STA106 Project I Group 2

How the government allocate the healthcare in the different region

By: Xuanting Guo, Yun-Hsiu Lu, Yaqing Pi, Angel Perez

Instructor: Maxime Guiffo Pouokam

Date: 2023/08/24

I. Introduction

In the rapidly evolving healthcare landscape, understanding patient length of stay at hospitals is crucial for optimizing hospital operations, patient care, and resource allocation. While different factors can influence a patient's duration of stay, one potential influencer is the geographical region of the hospital. Given the diverse socio-economic, cultural, and healthcare infrastructure variations across regions, it becomes pertinent to examine if and how these regional distinctions are interrelated with the length of stay.

In this report, we expect to see statistically significant differences among the regions. If we successfully find the variance for the regions, we can further make some suggestions for the government and local hospitals. We are going to analyze whether we should consider the region as a factor that can cause the length of staying at hospital for the patient. We choose to use the "Senic" data set as our data resources. "Senic" is an observational study that investigated the different regions with the patient's length of staying at hospital. The dataset compiles a variety of random samples that includes four different regions hospitals that are what we are going to consider as our factor levels; which were named NE (North East), W (West), S(South) and NC (North Central). To find out if the average value of each group is significantly different from others, we firstly assume they are equal to each other, and then calculate how likely it is going to happen given sample data. Since the data set only contained the regions, we decided to use Single-Factor Anova to analyze it. We are trying to investigate the difference in the regions with the length of patient staying. With using Anova, We are able to know the difference between each region. We can know which region has the worst performance on length of staying at hospital. Also, by comparing the average value of different regional groups, we want to get some statistical inference on their difference, and see if any of them can be combined.

II. Summary of your data.

II.1. Graphs by group

II.1.1. Boxplot of length by region

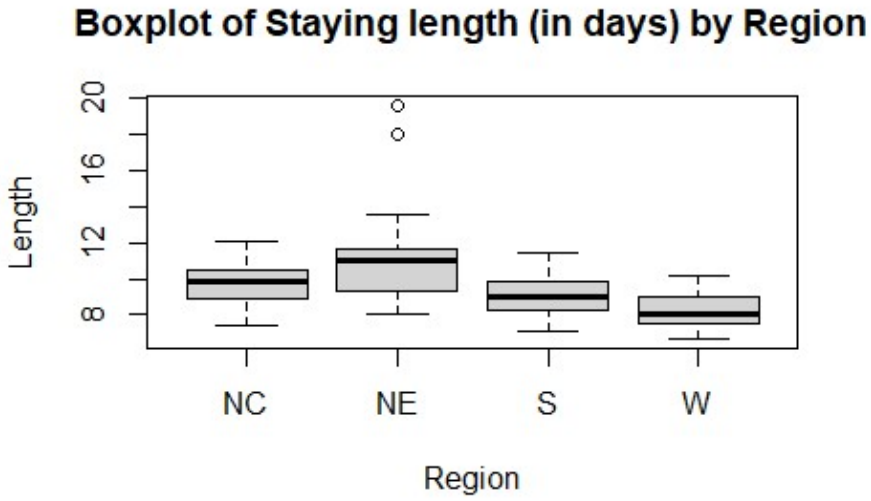

Figure 1 Boxplot of Staying length by region

The boxplot provides a visual representation of the distribution of length of stay for patients at the hospital (in days) across different regions. First, the median lines within each box highlight the central tendency. We can see that the median stay for all regions seems to be close to 9 days, except for the patients in the North East. It seems like patients in the North East tend to have a longer recovery time in the hospitals. As for the patients in the West, they tend to have a shorter stay with an average of 8.22 days. Second, we can observe variations in the length of stay within each group as indicated by the range covered by the boxes. For the northeast region, there is a fairly large variance in the data. With higher variance, it means that the variances are spread out from the median farther compared to the other regions. From the three box plots, we

conclude that patients in the Northeast have a positive tendency on average days staying in the hospitals, compared to the other regions. Reversely, the patients in the West have a negative effect on average recovery staying in the hospitals. This means that patients in the West tend to recover faster than those in the north east.

