

Proyecto Final: Análisis de Correspondencia Simple

Ana Karen Martínez Marín

2022-06-06

“Incidencias de violencia atendidas en el Centro de Justicia para las Mujeres (CJM) de la zona Central y Noreste del país”

Introducción

Como se ha visto actualmente en nuestro país se ha notado un aumento en violencia contra la mujer, ya sea física, emocional, psicológica, sexual, por mencionar algunos. Esto es un problema social que debe de ser principalmente prevenido y en su defecto atendido, brindar atención y apoyo a aquellas mujeres que han sido víctimas de esta violación hacia sus derechos humanos.

Por fortuna a lo largo del territorio mexicano actualmente existen 55 *Centros de Justicia para las Mujeres (CJM)* distribuidos en las 31 entidades federativas según datos del (Instituto Nacional de Estadística y Geografía (INEGI), 2021). Éstos son instancias que concentran servicios interinstitucionales y especializados de atención hacia mujeres en situación de violencia al igual que sus hijas e hijos.

Por lo que en este estudio se pretende identificar las entidades federativas del Centrosur y Noreste del país que tienen mayor atención a mujeres en situación de violencia y qué tipo de violencia fue más registrado, por medio de un Análisis de Correspondencia Simple (CA).

Descripción de los datos

La base de datos utilizada para este análisis fue tomada de la página de INEGI en la sección de datos abiertos-microdatos. Es una recopilación de información de los Centros de Justicia para las Mujeres (CJM) del país. Donde proporciona información sobre la infraestructura, recursos económicos, capital humano, instituciones o dependencias representadas en los Centros, espacios físicos, áreas de atención, registro de información y estadísticas de mujeres atendidas por situación de violencia. La información cubre tres periodos: del 1 de enero al 31 de diciembre de 2019, del 1 de enero al 31 de diciembre de 2020 y del 1 de enero al 30 de junio de 2021.

Se trabajó con los estados del país de la zona Centrosur (Ciudad de México, México y Morelos) y Noreste (Coahuila, Nuevo León y Tamaulipas), ya que son los estados con mayor número de registro y así alcanzar con el objetivo del estudio. Cabe resaltar que para el caso de Tamaulipas no había registro de ningún CJM.



Se trabajó con la información de incidencias de violencia en mujeres atendidas de los estados anteriormente mencionados donde hay Centros de Justicia para las Mujeres, sólo del periodo más actual que es del 1 de enero al 30 de junio de 2021.

Los tipos de violencias registrados son:

- **Violencia emocional.**
- **Violencia Física.**
- **Violencia Sexual.**
- **Violencia Económica.**
- **Violencia Patrimonial.**

Para el análisis de esta base se hizo uso del software estadístico R.

Exploración de la matriz

Matriz de datos

```
library(readr)
CJM <- read.csv("CJM.csv")
```

Dimensión

```
dim(CJM)
```

```
## [1] 161072      2
```

La matriz tiene 161072 filas y 2 columnas (estados y tipos de violencia).

Nombre de las variables

```
colnames(CJM)
```

```
## [1] "Estados"      "TipoViolencia"
```

Tipo de variables

```
str(CJM)
```

```
## 'data.frame':  161072 obs. of  2 variables:
## $ Estados      : chr  "Ciudad de Mexico" "Ciudad de Mexico" "Ciudad de Mexico" "Ciudad de Mexico" .
## $ TipoViolencia: chr  "Emocional" "Emocional" "Emocional" "Emocional" ...
```

Las dos variables son de tipo carácter, por lo que la base de datos cumple con los requisitos para poder aplicar un Análisis de Correspondencia Simple.

Presencia de NA'S

```
anyNA(CJM)
```

```
## [1] FALSE
```

No hay datos faltantes, por lo que no hay ningún problema para seguir con el análisis.

