

# Homework 3

## Applied Multivariate Analysis

Emorie Beck

September 22, 2018

## 1 Workspace

### 1.1 Packages

```
library(car)
library(knitr)
library(psych)
library(kableExtra)
library(multcomp)
library(lme4)
library(plyr)
library(tidyverse)
library(MVN)
```

### 1.2 data

The file, Set\_5.csv, contains data from a study in which college students completed the NEO-PI Personality Inventory. This 240-item scale purportedly measures the Big Five personality dimensions, assumed to be fairly independent. The inventory is scored on 6 subscales per dimension, listed below. The file contains the subscale scores, rather than the individual items, which should help reduce the impact of the small sample size.

Neuroticism: Anxiety Neuroticism: Angry\_Hostility Neuroticism: Depression Neuroticism: Self\_Consciousness  
Neuroticism: Impulsiveness Neuroticism: Vulnerability Extraversion: Warmth Extraversion: Gregariousness  
Extraversion: Assertiveness Extraversion: Activity Extraversion: Excitement\_Seeking Extraversion: Positive\_Emotions  
Openness: Fantasy Openness: Aesthetics Openness: Feelings Openness: Actions Openness: Ideas  
Openness: Values Agreeableness: Trust Agreeableness: Straightforwardness Agreeableness: Altruism  
Agreeableness: Compliance Agreeableness: Modesty Agreeableness: Tender\_Mindedness Conscientiousness: Competence  
Conscientiousness: Order Conscientiousness: Dutifulness Conscientiousness: Achievement\_Striving  
Conscientiousness: Self\_Discipline Conscientiousness: Deliberation

```
wd <- "https://github.com/emoriebeck/homeworks/raw/master/multivariate/homeworks/homework5"

dat <- sprintf("%s/Set_5(2).csv", wd) %>%
  read.csv(., stringsAsFactors = F)

head(dat)

##   ID Anxiety Angry_Hostility Depression Self_Consciousness Impulsiveness
## 1  2    2.625             2.000        1.750           2.250000         2.625
```

```
## 2 3 3.625 2.875 3.000 3.500000 4.250
## 3 4 3.000 2.750 2.625 2.875000 3.000
## 4 5 4.375 3.125 4.500 4.000000 3.875
## 5 6 3.500 2.875 3.000 2.571429 3.625
## 6 7 4.000 4.125 2.875 2.375000 4.000
## Vulnerability Warmth Gregariousness Assertiveness Activity
## 1 2.166667 4.666667 4.000 3.000000 4.833333
## 2 2.125000 4.500000 2.750 2.625000 3.000000
## 3 2.875000 3.750000 3.125 2.375000 3.250000
## 4 3.750000 3.250000 2.250 2.500000 1.875000
## 5 2.750000 3.750000 3.125 3.285714 3.500000
## 6 3.125000 3.500000 2.625 3.375000 3.125000
## Excitement_Seeking Positive_Emotions Fantasy Aesthetics Feelings
## 1 3.500 4.750 3.857143 3.571429 4.666667
## 2 2.875 3.500 3.500000 4.125000 3.625000
## 3 3.875 3.375 3.375000 3.500000 3.250000
## 4 2.750 2.625 3.000000 3.750000 4.250000
## 5 3.750 3.625 3.125000 1.625000 3.125000
## 6 2.000 3.375 3.500000 2.000000 3.250000
## Actions Ideas Values Trust Straightforwardness Altruism Compliance
## 1 2.571429 4.400 4.600 5.000 2.166667 4.833333 2.750
## 2 3.000000 3.875 3.125 3.250 3.750000 3.625000 3.125
## 3 2.375000 4.125 3.500 3.250 3.125000 4.000000 3.750
## 4 3.375000 2.750 4.125 3.000 3.428571 3.875000 4.000
## 5 2.750000 2.500 3.625 3.375 3.250000 4.125000 3.625
## 6 2.625000 1.125 3.625 2.500 2.875000 3.000000 2.250
## Modesty Tender_Mindedness Competence Order Dutifulness
## 1 4.000 3.833333 4.50 3.625 3.285714
## 2 2.625 3.250000 3.00 2.250 3.875000
## 3 2.750 3.250000 3.75 3.250 3.750000
## 4 4.125 3.750000 2.75 3.000 2.875000
## 5 3.375 3.375000 3.75 4.000 3.750000
## 6 2.625 3.375000 3.00 3.625 2.625000
## Achievement_Striving Self_Discipline Deliberation
## 1 4.333333 4.250 2.875
## 2 2.750000 3.750 3.500
## 3 3.375000 3.375 3.125
## 4 2.875000 2.625 3.250
## 5 3.375000 2.875 3.375
## 6 3.000000 2.625 2.625
```

```
source <- tribble(
  ~Factor, ~Facet,
  "Neuroticism", "Anxiety",
  "Neuroticism", "Angry_Hostility",
  "Neuroticism", "Depression",
  "Neuroticism", "Self_Consciousness",
  "Neuroticism", "Impulsiveness",
  "Neuroticism", "Vulnerability",
  "Extraversion", "Warmth",
  "Extraversion", "Gregariousness",
  "Extraversion", "Assertiveness",
  "Extraversion", "Activity",
```

```

"Extraversion", "Excitement_Seeking",
"Extraversion", "Positive_Emotions",
"Openness", "Fantasy",
"Openness", "Aesthetics",
"Openness", "Feelings",
"Openness", "Actions",
"Openness", "Ideas",
"Openness", "Values",
"Agreeableness", "Trust",
"Agreeableness", "Straightforwardness" ,
"Agreeableness", "Altruism",
"Agreeableness", "Compliance",
"Agreeableness", "Modesty",
"Agreeableness", "Tender_Mindedness",
"Conscientiousness", "Competence",
"Conscientiousness", "Order",
"Conscientiousness", "Dutifulness",
"Conscientiousness", "Achievement_Striving",
"Conscientiousness", "Self_Discipline",
"Conscientiousness", "Deliberation"
)

dat <- dat %>% select(ID, source$Facet)

