Source Visual Outline

```
let's import necessary packages
```{r}
library(psych)
library(dplyr)
```

let's import our dataset
```{r}
data=read.csv(file.choose(),header=TRUE)
data
```
```

Description: df [56 × 14]

Attractive <int>	Classy <int>	Beautiful <int>	Elegant <int>	Sexy <int>	Dependable <int>	Honest <int>	Reliable <int>	Trustworthy <int>	Sincere <int>
4	4	5	5	5	5	4	4	3	4
5	5	5	5	5	3	4	3	3	3
5	3	5	4	5	3	3	3	3	3
4	5	4	4	4	4	4	4	4	4
5	5	5	5	4	3	2	2	2	2
5	5	5	5	5	2	2	2	2	2
5	2	4	4	5	1	2	1	1	1
5	5	5	5	5	4	4	4	4	4
5	4	5	4	5	3	3	3	3	3
5	5	5	5	5	5	5	4	2	2

1-10 of 56 rows | 1-10 of 14 columns

we will check for missing values in our data

```{r}
is.null(data)
```

```
[1] FALSE
```

this will provide us the name of all the columns in the data

```{r}
names(data)
```

```
 [1] "Attractive"    "Classy"        "Beautiful"     "Elegant"       "Sexy"          "Dependable"
 [7] "Honest"        "Reliable"      "Trustworthy"   "Sincere"       "Expect"        "Experienced"
[13] "Knowledgeable" "Qualified"
```

The KMO function in the psych package produces an overall measure of sampling adequacy and MSA for each item.

```{r}
KMO(data)
```

```
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = data)
Overall MSA =  0.78
MSA for each item =
   Attractive        Classy     Beautiful       Elegant          Sexy    Dependable        Honest      Reliable
         0.58          0.83          0.56          0.71          0.49          0.85          0.84          0.88
  Trustworthy       Sincere        Expect   Experienced Knowledgeable     Qualified
         0.77          0.82          0.89          0.76          0.72          0.75
```

The overall KMO for data is 0.78, which is acceptable and this suggest that data is appropriate for factor analysis.

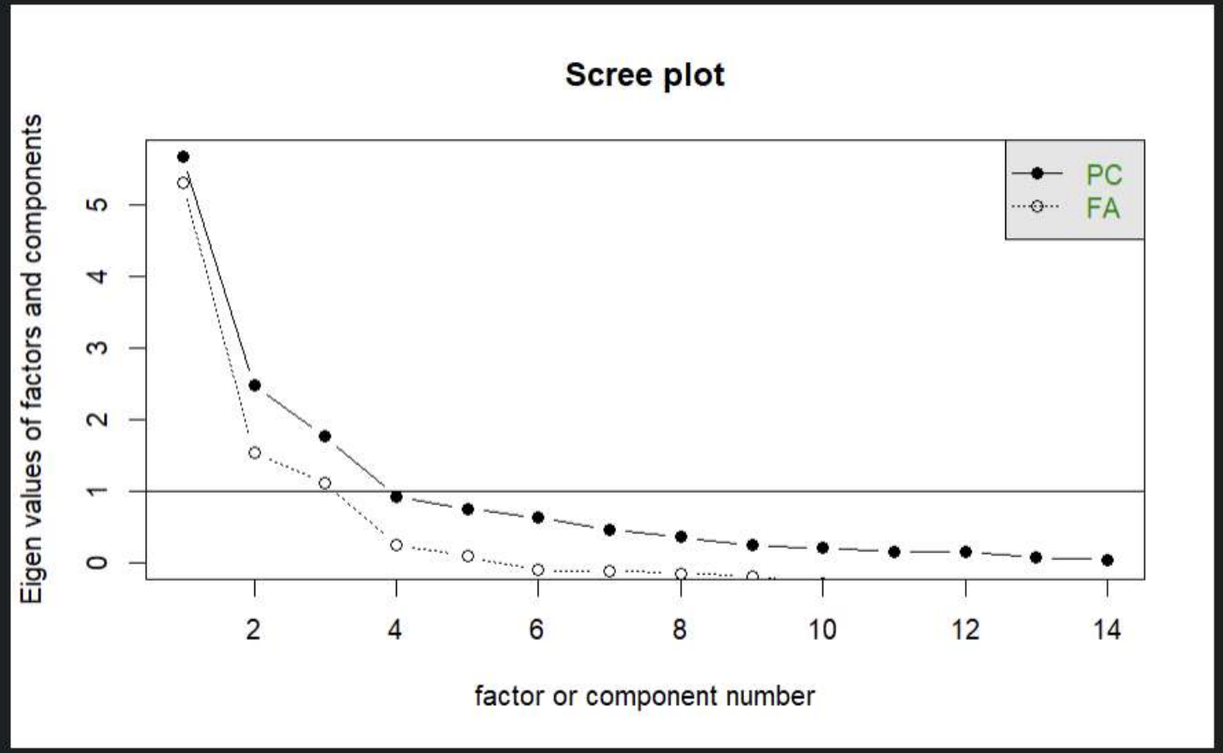
```{r}
cortest.bartlett(data)
```

```
R was not square, finding R from data
$chisq
[1] 573.3409

$p.value
[1] 4.754285e-71

$df
[1] 91
```

