# Botanical_fanfeng

FAN FENG

## Summary

This report will mainly explore the relationship between valence, gender, age and the outcome variables.We think there maybe a difference in valence between genders under same condition.

Similarly, we expect to see that age will affect the pro-environmental behaviours.Investigating whether education affect pro-environmental behaviour and if it is connected to valence and various demographic variables as well as the previous proposal said.

## Introduction

To understand how education and different walking affect pro-environmental behaviour, and to understand if valence and various demographic variables are associated with pro-environmental behaviour.

We use linear regression to fit the models, and use hypothesis tests to determine if the variable has a significant linear relationship.

We seprate the dataset into within control and without control.We fit the model base on these two different dataset and use different demographic variables to see their effect on pro-environmental actions.

## Data Description

This report plot the histogram of Age and Z_beh_Score of the data set and plot the distribution of gender and valence as well.

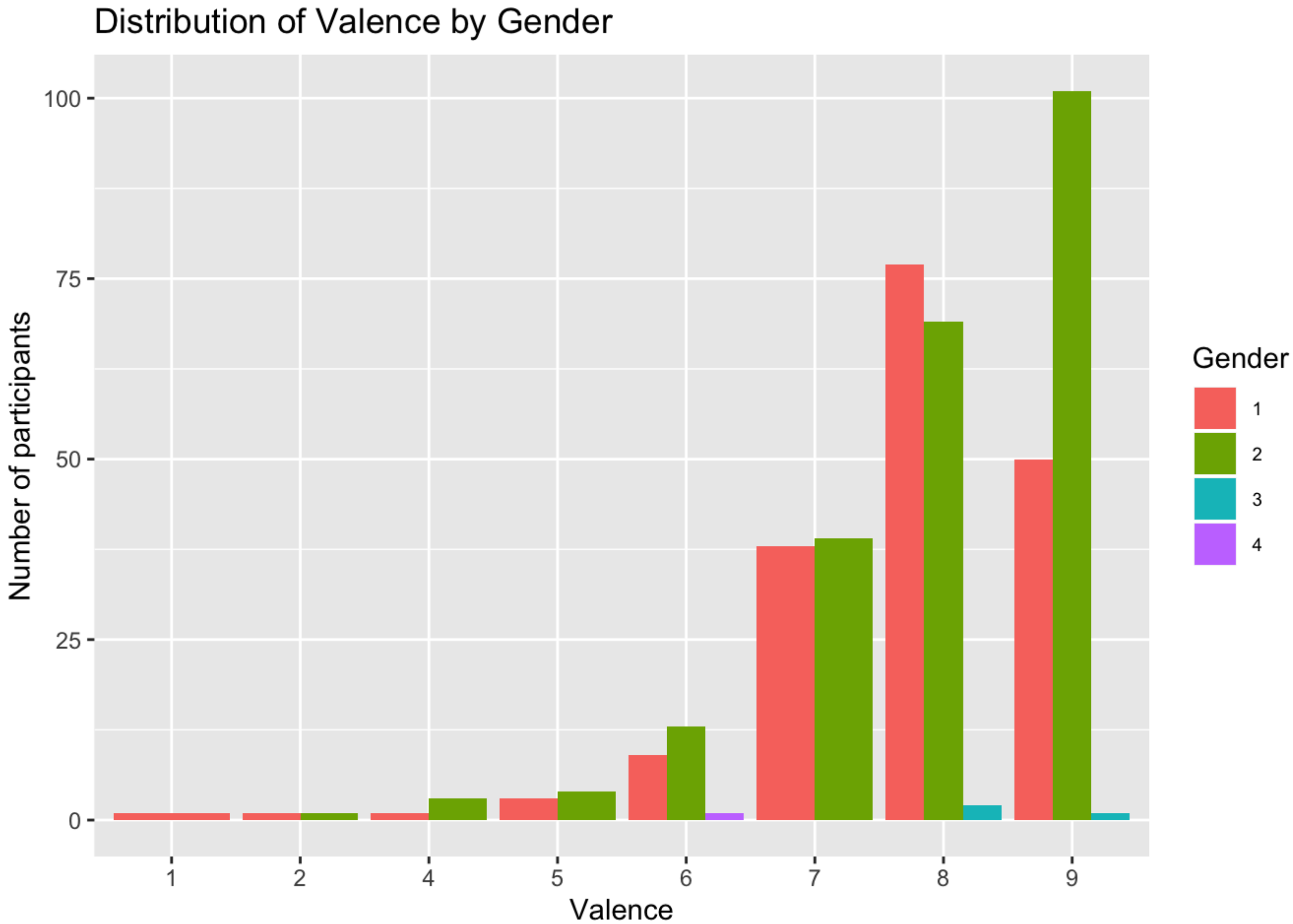
I use QQplot to check the normality of response variable and got the result of the response variable is not normal.We decided to use multiple regression instead of two way Anova which we mentioned in the proposal.

