STA 106 Project I Group 11

By: Marsha He, Dongmin Wu, Abner Cabrera, Cina Abbasi Baharanchi

Instructor: Maxime Guiffo Pouokam

2024-08-20

Picture of Sparrow and its nest

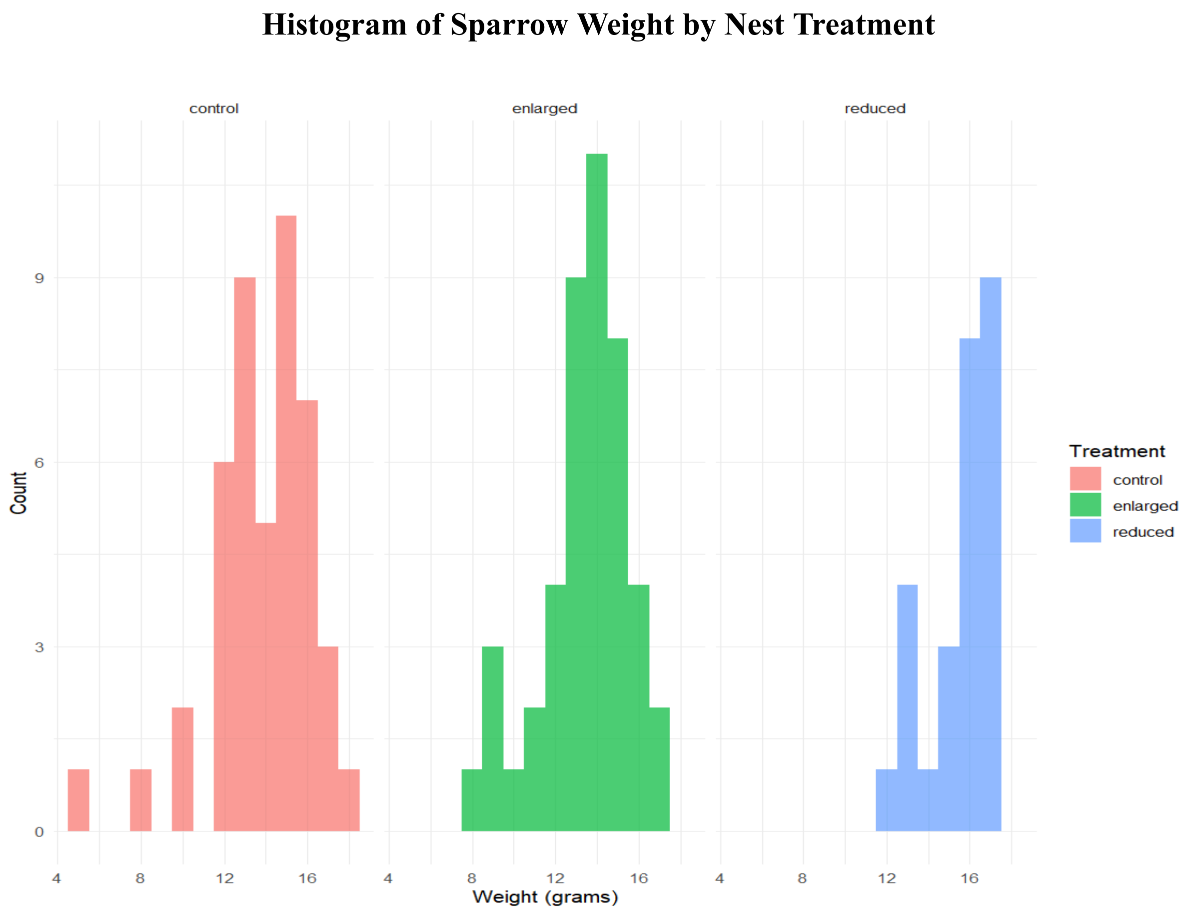
I. Introduction

The “sparrow” dataset contains data from an observational study conducted to investigate whether different nest sizes on Kent Island influence the size of sparrows they attract. Our independent variable is the type of nest, with its three factor levels of control(not manipulated), enlarged(manipulated to be a larger nest than normal), or reduced(manipulated to be smaller than normal) being used to measure the difference of weight of sparrows, which is our dependent variable.

