Pyrolysis of Biochar and growth effects on plants

Background:

The inclusion of biochar is a widely used method to increase plant growth and crop yield in the agricultural industry; this method is based around the pyrolysis of organic matter such as old plant material or manure from animals to release additional nutrients into the soil. Some of the main nutrients that are needed to accelerate plant growth are nitrogen and phosphorus which are present in both manure and the pyrolyzed organic matter. The main study we will be investigating and working on is "Biochar affects growth and shoot nitrogen in four crops for two soils" by David Olszyk. This research was based around four different crops and two different soil types. The four crops that were investigated were carrots, lettuce, soybeans and sweet corn, while the two soil types were sandy loam (Coxville soil) and loamy sand (Norfolk soil). Along with these soil and crop types, a few different types of biochar were investigated, Pine chips (PC), poultry litter (PL), swine solids (SS), switchgrass (SG) and a couple blends of PC and SG in ratios of 50/50 and 80/20. These biochar precursors were pyrolyzed at three different temperatures (350, 500, 700 C) to determine what kinds of nutrients and to what extent they were added. ***For this project we want to investigate which biochar precursors yield the highest number of specific elements when pyrolyzed and what is the optimal temperature for a maximum release of specific elements along with which biochar precursor will induce the most plant growth.*** This information is important to know for bioaccumulation purposes, if there are increased amounts of heavy metals and radioactive materials in plants and crops that will be eaten later, those materials will be transferred higher and higher up the food chain.

Exploratory Data Analysis and Data Preprocessing:

The data that will be used for this project is based off a provided .CSV file labeled, "Biochar Chemical Analysis SH Unit Change 012920". The data begins with 18 samples, 3 of each biochar precursor, which was pyrolyzed at each temperature interval, next to these are the individual values for each element that was measured during this process, the elements are as follows along with their units of measurement:

"Molybdenum (PPM), Copper (PPM), Copper (mg/kg (DO)), Lead(PPM), Zinc(PPM), Zinc(mg/kg(DO)), Argon(PPB), Nickle(PPM), Cobalt(PPM), Manganese(PPM), Manganese(mg/kg(DO)), Iron(%), Iron(g/kg(DO)), Arsenic(PPM), Uranium(PPM), Gold(PPM), Thorium(PPM), Strontium(PPM), Cadmium(PPM),

Antimony(PPM), Bismuth(PPM), Vanadium(PPM), Calcium(%), Calcium(g/kg(DO)), Phosphorus(%), Phosphorus(g/kg(DO)), Lanthanum(PPM), Chromium(PPM), Magnesium(%), Magnesium(g/kg(DO)), Barium(PPM), Titanium(PPM), Boron(PPM), Aluminum(%), Aluminum(g/kg(DO)), Sodium(%), Potassium(%), Potassium(g/kg(DO)), Tungsten(PPM), Scandium(PPM), Thallium(Tl), Sulfur(%), Sulfur(g/kg(DO)), Mercury(PPB), Selenium(PPM), Tellurium(PPM), Gallium(PPM).”

All of these elements were measured in this study allowing us to draw many possible conclusions from the provided data; however, we can break this down into another problem besides which biochar precursor adds the most nutrients. We can also investigate which one adds the highest number of heavy metals, the most radioactive materials, and so on. This additional information for all these elements can allow us to push this project in a different direction than the main paper associated with it. The provided .CSV also provides the Minimal Detectable Limit for each of these listed elements. Some preprocessing of this data may be needed, such as putting it all on the same scale as many of the elements that were measured were in PPM/PPB and others are reported in g/kg, while this may not seem like a big problem it can lead to one as g/kg are orders of magnitude larger than ppm (mg/kg). A conversion of the measurements will need to be done to better scale these values. Once these conversions are complete, we can move on to modeling and analysis of these values to discover the optimal temperature and biochar precursor that yields the highest concentrations of our desired elements. For this dataset it was decided to reduce the number of rows significantly to only observe desired chemicals for each biochar trial, the new set of column labels are as follows:

“Sample, Feedstock, Temperature, Type, Mo, Cu, , Pb, Zn, Zn(DO), Ni, Co, Mn, Mn(DO), Fe, Fe(DO), Th, Sr, Cd, V, Ca, Ca(DO), P(%), P(DO), La, Cr, Mg(%), Mg(DO), Ba, Ti, B, Al(%), Al(DO), Na(%), K(%), K(DO), W”

The second set of data has a few different sheets associated with it; these sheets contain the concentrations of specific elements detected in each of the plants along with the biochar treatment, soil type and the temperature that pyrolysis took place at for each of the cases. There is a sheet for 4 different plant types, Carrots, Corn, Lettuce and beans. These four sheets contained all of the data that would be needed for the ANOVA analysis, however there were repeated columns that were used for calculations, the following list are the column labels from the original sheets:

