Faculty of Engineering and Materials Sciences
German University in Cairo

Effect of Chocolate Consumption on Heart disease
(Cardiovascular disease)

MATS 902

DOE

Project

Authors:

Ahmed EssamEldin Alrefaie	ID: 46-5804 (T-03)
Mohamed Ahmed Salaheldin	ID: 46-20529 (T-03)
Mohamed Ahmed Shaaban	ID: 46-10273 (T-03)
Hassan Hatem Hakim	ID: 46-4672 (T-02)
Omar Nasser El-Nabulsi	ID: 46-10512 (T-05)
Ahmed Shousha	ID:46-7700 (T-03)

Supervised By:

Assoc. Prof. Dr. Lamia A. Shihata

Eng.Alaa Zarif

Contents

Chapter 1

Introduction

1.1 Motivation

Recent studies show that there is a drastic increase in cardiovascular health problems in recent years. Anti-oxidants are a great way to help prevent such problems. There is some speculation that dietary flavonoids from chocolate, in particular epicatechin, may promote cardiovascular health as a result of direct antioxidant effects or through antithrombotic mechanisms. Thus the experiment aims to show the effect of different types of chocolate on the amount of antioxidants in plasma

1.2 Experiment

[1]Our chosen experiment was the experiment conducted on 28th of August 2003 in nature. The paper is named “Plasma Antioxidants from Chocolate”. The experiment has 1 factor which is chocolate, and they were studying its effect on cardiovascular health. All the data acquired for this project is gathered from Douglas Montgomery’s textbook “Design and analysis of experiments”.

The 3 treatments for this experiment were: 100 g of dark chocolate, 100 g of dark chocolate with 200 mL of full-fat milk, and 200 g of milk chocolate. The sample size was 12 subjects of which 7 were women and 5 were men. The average age range was 32.2 ± 1 year and range (25–35 years). All were non-smokers, had low blood lipid levels, were not taking drugs or vitamins supplements and they averaged a weight of 65.8 ± 3.1 Kg (46.0 Kg to 86.6 Kg range) with a BMI of 21.9 ± 0.4 Kg/m^2 (18.6 Kg/m^2 to 23.6 Kg/m^2 range). On different days, the subjects consumed 100 grams of each type of chocolate. After 1 hour of consumption their total anti-oxidant capacity of their plasma was measured using FRAP assay. epicatechin, may promote cardiovascular health as a result of direct antioxidant effects.

1.2.1 Dataset

	Subjects (Observations)											
Factor	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
DC	118.8	122.6	115.6	113.6	119.5	115.9	115.8	115.1	116.9	115.4	115.6	107.9
DC + MK	105.4	101.1	102.7	97.1	101.9	98.9	100.0	99.8	102.6	100.9	104.5	93.5
MC	102.1	105.8	99.6	102.7	98.8	100.9	102.8	98.7	94.7	97.8	99.7	98.6

Chapter 2

Grouped Data Findings

2.1 Grouped Dataset

The following table shows the dataset gathered from the paper [2]

118.8	122.6	115.6	113.6	119.5	115.9	115.8	115.1	116.9	115.4	115.6	107.9
105.4	101.1	102.7	97.1	101.9	98.9	100.0	99.8	102.6	100.9	104.5	93.5
102.1	105.8	99.6	102.7	98.8	100.9	102.8	98.7	94.7	97.8	99.7	98.6

