

TFM

David Fernández

2024-12-21

I. INTRODUCCIÓN

La ciberdelincuencia o delitos informáticos, es un tema que se encuentra a la orden del día, siendo, con el paso del tiempo, cada vez mayor su presencia. Es tal la importancia de este fenómeno, que llega a causar la desesperación en la población debido a los innumerables delitos cometidos en este sector, lo cual implica un constante estado de alerta en la sociedad.

Respecto a este fenómeno, existen innumerables definiciones al respecto siendo la de mayor interés la que nos ofrece el Convenio de Budapest en su preámbulo, el cual lo define como “todo acto dirigido contra la confidencialidad, la integridad y la disponibilidad de los sistemas informáticos, de las redes y de los datos, así como el uso fraudulento de tales sistemas, redes y datos”. Además, define sistema informático en su artículo 1, apartado a), como “todo dispositivo aislado o conjunto de dispositivos interconectados o relacionados entre sí, cuya función, o la de alguno de sus elementos, sea el tratamiento automatizado de datos en ejecución de un programa.”

El Convenio de Budapest consiste en un convenio sobre la ciberdelincuencia, el cual fue creado en Budapest, concretamente el 23 de noviembre de 2001, de ahí su denominación. Se realizó debido a la imperiosa necesidad de abordar este tema, para aplicar una política penal común. A este tratado se han adherido numerosos países, entre ellos España, sobre el que tratará este análisis.

Miró Llinares lo define como el delito cometido en el ámbito de intercomunicación social que es el ciberespacio. Este ciberespacio, es considerado un gran factor de riesgo debido a que en él no existe una vigilancia que permita amedrentar al delincuente de cometer un hecho, como ocurre en un espacio físico, lo cual se traduce en un aumento drástico de oportunidades criminales y, por ende, de posibilidades de contacto entre el agresor y la víctima (Pérez, 2013).

Según Cerezo y García (2020), este tipo de comportamientos delictivos se realizan con 3 objetivos claros: los dispositivos informáticos, los datos y la información informatizada. Además, comentan que existe una evidente falta de concienciación de los usuarios lo cual facilita enormemente la comisión de estos delitos, recalcando también, que, para la comisión de estos, tampoco es necesario disponer de muchos conocimientos, es decir, no hace falta ser un experto en la materia.

El perfil del ciberdelincuente, suele atribuirse al individuo que posee un dominio del medio informático, sin que ello lleve aparejado estar relacionado con una clase social o un segmento de la población en específico. Según datos ofrecidos por Digiware, proveedor experto en la generación de estrategias integrales en seguridad de la información de Latinoamérica, aproximadamente el 76% de los ciberdelincuentes son hombres, cuyas edades oscilan entre los 14 años (8%) hasta los 50 (11%), situándose la edad media en 35 años (43%). El perfil del ciberdelincuente suele ser el de una persona con un coeficiente intelectual medio, que le permita tener conocimientos informáticos, además de poseer un medio informático con el cual tenga acceso a la red para cometer los actos ilícitos (Sánchez, 2019).

En lo que respecta al perfil de la víctima, no existe un perfil claro. Sánchez (2019), recoge en su estudio que según Pratt et. al, lo relevante no son tanto los datos demográficos como el actuar de las víctimas, es decir, es la propia actividad de las víctimas en la red la que determinará los riesgos asociados a su conducta. La edad es muy variable al igual que su poder adquisitivo (Sánchez, 2019).

Los ciberdelitos podemos clasificarlos en distintos grupos penales que a su vez se subdividen en distintas tipologías penales. Observemos un pequeño esquema sacado del propio Portal de Criminalidad donde se muestra esto: (ver Figura 1).

DENOMINACIÓN	CÓDIGO PENAL ESPAÑOL	TIPO HECHO SEC	VARIABLES SEC A UTILIZAR
Acceso e interceptación ilícita	Art. CP 197 A 201. Descubrimiento y revelación de secretos Art. CP 278 a 286. Delitos relativos al mercado y los consumidores (espionaje industrial)	DESCUBRIMIENTO/REVELACIÓN DE SECRETOS	Medio Empleado: Internet/informática, Telefonía/comunicaciones, Intranet y otras redes, páginas de streaming, redes de archivos compartidos P2P, páginas de descargas directas, páginas de enlaces, blogs y correos electrónicos, redes sociales.
		ACCESO ILEGAL INFORMÁTICO	Ninguna
		OTROS RELATIVOS AL MERCADO/CONSUMIDORES	Medio Empleado: Internet/informática, Telefonía/comunicaciones, Intranet y otras redes, páginas de streaming, redes de archivos compartidos P2P, páginas de descargas directas, páginas de enlaces, blogs y correos electrónicos, redes sociales.
Interferencia en los datos y en el sistema	Arts. 263 a 267 y 625.1. Daños y daños informáticos	DAÑOS	Medio Empleado: Internet/informática, Telefonía/comunicaciones, Intranet y otras redes, páginas de streaming, redes de archivos compartidos P2P, páginas de descargas directas, páginas de enlaces, blogs y correos electrónicos, redes sociales.
		ATAQUES INFORMÁTICOS	Ninguna
Falsificación informática	Arts CP 388-389, 399 bis, 400 y 401	FALSIFICACIÓN DE MONEDA, SELLOS Y EFECTOS TIMBRADOS FABRICACIÓN TENENCIA DE ÚTILES PARA FALSIFICAR USURPACIÓN DEL ESTADO CIVIL	Medio Empleado: Internet/informática, Telefonía/comunicaciones, Intranet y otras redes, páginas de streaming, redes de archivos compartidos P2P, páginas de descargas directas, páginas de enlaces, blogs y correos electrónicos, redes sociales.
Fraude Informático	Arts. CP 248 a 251 y 623.4	ESTAFAS BANCARIAS ESTAFAS CON TARJETAS DE CRÉDITO, DÉBITO Y CHEQUES DE DÉBITO OTRAS ESTAFAS	Medio Empleado: Internet/informática, Telefonía/comunicaciones, Intranet y otras redes, páginas de streaming, redes de archivos compartidos P2P, páginas de descargas directas, páginas de enlaces, blogs y correos electrónicos, redes sociales.
Delitos sexuales	Arts. CP 181, 183.1, 183.bis, 184, 185, 186, 189	EXHIBICIONISMO	Medio Empleado: Internet/informática, Telefonía/comunicaciones, Intranet y otras redes, páginas de streaming, redes de archivos compartidos P2P, páginas de descargas directas, páginas de enlaces, blogs y correos electrónicos, redes sociales.
		PROVOCACIÓN SEXUAL	
		ACOSO SEXUAL	
		ABUSO SEXUAL	
		CORRUPCIÓN DE MENORES/INCAPACITADOS	
		PORNOGRAFÍA DE MENORES	
		DELITO DE CONTACTO MEDIANTE TECNOLOGÍA CON MENOR DE 13 AÑOS CON FINES SEXUALES	Ninguna
Contra la propiedad Industrial/Intelectual	Arts 270 a 277 y 623.5 del CP (Contra la propiedad intelectual y contra la propiedad industrial)	DELITOS CONTRA LA PROPIEDAD INTELECTUAL	Medio Empleado: Internet/informática, Telefonía/comunicaciones, Intranet y otras redes, páginas de streaming, redes de archivos compartidos P2P, páginas de descargas directas, páginas de enlaces, blogs y correos electrónicos, redes sociales.
		DELITOS CONTRA LA PROPIEDAD INDUSTRIAL	Medio Empleado: Internet/informática, Telefonía/comunicaciones, Intranet y otras redes, páginas de streaming, redes de archivos compartidos P2P, páginas de descargas directas, páginas de enlaces, blogs y correos electrónicos, redes sociales.
Contra el honor	Arts. 205 a 210 y 620.2 del Código Penal	CALLUMNIAS	Medio Empleado: Internet/informática, Telefonía/comunicaciones, Intranet y otras redes, páginas de streaming, redes de archivos compartidos P2P, páginas de descargas directas, páginas de enlaces, blogs y correos electrónicos, redes sociales.
		INJURIAS	Medio Empleado: Internet/informática, Telefonía/comunicaciones, Intranet y otras redes, páginas de streaming, redes de archivos compartidos P2P, páginas de descargas directas, páginas de enlaces, blogs y correos electrónicos, redes sociales.
Amenazas y coacciones	Arts 169 a 172 y 620 del C.Penal	AMENAZAS	Medio Empleado: Internet/informática, Telefonía/comunicaciones, Intranet y otras redes, páginas de streaming, redes de archivos compartidos P2P, páginas de descargas directas, páginas de enlaces, blogs y correos electrónicos, redes sociales.
		AMENAZAS A GRUPO ÉTNICO CULTURAL O RELIGIOSO COACCIONES	Medio Empleado: Internet/informática, Telefonía/comunicaciones, Intranet y otras redes, páginas de streaming, redes de archivos compartidos P2P, páginas de descargas directas, páginas de enlaces, blogs y correos electrónicos, redes sociales.

Figure 1: Esquema de tipo de ciberdelitos. Tomado del Portal Estadístico de Criminalidad

En nuestro Código Penal (CP) podemos apreciar distintos apartados donde se van haciendo referencia expresa a las distintas tipologías. Vamos a comentarlas en orden de aparición.

En primer lugar, nos encontramos con el artículo 197 CP, el cual nos habla sobre el apoderamiento y modificación de datos en perjuicio de un tercero, descubrimiento de secretos o vulneración de la intimidad sin consentimiento, etc. En todo caso estamos hablando de datos personales conseguidos por la intromisión en un sistema informático y en su posible divulgación o manipulación. Además, en este artículo encontramos con diferentes apartados. El primero de ellos, el artículo en sí, nos habla de que sin la autorización del dueño de dicho sistema se facilite el acceso a otra persona o se mantenga en él contra su voluntad, y en su segundo apartado, el 'bis', sobre quien utilice un instrumento técnico para interceptar las transmisiones no públicas de datos. En el artículo 197 ter ya nos habla de la facilitación de un programa informático a un tercero, cuyo programa este adaptado para cometer dichos delitos o una contraseña o código de acceso a sistemas de información. Los apartados quarter y quinquies no nos dicen nada respecto al tema abordado.

Posteriormente, nos encontramos con el artículo 183 CP, el cual versa sobre el contacto con menores de 16 años mediante cualquier tecnología de la información para cometer cualquier delito de los artículos 181 y 189 CP, además de embaucar a facilitar material pornográfico tanto suyas como en las que aparezca un menor (artículo 189 CP). En su apartado 'bis' nos habla de la inexistencia del delito con consentimiento del menor y edad próxima del autor a la víctima, salvo que se de alguna circunstancia del artículo 178 CP.

El siguiente artículo que tenemos es el 264 CP y su apartado 'bis'. Estos artículos nos hablan sobre el fraude informático que conlleve daños contra el patrimonio mediante la alteración de datos. Además, encontramos el apartado 'ter' el cual nos comenta lo mismo que el 197 ter, anteriormente mencionado.

A continuación, tenemos el artículo 270.2 CP el cual trata de penalizar a quien facilite activamente el acceso no autorizado a contenido protegido por derechos de autor en línea.

Por otra parte, el artículo 510 CP hace referencia a los delitos de odio donde en su apartado 3 nos comenta la realización de estos delitos mediante medios de comunicación social, de internet o tecnologías de la información previendo una pena en su mitad superior para dichos casos.

También tenemos el artículo 575 CP el cual nos habla en su apartado 2 sobre el uso de medios de comunicación o medios accesibles por internet para incitar a que las personas se incorporen a grupos terroristas u organizaciones criminales.

Por último, en el artículo 578 CP el cual versa sobre enaltecer cualquier acto terrorista o autor de ellos o también a la humillación de las víctimas de dichos delitos remarcando en su apartado número 2 que será en su mitad superior la pena si se realiza mediante el uso de medios de comunicación, internet, etc.

Por ende, vemos que no nos encontramos con un único delito que haga referencia a todos estos tipos de delitos, sino que son modalidades previstas en cada precepto penal.

A pesar de que estos tipos delictivos estén actualmente en pleno auge y por tanto se los encuadre principalmente en este último lustro, los ciberdelitos existen desde hace mucho tiempo, en concreto el primer ciberdelito ocurrido fue en el año 1824 el cual consistió en un hackeo, realizado por 2 personas, del sistema telegráfico francés. (Jiménez, 2022).

Durante todo este tiempo los ciberdelitos realizados no eran tal y como los conocemos hoy en día. No fue hasta la década de los 80 cuando se comenzó a ver algo similar, es decir, un hacker que copia, manipula y sustrae información de otro sistema informático. (Jiménez, 2022).

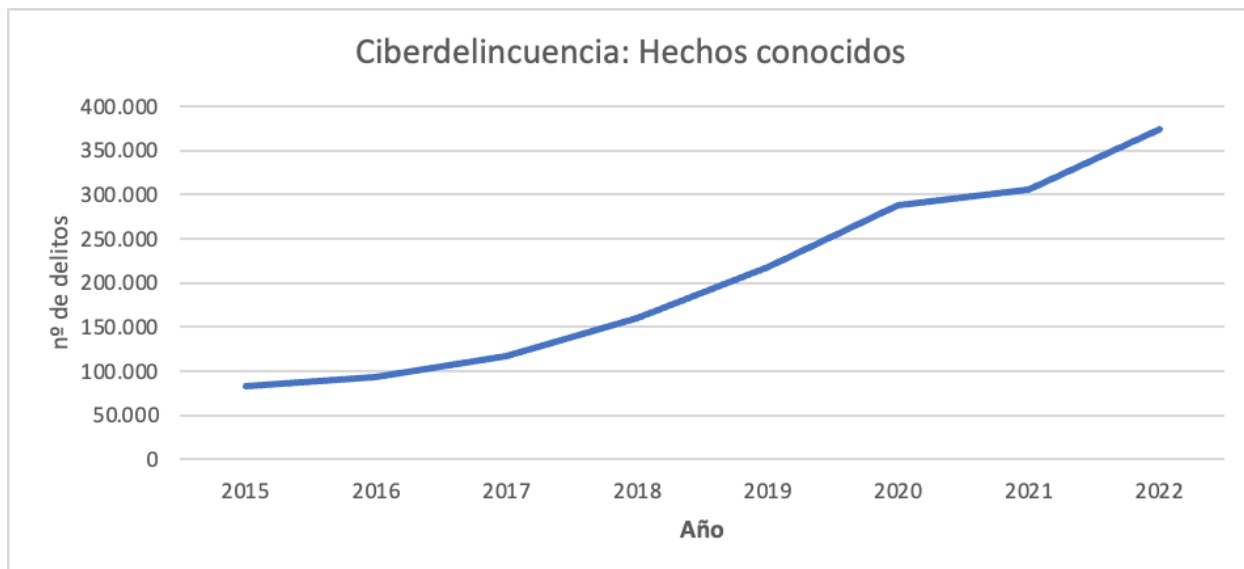
Desde los inicios de la revolución tecnológica, iniciada por la creación de los teléfonos y explotada por la creación del internet, este tipo de delitos han ido aumentando progresivamente hasta el día de hoy. Se dice que a raíz de la pandemia acontecida estos años atrás, los ciberdelincuentes salieron más a la luz y aprovecharon esa oportunidad de encontrarse todo el mundo encerrado en sus casas, lo cual retroalimenta el uso masivo de las TIC's y, por ende, facilita la comisión de estos delitos. Por tanto, es interesante conocer si realmente existe esa diferencia que nos permita afirmar que el año del COVID fue un gran impulso para la comisión de ciberdelitos y, además, en caso de que esto realmente haya ocurrido, si tras finalizar la pandemia, los números de los ciberdelitos siguen manteniéndose en esos picos; si siguen aumentando; o si, por el contrario, han comenzado a disminuir al volver a la "vida normal".

Cabe destacar la gran cifra negra existente en estos tipos de delitos. Al nombrar a la cifra negra nos referimos a todos esos actos delictivos que no quedan registrados en las estadísticas oficiales debido a que no son denunciados por diversas razones como puede ser la desconfianza hacia las autoridades o el desconocimiento propio de las personas de ser víctimas, entre otros.

Metafóricamente nos referimos a ella como un iceberg, donde la punta de este, es decir, lo que sobresale del agua, sería las estadísticas con las que contamos, mientras que el restante de ese enorme bloque de hielo sumergido bajo el agua, sería todos esos delitos que no tenemos constancia de que realmente han sucedido.

Si tenemos en cuenta los datos del Portal Estadístico de Criminalidad, gráficamente podemos observar un incremento paulatino de los ciberdelitos hasta el año 2020, año en el cual, a pesar de aumentar, lo hace en menor medida de lo que venía haciéndolo, para después, en el año siguiente, aumentar de forma más drástica. (Ver figura 2).

Hemos tomado como año de inicio el 2015, ya que es a partir de ese momento cuando se comienzan a incluir en las estadísticas oficiales los datos proporcionados por los Mossos d'Esquadra y de la Ertzaintza.



Por tanto, teniendo en cuenta la información que proporciona el SEC (Servicio Estadístico de Criminalidad) se pone de manifiesto que la cibercriminalidad en España no sólo ha aumentado en este período en términos absolutos, sino que el porcentaje que supone con respecto al total de delitos cometidos en nuestro país también está en aumento (ver Tabla 1).

```
library(knitr)
Tabla1 = data.frame(
  Año = "%",
  `2015` = "4,1%",
  `2016` = "4,60%",
  `2017` = "5,70%",
  `2018` = "7,50%",
  `2019` = "9,90%",
  `2020` = "16,30%",
  `2021` = "16,60%",
  `2022` = "16,16%"
)
colnames(Tabla1) <- c("**Año**", "**2015**", "**2016**", "**2017**", "**2018**", "**2019**", "**2020**", "**2021**", "**2022**")
kable(Tabla1, caption = "***Porcentaje que representan los ciberdelitos con respecto al total de delitos***")
```

Table 1: *Porcentaje que representan los ciberdelitos con respecto al total de delitos en España.* Datos tomados del SEC (elaboración propia)

Año	2015	2016	2017	2018	2019	2020	2021	2022
%	4,1%	4,60%	5,70%	7,50%	9,90%	16,30%	16,60%	16,16%

No obstante, el incremento no ha sido proporcional en todos los grupos penales (ver figura 2). Es el fraude informático el que mayor incremento ha tenido desde 2015.

En la figura 3, se presentan de forma gráfica los mismos datos, pero eliminando la información sobre el fraude informático, y se aprecia cómo son los ciberdelitos de falsificación informática y los delitos sexuales los que más aumentan, sobre todo en la segunda mitad del periodo estudiado. También lo hacen los de amenazas y coacciones, si bien descienden en el último año de estudio.

En el caso particular de Castilla la Mancha, vemos cómo la tendencia se asemeja a la nacional. En la figura

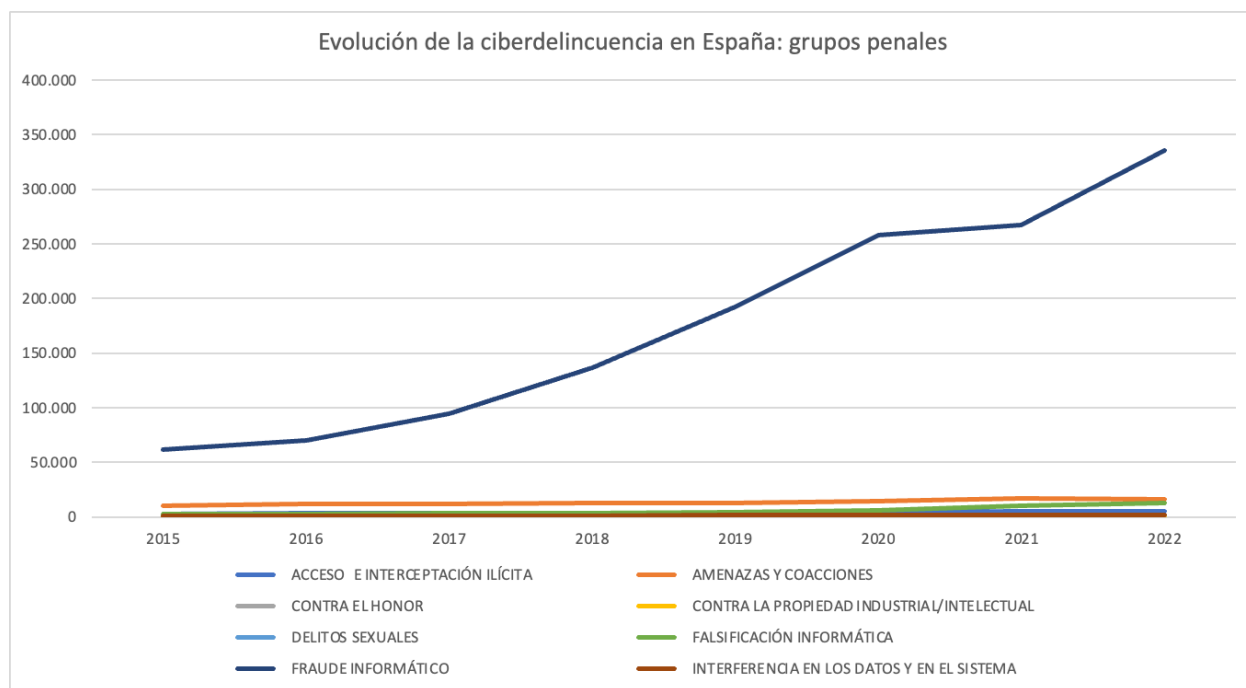


Figure 2: Gráfico de los hechos conocidos de ciberdelincuencia en España durante los últimos 8 años. Datos tomados del SEC (elaboración propia)

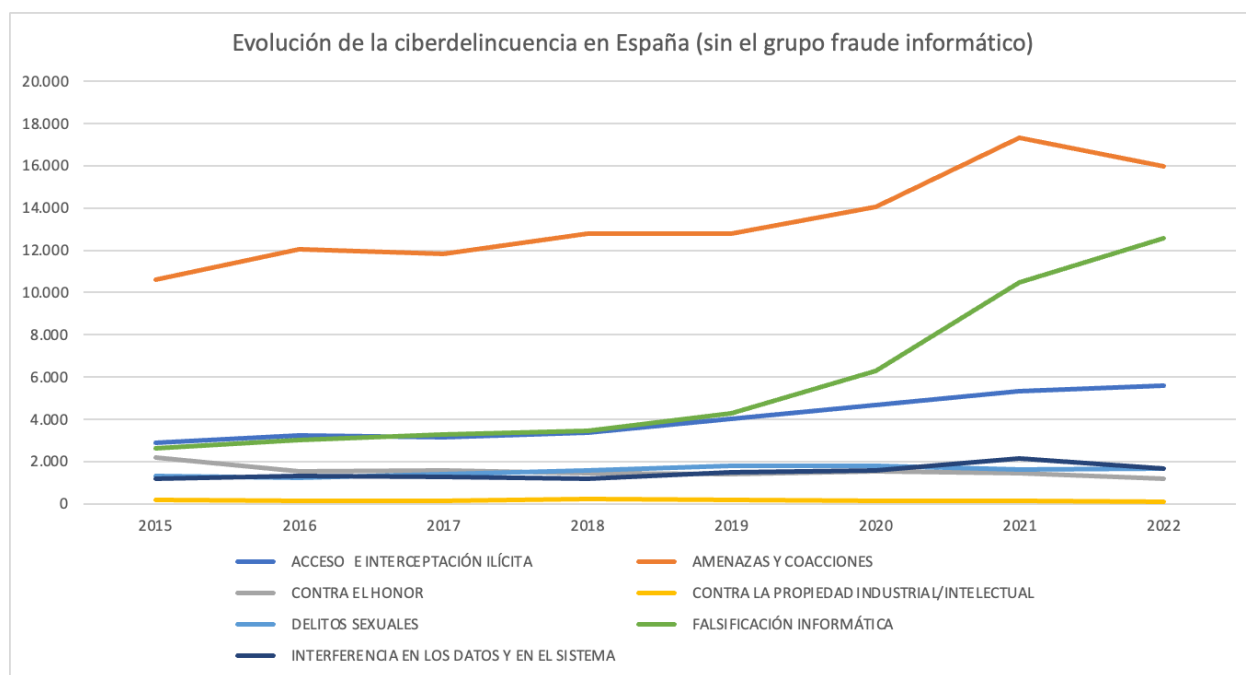


Figure 3: Gráfico de los hechos conocidos de ciberdelincuencia en España durante los últimos 8 años (sin fraude informático). Datos tomados del SEC (elaboración propia)

2, se destaca el enorme incremento del grupo delictivo de fraude informático también en esta comunidad, sin tener una fase de meseta entre 2020-21 (como ocurría a nivel nacional). (Figura 4)

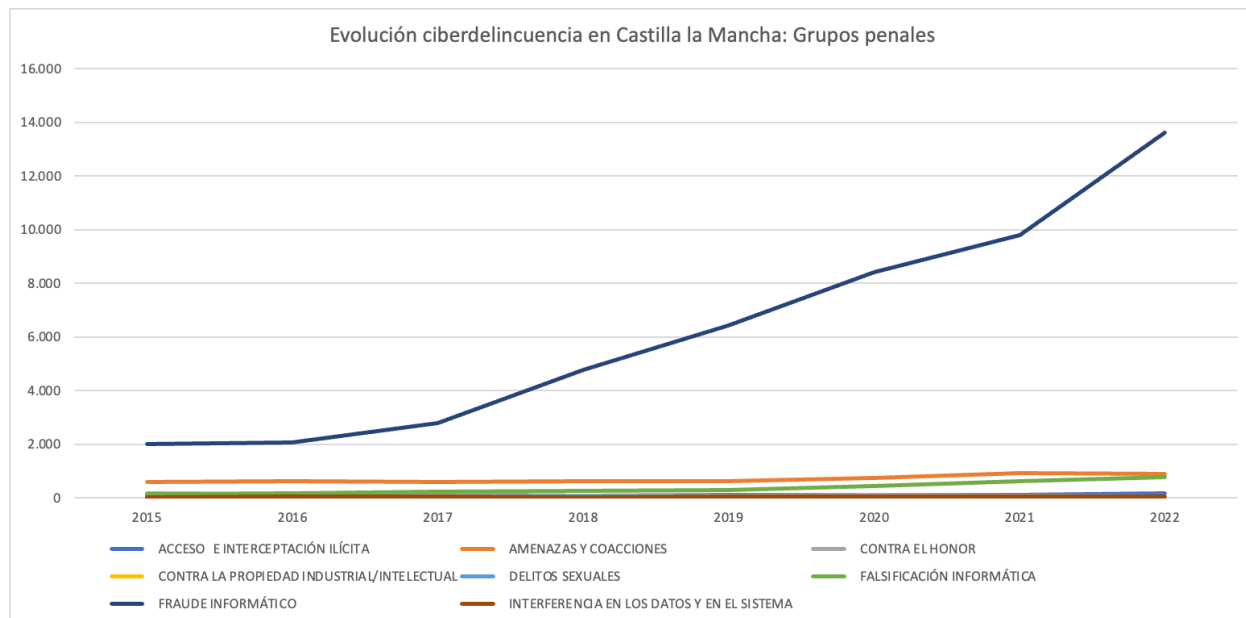


Figure 4: Gráfico de los hechos conocidos de ciberdelincuencia en CLM durante los últimos 8 años. Datos tomados del SEC (elaboración propia)

Y si presentamos el gráfico eliminando los datos relativos al fraude informático, apreciamos la misma tendencia que a nivel nacional, sobre todo en los ciberdelitos de amenazas y coacciones y ciberdelitos sexuales. (Figura 6)

Dado el gran desconocimiento hacia este sector, en este trabajo analizamos qué ha ocurrido en España durante el período referido, 2015-2022, diferenciando entre Comunidades Autónomas y particularizando en el posicionamiento de la comunidad autónoma de Castilla-La Mancha (CLM) con respecto al resto.

Para ello analizaremos los datos tomados del Portal Estadístico de Criminalidad, relativos a hechos conocidos y hechos esclarecidos de infracciones de cada grupo penal en cada comunidad autónoma. Se trabajará con los datos de los años 2015 a 2022 debido a que es a partir de 2015 cuando empieza a tener en cuenta los datos de la Ertzaintza y de los Mossos d'Esquadra, como hemos dicho con anterioridad.

II. OBJETIVOS

En este trabajo tenemos dos objetivos principales:

1. **Analizar la evolución de la ciberdelincuencia en España en el periodo 2015-2022**, y establecer en qué punto se encuentra CLM con relación al resto de CC.AA.
 - 1.1. Estudiar la evolución de la tasa de hechos conocidos de los diferentes grupos penales en las diferentes CC.AA.
 - 1.2. Estudiar la evolución del índice de esclarecimiento de los diferentes grupos penales en las diferentes CC.AA.
2. **Analizar la situación actual de las diferentes provincias de CLM**, en cuanto a su tipología ciberdelictiva.

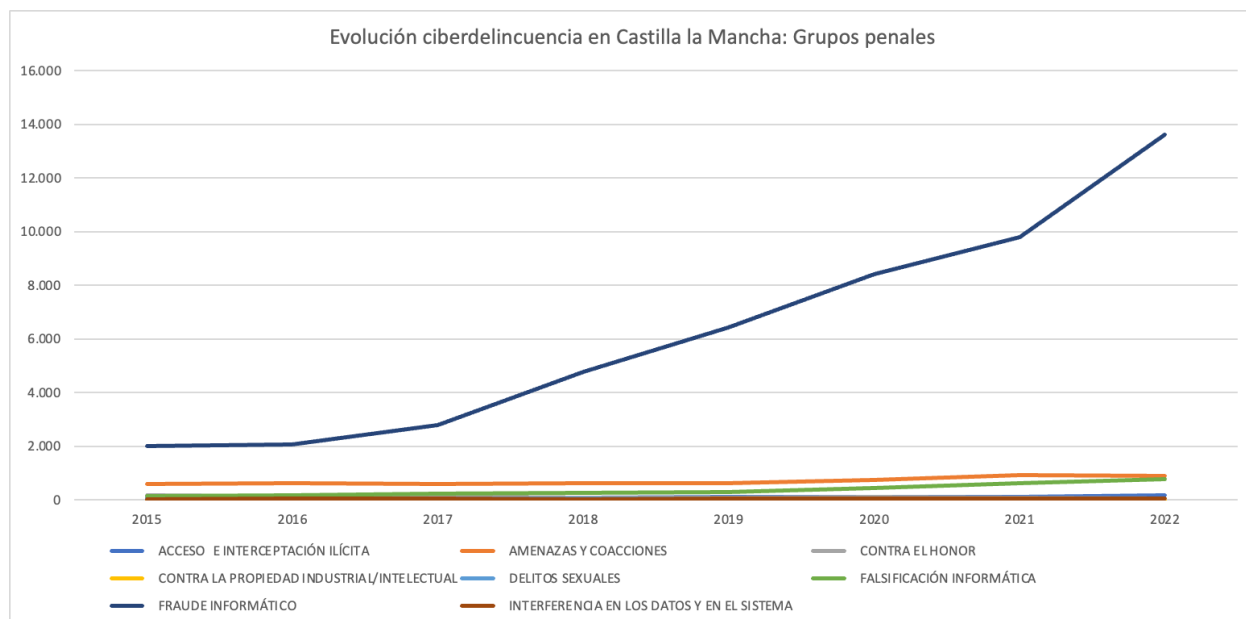


Figure 5: Gráfico de los hechos conocidos de ciberdelincuencia en CLM durante los últimos 8 años (sin fraude informático). Datos tomados del SEC (elaboración propia)

III. MATERIAL Y MÉTODOS

A. MATERIAL

Para dar respuesta a los objetivos planteados se trabajará con datos publicados en el Portal Estadístico de Criminalidad, por el Sistema Estadístico de Criminalidad (SEC) (figura 6) dependiente del Ministerio del Interior (MIR), que recoge periódicamente información sobre. datos de infracciones penales registrados por las fuerzas y cuerpos de seguridad del Estado (Guardia Civil, Cuerpo Nacional de Policía, Policía Foral de Navarra y distintos cuerpos de Policía Local, y a partir de 2015 datos proporcionados también por los Mosos d'Esquadra y la Ertzaintza). Se creó en 2007 para que los ciudadanos tuviesen conocimiento de las actividades delictivas en nuestro país.

