The dependence of rust removal on white vinegar concentration, salt amount and time, for rust found on pennies.

STATS345 Final Project

Thomas Wingate, Sylvia Li

November 29, 2020

Experiment

The goal of our experiment was to determine which of the three factors (Vinegar concentration, salt, time) had the most effect on rust removal. White vinegar has been observed to have rust removing properties especially when salt was added (Y.F et al., 2018). White vinegar is preferred to distilled vinegar as white vinegar contains 5-20% acetic acid whereas distilled vinegar is milder and contains only 5-8% of acetic acid. The increase in acidity helps dissolve the rust easier and the added salt increases the acidity and allows the solution to break down the rust easier. Our experiment was done with a *completely randomized design* as our subjects (pennies) were assigned to groups at complete random; We decided on this design as measuring rust on each penny and grouping the subjects based on that quality would be too difficult with the lack of instruments at hand.

- <u>*Completely randomized design*</u> *as ob*jects or subjects are assigned to groups completely at random.

Statistical Tests/Methods

We tested and observed the effects of factors by creating graphs, using Lenth's Method, and by doing multiple Levene tests to determine the dependence of rust removal on the concentration of salt, vinegar and time spent in solution.

<u>Materials and Methods</u>

Solution/trial set-up

To determine the effects of vinegar and salt on the rust found on pennies, we took 300 rusty pennies, and 6 clear cups. The cups were labelled with their designated concentration of white vinegar. The concentration of acetic acid was varied in each cup by diluting it with water, a 1:10 dilution, 1:4 dilution, 1:2 dilution, 3:4 dilution, and finally, one where there was no water added at all to be used as a control. For each trial, salt was added in the cups; The first trial had no salt, second had 0.2 grams of salt, 0.5 grams of salt for the next, fourth trial had 0.75 grams, and the last trial had 1 gram of salt. A total of 5 trials each with 6 cups was used. Once a trial is finished the 6 cups were rinsed and dried out for the next trials usage.

Removing the rust

To observe the effects on rust the 10 pennies were submerged in the solution in each cup. The effects were observed for a total of 45 minutes, checking after 1,2,5,10,15,30, and 45 minutes. At each time interval the pennies were taken out and each penny was given a score ore 0 or 1 depending on whether all the rust was removed.

Analysis

To determine the dependence, we looked at the overall effects for each of the trials by creating bar graphs and boxplots comparing the number of pennies cleaned to the amount of vinegar, salt and time spent. We first visualized the dependence of rust removal on salt by creating bar graphs and boxplots. As seen in the Figure 1 panel of bar graphs the amount of salt increased the number of pennies cleaned where the difference between 0g salt and 1g of salt saw an increase from 1 total penny cleaned to 20 total pennies cleaned. The increase in salt also allowed concentrations of vinegar starting from 1:10 to clean pennies whereas when no salt was present the only trial which cleaned any pennies was the one which had *only vinegar*. The positive correlation found between an increase acidity and rust removal was found in a similarly done experiment (Y.F et al., 2018) and found in experiments observing