2.2 Measures Of Central Tendency

Statistics

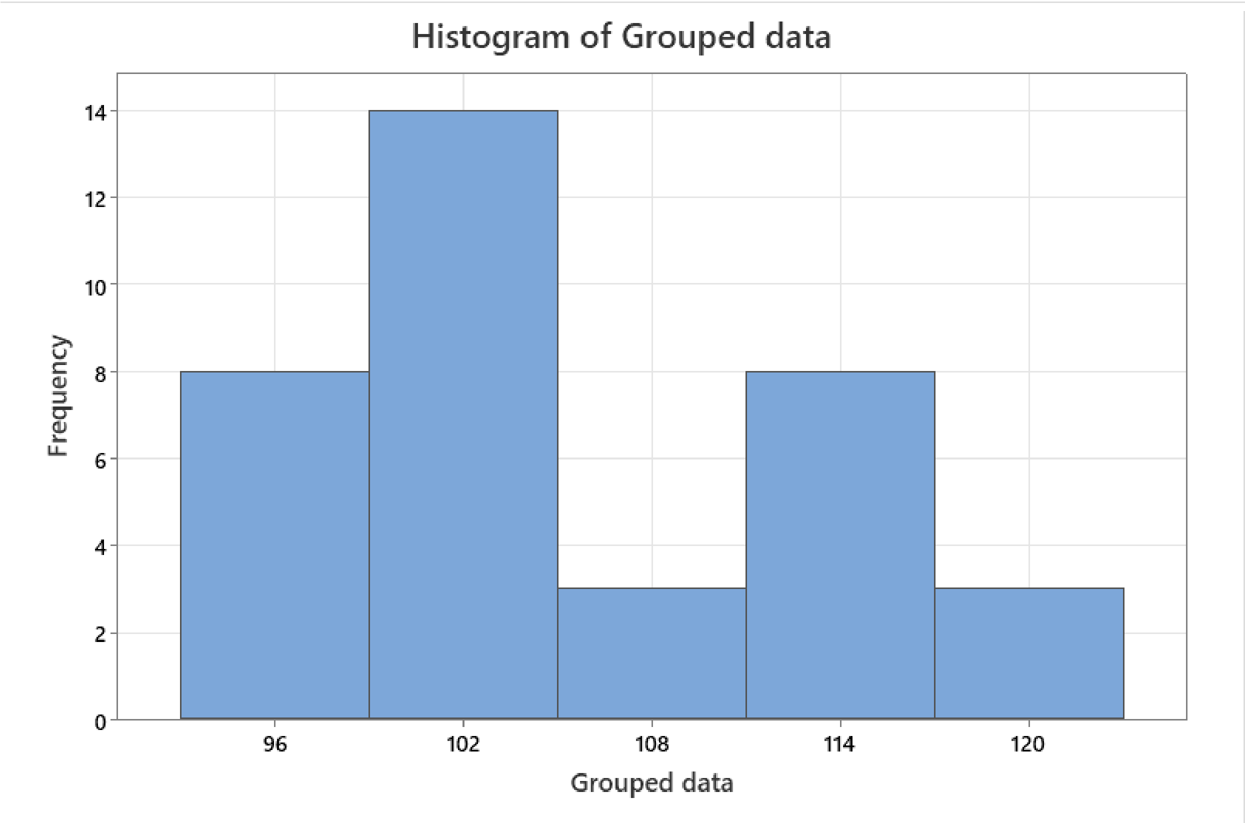
Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Grouped data	36	0	105.65	1.35	8.10	93.50	99.62	102.65	115.33	122.60