En este trabajo utilizaremos los datos sobre ciberdelincuencia, que comprenderán los años 2015 hasta 2022 (por la razón que comentamos en la introducción y debido a que, a fecha 25 de mayo de 2024, 2022 es el último año en el que se han publicado datos relacionados con dichos delitos).



Figure 6: Cabecera del Portal Estadístico de Criminalidad/Cibercriminalidad.

En primer lugar, con el fin de dar respuesta al primero de los objetivos se estudiará la evolución temporal de

los hechos conocidos (ver tabla 1) por grupos de ciberdelitos en las diferentes comunidades autónomas, para poder estudiar la estabilidad o no, a lo largo del periodo, de las estructuras de covariación entre los grupos penales, y las similitudes de las comunidades, y así poder establecer en qué punto se encuentra CLM con relación al resto de CC.AA. A continuación, analizaremos también los hechos esclarecidos de estos grupos penales.

De este modo los datos se estructurarán en una tabla de tres vías. Cada tabla (k) se corresponderá con los datos de un año de la serie (tendremos por tanto, $K = 8$ tablas, cada una correspondiente a un año), las filas (individuos) de cada tabla serán las distintas comunidades autónomas de España ($I = 19$; 17 comunidades más las dos ciudades autónomas de Ceuta y Melilla) y las columnas (variables) serán los 8 grupos penales siendo estos ($J = 8$): Acceso e Interpretación Ilícita (AII) (X1), Amenazas y Coacciones (AYC) (X2), Contra el Honor (HON) (X3), Contra la Propiedad Industrial/Intelectual (CPI/I) (X4), Delitos Sexuales (SEX) (X5), Falsificación Informática (FALIN) (X6), Fraude Informático (FRAIN) (X7) e Interferencias en los Datos y en el Sistema (IDS) (X8). Por tanto, cada tabla (matriz de datos) tiene dimensiones 19×8 ($X_{19 \times 8}$). (Ver tabla 2). El esquema de la estructura de los datos se representa en la figura 8.

```
Tabla2 = data.frame(
  Variable = c("$X_1$", "$X_2$", "$X_3$", "$X_4$", "$X_5$", "$X_6$", "$X_7$", "$X_8$"),
  Grupo_Penal = c("Acceso e Interpretación Ilícita", "Amenazas y Coacciones", "Contra el Honor", "Contra la Propiedad Industrial/Intelectual", "Delitos Sexuales", "Falsificación Informática", "Fraude Informático", "Interferencias en los Datos y en el Sistema"),
  Código = c("AII", "AYC", "HON", "CPI/I", "SEX", "FALIN", "FRAIN", "IDS")
)
kable(Tabla2, caption = "***Grupos Penales***", align = c("c", "c", "c"))
```

Table 2: *Grupos Penales*

Variable	Grupo_Penal	Código
X_1	Acceso e Interpretación Ilícita	AII
X_2	Amenazas y Coacciones	AYC
X_3	Contra el Honor	HON
X_4	Contra la Propiedad Industrial/Intelectual	CPI/I
X_5	Delitos Sexuales	SEX
X_6	Falsificación Informática	FALIN
X_7	Fraude Informático	FRAIN
X_8	Interferencias en los Datos y en el Sistema	IDS

Para hacer los datos de hechos conocidos comparables entre las diferentes comunidades autónomas, hemos calculado la tasa de hechos conocidos respecto al total de la población de cada CC.AA. (n° de hechos / 100.000 hbs.) debido a que las poblaciones son muy distintas de unas comunidades a otras.

También utilizaremos los datos que nos proporciona el SEC sobre los hechos esclarecidos en cada CCAA, como hemos dicho, para lo cual calculamos el Índice de esclarecimiento que es el cociente entre los hechos esclarecidos y los conocidos multiplicado por 100, que es una medida de la eficacia policial. Hay que señalar que el SEC no incluyen datos de la Ertzaintza en toda la serie histórica, por lo que los resultados de este índice en el país vasco están sesgados en el sentido de que será más alto que el que aparece en los datos.

Para conseguir el segundo objetivo utilizaremos los datos de 2022 sobre hechos conocidos de las tipologías ciberdelictivas (ver tabla 3) de las provincias de CLM. También calcularemos las tasas por provincias. La estructura de los datos es como sigue (ver figura 8).

```
Tabla3 = data.frame(
  Variable = c("$X_1$", "$X_2$", "$X_3$", "$X_4$", "$X_5$", "$X_6$", "$X_7$", "$X_8$", "$X_9$", "$X_{10}$"),
  Grupo_Penal = c("Abuso Sexual", "Acceso Ilegal Informático", "Acoso Sexual", "Amenazas", "Amenazas a la Integridad", "Delitos Sexuales", "Falsificación Informática", "Fraude Informático", "Interferencias en los Datos y en el Sistema", "Robo de Datos"),
  Código = c("ABSEX", "AIINF", "ACSEX", "AME", "AGECTR", "ATQINF", "CALUM", "COAC", "PINDUS", "PINTEL")
)
kable(Tabla3, caption = "***Tipologías Penales***", align = c("c", "c", "c"))
```

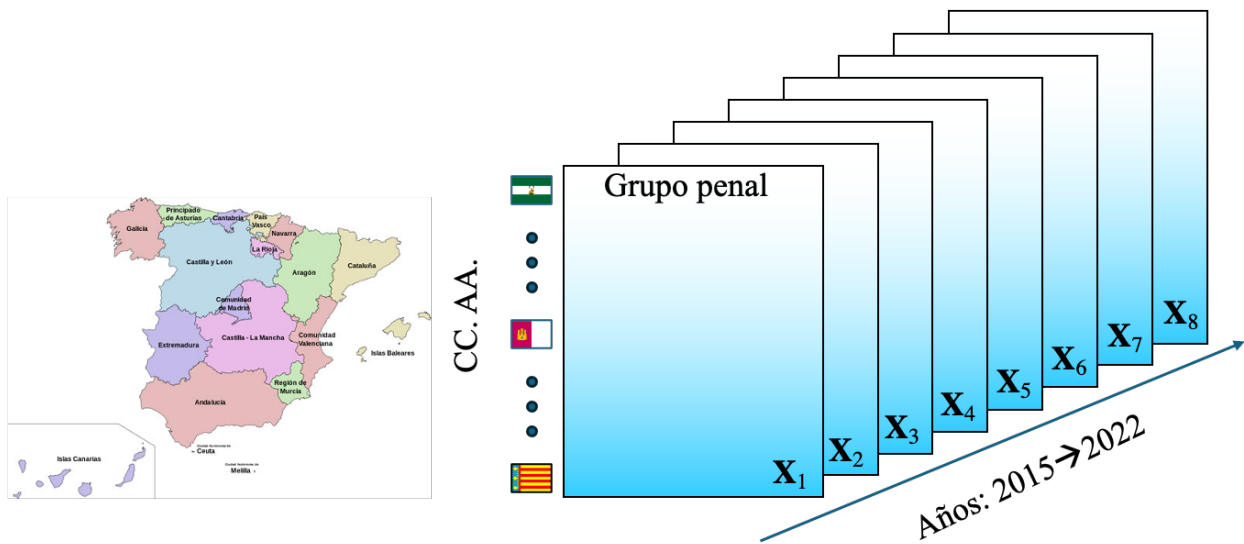



Figure 7: Estructura de los datos en análisis: Hechos conocidos/100.000 habs.



Figure 8: Estructura de los datos en análisis, para la tasa de hechos conocidos en CLM

Table 3: *Tipologías Penales*

Variable	Grupo_Penal	Código
X_1	Abuso Sexual	ABSEX
X_2	Acceso Ilegal Informático	AIINF
X_3	Acoso Sexual	ACSEX
X_4	Amenazas	AME
X_5	Amenazas a Grupos Étnico Cultural o Religioso	AGECTR
X_6	Ataques Informáticos	ATQINF
X_7	Calumnias	CALUM
X_8	Coacciones	COAC
X_9	Contra la Propiedad Industrial	PINDUS
X_{10}	Contra la Propiedad Intelectual	PINTEL
X_{11}	Corrupción de Menores/con Discapacidad/Diversidad Funcional	CMD/D
X_{12}	Daños	DAÑOS
X_{13}	Delito de Contacto MEDiante Tecnologías con Menor de 16 años con Fines Sexuales	TEC16SEX
X_{14}	Descubrimiento/Revelación de Secretos	DES/REVSEC
X_{15}	Estafa Bancaria	ESTBAN
X_{16}	Estafas con Tarjetas de Crédito, Débito y Cheques de Viaje	ESTTARYCHEQ
X_{17}	Estafas Informáticas	ESTINF
X_{18}	Exhibicionismo	EXHI
X_{19}	Fabricación de Moneda, Sellos y Efectos Timbrados	FDSET
X_{20}	Injurias	INJURIAS
X_{21}	Otras Estafas	OTRASEST
X_{22}	Otros Relativos al MERCado/Consumidores	ORM/C
X_{23}	Pornografía de Menores	PORME
X_{24}	Provocación Sexual	PROSEX
X_{25}	Usurpación de Estado Civil	USUESTCIV

Las tipologías delictivas de Amenazas a Grupo Étnico Cultural o Religioso (AGECTR); Contra la Propiedad Industrial (PINDUS); Contra la Propiedad Intelectual (PINTEL); Daños (DAÑOS); Exhibicionismo (EXHI), Fabricación de Moneda, Sellos y Efectos de Timbrados (FDSET); Otros Relativos al Mercado/Consumidores (ORM/C) Y Provocación Sexual (PROSEX), han sido eliminadas del estudio porque tenían pocos hechos conocidos o ninguno, por lo que no nos serían de utilidad y solo podrían perjudicar alterando al resto de resultados.

B. METODOLOGÍA

En este trabajo se utilizarán dos técnicas multivariantes para el análisis de datos: el JK-Biplot y el Análisis Triádico Parcial (PTA) (también llamado X-STATIS). El primero permite analizar los datos estructurados en una matriz de dos vías. En nuestro caso lo utilizaremos para analizar los datos de 2022, correspondientes a los hechos conocidos de tipologías delictivas de CLM. El segundo, que es un análisis de datos de tres vías, lo utilizaremos para analizar los datos de la tasa de hechos conocidos y los datos del índice de esclarecimiento de las CC. AA entre 2015-2022, que como hemos visto en la figura 8 presentan una estructura de cubo de datos.

1. MÉTODOS BIPLLOT

Gabriel definió el Biplot en 1971, como una representación gráfica de una matriz Z de n filas y m columnas mediante marcadores a_1, a_2, \dots, a_n para sus filas y b_1, b_2, \dots, b_m para sus columnas. El prefijo “bi” indica que es una visualización conjunta de filas y columnas, a diferencia de la mayoría de otros gráficos que solo muestran filas o columnas por separado.

Por tanto, tal como refiere Galindo (en 1986), se trata de una representación gráfica de datos multivariantes, que permite representar conjuntamente filas (individuos) y columnas (variables) de una matriz mediante productos internos de vectores. Es útil para describir datos y mostrar resultados de modelos formales.

En un Biplot, las variables se representan con vectores, cuanto mayor es la longitud del mismo mayor es la variabilidad de la variable que representa, y los individuos se representan con puntos en el plano.

Gabriel (1971) propuso dos tipos de Biplots: el GH-Biplot, con alta calidad de representación para las variables y menor para los individuos, y el JK-Biplot, que representa bien los individuos y con menor calidad de representación las variables.

Galindo (1985; 1986) presentó una alternativa a los Biplots clásicos de Gabriel, demostrando que es posible representar filas y columnas con la misma calidad en un solo sistema de coordenadas, y denominó a esta técnica HJ-Biplot.

La fundamentación teórica de los Biplot clásicos, se basa en la descomposición de la matriz de datos \mathbf{Z} de orden $(n \text{ individuos} \times D \text{ variables})$, y de rango s , por una de bajo rango $(q < s)$

$$\mathbf{Z} = \mathbf{A}\mathbf{B}',$$

donde \mathbf{A} es una matriz de orden $n \times s$ y \mathbf{B} una matriz de orden $D \times s$.

Las n filas de \mathbf{A} y las D de \mathbf{B} proporcionan las coordenadas de los n individuos y de las D variables en un espacio euclídeo s -dimensional (con la misma dimensión que el rango r de \mathbf{Z}).

La descomposición de \mathbf{Z} en el producto de las matrices \mathbf{A} y \mathbf{B}' puede realizarse de infinitas maneras distintas, cumpliéndose sea cual sea la descomposición que se realice, que el producto escalar ordinario del i -ésimo punto fila de \mathbf{A} con el j -ésimo punto columna de \mathbf{B}' con respecto al origen, es igual al (i, j) -ésimo elemento z_{ij} de la matriz \mathbf{Z} .

Habitualmente el interés se centra en encontrar una representación de dimensión 2 ó 3 de \mathbf{Z} y se parte de la descomposición en valores singulares (SVD) de \mathbf{Z} :

$$\mathbf{Z} = \mathbf{A}\mathbf{B}' = \mathbf{U}\mathbf{\Sigma}\mathbf{V}'$$

donde:

$\mathbf{\Sigma}$ es una matriz diagonal de orden $s \times s$ construida a partir de los valores singulares no nulos de \mathbf{Z} ($\sqrt{\lambda_1} > \sqrt{\lambda_2} > \dots > \sqrt{\lambda_s} > 0$).

\mathbf{U} es una matriz de orden $n \times s$ cuyas s columnas ($\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_s$) constituyen un conjunto ortonormal de vectores de \mathbb{R}^n .

\mathbf{V} es una matriz de orden $s \times D$ cuyas s columnas ($\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_s$) constituyen un conjunto ortonormal de vectores de \mathbb{R}^D .

Debe cumplirse que $\mathbf{U}'\mathbf{U} = \mathbf{V}'\mathbf{V} = \mathbf{I}$, es decir, las columnas de \mathbf{U} y \mathbf{V} son ortonormales para garantizar así la unicidad de la factorización.

Para el JK-Biplot, que es el que utilizaremos en esta investigación, se impone la métrica $\mathbf{B}\mathbf{B}' = \mathbf{I}$ en el espacio de las filas de la matriz \mathbf{Z} , luego $\mathbf{Z}'\mathbf{Z} = \mathbf{A}\mathbf{A}'$. Se seleccionan los primeros s valores singulares más grandes $\mathbf{\Sigma}$ y las correspondientes columnas de \mathbf{U} y de \mathbf{V} . Las coordenadas de los individuos en el Biplot se obtienen como $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}$ y las coordenadas de las variables $\mathbf{B} = \mathbf{V}$. Por lo tanto, este Biplot preserva la métrica usual entre las filas, pero no entre las columnas, obteniéndose alta calidad de representación para las filas.

En la interpretación del JK-Biplot, los individuos están representados por puntos. La proximidad de dos puntos indica una similitud de los mismos en las variables medidas. Las variables están representadas por vectores cuya dirección y longitud indican la contribución y la variabilidad de la variable respectiva. La proyección de un punto de fila sobre un vector que representa a una variable indica el valor relativo de esa variable para ese individuo (ver figura 9)

El JK-Biplot ofrece varias ventajas, incluyendo una alta calidad de representación para las filas y la facilidad de visualización de las relaciones entre individuos y variables. Sin embargo, su enfoque en las columnas

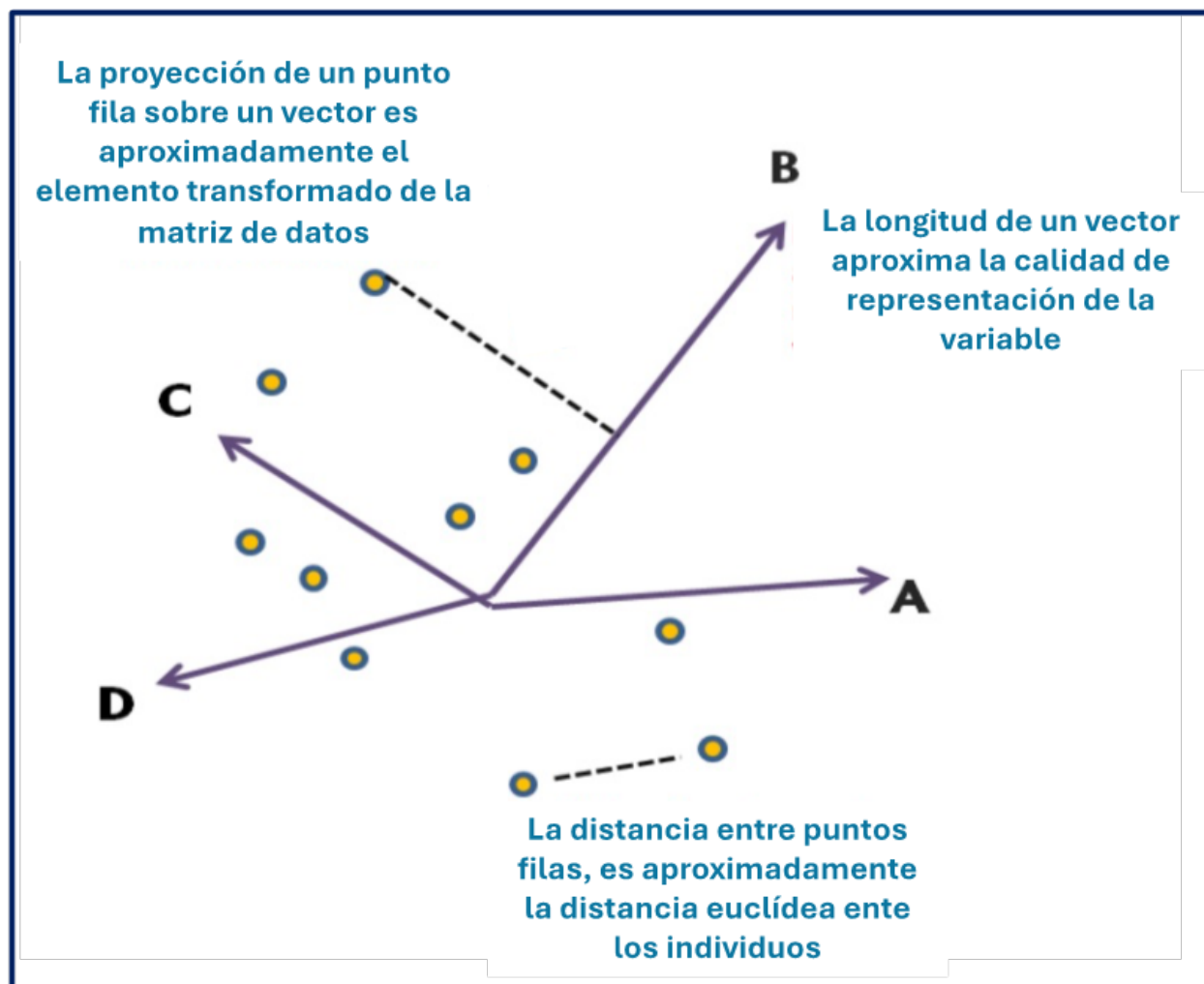


Figure 9: Interpretación gráfica del JK-Biplot

puede llevar a una representación menos precisa de las variables en comparación con otros tipos de Biplots. Por ejemplo, el GH-Biplot prioriza la calidad de representación de las columnas (variables), mientras que el HJ-Biplot equilibra la representación tanto de filas como de columnas, como hemos dicho con anterioridad.

En nuestro estudio las filas serán las provincias de CLM y las columnas las tipologías ciber delictuales.

2. X-STATIS

La otra técnica que se va a utilizar es la del Análisis Triádico Parcial (PTA, por sus siglas en inglés), también denominado X-STATIS. De aquí en adelante nos referiremos a ella como PTA.

El PTA consiste en una herramienta de exploración de datos de 3 vías (en los que tenemos K tablas de dos vías), desarrollada por Jaffrenou (1978), la cual, al inicio de su existencia, se denominó Análisis Triádico gracias a Thioulouse y Chessel (1987), y más tarde Kroonenberg se refirió a ella como Análisis Triádico Parcial. Se trata de una técnica de la familia de las técnicas STATIS (Structuration des Tableaux à Trois Indices de la Statistique) (siendo un caso particular cuando tanto las filas como las columnas son las mismas para todas las K tablas). Tal como puede consultarse en Abdi et al. (2012, p. 125), existen 3 etapas en el método, al igual que cualquiera de las técnicas STATIS (Lavit, 1994): análisis de la la interestructura, del compromiso y de la intraestructura.

Los pasos a seguir serían los siguientes:

Interestructura

En esta fase es en la que analizamos la estructura de los datos y las relaciones entre las distintas tablas. El objetivo es, además, obtener los pesos óptimos que permitan calcular una representación común de las observaciones. La principal herramienta en esta fase es calcular el coeficiente de correlación vectorial (RV) entre tablas, que mide la similitud entre dos matrices de datos evaluando a su vez la relación entre las mismas. El análisis de la matriz de correlaciones vectoriales entre tablas, mediante un análisis de componentes principales, nos permite visualizar gráficamente la relación entre ellas, así como obtener los pesos que cada tabla tendrán para la formación de la “tabla media” (o tabla compromiso), calculando la media ponderada de las tablas en estudio, utilizando como pesos las componentes de cada tabla en el primer vector propio resultado del PCA de la matriz de correlaciones vectoriales.

En nuestro caso, se calculará la matriz RV entre los diferentes años. Esta matriz RV tendrá dimensiones 8x8.

Compromiso

Utilizando los pesos obtenidos en el paso anterior, se calcula una tabla de compromiso, tabla de representación común que sintetiza la información contenida en las distintas tablas. Esta tabla nos proporciona una visión general de las tendencias y patrones comunes de las K tablas. Esta matriz de datos se analiza mediante unas componentes principales, de forma que se obtiene una representación en baja dimensión (representación del compromiso) en la que se pueden visualizar, e interpretar como en un Biplot, las relaciones entre las “filas medias” y “columnas medias”. En nuestro estudio, la matriz compromiso será una matriz de “provincias x grupos delictivos”.

Intraestructura

En esta fase se proyectan las filas y las columnas de las tablas originales sobre el compromiso, de manera que se pueden estudiar las “trayectorias” de las mismas y su evolución. Por lo tanto, podremos estudiar la evolución de las CC.AA. y de grupos de ciberdelitos cometidos a lo largo del período de estudio.

El PTA, a diferencia de otros métodos como el STATIS, o el STATIS-DUAL, como hemos visto, no trabaja con operadores, lo cual indica que trabajamos directamente con las matrices de datos, y, por ende, con la información original, sin transformaciones adicionales. Aquí trabajamos con los mismos individuos (I) (19 CC. AA) y variables (J) (8 grupos penales) en las K tablas (8 años).

Otra forma de ver este procedimiento, es tal como propone Jaffrenou (1978). Lo que ocurre es que cada matriz X_k traspuesta se “desdobla” como vector y se posiciona en una nueva matriz, consiguiendo así una

nueva matriz bidimensional (Vicente-Galindo, 2013). (Figura 10)

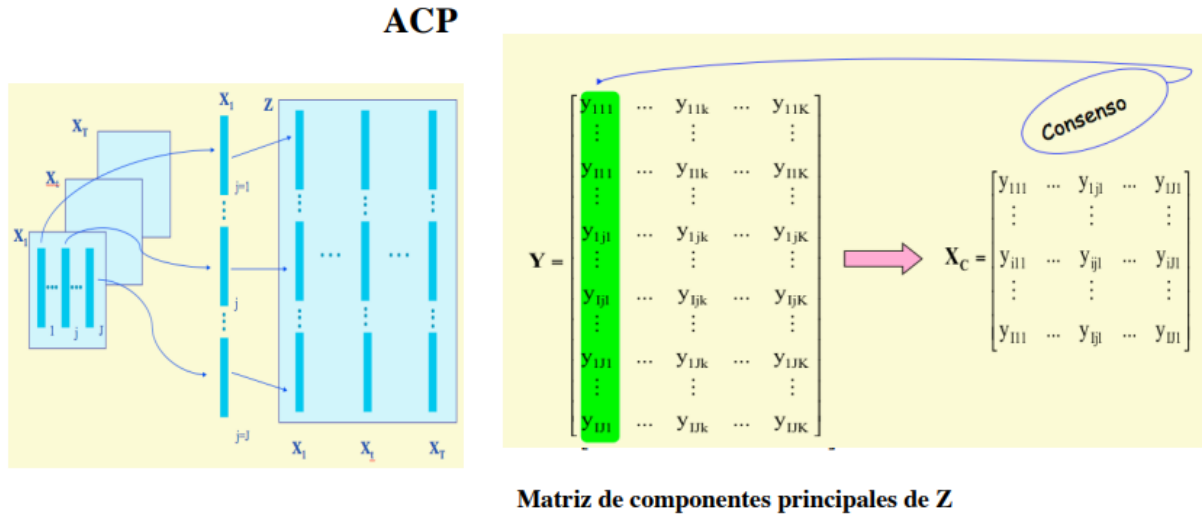


Figure 10: Esquema de transformación de la tabla en el PTA (tomado de Vicente-Galindo, 2013)

Las K matrices se estiran, lo que significa que cada matriz pasa a ser una columna de la nueva tabla. Esta tabla es la empleada al realizar el análisis la cual nos permite comparar estas K tablas de manera adecuada.

Es decir, el PTA es una técnica de análisis de datos que nos permite analizar datos estructurados en distintas tablas permitiéndonos comprender en profundidad sus relaciones y analizar patrones en datos de 3 vías. Su capacidad para trabajar directamente con matrices que tienen los mismos individuos y las mismas variables lo hacen muy útil a la hora de realizar estudios longitudinales y comparativos.

Esta técnica es muy relevante en este tipo de estudios de cibercriminalidad para analizar la tendencia y la estabilidad o no de los patrones delictivos a lo largo del tiempo.

C. SOFTWARES

Para el tratamiento de los datos hemos utilizado el programa denominado Excel, con el cual hemos adecuado los datos para posteriormente introducirlos en el programa que emplearemos para realizar el análisis. Se ha utilizado los programas estadístico R y Rstudio, complementación uno del otro, para realizar los distintos análisis. Los paquetes que han sido utilizados han sido los siguientes:

- readxl, para la lectura de los excels, donde aparecen nuestros datos
- ade4 y ade4TkGUI, para el propio análisis de los de los datos
- adegraphics, que complementa a ade4, para hacer posible la visualización del análisis mediante gráficos
- corrplot, para la visualización gráfica de las correlaciones
- ggplot2, que crea también gráficas, en nuestro caso empleado para la descriptiva del estudio
- gridExtra, nos permite visualizar varios gráficos de forma simultánea
- FactoMineR, que permite realizar técnicas de análisis exploratorio de los datos, con los que a su vez podemos realizar gráficos como los biplots

IV. RESULTADOS

A. Análisis descriptivo de las tasas de hechos conocidos de los diferentes grupos penales

Comenzaremos realizando un análisis estadístico descriptivo de la tasa de hechos conocidos. (Ver summary)

```
library(readxl)
datos=read_excel('Grupos penales.xls')
AIIo <- read_excel("Grupos penales.xls",
                  sheet = "AII")

AYCo <- read_excel("Grupos penales.xls",
                  sheet = "AYC")

SEXo <- read_excel("Grupos penales.xls",
                  sheet = "SEX")

HONo <- read_excel("Grupos penales.xls",
                  sheet = "HON")

CPIo <- read_excel("Grupos penales.xls",
                  sheet = "CPI")

FALINo <- read_excel("Grupos penales.xls",
                  sheet = "FALIN")

FRAINo <- read_excel("Grupos penales.xls",
                  sheet = "FRAIN")

IDSo <- read_excel("Grupos penales.xls",
                  sheet = "IDS")
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.3.3

library(gridExtra)
AII=AIIo[,2:9]
AII_reversed <- AII[, rev(seq_len(ncol(AII)))]
AII_reversed = as.data.frame(AII_reversed)
rownames(AII_reversed)=AIIo$CCAA

AYC=AYCo[,2:9]
AYC_reversed <- AYC[, rev(seq_len(ncol(AYC)))]
AYC_reversed = as.data.frame(AYC_reversed)
rownames(AYC_reversed)=AYCo$CCAA

SEX=SEXo[,2:9]
SEX_reversed <- SEX[, rev(seq_len(ncol(SEX)))]
SEX_reversed = as.data.frame(SEX_reversed)
rownames(SEX_reversed)=SEXo$CCAA

HON=HONo[,2:9]
HON_reversed <- HON[, rev(seq_len(ncol(HON)))]
HON_reversed = as.data.frame(HON_reversed)
```

```

rownames(HON_reversed)=HONo$CCAA

CPI=CPIo[,2:9]
CPI_reversed <- CPI[, rev(seq_len(ncol(CPI)))]
CPI_reversed = as.data.frame(CPI_reversed)
rownames(CPI_reversed)=CPIo$CCAA

FALIN=FALINo[,2:9]
FALIN_reversed <- FALIN[, rev(seq_len(ncol(FALIN)))]
FALIN_reversed = as.data.frame((FALIN_reversed))
rownames(FALIN_reversed)=FALINo$CCAA

FRAIN=FRAINo[,2:9]
FRAIN_reversed <- FRAIN[, rev(seq_len(ncol(FRAIN)))]
FRAIN_reversed = as.data.frame(FRAIN_reversed)
rownames(FRAIN_reversed)=FRAINo$CCAA

IDS=IDSo[,2:9]
IDS_reversed <- IDS[, rev(seq_len(ncol(IDS)))]
IDS_reversed = as.data.frame(IDS_reversed)
rownames(IDS_reversed)=IDSo$CCAA

summary(datos[,3:10])

```

##	AII	AYC	HON	CPI/I
##	Min. : 0.000	Min. : 4.289	Min. : 0.000	Min. : 0.0000
##	1st Qu.: 5.783	1st Qu.: 25.377	1st Qu.: 2.717	1st Qu.: 0.1478
##	Median : 7.734	Median : 31.445	Median : 3.975	Median : 0.2343
##	Mean : 8.927	Mean : 31.314	Mean : 4.024	Mean : 0.3118
##	3rd Qu.: 11.083	3rd Qu.: 37.866	3rd Qu.: 5.015	3rd Qu.: 0.3638
##	Max. : 25.641	Max. : 68.250	Max. : 14.008	Max. : 1.6670
##	SEX	FALIN	FRAIN	IDS
##	Min. : 0.000	Min. : 0.9512	Min. : 33.71	Min. : 0.000
##	1st Qu.: 2.474	1st Qu.: 5.3169	1st Qu.: 137.68	1st Qu.: 2.061
##	Median : 3.093	Median : 9.2413	Median : 315.98	Median : 2.848
##	Mean : 3.215	Mean : 12.1128	Mean : 348.60	Mean : 3.097
##	3rd Qu.: 3.748	3rd Qu.: 16.6844	3rd Qu.: 552.09	3rd Qu.: 3.830
##	Max. : 13.754	Max. : 43.3979	Max. : 928.60	Max. : 11.974

Podemos ver que hay en muchos grupos penales en los que el valor mínimo toma el valor 0, es decir, que no se conoce ningún hecho delictivo en alguna de las comunidades autónomas. Por otra parte, observamos que hay una clara diferencia cuantitativa de los delitos de Fraude informático, respecto al resto de grupos, siendo su tasa mínima es de 33.71.