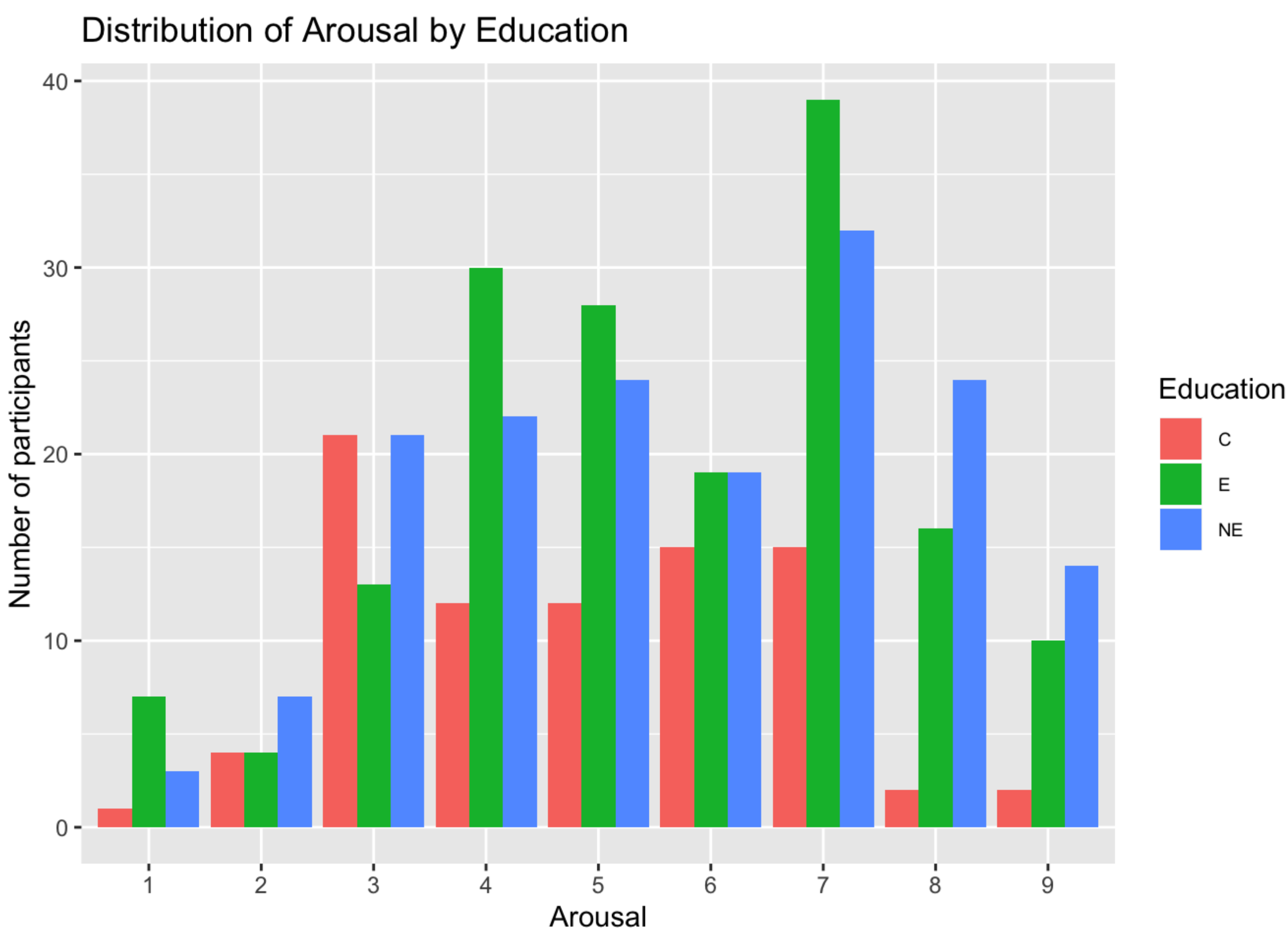
We also output the graph of distribution of arousal by education and the boxplot of it.In order to check the relationship between them.

```
## ── Attaching packages ─────────────────────────────────────────────────
── tidyverse 1.3.0 ──
```

