STA 3000 Project Report - Chick Weights

Team Members:

Kobe Philbert-Willis - Role: Abstract, Introduction, Conclusion, and References

Mark Pedraza - Role: Analysis and Results, Code, Visuals, Appendices

Fernando Rey - Role: Methodology

Carlo Capan - Role: Data Description

5-20-25

Abstract

The primary goal of this report is to figure out which diet helps the baby chicks grow the fastest. Our team used the ChickWeight dataset in R. We analyzed it using linear regression modelling, t-tests, and confidence intervals. Two models were run, one with the diets and one without, as a control group. The model with the diet implemented showed a more apparent connection. Growth curves and boxplots were utilized in the appendices to make our findings easier to understand. The best diet for growing the biggest chicks was Diet 3 after 21 days, although not every statistical comparison was significant. Our team used the confidence intervals to understand the weight range and Diet 3 had the highest overall. In conclusion, if you want the heaviest chicks possible within 21 days, your best bet is Diet 3. Although it should be noted that there were violations in the model, this conclusion should be taken with a level of caution.

Introduction

In this report, our task is to discover which of the four diets helps baby chicks grow the most. We are doing this since we're a consulting group interested in helping our client farms raise the biggest chicks possible to increase their profits as much as possible. The dataset we utilized is called ChickWeight, which was built into R. We believe this dataset represents baby chicks' weight and simulates their weights on the four diets. Our objective was to test the before and after for a 21-day period to determine which diet was the most effective at allowing the chicks to gain weight. Our methodology includes running two linear regression models to compare the results with time and diet. Additionally, we performed t-tests to compare the diets and the confidence intervals to find the average weight of the chicks by day 21. We hypothesize that one of these diets will be highly effective and significantly outperform the others.

Data Description

The data used in this project is the ChickWeight dataset, a built-in dataset in R, which tracks the weight gain of baby chicks over time under different feeding conditions. This dataset is a reliable and clean resource for evaluating the impact of diet on chick growth and is commonly used in statistical modeling exercises.
The dataset consists of 578 observations with four key variables:

- Weight: The dependent variable, measured in grams, represents a chick's weight.
- Time: A numerical variable that indicates the day of observation, ranging from Day 0 to Day 21.
- Chick: A categorical identifier for each chick, numbered 1 through 50.
- Diet: A categorical variable indicating the type of feed given to each chick. There are four diets in total, labeled as 1, 2, 3, and 4.

Chicks were randomly assigned to four diets and monitored for 21 days. The data, in a long format with each row representing a chick's weight measurement at a specific time, allows for modeling weight changes and evaluating the impact of different diets on weight gain. The complete dataset required no cleaning or imputation, and variables were used directly in the analysis, with interaction terms (Time × Diet) created for further modeling. The sample size and variability in weight and time provided sufficient power for regression modeling and hypothesis testing.

Analysis and Results

Using R, we created a linear regression model, confidence intervals, and a two-sample t-test and visuals for model diagnostics and general data analytics. With these, we can finally check the effects all the different types of diets had on the chicks over time and whether any of these statistical tests provided us with useful insight.

Linear Regression

Starting with linear regression, we made weight the response variable, with diets and time as predictor variables. With just this, we created a pretty good model. In the summary output (*Appendix 1.1*), no variable is insignificant, which means they are all linearly correlated with weight. Our Adjusted R-squared is 74.35%, which means that 74.35% of the variance in weight is accounted for by time and diet. The ANOVA F-test with a p-value of 2.2e-16 indicates an overall linear correlation within the model.

All in all, it's relatively good. However, if we make an interaction term (Time*Diet), we make this model slightly better. We do this since there is likely a relationship that exists between the two predictors that is not accounted for when we use each by itself. Lastly, the residual standard error is 35.99 with 570 degrees of freedom.

We get a slightly better model when creating this new model with the interaction term. In the summary output (*Appendix 1.2*), the Adjusted R-Squared goes up to 77.02%, and the ANOVA F-test still tells us that an overall linear correlation exists. However, since there is a new term created for each Diet, we now have Time + Diet1-3 instead of Time + Diet1-3 + Time*Diet1-3. Because of this, the p-value for all solem Diet predictors jumps significantly. However, the model is still good, as shown by the Adjusted R-Squared and the F-test. We can also see improvement in a slightly reduced residual standard error with 34.07 on 570 degrees of freedom. This means the model is better fitted to our data so we will use this model for future inferences and visualizations.

Visuals in Linear Regression

In the visuals, we can make some inferences. For instance, since we have a categorical variable (Diet), and interaction terms, we can find multiple regression lines in our model (*Appendix 4.1, 4.2*). With these graphs, we can see that the slopes for the regression lines are all different. This indicates that each diet contributes to the chick's weight differently over time. Diet 3 increases weight more rapidly over time than the other diets.