“crop, stakeno, Color, soil, Treatment, feedstock, temperature, trtno, repno, species, tcn, tcngkg, log10tcngkg, arctcn, oc, arcoc, C/N ratio, leafaluminum,

leaflog10aluminum, leafboron, leaflog10boron, leafcalcium, leaflog10calcium, leafcopper, leaflog10copper, leafiron, leaflog10iron, leafpotassium, leaflog10potassium, leafmagnesium, leaflog10magnesium, leafmanganese, leaflog10manganese, leafmolybenum, leaflog10molybdenum, leafsodium, leaflog10sodium, leafphosphorus, leaflog10phosphorus, leafsulfur, leaflog10sulphur, leafzinc, leaflog10zinc, revtrdw, revtrdw, revdrdw, log10revdiffrootdw, revisedrootdw, log10totrootdw, ratiotapdiff, arsintapdiff, shdw, logshootdw, shoot/total root ratio, arsinshoottotal, Ngshoot, log10Ngshoot, Alshoot, log10Algshoot, Bshoot, log10Bgshoot, Cashoot, log10Cashoot, Cushoot, log10Cushoot, Feshoot, log10Feshoot, Kshoot, log10Kgshoot, Mgshoot, log10Mgshoot, Mnshoot, log10Mnshoot, Moshoot, log10Moshoot, Nashoot, log10Nashoot, Pshoot, log10Pshoot, Sshoot, log10Sshoot, Znshoot, log10Znshoot, taprootlaboratory, traluminum, trlog10aluminum, trboron, trlog10boron, trcalcium, trlog10calcium, triron, trlog10iron, trlog10iron2, trpotassium, trlog10potassium, trpotassiummgg, trlog10potassiummgg, trmagnesium, trlog10magnesium, trmanagese, trlog10manganese, trmolybdenum, trlog10molybdenum, trsodium, trlog10sodium, trphosphorus, trlog10phosphorus, trsulphur, trlog10sulphur, trzinc, trlog10zinc, gKtaproot, log10gKtaproot, gPtaproot, log10gPtaproot, gCataproot, log10gCataproot, Mggtaproot, log10Mgtaproot, Fegtaproot, log10Fegtaproot, Mngtaproot, log10Mngtaproot, Zngtaproot, log10Zngtaproot"

For our study many of these were removed the following column names are the ones that we kept for this study:

"Root Dry weight (g), Shoot Dry weight (g), Total Dry weight (g), crop, stakeno, soil, feedstock, temperature, Total combustible nitrogen in % , total organic in %, % C/% N ratio, Aluminum concentration in µg/g, Calcium concentration in µg/g, Iron concentration in µg/g, Potassium concentration in µg/g, Magnesium concentration in µg/g, Manganese concentration in µg/g, Sodium concentration in µg/g, Phosphorus concentration in µg/g, Zinc concentration in µg/g + 0.1 because of 0's."

These kept datapoints allow us to more easily interpret and run our own analysis on the data. The removal of the original data should not affect our results as a majority of the removed columns contained calculations for other measurements that we will not be using for this study, allowing us to reduce the total amount of noise present in the data. To properly run future analysis on this data we need to have as many usable records as possible, in the original data sets there was an abundance of 0 or missing values for the detection of elements. To work around this problem, we went back and replaced these

values with a close to 0 value of .0001, this small value will allow us to still utilize the data from each observation while reducing the total amount of possible skew that can develop from the inclusion of these non-zero values.

Methods:

We performed a normality check of the data using Q-Q plots and histograms.

If the data points aligned with the regression line—regardless of the presence of extreme outliers—the data was considered normally distributed. If the points were inconsistently scattered along the line with the presence of large number of outliers, then it was discarded. Variables (mineral elements) that did not meet the normality criteria were excluded from further analysis. Histogram, a visual display of the distribution of a variable, were also examined to assess skewness. As a general rule, if the histograms are more spread toward either end, the data is said to be skewed. If the histogram follows a bell-shaped pattern, it is called a normal distribution. If both the histograms and Q-Q plots indicated skewness, those variables were not included in the analysis.

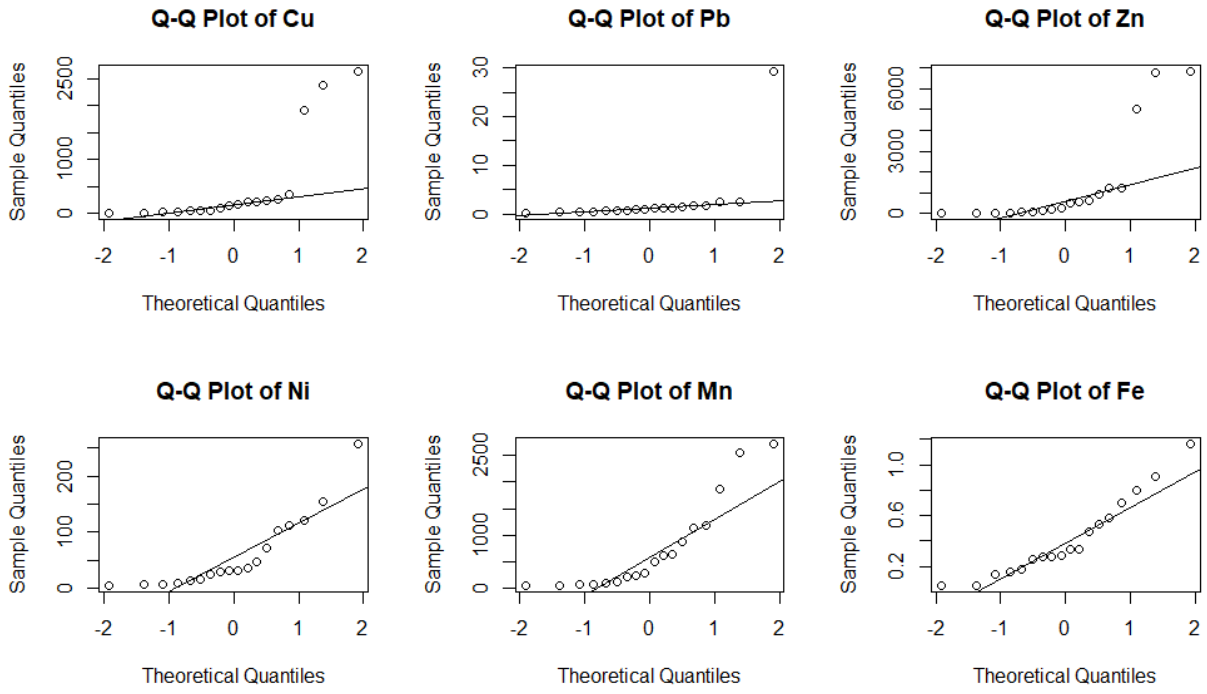

Figure One: Normality check using Q-Q Plots of elements released from biochar, elements present in this figure are copper, lead, zinc, nickel, manganese, and iron.

