Item Response Theory (IRT) for Job Satisfaction Survey

Daisy Shi

11/9/2021

```
#https://www.kaggle.com/annettecatherinepaul/irt-analysis-using-job-satisfaction-survey/notebook#Modelling-using-GRM
```

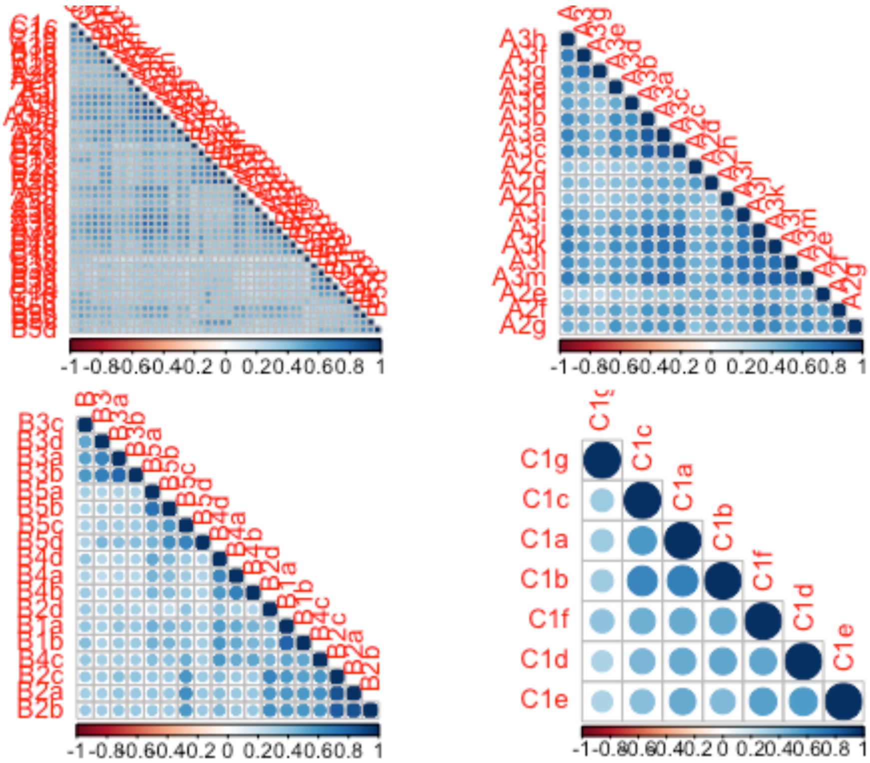
```
data<-read.csv("/Users/daisyshi/Desktop/Survey for Job Satisfaction/Survey_data.csv")
head(data)
##   A1a A1b A2a A2b A2c A2d A2e A2f A2g A2h A3a A3b A3c A3d A3e A3f A3g A3h A3i
## 1   2   5   1   1   1   2   1   1   3   1   6   7   6   5   4   3   3   4   1
## 2   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
## 3   6   6   3   3   3   4   2   4   4   4   7   5   7   4   5   4   2   5   7
## 4   7   7   7   7   4   7   1   7   7  NA   7   7   7   6   7   7   7   7   7
## 5   7   7   5   7   5   7   1   5   7   6   6   6   6   4   4   4   3   4   6
## 6   1   3   2   1   2   5   2   3   2   1   3   2   2   2   4   2   3   1   1
##   A3j A3k A3l A3m A4 A4ai A4aii A4aiii A4aiv A4av A4avi A4avii B1a B1b B2a B2b
## 1   3   3   3   4  3   NA    NA     NA    NA   NA    NA     NA   2   1   1   1
## 2   1   1   1   1  3   NA    NA     NA    NA   NA    NA     NA   1   1   1   1
## 3   7   6   6   7  3   NA    NA     NA    NA   NA    NA     NA   7   7   1   1
## 4   7   7   7   7  3   NA    NA     NA    NA   NA    NA     NA   7   7   7   7
## 5   5   6   7   7  1   NA     1     NA    NA   NA    NA     NA   7   5   1   1
## 6   3   3   2   2  3   NA    NA     NA    NA   NA    NA     NA   2   5   2   2
##   B2c B2d B3a B3b B3c B3d B4a B4b B4c B4d B5a B5b B5c B5d B6 B7a B7b B7c B7d
## 1   1   2   1   1   1   1   5   3   3   4   2   1   1   1  3  NA  NA  NA  NA
## 2   1   1   1   1   1   1   1   1   1   1   1   1   1   1  1  NA  NA  NA  NA
## 3   1   2   4   1   2   4   4   5   1   4   4   4   2   4  2  NA  NA  NA  NA
## 4   7   7   7   7   7   7   7   7   7   7   4   1   7   7  2  NA  NA  NA  NA
## 5   1   1   4   4   4   2   7   7   4   7   4   4   1   1  3  NA  NA  NA  NA
## 6   2   8   2   2   2   1   1   2   2   2   2   1   1   1  1  NA  NA  NA   1
##   B7e B8 B8ai B8aii B8aiii B8aiv B8av B8avi B8avii B9 C1a C1b C1c C1d C1e C1f
## 1  NA  1    1     1      1     1    1     1     NA  2   1   4   4   1   1   2
## 2   1  3   NA    NA     NA    NA   NA    NA     NA  1   1   1   1   1   1   1
## 3  NA  1    1     1      1     1   NA    NA     NA  2   3   7   7   1   2   4
## 4  NA  2   NA    NA     NA    NA   NA    NA     NA  5   2   7   7   7   4   4
## 5  NA  1    1     1      1     1    1     1      1  4   2   5   6   1   1   3
## 6  NA  1    1    NA      1     1    1     1     NA  3   1   2   2   2   2   2
##   C1g C2 C3 C4a C4b C4c C4d C4e C4f C4g C4h C4i C4j C4k C4l c4m C4n C4o C4ii C5
## 1   1  1  1   1   1   1   1   1   1   1   2   1   1   1   1   1   1   1    1  1
## 2   1  1  1   3   3   1   1   1   1   1   1   1   1   1   1   1   1   1    2  1
## 3   4  1  1   4   4   2   3   2   2   1   1   1   1   1   1   3   1   1    2  1
## 4   4  1  1   5   5   3   5   5   5   1   2   1   1   1   5   5   1   1    1  1
## 5   8  1  1   4   3   1   2   2   3   1   2   1   1   1   1   2   1   2    2  1
## 6   2  1  1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   NA  1
##   C6 C6ai C6aii C6aiii C6aiv C6av C6avi C6avii C6aviii C6aix C6ax C6axi C7 C8
## 1  2   NA    NA     NA    NA   NA    NA     NA      NA    NA   NA    NA  1  2
## 2  2   NA    NA     NA    NA   NA    NA     NA      NA    NA   NA    NA  1  1
## 3  2   NA    NA     NA    NA   NA    NA     NA      NA    NA   NA    NA  1  1
## 4  1    1    NA     NA    NA    1     1      1       1     1    1    NA  2  2
## 5  2   NA    NA     NA    NA   NA    NA     NA      NA    NA   NA    NA  1  2
## 6  2   NA    NA     NA    NA   NA    NA     NA      NA    NA   NA    NA  1  2
##   C9 D3 D8 D14 AgencySize
## 1 NA  2  4   2          3
## 2  1  2  3   2          3
## 3  2  1  3   2          3
## 4 NA  1  2   2          3
## 5 NA  1  2   2          3
## 6 NA  2  3   2          3
a <- data[,c(5:23)]
b <- data[, c(32:49)]
c <- data[, c(65:71)]
combine <- cbind(a,b,c)
names(combine)
##  [1] "A2c" "A2d" "A2e" "A2f" "A2g" "A2h" "A3a" "A3b" "A3c" "A3d" "A3e" "A3f"
## [13] "A3g" "A3h" "A3i" "A3j" "A3k" "A3l" "A3m" "B1a" "B1b" "B2a" "B2b" "B2c"
## [25] "B2d" "B3a" "B3b" "B3c" "B3d" "B4a" "B4b" "B4c" "B4d" "B5a" "B5b" "B5c"
## [37] "B5d" "C1a" "C1b" "C1c" "C1d" "C1e" "C1f" "C1g"
# Missing Values Treatment
```

```
# Now, lets take a deeper look into the missing values.

#check for missing values in each series
sapply(combine, function(x) sum(is.na(x)))
## A2c A2d A2e A2f A2g A2h A3a A3b A3c A3d A3e A3f A3g A3h A3i A3j A3k A3l A3m B1a
##  26  19  13  26  30  16  11  17  48  34  36  31  33  18  21  22  30  24  15  14
## B1b B2a B2b B2c B2d B3a B3b B3c B3d B4a B4b B4c B4d B5a B5b B5c B5d C1a C1b C1c
##  51  11  16  18  18  23  36  28  28  28  25  35  28  12  26  19  40   9  13  22
## C1d C1e C1f C1g
##  21  26  22  22
sapply(a, function(x) sum(is.na(x)))
## A2c A2d A2e A2f A2g A2h A3a A3b A3c A3d A3e A3f A3g A3h A3i A3j A3k A3l A3m
##  26  19  13  26  30  16  11  17  48  34  36  31  33  18  21  22  30  24  15
sapply(b, function(x) sum(is.na(x)))
## B1a B1b B2a B2b B2c B2d B3a B3b B3c B3d B4a B4b B4c B4d B5a B5b B5c B5d
##  14  51  11  16  18  18  23  36  28  28  28  25  35  28  12  26  19  40
sapply(c, function(x) sum(is.na(x)))
## C1a C1b C1c C1d C1e C1f C1g
##   9  13  22  21  26  22  22
#Impute the missing values, the imputation will take some time
library("mice")
##
## Attaching package: 'mice'
## The following object is masked from 'package:stats':
##
##     filter
## The following objects are masked from 'package:base':
##
##     cbind, rbind
library(randomForest)
## randomForest 4.6-14
## Type rfNews() to see new features/changes/bug fixes.
library(mice)

set.seed(123)
imp <- mice(combine, method = "rf", m = 5)
##
##  iter imp variable
##   1   1  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   1   2  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   1   3  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   1   4  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   1   5  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   2   1  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   2   2  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   2   3  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   2   4  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   2   5  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   3   1  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   3   2  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   3   3  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
```

```
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   3   4  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   3   5  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   4   1  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   4   2  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   4   3  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   4   4  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   4   5  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   5   1  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   5   2  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   5   3  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   5   4  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
