MT5762 PROJECT 1

ANDREAS MICHALIS LOUCAS

200029834

2020

## QUESTION 1[Which 5 elements did you choose, and why?]

The 5 elements chosen to undertake the project are [Sc,Ti,Se,Sm,Tm]. These five elements have been chosen because at a first glimpse the groups for each element looked similar. We can visualize this characteristics from the APPENDIX [BOXPLOT[1-5]](#_BOXPLOTS_[1-5]) .

## QUESTION 2[Do the data indicate differences in the elemental composition of Cannabis leaves grown in different soil types?]

From the APPENDIX [BOXPLOT[1-5]](#_BOXPLOTS_[1-5]) we cannot determine an obvious answer if there are differences in elemental composition since some soils data overlap with each other.

Our data has to fulfil the F-distribution assumptions to achieve precise results. Data must follow independence, normality and have constant spread between them.

For the normality assumption we need to test hypothesis **Ho:** element follows normal distribution and **H1:** element does not follow normal distribution. We tested the hypothesis using a Shapiro-Wilk test and elements [Sc,Ti,Sm,Tm] are normally distributed, but [Se] is not. Results can be found in APPENIX [TABLE[SHAPIRO]](#_TABLE[SHAPIRO]) .

For constant spread assumption we need to test hypothesis **Ho:** σ(mb)^2 = σ(bhb)^2 = σ(pm)^2 = σ(nth)^2 and **H1:** at least 1 variance is not equal with at least 1 other variance. We tested the hypothesis using a Levene’s Test and element [Sc,Ti,Se] satisfies it, but elements [Sm,Tm] do not satisfy it at p-value = 0.05, but satisfy it if the p-value = 0.01 . Results can be found in APPENDIX [TABLE[LEVENE]](#_TABLE[LEVENE])

The ANOVA hypothesis for each of our 5 elements is, **Ho:** μ(mb) = μ(bhb) = μ(pm) = μ(nth) and **H1:** exists at least 1 μ not equal with at least 1 other μ.

The results from the ANOVA test can be found in APPENDIX [TABLE[ANOVA]](#_TABLE[ANOVA]). Our result states that for all 5 elements there exists at least 1 mean group that is not equal with at least one other mean group. Our conclusion from this, is that data indicate differences in the elemental composition of Cannabis leaves grown in different soil types. We should not rely on these results, since all of our assumptions were not met and maybe we have returned false results for some elements. All tests were made considering a 5% type-I-error. The means that influence the rejection of **Ho** can be discovered from the APPENDIX [GRAPH[1-5]](#_GRAPHS[1-5]). If the range between the lower and upper boundary does not contain zero. Then our means between the two groups are significantly different. If zero exists in the lower and upper boundary, than the means of the two groups are not significantly different.

# QUESTION 3[Are some of the elements related to one another in terms of their levels in the sampled leaves?]

We will explore if elements have a relation with each other from the cannabis leave results. The method we will apply to compute such procedure is by using the correlation coefficient ‘ρ’ measure. Our pairs for the examination are [Sc-Ti, Sm-Tm, Sm-Sc, Tm-Ti, Se-Sc]. Results of the correlation coefficient for these pairs can be found in APPENDIX [TABLE[CORRELATION]](#_TABLE[CORRELATION]). Our results state that most of our pairs, have a positive linear relationship. This indicates that each element’s value in cannabis leaves depends and influences another elements value. Our strongest pair from our tests is Sm - Sc with a value approximately 0.77 and a perfect value of positive linear relation is 1. Our weakest relationship which is close to a no relation value is Sc - Tm with a value of 0.314 .

## QUESTION 4[Results of this experiment ultimately allow the determination of what soil the plants were grown in, just from the elemental composition of the leaves?]

