



Práctica 2: análisis de correspondencias

ENUNCIADO 35 – SAMUR

Manuela Ibáñez | Ángel Vega | octubre 2025

Tabla de contenido

Enunciado de la práctica:.....	3
1. Descripción del conjunto de datos.....	4
2. Análisis de Correspondencias (AC) inicial: distritos vs. Hospitales.....	4
1) Tabla de contingencia principal.....	7
2) Tabla de distribución por franja horaria	8
3) Inertia and Chi-Square Decomposition.....	9
4) Summary Statistics	10
a) Row Points.....	10
b) Column Points	10
5) Partial Contributions to Inertia e Indices of Coordinates That Contribute Most row / columns	11
6) Squared cosines row/columns/supplementary.....	13
7) Gráfico factorial del análisis de correspondencias	14
3. Repetición del Análisis de Correspondencias sin el distrito Pozuelo.....	15
1) Inertia and Chi-Square Decomposition S.P.....	15
2) Summary Statistics for the Row Points S.P.....	15
3) Gráfico del Análisis de Correspondencias S.P	16
4. Agrupación de hospitales y distritos y comprobación de frecuencias.....	17
1) Análisis de frecuencias	18
5. Análisis final de Correspondencias final	19
1) Inertia and Chi-Square Decomposition FINAL.....	19
2) Summary Stats for the Row Points FINAL	19
3) Summary Stats for the Column Points FINAL.....	20
4) Partial Contributions to Inertia for the Row Points FINAL	20
5) Squared Cosines for the Row Points FINAL	21
6) Squared Cosines for the Row Points FINAL	21
6. Conclusión.....	22

Enunciado de la práctica:

El conjunto de datos **Samur** contiene información de las actuaciones del Samur a lo largo del año 2023. Las variables categóricas son la variable horaria, hora aviso, codigo, distrito, hospital, intervalo, mes.

Realizar un análisis de correspondencias donde las filas sean los distritos y las columnas el hospital. Utilizar la variable horaria como suplementaria y representarlas todas. Repetir el análisis eliminando el distrito de Pozuelo.

Pensar en considerar algunos hospitales por su baja presencia como suplementarios para que el análisis sea más robusto. Podéis agrupar también algunos distritos si lo consideras necesario y lo justificáis.

Incluir comentarios oportunos sobre las distintas modalidades suplementarias, incluyendo información del valor-test.

1. Descripción del conjunto de datos

Cargamos los datos del fichero de nuestra base de datos samur.sas7bdat y examinamos su estructura con el fin de identificar las variables categóricas relevantes al estudio.

Alphabetic List of Variables and Attributes					
#	Variable	Type	Len	Format	Informat
11	HORARIO	Char	11		
10	HORA_AVISO	Num	8		
3	codigo	Char	30	\$30.	\$30.
4	distrito	Char	30	\$30.	\$30.
2	hora_intervencion	Num	8	TIME8.	TIME8.
1	hora_llamada	Num	8	TIME8.	TIME8.
5	hospital	Char	30	\$30.	\$30.
8	intervalo	Num	8		
7	mes	Char	8		
9	tiempo_INTERVENCION	Num	8		
6	year	Num	8		

Data Set Name	TRABAJO.SAMUR	Observations	275148
Member Type	DATA	Variables	11
Engine	V9	Indexes	0
Created	19/10/2025 19:00:24	Observation Length	160
Last Modified	19/10/2025 19:00:24	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	WINDOWS_64		
Encoding	wlatin1 Western (Windows)		

Nuestra base de datos contiene 27.548 observaciones y 11 variables correspondientes a las actuaciones del SAMUR durante el año 2023.

Entre las variables se encuentran dos grupos de variables: categóricas (*HORARIO*, *codigo*, *distrito*, *hospital*, *intervalo* o *mes*) y numéricas (*hora_aviso*, *hora_llamada*, *hora_intervencion*, *tiempo_intervencion* y *year*).

Las variables permiten analizar perfectamente la distribución de intervenciones según los distintos distritos, hospitales y las franjas horarias. Proporcionan una estructura perfecta para aplicar el Análisis de Correspondencias entre las variables de ***DISTRITO*** (que serán las filas) y la variable ***HOSPITAL*** (que será la columna), además consideraremos ***HORARIO*** como variable suplementaria.

2. Análisis de Correspondencias (AC) inicial: distritos vs. Hospitales

Realizamos el análisis principal con todas las observaciones y con la variable HORARIO como variable suplementaria.

En el análisis de correspondencias partimos del siguiente contraste:

$$\left\{ \begin{array}{l} H_0: \text{Las variables "Distrito" y "Hospital" son independientes.} \\ H_1: \text{Existe asociación entre las categorías de "Distrito" y "Hospital"} \end{array} \right.$$

AC - Distritos vs Hospitales (SAMUR 2023)

The CORRESP Procedure

Contingency Table																				
	Alcalá de Henares (Ppe. de Ast)	Central de la Defensa	Centros Municipales (Asist. y)	Clínico San Carlos	Concepción (Fund. J. Díaz)	Doce de Octubre	Doctor Rodríguez Lafora	Getafe	Gregorio Marañón	INFANTA LEONOR	La Paz	La Princesa	Militares (Aire, Polic. Naval)	Moncloa (ASISA)	Niño Jesús (Infantil)	Otros Centros Públicos	Puerta de Hierro	Ramón y Cajal	HOSPITALIZACIÓN SIN	Sum
ARGANZUELA	0	58	0	410	1518	1512	0	0	1118	98	84	78	0	4	52	0	0	50	10092	15050
BARAJAS	0	2	0	14	22	20	0	0	60	18	118	60	0	0	4	0	0	948	2228	3492
CARABANHEL	0	2310	4	530	180	3088	0	2	140	48	28	22	0	10	20	0	2	18	12404	18780
CENTRO	2	110	4	3398	4832	818	0	0	1734	142	338	484	2	6	76	0	0	240	27734	39718
CHAMARTÍN	0	10	0	128	142	72	0	0	330	24	1400	1678	0	0	52	0	0	370	8132	12338
CHAMBERÍ	0	20	0	2522	842	58	0	0	276	30	302	288	0	8	28	0	0	130	8930	13432
CIUDAD LINEAL	0	10	0	40	68	98	0	0	940	60	232	1030	0	0	88	0	0	1732	8722	13018
FUENCARRAL	0	8	0	102	148	38	0	0	82	28	3420	54	0	2	12	0	0	604	7338	11830
FUERA TÉRMINO MUNICIPAL	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	2	8	12
HORTALEZA	0	8	0	28	38	44	0	0	128	28	344	110	0	0	20	0	0	2238	5688	8688
LATINA	0	2008	8	2178	314	894	2	0	102	32	40	22	0	2	40	0	4	32	9832	15308
MONCLOA	0	84	0	2002	2122	198	2	0	188	20	344	88	0	0	24	0	8	92	9088	14230
MORATALAZ	0	0	0	10	8	48	0	2	1898	98	14	42	0	0	42	0	0	42	3484	5484
POZUELO	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	10
RETIRO	0	12	0	84	112	148	0	0	2794	118	50	180	0	2	98	0	0	38	5854	9488
SALAMANCA	0	8	2	182	212	78	0	2	1338	44	158	2852	0	0	72	0	0	152	9204	14278
SAN BLAS	0	10	2	42	50	108	0	0	722	128	178	158	0	0	52	0	0	2188	7772	11402
TETUÁN	0	18	0	348	248	82	0	0	150	48	3888	188	0	0	22	0	0	210	9848	14802
USERA	0	62	0	42	50	3528	0	0	84	28	12	10	0	0	28	0	0	14	7882	11718
VALLERÍA S. PTE.	0	8	0	38	62	290	0	0	2958	2588	52	82	0	0	68	2	0	38	11940	18082
VICALVARO	2	4	0	20	18	48	0	0	1094	300	28	34	0	0	8	0	0	52	2874	4478
VILLA DE VALLECAS	0	8	0	10	20	188	0	2	442	1888	32	20	0	0	18	0	0	20	4244	6888
VILLA VERDE	0	10	0	18	14	3144	0	0	80	48	22	8	0	0	14	0	0	10	5712	9058
Sum	4	4758	20	12118	10798	14248	4	8	18412	5758	10854	7424	2	34	828	2	12	9220	178998	271494

