Customer Satisfaction Index

Shankar MM, Mentor - Cares,Bangalore

09/30/2018

note: Sepcial thanks to Dr. Kantaraj Tavane, Research Scholar from VTU, who gave permission to use partial data for demo purpose

Exploratory factor analysis (EFA)

This section dealt with how to use EFA analysis especially using perceptual data which measures various constructs such as customer satisfaction, employee commitment or investor bias which is normally captured in the form of rating scale. Though PCA and EFA both are used for dimension reduction, but theoretically and philosophically both are different, especially EFA dealt with latent but PCA is able to reduce data in the form of images, pictures, videos and financial data etc., former may not concern interpretation of the model, but latter is very much used to not only to reduce the dimension but bring meaning out of dimension formation.

Steps followed in EFA

- Converting raw data into correlation matrix
- Applying Eigen value on correlation matrix to determine number of factors, especially in academic research, when eigen value is greater than 1, it become factor
- Estimation method is Maximum likelihood and rotation is based on oblimin than varimax, because most of the time, there is correlation among the derived factors.
- Assumptions check, 1. KMO test for measure of sampling adequacy and 2. Bartlet test of sphericity for factor correlability.
- Factor loadings are kept minimum of .35
- Factors are named based on domain knowledge with help of prior literature

```r
str(csi)  # to display information about data set
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame':    150 obs. of  42 variables:
##  $ ID           : num  1 2 3 4 5 6 7 8 9 10 ...
##  $ Gender       : chr  "Female" "Male" "Male" "Male" ...
##  $ Location     : chr  "City" "City" "City" "Home town" ...
##  $ Profession   : chr  "PVT" "PVT" "PVT" "PVT" ...
##  $ Age          : chr  "25-32 yrs" "25-32 yrs" "25-32 yrs" "18-25 yrs" ...
##  $ Qualification: chr  "Others" "Others" "Others" "PG" ...
##  $ Experience   : chr  "5-10 yrs" "LT 2 yrs" "5-10 yrs" "LT 2 yrs" ...
##  $ Freq_visit   : chr  "Once in 2 Months" "Once in 3 months" "once in GT 6 months" "Once in a Month"
##  $ Whom         : chr  "Others" "Others" "Others" "Children" ...
##  $ Occasion     : chr  "Others" "Discount" "Others" "Wedding" ...
##  $ price        : num  7 1 7 5 2 7 5 5 2 10 ...
##  $ fq           : num  1 6 1 6 1 1 2 2 5 3 ...
##  $ su           : num  8 7 9 7 7 9 3 3 10 6 ...
##  $ vom          : num  9 10 8 8 6 8 4 4 9 7 ...
##  $ poo          : num  6 8 10 4 10 10 6 6 8 8 ...
##  $ ff           : num  5 9 2 10 9 2 8 8 6 1 ...
##  $ cc           : num  2 5 6 9 8 6 7 7 4 9 ...
##  $ br           : num  10 2 3 3 3 3 9 9 7 5 ...
##  $ dur          : num  3 3 4 1 4 4 1 1 1 2 ...
##  $ sty          : num  4 4 5 2 5 5 10 10 3 4 ...
```

```
##  $ 1.col        : num  10 9 5 9 9 8 9 9 8 5 ...
##  $ 2.tex        : num  7 5 6 2 8 9 8 8 5 7 ...
##  $ 3.hf         : num  8 10 5 4 7 7 4 4 9 6 ...
##  $ 4.fbpat      : num  9 9 5 7 5 4 5 5 8 10 ...
##  $ 5.sty        : num  7 6 5 8 9 5 9 9 9 8 ...
##  $ 6.vom        : num  8 6 3 7 4 10 3 3 4 10 ...
##  $ 7.cc         : num  10 5 3 5 10 7 5 5 9 7 ...
##  $ 8.cct        : num  9 4 3 3 5 6 4 4 8 9 ...
##  $ 9.seams      : num  8 9 4 6 8 5 6 6 9 10 ...
##  $ 10.care      : num  9 5 4 10 6 6 5 5 3 5 ...
##  $ 11.trnd      : num  8 10 5 9 4 4 2 2 2 9 ...
##  $ 12.drapes    : num  6 5 7 3 9 5 6 6 1 6 ...
##  $ 13.fit       : num  9 6 7 3 3 3 2 2 7 8 ...
##  $ 14.conf      : num  10 4 7 6 8 1 9 9 10 10 ...
##  $ 15.suits     : num  9 9 7 6 6 2 2 2 9 10 ...
##  $ 16.dur       : num  9 10 6 8 8 2 9 1 10 8 ...
##  $ 17.season    : num  8 9 7 9 9 2 5 3 8 10 ...
##  $ 18.limp      : num  7 7 5 9 6 3 4 6 5 7 ...
##  $ 19.puck      : num  6 9 7 9 6 3 7 2 3 9 ...
##  $ 20.shri      : num  6 10 6 8 3 4 3 3 5 10 ...
##  $ 21.cf        : num  7 8 8 8 1 4 6 2 1 9 ...
##  $ 22.m/c wash  : num  6 3 7 9 1 4 8 4 1 10 ...
```

```r
csi1 <- csi[,21:42] # rating scale items is only selected for EFA model
kable(describe(csi1),caption = "Descriptive statistics of rating items", digits = 2)
```

