NAME: Gayatri Walke

Roll No.: 3474

Assignment 8: Case Study

Problem Statement: To study effect of various factors such as Gender, Age, height. Pre-weight on Weight Loss Dataset:

Person	gender	Age	Height	pre.weight	Diet	weight6weeks
25	0	41	171	60	2	60
26	0	32	174	103	2	103
1	0	22	159	58	1	54.2
2	0	46	192	60	1	54
3	0	55	170	64	1	63.3
4	0	33	171	64	1	61.1
5	0	50	170	65	1	62.2
6	0	50	201	66	1	64
7	0	37	174	67	1	65
8	0	28	176	69	1	60.5
9	0	28	165	70	1	68.1
10	0	45	165	70	1	66.9
11	0	60	173	72	1	70.5
12	0	48	156	72	1	69
13	0	41	163	72	1	68.4
14	0	37	167	82	1	81.1
27	0	44	174	58	2	60.1
28	0	37	172	58	2	56
29	0	41	165	59	2	57.3
30	0	43	171	61	2	56.7
31	0	20	169	62	2	55
32	0	51	174	63	2	62.4
33	0	31	163	63	2	60.3
34	0	54	173	63	2	59.4
35	0	50	166	65	2	62
36	0	48	163	66	2	64
37	0	16	165	68	2	63.8

38	0	37	167	68	2	63.3
39	0	30	161	76	2	72.7
40	0	29	169	77	2	77.5
52	0	51	165	60	3	53
53	0	35	169	62	3	56.4
54	0	21	159	64	3	60.6
55	0	22	169	65	3	58.2
56	0	36	160	66	3	58.2
57	0	20	169	67	3	61.6
58	0	35	163	67	3	60.2
59	0	45	155	69	3	61.8
60	0	58	141	70	3	63
61	0	37	170	70	3	62.7
62	0	31	170	72	3	71.1
63	0	35	171	72	3	64.4
64	0	56	171	73	3	68.9
65	0	48	153	75	3	68.7
66	0	41	157	76	3	71
15	1	39	168	71	1	71.6
16	1	31	158	72	1	70.9
17	1	40	173	74	1	69.5
18	1	50	160	78	1	73.9
19	1	43	162	80	1	71
20	1	25	165	80	1	77.6
21	1	52	177	83	1	79.1
22	1	42	166	85	1	81.5
23	1	39	166	87	1	81.9
24	1	40	190	88	1	84.5
41	1	51	191	71	2	66.8
42	1	38	199	75	2	72.6
43	1	54	196	75	2	69.2
44	1	33	190	76	2	72.5
45	1	45	160	78	2	72.7
46	1	37	194	78	2	76.3
47	1	44	163	79	2	73.6
48	1	40	171	79	2	72.9
49	1	37	198	79	2	71.1
50	1	39	180	80	2	81.4

51	1	31	182	80	2	75.7
67	1	36	155	71	3	68.5
68	1	47	179	73	3	72.1
69	1	29	166	76	3	72.5
70	1	37	173	78	3	77.5
71	1	31	177	78	3	75.2
72	1	26	179	78	3	69.4
73	1	40	179	79	3	74.5
74	1	35	183	83	3	80.2
75	1	49	177	84	3	79.9
76	1	28	164	85	3	79.7
77	1	40	167	87	3	77.8
78	1	51	175	88	3	81.9

Diet data set

This data set contains information on 78 people using one of three diets. The dataset is primarily used for ANOVA.

Variable name	Variable	Data type
Person	Participant number	
gender	Gender, 1 = male, 0 = female	Binary
Age	Age (years)	Scale
Height	Height (cm)	Scale
preweight	Weight before the diet (kg)	Scale
Diet	Diet	Binary
weight10weeks	Weight after 6 weeks (kg)	Scale