Resumen de la matriz

```
summary(CJM)
```

```
##      Estados      TipoViolencia
## Length:161072    Length:161072
## Class :character Class :character
## Mode  :character Mode  :character
```

Tratamiento de la matriz

```
library(gplots)
library(FactoMineR)
library(factoextra)
library(graphics)
```

Tabla de contingencia

Se convierte la matriz a una tabla de contingencia, para obtener la frecuencia de cada uno de los tipos de violencia por estado.

```
#1. Convierte la matriz como una tabla
base=table( CJM$Estados, CJM$TipoViolencia)
base
```

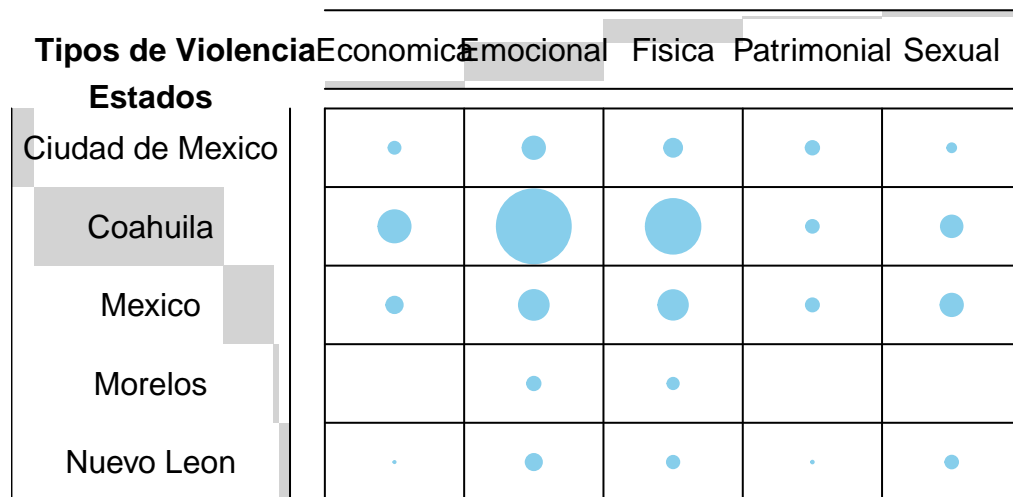
```
##
##      Economica Emocional Fisica Patrimonial Sexual
## Ciudad de Mexico      1526      5401      3459      1940      836
## Coahuila              11143      59159      32235      1744      5007
## Mexico                2883      9565      9370      1830      5389
## Morelos                0        1923      1367         0         0
## Nuevo Leon             47        2811      1636         74      1727
```

Gráfica de la tabla de contingencia

- Se presenta el gráfico de la tabla de contingencia:

```
# 2. Graph
balloonplot(t(base), main = "Incidencias de violencia atendidas en el CJM
del Centrosur y Noreste", xlab = "Tipos de Violencia",
ylab = "Estados",
label = FALSE, show.margins = FALSE)
```

Incidencias de violencia atendidas en el CJM del Centrosur y Noreste



Como se puede observar el estado más destacado es **Coahuila**, esto quiere decir que hubo mayor atención en los Centros de Justicia para las Mujeres con tipo de violencia emocional, seguido de la violencia física, económica, sexual y patrimonial, respectivamente.

La segunda entidad federativa con mayor frecuencia es **México** donde el tipo de violencia atendido mayormente fue violencia emocional, seguido de física y sexual, principalmente.

Por último la **capital de México** está en tercer lugar con mayor atención en el CJM, donde el tipo de violencia destacada es la emocional, seguida de la física y patrimonial.