II.1.1. Histogram of Length Staying by region

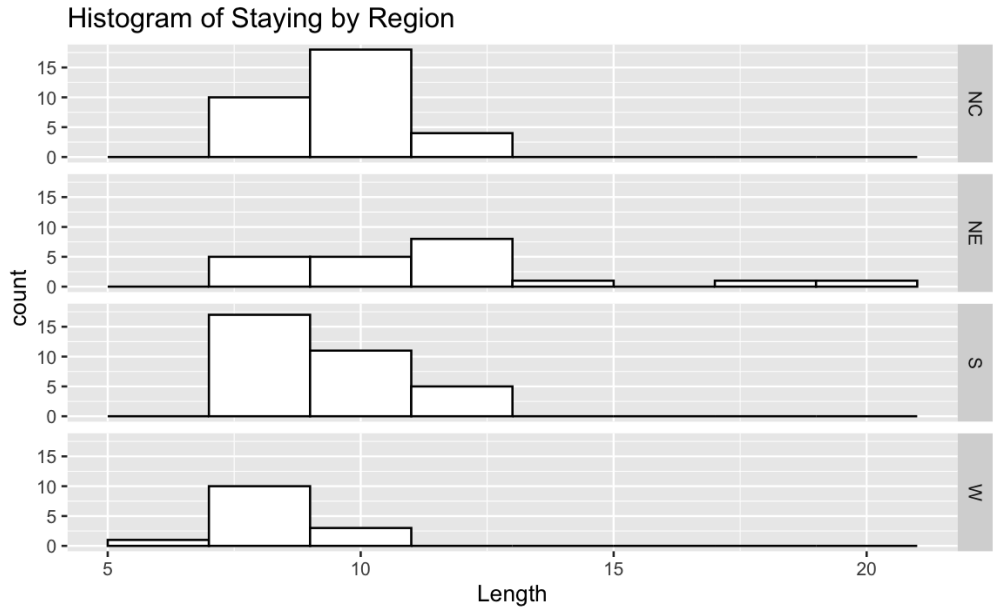

Figure 2 Histogram of Staying by Region

From the histogram, we can observe that the group has a similar distribution. In the NE group, it has two extreme values. The two extreme values are located at about length equal to 16 and 20. In the NC Region, most people stay for about 10 days. In the NE Region, most people stay for about 12 days. In the S region, most people stay for about 8 days. In the W region, most people stay for about 8 days. Most participants are staying for about 8-10 days.

II.2. Summary Data Values

1. The table of sample means follows:

Region	**NC**	**NE**	**W**	**S**	**Overall**
Means	9.6834	11.1943	9.2030	8.2186	9.6371
Std. dev	1.1929	2.9367	1.2695	1.0280	1.9261
Sample size	32	21	33	14	100

Table 1 Summary of statistics for Senic

In this table, we summarize all the data and group by region. We also include the overall data in the last column. We have four regions in this data, NE (North East),W(West),NC(North Central),and S(South). Through the table provided to us in this dataset, it originally has 100 samples. 32 of the participants from group NC, 21 of the participants from group NE, 33 of the participants from group W, and 14 of the participants from group S. The mean value of NE is relatively high compared with the other three groups. The group S has a relatively small mean value. For group NE, the participants' average time that they usually stay for is 11.19 days. For group NC, the average stay length is 9.68days. For group W, the average stay length is 9.20 days. For group S the average stay length is 8.21 days. In general, the overall stay length is 9.63 days.The standard deviation tells us a participant's average stay length can be spread from the mean. For group NC, the The standard deviation tells us a participants average stay length can be spared from the mean by 1.19 days. For group NE, The standard deviation tells us a participant's average stay length can be spread from the mean by 2.93 days. For group W, The

standard deviation tells us a participant's average stay length can be spread from the mean by 1.2695 days. For group S, The standard deviation tells us a participant's stay length can spread from the mean by 1.0280 days. In general, the overall standard deviation tells us a participant's stay length can spread from the mean by 1.9261 days.

III. Diagnostics

Diagnostic checks in a single-factor ANOVA help assess whether the assumptions underlying the ANOVA analysis are met. All of the test statistics and CIs are based on three assumptions:

i. Assumption of Normality: All observations (Y_{ij}) are randomly sampled.

ii. Assumption of Independence: All groups are independent.

iii. Assumption of the mean of zero and constant variance: $\varepsilon_{ij} \sim N(0, \sigma_{\varepsilon})\ for\ all\ i, j$

These assumptions are important because they ensure the validity and reliability of the statistical results. If the assumptions are seriously violated, it may affect the reliability of the ANOVA results. It leads to erroneous conclusions and misleading interpretations. That is, transformations of the data could be considered in some cases. Therefore, it's essential to check these assumptions before interpreting ANOVA results. To approach a good ANOVA analysis, we will first focus on checking these assumptions via diagnostic checks, such as examining residuals, normality plots, and tests for homogeneity of variances (constant variances), to help assess whether these assumptions are satisfied.

