

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

```
csi <- read_xlsx("csi copy.xlsx",sheet = "Sheet1")
str(csi) # to display information about data set
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame': 546 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 ...
```

```
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	546	6.50	2.86	7	6.72	2.97	1	10	9	-0.40	-0.98	0.12
2.tex	2	546	6.26	2.38	7	6.39	2.97	1	10	9	-0.39	-0.71	0.10
3.hf	3	546	6.41	2.46	7	6.55	2.97	1	10	9	-0.45	-0.66	0.11
4.fbpat	4	546	6.48	2.44	7	6.61	2.97	1	10	9	-0.42	-0.66	0.10
5.sty	5	546	6.62	2.43	7	6.79	2.97	1	10	9	-0.55	-0.56	0.10
6.vom	6	546	6.60	2.36	7	6.74	2.97	1	10	9	-0.46	-0.61	0.10
7.cc	7	546	6.69	2.38	7	6.86	2.97	1	10	9	-0.52	-0.48	0.10
8.cct	8	546	6.35	2.37	7	6.47	2.97	1	10	9	-0.41	-0.66	0.10
9.seams	9	546	6.56	2.51	7	6.75	2.97	1	10	9	-0.55	-0.48	0.11
10.care	10	546	6.37	2.40	7	6.50	2.97	1	10	9	-0.42	-0.61	0.10
11.trnd	11	546	6.64	2.46	7	6.81	2.97	1	10	9	-0.50	-0.58	0.11
12.drapes	12	546	6.27	2.41	7	6.40	2.97	1	10	9	-0.41	-0.63	0.10
13.fit	13	546	6.53	2.53	7	6.69	2.97	1	10	9	-0.44	-0.69	0.11
14.conf	14	546	6.84	2.60	7	7.11	2.97	1	10	9	-0.69	-0.45	0.11
15.suits	15	546	6.87	2.57	7	7.12	2.97	1	10	9	-0.72	-0.43	0.11
16.dur	16	546	6.68	2.55	7	6.91	2.97	1	10	9	-0.58	-0.46	0.11
17.season	17	546	6.51	2.41	7	6.65	2.97	1	10	9	-0.44	-0.47	0.10
18.limp	18	546	6.11	2.30	6	6.20	2.97	1	10	9	-0.27	-0.54	0.10
19.puck	19	546	5.79	2.41	6	5.85	2.97	1	10	9	-0.21	-0.66	0.10
20.shri	20	546	5.72	2.42	6	5.78	2.97	1	10	9	-0.14	-0.82	0.10
21.cf	21	546	5.75	2.52	6	5.83	2.97	1	10	9	-0.23	-0.87	0.11
22.m/c wash	22	546	5.65	3.02	6	5.68	2.97	1	10	9	-0.22	-1.22	0.13

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	0.63	0.68	0.57	0.50	0.43	0.44	0.35	0.38	0.35	0.40	0.41	0.38
2.tex	0.63	1.00	0.59	0.58	0.49	0.42	0.42	0.40	0.36	0.38	0.42	0.46	0.38
3.hf	0.68	0.59	1.00	0.52	0.52	0.43	0.46	0.40	0.40	0.43	0.42	0.46	0.38
4.fbpat	0.57	0.58	0.52	1.00	0.51	0.50	0.46	0.43	0.40	0.40	0.46	0.44	0.38
5.sty	0.50	0.49	0.52	0.51	1.00	0.40	0.51	0.37	0.42	0.38	0.41	0.44	0.38
6.vom	0.43	0.42	0.43	0.50	0.40	1.00	0.44	0.39	0.38	0.37	0.41	0.40	0.38
7.cc	0.44	0.42	0.46	0.46	0.51	0.44	1.00	0.45	0.46	0.41	0.43	0.42	0.38
9.seams	0.35	0.40	0.40	0.43	0.37	0.39	0.45	1.00	0.46	0.46	0.44	0.46	0.40
11.trnd	0.38	0.36	0.40	0.40	0.42	0.38	0.46	0.46	1.00	0.50	0.50	0.53	0.40
13.fit	0.35	0.38	0.43	0.40	0.38	0.37	0.41	0.46	0.50	1.00	0.49	0.58	0.50
14.conf	0.40	0.42	0.42	0.46	0.41	0.41	0.43	0.44	0.50	0.49	1.00	0.60	0.48
15.suits	0.41	0.46	0.46	0.44	0.44	0.40	0.42	0.46	0.53	0.58	0.60	1.00	0.50
16.dur	0.35	0.31	0.37	0.38	0.31	0.35	0.36	0.44	0.41	0.51	0.48	0.52	1.00
17.season	0.36	0.35	0.40	0.40	0.34	0.37	0.35	0.47	0.41	0.43	0.42	0.55	0.50
19.puck	0.17	0.25	0.27	0.29	0.24	0.22	0.25	0.28	0.33	0.26	0.21	0.31	0.28
20.shri	0.13	0.26	0.15	0.25	0.22	0.21	0.19	0.17	0.22	0.19	0.18	0.24	0.28
21.cf	0.13	0.22	0.17	0.20	0.15	0.20	0.18	0.18	0.27	0.19	0.18	0.26	0.28
22.m/c wash	-0.02	0.04	0.00	0.09	0.10	0.15	0.11	0.11	0.08	0.05	0.08	0.11	0.10