Prueba Chi-cuadrado

Se utiliza la prueba de chi-cuadrado para evaluar si existe una dependencia significativa entre las categorías de filas y columnas, basándose en la prueba de hipótesis:

- H_0 : No hay dependencia entre filas y columnas.
- H_1 : Sí hay dependencia entre filas y columnas.

```
# 3. Prueba Chi-cuadrado
```

```
chisq.test(base)
```

```
##  
## Pearson's Chi-squared test  
##  
## data: base  
## X-squared = 19269, df = 16, p-value < 2.2e-16
```

Se obtuvo un valor bastante alto de **X-squared = 19269** y un **p-value < 0.05**, por lo tanto se rechaza H_0 y se puede que con un nivel de significancia del 0.05 sí hay asociación entre la variable “Estados” y “TipoViolencia”.

Metodología de análisis

Descripción del método utilizado

Análisis de Correspondencia (CA) Un Análisis de Correspondencia “es una técnica que nos permite representar las categorías de dos o más variables cualitativas en un espacio de pequeñas dimensiones (...) la finalidad es reducir la información original sin que ello implique perder mucha información” (Rodríguez-Jaume & Mora Catalá, 2001, p. 43).

Este análisis proporciona una solución para resumir y visualizar conjuntos de datos en gráficos de dos dimensiones. Para observar gráficamente la asociación entre elementos de fila y columna en la tabla de contingencia. Esto con el objetivo de identificar las variables más importantes que explican las variaciones en un conjunto de datos.

Resultados

- Como ya se evaluó la prueba chi-cuadrada, se procede a construir el Análisis de Correspondencia Simple.

```
AC<-CA(base, graph =FALSE)
```

```
AC
```

```
## **Results of the Correspondence Analysis (CA)**
## The row variable has 5 categories; the column variable has 5 categories
## The chi square of independence between the two variables is equal to 19269.29 (p-value = 0 ).
## *The results are available in the following objects:
##
##      name                description
## 1  "$eig"                "eigenvalues"
## 2  "$col"                "results for the columns"
## 3  "$col$coord"          "coord. for the columns"
## 4  "$col$cos2"           "cos2 for the columns"
## 5  "$col$contrib"         "contributions of the columns"
## 6  "$row"                "results for the rows"
## 7  "$row$coord"          "coord. for the rows"
## 8  "$row$cos2"           "cos2 for the rows"
## 9  "$row$contrib"         "contributions of the rows"
## 10 "$call"               "summary called parameters"
## 11 "$call$marge.col"     "weights of the columns"
## 12 "$call$marge.row"     "weights of the rows"
```

- Se visualizan los valores propios del análisis.

