

## Lab ACP

Edwin Sanchez    Stephanie Tamayo    Andres Felipe Torres    Fredy  
Urrea    Sergio Velasquez    Manuel Espitia

2025-05-18

# Section 1

## Introduccion

# Introduccion

## Carga de archivos, organizacion de los datos y renombrar variables

```
CIUDADES <- readxl::read_excel("ciudades original-filtrado-con etiquetas (1).xlsx",  
                               sheet = 1)  
  
RHINT <- CIUDADES[,c("CIUDADES", "RH_1", "RH_2", "RH_5", "RH_6", "RH_7",  
                    "RH_8", "RH_9", "RH_10", "RH_11", "RH_12", "RH_13",  
                    "RH_14", "RH_15", "RH_16", "INT_66", "INT_67",  
                    "INT_68", "INT_69", "INT_70")]  
  
colnames(RHINT)[2:20] <- c("PC", "TCP", "AA", "CBPS", "CBES", "RAP", "CC",  
                          "CPT", "CS", "MI", "DCFS", "VI", "H", "S", "BCPIB",  
                          "CI", "TOE", "DP", "DM")
```



## Section 2

### **Punto 2.**

## 2. Preparación de datos

Seleccionamos las variables activas (internacionalización) y las variables

```
base_acp <- RHINT[, c("BCPIB", "CI", "TOE", "DP", "DM",  
                      "PC", "TCP", "AA", "CBPS", "CBES", "RAP", "CC",  
                      "CPT", "CS", "MI", "DCFS", "VI", "H", "S")]
```

### 3. Análisis de Componentes Principales (ACP)

```
res.pca <- PCA(base_acp, scale.unit = TRUE, quanti.sup = 6:19, graph = FALSE)
```

## 4. Resultados

### a) Valores propios

```
eig.val <- round(res.pca$eig, 3)
kable(eig.val, format = "latex", booktabs = TRUE,
      caption = "Valores propios y varianza explicada") %>%
  kable_styling(latex_options = c("striped", "hold_position"))
```

**Table 1:** Valores propios y varianza explicada

	eigenvalue	percentage of variance	cumulative percentage of variance
comp 1	3.110	62.199	62.199
comp 2	1.162	23.238	85.437
comp 3	0.476	9.520	94.957
comp 4	0.252	5.043	100.000
comp 5	0.000	0.000	100.000



## b) Correlaciones de variables activas

```
kable(round(res.pca$var$cor, 3), format = "latex",  
      booktabs = TRUE,  
      caption = "Correlación de variables activas con los factores") %>%  
kable_styling(latex_options = c("striped", "hold_position"))
```

**Table 2:** Correlación de variables activas con los factores

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
BCPIB	-0.699	0.672	0.218	0.113	0
CI	0.924	-0.352	0.104	-0.103	0
TOE	0.802	0.285	0.523	-0.037	0
DP	0.876	0.178	-0.192	0.404	0
DM	0.596	0.688	-0.327	-0.253	0

### c) Correlaciones de variables ilustrativas

```
kable(round(res.pca$quanti.sup$cor, 3), format = "latex",  
      booktabs = TRUE,  
      caption = "Correlación de variables ilustrativas con los factores") %>%  
kable_styling(latex_options = c("striped", "hold_position"))
```

