

18 de abril

sergio

2022-04-23

```
#ANALISIS FACTORIAL
```

```
x<-as.data.frame(state.x77)
```

#Quitar los espacios de los nombres

```
colnames(x) [4] = "Life.Exp"  
colnames(x) [6] = "HS.Grad"
```

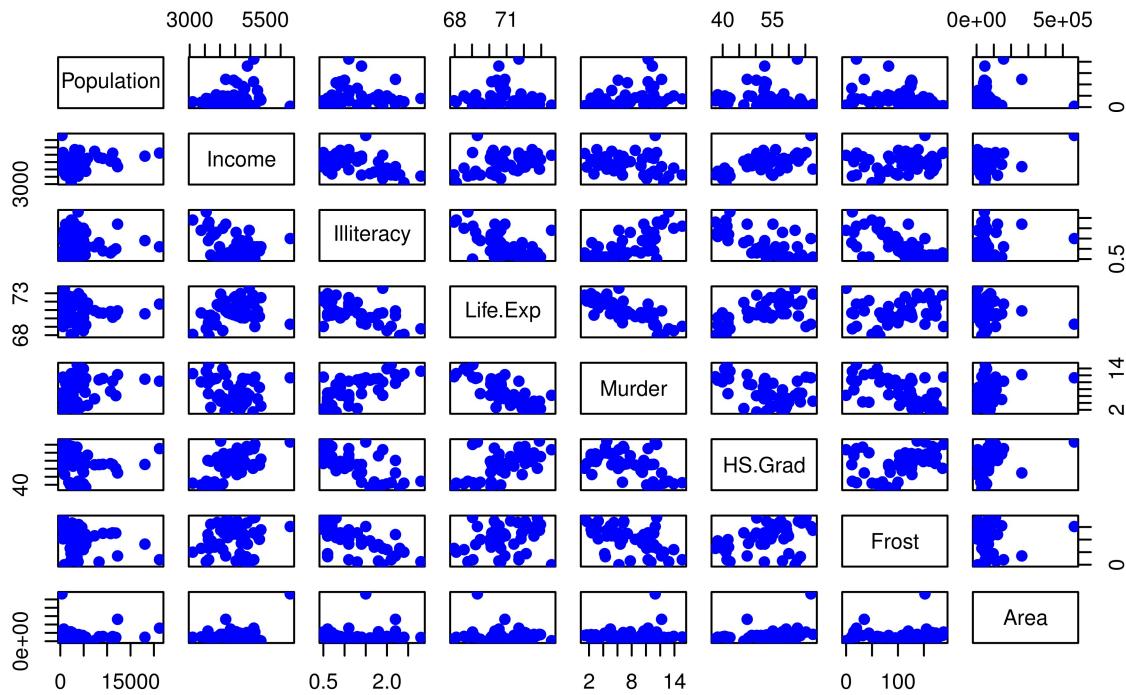
Separa n (estados) y p (variables)

```
n<-dim(x) [1]  
p<-dim(x) [2]
```

Generacion de un scarter plot para la visualización de variables originales

```
pairs(x, col="blue", pch=19, main="matriz original")
```

matriz original



```
# Transformación de alguna variables ##1.- Aplicamos logaritmo para las columnas 1,3 y 8
```