Supplementary Columns						
	MAÑANA	MEDIA NOCHE	MEDIO_DIA	NOCHE	TARDE	TEMPRANO
ARGANZUELA	3480	1962	3336	1000	3486	1786
BARAJAS	884	426	812	248	684	438
CARABANCHEL	4114	2886	2826	1476	4292	1866
CENTRO	6932	6098	8116	6068	8500	4002
CHAMARTIN	2044	1456	2724	1018	2686	1518
CHAMBERI	2890	1744	2622	2112	2466	1598
CIUDAD LINEAL	3554	1592	2702	844	2910	1414
FUENCARRAL	3506	1136	2840	552	2488	1308
FUERA TERMINO MUNICIPAL	4	2	0	0	0	6
HORTALEZA	2446	890	2042	446	1970	874
LATINA	3972	2026	3190	982	3450	1686
MONCLOA	3374	1676	3196	1018	3286	1680
MORATALAZ	1570	622	1056	380	1320	516
POZUELO	0	0	6	0	2	2
RETIRO	2566	978	2366	484	2102	990
SALAMANCA	3460	1828	3266	1330	3040	1354
SAN BLAS	2942	1416	2590	680	2586	1188
TETUAN	3566	2020	3080	1416	3024	1696
USERA	2704	1734	2362	1124	2552	1242
VALLECAS PTE.	4298	2724	3686	1536	4028	1810
VICALVARO	1184	560	948	218	1042	526
VILLA DE VALLECAS	1684	880	1540	352	1592	820



Mapa de Madrid con los hospitales de la base de datos reflejados

1) Tabla de contingencia principal

La primera tabla muestra la tabla de contingencia principal entre los distritos (las filas) y los hospitales (columnas). Esta tabla refleja el número de traslados realizados desde cada uno de los distritos hacia los distintos hospitales.

En la tabla hemos subrayado con color rojo los hospitales y distritos con más concentración de casos, en verde los que menos.

Más concentración de casos: vemos como los hospitales del *Doce de octubre*, *Gregorio Marañón*, *La Paz* y el *Clinico San Carlos* absorben gran parte de las derivaciones. Estos primeros tres hospitales tienen una media de casi 400.000 personas adscritas, siendo la Paz con 528 922 personas adscritas al 31 de diciembre de 2021, el que más. Esto explica a qué se puede deber esta concentración.

En cuanto a los distritos, *Carabanchel*, *Centro*, *Latina* y *Puente de Vallecas* son los que presentan un mayor volumen de actuaciones, superando las 14.000 intervenciones cada uno.

También podemos apreciar valores elevados en la categoría *Sin hospitalización*, lo que indica un número importante de intervenciones que se resolvieron in situ.

Menor concentración de casos: en los hospitales y distritos periféricos es donde encontramos menor concentración de casos, como en los hospitales *Príncipe de Asturias* en Alcalá de Henares o el hospital *Doctor Rodríguez Lafora* en Valdeletras. Los distritos con menos intervenciones son *Barajas*, *Pozuelo* y *Vicálvaro* con menos de 4500 casos cada uno.



En esta imagen podemos observar las localizaciones de los distritos *Latina*, *Carabanchel*, *Centro* y *Puente de Vallecas*, que son precisamente los que presentan los valores más elevados de derivaciones al *Hospital 12 de octubre* (subrayado también). Esta correlación es coherente dado la proximidad geográfica, además son las áreas de referencia tradicional del 12 de octubre.

2) Tabla de distribución por franja horaria

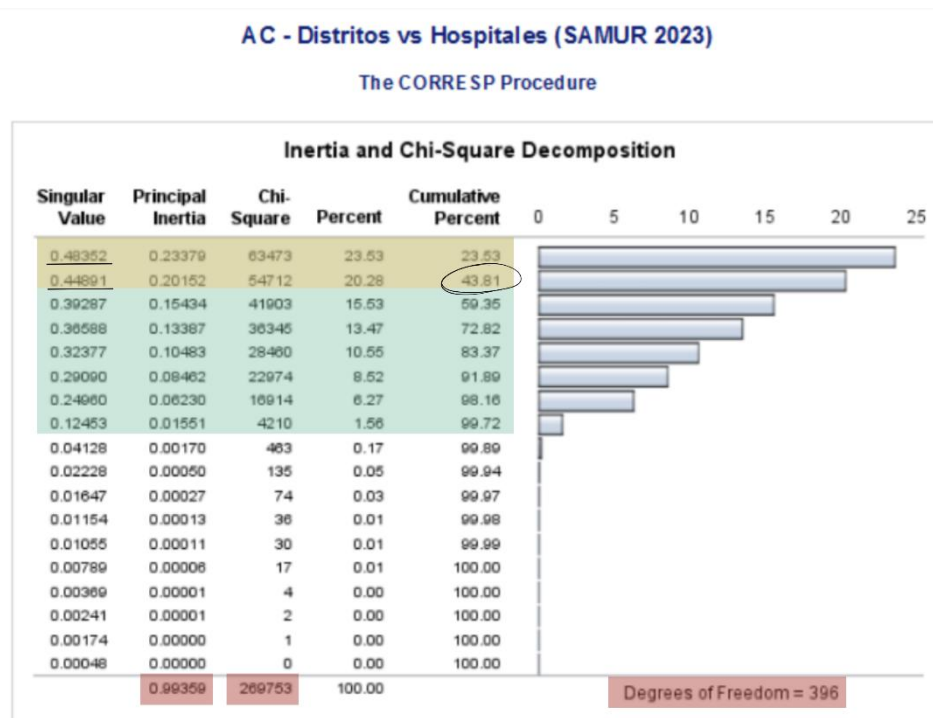
Esta tabla muestra la distribución de las intervenciones por franja horaria, la variable *HORARIO*, que actúa como suplementaria, y por distrito.

Hay una clara concentración de actuaciones en franjas de mañana y tarde. Destaca en especial el distrito *CENTRO*, que en todas las franjas horarias presenta valores muy superiores al resto de la tabla y que a mediodía y tarde obtiene las máximas cifras de 8500 y 8116 intervenciones. El distrito Centro concentra la gran parte de la actividad laboral, administrativa, turística y comercial de la ciudad, por lo que durante el día su población se multiplica por el desplazamiento de muchos de los trabajadores, estudiantes y turistas. Estas horas del mediodía y tarde coinciden con los picos de movimiento peatonal, tráfico, ocio y restauración. Este aumento tanto de movilidad como de interacción social incrementa mucho la probabilidad de incidencias médicas.

En otros distritos como *Arganzuela*, *Latina*, *Vallecas* y *Salamanca* se observan también cifras elevadas sobre todo en las columnas de mañana, mediodía y tarde.

