# Principle Component Analysis on the "Various Factors of Car"

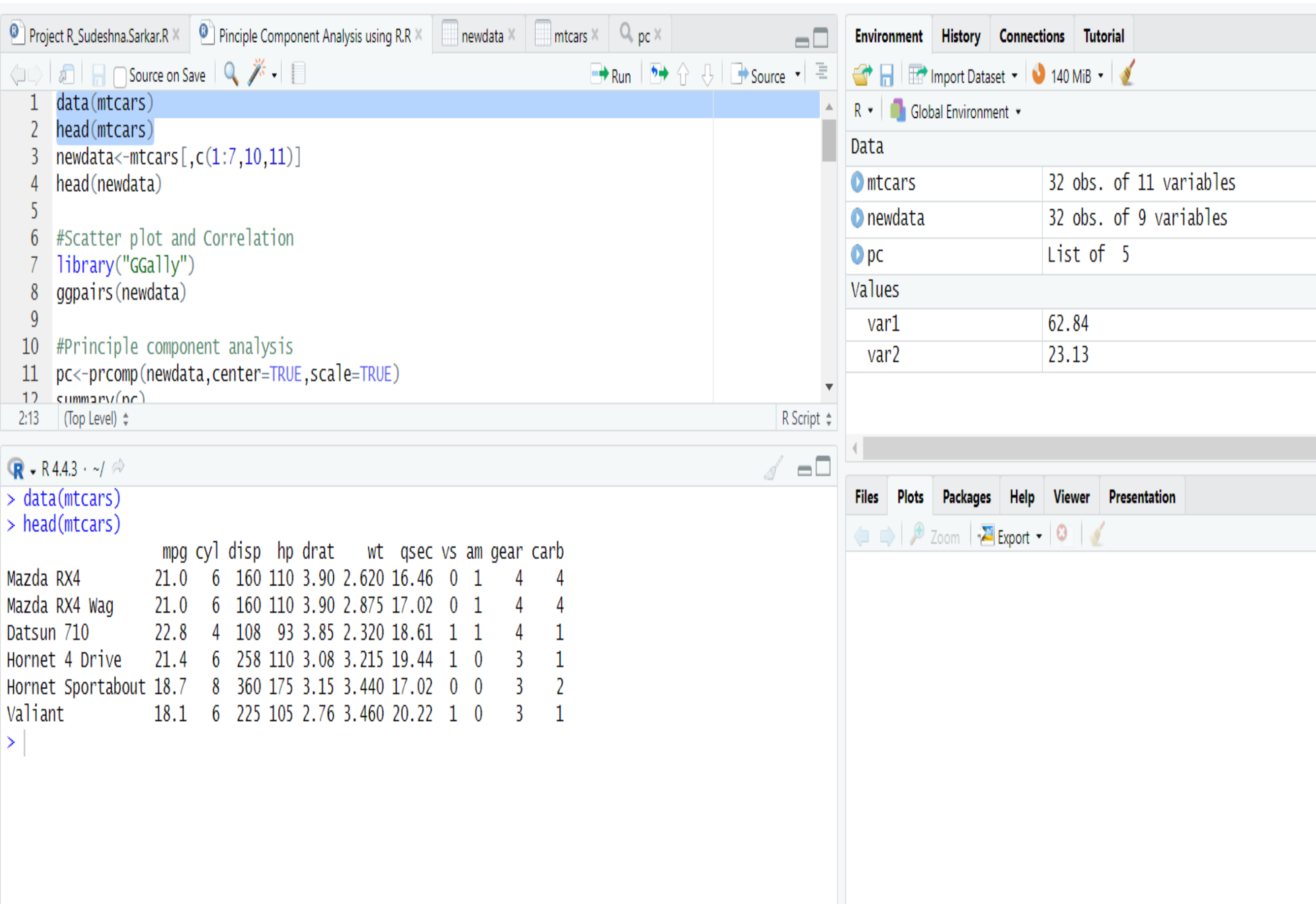
### What is PCA?

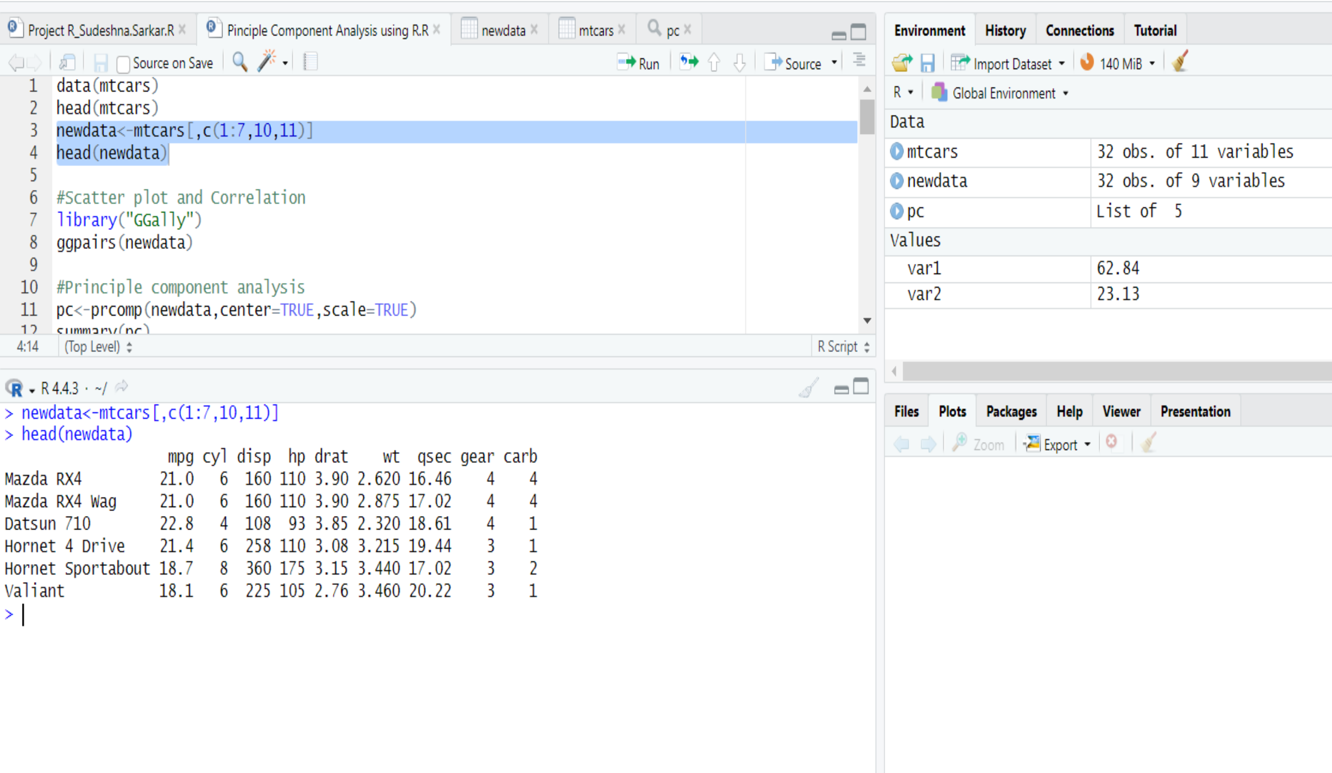
⇨ **Principal Component Analysis (PCA)** is a statistical technique used to **reduce the number of variables** in a dataset while keeping the most important information.