```

## 2 The Question

Given what you have learned up through exploratory factor analysis, analyze the data in the way you think is appropriate and form conclusions about the claimed number of dimensions and their independence.

### 2.1 Do we need FA?

#### 2.1.1 Correlations

```

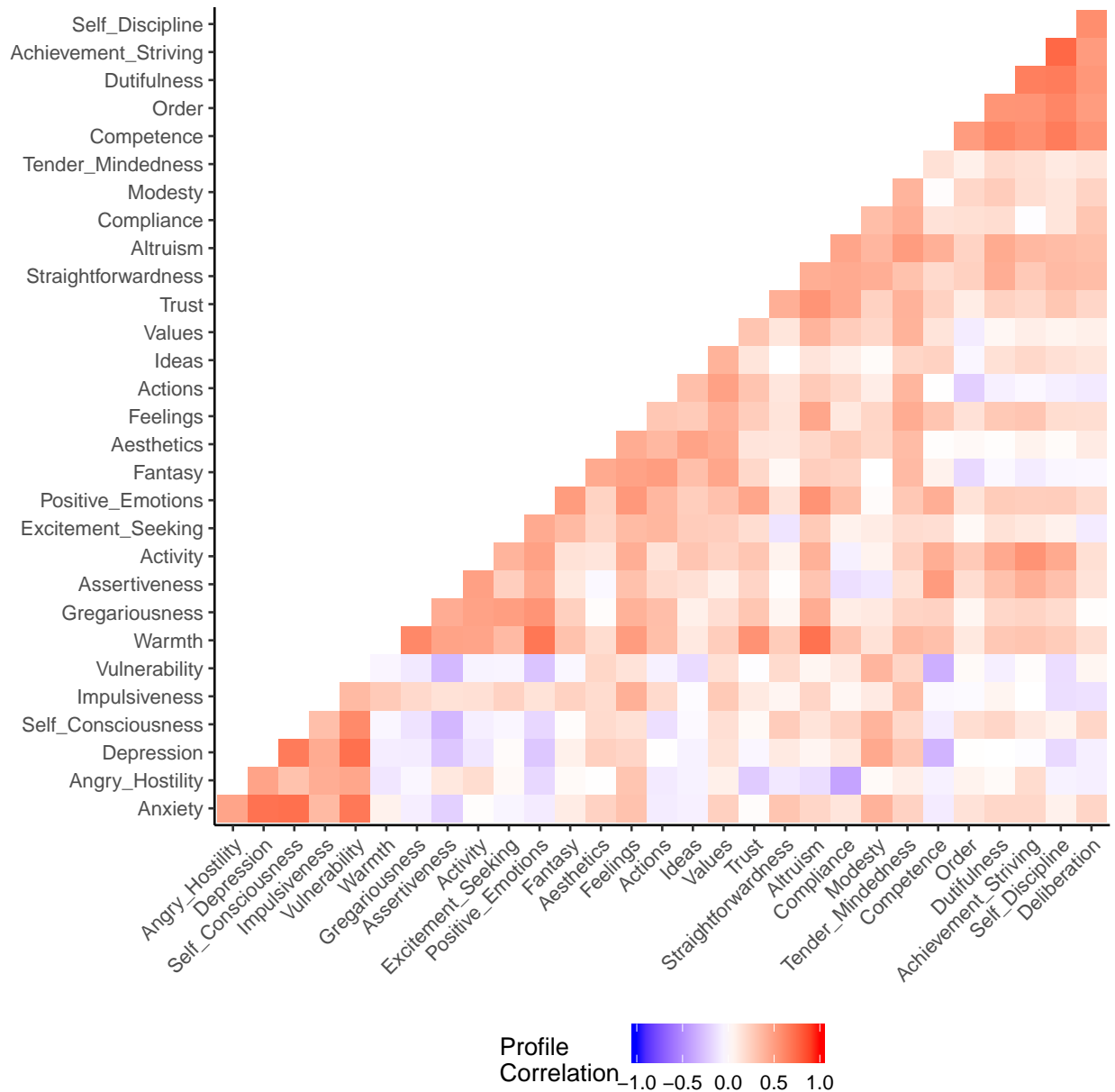
r <- r_long <- dat %>% select(-ID) %>% cor()
r_long[upper.tri(r_long, diag = T)] <- NA
order <- colnames(dat)[-1]

r_long <- r_long %>% data.frame %>%
  mutate(V1 = rownames(.)) %>%
  gather(key = V2, value = r, -V1, na.rm = T) %>%
  mutate(V1 = factor(V1, levels = order),
         V2 = factor(V2, levels = order))

r_long %>%
  ggplot(aes(x = V1, y = V2, fill = r)) +
  geom_raster() +
  scale_fill_gradient2(low = "blue", high = "red", mid = "white",
                      midpoint = 0, limit = c(-1,1), space = "Lab",
                      name="Profile\nCorrelation") +
  theme_classic() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1),
        legend.position = "bottom",

```

```
axis.title = element_blank())
```



There appear to be intercorrelations among the variables.

## 2.1.2 KMO

```
(KMO1 <- KMO(r))

## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = r)