Por último, en los distritos de *Pozuelo* y “*Fuera el término municipal*” no hay casi intervenciones a lo largo del día, las cifras descienden hasta 6 y 2. Esto también se puede deber a lo explicado antes, gran parte de la población que vive en distritos periféricos se moviliza al centro de Madrid para trabajar o estudiar durante el día, por lo que no hay “ciudadanos para intervención” restantes.

3) Inertia and Chi-Square Descomposition



Esta tabla resume la fuerza de la asociación entre las variables y cuánto aporta cada eje factorial a explicarla.

El estadístico que utilizamos para el contraste de independencia presentado anteriormente es:

$$\chi^2 = N \sum_{i=1}^n \sum_{j=1}^p \frac{(f_{ij} - f_{i.}f_{.j})^2}{f_{i.}f_{.j}}$$

Y bajo la hipótesis nula, el estadístico sigue aproximadamente una distribución de $\chi^2_{(n-1)(p-1)}$

La tabla de descomposición de la inercia muestra que el estadístico Chi-cuadrado total es 269753 con 396 grados de libertad, por lo que se rechaza la hipótesis de independencia ($p < 0.0001$). Confirmamos que existe relación significativa entre distritos y hospitales.

Además, la inercia total, que es 0.9936 refleja una asociación notable.

Las dos primeras dimensiones (subrayadas en amarillo) concentran aproximadamente el 44% de la variabilidad total (la primera un 23.5% y la segunda un 20.3%), lo cual justifica su representación en el plano factorial. El análisis bidimensional es suficiente para captar la estructura principal de relaciones entre distritos y hospitales.

4) Summary Statistics

	Quality	Mass	Inertia
ARGANZUELA	0.3349	0.0554	0.0138
BARAJAS	0.2745	0.0129	0.0238
CARABANCHEL	0.6408	0.0691	0.0739
CENTRO	0.0718	0.1463	0.0437
CHAMARTIN	0.4457	0.0454	0.0324
CHAMBERI	0.1014	0.0495	0.0300
CIUDAD LINEAL	0.3940	0.0479	0.0262
FUENCARRAL	0.6026	0.0436	0.0771
FUERA TERMINO MUNICIPAL	0.0000	0.0000	0.0001
HORTALEZA	0.2890	0.0319	0.0536
LATINA	0.4094	0.0564	0.0635
MONCLOA	0.1506	0.0524	0.0344
MORATALAZ	0.4345	0.0201	0.0253
POZUELO	0.0000	0.0000	0.0000
RETIRO	0.3984	0.0349	0.0380
SALAMANCA	0.1244	0.0526	0.0667
SAN BLAS	0.2581	0.0420	0.0374
TETUAN	0.5416	0.0545	0.0689
USERA	0.4666	0.0432	0.0616
VALLECAS PTE.	0.8995	0.0666	0.0714
VICÁLVARO	0.7421	0.0165	0.0141
VILLA DE VALLECAS	0.5584	0.0253	0.0804
VILLAVEDE	0.4401	0.0334	0.0639

	Quality	Mass	Inertia
Alcalá de Henares (Ppe. de Ast.)	0.0000	0.0000	0.0002
Central de la Defensa	0.3782	0.0175	0.0985
Centros Municipales (Asist. y	0.2019	0.0001	0.0002
Clínico San Carlos	0.2188	0.0446	0.0709
Concepción (Fund. J. Díaz)	0.0842	0.0398	0.0610
Doce de Octubre	0.5792	0.0525	0.1502
Doctor Rodríguez Lafora	0.0000	0.0000	0.0001
Getafe	0.0000	0.0000	0.0002
Gregorio Marañón	0.6122	0.0605	0.0956
INFANTA LEONOR	0.6418	0.0212	0.1359
La Paz	0.5910	0.0400	0.1555
La Princesa	0.1505	0.0273	0.0950
Militares (Aire, Polic. Naval,	0.0000	0.0000	0.0000
Moncloa (A SISA)	0.2270	0.0001	0.0003
Niño Jesús (Infantil)	0.3088	0.0030	0.0015
Otros Centros Públicos	0.0000	0.0000	0.0001
Puerta de Hierro	0.1794	0.0000	0.0003
Ramón y Cajal	0.3034	0.0340	0.1336
SIN HOSPITALIZACIÓN	0.0411	0.6593	0.0008

Estas tablas resumen cómo de bien está representada cada categoría (de las filas o de las columnas) dentro del análisis.

Para la interpretación nos hemos fijado en los distritos y los hospitales con valores altos en las columnas de Quality e Inertia, que indican lo bien representados gráficamente que están y cuánto más contribuyen a la estructura del AC.

a) Row Points

En la tabla de las filas, se destacan varios distritos como *Tetuán*, *Usera*, *Puente de Vallecas*, *Fuencarral*, *Carabanchel* o *Villaverde* donde no solo hay puntuaciones altas en Quality, sino que también en Inertia. En distritos como *Salamanca*, solo hay puntuación alta en la columna de Inertia.

Los distritos de *Pozuelo* y “*Fuera del término municipal*” tienen valores nulos en ambas columnas.

b) Column Points

Encontramos puntuaciones altas para Quality en los hospitales de *Gregorio Marañón*, *La Paz* e *Infanta Leonor*. Hospitales como el *Ramón y Cajal*, *Doce de Octubre* y el *Infanta Leonor*, también tienen puntuaciones altas en Inertia.

Getafe, *Militares*, *Príncipe de Asturias* y “*Otros centros públicos*” tienen valores casi nulos para las columnas.

5) Partial Contributions to Inertia e Indices of Coordinates That Contribute Most row / columns

	Dim1	Dim2
ARGANZUELA	0.0190	0.0008
BARAJAS	0.0249	0.0033
CARABANCHEL	0.1905	0.0124
CENTRO	0.0093	0.0046
CHAMARTIN	0.0492	0.0142
CHAMBERI	0.0020	0.0127
CIUDAD LINEAL	0.0438	0.0001
FUENCARRAL	0.1024	0.1104
FUERA TERMINO MUNICIPAL	0.0001	0.0000
HORTALEZA	0.0578	0.0093
LATINA	0.0903	0.0234
MONCLOA	0.0067	0.0178
MORATALAZ	0.0037	0.0500
POZUELO	0.0000	0.0000
RETIRO	0.0047	0.0691
SALAMANCA	0.0308	0.0054
SAN BLAS	0.0410	0.0000
TETUAN	0.0745	0.0977
USERA	0.1213	0.0011
VALLECAS PTE.	0.0055	0.2999
VICALVARO	0.0026	0.0485
VILLA DE VALLECAS	0.0015	0.2187
VILLVERDE	0.1189	0.0006

	Dim1	Dim2	Best
ARGANZUELA	0	0	1
BARAJAS	0	0	1
CARABANCHEL	1	0	1
CENTRO	0	0	1
CHAMARTIN	1	0	1
CHAMBERI	0	0	2
CIUDAD LINEAL	0	0	1
FUENCARRAL	2	2	2
FUERA TERMINO MUNICIPAL	0	0	1
HORTALEZA	1	0	1
LATINA	1	0	1
MONCLOA	0	0	2
MORATALAZ	0	2	2
POZUELO	0	0	2
RETIRO	0	2	2
SALAMANCA	0	0	1
SAN BLAS	0	0	1
TETUAN	2	2	2
USERA	1	0	1
VALLECAS PTE.	0	2	2
VICALVARO	0	0	2
VILLA DE VALLECAS	0	2	2
VILLVERDE	1	0	1

Estas dos tablas de contribuciones parciales muestran qué distritos definen cada dimensión del análisis de correspondencias.

