



UNIVERSIDADE DA CORUÑA

Universidade de Vigo

Proyecto Fin de Máster

---

# Estudio de Mesas Testigo en las Elecciones al Parlamento de Galicia

---

Alumna: Noema Afonso Casalderrey

Director: Salvador Naya Fernández

**Máster en Técnicas Estadísticas**

Universidade de Santiago de Compostela

Curso 2013 - 2014



Departamento de Matemáticas



**Salvador Naya Fernández**  
Escuela Politécnica Superior  
C/ Mendizábal, s/n  
15403 Ferrol, SPAIN.  
E-mail: [salva@udc.es](mailto:salva@udc.es)

Salvador Naya Fernández, Titular de Universidad del Departamento Matemáticas de la Universidad de A Coruña.

**Autoriza** la presentación del siguiente proyecto realizado bajo mi dirección y supervisión:

Trabajo Fin de Máster del Máster en Técnicas Estadísticas titulado “*Estudio de Mesas Testigo en las Elecciones al Parlamento de Galicia*”, elaborado por Noema Afonso Casalderrey.

Santiago de Compostela, 30 de junio de 2014



Fdo.: Salvador Naya Fernández



# Agradecimientos

Mis más sinceros agradecimientos al profesor Salvador Naya por su dedicación y por haberme descubierto el apasionante mundo de la investigación social.

Y a Manuel, compañero de fatigas.



# Índice

<b>1</b>	<b>Introducción</b>	<b>3</b>
<b>2</b>	<b>Estudio de Mesas Testigo en Galicia</b>	<b>7</b>
2.1	Elección de las variables . . . . .	7
2.1.1	Depuración de los datos . . . . .	8
2.2	Procedimiento de selección . . . . .	10
2.2.1	Medidas de asociación entre datos composicionales . . . . .	10
<b>3</b>	<b>Estudios Demoscópicos</b>	<b>19</b>
3.1	Estudio demoscópico de Galicia . . . . .	21
3.1.1	Contextualización . . . . .	21
3.1.2	Votos CERA . . . . .	22
3.1.3	Selección de las mesas testigo . . . . .	22
3.1.4	Selección del municipio testigo . . . . .	27
3.1.5	Baiona, ¿ciudad testigo de Galicia? . . . . .	29
3.2	Municipios parecidos a A Coruña . . . . .	35
3.3	Municipios parecidos a Lugo . . . . .	36
3.4	Municipios parecidos a Ourense . . . . .	37
3.5	Municipios parecidos a Pontevedra . . . . .	38
<b>4</b>	<b>Investigación Social</b>	<b>39</b>
4.1	Un ejemplo práctico:	
	Predicción electoral mediante la encuesta . . . . .	39
4.1.1	Selección del tamaño muestral . . . . .	40
4.1.2	Selección de los votantes a encuestar . . . . .	41
4.1.3	Ventajas e inconvenientes de la encuesta . . . . .	44
<b>5</b>	<b>Web Interactiva: Shiny</b>	<b>45</b>
5.1	Descripción de la web . . . . .	46
5.2	Ejecución de la web . . . . .	53
<b>A</b>	<b>Resultados</b>	<b>55</b>
	<b>Bibliografía</b>	<b>59</b>





# Capítulo 1

## Introducción

Toda entidad que esté dirigida a un público, desea conocer las opiniones, actitudes, tendencias, comentarios, inquietudes, temores del consumidor. Pero disponer de esta información no es tarea fácil. Lo ideal es poder extrapolar los resultados de una muestra pequeña a una gran población de estudio.

La idea de este trabajo surge una vez que se ha publicado la existencia de un pueblo testigo, como es Betanzos (“*Betanzos, magic town*”). El objetivo del proyecto es analizar otros pueblos similares de Galicia con el fin de predecir futuros resultados electorales. Concretamente, se buscan ejemplos de pueblos/ciudades testigo para cada una de las cuatro provincias gallegas. Para ello, se realiza un estudio sociodemográfico de Galicia, que se centra en los resultados electorales de las elecciones al parlamento gallego de los últimos años (2005, 2009 y 2012).

En 1977 se celebraron las primeras elecciones generales de la democracia por lo que nadie sabía con seguridad lo que iba a suceder. Las encuestas realizadas por el Instituto Gallup, las *macroencuestas*, eran las que mejor funcionaban. Se hacían entre seis meses y un año antes de cada elección y sus objetivos prioritarios eran el conocimiento de las demandas y orientaciones del electorado, las causas por las que pensaban acercarse a las urnas o abstenerse, los valores que predominaban en una sociedad recién salida de una dictadura y el modelo democrático hacia el que querían caminar. En un segundo plano quedaba el pronóstico electoral.