2.3 Frequency Distribution Table

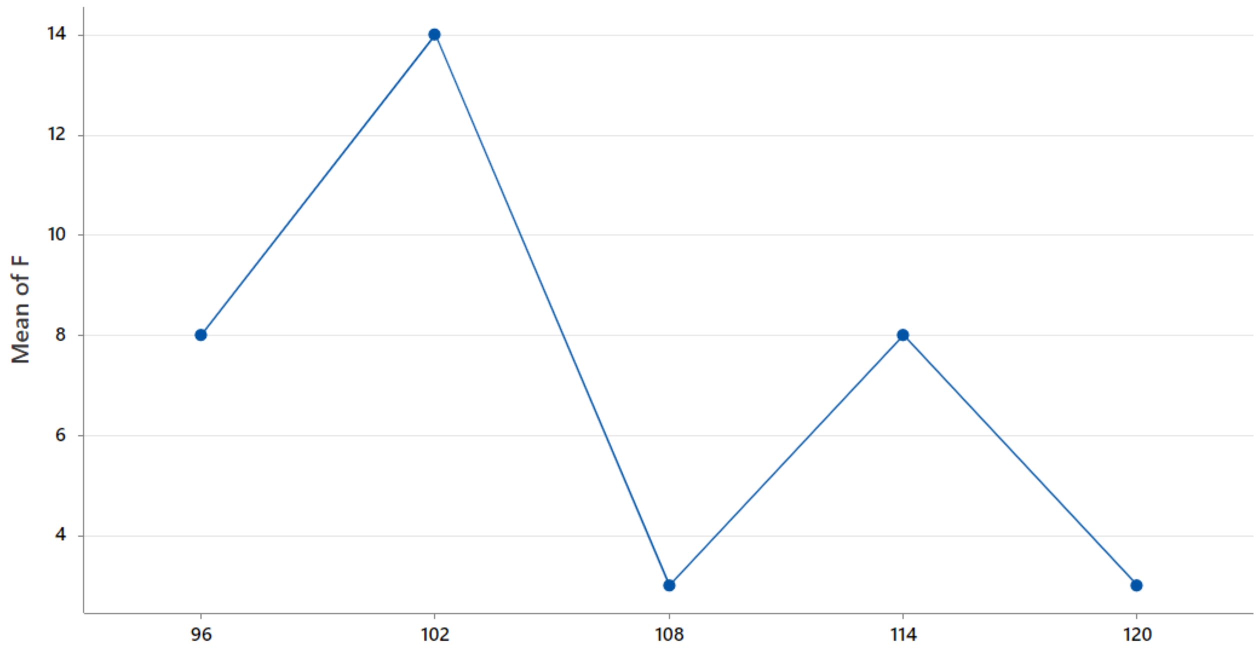
Range	Class Width	F	CF	RF	RCF	Xm	F.Xm	F.Xm^2
93.5-98.5	93-99	8	8	22.2%	22.2%	96	768	73728
99.5-104.5	99-105	14	22	38.8%	61%	102	1428	145656
105.5-110.5	105-111	3	25	8.33%	69.3%	108	324	34992
111.5-116.5	111-117	8	33	22.2%	91.5%	114	912	103968
117.5-116.5	117-123	3	36	8.33%	100%	120	360	43200
Total		36		100%			3792	401544

A frequency distribution shows how often each different value in a set of data occurs. A histogram is the most commonly used graph to show frequency distributions. It looks very much like a bar chart, but there are important differences between them. This helpful data collection and analysis tool is considered one of the seven basic quality tools.