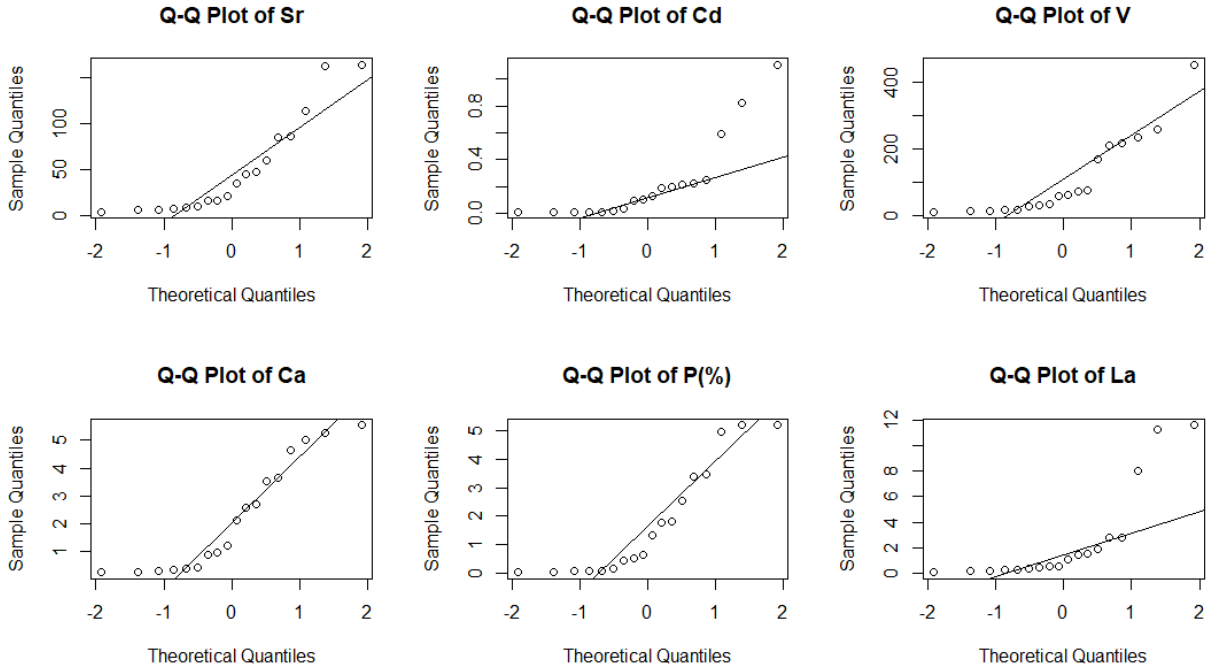

Figure Two: Normality check using Q-Q Plots of elements released from biochar, elements present in this figure are Strontium, Cadmium, Vanadium, Calcium, percent phosphorus, and lanthanum.