Table 1: Descriptive statistics of rating items

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
1.col	1	150	5.93	3.03	6.5	6.04	3.71	1	10	9	-0.36	-1.14	0.25
2.tex	2	150	5.97	2.54	6.0	6.01	2.97	1	10	9	-0.19	-0.97	0.21
3.hf	3	150	6.07	2.59	6.0	6.17	2.97	1	10	9	-0.31	-0.91	0.21
4.fbpat	4	150	6.25	2.65	6.5	6.40	2.22	1	10	9	-0.41	-0.79	0.22
5.sty	5	150	6.59	2.68	7.0	6.79	2.97	1	10	9	-0.54	-0.87	0.22
6.vom	6	150	6.23	2.52	6.0	6.29	2.97	1	10	9	-0.16	-0.97	0.21
7.cc	7	150	6.58	2.61	7.0	6.78	2.97	1	10	9	-0.56	-0.73	0.21
8.cct	8	150	6.21	2.53	7.0	6.33	2.97	1	10	9	-0.39	-0.78	0.21
9.seams	9	150	6.50	2.72	7.0	6.73	2.97	1	10	9	-0.61	-0.69	0.22
10.care	10	150	6.21	2.54	6.0	6.35	2.97	1	10	9	-0.34	-0.66	0.21
11.trnd	11	150	6.57	2.63	7.0	6.77	2.97	1	10	9	-0.54	-0.80	0.21
12.drapes	12	150	6.20	2.51	6.0	6.34	2.97	1	10	9	-0.36	-0.69	0.20
13.fit	13	150	6.37	2.57	7.0	6.51	2.97	1	10	9	-0.39	-0.87	0.21
14.conf	14	150	6.72	2.76	7.0	7.00	2.97	1	10	9	-0.68	-0.60	0.23
15.suits	15	150	6.55	2.69	7.0	6.76	2.97	1	10	9	-0.52	-0.83	0.22
16.dur	16	150	6.51	2.75	7.0	6.74	2.97	1	10	9	-0.53	-0.75	0.22
17.season	17	150	6.27	2.48	7.0	6.42	2.97	1	10	9	-0.46	-0.58	0.20
18.limp	18	150	5.85	2.53	6.0	5.88	2.97	1	10	9	-0.07	-0.97	0.21
19.puck	19	150	5.80	2.67	6.0	5.84	2.97	1	10	9	-0.17	-0.98	0.22
20.shri	20	150	5.67	2.65	6.0	5.72	2.97	1	10	9	-0.09	-1.06	0.22
21.cf	21	150	5.82	2.72	6.0	5.94	2.97	1	10	9	-0.30	-1.07	0.22
22.m/c wash	22	150	6.04	3.02	7.0	6.17	2.97	1	10	9	-0.36	-1.22	0.25