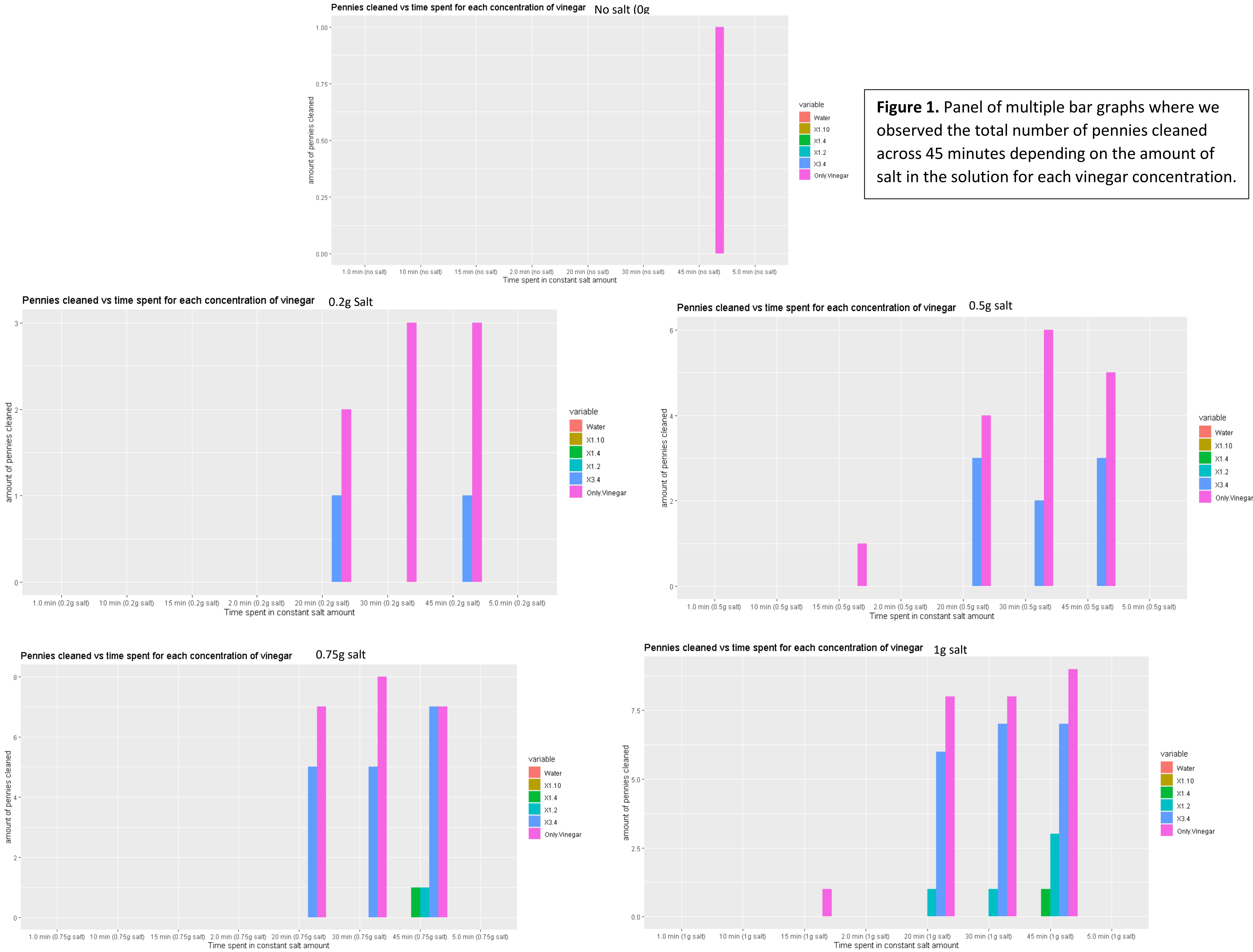

Figure 1. Panel of multiple bar graphs where we observed the total number of pennies cleaned across 45 minutes depending on the amount of salt in the solution for each vinegar concentration.

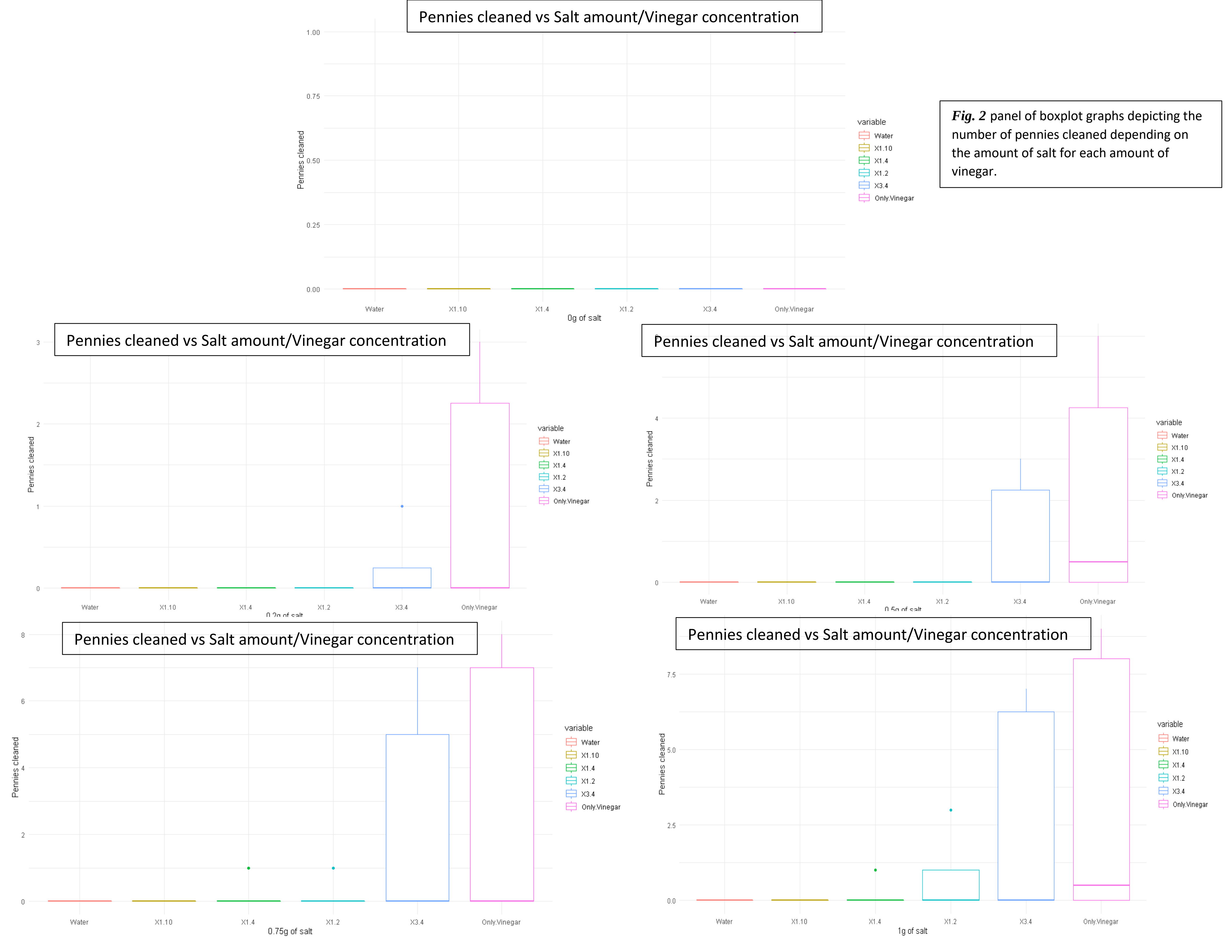

Fig. 2 panel of boxplot graphs depicting the number of pennies cleaned depending on the amount of salt for each amount of vinegar.