We are interested in this question because understanding the relationship between nest size and sparrow size is significant. It can provide insights into avian ecology, particularly how environmental factors like nest size may impact the growth and development of sparrows. To answer this question, we would hypothetically test the equal means by utilizing a 5% significance level and an F-test of equal means for the Single Factor Anova group means. The F-test statistic will give us an idea regarding our data, where the p-value is the probability that, upon collecting our data, we may find that the average weight of sparrows associated with each nest type was in fact the same. If our observed p-value is lower than the 5% significance level, we can then suggest that it is highly likely that at least one of the nest types has a different average sparrow weight from the rest. If the p-value is larger than the significance level, we cannot reject the claim that the average sparrow weight for each nest type is the same . This test, known as ANOVA, is performed to determine the possibility that one of the nest types has a statistically different average sparrow weight from the other, and with this information, we can proceed to our next target.

Next, we would like to see how the sparrow weight is the same or different in the treatment of each nest. In this regard, we will build confidence intervals on the same scale, which will show how much weight each type of sparrow has on average in all nest treatments. Accordingly, the intervals of confidence will be three: the first - comparing the average weight of sparrows in the control nest to those in the enlarged nest, the second - comparing the average weight of sparrows in the control nest to those in the reduced nest, and the third - identifying the nest treatment that tends to attract the largest sparrows. These confidence intervals will indicate if the effect of all nests on sparrow size is different from the control node. Since there is a need to have several simultaneous confidence intervals, applying Turkey's and Bonferrioni's multiplier will be a more general approach. Turkey's multiplier is specifically applied to analyze paired comparisons, while the Bonferroni multiplier is useful when different types of comparisons are being made. Through a T-test multiplier comparison, we can identify the confidence level that provides the narrowest and appropriate limits with 95% power, enabling correct interpretation of the differences in body weights of the sparrows in the nest treatment. With the help of these confidence intervals, we will also determine which nest treatment, either enlarged, reduced, or controlled, seemed to be more mesmerizing and least appealing. Additionally, we would analyze how each nest treatment is affecting the overall average sparrow weight across various nests. A suitable way to do this would be through the utility of the Greek letter gamma(γi), which allows us to get the average effect that a specific nest treatment's sparrow weight has on the overall sparrow population on Kent Island. By comparing and contrasting the impacts of each nest treatment, we can have evidence of which nest type has the overall best or worst effect on sparrow size compared to the other nest types.