**Table 3:** Correlación de variables ilustrativas con los factores

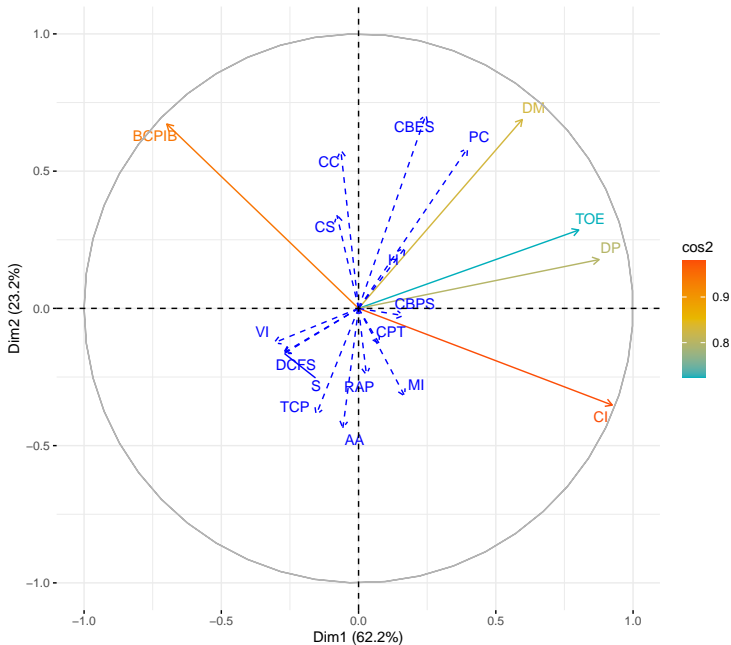
	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
PC	0.397	0.583	0.062	-0.436	0
TCP	-0.154	-0.387	-0.040	-0.305	0
AA	-0.057	-0.435	-0.067	-0.231	0
CBPS	0.163	-0.026	-0.277	0.112	0
CBES	0.246	0.706	-0.038	-0.462	0
RAP	0.029	-0.242	-0.200	-0.066	0
CC	-0.065	0.577	-0.155	-0.334	0
CPT	0.076	-0.128	-0.164	-0.040	0
CS	-0.079	0.340	-0.066	0.086	0
MI	0.168	-0.319	0.163	-0.313	0
DCFS	-0.270	-0.165	0.103	-0.038	0
VI	-0.308	-0.125	-0.017	-0.108	0
H	0.169	0.219	0.094	0.094	0
S	-0.273	-0.162	0.137	0.127	0

## 5. Visualizaciones

### a) Mapa de variables

```
fviz_pca_var(res.pca, col.var = "cos2",  
             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),  
             repel = TRUE, title = "Variables activas e ilustrativas")
```

# Variables activas e ilustrativas



## Punto 3.

Utilizar la funcion PCA del paquete FactoMiner para realizar un ACP de las variables de recursos humanos (RH) utilizando como ilustrativas las que le correspondieron al grupo

### Carga de archivos y ejecucion del ACP

```
# Estandrizacion y asignacion de nombres de la filas

RHINTce <- scale(RHINT[,2:20], scale = TRUE, center = TRUE)
colnames(RHINTce) <- colnames(RHINT)[2:20]
rownames(RHINTce) <- RHINT$CIUDADES

require(FactoMineR)

#ACP sin grafica
pcaRHINTce <- PCA(RHINTce, quanti.sup = c(15:19), graph = F, ncp = 6)
#pcaRHINTce
```

## Valores propios y varianza explicada

**Table 4:** Valores propios, porcentaje de varianza y varianza acumulada

	Valor propio	% Varianza	% Acumulado
comp 1	3.5715	25.5104	25.5104
comp 2	2.2390	15.9932	41.5036
comp 3	2.0716	14.7971	56.3007
comp 4	1.6027	11.4479	67.7486
comp 5	1.2583	8.9880	76.7366
comp 6	1.1918	8.5128	85.2494
comp 7	0.7712	5.5087	90.7581
comp 8	0.4614	3.2959	94.0539
comp 9	0.3276	2.3401	96.3941
comp 10	0.2153	1.5376	97.9316
comp 11	0.1288	0.9199	98.8515
comp 12	0.0942	0.6726	99.5241
comp 13	0.0538	0.3845	99.9086
comp 14	0.0128	0.0914	100.0000

### **Análisis de los valores propios y varianza explicada:**

Los resultados del ACP revelan que los primeros seis componentes presentan valores propios mayores a 1, explicando en conjunto el 85.25% de la varianza total. El primer componente principal captura el 25.51% de la variabilidad, mostrando la mayor influencia individual, mientras que los dos primeros componentes acumulan el 41.50%, lo que permite una representación bidimensional inicial aunque limitada. La inclusión progresiva de componentes hasta el sexto

