

Assignment1

George Paul

2024-01-22

```
# install.packages('readxl')
library(readxl)

# install.packages("vioplot")
library("vioplot")
```

```
## Loading required package: sm

## Package 'sm', version 2.2-5.7: type help(sm) for summary information

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##   as.Date, as.Date.numeric
```

```
# install.packages("ggplot2")
library(ggplot2)
```

```
rm(list = setdiff(ls(), lsf.str()))
```

```
excel_path <- 'D:\\FILES\\BRSMa1.xlsx'
```

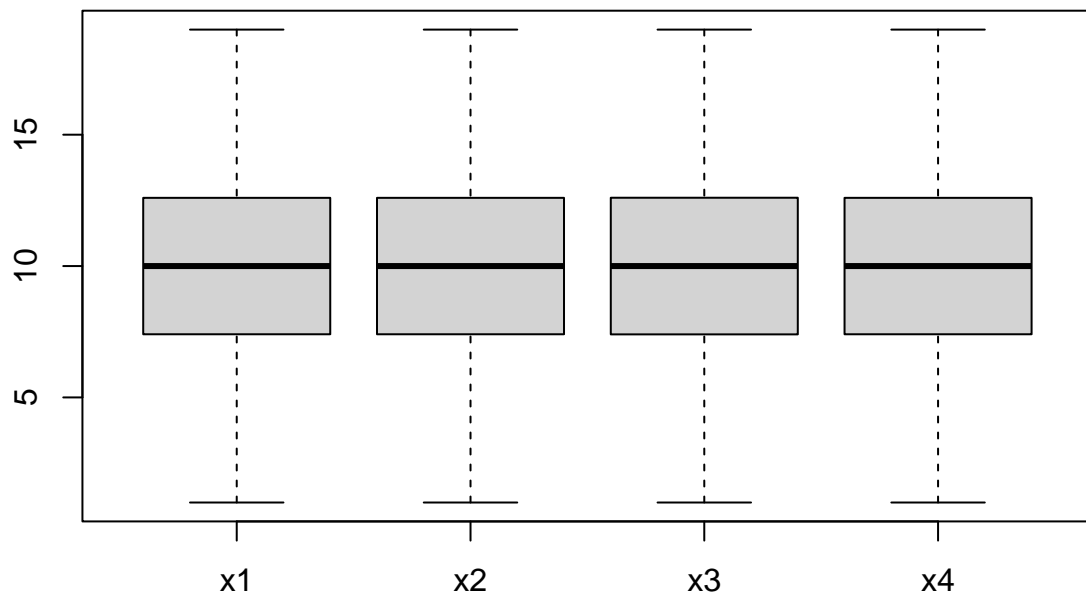
Question 1

```
q1data <- read_excel(excel_path)
q1data
```

