



Universidad
Zaragoza

Trabajo Fin de Grado en Administración y Dirección de Empresas

Paso del bipartidismo a la pluralidad política:
¿Cómo se comportan los individuos?

Autora

ROMINA CAPAPEY MARTÍNEZ

Directores

Jesús Ángel Miguel Álvarez y Pilar Gargallo Valero

Facultad de Economía y Empresa

Universidad de Zaragoza

Octubre 2017

Resumen

Este trabajo trata de estudiar el impacto que ha tenido la aparición de nuevos partidos políticos, analizando el comportamiento de los individuos en su decisión entre el paso del bipartidismo a la pluralidad política. En concreto, analizamos, mediante una encuesta a vecinos de la ciudad de Zaragoza, el voto declarado en las últimas elecciones generales del 2011, 2015 y 2016, junto a otras variables de interés. El objetivo es estudiar el cambio de la escena política española, detectando grupos de personas que han variado o no su intención de voto y los motivos de esta influencia.

Palabras clave: Análisis factorial, Análisis clúster, Intención de voto, Cambio político

From bipartisanship to political plurality: How do individuals behave?

Abstract

This paper tries to analyze the impact of the emergence of new political parties, therefore, we analyze the transition from bipartisanship to political plurality. Specifically, we analyzed through a survey of neighbours of the city of Zaragoza the vote declared in the last general elections of 2011, 2015 and 2016, along with other variables of interest. The aim is to study the change in the Spanish political scene, detecting groups of people who have changed their intention to vote and the reasons for this influence.

Keywords: Factor Analysis, Cluster Analysis, Intention to vote, Political change

INDICE

1.- INTRODUCCIÓN	7
2.- CONJUNTO DE DATOS.....	8
2.1.- Diseño muestral.....	8
2.2.- Diseño de la encuesta	10
3.- ANÁLISIS EXPLORATORIO DE DATOS	10
4.- ANÁLISIS FACTORIAL.....	11
4.1.- Formulación del problema	12
4.2.- Análisis de la matriz de correlación	12
4.3.- Extracción de factores:	16
4.4.- Determinación del número de factores:	16
4.5.- Interpretación de los factores	18
4.6.- Rotación de factores	19
4.7.- Cálculo de Puntuaciones Factoriales.....	21
5.- ANÁLISIS CLÚSTER.....	21
5.1.- Formulación del problema:	21
5.2.- Métodos de clasificación:.....	22
5.2.1.- <i>Clúster Jerárquico:</i>	22
5.2.2.- <i>Clúster de K-medias:</i>	24
5.2.3.- <i>Clúster en dos pasos o Clúster bietápico:</i>	24
6.- CONCLUSIONES.....	37
7.- BIBLIOGRAFÍA.....	40
ANEXO I: ENCUESTA.....	41
ANEXO II: ANÁLISIS EXPLORATORIO DE DATOS.....	46

INDICE DE FIGURAS

<i>Figura 1: Diagrama de sectores de la variable “sexo” dentro de la muestra</i>	<i>9</i>
<i>Figura 2: Gráfico de sedimentación</i>	<i>18</i>
<i>Figura 3: Resumen del modelo bietápico.....</i>	<i>25</i>
<i>Figura 4: Diagrama de sectores de tamaños de clústeres</i>	<i>26</i>
<i>Figura 5: Importancia del predictor.....</i>	<i>27</i>
<i>Figura 6: Comparación de clústeres en las Elecciones Generales de 2011, 2015 y 2016.....</i>	<i>28</i>
<i>Figura 7: Barras de error de la variable “Elecciones 2011” por clúster.....</i>	<i>28</i>
<i>Figura 8: Barras de error de la variable “Elecciones 2015” por clúster.....</i>	<i>28</i>
<i>Figura 9: Barras de error de la variable “Elecciones 2016” por clúster.....</i>	<i>29</i>
<i>Figura 10: Diagrama de barras de la variable “Sexo” por clúster.....</i>	<i>29</i>
<i>Figura 11: Diagrama de barras de la variable “Edad” por clúster.....</i>	<i>30</i>
<i>Figura 12: Diagrama de barras de la variable “Situación laboral” por clúster.....</i>	<i>31</i>
<i>Figura 13: Diagrama de barras de la variable “Nivel máximo de estudios” por clúster.....</i>	<i>31</i>
<i>Figura 14: Diagrama de barras de la variable “Ingresos mensuales” por clúster.....</i>	<i>32</i>
<i>Figura 15: Diagrama de barras de la variable “Representación política” por clúster.....</i>	<i>33</i>
<i>Figura 16: Diagrama de barras de la variable “Posición ideológica” por clúster.....</i>	<i>33</i>
<i>Figura 17: Diagrama de barras de la variable “Religión” por clúster.....</i>	<i>34</i>
<i>Figura 18: Diagrama de barras de la variable “Aparición de Podemos” por clúster.....</i>	<i>35</i>
<i>Figura 19: Barras de error de la variable “Aparición de Podemos” por clúster.....</i>	<i>35</i>
<i>Figura 20: Diagrama de barras de la variable “Aparición de Ciudadanos” por clúster.....</i>	<i>35</i>
<i>Figura 21: Barras de error de la variable “Aparición de Ciudadanos” por clúster.....</i>	<i>35</i>
<i>Figura 22: Barras de error del Factor “Estado de Bienestar” por clúster.....</i>	<i>36</i>
<i>Figura 23: Barras de error del Factor “Preocupación Social” por clúster.....</i>	<i>36</i>
<i>Figura 24: Barras de error de la variable “Igualdad de derechos y oportunidades” por clúster.....</i>	<i>36</i>

INDICE DE TABLAS

<i>Tabla 1: Población original estratificada por junta municipal y sexo</i>	<i>8</i>
<i>Tabla 2: Frecuencias de la variable “Sexo” dentro de la muestra</i>	<i>8</i>
<i>Tabla 3: Muestreo estratificado por junta municipal y sexo</i>	<i>9</i>
<i>Tabla 4: Matriz de correlaciones.....</i>	<i>13</i>
<i>Tabla 5: Prueba de KMO y Bartlett.....</i>	<i>13</i>
<i>Tabla 6: Matriz anti imagen.....</i>	<i>15</i>
<i>Tabla 7: Comunalidades</i>	<i>15</i>
<i>Tabla 8: Varianza total explicada.....</i>	<i>17</i>
<i>Tabla 9: Matriz de componente</i>	<i>20</i>
<i>Tabla 10: Matriz de componente rotado.....</i>	<i>20</i>

1.- INTRODUCCIÓN

Desde el inicio de la democracia en 1978, y especialmente desde 1989, dos grandes formaciones políticas han gobernado España con la mayor parte del apoyo electoral y prácticamente sin la posibilidad de acceso a otras opciones.

En la última década, hemos sido testigos de grandes cambios en el panorama político nacional. Nuestro sistema político, tradicionalmente bipartidista, se ha convertido en un sistema pluripartidista en el que coexisten actualmente cuatro fuerzas políticas nacionales.

Hechos relevantes como la crisis económica que ha afectado a la economía española con mayor intensidad que a sus socios europeos, sumados a las continuas noticias en los medios de comunicación sobre corrupción política, provocaron el nacimiento de nuevos movimientos que ansiaban un cambio ya que la sociedad no se sentía identificada con sus gestores políticos.

Dicho clima de descontento social ha conducido a que los partidos tradicionales no gocen de suficiente apoyo por sí solos y, en consecuencia, se vean obligados a propiciar un clima de diálogo y entendimiento entre los nuevos partidos emergentes, las diversas corrientes políticas y entre la sociedad.

Además de las evidencias anteriores, ¿es posible que algo más haya cambiado? Se estima que al final de cada legislatura, debido al número de jóvenes que se incorporan como nuevos votantes y el número de personas mayores que fallecen, el electorado se renueva en un 10%. Por tanto, ¿han influido en el cambio político de los últimos años las diversas características de los españoles?

Con todo esto en mente, el objeto del presente estudio consiste en identificar las características de los votantes que en un inicio se encontraban reunidos en un gran bloque político y, que posteriormente, se disgregaron con sus votos en un mayor número de partidos. Concretamente, se pretende formar grupos de personas con características socio-demográficas similares y, a través de ello, poder observar la evolución de su comportamiento político.

Para llevar a cabo este estudio, se realizó una encuesta a 196 zaragozanos en la que se les pedía que respondieran a cuestiones políticas, religiosas y características personales que más tarde fueron tratadas mediante las técnicas estadísticas multivariantes de Análisis Factorial y Análisis Clúster.

A continuación, se explicará cómo se ha llevado a cabo la realización de las técnicas estadísticas citadas con sus posteriores resultados y conclusiones.

2.- CONJUNTO DE DATOS

2.1.- Diseño muestral

La muestra de individuos analizada representa a la población de Zaragoza, la cual está compuesta por un número total de 670.571 habitantes y de los que 323.911 son hombres y 346.660 son mujeres (Ver Tabla 1). Ante la imposibilidad de encuestar a toda la población se realizó un muestreo estratificado con afijación proporcional por género y junta municipal, fijando un error del 7% y un nivel de confianza del 95%. El tamaño total de la muestra fue de 196 individuos, de los cuales 95 fueron hombres y 101 mujeres (Ver Tabla 2 y Figura 1).

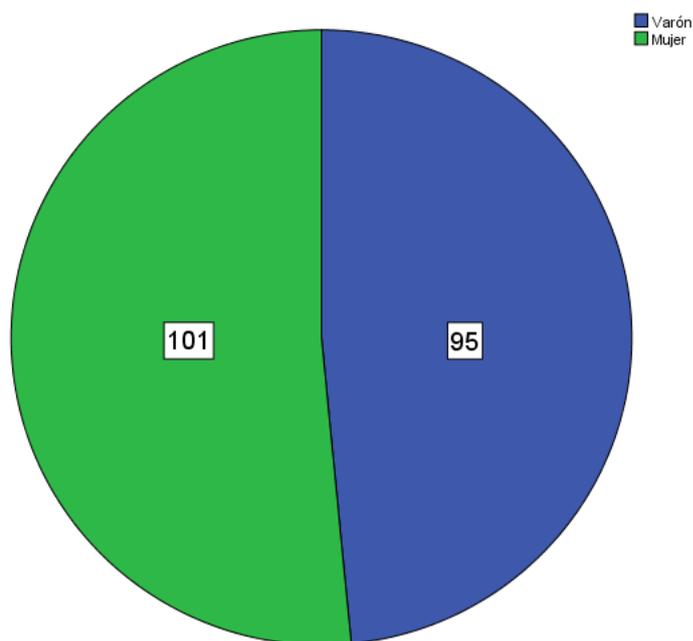
Tabla 1: Población original estratificada por junta municipal y sexo

	Hombres	Mujeres	Total
Casco Histórico	22.837	24.042	46.879
Centro	23.954	29.457	53.411
Delicias	53.090	57.655	110.745
Universidad	23.272	27.049	50.321
Casablanca	20.808	20.632	41.440
San José	31.844	35.520	67.364
Las Fuentes	21.050	21.794	42.844
La Almozara	11.982	12.746	24.728
Miralbueno	5.879	6.008	11.887
Oliver-Valdefierro	15.762	15.803	31.565
Torrero-La Paz	18.822	19.517	38.339
ACTUR-Rey Fernando	29.216	29.923	59.139
El Rabal	38.772	39.764	78.536
Santa Isabel	6.623	6.750	13.373
TOTAL	323.911	346.660	670.571

Tabla 2: Frecuencias de la variable “Sexo” dentro de la muestra

		Sexo			
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Varón	95	48,5	48,5	48,5
	Mujer	101	51,5	51,5	100,0
	Total	196	100,0	100,0	

Figura 1: Diagrama de sectores de la variable “sexo” dentro de la muestra



Para alcanzar una mayor representatividad de la muestra, el muestreo fue realizado en proporción al número de habitantes que residían en cada una de las juntas municipales y al sexo (Ver Tabla 3).