Exploratory Factor Analysis

Normally, expected correlation score is ranged between .10 to .70, any thing below .10 or greater than .70, need careful investigation further. based on eigen value greate than 1, 3 factors are derived, Assumpitons test such as KMO and Bartlett test are done. both are meeting the threshold values. Scree plot and factor loadings are obtained.

```
satcor<- corr.test(csi1[,-c(8,10,12,18)])  # further items are eliminated in EFA
satcor<-satcor$r  # storing correlation matrix
kable(satcor,digits = 2,caption = "Correlation matrix")
```

Table 2: Correlation matrix

	1.col	2.tex	3.hf	4.fbpat	5.sty	6.vom	7.cc	9.seams	11.trnd	13.fit	14.conf	15.suits	16.dur
1.col	1.00	0.71	0.69	0.62	0.57	0.42	0.41	0.33	0.42	0.32	0.51	0.43	0.32
2.tex	0.71	1.00	0.68	0.55	0.54	0.37	0.39	0.32	0.43	0.44	0.47	0.43	0.30
3.hf	0.69	0.68	1.00	0.52	0.49	0.43	0.45	0.44	0.52	0.51	0.50	0.49	0.37
4.fbpat	0.62	0.55	0.52	1.00	0.60	0.50	0.43	0.49	0.45	0.41	0.51	0.51	0.35
5.sty	0.57	0.54	0.49	0.60	1.00	0.47	0.54	0.43	0.49	0.47	0.53	0.55	0.39
6.vom	0.42	0.37	0.43	0.50	0.47	1.00	0.48	0.41	0.42	0.36	0.36	0.39	0.34
7.cc	0.41	0.39	0.45	0.43	0.54	0.48	1.00	0.49	0.52	0.44	0.50	0.50	0.40
9.seams	0.33	0.32	0.44	0.49	0.43	0.41	0.49	1.00	0.52	0.48	0.55	0.48	0.46
11.trnd	0.42	0.43	0.52	0.45	0.49	0.42	0.52	0.52	1.00	0.62	0.56	0.65	0.49
13.fit	0.32	0.44	0.51	0.41	0.47	0.36	0.44	0.48	0.62	1.00	0.60	0.70	0.55
14.conf	0.51	0.47	0.50	0.51	0.53	0.36	0.50	0.55	0.56	0.60	1.00	0.71	0.48
15.suits	0.43	0.43	0.49	0.51	0.55	0.39	0.50	0.48	0.65	0.70	0.71	1.00	0.54
16.dur	0.32	0.30	0.37	0.35	0.39	0.34	0.40	0.46	0.49	0.55	0.48	0.54	1.00
17.season	0.43	0.45	0.49	0.48	0.49	0.41	0.37	0.47	0.54	0.45	0.54	0.58	0.51
19.puck	0.43	0.45	0.45	0.43	0.47	0.39	0.37	0.30	0.44	0.38	0.35	0.46	0.35
20.shri	0.23	0.38	0.29	0.40	0.41	0.39	0.37	0.27	0.46	0.36	0.23	0.43	0.27
21.cf	0.33	0.45	0.30	0.40	0.39	0.38	0.37	0.31	0.51	0.43	0.29	0.41	0.38
22.m/c wash	0.23	0.36	0.30	0.38	0.40	0.43	0.34	0.38	0.39	0.39	0.34	0.30	0.44

```
ev <- eigen(satcor) # get eigenvalues
corMat<-ev$values

kable(data.frame(sn=seq(1,length(corMat),1),corMat),caption = "Eigen values",col.names = c("SN","Eigen")
```

Table 3: Eigen values

SN	Eigen
1	8.62
2	1.50
3	1.32
4	0.87
5	0.74
6	0.63
7	0.57
8	0.54
9	0.51
10	0.44
11	0.40
12	0.39
13	0.35
14	0.31

SN	Eigen
15	0.27
16	0.20
17	0.18
18	0.17

```
kmo<-KMO(satcor)
paste("Overall kmo:",round(kmo$MSA,3))
```

```
## [1] "Overall kmo: 0.916"
```

```
kable(kmo$MSAi,digits = 3,caption = "Measure of sample Adequacy",col.names = "KMO item wise")
```