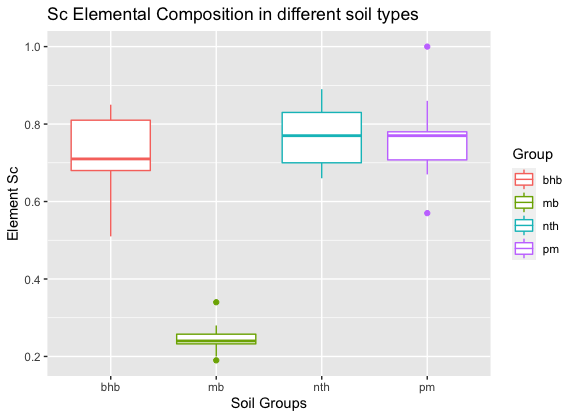
When considering this experiment on a global scale, I think this cannot be achieved for the given reasons. Firstly, we only consider four soil types in this experiment in one country. In reality there are thousand different soil types when considering all the countries. Secondly, each country has different weather and climate circumstances which is an important variable for plant growth. Thirdly, our element samples that we considered in this experiment did not fulfil our assumption fully. This might have influenced our results and we might have achieved untrue outcomes. Fourthly, we tested all the Hypothesis considering a 5% type-I-error, this implies that there is a 5% probability we accepted the wrong hypothesis and concluded to false assumptions. Lastly, there are different methods in growing plants. It can be done indoors-outdoors and can be chemically strengthen or allowed to grow on its own. These variables might influence and change the element composition in the cannabis leaves. For these reasons, I think this is a good start, but we still have many scenarios and variables we have to consider to achieve more satisfactory results.

