PCA Motorcycle

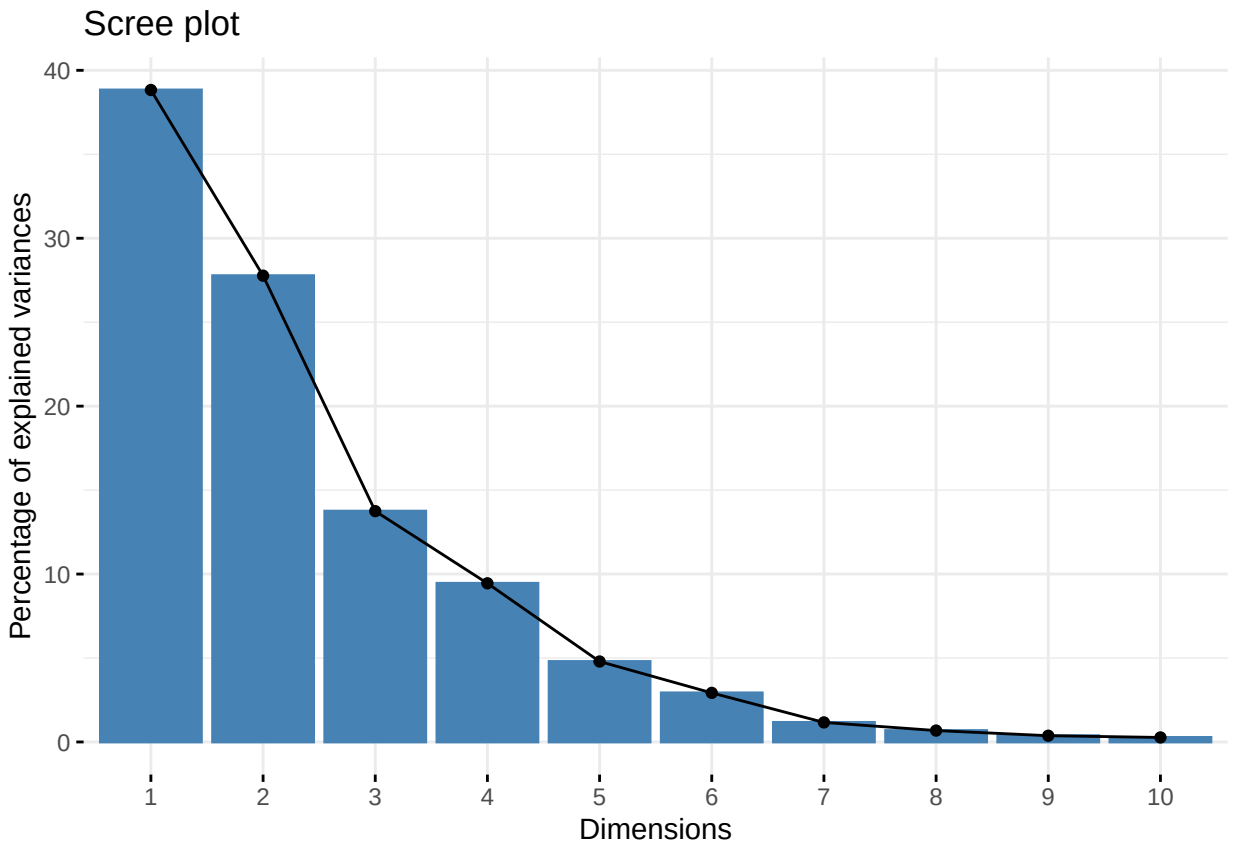
2023-04-02

```
##Analysis for a 2 wheeler producer to understand the major Factors influencing Consumer preferences fo
##The data used for this analysis is collected by a two wheeler producer ##specifically for this manage

##Following are the statements used for this purpose:
## 1. I use a 2W because it is affordable
## 2. It gives me a sense of freedom to own a 2W
## 3. Low maintenance cost makes it economical in the long run
## 4. 2W is man's vehicle essentially
## 5. I feel powerful on my 2W
## 6. Some of my friends who don't have a 2W are jealous of me
## 7. I feel good when I see ads for my 2W
## 8. My 2W gives me comfortable ride
## 9. I think 2W are a safe way to travel
##10. 3 people should be legally allowed to travel on a 2W

library(readxl)
library(factoextra)
```

```
## Loading required package: ggplot2
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
##Deciding the number of factors to be extracted using Principal Component Analysis (PCA):
motorcycle = read_excel("Motorcycle Preferences Data.xls")
pca_model = prcomp(motorcycle, scale = TRUE)
fviz_eig(pca_model)
```




```
##After looking at this plot, we can see that either 3 or 4 components should be ##enough to explain mo

##Extracting the factors:

fact_model_1 = factanal(motorcycle, factors = 3)
fact_model_1
```

```
##
## Call:
## factanal(x = motorcycle, factors = 3)
##
## Uniquenesses:
##      Affordable Sens_of_ Freedm      Maint_Cost        Mans_Veh        Powerful
##           0.814           0.417           0.722           0.073           0.014
##       Frnds_Jls   Ads_Feel_Good     Comfortable          Safety         Legal_3
##           0.005           0.060           0.069           0.216           0.977
##
## Loadings:
##                 Factor1 Factor2 Factor3
## Affordable       0.178   0.393
## Sens_of_ Freedm         -0.193  -0.734
## Maint_Cost      -0.118   0.397   0.326
## Mans_Veh         0.962
## Powerful         0.933           0.340
## Frnds_Jls        0.980          -0.181
## Ads_Feel_Good    0.944           0.219
## Comfortable     -0.262   0.898   0.234
```

```
## Safety                   0.870   0.166
## Legal_3                          0.112
##
##                Factor1 Factor2 Factor3
## SS loadings      3.779   1.918   0.937
## Proportion Var   0.378   0.192   0.094
## Cumulative Var   0.378   0.570   0.663
##
## Test of the hypothesis that 3 factors are sufficient.
## The chi square statistic is 21.94 on 18 degrees of freedom.
## The p-value is 0.235
```

```
##After looking at the results, specifically, at the proportion of variance ##explained by the each fac