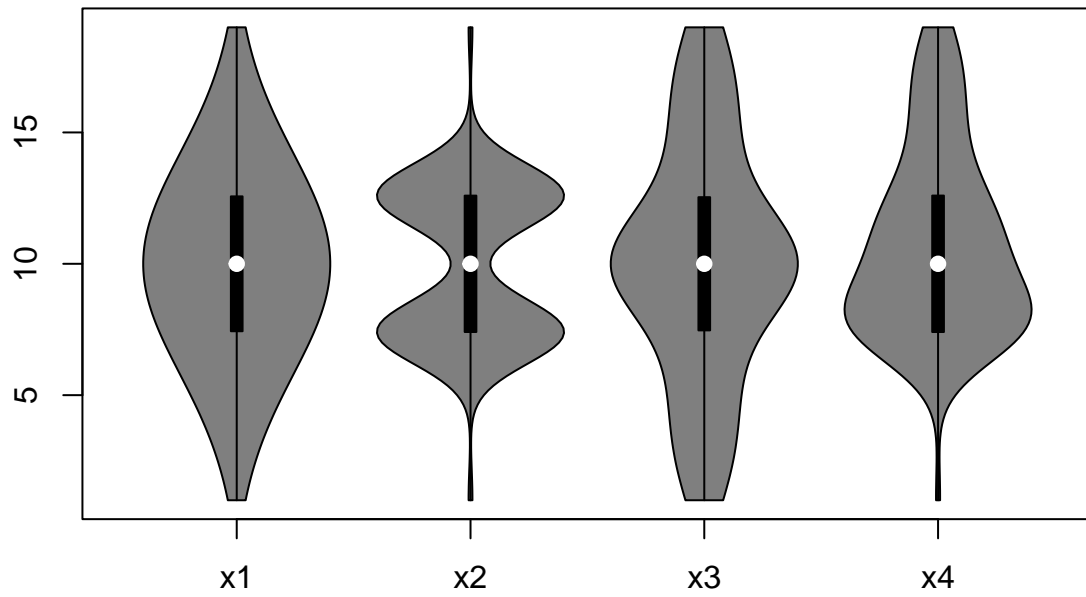
```
## # A tibble: 100 x 4
##       x1     x2     x3     x4
##   <dbl> <dbl> <dbl> <dbl>
## 1  1     1     1     1
## 2  2  2.02  7.10  1.26  7.40
```

```
## 3 2.68 7.16 1.52 7.40
## 4 3.18 7.19 1.78 7.40
## 5 3.59 7.21 2.04 7.40
## 6 3.93 7.23 2.31 7.40
## 7 4.24 7.25 2.57 7.40
## 8 4.51 7.26 2.83 7.40
## 9 4.76 7.27 3.09 7.40
## 10 4.99 7.28 3.35 7.40
## # i 90 more rows
```

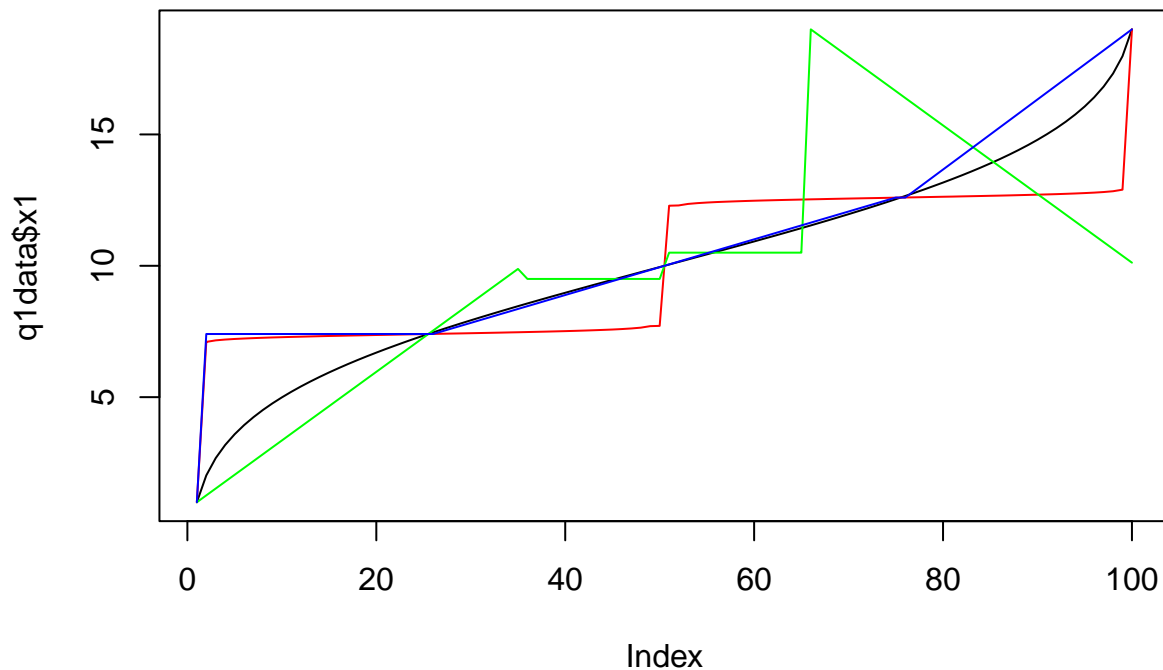
```
boxplot(q1data)
```



```
vioplot(q1data)
```



```
plot(q1data$x1, type = "l") +  
  lines(q1data$x2, type = "l", col = "red") +  
  lines(q1data$x3, type = "l", col = "green") +  
  lines(q1data$x4, type = "l", col = "blue")
```



```
## integer(0)
```

The box plot for this data is entirely misleading - showing the same plots for all data groups.