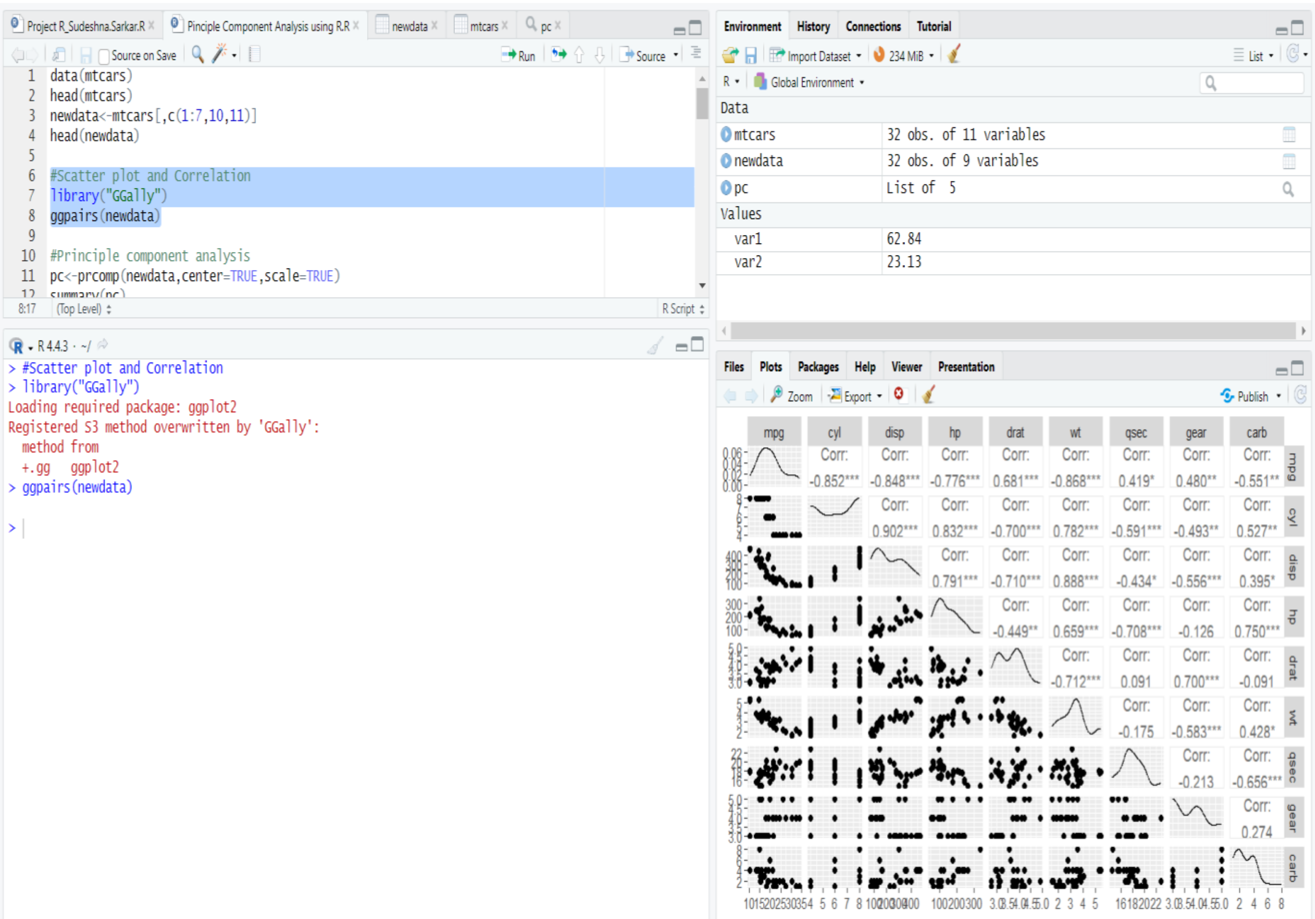
## 📌 Key Points:

- PCA is a **dimensionality reduction** method.
- It helps in **data visualization**, **compression** and **improving machine learning models**.
- It creates new features that are **uncorrelated** and **ordered by importance**.
- It is widely used in **data science**, **image processing** and **pattern recognition**.

### 🧠 Example:




Project R_Sudeshna.Sarkar.R
Principle Component Analysis using R.R
newdata
mtcars
pc
Source on Save
Run
Source
data(mtcars)
head(mtcars)
newdata<-mtcars[,c(1:7,10,11)]
head(newdata)
#Scatter plot and Correlation
library("GGally")
ggpairs(newdata)
#Principle component analysis
pc<-prcomp(newdata,center=TRUE,scale=TRUE)
R Script
R 4.4.3 · ~/
> newdata<-mtcars[,c(1:7,10,11)]
> head(newdata)
mpg cyl disp hp drat wt qsec gear carb
Mazda RX4 21.0 6 160 110 3.90 2.620 16.46 4 4
Mazda RX4 Wag 21.0 6 160 110 3.90 2.875 17.02 4 4
Datsun 710 22.8 4 108 93 3.85 2.320 18.61 4 1
Hornet 4 Drive 21.4 6 258 110 3.08 3.215 19.44 3 1
Hornet Sportabout 18.7 8 360 175 3.15 3.440 17.02 3 2
Valiant 18.1 6 225 105 2.76 3.460 20.22 3 1
Environment
History
Connections
Tutorial
Import Dataset
140 MiB
Global Environment
Data
mtcars 32 obs. of 11 variables
newdata 32 obs. of 9 variables
pc List of 5
Values
var1 62.84
var2 23.13
Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export

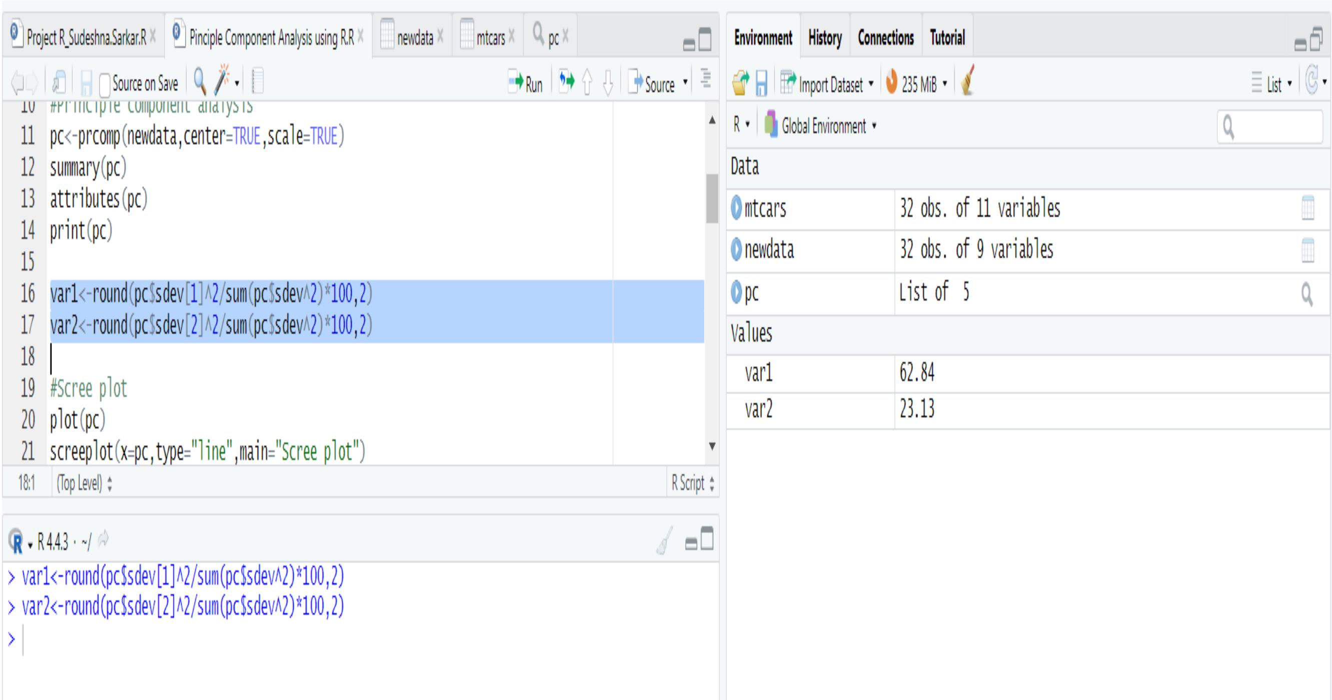

data(mtcars)
head(mtcars)
newdata<-mtcars[,c(1:7,10,11)]
head(newdata)
#Scatter plot and Correlation
library("GGally")
ggpairs(newdata)
#Principle component analysis
pc<-prcomp(newdata,center=TRUE,scale=TRUE)
R 4.4.3 · ~/
> #Scatter plot and Correlation
> library("GGally")
Loading required package: ggplot2
Registered S3 method overwritten by 'GGally':
method from
+.gg ggplot2
> ggpairs(newdata)
Environment
History
Connections
Tutorial
Import Dataset
234 MiB
Global Environment
Data
mtcars 32 obs. of 11 variables
newdata 32 obs. of 9 variables
pc List of 5
Values
var1 62.84
var2 23.13
Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export
Publish
mpg
cyl
disp
hp
drat
wt
qsec
gear
carb