A continuación, se presenta la información de forma gráfica para poder visualizar mejor la evolución de cada grupo delictivo en cada CC.AA. a lo largo del período 2015-2022, y así observar a simple vista en qué comunidades autónomas se cometen en mayor o menor medida ciertos tipos de delitos y en qué años.

```

barplot(as.matrix(t(AII_reversed)), beside = TRUE,
        main = "Acceso e interpretación ilícita (AII)",
        xlab = "CCAA", ylab = "Nº de hechos conocidos /100.000 hbts.",
        col = rainbow(ncol(AII_reversed)))

```

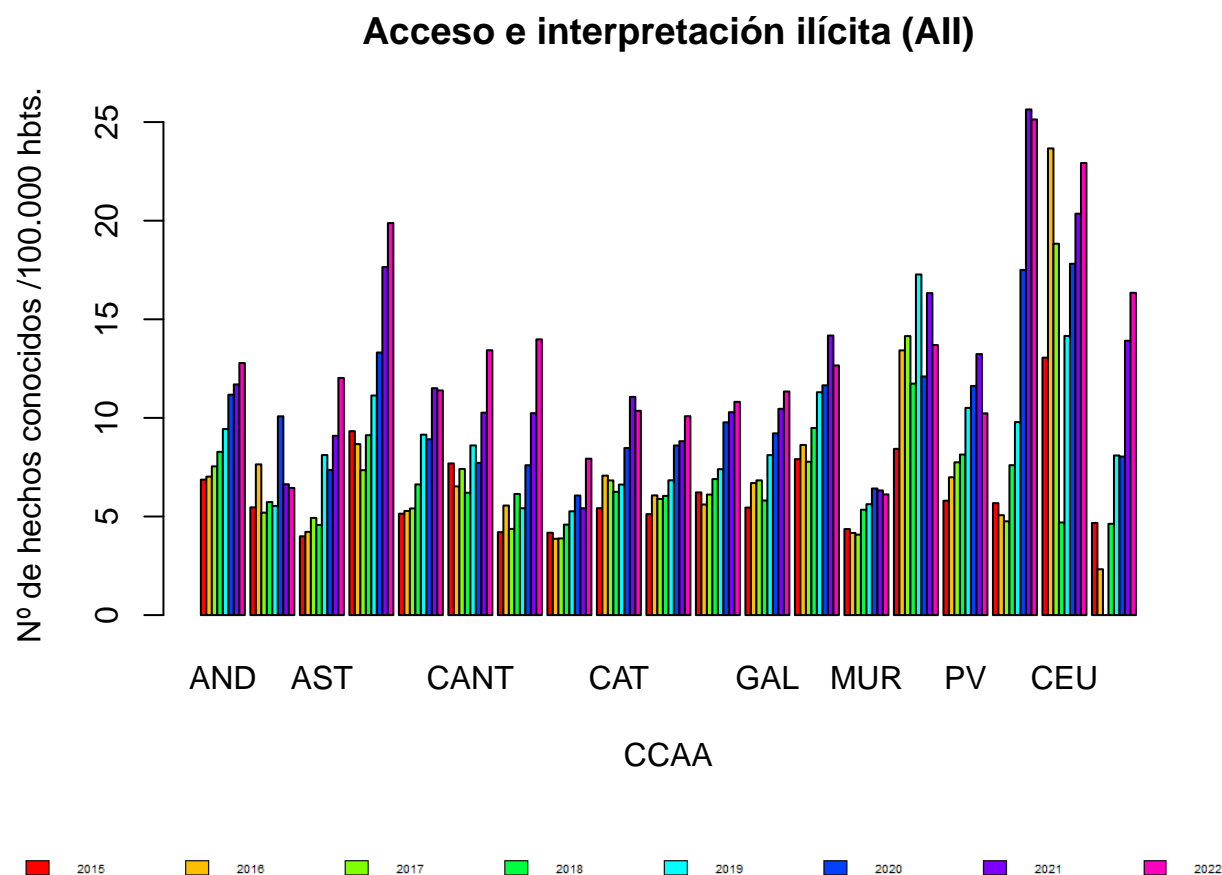



Figure 11: Acceso e Interceptación Ilícita. N° hechos conocidos/100.000 hbts

En el primer grupo penal, Acceso e Interceptación Ilícita (ver Figura 11), podemos observar 2 comunidades que destacan sobre el resto, estas son La Rioja y Ceuta. Lo más llamativo de ambas es que, respecto a la primera, vemos un gran incremento de la comisión de dicho delito en el 2020, seguido de otro en el 2021. Por otra parte, en Ceuta, vemos que este delito ya era el que más se cometía desde el 2015, ocurriendo un gran descenso en 2018, donde se convierte en la comunidad que menos ocurre y volviendo a situarse en la número 1 de nuevo al siguiente año. Por el lado contrario, tenemos a Murcia, donde, en general, es la comunidad que tiene menor nº de hechos conocidos por cada 100.000 habs. Observamos que CLM es una de las comunidades con las tasas más bajas de este grupo de ciberdelito, prácticamente en todos los años.

```

barplot(as.matrix(t(AYC_reversed)), beside = TRUE,
        main = "Amenazas y Coacciones (AYC)",
        xlab = "CCAA", ylab = "Nº de hechos conocidos /100.000 hbts.",
        col = rainbow(ncol(AYC_reversed)))
  
```

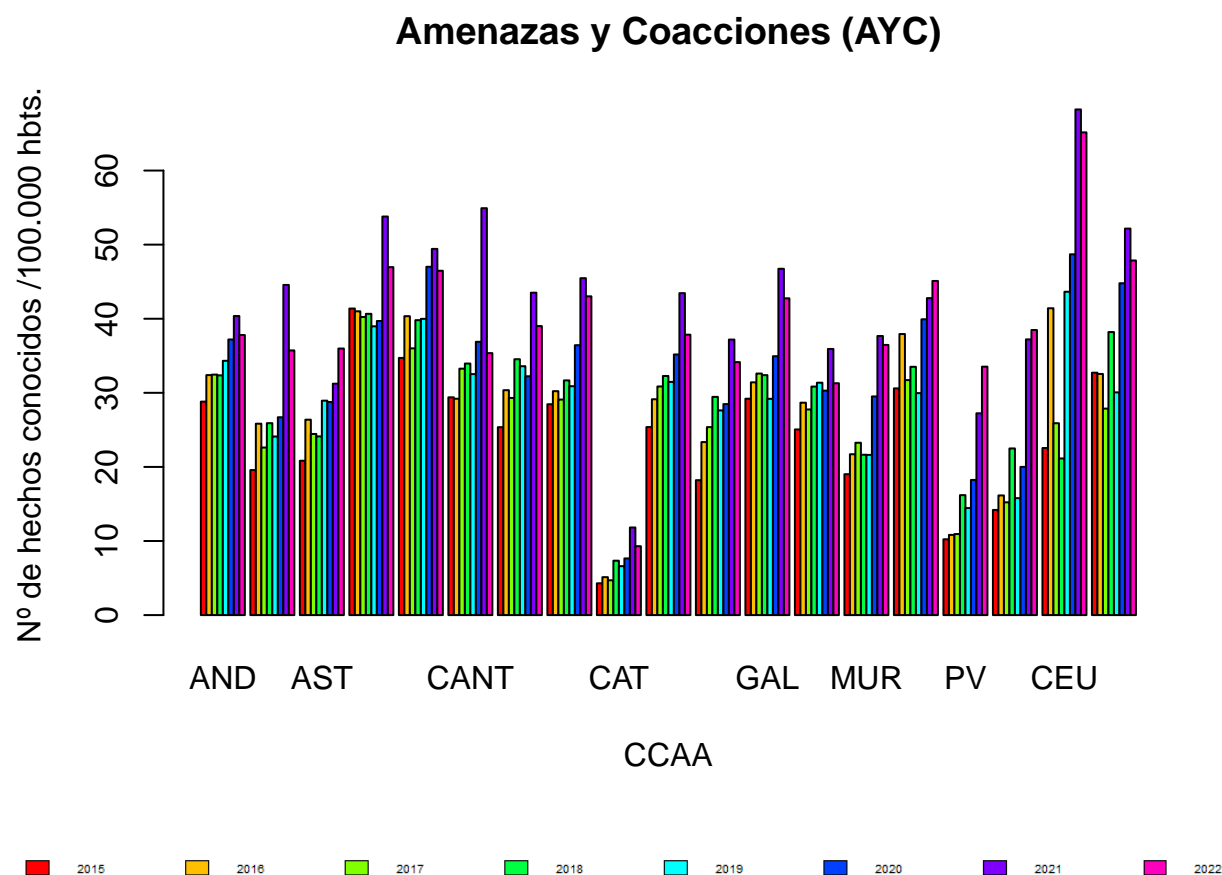


Figure 12: Amenazas y Coacciones. Nº hechos conocidos/100.000 habs.

Respecto al grupo de Amenazas y Coacciones, Cataluña es la comunidad con una frecuencia bastante inferior al resto de comunidades. Del otro lado, tenemos principalmente a Ceuta, que presenta una mayor tasa. (Ver figura 12). En cuanto a este grupo delictivo, CLM sigue una tendencia similar a la mayoría de las CC.AA. aumentando sobre todo en los tres últimos años.

```
barplot(as.matrix(t(HON_reversed)), beside = TRUE,
        main = "Contra el Honor (HON)",
        xlab = "CCAA", ylab = "Nº de hechos conocidos /100.000 hpts.",
        col = rainbow(ncol(HON_reversed)))
```

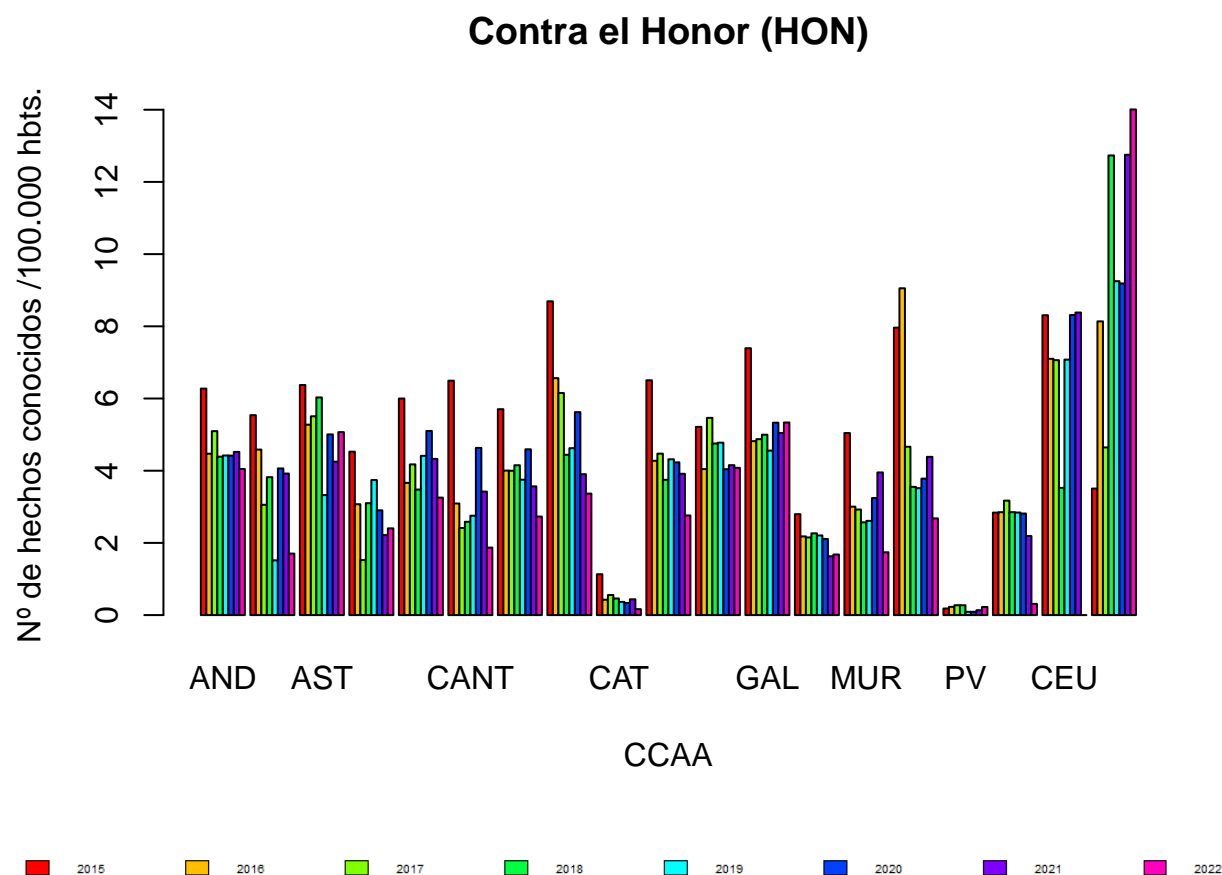


Figure 13: Contra el Honor. Nº hechos conocidos/100.000 hbs.

En los delitos contra el Honor, encontramos varias peculiaridades. La primera de ellas que apenas se cometen dichos delitos en dos comunidades, Cataluña y País Vasco, a quien se suma, respecto al este último años 2022, La Rioja. En cuanto a la mayor tasa, Ceuta y Melilla son las que se encuentran con un mayor número de delitos cometidos de este tipo, fundamentalmente en los últimos años. En Navarra y CLM ha disminuido este tipo de delitos a lo largo del periodo, y vemos como en Madrid este tipo de delitos no son tan frecuentes. (Ver figura 13).

```
barplot(as.matrix(t(SEX_reversed)), beside = TRUE,
        main = "Sexuales (SEX)",
        xlab = "CCAA", ylab = "Nº de hechos conocidos /100.000 hbs.",
        col = rainbow(ncol(SEX_reversed)))
```

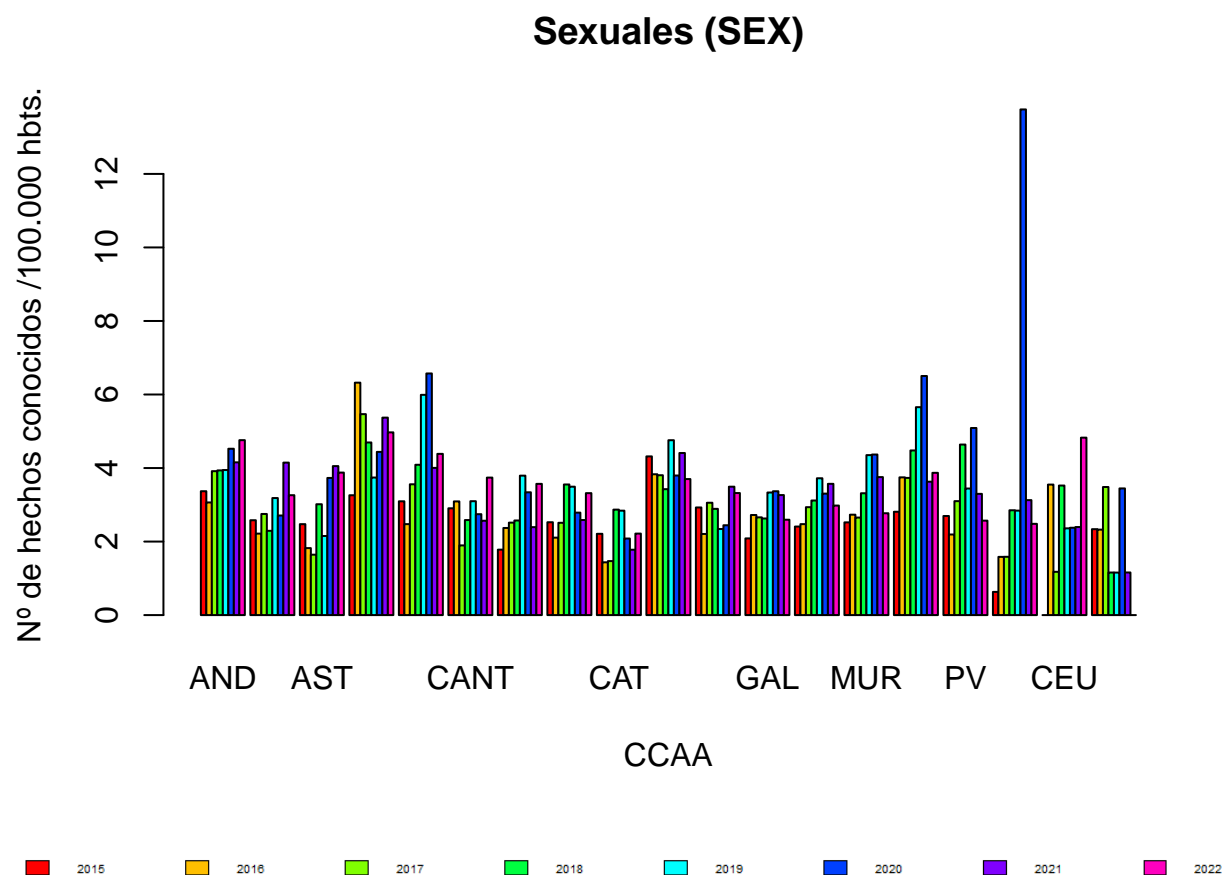


Figure 14: Sexuales. Nº hechos conocidos/100.000 habs.

En cuanto a delitos sexuales, tenemos una tendencia estable en prácticamente todas las comunidades, bien es cierto que cabe destacar un par de datos. En primer lugar, el despunte de La Rioja en el año 2019; en segundo lugar, la nula tasa de Ceuta en el 2015 junto a la escasa de La Rioja en el mismo año; y, en tercer lugar, la nula tasa de Melilla en 2020. (Ver figura 14). CLM no destaca especialmente en cuanto a este tipo de ciberdelitos en relación al esto de CC.AA. ya que sus tasas están en torno a los valores medios.

```
barplot(as.matrix(t(CPI_reversed)), beside = TRUE,
        main = "Contra la Propiedad Industrial/Intelectual (CPI/I)",
        xlab = "CCAA", ylab = "Nº de hechos conocidos /100.000 hbs.",
        col = rainbow(ncol(CPI_reversed)))
```

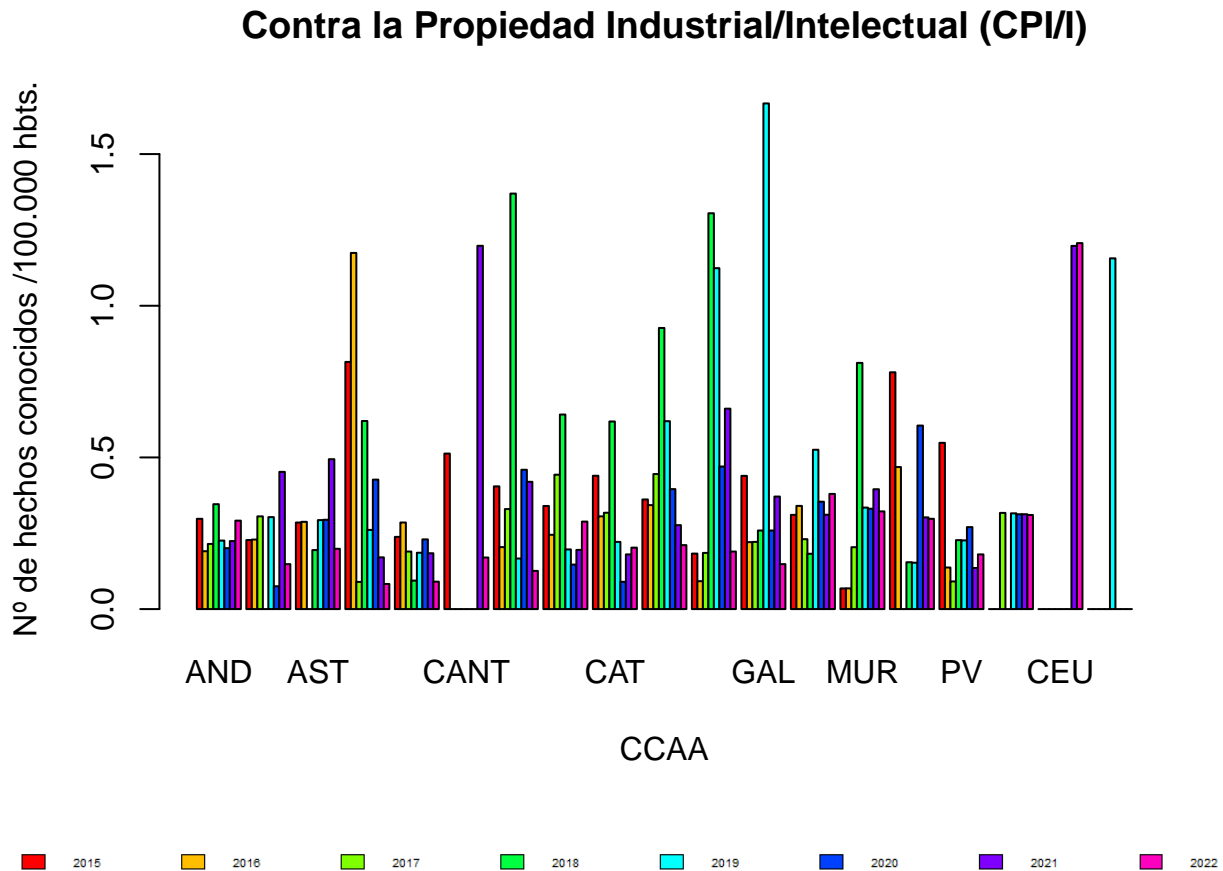


Figure 15: Contra la Propiedad Industrial/Intelectual. Nº hechos conocidos/100.000 habs.

Los delitos Contra la Propiedad Industrial/Intelectual, son los más inestables de todos. Encontramos varios despuntes de varias comunidades en distintos años siendo pocos años los que se mantienen en cifras prácticamente constantes, como sucede en Andalucía. En este tipo de delitos, destacan las comunidades de Cantabria, Ceuta y Melilla. En estas comunidades encontramos valores nulos en muchos de los años y en el resto valores muy altos, lo cual llama la atención. (Ver figura 15). CLM, si bien no destaca del resto de comunidades, al igual que la mayoría de ellos presenta una tasa más alta en los años 2017 y 2018 en este tipo de ciberdelitos.

Si atendemos a los delitos de Falsificación Informática, podemos observar como en todas las comunidades hay una tendencia creciente muy similar, empezando a aumentar más drásticamente la comisión de estos delitos a partir del año 2019. (Ver figura 16 a continuación). Igual ocurre con los delitos de Fraude Informático. (Ver figura 17 a continuación). CLM no se aparta de esa tendencia, y presenta un comportamiento similar al resto de CC.AA. situándose en tasas medias con respecto al resto.

```
barplot(as.matrix(t(FALIN_reversed)), beside = TRUE,
        main = "Falsificación Informática (FALIN)",
        xlab = "CCAA", ylab = "Nº de hechos conocidos /100.000 habs.",
        col = rainbow(ncol(FALIN_reversed)))
```

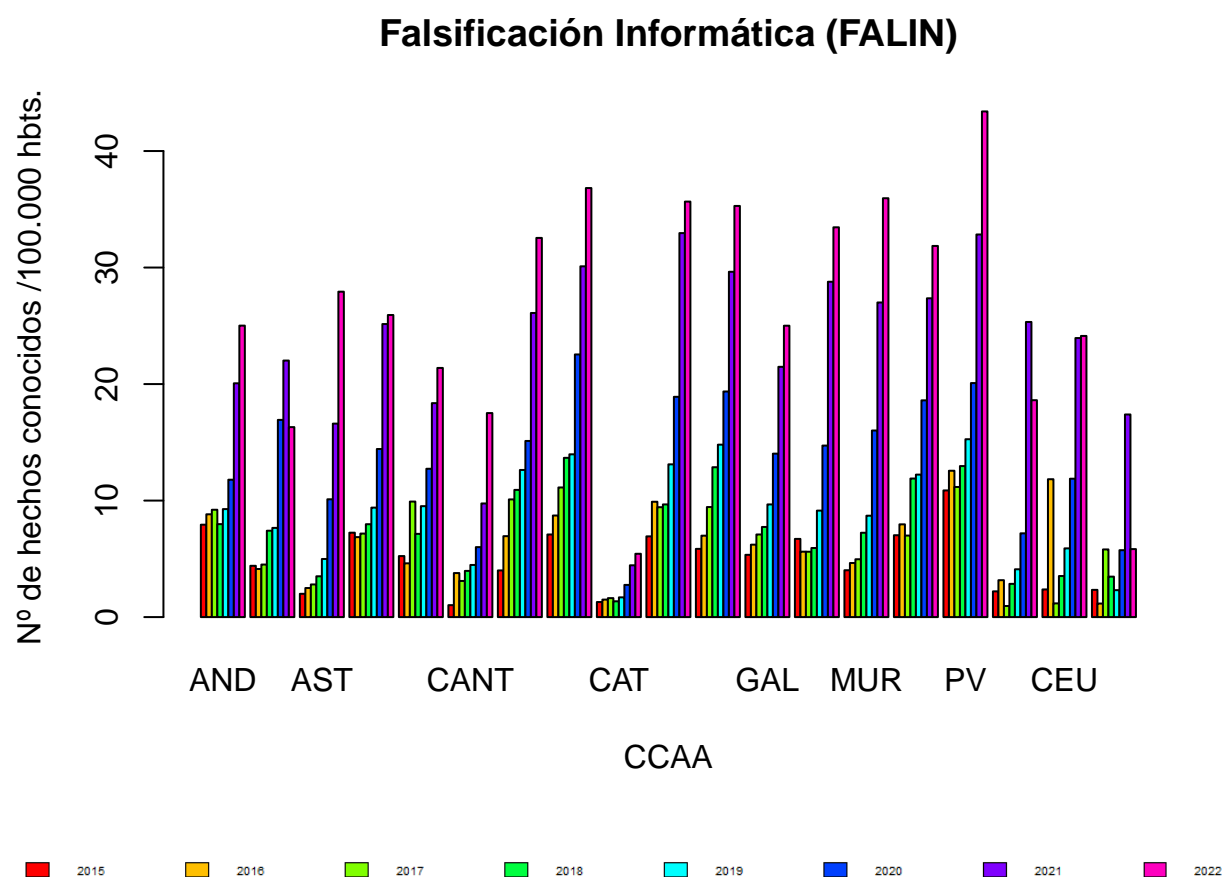


Figure 16: Falsificación Informática. Nº hechos conocidos/100.000 hbs.

```
barplot(as.matrix(t(FRAIN_reversed)), beside = TRUE,
        main = "Fraude Informático (FRAIN)",
        xlab = "CCAA", ylab = "Nº de hechos conocidos /100.000 hbs.",
        col = rainbow(ncol(FRAIN_reversed)))
```

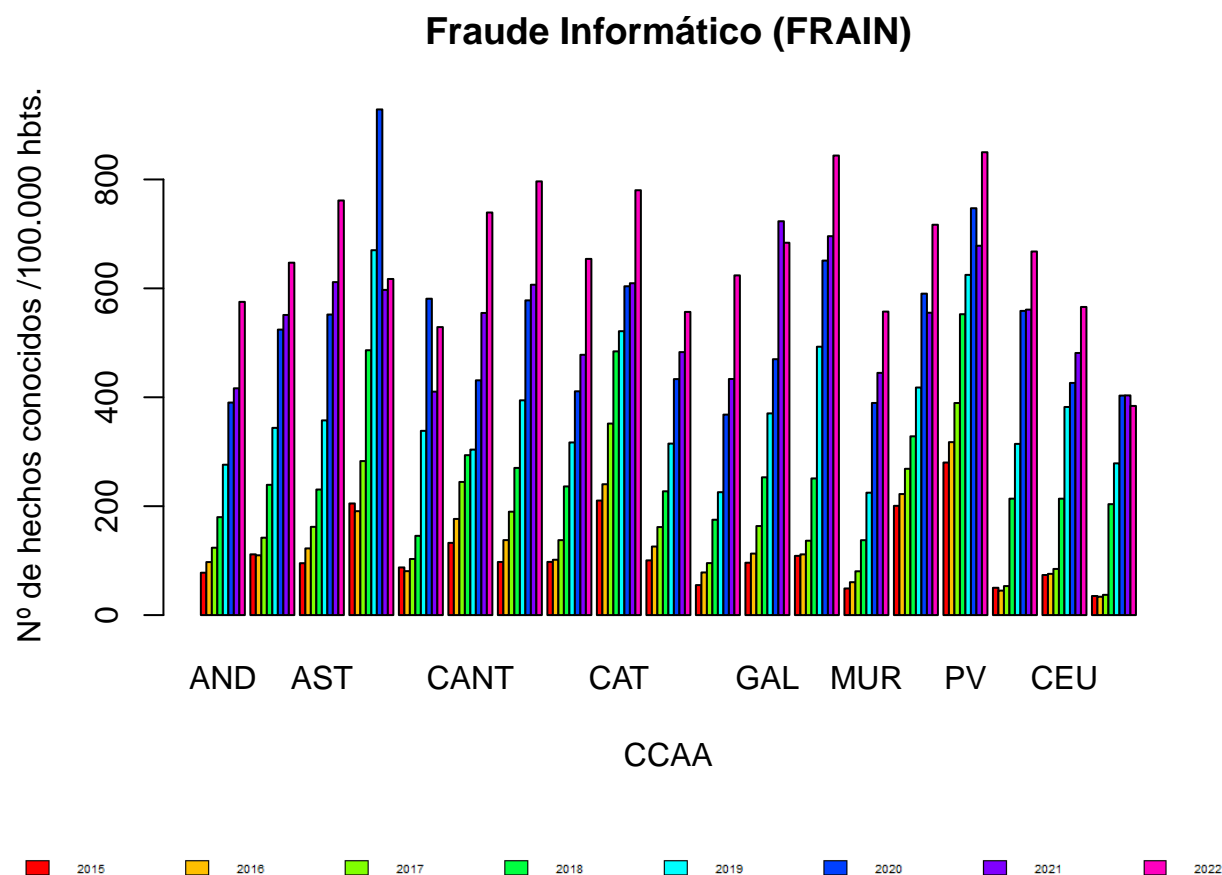


Figure 17: Fraude Informático. Nº hechos conocidos/100.000 habs.