##Let's try 4 factors and see if this can indeed be improved:

fact_model_2 = factanal(motorcycle, factors = 4)
fact_model_2
```

```
##
## Call:
## factanal(x = motorcycle, factors = 4)
##
## Uniquenesses:
##      Affordable Sens_of_ Freedm      Maint_Cost        Mans_Veh        Powerful
##           0.005           0.402           0.500           0.063           0.017
##       Frnds_Jls   Ads_Feel_Good     Comfortable          Safety         Legal_3
##           0.005           0.059           0.216           0.005           0.687
##
## Loadings:
##                 Factor1 Factor2 Factor3 Factor4
## Affordable       0.125   0.274   0.950
## Sens_of_ Freedm -0.113  -0.199  -0.153  -0.723
## Maint_Cost      -0.118   0.390   0.494   0.300
## Mans_Veh         0.963
## Powerful         0.946                   0.290
## Frnds_Jls        0.965                  -0.224
## Ads_Feel_Good    0.947           0.103   0.174
## Comfortable     -0.223   0.826   0.125   0.191
## Safety                   0.992
## Legal_3                 -0.115   0.531   0.121
##
##                Factor1 Factor2 Factor3 Factor4
## SS loadings      3.752   1.966   1.488   0.836
## Proportion Var   0.375   0.197   0.149   0.084
## Cumulative Var   0.375   0.572   0.721   0.804
##
## Test of the hypothesis that 4 factors are sufficient.
## The chi square statistic is 9.87 on 11 degrees of freedom.
## The p-value is 0.542
```

```
##Here, similar to the earlier result, even if there are 4 factors Factor 4 ##accounts only for 8.4 % v

##Understanding what each factor represents:
##Now, let us try to interpret these 3 factors based on their constituent loadings.
```

```
fact_model_1$loadings
```

```
##
## Loadings:
##                 Factor1 Factor2 Factor3
## Affordable       0.178   0.393
## Sens_of_ Freedm          -0.193  -0.734
## Maint_Cost      -0.118   0.397   0.326
## Mans_Veh         0.962
## Powerful         0.933           0.340
## Frnds_Jls        0.980          -0.181
## Ads_Feel_Good    0.944           0.219
## Comfortable     -0.262   0.898   0.234
## Safety                   0.870   0.166
## Legal_3                          0.112
##
##                Factor1 Factor2 Factor3
## SS loadings      3.779   1.918   0.937
## Proportion Var   0.378   0.192   0.094
## Cumulative Var   0.378   0.570   0.663
```

```
##As can be seen from the output, Factor 1 has high loading values for the ##variables "Mans_Veh", "Pow

##For the Factor 2, the loading values for the variables "Comfortable" and ##"Safety" are high. This Fa

##For the Factor 3, the loading values for the variables "Affordable", ##"Maint_Cost", and "Legal_3" ar

knitr::opts_chunk$set(echo = TRUE)
```