##   5   5  A2c  A2d  A2e  A2f  A2g  A2h  A3a  A3b  A3c  A3d  A3e  A3f  A3g  A3h  A3i  A3j
 A3k  A3l  A3m  B1a  B1b  B2a  B2b  B2c  B2d  B3a  B3b  B3c  B3d  B4a  B4b  B4c  B4d  B5a
 B5b  B5c  B5d  C1a  C1b  C1c  C1d  C1e  C1f  C1g
final <- complete(imp)
sapply(final, function(x) sum(is.na(x)))
## A2c A2d A2e A2f A2g A2h A3a A3b A3c A3d A3e A3f A3g A3h A3i A3j A3k A3l A3m B1a
##   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
## B1b B2a B2b B2c B2d B3a B3b B3c B3d B4a B4b B4c B4d B5a B5b B5c B5d C1a C1b C1c
##   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
## C1d C1e C1f C1g
##   0   0   0   0
#Extract the clean sets
clean_a <- final[,1:19]
clean_b <- final[,20:37]
clean_c <- final[,38:44]

#Check if correct extraction has been done
names(clean_a)
##  [1] "A2c" "A2d" "A2e" "A2f" "A2g" "A2h" "A3a" "A3b" "A3c" "A3d" "A3e" "A3f"
## [13] "A3g" "A3h" "A3i" "A3j" "A3k" "A3l" "A3m"
names(clean_b)
##  [1] "B1a" "B1b" "B2a" "B2b" "B2c" "B2d" "B3a" "B3b" "B3c" "B3d" "B4a" "B4b"
## [13] "B4c" "B4d" "B5a" "B5b" "B5c" "B5d"
names(clean_c)
## [1] "C1a" "C1b" "C1c" "C1d" "C1e" "C1f" "C1g"
#Exploratory Data Analysis
#In this section we will first look into the frequency of responses. Post which, we will look into
#how the correlations between the items looks like. If the items arent correlated we need to #assess
the fit of it in our further analysis.

#Load the necessary libraries
library(ltm)
## Loading required package: MASS
## Loading required package: msm
## Loading required package: polycor
library(corrplot)
## corrplot 0.84 loaded
library(psych)
##
## Attaching package: 'psych'
```

```
## The following object is masked from 'package:ltm':
##
##     factor.scores
## The following object is masked from 'package:polycor':
##
##     polyserial
## The following object is masked from 'package:randomForest':
##
##     outlier
#To understand the frequency spread
description <- descript(final)
description$perc * 100
##             1        2         3          4          5          6          7
## A2c 46.225200 32.44009 11.079619  4.096882  2.6797217  1.777892  1.5459933
## A2d 29.399639 32.36279 15.588766  4.689513  6.6477712  5.204844  5.7459418
## A2e 29.966503 30.48183 15.614532  6.853904  5.7717083  5.488276  5.7974749
## A2f 18.474620 24.01443 20.716310  9.688225  8.6575625  7.472301 10.7961865
## A2g 12.831744 17.93352 15.279567 12.084514 11.1053852 10.332389 18.9641845
## A2h 37.284205 26.20459 13.166710 14.274671  3.3238856  2.834321  2.8343211
## A3a 13.501675 26.69415 18.629219  8.631796  9.8428240 10.177789 12.0845143
## A3b  8.554496 21.51507 19.994847  9.095594 13.1151765 11.285751 15.8464313
## A3c 13.398609 23.78253 17.675857 10.976552 10.0489565 10.693120 12.9605772
## A3d 13.372842 24.81319 21.695439 13.166710  9.1728936  7.961865  8.2195310
## A3e 11.131152 20.53594 15.253801 15.305334 11.5176501  9.507859 13.7335738
## A3f 22.880701 27.51868 18.165421  9.610925  8.2710642  5.694409  7.6011337
## A3g 30.043803 28.65241 17.675857  6.750837  5.9520742  4.509147  5.5398093
## A3h 23.164133 25.43159 18.165421 12.651379  7.3950013  5.926308  6.8281371
## A3i 19.608348 23.39603 19.427982 17.598557  6.6735377  5.952074  7.1115692
## A3j 13.244009 21.41201 21.669673 15.872198 10.7188869  7.188869  9.6366916
## A3k 15.021902 22.08194 19.917547 15.846431  9.9458902  7.523834  9.4047926
## A3l 25.818088 23.85983 16.722494 17.985055  4.8698789  4.689513  5.7974749
## A3m 19.608348 21.66967 16.593661 16.078330  7.7814996  7.498068 10.5127544
## B1a 17.495491 26.82298 18.191188  6.235506 10.1262561  8.940995 11.7237825
## B1b 18.268488 25.68926 17.804690  8.142231  9.7397578  7.858799 11.9556815
## B2a 40.865756 26.25612 12.316413  4.972945  4.8956455  4.251482  6.0293739
## B2b 35.995877 26.15305 13.166710  7.575367  5.4367431  4.483381  6.0036073
## B2c 34.192219 24.19480 14.326205  6.776604  6.2097398  4.560680  8.7606287
## B2d 21.334708 19.86601 12.110281 13.733574  8.1937645  5.694409 10.1777892
## B3a 29.141974 33.05849 16.258696  7.498068  6.3385725  3.890750  2.6024221
## B3b 40.015460 30.86833 13.553208  6.158207  3.7361505  2.447823  2.0613244
## B3c 26.178820 25.32852 16.361762 14.815769  4.3545478  3.581551  4.1226488
## B3d 44.447307 27.62175 13.630508  7.137336  2.5251224  1.442927  1.6490595
## B4a 21.489307 23.83406 17.340892  7.961865  8.0906983  7.188869 12.6513785
## B4b 25.405823 28.60088 19.556815 10.023190  5.8747745  3.916516  5.2563772
## B4c 28.858542 25.48312 14.326205 11.594950  4.7925792  4.354548  8.5029632
## B4d 14.944602 20.94821 17.495491 12.831744  9.0698274  7.833033 14.2746715
## B5a 27.467148 27.13218 17.624324 15.923731  3.4269518  2.602422  3.3496522
## B5b 38.804432 28.05978 14.506570  9.868591  2.9631538  1.855192  2.6024221
## B5c 55.011595 22.93223  8.245298  7.034270  1.4944602  1.211028  2.4220562
## B5d 47.771193 29.37387 10.332389  6.209740  2.4220562  1.185262  1.6748261
## C1a 44.292708 27.36408 12.805978  7.214635  2.8085545  2.009791  2.8858542
## C1b 30.713734 25.35429 12.574079 10.383922  6.5447050  4.895645  7.2919351
## C1c 28.188611 24.16903 13.037877 12.136047  5.5913424  3.710384  6.3128060
## C1d 53.800567 22.59727  8.606029  6.209740  2.5766555  1.726359  3.4011853
## C1e 42.076784 31.87323 12.960577  6.183973  2.6797217  1.571760  1.8551920
## C1f 48.647256 24.16903  8.940995  6.725071  3.7876836  2.138624  3.0662200
## C1g 64.596753 10.22932  2.757021  6.776604  0.3864983  0.231899  0.8760629
##              8
## A2c  0.15459933
## A2d  0.36073177
## A2e  0.02576656
## A2f  0.18036589
## A2g  1.46869364
## A2h  0.07729967
## A3a  0.43803144
## A3b  0.59263077
## A3c  0.46379799
## A3d  1.59752641
## A3e  3.01468694
## A3f  0.25766555
## A3g  0.87606287
## A3h  0.43803144
## A3i  0.23189900
## A3j  0.25766555
## A3k  0.25766555
## A3l  0.25766555
```

```
## A3m  0.25766555
## B1a  0.46379799
## B1b  0.54109766
## B2a  0.41226488
## B2b  1.18526153
## B2c  0.97912909
## B2d  8.88946148
## B3a  1.21102809
## B3b  1.15949498
## B3c  5.25637722
## B3d  1.54599330
## B4a  1.44292708
## B4b  1.36562742
## B4c  2.08709096
## B4d  2.60242206
## B5a  2.47358928
## B5b  1.33986086
## B5c  1.64905952
## B5d  1.03066220
## C1a  0.61839732
## C1b  2.24169029
## C1c  6.85390363
## C1d  1.08219531
## C1e  0.79876321
## C1f  2.52512239
## C1g 14.14583870
#To understand the correlation statistics
relation <- cor(final, method = "spearman")
relationa <- cor(clean_a, method = "spearman")
relationb <- cor(clean_b, method = "spearman")
relationc <- cor(clean_c, method = "spearman")
par(mfrow=c(2,2))
corrplot(relation, type = "lower", order = "hclust")
corrplot(relationa, type = "lower", order = "hclust")
corrplot(relationb, type = "lower", order = "hclust")
corrplot(relationc, type = "lower", order = "hclust")
```