III.1. remove outliers via semi-studentized/ studentized residuals

Outliers can sometimes contain valuable information and occupy an amount of data. It is reasonable to not capture all types of outliers to prevent the damaged impact of the removal. Since studentized residuals are a powerful tool for outlier detection and removal, using studentized residuals is a valuable approach.

$$e_{ij}^{*} = \frac{e_{ij}}{\sqrt{MSE}} \sim t_{1-\frac{\alpha}{2g}, n_T - a}$$

$$where\ g\ is\ the\ number\ of\ the\ adjustment\ we\ need\ to\ detect.$$

To identify outliers, we estimate the residuals and derive the Mean Squared Error (MSE) and studentized residuals. Studentized residuals are standardized by dividing them by their estimated standard deviation. These standardized residuals serve to quantify the magnitude of deviation of each observation from the expected values. In simpler terms, they allow us to assess the relative size of each residual, providing insights into outliers and the distribution's divergence from normality.

Considering the presence of multiple residuals simultaneously, adjustments to the value of the significant level are necessary. Falling to account for the total number of observations can lead to issues of asymmetry, subsequently causing an increase in the significance level. Such a scenario is undesirable as it makes it difficult to correctly reject the null hypothesis, and may result in Type I error. To mitigate these problems, we adjust the significance level by dividing by twice the total number of observations. The value of g is n_T because we are detecting all residuals simultaneously.

In the context of outlier detection, a standardized residual value significantly exceeding 2 suggests an outlier for the corresponding observation. By applying the standardized residuals

method, we pinpoint anomalies within samples #67 and #90, both belonging to the North East region. These observations are identified as outliers. This observation aligns with our earlier identification of two outliers within the North East region using boxplots. Upon closer examination, we note that the residuals for #67 and #90, namely 8.365714286 and 6.745714286, are considerably greater than those of the other residuals.

It is well-established that the distribution of standardized residuals closely approximates a standard normal distribution with a mean of 0 and a standard deviation of 1. With regard to assessing normality, the removal of outliers from the dataset is a common practice. Consequently, we obtained a new dataset comprising 98 observations after the exclusion of outliers.

III.2. Assessing normality

III.2.1 Q-Q Plot (Quantile-Quantile Plot):

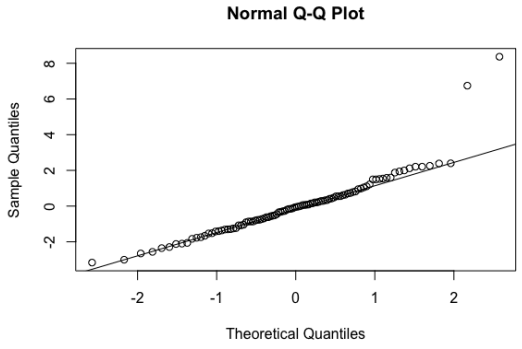

Figure 3 QQ Plot without remove outliers

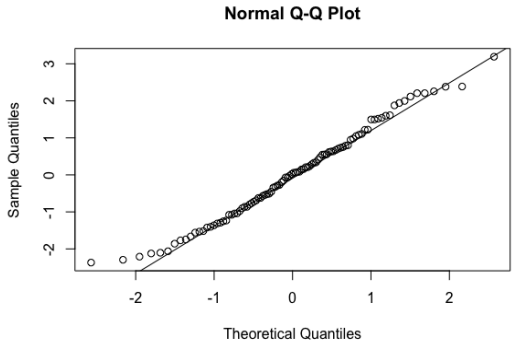

Figure 4 QQ Plot with remove outlier

Figure 3 graphs the original data set. It shows there are two extreme values. We consider them as outliers which did not closely follow the quantile line.

A Q-Q plot is a graphical tool used to assess whether a dataset follows a normal distribution. The idea behind the Q-Q plot is to compare the quantile of our dataset against the quantiles of a theoretical normal distribution. In our Q-Q plot, we find that the new data points closely follow a straight line when we remove outliers. It indicates that our new data is an approximately normal distribution. That is, we are graphically satisfied with the assumption of normality.

