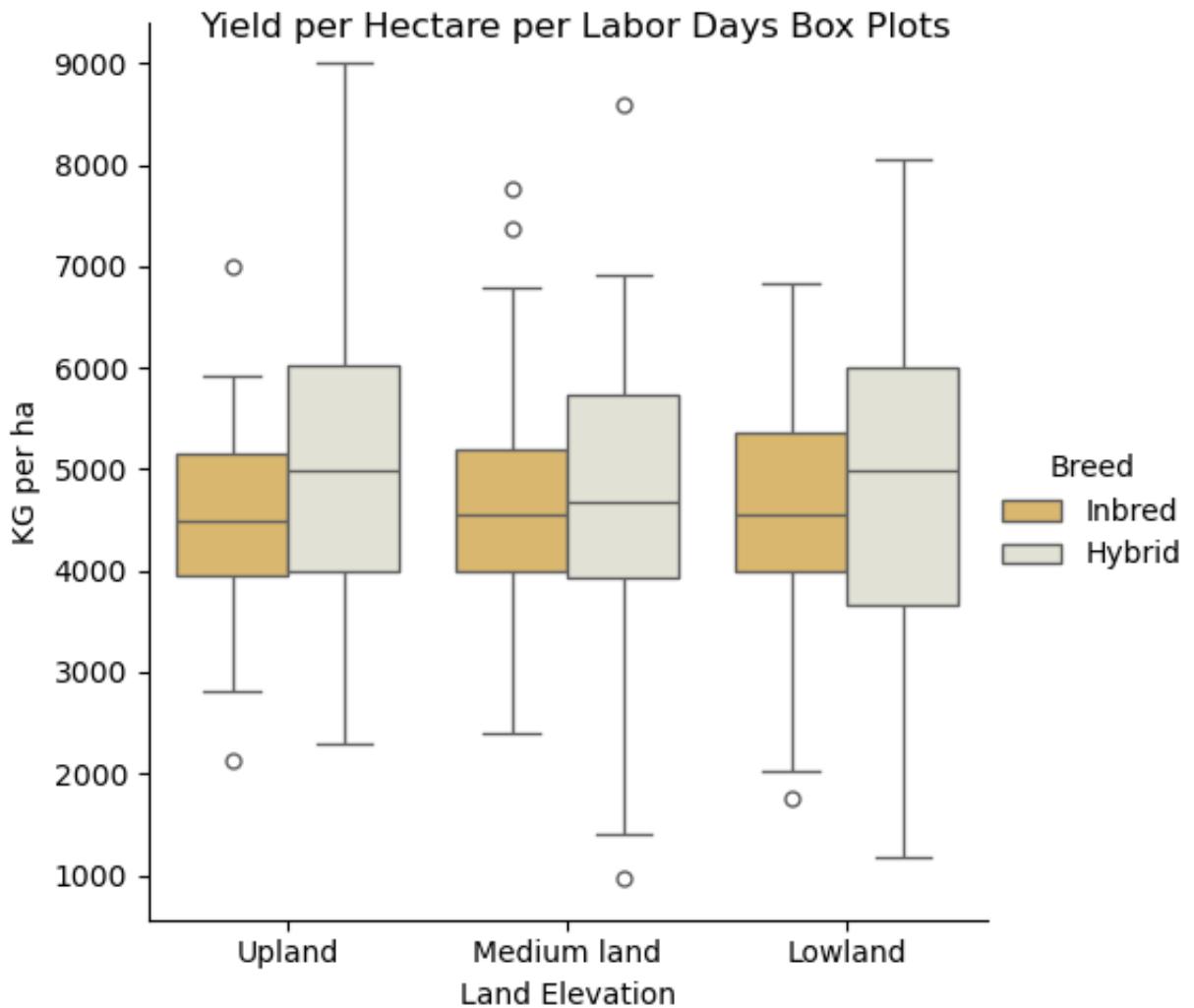


Figure 16. Mean Yield per Hectare Per Breed and Per Land Elevation Boxplots

As shown by the whiskers in the box plots, although medium land and low land have nearly the same variance, the pairs in the upland elevation has a significantly different variance from the others. Hence, while this violates the assumption of ANOVA, given that the samples for each pair would be truncated to 22, the loss in statistical power from unequal variances would be mitigated and lessened as well [7].

Mean Yield per Hectare per Labor days of Hybrid Rice vs. Inbred Rice

Given that the groups are not based on other conditions than their breed, there would be no need to check the frequencies of each group, as this has been discussed previously in the first figures.

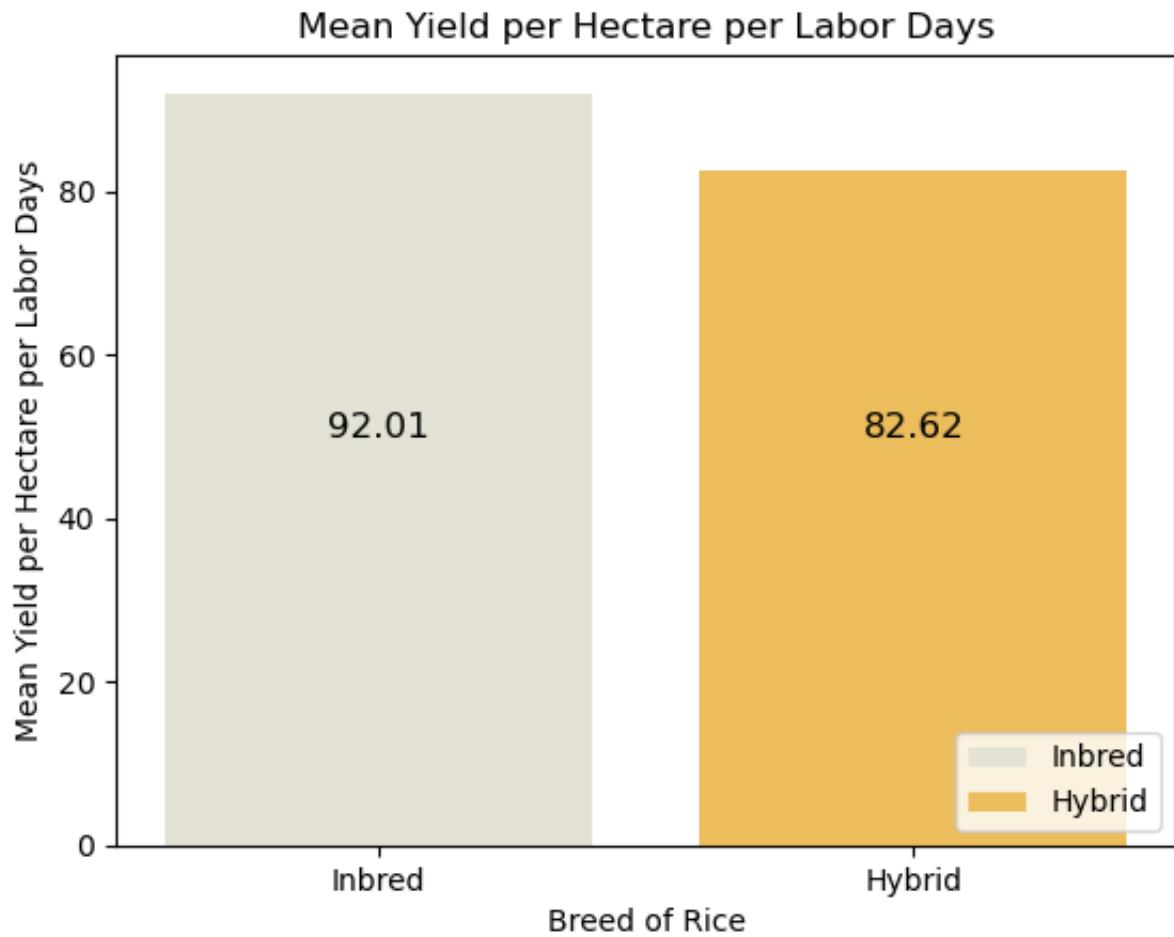


Figure 17. Mean Yield per Hectare per Labor Days of Hybrid vs. Inbred Rice

It can be seen from this figure that if viewed as it is, the mean yield per hectare per labor days (or the amount of yield per hectare for each day spent for labor) is higher for inbred rice as compared to hybrid rice. This necessitates further exploration on whether there is indeed a significant difference for the amount of yield per hectare for each labor day for each breed.