```
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	7.17
2	1.90
3	1.24
4	0.85
5	0.75
6	0.66
7	0.62
8	0.56
9	0.54
10	0.53
11	0.50
12	0.46

SN	Eigen
13	0.45
14	0.42
15	0.41
16	0.34
17	0.32
18	0.29

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

## [1] "Overall kmo: 0.932"

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

Table 4: Measure of sample Adequacy

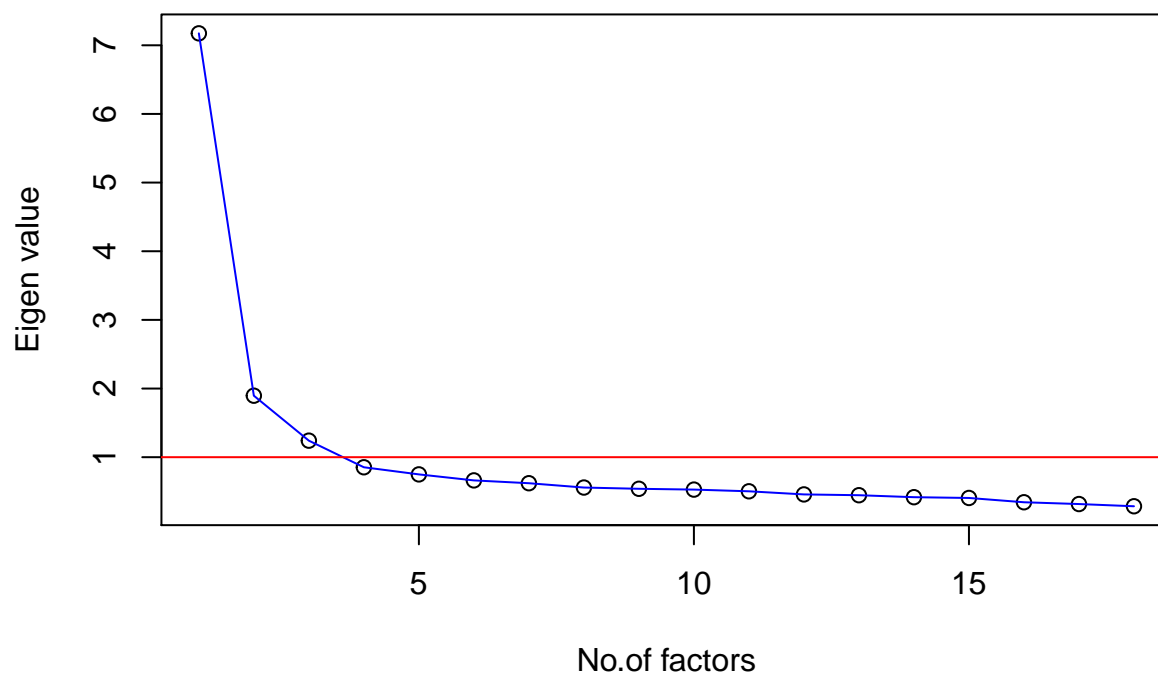
KMO item wise	
1.col	0.903
2.tex	0.933
3.hf	0.930
4.fbpat	0.956
5.sty	0.956
6.vom	0.966
7.cc	0.958
9.seams	0.959
11.trnd	0.954
13.fit	0.946
14.conf	0.953
15.suits	0.938
16.dur	0.926
17.season	0.933
19.puck	0.870
20.shri	0.854
21.cf	0.834
22.m/c wash	0.798