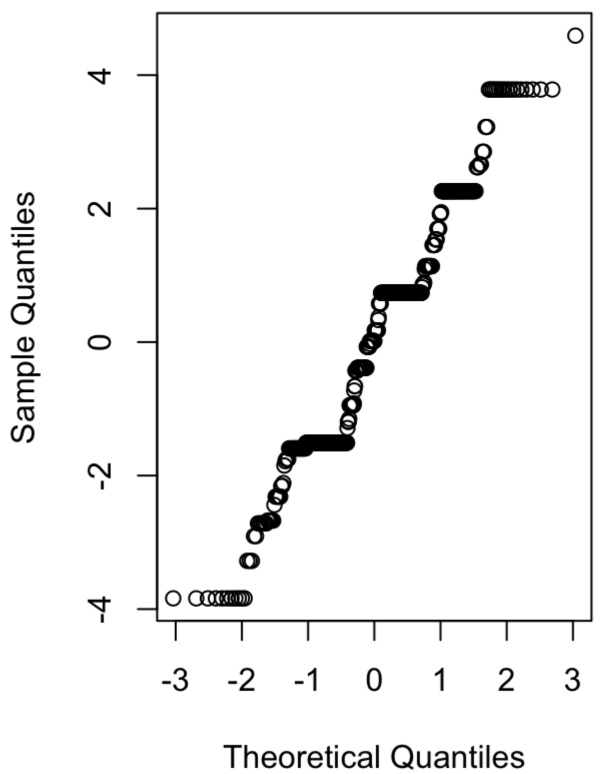
```
## ✓ ggplot2 3.2.1     ✓ purrr   0.3.3
## ✓ tibble  2.1.3     ✓ dplyr   0.8.3
## ✓ tidyr   1.0.0     ✓ stringr 1.4.0
## ✓ readr   1.3.1     ✓ forcats 0.4.0
```

```
## ── Conflicts ──────────────────────────────────────────────────────────
tidyverse_conflicts() ──
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

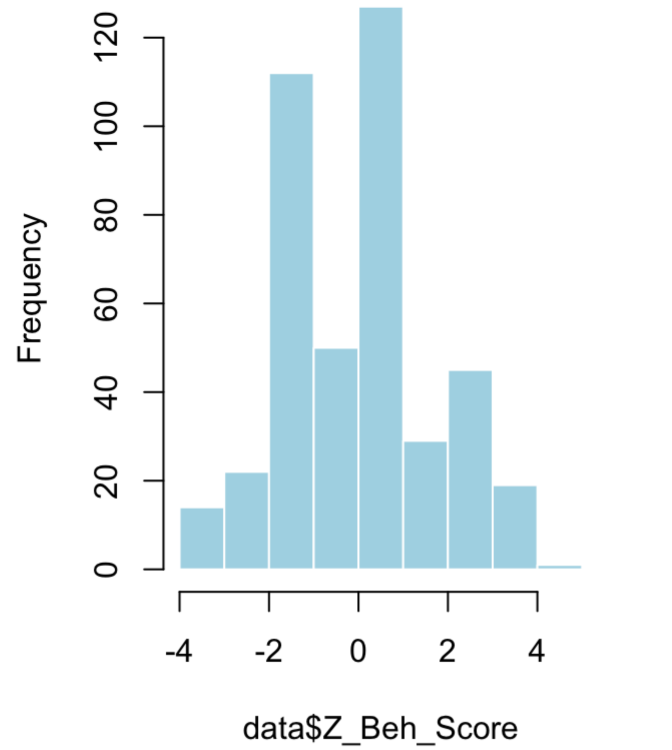

Distribution of Valence by Gender
Number of participants
0
25
50
75
100
1
2
4
5
6
7
8
9
Valence
Gender
1
2
3
4


Distribution of Arousal by Education
Number of participants
0
10
20
30
40
1
2
3
4
5
6
7
8
9
Arousal
Education
C
E
NE

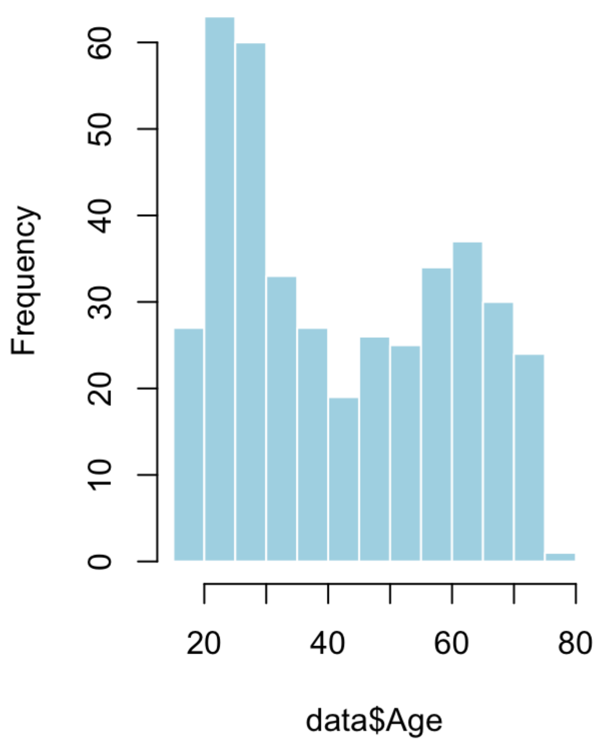
Normal Q-Q Plot

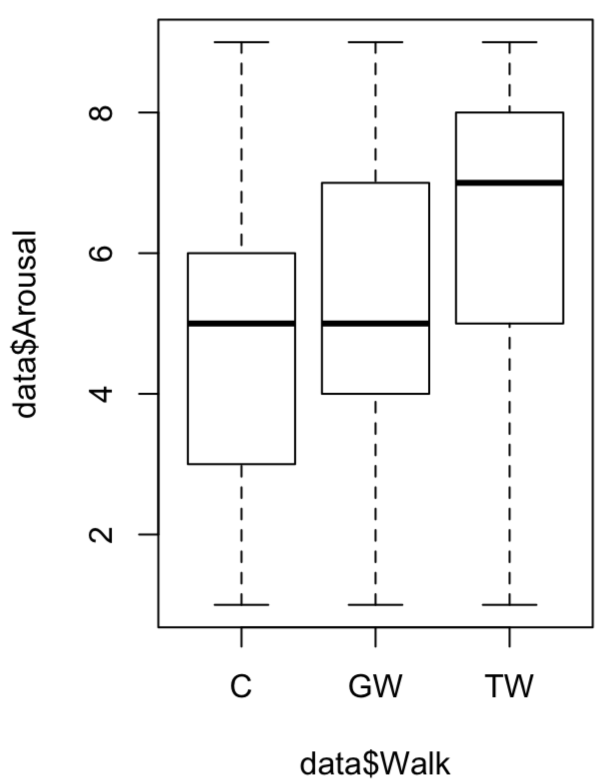



Histogram of data$Z_Beh_Score




Histogram of data$Age



# Method

Two way anova have many assumption restrict and what we checked before showed that two way anova is not suitable to use.We decided to use multiple regression.This method is to find linear relationship between our reponse variables and explanatory variables.The result of this method can tell us whether this two or more variables have significant relationship or not.

# Result

Table1 and Table2 showed that valence and political have a significant relationship with pro_envrionmental actions under conditions within the control.

Table3 and Table4 showed that valence do not have significant relationship with reponse variable and political is still significant related but the level is lower.

Table5 showed that we do not have enough evidence to say there have interaction between arousal and education.

Table6 and 7 is the combination of the table.

We use barplot to compare the effection of different walks.From the graph,we can find that tree walk do have a higher mean than ground walk.We use barplot to compare the effection of education.From the graph,we can find that people who get educated do have a higher mean than those who didn't.

We use linear model fit the walk and education seprately.The results showed that there is no significant relationship between walk and response variable and also no relationship between education and pro-environmental actions.

We can see maybe the difference showed in the boxplot is cause by the high variance.