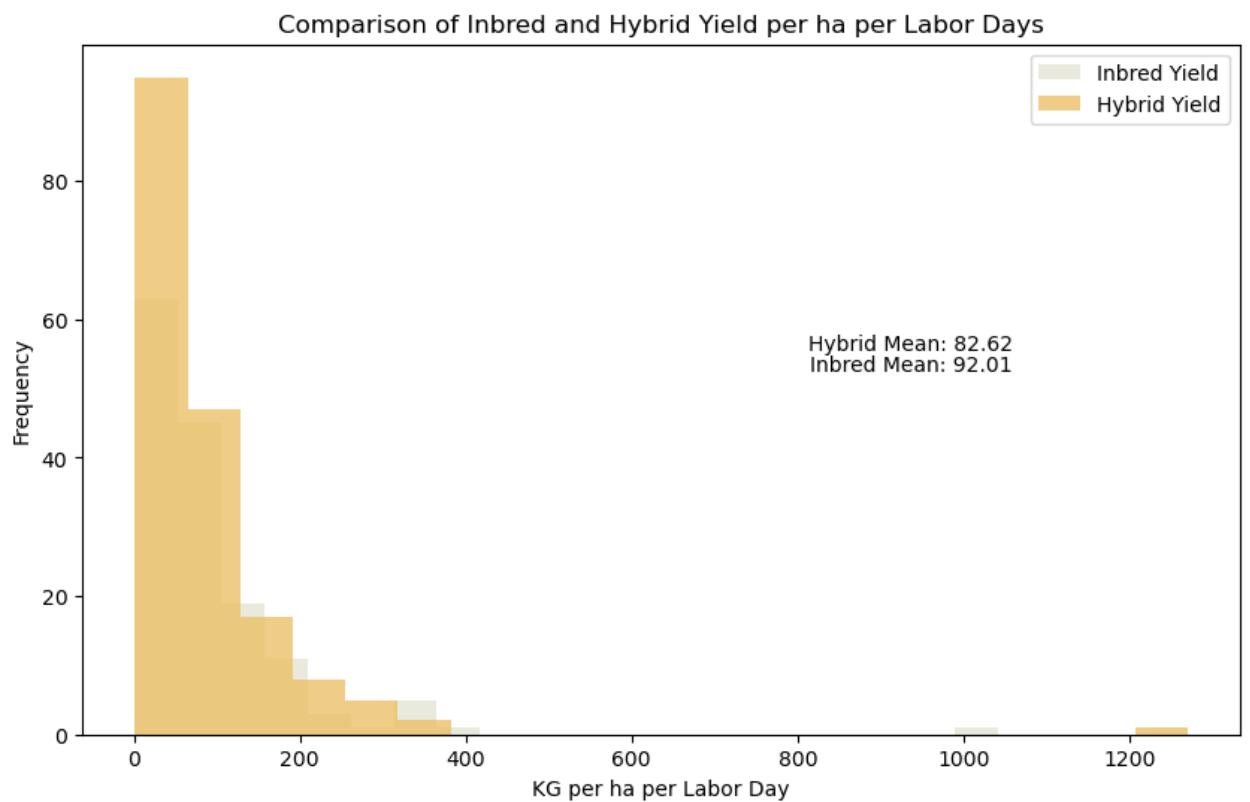


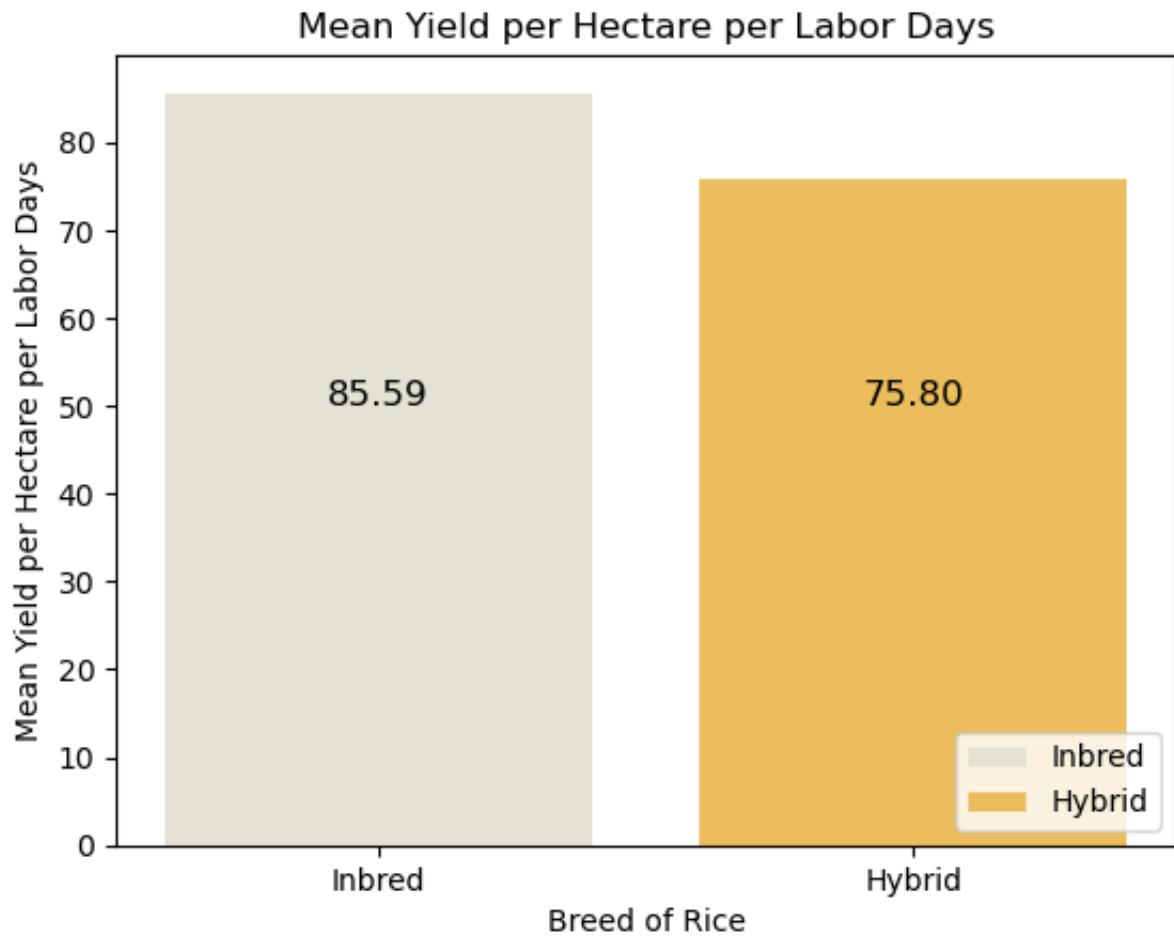
Figure 18. Mean Yield per Hectare per Labor Days of Hybrid vs. Inbred Rice Histogram

From this figure, it can be said that the distributions for both groups are skewed to the right, essentially signifying that the kilogram yield per hectare for one labor day spent is on average, around the left, lower side of the histogram (smaller values). Given this similar type of distribution for both groups, Mann-Whitney U test which does not necessitate normality but necessitates similarity in distributions would be most useful in testing the associated hypothesis.



Figure 19. Mean Yield per Hectare per Labor Days of Hybrid vs. Inbred Rice Box Plot

Given that there are two extreme outliers in total for both groups, these would be removed when doing the hypothesis testing, as they would skew the means of both groups. Moreover, it can be seen here that there is indeed a right skew and that there are non-extreme outliers, where the former further supports the use of the Mann-Whitney U test. To show their effects on the mean, a bar plot showing the means of both groups without the outliers would be shown.



Even if the outliers were removed, it can be said that the mean yield per hectare for a one labor day is higher for inbred rice than hybrid rice. This may indicate (not conclusively) a significant difference between their mean yield per hectare per labor days.

Cost of Hybrid Rice vs. Inbred Rice

Lastly, given that the groups are not based on other conditions than their breed, there would be no need to check the frequencies of each group, as this has been discussed previously in the first figures.

