5-1 Project Two

DAT-475

Ashley Littles

Southern New Hampshire University

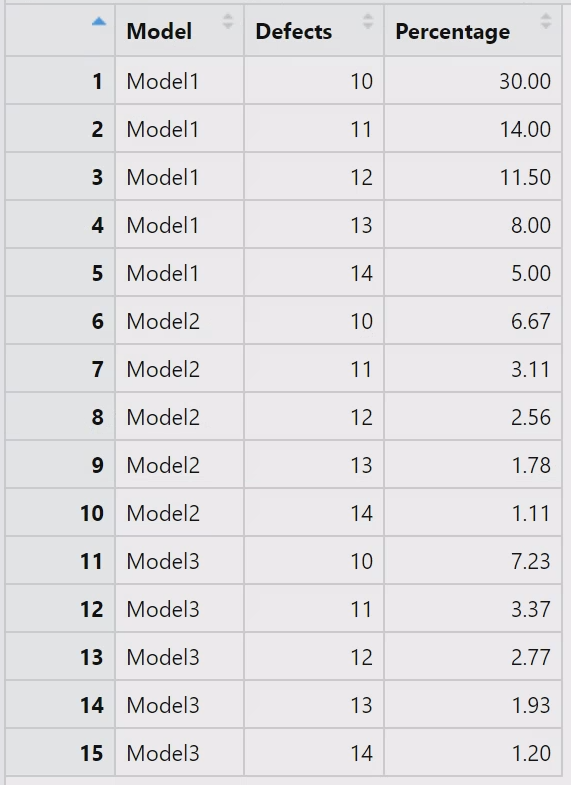
I am creating this report for my immediate supervisor to display the results from the one-way ANOVA process for my hypothesis testing. We are working alongside a manufacturing company in Tijuana, Mexico who is having issues regarding increased defects with the assembly's that are made of electronic boards and Thru-Holes. The hypothesis is that there is a statistical difference in the percentage of defects between the three different production lines that is significant. To create the valid hypothesis test for this case study, we first made sure to define the hypothesis and then determined the significance level of 0.05 as one of the parameters. The other parameters are the dependent and independent variables which are the number percentage of defects and the model type. The one-way ANOVA test will be used with these parameters. I chose these parameters because the information we need to gather relates to if the different models come out with different percentages of defects that are significant.

The null and alternative hypothesis will help lead us to resolving the identified problem. The null and alternative hypothesis statements are as follows:

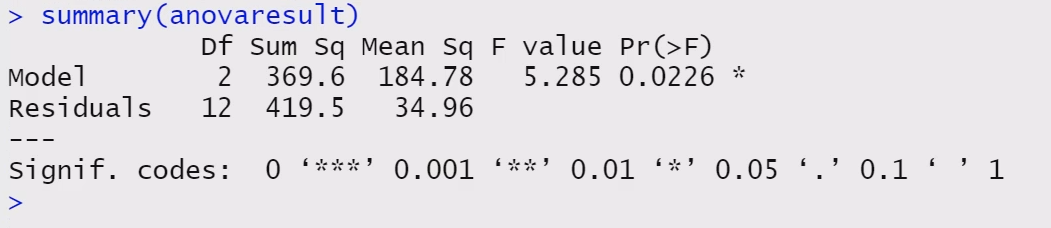
**Null hypothesis** (H0): There is not a significant difference of the defect percentages between the production lines Model 1=Model 2=Model 3.

**Alternative hypothesis** (H1): There is a significant difference in percentage of defects between the models.

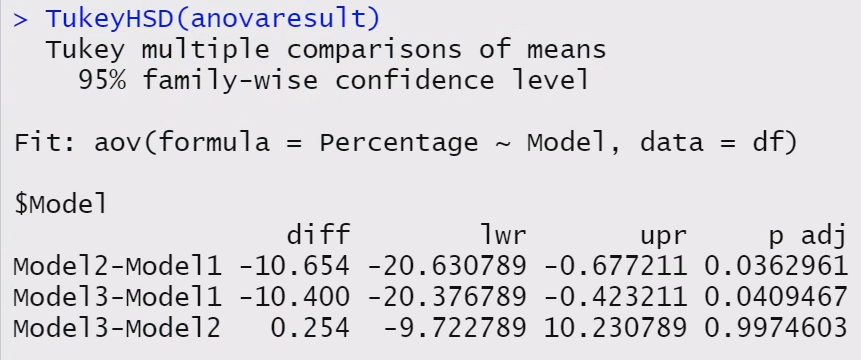
Here is the sample evidence data set displaying the three models and the number of defects along with the percentage of those defects:



To complete the hypothesis test with the sample evidence we used R. I first imported the data into the R script using the df <- read.csv function. Once the data set was imported, I was able to run a one-way ANOVA test on the sample evidence data set. To perform a one-way ANOVA test I used the aov() function.



Here are the results after running the one-way ANOVA test. We can see that the F-statistic is 5.285 and the p-value is 0.0226. With these results we can say that there is a significant difference between the defect percentages across the three models. I then used the TukeyHSD() function to provide a comparison of the means between the models.



This analysis shows a significant difference between the Model 1 compared to the other two models. We do not see any significant differences between Model 2 and Model 3. With the information gathered, we can come to the conclusion that the null hypothesis will have to be rejected. The alternative hypothesis would be accepted. So, there is a significant difference in percentage of defects between the models specifically Model 1 compared to Model 2 and Model 3. With this information proceeding forward, we should take a deeper look into the manufacturing process of Model 1 and what could be the cause of the defects.

**References**

Midway, S. (n.d.) *“Chapter 7 Understanding ANOVA in R.”* <https://bookdown.org/steve_midway/DAR/understanding-anova-in-r.html>

finnstats (August 28, 2021) “How to Perform Tukey HSD Test in R.” r-bloggers <https://www.r-bloggers.com/2021/08/how-to-perform-tukey-hsd-test-in-r/>