acidity in passion fruit (Jiang-xue et al., 2019). The boxplots (Fig. 2 panel) helped us observe the same positive correlation but on a more specific level where we only observed salt concentration alongside vinegar concentration, with the range quantiles of the bar correlating to the sum of pennies cleaned by time and where a larger box range shows that more pennies were cleaned faster. The positive correlation is seen and agrees with the bar graphs shown; The average amount of pennies cleaned rises with an increase in salt. Time was also a part of our trials and to observe the effects of time on the number of pennies cleaned we also used bar graph and box plots. Bar graphs and boxplots showed the strong positive correlation and dependence of time.

When bar graphs were used to observe the number of pennies cleaned depending on time (Figure 3), we saw that regardless of time spent the removal of rust is highly dependent on the concentration of salt the penny is submerged in. When left for only 1 through 10 minutes we saw no pennies being leaned of all the rust (they all outputted the same graph) and only after 15 minutes did we see any penny being cleaned and when pennies were cleaned the largest amount of pennies cleaned were always cleaned by the higher concentrations of vinegar, suggesting that regardless of time the rust removal is dependent on the concentration of vinegar; Both the bar graphs and the boxplot graphs showed that the concentration

```
Significant Factors Selected by Lenth's Method:
 Concentration3/4 ConcentrationOnly Vinegar
$pse
[1] 0.15

$sme
[1] 2.077935

> par(mfrow=c(1,1))
> LenthPlot(g3, las=3)
    alpha       PSE        ME       SME
0.0500000 0.1500000 0.7863398 2.0779351
```

Table 1. Significant factors test done by Lenths method shows that the 2 highest concentrations of salt had the most effect on the amount of pennies cleaned which agrees with our graphs (Figure 1, Figure 2)

of vinegar had a strong effect on the number of pennies cleaned regardless of time and salt amount is agreed upon when a test for significant effects was done This data is agreed upon by our significance tests done (Table 1, Table 2). Time and salt had an effect but at a much smaller amount compared to vinegar concentration.

```
Significant Factors Selected by Lenth's Method:
 Concentration3/4 ConcentrationOnly Vinegar Concentration3/4:Time15 ConcentrationOnly Vinegar:Time15 Concentration3/4:Time10 ConcentrationOnly Vinegar:T
ime10 Concentration3/4:Time5 ConcentrationOnly Vinegar:Time5 Concentration3/4:Time2 ConcentrationOnly Vinegar:Time2 Concentration3/4:Time1 Concentration
Only Vinegar:Time1
$pse
[1] 1.2

$sme
[1] 4.789815

> par(mfrow=c(1,1))
> LenthPlot(g5, las=3)
   alpha      PSE       ME      SME
0.050000 1.200000 2.548292 4.789815
```

Table 2. Significant factors test done by Lenths method shows that the 2 highest concentrations of salt had the most effect on the number of pennies cleaned *regardless of time spent* which agrees with our graphs (Figures 1,2); Although, still showed significance and dependence.

```
Significant Factors Selected by Lenth's Method:
 Concentration3/4 ConcentrationOnly Vinegar Concentration3/4:Salt0.5 ConcentrationOnly Vinegar:Salt0.5 Concentration3/4:Salt0.
2 ConcentrationOnly Vinegar:Salt0.2 Concentration3/4:Salt0 ConcentrationOnly Vinegar:Salt0
$pse
[1] 0.375

$sme
[1] 1.601341

> par(mfrow=c(1,1))
> LenthPlot(g4, las=3)
   alpha      PSE       ME      SME
0.050000 0.375000 0.839474 1.601341
```

Table 3. Significant factors test done by Lenths method shows that the 2 highest concentrations of salt had the most effect on the number of pennies cleaned regardless of salt amount which agrees with our graphs (Figures 1,2); Although still showed significance and dependence.

