

R para Componentes Principales

Paquetes o librerías sugeridos

```
- library Stats
- library FactoMine R
- library psych
- library factoextra
- library ggplot2
- library(ggrepel)
- library(pastecs)
- library(corrplot)
```

- Ingresar los datos originales de las variables como columnas
- Centrar los datos restando a cada observación su valor medio, o tipificar los datos dividiendo los datos centrados por su desvío
- Armar una matriz con los datos centrados (X_m) y sin centrar (X)
-

```
# Si quieren hacer análisis descriptivo de las variables
```

```
#=====
```

En el ejemplo de clase de los países europeos

Se puede usar paquete pastecs

```
#Descriptivos
```

```
stat.desc(datos,basic=FALSE)
```

```
#Matriz correlaciones
```

Se puede usar paquete corrplot

```
R<-cor(datos[2:10], method="pearson")
```

```
corrplot(R, p.mat = res1[[1]], sig.level=0.05)
```

```
#=====
```

Para el Test de Bartlett

```
#=====
```

Usar del paquete psych

```
cortest.bartlett( R, n= número de observaciones
```

```
#=====
```

```
# Estimar PCA (Análisis de Componentes Principales)
```

```
#=====
```

```
fit<-PCA(datos[2:10],scale.unit=TRUE,ncp=9,graph=TRUE)
```

```
head(fit)
```

```
# Scree plot
fviz_eig(fit, geom="line")+
  theme_grey()
```

- Autovalores y autovectores se obtienen con la función `svd (V)` del paquete `stats`
- Para hallar las puntuaciones factoriales o proyecciones sobre los ejes
Se ejecuta con `prcomp` del paquete `stats`
- Para hallar las cargas factoriales se puede usar la función `PCA` del paquete `Facto Mine R`

```
# Representación gráfica variables
#=====
```

Se puede usar el paquete `ggplot2`

```
datos.grafico2<-data.frame(fit$var$coord[,1:2])
```

```
ggplot(datos.grafico2)+
  geom_point(aes(x=Dim.1, y=Dim.2, colour="darkred"))+
  geom_text_repel(aes(x=Dim.1,
y=Dim.2), label=rownames(datos.grafico2))+
  geom_vline(xintercept = 0, colour="darkgray")+
  geom_hline(yintercept = 0, colour="darkgray")+
  labs (x="Dimensión 1 (40.01%)", y="Dimensión 2 (23.46%)")+
  theme(legend.position="none")
```

O usar el paquete `factoextra`

```
fviz_pca_var(fit, col.var="contrib")+
scale_color_gradient2(low="white", mid="blue", high="red",
  midpoint=10.0)+
theme_gray()
```

```
#=====
# Representación conjunta de variables y objetos
#=====
```

```
library(ggbiplot)
```

```
ggbiplot(fit, obs.scale = 1, var.scale = 1, circle=TRUE)+  
  scale_color_discrete(name = '')+  
  expand_limits(x=c(-8,4), y=c(-2.5, 2.5))+  
  labs (x="Dimensión 1 (40.01%)", y="Dimensión 2 (23.46%)")+  
  geom_text_repel(aes(x=datos.grafico$Dim.1, y=datos.grafico$Dim.2),  
    label=datos.grafico$pais, size=3)
```