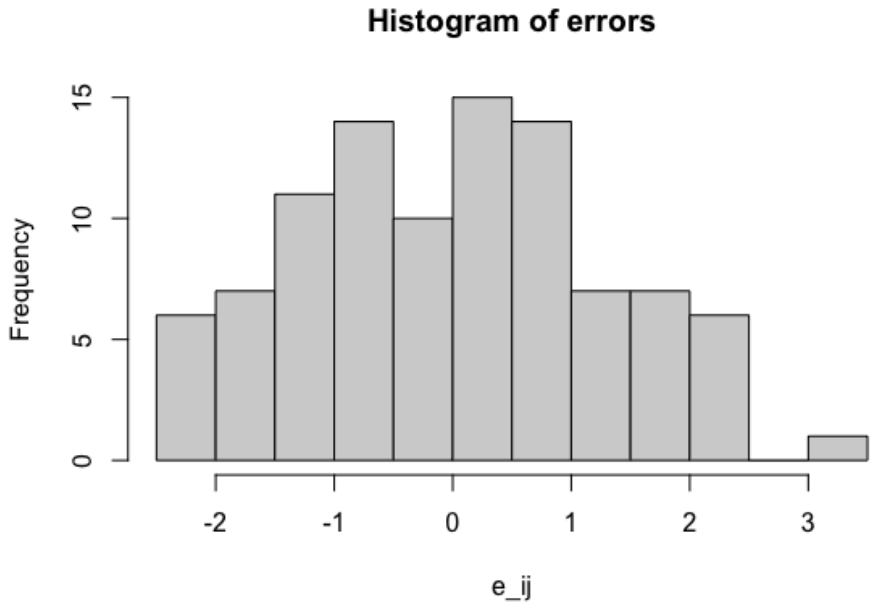

Figure 5 Histogram of Errors

III.2.2. Shapiro-Wilk test (SW test)

Since the Q-Q plot is subjective, the Shapiro-Wilk test is a formal statistical test used to assess whether a dataset significantly deviates from a normal distribution. The concept of the test is to compute a test statistic that assesses how well the sample data matches a normal distribution. Then, the test statistic is compared to critical values to determine whether the data is significantly non-normal. The hypotheses are:

H_0: *The data is normally distributed.*

H_A: *The data is not normally distributed.*

The test statistic, the p-value for Shapiro-Wilk test, is 0.327816[1]. The p-value associated with the test is greater than a chosen significance level which we assume has a given alpha of 0.05. Thus, we fail to reject the null hypothesis and conclude the data is normally distributed.

In practice, it's recommended to use both graphical and mathematical methods together for a comprehensive assessment of normality. Since both methods suggest that our data is approximately normally distributed, we can proceed with confidence.

[1] The exact calculation of the p-value for the Shapiro-Wilk test is shown in the appendix.

III.3. Assessing constant variance(homogeneity)