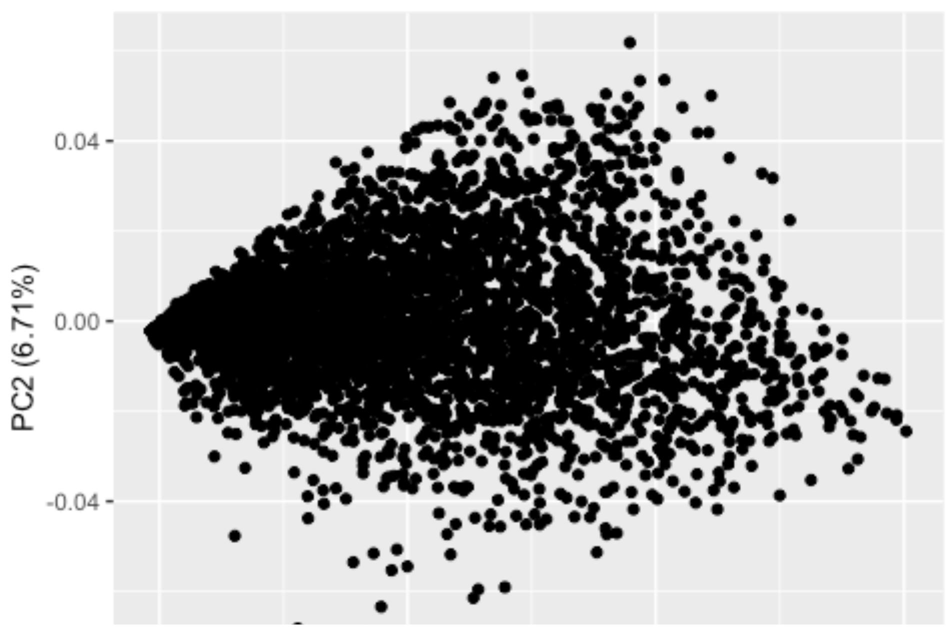
#We can observe that all the plots exhibit strong positive correlations. Hence we can use the #whole dataset for our analysis without further cleaning on the grounds of correlation.

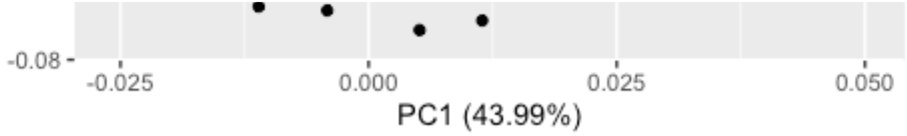
#Pre Model Checks
#Before producing to model building, we need to understand if our analysis should take a multi dimensional or uni dimensional approach. This can be visualized using principal components and the scree plot which uses a plot of successive eigen vectors. Ideally if the first principal component (PC1) can explain atleast 40% of the variaition in the data, it should be good enough to take the unidimensionality approach in IRT.

#But sometimes we are not okay with this threshold which is assumed over years of analysis and research. Hence we can take a look at the factor model for better understanding of how the