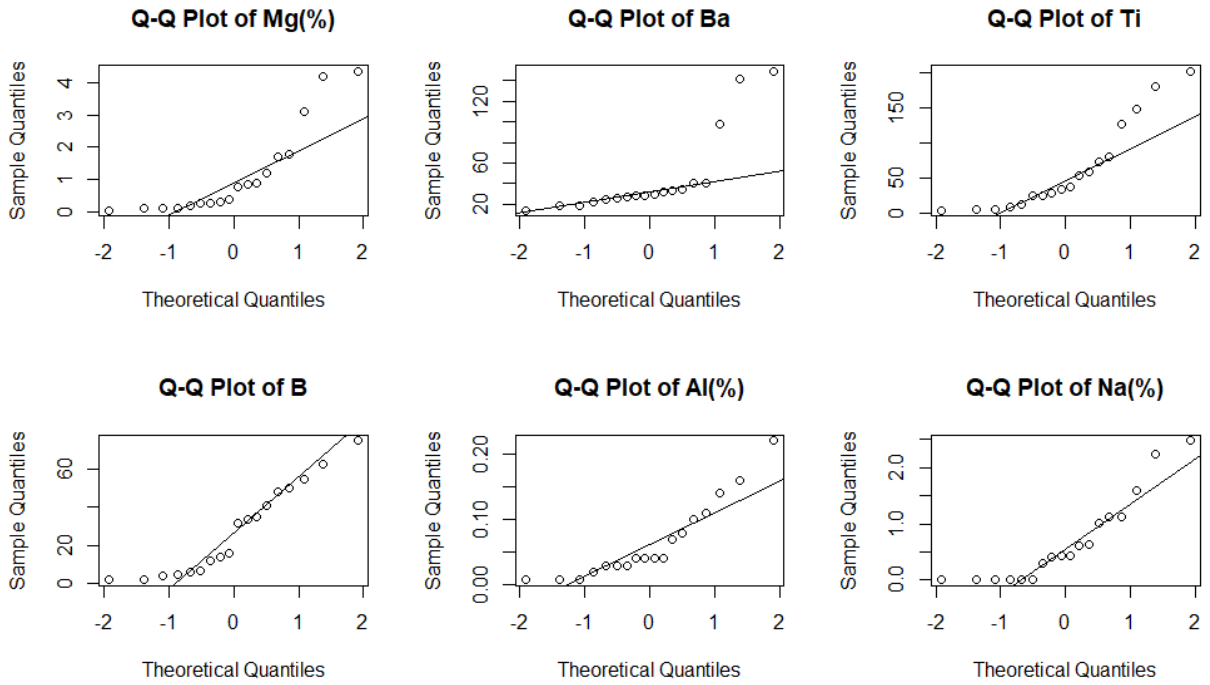

Figure Three: Normality check using Q-Q Plots of elements released from biochar,

elements present in this figure are percent magnesium, barium, titanium, boron, percent aluminum, and percent sodium.

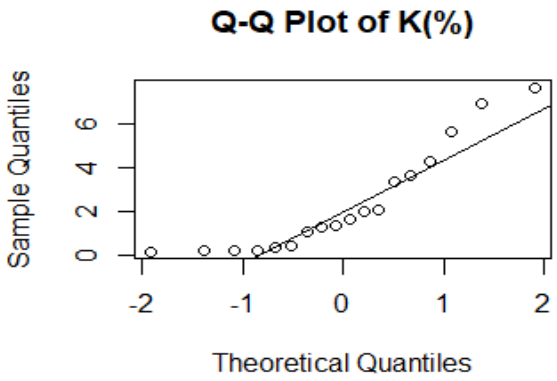

Figure Four: Normality check using Q-Q Plots of elements released from biochar, elements present in this figure are percent potassium.

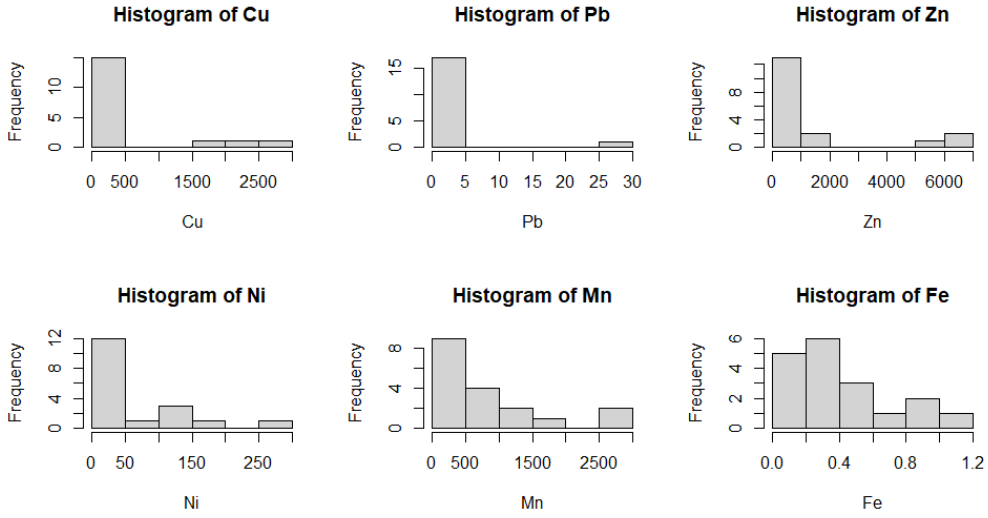

Figure Five: Skewness check using histograms of elements released from biochar, elements present in this figure are copper, lead, zinc, nickel, manganese, and iron.

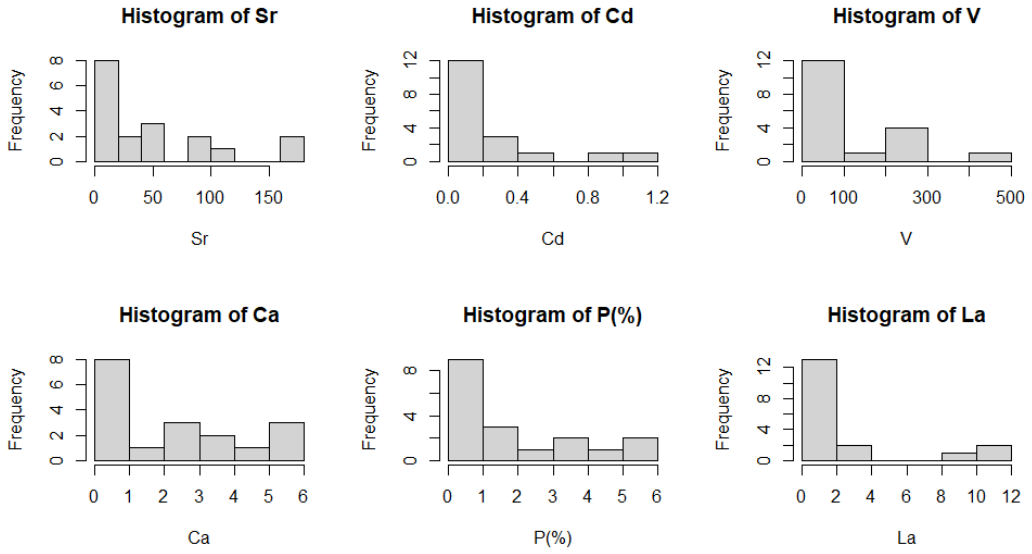

Figure Six: Skewness check using histograms of elements released from biochar, elements present in this figure are Strontium, Cadmium, Vanadium, Calcium, percent phosphorus, and lanthanum