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

chisq	DF	P value
4233.485	153	0

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

Factors Extracted

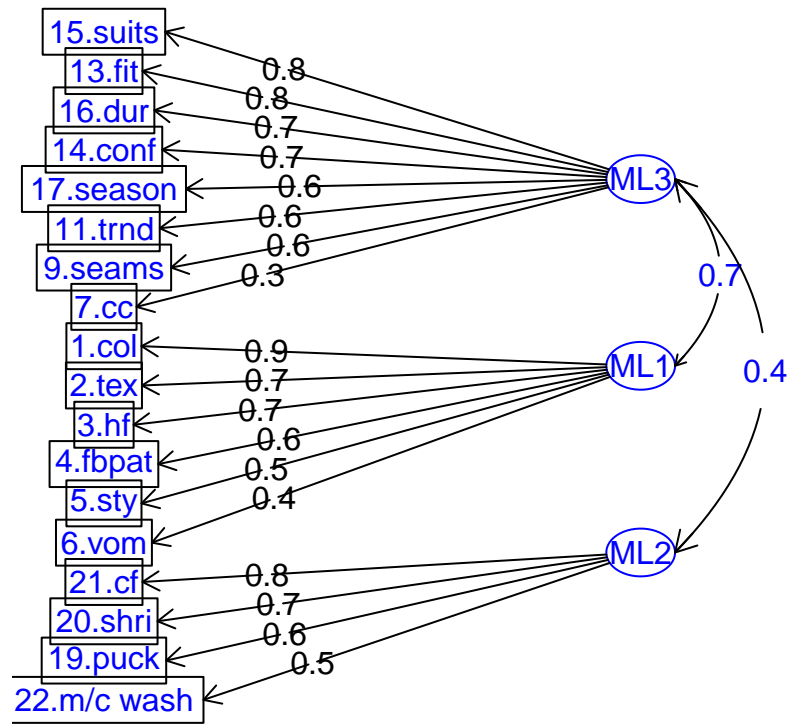


```
pa.out <- fa(satcor, nfactors= sum(corMat > 1.0), n.obs=nrow(csi), rotate="oblimin", scores="regression")
print(pa.out$loadings, sort=T, cutoff=0.3, digits=2)
```

```
##
## Loadings:
##          ML3   ML1   ML2
## 9.seams    0.58
## 11.trnd    0.60
## 13.fit     0.75
## 14.conf    0.68
## 15.suits   0.77
## 16.dur     0.73
## 17.season  0.60
## 1.col      0.90
## 2.tex      0.75
## 3.hf       0.73
## 4.fbpat    0.59
## 5.sty      0.51
## 19.puck    0.63
## 20.shri    0.71
## 21.cf      0.76
## 6.vom      0.35
## 7.cc       0.35 0.33
## 22.m/c wash 0.45
##
##          ML3   ML1   ML2
## SS loadings  3.49 2.79 1.75
## Proportion Var 0.19 0.16 0.10
## Cumulative Var 0.19 0.35 0.45
```

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

Factor Analysis



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.

```
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	546	6.87	2.57	7	7.12	2.97	1	10	9	-0.72	-0.43	0.11
13.fit	2	546	6.53	2.53	7	6.69	2.97	1	10	9	-0.44	-0.69	0.11
16.dur	3	546	6.68	2.55	7	6.91	2.97	1	10	9	-0.58	-0.46	0.11
14.conf	4	546	6.84	2.60	7	7.11	2.97	1	10	9	-0.69	-0.45	0.11
17.season	5	546	6.51	2.41	7	6.65	2.97	1	10	9	-0.44	-0.47	0.10
11.trnd	6	546	6.64	2.46	7	6.81	2.97	1	10	9	-0.50	-0.58	0.11
9.seams	7	546	6.56	2.51	7	6.75	2.97	1	10	9	-0.55	-0.48	0.11
7.cc	8	546	6.69	2.38	7	6.86	2.97	1	10	9	-0.52	-0.48	0.10