Los pronósticos electorales, al inicio de la democracia eran muy poco fiables, tanto por la falta de experiencia como por no conocer resultados anteriores de otras votaciones similares, pero a partir de 1977 las técnicas de marketing y encuestas progresaron a gran velocidad. Tanto es así, que cuando en 1985 se celebraron las segundas autonómicas gallegas, donde el recuento oficial todavía era muy lento, se fueron imponiendo varios métodos de predicción de resultados:

- Las israelitas: sondeos a pie de urnas realizados hasta tres horas antes del cierre.

- Los test de las cien primeras papeletas.
- Las mesas identificadas como muestra (que seguían en parte el ensayo de Robert L. Stoughton y Helen L. Merrell conocido por *Middletown*).
- Recuento adelantado de las actas o medias actas obtenidas por el grupo de interventores y apoderados de algunos partidos.

Uno de los métodos de predicción más revolucionarios en 1985, fue el de la llegada del voto de las primeras papeletas ideado por Xosé Luís Barreiro (vicepresidente de la Xunta de Galicia). Consistía en situar en cada mesa electoral un responsable que enviase, desde la cabina de teléfono más cercana, los primeros resultados electorales del recuento. El recuento válido seguía siendo el judicial, pero el proceso de Barreiro produjo una gran rapidez en la entrada de los datos (tanto parciales como totales) y la aleatoriedad de la entrada de estos datos hizo desaparecer la prioridad que residía en las mesas de las ciudades cercanas a los juzgados. Con todo ello, la mejora de las estimaciones fue extraordinaria.

Los estudios realizados por el matrimonio norteamericano Robert Lynd Stoughton y Helen Lynd Merrell durante los años veinte del siglo pasado, fueron el nacimiento para la ciencia sociológica de las ciudades que, por su composición social, representan a un estado. Es el primero de los llamados *Estudios Middletown* (1929).

Un estudio pionero en España es el realizado por el estadístico valenciano José Miguel Bernardo en 1984 sobre la aplicación de la metodología Bayesiana (el Teorema de Bayes permite obtener proporciones de votos a un partido *a posteriori* considerando ciertas imputaciones o probabilidades *a priori*). Las ideas de Bernardo se basan en encontrar mesas testigo, para ello analiza una gran cantidad de resultados anteriores buscando aquellas mesas que tienen menor “*distancia*” con los resultados finales. Esta técnica se realizó con gran éxito en distintas elecciones, tanto en España como en países extranjeros donde Bernardo trabajó como asesor para partidos políticos.

Hoy en día la prospección electoral sigue siendo imprescindible. Por lo que todo estudio electoral que ayude a conocer el desenvolvimiento de nuestra cultura democrática es necesario para este fin. Precisamente en el ensayo “*Betanzos, magic town*” (2013) escrito por el profesor Salvador Naya y la periodista Carmen Cotelo, basándose en los estudios realizados por el estadístico Bernardo, se investiga la posibilidad de que Betanzos sea la ciudad media (ciudad *middletown*) de Galicia.

Este proyecto de investigación, se ha inspirado en los estudios realizados por el matrimonio Lynd, por José Miguel Bernardo y por Salvador Naya. Se partirá de los resultados electorales de las últimas convocatorias al parlamento de Galicia y mediante técnicas estadísticas, como las distancias entre proporciones o el coeficiente de correlación de Pearson, buscará aquellas mesas y pueblos que mayor similitud tengan con

los resultados generales.

Como objetivo adicional se pretende hacer una aplicación Web usando la librería `shiny` de `RStudio` que facilite el análisis de resultados, tanto por provincias, como por municipios. Además, se espera que este proyecto facilite el trabajo de campo para futuros estudios demoscópicos en Galicia.



## Capítulo 2

# Estudio de Mesas Testigo en Galicia

El objeto de este apartado es presentar parte de la base teórica que aplicaremos en el estudio de las mesas electorales.

El fin del estudio es encontrar las mesas electorales y municipios que en su ámbito censal, mucho más pequeño y concreto que el de la comunidad gallega, reproduzcan con exactitud los resultados generales.

Para ello, partimos del estudio sobre determinación de mesas testigo [Bernardo, 1984] por lo que analizamos los resultados electorales de Galicia de los años 2005, 2009 y 2012 buscando aquellas mesas que más se semejen a los resultados totales.

Una forma de determinar la semejanza es midiendo la dependencia entre las variables. A mayor dependencia, más semejanza habrá entre ellas; y cuanto menor sea el grado de dependencia, menor relación tendrán.

