

Análisis PCA

Diego Alejandro Estrada Rivera 165352

Leer los datos y colocarlos en un data frame

```
INEGI <-  
read.csv(url("https://raw.githubusercontent.com/mauriciogtec/Propedeutico  
DataScience2017/master/Datos/DatosINEGI.csv"))
```

Visualizar un porción de los datos para asegurar su integridad

```
head(INEGI)
```

##	Estado	Poblacion	PIBpc	Secundarias
## 1	Aguascalientes	1184996	84.70	348
## 2	Baja California	3155070	83.07	595
## 3	Baja California Sur	637026	94.64	156
## 4	Campeche	822441	395.55	318
## 5	Coahuila de Zaragoza	2748391	106.05	553
## 6	Colima	650555	76.49	169

##	IndiceAprovechamientoSecundaria	PorcentajeAnalfabetas
## 1	85.3	3.26
## 2	86.0	2.57
## 3	87.9	3.21
## 4	78.5	8.31
## 5	75.9	2.63
## 6	81.8	5.13

##	DefuncionesGenerales	Nacimientos	Divorcios	Matrimonios
## 1	5267	26933	1417	6210
## 2	14756	62871	3294	17535
## 3	2739	12864	744	2856
## 4	3964	18343	1138	5267
## 5	15200	59966	3668	15384
## 6	3717	13422	658	3388

##	PorcentajePartosHospitales	PorcentajeAguaPotable
## 1	97.1	98.0
98.9		
## 2	65.7	93.3
95.9		
## 3	95.2	86.7
92.4		
## 4	87.0	89.5
90.3		
## 5	90.3	97.9
98.2		
## 6	98.5	97.9
98.5		

```
## PorcentajeElectricidad PorcentajeParedesSolidas PorcentajePisoTierra
## 1 99.2 92.3 1.7
## 2 98.5 77.0 3.3
## 3 96.7 90.3 5.8
## 4 96.8 80.7 4.7
## 5 99.1 84.8 1.6
## 6 99.0 94.7 4.5
```

Las variables que necesitamos hacer per capita para poder hacer un análisis son Secundarias, DefuncionesGenerales, Nacimientos, Divorcios, Matrimonios.

```
for (col in c("Secundarias", "DefuncionesGenerales", "Nacimientos",
"Divorcios", "Matrimonios")) {
INEGI[,col] <- INEGI[,col] / INEGI[, "Poblacion"]
}
```

Ahora se guardan en una matriz de datos las variables que queremos analizar (todas menos Estado y Población), y vamos a guardar Estado como row.names para usarlos como los nombres de las filas

```
X <- INEGI[, -(1:2)] # Remueve Las primeras dos columnas
head(X)
```

```
## PIBpc Secundarias IndiceAprovechamientoSecundaria
## 1 84.70 0.0002936719 85.3
## 2 83.07 0.0001885854 86.0
## 3 94.64 0.0002448880 87.9
## 4 395.55 0.0003866539 78.5
## 5 106.05 0.0002012086 75.9
## 6 76.49 0.0002597782 81.8
## PorcentajeAnalfabetas DefuncionesGenerales Nacimientos Divorcios
## 1 3.26 0.004444741 0.02272835 0.001195785
## 2 2.57 0.004676917 0.01992697 0.001044034
## 3 3.21 0.004299668 0.02019384 0.001167927
## 4 8.31 0.004819799 0.02230312 0.001383686
## 5 2.63 0.005530509 0.02181858 0.001334599
## 6 5.13 0.005713583 0.02063161 0.001011444
## Matrimonios PorcentajePartosHospitales PorcentajeAguaPotable
## 1 0.005240524 97.1 98.0
## 2 0.005557721 65.7 93.3
## 3 0.004483333 95.2 86.7
## 4 0.006404107 87.0 89.5
## 5 0.005597457 90.3 97.9
## 6 0.005207861 98.5 97.9
## PorcentajeAguaEntubada PorcentajeElectricidad
PorcentajeParedesSolidas
## 1 98.9 99.2
92.3
## 2 95.9 98.5
77.0
```

```
## 3          92.4          96.7
90.3
## 4          90.3          96.8
80.7
## 5          98.2          99.1
84.8
## 6          98.5          99.0
94.7
## PorcentajePisoTierra
## 1          1.7
## 2          3.3
## 3          5.8
## 4          4.7
## 5          1.6
## 6          4.5
```

```
row.names(X) <- INEGI$Estado
head(X)
```

```
##          PIBpc  Secundarias
IndiceAprovechamientoSecundaria
## Aguascalientes      84.70 0.0002936719
85.3
## Baja California      83.07 0.0001885854
86.0
## Baja California Sur  94.64 0.0002448880
87.9
## Campeche            395.55 0.0003866539
78.5
## Coahuila de Zaragoza 106.05 0.0002012086
75.9
## Colima              76.49 0.0002597782
81.8
##          PorcentajeAnalfabetas DefuncionesGenerales
## Aguascalientes          3.26      0.004444741
## Baja California          2.57      0.004676917
## Baja California Sur      3.21      0.004299668
## Campeche                 8.31      0.004819799
## Coahuila de Zaragoza      2.63      0.005530509
## Colima                   5.13      0.005713583
##          Nacimientos  Divorcios Matrimonios
## Aguascalientes      0.02272835 0.001195785 0.005240524
## Baja California      0.01992697 0.001044034 0.005557721
## Baja California Sur   0.02019384 0.001167927 0.004483333
## Campeche             0.02230312 0.001383686 0.006404107
## Coahuila de Zaragoza  0.02181858 0.001334599 0.005597457
## Colima               0.02063161 0.001011444 0.005207861
##          PorcentajePartosHospitales PorcentajeAguaPotable
## Aguascalientes          97.1          98.0
```