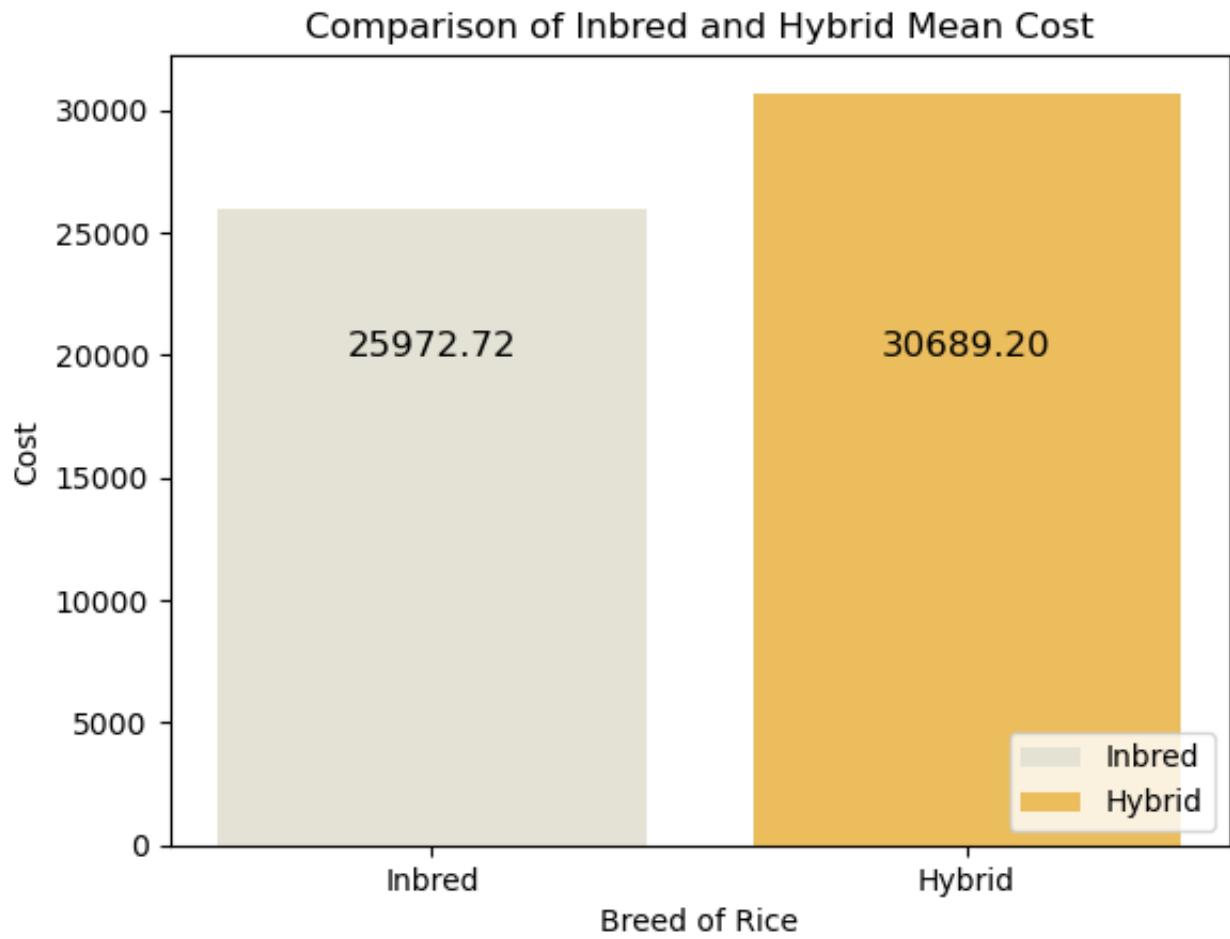


Figure 20. Comparison of Inbred and Hybrid Mean Cost

As seen above, the mean total costs for hybrid rice is significantly higher than the mean total costs for inbred rice. This would necessitate testing the hypothesis whether there is indeed a significant difference between the two groups.

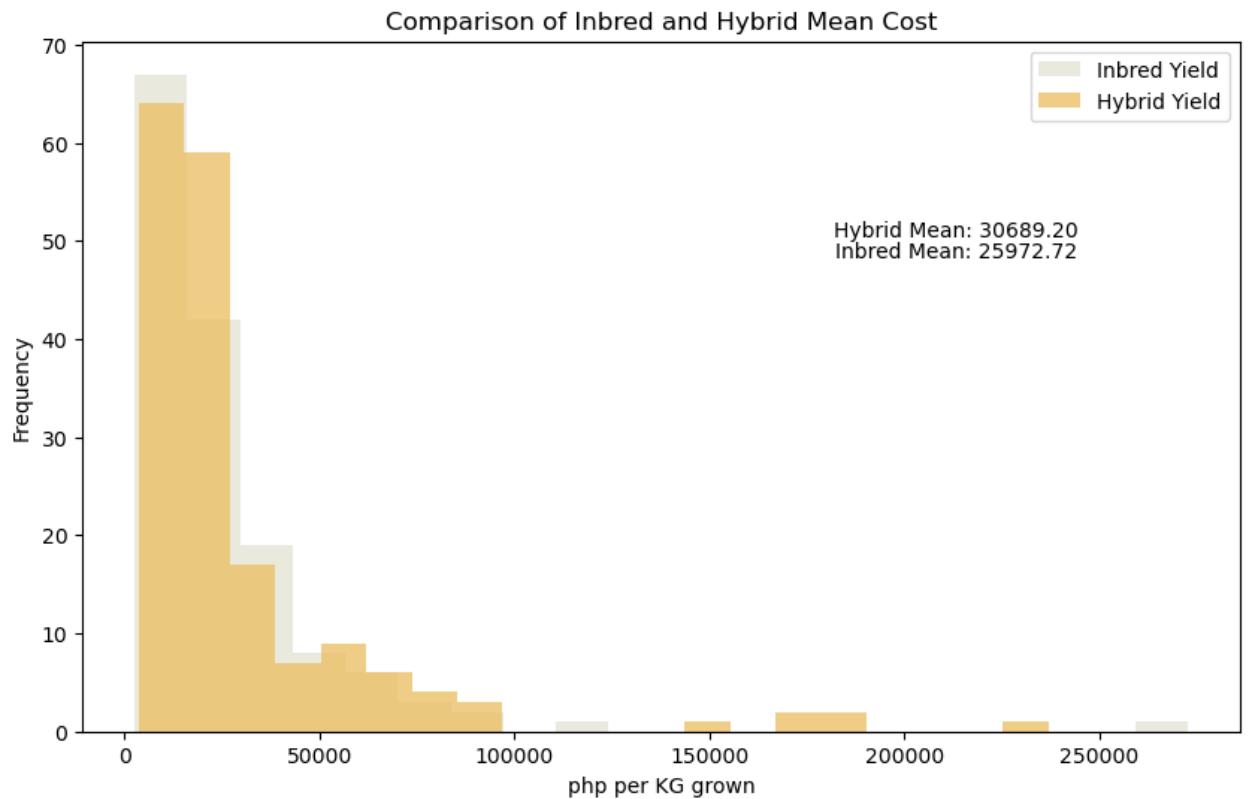


Figure 21. Comparison of Inbred and Hybrid Mean Cost Histogram

For the given figure, it is shown that the mean total costs have a right-skewed distribution, signifying the need for Mann-Whitney U test in proving the associated hypothesis. Moreover, it can be seen here that there are extreme outliers for both groups, signifying the need for their removal.