```

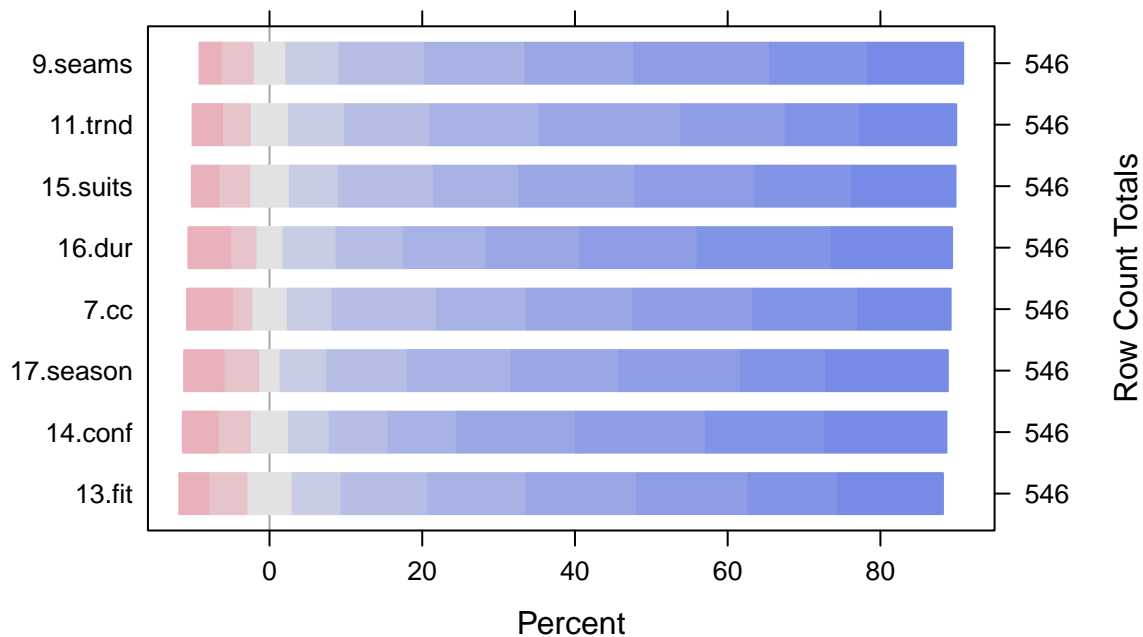
fac1 = csi1[,c(15,13,16,14,17,11,9,7)]
likt <- gather(fac1, measure, response)
likt$measure <- as.factor(likt$measure)
likt$response <- as.factor(likt$response)

df = table(likt$measure,likt$response) %>% as.data.frame.matrix()

colnames(df) = c("Least SAT 1","Rating 2","Rating 3","Rating 4","Rating 5",
                 "Rating 6","Rating 7","Rating 8", "Rating 9", "Highly SAT 10")
rownames(df) = names(fac1)
df <- tibble::rownames_to_column(df, var="Measure")
df1 = df
likert(Measure ~ ., data=df, ylab=NULL,
       ReferenceZero=3, as.percent=TRUE,
       positive.order=TRUE,
       main = "CSAT Attributes - Distribution",
       sub= list("First Hand Impressions"))

```

CSAT Attributes – Distribution



First Hand Impressions

```

kable(d2,digits = 2,caption = "Quality confirmation characteristics")

```

Table 7: Quality confirmation characteristics

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
1.col	1	546	6.50	2.86	7	6.72	2.97	1	10	9	-0.40	-0.98	0.12
2.tex	2	546	6.26	2.38	7	6.39	2.97	1	10	9	-0.39	-0.71	0.10

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
3.hf	3	546	6.41	2.46	7	6.55	2.97	1	10	9	-0.45	-0.66	0.11
4.fbpat	4	546	6.48	2.44	7	6.61	2.97	1	10	9	-0.42	-0.66	0.10
5.sty	5	546	6.62	2.43	7	6.79	2.97	1	10	9	-0.55	-0.56	0.10
6.vom	6	546	6.60	2.36	7	6.74	2.97	1	10	9	-0.46	-0.61	0.10