Por último, los delitos de Interferencia en los Datos y en el Sistema, podemos ver como en el año 2021 asciende la tasa en prácticamente todas las comunidades autónomas llegando a ser las más altas, o prácticamente, en comparación con el resto de los años. Destacar que, en Cataluña, al contrario, además de tener un menor número de estos delitos en todos los años, éstos han ido disminuyendo a lo largo del período. (Ver figura 18). En cuanto a CLM, vemos cómo este tipo de delitos, si bien han ido aumentando a lo largo del periodo, no es de las comunidades en la que la tasa sea más alta.

```
barplot(as.matrix(t(IDS_reversed)), beside = TRUE,
        main = "Interferencia en los Datos y en el Sistema (IDS)",
        xlab = "CCAA", ylab = "Nº de hechos conocidos /100.000 hbs.",
        col = rainbow(ncol(IDS_reversed)))
```

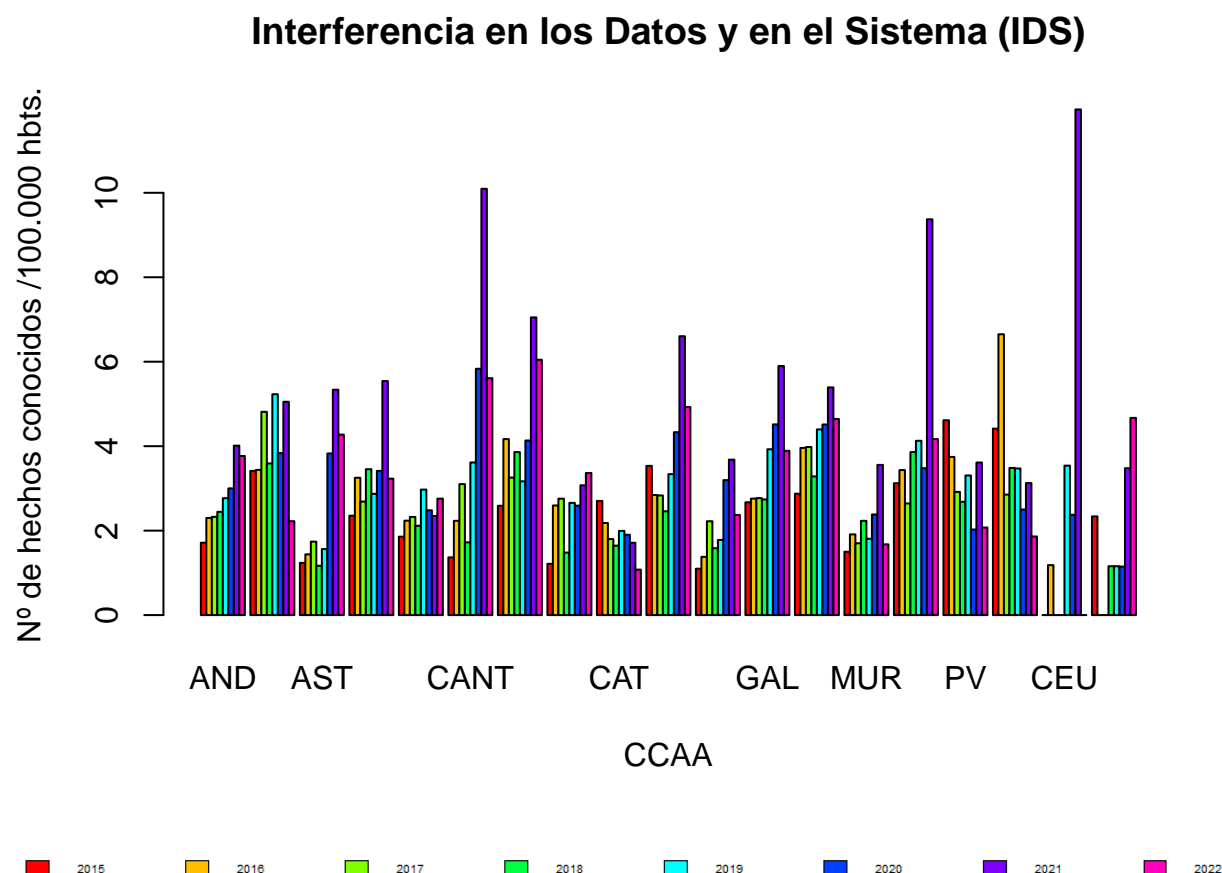


Figure 18: Interferencia en los Datos y en el Sistema. Nº hechos conocidos/100.000 habs.

Resumiendo, como características generales, podemos apreciar un aumento a lo largo de los años en todos los grupos penales, sobre todo a raíz de la pandemia COVID-19, siendo la más evidente en los delitos de Falsificación y Fraude informáticos. Bien es cierto que, en los delitos contra el Honor, parece ser lo contrario, aunque muy ligeramente. Por otra parte, tenemos el delito de Contra la Propiedad Industrial/Intelectual donde no encontramos un patrón claro. CLM no presenta un perfil diferenciador del resto de comunidades, ya que en general se sitúa en la media con respecto al resto de comunidades.

Además, vemos como el grupo penal de Fraude Informático, es, con creces, el que mayor número de veces se comete en todas las comunidades autónomas.

B. Análisis multivariante de la evolución de la tasa de hechos conocidos de los grupos penales

Una vez realizado el análisis descriptivo, se lleva a cabo un PTA, para integrar de forma simultánea la información con la que contamos. Se trata de una herramienta de exploración de datos de 3 vías (como ya hemos desarrollado en el apartado material y métodos).

Algunos de los resultados gráficos de este análisis se presentan a continuación.

```
library(ade4)
library(adegraphics)
```

```
## Registered S3 methods overwritten by 'adegraphics':
```



```

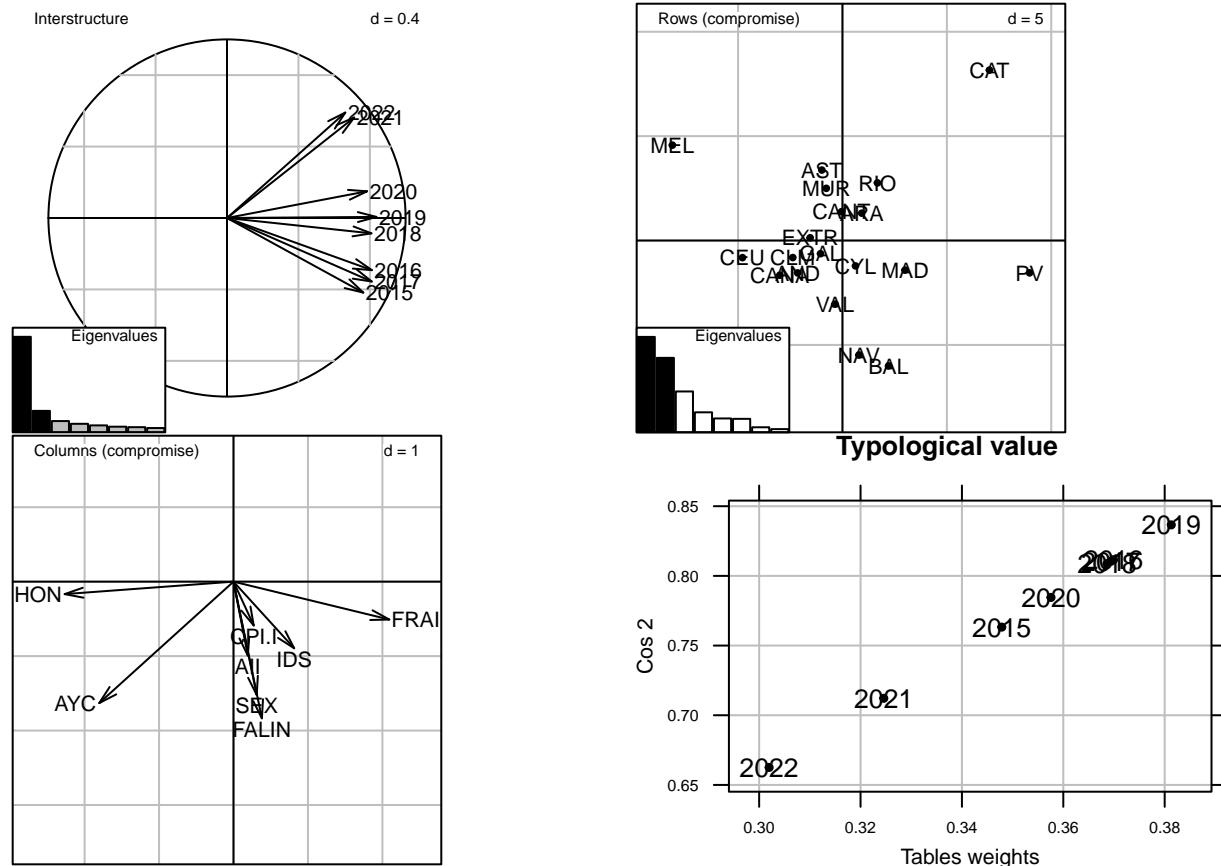
##      method      from
##      biplot.dudi  ade4
##      kplot.foucart ade4
##      kplot.mcoa   ade4
##      kplot.mfa    ade4
##      kplot.pta    ade4
##      kplot.sepan  ade4
##      kplot.statis ade4
##      scatter.coa   ade4
##      scatter.dudi  ade4
##      scatter.nipals ade4
##      scatter.pco   ade4
##      score.acm     ade4
##      score.mix     ade4
##      score.pca     ade4
##      screeplot.dudi ade4

##
## Attaching package: 'adegraphics'

## The following objects are masked from 'package:ade4':
##
##      kplotsepan.coa, s.arrow, s.class, s.corcircle, s.distri, s.image,
##      s.label, s.logo, s.match, s.traject, s.value, table.value,
##      triangle.class

library(ade4TkGUI)
datos$Año=as.factor(datos$Año)
wit <- withinpca(datos[,3:10], datos$Año, scannf = FALSE, scaling = "partial")
kta <- ktab.within(wit, colnames = rep(c('AND', 'ARA', 'AST', "BAL", 'CANA', 'CANT', 'CYL', 'CLM', 'CAT', 'C
kta1 = t(kta)
pta1 = pta(kta1, scann = FALSE)
plot(ppta1, plabels.bboxes.draw=FALSE)

```



En primer lugar, en la parte superior izquierda, tenemos representado los resultados del análisis de la interestructura. Cada año se representa mediante un vector. Se aprecian 3 grupos de años, siendo estos: el grupo de 2015, 2016 y 2017; el de 2018, 2019 y 2020 y por último el grupo formado por los dos últimos años de la serie 2021 y 2022.

Observamos cómo el año 2020 (que se corresponde con el año de la pandemia COVID-19) hace de “año de transición” separando los años 2021 y 2022 (“años post pandemia”) que se encuentran claramente más diferenciados del resto. Abajo a la izquierda vemos los valores propios de la interestructura, con el primer valor propio predominante, claramente separado del resto.

Esto podemos verlo además numéricamente. La interestructura se realiza empleando la matriz de correlaciones vectoriales (RV) entre las distintas matrices, que se presenta en la tabla 4:

```
RV = pta1$RV
rownames(RV) = c("**2015**", "**2016**", "**2017**", "**2018**", "**2019**", "**2020**", "**2021**", "**2022**")
kable(RV, caption = "***Matriz (RV) de coeficientes de correlación vectorial entre los años***")
```

Table 4: *Matriz (RV) de coeficientes de correlación vectorial entre los años*

	2015	2016	2017	2018	2019	2020	2021	2022
2015	1.0000000	0.7405602	0.6827485	0.5453542	0.5501594	0.4837208	0.3777697	0.3118432
2016	0.7405602	1.0000000	0.6872844	0.5835088	0.5828678	0.5371447	0.4287536	0.4327077
2017	0.6827485	0.6872844	1.0000000	0.6573730	0.6494718	0.5410854	0.3847290	0.3600687
2018	0.5453542	0.5835088	0.6573730	1.0000000	0.6781987	0.5846651	0.4585381	0.4672606
2019	0.5501594	0.5828678	0.6494718	0.6781987	1.0000000	0.6511629	0.5771260	0.4541879
2020	0.4837208	0.5371447	0.5410854	0.5846651	0.6511629	1.0000000	0.5692948	0.4910474

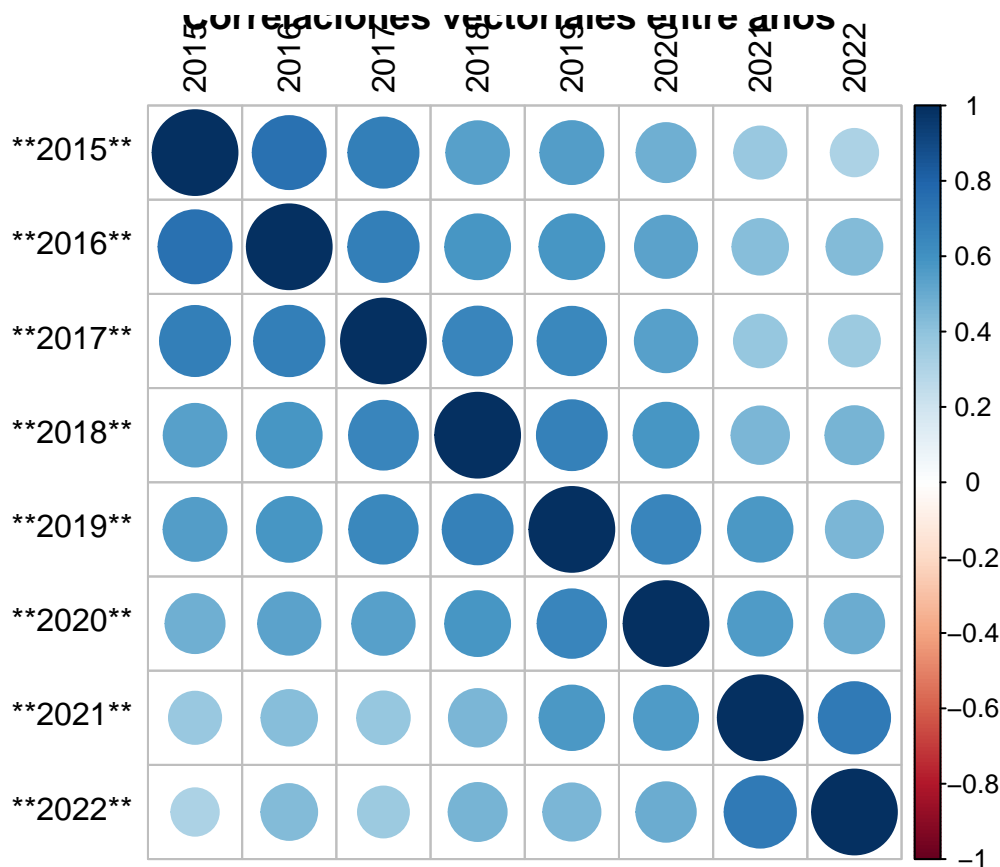
	2015	2016	2017	2018	2019	2020	2021	2022
2021	0.3777697	0.4287536	0.3847290	0.4585381	0.5771260	0.5692948	1.0000000	0.7057745
2022	0.3118432	0.4327077	0.3600687	0.4672606	0.4541879	0.4910474	0.7057745	1.0000000

Vemos que las correlaciones más altas las encontramos entre 2015-2016 y entre 2021 y 2022 (superando el 0.70), y las que más bajas las que se dan entre los tres primeros años con los dos últimos de la serie. En el siguiente gráfico, se representan estas correlaciones de forma gráfica para una interpretación más visual:

```
library(corrplot)
```

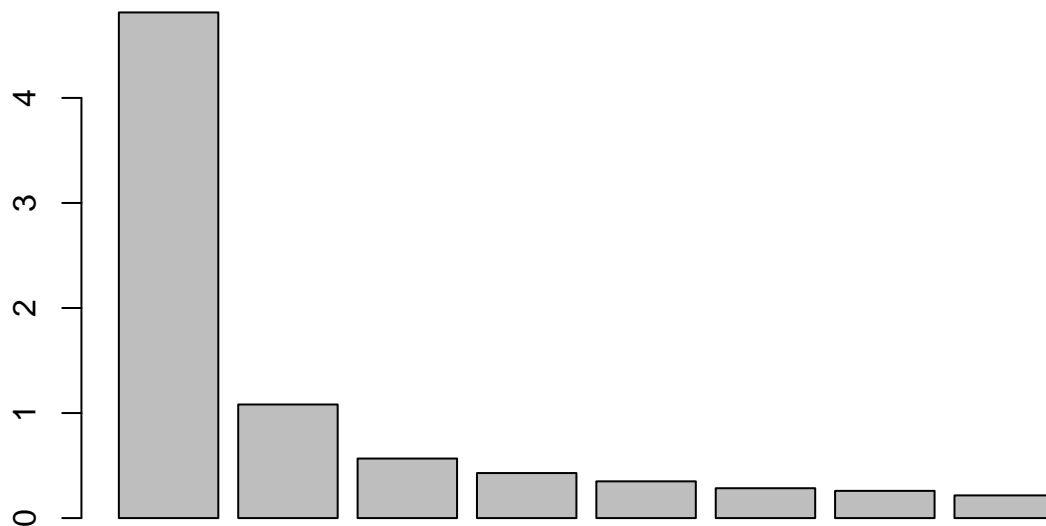
```
## corrplot 0.92 loaded
```

```
corrplot(RV, title = "***Correlaciones vectoriales entre años***", tl.col = "black")
```



Con este gráfico podemos ratificar lo mencionado anteriormente. Podemos ver claramente como a medida que van pasando los años, existe una menor correlación entre el primero año y el que se va sucediendo. A su vez vemos como los años que están contiguos mantienen siempre cierto grado de correlación, siendo el paso del año 2020 al 2021 en el que se produce el mayor salto.

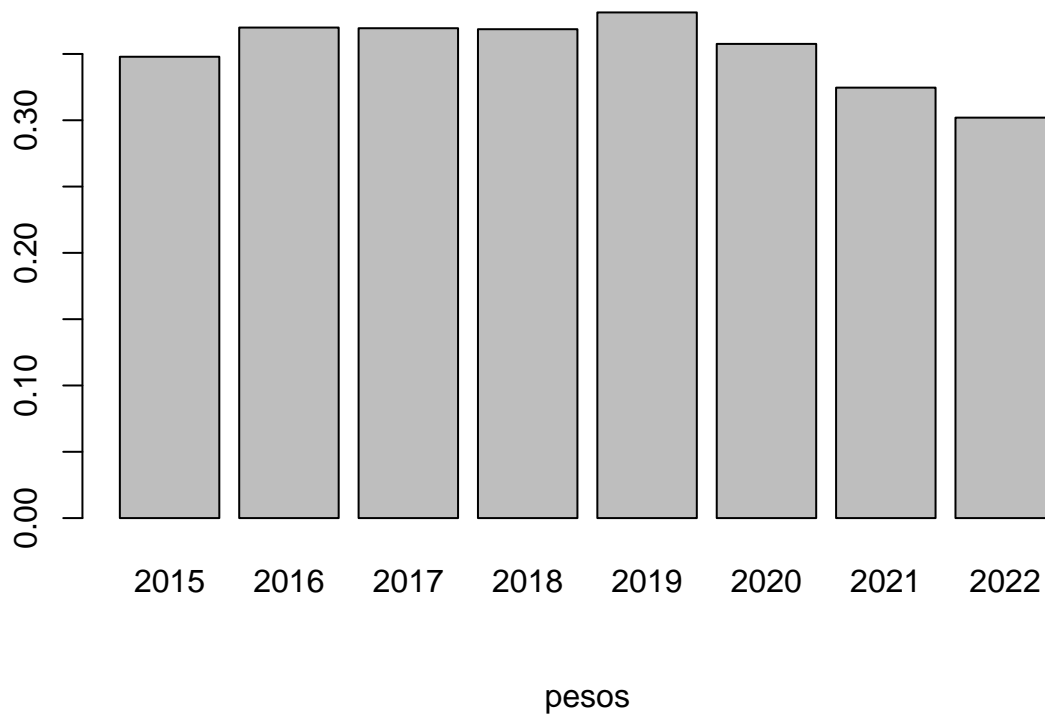
```
barplot(pta1$RV.eig, sub = "Valores Propios Interestructura")
```



Valores Propios Interestructura

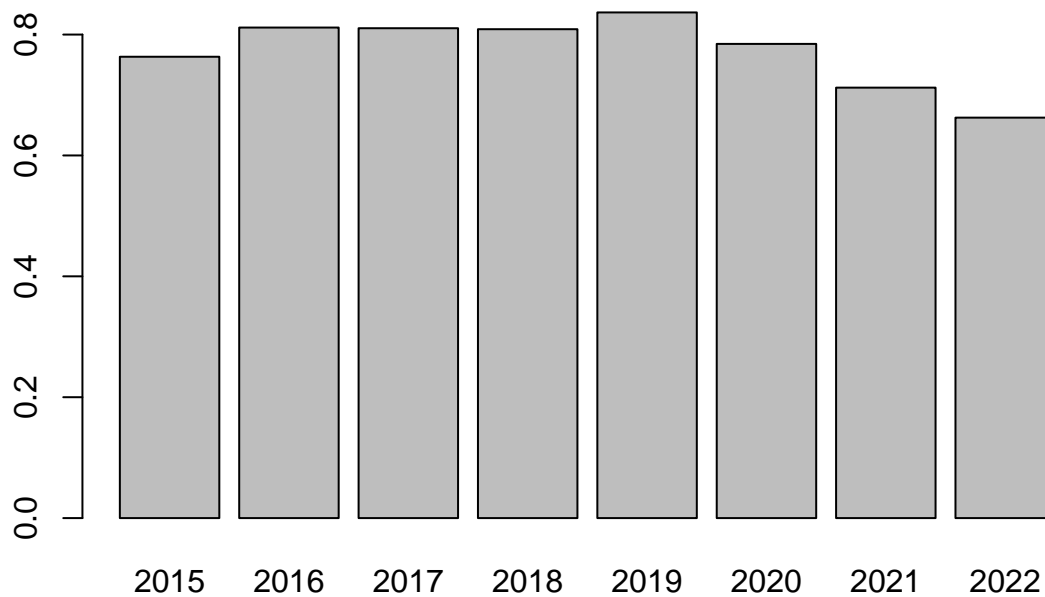
A continuación, se han representado los pesos que cada matriz de datos (cada año) aporta a la construcción del compromiso. Aquí podemos ver que el año 2022 es el que menos aporta en el análisis, seguido de año 2021, mientras que el que más aporta sería el año 2019.

```
barplot(pta1$tabw, names.arg=pta1$tab.names, sub="pesos")
```



Además, podemos obtener las calidades de representación (\cos^2) de los distintos años, siendo el 2019 el mejor representado, seguido del 2016 y 2017:

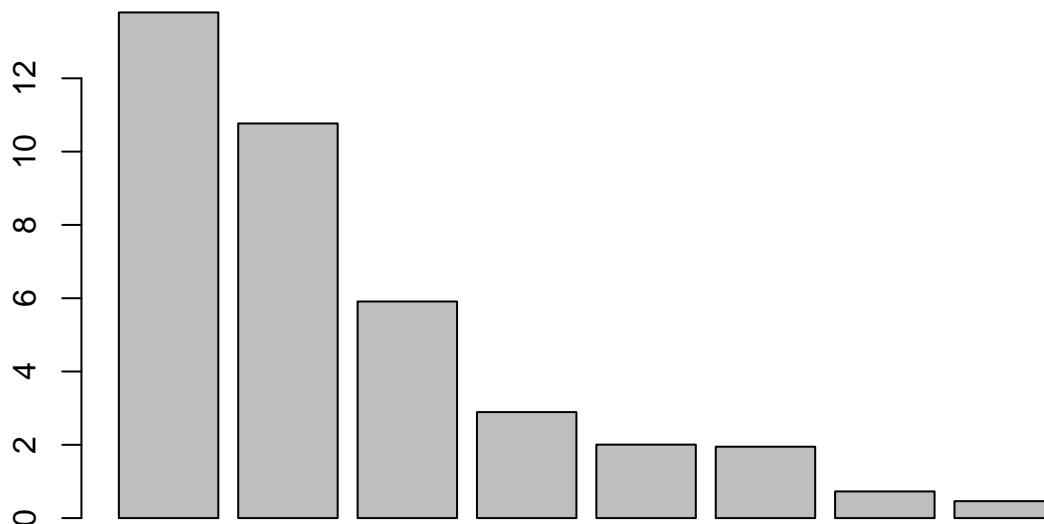
```
barplot(pta1$cos2,names.arg=pta1$tab.names, sub="cos2")
```



cos2

También, se presenta el gráfico de los valores propios del análisis de la interestructura, lo cual equivaldría a que el primer eje recoge un 60,1% de la información y un 13,5% el segundo, es decir, un total de 73,6% de la información.

```
barplot(pta1$eig, sub = "Valores Propios Compromiso")
```



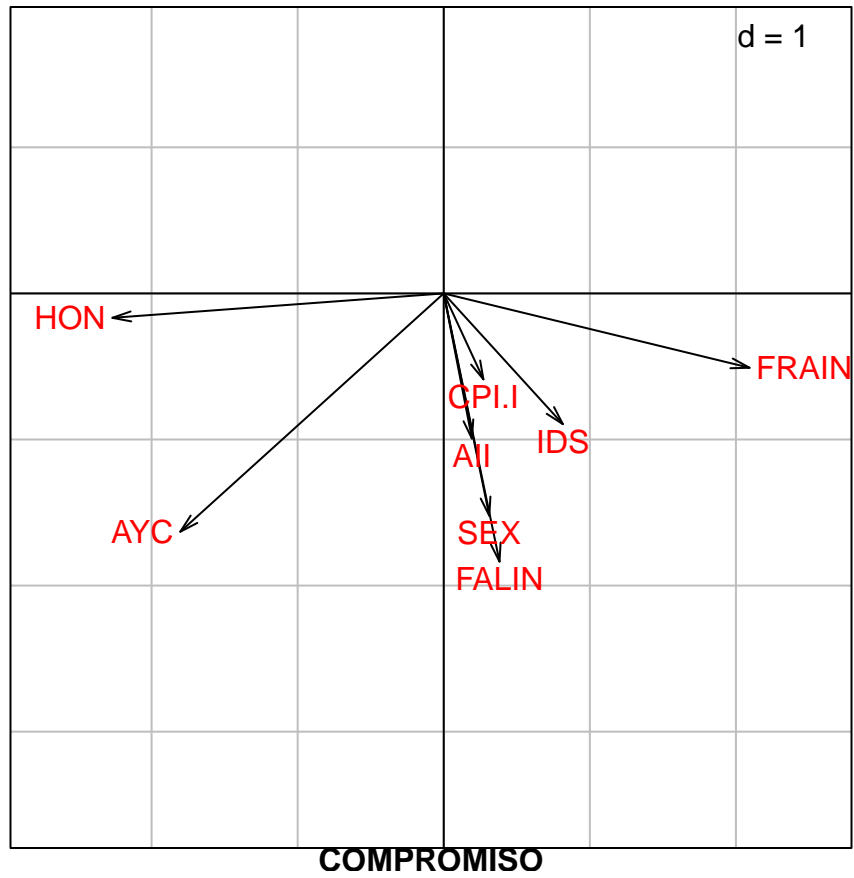
Valores Propios Compromiso

Al igual que con la interestructura, podemos ver la información recogida por los diferentes valores propios del análisis del compromiso. Los dos primeros ejes del compromiso recogen un 35,83% y un 27,96% respectivamente, lo que es un total del 63,79% de la información.

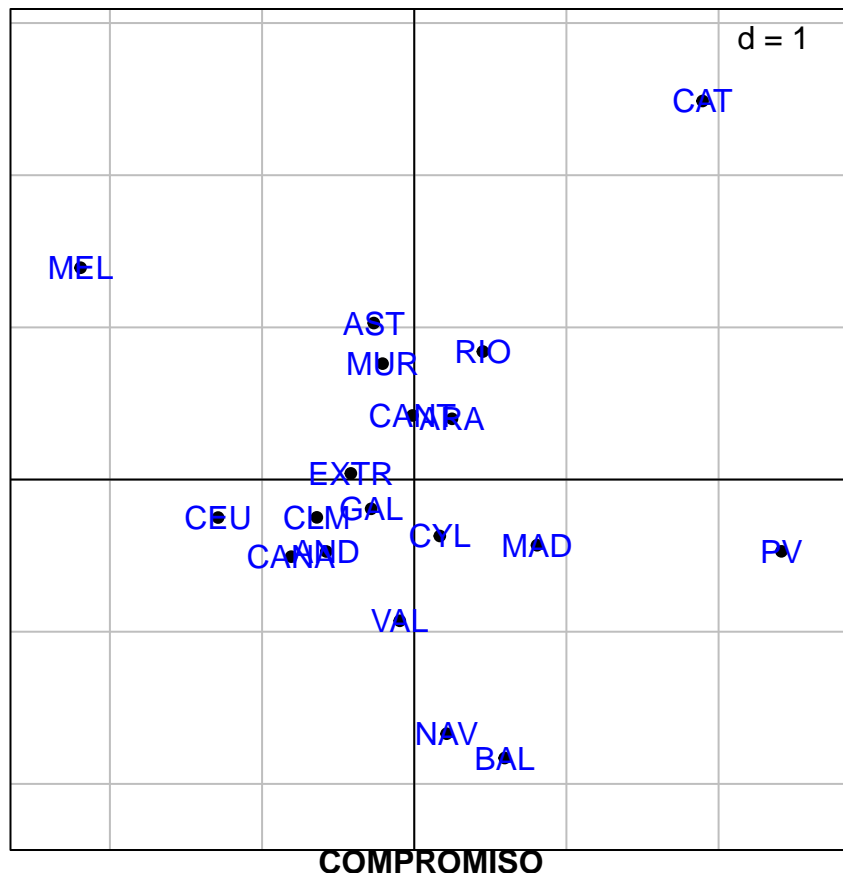
Volviendo al gráfico principal de dicho análisis, el de arriba a la derecha y el de abajo a la izquierda, se corresponden con la representación del compromiso. El primero de ellos hace referencia solo a los individuos (comunidades autónomas), mientras que el segundo a las variables (grupos delictivos).

Se pueden superponer ambos, obteniendo una representación biplot que nos permite realizar mejor la interpretación. Observémoslo de esa manera:(ver siguiente gráfico)

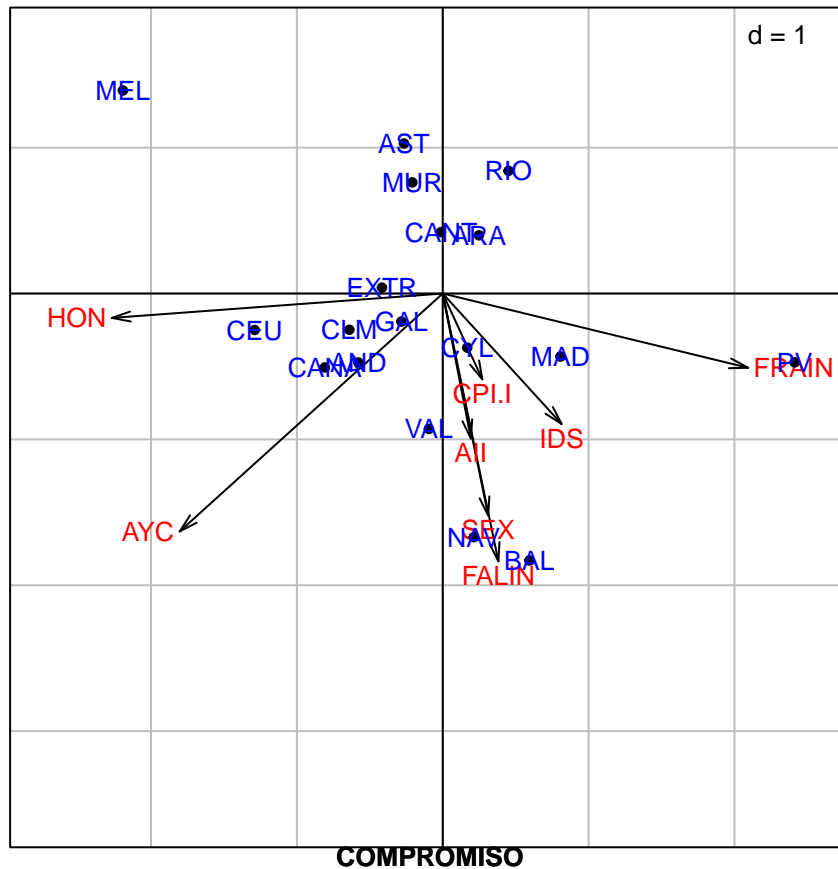
```
g0=s.arrow( pta1$co,plabels.boxes.draw=FALSE,sub="COMPROMISO",plabels.cex=1,plabels.col = "red")
```



```
g1=s.label(pta1$l1,plabels.draw=FALSE, sub="COMPROMISO",plabels.cex=1,plabels.col = "blue")
```

```
g01=superpose(g0,g1,plot=TRUE)
```



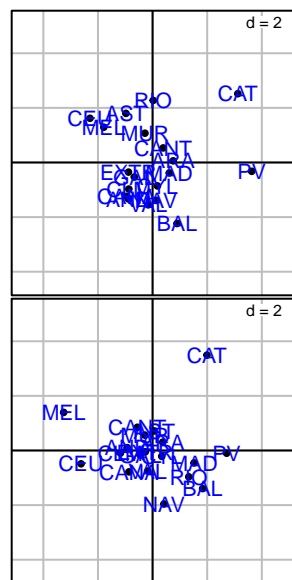
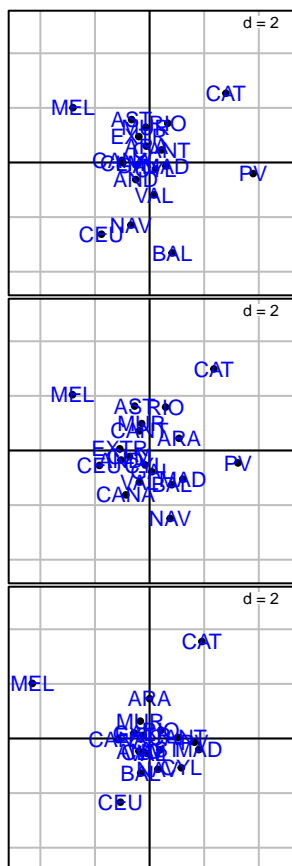
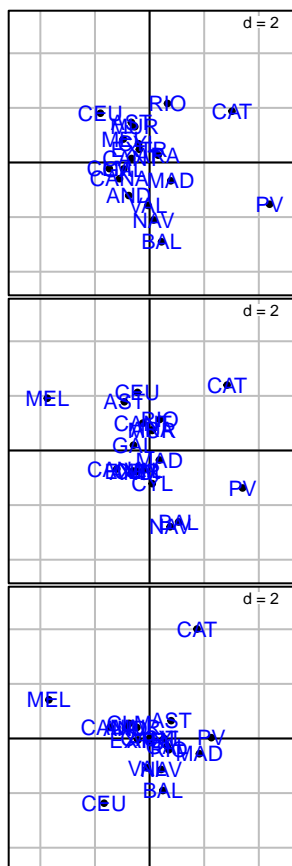
Por ejemplo, observamos como en Melilla (MEL) se cometen principalmente delitos contra el honor; en País Vasco (PV) fraudes informáticos, y en Navarra (NAV) y Baleares (BAL) falsificaciones informáticas (FALIN) y delitos sexuales (SEX). Es decir, en todas estas CC.AA. estos delitos presentan tasas mayores que la tasa media.