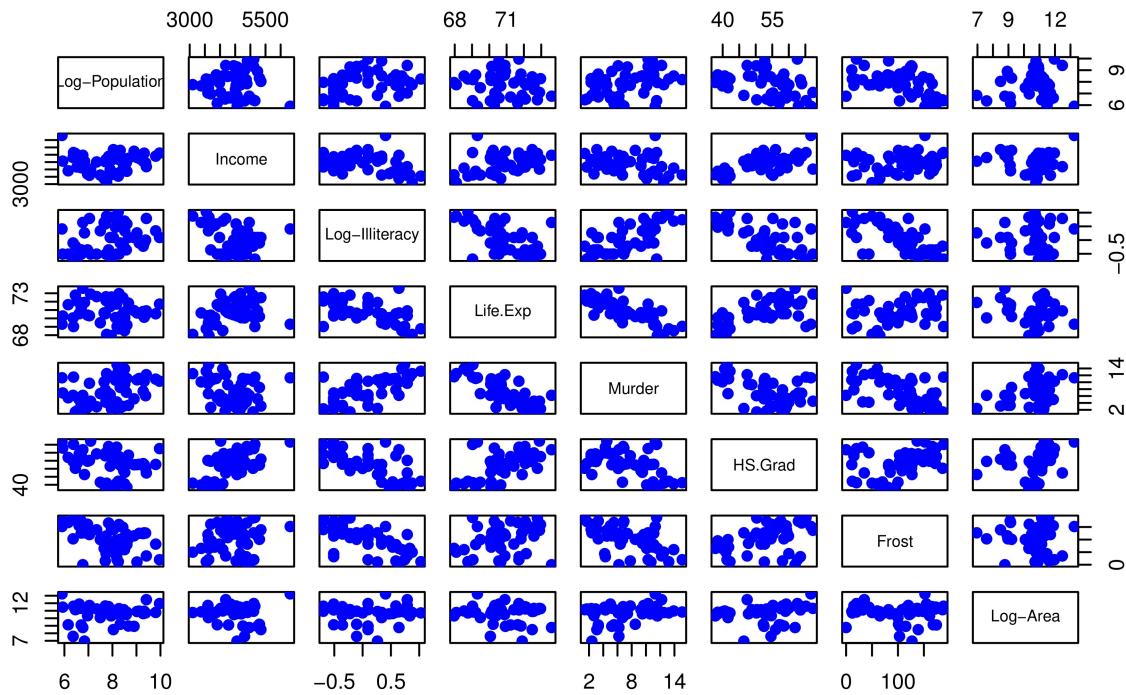
```
x[,1]<-log(x[,1])
colnames(x)[1]<-"Log-Population"

x[,3]<-log(x[,3])
colnames(x)[3]<-"Log-Illiteracy"

x[,8]<-log(x[,8])
colnames(x)[8]<-"Log-Area"

pairs(x,col="blue", pch=19, main="Matriz original")
```

Matriz original



Como las variables tiene diferentes unidades de medida, se va a implementar la matriz de correlaciones para estimar la matriz de carga

Reducción de la dimensionalidad

Análisis Factorial de componentes principales (PCFA)

#1.- Calcular la matriz de medias y de correlaciones

```
mu<-colMeans(x)
R<-cor(x)
```

2.- Reducción de la dimensionalidad mediante Análisis factorial de componentes principales.

1.- Calcular los valores y vectores propios.

```
eR<-eigen(R)
```

2.- Valores propios

```
eigen.val<-eR$values
```

3.- Vectores propios

```
eigen.vec<-eR$vectors
```

4.- Calcular la proporcion de variabilidad

```
prop.var<-eigen.val/sum(eigen.val)
```

5.- Calcular la proporcion de variabilidad acumulada

```
prop.var.acum<-cumsum(eigen.val)/sum(eigen.val)
```

Estimacion de la matriz de carga

se estima la matriz de carga usando los autovalores y autovectores. se aplica la rotación varimax Primera estimación de Lamda mayuscula se calcula multiplicando la matriz de los 3 primeros autovectores por la matriz diagonal formada por la raiz cuadrada de los primeros 3 autovalores.

```
L.est.1<-eigen.vec[,1:3] %*% diag(sqrt(eigen.val[1:3]))
```

Rotación varimax

```
L.est.1.var<-varimax(L.est.1)
```

Estimación de la matriz de los errores

```
##Estimación de la matriz de perturbaciones
```

```
Psi.est.1<-diag(diag(R-as.matrix(L.est.1.var$loadings))%*% t(as.matrix(L.est.1.var$loadings))))
```

##Se utiliza el método Análisis de factor principal (PFA) para estimación de autovalores y autovectores

```

RP<-R-Psi.est.1
eRP<-eigen(RP)

# Autovalores
eigen.val.RP<-eRP$values

# Autovectores
eigen.vec.RP<-eRP$vectors

# Proporcion de variabilidad
prop.var.RP<-eigen.val.RP/ sum(eigen.val.RP)

# Proporcion de variabilidad acumulada
prop.var.RP.acum<-cumsum(eigen.val.RP)/ sum(eigen.val.RP)

# Estimación de la matriz de cargas
# con rotación varimax
L.est.2<-eigen.vec.RP[,1:3] %*% diag(sqrt(eigen.val.RP[1:3]))

# Rotacion varimax
L.est.2.var<-varimax(L.est.2)

# Estimación de la matriz de covarianzas de los errores.
Psi.est.2<-diag(diag(R-as.matrix(L.est.2.var$loadings))%*% t(as.matrix(L.est.2.var$loadings)))

```

Obtencion de los scores de ambos métodos

PCFA

```
FS.est.1<-scale(x)%*% as.matrix(L.est.1.var$loadings)
```

PFA

```
FS.est.2<-scale(x)%*% as.matrix (L.est.2.var$loadings)
```

graficamos ambos scores

```

par(mfrow=c(2,1))
plot(FS.est.1[,1], FS.est.1[,2], xlab="primer factor",
      ylab="segundo factor", main="scores con factor I y II con PCFA",
      pch=19, col="blue")

```

scores con factor I y II con PCFA