III.3.1. Plotting errors vs. groups

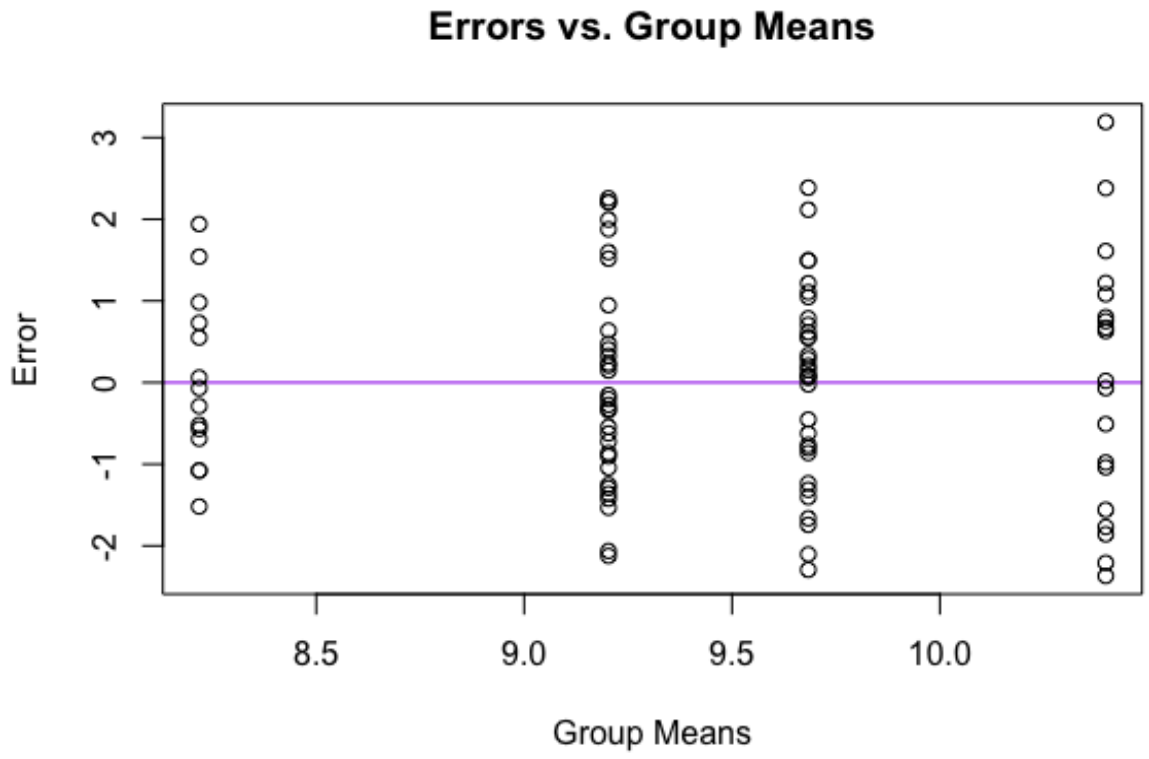

Figure 6 Error vs. Group Means

Plotting errors(residuals) by groups is a graphical technique that visually examines whether the spread of residuals remains consistent across different levels of a categorical variable. In more detail, create a scatter where the x-axis represents the groups and the y-axis represents the residuals. Look for patterns in our scatter plot. The spread of residuals appears relatively consistent across groups. It suggests constant variance. Notice that the variances of North East region are slightly larger than the other regions.

III.3.2.Brown-Forsythe test(BF test)

The Brown-Forsythe test is a statistical method that quantifies differences in group variances by analyzing the absolute deviations of observations from their respective group medians. The hypotheses are:

$$H_0: \textit{The groups' variances are equal to each other.}$$

$$H_A: \textit{At least one group variance is not equal.}$$

Evaluating our data, the p-value for the Brown-Forsythe test associated with the ANOVA on absolute deviations is 0.2668[2]. As a result, the p-value is higher than the chosen significance level of 0.05. We fail to reject the null hypothesis and conclude the variances are approximately equal.

In summary, a combination of visual inspection and formal testing enhances our assessment of the assumption of constant variance. They both told that the assumption of constant variance is satisfied.

IV. Analysis.

IV.1. Fitted model

We build a fitted model after removing the outliers from the diagonistisc part. In this analysis, we will use the group mean model to determine if there are significant differences between different geographical groups and how they are related to the length of patients staying in hospital.

[2] The exact calculation of the p-value for the BF test is shown in appendix.

$$Y_{ij} = \mu_{ij} + \varepsilon_{ij} \quad where\ \varepsilon_{ij} \sim N(0, \sigma_{\varepsilon})$$

Our primary goal here is to check the difference between length of stay in the following four groups, and they are NC, NE, W, and S.Here μi represents each the average of each patient's length of stay in hospital, meaning that i = NC, NE, W, S for the four different geological regions. In this model, ε_{ij} represents the individual error for any given jth value in the ith group.

Because outliers significantly impact how we analyze our data, we want to find out those unusual values and remove them beforehand.

According to the formula, we have $\widehat{\gamma_i} = \bar{Y}_{i.} - \bar{\bar{Y}},\ for\ i = NC, NE, S, W$

	NC	**NE**	**S**	**W**
ɣ$_i$	0.3074	1.023	-0.173	-1.1574

Table 2 ɣ with region

We found the γ_{NC} is 0.3074, γ_{NE} is 1.023, γ_S is -0.173, and γ_w is -1.1574. We add them all together and get the result γ_i is -1.776e-15. It's already close to 0. Thus, we satisfy the constraint that the sum of the ɣ is zero.