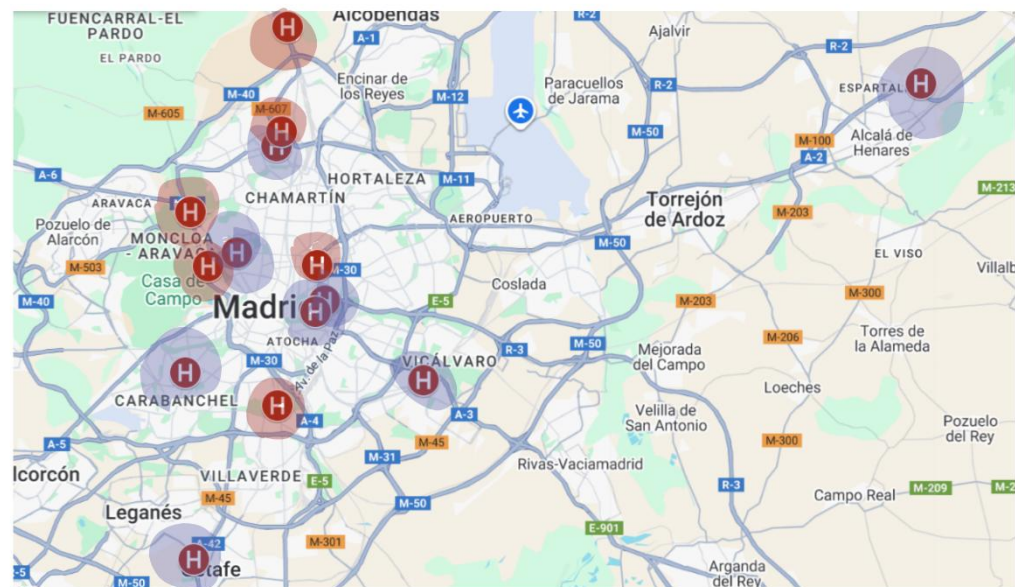
En la primera dimensión (subrayado azul), las mayores contribuciones corresponden a *Fuencarral* (0.1024), *Carabanchel* (0.1905), *Usera* (0.1213), *Villaverde* (0.1189) y *Latina* (0.0903). En la segunda dimensión (subrayado morado), *Puente de Vallecas* (0.2187), *Villa de Vallecas* (0.2999) y *Tetuán* (0.977) suman las mayores contribuciones.

Hay algunos distritos como *Pozuelo* o “*Fuera de Término municipal*” que directamente no presentan contribuciones.

Los índices de coordenadas reafirman estos resultados, asignando el Eje 1 principalmente a los distritos como *Latina*, *Carabanchel*, *Chamartín*, *Usera*, *Villaverde* etc, y el Eje 2 a las zonas como *Villa* y *Puente de Vallecas*, *Tetuán*, *Moratalaz*, *Moncloa*...



Indices of the Coordinates That Contribute Most to Inertia for the Column Points			
	Dim1	Dim2	Best
Alcalá de Henares (Ppe. de Ast	0	0	2
Central de la Defensa	1	0	1
Centros Municipales (Asist. y	0	0	1
Clínico San Carlos	0	0	2
Concepción (Fund. J. Díaz)	0	0	2
Doce de Octubre	1	0	1
Doctor Rodríguez Lafora	0	0	1
Getafe	0	0	2
Gregorio Marañón	0	2	2
INFANTA LEONOR	0	2	2
La Paz	2	2	2
La Princesa	0	0	1
Militares (Aire,Polic. Naval,	0	0	1
Moncloa (ASISA)	0	0	1
Niño Jesús (Infantil)	0	0	2
Otros Centros Públicos	0	0	2
Puerta de Hierro	0	0	1
Ramón y Cajal	1	0	1
SIN HOSPITALIZACION	0	0	2



6) Squared cosines row/columns/supplementary

Squared Cosines for the Row Points			Squared Cosines for the Column Points			Squared Cosines for the Supplementary Column Points		
	Dim1	Dim2		Dim1	Dim2		Dim1	Dim2
ARGANZUELA	0.3238	0.0111	Alcalá de Henares (Ppe. de Ast	0.0000	0.0000	MAÑANA	0.1255	0.0008
BARAJAS	0.2481	0.0283	Central de la Defensa	0.3350	0.0432	MEDIA NOCHE	0.3277	0.0185
CARABANHEL	0.0087	0.0341	Centros Municipales (Asist. y	0.1833	0.0385	MEDIO_DIA	0.2038	0.0011
CENTRO	0.0502	0.0215	Clínico San Carlos	0.1100	0.1088	NOCHE	0.0344	0.0231
CHAMARTIN	0.3589	0.0888	Concepción (Fund. J. Díaz)	0.0399	0.0443	TARDE	0.0480	0.1492
CHAMBERI	0.0158	0.0857	Doce de Octubre	0.5752	0.0040	TEMPRANO	0.0134	0.0510
CIUDAD LINEAL	0.3931	0.0009	Doctor Rodríguez Lafora	0.0000	0.0000			
FUENCARRAL	0.3123	0.2903	Getafe	0.0000	0.0000			
FUERA TERMINO MUNICIPAL	0.0000	0.0000	Gregorio Marañón	0.0583	0.5559			
HORTALEZA	0.2538	0.0352	INFANTA LEONOR	0.0097	0.8320			
LATINA	0.3348	0.0748	La Paz	0.3030	0.2880			
MONCLOA	0.0455	0.1050	La Princesa	0.1504	0.0000			
MORATALAZ	0.0344	0.4001	Militares (Aire, Polic. Naval,	0.0000	0.0000			
POZUELO	0.0000	0.0000	Moncloa (ASISA)	0.1805	0.0465			
RETIRO	0.0292	0.3892	Niño Jesús (Infantil)	0.0995	0.2093			
SALAMANCA	0.1079	0.0185	Otros Centros Públicos	0.0000	0.0000			
SAN BLAS	0.2581	0.0000	Puerta de Hierro	0.1251	0.0542			
TETUAN	0.2543	0.2875	Ramón y Cajal	0.2788	0.0246			
USERA	0.4832	0.0038	SIN HOSPITALIZACION	0.0131	0.0280			
VALLECAS PTE.	0.0181	0.8514						
VICALVARO	0.0438	0.6983						
VILLA DE VALLECAS	0.0045	0.5520						
VILLAVERDE	0.4380	0.0021						

Los cosenos al cuadrado indican la proporción de la varianza de un punto (distrito u hospital) que queda explicada por cada eje factorial.