Model Diagnostics

However, we have to check for model diagnostics. In the Q-Q plot (*Appendix 4.5),* we unfortunately see a violation of normality. We also find a violation of normality from the Shapiro-Wilk test (*Appendix 4.4*). In the Residuals vs Fitted Values plot (*Appendix 4.3*), we also see a violation of equal variance. Luckily, though, there was no violation of independence or linearity. There were no notable outliers in any plot. These violations mean the inferences we make from this model can be invalid. So, although the model is fitted well, we must also take it with a grain of salt.

Confidence Intervals

With confidence intervals, we will examine the 95% confidence interval for our linear regression model coefficients and the mean weight on day 21 for each diet. The regression confidence intervals will help us determine if our model is valid, and the mean weight ones will help show which diet correlates to a higher mean weight by the end of the trial.

95% CI for Linear Regression Model Coefficients

First, let's review the 95% confidence intervals created for the coefficients for our linear regression (*Appendix 2.2*). At a glance, they do appear fairly tight around each coefficient. This means that we are confident that each coefficient fits closely. But notably, for each solem diet coefficient (no interaction), each confidence interval includes 0. Unfortunately, this tells us the same story as their respectively high p-values. That is, each coefficient has no linear correlation to the response if all other coefficients remain constant. Although, as mentioned previously, this is not an issue. Since we have interaction terms, we are still accounting for diet and time. Also, the summary output of the model shows promising statistics such as the adjusted R-squared and F-test, as mentioned previously. So overall, these confidence intervals show tight fits, but also signs of no correlation between specific predictors.

95% CI for Mean Weight on Day 21

For the next set of confidence intervals, we took a 95% confidence interval for the mean weight of each diet on day 21 (*Appendix 2.1*). Below is the interpretation of the confidence interval for each diet.

- Diet 1 - We are 95% confident that the mean weight of chicks on diet 1, day 21, is between 146.47 gm and 209.03 gm

- Diet 2 - We are 95% confident that the mean weight of chicks on diet 2, day 21, is between 158.8 gm and 270.6 gm

- Diet 3 - We are 95% confident that the mean weight of chicks on diet 3, day 21, is between 219.06 gm and 321.53 gm

- Diet 4 - We are 95% confident that the mean weight of chicks on diet 4, day 21, is between 205.24 gm and 271.88 gm

From these intervals, we notice that diet 3 has the highest minimum and maximum range, followed by diet 4, then 2, and finally 1, which has the lowest interval. Since diet 3 has the highest interval, it likely has the biggest effect on chick weight compared to the other diets by the time the trial finishes. However, there is an interesting problem with this conclusion that we will test when we check our Two-Sample T-tests.

Two-Sample T-Tests

Diet 3 has the highest confidence interval, and diet 4 comes in second; we need to make sure they are significantly different to differentiate them. To do this, we took a two-sample T-test to see if the mean difference in weight between diet 3 and diet 4 on day 21 was significant (*Appendix 3.1*). Surprisingly, it was not. Below is the interpretation of the generated 95% confidence interval from the two-sample t-test.

- We are 95% confident that the actual difference in mean weight of the chicks under diet 3 and diet 4 is between -25.51 and 89.0.

The confidence interval here includes 0, which suggests the difference in mean is not statistically significant at the 0.05 level. We can also see this with the high p-value of 0.2557. So if we were to conduct hypothesis testing (H0: Difference in means = 0, HA: Difference in means ≠ 0, alpha = 0.05), we would fail to reject the null hypothesis. However, we must keep in mind that failing to reject the null hypothesis doesn't necessarily mean the null hypothesis is true. From all this information, we can say that although there is a difference in the means on day 21 (diet three is higher by 31.74 gms), it is not statistically significant.

General Graph Analytics

In section *5* of the *Appendix*, there are visuals of how the data generally looks. In *5.1) Average Chick Weight Over Time by Diet*, we notice that diet 4 increases the weight the most up until day 11, where diet 3 surpasses it in weight and ends as the diet that increases chicks' weight the most on average. This is similar to the linear regression plots we saw earlier, where diet 3 showed the most potential. There is no other instance of lines going past each other. So at the end, the chicks with the highest average weight to the lowest by diet, diet 3, 4, then 1.

Another interesting visualization is the *5.3) Box-plots of Weight by Diet on Day 21*. This plot indicates that more than 50% of the chicks on diet 3 had a higher weight than chicks on other diets. This is shown through the entirety of quartiles 3 and 4, as well as the median, being above all other quartiles/medians from diets 1, 2, and 4.

The last visualization in the *Appendix* is *5.4) Mean Weight with 95% Confidence Interval Over Time by Diet*. Here, we see the range of a 95% confidence interval instead of a line for each diet over time. Although hard to see, we notice that the confidence interval for diet 3 is higher than all other diets. But not in its entirety, there's an overlap between diets 3 and 4. Generally, there's plenty of overlap between the confidence intervals, but in order of highest interval to lowest, it follows diet 3, 4, 2, then finally, 1.