**Table 5:** Correlaciones variables factor primeras 6 dimensiones

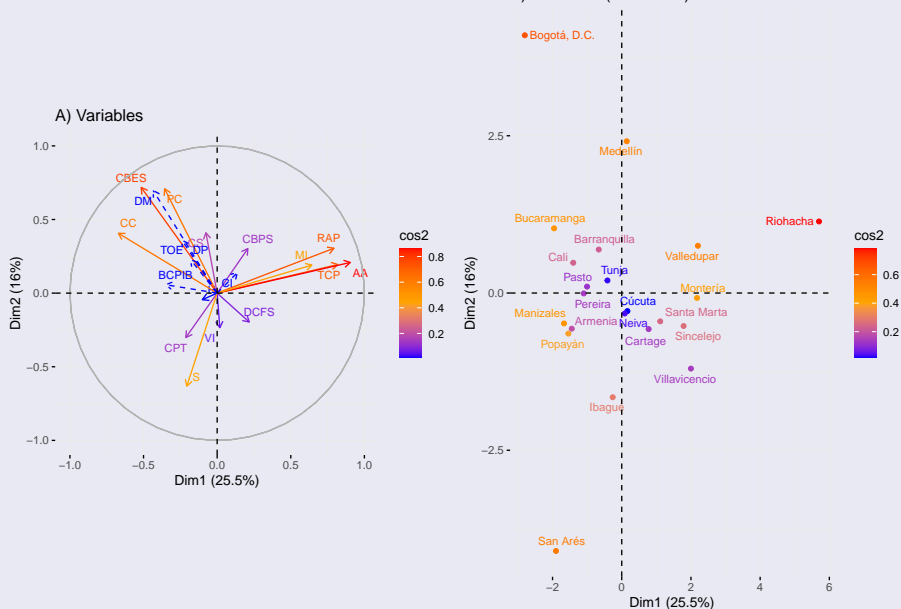
	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Dim.6
PC	-0.3569	0.7088	-0.0370	0.2907	0.0014	0.4134
TCP	0.8193	0.1945	-0.1064	0.3789	0.1926	0.0511
AA	0.9042	0.2083	-0.2736	-0.0455	0.0122	0.0645
CBPS	0.2085	0.3028	0.6300	-0.5551	-0.1747	-0.1542
CBES	-0.5160	0.7179	0.1057	0.0908	0.1345	0.2576
RAP	0.7936	0.3049	0.2618	-0.0864	-0.2436	0.1866
CC	-0.6709	0.4063	0.1403	0.1353	0.2379	-0.0390
CPT	-0.2128	-0.3014	-0.1846	-0.5536	0.3631	0.4263
CS	-0.0764	0.4103	-0.2834	0.2444	0.0393	-0.7590
MI	0.6433	0.1904	-0.2198	0.0607	0.6197	0.0650
DCFS	0.2180	-0.1962	0.7679	0.4088	0.2387	0.1102
VI	0.0181	-0.2347	0.8007	0.3024	0.0763	-0.0669
H	-0.1010	-0.0445	0.1945	-0.3387	0.6877	-0.3074
S	-0.2064	-0.6319	-0.2376	0.5359	0.0577	0.1202

## Análisis de correlaciones variables-factor en las primeras seis dimensiones:

El análisis de las correlaciones revela patrones significativos en la estructura de los datos. En la Dimensión 1, la Tasa de Ocupación (0.904) y Tasa de Crecimiento Poblacional (0.819) muestran las correlaciones más altas, sugiriendo que este eje representa principalmente dinamismo económico y laboral. Contrariamente, la *Calidad de los Colegios* (-0.671) y *Cobertura en Educación Superior* (-0.516) presentan correlaciones negativas, indicando una posible tensión

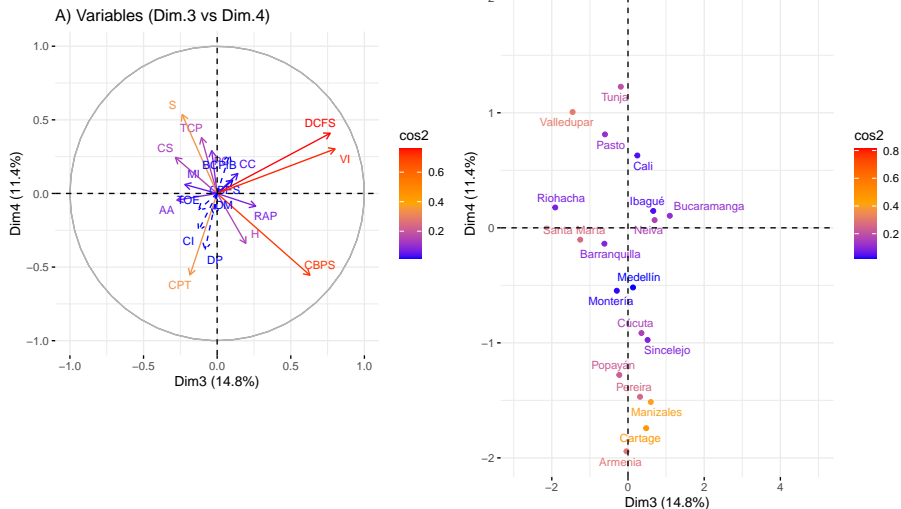
# Variables y objetos graficados en distintas dimensiones