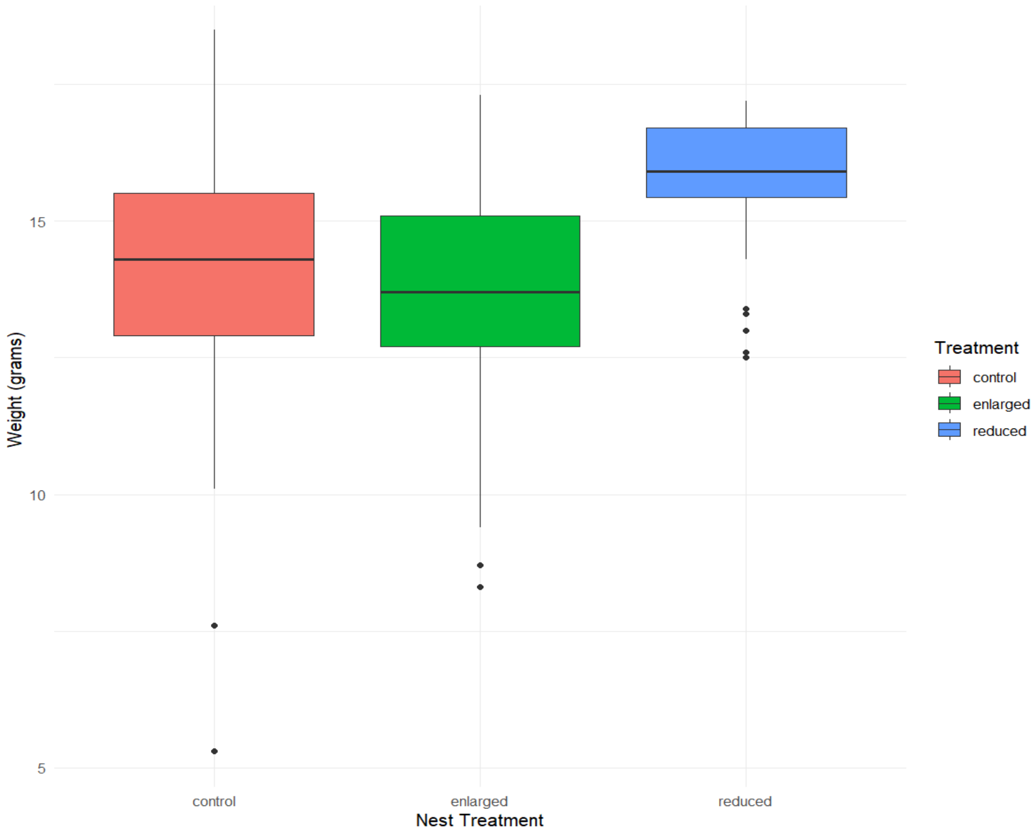
II. Summary of Data



The histograms display the distribution of sparrow weights for each treatment group:

- **Control group**: The weights are approximately normally distributed, centered around 14 grams. The graph is slightly skewed towards the lighter sparrows. The majority of sparrows in the group weigh between 12 and 16 grams.
- **Enlarged Group**: The weight distribution in the enlarged nests is narrower compared to the control group. The distribution is slightly skewed towards heavier sparrows, with most weights concentrated between 12 and 16 grams.
- **Reduced Group**: The reduced nest treatment's distribution is clearly shifted towards heavier sparrows, with most weights falling between 14 and 18 grams. There are fewer lighter sparrows in this group.

Boxplot of Sparrow Weight by Nest Treatment



The boxplots illustrate the central tendencies and variabilities in sparrow weights among the three treatments:

- **Control Group**: The median weight is approximately 14 grams, with a broader interquartile range (IQR) indicating moderate variability. Some outliers are present on the lower end, reflecting a few significantly lighter sparrows.
- **Enlarged Group**: This group has a lower median weight (around 13.5 grams) and smaller IQR than the control group, which means the weights are more closed to mean. The distribution shows fewer outliers.
- **Reduced Group**: The median weight in this group is the highest around 16 grams, with the smallest IQR which indicates the least variability in weights. The reduced treatment also shows some outliers on the heavier side, which suggests that these nests tend to attract larger sparrows.

Summary Table

	Control	Enlarged	Reduced	Overall
Means	**13.9**	**13.5**	**15.6**	**14.13**
Std. Dev	**2.42**	**2.10**	**1.46**	**2.24**
Sample Size	**45**	**45**	**26**	**116**

The summary table provides the mean, standard deviation, and sample size for each treatment group:

- **Control Group**: Mean weight is 13.9 grams with a standard deviation of 2.42 grams, based on 45 observations.
- **Enlarged Group**: Mean weight is slightly lower at 13.5 grams, with a standard deviation of 2.10 grams, also from 45 observations.
- **Reduced Group**: Mean weight is notably higher at 15.6 grams, with a lower standard deviation of 1.46 grams, based on 26 observations.
- The **Overall** mean weight of sparrows is 14.13 grams, with a standard deviation of 2.24 grams, from a total of 116 observations.

Summary:

The data suggests that the reduced nest treatment has the heaviest sparrows, with a higher mean weight and less variability in weights compared to the control and enlarged treatments. The control and enlarged treatments, while having similar mean weight, show more variability and include some lighter sparrows, particularly in the control group.

This pattern could imply that smaller nests (reduced treatment) are more likely to attract or support larger sparrows on Kent Island. The following analysis and hypothesis testing will help to confirm whether these observed differences are statistically significant.

III. Diagnostics

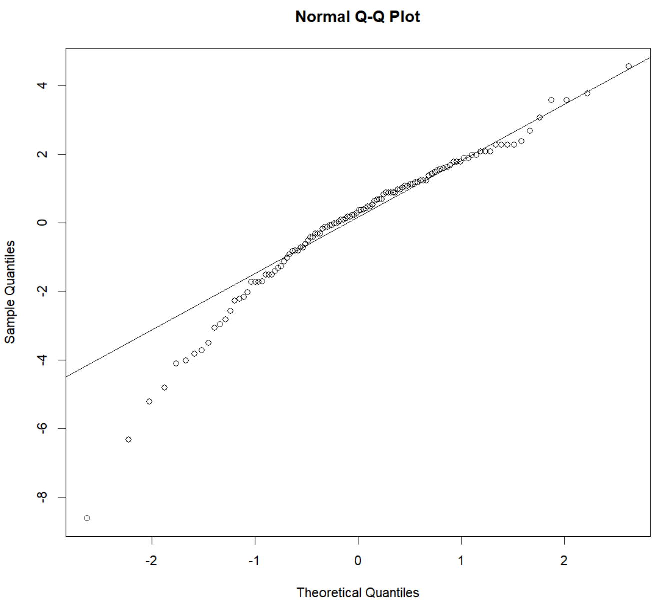
Assumptions:

1. **Independence of Observations:** The sparrow weights within each treatment group should be independent of each other. Which means the weight of one sparrow does not influence the weight of another.
2. **Normality of Residuals:** The residuals, or differences between the observed and predicted values should be normally distributed. This is crucial for the accuracy of p-values and confidence intervals in hypothesis testing.
3. **Homogeneity of Variances:** The variance of sparrow weights across the different treatment groups should be approximately equal.

Diagnostic Checks:

Q-Q Plot Analysis

1. Before Removing Outliers:

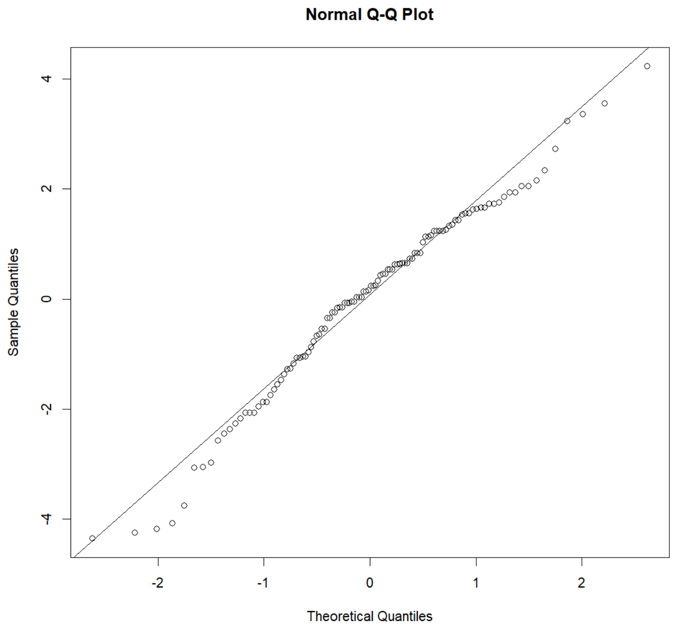


The points on the Q-Q plot do not fall approximately along the reference line, showing a noticeable deviation from the line. The points in the lower and upper ends of the distribution curve away from the line, indicating the residuals are not perfectly normally distributed.

By Shapiro-Wilk normality test, W=0.94055, p-value=0.00006245

Since the p-value is less than the significance level of 0.05, we reject the null hypothesis that the residuals are normally distributed. Which also confirms the observation from the QQ plot.

2. After Removing Outliers:



The Q-Q plot after removing the outliers shows a much closer alignment of the points to the diagonal line with only slight deviations at the ends. This improvement suggests that removing the outliers has enhanced the normality of the residuals.