IV.2. Hypothesis test

Our goal in this hypothesis test is to determine if the mean weight loss varies significantly between region groups NC, NE, W, S. As such, our null hypothesis, H0, in this case will be:

H_0: $\mu_{NC} = \mu_{NE} = \mu_W = \mu_S$

Here, H_0 is claiming that the average length of stay in hospital in different regional groups are all the same, and there is no difference. This sets a basis for what our calculations will be trying to support or reject. Alternatively, if we reject this null, that would mean that at least one of our means for weight loss across regional groups NC, NE, W or S are different. Thus, our alternative hypothesis in this case will be:

H_A: At least one of the mean length of stay across regional groups NC, NE, W and S are different.

We will now calculate the test statistic for the hypothesis test, and we will do this using an ANOVA table. For single factor ANOVA, the test statistic is equal to the MSA/MSE, and the higher the value, the more unlikely the null hypothesis is true. In fact, if the value of the test statistic is one, that means there is absolutely no difference between the group means. We will now calculate the value of the test statistic and note the corresponding p-value. Here is the output of the ANOVA table:

	Df	**Sum Sq**	**Mean Sq**	**F value**	**Pr(>F)**
Region	3	42.09581	14.031936	8.542816	4.486619e-05
Residuals	94	154.39897	1.642542	NA	NA

Table 3 Anova Table

As seen here, the value of our test-statistics F_s in this case is 8.5428. The p-value that corresponds with this test statistic in this case is .0000448. This means that if in fact the null was true and there was no significant difference in mean weight loss across geographical groups NC, NE, W and S, and we would observe our data or more extreme data .004% of the time.

IV.3. Simultaneous Confidence Intervals

$$\overline{Y_{i\cdot}} - \overline{Y_{i'\cdot}} \pm M\sqrt{MSE(\frac{1}{n_i} + \frac{1}{n_i'})}$$

IV.3.1. Multipliers (M)

Since we are interested in making multiple confidence intervals for an ANOVA model, we consider using the simultaneous confidence interval rather than using the single confidence interval. Detecting the model with 6 pairwise at the same time, we use the simultaneous one to adjust to prevent the increase in the Type I error. As for the property of the pairwise, we consider Tukey, Scheffe, and Bonferroni multipliers to adjust. Here are the results of our multiplier calculation:

	Tukey	**Scheffe**	**Bonferroni**
Multiplier[3]	2.615589	2.846813	2.695239

Table 4 Multiplier with Tukey, Scheffe, and Bonferroni

The smallest value for a multiplier, in this case, is Tukey. We will be using the Tukey multiplier with a value of 2.614607 to construct our 6 simultaneous pairwise confidence intervals.

IV.3.2. Confidence intervals

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

[3] The exact calculation of the multipliers are shown in appendix.

are as follows: We can differentiate the provided simultaneous confidence intervals based on whether they cross zero or not:

Means being compared	**95% C.I.**
$\mu_{NC} - \mu_{NE}$	-1.6863799 0.2553602
$\mu_{NC} - \mu_{W}$	-0.3512657 1.3120801
$\mu_{W} - \mu_{S}$	-0.08473326 2.05365101
$\mu_{NE} - \mu_{S}$	0.999664 3.361088
$\mu_{NE} - \mu_{W}$	0.2305428 2.1612913
$\mu_{NC} - \mu_{S}$	0.3907092 2.5390230

Table 5 Confidence Interval with groups

Confidence intervals for μNC-μNE, μNC-μW, and μW-μS cross zero, suggesting that there is no statistically significant difference in the average lengths of stay between these region pairs at the 95% confidence level. In contrast, the intervals for μNE-μS, μNE-μW, and μNC-μS are entirely above zero, indicating that the North East region has a significantly longer average length of stay than the South and West regions, and the North Central region has a longer average stay than the South region at the 95% confidence level.

The first three intervals contain zero, meaning that there's a possibility that there might be no difference in the average length of stay between these two regions. For the rest three, as these

intervals do not include zero, there's a significant difference between these regions, with the South likely having a longer average stay compared to the West.

IV.4. Power calculation

Power is a significant statistical term that calculates the probability of rejecting the null hypothesis given that the null hypothesis is false. 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:

$$\phi = \frac{1}{\sigma_\varepsilon} \cdot \sqrt{\frac{\sum_i n_i(\mu_i - \mu_\cdot)^2}{a}}$$

The φ parameter measures how different Fs under the alternative hypothesis is to Fs under the null hypothesis. Notably, the σε 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 σ_ε with 1.642542 in our calculation. Another important value that is necessary to calculate power is an alpha value. In our case we chose alpha = 0.05. Thus, with values φ and alpha, and using the appropriate degrees of freedom as below:

d.f.{num} = 3, and d.f.{denom} = 94

As such, the resulting power of this test is about 0.9924617, meaning that we have a probability of 0.9924617 to correctly reject the null given that the null 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 lengths of patient's stay in the hospital, regardless of region, are the same,

when, in reality, at least one of the average population weight loss values per diet is not equal is 0.9924617.