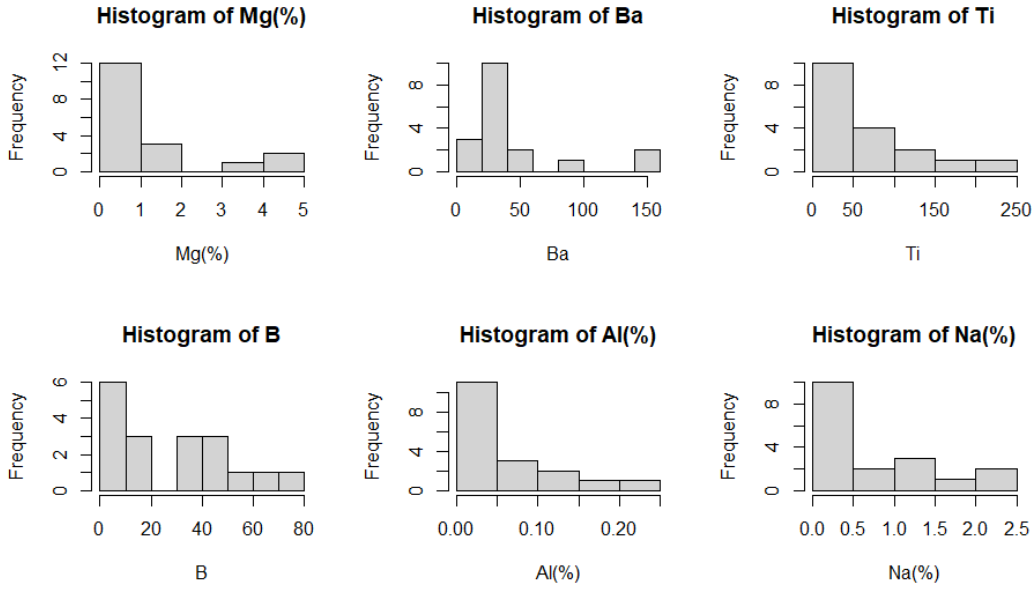

Figure Seven: Skewness check using histograms of elements released from biochar, elements present in this figure are percent magnesium, barium, titanium, boron, percent aluminum, and percent sodium.

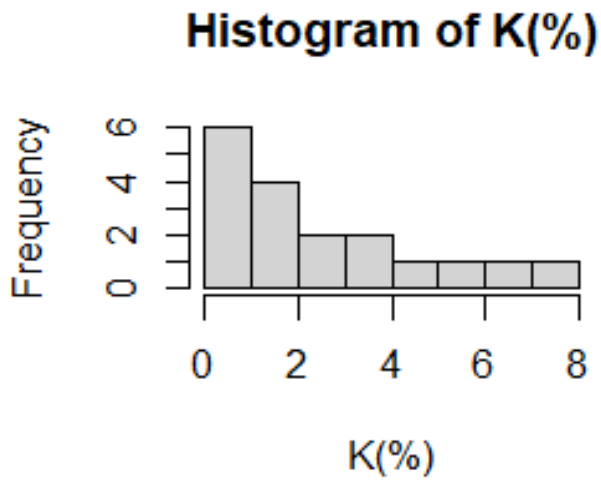

Figure Eight: Skewness check using histograms of elements released from biochar, elements present in this figure are percent potassium.

Statistical Methods Used:

The first method we used to analyze the data was Principal Component Analysis (PCA), which is an unsupervised statistical method. This method was used on the first set of data. The goal of this method was to reduce the number of variables we were working with so calculations and interpretation of the data could be more easily manageable. Along with reducing the size of the data distribution, PCA breaks the reduced data down into a few different components to summarize the data. To achieve this, we used a scree plot and chose the number of components based on where the elbow was formed, which for this set of data was 2 components. This means that the majority of our data for this dataset can be described through those two components which in this case represent the abundance of nutrients in the biochar and the saltiness produced by the biochar. The second method was K means clustering which is a unsupervised statistical method used to group similar acting data points into predetermined groups. This is done by analyzing the records and finding similar trends in their data and grouping them into an overall group that defines their attributes. For this project after the PCA analysis was completed, groupings for the K-means were formed from the two principal components, this resulted in 4 defined groupings, each with distinct but measurable characteristics. This method was done to better understand what the data represented and how the attributes from each of the

biochar's would stack up against each other. Both PCA and K means were used on the first set of data to get a better understanding of which elements and what kind of conditions would be necessary for increased plant growth. The third method of statistics that was used was ANOVA analysis; this was used on the second set of data once the attributes that contributed to plant growth were found from the PCA and K means analysis. This method was used to test the significance of a few different variables in plant growth. This method uses analysis of variance to better understand which variables are significant to a response variable. For this method the variance of the variables must be calculated followed by an f-test which is used to compare the variances between the variables followed by a p-value which indicates the level of significance for each analyzed variable. The level of significance represents the influence of a specific variable has on the outcome that is being looked at, which in this case is plant growth. These methods are further delved into in the results section of this paper where figures and interpretation of what each statistical method yielded is discussed.