By Shapiro-Wilk normality test, W=0.97807, p-value=0.0621

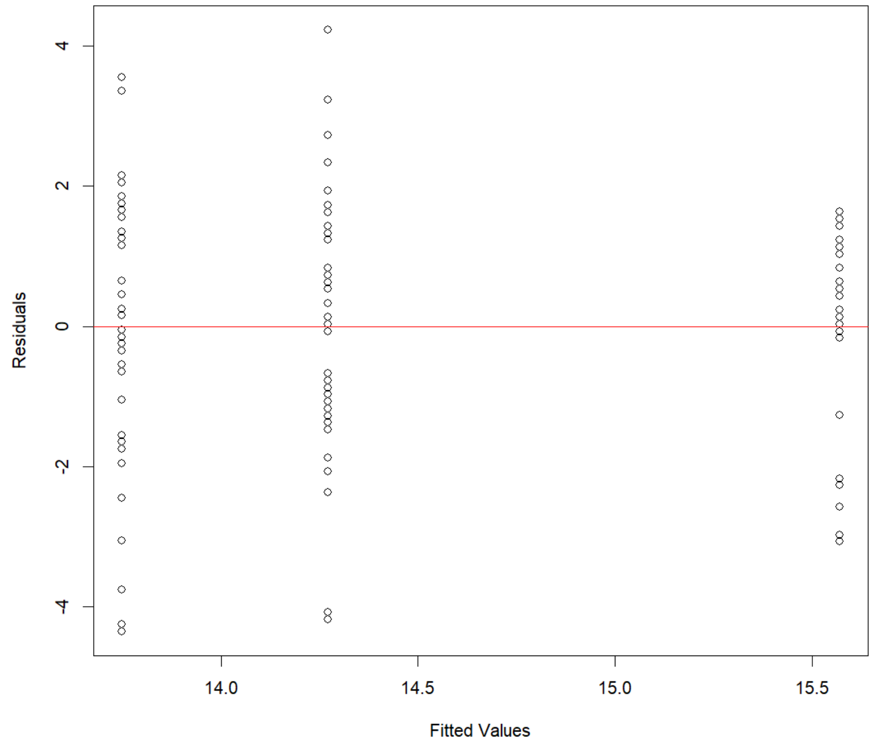
The p-value is now above 0.05, indicating that we fail to reject the null hypothesis, which means the residuals can be considered normally distributed.

Brown-Forsythe Test

Degree of freedom	F value	p-value
(2, 109)	1.0434	0.3558

p-value interpretation: The p-value is much higher than 0.05, which means we fail to reject the null hypothesis. This suggests that there is no significant evidence to conclude that the variances across the treatment groups are different.

Residuals vs. Fitted Values Plot



The spread of residuals are consistent across the range of fitted values, which supports the assumption of homoscedasticity. There is no noticeable evidence or other systematic patterns that would indicate increasing or decreasing variance.

The residuals appear to be randomly scattered, there are no systematic patterns or trends in the plot, which suggests the independence of errors. Residuals are behaving as expected, which also supports the validity of the model.

Conclusion on Assumptions:

- Independence: Assumed to be met based on the study design.
- Normality: The assumption of normality was initially violated but was addressed by removing outliers. Removing outliers resulted in normally distributed residuals.
- Homogeneity of Variance: Was confirmed by the brown-Forsythe test and the residuals vs fitted values plot.

IV. Analysis of Data

Model Fit

We will be using the model:

$Yij = \mu i + \epsilon ij$

We are looking to find a change in weight of sparrows between a control group, enlarged nest group, reduced nest group. Mainly comparing the weight between these 3 groups to see if there is similarity or 1 different group. In this μi is the average weight of the group chosen. ANOVA is the way we will estimate averages for each group. i = A,B,C, for the 3 groups. In the model, ϵij is for individual error for any given jth value in the ith group.

Hypothesis Test

Our main goal is to find a differentiation in our group means to figure out if weight varies. Our Null and Alternative would be as follows:

H_0: all groups equal ($\mu A = \mu B = \mu C$)

H_A: at least one group different

The null here is claiming there would be no difference in weight average. This is our basis for what we are trying to support or reject. The alternative if null is rejected would be if one is different from the others in weight based on the mean.

We will be calculating the test statistic for the hypothesis by using the ANOVA below. The test statistic is equal to the F test (MSA/MSE). If the test statistic is equal to 1 there is no difference between the groups. We also get the P-value from ANOVA.

Anova Table

	Df	Sum Sq	Mean Sq	F	Pr(>F)
Treatment	**2**	**72.74**	**36.372**	**8.1288**	**0.0005031**
Residuals	**113**	**505.62**	**4.474**	**n/a**	**n/a**

From the ANOVA we find that the test statistic comes out to 8.1288. The P-value comes out to 0.0005031. Using this data later we will conclude our findings.