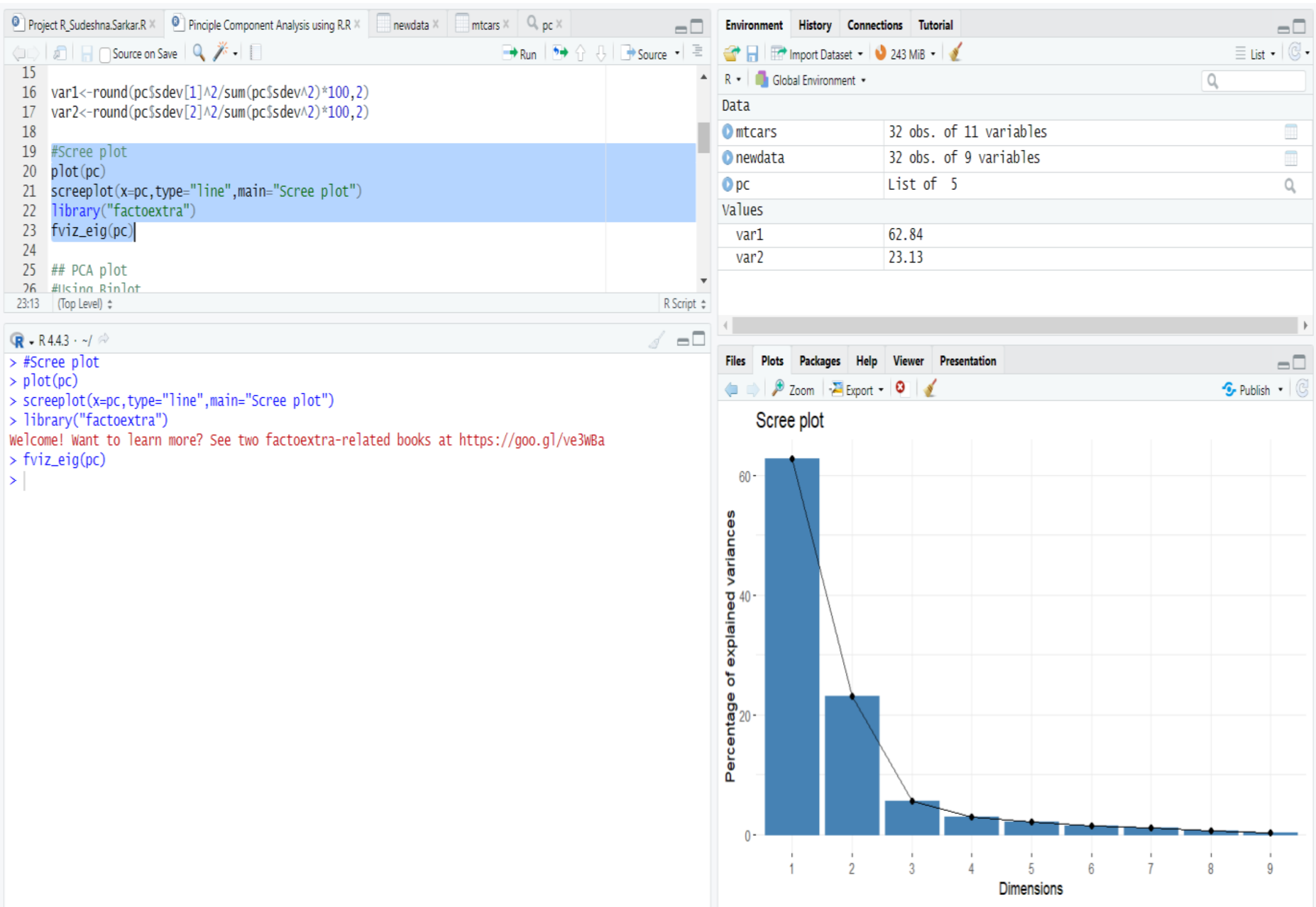
```r
#Scatter plot and Correlation
library("GGally")
ggpairs(newdata)

#Principle component analysis
pc<-prcomp(newdata,center=TRUE,scale=TRUE)
summary(pc)
attributes(pc)
print(pc)

var1<-round(pc$sdev[1]^2/sum(pc$sdev^2)*100,2)
```

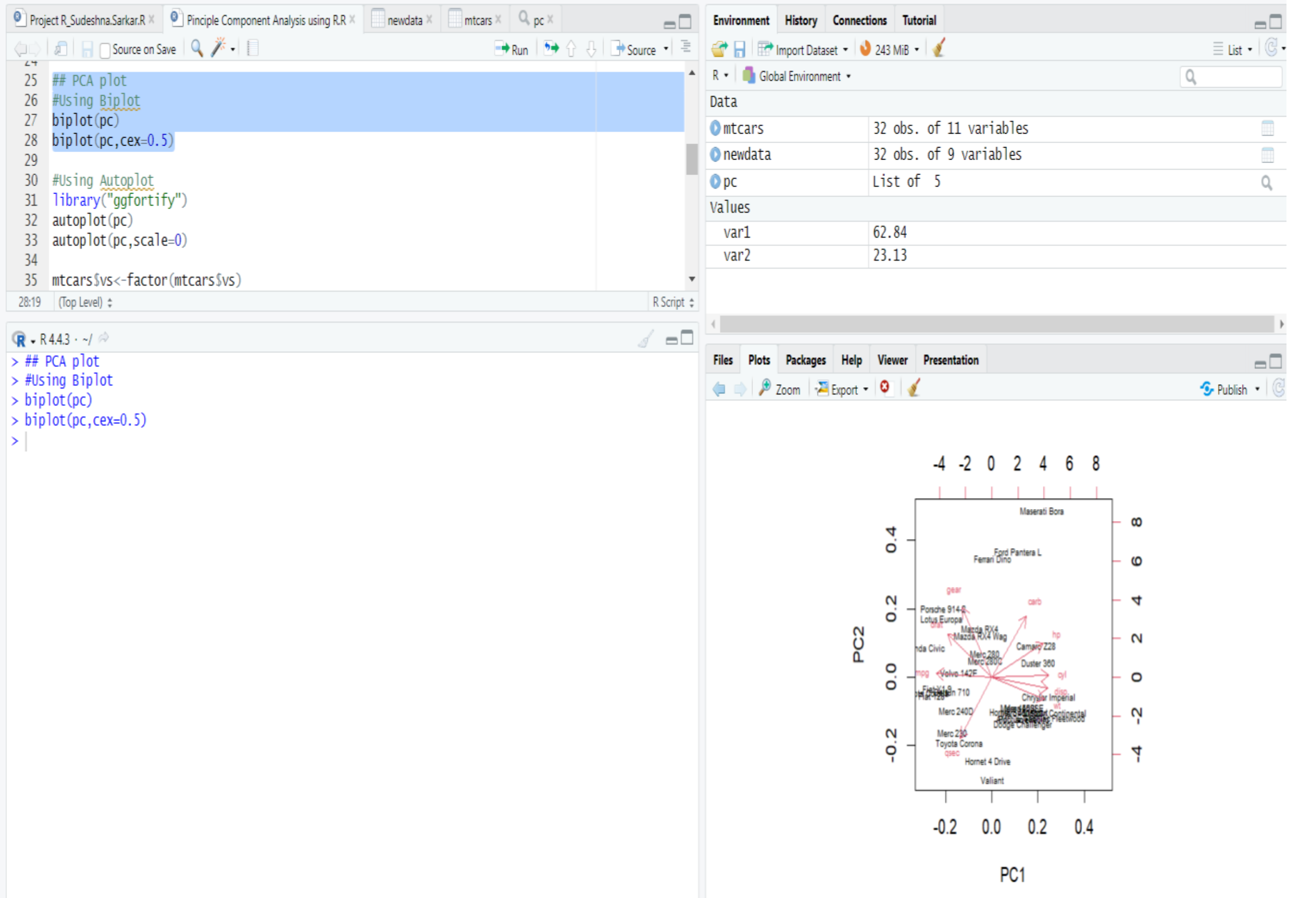
```
> #Principle component analysis
> pc<-prcomp(newdata,center=TRUE,scale=TRUE)
> summary(pc)
Importance of components:
                          PC1    PC2     PC3     PC4     PC5     PC6     PC7    PC8     PC9
Standard deviation     2.3782 1.4429 0.71008 0.51481 0.42797 0.35184 0.32413 0.2419 0.14896
Proportion of Variance 0.6284 0.2313 0.05602 0.02945 0.02035 0.01375 0.01167 0.0065 0.00247
Cumulative Proportion  0.6284 0.8598 0.91581 0.94525 0.96560 0.97936 0.99103 0.9975 1.00000
>
```