Tabla 3: Muestreo estratificado por junta municipal y sexo

	Hombres	Mujeres	Total
Casco Histórico	7	7	14
Centro	7	8	15
Delicias	15	17	32
Universidad	7	8	15
Casablanca	6	6	12
San José	9	10	19
Las Fuentes	6	6	12
La Almozara	4	4	8
Miralbueno	2	2	4
Oliver-Valdefierro	5	5	10
Torrero-La Paz	6	6	12
ACTUR-Rey Fernando	8	9	17
El Rabal	11	11	22
Santa Isabel	2	2	4
TOTAL	95	101	196

2.2.- Diseño de la encuesta

Una vez establecido nuestro diseño muestral, se realizó una encuesta anónima en cada una de las juntas municipales hasta obtener el conjunto de datos muestrales deseados. El cuestionario que le fue entregado a cada uno de los individuos constaba de dos partes (Véase Anexo I: Encuesta):

- 1ª Parte: Cuestiones personales de índole política y religiosa, y opinión personal sobre temas económicos y sociales generales. Al tratarse de una serie de preguntas que los individuos podían considerar de carácter privado y resultar problemáticas para obtener respuestas, fueron presentadas de forma que ellos las percibieran de menor a mayor grado de intromisión en su intimidad.
- 2ª Parte: Datos de clasificación. Este apartado está formado por preguntas sobre datos personales de los que obtendremos las variables para poder interpretar los resultados obtenidos como, por ejemplo, en función de la edad, situación laboral o nivel de ingresos.

3.- ANÁLISIS EXPLORATORIO DE DATOS

Antes de comenzar la aplicación de las técnicas estadísticas multivariantes en las que se centra este estudio, sería adecuado dar una visión general del análisis unidimensional de las variables a tratar y que se desarrolla de forma más extensa en el Anexo II.

Así, respecto a los datos de carácter personal, podemos afirmar que la mayoría de los individuos encuestados están empleados, sus edades se sitúan entre los 26 y 55 años, poseen estudios universitarios y de grado superior o formación profesional y sus ingresos netos mensuales oscilan entre 900 € y 1.800 €.

Respecto a las cuestiones o variables de índole político, la inmensa mayoría declaró no sentirse representada por los partidos políticos o solamente en algunos aspectos y definieron su ideología como de centro-derecha o izquierda. Además, otorgaron una puntuación media de 5,4 sobre 10 a la cuestión sobre la aparición de Podemos, y una puntuación media de 9,02 sobre 10 a la aparición de Ciudadanos.

En cuanto a su voto a lo largo de las tres últimas convocatorias de elecciones generales, se observa que en el año 2011 el 50% de los encuestados votó a PP o PSOE y que en los años 2015 y 2016 hacía falta la suma de, al menos, tres partidos políticos para alcanzar dicho porcentaje.

Los encuestados también debían puntuar una serie de cuestiones de tipo económico y social según la importancia que para ellos tenían, y aunque con carácter general todas ellas eran de alta importancia para la muestra, hubo diferencias entre unas cuestiones y otras.

Las cuestiones que obtuvieron una puntuación media en grado de importancia ligeramente menor son aquellas relativas a la redistribución de la riqueza entre la sociedad, ayudas a los inmigrantes y refugiados de conflictos bélicos, conservación del medioambiente, respeto a las distintas opciones sexuales y reconocimiento a los matrimonios homosexuales y acceso a la vivienda y políticas antidesahucios.

Las cuestiones que obtuvieron las puntuaciones medias más altas en grado de importancia fueron: Igualdad de derechos y oportunidades de todos los españoles, corrupción política y fraude fiscal, políticas para la reducción del desempleo y mejoras en la calidad del empleo, mejora y refuerzo de la sanidad y la educación públicas, y refuerzo de las medidas para resolver los problemas económicos.

Por último, dentro de las cuestiones de materia religiosa, más del 70% de la muestra se definió como católico o no practicante.

4.- ANÁLISIS FACTORIAL

El Análisis Factorial es una técnica estadística multivariante que, ante un conjunto de numerosas variables, tiene como finalidad agruparlas dentro de un número más reducido de dimensiones que se denominan “Factores”.

Dentro de cada uno de los factores, se encontrarán agrupadas aquellas variables entre las que exista una determinada interrelación, de modo que los factores sean capaces de explicar en buena medida la información que todas ellas aportan, en lugar de obtener dicha información atendiéndolas una a una, lo que daría lugar a una labor mucho más tediosa.

El Análisis Factorial, por tanto, es una técnica de reducción de datos cuyo objetivo es explicar, sin mucha distorsión, la información contenida en la matriz de datos a través de un número menor de dimensiones a las que llamaremos factores.

Existen dos tipos de Análisis Factorial: exploratorio y confirmatorio. En el Análisis Factorial Exploratorio, que es el tipo de técnica que abordaremos en este apartado, el número de dimensiones subyacentes o factores no es conocido hasta que no finaliza el proceso de análisis, mientras que en el Análisis Factorial Confirmatorio, el número de factores ya se ha fijado previamente y su objetivo es llevar a cabo su corroboración.

A continuación, explicaremos cada una de las fases que han de llevarse a cabo para la correcta realización de nuestro Análisis Factorial:

4.1.- Formulación del problema

En el presente apartado, el conjunto de variables que trataremos de reducir en número son diez cuestiones de interés general en la sociedad que les fueron presentadas a los individuos de la muestra para que las puntuaran según su grado de preocupación e importancia.

El análisis estadístico univariante de cada una de estas diez variables cuantitativas ha sido realizado en el apartado “Análisis Exploratorio de Datos”, pero también las enumeraremos a continuación con el fin de entender correctamente su significado.

- **Variable 1:** Igualdad de derechos y oportunidades de todos los españoles.
- **Variable 2:** Redistribución de la riqueza entre la sociedad.
- **Variable 3:** Ayuda a la inmigración y refugiados de conflictos bélicos.
- **Variable 4:** Conservación del medioambiente.
- **Variable 5:** Respeto a las distintas opciones sexuales y reconocimiento de los matrimonios homosexuales.
- **Variable 6:** Corrupción política y fraude fiscal.
- **Variable 7:** Políticas para la reducción del desempleo y mejoras en la calidad del empleo.
- **Variable 8:** Mejora y refuerzo de la sanidad y la educación públicas.
- **Variable 9:** Refuerzo de las medidas para resolver los problemas económicos.
- **Variable 10:** Acceso a la vivienda y políticas anti desahucios.

La variable 1 se ha eliminado puesto que la mayoría de los entrevistados puntuaban entre 9 y 10, por lo tanto, no presenta diferencias entre los individuos analizados.

4.2.- Análisis de la matriz de correlación

El análisis de la matriz de correlación de estas diez variables tiene como finalidad observar el grado de interrelación que existe entre ellas. Si dicha interrelación entre las variables es alta, será adecuado llevar a cabo la técnica de Análisis Factorial ya que eso significará que los Factores obtenidos reflejan y explican correctamente la información que contienen. En caso contrario, si se obtienen correlaciones o interrelaciones muy bajas, el Análisis Factorial resultará una técnica inapropiada para alcanzar su objetivo.

Para conocer si las variables objeto del estudio son adecuadas para la aplicación de la técnica, es posible utilizar varios métodos que mostraremos a continuación:

- Test de esfericidad de Bartlett: Contrasta, bajo la hipótesis de normalidad multivariante, si la matriz de correlaciones es la matriz identidad. Es decir, si se aceptara la hipótesis nula del contraste, significaría que no existe interrelación entre las variables ya que su correlación sería igual a cero y, por tanto, no sería adecuada la aplicación de un Análisis Factorial.

En la Tabla 4 se muestra la matriz de correlaciones de las variables en la que se observa que no se trata de una matriz identidad, por tanto, existe correlación entre las variables.

En la Tabla 5 se observa que el resultado del test de esfericidad de Bartlett ha dado lugar a un valor muy alto que nos lleva a rechazar la hipótesis nula y, como acabamos de explicar, a afirmar que existe correlación entre las variables.

Tabla 4: Matriz de correlaciones

Matriz de correlaciones										
		Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 9	Variable 10
Correlación	Variable 2	1,000	,557	,303	,324	,224	,298	,392	,321	,460
	Variable 3	,557	1,000	,481	,562	,385	,287	,330	,305	,464
	Variable 4	,303	,481	1,000	,507	,326	,418	,354	,477	,440
	Variable 5	,324	,562	,507	1,000	,426	,383	,370	,430	,483
	Variable 6	,224	,385	,326	,426	1,000	,581	,346	,523	,387
	Variable 7	,298	,287	,418	,383	,581	1,000	,583	,769	,614
	Variable 8	,392	,330	,354	,370	,346	,583	1,000	,617	,609
	Variable 9	,321	,305	,477	,430	,523	,769	,617	1,000	,597
	Variable 10	,460	,464	,440	,483	,387	,614	,609	,597	1,000
	Sig. (unilateral)	Variable 2		,000	,000	,000	,001	,000	,000	,000
Variable 3		,000		,000	,000	,000	,000	,000	,000	,000
Variable 4		,000	,000		,000	,000	,000	,000	,000	,000
Variable 5		,000	,000	,000		,000	,000	,000	,000	,000
Variable 6		,001	,000	,000	,000		,000	,000	,000	,000
Variable 7		,000	,000	,000	,000	,000		,000	,000	,000
Variable 8		,000	,000	,000	,000	,000	,000		,000	,000
Variable 9		,000	,000	,000	,000	,000	,000	,000		,000
Variable 10		,000	,000	,000	,000	,000	,000	,000	,000	

Tabla 5: Prueba de KMO y Bartlett

Prueba de KMO y Bartlett	
Medida Kaiser-Meyer-Olkin de adecuación de muestreo	,861
Prueba de esfericidad de Bartlett	Aprox. Chi-cuadrado
	gl
	Sig.
	779,067
	36
	,000

- Medida de adecuación de la muestra KMO: El índice KMO es una medida propuesta por Kaiser, Meyer y Olkin, que realiza una comparación entre los coeficientes de correlación observados y los coeficientes de correlación parciales.

El coeficiente de correlación parcial indica la relación entre dos variables eliminando la influencia del resto. Estima la correlación entre sus factores únicos y, ésta debería ser lo más próxima a cero posible ya que los factores únicos deberían estar incorrelados entre sí.

Por tanto, el índice KMO toma valores entre 0 y 1, y cuanto mayor sea su valor, más adecuado resultará realizar un Análisis Factorial. Kaiser, Meyer y Olkin aconsejan que si $KMO \geq 0,75$ la idea de realizar un Análisis Factorial es buena, si $0,5 \leq KMO < 0,75$ la idea es aceptable y si $KMO < 0,5$ es inaceptable.

En la Tabla 5 expuesta anteriormente, se observa que el valor que en nuestro caso toma dicho índice es igual a 0,861 por lo que sería posible llevar a cabo un Análisis Factorial.

- Medida de adecuación de la muestra MSA: Con esta medida se pretende analizar la adecuación de cada variable al Análisis Factorial de manera individual. Un valor bajo de MSA en una determinada variable significaría incompatibilidad de ésta con el análisis, por lo que nos serviría para identificar qué variables podrían eliminarse para desarrollar correctamente nuestra técnica.

En nuestro caso, las MSA de cada variable es muy alta (diagonal de la matriz de correlación anti-imagen), indicando su adecuación al análisis factorial.

Igualmente, podemos observar en la matriz anti-imagen que aparece en la Tabla 6, los coeficientes de correlación parciales de cada variable que nos aportarían la misma información.

Aunque en el caso de esta matriz los valores aparecen en negativo, como se ha explicado anteriormente, nos interesa que los coeficientes de correlación parciales sean lo más próximos a 0 posible, ya que eso significaría que las variables vendrán explicadas en mayor medida por sus factores comunes, y no por sus factores únicos o específicos.