Confidence Intervals

We will also be constructing 95% simultaneous confidence intervals for pairwise comparisons across groups. The pairwise confidence intervals that we will be analyzing are as follows:

$\mu A - \mu B$

$\mu A - \mu C$

$\mu B - \mu C$

The first pairwise confidence interval we will construct is $\mu A - \mu B$. The results of this confidence interval are as follows:

$\mu A - \mu B = (-2.6771 , -0.6124)$

The next pairwise confidence interval we will construct is $\mu A - \mu C$. The results of this confidence interval are as follows:

$\mu A - \mu C = (-0.4746 , 1.2923)$

The next pairwise confidence interval we will construct is $\mu B - \mu C$. The results of this confidence interval are as follows:

$\mu B - \mu C = (-3.0860 , -1.0213)$

Factor Effects Model

In order to conduct our analysis of the data, we developed a single factor ANOVA factor effect model. This model follows the following equation: $Yij = \mu. + \gamma i + \epsilon ij$.

The Factor effect model is specifically useful for use because it investigates the differences in overall mean and factor level means. In this equation, $\mu.$ represents the ground mean which is the common mean for all groups. γi represents the "effect" by acting as a mathematical term that calculates how the overall mean is affected by a subject following i. Lastly, ϵij represents error.

To calculate the value of the gammas, we will be subtracting the mean of the group means, or Y(double bar) , from each respective group mean. In this case, the respective values of the Gammas are as follows:

	Control	Enlarged	Reduced
Gamma	**-0.4120**	**-0.8209**	**1.2328**

Power Calculation

Power is a significant term that finds the probability of rejecting the null hypothesis given that the null hypothesis is wrong. In other words, power is the probability of not making a Type II error. To determine power, we must first calculate the ϕ parameter. The equation for this parameter is as follows:

$\phi = sqrt((1/\sigma\epsilon) * p\ (\Sigma n\ i = 1ni * (\mu i - \mu.)\ 2)/a)$

```
power = 0.931603
```

The ϕ parameter measures how different Fs under H_A is to Fs under the H_0. Notably, the $\sigma\epsilon$ term in the equation represents the standard error which is often approximated or estimated using the square root of the Mean Squared Error(MSE). As such, we replaced $\sigma\epsilon$ with 5.377458 in our calculation. Another important value we chose was alpha = 0.05 .

As such, the resulting power of this test is about 0.9316, meaning that we have a probability of 0.9316 to correctly reject H_0 given that the H_0 is false. In the context of our problem and dataset, this means that the probability of correctly rejecting the null hypothesis, that all the average population weight values regardless are the same, when at least one of the the average population weight loss values is not equal is 0.8778.

V. Interpretation

The analysis of the sparrow dataset reveals significant differences in sparrow weights across the three nest treatments—Control, Enlarged, and Reduced. The ANOVA results indicate a statistically significant effect of nest treatment on sparrow weight ($F(2,113)=8.13, p=0.0005031$), allowing us to reject the null hypothesis that the mean weights are equal across all treatments.

The sparrows in the reduced nest treatment have the highest mean weight of 15.6 grams with the smallest variability, as indicated by a standard deviation of 1.46 grams. This suggests that smaller nests are more likely to attract larger sparrows. In contrast, the control and enlarged nests have lower mean weights of 13.9 grams and 13.5 grams, respectively, with the control group showing the highest variability (standard deviation of 2.42 grams). The factor effects (γ) for each treatment further highlight these differences, with the reduced nests showing a positive effect (γ reduced=1.233 grams), indicating heavier sparrows, while the control and enlarged nests show negative effects (γ control= −0.412 grams and

γ enlarged= −0.82 grams), indicating lighter sparrows relative to the overall mean weight of 14.13 grams.

The reduced group's weight distribution is skewed towards heavier sparrows, with most weights concentrated between 14 and 18 grams. The control and enlarged groups exhibit broader weight distributions, with the control group having a median weight around 14 grams and the enlarged group around 13.5 grams. These differences in distribution suggest that reduced nests are more effective at attracting larger sparrows, while the other treatments attract a more varied population. The confidence intervals for pairwise comparisons between treatments further support these observations. For instance, the 95% confidence interval for the difference in mean weights between the control and enlarged groups is (−2.6771,−0.6124) grams, confirming that sparrows in the control group are significantly heavier than those in the enlarged group. Similarly, the confidence interval between the reduced and enlarged groups (−3.0860,−1.0213) grams indicates that sparrows in the reduced nests are significantly heavier than those in the enlarged nests.