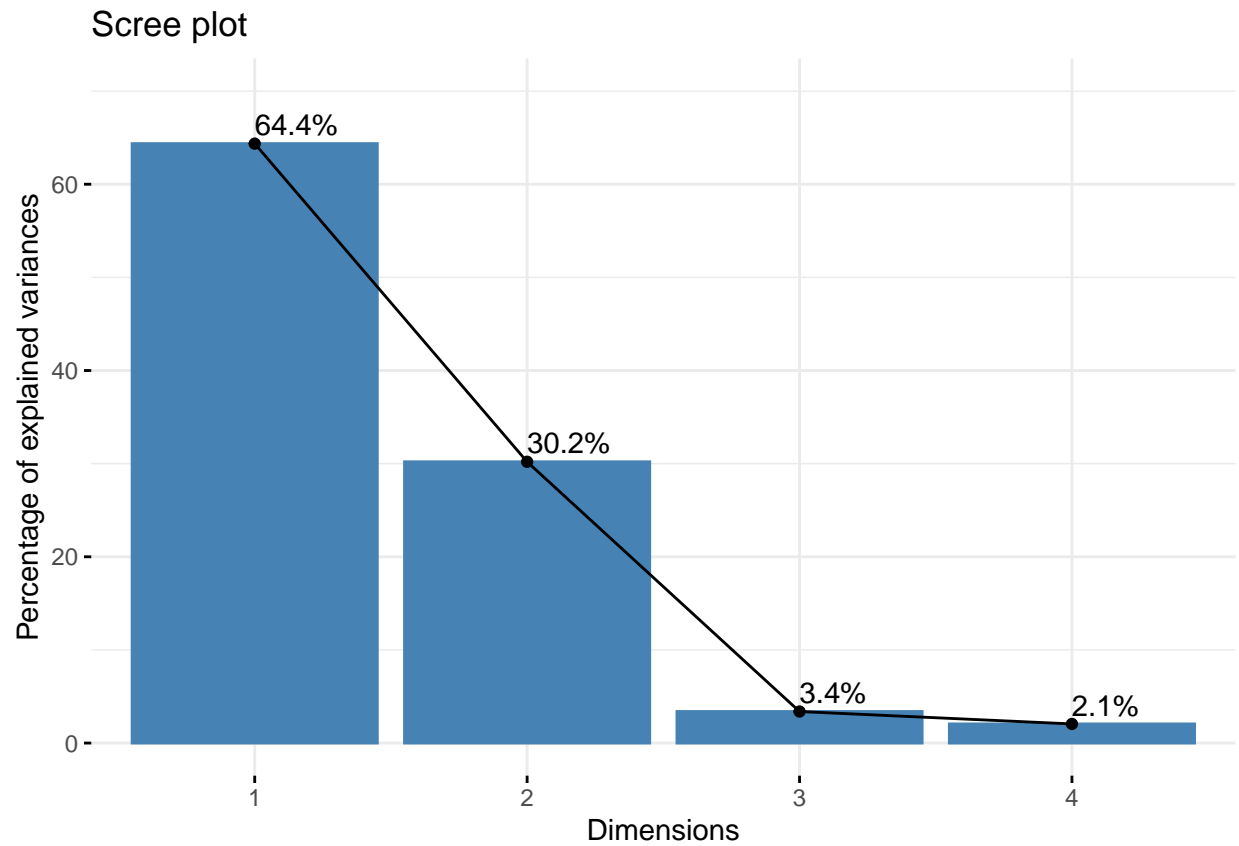
```
eig.val <- get_eigenvalue(AC)
```

```
eig.val
```

```
##      eigenvalue variance.percent cumulative.variance.percent
## Dim.1 0.076995185      64.360290      64.36029
## Dim.2 0.036128893      30.200149      94.56044
## Dim.3 0.004050501       3.385815      97.94625
## Dim.4 0.002456928       2.053746     100.00000
```

Para este caso la primera dimensión representa el 64.3% de los datos y con la segunda se tiene una varianza acumulada del 94.5% de los datos.