B) Ciudades (Individuos)

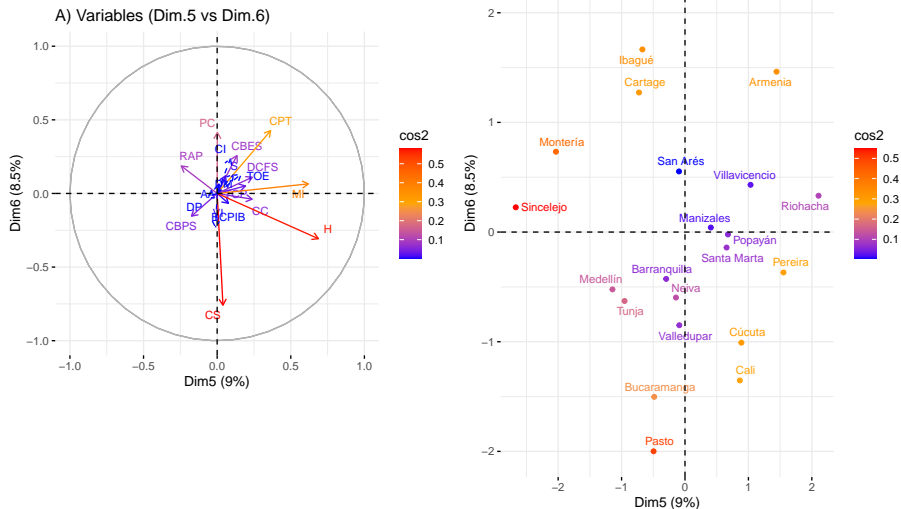


**Figure 1:** Figura 1. Variables y ciudades en las Dimensiones 1 y 2





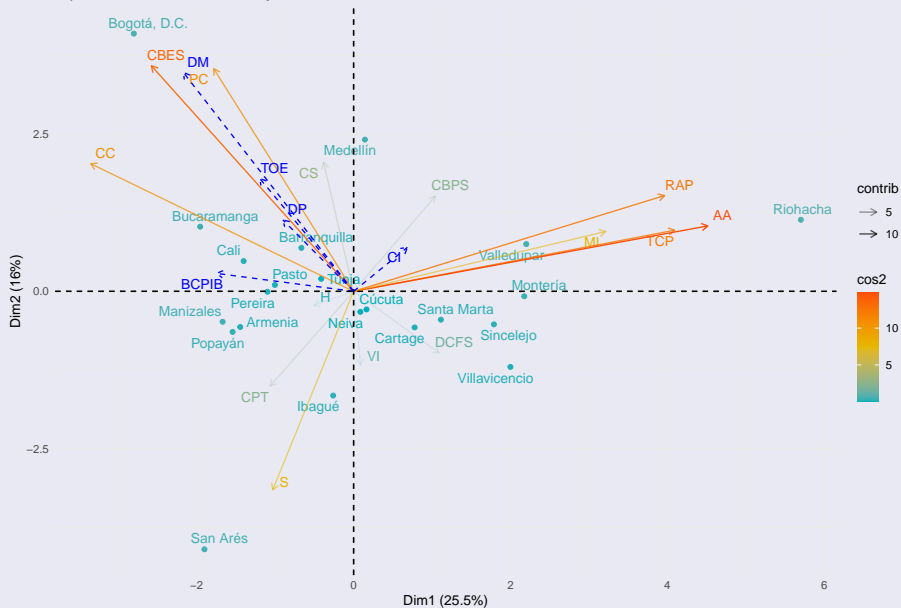
**Figure 2:** Figura 1. Variables y ciudades en las Dimensiones 3 y 4