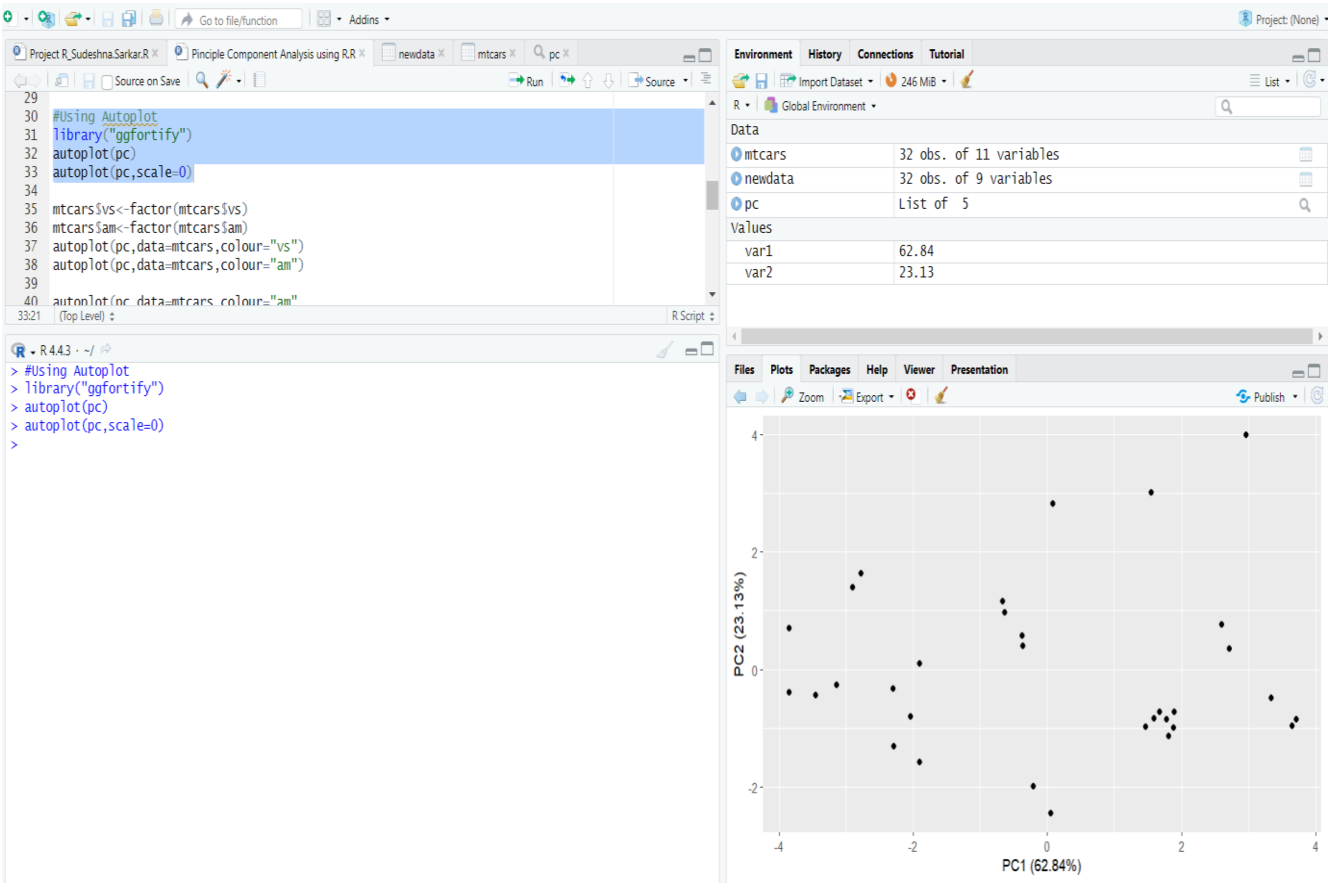
```r
pc<-prcomp(newdata,center=TRUE,scale=TRUE)
summary(pc)
attributes(pc)
print(pc)

var1<-round(pc$sdev[1]^2/sum(pc$sdev^2)*100,2)
var2<-round(pc$sdev[2]^2/sum(pc$sdev^2)*100,2)

#Scree plot
plot(pc)
screeplot(x=pc,type="line",main="Scree plot")
```

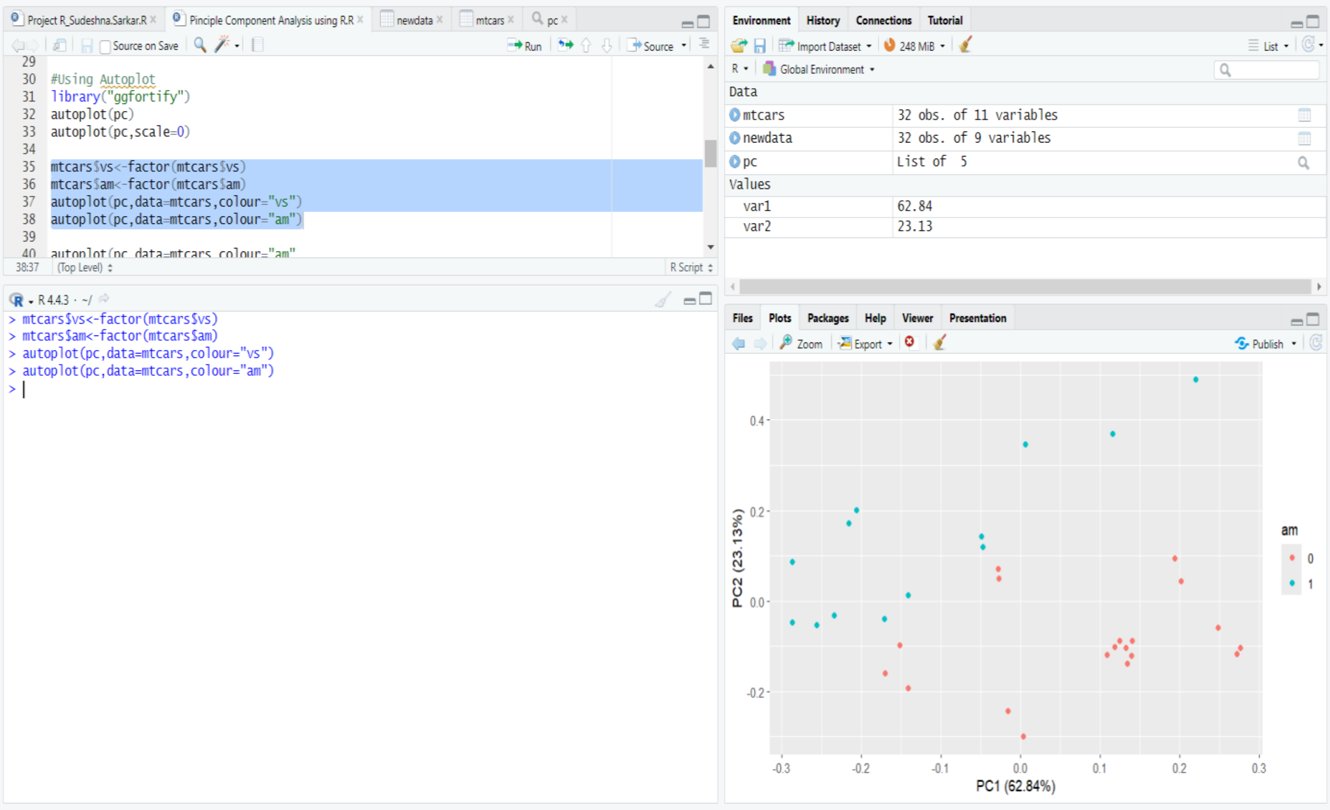
```
> print(pc)
Standard deviations (1, .., p=9):
[1] 2.3782219 1.4429485 0.7100809 0.5148082 0.4279704 0.3518426 0.3241326 0.2418962 0.1489644

Rotation (n x k) = (9 x 9):
            PC1         PC2         PC3          PC4        PC5         PC6         PC7
mpg  -0.3931477  0.02753861 -0.22119309 -0.006126378 -0.3207620  0.72015586  0.38138068
cyl   0.4025537  0.01570975 -0.25231615  0.040700251  0.1171397  0.22432550  0.15893251
disp  0.3973528 -0.08888469 -0.07825139  0.339493732 -0.4867849 -0.01967516  0.18233095
hp    0.3670814  0.26941371 -0.01721159  0.068300993 -0.2947317  0.35394225 -0.69620751
drat -0.3118165  0.34165268  0.14995507  0.845658485  0.1619259 -0.01536794 -0.04767957
wt    0.3734771 -0.17194306  0.45373418  0.191260029 -0.1874822 -0.08377237  0.42777608
qsec -0.2243508 -0.48404435  0.62812782 -0.030329127 -0.1482495  0.25752940 -0.27622581
gear -0.2094749  0.55078264  0.20658376 -0.282381831 -0.5624860 -0.32298239  0.08555707
carb  0.2445807  0.48431310  0.46412069 -0.214492216  0.3997820  0.35706914  0.20604210
             PC8         PC9
mpg   0.12465987  0.11492862
cyl  -0.81032177  0.16266295
disp  0.06416707 -0.66190812
hp    0.16573993  0.25177306
drat -0.13505066  0.03809096
wt    0.19839375  0.56918844
qsec -0.35613350 -0.16873731
gear -0.31636479  0.04719694
carb  0.10832772 -0.32045892
```