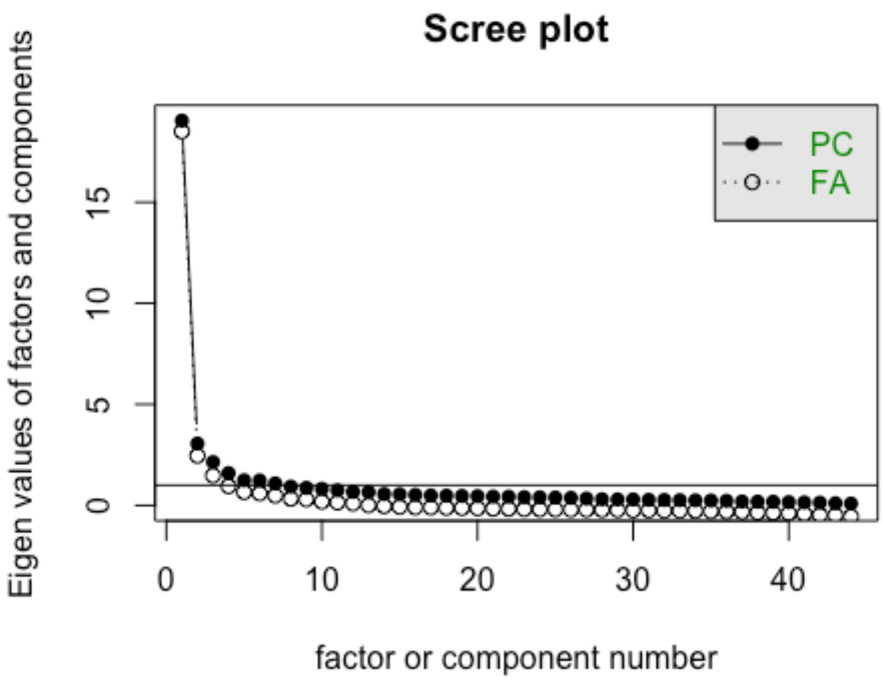
research. Hence we can take a look at the factor model for better understanding of how the underlying structure of our dataset looks like.

```
#load the libraries
library(ggfortify)
## Loading required package: ggplot2
##
## Attaching package: 'ggplot2'
## The following objects are masked from 'package:psych':
##
##     %+%, alpha
## The following object is masked from 'package:randomForest':
##
##     margin
library(ggplot2)

#First let's look into the principal components to see if there is any exhbition of unidimensionality
par(mfrow=c(1,1))
pc <- prcomp(final)
summary(pc)
## Importance of components:
##                           PC1     PC2     PC3     PC4     PC5     PC6     PC7
## Standard deviation     8.1625 3.18830 2.69250 2.27719 2.14029 2.13064 1.95117
## Proportion of Variance 0.4399 0.06712 0.04786 0.03424 0.03024 0.02997 0.02514
## Cumulative Proportion  0.4399 0.50701 0.55487 0.58911 0.61935 0.64933 0.67446
##                            PC8     PC9    PC10    PC11   PC12    PC13    PC14
## Standard deviation     1.89681 1.73096 1.69919 1.65493 1.5371 1.52444 1.45354
## Proportion of Variance 0.02375 0.01978 0.01906 0.01808 0.0156 0.01534 0.01395
## Cumulative Proportion  0.69822 0.71800 0.73706 0.75514 0.7707 0.78609 0.80004
##                           PC15    PC16    PC17    PC18    PC19    PC20    PC21
## Standard deviation     1.36068 1.34586 1.32401 1.30314 1.23606 1.21398 1.18954
## Proportion of Variance 0.01222 0.01196 0.01157 0.01121 0.01009 0.00973 0.00934
## Cumulative Proportion  0.81226 0.82422 0.83579 0.84700 0.85709 0.86682 0.87617
##                           PC22    PC23    PC24    PC25    PC26    PC27    PC28
## Standard deviation     1.17337 1.13903 1.10473 1.07359 1.05038 1.00682 0.98091
## Proportion of Variance 0.00909 0.00857 0.00806 0.00761 0.00728 0.00669 0.00635
## Cumulative Proportion  0.88526 0.89382 0.90188 0.90949 0.91677 0.92347 0.92982
##                           PC29    PC30    PC31    PC32    PC33    PC34    PC35
## Standard deviation     0.96295 0.95467 0.94329 0.93328 0.92441 0.89346 0.85796
## Proportion of Variance 0.00612 0.00602 0.00587 0.00575 0.00564 0.00527 0.00486
## Cumulative Proportion  0.93594 0.94196 0.94783 0.95358 0.95923 0.96450 0.96936
##                          PC36    PC37    PC38    PC39    PC40    PC41    PC42
## Standard deviation     0.8441 0.79564 0.77716 0.75169 0.72237 0.69031 0.67954
## Proportion of Variance 0.0047 0.00418 0.00399 0.00373 0.00345 0.00315 0.00305
## Cumulative Proportion  0.9741 0.97824 0.98223 0.98596 0.98940 0.99255 0.99560
##                           PC43    PC44
## Standard deviation     0.60588 0.54723
## Proportion of Variance 0.00242 0.00198
## Cumulative Proportion  0.99802 1.00000
#We can see that approximately 44% of the data is explained by the first principal component. Now let's look into the scree plot for verification of the same.

autoplot(pc, data = final)
```



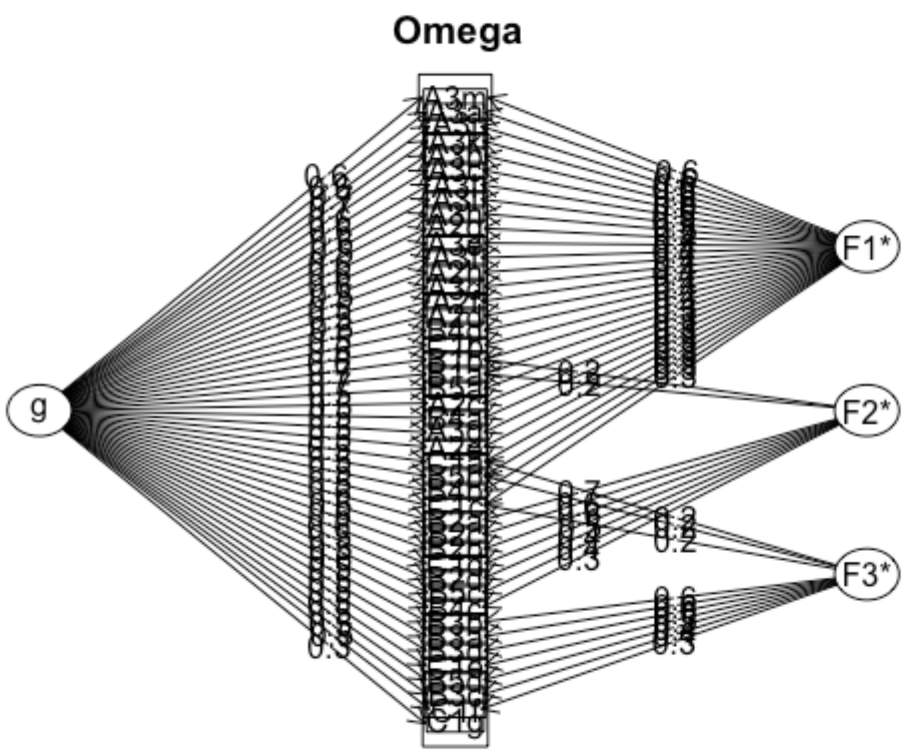



```
#Scree plot
library(psych)
scree(final)
```

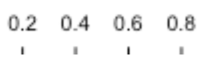


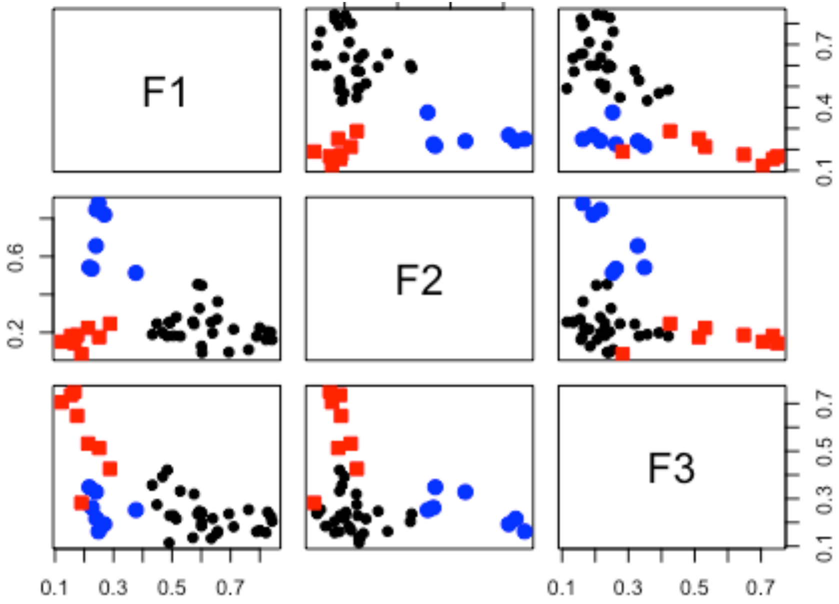
#The scree plot suggests unidimensionality approach as well. Additionally lets take a look at the factor plot to understand the underlying structure of our dataset.

```
#factor plot
plot(omega(final))
## Loading required namespace: GPArotation
```