The violin plot does away with most problems giving a much better idea of the distributions but perhaps x1 and x3 seem too similar even though they are very different.

The line plot is also a valid method of showing the data since the data seems to be of a somewhat continuous nature and does away with the problems of the violin plot.

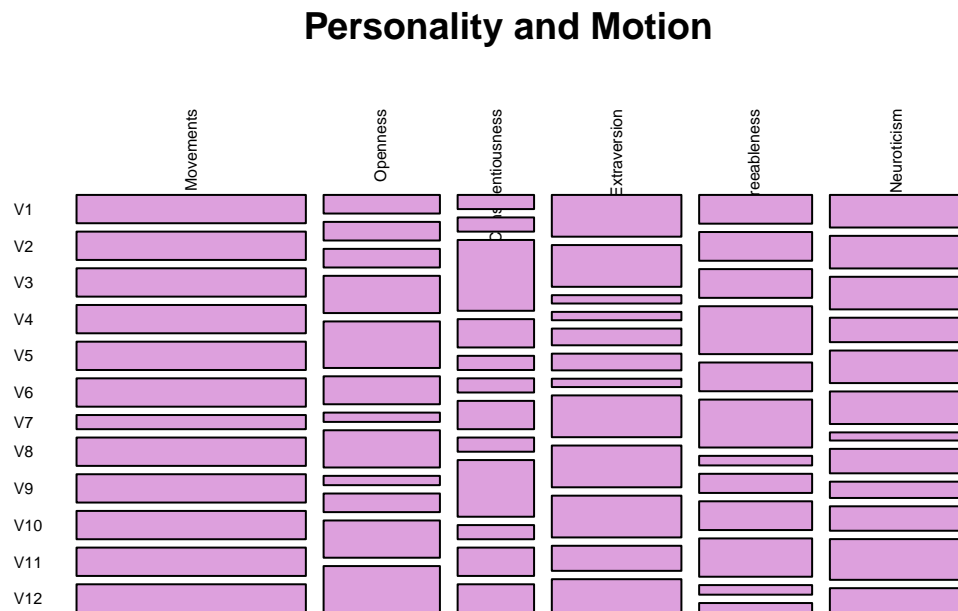
Question 2

```
q2data <- read_excel(excel_path, sheet = "Movement Personality Results")
q2data
```