### 2.1. Elección de las variables

A través de la página web <http://abertos.xunta.es> que la Xunta de Galicia ha creado para facilitar el acceso a determinados datos al público, recogemos los resultados oficiales de las elecciones al parlamento de Galicia.

Estos archivos (.csv) contienen los resultados de todas las mesas electorales de las elecciones al parlamento de Galicia de los años 2005, 2009 y 2012. El contenido de los archivos, está compuesto por una serie de datos distribuidos en forma de tabla  $n \times m$ . Donde las filas ( $n$ ) representan a todas las mesas electorales y las columnas ( $m$ ) representan a los datos sobre dicha mesa (código provincia, provincia, código municipio,

municipio, distrito, sección, mesa, censo, total votos, total abstención, votos en blanco, votos nulos, votos a candidaturas, votos válidos y el número de votos que recibieron cada partido político que participó ese año).

### 2.1.1. Depuración de los datos

En primer lugar, unificamos los tres archivos en uno. Añadimos una columna con el año de la elección. Las columnas que identifican a la mesa electoral permanecen sin variación (código provincia, provincia, código municipio, municipio, distrito, sección y mesa).

A continuación, nos centramos en las columnas de los partidos políticos. Vamos a tener en cuenta a los partidos que superan el 5% de los votos válidos, ya que es el umbral de voto exigido para entrar en el reparto de parlamentarios en la comunidad gallega. Los partidos que no alcancen dicho umbral los unificamos en una sola columna llamada OTROS.

#### • Galicia 2005:

Votos válidos: 1.672.556  $\xrightarrow{5\%}$  83.627,8  
 PP = 756.562 (> 83.627,8)  
 PSOE = 555.603 (> 83.627,8)  
 BNG = 311.954 (> 83.627,8)  
 OTROS = 27.525

#### • Galicia 2009:

Votos válidos: 1.690.975  $\xrightarrow{5\%}$  84.548,75  
 PP = 789.427 (> 84.548,75)  
 PSOE = 524.488 (> 84.548,75)  
 BNG = 270.712 (> 84.548,75)  
 OTROS = 78.277

#### • Galicia 2012:

Votos válidos: 1.443.848  $\xrightarrow{5\%}$  72.192,4  
 PP = 661.281 (> 72.192,4)  
 PSOE = 297.584 (> 72.192,4)  
 BNG = 146.027 (> 72.192,4)  
 AGE = 200.828 (> 72.192,4)  
 OTROS = 99.680

De este modo, tenemos 5 columnas en la tabla de datos:

-**PP:** Partido Popular

-**PSOE:** Partido dos Socialistas de Galicia - Partido Socialista Obrero Español

-**BNG:** Bloque Nacionalista Galego

-**AGE:** Alternativa Galega de Esquerda (Esquerda Unida-ANOVA)

-**OTROS:** La suma de todos los partidos y coaliciones que no han llegado al

umbral del 5 % de los votos válidos exigidos para optar a la representación parlamentaria.

Ahora, seleccionamos las variables imprescindibles de las siguientes: censo, total votos, total abstención, votos en blanco, votos nulos, votos a candidaturas y votos válidos. Veamos la relación entre dichas variables:

Censo := Total Votos + Total Abstención

Total Votos := Censo - Total Abstención

Total Abstención := Censo - Total Votos

Votos en Blanco := Votos Válidos - Votos a Partidos

Votos Nulos := Total Votos - Votos a Partidos - Votos en Blanco

Votos a Candidaturas := Votos a Partidos

Votos Válidos := Votos a Partidos + Votos en Blanco

donde con Votos a Partidos nos referimos a la suma de todos los votos que han recibido los partidos políticos (PP + PSOE + BNG + AGE + OTROS).

Finalmente, concluimos que las variables a tener en cuenta son:

-**Total Votos:** acto de votar (nulo, en blanco o a un partido).

-**Total Abstención:** acto de no votar, de no participar con el voto.

-**Votos en Blanco:** votar introduciendo un sobre vacío, manifestando que ninguna de las opciones presentadas es adecuada.

Por último, convertimos las columnas de los votos a partidos en proporciones, es decir, en datos composicionales. De este modo, ya están listos para trabajar con ellos.