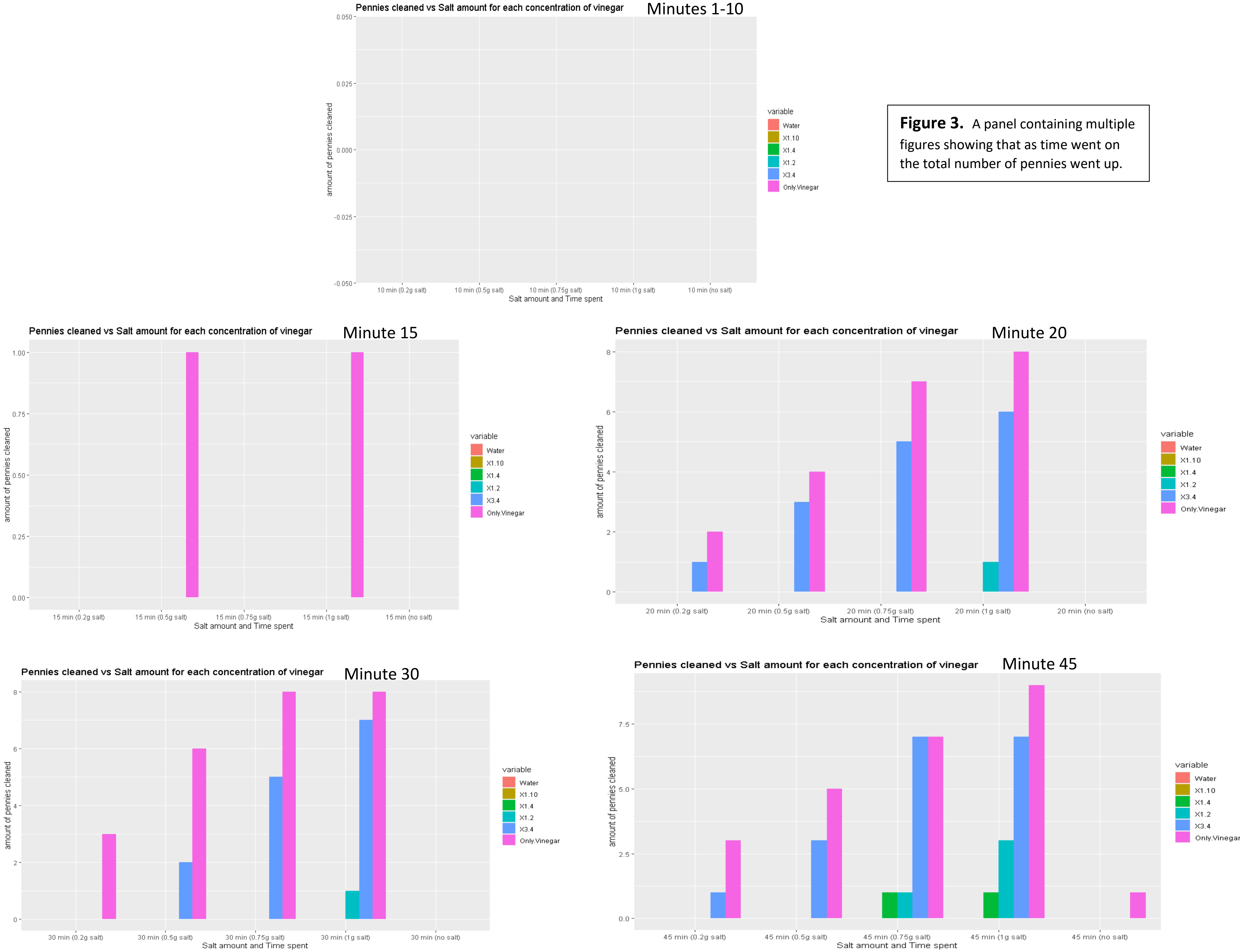

Figure 3. A panel containing multiple figures showing that as time went on the total number of pennies went up.

The stronger dependence of rust removal on vinegar concentration is also agreed upon when Levene's tests were done on the factors of salt, time, and vinegar concentration (Table 4).