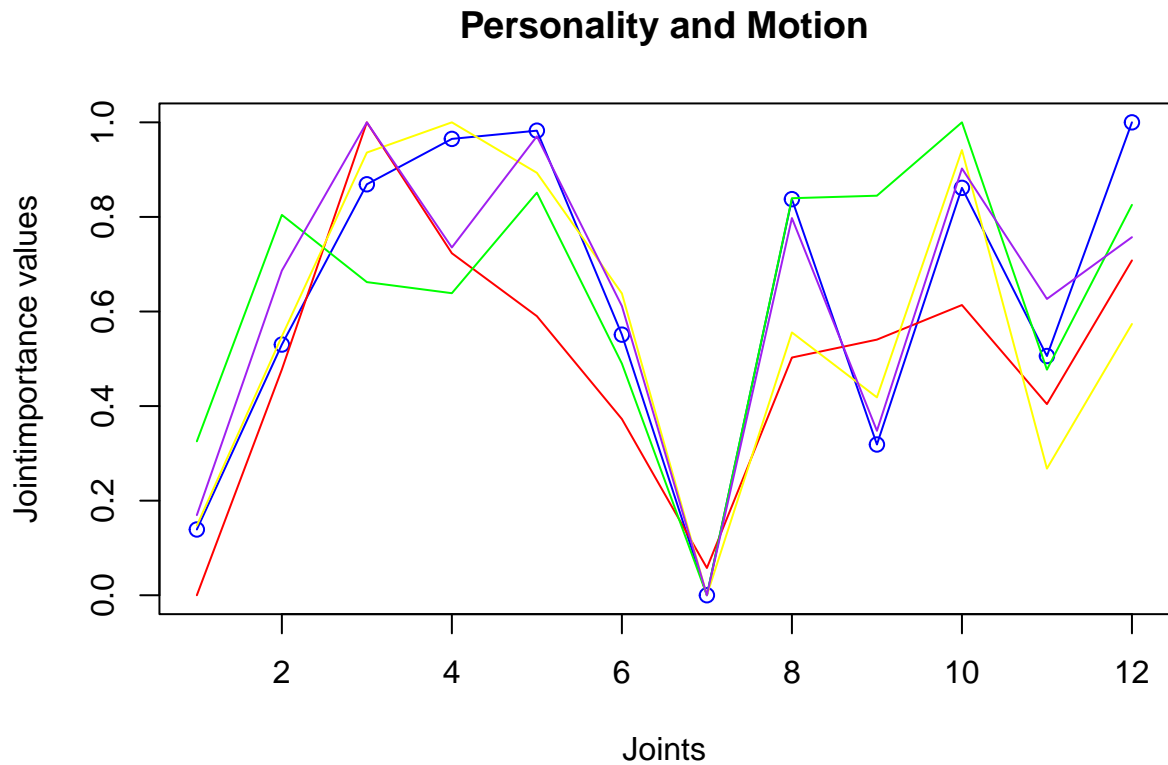
```
## # A tibble: 12 x 6
##   Movements Openness Conscientiousness Extraversion Agreeableness Neuroticism
##   <chr>      <dbl>          <dbl>          <dbl>          <dbl>          <dbl>
## 1 Root        0.139            0            0.325          0.147          0.169
## 2 Hips        0.530            0.477          0.804          0.548          0.686
## 3 Knee        0.869            1            0.662          0.936          1
## 4 Ankle       0.965            0.723          0.639          1            0.735
## 5 Toe         0.982            0.590          0.851          0.893          0.970
```

```
## 6 Torso      0.551      0.373      0.490      0.638      0.612
## 7 Neck       0      0.0576     0      0      0
## 8 Head      0.838      0.503      0.840      0.556      0.798
## 9 Shoulder  0.319      0.541      0.845      0.418      0.348
## 10 Elbow    0.861      0.614      1      0.941      0.902
## 11 Wrist    0.506      0.404      0.477      0.268      0.627
## 12 Finger   1      0.708      0.826      0.574      0.757
```

```
mosaicplot(as.data.frame(t(q2data)), main = "Personality and Motion", color = "plum", las = 2, cex.axis
```



```
plot(q2data$Openness, type = "o", col = "blue", xlab = "Joints", ylab = "Jointimportance values", main =
lines(q2data$Conscientiousness, col="red") +
lines(q2data$Extraversion, col="green") +
lines(q2data$Agreeableness, col="yellow") +
lines(q2data$Neuroticism, col="purple")
```



```
## integer(0)
```

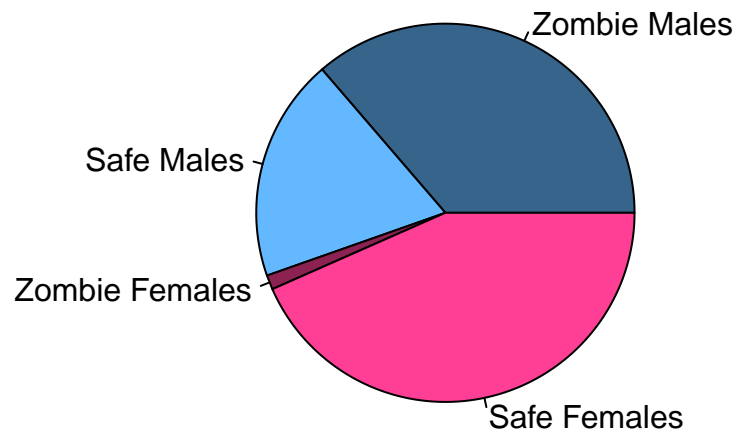
Since we want to know how much each joint's value contributes to that personality trait, a mosaic plot with tiles scaled according to contribution is a great choice

An inferior way to present this data is shown above as a line plot. The data is not continuous nor does it show the relation each point will have with their counterparts in the same joint or personality trait.