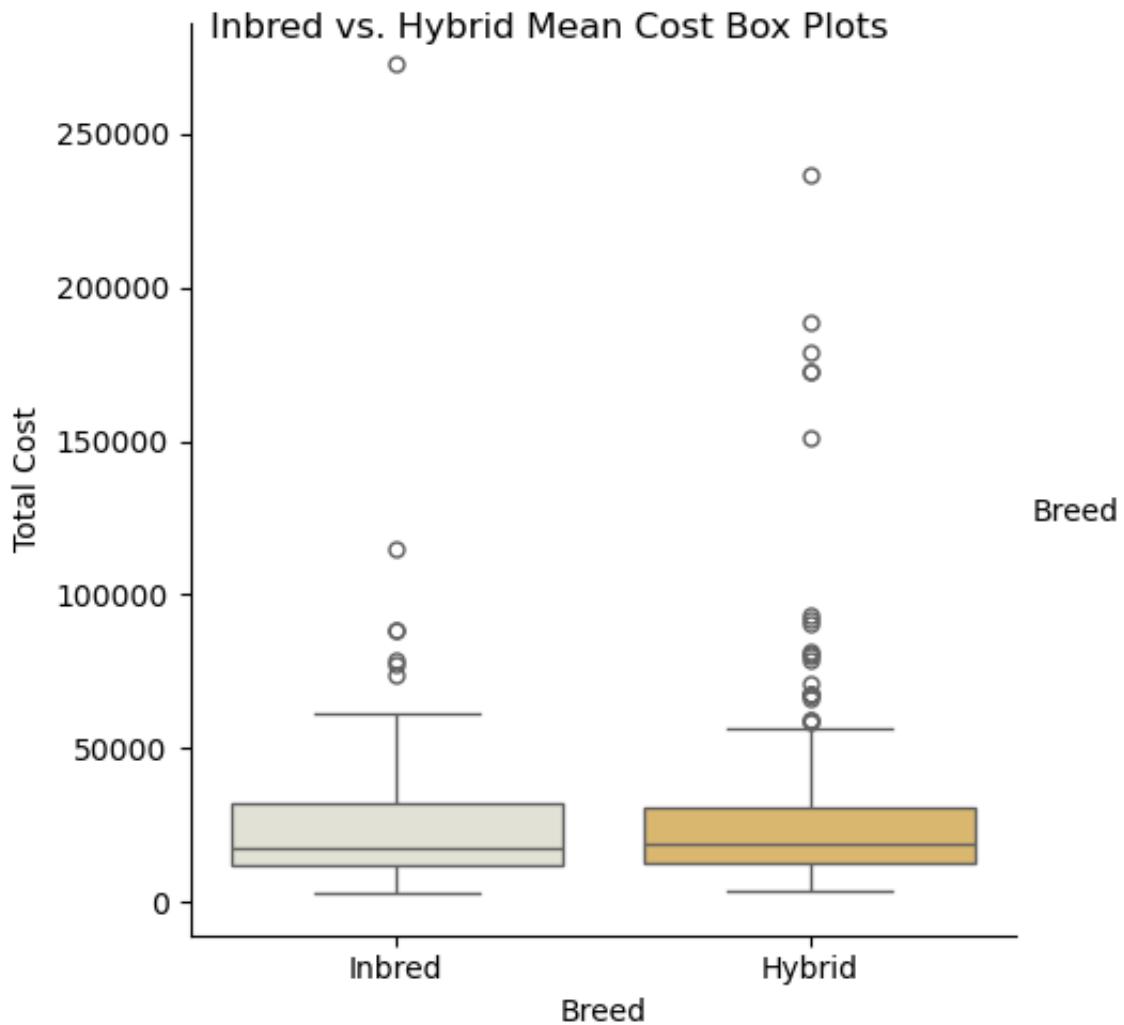


Figure 22. Comparison of Inbred and Hybrid Mean Cost Box Plots

The figure above shows the number of extreme outliers per group, which would be removed during hypothesis testing. Moreover, another bar plot showing the means without the extreme outliers would be displayed and discussed.

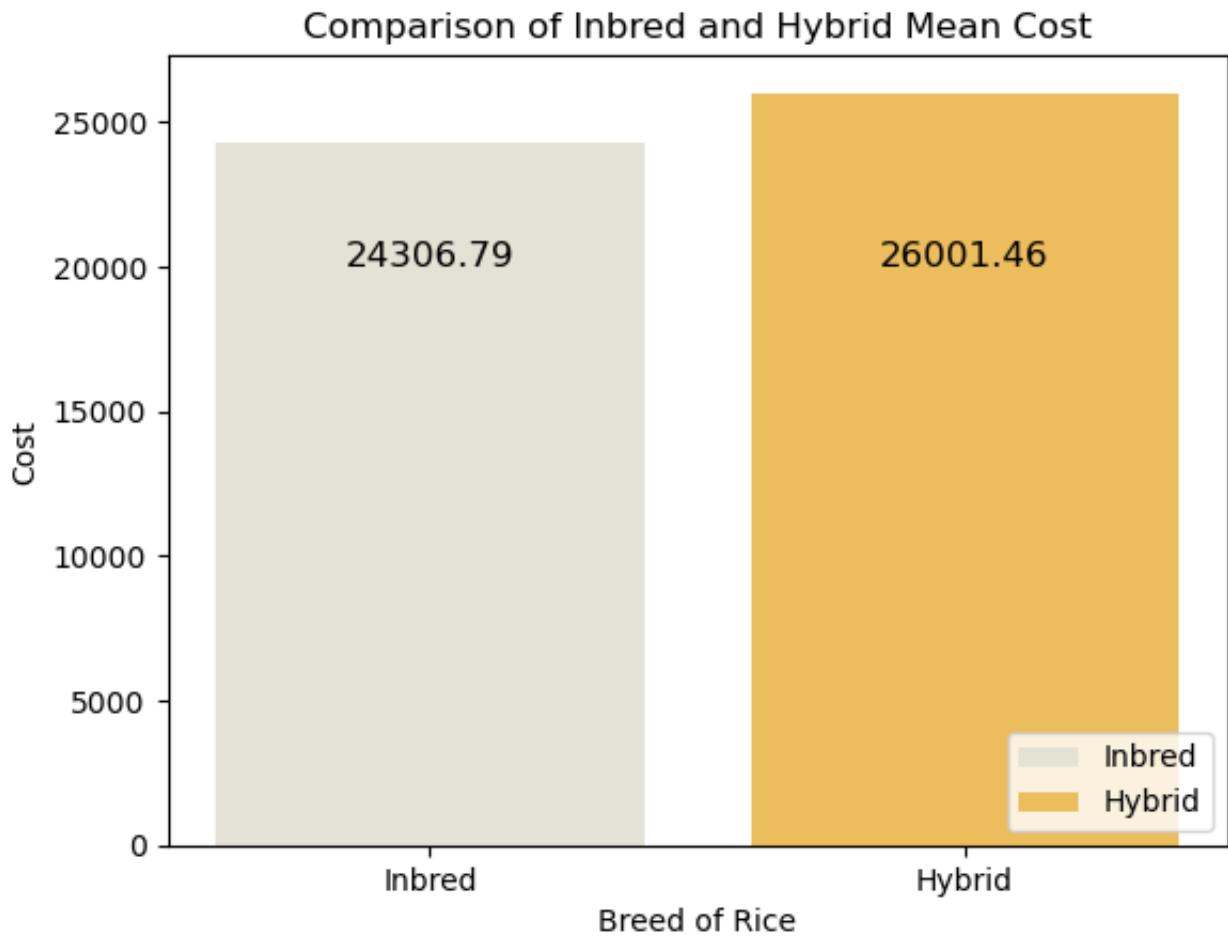


Figure 23. Comparison of Inbred and Hybrid Mean Cost After Removing The Outliers

The figure above still shows that there may be indeed a significant difference between the groups where mean total cost of hybrid rice still outweighs the mean total cost of inbred rice. This further supports the need for hypothesis testing of the associated hypothesis.

Gross Profit of Hybrid Rice vs. Inbred Rice

Again, given that the groups are not based on other conditions than their breed, there would be no need to check the frequencies of each group, as this has been discussed previously in the first figures.