2.3.1 Histogram

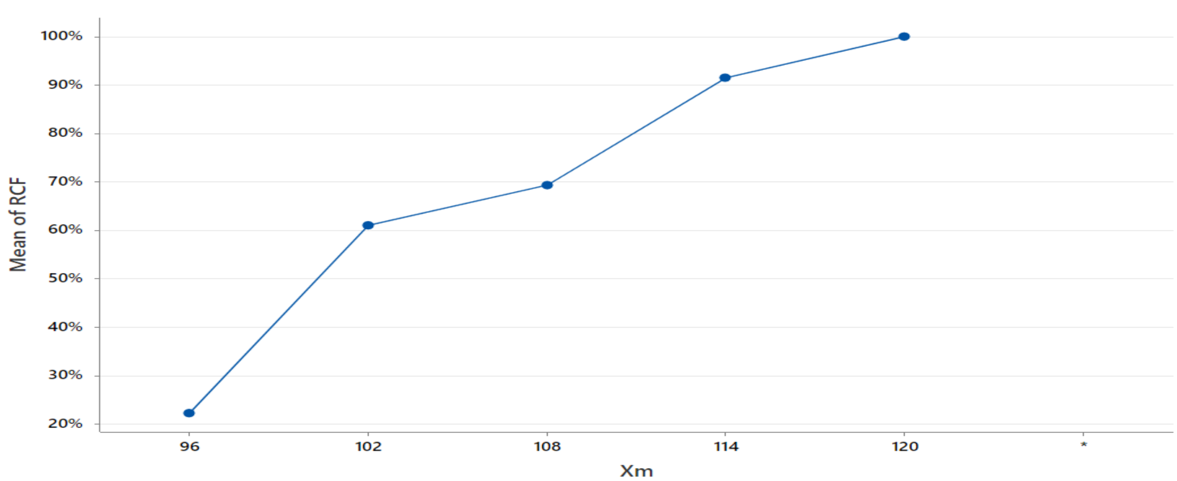


2.3.2 Frequency Polygon



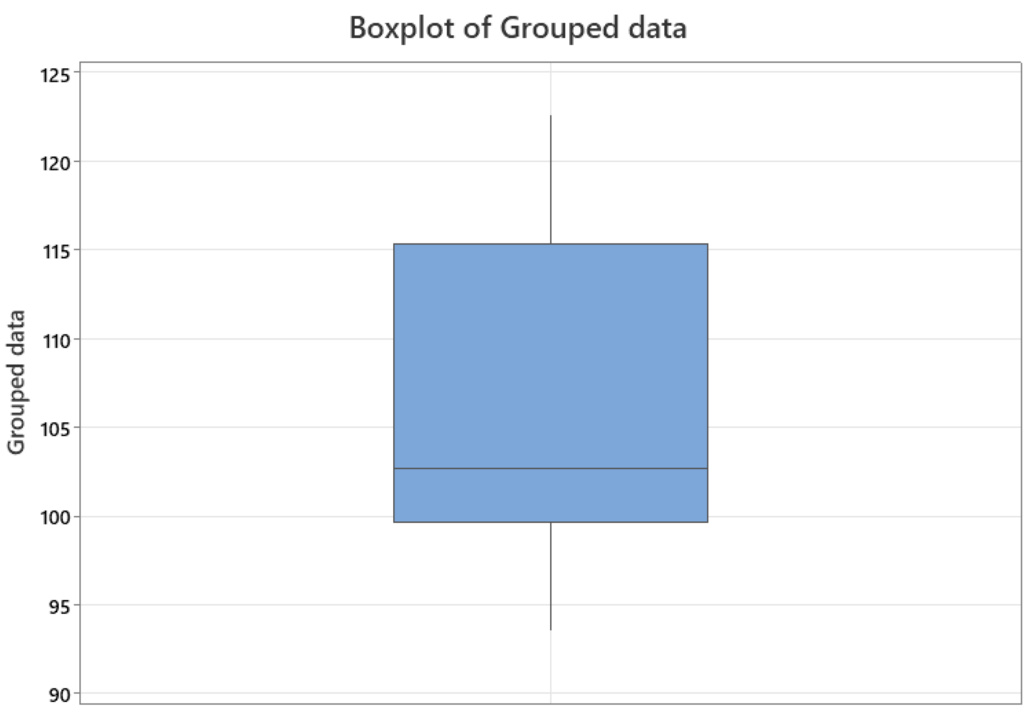
2.3.3 Ogive Curve

The Ogive is a graph of a cumulative distribution, which explains data values on the horizontal plane axis and either the cumulative relative frequencies, the cumulative frequencies or cumulative per cent frequencies on the vertical axis.



2.3.4 Box Plot

Boxplot is a method for graphically demonstrating the locality, spread and skewness groups of numerical data through their quartiles.[1] In addition to the box on a box plot, there can be lines (which are called whiskers) extending from the box indicating variability outside the upper and lower quartiles, thus, the plot is also termed as the box-and-whisker plot and the box-and-whisker diagram. Outliers that differ significantly from the rest of the dataset[2] may be plotted as individual points beyond the whiskers on the box-plot.



2.4 Mean Estimation

Data	N	Mean	Standard deviation	Degrees of freedom
Chocolate vs BP	36	105.65	8.1	36-1=35

1) Two Sided Interval
   a) 80% α=0.2 $\alpha/2$=0.1 t_{Value}=1.31

      $103.48 \leq \mu \leq 107.26$

      At a confidence level of 80% the population mean is between 103.48 and 107.26

   b) 95% α=0.05 $\alpha/2$=0.025 t_{Value}=2.04

      $102.4 \leq \mu \leq 108.31$

      At a confidence level of 95% the population mean is between 102.4and 108.31

   c) 99% α=0.01 $\alpha/2$=0.005 t_{Value}=2.75

      $101.41 \leq \mu \leq 109.33$

      At a confidence level of 99% the population mean is between 101.41 and 109.33

2) Lower Bound
   a) 99% α=0.05 t_{Value}=1.697

      $103.48 \leq \mu$

      At a confidence level of 99% the he population mean is above 103.48

2.5 Variance Estimation

1) Lower Bound
   a) 99% α=0.01 χ_{alpha} =50.89

      $51.34 \leq \sigma^2 \longrightarrow 7.17 \leq \sigma$

      At a confidence level of 99% the population variance is above 51.34

2) Two Sided Interval
   a) 95% α=0.05 $\alpha/2$=0.025 1- $\alpha/2$ = 0.975 χ_{alpha} =43.77 $\chi_{1\text{-}alpha}$=16.79