También podemos ver altas correlaciones en los grupos penales de delitos sexuales, falsificación informática, contra la propiedad industrial/intelectual, inferencia en datos y acceso e interpretación ilícita. Bien es cierto que el grupo de delitos contra la propiedad industrial/intelectual aparece mal representado en este gráfico y se debería interpretar la información respecto a otra dimensión no representada.

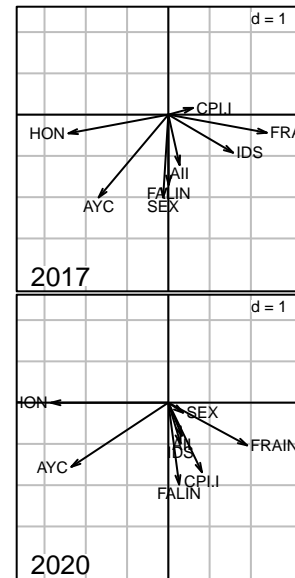
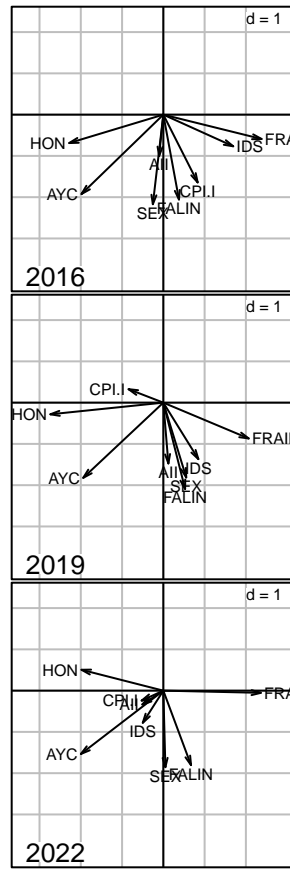
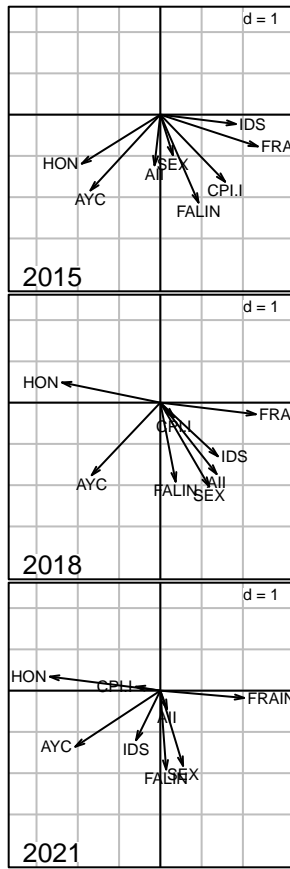
Además, existe independencia de correlación entre los grupos de la propiedad industrial/intelectual (CPI.I) y los de honor (HON) y de este último y delitos sexuales (SEX) y falsificaciones informáticas (FALIN). (vemos que presentan ángulos aproximadamente rectos entre sí).

Por último, en el análisis de la intraestructura, representamos las trayectorias, tanto de las comunidades, como de los grupos penales. Podemos observarlas en la siguiente imagen:

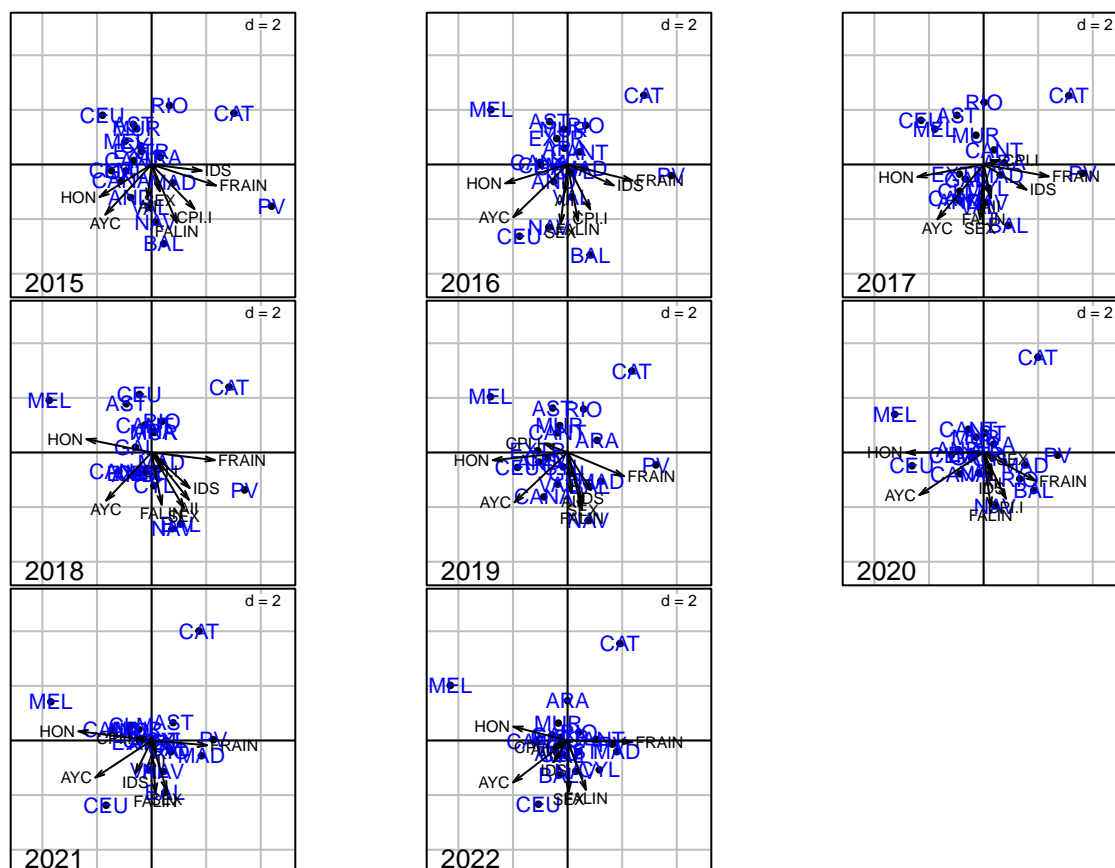
```
la1=s.label(pta1$Tli,facets=pta1$TL[, 1], psub.cex=0, labels=pta1$TL[,2], plabels.cex=1.2, plabels.col=
```



```
ar1=s.arrow(pta1$Tco*3, facets=pta1$TC[, 1], psub.cex=1.5, labels=pta1$TC[,2], plot=TRUE, plabels.bboxes
```



```
s1=superpose(la1,ar1)
plot(s1)
```



Respecto a los grupos penales, a pesar de que muchos de ellos permanecen estables, hay otros como el grupo de IDS que poco a poco se van desplazando, situándose finalmente en una posición muy diferente a la inicial.

Por otra parte, tenemos el grupo de CPI/I que no lleva una evolución paulatina, sino que va variando drásticamente, principalmente entre los años 2017 y 2019.

Otro caso curioso es el del grupo de delitos sexuales (SEX), el cual mantiene una posición más o menos estable durante todos los años salvo en 2020.

El resto de grupos delictivos se mantienen constantes.

Respecto a las CCAA, muchas de ellas, como por ejemplo Cataluña (CAT) y Melilla (MEL), que son las más visibles en los biplots, no varían apenas durante el transcurso del tiempo.

Ceuta (CEU), sin embargo, vemos cómo se va desplazando hacia la parte inferior del gráfico, observando que en el año 2016 donde hace un cambio brusco hacia esa zona volviendo a su situación original al año siguiente. (Ver imagen anterior). Por tanto, durante los últimos años de estudio (además de en 2016), aumentan los delitos de amenazas y coacciones (AYC), falsificaciones informáticas (FALIN) y delitos sexuales (SEX), fundamentalmente.

En la Rioja (RIO) sucede algo parecido, aumentando a partir el 2020, los delitos de fraude informático, interferencia con datos y sistemas, así como los delitos sexuales y falsificaciones informáticas.

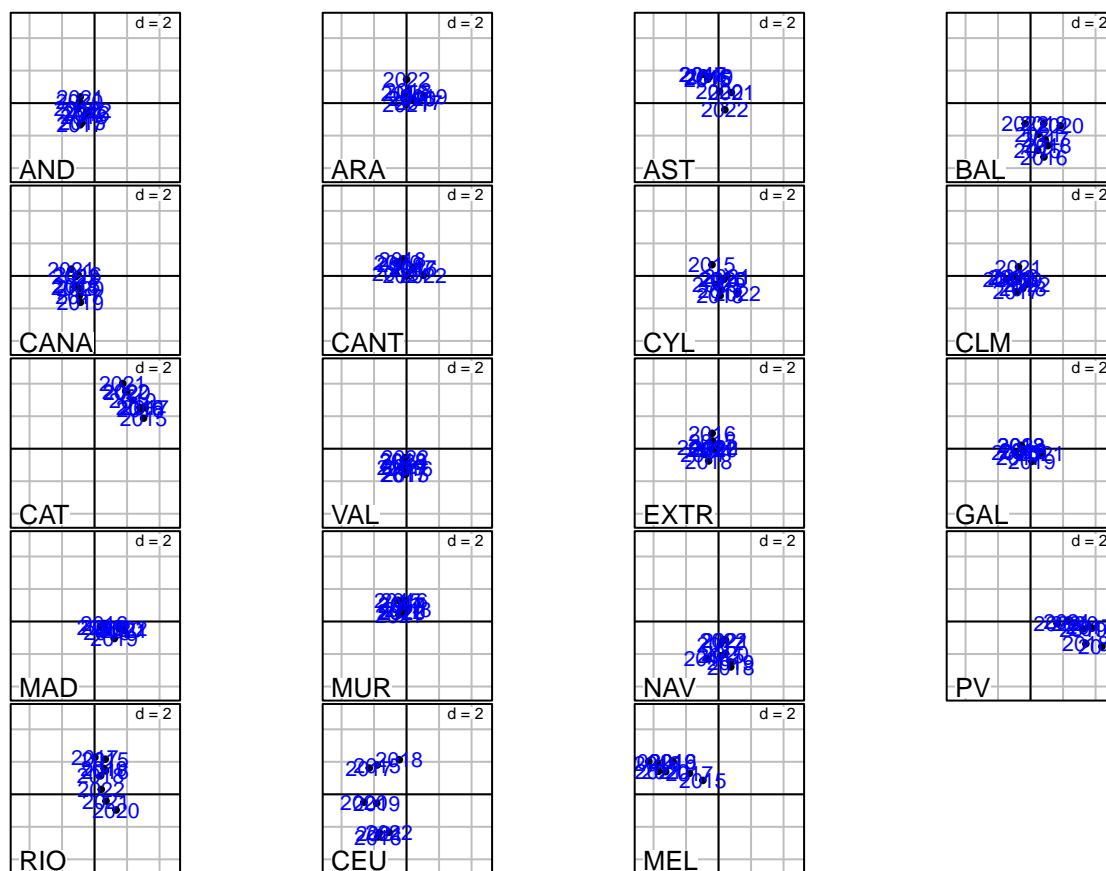
A lo largo de los años, vemos como la sintonía general de las CCAA apunta en mayor medida hacia donde apuntan los grupos penales, la parte inferior del gráfico, acentuándose aún más en los últimos 2 años, lo cual significa un aumento de los hechos conocidos de ciberdelitos durante este periodo de tiempo. Esto puede corresponderse a distintas posibilidades. Una de ellas es que, a medida que pasa el tiempo, van aumentando las formas de comisión de dichos delitos. Otra opción consistiría en la falta de concienciación de la sociedad al respecto, lo cual implica una gran vulneración. También, podrían darse las dos alternativas

simultáneamente. Esa acentuación final, podría deberse a la aparición del COVID-19, ya que a raíz de ese acontecimiento observamos un aumento de las cifras.

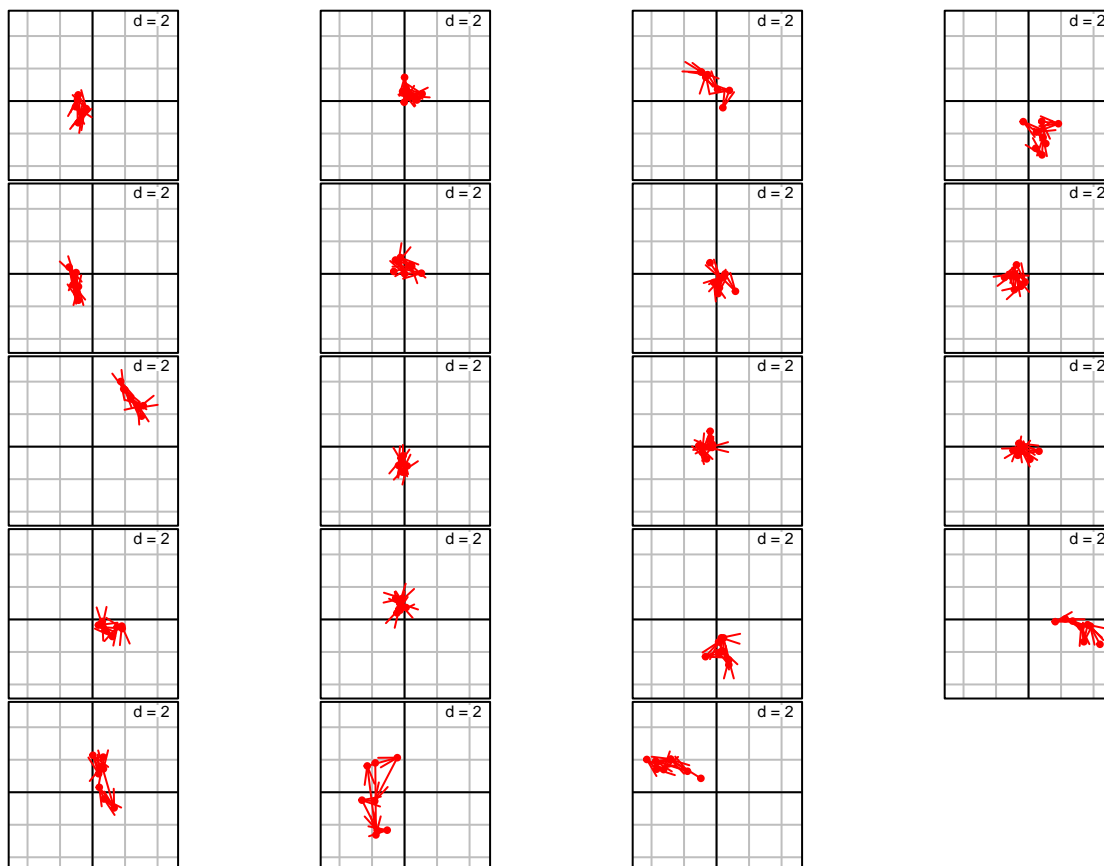
Es por ello necesario un plan de prevención que trate de regular este sector ya que cada vez son más las personas afectadas, en otras palabras, este tipo de delitos cada vez suponen un porcentaje más amplio en relación al total de delitos.

Haciendo referencia a CLM vemos como su trayectoria es envolvente (ver siguiente gráfico), lo que indica que a lo largo del periodo estudiado su posición no ha cambiado prácticamente. Vemos que su posición en el gráfico se corresponde con tasas algo mayores que la media en delitos contra el honor (HON) y AYC en prácticamente todos los años, y también algo mayores en delitos sexuales (SEX) y FALIN y algo menores que la media en FRAIN e IDS. Por tanto, podemos decir que CLM es una comunidad que no destaca de manera prominente del resto de CC.AA.

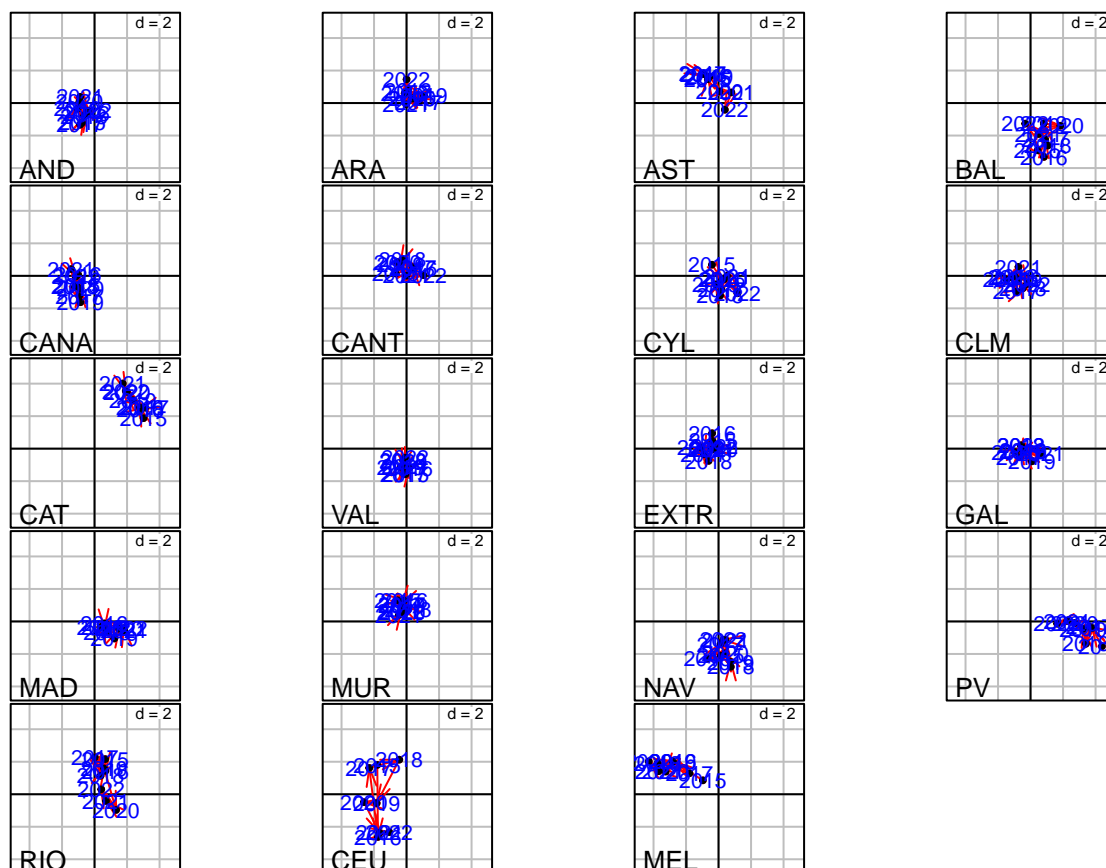
```
la1.1=s.label(pta1$Tli,facets=pta1$TL[, 2], psub.cex=1.5, labels=pta1$TL[,1], plabels.cex=1.2, plabels.
```



```
tr1=s.trajectory(pta1$Tli, facets=pta1$TL[, 2], plabels.cex = 0, psub.cex=0,col="red", plot=TRUE)
```



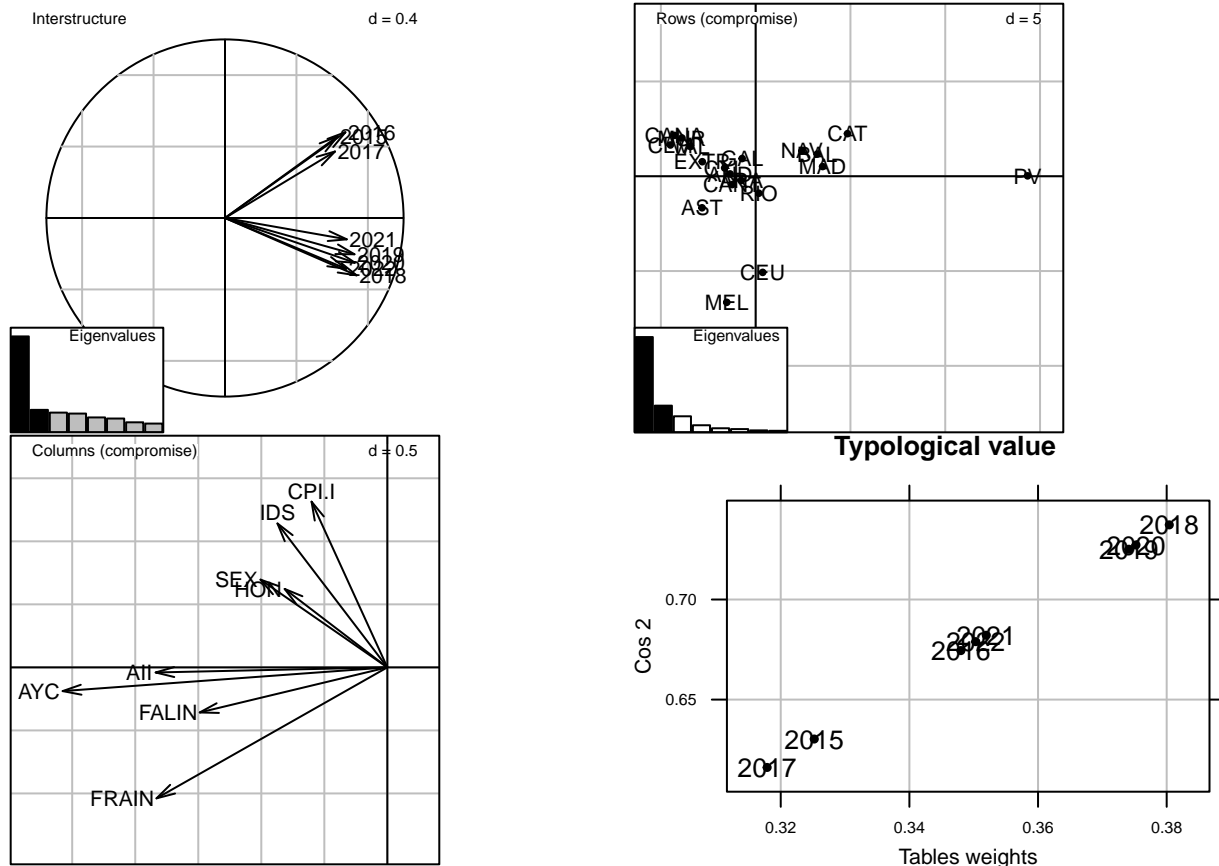
```
s2=superpose(tr1,la1.1)
plot(s2)
```



C. Análisis multivariante de la evolución del índice de esclarecimiento (IE) de los ciberdelitos por grupos penales.

A continuación, se presentan los resultados del análisis triádico parcial del índice de esclarecimiento. Recordemos que el índice de esclarecimiento es el cociente entre los hechos esclarecidos y los conocidos multiplicado por 100 y se trata de una medida de la eficacia policial.

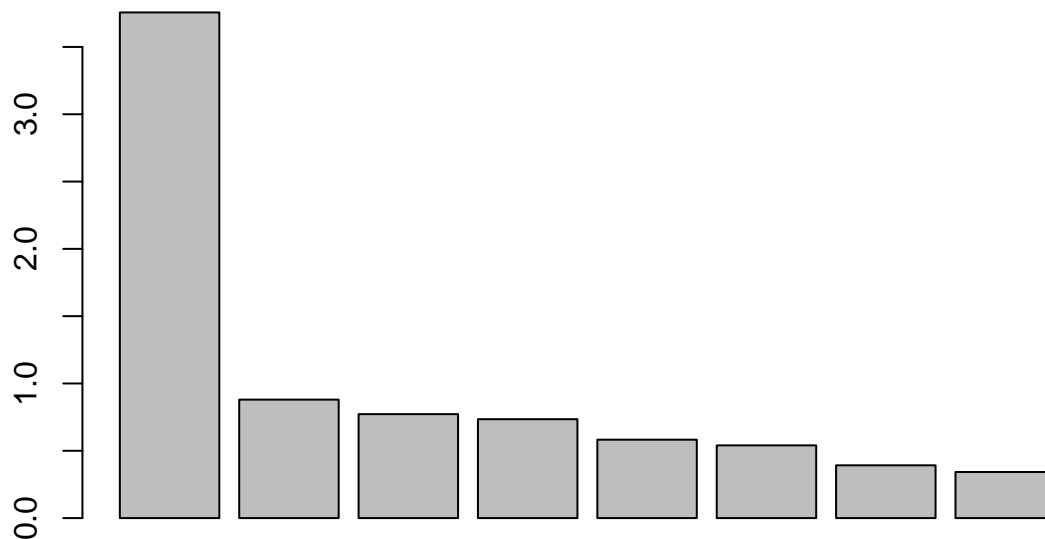
```
datos_IE=read_excel('conocidos, esclarecidos e IE.xlsx')
wit_IE = withinpca(datos_IE[,3:10], datos$Año, scannf = FALSE, scaling = "partial")
kta_IE <- ktab.within(wit_IE, colnames = rep(c('AND', 'ARA', 'AST', "BAL", 'CANA', 'CANT', 'CYL', 'CLM', 'CAT', 'VAL', 'EXTR', 'GAL', 'MAD', 'MUR', 'NAV', 'PV', 'RIO', 'CEU', 'MEL'), 19))
kta2 = t(kta_IE)
pta2 = pta(kta2, scann = FALSE)
plot(pta2, plabels.boxs.draw=FALSE)
```

En este caso, respecto a la interestructura de los datos, sí que podemos apreciar claramente dos grupos, siendo uno de ellos el formado por los tres primeros años de la serie 2015, 2016 y 2017, y el otro el compuesto por el resto de años. Esto puede deberse a que a lo largo de los años hemos ido progresando en cuanto a la resolución de dichos delitos, aumentando con ello el índice de esclarecimiento. (ver gráfico anterior en su aprte superior izquierda).