```

fac2 = csi1[,c(1:6)]

likt <- gather(fac2, measure, response)
likt$measure <- as.factor(likt$measure)
likt$response <- as.factor(likt$response)

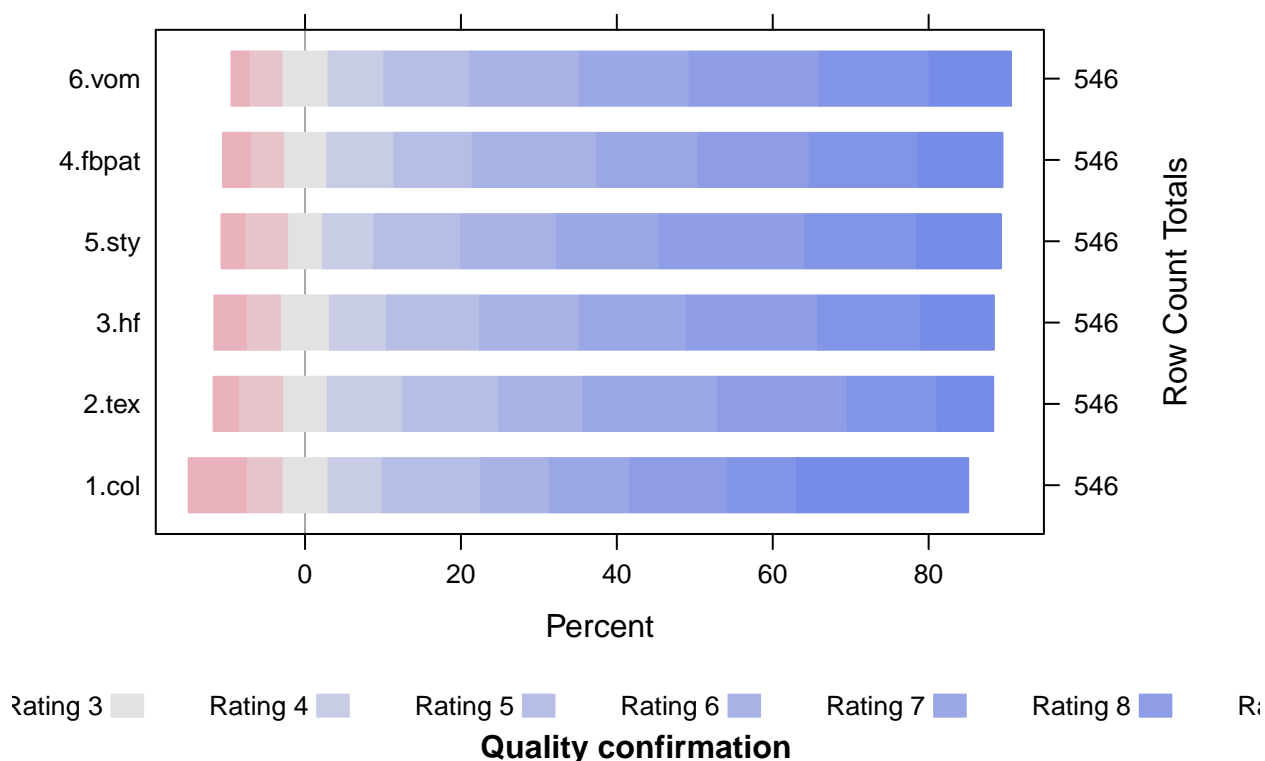
df = table(likt$measure,likt$response) %>% as.data.frame.matrix()

colnames(df) = c("Least SAT 1","Rating 2","Rating 3","Rating 4","Rating 5",
                 "Rating 6","Rating 7","Rating 8", "Rating 9", "Highly SAT 10")
rownames(df) = names(fac2)

df <- tibble::rownames_to_column(df, var="Measure")
df2 = df
likert(Measure ~ ., data=df, ylab=NULL,
       ReferenceZero=3, as.percent=TRUE,
       positive.order=TRUE,
       main = "CSAT Attributes - Distribution",
       sub= list("Quality confirmation"))

```


CSAT Attributes – Distribution