Results:

-Method 1:

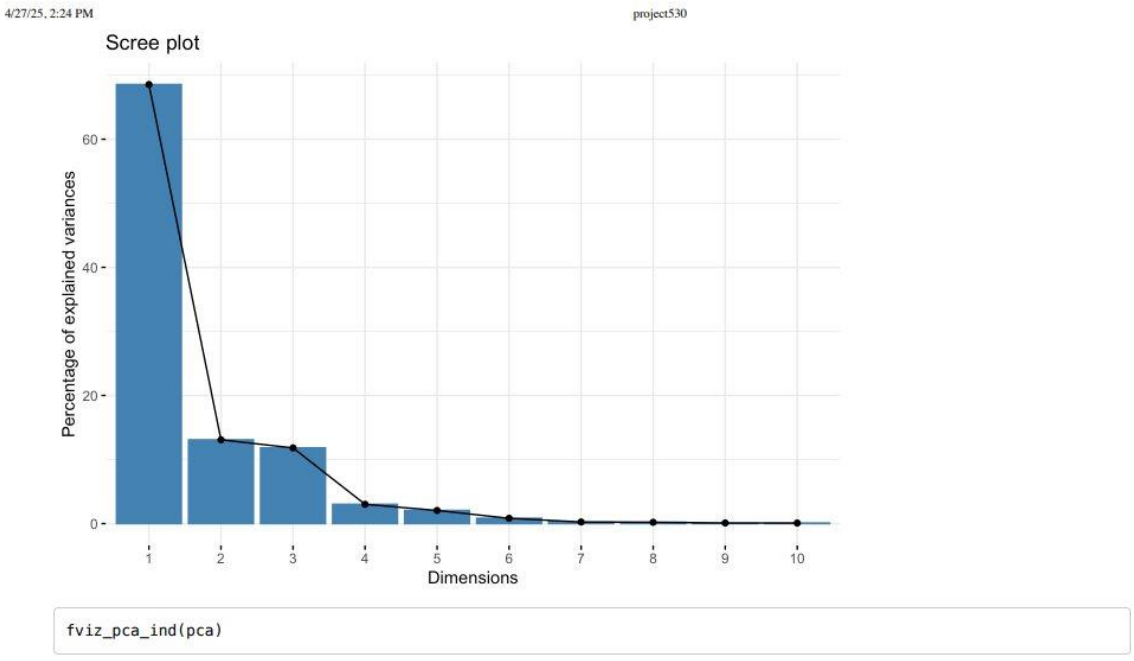

Figure Nine: This figure is the scree plot that was generated for the PCA method. From this plot we can see that there are three major components however using two should be fine

as the elbow of the plot forms lands on 2 components. The following writing will break down the interpretation of this data.

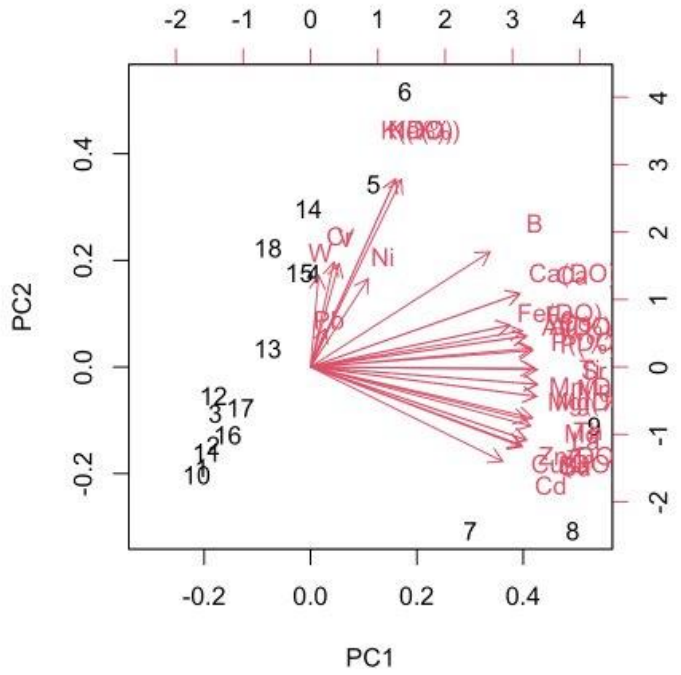

Figure Ten: This is the Bivariate plot for the PCA analysis. This plot allows us to see the score of datapoints we are looking at. It is broken down into two PC values, PC1 representing the availability of nutrients in the biochar, and PC2 representing the salt content that is produced from the biochar after it is combusted.

PCA: From the principal component analysis that was done for the first set of data, it was concluded that there were two major principal components. Principal component one explains the score for the nutrients/metals that are present in the biochar's after pyrolysis such as Mn, P, Sr, Ti, and Zn. A high score in this component indicates an increased concentration of nutrients in that biochar, a low score in this component indicates a low concentration of these elemental nutrients. The second principal explains more about the soluble elements, which are more closely associated with solubility, resulting in a higher presence of Na, K, Ba. A high score in this component indicates that a biochar has a high concentration of soluble elements and is therefore a more salty or alkaline biochar. The opposite of this is true for a low score in this component, a low score will indicate a lack of these elements and a lower chance that the biochar will be salty or alkaline. These two principal components help to reflect the grouping we saw in the K means clustering approach, which was broken down into nutrient load and salt content.

We can see a similar trend between these two methods that having a lower salt content and a higher nutrient content will help to promote growth in all plants. There were similar groupings in this method as well with a high nutrient and high salt group (samples 4,5,6), a high nutrient and low salt group (samples 7,8,9), a low nutrient and high salt group (samples 13, 14, 15, 18) and a low nutrient and low salt group (samples 1,2,3,10,11,12,16,17).