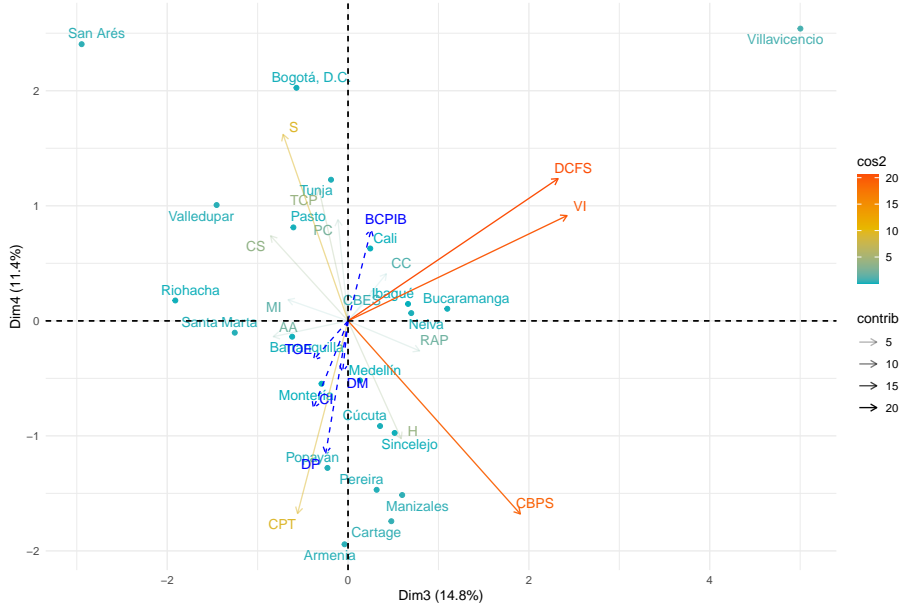
**Figure 3:** Figura 4. Variables y ciudades en las Dimensiones 5 y 6

# Biplots

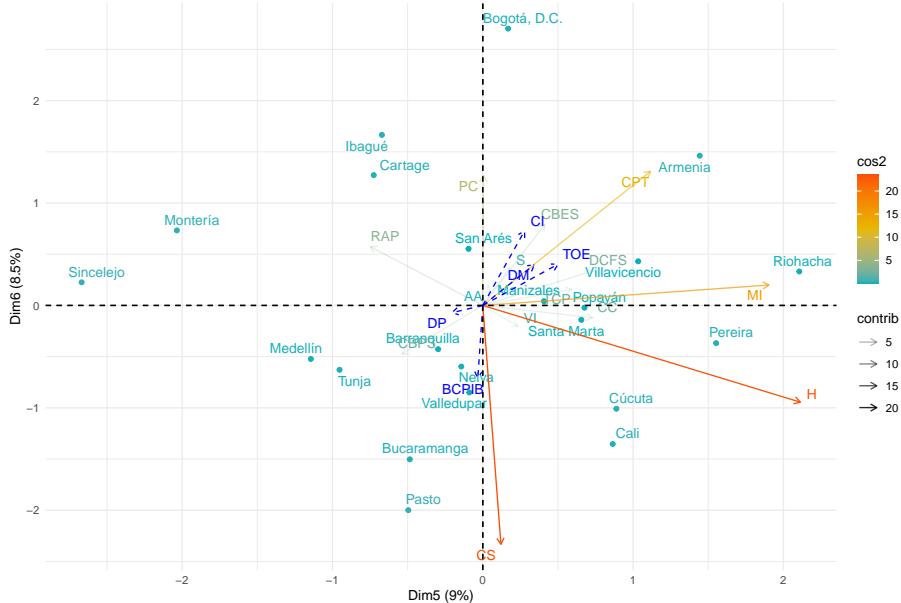
Biplot ACP – Dimensiones 1 y 2



**Figure 4:** Figura 4. Biplot: Variables y ciudades en las Dimensiones 1 y 2



**Figure 5:** Figura 4. biplot Dimensiones 3 y 4



**Figure 6:** Figura 4. biplot Dimensiones 5 y 6

# Cosenos cuadrados de los objetos con los factores

**Table 6:** Resultados del PCA: Coordenadas, Contribución y Calidad de Representación

Variable	Coord_Dim1	Coord_Dim2	Contrib_Dim1	Contrib_Dim2	Cos2_Dim1	Cos2_Dim2
PC	-0.3569	0.7088	3.57	22.44	0.1274	0.5024
TCP	0.8193	0.1945	18.80	1.69	0.6713	0.0378
AA	0.9042	0.2083	22.89	1.94	0.8175	0.0434
CBPS	0.2085	0.3028	1.22	4.09	0.0435	0.0917
CBES	-0.5160	0.7179	7.46	23.02	0.2663	0.5154
RAP	0.7936	0.3049	17.63	4.15	0.6298	0.0930
CC	-0.6709	0.4063	12.60	7.37	0.4501	0.1651
CPT	-0.2128	-0.3014	1.27	4.06	0.0453	0.0908
CS	-0.0764	0.4103	0.16	7.52	0.0058	0.1684
MI	0.6433	0.1904	11.59	1.62	0.4138	0.0363
DCFS	0.2180	-0.1962	1.33	1.72	0.0475	0.0385
VI	0.0181	-0.2347	0.01	2.46	0.0003	0.0551
H	-0.1010	-0.0445	0.29	0.09	0.0102	0.0020
S	-0.2064	-0.6319	1.19	17.83	0.0426	0.3993

## Análisis de las Tablas de Coordenadas, Contribución y Calidad de Representación (cos2):

El análisis de las tablas 3, 4 y 5 revela la manera en que cada variable original contribuye y se proyecta sobre las dimensiones principales extraídas del Análisis de Componentes Principales (ACP).

