STAT 151 X01

Group #23

Lab 4  
**Julian STAMER**

Robert JOSEPH

Nolan Van EGMOND

1)

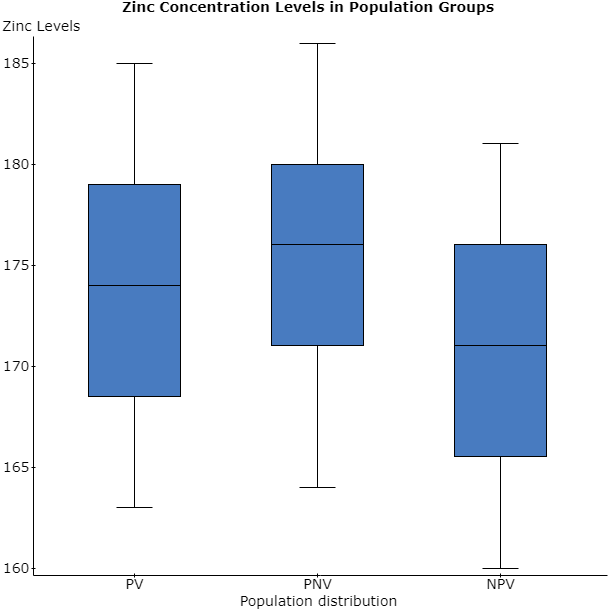
a) The 36 subjects in the study responded voluntarily, therefore this was voluntary sampling and the outcomes cannot be generalized to any population whatsoever.

1. This study is an example of an observational study.

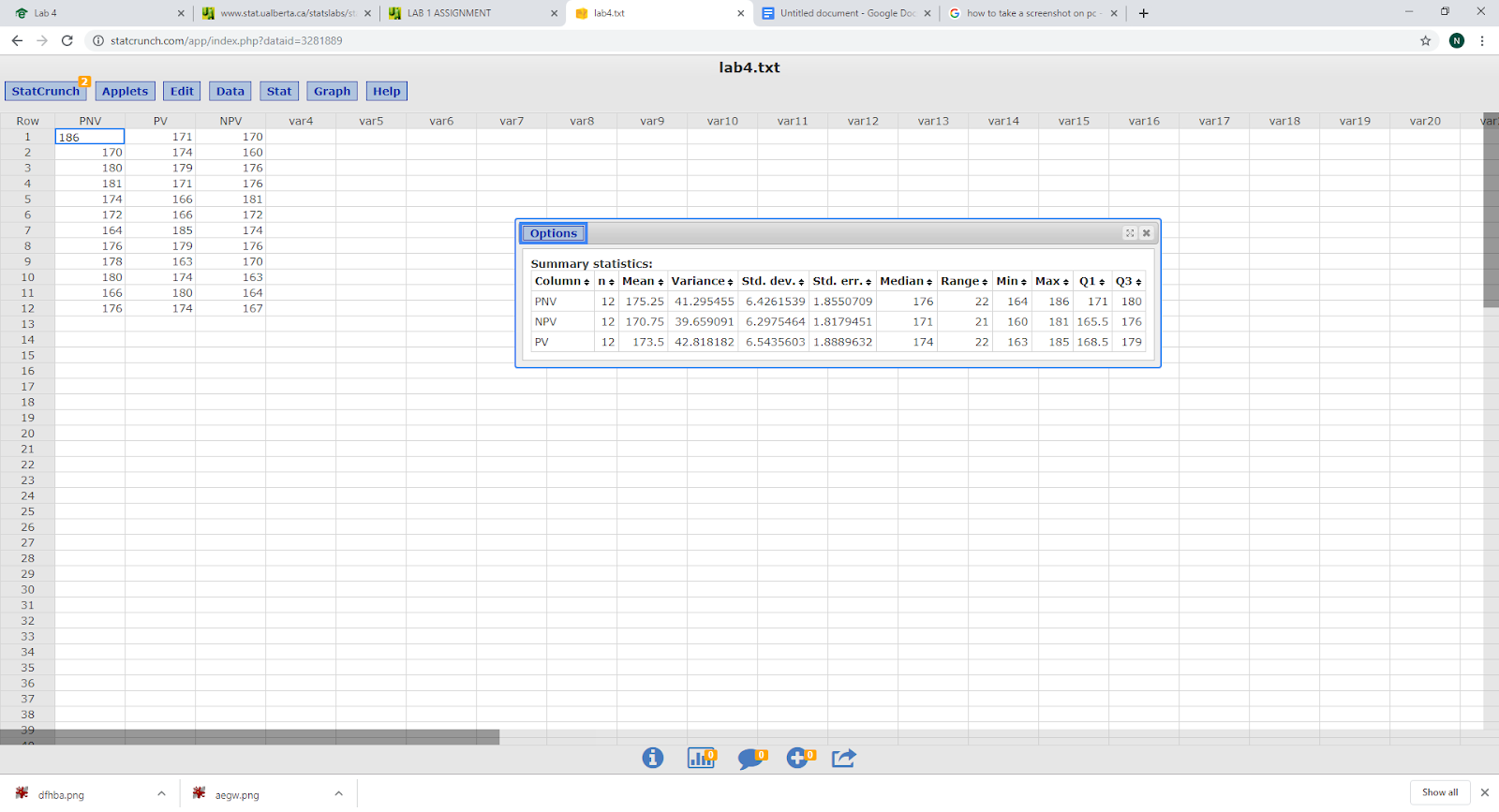
Therefore, this study cannot be used to draw conclusions since the participants were volunteers and because the number of participants are too small to conclude anything.