Methodology

To evaluate the impact of different diets on chick weight gain over time, we used a combination of statistical modeling, hypothesis testing, and graphical analysis using the ChickWeight dataset in R. Our objective was to determine which diet leads to the most significant weight gain by Day 21 and to assess the statistical significance of any observed differences.

We began with multiple linear regression, selecting weight as the response variable and using time and diet as predictor variables. To capture potential interaction effects, specifically how the impact of time on weight might differ by diet, we extended the model to include an interaction term (Time × Diet). The model with interaction was preferred for inference and visualization based on improved fit indicators, such as a higher adjusted R-squared and lower residual standard error.

To assess the reliability and precision of model estimates, we calculated 95% confidence intervals for the regression coefficients and the mean weight at Day 21 for each diet. Confidence intervals were chosen to provide a range of values for population parameters and to compare average weights across diets visually.

To determine whether observed differences in mean weights between diets were statistically meaningful, we performed two-sample t-tests. Specifically, we compared Diets 3 and 4 with the highest mean weights on Day 21. The two-sample t-test was appropriate here because it tests the difference in means between two independent groups under the assumption of approximate normality.

We conducted diagnostic checks on our regression models to verify key assumptions such as normality of residuals, equal variance, and independence. These included visual tools like Q-Q plots, residual vs. fitted plots, and formal tests like the Shapiro-Wilk test.

Additionally, we used data visualization techniques to complement our statistical findings. Growth curves, box plots, and mean weight plots with confidence bands over time helped identify trends.

These combined methods allowed us to gather our conclusions from multiple angles. Ultimately, this approach provided a detailed and statistically grounded evaluation of diet effectiveness in promoting chick growth.

Results

From all the visualizations and statistics calculated, we see a clear pattern in weight distribution among the chicks for all the diets. By the end of the trial on day 21, diet 3 will increase chick weight the most, followed by diet 4, then diet 2, and finally diet 1. However, by day 21, the mean difference between diets 3 and 4 is not statistically significant at alpha = 0.05. This means they are relatively close. So even though we can visually see a slightly better outcome for weight through the graphs, we also must take their difference with a grain of salt.

Conclusion

In conclusion, there's a consistent trend in weight gain among the chicks regarding diet. This was discovered through extensive analysis using regression models, t-tests, and confidence intervals. Our team found that Diet 3 led to the highest average weights after 21 days. A trend noticed was Diet 4 performing extremely well until being overtaken at Day 11. A possible solution for an even more effective diet could be seeing if these diets could be intermixed to produce even better results. Diet 1 shouldn't be utilized under any circumstances as it had the lowest weight gain out of the diets.

Furthermore, the difference between Diet 3 and 4 wasn't statistically significant when considering t-tests. Our team gained the greatest insight from the interaction regression model, which shows how diets affect weight gain overtime. One major issue should be noted: there were violations in this model, so the results should be taken cautiously. Despite this, the pattern was clear, and Diet 3 is the best diet to put the baby chicks on for maximum weight gain.

References

"Chickweight: Weight versus Age of Chicks on Different Diets." RDocumentation, www.rdocumentation.org/packages/datasets/versions/3.6.2/topics/ChickWeight. Accessed 20 May 2025.

Appendix

1.1) Summary Output of Linear Regression Model without Interaction Terms

```
Residuals:
     Min       1Q   Median       3Q      Max
-136.851  -17.151   -2.595   15.033  141.816

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  10.9244     3.3607   3.251  0.00122 **
Time          8.7505     0.2218  39.451  < 2e-16 ***
Diet2        16.1661     4.0858   3.957 8.56e-05 ***
Diet3        36.4994     4.0858   8.933  < 2e-16 ***
Diet4        30.2335     4.1075   7.361 6.39e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 35.99 on 573 degrees of freedom
Multiple R-squared:  0.7453,    Adjusted R-squared:  0.7435
F-statistic: 419.2 on 4 and 573 DF,  p-value: < 2.2e-16
```

1.2) Summary Output of Linear Regression Model with Interaction Terms

```
Residuals:
     Min       1Q   Median       3Q      Max
-135.425  -13.757   -1.311   11.069  130.391

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  30.9310     4.2468   7.283 1.09e-12 ***
Time          6.8418     0.3408  20.076  < 2e-16 ***
Diet2        -2.2974     7.2672  -0.316  0.75202
Diet3       -12.6807     7.2672  -1.745  0.08154 .
Diet4        -0.1389     7.2865  -0.019  0.98480
Time:Diet2    1.7673     0.5717   3.092  0.00209 **
Time:Diet3    4.5811     0.5717   8.014 6.33e-15 ***
Time:Diet4    2.8726     0.5781   4.969 8.92e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 34.07 on 570 degrees of freedom
Multiple R-squared:  0.773,     Adjusted R-squared:  0.7702
F-statistic: 277.3 on 7 and 570 DF,  p-value: < 2.2e-16
```