Code:

```
diet =read.table("Diet_R.csv", header = TRUE, sep = ",") summary(diet)


#CALCULATE WEIGHT LOSS

diet$weight.loss = diet$pre.weight - diet$weight6weeks


# WEIGHT LOSS PLOT

plot(diet$Person,diet$pre.weight,ylab="weight",pch = 8,col="red")

points(diet$Person,diet$weight6weeks,pch = 8,col="green")


#histogram of weight loss hist(diet$weight.loss,xlab = "Weight loss",      col =

"green",border = "red", breaks = 5 ,main="weight loss")


#effects of various factors #DIET
```

```
summary(aov(weight.loss~diet$Diet,data=diet))

#BOX PLOT - DIET Which diet was best for losing weight?
boxplot(weight.loss~diet$Diet,data=diet,
      col=rainbow(3),ylab = "Weight loss (kg)", xlab = "Diet type" )

#GENDER
summary(aov(weight.loss~diet$gender,data=diet))

#BOX PLOT - GENDER Are there gender differences for weight lost?
boxplot(weight.loss~diet$gender,data=diet,        col=rainbow(2),ylab = "Weight loss
(kg)", xlab = "Gender" )

#AGE
summary(aov(weight.loss~diet$Age,data=diet)
#HEIGHT
summary(aov(weight.loss~diet$Height,data=diet))
#effect of diet and gender
summary(aov(weight.loss~diet$Diet*diet$gender,data=diet))
#effect of age and height
summary(aov(weight.loss~diet$Age*diet$Height,data=diet))

#corelation analysis library(corrplot)
corrplot(cor(diet), method="circle")

#linear regression
lifelm=lm(diet$weight.loss~diet$gender+diet$Age+diet$Height, data=df)
summary(lifelm) #height
plot(diet$Height,diet$weight.loss,pch = 8,     col="green",ylab = "Weight
loss (kg)", xlab = "height") abline(lm(diet$weight.loss~diet$Height))
#age
plot(diet$Age,diet$weight.loss,pch = 8,     col="yellow",ylab = "Weight
loss (kg)", xlab = "age") abline(lm(diet$weight.loss~diet$Age))

#Fisher's, Welch's and Kruskal-Wallis one-way ANOVA,  diet.fisher  =
aov(weight.loss~diet$Diet,data=diet)        diet.welch                    =
oneway.test(weight.loss~diet$Diet,data=diet)          diet.kruskal          =
kruskal.test(weight.loss~diet$Diet,data=diet)
```

summary(diet.fisher) print(diet.welch)

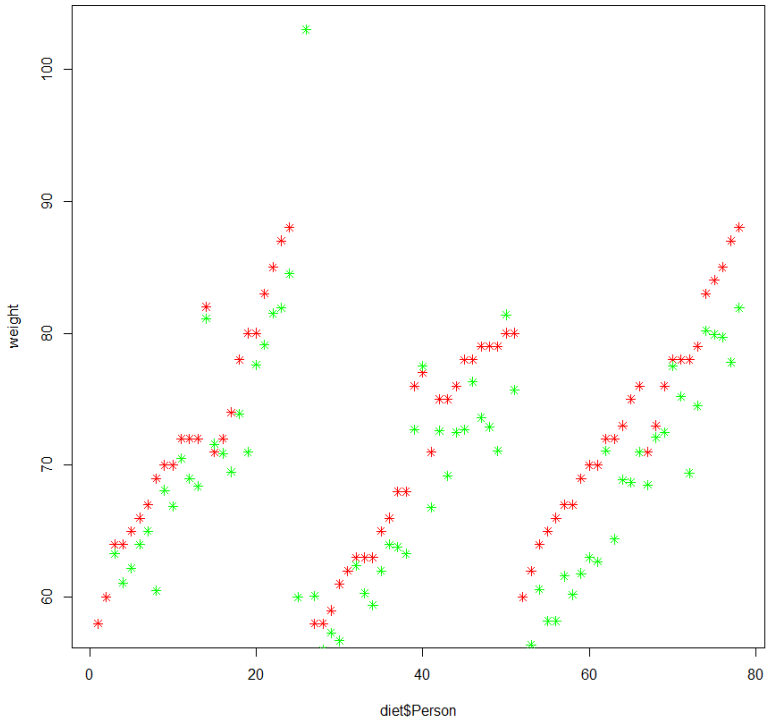
print(diet.kruskal)

Output:

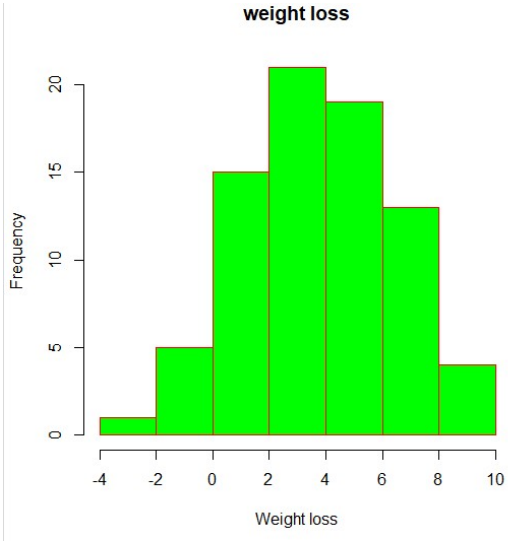
1. Summry & Weight Loss Calculation

```
> summary(diet)
     Person          gender           Age            Height        pre.weight          Diet
 Min.   : 1.00   Min.   :0.0000   Min.   :16.00   Min.   :141.0   Min.   : 58.00   Min.   :1.000
 1st Qu.:20.25   1st Qu.:0.0000   1st Qu.:32.25   1st Qu.:164.2   1st Qu.: 66.00   1st Qu.:1.000
 Median :39.50   Median :0.0000   Median :39.00   Median :169.5   Median : 72.00   Median :2.000
 Mean   :39.50   Mean   :0.4231   Mean   :39.15   Mean   :170.8   Mean   : 72.53   Mean   :2.038
 3rd Qu.:58.75   3rd Qu.:1.0000   3rd Qu.:46.75   3rd Qu.:174.8   3rd Qu.: 78.00   3rd Qu.:3.000
 Max.   :78.00   Max.   :1.0000   Max.   :60.00   Max.   :201.0   Max.   :103.00   Max.   :3.000
  weight6weeks
 Min.   : 53.00
 1st Qu.: 61.85
 Median : 68.95
 Mean   : 68.68
 3rd Qu.: 73.83
 Max.   :103.00
> #CALCULATE WEIGHT LOSS
> diet$weight.loss = diet$pre.weight - diet$weight6weeks
> # WEIGHT LOSS PLOT
> plot(diet$Person,diet$pre.weight,ylab="weight",pch = 8,col="red")
> points(diet$Person,diet$weight6weeks,pch = 8,col="green")
> summary(diet)
     Person          gender           Age            Height        pre.weight          Diet
 Min.   : 1.00   Min.   :0.0000   Min.   :16.00   Min.   :141.0   Min.   : 58.00   Min.   :1.000
 1st Qu.:20.25   1st Qu.:0.0000   1st Qu.:32.25   1st Qu.:164.2   1st Qu.: 66.00   1st Qu.:1.000
 Median :39.50   Median :0.0000   Median :39.00   Median :169.5   Median : 72.00   Median :2.000
 Mean   :39.50   Mean   :0.4231   Mean   :39.15   Mean   :170.8   Mean   : 72.53   Mean   :2.038
 3rd Qu.:58.75   3rd Qu.:1.0000   3rd Qu.:46.75   3rd Qu.:174.8   3rd Qu.: 78.00   3rd Qu.:3.000
 Max.   :78.00   Max.   :1.0000   Max.   :60.00   Max.   :201.0   Max.   :103.00   Max.   :3.000
  weight6weeks     weight.loss
 Min.   : 53.00   Min.   :-2.100
 1st Qu.: 61.85   1st Qu.: 2.000
 Median : 68.95   Median : 3.600
 Mean   : 68.68   Mean   : 3.845
 3rd Qu.: 73.83   3rd Qu.: 5.550
 Max.   :103.00   Max.   : 9.200
```