```
#with control
fit1<- lm(Z_Beh_Score~as.factor(Condition)+Valence,data=data)
tidy(fit1) %>%knitr::kable(caption = "valence and condition under control")
```

valence and condition under control

| term | estimate | std.error | statistic | p.value |
|---|---:|---:|---:|---:|
| (Intercept) | -1.7179738 | 0.6041078 | -2.8438200 | 0.0046807 |
| as.factor(Condition)2 | 0.0585779 | 0.2708045 | 0.2163108 | 0.8288531 |
| as.factor(Condition)3 | -0.1057621 | 0.2689987 | -0.3931696 | 0.6943986 |
| as.factor(Condition)4 | 0.0653562 | 0.2639639 | 0.2475953 | 0.8045715 |
| as.factor(Condition)5 | -0.1087796 | 0.2665257 | -0.4081394 | 0.6833844 |
| Valence | 0.2179693 | 0.0728480 | 2.9921113 | 0.0029379 |

```
fit2<- lm(Z_Beh_Score~as.factor(Condition)+Political,data=data)
tidy(fit2) %>%knitr::kable(caption = "political and condition under control")
```

political and condition under control

| term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|
| (Intercept) | 0.4859961 | 0.2497795 | 1.9457006 | 0.0524181 |
| as.factor(Condition)2 | 0.0895763 | 0.2758166 | 0.3247676 | 0.7455334 |
| as.factor(Condition)3 | 0.0020687 | 0.2737245 | 0.0075577 | 0.9939738 |
| as.factor(Condition)4 | 0.0875259 | 0.2692667 | 0.3250530 | 0.7453175 |
| as.factor(Condition)5 | -0.1147560 | 0.2738819 | -0.4189982 | 0.6754511 |
| Political | -0.1771803 | 0.0564881 | -3.1365954 | 0.0018406 |

```
#without control
fit4<- lm(Z_Beh_Score~as.factor(Walk)+as.factor(Education)+Valence,data=data_m)
tidy(fit4) %>%knitr::kable(caption = "valence and condition without control")
```

valence and condition without control

| term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|
| (Intercept) | -1.2102553 | 0.6518981 | -1.8565101 | 0.0642774 |
| as.factor(Walk)TW | 0.0328684 | 0.1903667 | 0.1726584 | 0.8630264 |
| as.factor(Education)NE | 0.1049089 | 0.1898476 | 0.5525955 | 0.5809168 |
| Valence | 0.1456777 | 0.0779643 | 1.8685186 | 0.0625811 |