The Levene tests agreed with our Lenth tests and observed differences in bar/box plot graphs.

```
> leveneTest(Cleaned ~ Concentration, data = interactionData)
Levene's Test for Homogeneity of Variance (center = median)
       Df F value    Pr(>F)    
group   5  10.227 7.054e-09 ***
      234                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> leveneTest(Cleaned ~ Salt)
Levene's Test for Homogeneity of Variance (center = median)
       Df F value   Pr(>F)   
group   4  3.7682 0.005435 **
      235                    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> leveneTest(Cleaned~Time)
Levene's Test for Homogeneity of Variance (center = median)
       Df F value    Pr(>F)    
group   7   6.505 5.162e-07 ***
      232                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 4. Multiple Levene's test were done to assess the equality of variances for a variable calculated for two or more groups which tests the null hypothesis that the population variances are equal. When concentration of vinegar was tested the P-value allows a very strong rejection of the null hypothesis. When salt was tested, we could only reject with a significance of 0.05. When time was tested, we rejected the null hypothesis *strongly* compared to salt, but less than concentration.

Concentration of vinegar rejected the null hypothesis at a higher significance compared to salt and time suggesting a higher effect. Time spent rejected the null hypothesis at a higher significance than salt suggesting a stronger effect than salt but less than concentration. Salt rejected the null hypothesis at the smallest significance between the 3 factors suggesting that salt amount did have an effect on the rust removal but not as much as time and vinegar concentration. Interaction plots also show the strong effect of vinegar concentration compared to time and salt amount (Figure 4.) Where we can see that there is a dependence on all 3 factors for time, salt, and vinegar amount. We can see that vinegar concentration had a much larger effect when comparing the interaction plots (Fig. 5) where when vinegar concentration was at ¾ and *only vinegar*, the number of pennies cleaned was always considerably higher regardless of salt or time. This agrees with our past observations from significant factor tests, and Levene tests. When comparing

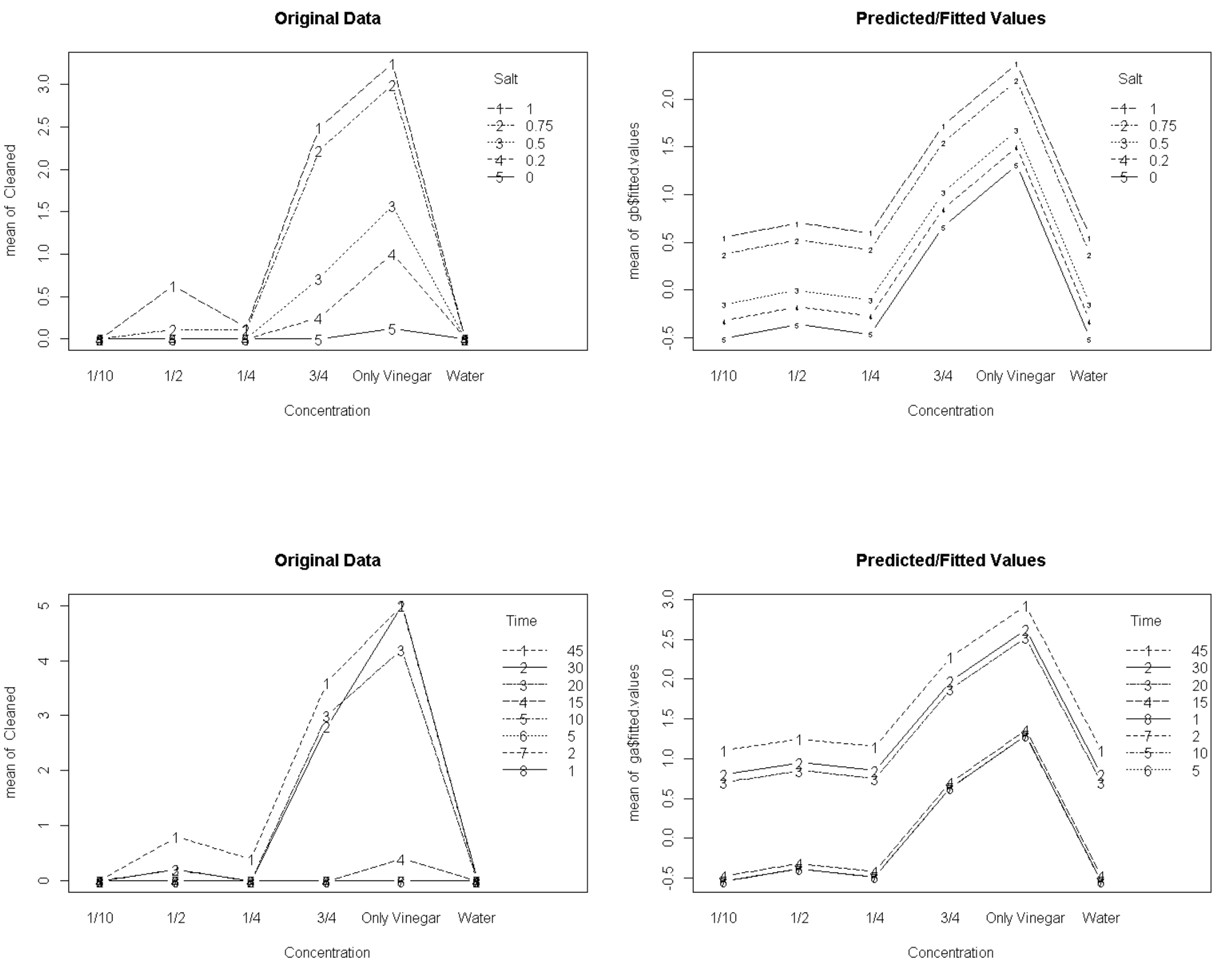

Figure 4. Interaction plots showing the effects of concentration, salt, and time on the average amount of pennies cleaned. We see in the first interaction plot that the number of pennies cleaned when salt was considered raised to a mean of 3. Which is less than when time was considered which had a bigger effect.