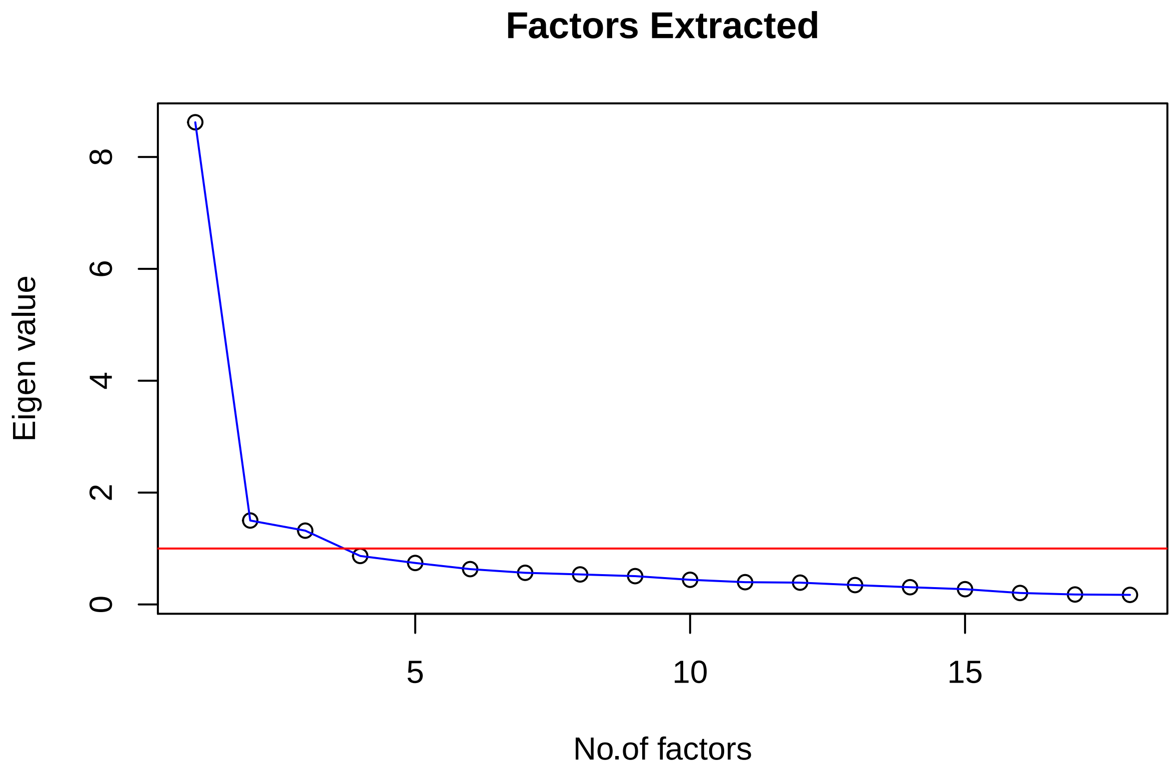
Table 4: Measure of sample Adequacy

	KMO item wise
1.col	0.858
2.tex	0.902
3.hf	0.901
4.fbpat	0.947
5.sty	0.959
6.vom	0.951
7.cc	0.950
9.seams	0.938
11.trnd	0.949
13.fit	0.913
14.conf	0.924
15.suits	0.915
16.dur	0.940
17.season	0.935
19.puck	0.889
20.shri	0.889
21.cf	0.852
22.m/c wash	0.886

```
barlettest <- cortest.bartlett(satcor,nrow(csi1))
barletresult <- data.frame(barlettest$chisq,barlettest$df,barlettest$p.value)
kable(barletresult,col.names = c("chisq","DF","P value"))
```