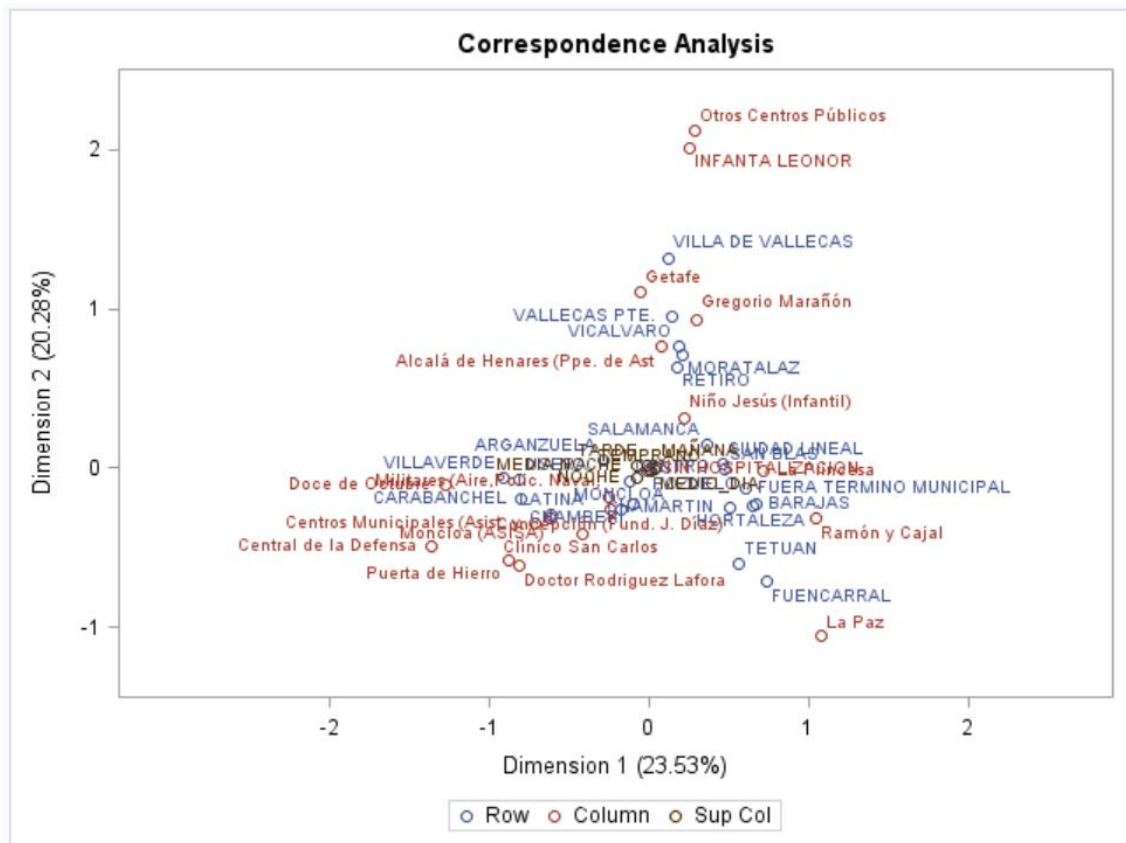
En las tablas vemos cómo hay distritos y hospitales que se explican mejor que otros, coinciden además con varios de los que tenían las puntuaciones más altas en las anteriores tablas, si un distrito tiene mucha carga para explicar un factor, va a estar bien representado en ese mismo factor, como pasa en la dimensión 1 con *Usera*, *Tetuán*, *Carabanchel*... o en la dimensión 2 con *Vallecas*, *Fuencarral*, *Vicálvaro* etc.

En la tabla de los cosenos de los hospitales vemos que hay muchos hospitales que se describen mejor en el Eje 1, como el *Doce de Octubre*, *Ramón y Cajal*, *La Paz* etc. En el Eje 2 el *Gregorio Marañón*, *Infanta Leonor* y el *Niño Jesús* se explican también con valores muy altos.

Para nuestras variables suplementarias, en la dimensión 1 se explican bien el horario de mañana-mediodía-medianoche y para la dimensión 2 las variables tarde y temprano.

Hay tanto hospitales como distritos y horarios que no se explican bien en ninguna de las dimensiones.

7) Gráfico factorial del análisis de correspondencias



El gráfico factorial de correspondencias representa en un mismo plano a los distritos (en azul) y a los hospitales (en rojo), proyectados sobre las dos primeras dimensiones, explicando conjuntamente el 43.8% de la inercia total.

La Dimensión 1 (con un 23.53% de representación), en el eje horizontal, diferencia dos grandes grupos. Los distritos situados a la derecha (*Chamartín, Tetuán, Fuencarral, Ciudad Lineal*) están asociados a los hospitales del norte y alta capacidad, como *La Paz*, el *Ramón y Cajal* o el *Gregorio Marañón*. Los distritos del lado izquierdo (*Carabanchel, Usera, Arganzuela, Latina* etc.) están mas próximos a los hospitales del sur de Madrid, como *Puerta de Hierro, el Clínico San Carlos o el Doce de Octubre*.

La Dimensión 2 (20.28%), en el eje vertical, separa otras dos aglomeraciones. En la parte superior del eje, *Puente de Vallecas, Vicálvaro, Villa de Vallecas...* están asociados a centros como la *Infanta Leonor* y otros centros públicos. En la parte inferior del mapa, *Fuencarral y La Paz* representan un perfil opuesto dentro de la red hospitalaria.

Las modalidades situadas cerca del punto de origen (como *Centro, Salamanca, Moncloa o Barajas*) tienen unos perfiles más equilibrados y no se asocian de forma marcada con ningún hospital en concreto.

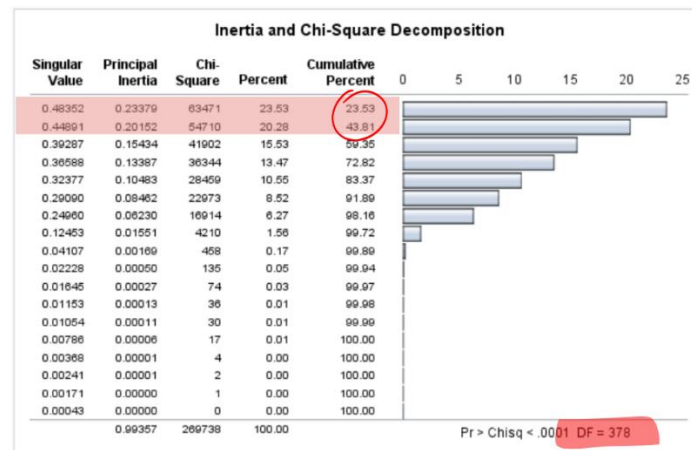
El agrupamiento visible de los distritos del sur, (*Carabanchel, Usera, Latina*) con hospitales situados en esa zona sur de Madrid, sugiere básicamente un patrón geográfico. *Pozuelo* no aparece alejado ni asociado a ningún grupo.

3. Repetición del Análisis de Correspondencias sin el distrito Pozuelo

Se repite el análisis sin contar con el distrito de Pozuelo, queremos observar su influencia en la estructura de correspondencias.

Se presentan a continuación las tablas donde se ve mejor la diferencia respecto al AC completo.

1) Inertia and Chi-Square Decomposition S.P



En la tabla de Inercia y descomposición Chi-Cuadrado observamos a simple vista que han bajado los grados de libertad de 396 a 378. Los porcentajes explicados de las dos primeras dimensiones no han variado prácticamente. La estructura general se mantiene, pero el total de inercia disminuye un poco, confirmando que *Pozuelo* apenas aportaba información a la relación entre distritos y hospitales.