Para mayor comprensión sobre la manipulación de los datos, en la Figura 2.1 podemos observar las 14 primeras mesas de la muestra.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	AÑO	C. PROV.	PROVINCIA	C. MUN.	MUNICIPIO	DISTRITO	SECCIÓN	MESA	T.VOTOS	T.ABST.	V.BLANCO	PP	PSOE	BNG	OTROS	AGE
2	2005	15	A Coruña	1	Abegondo	1	1	U	518	251	5	0,57	0,25	0,17	0,01	0
3	2005	15	A Coruña	1	Abegondo	1	2	A	225	117	2	0,6	0	0,26	0,14	0
4	2005	15	A Coruña	1	Abegondo	1	2	B	264	142	5	0,6	0,19	0,19	0,02	0
5	2005	15	A Coruña	1	Abegondo	2	1	U	455	252	4	0,49	0,3	0,21	0	0
6	2005	15	A Coruña	1	Abegondo	2	2	A	275	187	1	0,7	0,22	0,08	0	0
7	2005	15	A Coruña	1	Abegondo	2	2	B	224	133	3	0,55	0,23	0,22	0	0
8	2005	15	A Coruña	1	Abegondo	3	1	A	587	275	10	0,56	0,29	0,13	0,01	0
9	2005	15	A Coruña	1	Abegondo	3	1	B	309	141	3	0,42	0,27	0,29	0,03	0
10	2005	15	A Coruña	1	Abegondo	3	2	U	553	230	5	0,51	0,27	0,21	0,02	0
11	2005	15	A Coruña	2	Ames	1	1	A	336	114	2	0,54	0,31	0,14	0,02	0
12	2005	15	A Coruña	2	Ames	1	1	B	339	140	6	0,51	0,28	0,18	0,02	0
13	2005	15	A Coruña	2	Ames	1	3	A	183	51	2	0,39	0	0,48	0,12	0
14	2005	15	A Coruña	2	Ames	1	3	B	239	93	7	0,63	0,25	0,11	0	0

Figura 2.1: Primeros datos de la muestra.

## 2.2. Procedimiento de selección

En el análisis de las elecciones es necesario determinar la relación de cada una de las mesas electorales con el resultado total.

Los objetivos de dicho análisis son:

a) Determinar si las dos variables (mesa y resultado total) están correlacionadas, es decir, si los valores de una variable tienden a ser más altos o más bajos para valores más altos o más bajos de la otra variable.

b) Encontrar posibles mesas y municipios testigo de la comunidad gallega.

c) Valorar el nivel de concordancia entre los valores de las dos variables.

### 2.2.1. Medidas de asociación entre datos composicionales

Los datos composicionales son realizaciones de vectores aleatorios de suma constante. Así, el vector  $x = (x_1, x_2, \dots, x_q)'$  es un dato composicional si  $x_1, x_2, \dots, x_q > 0$  y  $\sum_{i=1}^q x_i = k$  con  $k$  constante.

El problema de encontrar parecidos entre proporciones ha sido y es una fuente de preocupación para muchos científicos. En 1897, Karl Pearson manifestó la inadecuación de los métodos estadísticos clásicos para el estudio de los mismos. Es indiscutible el interés por poseer herramientas adecuadas para su análisis, ya que estas medidas aparecen en numerosas ocasiones en las ciencias aplicadas (sociología, ciencias de la tierra, medicina, química, ingeniería, economía...).

El origen de todos los problemas puede encontrarse en la restricción de la suma constante, pues dificulta la aplicación de los procedimientos estadísticos habituales que se utilizan para datos que no presentan tal restricción.

Las principales dificultades que presenta trabajar con datos composicionales son las siguientes:

- La imposibilidad de interpretar correctamente las covarianzas y los coeficientes de correlación.
- La incoherencia en relación a las subcomposiciones.
- La distancia euclidiana no es una medida de diferencia adecuada entre este tipo de datos.
- La ausencia de familias paramétricas suficientemente flexibles para modelar los



conjuntos de datos composicionales.

Estas dificultades manifiestan la necesidad de elaborar nuevas metodologías compatibles con el carácter composicional de los datos. En 1982, Aitchison presenta una manera de evitar la restricción de la suma constante, conocida como la *geometría de Aitchison del simplex*. El simplex de  $q$  partes,  $S^q$ , es el espacio formado por los vectores cuyas  $q$  componentes son positivas y la suma de todas ellas es 1.

$$S^q = \{[x_1, x_2, \dots, x_q] : x_i > 0 (i = 1, \dots, q); x_1 + x_2 + \dots + x_q = 1\} \quad (2.1)$$

Aitchison argumenta que las dificultades de interpretación vienen dadas porque nuestro interés se centra en las magnitudes absolutas de las partes  $x_1, x_2, \dots, x_q$  de una composición. Cuando lo que deberíamos hacer es centrarnos en la magnitud relativa de las partes, es decir, en los cocientes  $\frac{x_i}{x_j} \forall i, j = 1, 2, \dots, q$  con  $i \neq j$ . Estamos ante un principio fundamental de los datos composicionales que Aitchison denomina *invarianza por cambios de escala*. Así problemas de las falsas correlaciones, con los cocientes se desvanecen.