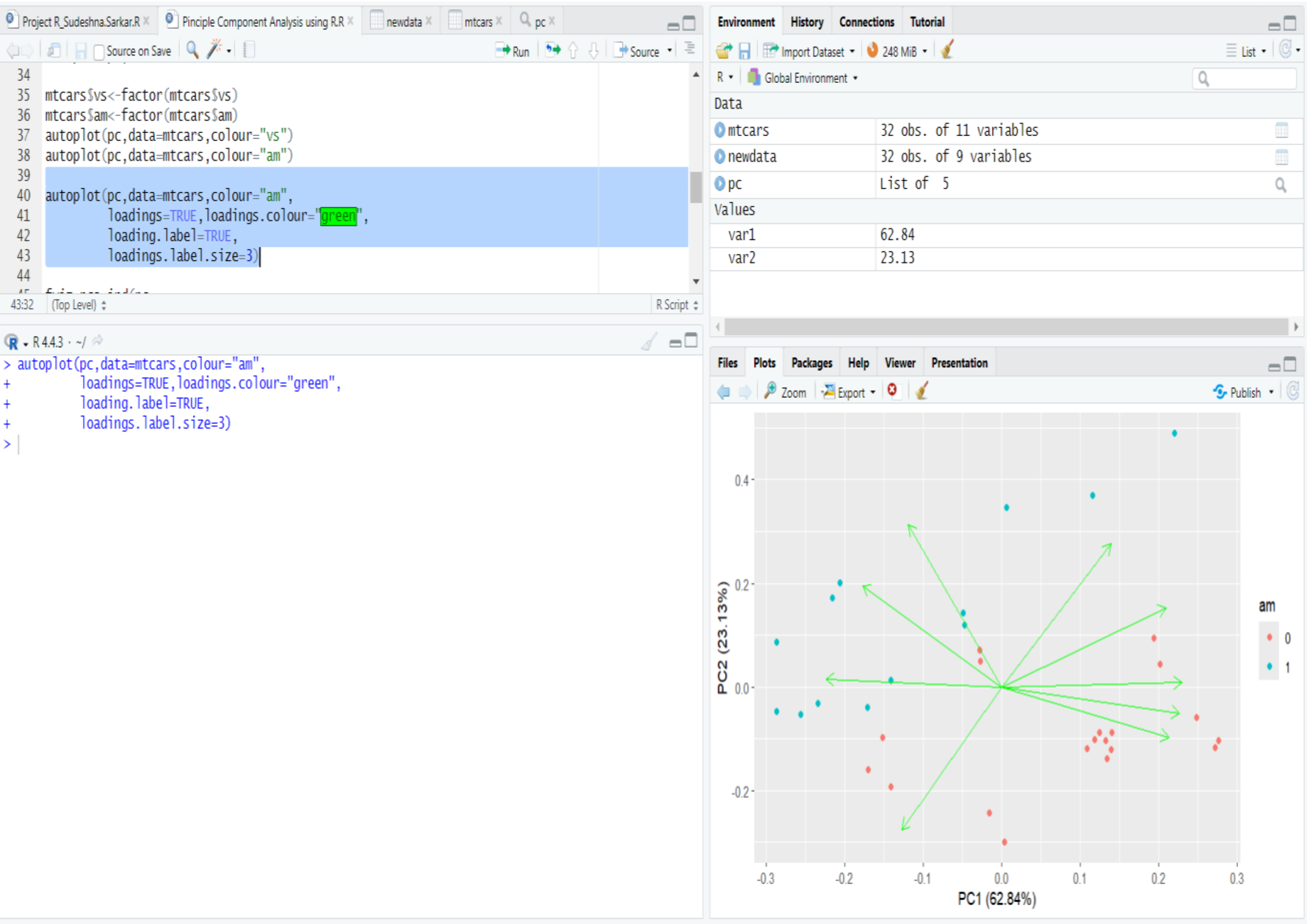

Project R_Sudeshna.Sarkar.R
Pinciple Component Analysis using R.R
newdata
mtcars
pc
Source on Save
Run
Source
#Principle component analysis
pc<-prcomp(newdata,center=TRUE,scale=TRUE)
summary(pc)
attributes(pc)
print(pc)
var1<-round(pc$sdev[1]^2/sum(pc$sdev^2)*100,2)
var2<-round(pc$sdev[2]^2/sum(pc$sdev^2)*100,2)
#Scree plot
plot(pc)
screeplot(x=pc,type="line",main="Scree plot")
18:1
(Top Level)
R Script
R 4.4.3 · ~/
> var1<-round(pc$sdev[1]^2/sum(pc$sdev^2)*100,2)
> var2<-round(pc$sdev[2]^2/sum(pc$sdev^2)*100,2)
Environment
History
Connections
Tutorial
Import Dataset
235 MiB
List
Global Environment
Data
mtcars
32 obs. of 11 variables
newdata
32 obs. of 9 variables
pc
List of 5
Values
var1
62.84
var2
23.13