-Method 2:

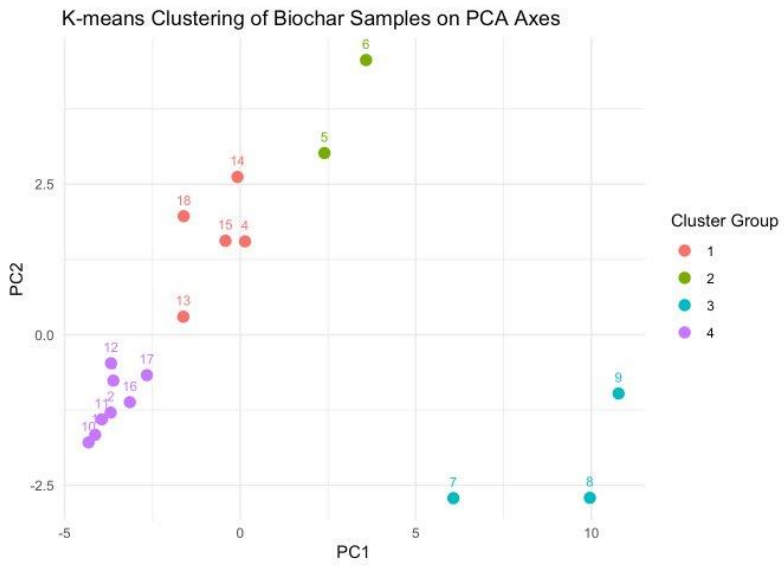

Figure Eleven: The figure above represents the clusters that were found during the k means clustering analysis. They are broken down into 4 groups, which also reflect the interpretation of data we found during the PCA analysis, which was used to create the clusters. The writing below will further breakdown and interpret what each of these clusters represent.

K Means Clustering: From the K means clustering of the first set of data, 4 main groups were observed. Group one is characterized as having moderate nutrients along with having a high salt content. This group contains samples 4,13,14,15,18. These conditions will make it harder for plants to grow as high levels of salt can cause it to be harder for plants to absorb water, which in turn will make the growing processes for plants harder, even with a moderate concentration of nutrients. Group Two is characterized as a biochar with a high nutrient presence with a high salt presence as well. This group consists of samples of 5 and 6. These conditions are also good for plant growth. However, the high salt content may make it harder for plants to absorb water and therefore some of the nutrients may not be absorbed as effectively. Group Three, which can be characterized as a biochar with a very

high nutrient presence with low salt content, contains only sample 7,8,9 these conditions are ideal for plant growth as a high concentration of nutrients will allow the plant to grow better and the lack of salt will allow water to be more readily available for the plant to use. The fourth and final group that was found through K-means clustering can be characterized as having a low nutrient content along with a low salt content, and contains samples 1,2,3,10,11,12,16,17 this combination of nutrients and salt content isn't the best but it also isn't the worst for plant growth, these plants should grow fine however they will likely not have as easy of a time growing compared to other groups. This grouping method allowed us to break the samples down into multiple groups (4) to more easily visualize and interpret the contents of the biochar that were produced during the pyrolysis process.

-Method 3:

ANOVA: From the ANOVA analysis we were able to produce 4 sets of responses, 1 for each type of plant. The results along with some interpretations of what the results mean are as follows:

```
carrot$soil <- as.factor(carrot$soil)
carrot$feedstock <- as.factor(carrot$feedstock)
carrot$temperature <- as.factor(carrot$temperature)
anova_model_carrot <- aov(total_dry ~ soil + feedstock + temperature, data = carrot)
summary(anova_model_carrot)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## soil          1   7.75   7.750   16.70 6.21e-05 ***
## feedstock     6  52.80   8.800   18.96  < 2e-16 ***
## temperature   2   2.47   1.234    2.66   0.0723 .
## Residuals   212  98.39   0.464
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 6 observations deleted due to missingness
```

Figure Twelve: This represents the outcome of the ANOVA analysis that was done on the second set of data for the carrots plant type. From this we can see that both the soil type and the biochar are significant variables that will describe how the plant growth will function.

```
soybean$soil <- as.factor(soybean$soil)
soybean$feedstock <- as.factor(soybean$feedstock)
soybean$temperature <- as.factor(soybean$temperature)
anova_model_soy <- aov(total_dry ~ soil + feedstock + temperature, data = soybean)
summary(anova_model_soy)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## soil          1 157.56  157.56 245.127 <2e-16 ***
## feedstock     6 107.02   17.84  27.751 <2e-16 ***
## temperature   2   4.06    2.03   3.155 0.0446 *
## Residuals   218 140.12    0.64
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure Thirteen: This represents the outcome of the ANOVA analysis that was done on the second set of data for the soybean plant type. From this we can see that both the soil type and the biochar are significant variables that will describe how the plant growth will function. This is like the carrot plant type and may be due to the tendency of these plants both growing directly in the soil, and therefore soil having a larger effect on the growth of these plants.