```
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	546	5.79	2.41	6	5.85	2.97	1	10	9	-0.21	-0.66	0.10
20.shri	2	546	5.72	2.42	6	5.78	2.97	1	10	9	-0.14	-0.82	0.10
21.cf	3	546	5.75	2.52	6	5.83	2.97	1	10	9	-0.23	-0.87	0.11
22.m/c wash	4	546	5.65	3.02	6	5.68	2.97	1	10	9	-0.22	-1.22	0.13

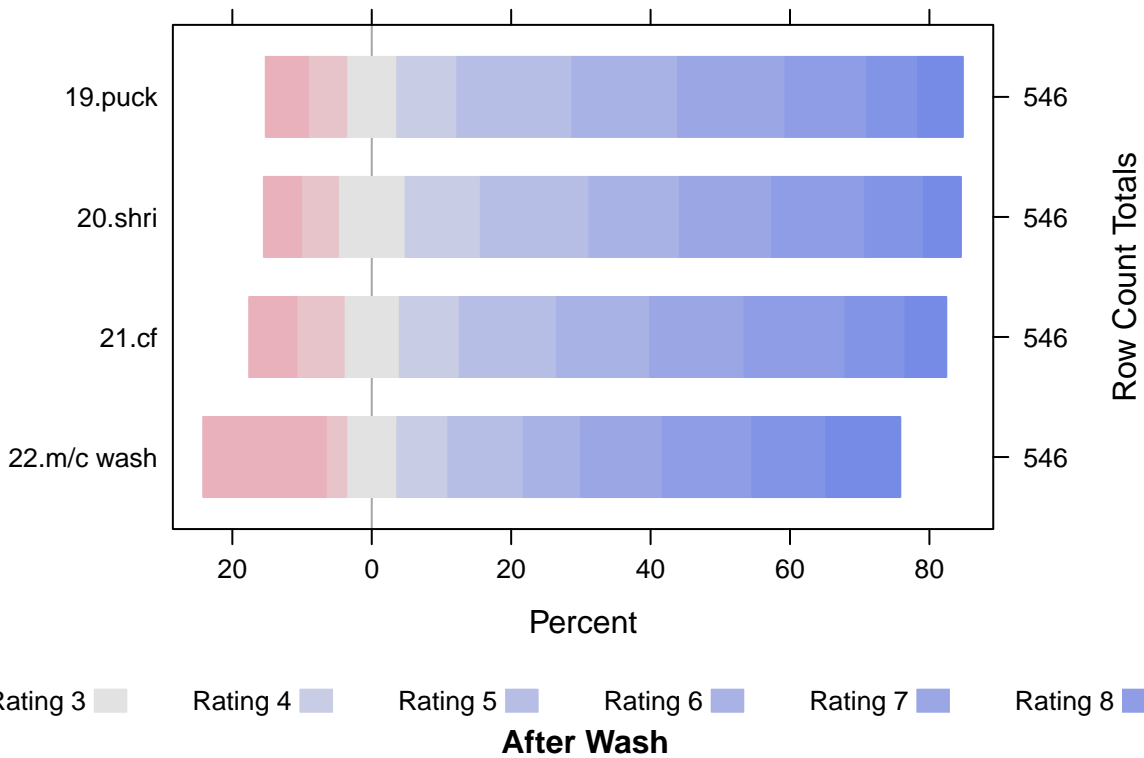
```
fac3 = csi1[,c(19:22)]
likt <- gather(fac3, measure, response)
likt$measure <- as.factor(likt$measure)
likt$response <- as.factor(likt$response)

df = table(likt$measure,likt$response) %>% as.data.frame.matrix()

colnames(df) = c("Least SAT 1","Rating 2","Rating 3","Rating 4","Rating 5",
                 "Rating 6","Rating 7","Rating 8", "Rating 9", "Highly SAT 10")
rownames(df) = names(fac3)
df <- tibble::rownames_to_column(df, var="Measure")
df3 = df
likert(Measure ~ ., data=df, ylab=NULL,
       ReferenceZero=3, as.percent=TRUE,
```

```
positive.order=TRUE,
main = "CSAT Attributes - Distribution",
sub= list("After Wash"))
```