En este caso la información recogida por las dos primeras componentes principales del análisis de la interestructura es de 46,95% y 11% respectivamente, es decir, un total de 57,95% en el plano. (ver en el siguiente gráfico)

```
barplot(pta2$RV.eig, sub = "Valores Propios Interestructura del índice de esclarecimiento")
```



Valores Propios Interestructura del índice de esclarecimiento

Los mayores pesos que nos encontramos son 0.38, 0.374 y 0.375, correspondiente, al igual que en las calidades de representación, a los años 2018, 2019 y 2020, respectivamente. Respecto a los que tienen los menores pesos, observamos a continuación, que también lo son los años 2015 y 2017 con los valores 0.32 y 0.31, respectivamente.

```
barplot(pta2$tabw,names.arg=pta2$tab.names, sub="Gráfico de los pesos de cada matriz para la formación de
```

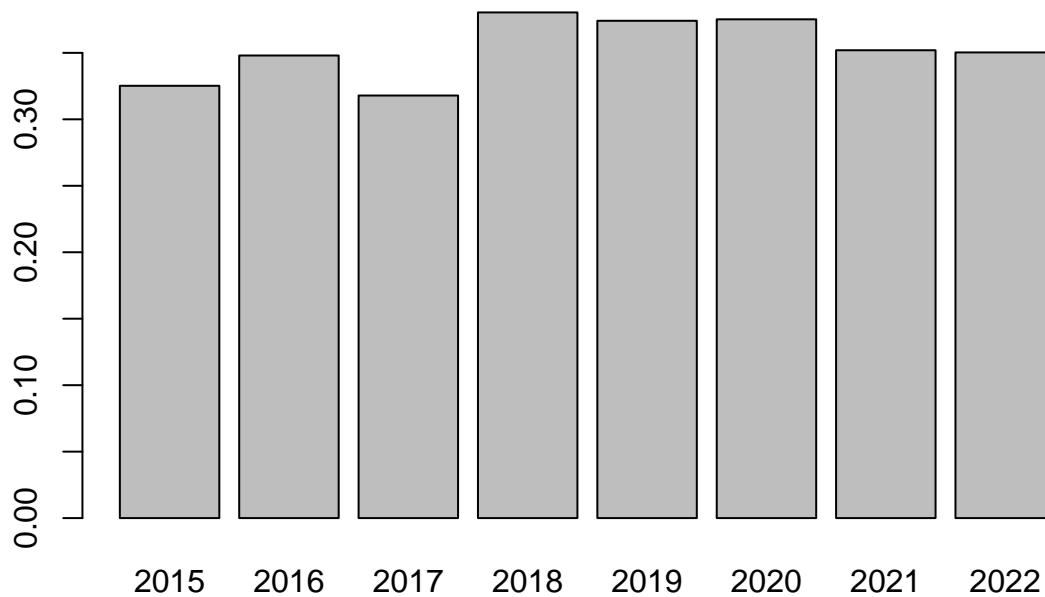


Gráfico de los pesos de cada matriz para la formación del compromiso del índice de esclare

En el siguiente gráfico se representan los cosenos al cuadrado. En esta podemos ver 3 grupos bien diferenciados. Los años 2018, 2019 y 2020 son los que mejor vienen representados por el compromiso, siendo el año 2017, seguido del 2015, los que menos.

```
barplot(pta2$cos2,names.arg=pta2$tab.names, sub="Gráfico de las calidades de representación del índice c
```

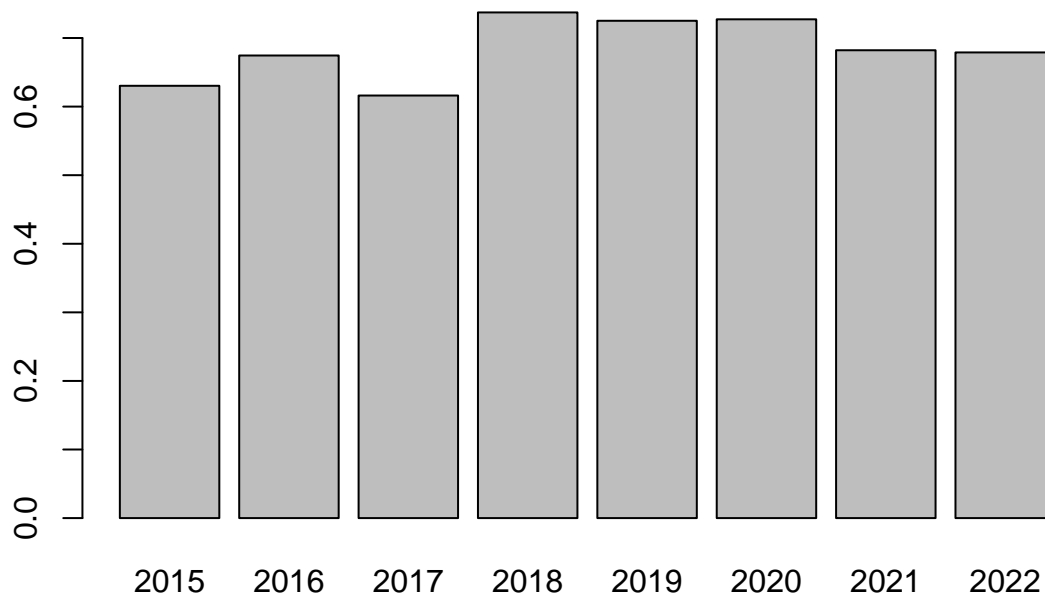
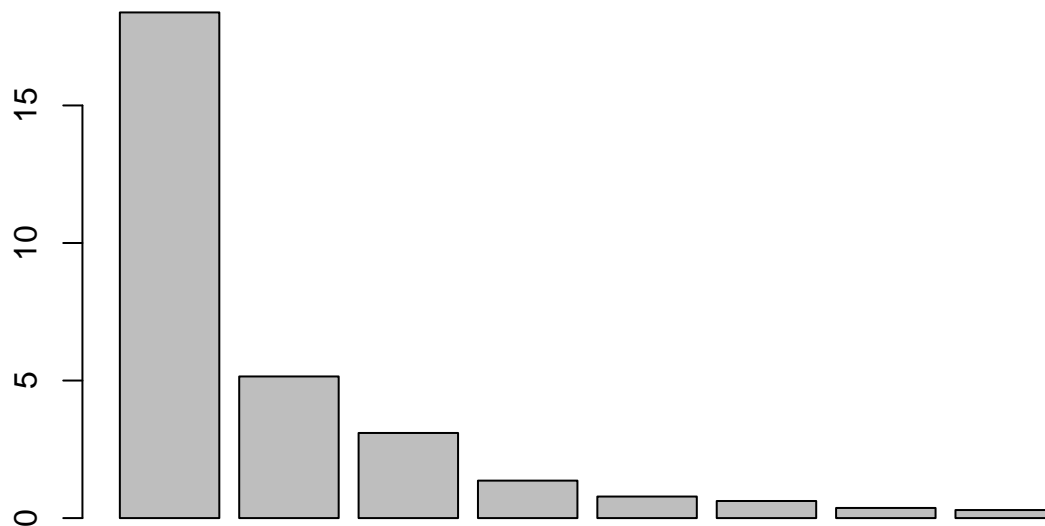


Gráfico de las calidades de representación del índice de esclarecimiento

Vemos que los años 2018, 2019 y 2020 son los que presentan una mayor calidad de representación, con los valores 0.737, 0.725 y 0.727, respectivamente. Por otra parte, los años 2015 y 2017 son los que presentan las peores con los valores 0.63 y 0.61, respectivamente.

```
barplot(pta2$eig, sub = "Valores Propios Compromiso")
```

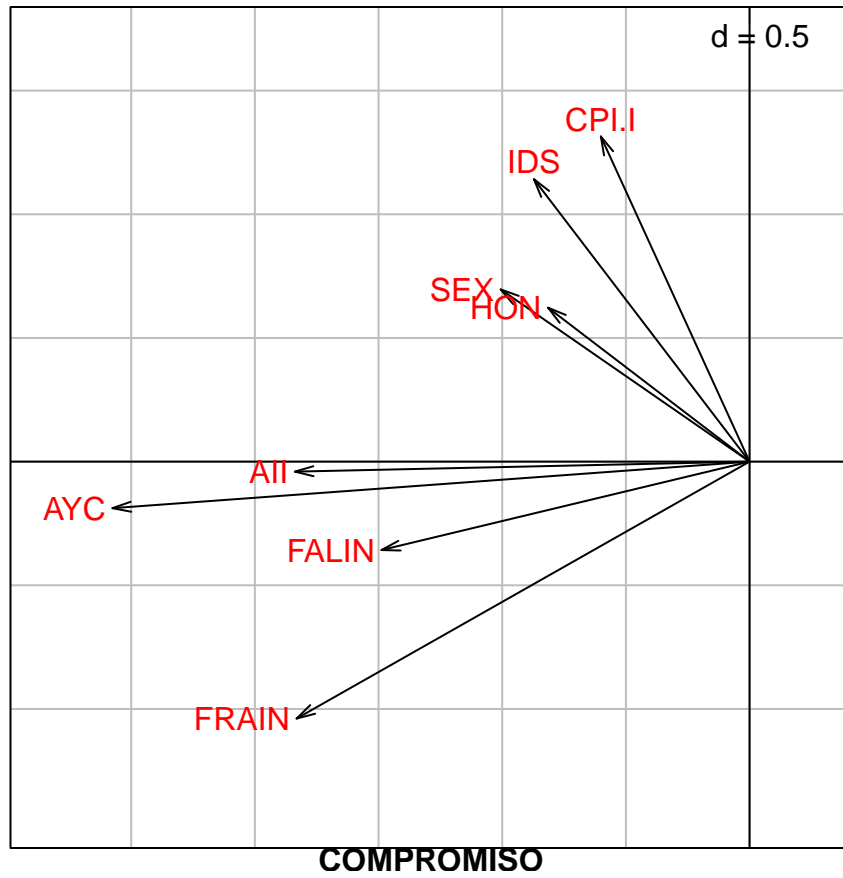


Valores Propios Compromiso

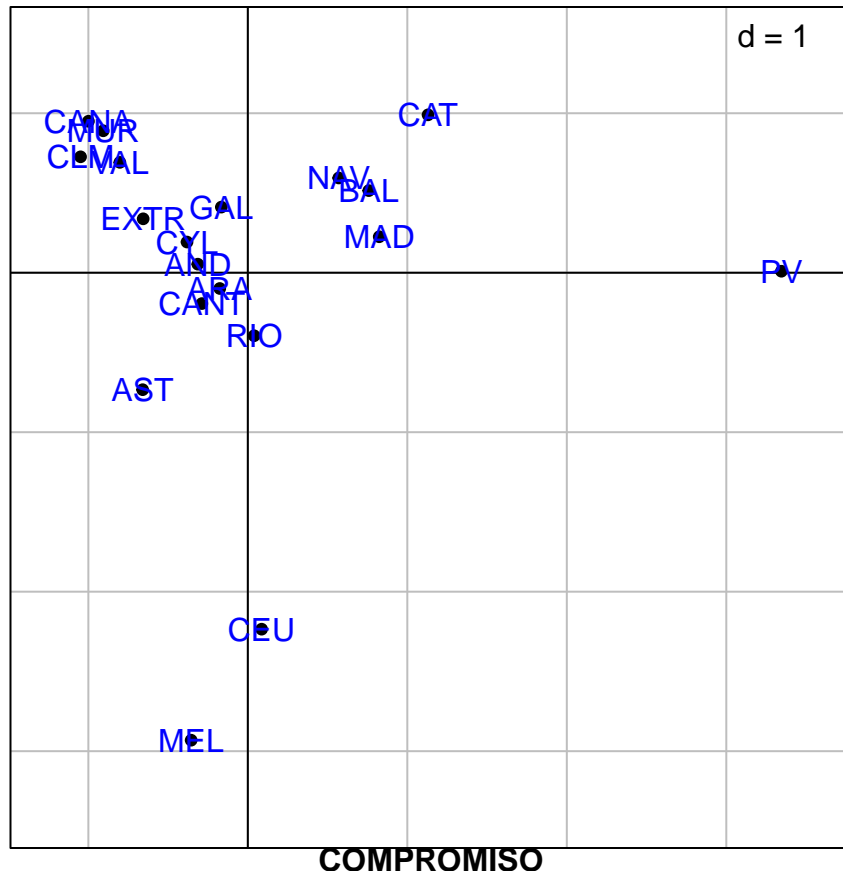
En cuanto a la información recogida por las dos primeras componentes principales, tenemos que la primera recoge un 31.26% de la información y la segunda un 17,16%, es decir, un total de 48,42% de la información es la recogida por el análisis de la matriz compromiso. (ver gráfico anterior)

Para observar mejor el resultado el análisis del compromiso, se presenta el gráfico, un biplot donde superpongo las comunidades y los grupos penales: (Ver siguiente gráfico)

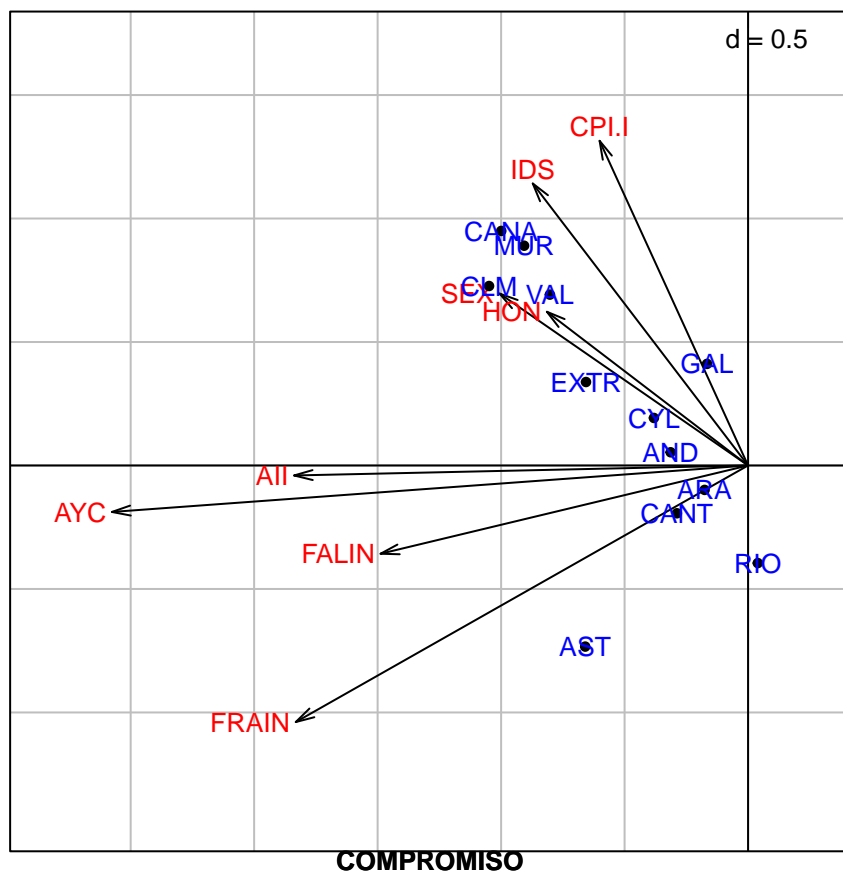
```
g2=s.arrow(pta2$co,plabels.draw=FALSE,sub="COMPROMISO",plabels.cex=1,plabels.col = "red")
```



```
g3=s.label(pta2$11,plabels.draw=FALSE, sub="COMPROMISO",plabels.cex=1,plabels.col = "blue")
```



```
g23=superpose(g2,g3,plot=TRUE)
```



Si atendemos al gráfico, en los grupos penales, podemos apreciar también dos agrupaciones. La primera sería el compuesto por los grupos CPI.I, IDS, SEX y HON, los cuales están muy relacionados entre sí; y la otra por ALL, AYC, FALIN y FRAIN, que también guardan una alta correlación entre sí.

Observamos total independencia de relación entre los grupos CPI.I y FALIN al igual que entre SEX y HON con FRAIN.

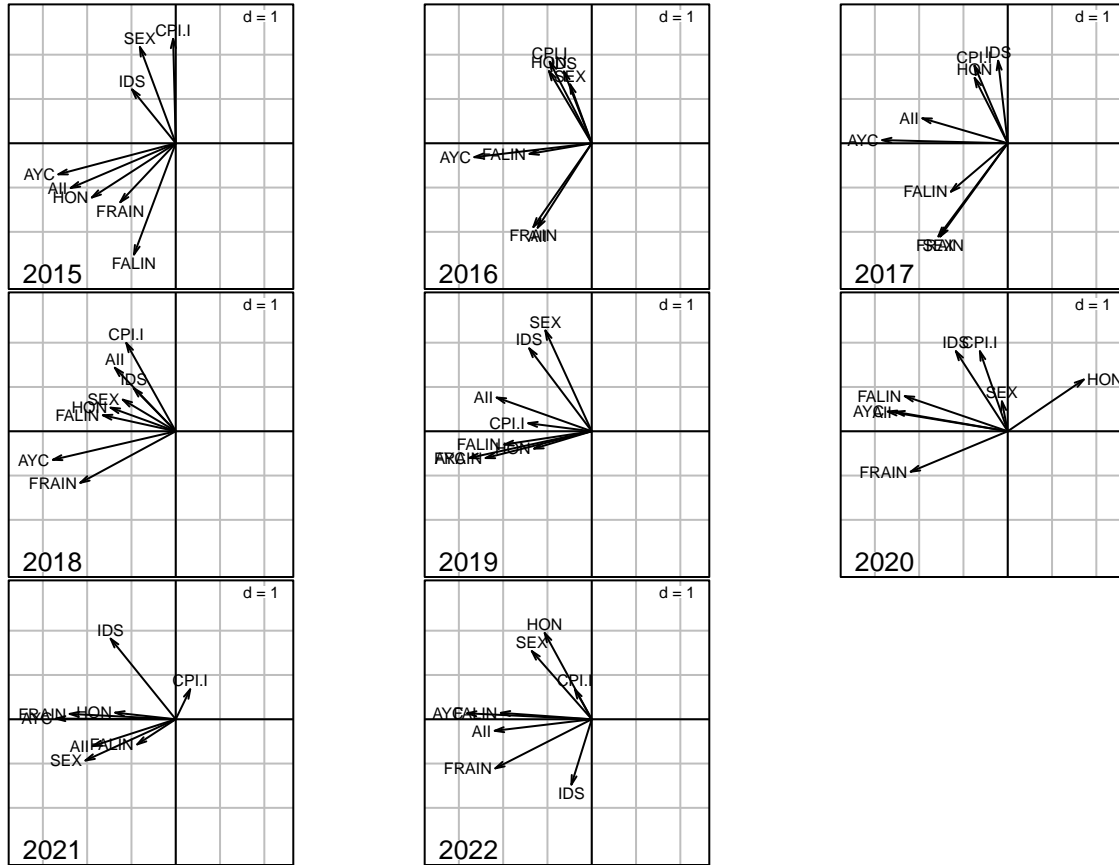
En Madrid tenemos valores por debajo de la media en todos los grupos penales. Valores similares encontramos en Cataluña, Navarra, Baleares y País Vasco, aunque este último ya hemos dicho que está mal representado en este caso ya que no se incluyen datos de la Ertzaintza en toda la serie histórica, por eso aparece en el extremo derecho del eje 1 siendo la comunidad autónoma con menor índice de esclarecimiento en términos medios de toda la serie temporal.

Por otra parte, en comunidades como Canarias, Murcia, Castilla la Mancha y Valencia, encontramos valores por encima de la media en todos los grupos penales.

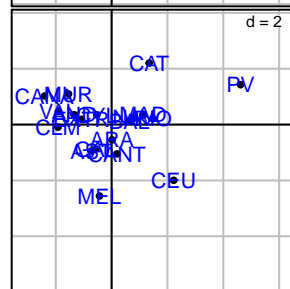
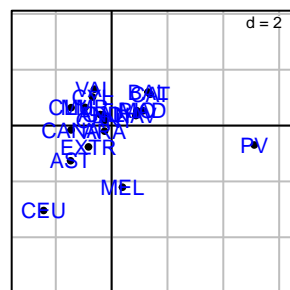
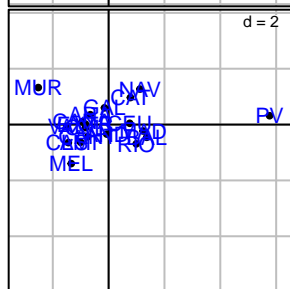
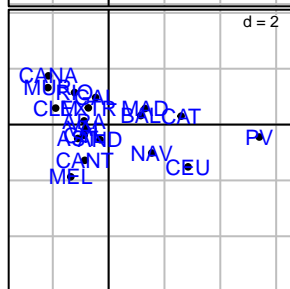
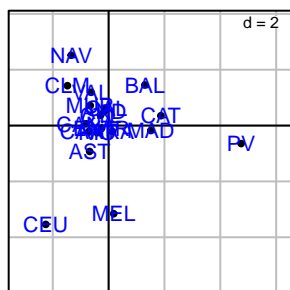
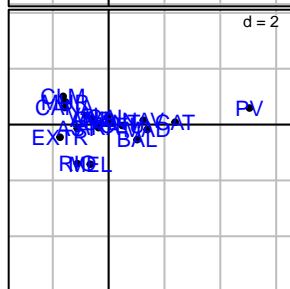
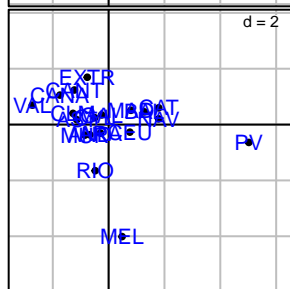
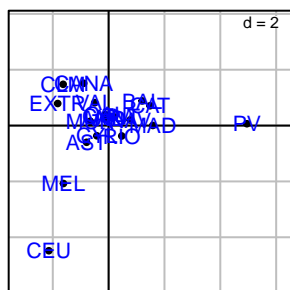
Ceuta y Melilla destacan, sobre todo, por tener índices de esclarecimiento más altos que el resto de comunidades de los delitos que se sitúan en el tercer cuadrante: FRAIN, FALIN, AYC y ALL, y menores son los que menor índice de esclarecimiento tienen el resto de cibercriminales (a excepción del País Vasco, que ya hemos comentado).

Observemos a continuación las trayectorias de este nuevo análisis. (ver Figura a continuación)

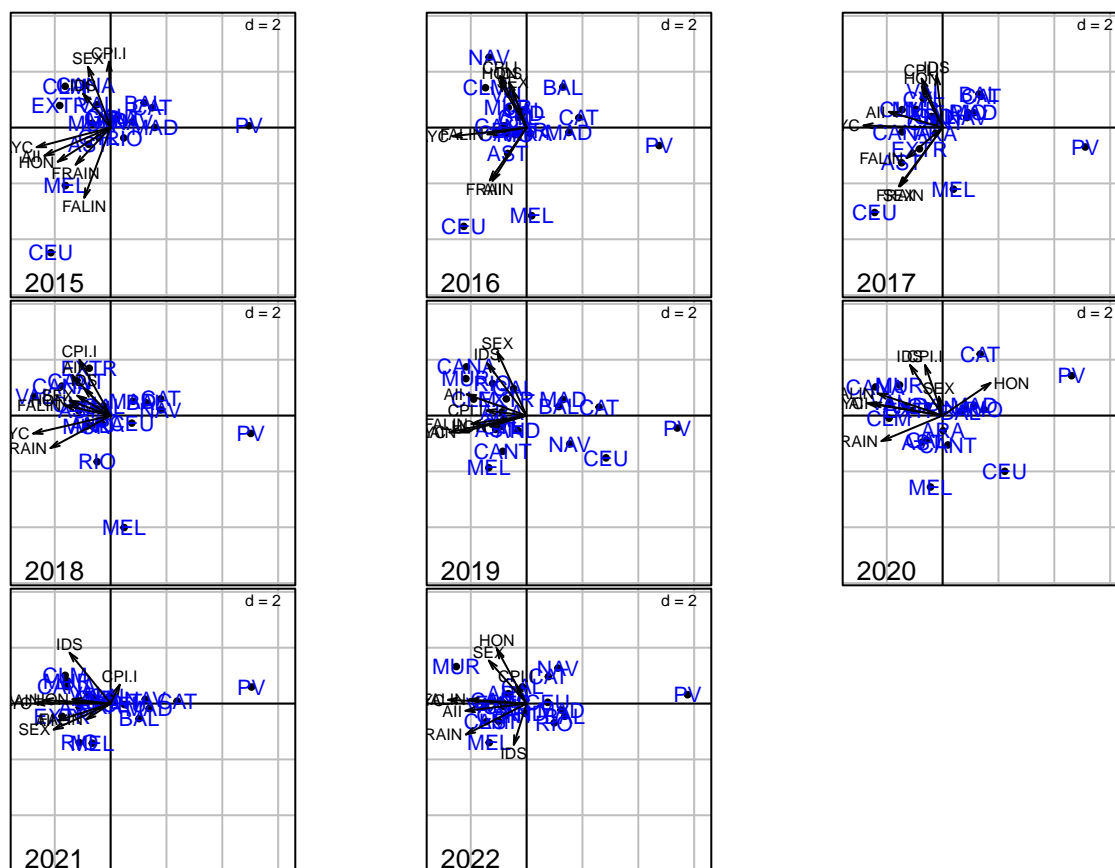

```
ar2=s.arrow(pta2$Tco*3, facets=pta2$TC[, 1], psub.cex=1.5, labels=pta2$TC[,2], plot=TRUE, plabels.bboxes
```



```
la2=s.label(pta2$Tli,facets=pta2$TL[, 1], psub.cex=0, labels=pta2$TL[,2], plabels.cex=1.2, plabels.col=
```



```
s3=superpose(la2,ar2)
plot(s3)
```



En cuanto a los grupos penales, encontramos que el grupo SEX que varía bruscamente de posición en el año 2021. Lo mismo ocurre con el de CPI/I en 2019 y con el de IDS en 2022.

El grupo HON va desplazándose durante todos los años sin encontrar un patrón claro.

El resto de grupos parecen mantenerse estables. Respecto a las CCAA, dentro de lo que podemos apreciar, observamos la estabilidad de País Vasco, por las razones que ya hemos comentado con anterioridad.

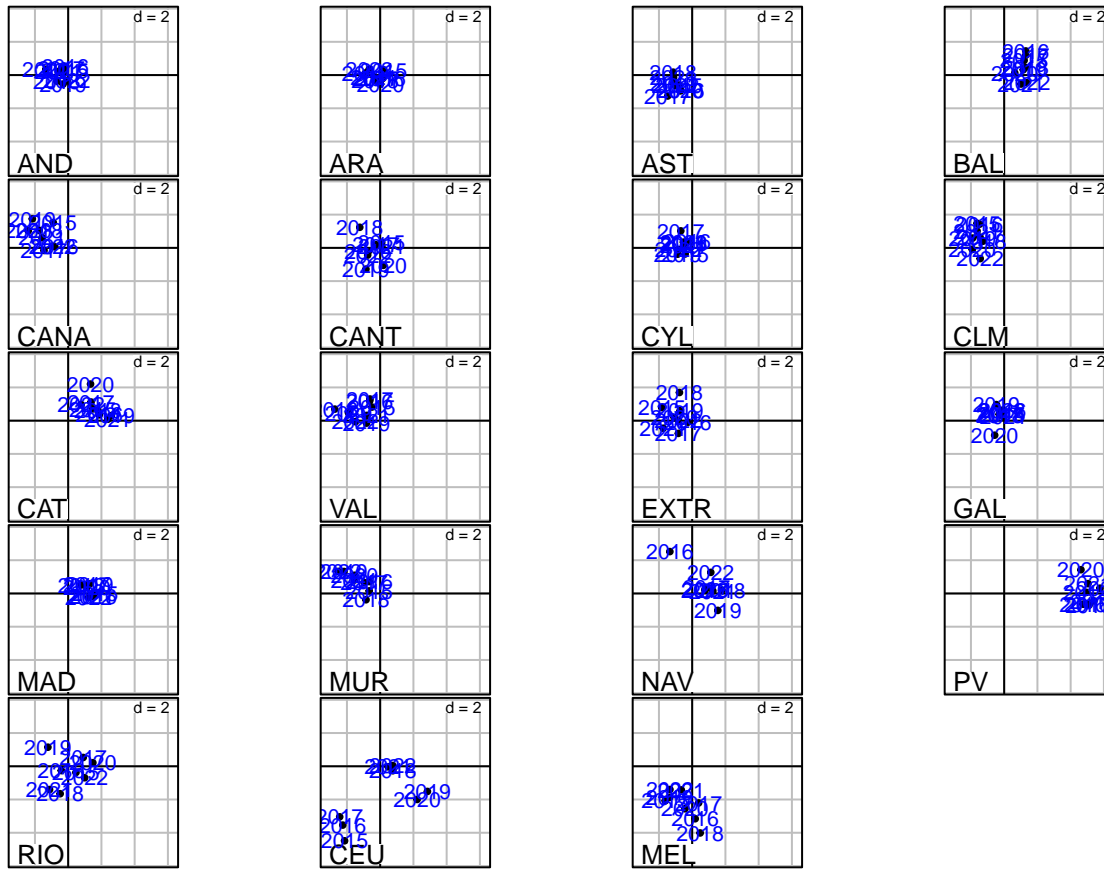
Ceuta se va desplazando durante los 3 primeros años por la parte inferior del gráfico y a raíz de año 2018, incluido, se traslada a la parte central. A simple vista parece ser la comunidad más irregular de todas, pudiendo ser debido a las cifras nulas de ciertos delitos en algunos años. (Ver siguiente gráfico). Por tanto, en los tres primeros años el índice de esclarecimiento de los delitos de FRAIN, FALIN, AYC y AII, era mayor que en los últimos años, donde el IE (índice de esclarecimiento) de esta comunidad es menor que la media para todos los delitos.

El resto de Comunidades parecen mantenerse estables por el centro del gráfico, siendo las que más lo hacen Madrid, Aragón, Andalucía y Castilla y León. (Ver siguiente gráfico).

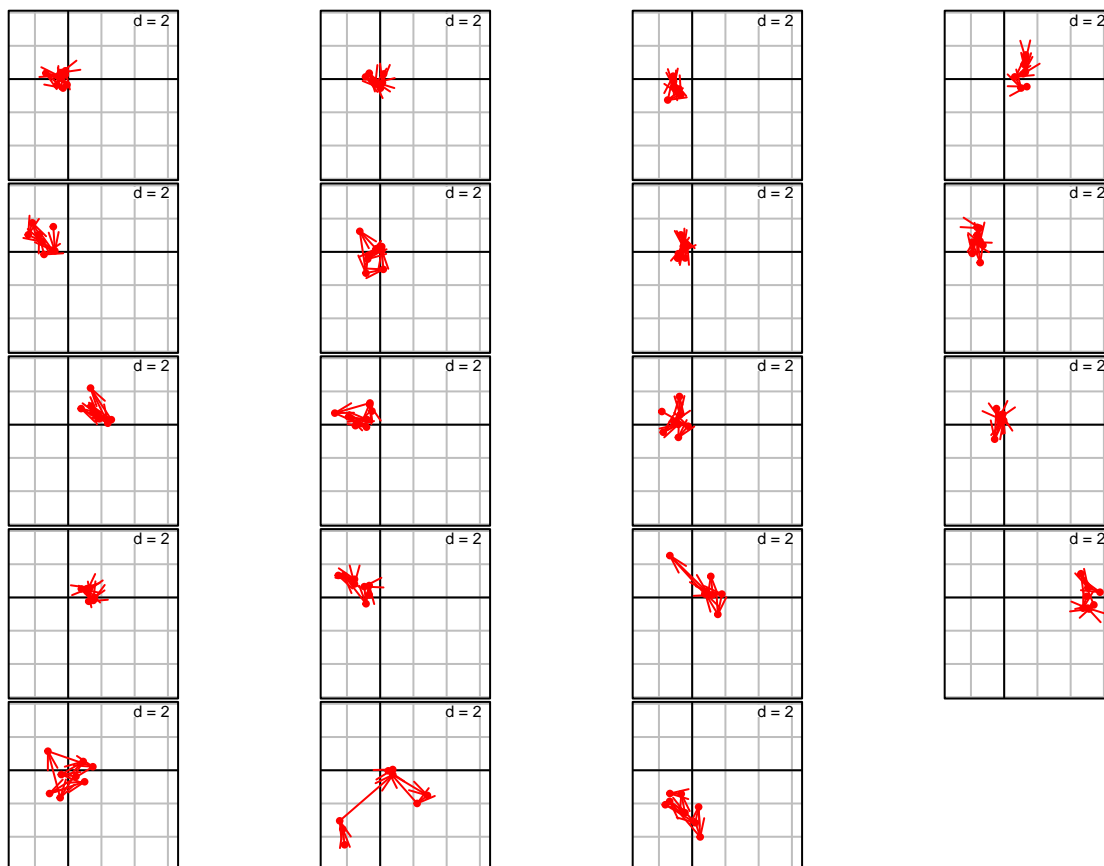
Como resumen, podemos comentar que el índice de esclarecimiento es constante, salvo en los años 2019 y 2020, principalmente el primero, vemos las CCAA más desagrupadas hacia la zona derecha del gráfico, zona contraria a donde apuntan las variables.