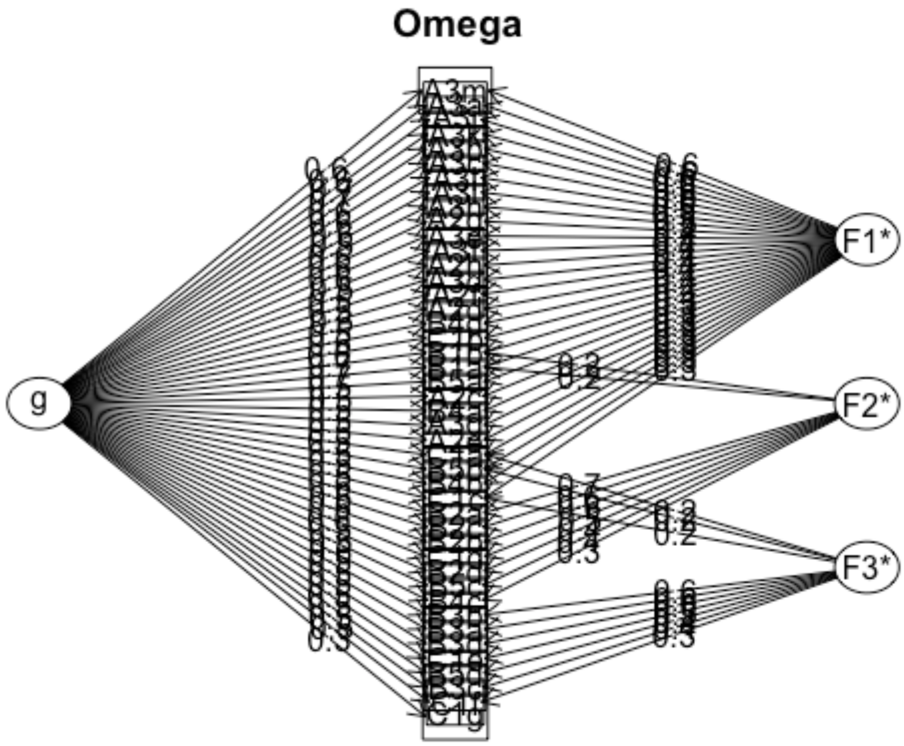



```
## Use omega.diagram to see the  hierarchical structure
```





omega(final)




```
## Omega
## Call: omegah(m = m, nfactors = nfactors, fm = fm, key = key, flip = flip,
##     digits = digits, title = title, sl = sl, labels = labels,
##     plot = plot, n.obs = n.obs, rotate = rotate, Phi = Phi, option = option,
##     covar = covar)
## Alpha:                 0.97
## G.6:                   0.98
## Omega Hierarchical:    0.72
## Omega H asymptotic:    0.74
## Omega Total            0.97
##
## Schmid Leiman Factor loadings greater than  0.2
##        g   F1*   F2*   F3*   h2   u2   p2
## A2c 0.48  0.33             0.34 0.66 0.67
## A2d 0.51  0.37             0.41 0.59 0.64
## A2e 0.50  0.30             0.36 0.64 0.71
## A2f 0.62  0.42             0.59 0.41 0.66
## A2g 0.57  0.43             0.52 0.48 0.63
## A2h 0.50  0.42             0.43 0.57 0.57
## A3a 0.62  0.58             0.73 0.27 0.53
## A3b 0.61  0.55             0.68 0.32 0.55
## A3c 0.64  0.55             0.72 0.28 0.57
## A3d 0.49  0.42             0.41 0.59 0.58
## A3e 0.56  0.43             0.49 0.51 0.63
```

```
## A3f 0.54  0.42           0.48 0.52 0.61
## A3g 0.45  0.31           0.32 0.68 0.65
## A3h 0.59  0.48           0.59 0.41 0.60
## A3i 0.55  0.49           0.55 0.45 0.55
## A3j 0.68  0.58           0.79 0.21 0.58
## A3k 0.68  0.57           0.78 0.22 0.59
## A3l 0.60  0.54           0.66 0.34 0.55
## A3m 0.65  0.59           0.78 0.22 0.54
## B1a 0.66  0.33  0.24      0.61 0.39 0.72
## B1b 0.65  0.35  0.24      0.60 0.40 0.70
## B2a 0.64        0.67      0.86 0.14 0.48
## B2b 0.64        0.61      0.78 0.22 0.52
## B2c 0.64        0.64      0.82 0.18 0.51
## B2d 0.51        0.37      0.41 0.59 0.63
## B3a 0.50             0.59 0.60 0.40 0.42
## B3b 0.49             0.60 0.61 0.39 0.40
## B3c 0.45             0.38 0.36 0.64 0.58
## B3d 0.45             0.57 0.54 0.46 0.38
## B4a 0.53  0.32            0.39 0.61 0.71
## B4b 0.50  0.26            0.34 0.66 0.73
## B4c 0.58        0.33      0.47 0.53 0.71
## B4d 0.61  0.36            0.52 0.48 0.71
## B5a 0.54  0.33            0.42 0.58 0.68
## B5b 0.54  0.27       0.22 0.41 0.59 0.70
## B5c 0.54        0.37      0.46 0.54 0.63
## B5d 0.46             0.39 0.38 0.62 0.57
## C1a 0.55  0.28       0.25 0.44 0.56 0.68
## C1b 0.59  0.36            0.49 0.51 0.70
## C1c 0.50  0.25       0.20 0.35 0.65 0.70
## C1d 0.60        0.46      0.60 0.40 0.61
## C1e 0.48             0.51 0.49 0.51 0.47
## C1f 0.47             0.28 0.32 0.68 0.68
## C1g 0.27             0.20 0.12 0.88 0.61
##
## With eigenvalues of:
##    g  F1*  F2*  F3*
## 13.6  5.3  2.1  2.0
##
## general/max  2.58   max/min =   2.59
## mean percent general =  0.61    with sd =  0.09 and cv of  0.15
## Explained Common Variance of the general factor =  0.59
##
## The degrees of freedom are 817  and the fit is  6.39
## The number of observations was  3881  with Chi Square =  24692.39  with prob <  0
## The root mean square of the residuals is  0.04
## The df corrected root mean square of the residuals is  0.05
## RMSEA index =  0.087  and the 10 % confidence intervals are  0.086 0.088
## BIC =  17940.83
##
## Compare this with the adequacy of just a general factor and no group factors
## The degrees of freedom for just the general factor are 902  and the fit is  13.99
## The number of observations was  3881  with Chi Square =  54067.38  with prob <  0
## The root mean square of the residuals is  0.14
## The df corrected root mean square of the residuals is  0.15
##
## RMSEA index =  0.123  and the 10 % confidence intervals are  0.122 0.124
## BIC =  46613.39
##
## Measures of factor score adequacy
##                                                  g  F1*  F2*  F3*
## Correlation of scores with factors            0.86 0.78 0.83 0.83
## Multiple R square of scores with factors      0.74 0.61 0.68 0.68
## Minimum correlation of factor score estimates 0.48 0.21 0.36 0.37
##
##  Total, General and Subset omega for each subset
##                                                  g  F1*  F2*  F3*
## Omega total for total scores and subscales    0.97 0.96 0.90 0.84
## Omega general for total scores and subscales  0.72 0.63 0.53 0.43
## Omega group for total scores and subscales    0.20 0.33 0.37 0.41
```

This plot explains the underlying structure of the dataset. We can see that the general load factor is distributed among three factors as confirmed by the three series. However, IRT analysis ignores these three factors and tries to formulate a cumulative relationship known as the latent trait which in pur case is the job satisfaction. Additionally, it is interesting to note tha B2a seems to have a higher load than other loadings which is a shift away from the general trend. This could depict that this particular item is too generalized leading to general trend. This could depict that this particular item is too generalized leading to ineffective discrimination and thus rendering the contribution of that item, not much useful. Hence one should try and avoid such items or look into factor treatments. Additionally the qualitative aspects of the item can also help you determine the overall use of the item. While removing the item, you need to ensure that the newly formed structure is viable for modelling and that it converges. However for our analysis we will be using the full dataset as our aim is to provide a comprehensive analysis of the whole dataset.