      $51.34 \leq \sigma^2 \leq 155.61 \longrightarrow 7.17 \leq \sigma \leq 12.47$

      At a confidence level of 95% the population variance is between 51.34 and 155.61

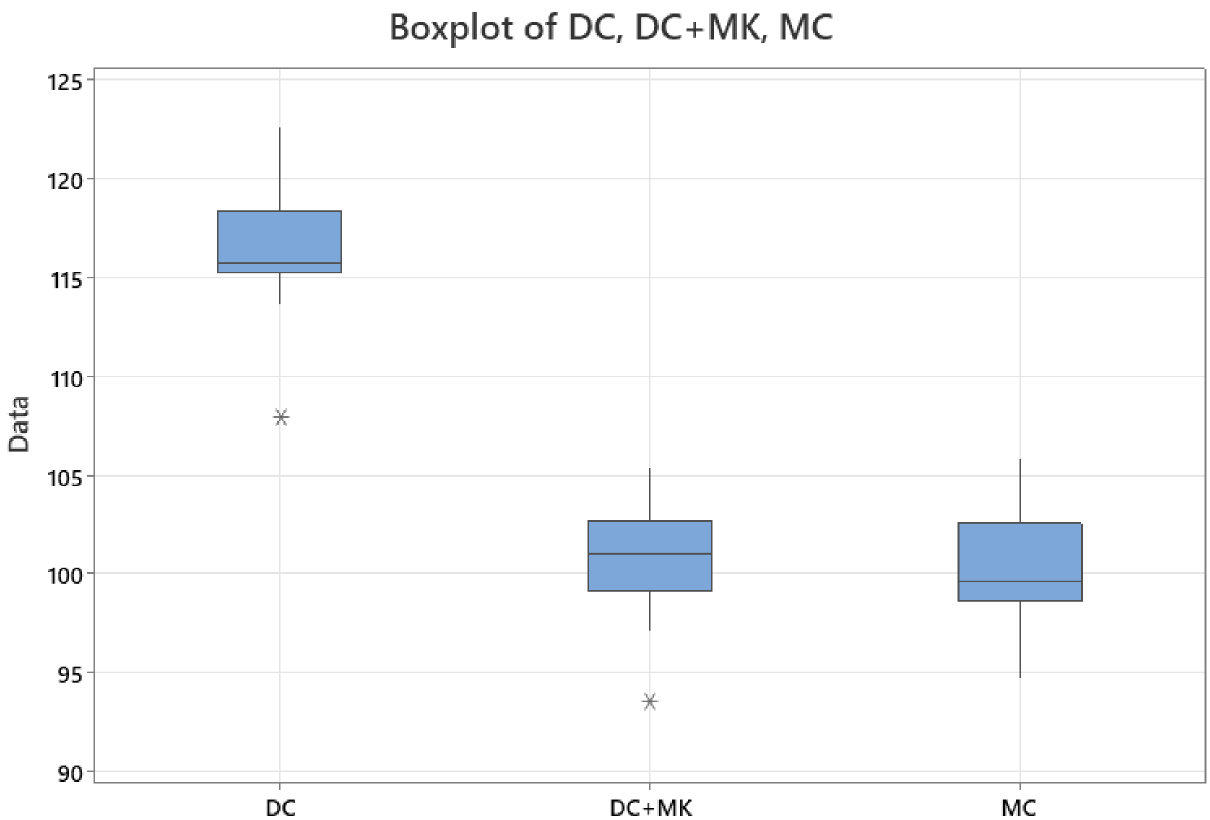
Chapter 3

Data separated into Treatments

3.1 Measures Of Central Tendency

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
DC	12	24	116.06	1.02	3.53	107.90	115.17	115.70	118.32	122.60
DC+MK	12	24	100.70	0.934	3.24	93.50	99.13	101.00	102.68	105.40
MC	12	24	100.18	0.834	2.89	94.70	98.62	99.65	102.55	105.80

3.2 Box Plots



3.3 Anova

Analysis of variance, or ANOVA, is a statistical method that separates observed variance data into different components to use for additional tests. A one-way ANOVA is used for three or more groups of data, to gain information about the relationship between the dependent and independent variables.