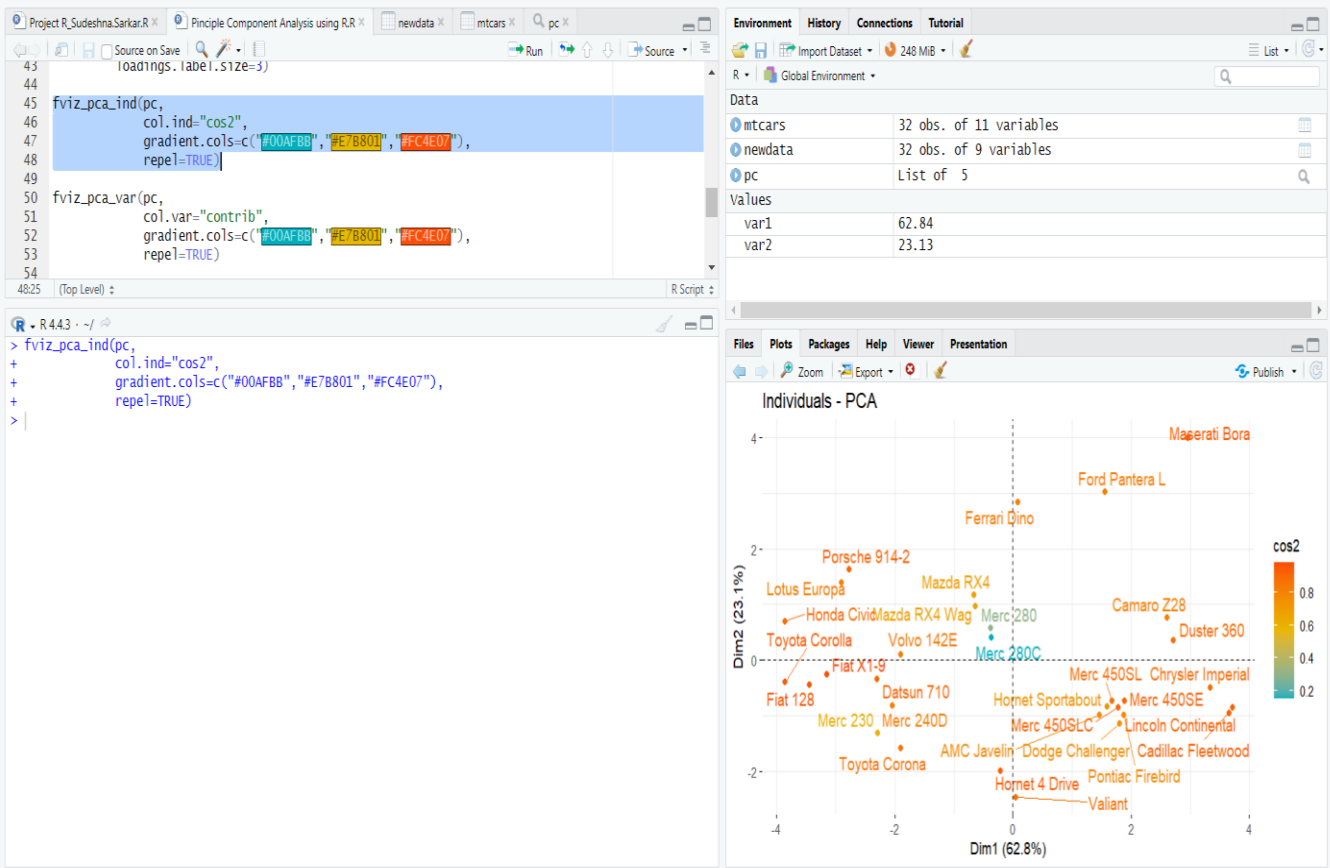

Project R_Sudeshna.Sarkar.R
Pinciple Component Analysis using R.R
newdata
mtcars
pc
Source on Save
Run
Source
var1<-round(pc$sdev[1]^2/sum(pc$sdev^2)*100,2)
var2<-round(pc$sdev[2]^2/sum(pc$sdev^2)*100,2)
#Scree plot
plot(pc)
screeplot(x=pc,type="line",main="Scree plot")
library("factoextra")
fviz_eig(pc)
## PCA plot
#Using Biplot
23:13
(Top Level)
R Script
R 4.4.3 · ~/
> #Scree plot
> plot(pc)
> screeplot(x=pc,type="line",main="Scree plot")
> library("factoextra")
Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
> fviz_eig(pc)
Environment
History
Connections
Tutorial
Import Dataset
243 MiB
List
Global Environment
Data
mtcars
32 obs. of 11 variables
newdata
32 obs. of 9 variables
pc
List of 5
Values
var1
62.84
var2
23.13
Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export
Publish
Scree plot
Percentage of explained variances
Dimensions
0
20
40
60
1
2
3
4
5
6
7
8
9


Project R_Sudeshna Sarkar.R
Pinciple Component Analysis using R.R
newdata
mtcars
pc
Source on Save
Run
Source
## PCA plot
#Using Biplot
biplot(pc)
biplot(pc,cex=0.5)
#Using Autoplot
library("ggfortify")
autoplot(pc)
autoplot(pc,scale=0)
mtcars$vs<-factor(mtcars$vs)
28:19
(Top Level)
R Script
R 4.4.3 · ~/
> ## PCA plot
> #Using Biplot
> biplot(pc)
> biplot(pc,cex=0.5)
>
Environment
History
Connections
Tutorial
Import Dataset
243 MiB
List
Global Environment
Data
mtcars 32 obs. of 11 variables
newdata 32 obs. of 9 variables
pc List of 5
Values
var1 62.84
var2 23.13
Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export
Publish
PC1
PC2
Maserati Bora
Ford Pantera L
Ferrari Dino
Valiant
Hornet 4 Drive
Toyota Corona
Duster 360
Camaro Z28
Chrysler Imperial
Lotus Europa