Additionally, the power of the test, calculated to be 0.9316, indicates a high probability (93.16%) of correctly rejecting the null hypothesis when it is false, reinforcing the robustness of the ANOVA results. Diagnostics performed on the residuals further confirm the assumptions necessary for the validity of the ANOVA. The Shapiro-Wilk test results show that, after removing outliers, the residuals follow an approximately normal distribution ($W= 0.97807$, $p= 0.621$), which supports the reliability of the ANOVA analysis.

VI. Conclusion

Based on our data analysis, we observed that the weight of sparrows in the reduced nest tends to have the highest average values. Our statistical data analysis through the hypothesis testing of equal sample means indicates that at least one nest treatment average weight of sparrows is different from the other. However, it is important to note that the hypothesis test does not necessarily determine the specific nest treatments that stand out.

By further analyzing, the result of confidence intervals shows the fact that there is no significant difference between the true average sparrow weight in the control and enlarged nests. But still, there is conclusive evidence to make a statement that sparrows in the reduced nests are heavier in comparison with the control nest. Moreover, the reduced nests result in sparrows that are generally heavier than those in the enlarged nests. Therefore, on the strength of confidence intervals inferences, it is evident that the reduced nests attract the largest sparrows in comparison to the other nest types. More specifically, the means that differ significantly from each other are those of the reduced nests compared with both the control and enlarged nests.

The factor effect model was beneficial because it allowed us to look at the factor level means (as each nest type) and see how the weight contributions differed in general. This analysis helped us to recognize that the sparrows in the reduced nests contributed 1.2328 more grams to the mean weight from the overall mean. On the contrary, the random enlarged nests shrank the overall mean weight by 0.8209 grams. The results become even clearer again that the reduced nest type is the one that has the biggest simple impact as it seems to show a positive trend on the overall mean weight.

At last, we need to investigate the power of our test as well. As a consequence of the fact that the null hypothesis was rejected in our initial testing process, we concluded that the alternative hypothesis must be true. To confirm that the null is indeed not true, we need to calculate the power of our test or the probability that the null is rejected when the alternative is true in actuality. After the calculation of power had been completed, we came to the conclusion that it was fairly strong to support our decision to reject the null hypothesis, given that the alternative is true.

Appendix

```r
#library
library(ggplot2)
library(dplyr)

#load data
data <- read.csv("E:/Study/STA106/project1/sparrow.csv") #load data

#histogram
ggplot(data, aes(x = Weight, fill = Treatment)) +
  geom_histogram(binwidth = 1, alpha = 0.7, position = 'identity') +
  facet_wrap(~Treatment) +
  labs(title = "Histogram of Sparrow Weights by Nest Treatment",
       x = "Weight (grams)", y = "Count") +
  theme_minimal()

#boxplot
ggplot(data, aes(x = Treatment, y = Weight, fill = Treatment)) +
  geom_boxplot() +
  labs(title = "Boxplot of Sparrow Weights by Nest Treatment",
       x = "Nest Treatment", y = "Weight (grams)") +
  theme_minimal()

#summary table
summary <- data %>%
  group_by(Treatment) %>%
  summarise(
    Mean = mean(Weight),
    SD = sd(Weight),
    SampleSize = n()
  )
print(summary_table)

#overall summary table
overallSummary <- data %>%
  summarise(
    Mean = mean(Weight),
    SD = sd(Weight),
    sampleSize = n()
  )
print(overallSummary)

# fit ANOVA model
model <- aov(Weight ~ Treatment, data = data)

# qq plot for residuals
qqnorm(residuals(model))
qqline(residuals(model))

#shapiro-Wilk test
shapiro.test(residuals(model))


#identify outliers
outliers <- boxplot.stats(data$Weight)$out

# filter out outliers
dataNoOutliers <- data[!(data$Weight %in% outliers),]
```

```r
# model no outliers
modelNoOutliers <- aov(Weight ~ Treatment, data = dataNoOutliers)

#qq plot no outliers
qqnorm(residuals(modelNoOutliers))
qqline(residuals(modelNoOutliers))
shapiro.test(residuals(modelNoOutliers))


install.packages("car")
library(car)

str(data)
str(dataNoOutliers)

# levenetest
leveneTest(Weight ~ Treatment, data = dataNoOutliers, center = median)


modelNoOutliers <- aov(Weight ~ Treatment, data = dataNoOutliers)


# residuals vs fitted
plot(fitted(modelNoOutliers), residuals(modelNoOutliers),
     xlab = "Fitted Values", ylab = "Residuals",
     main = "Residuals vs Fitted Values (No Outliers)")
abline(h = 0, col = "red")
```{r}