Tabla 6: Matriz anti imagen

		Matrices anti-imagen								
		Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 9	Variable 10
Covarianza anti-imagen	Variable 2	,615	-,228	,020	,040	,043	,001	-,075	-,018	-,083
	Variable 3	-,228	,465	-,129	-,158	-,112	,038	-,002	,044	-,060
	Variable 4	,020	-,129	,609	-,127	,038	-,026	,014	-,087	-,022
	Variable 5	,040	-,158	-,127	,549	-,089	,024	-,018	-,035	-,072
	Variable 6	,043	-,112	,038	-,089	,576	-,146	,031	-,058	,039
	Variable 7	,001	,038	-,026	,024	-,146	,323	-,053	-,155	-,094
	Variable 8	-,075	-,002	,014	-,018	,031	-,053	,509	-,102	-,123
	Variable 9	-,018	,044	-,087	-,035	-,058	-,155	-,102	,332	-,034
	Variable 10	-,083	-,060	-,022	-,072	,039	-,094	-,123	-,034	,437
	Correlación anti-imagen	Variable 2	,825 ^a	-,425	,032	,068	,072	,002	-,134	-,040
Variable 3		-,425	,780 ^a	-,242	-,312	-,217	,099	-,004	,112	-,132
Variable 4		,032	-,242	,900 ^a	-,220	,064	-,059	,025	-,194	-,043
Variable 5		,068	-,312	-,220	,882 ^a	-,157	,057	-,034	-,081	-,148
Variable 6		,072	-,217	,064	-,157	,860 ^a	-,339	,058	-,132	,077
Variable 7		,002	,099	-,059	,057	-,339	,831 ^a	-,131	-,475	-,249
Variable 8		-,134	-,004	,025	-,034	,058	-,131	,911 ^a	-,248	-,262
Variable 9		-,040	,112	-,194	-,081	-,132	-,475	-,248	,856 ^a	-,089
Variable 10		-,161	-,132	-,043	-,148	,077	-,249	-,262	-,089	,909 ^a

a. Medidas de adecuación de muestreo (MSA)

De todos modos, podría resultar arriesgado eliminar variables del análisis utilizando únicamente las medidas que hemos expuesto. Sería adecuado atender a otros indicadores como, por ejemplo: los residuos del modelo, la interpretabilidad de los factores obtenidos o las comunalidades.

Las comunalidades nos indican qué proporción de la varianza explicada de cada variable se debe a causa de los factores comunes. En la Tabla 7 aparecen las comunalidades de las variables de nuestro estudio y en ellas observamos que en la mayoría de ellas toman valores altos o medio-altos.

Tabla 7: Comunalidades

Comunalidades		
	Inicial	Extracción
Variable 2	1,000	,550
Variable 3	1,000	,799
Variable 4	1,000	,488
Variable 5	1,000	,576
Variable 6	1,000	,471
Variable 7	1,000	,826
Variable 8	1,000	,589
Variable 9	1,000	,804
Variable 10	1,000	,644

Método de extracción: análisis de componentes principales.

4.3.- Extracción de factores:

Una vez que hemos confirmado que los datos de los que disponemos son adecuados para llevar a cabo la técnica de Análisis Factorial, deben estimarse los factores y determinar el número de ellos que representarán a nuestras variables originales.

Existen numerosos métodos para estimar los factores, pero en nuestro análisis emplearemos el “Método de los Componentes Principales” que consiste en estimar las puntuaciones factoriales mediante las puntuaciones tipificadas de los **k** primeros componentes principales y la matriz de cargas factoriales mediante las correlaciones de las variables originales con dichos componentes.

4.4.- Determinación del número de factores:

Entre aquellos componentes o factores que explican la información original de la que disponemos, generalmente un número más reducido de ellos es capaz de contener casi toda esta información. Por tanto, uno de los problemas es decidir cuál va a ser ese número exacto de factores que queremos conservar.

Entre algunos de los criterios más utilizados para llevar a cabo esta tarea se encuentran:

- Determinación “a priori”: Consiste en plantear desde el momento inicial cuál es el número de factores que hay y cuáles son. Este criterio resulta fiable cuando el investigador conoce a fondo el tema que desea analizar, así como los datos y las variables.
- Criterio de división a la mitad: La muestra es dividida en dos partes iguales al azar intentado que no existan diferencias significativas. Se realiza un Análisis Factorial en cada una de ellas y se conservan aquellos factores en los que se observe una alta correspondencia de cargas factoriales en las dos muestras.
- Pruebas de significación: Consiste en aplicar contrastes de hipótesis de modelos anidados para seleccionar el número de factores.

En nuestro estudio, para determinar el número de factores utilizaremos los siguientes métodos:

- Regla de Kaiser: Consiste en calcular los valores propios de la matriz de correlaciones de las variables originales y, tomar como número de factores el número de valores

propios superiores a la unidad. Esta regla tiende a infra-estimar el número de factores por lo que puede servir para establecer un número mínimo de ellos.

En la Tabla 8 observamos que existen dos auto-valores que son superiores a la unidad, por lo que en este momento podríamos confirmar la existencia de, al menos, dos factores.

Tabla 8: Varianza total explicada

Componente	Autovalores iniciales			Sumas de extracción de cargas al cuadrado			Sumas de rotación de cargas al cuadrado		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	4,580	50,886	50,886	4,580	50,886	50,886	3,217	35,744	35,744
2	1,167	12,968	63,854	1,167	12,968	63,854	2,530	28,110	63,854
3	,849	9,432	73,286						
4	,662	7,354	80,640						
5	,505	5,612	86,252						
6	,379	4,213	90,465						
7	,357	3,966	94,431						
8	,287	3,192	97,623						
9	,214	2,377	100,000						

Método de extracción: análisis de componentes principales.

- Criterio del porcentaje de la varianza: Consiste en determinar el número de factores como el mínimo número de componentes necesarios para que el porcentaje acumulado de varianza explicada sea un 75% u 80%.

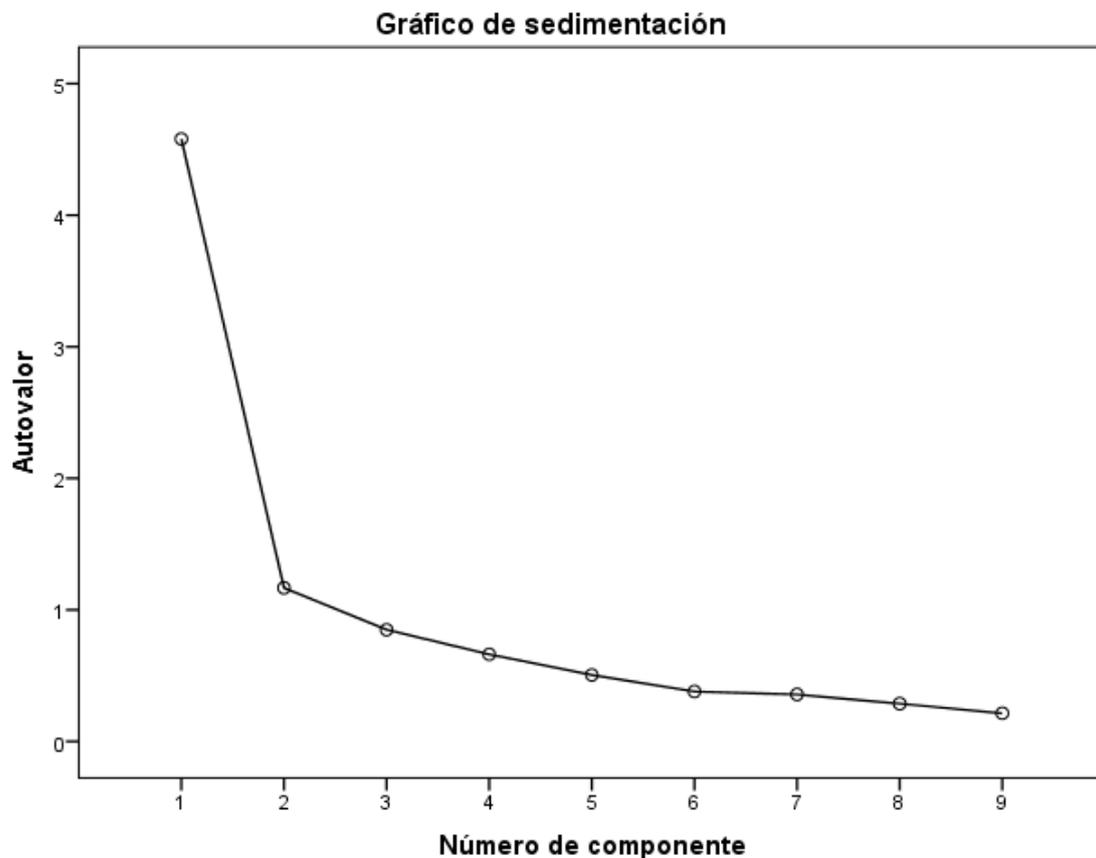
En nuestro análisis determinaremos que el porcentaje acumulado de varianza explicada sea aproximadamente un 64% ya que, como se observa en la Tabla 8 que acabamos de mostrar, entre los dos primeros componentes explican dicho porcentaje de varianza acumulada, mientras que los restantes sólo lo hacen en pequeñas proporciones.

Por tanto, a través de este criterio, confirmaríamos la elección de un número de factores igual a dos.

- Gráfico de sedimentación: Representa gráficamente el valor propio de cada uno de los componentes y suelen distinguirse visualmente aquellos cuyas varianzas son altas de aquellos que son bajas. El punto de inflexión que establezca la separación entre ellos, determinará que puede seleccionarse como número de factores el número de componentes que se encuentren antes de este punto de inflexión.

En la Figura 2 se observa que el punto de inflexión del gráfico se sitúa en el auto-valor correspondiente al segundo componente, lo que volvería a confirmar que el número de factores que seleccionaríamos sería igual a dos.

Figura 2: Gráfico de sedimentación



4.5.- Interpretación de los factores

La interpretación de los factores está basada en sus correlaciones estimadas respecto a las variables del inicio. En primer lugar, han de identificarse las variables cuyas correlaciones con el factor son las más elevadas en valor absoluto y, posteriormente se debe asignar un nombre a los factores relacionado con la estructura de sus correlaciones con las variables de modo que resulte claro percibir qué significa cada uno de los factores.

Dos estrategias más pueden ayudar a interpretar los factores: a) ordenarlos y b) eliminar las cargas bajas. Se puede ordenar la matriz factorial de tal forma que las variables con cargas altas para un mismo factor aparezcan juntas. La eliminación de cargas factoriales bajas también facilita la interpretación de los resultados al suprimir información redundante y es el investigador quién debe decidir a partir de qué valores deben eliminarse dichas cargas factoriales.

4.6.- Rotación de factores

La matriz de cargas factoriales es un elemento fundamental para interpretar el significado de los factores, pero a través de los métodos de extracción de factores del apartado 3, es inusual que obtengamos matrices de cargas factoriales fácilmente interpretables.

Para solucionar este problema y obtener una matriz de cargas factoriales cuya interpretación resulte más sencilla, existen procedimientos de “Rotación de factores”. Dichos métodos intentan aproximar la solución obtenida al “Principio de Estructura Simple”, por el que la matriz de cargas factoriales debe reunir las siguientes características:

- 1) Cada factor debe tener unos pocos pesos altos y los otros próximos a cero.
- 2) Cada variable no debe estar saturada más que en un factor.
- 3) No deben existir factores con la misma distribución, es decir, dos factores distintos deben presentar distribuciones diferentes de cargas altas y bajas.

De este modo, cada factor tendrá una correlación alta con un grupo de variables y una correlación baja con el resto de ellas. Al examinar las características del grupo de variables asociado a cada factor, podremos identificar sus rasgos comunes y dar paso a la denominación del factor que responda a esos rasgos comunes.

Existen dos formas básicas de realizar la rotación de factores: Rotación Ortogonal, a través de los métodos Varimax, Quartimax o Equamax y Rotación Oblicua, a través de los métodos Oblimin o Promax. En todos ellos, la comunalidad de la variable no se modifica, aunque sí lo hace la varianza explicada por cada factor.