2)

a)



b) There does seem to be a difference in means for the three populations. It seems that non-pregnant vegetarian women have smaller levels of zinc with a mean of 170.75 while pregnant non-vegetarian women have the most with a mean of 175.25. There do not appear to be any outliers in any of the populations.

c) 

The group with highest mean Zinc levels is the Pregnant Non-vegetarian group

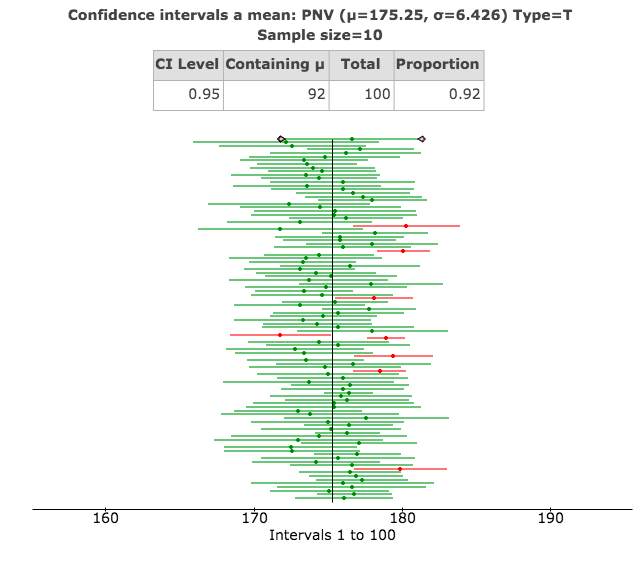
The group with lowest mean Zinc levels is the Non pregnant vegetarian group

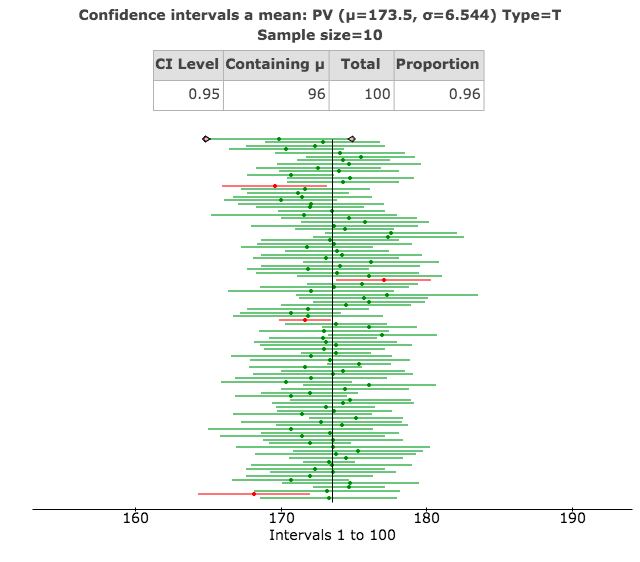
There are no substantial differences in the standard deviations between these groups.