También la incoherencia en relación a las composiciones desaparece, esto es, la magnitud relativa de las partes de una composición no cambia en relación a la magnitud relativa de las partes de la composición original.

El estudio de Aitchison consiste en la transformación de los datos composicionales al espacio real multivariante. Esta estrategia se remonta al trabajo de McAlistar (1879) que desarrolló los pilares de la ley lognormal univariante tomando el algoritmo de los datos. Entonces, si transformamos los cocientes mediante logaritmos, posteriormente podemos aplicar cualquier técnica estadística clásica.

Existen distintas formas de transformar los datos, todas basadas en los logaritmos cocientes entre las partes de un dato composicional. Veámoslas:

- La transformación logcociente aditiva (**alr**) de  $x \in S^q$  a  $y \in \mathbb{R}^{q-1}$ :  
Es una transformación biyectiva pero no simétrica en las partes de  $x$ . La transformación **alr** viene definida por:

$$y = alr(x) = \left( \ln \left( \frac{x_1}{x_q} \right), \ln \left( \frac{x_2}{x_q} \right), \dots, \ln \left( \frac{x_{q-1}}{x_q} \right) \right)' \quad (2.2)$$

- La transformación logcociente centrada (**clr**) de  $x \in S^q$  a  $z \in \mathbb{R}^q$ :  
Introducida por Aitchison en 1986, es una transformación biyectiva y simétrica entre las partes. La transformación **clr** está definida por:

$$z = clr(x) = \left( \ln \left( \frac{x_1}{g(x)} \right), \ln \left( \frac{x_2}{g(x)} \right), \dots, \ln \left( \frac{x_q}{g(x)} \right) \right)' \quad (2.3)$$

donde  $g(x)$  es la media geométrica de las  $q$  partes de  $x$ . Asimismo, la suma de las componentes del vector transformado es igual a cero (es ortogonal al vector de unidades) y su imagen es el hiperplano de  $\mathbb{R}^q$  pasando por el origen.

Llegados a este punto, surge un nuevo problema, la matriz de covarianzas del vector transformado es singular. Para ello, Aitchison tiene una estrategia: si existe simetría en el tratamiento de sus componentes, aplica la transformación **clr**. Pero si pretende modelar conjuntos de datos composicionales con distribuciones multivariantes, aplica la transformación **alr**. De este modo evita trabajar con distribuciones degeneradas.

En definitiva, al trabajar con datos composicionales es necesario desarrollar nuevas metodologías, ya que al ser la restricción de la suma de las partes de un dato composicional igual a una constante, esto provoca que los métodos estadísticos clásicos sean inadecuados. Las nuevas metodologías se fundamentan en la línea iniciada por Aitchison (1982) que introduce las transformaciones **alr** y **clr** con el objeto de trasladar el problema a un espacio real multidimensional sin restricciones.

A continuación, veamos alguno de los métodos clásicos de inferencia sobre proporciones.

### Distancia de Aitchison:

Una de las dificultades que presenta trabajar con datos composicionales es que la distancia euclidiana no es una medida de diferencia adecuada entre este tipo de datos. Aitchison introduce una nueva métrica  $d_a : S^q \times S^q \rightarrow \mathbb{R}^+$  que recibe el nombre de *distancia Aitchison* y se define como sigue:

$$d_a = \left[ \sum_{i=1}^q \left( \ln \frac{x_i}{g(X)} - \ln \frac{y_i}{g(Y)} \right)^2 \right]^{1/2} = \left[ \frac{1}{q} \sum_{i < j} \left( \ln \frac{x_i}{x_j} - \ln \frac{y_i}{y_j} \right)^2 \right]^{1/2} \quad (2.4)$$

donde  $g$  es la media geométrica de las partes de la composición.

La distancia de Aitchison satisface los axiomas usuales de distancia:

- i)  $d_a(x, y) \geq 0 \quad \forall x, y \in S^q$
- ii)  $d_a(x, y) = 0 \Leftrightarrow x = y, x \in S^q$
- iii)  $d_a(x, y) = d_a(y, x) \quad \forall x, y \in S^q$  (simetría)

iv)  $d_a(x, z) \leq d_a(x, y) + d_a(y, z) \forall x, y, z \in S^q$  (desigualdad triangular)

### Coefficiente de correlación de Pearson:

El coeficiente de correlación de Pearson, pensado para variables cuantitativas, es un valor estadístico que mide la relación lineal entre dos variables. Insistimos en que la relación que se mide es lineal, puesto que puede haber variables fuertemente relacionadas de otra forma que no sea la lineal, en cuyo caso no procedería aplicar la correlación de Pearson. Por ejemplo, la relación entre población y tiempo no es una relación lineal sino exponencial. En este caso (y en otros muchos) no es conveniente utilizar la correlación de Pearson.