La Tabla 9 nos muestra la matriz de componente, que nos indica que los factores no cumplen el principio de estructura simple ya que hay variables que se encuentran saturadas en más de un factor.

La Tabla 10 nos muestra la matriz de componente rotado, obtenida tras aplicar la técnica de rotación factorial Varimax. Tras la obtención de este resultado, eliminaremos las cargas factoriales que sean menores a 0,5, lo que da lugar a que ninguna variable se encuentre saturada en más de un factor y sea así más sencilla su interpretación

Tabla 9: Matriz de componente

Matriz de componente^a

	Componente	
	1	2
Variable 9	,805	
Variable 10	,801	
Variable 7	,789	-,451
Variable 8	,730	
Variable 5	,694	
Variable 4	,666	
Variable 3	,663	,599
Variable 6	,654	
Variable 2	,584	,457

Método de extracción: análisis de componentes principales.

a. 2 componentes extraídos.

Tabla 10: Matriz de componente rotado

Matriz de componente rotado^a

	Componente	
	1	2
Variable 7	,897	
Variable 9	,874	
Variable 8	,715	
Variable 10	,649	
Variable 6	,638	
Variable 3		,883
Variable 2		,723
Variable 5		,677
Variable 4		,584

Método de extracción: análisis de componentes principales.

Método de rotación: Varimax con normalización Kaiser.

a. La rotación ha convergido en 3 iteraciones.

De esta forma, nuestra dimensión queda reducida a 2 factores diferenciados en los que se agrupan nuestras variables originales y que explicamos a continuación:

- **Factor 1:** que puede denominarse “*Estado de Bienestar*” y recoge el grupo de variables 6, 7, 8, 9 y 10, referidos a aquellos aspectos económicos relevantes entre la opinión ciudadana como: corrupción, desempleo, sanidad, educación, etc.
- **Factor 2:** que puede denominarse “*Preocupación Social*” y en el que se encuentran recogidas las variables 2, 3, 4 y 5, relacionadas con aspectos sociales como, por ejemplo, la inmigración o la conservación del medioambiente.

4.7.- Cálculo de Puntuaciones Factoriales

Tras determinarse cuáles son los factores rotados, habrá de calcular la “Matriz de puntuaciones factoriales”. Dichas puntuaciones factoriales nos permitirían conocer qué sujetos serían los más raros o extremos, dónde se ubican ciertos grupos de la muestra o conocer en qué factor sobresalen unos sujetos y en qué factor no lo hacen.

Para el cálculo de dicha matriz de puntuaciones factoriales, el programa utilizado SPSS nos permite elegir entre tres de los métodos de estimación más utilizados: Regresión, Bartlett y Anderson-Rubin.

5.- ANÁLISIS CLÚSTER

El Análisis Clúster es una técnica multivariante de clasificación cuyo objetivo es la obtención de conglomerados o grupos de objetos. En este tipo de técnica, es fundamental que los objetos que componen un grupo compartan características muy similares entre ellos y, a su vez, que sean diferenciadoras del resto de grupos obtenidos. Por ello, podemos decir que debe existir un alto grado de homogeneidad interna y heterogeneidad externa entre clústeres o grupos.

A continuación, explicaremos de forma abreviada cada uno de estos métodos de clasificación y los criterios que se deben adoptar para implementarlos en una base de datos:

5.1.- Formulación del problema:

Para llevar a cabo el primer paso de nuestro análisis, es muy importante decidir si lo que deseamos obtener son agrupaciones de variables u objetos, o ambos. En nuestro caso, nos interesa conocer la variedad de grupos de votantes (objetos) de la muestra a través de un comportamiento similar en las elecciones generales de 2011, 2015 y 2016.

5.2.- Métodos de clasificación:

Para la asignación de los objetos de la muestra a los distintos grupos o clústeres, el programa SPSS nos ofrece tres tipos de métodos entre los que deberemos elegir en función de la base de datos que dispongamos: Clúster jerárquico, Clúster de k-medias y Clúster bietápico.

Si desde un primer momento queremos obtener un número fijo y determinado de grupos, habría de elegirse el método de k-medias ya que con este método dicho número de grupos se decide en el momento anterior de comenzar el análisis.

Por otro lado, si no conocemos el número de grupos a obtener y nuestro tamaño muestral es elevado, debemos elegir el método de Clúster bietápico. Si nuestro tamaño muestral es reducido, en ese caso escogeríamos el método de Clúster jerárquico.

5.2.1.- Clúster Jerárquico:

Este procedimiento se caracteriza por desarrollar una jerarquía o estructura en forma de árbol para la formación de grupos y puede realizarse de dos modos diferentes: Clúster aglomerativo o Clúster divisivo.

En el Clúster aglomerativo, los individuos de la muestra comienzan siendo un grupo cada uno de ellos. A lo largo del proceso van uniéndose aquellos grupos que son más similares entre sí hasta que en el último momento todos los grupos acaban formando uno solo que comprende a todos. Por ello, en este procedimiento, el analista debe decidir en qué paso paraliza el proceso. En el Clúster divisivo sucede lo contrario, todos los individuos comienzan formando parte de un solo grupo hasta que al final cada uno de ellos es un grupo individual.

Para llevar a cabo este análisis, habrán de tenerse en cuenta las siguientes cuestiones:

- *Selección del criterio para determinar la semejanza o distancia entre individuos ó medidas de asociación:*

La diferencia en el uso de criterios de semejanza y de distancia se encuentra en que los primeros miden cuán similares son dos individuos, mientras que los segundos miden el recorrido que existe entre ellos. Si se opta por una medida de distancia, como por ejemplo la distancia euclídea, los grupos formados contendrán individuos similares por existir una distancia pequeña entre ellos. Si por el contrario, se opta por una medida de semejanza, como el coeficiente de correlación, nos encontraríamos con grupos formados por individuos que comparten una similaridad alta entre ellos.

- *Selección del criterio de unión de grupos a lo largo del proceso ó medidas de conglomeración:*

A continuación, mencionaremos algunos de estos criterios de unión de grupos:

- Método del vecino más próximo ó enlace simple: Mide la proximidad entre dos grupos calculando la distancia entre los individuos más próximos o la similitud entre los individuos más similares.
 - Método del vecino más alejado o enlace completo: Mide la proximidad entre dos grupos calculando la distancia o la similitud entre los individuos más lejanos.
 - Método medio entre grupos o enlace medio intergrupos: Mide la proximidad entre dos grupos calculando la media de las distancias o similitudes entre individuos de ambos grupos.
 - Método medio dentro de los grupos o enlace medio intragrupos: Mide la proximidad entre dos grupos calculando la distancia media existente entre los miembros del grupo unión de los dos grupos.
 - Método de agrupación de centroides: Mide la proximidad entre dos grupos calculando la distancia entre sus medias para todas las variables.
 - Método de agrupación de medianas: Mide la proximidad entre dos grupos calculando la distancia entre sus medianas.
 - Método de Ward: Agrupa de forma jerárquica los elementos minimizando así la distancia euclídea al cuadrado.
- *Determinación del número de clústeres necesarios para representar nuestra base de datos:*

Para decidir el número de clústeres con el que realizar el análisis existen varias opciones. Por un lado, el programa SPSS proporciona las distancias a las que se forma cada grupo, así que un modo de determinar el número de grupos consiste en localizar en qué iteraciones dichas distancias pegan grandes saltos. Por otra parte, se puede calcular la tasa de variación entre los coeficientes de conglomeración que se han obtenido a lo largo de cada etapa y detenerse cuando el incremento en la tasa de variación se observe que es claramente elevado a la anterior.

5.2.2.- Clúster de *K*-medias:

Este método se caracteriza por conocerse de antemano el número de grupos o clústeres que se quieren obtener y se basa en el cálculo de distancias entre los individuos del conjunto de datos.

El proceso consiste en lo siguiente: una vez decidido un número *K* de grupos, se seleccionan a los *K* individuos o semillas más distantes entre sí. Posteriormente, se asigna cada observación de la muestra al grupo con semilla más cercana y se recalculan los puntos semillas de cada grupo. El proceso se detiene cuando las semillas apenas cambian o si los grupos obtenidos en dos iteraciones consecutivas son los mismos.

Un problema que presenta este tipo de método de clasificación, es que es muy sensible a los casos atípicos. Si existe algún atípico en nuestra base de datos, es probable que sea seleccionado como semilla inicial debido a la distancia a la que se encuentra de los demás por lo que su clúster no será realmente representativo. Para evitar este problema, sería necesario analizar previamente la base de datos de la que se dispone para eliminar estos casos atípicos.

5.2.3.- Clúster en dos pasos o Clúster bietápico:

Este último tipo de clúster, que es el que utilizaremos para llevar a cabo nuestro análisis en el presente estudio, se caracteriza por haber sido diseñado para manejar grandes bases de datos debido a que no se requiere un profundo análisis de ellos para su realización y por no conocer a priori el número de clústeres que se van a obtener. En este tipo de método, el algoritmo se basa en una medida de distancia en la que supone independencia entre las variables. El procedimiento está compuesto por dos etapas:

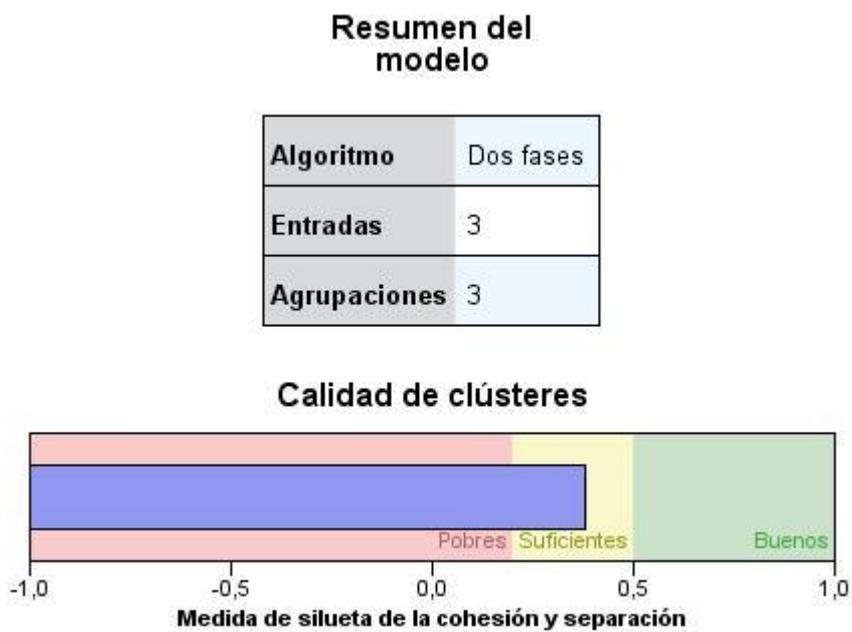
1ª.- Preclasificación: La primera etapa consiste en una clasificación secuencial en la que se forman pequeños conglomerados de individuos con el objetivo de reducir el tamaño de la matriz de distancias entre todos los pares de individuos posibles. Al finalizar esta preclasificación, estos “preclústeres” ya son tratados como una entidad por lo que la matriz de distancias se ve reducida al no depender de todos los individuos sino del número de “preclústeres”.

2ª.- Refinamiento: En el segundo paso se realiza un procedimiento jerárquico aglomerativo utilizando la clasificación de preclústeres obtenida en la primera etapa. El algoritmo selecciona el número final de clústeres obtenido basándose en el criterio de Schwarz (BIC) o en el de Akaike (AIC) encontrando el mayor incremento en la distancia entre los dos grupos más cercanos en cada paso del algoritmo jerárquico.

A continuación, mostraremos los resultados del análisis clúster sobre nuestro estudio, para el que se ha utilizado este último método en dos etapas.

En la Figura 3 se muestra el resumen del modelo y el gráfico de calidad de los clústeres. En la tabla de resumen del modelo se observa que se han obtenido 3 clústeres basados en las 3 variables referidas al voto de los individuos de la muestra en las elecciones de 2011, 2015 y 2016, mientras que el siguiente gráfico indica que la calidad de los clústeres es suficiente para poder proseguir con el análisis.