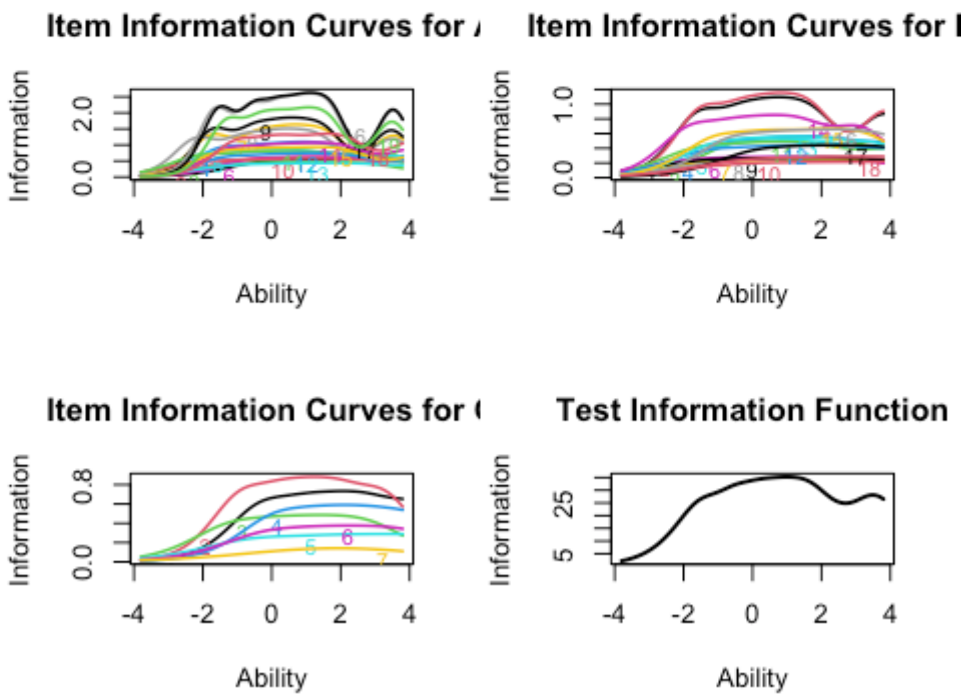
##Modelling using GRM Since we are using likert scale ranging from 1 to 8, from strongly agree to strongly disagree with 8 being N/A. A Graded Response Model would be ideal for our analysis. The ltm package offers grm functionalities.

```
#Fit the GRM model and check for convergence
library(ltm)
fit <- grm(final, IRT.param = TRUE)
fit$convergence
## [1] 1
fit
##
## Call:
## grm(data = final, IRT.param = TRUE)
##
## Coefficients:
##      Extrmt1  Extrmt2  Extrmt3  Extrmt4  Extrmt5  Extrmt6  Extrmt7  Dscrmn
## A2c   -0.279   1.242    2.088    2.587    3.090    3.664    5.517   1.328
## A2d   -1.027   0.423    1.165    1.449    1.972    2.594    4.780   1.386
## A2e   -1.064   0.343    1.188    1.633    2.116    2.825    8.993   1.211
## A2f   -1.481  -0.375    0.439    0.833    1.270    1.782    4.428   1.845
## A2g   -1.968  -0.870   -0.203    0.294    0.770    1.306    3.465   1.618
## A2h   -0.632   0.454    1.097    2.136    2.594    3.306    6.326   1.407
## A3a   -1.743  -0.399    0.292    0.621    1.027    1.591    3.555   2.257
## A3b   -2.175  -0.826   -0.019    0.308    0.813    1.362    3.513   2.172
## A3c   -1.701  -0.518    0.131    0.542    0.935    1.493    3.484   2.404
## A3d   -2.068  -0.561    0.416    1.040    1.561    2.214    3.742   1.341
## A3e   -2.105  -0.788   -0.130    0.507    1.024    1.565    3.015   1.588
## A3f   -1.318  -0.074    0.737    1.236    1.765    2.290    4.549   1.521
## A3g   -1.082   0.358    1.299    1.751    2.258    2.824    4.963   1.165
## A3h   -1.237  -0.107    0.622    1.194    1.640    2.154    4.071   1.824
## A3i   -1.461  -0.362    0.400    1.188    1.612    2.157    4.510   1.714
## A3j   -1.636  -0.584    0.193    0.764    1.225    1.657    3.536   2.884
## A3k   -1.524  -0.495    0.204    0.799    1.241    1.701    3.521   2.896
## A3l   -1.038  -0.069    0.564    1.396    1.710    2.191    4.062   2.065
## A3m   -1.309  -0.350    0.242    0.849    1.197    1.641    3.497   2.616
## B1a   -1.552  -0.309    0.398    0.670    1.133    1.712    3.859   1.844
## B1b   -1.495  -0.318    0.406    0.745    1.213    1.690    3.895   1.893
## B2a   -0.497   0.666    1.328    1.671    2.100    2.656    5.049   1.313
## B2b   -0.733   0.432    1.090    1.577    2.016    2.542    4.175   1.313
## B2c   -0.832   0.300    0.991    1.388    1.854    2.312    4.506   1.256
## B2d   -1.814  -0.515    0.155    0.961    1.504    1.966    3.014   0.929
## B3a   -1.390   0.529    1.629    2.324    3.195    4.107    5.608   0.871
## B3b   -0.639   1.087    2.115    2.839    3.493    4.268    5.458   0.916
## B3c   -1.520   0.012    0.966    2.041    2.505    2.972    3.697   0.914
## B3d   -0.462   1.231    2.426    3.478    4.092    4.617    5.537   0.813
## B4a   -1.564  -0.265    0.523    0.926    1.396    1.927    4.186   1.228
## B4b   -1.323   0.152    1.155    1.792    2.302    2.823    4.344   1.171
## B4c   -1.057   0.107    0.770    1.420    1.763    2.173    3.644   1.311
## B4d   -1.791  -0.609    0.130    0.665    1.087    1.535    3.132   1.624
## B5a   -1.067   0.147    0.946    1.939    2.285    2.674    3.365   1.440
## B5b   -0.583   0.642    1.416    2.234    2.611    2.963    3.890   1.427
## B5c    0.135   1.330    1.966    2.781    3.019    3.253    4.307   1.167
## B5d   -0.186   1.490    2.404    3.288    3.908    4.698    5.425   0.940
## C1a   -0.302   0.854    1.558    2.165    2.549    2.902    4.265   1.506
## C1b   -0.884   0.194    0.732    1.264    1.683    2.099    3.262   1.647
## C1c   -1.148   0.095    0.760    1.465    1.853    2.168    2.826   1.222
## C1d    0.094   1.123    1.687    2.272    2.645    2.968    4.099   1.346
## C1e   -0.539   1.284    2.340    3.187    3.745    4.281    5.445   0.941
## C1f   -0.128   1.121    1.725    2.332    2.805    3.151    3.965   1.073
## C1g    0.940   1.773    2.034    2.757    2.772    2.791    2.896   0.679
##
## Log.Lik: -247014.3
```

If the model has converged, you should see an output of 0. With convergence we can proceed with our unidimensionality approach towards modelling. Now let's examine the fit. Ideally we want to see high descrimination power to conclude towards good analysis.

```
par(mfrow=c(2,2))
```

```
#Item Information Curves
plot(fit, type ="IIC", lwd = 1.5, item = 1:19, main = "Item Information Curves for A")
plot(fit, type ="IIC", lwd = 1.5, item = 20:37, main = "Item Information Curves for B")
plot(fit, type ="IIC", lwd = 1.5, item = 38:44, main = "Item Information Curves for C")

#Test Information Plot
plot(fit, type = "IIC", item = 0, lwd = 2)
```

