

## Divyansh Tiwari (2020111002)

2023-02-06

### ADVERT RATING: OUTLIER DETECTION

The following is the code for the Question 1

```
library(readxl)
library(ggplot2)
library(tidyr)
library(reshape2)

##
## Attaching package: 'reshape2'

## The following object is masked from 'package:tidyr':
##
##      smiths

mydataq1 <- read_excel("./Reliability_Activity.xlsx", 1)

# print(mydataq1)

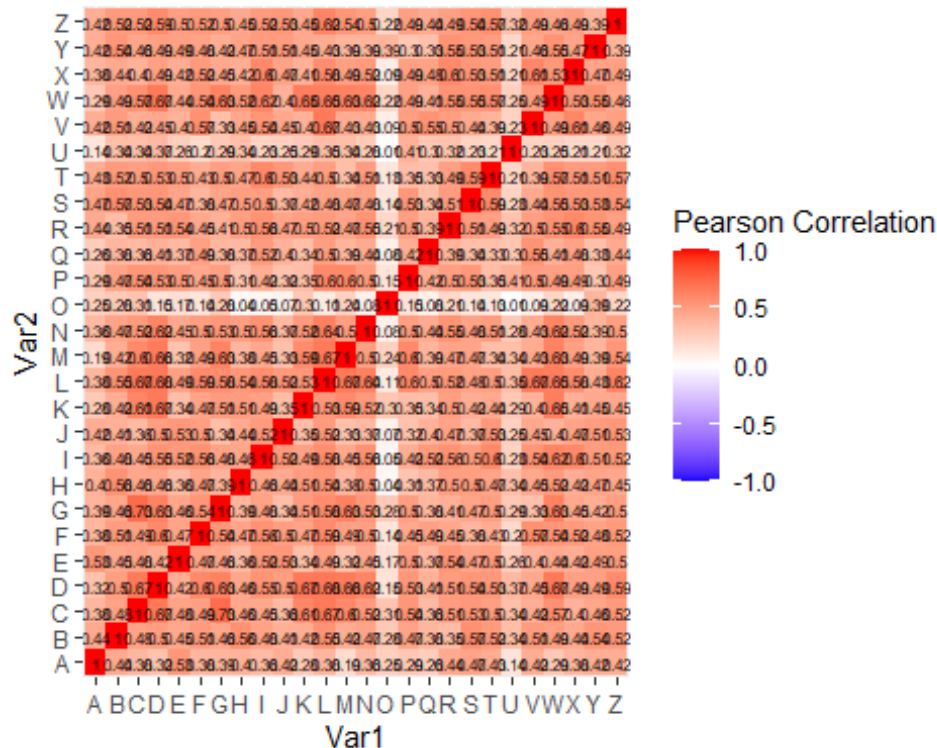
cormat <- round(x = cor(mydataq1, method="pearson"), digits = 2)

# print(cormat)

melted_cormat <- melt(cormat)

# print(melted_cormat)

ggplot(data=melted_cormat, aes(x=Var1, y=Var2, fill=value)) +
  geom_tile() +
  geom_text(aes(label = value), size = 2) +
  scale_fill_gradient2(low = "blue", high="red", limit = c(-1, 1),
name="Pearson Correlation") +
  theme(axis.text.x = element_text(angle = 0))
```



From the heatmap that I have plotted above I can clearly see that the correlation values are the least (close to 0) for the Participant 'O' which is shown in the lighter shade of red which depicts the least correlation values.

So, the participant who has provided random ratings is the participant 'O'.

## RELIABLE JOB: INTERNAL CONSISTENCY

The following is the code for the second question.

```
library("ltm")

## Loading required package: MASS

## Loading required package: msm

## Loading required package: polycor

mydataq2 <- read_excel("./Reliability_Activity.xlsx", 2)

# print(head(mydataq2))
js <- subset(mydataq2, select = c("JS1", "JS2", "JS3", "JS4"))
cronbach.alpha(js, CI=TRUE, standardized = TRUE)
```

```
##
## Standardized Cronbach's alpha for the 'js' data-set
##
## Items: 4
## Sample units: 30
## alpha: 0.871
##
## Bootstrap 95% CI based on 1000 samples
## 2.5% 97.5%
## 0.754 0.930
```

Given above is the cronbach alpha value for the job satisfaction questionnaire. From the alpha value, we can infer that the questionnaire for assessing job satisfaction is internally consistent and is reliable.

```
jp <- subset(mydataq2, select = c("JP1", "JP2", "JP3", "JP4"))
cronbach.alpha(jp, CI=TRUE, standardized = TRUE)

##
## Standardized Cronbach's alpha for the 'jp' data-set
##
## Items: 4
## Sample units: 30
## alpha: 0.545
##
## Bootstrap 95% CI based on 1000 samples
## 2.5% 97.5%
## 0.185 0.724
```

Given above is the cronbach alpha value of the job performance questionnaire. From the alpha value, we can infer that the questionnaire is not acceptable as it has a poor internal reliability among the set of questions.

```
except_avgs <- subset(mydataq2, select = c("JS1", "JS2", "JS3", "JS4", "JP1",
"JP2", "JP3", "JP4"))
print(cronbach.alpha(except_avgs, CI=TRUE, standardized = TRUE))

##
## Standardized Cronbach's alpha for the 'except_avgs' data-set
##
## Items: 8
## Sample units: 30
## alpha: 0.643
##
## Bootstrap 95% CI based on 1000 samples
## 2.5% 97.5%
## 0.412 0.766
```

Given above is the cronbach alpha value of the entire questionnaire. This is done to check the internal reliability between the two subset of the questionnaires and we find that the questionnaire's reliability is questionable and cannot be relied upon with this set.

Although, this is not a very good thing to calculate the cronbach alpha value for 2 set of questionnaire which are aimed at assessing different thing, I still gave it a try.

---

---