#load library
library(ggplot2)
library(dplyr)

#load data
data <- read.csv("~/STATS 108 R/sparrow.csv")
Weight = sparrow$Weight
Treatment = sparrow$Treatment

#histogram
ggplot(data, aes(x = Weight, fill = Treatment)) +
  geom_histogram(binwidth = 1, alpha = 0.7, position = 'identity') +
  facet_wrap(~Treatment) +
  labs(title = "Histogram of Sparrow Weights by Nest Treatment",
       x = "Weight (grams)", y = "Count") +
  theme_minimal()

#boxplot
ggplot(data, aes(x = Treatment, y = Weight, fill = Treatment)) +
  geom_boxplot() +
  labs(title = "Boxplot of Sparrow Weights by Nest Treatment",
       x = "Nest Treatment", y = "Weight (grams)") +
  theme_minimal()

#summary table
summary_table <- data %>%
  group_by(Treatment) %>%
  summarise(
    Mean = mean(Weight),
```
```

```
    Std_Dev = sd(Weight),
    Sample_Size = n()
  )
print(summary_table)

#overall summary table
overall_summary <- data %>%
  summarise(
    Mean = mean(Weight),
    SD = sd(Weight),
    sampleSize = n()
  )
print(overall_summary)
```

```{r}
give.me.power = function(ybar,ni,MSE,alpha){
  a = length(ybar) # Finds a
  nt = sum(ni) #Finds the overall sample size
  overall.mean = sum(ni*ybar)/nt # Finds the overall mean
  phi = (1/sqrt(MSE))*sqrt( sum(ni*(ybar - overall.mean)^2)/a)
  phi.star = a *phi^2 #Finds the value of phi we will use for R
  Fc = qf(1-alpha,a-1,nt-a) #The critical value of F, use in R's
function
  power = 1 - pf(Fc, a-1, nt-a, phi.star)# The power, calculated using
a non-central F
  return(power)
}

group.means =  by(sparrow$Weight,sparrow$Treatment,mean)
group.nis = by(sparrow$Weight,sparrow$Treatment,length)
the.model = lm(Weight ~ Treatment, data = sparrow)
anova.table = anova(the.model)
MSE = anova.table[2,3]
```

```{r}
the.power = give.me.power(group.means,group.nis,MSE,0.05)
the.power
overall.mean = sum(group.means*group.nis)/sum(group.nis)
effect.size = sqrt( sum( group.nis/sum(group.nis) *(group.means
-overall.mean)^2 )/MSE)
library(pwr)
pwr.anova.test(k = 3, f = effect.size, sig.level = 0.05, power = 0.95)
```
```{r}
give.me.CI = function(ybar,ni,ci,MSE,multiplier){
  if(sum(ci) != 0 & sum(ci !=0 ) != 1){
    return("Error - you did not input a valid contrast")
  } else if(length(ci) != length(ni)){
    return("Error - not enough contrasts given")
  }
  else{
    estimate = sum(ybar*ci)
    SE = sqrt(MSE*sum(ci^2/ni))
    CI = estimate + c(-1,1)*multiplier*SE
    result = c(estimate,CI)
    names(result) = c("Estimate","Lower Bound","Upper Bound")
    return(result)
```

```
  }
}

t.value = qt(1-0.05/2, sum(group.nis) - length(group.nis))
ci.1 = c(1,0,-1)
ci.2 = c(1,-1,0)
ci.3 = c(0,1,-1)

CI1 = give.me.CI(group.means,group.nis,ci.1,MSE,t.value)
CI2 = give.me.CI(group.means,group.nis,ci.2,MSE,t.value)
CI3 = give.me.CI(group.means,group.nis,ci.3,MSE,t.value)
CI1
CI2
CI3

gammai = group.means - mean(group.means)
```
```
```{r}
library(pwr)

k <- 3 #number of groups
n <- (45 + 45 + 26) / 3 # average number of group members
f <- sqrt((72.74 / (72.74 + 505.62))) # effect size (Cohen's f)
alpha <- 0.05 # significance level
power <- pwr.anova.test(k = k, n = n, f = f, sig.level = alpha)
print(power)
```
```