One-way ANOVA: DC, DC+MK, MC

Method

Null hypothesis	All means are equal
Alternative hypothesis	Not all means are equal
Significance level	$\alpha = 0.05$

Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
Factor	3	DC, DC+MK, MC

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	2	1987.1	993.53	89.12	0.000
Error	33	367.9	11.15		
Total	35	2354.9			

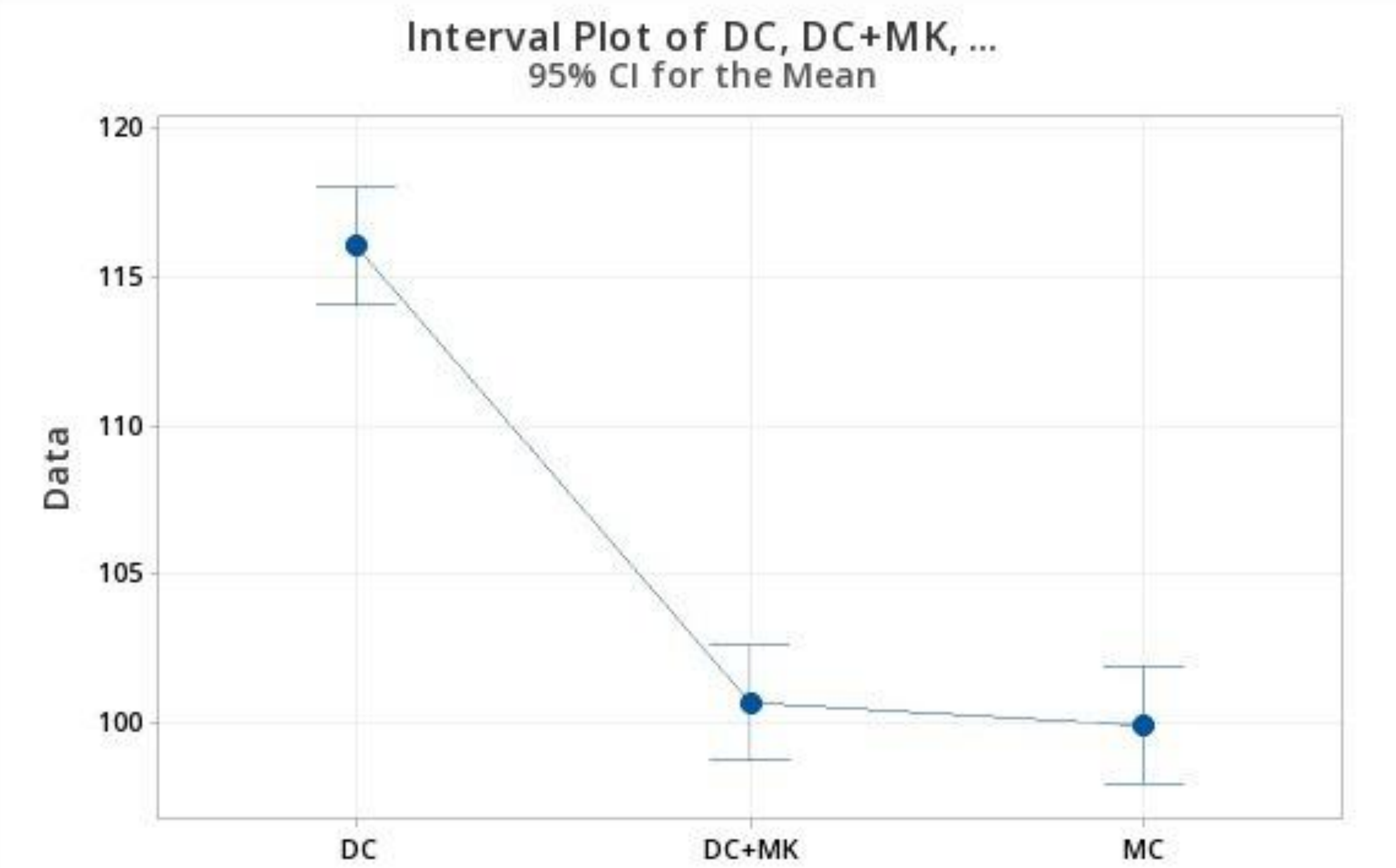
Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
3.33890	84.38%	83.43%	81.41%

Means

Factor	N	Mean	StDev	95% CI
DC	12	116.06	3.53	(114.10, 118.02)
DC+MK	12	100.700	3.235	(98.739, 102.661)
MC	12	99.925	3.240	(97.964, 101.886)

Pooled StDev = 3.33890



The pooled standard deviation is used to calculate the intervals.