## Overall MSA = 0.86
## MSA for each item =
##           Anxiety      Angry_Hostility      Depression
##           0.85         0.74             0.81
```

##	Self_Consciousness	Impulsiveness	Vulnerability
##	0.89	0.87	0.85
##	Warmth	Gregariousness	Assertiveness
##	0.88	0.88	0.87
##	Activity	Excitement_Seeking	Positive_Emotions
##	0.88	0.85	0.89
##	Fantasy	Aesthetics	Feelings
##	0.84	0.80	0.90
##	Actions	Ideas	Values
##	0.85	0.78	0.84
##	Trust	Straightforwardness	Altruism
##	0.90	0.80	0.91
##	Compliance	Modesty	Tender_Mindedness
##	0.79	0.85	0.90
##	Competence	Order	Dutifulness
##	0.89	0.88	0.90
##	Achievement_Striving	Self_Discipline	Deliberation
##	0.89	0.87	0.86

The MSA for each item range from 1 to 1, with a mean of 1, indicating strong evidence for using a data reduction technique.

### 2.1.3 Bartlett's Test

```
(CB_1 <- corstest.bartlett(R=r,n=nrow(dat)))

## $chisq
## [1] 3707.648
##
## $p.value
## [1] 0
##
## $df
## [1] 435
```

In addition, the  $\chi^2$  value of the Bartlett test ( $\chi^2(435) = 3707.65$ ), which indicates that the correlation matrix departs significantly from an identity matrix (independence among indicators).

## 2.2 How Many Factors?

Now that we have seen evidence suggesting that we should conduct a CFA or PCA, we need to determine how many factors we should extract.