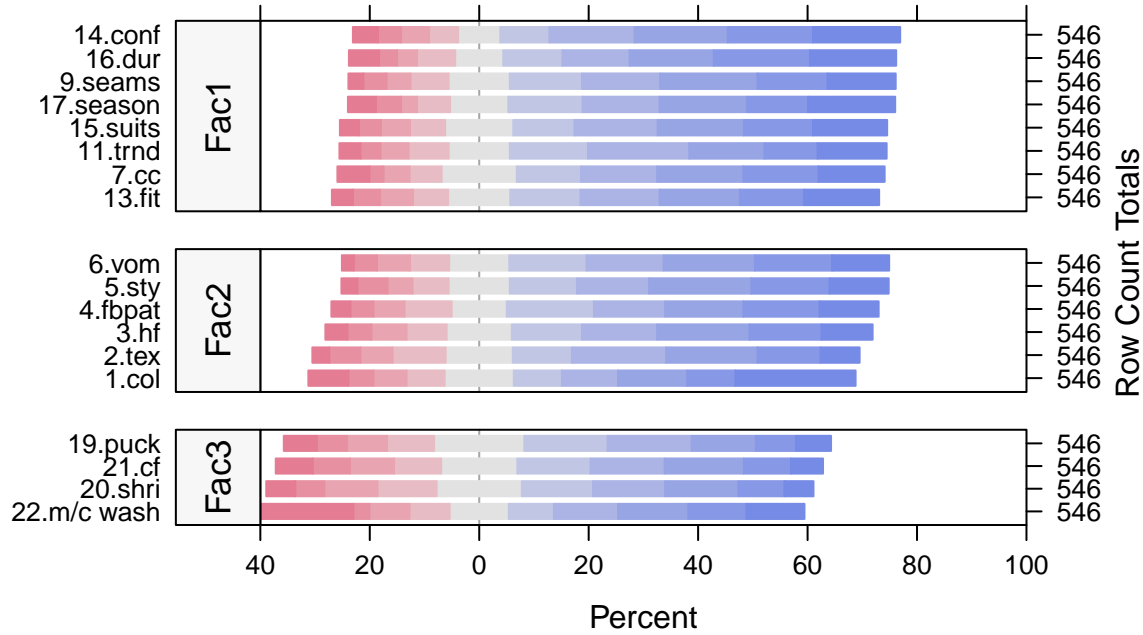
CSAT Attributes – Distribution



```
dfscore = data.frame(rbind(df1,df2,df3),Type = c(rep("Fac1",8),rep("Fac2",6),rep("Fac3",4)))

likert(Measure ~ . | Type, data=dfscore,
  layout=c(1,3),
  scales=list(y=list(relation="free")),
  between=list(y=1),
  strip.left=strip.custom(bg="gray97"),
  strip=FALSE,
  par.strip.text=list(cex=1.1, lines=2),
  ylab=NULL,
  cex=1.2,
  ReferenceZero=5,
  as.percent=TRUE,
  positive.order=TRUE,
  main = list("Cust Satisfaction Survey Report"),
  sub= list("Satisfaction Rating"),
  xlim=c(-40,-20,0,20,40,60,80,100),resize.height.tuning=2)
```

Cust Satisfaction Survey Report



```
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	546	6.66	1.83	7.00	6.81	1.85	1.00	10	9.00	-0.71	0.11	0.08
fac2	2	546	6.48	1.93	6.67	6.61	2.10	1.17	10	8.83	-0.51	-0.42	0.08
fac3	3	546	5.73	1.91	5.75	5.75	1.85	1.00	10	9.00	-0.12	-0.29	0.08

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

```
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 Descriptiv Analysis",digits = 2)
```

Table 10: Table Reliability and Descriptiv Analysis

	Constructs Name	Cronbach Alpha	Mean	Std	No of items
	First Hand Impressions	0.88	6.66	1.83	8
1	Quality confirmation characterstics	0.87	6.48	1.93	6
2	After Wash Effectives	0.72	5.73	1.91	4