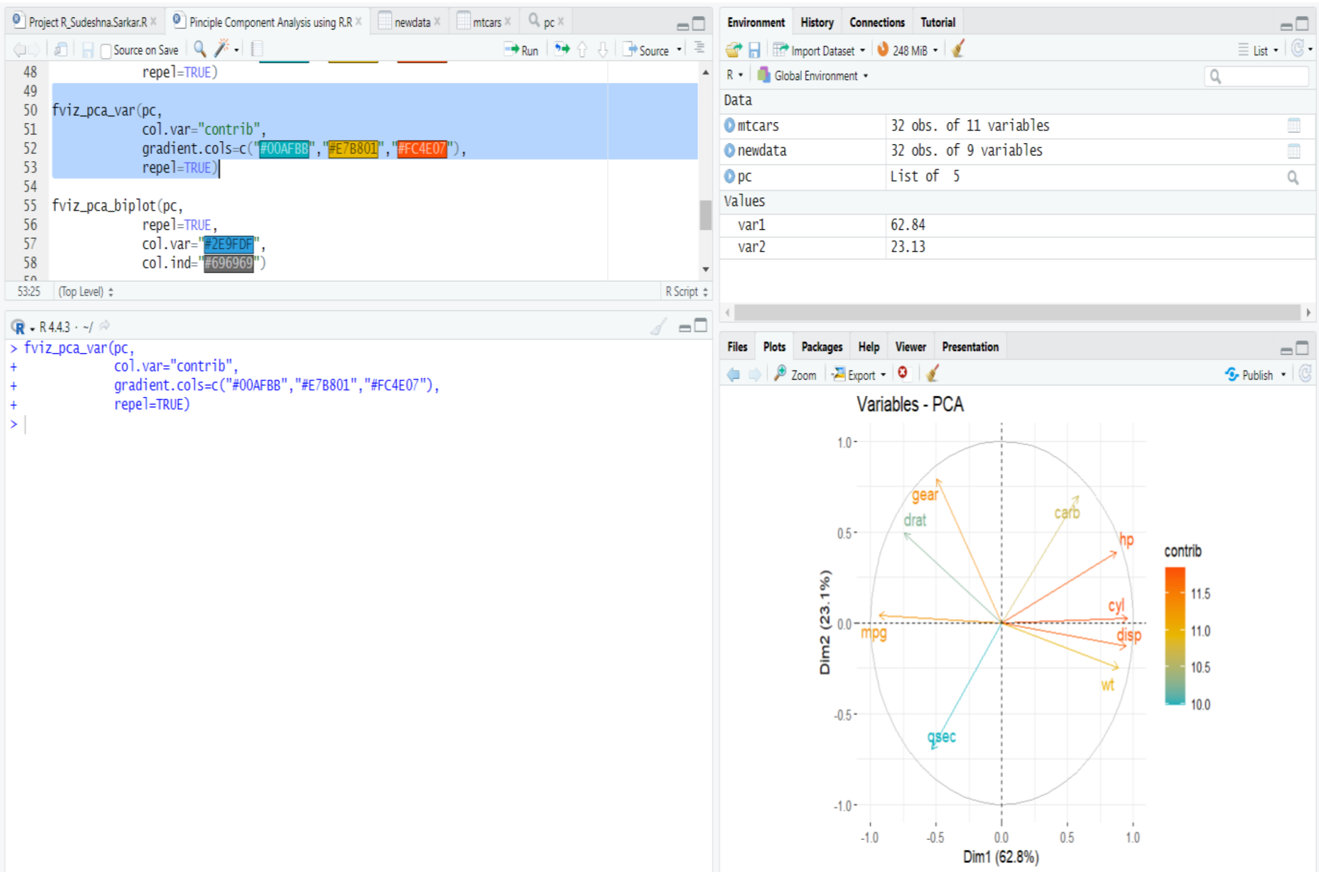

Go to file/function
Addins
Project: (None)
Project R_Sudeshna Sarkar.R
Pinciple Component Analysis using R.R
newdata
mtcars
pc
Source on Save
Run
Source
#Using Autoplot
library("ggfortify")
autoplot(pc)
autoplot(pc,scale=0)
mtcars$vs<-factor(mtcars$vs)
mtcars$am<-factor(mtcars$am)
autoplot(pc,data=mtcars,colour="vs")
autoplot(pc,data=mtcars,colour="am")
33:21
(Top Level)
R Script
R 4.4.3 · ~/
> #Using Autoplot
> library("ggfortify")
> autoplot(pc)
> autoplot(pc,scale=0)
>
Environment
History
Connections
Tutorial
Import Dataset
246 MiB
List
Global Environment
Data
mtcars 32 obs. of 11 variables
newdata 32 obs. of 9 variables
pc List of 5
Values
var1 62.84
var2 23.13
Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export
Publish
PC2 (23.13%)
PC1 (62.84%)


Project R_Sudeshna.Sarkar.R
Pinciple Component Analysis using R.R
newdata
mtcars
pc
Source on Save
Run
Source
#Using Autoplot
library("ggfortify")
autoplot(pc)
autoplot(pc,scale=0)
mtcars$vs<-factor(mtcars$vs)
mtcars$am<-factor(mtcars$am)
autoplot(pc,data=mtcars,colour="vs")
autoplot(pc,data=mtcars,colour="am")
(Top Level)
R Script
R 4.4.3 · ~/
> mtcars$vs<-factor(mtcars$vs)
> mtcars$am<-factor(mtcars$am)
> autoplot(pc,data=mtcars,colour="vs")
> autoplot(pc,data=mtcars,colour="am")
Environment
History
Connections
Tutorial
Import Dataset
248 MiB
List
Global Environment
Data
mtcars
32 obs. of 11 variables
newdata
32 obs. of 9 variables
pc
List of 5
Values
var1
62.84
var2
23.13
Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export
Publish
PC2 (23.13%)
PC1 (62.84%)
am
0
1

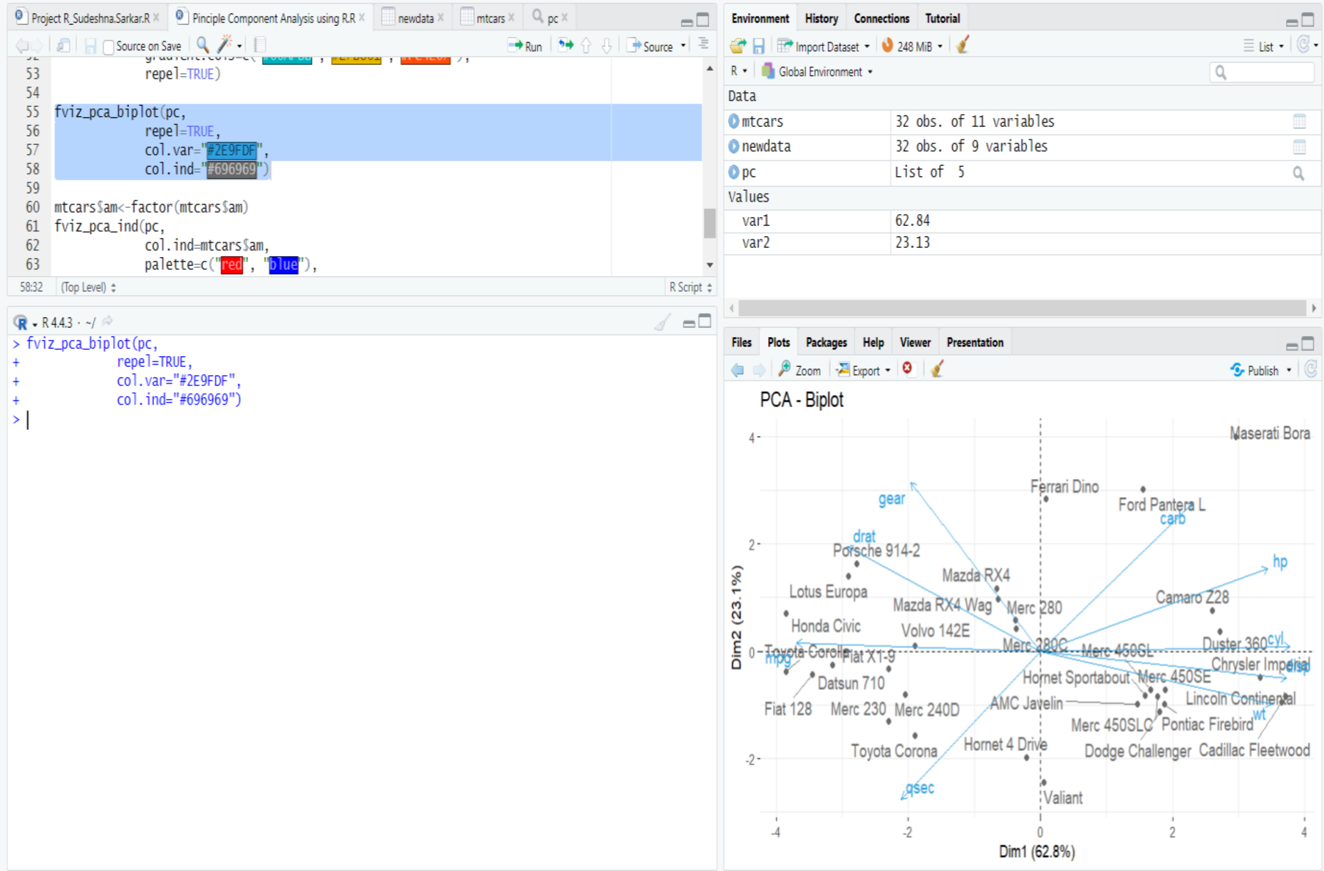

Project R_Sudeshna.Sarkar.R
Pinciple Component Analysis using R.R
newdata
mtcars
pc
Source on Save
Run
Source
mtcars$vs<-factor(mtcars$vs)
mtcars$am<-factor(mtcars$am)
autoplot(pc,data=mtcars,colour="vs")
autoplot(pc,data=mtcars,colour="am")
autoplot(pc,data=mtcars,colour="am",
loadings=TRUE,loadings.colour="green",
loading.label=TRUE,
loadings.label.size=3)
(Top Level)
R Script
R 4.4.3 · ~/
> autoplot(pc,data=mtcars,colour="am",
+ loadings=TRUE,loadings.colour="green",
+ loading.label=TRUE,
+ loadings.label.size=3)
Environment
History
Connections
Tutorial
Import Dataset
248 MiB
List
Global Environment
Data
mtcars
32 obs. of 11 variables
newdata
32 obs. of 9 variables
pc
List of 5
Values
var1
62.84
var2
23.13
Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export
Publish
PC2 (23.13%)
PC1 (62.84%)
am
0
1