## Baja California	65.7	93.3
## Baja California Sur	95.2	86.7
## Campeche	87.0	89.5
## Coahuila de Zaragoza	90.3	97.9
## Colima	98.5	97.9
##	PorcentajeAguaEntubada	PorcentajeElectricidad
## Aguascalientes	98.9	99.2
## Baja California	95.9	98.5
## Baja California Sur	92.4	96.7
## Campeche	90.3	96.8
## Coahuila de Zaragoza	98.2	99.1
## Colima	98.5	99.0
##	PorcentajeParedesSolidas	PorcentajePisoTierra
## Aguascalientes	92.3	1.7
## Baja California	77.0	3.3
## Baja California Sur	90.3	5.8
## Campeche	80.7	4.7
## Coahuila de Zaragoza	84.8	1.6
## Colima	94.7	4.5

Instalamos y llamamos la libreria "FactoMineR" que necesitaremos para hacer el análisis de componentes principales

```
#install.packages("FactoMineR")
```

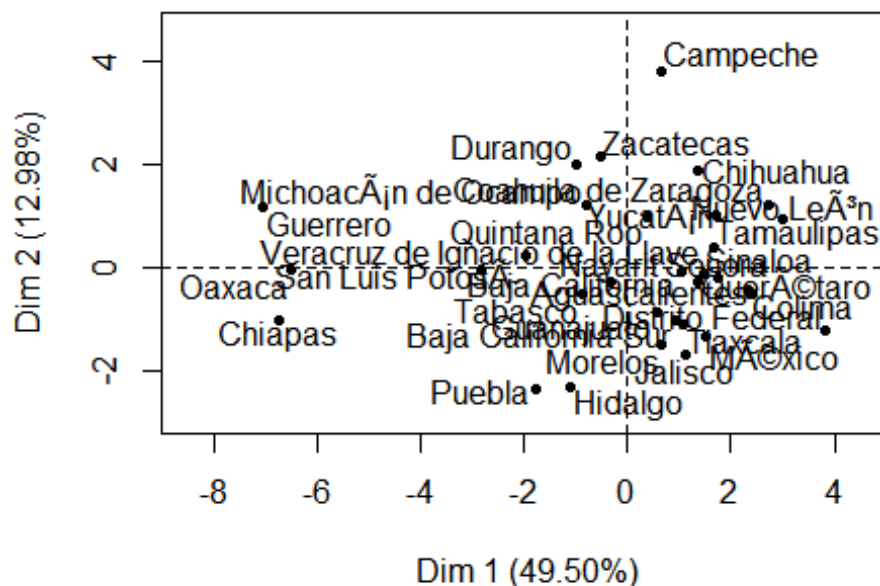
```
library(FactoMineR)
```

```
## Warning: package 'FactoMineR' was built under R version 3.4.1
```

Realizamos el análisis PCA

```
model <- PCA(X)
```

Individuals factor map (PCA)



Variables factor map (PCA)