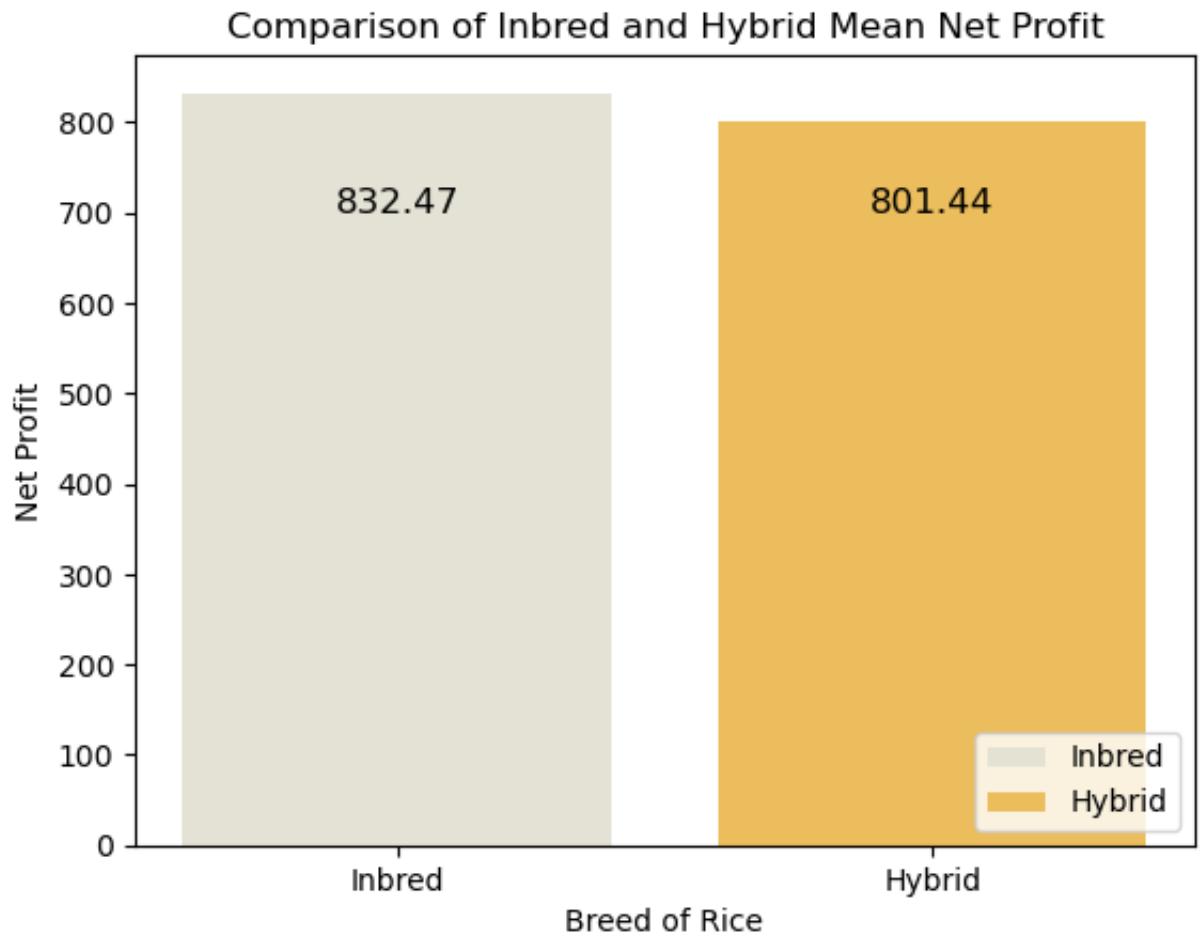


Figure 24. Comparison of Inbred and Hybrid Mean Net Profit

As seen here, the net profit which was acquired by multiplying both the kilograms sold and the price per kilogram and subtracting this product by the cost per yield multiplied by the kilogram sold is different for Inbred Rice and Hybrid Rice by a small margin. Still, in order to confirm whether there is a small significance between the two as profitability is one of the key factors to choosing which rice breed to plant, a hypothesis test would be done, and which hypothesis test to use would be explored in the following figures.

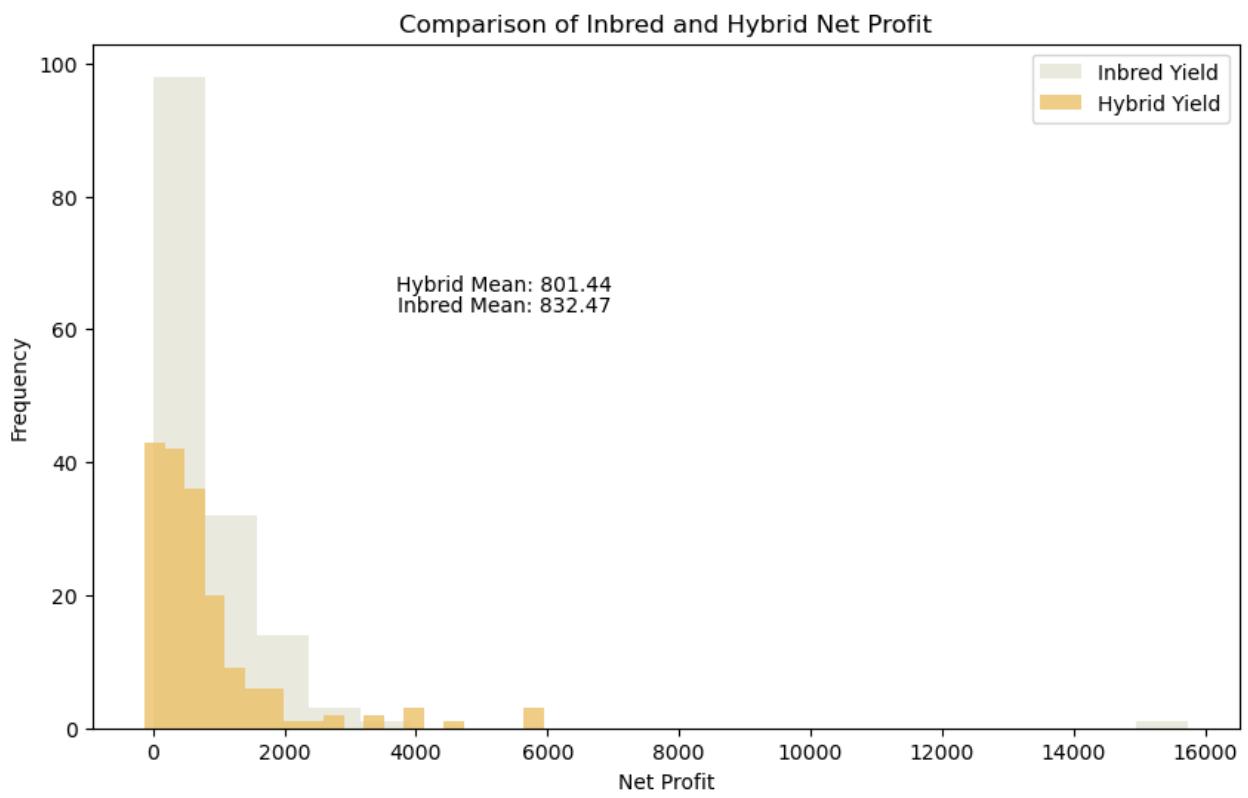


Figure 25. Comparison of Inbred and Hybrid Net Profit Histogram

As seen in this histogram, the net profit in both groups follows a right-skewed distribution, and so, similar to labor efficiency, the hypothesis test to be used would be Mann-Whitney U test which does not necessitate normality and only necessitates similar shapes for their distributions.