En cuanto a CLM vemos como la trayectoria es bastante estable, con movimientos intermitentes a lo largo del periodo, y se observa como en 2022 el IE de esta comunidad ha disminuido para los delitos que se encuentran en el segundo cuadrante (SEX, HON, IDS y CPI/I)

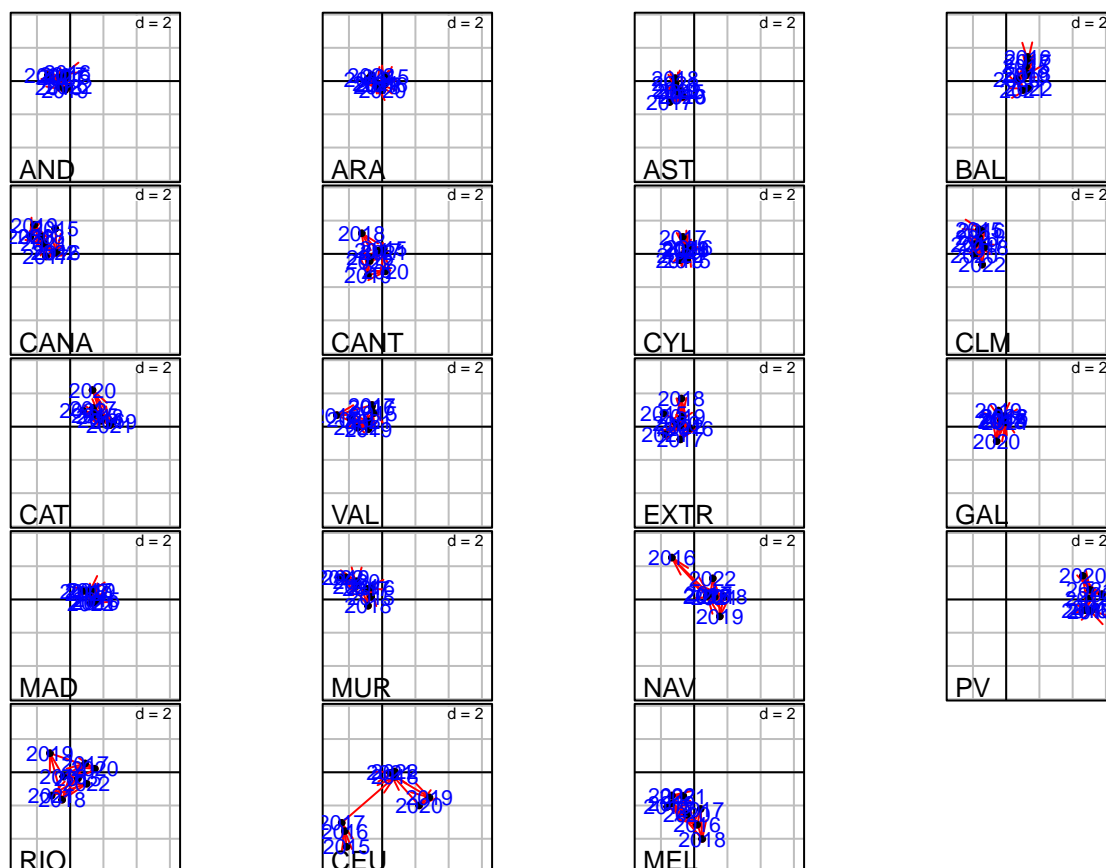
```
la2.2=s.label(pta2$Tli,facets=pta2$TL[, 2], psub.cex=1.5, labels=pta2$TL[,1], plabels.cex=1.2, plabels.
```



```
tr2=s.traject(pta2$Tli, facets=pta2$TL[, 2], plabels.cex = 0, psub.cex=0,col="red", plot=TRUE)
```



```
s4=superpose(tr2,la2.2)
plot(s4)
```



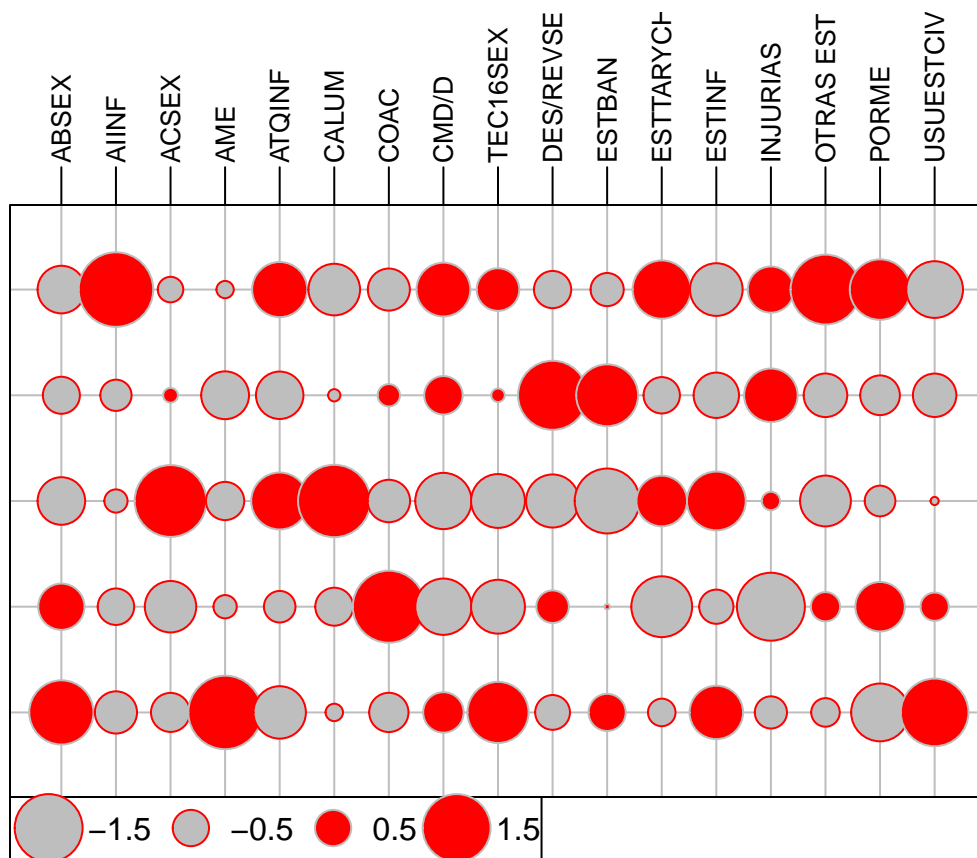
D. Análisis multivariante de la situación actual (2022) de ciberdelincuencia en la provincias de CLM en cuanto a su tipología penal

Como parte final de este análisis, nos centraremos en Castilla-La Mancha, donde intentaremos obtener la máxima información relativa a los hechos conocidos, a diferencia del resto del análisis, sobre las distintas tipologías delictivas en el año 2022.

Antes de mostrar presentar los resultados, cabe mencionar que algunas tipologías como el exhibicionismo, ha sido eliminadas del estudio dado que no existe ningún caso registrado o hay una cantidad ínfima de ellos. Este hecho ya lo comentamos en el apartado MATERIAL (en material y métodos). También se han calculado las tasas de cada tipología delictiva en cada provincia.

Primero de todos, mostraremos una representación gráfica de los datos originales (pero estandarizados).

```
library(readxl)
datosclm = read_excel("datosclm.xlsx")
datos1=datosclm[2:18]
table.value(scale(datos1), symbol="circle", col=c("grey", "red"))
```



En la imagen anterior, encontramos una distribución más o menos homogénea en cuanto a la comisión de los delitos en las distintas provincias, es decir, no hay una provincia en la que ocurra en mayor medida todos los delitos o una gran parte de ellos, sino que en cada provincia destacan algunas tipologías delictivas que en otras no.

El análisis JK-Biplot de la matriz de tasas (provincias x tipologías delictivas), proporcionó los siguientes resultados (ver figura 20 y la imagen siguiente)

La absorción de inercia en el primer plano alcanza un porcentaje del 63,7%, y con las tres primeras dimensiones tenemos un 84,41%. En este análisis retendremos la información relativa a estas tres primeras dimensiones.

Eigenvalues

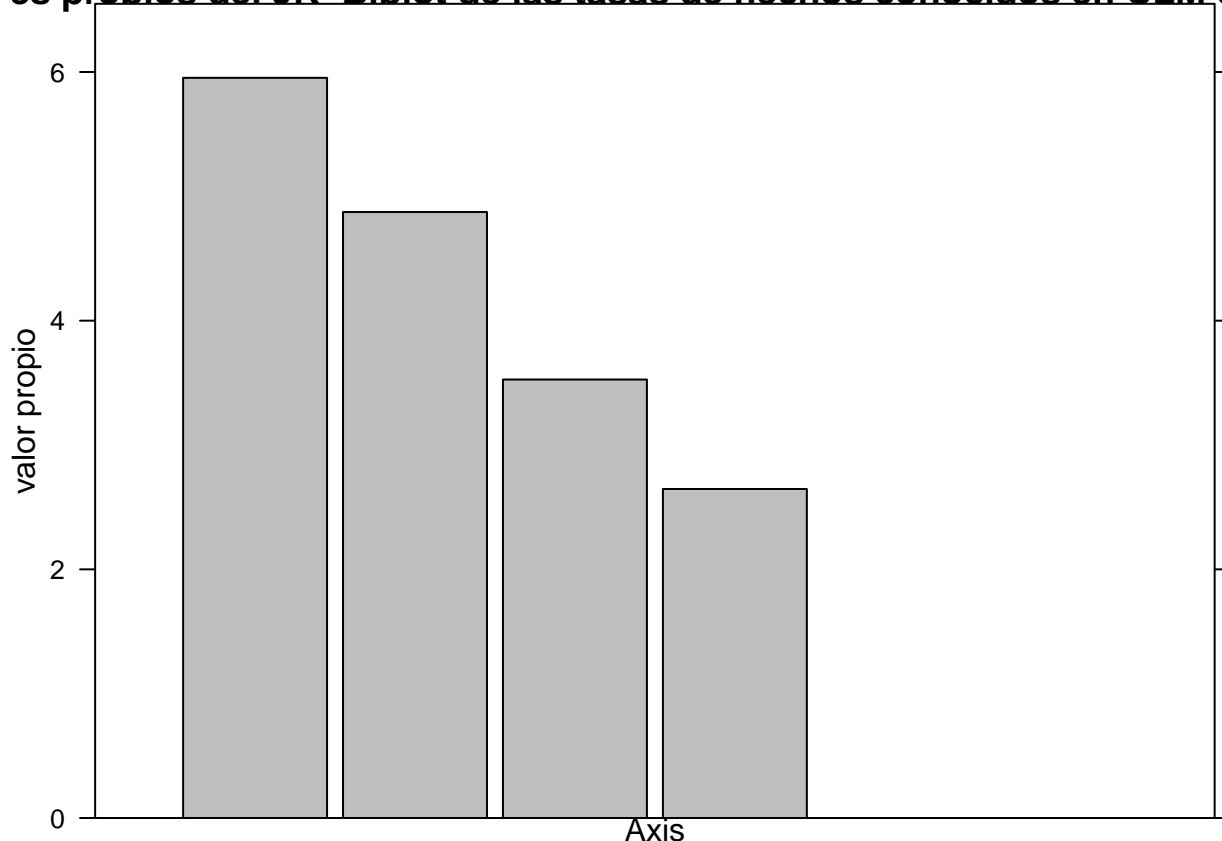
	Dim.1	Dim.2	Dim.3	Dim.4
<u>Variance</u>	5.951	4.876	3.522	2.651
<u>% of var.</u>	35.005	28.681	20.721	15.594
<u>Cumulative % of var.</u>	35.005	63.685	84.406	100.000

Figure 19: Resultados del análisis JK-Biplot de las tareas de hechos conocidos en CLM en 2022

```
pca1=dudi.pca(datos1,scale=TRUE,scannf=FALSE, nf = 5)
```

```
screplot(pca1, main= "Valores propios del JK-Biplot de las tasas de hechos conocidos en CLM en 2022", y
```

es propios del JK-Biplot de las tasas de hechos conocidos en CLM en



```
pca1$eig
```

```
## [1] 5.953220 4.873995 3.526530 2.646256
```

A continuación, obtendremos las calidades de representación (\cos^2) tanto de las variables como de los individuos de las distintas dimensiones, para así poder hacer un mejor análisis. Empecemos por las variables (tipologías delictivas): (Ver los siguientes gráficos):

```
library(FactoMineR)
```

```
##
## Attaching package: 'FactoMineR'
## The following object is masked from 'package:ade4':
##
##      reconst
```

```
library(factoextra)
```

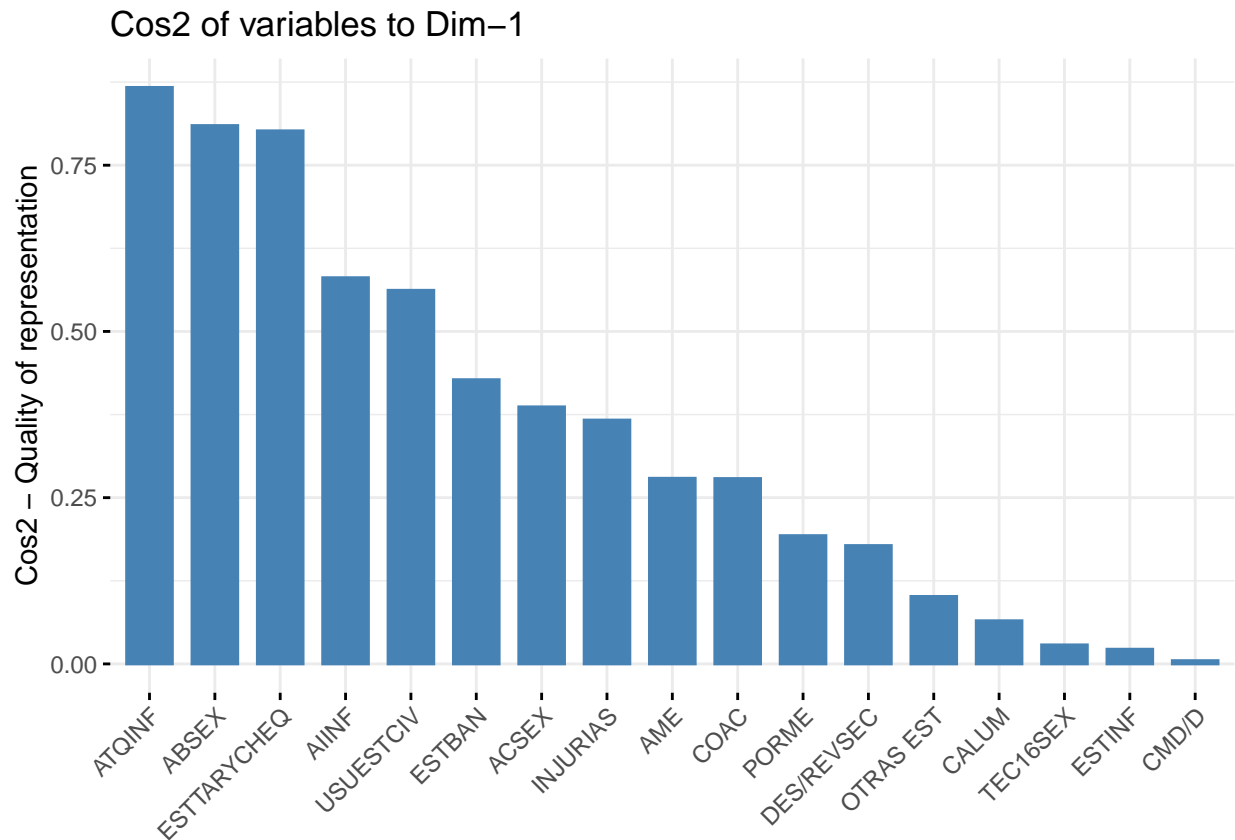
```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
library(Factoshiny)
```

```
## Warning: package 'Factoshiny' was built under R version 4.3.3
## Loading required package: shiny
## Loading required package: FactoInvestigate
## Warning: package 'FactoInvestigate' was built under R version 4.3.3
```

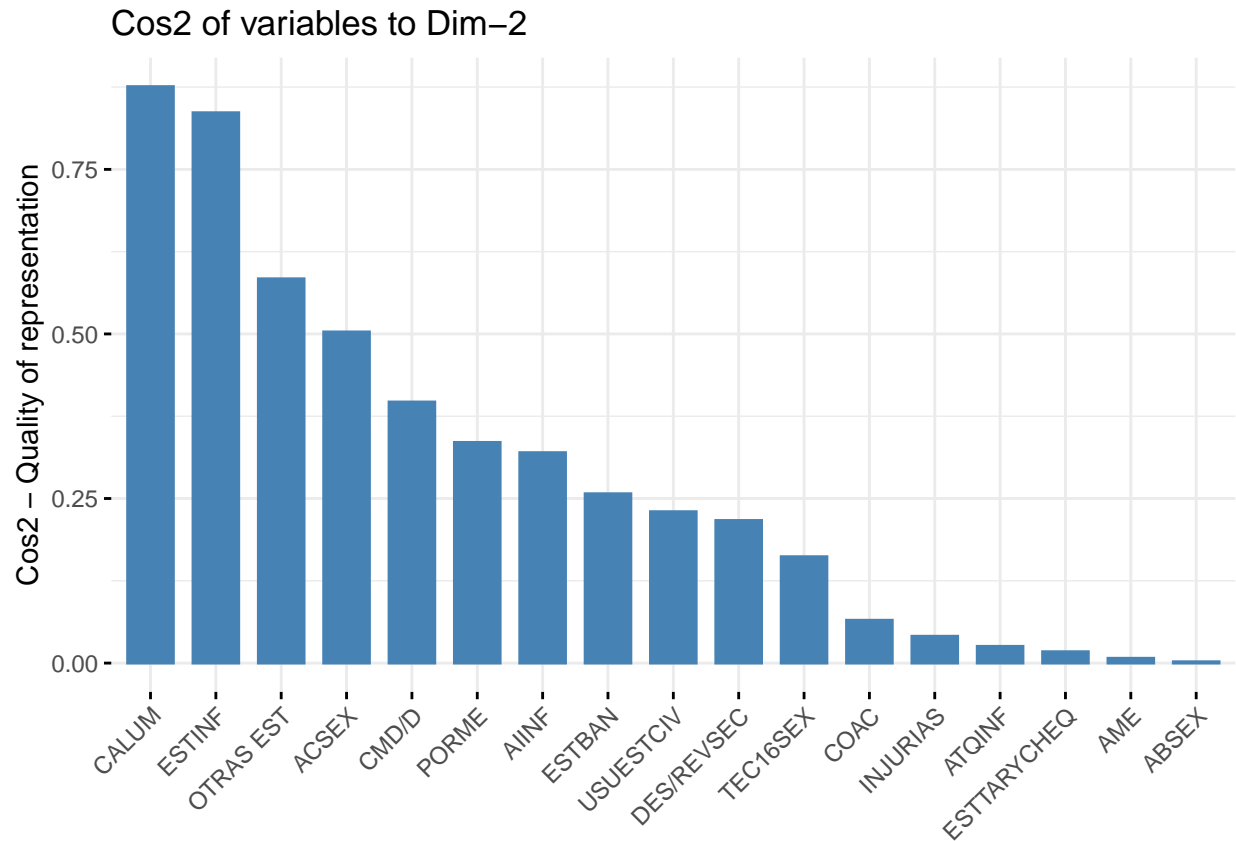


```
res.PCA = PCA(datos1,graph=FALSE)
fviz_cos2(res.PCA,
  choice = c("var"),
  axes = 1,
  fill = "steelblue",
  color = "steelblue",
  sort.val = c("desc"),
  top = Inf,
  xtickslab.rt = 45)
```



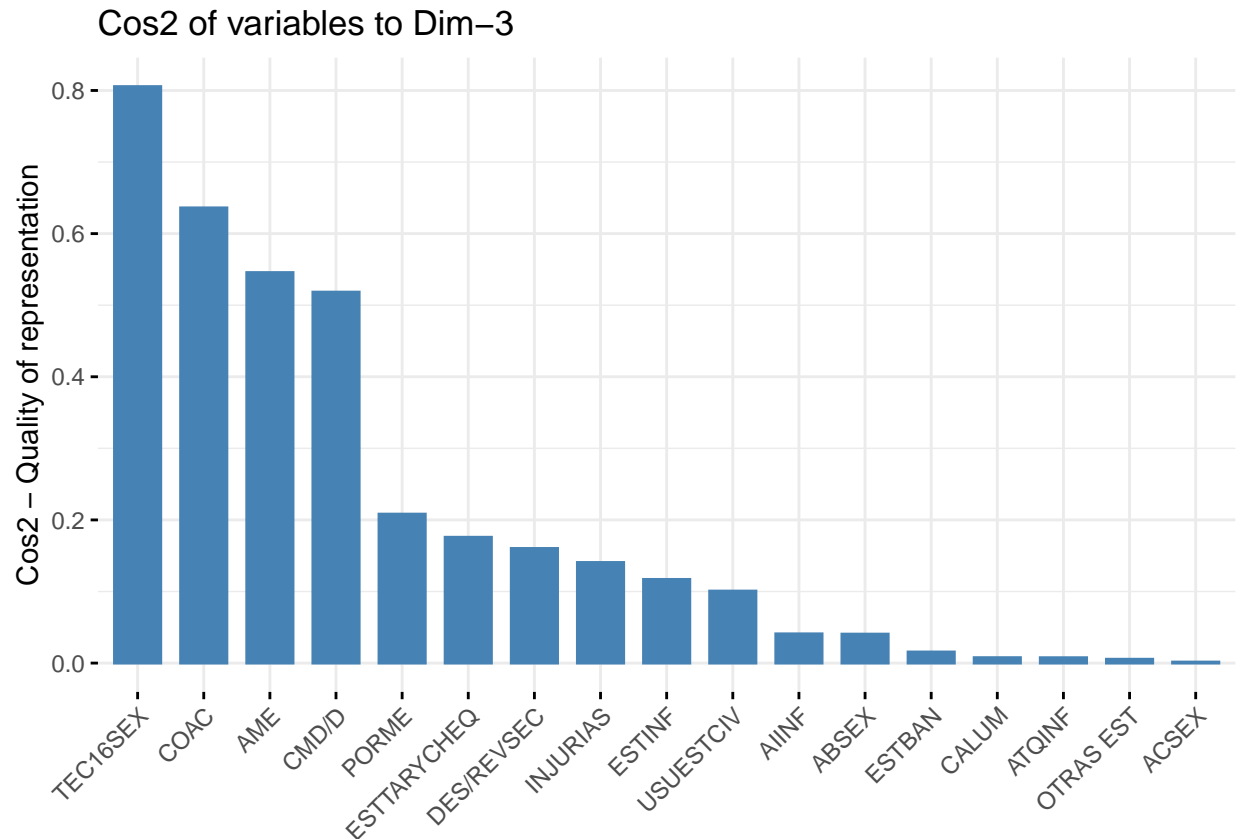
El gráfico anterior nos muestra que los Ataques informáticos (ATQINF), los Abusos sexuales (ABSEX) y son las Estafas con tarjeta de crédito, débito y cheques de viaje (ESTTARYCHEQ) son las tipologías que tienen mayor calidad de representación en la 1ª dimensión, siendo casi nula en el caso de las Calumnias (CALUM), Estafas informáticas (ESTINF), Corrupción de menores/con discapacidad/diversidad funcional (CMD/D) y Delito de contacto mediante tecnología con menores de 16 años con fines sexuales (TEC16SEX).

```
fviz_cos2(res.PCA,
  choice = c("var"),
  axes = 2,
  fill = "steelblue",
  color = "steelblue",
  sort.val = c("desc"),
  top = Inf,
  xtickslab.rt = 45)
```



Calumnias (CALUM) y Estafas informáticas (ESTINF) son las que presentan la mejor calidad de representación en esta 2ª dimensión. Por otra parte, Coacciones (COAC), Injurias, Estafas con tarjeta de crédito, débito y cheques de viaje (ESTTARYCHEQ), Amenazas (AME) y Abusos sexuales (ABSEX), aparecen muy mal representadas. (Ver gráfico anterior)

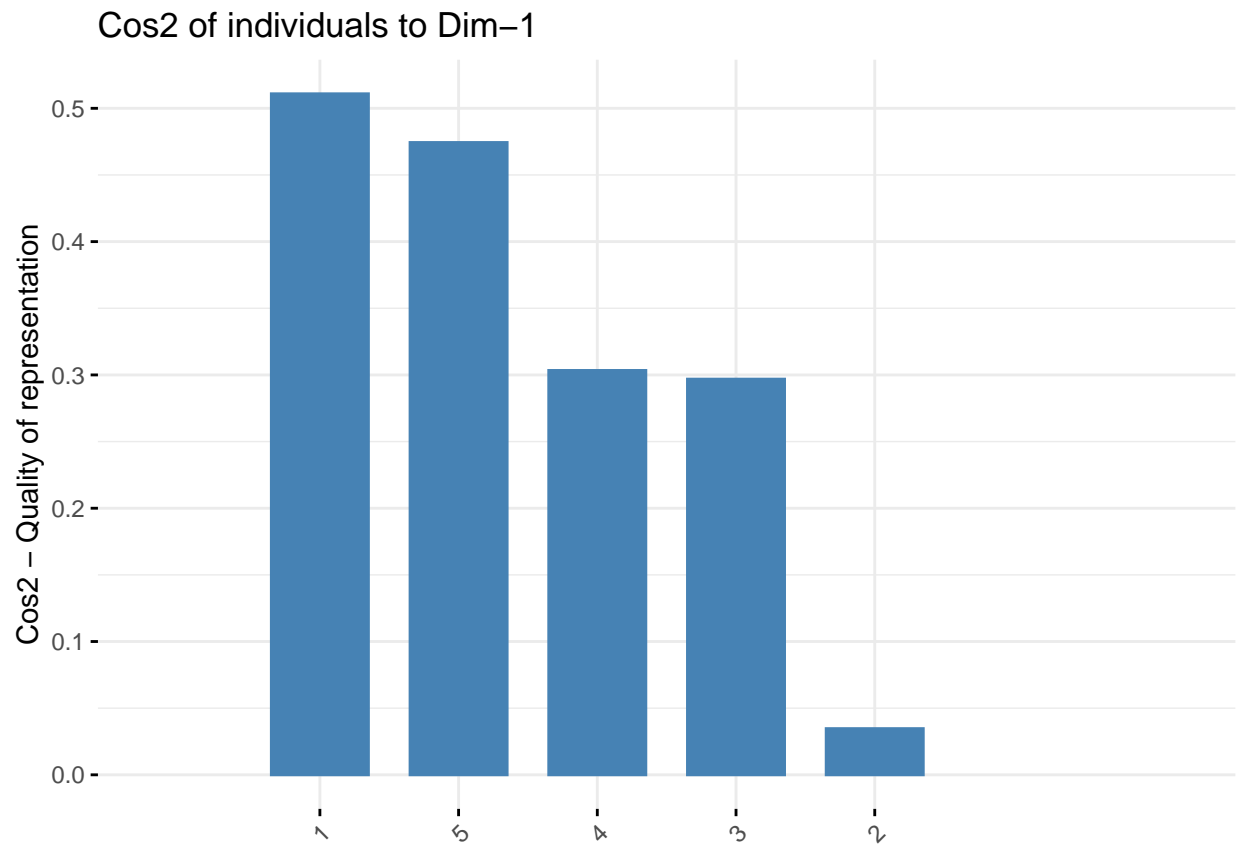
```
fviz_cos2(res.PCA,
  choice = c("var"),
  axes = 3,
  fill = "steelblue",
  color = "steelblue",
  sort.val = c("desc"),
  top = Inf,
  xtickslab.rt = 45)
```



Delito de contacto mediante tecnología con menores de 16 años con fines sexuales (TEC16SEX) y Coacciones (COAC) son las tipologías que presentan mayor correlación con la 3ª dimensión, siendo una gran cantidad de ellas la que presentan escasa o nula. (Ver gráfico anterior)

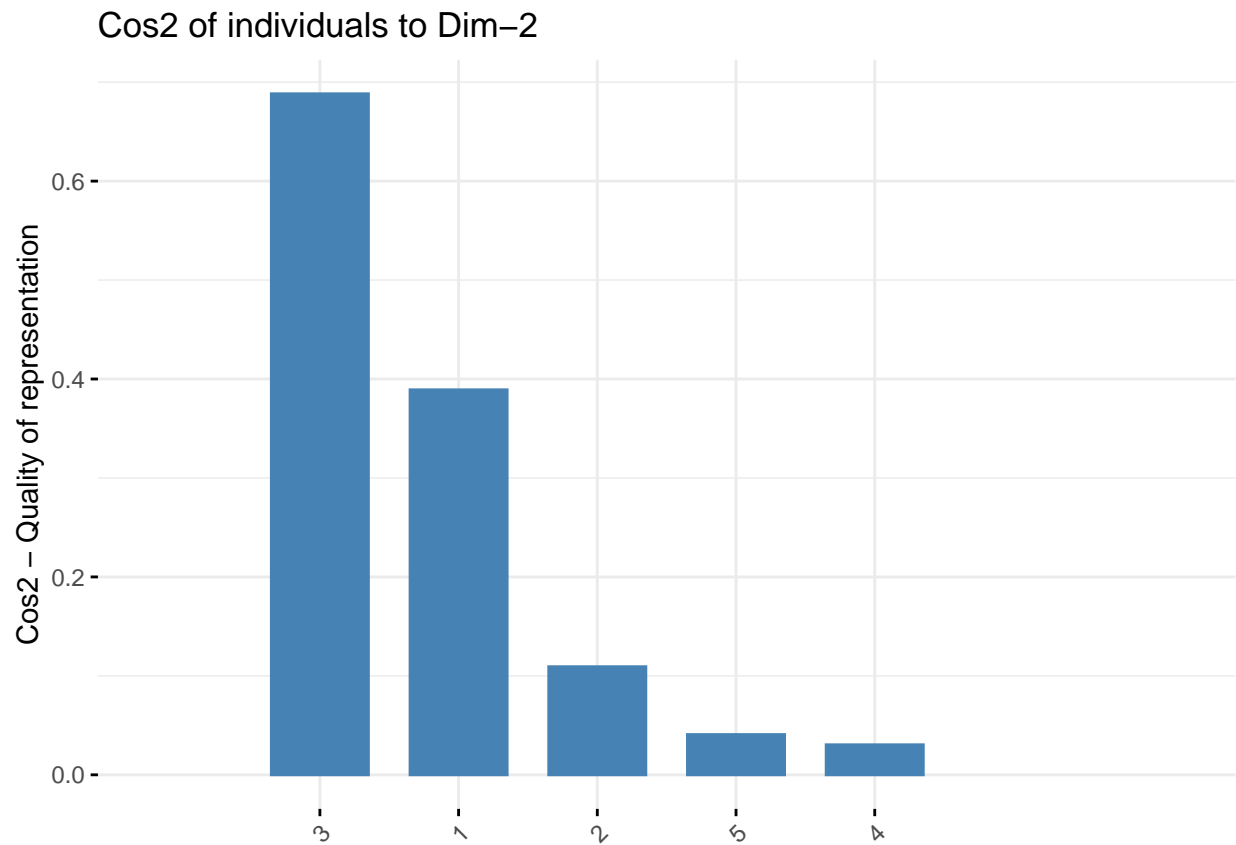
Si analizamos las calidades de representación para las provincias: (Ver siguientes 3 gráficos)

```
fviz_cos2(res.PCA,
  choice = c("ind"),
  axes = 1,
  fill = "steelblue",
  color = "steelblue",
  sort.val = c("desc"),
  top = Inf,
  xtickslab.rt = 45, xlim=c(0,7))
```



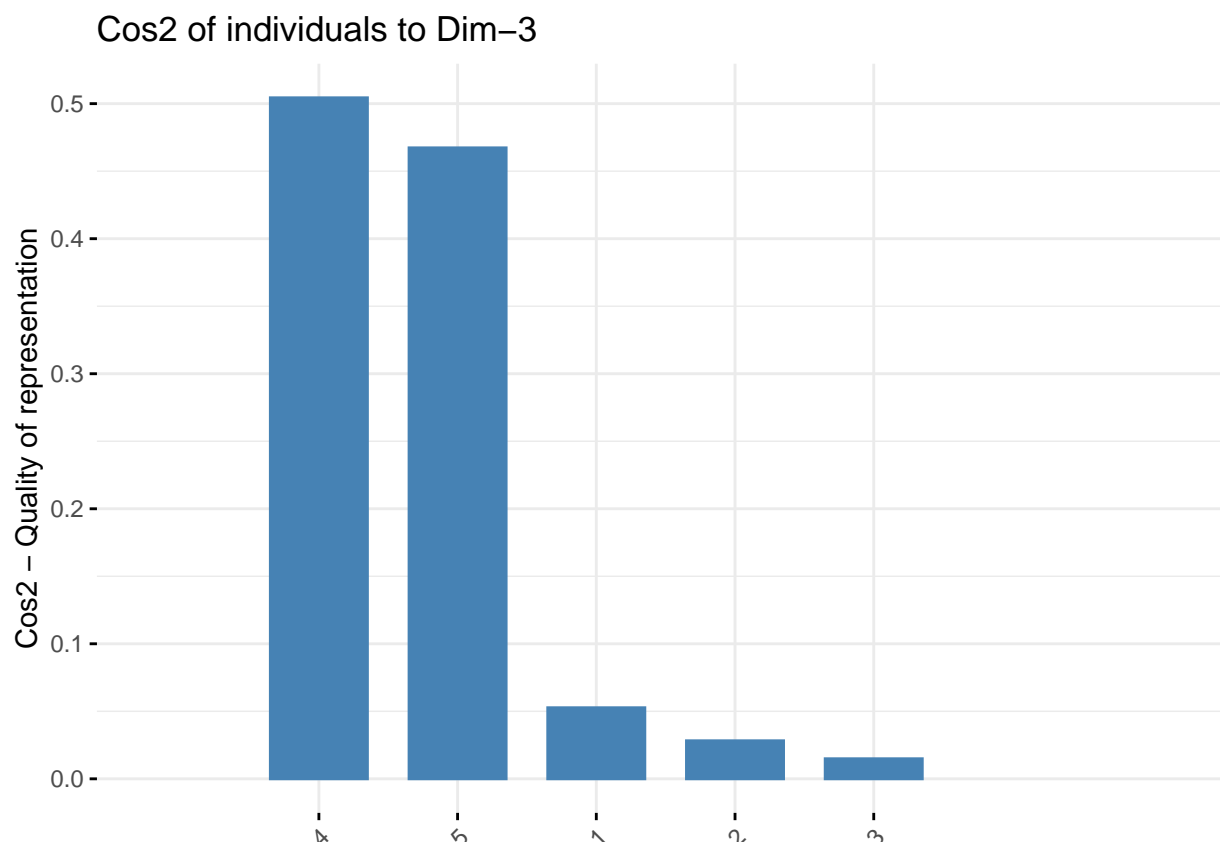
En este gráfico, observamos que Albacete y Toledo con las provincias que tienen una mayor correlación con la primera dimensión y Ciudad Real la que menor.

```
fviz_cos2(res.PCA,
  choice = c("ind"),
  axes = 2,
  fill = "steelblue",
  color = "steelblue",
  sort.val = c("desc"),
  top = Inf,
  xtickslab.rt = 45, xlim=c(0,7))
```



En cuanto a la 2º dimensión, observamos en este gráfico que Cuenca es la provincia que mejor viene representada, siendo Toledo y Guadalajara las que peor.

```
fviz_cos2(res.PCA,  
  choice = c("ind"),  
  axes = 3,  
  fill = "steelblue",  
  color = "steelblue",  
  sort.val = c("desc"),  
  top = Inf,  
  xtickslab.rt = 45, xlim=c(0,7))
```



En la 3ª dimensión, son Guadalajara, seguida de Toledo, las provincias con mayor correlación. El resto de ellas vienen muy mal representadas. (Ver gráfico anterior).