2) Summary Statistics for the Row Points S.P

Summary Statistics for the Row Points			
	Quality	Mass	Inertia
ARGANZUELA	0.3349	0.0554	0.0138
BARAJAS	0.2745	0.0129	0.0238
CARABANCHEL	0.6408	0.0691	0.0739
CENTRO	0.0718	0.1463	0.0437
CHAMARTIN	0.4457	0.0454	0.0324
CHAMBERI	0.1014	0.0495	0.0300
CIUDAD LINEAL	0.3940	0.0479	0.0262
FUENCARRAL	0.6026	0.0436	0.0771
FUERA TERMINO MUNICIPAL	0.0000	0.0000	0.0001
HORTALEZA	0.2890	0.0319	0.0536
LATINA	0.4094	0.0564	0.0635
MONCLOA	0.1506	0.0524	0.0344
MORATALAZ	0.4345	0.0201	0.0253
POZUELO	0.0000	0.0000	0.0000
RETIRO	0.3984	0.0349	0.0380
SALAMANCA	0.1244	0.0526	0.0667
SAN BLAS	0.2581	0.0420	0.0374
TETUAN	0.5418	0.0545	0.0689
USERA	0.4668	0.0432	0.0616
VALLECAS PTE.	0.8695	0.0666	0.0714
VICALVARO	0.7421	0.0165	0.0141
VILLA DE VALLECAS	0.5564	0.0253	0.0804
VILLAVEDE	0.4401	0.0334	0.0639

Summary Statistics for the Row Points			
	Quality	Mass	Inertia
ARGANZUELA	0.3349	0.0554	0.0138
BARAJAS	0.2745	0.0129	0.0238
CARABANCHEL	0.6408	0.0691	0.0739
CENTRO	0.0718	0.1463	0.0437
CHAMARTIN	0.4457	0.0454	0.0324
CHAMBERI	0.1014	0.0495	0.0300
CIUDAD LINEAL	0.3940	0.0479	0.0262
FUENCARRAL	0.6026	0.0436	0.0771
FUERA TERMINO MUNICIPAL	0.0000	0.0000	0.0001
HORTALEZA	0.2890	0.0319	0.0536
LATINA	0.4094	0.0564	0.0635
MONCLOA	0.1506	0.0524	0.0344
MORATALAZ	0.4345	0.0201	0.0253
RETIRO	0.3984	0.0349	0.0380
SALAMANCA	0.1244	0.0526	0.0667
SAN BLAS	0.2581	0.0420	0.0374
TETUAN	0.5418	0.0545	0.0689
USERA	0.4668	0.0432	0.0616
VALLECAS PTE.	0.8695	0.0666	0.0714
VICALVARO	0.7421	0.0165	0.0141
VILLA DE VALLECAS	0.5564	0.0253	0.0804
VILLAVEDE	0.4401	0.0334	0.0639

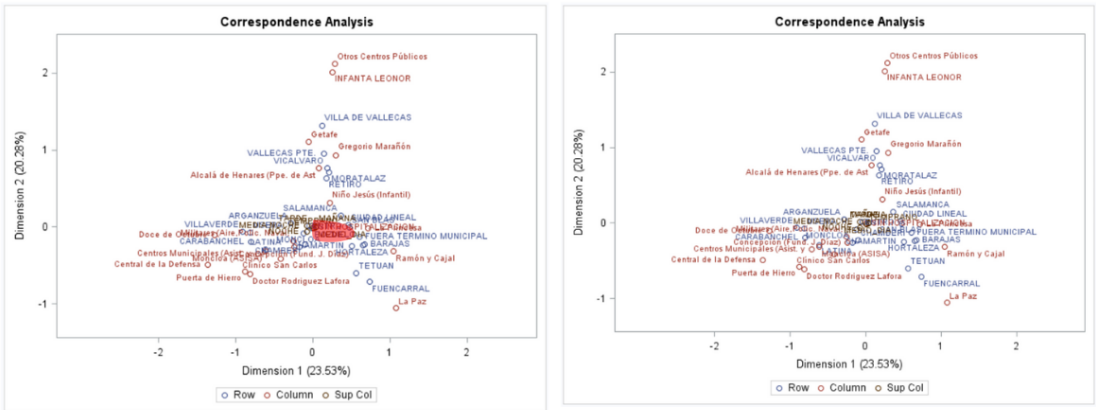
Tras eliminar el distrito de *Pozuelo*, los valores de Quality, Mass e Inertia para el resto de los distritos permaneces totalmente idénticos. Esto confirma que *Pozuelo* no aportaba nada de información al modelo, su presencia no afecta a la estructura factorial ni a la distribución de inercia entre las distintas modalidades.

Partial Contributions to Inertia for the Row Points			Partial Contributions to Inertia for the Row Points		
	Dim1	Dim2		Dim1	Dim2
ARGANZUELA	0.0190	0.0008	ARGANZUELA	0.0190	0.0008
BARAJAS	0.0249	0.0033	BARAJAS	0.0249	0.0033
CARABANCHEL	0.1905	0.0124	CARABANCHEL	0.1905	0.0124
CENTRO	0.0093	0.0046	CENTRO	0.0093	0.0046
CHAMARTIN	0.0492	0.0142	CHAMARTIN	0.0492	0.0142
CHAMBERI	0.0020	0.0127	CHAMBERI	0.0020	0.0127
CIUDAD LINEAL	0.0438	0.0001	CIUDAD LINEAL	0.0438	0.0001
FUENCARRAL	0.1024	0.1104	FUENCARRAL	0.1024	0.1104
FUERA TERMINO MUNICIPAL	0.0001	0.0000	FUERA TERMINO MUNICIPAL	0.0001	0.0000
HORTALEZA	0.0578	0.0093	HORTALEZA	0.0578	0.0093
LATINA	0.0903	0.0234	LATINA	0.0903	0.0234
MONCLOA	0.0067	0.0178	MONCLOA	0.0067	0.0178
MORATALAZ	0.0037	0.0500	MORATALAZ	0.0037	0.0500
POZUELO	0.0000	0.0000	RETIRO	0.0047	0.0691
RETIRO	0.0047	0.0691	SALAMANCA	0.0305	0.0054
SALAMANCA	0.0305	0.0054	SAN BLAS	0.0410	0.0000
SAN BLAS	0.0410	0.0000	TETUAN	0.0745	0.0977
TETUAN	0.0745	0.0977	USERA	0.1213	0.0011
USERA	0.1213	0.0011	VALLECAS PTE.	0.0055	0.2999
VALLECAS PTE.	0.0055	0.2999	VICALVARO	0.0026	0.0485
VICALVARO	0.0026	0.0485	VILLA DE VALLECAS	0.0015	0.2187
VILLA DE VALLECAS	0.0015	0.2187	VILLAVEVERDE	0.1189	0.0006
VILLAVEVERDE	0.1189	0.0006			

Lo mismo ocurre en tablas como Partial Contributions to Inertia for the Row Points, el distrito no tiene ninguna influencia alguna en el modelo.

3) Gráfico del Análisis de Correspondencias S.P

Con



Con el gráfico intacto terminamos de confirmar que haber quitado le distrito de *Pozuelo* no cambia ni influye en absoluto en el modelo.

4. Agrupación de hospitales y distritos y comprobación de frecuencias

En el siguiente apartado agrupamos los hospitales con baja frecuencia como suplementarios y combinamos los distritos con características similares para mejorar la estabilidad del análisis

Tras agruparlos, se analizan las frecuencias de los hospitales (agrupados) para justificar si incluimos o excluimos algunas categorías.

Hemos agrupado los distritos en 3 grandes zonas, basadas en el análisis inicial (excluyendo *Pozuelo*).

Resulta que si proyectábamos en un mapa de Madrid los resultados de a qué dimensión pertenecía cada distrito, la dimensión 2 hacía un “sándwich” a la dimensión 1, los distritos de esta dimensión quedaban distribuidos en las zonas más al norte y más al sur de Madrid, mientras que en la dimensión 1 hacia una recta diagonal de suroeste a noreste de la capital.



La dimensión 1 está coloreada en azul y la dimensión 2 en morado.