chisq	DF	P value
1596.524	153	0

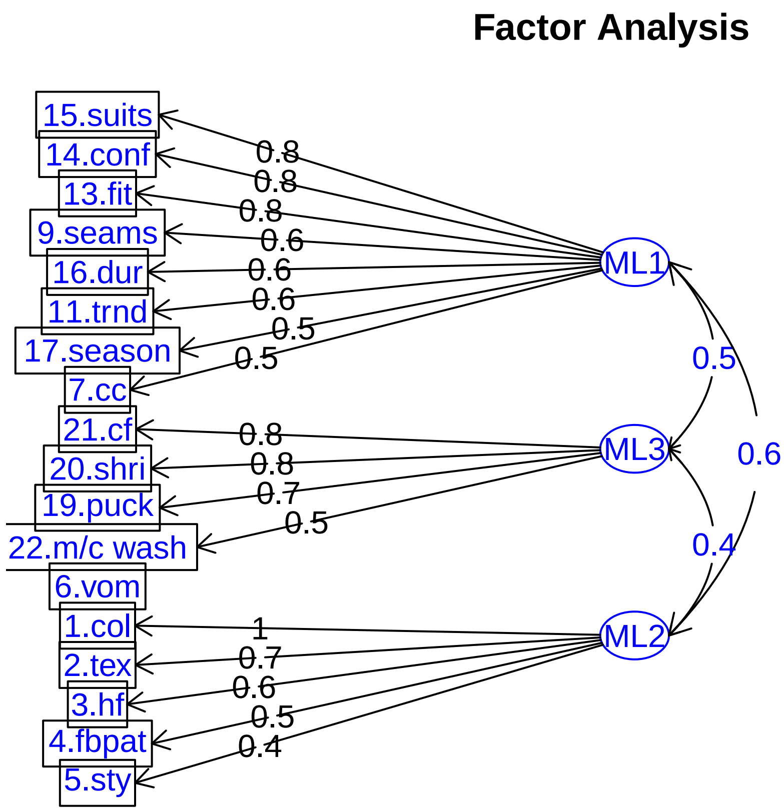
```
plot(corMat,main = "Factors Extracted",ylab = "Eigen value",xlab="No.of factors")
lines(corMat,col="blue")
abline(h=1,col="red")
```




```r
pa.out <- fa(satcor, nfactors= sum(corMat > 1.0), n.obs=nrow(csi), rotate="oblimin", scores="regression
print(pa.out$loadings,sort=T, cutoff=0.3, digits=2)
```

```
##
## Loadings:
##              ML1   ML3   ML2
## 9.seams       0.64
## 11.trnd       0.63
## 13.fit        0.81
## 14.conf       0.81
## 15.suits      0.84
## 16.dur        0.64
## 19.puck             0.68
## 20.shri             0.79
## 21.cf               0.84
## 22.m/c wash         0.51
## 1.col                     0.96
## 2.tex                     0.70
## 3.hf                      0.61
## 4.fbpat                   0.49
## 5.sty         0.33        0.37
## 6.vom
## 7.cc          0.47
## 17.season     0.49
##
##                  ML1  ML3  ML2
## SS loadings     4.01 2.36 2.35
## Proportion Var 0.22 0.13 0.13
## Cumulative Var 0.22 0.35 0.48
```

```r
fa.diagram(pa.out,col="blue")
```



Factor Mean score

To arrive single score for each factor, mean is used to get the item parceling score and factors are named based on the domain experts suggestions.

```r
d1 <- describe(csi1[,c(15,13,16,14,17,11,9,7)])
d2<- describe(csi1[,c(1:6)])
d3<- describe(csi1[,c(19:22)])

csi1$fac1 <- rowMeans(csi1[,c(15,13,16,14,17,11,9,7)])
csi1$fac2 <- rowMeans(csi1[,c(1:6)])
csi1$fac3 <- rowMeans(csi1[,c(19:22)])

kable(d1,digits = 2,caption = "First Hand Impressions")
```