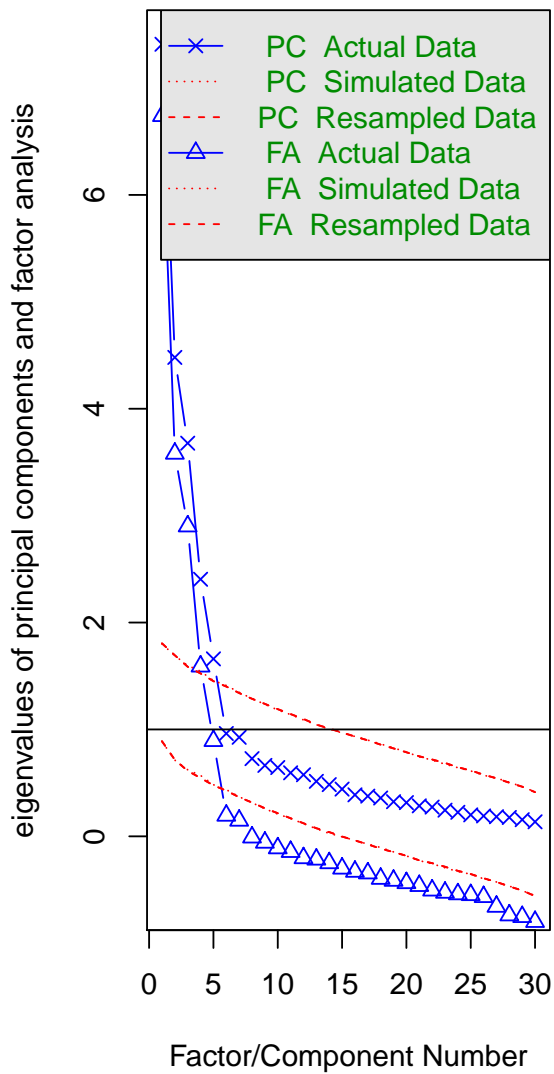
### 2.2.1 Parallel Analysis (Scree Test)

```
par(mfrow=c(1,2))
scree_1 <- fa.parallel(dat %>% select(-ID), fa="both")

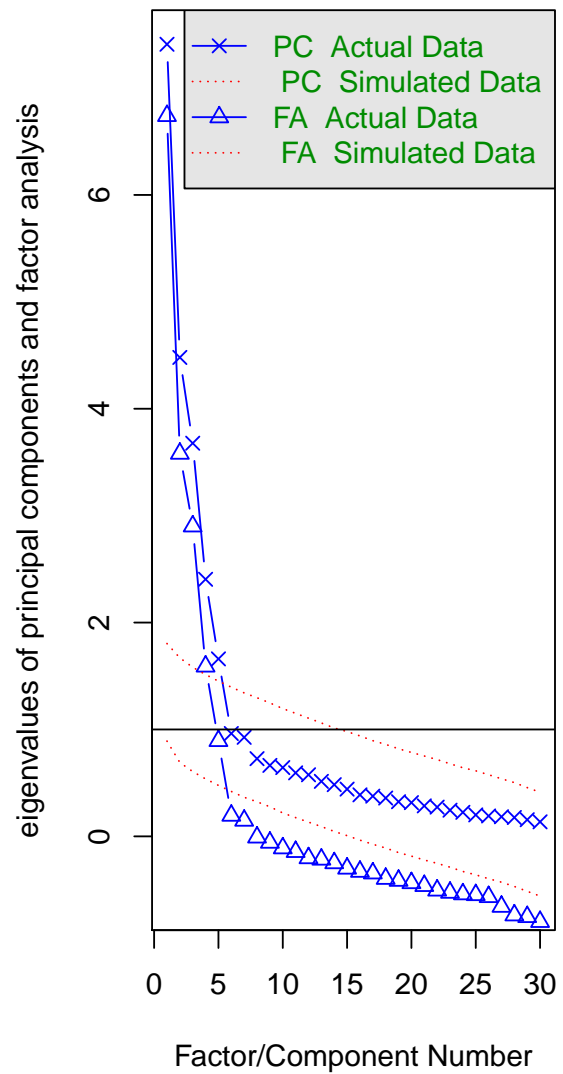
## Parallel analysis suggests that the number of factors = 5 and the number of components = 5

scree_2 <- fa.parallel(r, fa = "both", n.obs = nrow(dat))
```