Por esto, hemos decidido agrupar todos los distritos en 3 zonas, basadas en ese mismo mapa: la zona norte “NORTE” son los distritos de *Moncloa, Tetuán, Fuencarral, Chamberí etc.* en la parte superior del mapa, coloreado de morado, en la zona sur “SUR”, parte inferior-derecha del mapa, también coloreado de morado, los distritos de *Moratalaz, Vicálvaro, Retiro, Puente de Vallecas etc.* y por último la zona centro “CENTRO” haciendo la diagonal en el mapa y coloreado de azul lo constituyen los distrito de *Latina, Ciudad Lineal, Usera, Salamanca, Chamartín...*

En cuanto a los hospitales, los hemos agrupado también por distribución geográfica “ZONA NORTE”, “ZONA SUR”, “ZONA ESTE”, “ZONA OESTE” y “CENTRO” y además hemos agregado “SUPLEMENTARIO” que almacena los hospitales con menos cargas de la base, como *el Príncipe de Asturias, Rodríguez Lafora, Militares etc.*

1) Análisis de frecuencias

The FREQ Procedure

Frequency	Table of hospital_gr by distrito_gr				
hospital_gr	distrito_gr				Total
	CENTR	NORTE	SUPLE	SUR	
SIN_HOSP	62760	59640	12898	47068	182366
SUPLEMEN	6	2	0	10	18
ZONA_CEN	19050	7960	258	7962	35230
ZONA_EST	682	690	2660	2582	6614
ZONA_NOR	1646	17976	154	360	20136
ZONA_SUR	8840	4262	3358	14314	30774
Total	92984	90530	19328	72296	275138

La variable *SIN HOSPITALIZACIÓN* concentra casi un 66% del total de los casos, lo cual es coherente ya que muchas de las intervenciones se resuelven in situ en vez de hacer traslado hospitalario.

La zona centro concentra 35.230 casos, un 13% del total, la mayor proporción de ingresos hospitalarios se produce en hospitales del área central (*Clínico San Carlos, Central de la Defensa, La Princesa, Puerta de Hierro, etc.*).

En la zona sur los hospitales como *Doce de Octubre* y *Gregorio Marañón* son los principales destinos de los distritos del sur.

En la zona norte destaca *La Paz* y el *Ramón y Cajal* que reciben a los pacientes de distritos del norte, y en la zona este, con menor volumen, está vinculada a hospitales periféricos (como *La Infanta Leonor, El Niño Jesús*).

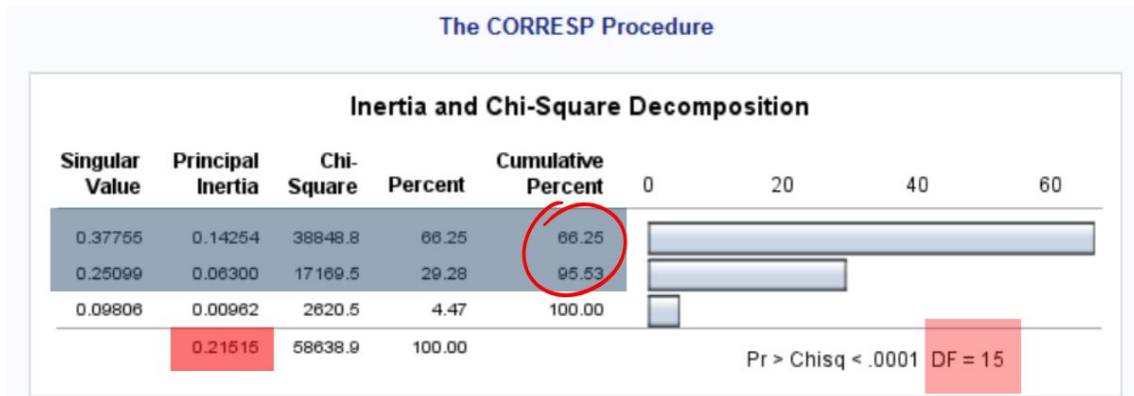
La variable suplementaria es residual, agrupando solo 18 casos, sin relevancia estadística.

Estos resultados confirman un patrón territorial coherente con la distribución geográfica de los servicios hospitalarios de Madrid y justifican la pertinencia del análisis de correspondencias posterior.

Este patrón sugiere que el análisis de correspondencias posterior mostrará un eje principal NORTE–SUR y un eje secundario CENTRO–PERIFERIA.

5. Análisis final de Correspondencias final

1) Inertia and Chi-Square Decomposition FINAL



En la tabla de la descomposición de la Inercia y Chi-Cuadrado, vemos que las dos primeras dimensiones concentran más de un 95% de representación, lo que indica que claramente el plano factorial bidimensional es más que suficiente para representar la relación entre las zonas de distrito y los hospitales agrupados.

La primera dimensión, con un 66%, representa el eje principal de diferenciación, con su gradiente geográfico norte-sur (mencionado anteriormente en la imagen coloreada del mapa) mientras que la segunda dimensión recoge el contraste “diagonal” cetro-periferia este – oeste de la ciudad.

La elevada proporción de inercia explicada y la significación del Chi-Cuadrado garantiza la estabilidad y la relevancia del análisis, validando la interpretación del mapa factorial que se presentará más adelante.

2) Summary Stats for the Row Points FINAL

Summary Statistics for the Row Points			
	Quality	Mass	Inertia
CENTR	0.9464	0.3372	0.1628
NORTE	0.9996	0.3296	0.4256
SUPLE	0.9451	0.0703	0.2411
SUR	0.8678	0.2629	0.1705

Todos los grupos presentan una excelente calidad de representación. Los grupos con mayor presencia en los datos son Centro y Norte, esto significa que son los distritos que aportan la mayor parte de los casos analizados. Respecto a la Inercia, NORTE es el grupo que más contribuye, definiendo fuertemente el eje primer eje factorial.

3) Summary Stats for the Column Points FINAL

Summary Statistics for the Column Points			
	Quality	Mass	Inertia
SIN_HOSP	0.2578	0.6596	0.0006
SUPLEMEN	0.0000	0.0001	0.0002
ZONA_CEN	0.9643	0.1292	0.1303
ZONA_EST	0.9363	0.0243	0.2195
ZONA_NOR	0.9985	0.0739	0.4945
ZONA_SUR	0.8402	0.1129	0.1550

En primer lugar, la calidad de representación (Quality) es alta en la mayoría de las categorías, lo que indica que el plano factorial bidimensional explica correctamente su posición. Para la masa (Mass), la mayor parte de los casos corresponde a *SIN_HOSP*, lo que significa que casi dos tercios de las asistencias del SAMUR no implican traslado hospitalario. En cuanto a la inercia, que mide la contribución de cada grupo hospitalario a la variabilidad total del análisis, destaca en especial la

Zona Norte. Salvo la variable residual *Suplementaria* todos los grupos tienen muy buena representación.

4) Partial Contributions to Inertia for the Row Points FINAL

Partial Contributions to Inertia for the Row Points		
	Dim1	Dim2
CENTR	0.0463	0.4214
NORTE	0.6228	0.0435
SUPLE	0.1149	0.5183
SUR	0.2160	0.0168

Partial Contributions to Inertia for the Column Points		
	Dim1	Dim2
SIN_HOSP	0.0000	0.0005
SUPLEMEN	0.0001	0.0000
ZONA_CEN	0.0098	0.4069
ZONA_EST	0.0991	0.4776
ZONA_NOR	0.7107	0.0782
ZONA_SUR	0.1803	0.0368

En la Dimensión 1 destacan los grupos de *NORTE* (0.6228), *SUR* (0.2160) y *ZONA_NOR* (0.7107), lo que indica que este eje representa principalmente un contraste geográfico norte–sur entre distritos y hospitales.