El coeficiente de correlación de Pearson es normalmente denotado como  $r$ . Sus valores absolutos oscilan entre 0 y 1. Esto es, si tenemos dos variables  $X$  e  $Y$ , y definimos el coeficiente de correlación de Pearson entre estas dos variables como  $r_{xy}$  entonces:

$$0 \leq |r_{xy}| \leq 1 \quad \Rightarrow \quad -1 \leq r_{xy} \leq 1 \quad (2.5)$$

Si observamos el signo del coeficiente de correlación de Pearson, tenemos que varía en el intervalo  $[-1, 1]$ . Cabe resaltar que la magnitud de la relación viene especificada por el valor numérico del coeficiente, el signo indica la dirección de dicho valor. Por lo tanto, tan fuerte es una relación de  $+1$  como de  $-1$  (relación perfecta positiva y relación perfecta negativa, respectivamente). Para continuar, veamos con más detalle estos conceptos.

- Si  $r = 1$ , existe una correlación positiva perfecta. El índice indica una dependencia total entre las dos variables denominada relación directa: cuando una de ellas aumenta, la otra también lo hace en proporción constante.

- Si  $0 < r < 1$ , existe una correlación positiva.

- Si  $r = 0$ , no existe relación lineal. Esto no implica necesariamente que las variables sean independientes: pueden existir todavía relaciones no lineales entre las dos variables.

- Si  $-1 < r < 0$ , existe una correlación negativa.

- Si  $r = -1$ , existe una correlación negativa perfecta. El índice indica una dependencia total entre las dos variables llamada relación inversa: cuando una de ellas aumenta, la otra disminuye en proporción constante.

Para una mayor comprensión de las posibles relaciones entre variables, en la Figura 2.2 disponemos de varios ejemplos de diagramas de dispersión con diferentes valores del coeficiente de correlación de Pearson.

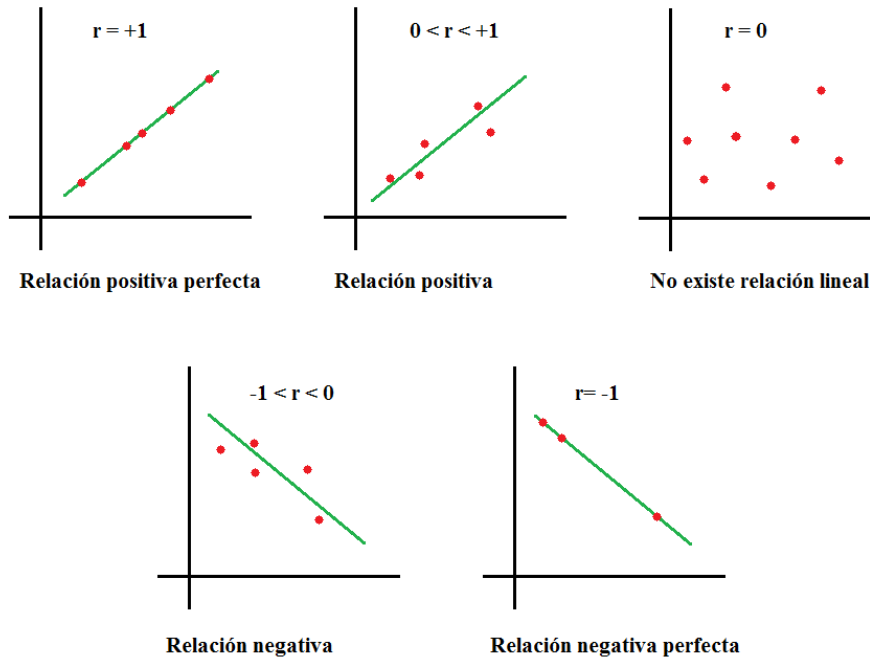


Figura 2.2: Ejemplos de diagramas de dispersión con diferentes valores del coeficiente de correlación de Pearson ( $r$ ).

El coeficiente de correlación de Pearson hace referencia a la media de los productos cruzados de las puntuaciones estandarizadas. Viene definido por la siguiente expresión:

$$r_{xy} = \frac{\sum Z_x Z_y}{N} \quad (2.6)$$

donde  $Z_x = \frac{X - \bar{X}}{S_x}$  y  $Z_y = \frac{Y - \bar{Y}}{S_y}$ .