2. Weight Loss Plot



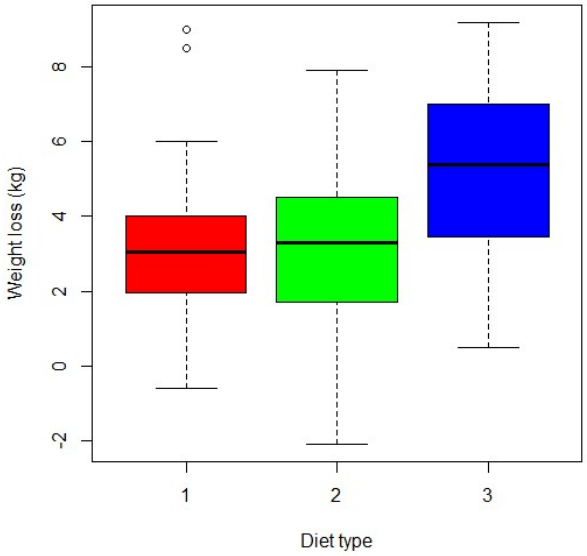
3. Histogram of Weight Loss



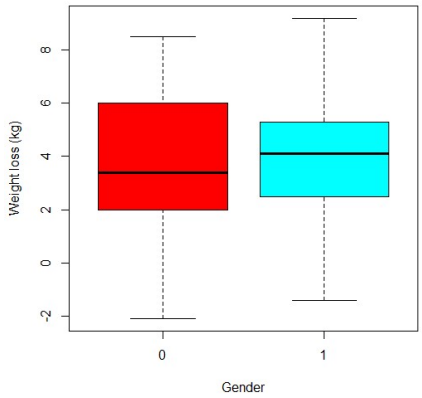
4. Effect of Diet

```
> summary(aov(weight.loss~diet$Diet,data=diet))
             Df Sum Sq Mean Sq F value  Pr(>F)
diet$Diet     1   45.8   45.78   7.639 0.00716 **
Residuals    76  455.5    5.99
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

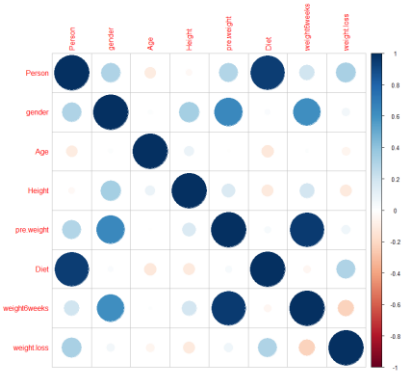
5. BOX PLOT - DIET Which diet was best for losing weight?



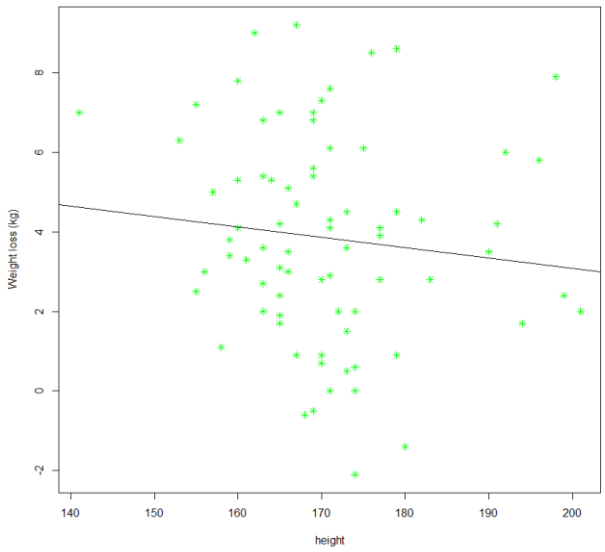
6. BOX PLOT – Gender



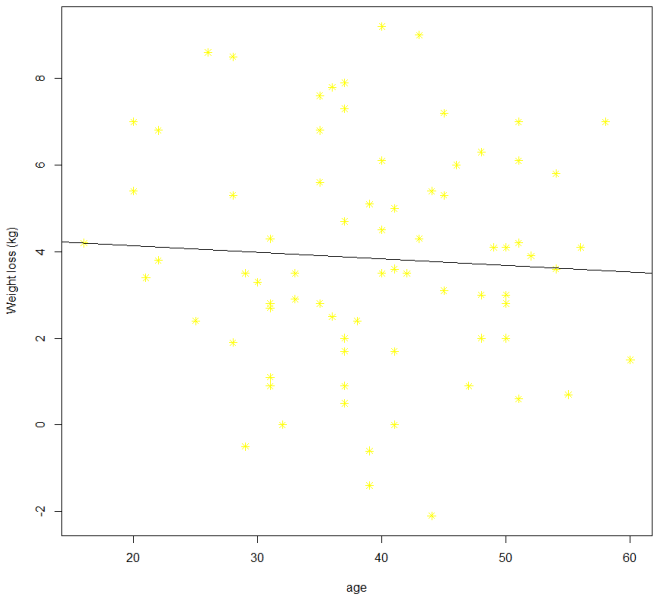
7. Correlation plot



8. Height Plot



9. Age Plot



Thus, we can infer that only diet type affects weight loss.