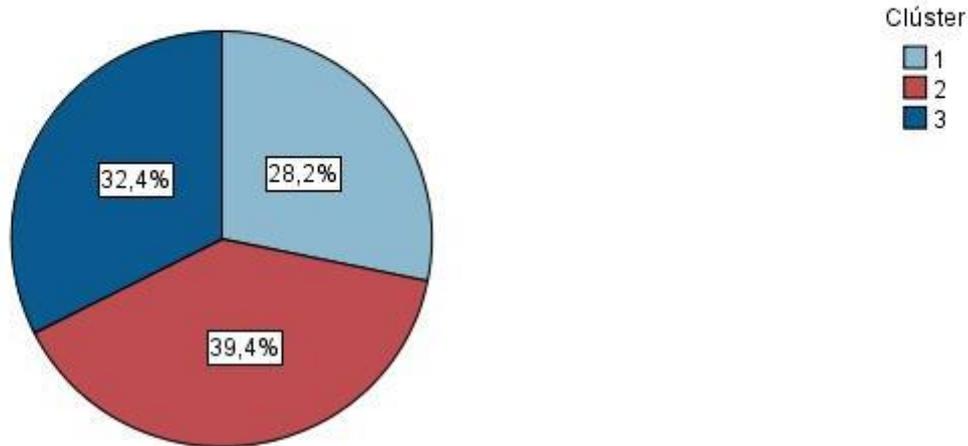
Figura 3: Resumen del modelo bietápico



En la Figura 4 se muestra el diagrama de sectores que refleja el tamaño de cada uno de los clústeres. El primer clúster contiene el 28,2% (53), el segundo contiene el 39,4% (74) y el tercero contiene el 32,4% (61) de los individuos encuestados.

Figura 4: Diagrama de sectores de tamaños de clústeres

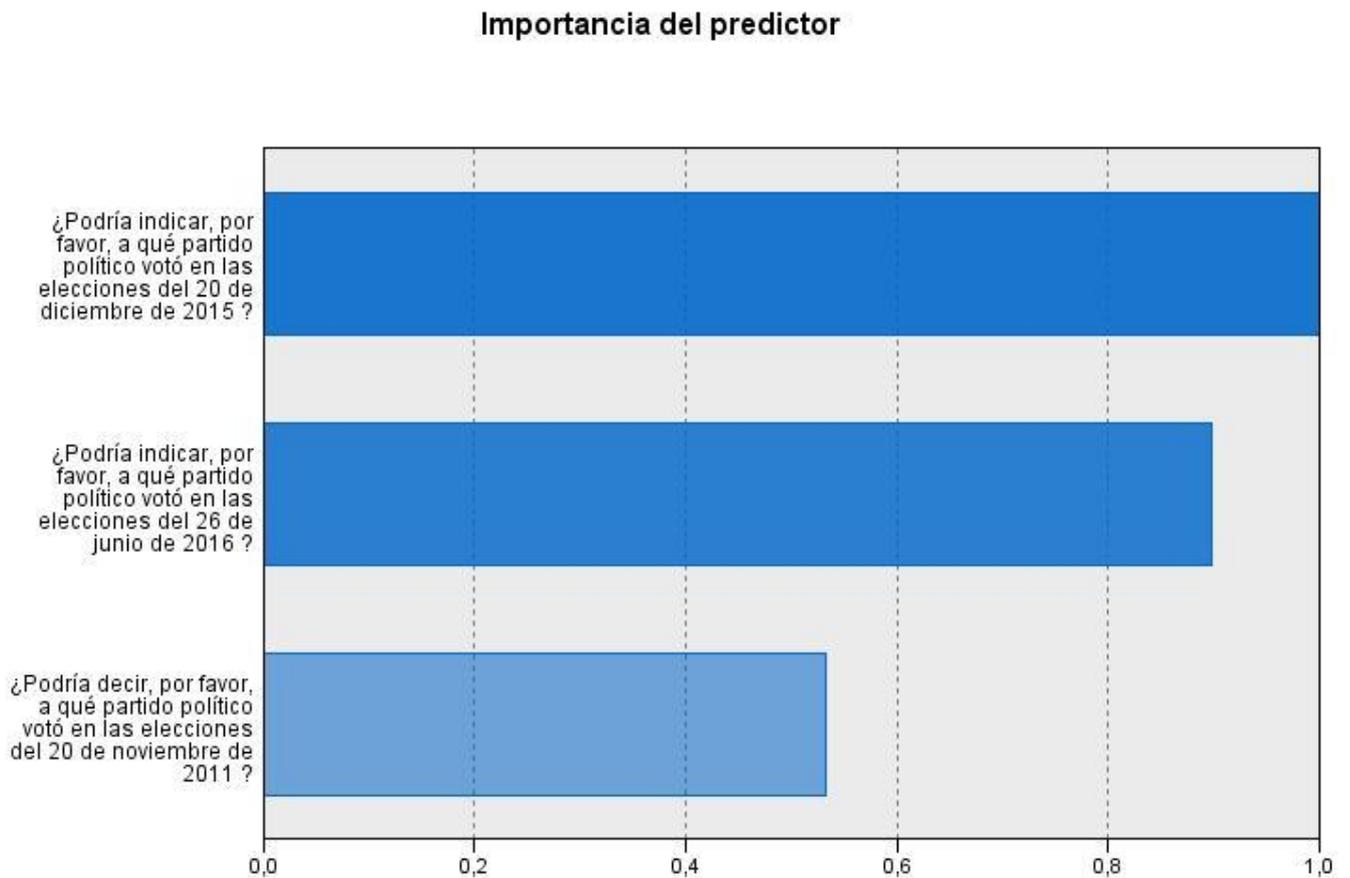
Tamaños de clústeres



Tamaño del clúster más pequeño	53 (28,2%)
Tamaño del clúster más grande	74 (39,4%)
Razón de tamaños: Del clúster más grande al clúster más pequeño	1,40

A continuación, en la Figura 5 se observa el peso o grado de importancia que han tenido cada una de las tres variables introducidas para la determinación de los clústeres obtenidos. La variable que más ha influido ha sido la respectiva a las elecciones del año 2015, debido a que es el momento en el que los individuos se reparten por primera vez entre un número mayor de partidos; y la variable con menos influencia se ha encontrado en la respectiva a las elecciones del año 2011, momento en el que los individuos se encontraban más agrupados en dos grandes partidos.

Figura 5: Importancia del predictor



La Figura 6 nos muestra una comparativa entre cada uno de los clústeres respecto a las variables introducidas relativas a las tres últimas convocatorias de elecciones generales. En el primer clúster, el programa SPSS nos indica que más del 94% de los individuos agrupados votaron al Partido Popular en las tres convocatorias consecutivas. El segundo clúster se encuentra formado por aquellos individuos que en mayor número no votaron en las elecciones de 2011, pero más tarde en los años 2015 y 2016 más del 32% de los individuos que componen el grupo votaron al partido Ciudadanos. Por último, casi el 40% del tercer clúster votó al Partido Socialista en el año 2011, pero la tendencia cambió en el año 2015 cuando el 37% del grupo se decantó por el partido Podemos y en el año 2016 este porcentaje creció hasta el 52%.

Figura 6: Comparación de clústeres en las Elecciones Generales de 2011, 2015 y 2016

Comparación de clústeres

■ 1 ■ 2 ■ 3



Las Figuras 7, 8 y 9 nos muestran las barras de error de cada una de las tres variables y en ellas puede observarse cómo se han distribuido de distinta manera cada uno de los clústeres respecto a ellas.

Figura 7: Barras de error de la variable “Elecciones 2011” por clúster

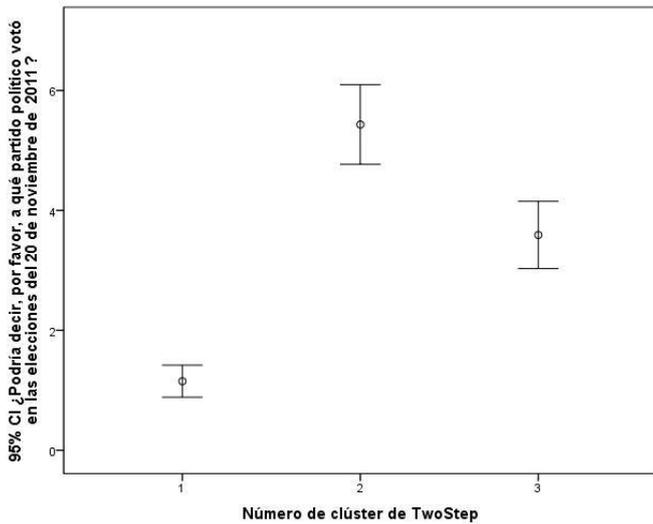


Figura 8: Barras de error de la variable “Elecciones 2015” por clúster

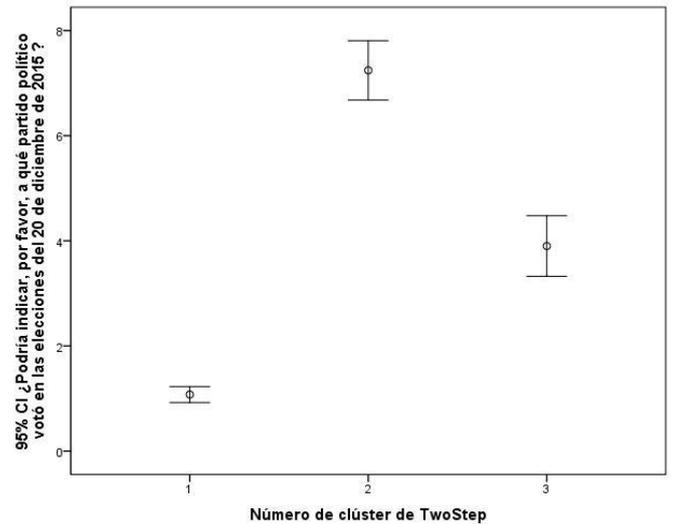
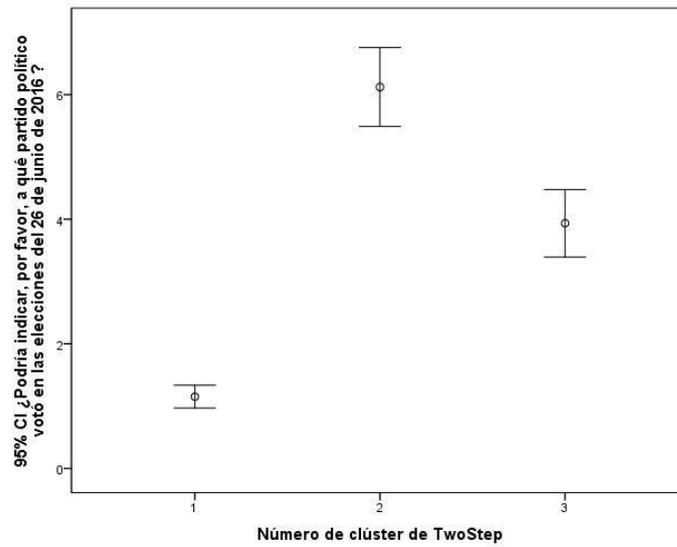


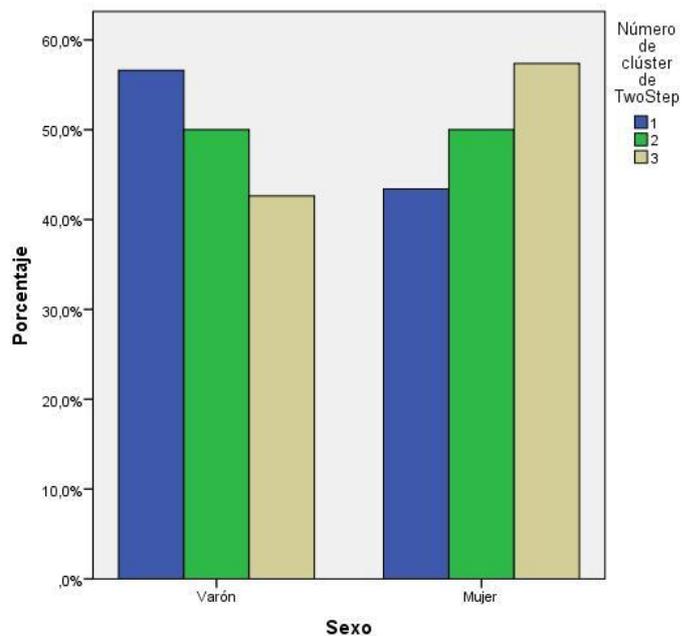
Figura 9: Barras de error de la variable “Elecciones 2016” por clúster



Una vez obtenidos los tres clústeres o grupos de individuos, realizaremos un análisis cruzado de dicha clasificación con los datos de las variables del estudio restantes, con el fin de conocer de qué forma están compuestos internamente los grupos.