Al operar con puntuaciones estandarizadas tenemos un índice libre de escala de medida. Estas puntuaciones muestran la posición en desviaciones tipo de un individuo respecto a su medida. Reflejan la medida en que dicho individuo se separa de la media. Su valor oscila, como ya hemos indicado, en términos absolutos, entre 0 y 1.

Podemos expresar la fórmula 2.6 de las formas siguientes:

- Las medias y las desviaciones tipo de X e Y son conocidas:

$$\begin{aligned}
r_{xy} &= \frac{\sum Z_x Z_y}{N} = \frac{\sum \left( \frac{X-\bar{X}}{S_x} \cdot \frac{Y-\bar{Y}}{S_y} \right)}{N} = \frac{\sum (X - \bar{X}) \cdot (Y - \bar{Y})}{N S_x S_y} \\
&= \frac{\sum (XY - X\bar{Y} - \bar{X}Y + \bar{X}\bar{Y})}{N S_x S_y} = \frac{1}{S_x S_y} \left( \frac{\sum XY}{N} - \frac{\bar{Y} \sum X}{N} - \frac{\bar{X} \sum Y}{N} + \frac{N \bar{X} \bar{Y}}{N} \right) \\
&= \frac{1}{S_x S_y} \left( \frac{\sum XY}{N} - \bar{X} \bar{Y} - \bar{X} \bar{Y} + \bar{X} \bar{Y} \right) = \frac{\frac{\sum XY}{N} - \bar{X} \bar{Y}}{S_x S_y} \quad (2.7)
\end{aligned}$$

- Puntuaciones directas (en caso de no disponer de la información de los estadísticos):

$$\begin{aligned}
r_{xy} &= \frac{\frac{\sum XY}{N} - \bar{X} \bar{Y}}{S_x S_y} = \frac{\frac{\sum XY}{N} - \frac{\sum X}{N} \frac{\sum Y}{N}}{\sqrt{\frac{\sum X^2}{N} - \left( \frac{\sum X}{N} \right)^2} \cdot \sqrt{\frac{\sum Y^2}{N} - \left( \frac{\sum Y}{N} \right)^2}} \\
&= \frac{N \sum XY - \sum X \sum Y}{\sqrt{N \sum X^2 - (\sum X)^2} \cdot \sqrt{N \sum Y^2 - (\sum Y)^2}} \quad (2.8)
\end{aligned}$$

- Puntuaciones diferenciales o centradas:

Definimos  $x = X - \bar{X}$  e  $y = Y - \bar{Y}$ , entonces

$$\begin{aligned}
r_{xy} &= \frac{\sum Z_x Z_y}{N} = \frac{\sum \left( \frac{X-\bar{X}}{S_x} \cdot \frac{Y-\bar{Y}}{S_y} \right)}{N} = \frac{\sum (X - \bar{X}) \cdot (Y - \bar{Y})}{N S_x S_y} \\
&= \frac{\sum (X - \bar{X}) \cdot (Y - \bar{Y})}{N \sqrt{\frac{\sum (X-\bar{X})^2}{N}} \cdot N \sqrt{\frac{\sum (Y-\bar{Y})^2}{N}}} = \frac{\sum xy}{N \sqrt{\frac{\sum x^2}{N}} \cdot N \sqrt{\frac{\sum y^2}{N}}} \\
&= \frac{\sum xy}{\sqrt{\sum x^2} \cdot \sqrt{\sum y^2}} \quad (2.9)
\end{aligned}$$

### Divergencia de Kullback-Leiber:

Sean  $X$  e  $Y$  dos variables aleatorias discretas, con distribuciones de probabilidad  $\{p_1, p_2, \dots, p_n\}$  y  $\{q_1, q_2, \dots, q_n\}$  respectivamente. Definimos la divergencia de Kullback-Leibler (KL) de  $X$  respecto de  $Y$  como:

$$D(X, Y) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i} \quad (2.10)$$

La divergencia de Kullback-Leibler entre dos distribuciones de probabilidad es no negativa, siendo nula únicamente si las dos distribuciones son idénticas. Nos permite conocer la diferencia o discrepancia que existe entre las dos distribuciones de probabilidad.

En la teoría de la información, la divergencia de Kullback-Leibler también es conocida como divergencia de la información, ganancia de la información o entropía relativa.

En nuestro caso, las variables serán:

- X: Proporción de los resultados totales en las elecciones.
- Y: Proporción de los datos obtenidos en cada una de las mesas electorales.