Parallel Analysis Scree Plots



Parallel Analysis Scree Plots

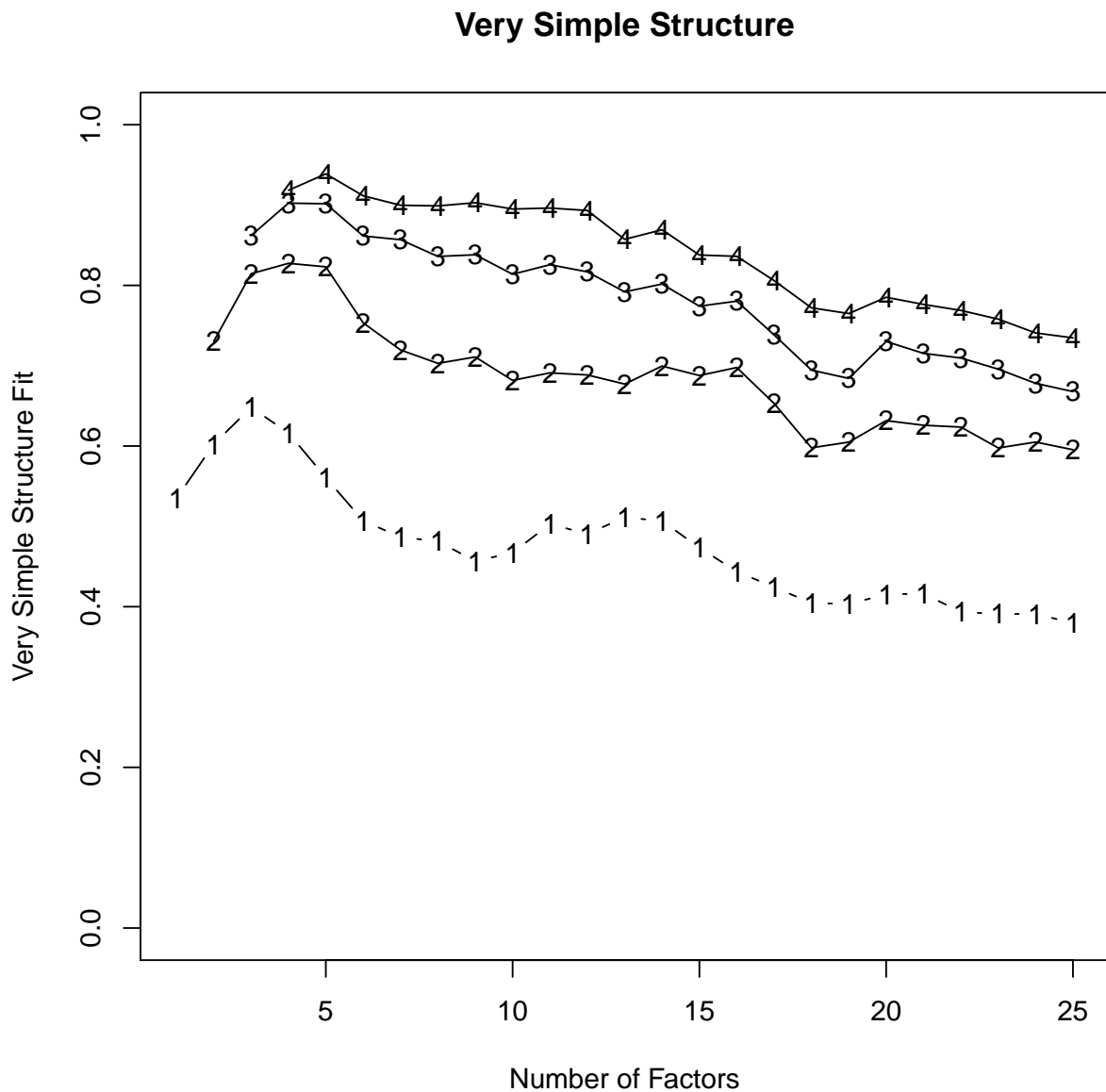


```
## Parallel analysis suggests that the number of factors = 5 and the number of components = 5
```

Parallel analysis suggests 5 principal components and 5 factors.

### 2.2.2 VSS

```
par(mfrow = c(1,1))
vss_1 <- vss(dat %>% select(-ID), n = 25, rotate = "none", fm = "pc")
```



VSS also suggests 5 factors.