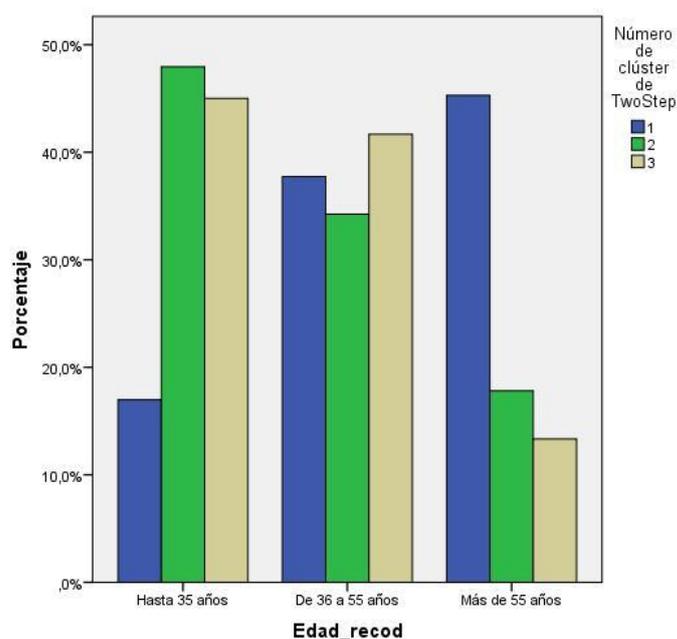
En la Figura 10 se observa que no hay grandes diferencias entre las proporciones de individuos que forman los clústeres en función del sexo. En el primer clúster, el porcentaje de hombres es ligeramente mayor al de mujeres y, sucede al contrario en el tercer clúster dónde el porcentaje de mujeres es ligeramente mayor al de hombres. Sin embargo, en el segundo clúster observamos que los individuos se distribuyen equitativamente en función del sexo.

Figura 10: Diagrama de barras de la variable “Sexo” por clúster



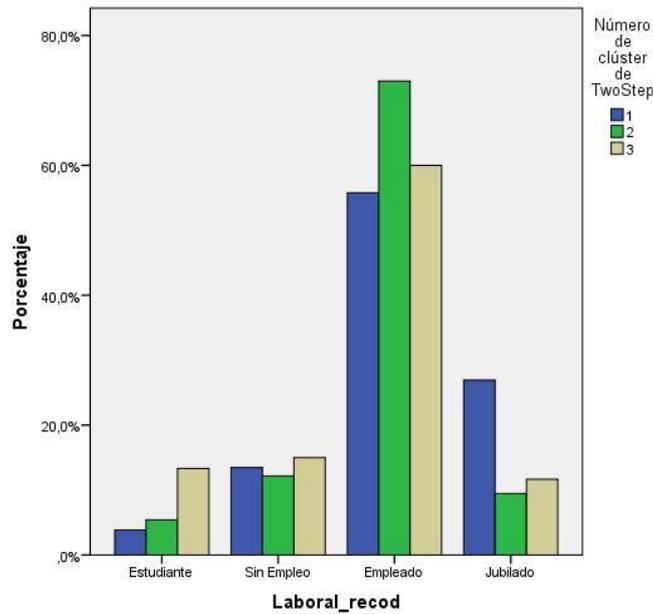
El gráfico correspondiente a la Figura 11 muestra claras diferencias relativas a la edad entre los tres clústeres, especialmente entre el primer clúster y los otros dos restantes. En el primer clúster se observa que la mayor proporción de individuos es la formada por aquellos que son mayores de 55 años y que las proporciones de individuos más jóvenes se vuelven más reducidas a medida que nos acercamos a edades menores. En el segundo y tercer clúster ocurre lo contrario. Los mayores porcentajes de individuos son jóvenes y se encuentran entre aquellos “Hasta 35 años”, reduciéndose a medida que aumenta la edad de los individuos.

Figura 11: Diagrama de barras de la variable “Edad” por clúster



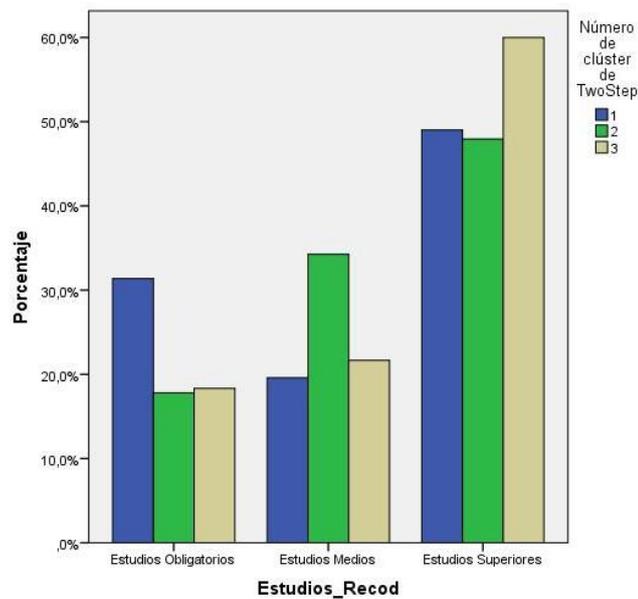
La Figura 12 nos muestra la composición de los tres clústeres en función de la situación laboral de los individuos. Podemos señalar que en los tres clústeres más del 50% de los individuos poseen empleo, aunque existe una diferencia evidente en el segundo clúster respecto al primero y tercero al contar con un porcentaje más elevado de individuos empleados. Si observamos a los individuos jubilados, destaca el mayor número de éstos provenientes del primer clúster. Sin embargo, si atendemos al segmento de estudiantes, es más elevada la proporción de individuos que pertenece al tercer clúster.

Figura 12: Diagrama de barras de la variable “Situación laboral” por clúster



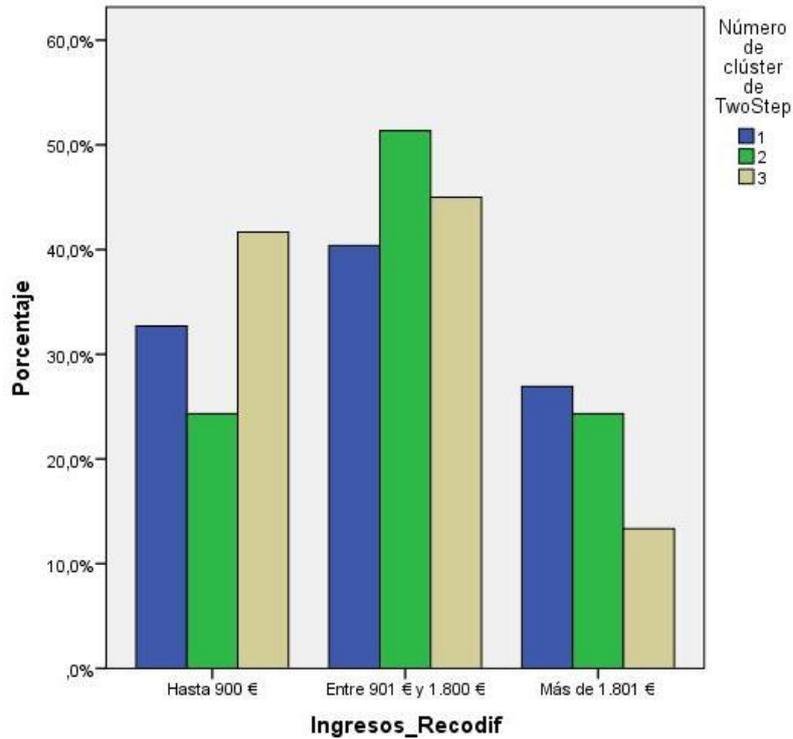
El siguiente gráfico correspondiente a la Figura 13, refleja el nivel máximo de estudios alcanzado por clúster. En este gráfico podemos destacar el elevado porcentaje de individuos con un nivel máximo de estudios que forman parte del tercer clúster en comparación con el primero y el segundo. También cabe señalar que en el primer clúster existe un amplio número de individuos, en contraste con los dos clústeres restantes, que únicamente cuenta con estudios obligatorios. Respecto al segundo clúster, se observa que las proporciones de individuos crecen progresivamente a medida que aumenta el nivel de enseñanza.

Figura 13: Diagrama de barras de la variable “Nivel máximo de estudios” por clúster



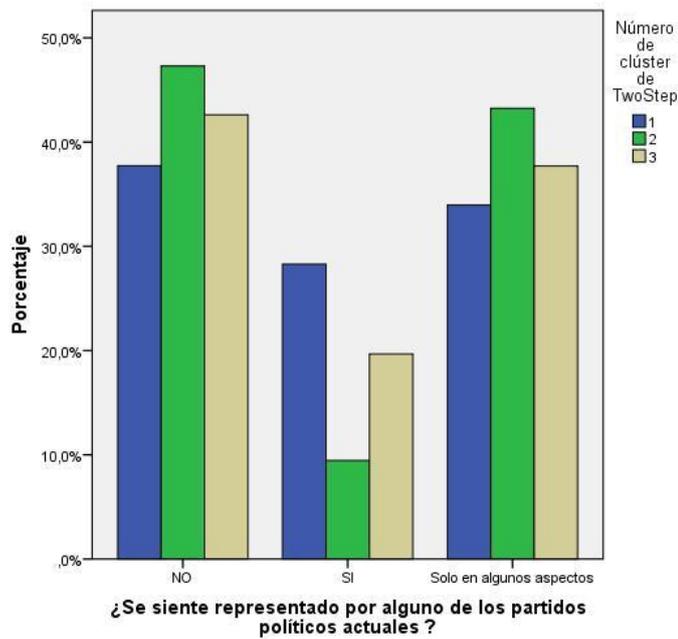
La Figura 14 muestra el nivel de ingresos netos mensuales por clúster. Si observamos el gráfico por tramos, podemos distinguir que el mayor porcentaje de individuos del primer tramo corresponde al tercer clúster. Más del 50% de los individuos del segundo clúster lo encontraríamos en el tramo correspondiente a “Entre 901 € y 1.800 €”, y el tercer tramo vemos que se encuentra liderado por individuos del primer clúster, aunque las proporciones de dicho tramo son menores en relación con el resto.