Table 6: First Hand Impressions

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
15.suits	1	150	6.55	2.69	7	6.76	2.97	1	10	9	-0.52	-0.83	0.22
13.fit	2	150	6.37	2.57	7	6.51	2.97	1	10	9	-0.39	-0.87	0.21
16.dur	3	150	6.51	2.75	7	6.74	2.97	1	10	9	-0.53	-0.75	0.22
14.conf	4	150	6.72	2.76	7	7.00	2.97	1	10	9	-0.68	-0.60	0.23
17.season	5	150	6.27	2.48	7	6.42	2.97	1	10	9	-0.46	-0.58	0.20
11.trnd	6	150	6.57	2.63	7	6.77	2.97	1	10	9	-0.54	-0.80	0.21
9.seams	7	150	6.50	2.72	7	6.73	2.97	1	10	9	-0.61	-0.69	0.22
7.cc	8	150	6.58	2.61	7	6.78	2.97	1	10	9	-0.56	-0.73	0.21

```r
kable(d2,digits = 2,caption = "Quality confirmation characterstics")
```

Table 7: Quality confirmation characterstics

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
1.col	1	150	5.93	3.03	6.5	6.04	3.71	1	10	9	-0.36	-1.14	0.25
2.tex	2	150	5.97	2.54	6.0	6.01	2.97	1	10	9	-0.19	-0.97	0.21
3.hf	3	150	6.07	2.59	6.0	6.17	2.97	1	10	9	-0.31	-0.91	0.21
4.fbpat	4	150	6.25	2.65	6.5	6.40	2.22	1	10	9	-0.41	-0.79	0.22
5.sty	5	150	6.59	2.68	7.0	6.79	2.97	1	10	9	-0.54	-0.87	0.22
6.vom	6	150	6.23	2.52	6.0	6.29	2.97	1	10	9	-0.16	-0.97	0.21

```r
kable(d3,digits = 2,caption = "After Wash Effectives")
```

Table 8: After Wash Effectives

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
19.puck	1	150	5.80	2.67	6	5.84	2.97	1	10	9	-0.17	-0.98	0.22
20.shri	2	150	5.67	2.65	6	5.72	2.97	1	10	9	-0.09	-1.06	0.22
21.cf	3	150	5.82	2.72	6	5.94	2.97	1	10	9	-0.30	-1.07	0.22
22.m/c wash	4	150	6.04	3.02	7	6.17	2.97	1	10	9	-0.36	-1.22	0.25

```r
kable(describe(csi1[,23:25]),caption = "Item parcelled Mean score - Factor wise",digits = 2)
```

Table 9: Item parcelled Mean score - Factor wise

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
fac1	1	150	6.51	2.03	6.62	6.61	2.22	1.00	10	9.00	-0.46	-0.44	0.17
fac2	2	150	6.17	2.11	6.33	6.24	2.47	1.17	10	8.83	-0.30	-0.77	0.17
fac3	3	150	5.83	2.26	6.00	5.86	2.22	1.25	10	8.75	-0.15	-0.81	0.18

Reliability Analysis

Though factor is derived, it is better to assess the cronbach alpha dimension wise, to check the measurement error is in permissible level, Nunally suggested to keep alpha at least more than .7

```r
fac1 <-psych::alpha(csi1[,c(15,13,16,14,17,11,9,7)])
fac2 <- psych::alpha(csi1[,c(1:6)])
fac3 <-psych::alpha(csi1[,c(19:22)])




alphascore <- data.frame(rbind(fac1$total,fac2$total,fac3$total))
nvar <- data.frame(rbind(fac1$nvar,fac2$nvar,fac3$nvar))

var_name<- c("First Hand Impressions","Quality confirmation characterstics","After Wash Effectives")
finalalpha<-round(alphascore[,c(1,7,8)],3)
cronbach<-data.frame(cbind(var_name,finalalpha,nvar))
colnames(cronbach) <- c("Constructs Name","Cronbach Alpha","Mean","Std","No of items")
kable(cronbach,caption = "Table Reliability and Descripitve Analysis",digits = 2)
```

Table 10: Table Reliability and Descripitve Analysis

	Constructs Name	Cronbach Alpha	Mean	Std	No of items
	First Hand Impressions	0.90	6.51	2.03	8
1	Quality confirmation characterstics	0.88	6.17	2.11	6
2	After Wash Effectives	0.83	5.83	2.26	4