En la Dimensión 2, los mayores aportes provienen de *SUPLEMENTARIO* (0.5183) en los distritos y de *ZONA_CEN* (0.4069) y *ZONA_EST* (0.4776) en los hospitales. Este segundo eje parece reflejar un contraste entre zonas centrales y periféricas.

En conjunto, los ejes del análisis diferencian claramente el gradiente norte–sur (Dim1) y la oposición centro–periferia (Dim2).

5) Squared Cosines for the Row Points FINAL

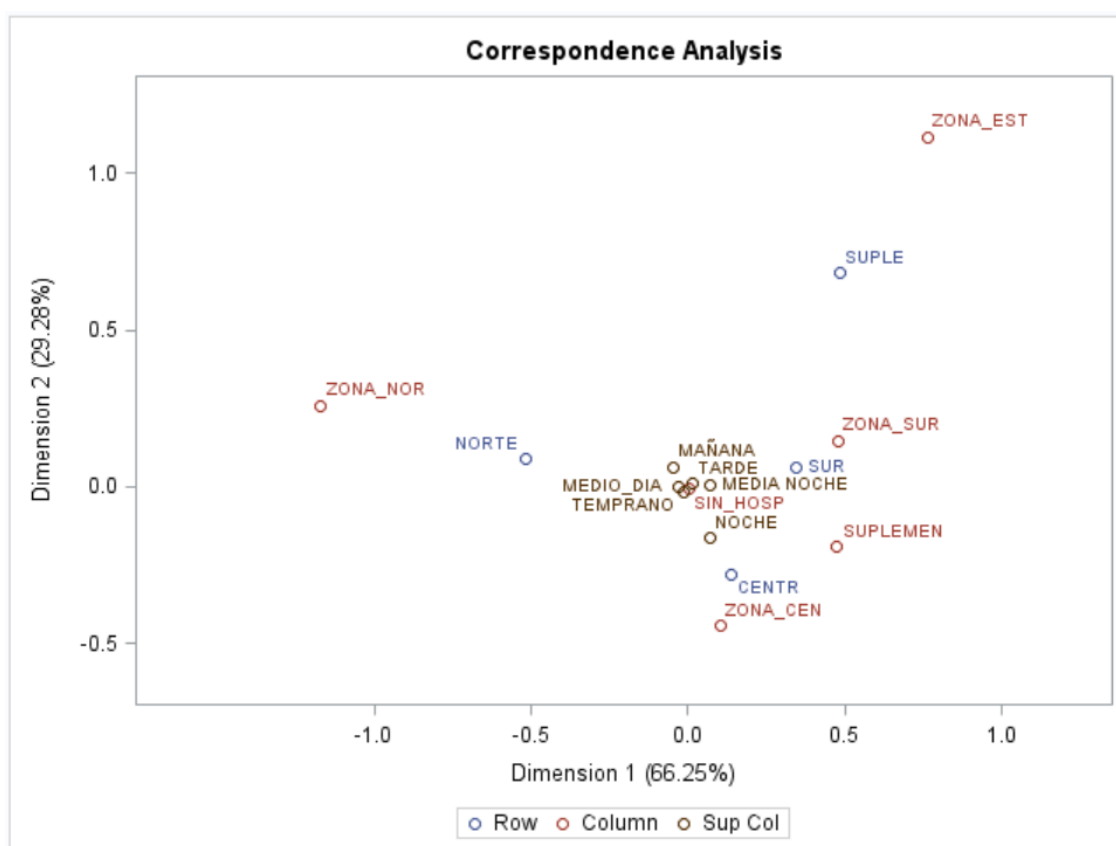
Squared Cosines for the Row Points			Squared Cosines for the Column Points			Squared Cosines for the Supplementary Column Points		
	Dim1	Dim2		Dim1	Dim2		Dim1	Dim2
CENTR	0.1883	0.7581	SIN_HOSP	0.0053	0.2525	MAÑANA	0.2599	0.4352
NORTE	0.9696	0.0299	SUPLEMEN	0.0000	0.0000	MEDIA NOCHE	0.9692	0.0082
SUPLE	0.3157	0.6294	ZONA_CEN	0.0497	0.9145	MEDIO_DIA	0.9774	0.0090
SUR	0.8391	0.0288	ZONA_EST	0.2991	0.6371	NOCHE	0.0978	0.5005
			ZONA_NOR	0.9522	0.0463	TARDE	0.4734	0.1918
			ZONA_SUR	0.7708	0.0695	TEMPRANO	0.2856	0.6912

La Dimensión 1 explica sobre todo el contraste norte–sur, ya que *NORTE* y *SUR* (distritos) y los hospitales *ZONA_NOR* y *ZONA_SUR* tienen los $\cos^2$ más altos en este eje.

La Dimensión 2 refleja un contraste centro–periferia, con *CENTRO* y *SUPLEMENTARIO* bien representados, junto a los hospitales *ZONA_CEN* y *ZONA_EST*.

En los horarios, *Media Noche* y *Medio Día* se asocian al eje 1, mientras que *Temprano* y *Noche* lo hacen al eje 2.

6) Squared Cosines for the Row Points FINAL



El mapa factorial muestra que la Dimensión 1 (66,25 %) refleja un eje geográfico norte–sur.

En el lado izquierdo se sitúan los distritos y hospitales del norte (*NORTE* y *ZONA_NOR*), mientras que a la derecha aparecen los grupos del sur (*SUR* y *ZONA_SUR*) y del este (*ZONA_EST*), junto con el grupo *SUPLEMENTARIO*.

La Dimensión 2 (29,28 %) diferencia principalmente las zonas periféricas (*SUPLE* y *ZONA_EST*, en la parte superior) de las centrales (*CENTR* y *ZONA_CEN*, en la parte inferior*), representando un contraste centro–periferia.

Los horarios se agrupan cerca del origen, lo que indica que no muestran un patrón diferenciado entre zonas distritales o hospitalarias.

En conjunto, el gráfico confirma que el eje principal del análisis es territorial, con una clara oposición norte–sur, mientras que el segundo eje refleja diferencias entre el centro urbano y las zonas periféricas.

6. Conclusión

El Análisis de Correspondencias aplicado a los datos del SAMUR ha permitido identificar una estructura clara en la relación entre las zonas de los distritos y los grupos hospitalarios de destino, incorporando además los horarios de atención como variables suplementarias.

Los resultados muestran una dependencia significativa entre las zonas distritales y hospitalarias ($\chi^2 = 58.639$; $p < 0.0001$), lo que indica que la distribución de los traslados no es aleatoria, sino que sigue un patrón territorial coherente.

Las dos primeras dimensiones del análisis explican el 95,53 % de la inercia total, lo que garantiza una representación fiel de la estructura de datos.

La Dimensión 1 (66,25 %) refleja el gradiente geográfico norte–sur, mientras que la Dimensión 2 (29,28 %) recoge un contraste centro–periferia.

En el plano factorial se observa una correspondencia clara entre los distritos del norte y los hospitales de referencia del norte (La Paz, Ramón y Cajal), así como entre los distritos del sur y los hospitales meridionales (Doce de Octubre, Gregorio Marañón).

Las zonas centrales se asocian con hospitales del área centro (Clínico San Carlos, Central de la Defensa), y las zonas periféricas del este con hospitales como el Infanta Leonor o el Niño Jesús.

En conjunto, el análisis confirma que la actuación del SAMUR presenta una organización territorial coherente con la red hospitalaria de la ciudad, y que la reducción de categorías mediante agrupaciones geográficas y hospitalarias ha permitido obtener un modelo estable, interpretable y representativo de la realidad asistencial de Madrid.