3.4 Post Anova

3.4.1 Fisher (LSD) TEST

$t_{(0.025,33)}=2.042$ $MS_E=10.43$ $n = 12$

$X_{DC}=116.06$ $X_{DC+MK}=100.7$ $X_{MC}=100.183$

$LSD = t\sqrt{\frac{2*MS_E}{n}} = 2.69$

$X_{DC} - X_{DC+MK} = 15.36 > LSD$ $(\mu_{DC} \neq \mu_{DC+MK})$

$X_{DC} - X_{MC} = 15.88 > LSD$ $(\mu_{DC} \neq \mu_{DC+MK})$

$X_{DC+MK} - X_{MC} = 0.52 < LSD$ $(\mu_{DC+MK} = \mu_{MC})$

3.4.2 The Scheffè Test:

$S_w^2=10.43$, C.V = 93.58 ,a=3, n = 12

$F' = (a-1)*C.V = 187.16$

$$F_{S1} = \frac{(X_{DC}-X_{DC+MK})^2}{S_w{}^2\times(\frac{1}{n_{DC}}+\frac{1}{n_{DC+MK}})} = 135.72 < F'$$

The Scheffè Test shows
<u>no Significant Difference</u>

$$F_{S2} = \frac{(X_{DC}-X_{MC})^2}{S_w{}^2\times(\frac{1}{n_{DC}}+\frac{1}{n_{MC}})} = 60.42 < F'$$

$$F_{S3} = \frac{(X_{DC+MK}-X_{MC})^2}{S_w{}^2\times(\frac{1}{n_{DC+MK}}+\frac{1}{n_{MC}})} = 0.064 < F'$$

3.5 Tukey Test

d.f = 33 K = 3 $q_{critical}$=3.49

$$q_1 = \frac{(X_{DC} - X_{DC+MK})^2}{\sqrt{\frac{S_W{}^2}{n}}}$$ = 16.48> $q_{critical}$ (Significant Difference)

$$q_2 = \frac{(X_{DC} - X_{MC})^2}{\sqrt{\frac{S_W{}^2}{n}}}$$ = 17.03> $q_{critical}$ (Significant Difference)

$$q_3 = \frac{(X_{DC+MK} - X_{MC})^2}{\sqrt{\frac{S_W{}^2}{n}}}$$ = 0.55< $q_{critical}$ (no Significant Difference)

Chapter 4

Analysis and Conclusion

The Data collected from the experiment shows that consuming dark chocolate is beneficial and increases antioxidants in blood plasma

Chapter 5

Recommendation and Future Work

Recommendations for future work is to attempt the study on larger samples and further explore the effect of dark chocolate in the blood

Chapter 5 Work on Paper

Grouped data: Range $= 122.6 - 93.5 = 29.1$

Class width $= \frac{29.1}{5} \simeq 6$

Range	class width	tally	F	CF	RF	RCF	Xm	F·Xm	F·Xm²
93.5 - 98.5	93 - 99	𝍸 |||	8	8	22.2%	22.2%	96	768	73728
99.5 - 104.5	99 - 105	𝍸 𝍸 ||||	14	22	38.8%	61%	102	1428	145656
105.5 - 110.5	105 - 111	|||	3	25	8.33%	69.3%	108	324	34992
111.5 - 116.5	111 - 117	𝍸 |||	8	33	22.2%	91.5%	114	912	103968
117.5 - 122.5	117 - 123	|||	3	36	8.33%	100%	120	360	43200
Total			36		100%			3792	401544

$\Rightarrow$ Mean $(\bar{x}) = \frac{\Sigma F \cdot X_m}{n} = \frac{3792}{36} = 105.3$

$\Rightarrow$ Median $= L_1 + \left(\frac{\frac{\Sigma F}{2} - (\Sigma f)}{f_{median}} \right) \times C$