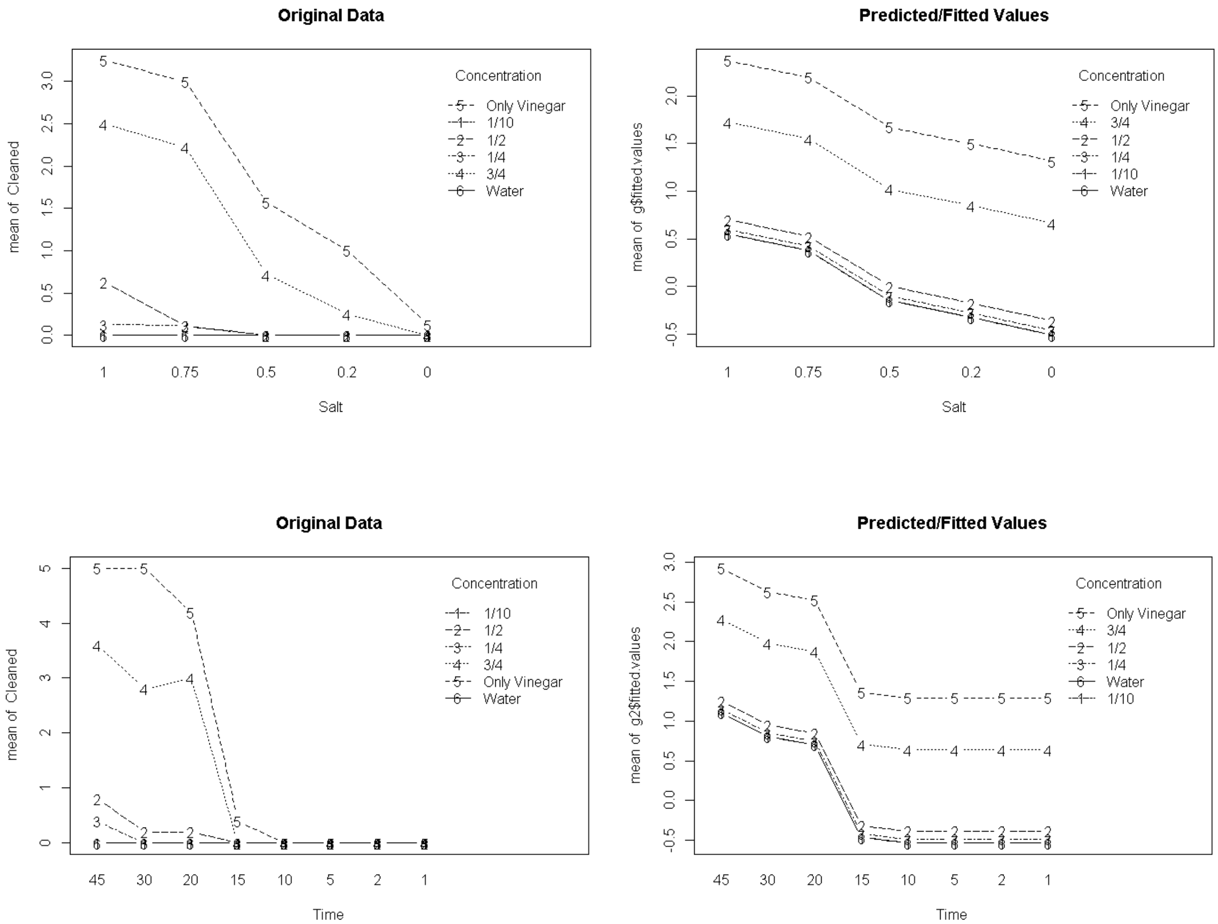

Figure 5. Interaction plots showing the effects of salt and time on a vinegar concentrations ability to clean pennies. When salt increased you can see a increase in pennies cleaned, similar to a linear rate. When the time effects on vinegar concentration was observed the interaction was much stronger than salt as seen in the large spike after 15 minutes, akin to a exponential-like increase.

interaction plots from figure 5 we can see that similar large effect of vinegar concentration, but the difference in effect of salt and time is more noticeable. Salt had an effect but at a much lesser amount compared to time. This observation agrees with past analysis of other data tables and graphs.

Conclusion

Our goal of determining the dependence of rust removal on salt, time, and vinegar concentration was determined by creating bar/boxplot graphs, doing Levene tests, and doing Lenth significant effect tests. Through these visualizations and tests it showed that vinegar concentration was the most important factor to the rust removal as agreed upon by our Lenth tests (Table 1-3) and interaction plots (Fig 4-5), alongside with our Levene tests (Table 4). The Lenth tests showed that the 2 highest concentrations of vinegar concentration had the most effect regardless of time and salt used; This effect was considerably higher compared to salt and time spent which is corroborated by the Levene tests (Table 4) which showed that the vinegar concentration rejected the hypothesis of equal population variances at a much greater significance than salt and time. Time rejected the hypothesis of equal population variances at a lower significance than vinegar concentration but higher than salt. Salt rejected the equal population variance to the lowest significance out of the three. The interaction plots showed a much greater interaction for vinegar concentration compared to salt and time which corroborates with the other tests and visualizations. Vinegar concentration was the biggest factor determining the effectiveness of removing rust from a penny, time would be the second biggest factor, and salt the least effecting factor.

References

Jiang-xue, L., Jia-li, J., Shi-ming, Z., Xiao-feng, D., & Xu-jie, D. (2019). Study on production of organic acid rust remover from passion fruitfermented by Aspergillus Niger. *Earth Environ. Sci.*

Y.F, O., P., S., Y., Z., X., S., S., P., S., K., K., G., & J., L. (2018). LASER CLEANING OF GREY CAST IRON AUTOMOTIVE BRAKE DISC: RUST REMOVAL AND IMPROVEMENT IN SURFACE INTEGRITY. *International Journal of Peening Science and Technology*, 155–180.

Appendix (CODE)