```
fit5<- lm(Z_Beh_Score~as.factor(Condition)+Political,data=data_m)
tidy(fit5) %>%knitr::kable(caption = "political and condition without control")
```

political and condition without control

| term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|
| (Intercept) | 0.4342026 | 0.2612577 | 1.6619710 | 0.0975327 |
| as.factor(Condition)2 | 0.0846631 | 0.2751881 | 0.3076555 | 0.7585518 |
| as.factor(Condition)3 | -0.0011345 | 0.2730437 | -0.0041549 | 0.9966875 |
| as.factor(Condition)4 | 0.0870084 | 0.2685553 | 0.3239869 | 0.7461672 |
| Political | -0.1589271 | 0.0627976 | -2.5307848 | 0.0118768 |

```
#interaction between arousal and education
fit6<- lm(Z_Beh_Score~as.factor(Walk)*as.factor(Education),data=data_m)
tidy(fit6) %>%knitr::kable(caption = "interaction between walk and education")
```

interaction between walk and education

| term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|
| (Intercept) | -0.0635679 | 0.1909783 | -0.3328540 | 0.7394559 |
| as.factor(Walk)TW | 0.0637050 | 0.2676904 | 0.2379800 | 0.8120443 |
| as.factor(Education)NE | 0.1471885 | 0.2661753 | 0.5529757 | 0.5806545 |
| as.factor(Walk)TW:as.factor(Education)NE | -0.1229167 | 0.3787074 | -0.3245690 | 0.7457130 |

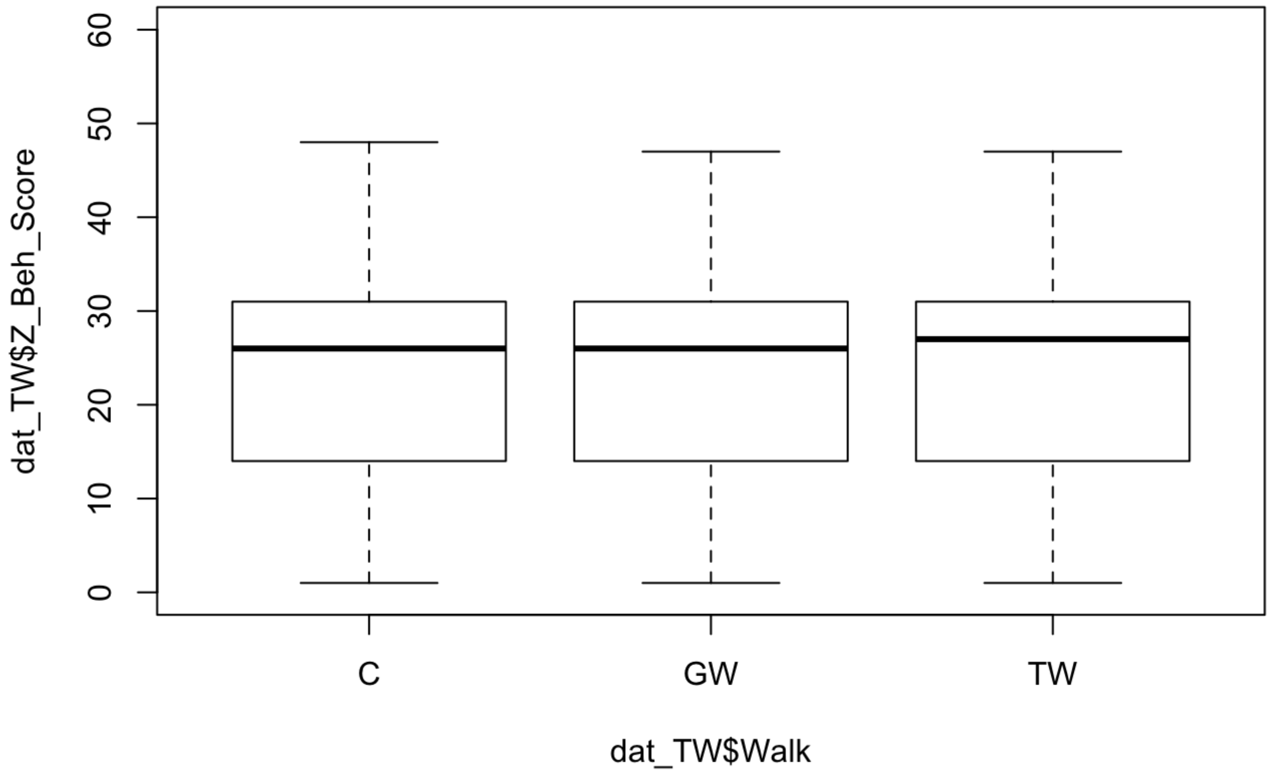
```
#full model without control
fitFull<- lm(Z_Beh_Score~as.factor(Condition)+Valence+Age+Political,data=data_m)
tidy(fitFull)
```