## 2.3 Exploratory Factor Analysis

```
fa_1 <- fa(dat %>% select(-ID), nfactors = 5, rotate = "none", scores = T)
fa_2 <- fa(dat %>% select(-ID), nfactors = 5, rotate = "varimax", scores = T)
fa_3 <- fa(dat %>% select(-ID), nfactors = 5, rotate = "oblimin", scores = T)

scores_1 <- fa_1$scores
scores_2 <- fa_2$scores
scores_3 <- fa_3$scores

# unrotated
cor(scores_1) %>% round(., 2)
```

```
##          MR1  MR2  MR3  MR4  MR5
## MR1    1.00  0.00  0.01  0.00 -0.01
## MR2    0.00  1.00  0.00  0.01  0.00
## MR3    0.01  0.00  1.00  0.01  0.02
## MR4    0.00  0.01  0.01  1.00  0.00
## MR5   -0.01  0.00  0.02  0.00  1.00

# varimax rotation
cor(scores_2) %>% round(., 2)

##          MR3  MR2  MR1  MR5  MR4
## MR3    1.00  0.00  0.04 -0.01  0.02
## MR2    0.00  1.00 -0.01  0.02  0.01
## MR1    0.04 -0.01  1.00  0.08  0.02
## MR5   -0.01  0.02  0.08  1.00  0.04
## MR4    0.02  0.01  0.02  0.04  1.00

# oblimin rotation
cor(scores_3) %>% round(., 2)

##          MR3  MR2  MR1  MR5  MR4
## MR3    1.00  0.02  0.32  0.11  0.09
## MR2    0.02  1.00  0.00  0.11  0.04
## MR1    0.32  0.00  1.00  0.46  0.07
## MR5    0.11  0.11  0.46  1.00  0.11
## MR4    0.09  0.04  0.07  0.11  1.00
```

The two models fit the data relatively well ( $RMSEA_{unrotated} = 0.0677843$ ,  $RMSEA_{varimax} = 0.0677843$ ,  $RMSEA_{oblimin} = 0.0677843$ ,  $TLL_{unrotated} = 0.8938015$ ;  $TLL_{varimax} = 0.8938015$ ;  $TLL_{oblimin} = 0.8938015$ ).

But we aren't just concerned with model fit. We are also generally interested in naming the factors.

There's no way for me to pretend I don't have expectations for how the data should come out. So let's look at the rotated and unrotated solutions and see if we managed to recover the Big 5.

```
fa_1$Structure %>% unclass %>%
  data.frame %>%
  mutate(Facet = rownames(.)) %>%
  full_join(source) %>%
  select(Factor, Facet, MR1, MR2, MR3, MR4, MR5) %>%
  mutate_at(vars(MR1:MR5), funs(round(., 2))) %>%
  mutate_at(vars(MR1:MR5), funs(cell_spec(., "latex",
    background = ifelse(., > .5, "yellow", "white")))) %>%
  mutate(Facet = str_remove_all(Facet, "_")) %>%
  kable(., "latex", escape = F, booktabs = T,
    caption = "Unrotated Solution") %>%
  kable_styling(full_width = F)
```

```
fa_2$Structure %>% unclass %>%
  data.frame %>%
  mutate(Facet = rownames(.)) %>%
  full_join(source) %>%
  select(Factor, Facet, MR1, MR2, MR3, MR4, MR5) %>%
  mutate_at(vars(MR1:MR5), funs(round(., 2))) %>%
  mutate_at(vars(MR1:MR5), funs(cell_spec(., "latex",
    background = ifelse(abs(.) > .5, "yellow", "white")))) %>%
```