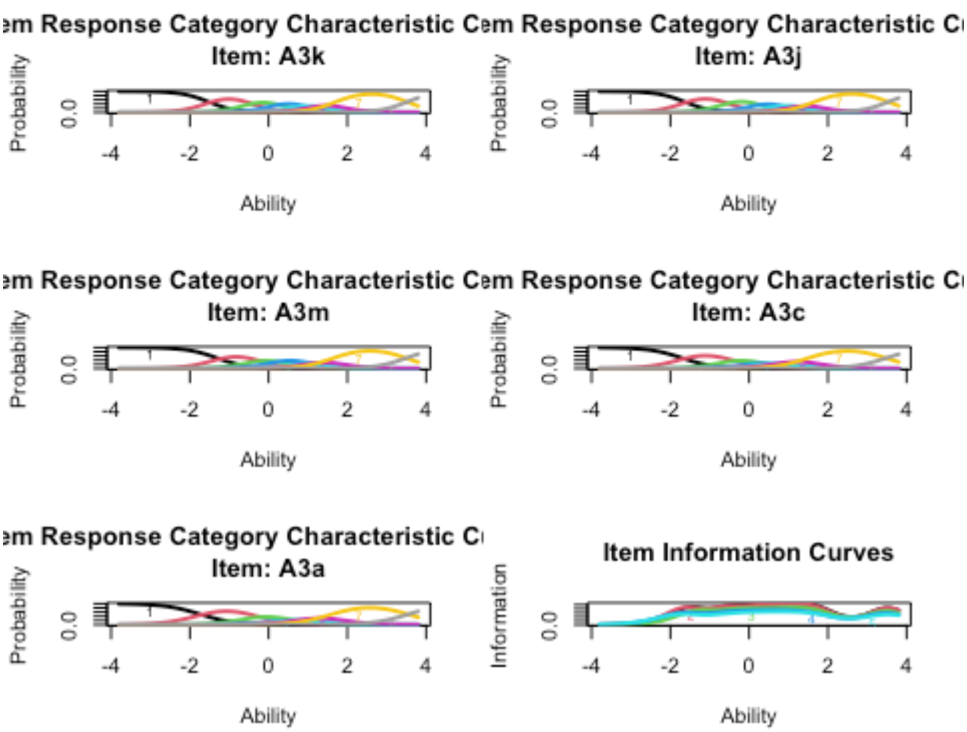


We can observe a high Information ability with series A, followed by B and least for C. The Test Information Function plot suggests the same as well. In general the dataset contirbutes well enough with respect to understanding the latent trait of job satisfaction.

Now let's take a look at the top performing items and the least performing items to compare and contrast. This is done by looking into the discrimination coefficients.

To understand the item description for each item, take a look at the source dataset.

```
#Item Response Category Characterstics Curves and Item Information Curves for top performing items
par(mfrow=c(3,2))
plot(fit, lwd = 2, item = c(17,16,19,9,7))
plot(fit, type = "IIC", item = c(17,16,19,9,7), lwd = 2)
```



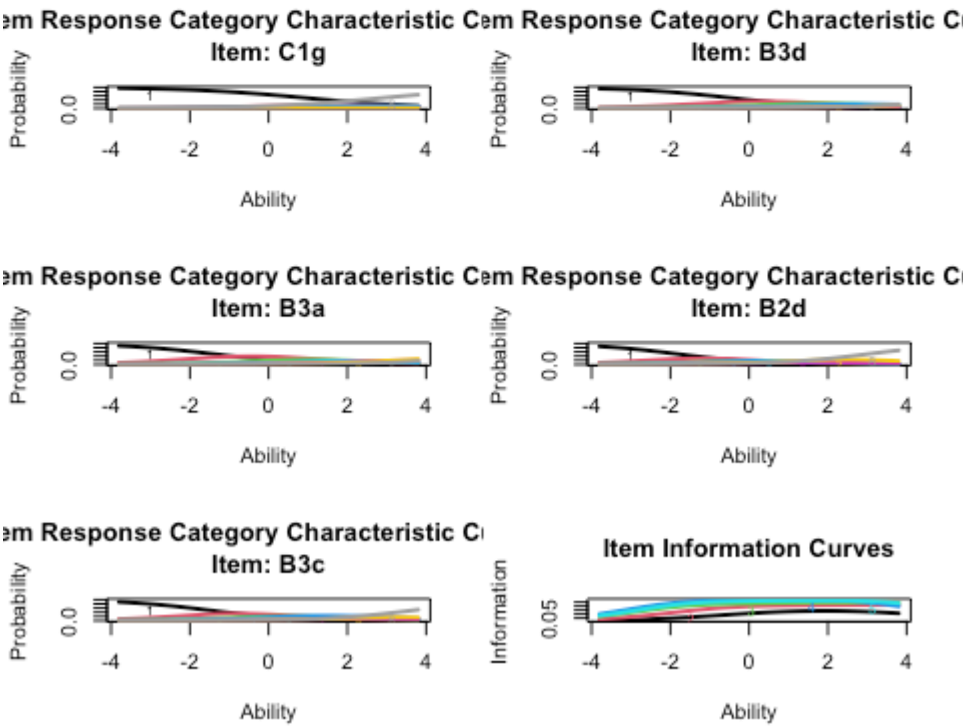
Let

us focus on the first plot. This is the plot with the highest descrimination power. Since our dataset ratings start from strongly agree and moves on to strongly disagree, this plot speaks of job dissatisfaction. In general responders strongly agree that they are satisfied contributing to the latent trait. As the ability increases the job dissatisfaction increases.

In general we can see how the item information curves for these five items are very high.

Now let us look into the lowest performing items.

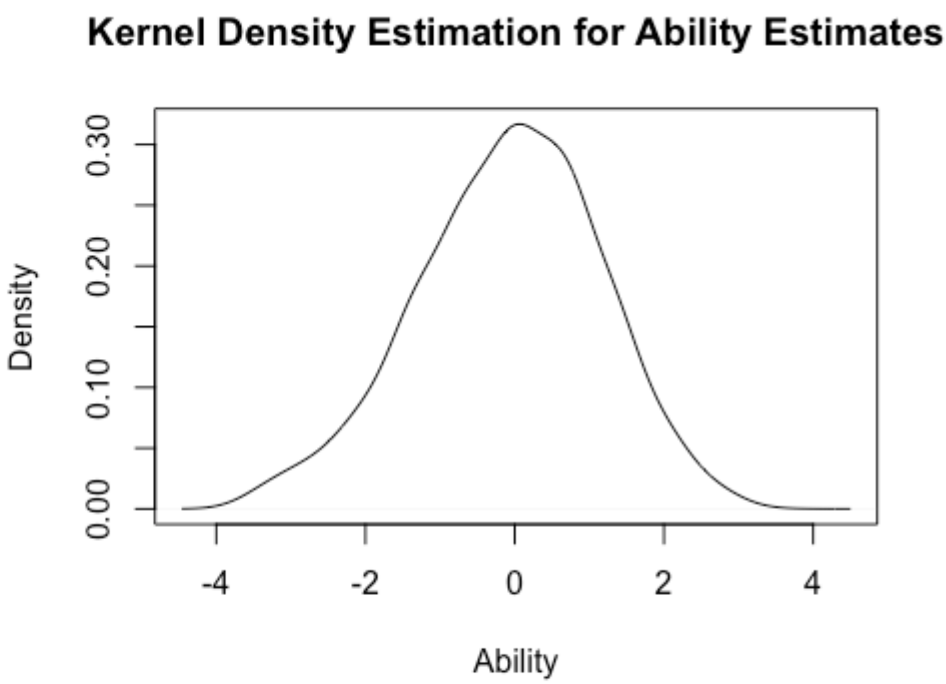
```r
#Item Response Category Characterstics Curves and Item Information Curves for least performing items
par(mfrow=c(3,2))
plot(fit, lwd = 2, item = c(44, 29, 26, 25, 28))
plot(fit, type = "IIC", lwd = 2, item = c(44, 29, 26, 25, 28))
```



We can see how these plots have a very low level steeps, which suggests the low information / descrimination of the item. This could be because of the over generalization of the item query. A qualitative analysis is required to assess the fit of this item in the survey.