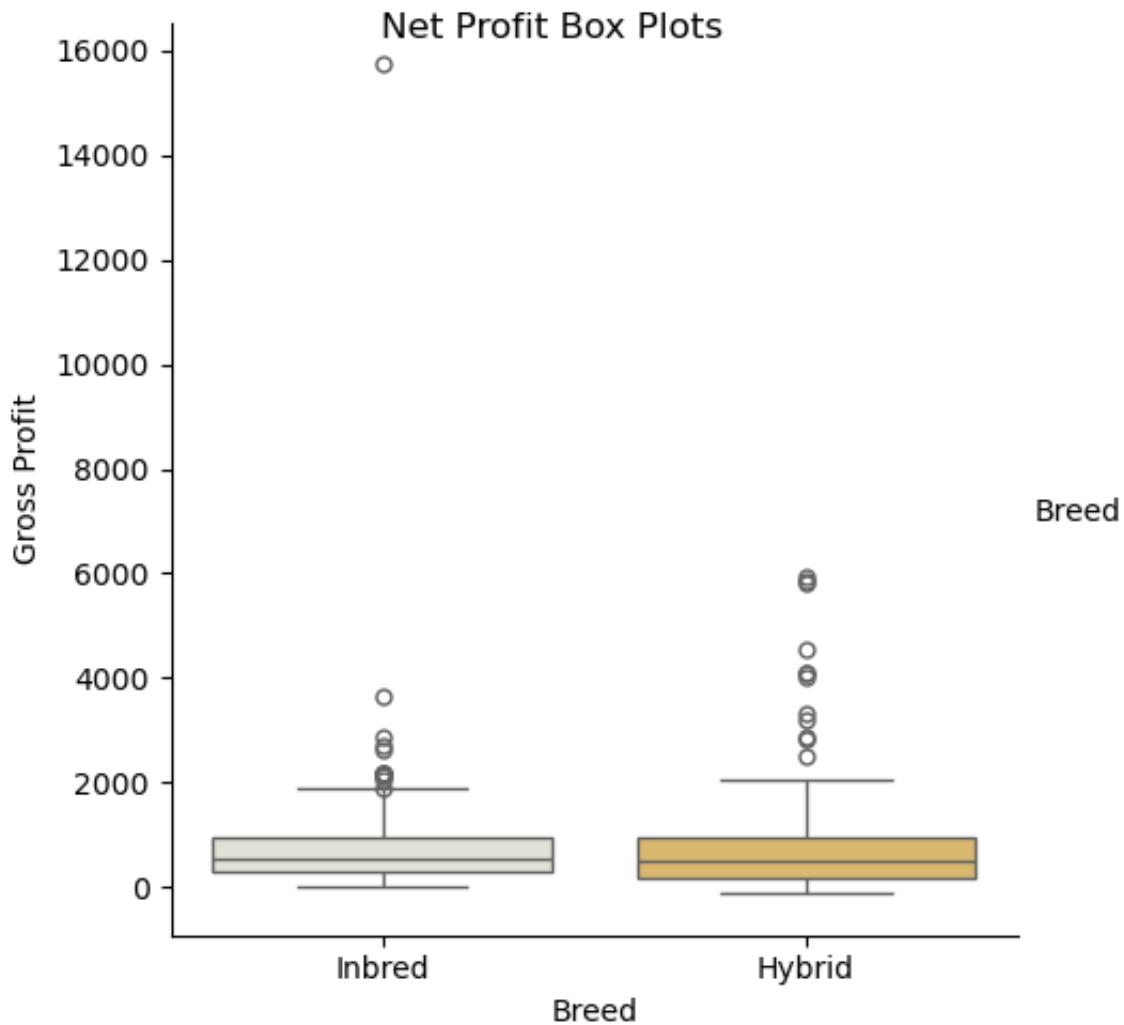


Figure 26. Comparison of Inbred and Hybrid Net Profit Box Plots/center>

Again, for the net profit, there is an extreme outlier for the inbred rice. Similarly, this outlier would be removed during hypothesis testing, and as seen in the box plot's whiskers, it signifies a right skewed distribution for both. Hence, this further supports using Mann-Whitney U test for this associated hypothesis. In order to see further its effect on the net profit for inbred rice, a bar plot showing the means of both groups, with the outlier removed would be shown.

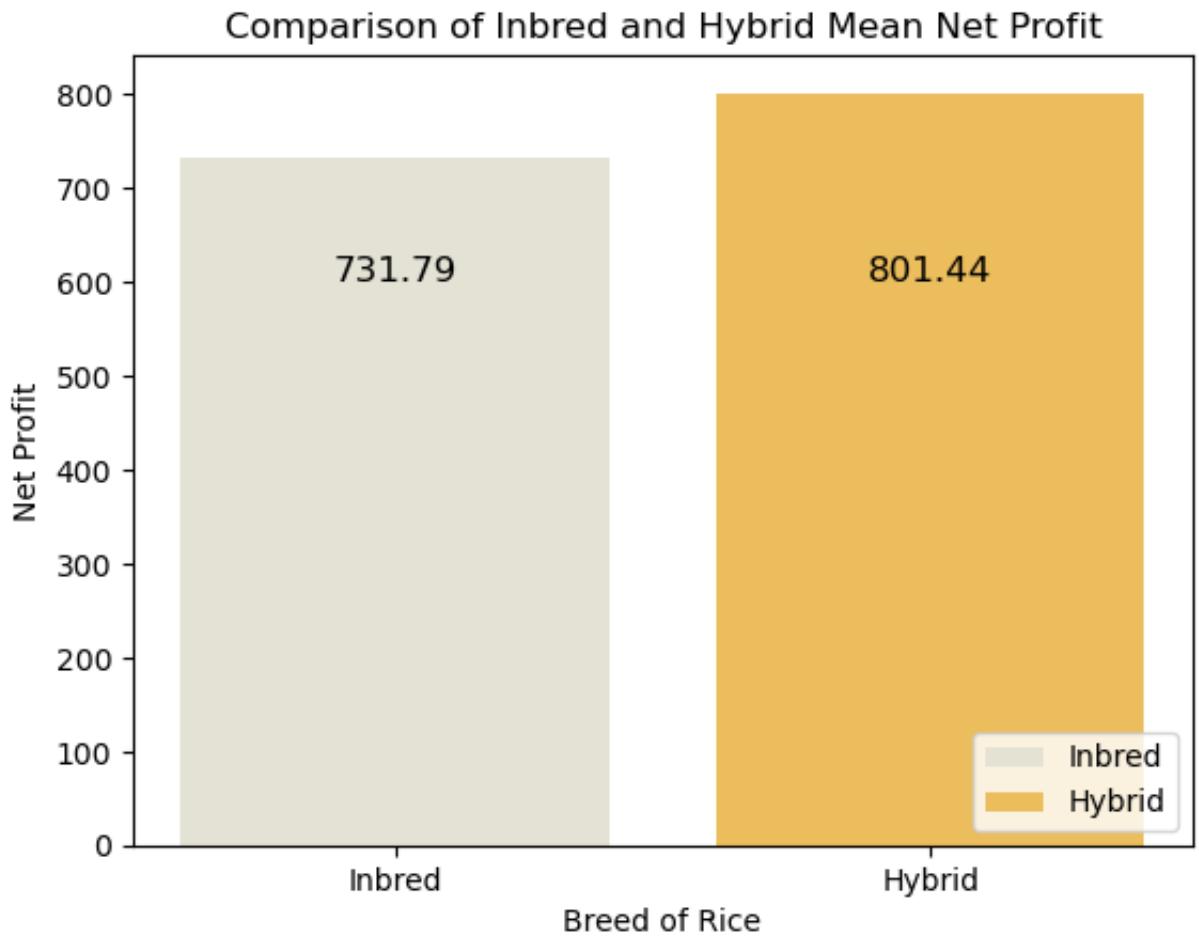


Figure 27. Comparison of Inbred and Hybrid Mean Net Profit After Removing The Outliers/center>

As seen here in the figure above, removing the outlier caused the means to differ significantly which means that one of the inbred rice farmers may be performing extremely well, which is not representative of most farmers and whose net profit may skew the mean of the inbred rice. Hence, it is best to remove the outlier when doing the hypothesis testing for these groups.

Hypothesis Testing

Hypotheses Overview and Discussion

The following are the hypotheses to be tested:

Table 4.

Hypotheses and Hypothesis Test

Research Question No	Null Hypothesis (H_0)	Alternative Hypothesis (H_a)	Hypothesis Test	Brief Explanation
				μ_1 is the mean yield per hectare

1	$\mu_1 = \mu_2$	$\mu_1 > \mu_2$	Z-test	of the hybrid variety while μ_2 is the mean yield per hectare of the inbred variety.
2	$\mu_1 = \mu_2$	$\mu_1 > \mu_2$	Z-test	μ_1 is the mean yield per hectare of the hybrid variety while μ_2 is the mean yield per hectare of the inbred variety, given a gravity irrigation system.
3	$\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$	At least one pair of means are significantly different from each other	One-way ANOVA	where μ_n is the mean yield per hectare for each breed-land elevation pair.
4	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	Mann-Whitney U Test	where μ_1 is the mean yield per hectare per labor day of the hybrid variety while μ_2 is the mean yield per hectare per labor day of the inbred variety.
5.1	$\mu_1 = \mu_2$	$\mu_1 > \mu_2$	Mann-Whitney U Test	where μ_1 is the mean total cost incurred from the hybrid variety while μ_2 is the mean total cost incurred from the inbred variety.
5.2	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	Mann-Whitney U Test	where μ_1 is the mean net profit derived from the hybrid variety while μ_2 is the net profit derived from the inbred variety.