Here the p value we get is very close to 0
In Bartlett's Test the value obtained should be less than 0.05
So we can say that the Bartlett's Test is statistically significant and we can proceed for Factor Analysis.

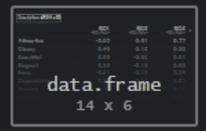
We plot a Scree plot to check how many factors to retain

```{r}
scree(data)
```



Here 3 factors have eigen value greater than 1
so based on eigen value criterion we decide to take 3 factors

```r
pca=pca(data,nfactors=3,rotate="varimax")
pca
```

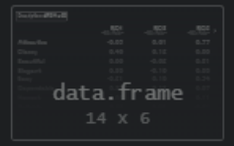


Description: df [14 x 6]

	RC1 <S3: AsIs>	RC3 <S3: AsIs>	RC2 <S3: AsIs>	h2 <dbl>	u2 <dbl>	com <dbl>
Attractive	-0.05	0.01	0.77	0.6034076	0.3965924	1.010371
Classy	0.40	0.12	0.60	0.5248341	0.4751659	1.830192
Beautiful	0.00	-0.02	0.91	0.8324223	0.1675777	1.001258
Elegant	0.30	-0.10	0.66	0.5388755	0.4611245	1.443065
Sexy	-0.21	0.10	0.54	0.3514287	0.6485713	1.378204
Dependable	0.89	0.18	0.07	0.8265412	0.1734588	1.089382
Honest	0.92	0.17	0.11	0.8898873	0.1101127	1.096023
Reliable	0.89	0.22	-0.01	0.8438009	0.1561991	1.127418
Trustworthy	0.90	0.18	0.00	0.8503598	0.1496402	1.077825
Sincere	0.90	0.14	-0.01	0.8256820	0.1743180	1.050900

1-10 of 14 rows

Previous 1 2 Next




```
Principal Components Analysis
Call: principal(r = r, nfactors = nfactors, residuals = residuals,
    rotate = rotate, n.obs = n.obs, covar = covar, scores = scores,
    missing = missing, impute = impute, oblique.scores = oblique.scores,
    method = method, use = use, cor = cor, correct = 0.5, weight = NULL)
Standardized loadings (pattern matrix) based upon correlation matrix

                        RC1  RC3  RC2
SS loadings            4.80 2.60 2.55
Proportion Var         0.34 0.19 0.18
Cumulative Var         0.34 0.53 0.71
Proportion Explained   0.48 0.26 0.26
Cumulative Proportion  0.48 0.74 1.00

Mean item complexity =  1.3
Test of the hypothesis that 3 components are sufficient.

The root mean square of the residuals (RMSR) is  0.07
 with the empirical chi square  45.8  with prob <  0.72

Fit based upon off diagonal values = 0.97
```

```r
print(pca$loadings,cutoff=0.5,sort=TRUE)
```

```
Loadings:
             RC1    RC3    RC2
Dependable    0.890
Honest        0.922
Reliable      0.891
Trustworthy   0.905
Sincere       0.897
Expect        0.550
Experienced          0.803
Knowledgeable        0.880
Qualified            0.869
Attractive                  0.775
Classy                      0.596
Beautiful                   0.912
Elegant                     0.663
Sexy                        0.544

                 RC1   RC3   RC2
SS loadings    4.797 2.601 2.548
Proportion Var 0.343 0.186 0.182
Cumulative Var 0.343 0.528 0.710
```