Let us look into the factor scores of these observations. For every unique response pattern, a latent score coefficient is calculated.

```r
factor <- ltm::factor.scores(fit, method = "EAP")
plot(factor)
```



We

can see that the peak is around 0. Ideally, the goal is to shift this peak towards the left, to achieve maximum job satisfaction.

##Model Evaluation To ensure that our model is contributing and performing better, we fit another model using the same datased by constraining the descrimination to be the same. Post which, we compared the group means between the two models to look into the overall performance of the model.

```
#fit the constrained model
fit_test <- grm(final, constrained = TRUE, IRT.param = TRUE)
fit_test$convergence
## [1] 0
fit_test
##
## Call:
## grm(data = final, constrained = TRUE, IRT.param = TRUE)
##
## Coefficients:
##      Extrmt1  Extrmt2  Extrmt3  Extrmt4  Extrmt5  Extrmt6  Extrmt7  Dscrmn
## A2c   -0.007   0.007    0.016    0.022    0.028    0.035    0.046   83.574
## A2d   -0.011   0.007    0.015    0.020    0.027    0.034    0.047   83.574
## A2e   -0.011   0.004    0.012    0.018    0.024    0.031    0.044   83.574
## A2f   -0.016  -0.004    0.006    0.012    0.018    0.025    0.041   83.574
## A2g   -0.019  -0.006    0.003    0.009    0.015    0.022    0.044   83.574
## A2h   -0.013  -0.005    0.002    0.012    0.018    0.025    0.036   83.574
## A3a   -0.017   0.003    0.013    0.018    0.024    0.032    0.049   83.574
## A3b   -0.019   0.001    0.012    0.018    0.025    0.033    0.053   83.574
## A3c   -0.018  -0.001    0.008    0.014    0.020    0.028    0.045   83.574
## A3d   -0.019  -0.006    0.005    0.012    0.018    0.026    0.039   83.574
## A3e   -0.020  -0.006    0.002    0.010    0.016    0.023    0.040   83.574
## A3f   -0.014   0.001    0.011    0.017    0.023    0.030    0.044   83.574
## A3g   -0.010   0.006    0.015    0.021    0.027    0.034    0.046   83.574
## A3h   -0.015  -0.003    0.005    0.013    0.019    0.026    0.040   83.574
## A3i   -0.020  -0.012   -0.003    0.006    0.012    0.019    0.033   83.574
## A3j   -0.021  -0.011    0.000    0.008    0.015    0.022    0.037   83.574
## A3k   -0.020  -0.010   -0.001    0.008    0.014    0.022    0.037   83.574
## A3l   -0.018  -0.010   -0.003    0.007    0.013    0.020    0.033   83.574
## A3m   -0.018  -0.008   -0.001    0.008    0.013    0.021    0.036   83.574
## B1a   -0.014   0.006    0.015    0.020    0.026    0.034    0.050   83.574
## B1b   -0.015   0.002    0.011    0.016    0.023    0.030    0.046   83.574
## B2a   -0.004   0.011    0.019    0.024    0.030    0.037    0.049   83.574
## B2b   -0.007   0.007    0.014    0.020    0.026    0.032    0.045   83.574
## B2c   -0.007   0.006    0.014    0.019    0.025    0.031    0.045   83.574
## B2d   -0.016  -0.005    0.000    0.007    0.012    0.017    0.027   83.574
## B3a   -0.011   0.007    0.016    0.023    0.029    0.037    0.047   83.574
## B3b   -0.008   0.008    0.017    0.023    0.030    0.037    0.047   83.574
## B3c   -0.015  -0.005    0.002    0.010    0.016    0.022    0.031   83.574
## B3d   -0.007   0.005    0.014    0.021    0.027    0.034    0.044   83.574
## B4a   -0.013   0.002    0.011    0.016    0.021    0.028    0.044   83.574
## B4b   -0.014  -0.001    0.009    0.016    0.022    0.028    0.040   83.574
## B4c   -0.012   0.000    0.007    0.014    0.019    0.025    0.038   83.574
## B4d   -0.018  -0.006    0.003    0.009    0.015    0.021    0.038   83.574
## B5a   -0.018  -0.009   -0.001    0.008    0.014    0.021    0.031   83.574
## B5b   -0.011   0.000    0.009    0.016    0.022    0.029    0.040   83.574
## B5c    0.000   0.011    0.018    0.025    0.031    0.037    0.048   83.574
## B5d   -0.005   0.011    0.019    0.026    0.032    0.039    0.049   83.574
## C1a   -0.007   0.006    0.014    0.021    0.027    0.034    0.045   83.574
## C1b   -0.010   0.003    0.010    0.016    0.022    0.028    0.040   83.574
## C1c   -0.012   0.000    0.006    0.013    0.018    0.023    0.033   83.574
## C1d    0.001   0.013    0.019    0.026    0.032    0.038    0.049   83.574
## C1e   -0.008   0.008    0.017    0.024    0.030    0.037    0.047   83.574
## C1f   -0.001   0.012    0.019    0.025    0.031    0.037    0.047   83.574
## C1g    0.009   0.015    0.019    0.025    0.028    0.033    0.038   83.574
##
## Log.Lik: -304983.1
#Now, we want to compare the constrained model with the unconstrained model.

#Fit test using comparison of group means
anova(fit_test, fit)
##
##  Likelihood Ratio Table
##               AIC      BIC   log.Lik       LRT df p.value
## fit_test 610584.2 612519.7 -304983.1
## fit      494732.6 496937.4 -247014.3 115937.62 43  <0.001
```

The p-value, when compared to a threshold value of 5%, is significant, suggesting the

The p-value, when compared to a threshold value of 5%, is significant, suggesting the negation of the null hypothesis that the model is insignificant.

We can see that the BIC and AIC of our model is considerably lower than the constrained model, suggesting improved model performance.

Conclusion and Recommendations From the results of this study, we are able to assess the effectiveness of the EPS in measuring job satisfaction. We also gain a better understanding of each question item's descriminating power by fitting a graded response model.

Public sector entities should focus on the results from the following question items as these highlight employee sentiment on job satisfaction better:

My agency inspires me to do the best in my job My agency motivates me to help it achieve its objectives I would recommend my agency as a great place to work My agency's senior leaders provide effective leadership I feel that my agency on the whole is well managed Public sector entities should further leverage the analysis of the following items with low discrimination results. This suggests a high chance of these items being generalised and could be important on a qualitative basis to the organisation. If they are deemed as irrelevant measures to the organisation, they can be excluded from further surveys owing to means of reduction of the same.

Purchasing decisions in my workplace are not influenced by gifts or incentives The people in my work group are committed to providing excellent customer service and making a positive difference to the community The people in my work group use their time and resources efficiently My immediate supervisor appropriately deals with employees who perform poorly In the past 12 months, my work group has implemented innovative processes or policies With the resulting test information curve, we determine that this subset of questions is able to distinguish employees which are dissatisfied with their jobs and are suitable for identifying action areas. As such, we recommend that PSC conducts an annual employee survey across all public sector entities using this as the preliminary study for a shortened evaluation. Feedback should be collected across the broader public sector to ensure comparability of data and better action planning to remediate issues or pain points.

With respect to the bifactor model as mentioned in the Exploratory Data Analysis section, we can see that there is an inherent latent trait wherein 3 factors are visualised and ignored by the IRT to focus on one latent trait. A possible approach would be to look for multidimensional IRT approaches for a better understanding of the contribution of these inherent traits as a factor or as a correlated factor. This can possibly open doors to more queries and thus formulating a better representative of the perceptions of the broader society.

With the analysis of the GRM modelling mentioned above, we can use items whose discrimination powers are relevant to the study and further use a subset approach towards modelling. We can conduct a similar analysis to evaluate the possible model performance of the same. If successful, we would have considerably lowered the items and thus looked into dimensionality reduction which is indeed important when compared to the lengthy survey.