The following are the hypotheses for each research question:

Research Question 1

Null Hypothesis: The mean yield per hectare for hybrid rice and inbred rice is equal.

Alternative Hypothesis: The mean yield per hectare for hybrid rice and inbred rice is

not equal, where the mean yield per hectare for hybrid rice is greater.

Rationale for Alternative Hypothesis: According to Biswas, et.al. [9], the yield potential for hybrid rice is significantly greater, and so this may be used as the rationale for setting the alternative hypothesis as hybrid rice yield per hectare as being greater in the alternative hypothesis.

Research Question 2

Null Hypothesis: Given similar gravity irrigation system, the mean yield per hectare for hybrid rice and inbred rice is equal.

Alternative Hypothesis: Given similar gravity irrigation system, the mean yield per hectare for hybrid rice and inbred rice is not equal, where the mean yield per hectare for hybrid rice is greater.

Rationale for Alternative Hypothesis: This is in line with the previous rationale where it was posited that hybrid rice has greater yield potential in general. Through this alternative hypothesis, it may be known whether this is true for a condition, that is, gravity irrigation system. [9]

Research Question 3

Null Hypothesis: Given different land elevations and rice breed pairs, the mean yield per hectare for hybrid rice and inbred rice is equal.

Alternative Hypothesis: There is at least one pair of breed-land elevation combination that has a significant difference of yield per hectare compared to the other combinations of breed-land elevation.

Rationale for Alternative Hypothesis: This is in line again with the previous rationale where it was posited that hybrid rice has greater yield potential in general. Through this alternative hypothesis, it may be known whether this is true for different conditions, that are, different land elevations. [9]

Research Question 4

Null Hypothesis: The mean yield per hectare per labor day for hybrid rice and inbred rice is equal.

Alternative Hypothesis: The mean yield per hectare per labor day for hybrid rice and inbred rice is not equal.

Rationale for Alternative Hypothesis: Given that there were no found studies positing the yield potential in terms of labor day, for this alternative hypothesis, it would be set as a two-tailed test, given the lack of assumptions.

Research Question 5

Null Hypothesis: The mean cost per yield for hybrid rice and inbred rice is equal.

Alternative Hypothesis: The mean cost per yield for hybrid rice and inbred rice is not equal, where the mean total cost incurred from the hybrid variety is greater.

Rationale for Alternative Hypothesis: According to Madeja, Nasibog, and Roxas [11], the farmers who ventured into hybrid rice varieties incurred the highest production costs. Thus, given this assumption, the alternative test would be hybrid rice having a greater mean total cost.

Null Hypothesis: The mean net profit per hectare of hybrid rice and inbred rice is equal.

Alternative Hypothesis: The mean net profit per hectare of hybrid rice and inbred rice is not equal.

Rationale for Alternative Hypothesis: According to Chenggapa, Gowda, and Janaiah, [12], in China, it was found that there is greater profitability in hybrid rice over inbred rice given a yield advantage of 15-20% and given the conditions in the study. Although the conditions may differ for the dataset used, given that it was found that hybrid rice is profitable in their study, the assumption can be used to support the alternative hypothesis where it is hypothesized that hybrid rice has a greater mean profit than inbred rice.

Hypothesis Tests Overview and Discussion

Z-TEST

In this report, we utilized the Z-test to compare the means of two independent groups—Hybrid and Inbred rice varieties—to assess whether the differences in their respective yields, costs, and other metrics were statistically significant. The Z-test is a statistical method that determines if there is a significant difference between sample means against a null hypothesis, which typically states that there is no difference. [13]

How the Z-test works is by calculating the Z-statistic, a value that indicates the number of standard deviations a data point is from the mean. If the Z-statistic falls within the range of the standard normal distribution (typically between -1.96 and +1.96 for a 95% confidence interval), the null hypothesis cannot be rejected, suggesting no significant difference between the groups. If the Z-statistic falls outside this range, the null hypothesis is rejected, indicating a statistically significant difference. [14]

The application of the Z-test in this context is supported by the assumption that the data samples are large enough for the Central Limit Theorem to apply, which assume that the sampling distribution of the sample mean approaches a normal distribution as the sample size grows, regardless of the population's distribution.

In the study, the Z-test is used to construct confidence intervals for the mean values of different variables, such as yield per hectare and gross earnings. These intervals provide a range within which the true mean of the population is expected to lie with a certain degree of confidence (95% in this case). By comparing these intervals between Hybrid and Inbred rice varieties, we can infer whether the observed differences in sample means likely reflect true differences in the population or are due to random variation.

ONE-WAY ANOVA

Analysis of Variance (ANOVA) was applied to evaluate the effect of land elevation on rice yield. ANOVA is a statistical method used to compare the means of three or more independent groups to see if at least one group mean is significantly different. [16] This is particularly important for agricultural studies where factors such as land elevation can have a profound impact on crop yields.

When the ANOVA results indicate significant differences, it suggests that elevation may play a crucial role in the productivity of rice varieties. However, ANOVA does not tell us which specific groups differ from each other.

MANN-WHITNEY U TEST

Aside from the Z-test, we also needed to use the Mann Whitney U test for determining the differences in two samples since some of the data was not normally distributed.

The Mann-Whitney U test is a non-parametric test that compares two independent samples to determine if they come from the same distribution. It is especially useful when the sample sizes are small or the data do not follow a normal distribution. [16]

The Mann-Whitney U test operates by ranking all data points from both groups together, regardless of which group they belong to, and then analyzing the sum of these ranks. If the ranks are significantly different between the two groups, the test indicates that the samples are likely drawn from different populations.

Results and Discussion

Results

Table 5.

Results of the Hypothesis Testing

Parameters	Hypothesis Test (0.05 level of significance)	Test Statistic	Test Statistic Value	p-value	Decision
0 Inbred and Hybrid Yield (KG/ha)	Z-test	Z-statistic	1.235843	0.108259	Fail to reject
1 Inbred and Hybrid Yield (KG/ha) (Gravity Irrigation)	Z-test	Z-statistic	1.222644	0.110732	Fail to reject
2 Inbred and Hybrid Yield (KG/ha) (per Land Elevation)	One-way ANOVA	F-statistic	1.032704	0.401377	Fail to reject
3 Yield per ha per Labor (KG/ha/labor)	Mann-Whitney U Test	Mann-Whitney U Statistic	11779.000000	0.187831	Fail to reject
4 Total costs (php)	Mann-Whitney U Test	Mann-Whitney U Statistic	13190.000000	0.228054	Fail to reject
5 Net Profit (php)	Mann-Whitney U Test	Mann-Whitney U Statistic	12033.000000	0.272897	Fail to reject

YIELD PER HA

Given the significance level of 0.05, the resulting p-value of 0.1082 suggests that we do not have sufficient evidence to reject the null hypothesis. Therefore, we fail to reject the null hypothesis, indicating that, at the 0.05 significance level, there is no significant difference in the mean yield per hectare for hybrid rice and inbred rice. This implies that the amount of rice produce harvested from one hectare of land is statistically considered to be equal, regardless of the breed of rice planted.

GRAVITY IRRIGATION

Given the significance level of 0.05, the resulting p-value of 0.1107 suggests that we do not have sufficient evidence to reject the null hypothesis. Therefore, we fail to reject the null hypothesis, indicating that, at the 0.05 significance level, given similar gravity irrigation system, there is no significant difference in the mean yield per hectare for hybrid rice and inbred rice. This implies that the amount of rice produce harvested from one hectare of land is statistically considered to be equal, regardless of the breed of rice planted, and regardless of irrigation type.

LAND ELEVATION

Given the significance level of 0.05, the resulting p-value of 0.4013 suggests that we do not have sufficient evidence to reject the null hypothesis. Therefore, we fail to

reject the null hypothesis, indicating that, at the 0.05 significance level, there is no significant difference in the mean yield per hectare for hybrid rice and inbred rice across different breed-land elevation pairs. This implies that the amount of rice produced harvested from one hectare of land, is statistically considered to be equal, regardless of the breed of rice planted or land elevation type.

LABOR

Given the significance level of 0.05, the resulting p-value of 0.1878 suggests that we do not have sufficient evidence to reject the null hypothesis. Therefore, we fail to reject the null hypothesis, indicating that, at the 0.05 significance level, there is no significant difference in the mean yield per hectare per labor days for hybrid rice and inbred rice. This implies that the labor productivity of planting either hybrid or inbred rice is the same.