2.1) Confidence Intervals for each diet on day 21

```
> # DIET 1 - Day 21
> # - We are 95% confident that the mean weight of chicks on Diet 1, Day 21 is between 146.4699 and 209.0301.
> subset_data <- subset(ChickWeight, Time == 21 & Diet == 1)
> t.test(subset_data$weight, conf.level = 0.95)$conf.int
[1] 146.4699 209.0301
attr(,"conf.level")
[1] 0.95
> # DIET 2 - Day 21
> # - We are 95% confident that the mean weight of chicks on Diet 2, Day 21 is between 158.8034 and 270.5966.
> subset_data <- subset(ChickWeight, Time == 21 & Diet == 2)
> t.test(subset_data$weight, conf.level = 0.95)$conf.int
[1] 158.8034 270.5966
attr(,"conf.level")
[1] 0.95
> # DIET 3 - Day 21
> # - We are 95% confident that the mean weight of chicks on Diet 3, Day 21 is between 219.0643 and 321.5357.
> subset_data <- subset(ChickWeight, Time == 21 & Diet == 3)
> t.test(subset_data$weight, conf.level = 0.95)$conf.int
[1] 219.0643 321.5357
attr(,"conf.level")
[1] 0.95
> # DIET 4 - Day 21
> # - We are 95% confident that the mean weight of chicks on Diet 4, Day 21 is between 205.2355 and 271.8756.
> subset_data <- subset(ChickWeight, Time == 21 & Diet == 4)
> t.test(subset_data$weight, conf.level = 0.95)$conf.int
[1] 205.2355 271.8756
attr(,"conf.level")
[1] 0.95
```

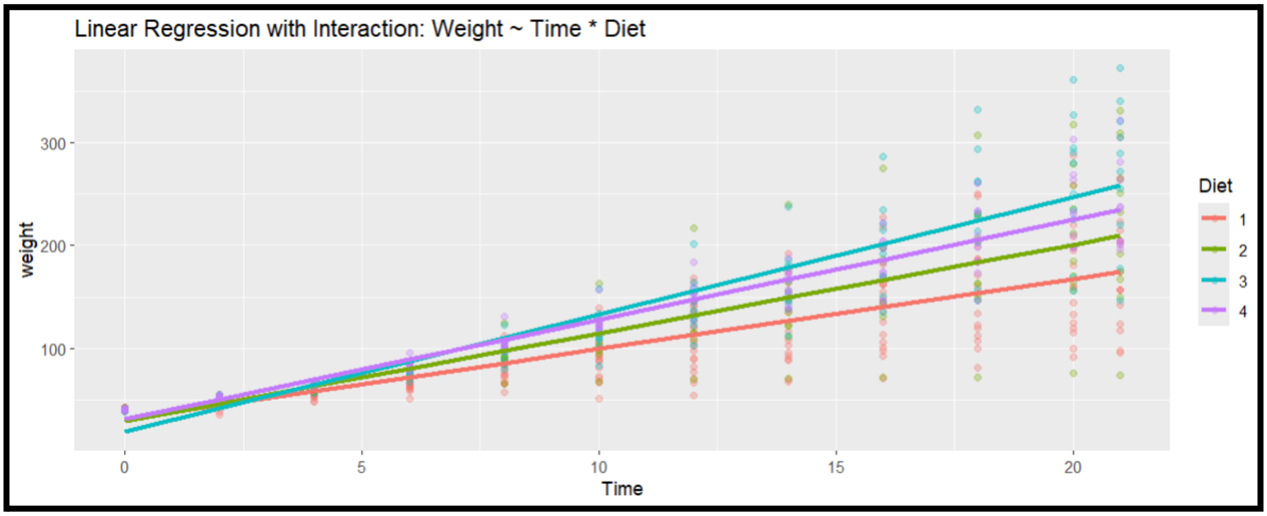
2.2) 98% Confidence Interval on the Linear Regression Model with Interaction Terms

```
> confint(mod.1, level = 0.98)
                    1 %      99 %
(Intercept)  21.0235225 40.838438
Time          6.0467568  7.636838
Diet2       -19.2511641 14.656395
Diet3       -29.6344344  4.273124
Diet4       -17.1374972 16.859776
Time:Diet2    0.4336904  3.100988
Time:Diet3    3.2474251  5.914722
Time:Diet4    1.5238768  4.221260
```