```
lettuce$soil <- as.factor(lettuce$soil)
lettuce$feedstock <- as.factor(lettuce$feedstock)
lettuce$temperature <- as.factor(lettuce$temperature)
anova_model_lettuce <- aov(total_dry ~ soil + feedstock + temperature, data = lettuce)
summary(anova_model_lettuce)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## soil          1   1.25   1.254   1.530  0.217
## feedstock     6 143.79  23.965  29.241 <2e-16 ***
## temperature   3   1.10   0.368   0.449  0.718
## Residuals   217 177.85   0.820
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure Fourteen: This represents the outcome of the ANOVA analysis that was done on the second set of data for the lettuce plant type. From this we can see that only the biochar feedstock is significant for this set of data. This allows us to infer that soil and temperature do not have much of an effect on the growth. Dissimilar to soybeans and carrots soil does not have any significance for this set of data.

```
corn$soil <- as.factor(corn$soil)
corn$feedstock <- as.factor(corn$feedstock)
corn$temperature <- as.factor(corn$temperature)
anova_model_corn <- aov(total_dry ~ soil + feedstock + temperature, data = corn)
summary(anova_model_corn)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## soil          1 2166.4  2166.4  567.38  < 2e-16 ***
## feedstock     8  669.9    83.7   21.93  < 2e-16 ***
## temperature   2  106.2    53.1   13.91 2.07e-06 ***
## Residuals   216  824.7     3.8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure Fifteen: This represents the outcome of the ANOVA analysis that was done on the second set of data for the corn plant type. From this we can see that all three tested variables are significant for plant growth. Which makes some sense as corn is a harder plant to grow in harsh environments, having soil, biochar type and temperature play a significant role in the growth of these corn plants

Overall we can see that the biochar feedstock holds the most significance across all plant types, leading to the conclusion that as more and more nutrients are introduced into the plants surrounding environment through the biochar the growth of the plant will also increase and relies heavily on what nutrients are available in the soil. Soil type is also significant in some of these cases which also makes sense, since each type of soil may contain different elements along with other differences that could impact growth such as sand content, and permeability. Temperature of combustion does not hold as much significance as we thought it would, this could be due to some elements that when pyrolyzed are not as sensitive to these temperature changes as other elements. Overall, this approach led us to the conclusion that the biochar will have a significant impact on the growth of all plant types and therefore should be selected based on which elements are needed to promote plant growth.

Conclusion and Discussion:

From the two clustering methods we see four distinct groups of biochar's that can be formed, this allows us to draw conclusions about what characteristics the biochar would need to have to lead to improved plant growth, our conclusion from these two methods are that if the biochar produces a high concentration of nutrients along with a lower concentration of salt, that biochar would be beneficial to plant growth and should be used for growing said plants. Using the ANOVA data, we can also conclude which biochar led to the most growth across plants due to the presence of a response variable. Due to the high

presence of plant required nutrients along the promotion of a lower salt content, Swine Solid biochar seems to be the most appropriate for inducing growth in plants across the board. Swine Solids contains an increased amount of desired nutrients that will promote plant growth and will also help to keep the salt content at lower levels allowing plants to properly absorb nutrients and water to promote quality plant growth. It also seems that combustion temperature does not hold very much significance when it comes to Swine Solids, as all three samples of Swine Solids produced significantly more nutrients than the other samples. This can be seen in both of the clustering methods along with the results of the ANOVA analysis, therefore leading us to the conclusion that in an ideal situation Swine Solids should be used as the biochar precursor to promote the growth of plants and they can be combusted at any of the three temperatures while still having a significant effect.

Sources:

1) *Essential Plant Nutrients*. Alabama Cooperative Extension System. (2024, December 9). https://www.aces.edu/blog/topics/farming/essential-plant-elements/
2) **Olszyk, D. M., Shiroyama, T., Novak, J. M., Cantrell, K. B., Sigua, G., Watts, D. W., & Johnson, M. G. (2020). *Biochar Affects Essential Nutrients of Carrot Taproots and Lettuce Leaves*. HortScience horts, 55(2), 261-271. Retrieved Apr 28, 2025, from https://doi.org/10.21273/HORTSCI14421-19**
3) **Olszyk, D., Shiroyama, T., Novak, J., Cantrell, K., Sigua, G., Watts, D., & Johnson, M. G. (2020). *Biochar affects growth and shoot nitrogen in four crops for two soils*. Agrosystems, Geosciences & Environment, 3(1). https://doi.org/10.1002/agg2.20067**

Data Citations:

4) **Olszyk, D., Shiroyama, T., Novak, J., Cantrell, K., Sigua, G., Watts, D., & Johnson, M. G. (2020). *Biochar Chemical Analysis SH Unit Change 012920.xlsx*.**

Agrosystems, Geosciences & Environment, 3(1). https://doi.org/10.1002/agg2.20067

5) **Olszyk, D., Shiroyama, T., Novak, J., Cantrell, K., Sigua, G., Watts, D., & Johnson, M. G. (2020). *Carrot SH Revised 012920.xls*. Agrosystems, Geosciences & Environment, 3(1). https://doi.org/10.1002/agg2.20067**
6) **Olszyk, D., Shiroyama, T., Novak, J., Cantrell, K., Sigua, G., Watts, D., & Johnson, M. G. (2020). *Corn SH Revised 012920.xls*. Agrosystems, Geosciences & Environment, 3(1). https://doi.org/10.1002/agg2.20067**
7) **Olszyk, D., Shiroyama, T., Novak, J., Cantrell, K., Sigua, G., Watts, D., & Johnson, M. G. (2020). *Lettuce SH 012920.xls*. Agrosystems, Geosciences & Environment, 3(1). https://doi.org/10.1002/agg2.20067**
8) **Olszyk, D., Shiroyama, T., Novak, J., Cantrell, K., Sigua, G., Watts, D., & Johnson, M. G. (2020). *Soybean SH Revised 012920.xls*. Agrosystems, Geosciences & Environment, 3(1). https://doi.org/10.1002/agg2.20067**