```
## # A tibble: 7 x 5
##   term                  estimate std.error statistic p.value
##   <chr>                    <dbl>     <dbl>     <dbl>   <dbl>
## 1 (Intercept)           -0.415     0.710      -0.585  0.559
## 2 as.factor(Condition)2  0.0674    0.277       0.244  0.808
## 3 as.factor(Condition)3  0.0674    0.276       0.244  0.807
## 4 as.factor(Condition)4  0.0695    0.268       0.260  0.795
## 5 Valence                0.125     0.0807      1.55   0.121
## 6 Age                   -0.00360   0.00559    -0.644  0.520
## 7 Political             -0.158     0.0639     -2.47   0.0142
```

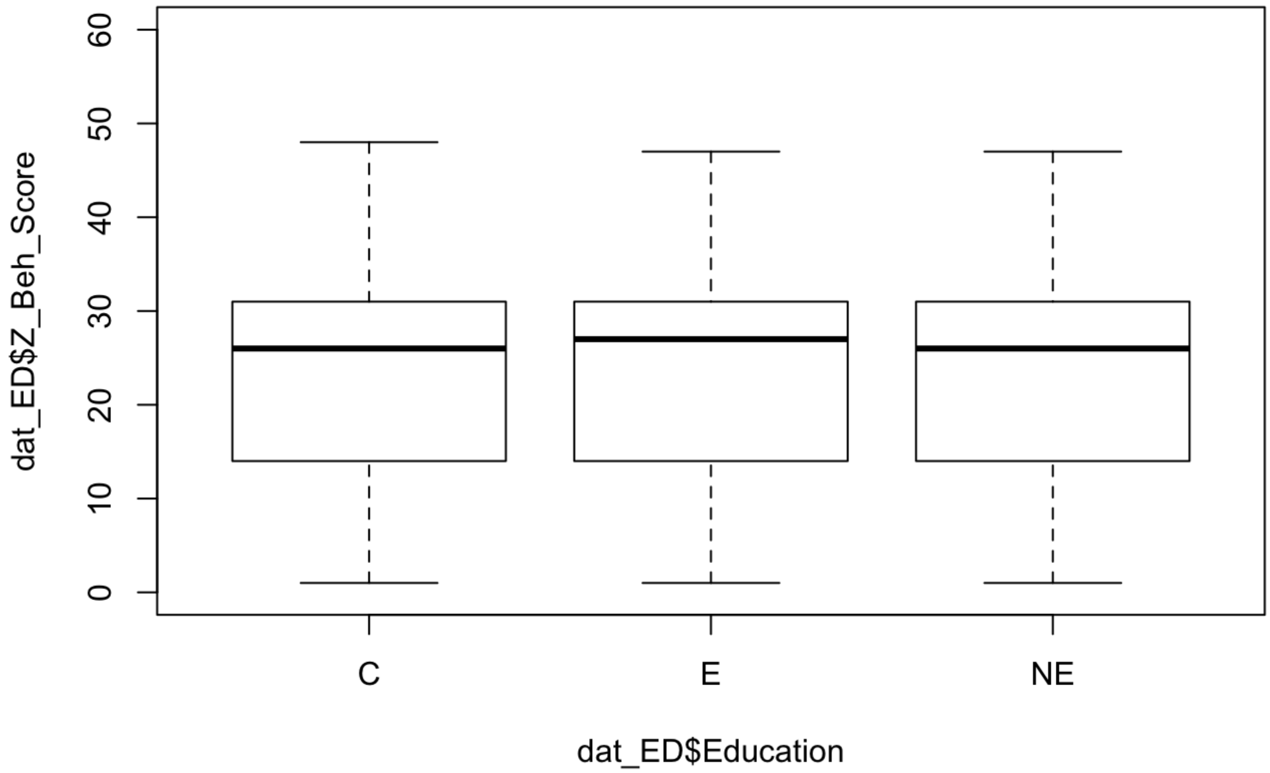
```
#full model with control
fitFull2<- lm(Z_Beh_Score~as.factor(Condition)+Valence+Age+Political,data=data)
tidy(fitFull2)
```