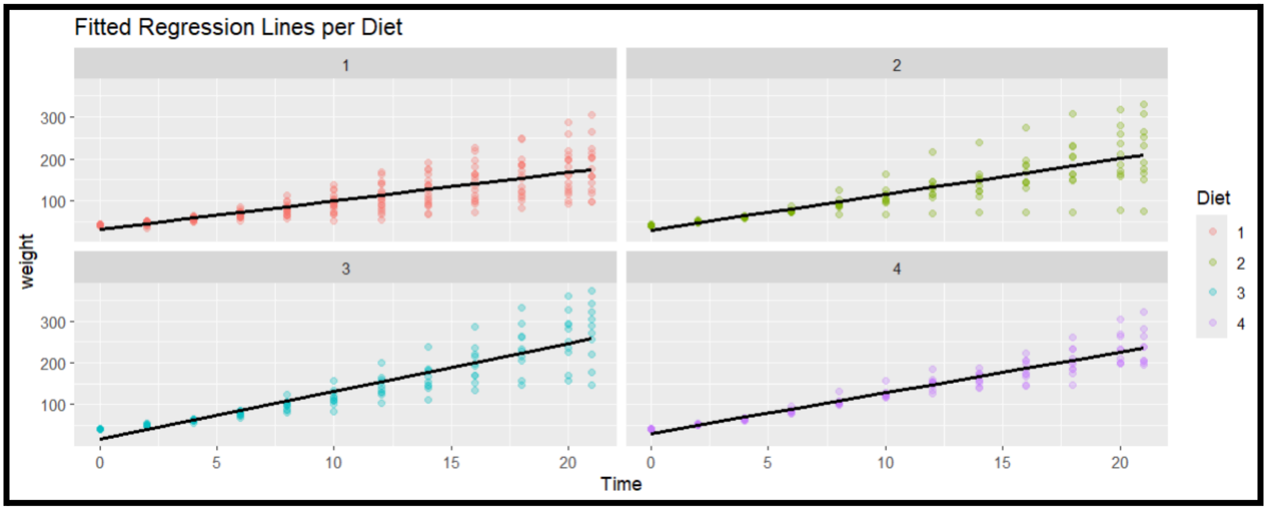
3.1) Two-Sample T-Test for Day 21 comparing Diets 3 and 4

```
data:  weight by Diet
t = 1.1816, df = 15.018, p-value = 0.2557
alternative hypothesis: true difference in means between group 3 and group 4 is not equal to 0
95 percent confidence interval:
 -25.51219  89.00108
sample estimates:
mean in group 3 mean in group 4
       270.3000        238.5556
```

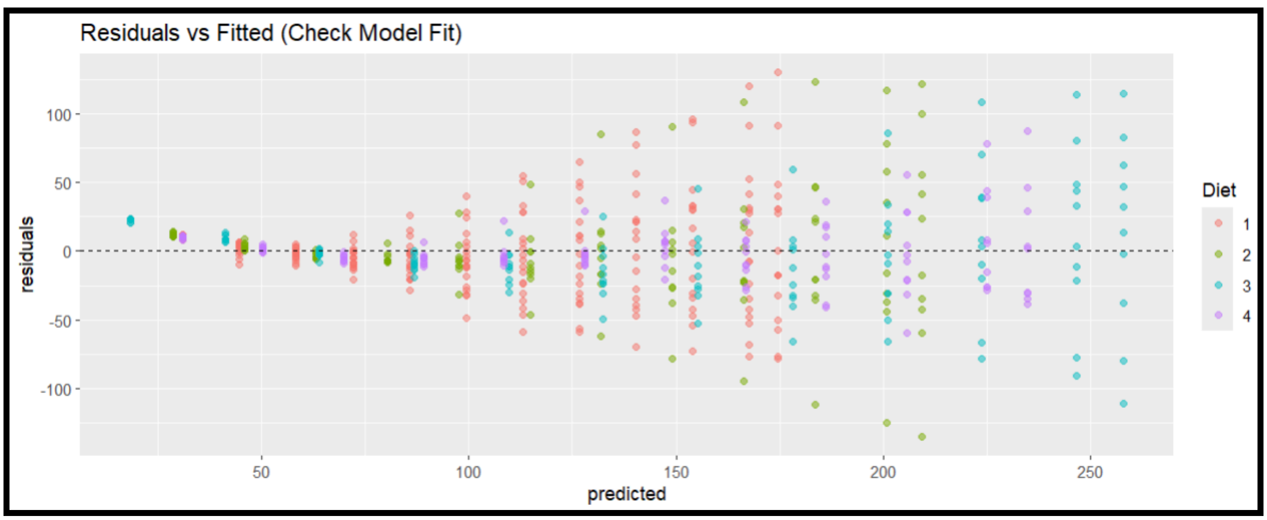
4.1) Linear Regression - Fitted Lines Separated by Diet