```
library(reshape2)

library(ggplot2)

library(tidyr)

library(BsMD)

library(car)


Alldata <- read.csv("data.csv", header = T, fill = T, stringsAsFactors = T)


########Section data by amount of salt

nosalt <- Alldata[1:8,]

salt2 <- Alldata[9:16,]

salt5 <- Alldata[17:24,]

salt75 <- Alldata[25:32,]

salt100 <- Alldata[33:40,]


#######Section data by time

minute1 <- Alldata[c(1,9,17,25,33),]

minute2 <- Alldata[c(2,10,18,26,34),]

minute5 <- Alldata[c(3,11,19,27,35),]

minute10 <- Alldata[c(4,12,20,28,36),]

minute15 <- Alldata[c(5,13,21,29,37),]

minute20 <- Alldata[c(6,14,22,30,38),]

minute30 <- Alldata[c(7,15,23,31,39),]

minute45 <- Alldata[c(8,16,24,32,40),]


########Simple plots to visualize data

####Evaluate time+salt and concentration of vinegar

#########BARPLOT VARYING SALT

df <- melt(minute30, id.vars = "Treatment")

ggplot(df, aes(x=Treatment, y = value, fill = variable)) + geom_bar(stat = 'identity', position = 'dodge') + labs(title = "Pennies
cleaned vs Salt amount for each concentration of vinegar", x = "Salt amount and Time spent", y = "amount of pennies cleaned")

##########BARPLOT VARYING TIME WITH CONSTANT SALT

df <- melt(salt100, id.vars = "Treatment")
```

```r
ggplot(df, aes(x=Treatment, y = value, fill = variable)) + geom_bar(stat = 'identity', position = 'dodge') + labs(title = "Pennies cleaned vs time spent for each concentration of vinegar", x = "Time spent in constant salt amount", y = "amount of pennies cleaned")


#####BOXPLOT

###########VARYing TIME WITH CONSTANT SALT

df <- melt(salt5)

plt <- ggplot(data = df, aes(x = variable, y = value, color = variable))

plt + geom_boxplot() + theme_minimal() + labs(x = "0.5g of salt", y = "Pennies cleaned")

###########VARYING SALT

df <- melt(minute45)

plt <- ggplot(data = df, aes(x = variable, y= value, color = variable))

plt + geom_boxplot() + theme_minimal() + labs(x = "45 minute", y = "Pennies cleaned")



############MODELS (EFFECT MODELS OR LINEAR MODELS OR INTERACTION PLOTS)##################

interactionData <- read.csv("intplotdata.csv", header = T, stringsAsFactors = T)

Treatment <- factor(interactionData$Treatment, levels = rev(unique(interactionData$Treatment)))

Concentration <- interactionData$Concentration

Cleaned <- interactionData$Cleaned

Time <- factor(interactionData$Minute, levels = rev(unique(interactionData$Minute)))

Salt <- factor(interactionData$Salt..g., levels = rev(unique(interactionData$Salt)))

#Interaction plot based on time AND salt

interaction.plot(Treatment, Concentration, Cleaned, col = 1:6, ylim = range(Cleaned),cex.axis = 0.7, cex = 0.5)

points(Treatment, Cleaned, pch = as.numeric(Cleaned))

#interaction based on time only

interaction.plot(Time, Concentration, Cleaned, col = 1:6, ylim = range(Cleaned))

points(Time, Cleaned, pch = as.numeric(Cleaned))

#Interaction based on salt only

interaction.plot(Salt, Concentration, Cleaned,  col = 1:6, ylim = range(Cleaned))

points(Salt, Cleaned, pch = as.numeric(Cleaned))


#####EFFECT MODELS: TREATMENT DIFFERENCE MODELS########

#####SALT INTERACTION

g <- lm(Cleaned ~ Salt + Concentration, contrasts = list(Salt = contr.sum, Concentration = contr.sum), data = interactionData)

summary(g)
```

```
par(mfrow = c(1,2))

with (interactionData, interaction.plot(Salt, Concentration, Cleaned, type = "b", main = "Original Data"))

with(interactionData, interaction.plot(Salt, Concentration, g$fitted.values, type = "b", main = "Predicted/Fitted Values"))

######TIME INTERACTION WITH CONCEN MODELS

g2 <- lm(Cleaned ~ Time + Concentration, contrasts = list(Time = contr.sum, Concentration = contr.sum), data =
interactionData)

summary(g2)

par(mfrow = c(1,2))

with (interactionData, interaction.plot(Time, Concentration, Cleaned, type = "b", main = "Original Data"))

with(interactionData, interaction.plot(Time, Concentration, g2$fitted.values, type = "b", main = "Predicted/Fitted Values"))

#####CONCENTRATION WITH TIME INTERACTION MODELS

ga <-lm(Cleaned ~ Concentration + Time, contrasts = list(Concentration = contr.sum, Time = contr.sum), data = interactionData)

summary(ga)

par(mfrow = c(1,2))

with (interactionData, interaction.plot(Concentration, Time, Cleaned, type = "b", main = "Original Data"))

with(interactionData, interaction.plot(Concentration, Time, ga$fitted.values, type = "b", main = "Predicted/Fitted Values"))