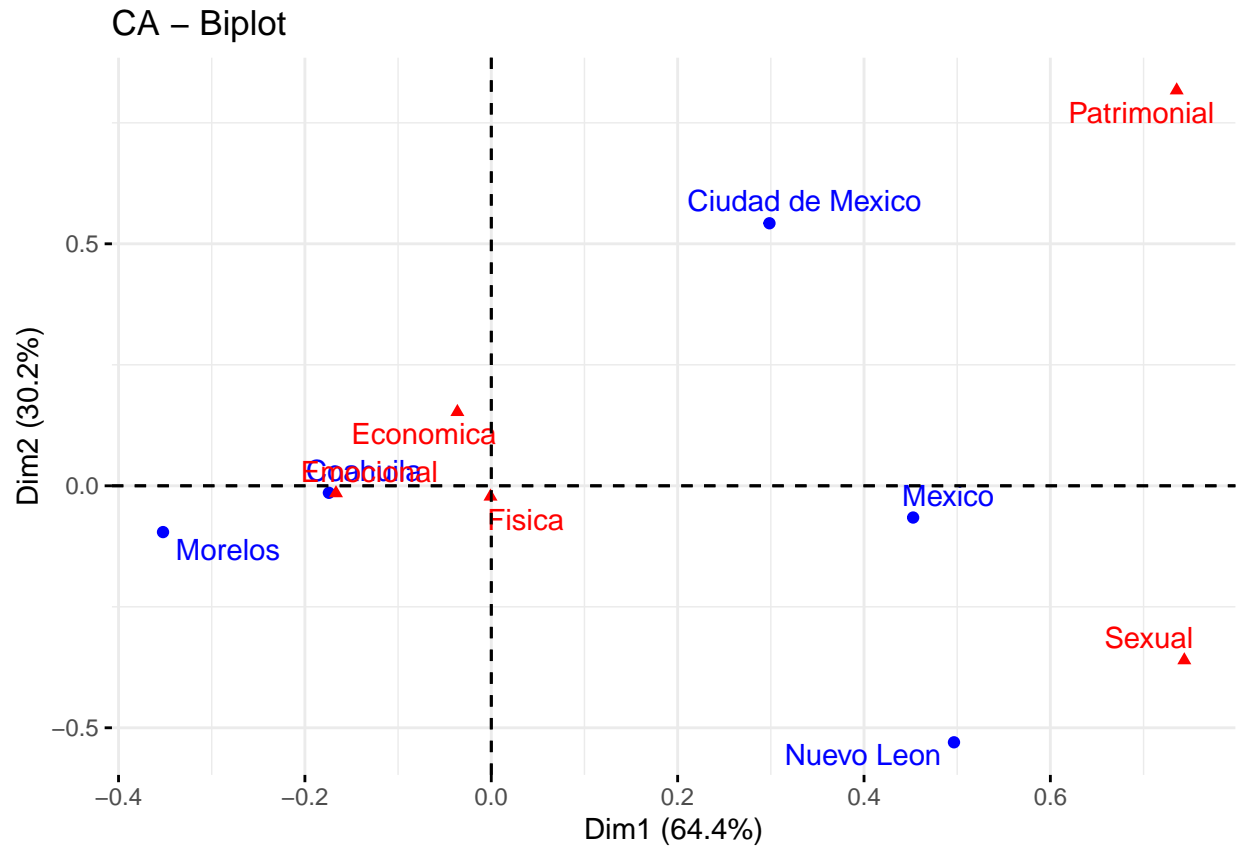
```
fviz_eig(AC, addlabels = TRUE, ylim = c(0,70))
```



Aquí se puede visualizar el “codo” que se forma, pero como se dijo anteriormente es recomendable tomar en cuenta las dos primera dimensiones por tener un valor alto y acumular una varianza del 94.6% de los datos.

Gráfico CA-Biplot

```
fviz_ca_biplot(AC, repel = TRUE)
```



En el gráfico se puede observar el agrupamiento de los estados con respecto a la ocurrencia del tipo de violencia registrados. Por ejemplo en México y Nuevo León hubo mayor atención en cuanto a violencia sexual. En el caso de Coahuila predominó la violencia emocional al igual que Morelos y para el caso de la Ciudad de México se identifica con el tipo de violencia patrimonial.

Conclusiones

Como ya se dijo, el estado con mayor atención a mujeres en situación de violencia del Noreste del país es Coahuila, siendo en mayor frecuencia el tipo de violencia emocional. Para el caso de la zona Centrosur fue el estado de México que tuvo mayor registro de atención a mujeres en situación de violencia mayormente de tipo emocional.

Cabe resaltar que en estas dos zonas el tipo de violencia más común del que fueron atendidas mujeres fue emocional en segundo lugar la violencia física, económica, sexual y patrimonial, respectivamente .

Referencias

- Enlace de acceso para consultar la base de datos
<https://www.inegi.org.mx/programas/cjm/2021/?ps=microdatos>
- Imagen tomada de la web: Regiones de México - Wikipedia, la enciclopedia libre
- INEGI, I. (2022). “Centro de Justicia para las Mujeres en México 2021” (abril de 2022).
- Rodríguez-Jaume, M. J., & Mora Catalá, R. (2001). Análisis de correspondencia.
- EL paquete de R que se utilizó para el análisis fue ***Correspondence Analysis (CA)***.