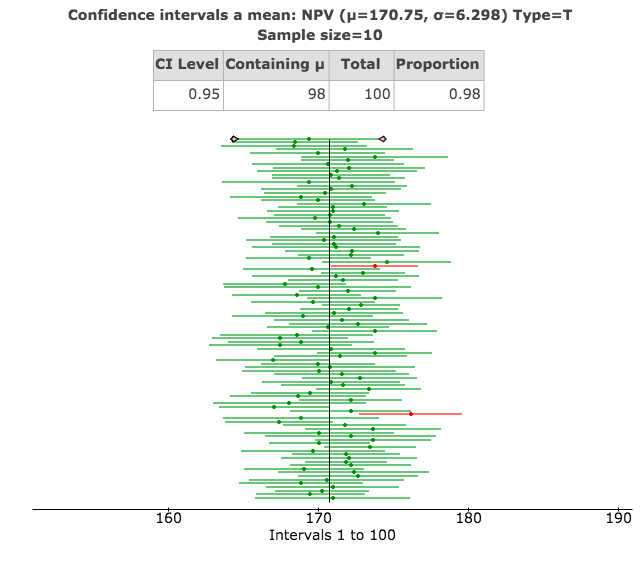
PV has the highest standard deviation being 6.54, PNV has 6.43 and NPV

is again the smallest with 6.29

d)







95% Confidence Interval:

PNV: 175.25 ± 3.64 = (172 to 179)

NPV: 170.75 ± 3.56 = (167 to 174)

PV: 173.5 ± 3.7 = (170 to 177)

All the confidence intervals are close and differ only slightly, so it is hard to make an inference on the difference.

3)

a)

**Two sample T hypothesis test:**

μ1 : Mean of PV

μ2 : Mean of PNV

μ1 - μ2 : Difference between two means

H0 : μ1 - μ2 = 0

HA : μ1 - μ2 < 0

(without pooled variances)

A screenshot of a cell phone

Description automatically generated

The null hypothesis is H0: μ1- μ2 = 0

The alternative hypothesis is HA: μ1- μ2 < 0

The value of the test statistic is - 0.66099088

Its distribution under the null hypothesis is 21.992792.

The P-value is 0.2577.

As the P -value is greater than alpha that is 0.2577>0.05 we do not reject H0 . There isn’t sufficient evidence to prove that PV tend to have a lower zinc level than PNV.

b)

**Two sample T confidence interval:**

μ1 : Mean of PV

μ2 : Mean of NPV

μ1 - μ2 : Difference between two means

(with pooled variances)

**95% confidence interval results:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Difference** | **Sample Diff.** | **Std. Err.** | **DF** | **L. Limit** | **U. Limit** |
| μ1 - μ2 | -1.75 | 2.6475403 | 22 | -7.2406625 | 3.7406625 |

Therefore, the 95 percent confidence interval for the difference (μ1 - μ2) is ( -7.2407669, 3.7407669)

As 0 lies in the confidence interval we therefore fail to reject the null hypothesis and therefore no evidence to prove that PV tend to have a lower zinc level than PNV.

It is also consistent with the result of the test in part a.

c)

The assumptions are:

* Random samples.
* Independent samples.
* Normally distributed along with n1 and n2 greater than 30.
* The distributions are approximately normal.

The sample does satisfy the less than 10% condition.

The sample also satisfies the normality condition

However, the sample does not satisfy the randomization condition since it is not a random sample as it is a voluntary sample.

4) a)

**Two sample T hypothesis test:**

μ1 : Mean of PV

μ2 : Mean of NPV

μ1 - μ2 : Difference between two means

H0: μ1 - μ2 = 0

HA : μ1 - μ2 ≠ 0

(without pooled variances)

**Hypothesis test results:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Difference** | **Sample Diff.** | **Std. Err.** | **DF** | **T-Stat** | **P-value** |
| μ1 - μ2 | 2.75 | 2.6216609 | 21.967771 | 1.0489533 | 0.3056 |

According to the output, the test statistic is 1.0489533 with a sample difference of 2.75 and the statistic follows a standard normal distribution. The *P-*value of the test is 0.3056 and is larger than α = 0.05, so we cannot conclude there is evidence in mean zinc levels between pregnant vegetarians and non-pregnant vegetarians.

b)

**Two sample T confidence interval:**

μ1: Mean of PV

μ2: Mean of NPV

μ1 - μ2: Difference between two means

(without pooled variances)

**95% confidence interval results:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Difference** | **Sample Diff.** | **Std. Err.** | **DF** | **L. Limit** | **U. Limit** |
| μ1 - μ2 | 2.75 | 2.6216609 | 21.967771 | -2.6874545 | 8.1874545 |

A 95% interval for the difference in mean zinc levels between pregnant vegetarians and non-pregnant vegetarians (-2.6874545 to 8.1874545). The interval contains zero, so there is no difference in proportions, and we fail to reject the null hypothesis. This finding is consistent with the result of the test in 4a.