Figura 14: Diagrama de barras de la variable “Ingresos mensuales” por clúster



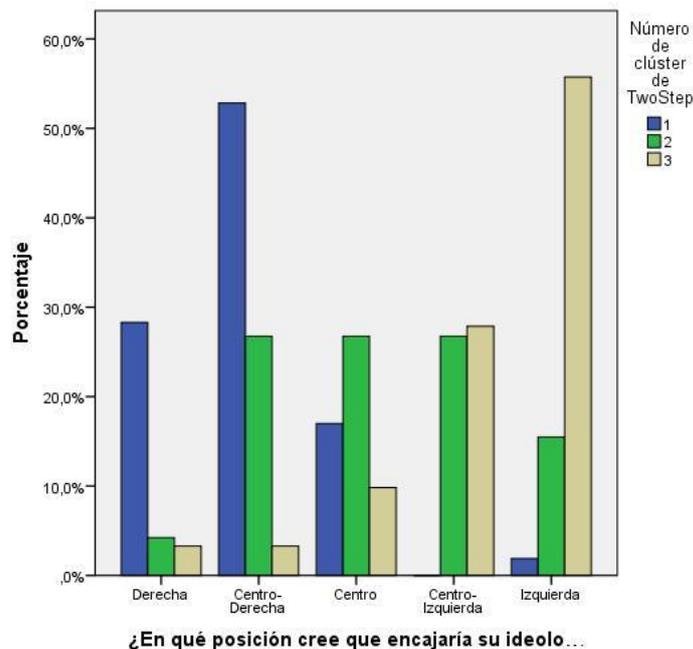
La Figura 15 muestra que existen diferencias entre los grupos en cuanto a si se sienten representados por los partidos políticos. El segundo clúster es el que contiene una proporción predominante de individuos “desencantados” con los partidos, seguido del tercer clúster; mientras que el primero sí que lo forma un mayor porcentaje de personas que afirman sentirse representadas políticamente.

Figura 15: Diagrama de barras de la variable “Representación política” por clúster



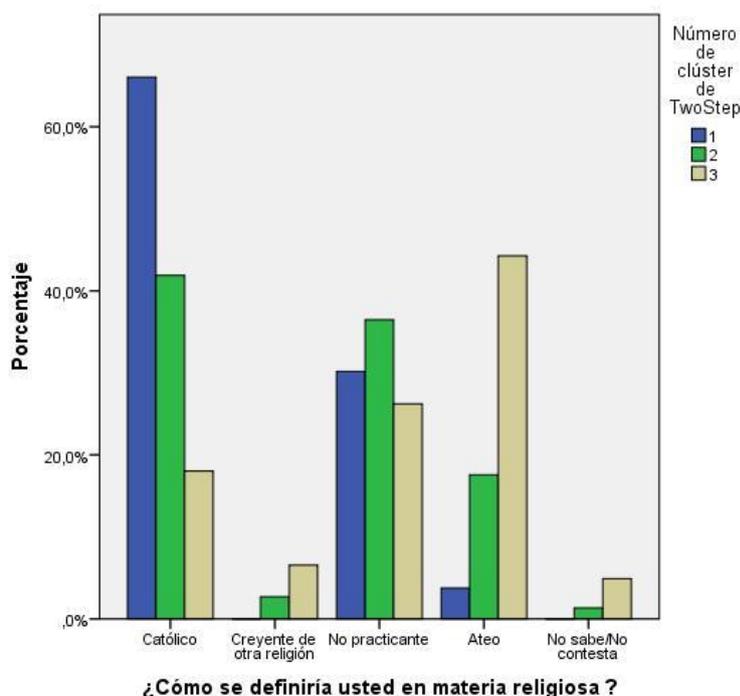
En la Figura 16 observamos que existen diferencias lógicas respecto a la variable de posición ideológica. El gráfico muestra con claridad cómo los individuos del primer clúster se han situado mayoritariamente entre las opciones “Derecha y Centro-Derecha”, mientras que aquellos que conforman el segundo clúster aparecen situados alrededor de la opción “Centro” y cómo los individuos del tercer clúster se han orientado hacia las posiciones de “Centro-Izquierda e Izquierda”.

Figura 16: Diagrama de barras de la variable “Posición ideológica” por clúster



La Figura 17 señala las diferencias existentes entre clústeres en materia religiosa. Se puede observar cómo la gran mayoría de individuos del primer clúster se definen como “Católicos” a diferencia del tercer clúster, que una amplia proporción se define como “Atea”. En último lugar, en el segundo clúster los individuos se reparten de forma similar entre las opciones “Católico” y “No practicante”.

Figura 17: Diagrama de barras de la variable “Religión” por clúster



A continuación, las Figuras 18 y 20 reflejan la valoración que han asignado cada uno de los clústeres a las variables “Aparición de Podemos” y “Aparición de Ciudadanos”, respectivamente. Asimismo, las Figuras 19 y 21 corresponden a los gráficos de las barras de error para estas variables ya que posiblemente su interpretación puede resultar más sencilla.

Respecto a la aparición de Podemos, destaca la valoración negativa que han asignado el segundo clúster y, con especial hincapié, el primer clúster donde más de la mitad de los individuos optó por la puntuación 0. El tercer clúster, como es lógico, atribuyó valoraciones altas y positivas a esta variable.

En cuanto a la aparición de Ciudadanos, podría decirse que el primer y segundo clúster han mantenido un comportamiento similar y sus valoraciones se sitúan en puntuaciones medias y altas, a diferencia del tercer clúster cuyas puntuaciones se encuentran en los valores medios y bajos.

Figura 18: Diagrama de barras de la variable “Aparición de Podemos” por clúster

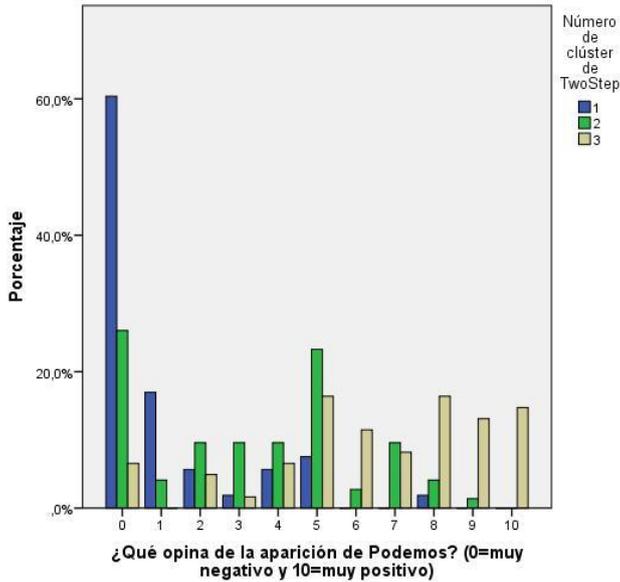


Figura 19: Barras de error de la variable “Aparición de Podemos” por clúster

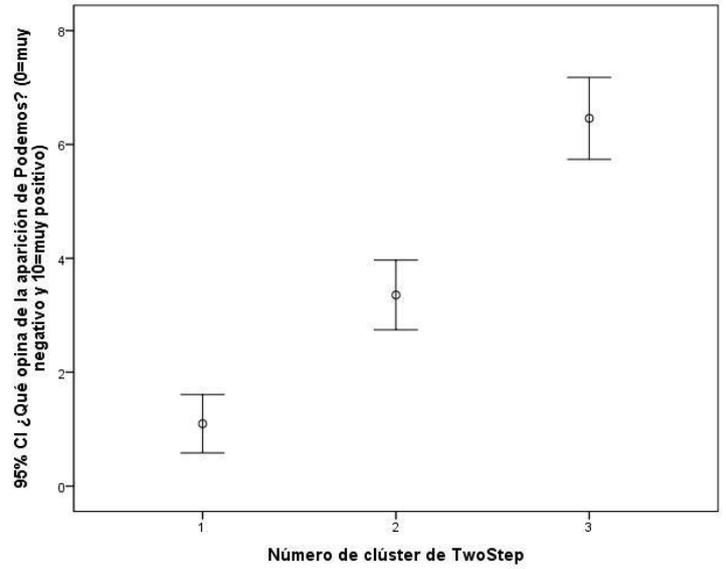


Figura 20: Diagrama de barras de la variable “Aparición de Ciudadanos” por clúster

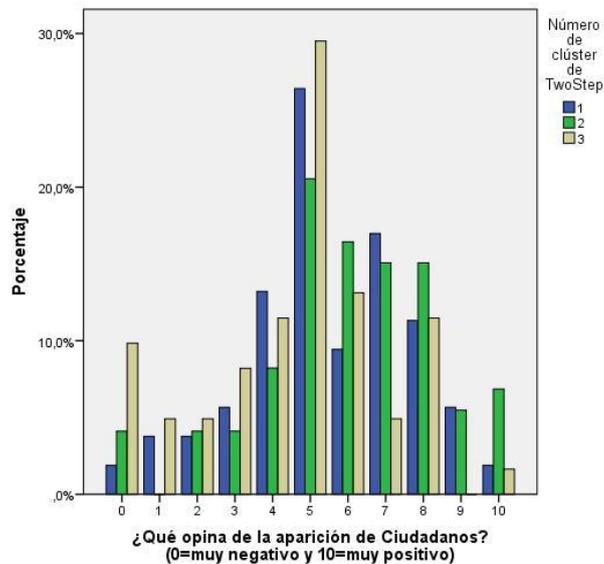
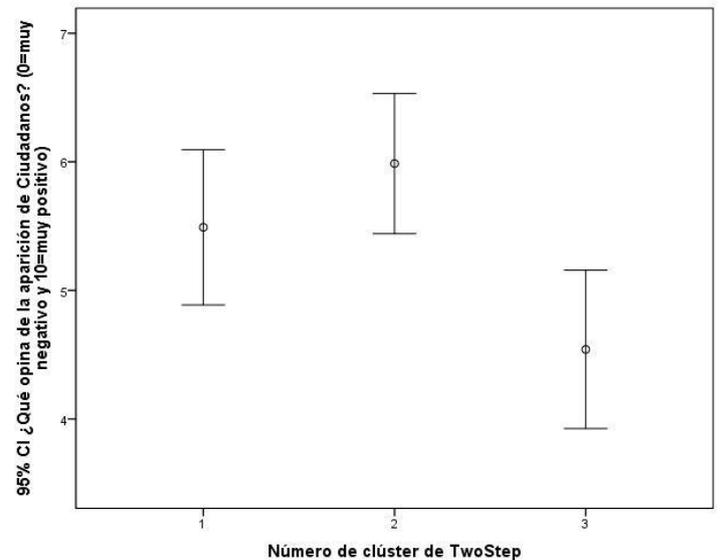


Figura 21: Barras de error de la variable “Aparición de Ciudadanos” por clúster



A continuación, las Figuras 22 y 23 corresponden a las barras de error por clúster de los factores “Estado de Bienestar” y “Preocupación Social”, respectivamente, obtenidos en el análisis factorial previo. La Figura 22 muestra cómo el factor “Estado de Bienestar” es valorado de forma similar en los tres clústeres, aunque cabe señalar cómo en el primer clúster existe un mayor rango de valoraciones bajas que en el resto de clústeres. Por el contrario, en cuanto al factor “Preocupación Social”, sí se observan diferencias evidentes entre clústeres en

la Figura 23, donde puede señalarse que el tercer clúster es el que mayor importancia le otorga a dicho factor, a diferencia del primero de ellos que es el que le da menor importancia.