4.2) Linear Regression Fitted Lines Separated by Diet, Points Also Separated



4.3) Residual vs Fitted Values

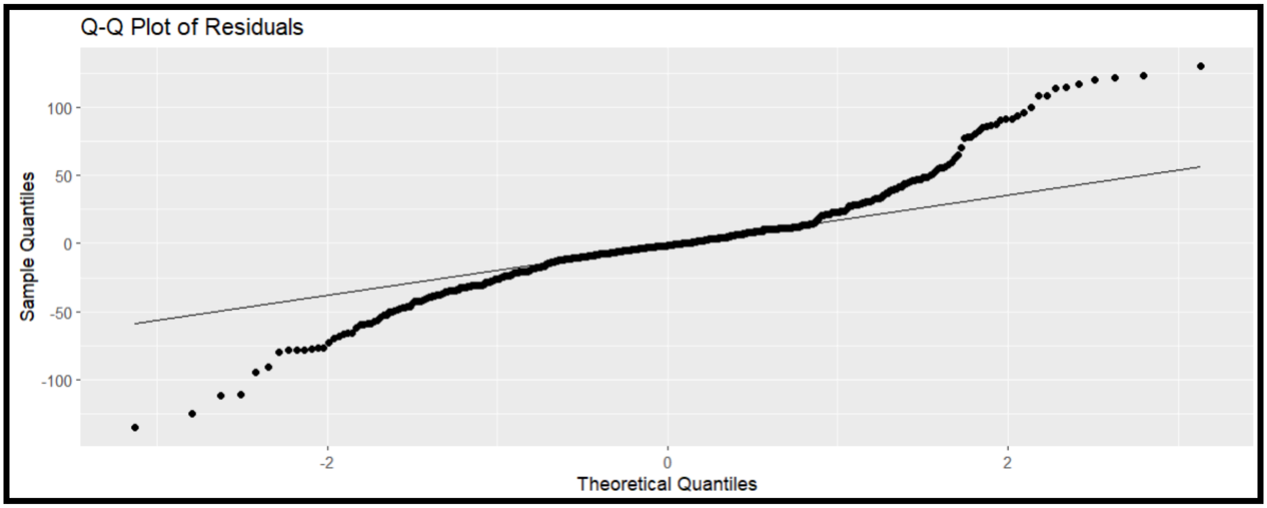


4.4) Shapiro-Wilk Normality Test

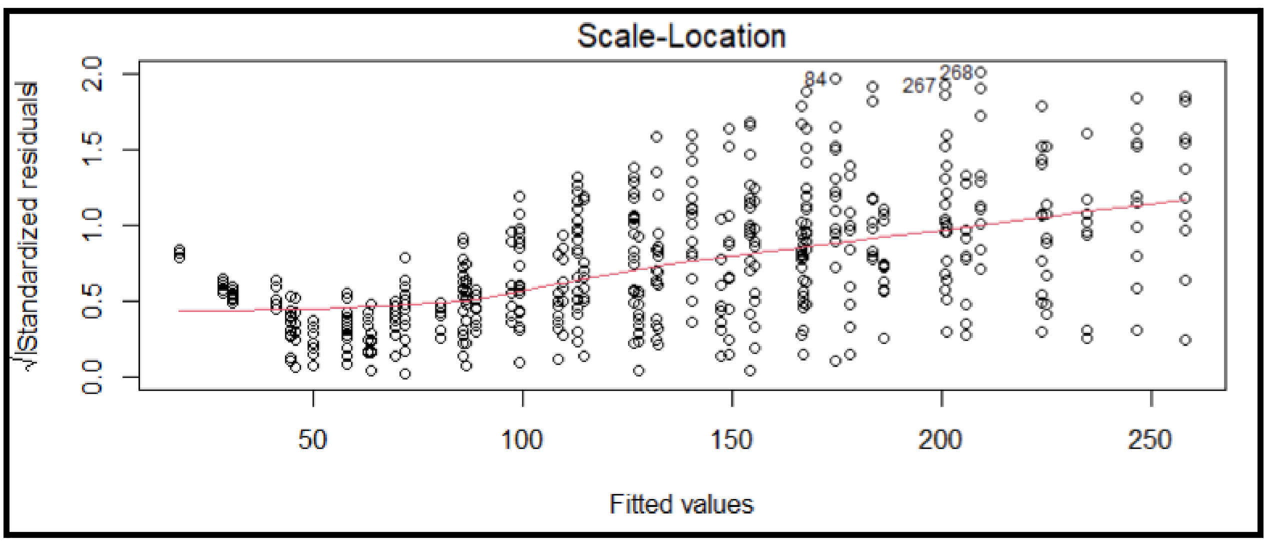
```
        Shapiro-Wilk normality test

data:  residuals(mod.1)
W = 0.92507, p-value = 2.252e-16
```