V. Interpretation.

From our ANOVA table, Table 3, with approximate p-value we are 99% confident that at least one region's average staying days is unequal to the other. But, we just reject the hypothesis that the averages of staying days among four regions are equal. We do not have sufficient evidence of which region is unequal to the other. As a result, to look forward at different regions' factor effects, we conduct further comparisons, such as the alternative hypothesis and pairwise confidence interval.

From the alternative hypothesis, γ captures the factor effects of levels of regions. Table 2, γ_{NE}=1.023, shows that patients in region NE have a positive effect on the days staying in the hospitals. That is, the patients recover longer if you stay in the region NE. On the contrasty, γ_{W} =-1.1574 shows that patients in region W tend to recover faster with short staying lengths. Through the value γ, we clarify the unequal means in regions NE and S. There are two kinds of factor effect in regions NE and S. Patients in region NE have poor recovery rates and thus tend to have an average of 1.023 days increasing the overall staying length in the hospitals. However, patients in Region S recover fast and tend to have an average of 1.1574 days decline in the overall stay length in the hospitals.

What's more, the results of the factor effects correspond to the results from the confidence intervals. Since we are interested in the case there is a significant difference between the two regions, we observe that the mean of NE is particularly prominent. From Table 5, the

95% confidence intervals suggest that the true value of μNE is greater than the true values of μS between 0.999664 to 3.361088 and μW between 0.2305428 to 2.1612913. It's another way to demonstrate region NE has a positive factor effect. Besides, with a negative factor effect in region S, the true value of μS is less than the true value of μNC between 0.3907092 and 2.5390230.

Last but not least, the largest difference is the average of region NE compared to region S, 3.361088. It makes sense that we know region NE has the largest positive effect and region S has the largest negative effect. In summary, our findings, with 95% confidence, illuminate a captivating narrative. Patients in the NE region, with their commendable extended stays, emerge as distinct main characters in this story. With these revelations in hand, we suggest healthcare administrators and decision-makers can leverage to optimize patient care and resource allocation.

Specifically, we suggest allocating more beds, medical equipment, and staffing in regional NE hospitals. Since extended stays might require additional monitoring devices and lower bed occupancy rates, the strategies prevent overcrowding and ensure that adequate resources are available for the expected patient load. Besides, recognize that their prolonged stays might indicate more complex medical conditions or a need for an extended recovery period. Allocate additional medical resources, nursing hours, and support services to cater to their specific needs. Ultimately, we look forward to improving the overall efficiency and effectiveness of healthcare services.

IV. Conclusion

There are four main takeaways:

1. We reject the null hypothesis, which means at least one group has a different population average number. In our null hypothesis test, we did not indicate there is a difference between each group.
2. By analyzing the confidence intervals, there is evidence at a 95% confidence level that patients in the NE region have significantly longer stays than those in the S and W regions.
3. Given a reasonably high value of the power (1-β), 0.9924617, it indicates that our test is effective in correctly detecting true effects when they exist. A high power value is desirable because it reduces the risk of Type II errors.
4. Meanwhile, we are good at balancing between Type I and Type II errors in hypothesis testing through trade-offs between α and β. In our case, with α=0.05, we're willing to accept a 5% chance of making a Type I error.