5) a) H0: Mean of PNV = Mean of PV = Mean of NPV

HA: At least one of the means of PNV, PV, NPV is different

   b)

**Analysis of Variance results:**

Data stored in separate columns.

**Column statistics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Column** | **n** | **Mean** | **Std. Dev.** | **Std. Error** |
| PNV | 12 | 175.25 | 6.4261539 | 1.8550709 |
| NPV | 12 | 170.75 | 6.2975464 | 1.8179451 |
| PV | 12 | 173.5 | 6.5435603 | 1.8889632 |

**ANOVA table**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **SS** | **MS** | **F-Stat** | **P-value** |
| Columns | 2 | 123.5 | 61.75 | 1.4966948 | 0.2387 |
| Error | 33 | 1361.5 | 41.257576 |  |  |
| Total | 35 | 1485 |  |  |  |

The pooled estimate of variance is 41.257576

Sum of squares due to treatments = 123.5

Sum of squares due to error = 1361.5

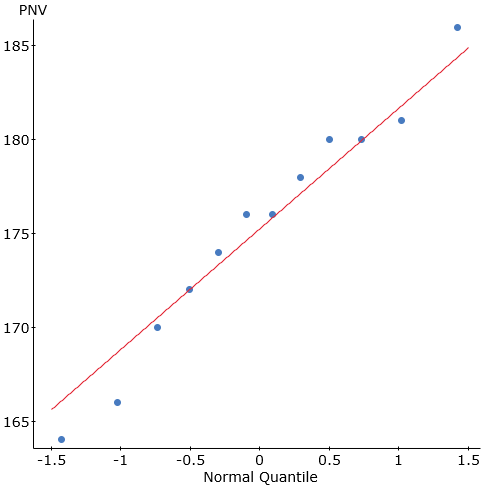
The value of the F-statistic is 1.4966948

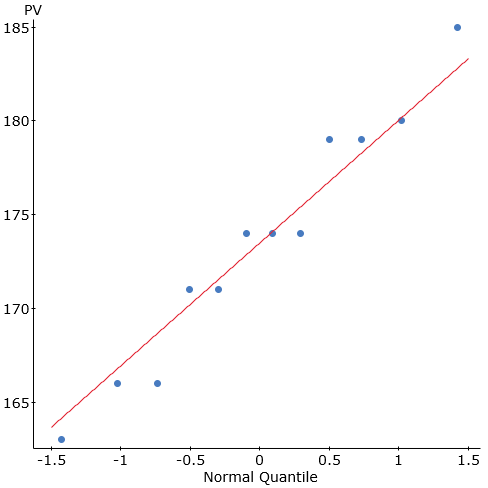
The p-value of the test is 0.2387

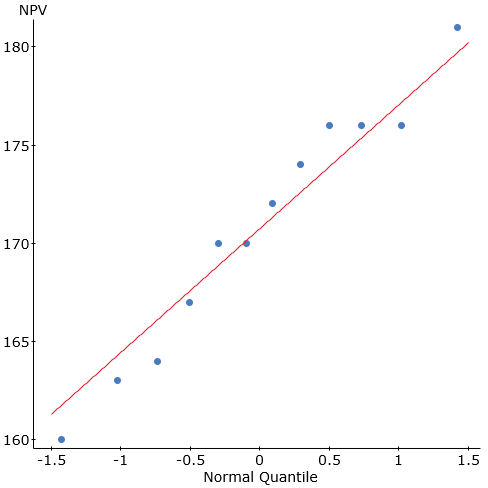
The distribution of the test statistic is by F-model with MST having degrees of freedom 2 and MSE having degrees of freedom 33

P-value > 𝛂 so do no reject H0 

c)







All we can observe all the QQ plots are normally distributed as they are close to the center line. Referring to the summary statistics and boxplots in Question 2, it appears that all three groups have roughly the same spread as the lines above and below the boxplot are roughly the same size and according to the summary statistics, the variance (39.66 to 42.82), standard deviation (6.30 to 6.54), and range (21 to 22) are all roughly the same for the three groups.