```
## # A tibble: 8 x 5
##   term                  estimate std.error statistic p.value
##   <chr>                    <dbl>     <dbl>     <dbl>   <dbl>
## 1 (Intercept)           -0.973     0.661      -1.47  0.141
## 2 as.factor(Condition)2  0.0888    0.276       0.322 0.748
## 3 as.factor(Condition)3  0.0472    0.275       0.172 0.864
## 4 as.factor(Condition)4  0.0625    0.267       0.234 0.815
## 5 as.factor(Condition)5 -0.103     0.272      -0.378 0.706
## 6 Valence                0.195     0.0749      2.61  0.00944
## 7 Age                   -0.00251   0.00500    -0.501 0.617
## 8 Political             -0.172     0.0570     -3.02  0.00272
```

```
dat_TW <- data %>% select(Walk, Z_Beh_Score) %>% mutate(Walk = as.factor(Walk), Z_Beh_Sc
ore = as.factor(Z_Beh_Score)) %>%
  na.omit()
dat_TW$Z_Beh_Score<-as.integer(dat_TW$Z_Beh_Score)
boxplot(dat_TW$Z_Beh_Score~dat_TW$Walk,ylim=c(0,60))
```




```
dat_ED <- data%>% select(Education, Z_Beh_Score) %>% mutate(Education = as.factor(Educat
ion), Z_Beh_Score = as.factor(Z_Beh_Score)) %>%
  na.omit()
dat_ED$Z_Beh_Score<-as.integer(dat_ED$Z_Beh_Score)
boxplot(dat_ED$Z_Beh_Score~dat_ED$Education,ylim=c(0,60))
```




```
fit7<- lm(Z_Beh_Score~as.factor(Education),data=data)
tidy(fit7)  %>%knitr::kable(caption = "Education related to pro-environmental actions")
```

Education related to pro-environmental actions

| term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|
| (Intercept) | -0.0869479 | 0.1885683 | -0.4610951 | 0.6449714 |
| as.factor(Education)E | 0.0558047 | 0.2316385 | 0.2409129 | 0.8097413 |
| as.factor(Education)NE | 0.1422036 | 0.2316385 | 0.6139032 | 0.5396148 |

```
fit8<- lm(Z_Beh_Score~as.factor(Walk),data=data)
tidy(fit8)  %>%knitr::kable(caption = "Education related to pro-environmental actions")
```

Education related to pro-environmental actions

| term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|
| (Intercept) | -0.0869479 | 0.188615 | -0.4609808 | 0.6450533 |
| as.factor(Walk)GW | 0.0991516 | 0.231233 | 0.4287952 | 0.6682942 |
| as.factor(Walk)TW | 0.0988531 | 0.232169 | 0.4257809 | 0.6704877 |

# Conclusion

After analyzing the data, we can conclude that valence and political have a significant relationship with pro_envrionmental actions under conditions within the control.And when we remove the control,we can see valence is not significant and the level of political is lower.

From the output,we can find that there is no interaction between arousal and education.We can also find from the graph that tree walk have higher response than ground walk and education do improve people's pro-environmental actions.

For the future analysis,we can use Eco_mean,nr and nep which represent how a person connects to nature as our response variables.