We retained 3 factors

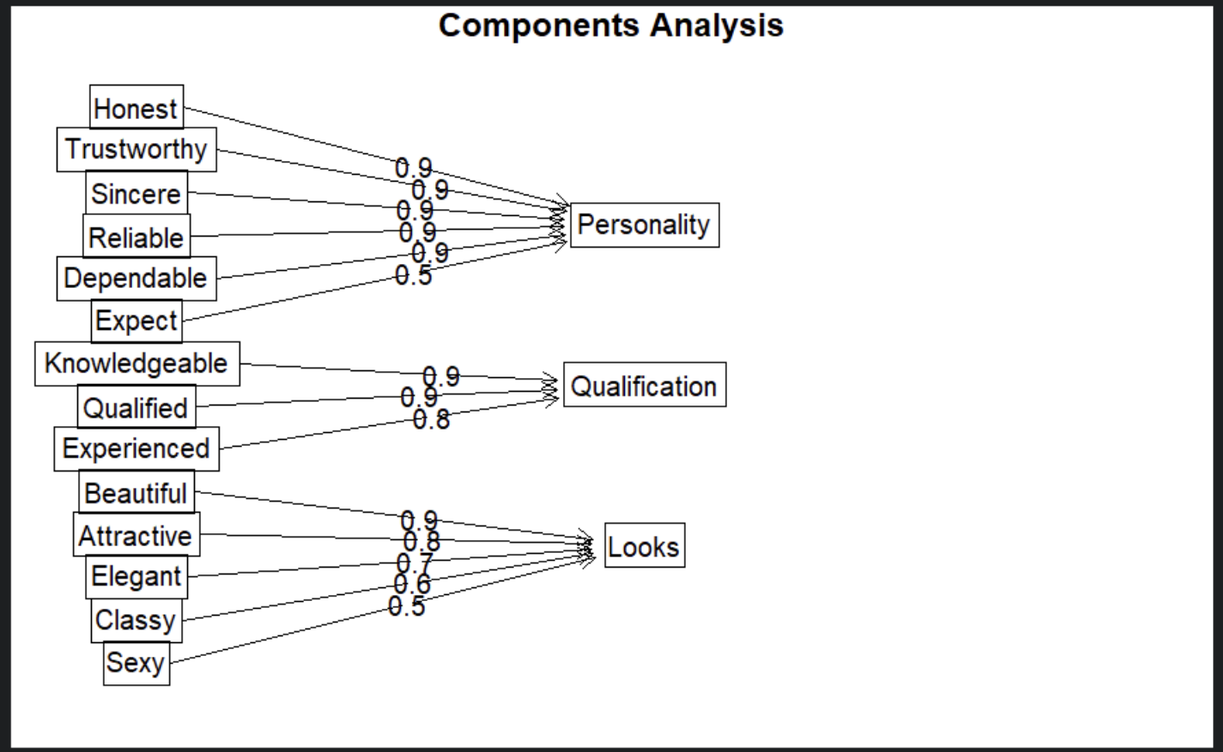
The first factor is something that is common to Dependable, Honest, Reliable, Trust-worthy, sincere & expert. It seems like a good name for this factor is “Personality”

The other two factors can be named as “Qualification” & “Looks”.

```r
colnames(pca$loadings)=c('Personality','Qualification','Looks')
print(pca$loadings,cutoff=0.5,sort=TRUE)
```

```
Loadings:
             Personality Qualification Looks
Dependable    0.890
Honest        0.922
Reliable      0.891
Trustworthy   0.905
Sincere       0.897
Expect        0.550
Experienced                0.803
Knowledgeable              0.880
Qualified                  0.869
Attractive                         0.775
Classy                             0.596
Beautiful                          0.912
Elegant                            0.663
Sexy                               0.544

               Personality Qualification Looks
SS loadings          4.797         2.601 2.548
Proportion Var       0.343         0.186 0.182
Cumulative Var       0.343         0.528 0.710
```

```{r}
fa.diagram(pca)
```



Conclusion:

Celebrity personality, qualification and looks are the three major qualities that have the greatest impact on advertising firms. When hiring celebrities to endorse items, the advertising industry should take into consideration these three elements.