Question 3.1

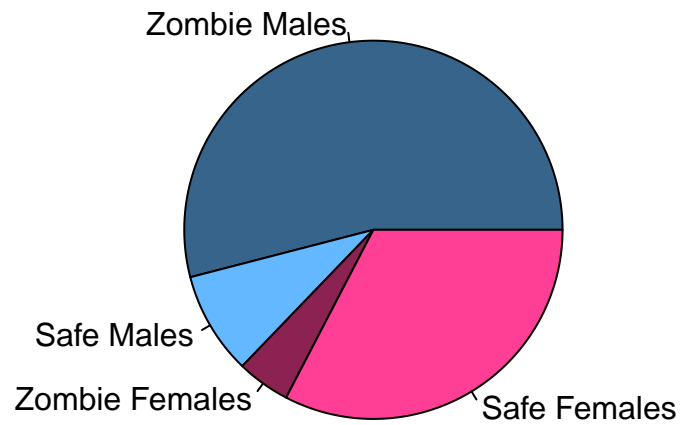
```
q31labels <- c("Zombie Males", "Safe Males", "Zombie Females", "Safe Females")
q31colors <- c("steelblue4", "steelblue1", "violetred4", "violetred1")
q31.1 <- c(118, 62, 4, 141)
pie(q31.1, labels = q31labels, col = q31colors, main="Safe Zone")
```

Safe Zone



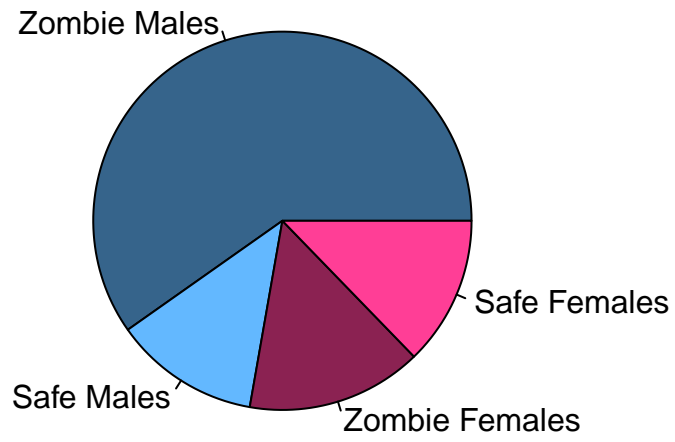
```
q31.2 <- c(154, 25, 13, 93)
pie(q31.2, labels = q31labels, col = q31colors, main="Contaminated City")
```

Contaminated City



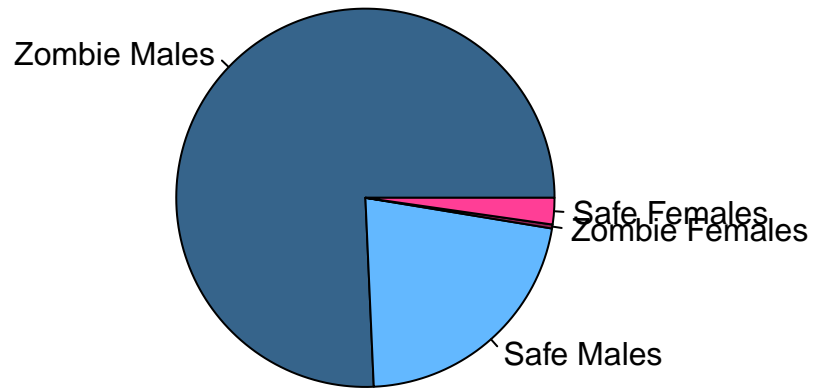
```
q31.3 <- c(422, 88, 106, 90)
pie(q31.3, labels = q31labels, col = q31colors, main="Rural Area")
```


Rural Area



```
q31.4 <- c(670, 192, 3, 20)
pie(q31.4, labels = q31labels, col = q31colors, main="Isolated Island")
```

Isolated Island



Pie charts can easily show you the fraction of people that survived in each zone and importantly, also the relation between being male or female in that zone.