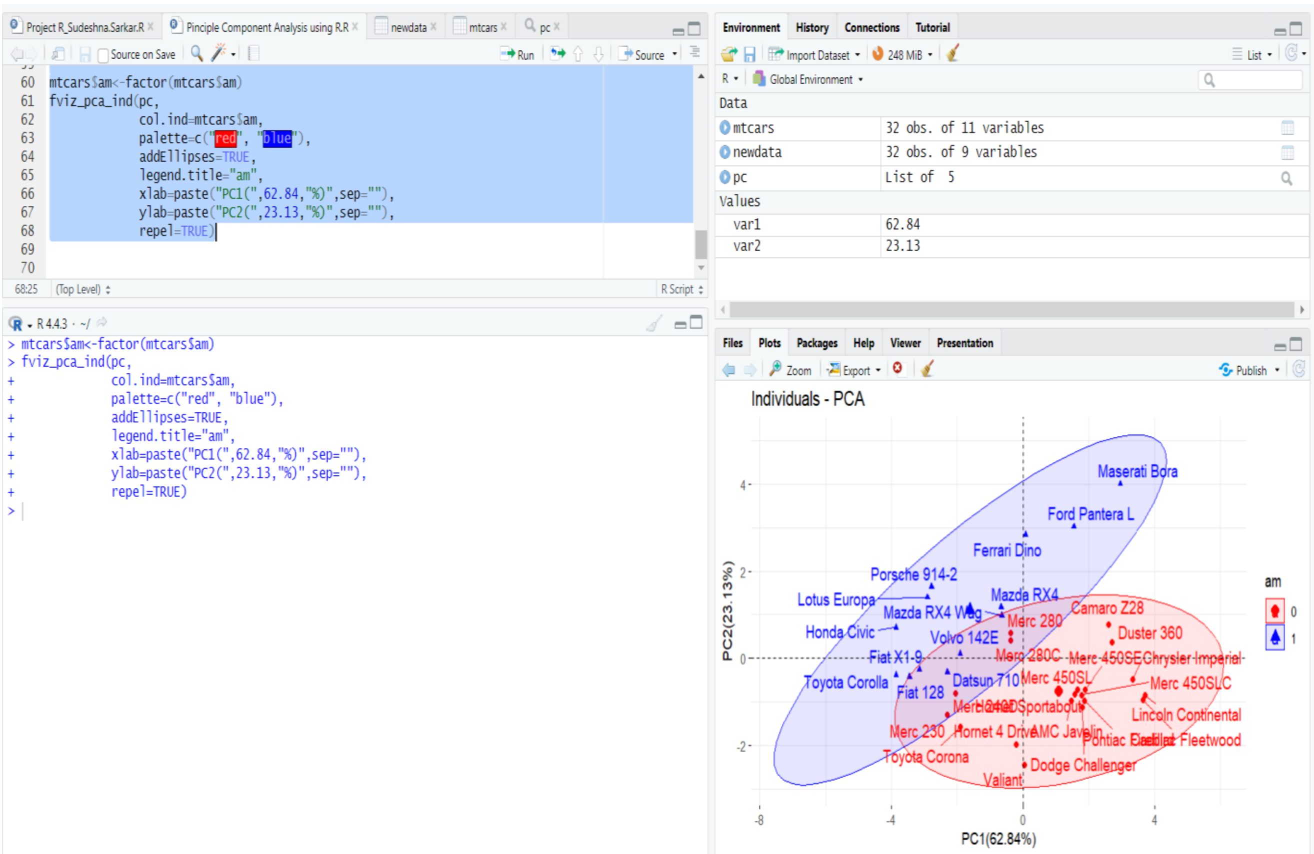


fviz_pca_ind(pc,
col.ind="cos2",
gradient.cols=c("#00AFBB","#E7B801","#FC4E07"),
repel=TRUE)
fviz_pca_var(pc,
col.var="contrib",
gradient.cols=c("#00AFBB","#E7B801","#FC4E07"),
repel=TRUE)
Environment History Connections Tutorial
Data
mtcars 32 obs. of 11 variables
newdata 32 obs. of 9 variables
pc List of 5
Values
var1 62.84
var2 23.13
Individuals - PCA
Dim1 (62.8%)
Dim2 (23.1%)
cos2
Maserati Bora
Ford Pantera L
Ferrari Dino
Porsche 914-2
Lotus Europa
Mazda RX4
Honda Civic
Mazda RX4 Wag
Merc 280
Camaro Z28
Duster 360
Toyota Corolla
Volvo 142E
Merc 280C
Fiat X1-9
Merc 450SL
Chrysler Imperial
Fiat 128
Datsun 710
Hornet Sportabout
Merc 450SE
Merc 230
Merc 240D
Merc 450SLC
Lincoln Continental
AMC Javelin
Dodge Challenger
Cadillac Fleetwood
Toyota Corona
Hornet 4 Drive
Pontiac Firebird
Valiant

fviz_pca_var(pc,
col.var="contrib",
gradient.cols=c("#00AFBB","#E7B801","#FC4E07"),
repel=TRUE)
fviz_pca_biplot(pc,
repel=TRUE,
col.var="#2E9FDF",
col.ind="#696969")
Environment History Connections Tutorial
Data
mtcars 32 obs. of 11 variables
newdata 32 obs. of 9 variables
pc List of 5
Values
var1 62.84
var2 23.13
Variables - PCA
Dim1 (62.8%)
Dim2 (23.1%)
contrib
gear
drat
carb
hp
cyl
mpg
disp
wt
qsec


Project R_Sudeshna.Sarkar.R
Pinciple Component Analysis using R.R
newdata
mtcars
pc
Source on Save
Run
Source
repel=TRUE)
fviz_pca_biplot(pc,
repel=TRUE,
col.var="#2E9FDF",
col.ind="#696969")
mtcars$am<-factor(mtcars$am)
fviz_pca_ind(pc,
col.ind=mtcars$am,
palette=c("red", "blue"),
58:32
(Top Level)
R Script
R 4.4.3 · ~/
> fviz_pca_biplot(pc,
+ repel=TRUE,
+ col.var="#2E9FDF",
+ col.ind="#696969")
Environment
History
Connections
Tutorial
Import Dataset
248 MiB
List
Global Environment
Data
mtcars
32 obs. of 11 variables
newdata
32 obs. of 9 variables
pc
List of 5
Values
var1
62.84
var2
23.13
Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export
Publish
PCA - Biplot
Dim2 (23.1%)
Dim1 (62.8%)


Project R_Sudeshna.Sarkar.R
Pinciple Component Analysis using R.R
newdata
mtcars
pc
Source on Save
Run
Source
mtcars$am<-factor(mtcars$am)
fviz_pca_ind(pc,
col.ind=mtcars$am,
palette=c("red", "blue"),
addEllipses=TRUE,
legend.title="am",
xlab=paste("PC1(",62.84,"%)",sep=""),
ylab=paste("PC2(",23.13,"%)",sep=""),
repel=TRUE)
68:25
(Top Level)
R Script
R 4.4.3 · ~/
> mtcars$am<-factor(mtcars$am)
> fviz_pca_ind(pc,
+ col.ind=mtcars$am,
+ palette=c("red", "blue"),
+ addEllipses=TRUE,
+ legend.title="am",
+ xlab=paste("PC1(",62.84,"%)",sep=""),
+ ylab=paste("PC2(",23.13,"%)",sep=""),
+ repel=TRUE)
Environment
History
Connections
Tutorial
Global Environment
Data
mtcars
32 obs. of 11 variables
newdata
32 obs. of 9 variables
pc
List of 5
Values
var1
62.84
var2
23.13
Files
Plots
Packages
Help
Viewer
Presentation
Individuals - PCA
am
PC2(23.13%)
PC1(62.84%)

### ✅ Key Insight:

This PCA result shows that the dimensionality of the dataset can be reduced from **9 variables** to **2 or 3 principal components** while still retaining most of the important information — helpful for visualization and analysis.

The Principal Component Analysis (PCA) on the dataset helps us understand the most important patterns in a simpler way. The first main component (PC1) explains about 62.86% of the total data variation and the second one (PC2) adds 17.61%. Together, they explain around 80.47% of the total information, which is quite a lot. This means we can focus on just these **two components instead of all 9 original variables**. The results also show how each original feature, like **mileage (`mpg`) or weight (`wt`), affects the components**. Overall, PCA makes the data easier to understand and analyse.