Por último, analicemos los resultados obtenidos del JK biplot, en el que recordemos que están mejor representadas los individuos, que en nuestro caso son las provincias de Castilla la Mancha.

Comencemos por el primero de todos, correspondiente a los 2 primeros ejes principales, comentaremos los resultados más relevantes: (Ver figura 21)

Tenemos una pareja de tipologías que son las que más correlación presentan en comparación con el resto, y estas son las Estafas con tarjeta de crédito, débito y cheques de viaje (ESTTARYCHEQ), junto a los Ataques informáticos (ATQINF). (Ver figura 21)

Respecto a la información que aportan, el Abuso sexual (ABSEX) en cuanto al 1º eje, es una de las variables que más información aporta de todo el conjunto de datos. A este eje podemos sumarle las Estafas con tarjeta de crédito, débito y cheques de viaje (ESTTARYCHEQ) y los Ataques informáticos (ATQINF), ambos siendo también de las variables que más variabilidad aportan, y las Amenazas (AME), la cual no es tan importante como el resto en cuanto a la información. (Ver figura 21)

Respecto al 2º eje, cabe destacar las Calumnias (CALUM) y las Estafas informáticas (ESTINF) por su gran aportación a dicho eje y en cuanto a su variabilidad, seguido de la Corrupción de menores/con discapacidad/diversidad funcional (CMD.D), el cual aporta también bastante información. (Ver figura 21)

Las 3 provincias que vienen bien representadas por estas dos dimensiones: Cuenca, Albacete y Toledo. Entre ellas no encontramos similitud, ya que cada una se sitúa en una parte distinta del gráfico, y a pesar de que Toledo, Ciudad Real y Guadalajara sí que estén aproximadas, salvo Toledo, las demás están mal representadas. (Ver figura 21)

Si atendemos a estas provincias, en Toledo es la provincia donde más ocurre Abusos sexuales (ABSEX), donde las cifras están por encima de la media en Calumnias (CALUM) y Estafas informáticas (ESTINF) y

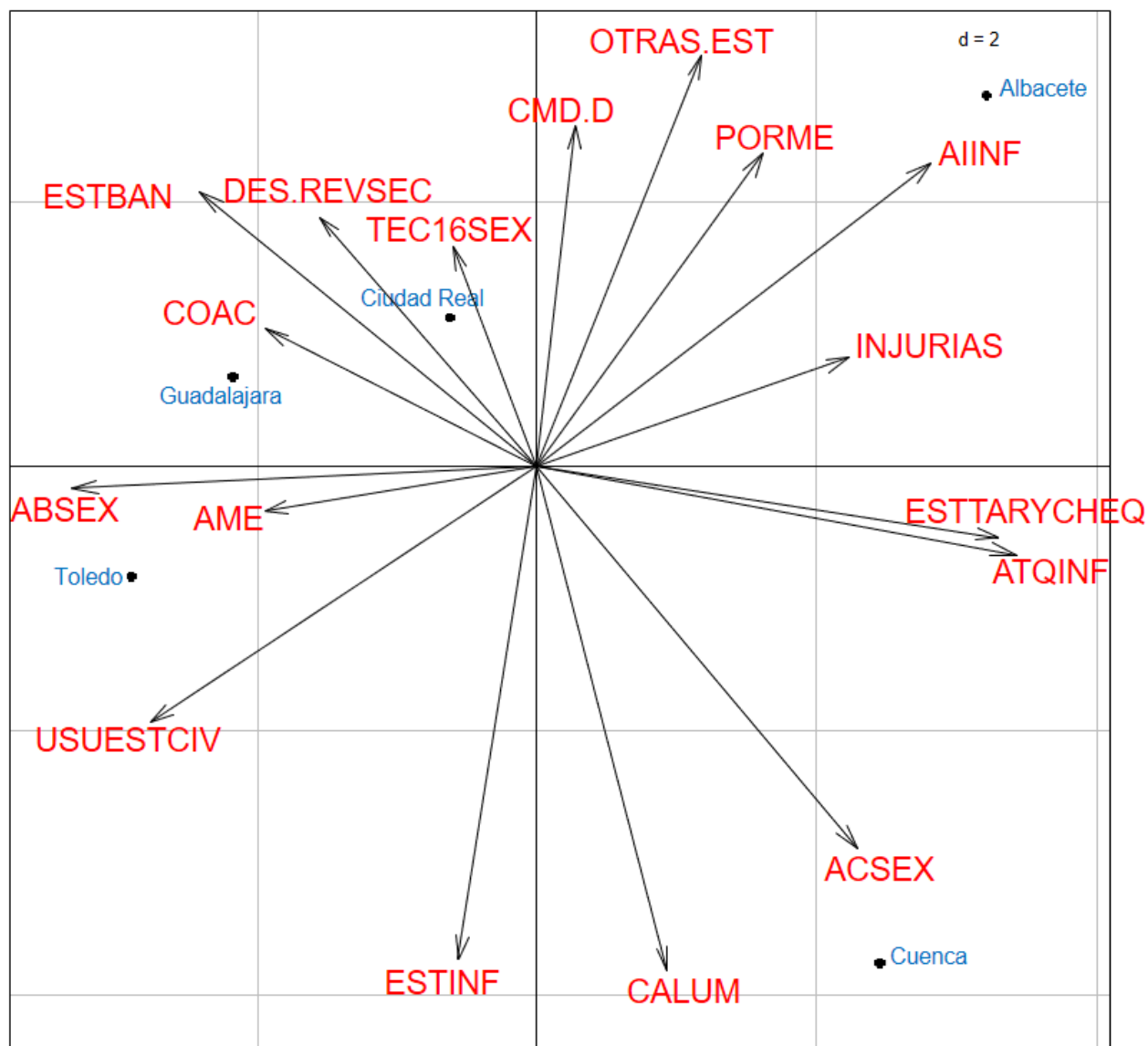


Figure 20: Biplot de las tipologías delictivas en Castilla-La Mancha en 2022. 1º y 2º dimensión

en donde menos ocurren Ataques informáticos (ATQINF) y Estafas con tarjeta de crédito, débito y cheques de viaje (ESTTARYCHEQ).

En Cuenca es donde más ocurren Calumnias (CALUM), Estafas informáticas (ESTFINF), Ataques informáticos (ATQINF) y Estafas con tarjeta de crédito, débito y cheques de viaje (ESTTARYCHEQ) y donde menos ocurren Otras estafas (OTRAS. EST). Y Albacete es el lugar donde más ocurren Otras estafas (OTRAS. EST), donde, ocurren también más Ataques informáticos (ATQINF) y Estafas con tarjeta de crédito, débito y cheques de viaje (ESTTARYCHEQ), donde ocurren Calumnias (CALUM) y Estafas informáticas (ESTINF) por debajo de la media, y donde ocurren muy pocos Abusos sexuales (ABSEX). (Ver figura 21)

Pasemos a analizar el siguiente gráfico en el que se tiene en cuenta la 1ª y la 3ª dimensión. (Ver figura 22)

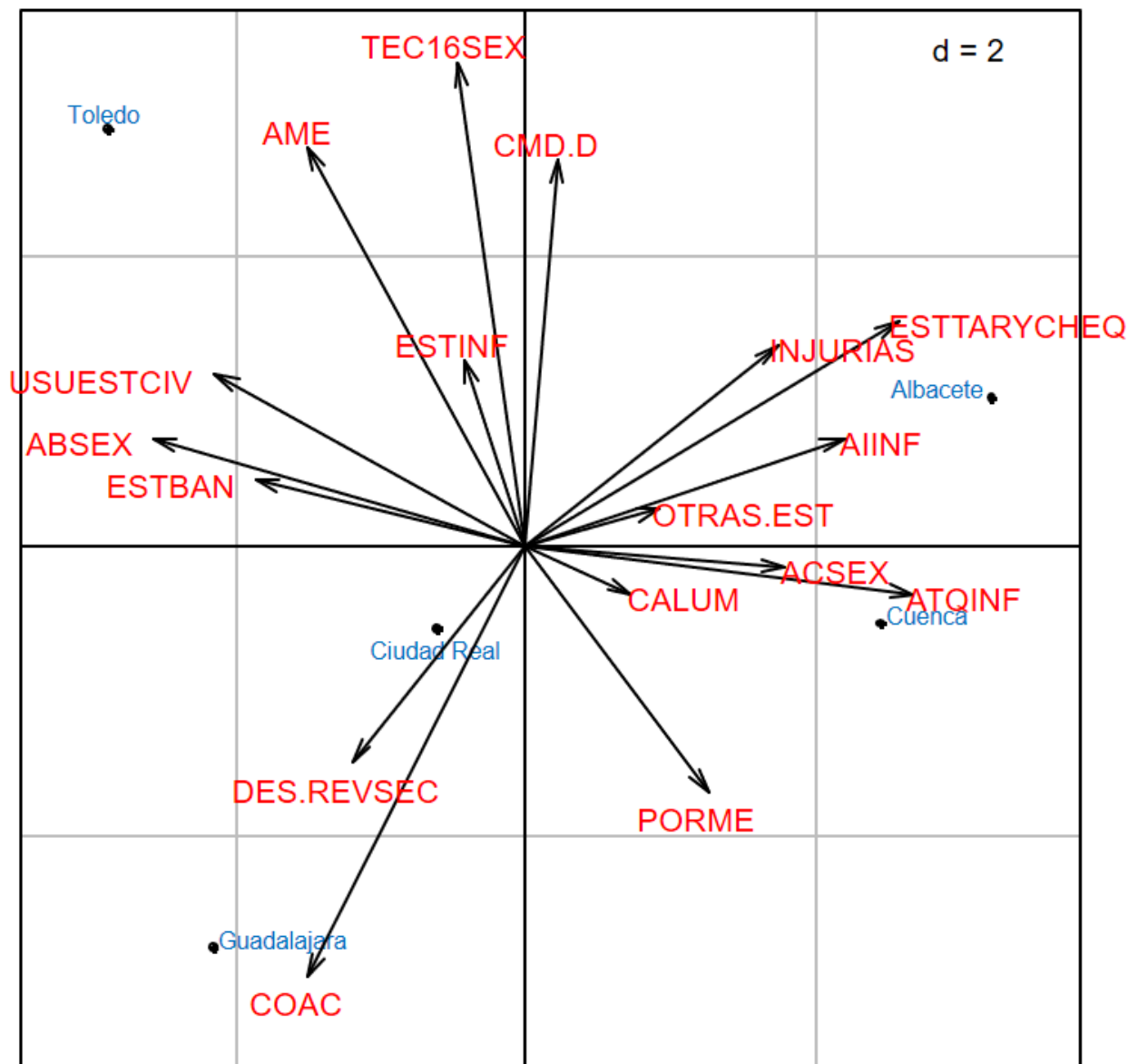


Figure 21: Biplot de las tipologías delictivas en Castilla-La Mancha en 2022. 1ª y 3ª dimensión

Nos encontramos con algunas tipologías, como las Calumnias (CALUM), Otras estafas (OTRAS. EST) y las Estafas informáticas (ESTINF), que aparecen muy mal representadas, sin embargo, en la figura 47, aparecían muy bien representadas. Eso quiero decir que no pertenecen al 1º eje, siendo Ataques informáticos (ATQINF)

y Abuso sexual (ABSEX) las que sí lo hacen. Por el contrario, las tipologías de Coacciones (COAC) y Delito de contacto mediante tecnología con menores de 16 años con fines sexuales (TEC16SEX), adquieren una mayor importancia ahora, cuando anteriormente eran las que menor variabilidad presentaban, lo cual es indicativo de que pertenecen a la 3ª dimensión (Ver figura 22).

Coacciones (COAC), Delito de contacto mediante tecnología con menores de 16 años con fines sexuales (TEC16SEX) y Corrupción de menores/con discapacidad/diversidad funcional (CMD.D) son tipologías que aportan al 3º eje, mientras que Ataque informático (ATQINF) en la tipología que más aporta al 1º. (Ver figura 22)

En cuanto al 3º eje, nos encontramos que divide las provincias igual que en la figura 31, sin embargo, este 1º eje ahora nos separa Toledo de Guadalajara y Ciudad Real, mientras que nos junta más Albacete y Cuenca. Si atendemos a las calidades de representación (Ver figuras 44 y 46), Toledo, Albacete y Guadalajara son las provincias que aparecen bien representadas, estando cada una en una parte distinta del biplot. Cabe mencionar aquí que en el anterior caso (biplot de las dimensiones 1-2) Guadalajara y Toledo estaban al lado, mientras que ahora, cambiando la perspectiva de las dimensiones, en concreto la 2ª por la 3ª, aparecen totalmente separadas. Por ende, en cuanto a estas 3 provincias, no encontramos similitudes. (Ver figura 22)

En Guadalajara es la provincia donde se comenten más delitos de Coacciones (COAC), los Ataques informáticos (ATQINF) ocurren por debajo de la media, y es donde menos ocurren los Delitos de contacto mediante tecnología con menores de 16 años con fines sexuales (TEC16SEX); en Toledo es donde más ocurren los casos de Delitos de contacto mediante tecnología con menores de 16 años con fines sexuales (TEC16SEX), de Corrupción de menores/con discapacidad/diversidad funcional (CMD.D) y de Abusos sexuales (ABSEX), ocurriendo estos últimos también bastante por Guadalajara, los delitos de Coacciones (COAC) ocurren por debajo de la media, y es el lugar más seguro para los Ataques informáticos (ATQINF); y en Albacete, los que más se cometen los Ataques informáticos (ATQINF) y las Estafas con tarjeta de crédito, débito y cheques de viaje (ESTTARYCHEQ), los Delitos de contacto mediante tecnología con menores de 16 años con fines sexuales (TEC16SEX) y los de Corrupción de menores/con discapacidad/diversidad funcional (CMD.D) ocurren por encima de la media, es donde más se ocasionan los Ataques informáticos (ATQINF), las Coacciones (COAC) se realizan por debajo de la media, y es el lugar más seguro respecto a los Abusos sexuales (ABSEX). (ver figura 22)

Analicemos, por último, la información en el plano 2-3. (Ver figura 23)

Amenazas (AME) y Coacciones (COAC) son las tipologías características de 3º eje, aportando, además, bastante información. Y Acoso sexual (ACSEX), Calumnias (CALUM), Estafas bancarias (ESTBAN) y Otras estafas (OTRAS.EST) de 2º eje. (Ver gráficos de calidades de representación de variables de 2º y 3º dimensión). En este caso Delito de contacto mediante tecnología con menores de 16 años con fines sexuales (TEC16SEX) no aparece tan pegada al eje de la 3ª dimensión debido a que también tiene algo de correlación con la 2ª, aunque bastante inferior, por ende, se inclina más hacia la 3ª. Sin embargo, en el caso del biplot de las dimensiones 1-3, la correlación de dicha tipología con la 1ª dimensión era casi nula, por ende, aparecía mucho más pegado al eje de la 3ª dimensión.

Además, nos encontramos con tipologías como Abuso sexual (ABSEX) y Ataques informáticos (ATQINF) que no están bien representadas en este plano ya que pertenecen a la 1ª.

Esta 3ª dimensión nos diferencia a las provincias de Toledo y Albacete del resto, mientras que la 2ª dimensión, nos separa Cuenca de Albacete y Ciudad Real. En cuanto a las calidades de representación, Cuenca, Guadalajara, Toledo y Albacete en cierta medida, son las mejor representadas. Todas ellas se encuentran totalmente separadas las unas de las otras, indicativo de que no guardan ningún tipo de similitud.

Respecto a la provincia de Toledo, nos encontramos con que es donde más ocurren Amenazas (AME) y menos Coacciones (COAC), Acosos sexuales (ACSEX) y Calumnias (CALUM) ocurren por encima de la media y Estafas bancarias (ESTBAN) y Otras estafas (OTRAS.EST) por debajo. En cuanto a Cuenca, es el lugar donde más ocurren Abusos sexuales (ABSEX) y Calumnias (CALUM) y donde menos Estafas bancarias (ESTBAN) y Otras estafas (OTRAS.EST), Coacciones (COAC) ocurren por encima de la media y Amenazas (AME) por debajo. En Guadalajara, las Coacciones (COAC) son los delitos más acontecidos y las Amenazas (AME) el que menos, las Estafas bancarias (ESTBAN) y Otras estafas (OTRAS.EST) ocurren

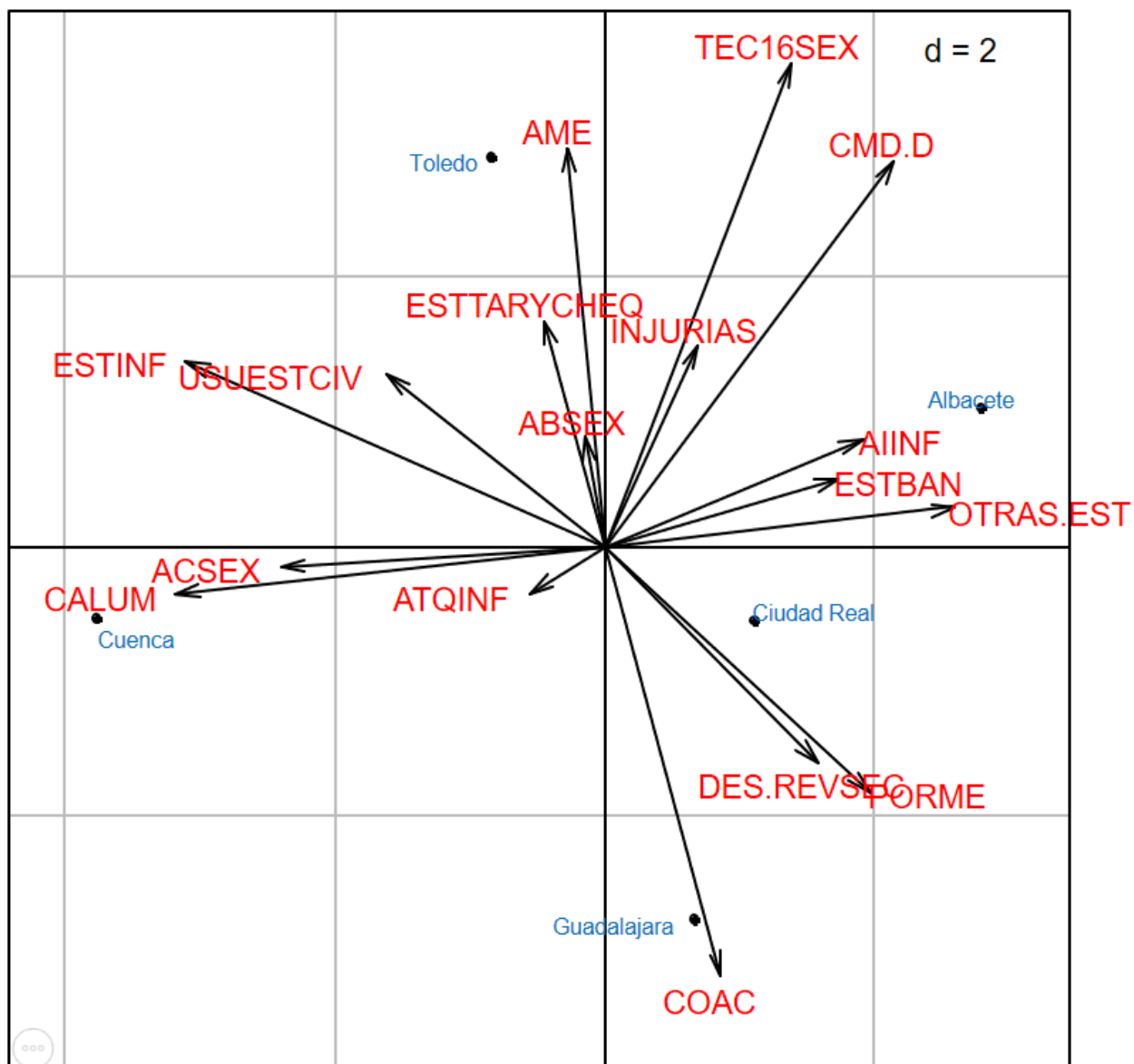


Figure 22: Biplot de las tipologías delictivas en Castilla-La Mancha en 2022. 2º y 3º dimensión

por encima de la media y los Acosos sexuales (ACSEX) y Calumnias (CALUM) por debajo. Por último, en Albacete es donde más Estafas bancarias (ESTBAN) y Otras estafas (OTRAS. EST) ocurren y donde menos Abusos sexuales (ABSEX) y Calumnias (CALUM), Las Coacciones (COAC) acontecen por debajo de la media y las Amenazas (AME) por encima.

V. CONCLUSIONES

Una vez realizados los análisis pertinentes, podemos sacar las siguientes conclusiones:

1. A lo largo de este periodo de tiempo (2015 – 2022), los ciberdelitos han ido aumentando paulatinamente, lo cual es indicativo de tres posibilidades: que cada vez haya más medios y nuevas formas empleadas por los ciberdelincuentes para cometerlos, que las personas aún no somos conscientes de dicho peligro y eso nos supone una gran vulneración, o ambas situaciones a la vez.

Es, por ende, necesario un plan de prevención que consiga, ya no disminuir los casos, sino pausar ese crecimiento, ya que cada vez, estos delitos, como hemos observado, abarcan un mayor porcentaje respecto al total.

2. Los años 2021 y 2022 aparecen claramente separados en la interestructura de los hechos conocidos. Estos años se corresponden con la “post pandemia” indicando ese aumento drástico de los ciberdelitos que hace que se encuentre diferenciado del resto de años. Completamente al contrario ocurre en el índice de esclarecimiento donde los 3 primeros años del periodo (2015, 2016 y 2017) son los que aparecen separados claramente del resto de años. Esto se debe a que con el transcurso del tiempo se ha ido mejorando y perfeccionando las técnicas que nos permiten esclarecer los delitos acontecidos permitiéndonos mejorar en gran medida las cifras.
3. Los grupos delictivos que más covarían entre sí, atendiendo al compromiso de los hechos delictivos son las Falsificación informáticas (FALIN), Delitos sexuales (SEX), Acceso e interpretación ilícita (AII), Contra la propiedad industrial/intelectual (CPI/I) e Inferencia en los datos y en el sistema (IDS), destacando en ellos las comunidades Baleares (BAL), Navarra (NAV), Madrid (MAD) y Castilla y León (CYL), principalmente. Más aislados encontramos los grupos de Delitos contra el Honor (HON) donde destaca las comunidades Ceuta (CEU) y Extremadura (EXT); las Amenazas y coacciones (AYC) destacando Canarias (CAN), Galicia (GAL), Castilla-La Mancha (CLM) y Andalucía (AND); y los Fraudes informáticos (FRAIN) destacando el País Vasco (PV). Respecto al índice de esclarecimiento, nos encontramos con dos grupos con gran covariación entre sí. El primero de ellos formado por los grupos delictivos de Acceso e interpretación ilícita (AII), Amenazas y coacciones (AYC), Falsificación informática (FALIN) y Fraude informático (FRAIN), donde son características Asturias (AST), Cantabria (CANT), Andalucía (AND), Aragón (ARA) y Castilla y León (CYL). El otro grupo que encontramos está formado por los grupos delictivos de Contra la propiedad industrial/intelectual (CPI/I), Interferencia en los datos y el sistema (IDS), Sexuales (SEX) y Contra el honor (HON), caracterizándose las comunidades Galicia (GAL), Canarias (CAN), Murcia (MUR) y Valencia (VAL).
4. En cuanto a la comunidad de Castilla-La Mancha:
 - 4.1. Atendiendo a los grupos penales, hemos observado que no destaca del resto de comunidades, es decir, se encuentra alrededor de la media de comisiones, tanto si nos centramos en los hechos conocidos como en los esclarecidos.
 - 4.2. Respecto a los hechos conocidos sí que observamos que presenta tasas algo mayores que la media en delitos contra el honor (HON) y Amenazas y Coacciones (AYC), en Delitos sexuales (SEX) y Falsificación Informática (FALIN) en prácticamente todos los años, y algo menores que la media en Fraude Informático (FRAIN) e Interferencia en los Datos y en el Sistema (IDS).
 - 4.3. En relación al índice de esclarecimiento, observamos que en el año 2022 ese índice disminuye para los grupos SEX, HON, IDS y CPLI, pero su sintonía general es constante a lo largo del periodo indicado.

4.4. En consonancia a las tipologías delictivas de ciberdelitos y las provincias de CLM, podemos apreciar que Cuenca se caracteriza por la comisión de Calumnias (CALUM) y Estafas informáticas (ESTINF); Guadalajara por las Coacciones (COAC); Toledo por los Abusos sexuales (ABSEX), las Amenazas (AME), los Delitos de contacto mediante tecnología con menores de 16 años con fines sexuales (TEC16SEX) y la Corrupción de menores/con discapacidad/diversidad funcional (CMD.D); y en Albacete los Ataques informáticos (ATQINF) y las Estafas con tarjeta de crédito, débito y cheques de viaje (ESTTARYCHEQ). Ciudad Real en cambio, no aparece bien representada en ninguna de las 3 dimensiones por lo que seguramente pertenezca a otra, por lo que no podemos sacar buenas conclusiones al respecto.

VI. BIBLIOGRAFÍA

- Abdi, H., Valentin, D. (2007). The STATIS method. In: N. Salkind (ed.), Encyclopedia of Measurement and Statistics (pp. 955–963).
- CEREZO DOMÍNGUEZ, A. I.; GARCÍA CORNEJO, R. (2020). La ciberdelincuencia en España: un estudio basado en las estadísticas policiales, Revista Peruana de Ciencias Penales, 1(34), 91–106. <https://doi.org/10.56176/rpcp.34.2022.3>
- de Europa, C. (2001). Convenio sobre la Ciberdelincuencia. Consultado en: http://www.agpd.es/portalwebAGPD/canaldocumentacion/legislacion/consejo_europa/convenios/common/pdfs/Convenio_Ciber.pdf.
- de España, G. (1995). Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal. Boletín Oficial del Estado, 281.
- Dirección General de Coordinación y Estudios, Secretaría de Estado de Seguridad. (2022). Informe sobre la cibercriminalidad en España 2022. Ministerio del Interior, Gobierno de España. NIPO 126-20-021-2.
- Gabriel, K. R. (1971). The biplot-graphic display of matrices with application to principal component analysis. Biometrika, 58(3), 453-467.
- Galindo, M. P. (1985). Contribuciones a la representación simultánea de datos multidimensionales. Tesis doctoral. Universidad de Salamanca.
- Galindo Villardon, M. (1986). Una alternativa de representación simultánea: HJ-Biplot. Qüestiió. 1986, vol. 10, núm. 1.
- Hotelling, H. (1933) Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24:417-441,498-520.
- Hotelling, H. (1936). Simplified calculation of principal components. Psychometrika, 1: 27-35.
- Jaffrenou, P.A. (1978). Sur L'Analyse des Familles Finies de Variables Vectorielles: Bases Algébriques et Applications à la Description Statistique. [Thèse de Troisième Cycle]. Université de Lyon, France.
- Jiménez Rozas, J. (2022). Ciberdelincuencia: Evolución y relación con la actual situación de pandemia. Nuevas modalidades y nuevas problemáticas.
- Kroonenberg, P. M. (1989). The analysis of multiple tables in factorial ecology. III. Three-mode principal component analysis: "Analyse triadique complete". ACTA OECOL. (OECOL. GEN.), 10(3), 245-256.
- Lavit, C. (1994). The ACT (STATIS method). Computational Statistics and Data Analysis, 18:97–115
- Lavit, C., Escoufier, Y., Sabatier, R., & Traissac, P. (1994). The act (statis method). Computational Statistics & Data Analysis, 18(1), 97-119.
- Ministerio del Interior de España (2022). Portal Estadístico de Criminalidad. Recuperado de <https://estadisticasdecriminalidad.ses.mir.es/publico/portalestadistico/>

- PÉREZ GIL, Julio. Recensión del libro de Fernando Miró Llinares, El cibercrimen. Fenomenología y criminología de la delincuencia en el ciberespacio. Revista Electrónica de Ciencia Penal y Criminología (en línea). 2013, núm. 15-r2, p. r2:1-r2:05.
- Pearson, K. (1901). On lines and planes of closets fit to systems of points in the space. Philosophical Magazine, 2: 559-572.
- Ruiz, S. A. (2021). La ciberdelincuencia como fenómeno jurídico. Su tratamiento procesal. Revista Aequitas: Estudios sobre historia, derecho e instituciones, (18), 371-402. (file:///C:/Users/david/Downloads/Dialnet-LaCiberdelincuenciaComoFenomenoJuridicoSuTratamien-8101080.pdf)
- Sánchez, J. F. E. (2019). Ciberdelincuencia. Aproximación criminológica de los delitos en la red. La Razón histórica: revista hispanoamericana de historia de las ideas políticas y sociales, (44), 153-173. (file:///C:/Users/david/Downloads/LRH%2044.8%20(2).pdf)
- Thioulouse, J., & Chessel, D. (1987). Multi-table analysis of ecological factors. I. From description of ecosystem status to that of ecosystem function using triadic analysis. ACTA OECOL., OECOL. GEN., 8(4), 463-480.
- Vicente-Galindo, P. (2013). Análisis de tablas de tres vías: recientes desarrollos del STATIS. Trabajo fin de master, 1-65.
- Villardón, J. L. V. (2002). Análisis de componentes principales. Cataluña: UOC, Departamento de Estadística, 32.