######CONCENTRATION WITH SALT INTERACTION MODEL

gb <-lm(Cleaned ~ Concentration + Salt, contrasts = list(Concentration = contr.sum, Salt= contr.sum), data = interactionData)

summary(gb)

par(mfrow = c(1,2))

with (interactionData, interaction.plot(Concentration, Salt, Cleaned, type = "b", main = "Original Data"))

with(interactionData, interaction.plot(Concentration, Salt, gb$fitted.values, type = "b", main = "Predicted/Fitted Values", cex =
0.5))




#######SIGNIFICANT EFFECTS######

qqnorm.lenth <- function (effects)

{

  abseffects<- 2*abs(effects)

  m <- length(abseffects)

  # Psuedo sd based all abs(effect)

  s0 <- 1.5*median(abseffects)

  non.sig <- abseffects < 2.5*s0

  # Refined estimate of sd

  pse <- 1.5*median(abseffects[non.sig])
```

```
  sme<-qt(1-(1-(1+0.95^(1/m))/2), m/3)*pse;
  sig <- abseffects>sme

  hqq<-qqnorm(effects, type = "n")
  text(hqq$x, hqq$y, labels = names (effects), col = sig + 1)
  qqline(effects)
  if(sum(sig) == 0) {cat("No Significant Effects Found by Lenth's Method. \n \n")}
  if(sum(sig) > 0)
  {cat ("Significant Factors Selected by Lenth's Method:\n",
      names(abseffects)[sig], "\n") }
  return(list(pse=pse, sme=sme))
}
#####SIGNIFICANT CONCENTRATIONS
g3 <- lm(Cleaned~(Concentration), data = interactionData)
effects <- g3$coefficients[-1]
qq<-qqnorm(effects, type = "n")
qqline(effects)
text(qq$x, qq$y, labels = names(effects))
qqnorm.lenth(effects)
par(mfrow=c(1,1))
LenthPlot(g3, las=3)
#####SIGNIFICANT CONCENTRATION AND SALT
g4 <- lm(Cleaned~(Salt +  Concentration)^2, data = interactionData)
effects2 <- g4$coefficients[-1]
qq2<-qqnorm(effects2, type = "n")
qqline(effects2)
text(qq2$x, qq2$y, labels = names(effects2))
qqnorm.lenth(effects2)
par(mfrow=c(1,1))
LenthPlot(g4, las=3)
######SIGNIFICANT CONCENTRATION AND TIME
g5 <- lm(Cleaned~(Concentration + Time)^2, data = interactionData)
effects3 <- g5$coefficients[-1]
qq3<-qqnorm(effects3, type = "n")
```

```
qqline(effects3)
text(qq3$x, qq3$y, labels = names(effects3))
qqnorm.lenth(effects3)
par(mfrow=c(1,1))
LenthPlot(g5, las=3)



#######ANOVA#############
leveneTest(Cleaned ~ Concentration, data = interactionData)
leveneTest(Cleaned ~ Salt, data = interactionData)
leveneTest(Cleaned~Time, data = interactionData)




#######Effect Comparison with Tukey Method##############
###TIME
TukeyHSD(aov(Cleaned ~ Time, interactionData))
plot(TukeyHSD(aov(Cleaned ~ Time, interactionData)))
###SALT
TukeyHSD(aov(Cleaned ~ Salt, interactionData))
plot(TukeyHSD(aov(Cleaned ~ Salt, interactionData)))
###General Treatment
TukeyHSD(aov(Cleaned ~ Treatment, interactionData))
plot(TukeyHSD(aov(Cleaned ~ Treatment, interactionData)))
####CONCENTRATION
TukeyHSD(aov(Cleaned ~ Concentration, interactionData))
plot(TukeyHSD(aov(Cleaned ~ Concentration, interactionData)))

####MODEL CHECKING##########
##SALT
par(mfrow = c(1,2))
qqnorm(g$residuals)
qqline(g$residuals)
shapiro.test(g$residuals)
plot(g$fitted.values, g$residuals, xlab = "fitted value", ylab = "residuals", col = Salt, pch = as.numeric(Concentration), main = "Residuals vs FItted in Interaction Model")
```

abline(h=0)

##TIME

par(mfrow = c(1,2))

qqnorm(g2$residuals)

qqline(g2$residuals)

shapiro.test(g2$residuals)

plot(g2$fitted.values, g2$residuals, xlab = "fitted value", ylab = "residuals", col = Time, pch = as.numeric(Concentration), main = "Residuals vs FItted in Interaction Model")

abline(h=0)

##Concentration with time

par(mfrow = c(1,2))

qqnorm(ga$residuals)

qqline(ga$residuals)

shapiro.test(ga$residuals)

plot(ga$fitted.values, ga$residuals, xlab = "fitted value", ylab = "residuals", col = Time, pch = as.numeric(Concentration), main = "Residuals vs FItted in Interaction Model")

abline(h=0)

####CONCENTRATION WITH SALT

par(mfrow = c(1,2))

qqnorm(gb$residuals)

qqline(gb$residuals)

shapiro.test(gb$residuals)

plot(gb$fitted.values, ga$residuals, xlab = "fitted value", ylab = "residuals", col = Time, pch = as.numeric(Concentration), main = "Residuals vs FItted in Interaction Model")

abline(h=0)