Appendix:

```
the.data = read.csv("C:/Users/vaim1/UC Davis/SAT 106A/senic.csv")

library(ggplot2)
boxplot(Length ~ Region, data = the.data, main = "Boxplot of Staying length (in days) by
Region")
ggplot(the.data, aes(x = Length)) + geom_histogram(binwidth = 2,color = "black",fill = "white")
+ facet_grid(Region~.) +ggtitle("Histogram of Staying by Region")

mean(the.data$Length)
sd(the.data$Length)
length(the.data$Length)
group.means = by(the.data$Length,the.data$Region,mean)
group.sds = by(the.data$Length,the.data$Region,sd)
group.nis = by(the.data$Length,the.data$Region,length)
the.summary = rbind(group.means,group.sds,group.nis)
the.summary = round(the.summary,digits = 4)
colnames(the.summary) = names(group.means)
rownames(the.summary) = c("Means","Std. Dev","Sample Size")
the.summary
the.means = aggregate(Length ~ Region, data = the.data, mean)
the.means

#Outlier QQ Plot
the.model_old= lm(Length ~ Region, data = the.data)
the.data$ei = the.model_old$residuals
qqnorm(the.model_old$residuals)
qqline(the.model_old$residuals)

# outliers
anova.table = anova(lm(Length ~ Region, data = the.data))
the.model= lm(Length ~ Region, data = the.data)
the.data$ei = the.model$residuals

nt =nrow(the.data)
a = length(unique(the.data$Region))
SSE = sum(the.data$ei^2)
MSE = SSE/(nt-a)
eij.star = the.model$residuals/sqrt(MSE)
```

```
alpha =0.05
t.cutoff = qt(1-alpha/(2*nt), nt-a)
C0.eij = which(abs(eij.star) > t.cutoff)
C0.eij

outliers = C0.eij
new.data = the.data[-outliers,]
new.data #remove the outliers
anova_result <- aov(Length ~ Region, data = new.data)
anova_summary <- capture.output(anova_result)
table_output <- capture.output(anova_summary)
writeLines(table_output, "project1.txt")

#No Outlier QQ Plot
#Histogram of Errors
#Error vs.Group Means
the.model= lm(Length ~ Region, data = new.data)
new.data$ei = the.model$residuals
qqnorm(the.model$residuals)
qqline(the.model$residuals)
plot(the.model$fitted.values, the.model$residuals, main = "Errors vs. Group Means",xlab =
"Group Means",ylab = "Error", abline(h = 0,col = "purple"))
hist(new.data$ei,main = "Histogram of errors",xlab = "e_ij")

#SW Test
sp.val = shapiro.test(the.model$residuals)$p.val
sp.val
message("The p-value for Shapiro-Wilk test is:", sp.val)

#BF Test
library(car)
the.BFtest = leveneTest(ei~Region, data=new.data, center=median)
p.val = the.BFtest[[3]][1]
p.val
message("The p-value for BF test is:", p.val)

#Calculate gamma
library(knitr)
#No Outlier Data
group.means = by(new.data$Length,new.data$Region,mean)
```

```
group.nis = by(new.data$Length,new.data$Region,length)
the.model = lm(Length ~ Region, data = new.data)

gammai = group.means - mean(group.means)
gammai
sum(gammai)
the.gammai = rbind(gammai)
the.gammai = round(the.gammai, digits = 4)
kable(the.gammai)
the.gammai

#tukey multipliers
Tuk1 = qtukey(1-alpha, a, nt-a)/sqrt(2)
Tuk1
# scheffe
sch = sqrt((a-1)*qf(1-alpha, a-1, nt-a))
sch
#bonferroni: I assume g=6 because I make 6 CIs total
g =6
Bon = qt(1-alpha/(2*g), nt-a)
Bon
#simultaneous CI
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)
  }
}
group.means = by(new.data$Length,new.data$Region,mean)
group.nis = by(new.data$Length,new.data$Region,length)
the.model = lm(Length ~ Region, data = new.data)
```

```
anova.table = anova(the.model)
anova.table
MSE = anova.table[2,3]
nt = sum(group.nis)
a = length(group.means)
alpha = 0.05

#Miu_1 =NC
#Miu_2 = NE
#Miu_3 = W
#Miu_4 = S
#compare mu_1, mu_2
c12 = c(1,-1,0,0)
give.me.CI(group.means,group.nis,c12,MSE,Tuk1)
#compare mu_2, mu_4
c24 = c(0,1,0,-1)
give.me.CI(group.means,group.nis,c24,MSE,Tuk1)
#compare mu_2, mu_3
c23 = c(0,1,-1,0)
give.me.CI(group.means,group.nis,c23,MSE,Tuk1)
#compare mu_1, mu_3
c13 = c(1,0,-1,0)
give.me.CI(group.means,group.nis,c13,MSE,Tuk1)
#compare mu_3, mu_4
c34 = c(0,0,1,-1)
give.me.CI(group.means,group.nis,c34,MSE,Tuk1)
#compare mu_1, mu_4
c14 = c(1,0,0,-1)
give.me.CI(group.means,group.nis,c14,MSE,Tuk1)
```