# APEENDIX

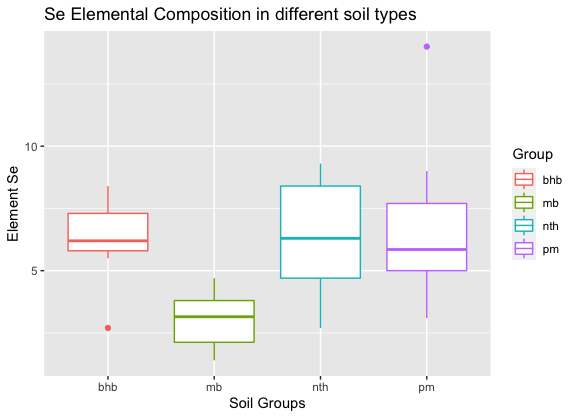
## GRAPHS-PLOTS

### BOXPLOTS [1-5]

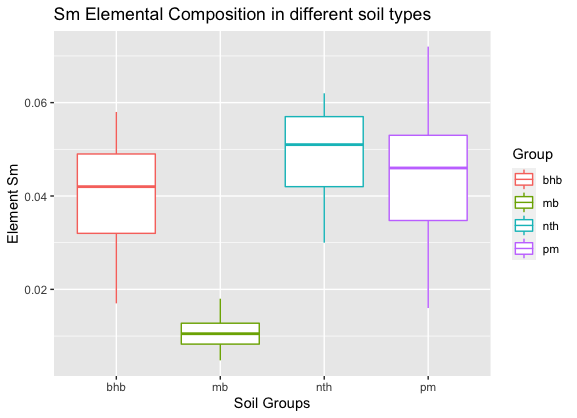
**BOXPLOT[1]**



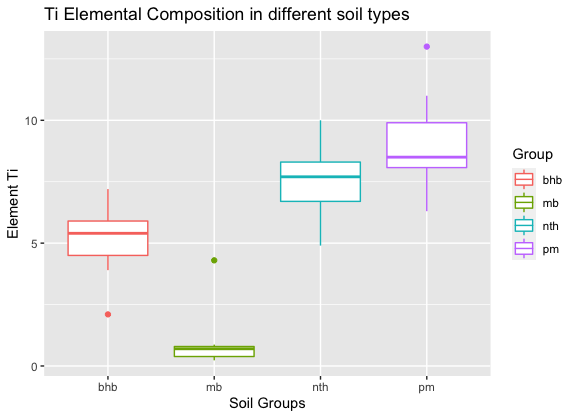
**BOXPLOT[2]**



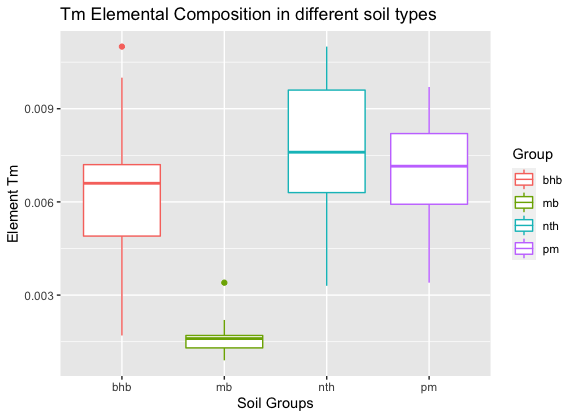
**BOXPLOT[3]**



**BOXPLOT[4]**



**BOXPLOT[5]**



### GRAPHS[1-5]

**GRAPH[1] Group Sc**

**Chart, box and whisker chart

Description automatically generated**

**GRAPH[2] Group Se**

**Chart, box and whisker chart

Description automatically generated**

**GRAPH[3] Group Se**

**Chart, box and whisker chart

Description automatically generated**

**GRAPH[4] Group Ti**

**Chart, box and whisker chart

Description automatically generated**

**GRAPH[5] Group Tm**

**Chart, box and whisker chart

Description automatically generated**

## TABLES

### TABLE[SHAPIRO]

|  |  |  |
| --- | --- | --- |
| ELEMENT | P-VALUE | Ho = Reject/Accept (Accepting for p-value > 0.05) |
| Sc | 0.30 | Large p-value. Accept |
| Ti | 0.10 | Large p-value. Accept |
| Se | 0.008 | Smaller p-value than 0.05. Reject |
| Sm | 0.94 | Large p-value. Accept |
| Tm | 0.81 | Large p-value. Accept |

### TABLE[LEVENE]

|  |  |  |
| --- | --- | --- |
| ELEMENT | P-VALUE | Ho = Reject/Accept (Accepting for p-value > 0.05) |
| Sc | 0.13 | Large p-value. Accept |
| Ti | 0.46 | Large p-value. Accept |
| Se | 0.18 | Large p-value. Accept |
| Sm | 0.03 | Smaller p-value than 0.05. Reject |
| Tm | 0.02 | Smaller p-value than 0.05. Reject |

### TABLE[ANOVA]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ELEMENT | Pr(>F) | F-value | P-Value | Ho=Reject/Accept |
| Sc | -10.56 | 107.80 | 2.7826 | F-value a lot larger than p-value. Reject Ho |
| Ti | -10.56 | 80.05 | 2.7826 | F-value a lot larger than p-value. Reject Ho |
| Se | 0.004 | 7.44 | 2.7826 | F-value a lot larger than p-value. Reject Ho |
| Sm | 12.45 | 23.65 | 2.7826 | F-value a lot larger than p-value. Reject Ho |
| Tm | 0.16 | 21.53 | 2.7826 | F-value a lot larger than p-value. Reject Ho |

### TABLE[CORRELATION]

|  |  |  |
| --- | --- | --- |
| PAIR OF ELEMENTS | VALUE OF ‘ρ’ | Relation |
| Sc with Ti | 0.757 | Positive Linear Relationship |
| Sm with Tm | 0.733 | Positive Linear Relationship |
| Sm with Sc | 0.765 | Positive Linear Relationship |
| Tm with Ti | 0.594 | Positive Linear Relationship |
| Sc with Tm | 0.315 | No relation-Positive Linear Relationship (Closer to no relation) |

# TABLES-SUMMARIES

# APPENDIX - CODE