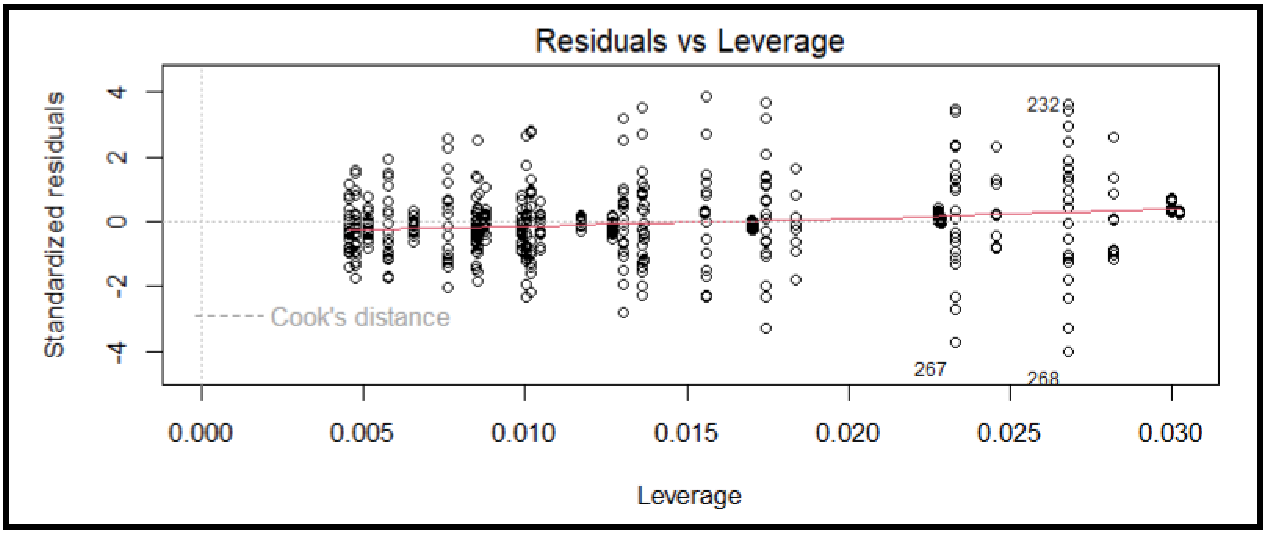
4.5) Q-Q Plot of Residuals



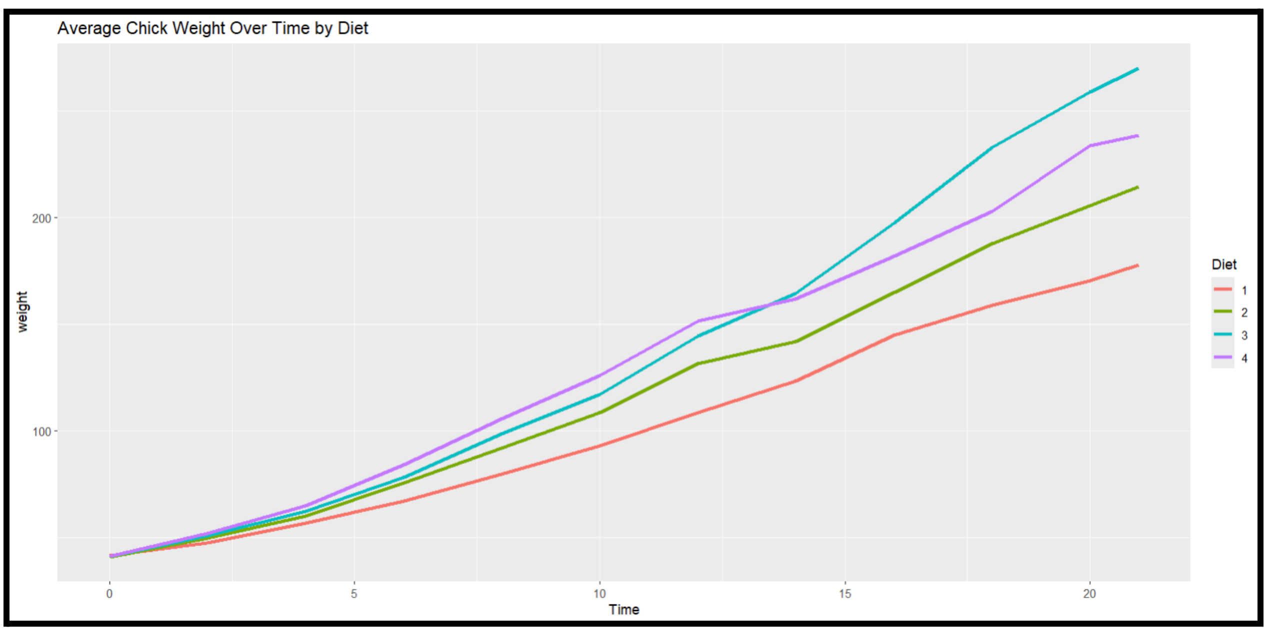
4.6) Scale-Location



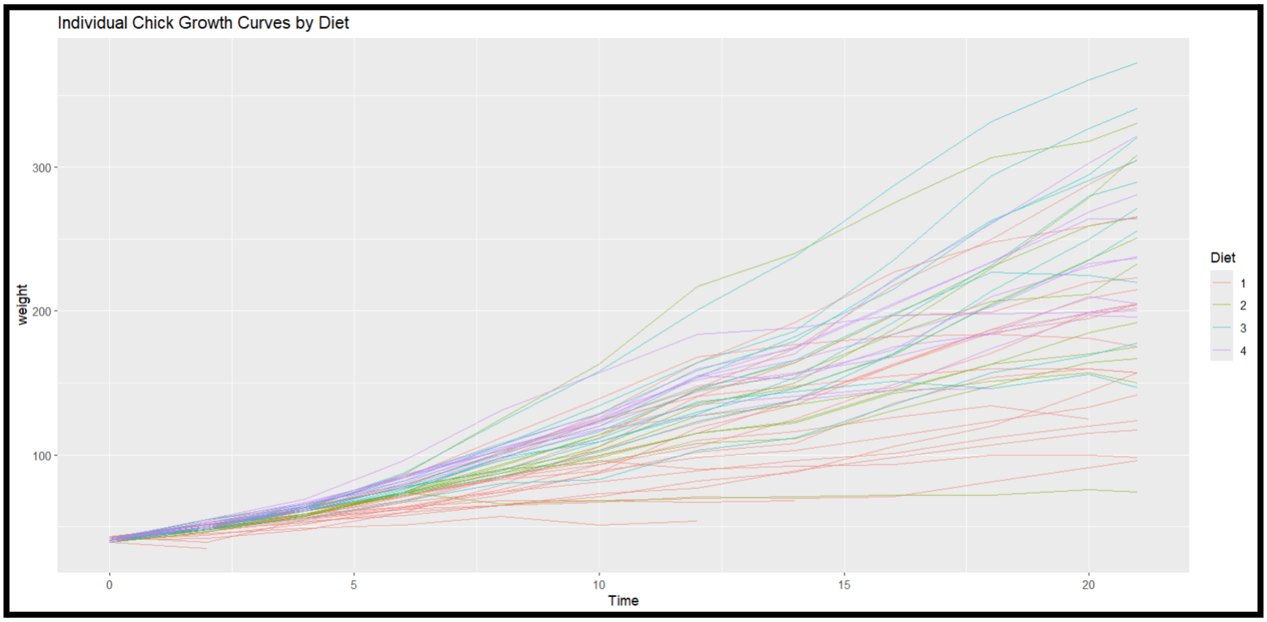
4.7) Residual vs Leverage