Table 1: Unrotated Solution

Factor	Facet	MR1	MR2	MR3	MR4	MR5
Neuroticism	Anxiety	0.26	0.78	0.25	0.09	0
Neuroticism	AngryHostility	0.03	0.47	0.12	0.69	0.02
Neuroticism	Depression	0.08	0.87	0.05	0.12	-0.05
Neuroticism	SelfConsciousness	0.17	0.72	0.28	-0.04	0.03
Neuroticism	Impulsiveness	0.29	0.43	-0.22	0.31	-0.18
Neuroticism	Vulnerability	0.05	0.81	0.12	0.07	-0.11
Extraversion	Warmth	0.73	-0.14	-0.24	-0.01	-0.38
Extraversion	Gregariousness	0.53	-0.2	-0.27	0.19	-0.34
Extraversion	Assertiveness	0.47	-0.42	-0.04	0.38	-0.12
Extraversion	Activity	0.62	-0.21	0.02	0.38	-0.02
Extraversion	ExcitementSeeking	0.42	-0.07	-0.35	0.22	-0.01
Extraversion	PositiveEmotions	0.7	-0.3	-0.32	-0.01	-0.08
Openness	Fantasy	0.41	0.1	-0.55	0	0.24
Openness	Aesthetics	0.37	0.3	-0.3	-0.1	0.46
Openness	Feelings	0.66	0.19	-0.22	0.27	0.07
Openness	Actions	0.38	0.01	-0.57	-0.09	0.12
Openness	Ideas	0.35	-0.09	-0.23	0.02	0.57
Openness	Values	0.47	0.24	-0.35	-0.05	0.27
Agreeableness	Trust	0.57	-0.06	-0.12	-0.31	-0.2
Agreeableness	Straightforwardness	0.43	0.19	0.26	-0.4	-0.1
Agreeableness	Altruism	0.78	0.03	-0.04	-0.21	-0.24
Agreeableness	Compliance	0.42	0.17	-0.06	-0.67	-0.05
Agreeableness	Modesty	0.36	0.43	0.15	-0.24	-0.09
Agreeableness	TenderMindedness	0.55	0.29	-0.19	-0.13	0
Conscientiousness	Competence	0.6	-0.42	0.35	0.07	0.15
Conscientiousness	Order	0.39	-0.08	0.59	0.03	0.05
Conscientiousness	Dutifulness	0.61	-0.11	0.52	0.04	0.03
Conscientiousness	AchievementStriving	0.6	-0.14	0.52	0.25	0.11
Conscientiousness	SelfDiscipline	0.57	-0.28	0.57	0	0.14
Conscientiousness	Deliberation	0.46	-0.06	0.52	-0.2	0.19

Table 2: Varimax Rotated Solution

Factor	Facet	MR1	MR2	MR3	MR4	MR5
Neuroticism	Anxiety	-0.05	0.84	0.16	0.1	0.07
Neuroticism	AngryHostility	0.14	0.62	0.07	-0.55	0
Neuroticism	Depression	-0.07	0.86	-0.13	0.03	0.08
Neuroticism	SelfConsciousness	-0.18	0.74	0.14	0.17	0.05
Neuroticism	Impulsiveness	0.4	0.5	-0.12	-0.07	0.13
Neuroticism	Vulnerability	-0.09	0.82	-0.1	0.08	-0.02
Extraversion	Warmth	0.77	0	0.15	0.37	0.1
Extraversion	Gregariousness	0.72	-0.06	0.07	0.09	0.05
Extraversion	Assertiveness	0.6	-0.22	0.34	-0.18	0.04
Extraversion	Activity	0.57	0	0.44	-0.13	0.17
Extraversion	ExcitementSeeking	0.5	-0.01	0.01	-0.04	0.31
Extraversion	PositiveEmotions	0.65	-0.21	0.19	0.25	0.33
Openness	Fantasy	0.32	0.03	-0.15	0.1	0.63
Openness	Aesthetics	0.01	0.21	0.02	0.14	0.69
Openness	Feelings	0.53	0.3	0.2	0	0.43
Openness	Actions	0.36	-0.07	-0.2	0.19	0.53
Openness	Ideas	0.04	-0.13	0.19	-0.05	0.67
Openness	Values	0.22	0.19	0	0.17	0.6
Agreeableness	Trust	0.39	-0.02	0.15	0.53	0.14
Agreeableness	Straightforwardness	0.03	0.22	0.32	0.55	0
Agreeableness	Altruism	0.53	0.13	0.3	0.54	0.16
Agreeableness	Compliance	0.01	0.07	0.07	0.78	0.21
Agreeableness	Modesty	0.03	0.44	0.16	0.41	0.07
Agreeableness	TenderMindedness	0.31	0.3	0.08	0.36	0.36
Conscientiousness	Competence	0.27	-0.24	0.74	0.06	0.12
Conscientiousness	Order	0.02	0.1	0.69	0.09	-0.11
Conscientiousness	Dutifulness	0.2	0.1	0.76	0.16	-0.02
Conscientiousness	AchievementStriving	0.24	0.11	0.81	-0.05	0.03
Conscientiousness	SelfDiscipline	0.12	-0.08	0.84	0.13	0.01
Conscientiousness	Deliberation	-0.08	0.05	0.68	0.28	0.07

```

mutate(Facet = str_remove_all(Facet, "_")) %>%
kable(., "latex", escape = F, booktabs = T,
      caption = "Varimax Rotated Solution") %>%
kable_styling(full_width = F)