COST & PROFIT

Given the significance level of 0.05, the resulting p-value of 0.2280 suggests that we do not have sufficient evidence to reject the null hypothesis. Therefore, we fail to reject the null hypothesis, indicating that, at the 0.05 significance level, there is no significant difference in the mean cost total cost for hybrid rice and inbred rice. This implies that the total cost for either breed is the same.

Given the significance level of 0.05, the resulting p-value of 0.2728 suggests that we do not have sufficient evidence to reject the null hypothesis. Therefore, we fail to reject the null hypothesis, indicating that, at the 0.05 significance level, there is no significant difference in the mean net profit for hybrid rice and inbred rice. This implies that the profitability of either breed is the same.

Discussion

The study reveals that while hybrid rice typically yields higher per hectare and may offer greater potential gross earnings, these advantages are not statistically significant in key aspects such as cost-effectiveness and labor efficiency when compared to inbred rice. This finding is crucial as it challenges the prevailing assumption of hybrid rice invariably leading to superior agronomic and economic outcomes.

Although from the Exploratory Data Analysis, hybrid rice shows a tendency for higher yields, this is accompanied by notable increases in both yield and cost variability. This suggests that the higher yields from hybrid rice come with greater risk, rendering it a less predictable choice for farmers. In contrast, inbred rice demonstrates more consistent labor efficiency and cost-effectiveness per kilogram as per its variance, highlighting its potential as a stable and reliable option for farmers, especially those with limited resources and risk management capabilities.

Contrasting the study with existing literature, such as Chengappa, Gowda, & Janaiah

[12], which reported a clear yield advantage for hybrid rice, our findings suggest that the benefits of hybrid rice might be context-dependent. This is supported by research by Chen, Mishra, Sha, & Yan [15], which found no yield superiority of hybrid rice over high-yielding inbred rice under specific conditions like higher planting density and reduced nitrogen application.

The greater variability associated with hybrid rice contradicts the expected consistent yield advantage often associated with hybrid varieties. In comparison, inbred rice's more consistent labor efficiency and cost-effectiveness align with findings pointing to high production costs involved in hybrid rice cultivation.

Interestingly, our study also indicates no significant difference in net profits between hybrid and inbred Rice, suggesting comparable profitability. This consistency in profitability underscores the need for a balanced approach to rice cultivation that weighs the benefits of higher yields against the risks and labor requirements.

Conclusion

In the study, we set out to find if there are significant differences in the productivity and profitability of planting either hybrid or inbred type of rice. With all hypothesis tests resulting to not significant, this suggests that the productivity and profitability of either type of rice is the same. Regardless of land elevation, irrigation type, or labor, the yield of both types is the same. And costs and net profits of either types is also the same.

In the current body of knowledge, hybrid rice is being put forward as the option for farmers which produces the most yields as compared to inbred rice [8] [9]. As the tests have shown, at least within the population of this dataset, which is for Isabela and Bukidnon farmers, this has proven to not be the case. According to the Rice Knowledge Bank [10], hybrid rice should produce more yields and are more resistant to diseases and insects. Though they may indeed increase yields, it is not a significant difference to what farmers are used to, which are the inbred type of rice.

Hence, the findings posit that projects involving cultivating hybrid or inbred rice should be decided and strategized well, catering to individual circumstances instead of a one-size-fits-all approach. In line with this, policymakers and agricultural advisors should provide context-specific support to farmers which include risk management training, market trend insights, and crop selection advice, empowering them to make informed decisions which are line with their contexts and goals.

Recommendations

In farming, the geography and specific location are significant factors that affect the production and yield for farmers. Data on rice production within different locations aside from the dataset is good to compare to see whether that has an effect on the profitability and productivity of different rice breeds. Furthermore, the conditions and assets of farmers are valuable factors to take into consideration on the yield as well. Irrigation, fertiliser, and specific cultivation methods are all features that can be specifically studied with larger data sets to better understand their impact to yield. Ultimately, the most important output in a study regarding rice production is to be able to create a guide for farmers on when it is ideal to cultivate a certain breed.

References

- [1] Cabrera, Ellanie. Socio-Economic Evaluation of Hybrid Rice Cultivation in the Philippines, Harvard Dataverse, V3 (2014). Retrieved from
https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi%3A10.7910%2FDVN%2F26246&fbclid=IwAR2y-A50Xg_zmCMFiDFaSXxUWtn5SxepsIkzSIZHhkZYYpRWvk5DkUiEp1c
<https://doi.org/10.7910/DVN/26246>
- [2] Department of Science and Technology, The Philippine Council for Agriculture, Aquatic, and Natural Resources Research and Development. Rice, Industry Strategic Science and Technology Program (nd). Retrieved from
<https://ispweb.pcaarrd.dost.gov.ph/isp-commodities/rice/#:~:text=The%20Philippines%20alone%20recorded%204.81,agricultur>
- [3] Department of Agriculture: Philippine Rice Research Institute. Palay Production (nd). Retrieved from <https://www.philrice.gov.ph/ricelytics/productions>
- [4] Department of Agriculture. Department of Agriculture 2022 Annual Report (2023). Retrieved from <https://www.da.gov.ph/media-resources/da-annual-reports/>
- [5] Minitab. (n.d.). Data Considerations for One-Way ANOVA. Retrieved from
<https://support.minitab.com/en-us/minitab/21/help-and-how-to/statistical-modeling/anova/how-to/one-way-anova/before-you-start/data-considerations/#:~:text=The%20actual%20sample%20size%20that,should%20be%20at>
- [6] 10.2.1 - ANOVA Assumptions | STAT 500. (n.d.). PennState: Statistics Online Courses. <https://online.stat.psu.edu/stat500/lesson/10/10.2/10.2.1>
- [7] Zach. (2019). How to Check ANOVA Assumptions. Statology. Retrieved from <https://www.statology.org/anova-assumptions/>

[8] Chengappa, P.G., Gowda, M.V.S., & Janaiah A(1994). Profitability of Hybrid Rice Cultivation: Evidence from Karnataka. Economic and Political Weekly. Retrieved from: <https://www.jstor.org/stable/4413711>

[9] Biswas J.K., Iftekharuddaula, K.M., Haque, M.M., Hasanuzzaman, M., & Pramanik, H.R. (2015). Comparative Performance of Hybrid and Elite Inbred Rice Varieties with respect to Their Source-Sink Relationship. The Scientific World Journal, 2015. <https://doi.org/10.1155/2015/326802>

[10] Rice Knowledge Bank. (n.d.). Hybrid Rice. Retrieved from <http://www.knowledgebank.irri.org/training/fact-sheets/crop-establishment/item/hybrid-rice-fact-sheet#:~:text=Increased%20yields%2C,resistance%20to%20diseases%20and%20insec>

[11] Madeja, G.C., Nasibog, L.D., & Roxas, C.A. (2015). Comparing the Profitability of Hybrid, Inbred, and Organic Rice Production. Semantic Scholar. Retrieved from <https://www.semanticscholar.org/paper/Comparing-the-Profitability-of-Hybrid%2C-Inbred%2C-and-Nasibog-Roxas/>

[12] Chengappa, P.G., Gowda, M.V.S., & Janaiah A(1994). Profitability of Hybrid Rice Cultivation: Evidence from Karnataka. Economic and Political Weekly. Retrieved from: <https://www.jstor.org/stable/4413711>

[13] Chen, J. (2023, August 20). Z-Test Definition: Its Uses in Statistics Simply Explained With Example. Retrieved from <https://www.investopedia.com/terms/z/z-test.asp#toc-what-is-a-z-test>

[14] Howell, E. (2023, February 02). Z-Test for Statistical Hypothesis Testing Explained. Retrieved from <https://builtin.com/data-science/z-test-statistics>

[15] Chen, F., Mishra, A., Sha, W., & Yan, Zedong. (2022). An economic assessment of adoption of hybrid rice: Micro-level evidence from southern China. Frontiers in Sustainable Food Systems. <https://doi.org/10.3389/fsufs.2022.1066657>

[16] Kenton, W. (2023, June 12). Analysis of Variance (ANOVA) Explanation, Formula, and Applications. Retrieved from <https://www.investopedia.com/terms/a/anova.asp>