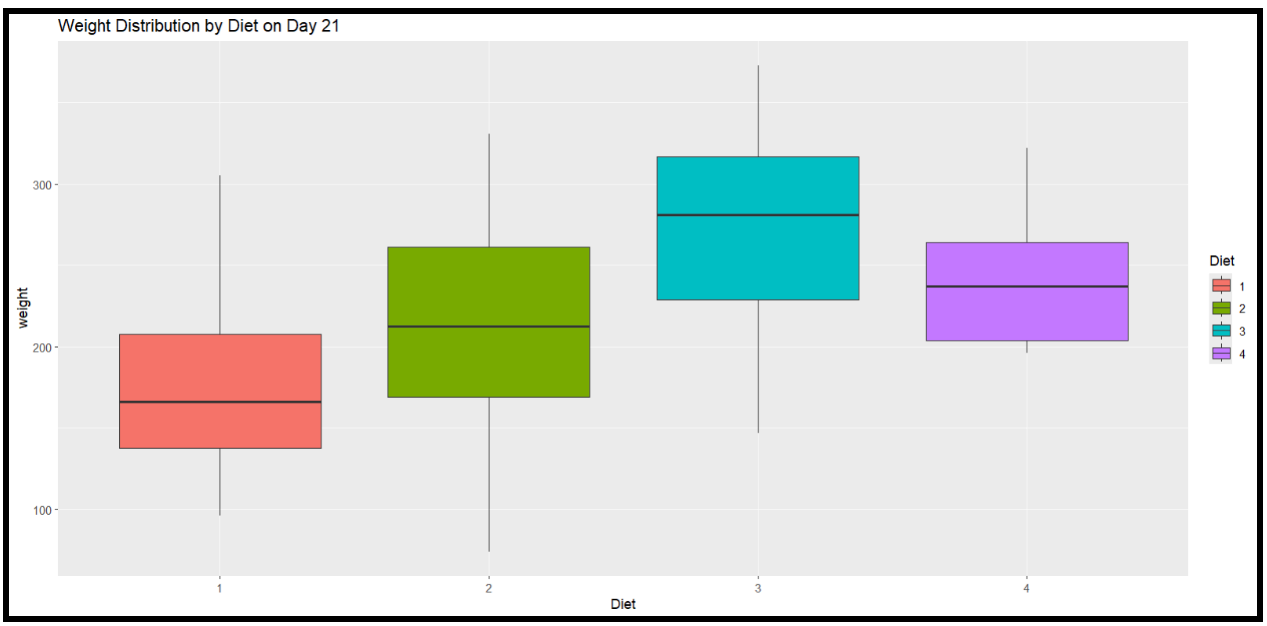
5.1) Average Chick Weight Over Time by Diet



5.2) Individual Chick Growth Curves by Diet



5.3) Box-plots of Weight by Diet on Day 21



5.4) Mean Weight with 95% Confidence Interval Over Time by Diet