En las dimensiones 1 y 2 (tabla 3), destacan por su alta contribución el Analfabetismo Absoluto (AA) en la Dim1 (22.89%), la Cobertura bruta en educación superior (CBES) con una fuerte

**Tabla 7:** Resultados del PCA: Coordenadas, Contribución y Calidad de Representación

Variable	Coord_Dim3	Coord_Dim4	Contrib_Dim3	Contrib_Dim4	Cos2_Dim3	Cos2_Dim4
PC	-0.0370	0.2907	0.07	5.27	0.0014	0.0845
TCP	-0.1064	0.3789	0.55	8.96	0.0113	0.1436
AA	-0.2736	-0.0455	3.61	0.13	0.0748	0.0021
CBPS	0.6300	-0.5551	19.16	19.23	0.3969	0.3082
CBES	0.1057	0.0908	0.54	0.51	0.0112	0.0082
RAP	0.2618	-0.0864	3.31	0.47	0.0685	0.0075
CC	0.1403	0.1353	0.95	1.14	0.0197	0.0183
CPT	-0.1846	-0.5536	1.64	19.13	0.0341	0.3065
CS	-0.2834	0.2444	3.88	3.73	0.0803	0.0598
MI	-0.2198	0.0607	2.33	0.23	0.0483	0.0037
DCFS	0.7679	0.4088	28.46	10.42	0.5896	0.1671
VI	0.8007	0.3024	30.95	5.70	0.6412	0.0914
H	0.1945	-0.3387	1.83	7.16	0.0378	0.1147
S	-0.2376	0.5359	2.72	17.92	0.0564	0.2872

En las dimensiones 3 y 4 (tabla 4), las variables de seguridad cobran protagonismo: Delitos contra la formación sexual (DCFS) y Violencia Intrafamiliar (VI) son las que más contribuyen a la Dim3 (28.46% y 30.95%, respectivamente), mientras que la variable Secuestros (S) tiene un rol importante en la Dim4 (17.92%). Esto sugiere que estas dimensiones capturan esencialmente aspectos relacionados con violencia y criminalidad, diferenciando ciudades según su situación en materia de seguridad.

## Figura 5.1: Resultados del PCA: Coordenadas, Contribución y Calidad de Representación

Variable	Coord_Dim5	Coord_Dim6	Contrib_Dim5	Contrib_Dim6	Cos2_Dim5	Cos2_Dim6
PC	0.0014	0.4134	0.00	14.34	0.0000	0.1709
TCP	0.1926	0.0511	2.95	0.22	0.0371	0.0026
AA	0.0122	0.0645	0.01	0.35	0.0001	0.0042
CBPS	-0.1747	-0.1542	2.42	2.00	0.0305	0.0238
CBES	0.1345	0.2576	1.44	5.57	0.0181	0.0663
RAP	-0.2436	0.1866	4.72	2.92	0.0593	0.0348
CC	0.2379	-0.0390	4.50	0.13	0.0566	0.0015
CPT	0.3631	0.4263	10.48	15.25	0.1319	0.1817
CS	0.0393	-0.7590	0.12	48.34	0.0015	0.5761
MI	0.6197	0.0650	30.52	0.35	0.3841	0.0042
DCFS	0.2387	0.1102	4.53	1.02	0.0570	0.0122
VI	0.0763	-0.0669	0.46	0.38	0.0058	0.0045
H	0.6877	-0.3074	37.59	7.93	0.4730	0.0945
S	0.0577	0.1202	0.26	1.21	0.0033	0.0144

Por su parte, en las dimensiones 5 y 6 (tabla 5), se destaca Mortalidad Infantil (MI), que explica el 48.34% de la Dim6 y tiene un cos2 alto (0.5761), reflejando una representación muy fiel en este plano. También Homicidios (H) muestra una alta contribución (37.97%) en la Dim5 y se representa con gran claridad ( $\text{cos2} = 0.4730$ ). Estas dimensiones permiten interpretar patrones más específicos o residuales, principalmente vinculados con indicadores críticos de salud pública y violencia letal.