Figura 22: Barras de error del Factor “Estado de Bienestar” por clúster

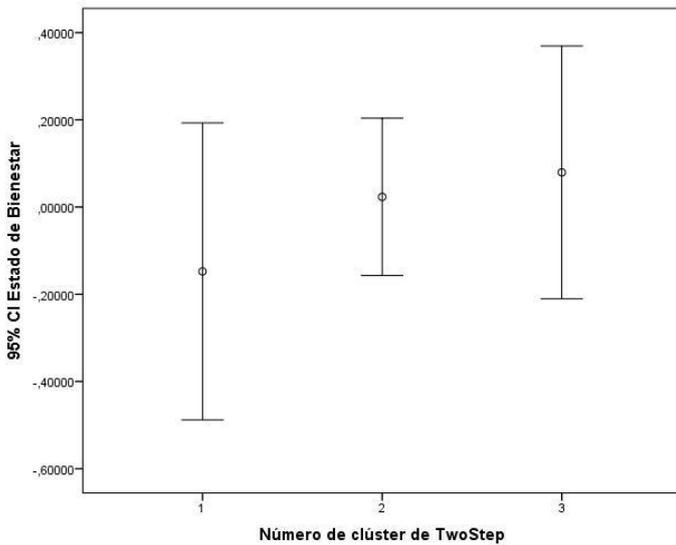
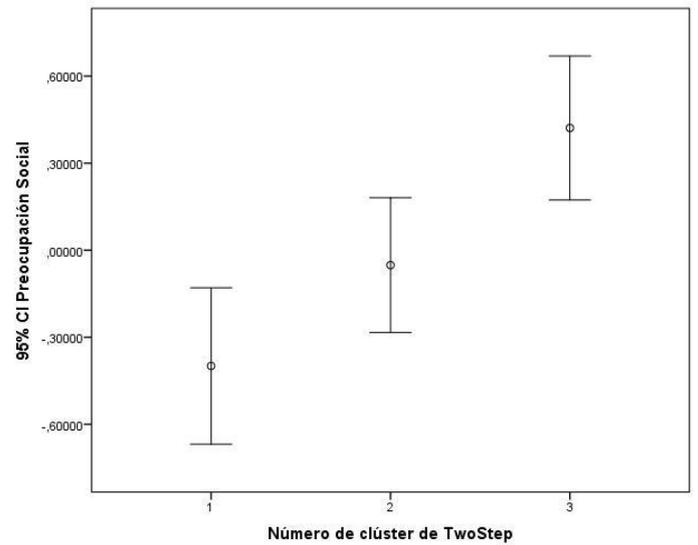
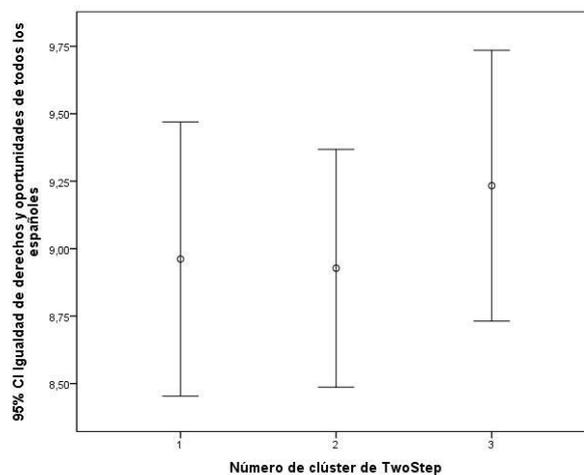


Figura 23: Barras de error del Factor “Preocupación Social” por clúster



Por último, la Figura 24 muestra las barras de error por clúster de la variable “Igualdad de Derechos y Oportunidades” que fue la única que no fue incluida dentro de los factores. Ello fue debido a que, como puede comprobarse en la figura, todos los individuos y, por tanto, todos los clústeres han tenido un comportamiento muy similar en lo que se refiere a esta variable a la que otorgaron un grado de puntuación muy alto.

Figura 24: Barras de error de la variable “Igualdad de derechos y oportunidades” por clúster



6.- CONCLUSIONES

La pretensión de este estudio ha sido analizar las características socio-demográficas de las personas con el fin de explicar a través de sus rasgos comunes un cambio en su comportamiento hacia el voto a los partidos políticos.

En primer lugar, para la obtención de los datos que se requerían en la elaboración del estudio, se entregó una encuesta a 196 zaragozanos con una serie de cuestiones que debían responder de forma anónima. Tras la introducción de todas las respuestas en el programa SPSS, se procedió a la aplicación de un Análisis Factorial entre un determinado grupo de variables, que los individuos debían puntuar según el grado de importancia que tenían para ellos, con el fin de simplificar la labor de estudio y comprensión de su información.

El resultado que aportó el Análisis Factorial fue la reducción de una serie de diez variables en únicamente dos factores que fueron denominados del siguiente modo:

- **“Factor 1: Estado de Bienestar”**, encargado de recoger las variables relativas a los aspectos económicos relevantes entre la opinión pública como: corrupción, desempleo, sanidad, educación, etc.
- **“Factor 2: Preocupación Social”**, relacionado con las variables de tipo social como por ejemplo: inmigración, medioambiente o reconocimiento de los matrimonios homosexuales.

A continuación, se dio paso a la aplicación de la segunda técnica estadística de nuestro estudio, el Análisis Clúster. El método utilizado en dicha técnica fue el de dos pasos o bietápico y se obtuvieron 3 clústeres o grupos de individuos diferenciados que, al cruzarlos con las variables del estudio y con los factores obtenidos anteriormente, aparecieron una serie de rasgos característicos que se exponen a continuación:

El **primer clúster** es el grupo de menor tamaño con 53 individuos de los cuales 30 son varones y 23 son mujeres. El rango de edad predominante en este clúster oscila entre 46 y 65 años. La mayoría se encuentran empleados aunque cabe destacar en este clúster que existe un subgrupo de personas jubiladas mucho mayor en comparación con los otros dos clústeres. El nivel máximo de estudios predominante es el de estudios universitarios, seguido de Educación Secundaria y un nivel de ingresos mensuales entre 900 € y 1.200 €. Se definen como católicos y con una ideología de centro-derecha. En este grupo se encuentra el mayor porcentaje de individuos que afirman sentirse representados políticamente. Además, en las tres convocatorias de elecciones estudiadas, se mantuvo fiel al Partido Popular, valoraron con un aprobado la aparición de Ciudadanos y de forma muy negativa la aparición de Podemos. En

general, a las cuestiones relativas al Estado de Bienestar les otorgan una importancia que oscila entre un nivel bajo y medio-alto, mientras que a las cuestiones relativas a la Preocupación Social, es el grupo que menor importancia les dan.

El segundo clúster es el grupo de mayor tamaño ya que está formado por 74 individuos cuyas edades oscilan en su mayor parte entre los 26 y los 45 años. En cuanto al género, no se aprecian diferencias, al tratarse de 37 varones y 37 mujeres. De todos ellos, la mayoría se encuentran empleados y predominan los individuos con estudios universitarios y de bachillerato. El nivel de ingresos predominante oscila entre 1.200 € y 1.800 €. Se definen principalmente como católicos o no practicantes y con una ideología de centro, centro-izquierda y centro-derecha. Casi todos los individuos han afirmado que no se sienten representados políticamente o únicamente en algunos aspectos y, en relación a su voto, se observa que este grupo no votó en las elecciones de 2011 para más tarde decantarse por el partido Ciudadanos en 2015 y 2016. La aparición de Podemos es valorada con una puntuación media de suspenso, mientras que la del partido Ciudadanos se sitúa entre el aprobado y el notable. Las cuestiones relativas al Estado de Bienestar las valoraron con puntuaciones medias-altas y a las referentes a la Preocupación Social se sitúan en una posición intermedia en comparación con los otros dos clústeres.

El tercer clúster se trataría del grupo de tamaño intermedio al contar con 61 individuos entre los cuales contamos con 26 varones y 35 mujeres. Comprende un mayor rango de edades, ya que el grupo se encuentra distribuido de forma muy equitativa entre los 18 y los 55 años. La mayoría de los individuos que componen el clúster se encuentran empleados, aunque se aprecia una presencia más numerosa de individuos en situación de desempleo y estudiantes que en los otros dos clústeres. El nivel máximo de estudios predominante en el grupo viene dado por los estudios universitarios o en grado superior y su nivel de ingresos mensuales oscila entre 1.200 € y 1.800 €. En su mayoría se definen como ateos o no practicantes y sitúan su ideología en la izquierda. Predomina el sentimiento de no representación o solo en algún aspecto por los partidos políticos y este grupo, que en el momento inicial entregó su voto al PSOE, en los años 2015 y 2016 dio su apoyo al partido Podemos. La aparición del partido Ciudadanos es valorada entre un aprobado y un notable, mientras que la del partido Podemos se encuentra entre el aprobado y el sobresaliente. Este grupo es el que mayor importancia da tanto a las cuestiones relativas al Estado de Bienestar como a las referentes a la Preocupación Social.

Una vez obtenida una visión general de cada uno de los grupos o clústeres, podemos deducir de ella las siguientes relaciones que podrían explicar los distintos comportamientos:

A pesar de la existencia de un descontento general de toda la muestra acerca de los representantes políticos, uno de los clústeres, el primero, ha sido el único en el que sus individuos se han mantenido fieles a uno de los dos partidos políticos tradicionales, en este caso, el Partido Popular. En este grupo destaca el hecho de que esté formado por personas de mayor edad que el resto de grupos, lo que explicaría que sus individuos sean menos propensos a apoyar un cambio que sea novedoso para ellos y que pueda generarles mayor incertidumbre. Por el contrario, los grupos en los que se da una presencia mayor de individuos más jóvenes, como por ejemplo los estudiantes, son aquellos cuyo voto ha sido dirigido hacia los nuevos partidos emergentes.

La variable renta no parece necesariamente determinante de un cambio de comportamiento ya que no se perciben grandes saltos o diferencias en los ingresos mayoritarios de los individuos de los distintos grupos. De igual forma se comporta la variable “Nivel educativo”, en la que se ha observado que el nivel máximo de estudios más frecuente era el universitario en todos los grupos.

Podría hablarse de un conjunto de variables que ejercerían su influencia como son: sexo, posición ideológica y grado de preocupación social. Si nos fijamos en la proporción de hombres y mujeres que conforman los clústeres, desde el primero hasta el último va incrementándose la presencia de mujeres hasta alcanzar la mayoría en el tercero. Lo mismo ocurre con el grado de preocupación social de los clústeres, que desde el primero hasta el tercero, dicho grado aumenta. También se ha observado cómo los clústeres se han posicionado ideológicamente en derecha, centro e izquierda para el primer, segundo y tercer clúster, respectivamente.

Por todo ello, y entendiendo que el primer clúster no cambió de voto y el segundo y tercero sí que lo hicieron, vemos cómo uno de los posibles perfiles de un individuo que ha cambiado de comportamiento podría verse como: mujer que se autodefine “de izquierdas” o “centro-izquierda” y con un alto grado de preocupación social; en contraposición al perfil de un hombre que se autodefine “de derechas” y con un grado de preocupación social medio-bajo.

La situación laboral de la muestra indica que, aunque la mayoría poseía empleo, el grupo en el que existía un mayor número de jubilados no cambió su voto a lo largo de las tres convocatorias de elecciones, frente a los grupos con mayor número de estudiantes y desempleados que sí lo hicieron. Aunque por el hecho que acabamos de nombrar, el de que la mayoría se encuentre empleada, es posible que la situación laboral de los individuos no sea un gran indicativo para un cambio de comportamiento.

Por último, en cuanto al aspecto religioso, se observa que existe una gran diferencia entre el número de católicos y ateos o no practicantes si se comparan el primer clúster con el tercero. A priori podría decirse que se trataría de una variable crucial en el comportamiento del votante pero, si observamos al segundo clúster, encontramos una mayor diversidad religiosa. Por ello, podría decirse que el aspecto religioso sería una variable que podría ir vinculada en mayor parte a una posición ideológica, o también a la edad, en lugar de tratarse de un determinante del comportamiento.

7.- BIBLIOGRAFÍA

Cifras de Zaragoza. Datos demográficos del padrón municipal (2015). Ayuntamiento de Zaragoza.

Centro de Investigaciones Sociológicas: www.cis.es

Hair, Joseph F. Multivariate data analysis [recurso electrónico] / Joseph F. Hair, Jr. 7th ed. Harlow: Pearson Education Ltd., [2014]

Peña Sánchez de Rivera, Daniel. Análisis de datos multivariantes / Daniel Peña [Reimp] Madrid [etc.]: McGraw-Hill, D.L. 2010

Pérez López, César. Técnicas de análisis de datos con SPSS 15 /César Pérez López. Madrid [etc.]: Pearson Educación, D.L. 2009

Rubiano Moreno, Jessica; Barreto Galeano, María Idaly. (Marzo, 2015). “Relación entre variables sociodemográficas, psicosociales y de marketing político en la conducta de voto y abstencionismo en jóvenes”. Revista Perspectiva Empresarial, Vol. 2, Nº 1, pp 61-74.

Unai Mezcu. “El bipartidismo se mantiene gracias al empuje del PP y se consolida como única suma estable”. Diario ABC (2016)

José María Pérez Medina. “¿Cómo y por qué acabó el bipartidismo en España?”. Diario El País (2016)