$= 99 + \left(\frac{18 - 8}{14} \right) \times 6 = 103.2$

$\Rightarrow$ Mode $= L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) \times C$

$= 99 + \left(\frac{(14-8)}{(14-8)+(14-3)} \right) \times 6 = 101.11$

$\Rightarrow S^2 = \frac{\Sigma F \cdot X_m^2 - \left[\frac{(\Sigma F \cdot X_m)^2}{n} \right]}{n-1} = \frac{401544 - \left(\frac{(3792)^2}{36} \right)}{35} = 60.5$

$S = \sqrt{60.5} = 7.78$

Mean and Varience Estimation

For this step we grouped All the data

Data	N	Mean	Standard Deviation	d.f
Chocb vs BP	36	105.37	8.64	35

Mean Estimation

i) Two sided Interval

a) 80% $\alpha = 0.2$ $\frac{\alpha}{2} = 0.1$ $t = 1.31$

$$103.48 \leq \mu \leq 107.26$$

b) 95% $\alpha = 0.05$ $\frac{\alpha}{2} = 0.025$ $t = 2.04$

$$102.43 \leq \mu \leq 108.31$$

c) 99% $\alpha = 0.01$ $\frac{\alpha}{2} = 0.005$ $t = 2.75$

$$101.41 \leq \mu \leq 109.33$$

ii) Lower Confidence bound

a) 95% $\alpha = 0.05$ $t = 1.697$

$$102.93 \leq \mu$$

Varience Estimation

i) Lower Bound

a) 99% $\alpha = 0.01$ $\chi = 50.89$

$$51.344 \leq \sigma^2 \rightarrow 7.17 \leq \sigma$$

ii) Two Sided interval

a) 95% $\alpha = 0.05$ $\frac{\alpha}{2} = 0.025$ $1 - \frac{\alpha}{2} = 0.975$

$\chi_{\alpha} = 43.77$ $\chi_{1-\alpha} = 16.79$

$$59.69 \leq \sigma^2 \leq 155.61 \rightarrow 7.72 \leq \sigma \leq 12.47$$

Post Anova: Fisher test (LSD)

$$LSD = t_{(0.025,\,36-3)}\sqrt{\frac{2(10.43)}{12}}$$

$$= 2.042 \times 1.318 = 2.69$$

$\bar{X}_{DC} = 116.06$, $\bar{X}_{DC+MK} = 100.7$, $\bar{X}_{MC} = 100.183$

$\bar{X}_{DC} - \bar{X}_{DC+MK} = 15.36 > LSD \; (\mu_{DC} \neq \mu_{DC+MK})$

$\bar{X}_{DC} - \bar{X}_{MC} = 15.877 > LSD \; (\mu_{DC} \neq \mu_{MC})$

$\bar{X}_{DC+MK} - \bar{X}_{MC} = 0.517 < LSD \; (\mu_{DC+MK} = \mu_{MC})$

The scheffé Test:

$S_w^2 = 10.43$, $C.V = 93.58$, $a = 3$

$f' = 2 \times 93.58 = 187.16$

$$F_{S_1} = \frac{(15.36)^2}{10.43 \times 2 \times 1/5} = 56.55 < f'$$

for Scheffe' test There is no Significant diffrence

$$F_{S_2} = \frac{(15.877)^2}{10.43 \times 2 \times 1/5} = 60.42 < f'$$

$$F_{S_3} = \frac{(0.517)^2}{10.43 \times 2 \times 1/5} = 0.064 < f'$$

Tukey:

$$q_1 = \frac{15.36}{\sqrt{10.43/12}} = 16.48$$

$k = 3$

$df = 33$

$$q_2 = \frac{15.88}{\sqrt{10.43/12}} = 17.03$$

$q_{critical} = 3.49$

from (q) tables for tukey tests found online

$$q_3 = \frac{0.517}{\sqrt{10.43/12}} = 0.55$$

$q_1 > q_c$ There is significant diffrence

$q_2 > q_c$ There is significant diffrence

$q_3 < q_c$ There is no significant diffrence

Chapter 6 References

[1] R. B. M. S. V. D. S. C. Mauro Serafini*, "Plasma antioxidants from chocolate," in *2003 Nature Publishing Group*, Glasgow, 2003.

[2] D. C. Montgomery, Design and Analysis of Experiments, 8th.