```

```

fa_3$Structure %>% unclass %>%
data.frame %>%
mutate(Facet = rownames(.)) %>%
full_join(source) %>%
select(Factor, Facet, MR1, MR2, MR3, MR4, MR5) %>%
mutate_at(vars(MR1:MR5), funs(round(., 2))) %>%
mutate_at(vars(MR1:MR5), funs(cell_spec(., "latex",
  background = ifelse(abs(.) > .5, "yellow", "white")))) %>%
mutate(Facet = str_remove_all(Facet, "_")) %>%
kable(., "latex", escape = F, booktabs = T,

```

Table 3: Oblimin Rotated Solution

Factor	Facet	MR1	MR2	MR3	MR4	MR5
Neuroticism	Anxiety	0.02	0.85	0.16	0.05	0.12
Neuroticism	AngryHostility	0.01	0.55	0.04	-0.61	0.02
Neuroticism	Depression	-0.05	0.87	-0.12	-0.01	0.1
Neuroticism	SelfConsciousness	-0.08	0.77	0.14	0.16	0.06
Neuroticism	Impulsiveness	0.37	0.48	-0.07	-0.18	0.24
Neuroticism	Vulnerability	-0.06	0.82	-0.1	0.04	0.01
Extraversion	Warmth	0.86	0	0.3	0.19	0.32
Extraversion	Gregariousness	0.72	-0.09	0.18	-0.06	0.23
Extraversion	Assertiveness	0.57	-0.26	0.4	-0.28	0.17
Extraversion	Activity	0.58	-0.03	0.5	-0.24	0.31
Extraversion	ExcitementSeeking	0.5	-0.02	0.08	-0.13	0.42
Extraversion	PositiveEmotions	0.75	-0.2	0.31	0.14	0.5
Openness	Fantasy	0.4	0.05	-0.08	0.06	0.69
Openness	Aesthetics	0.14	0.25	0.05	0.16	0.69
Openness	Feelings	0.59	0.3	0.28	-0.11	0.57
Openness	Actions	0.43	-0.05	-0.12	0.14	0.61
Openness	Ideas	0.13	-0.1	0.2	-0.01	0.66
Openness	Values	0.34	0.22	0.06	0.14	0.65
Agreeableness	Trust	0.54	0.01	0.26	0.44	0.28
Agreeableness	Straightforwardness	0.22	0.27	0.38	0.52	0.08
Agreeableness	Altruism	0.71	0.16	0.44	0.41	0.35
Agreeableness	Compliance	0.25	0.16	0.15	0.76	0.28
Agreeableness	Modesty	0.17	0.48	0.21	0.37	0.14
Agreeableness	TenderMindedness	0.45	0.33	0.16	0.28	0.47
Conscientiousness	Competence	0.37	-0.23	0.77	0.03	0.2
Conscientiousness	Order	0.11	0.11	0.69	0.08	-0.06
Conscientiousness	Dutifulness	0.32	0.11	0.8	0.11	0.09
Conscientiousness	AchievementStriving	0.31	0.1	0.83	-0.1	0.12
Conscientiousness	SelfDiscipline	0.25	-0.06	0.86	0.12	0.08
Conscientiousness	Deliberation	0.09	0.09	0.69	0.29	0.1

```
caption = "Oblimin Rotated Solution") %>%
kable_styling(full_width = F)
```

With the exception of the 2nd and 3rd factors, the unrotated solution doesn't resemble the expected solution. However, the indicators for each factor in the varimax rotated solution can clearly be identified as the Big 5 by content. Finally, in the oblimin rotated solution, we see that the factors can be fairly readily identified by content.

Another fun test is the order of extraction, which is typically Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience.

In this case, the order of extraction for the rotated solutions appears to be Extraversion, Neuroticism, Conscientiousness, Agreeableness, Openness.