### Test de correlación de Szekely:

Para corroborar lo calculado anteriormente (CCP y KL), realizamos el test de correlación de Szekely, introducido por María Rizzo y Gabor Szekely (2007). Este coeficiente de correlación es una medida multivariante de dependencia y se encuentra implementado en el Software R, concretamente en la librería `energy`.

El comando empleado es:

```
dcov.test(x, y, index=1.0, R=199)
```

Los argumentos están definidos por:

- `x`: Resultados de cada mesa electoral.
- `y`: Resultados totales de las elecciones.
- `index`: Exponente de la distancia euclídea en el intervalo  $(0, 2]$ .
- `R`: Número de réplicas.

El test de correlación de Szekely realiza una prueba no paramétrica de independencia multivariante. La decisión del test la obtenemos a través de permutaciones bootstrap, con R repeticiones. Este test está basado en las distancias euclidianas ( $\|x_i - x_j\|$ ). El argumento `index` es un exponente opcional en la distancia euclídea.

Los tamaños de las dos muestras tienen que concordar y no deben de contener valores perdidos. Los argumentos `X` e `Y` son, como ya hemos indicado, los datos electorales.

Las propiedades del test de Szekely:

1)  $R(X, Y) = 0 \Leftrightarrow X$  e  $Y$  son independientes.

2)  $0 \leq R(X, Y) \leq 1$ , cuanto más próximo sea su valor a 1, las variables son más dependientes, por el contrario, cuanto más se acerque a 0, las variables son independientes.

### Ley D´Hondt:

Otro problema de trabajar con datos composicionales en las predicciones electorales está en el reparto de escaños mediante la Ley D´Hondt (fórmula adoptada por la legislación española). Veamos en qué consiste y como se aplica la ley D´Hondt en la comunidad gallega:

*En cada circunscripción se excluyen las candidaturas que no hayan obtenido, al menos, el 5% de los votos válidos emitidos. Con el resto de las candidaturas, se ordenan de mayor a menor, en una columna, las cifras de votos obtenidos. Se divide el número de votos obtenidos por cada candidatura por 1, 2, 3, etc., hasta un número igual al de escaños correspondientes a la circunscripción. Los escaños se atribuyen a las candidaturas que obtengan los cocientes mayores, atendiendo a un orden decreciente.*

Estadísticamente, la ley D´Hondt se considera una función que depende de las proporciones de los votos. Es difícil estimar los escaños en una predicción debido a que la forma de reparto es sesgada.

Además, este procedimiento distorsiona la voluntad popular, distribuyendo los escaños de una forma que no respeta la representación proporcional, favoreciendo a los partidos mayoritarios. Un estudio reciente realizado por el profesor Bernardo, propone una fórmula alternativa a la ley D´Hondt que resulta más justa a la hora de repartir los escaños definida con el nombre de *solución correcta* (*Proportionality in parliamentary democracy: An alternative to Jefferson-d´Hondt rule. J. M. Bernardo(2004)*).

Las diferencias entre la solución correcta y la ley D´Hondt tienden a disminuir cuando el número de escaños aumenta. De forma análoga, las diferencias entre ellas aumentan cuando el número de escaños a repartir disminuye. La fórmula correcta de José Miguel Bernardo, se calcula con los siguientes pasos:

*Paso 1:* Partimos de los votos conseguidos por cada uno de los partidos con derecho a representación parlamentaria.

*Paso 2:* Repartimos los escaños correspondientes de modo proporcional a los votos obtenidos por cada uno de los partidos (solución ideal).

*Paso 3:* La repartición realizada en el paso 2 la redondeamos a los números

enteros más cercanos, tanto por defecto como por exceso.

*Paso 4:* Calculamos los errores a las aproximaciones enteras, esto es, la diferencia en valor absoluto de cada aproximación entera del paso 3 con su solución ideal del paso 2.

*Paso 5:* Finalmente, la solución correcta la obtenemos partiendo de los errores absolutos, del más pequeño al mayor, para asignar a cada partido la solución con mínimo error que sea compatible con el número total de escaños que deben ser distribuidos.

Para que la democracia actúe como tal, quizás se debería de contemplar la posibilidad de modificar la constitución para poder cambiar la Ley D´Hondt a otra más justa, como por ejemplo la fórmula correcta de Bernardo.

En el próximo capítulo, aplicaremos el coeficiente de correlación lineal de Pearson (CCP), la divergencia de Kullback-Leibler (KL) y el coeficiente de correlación de Szekely para seleccionar las mesas y municipios más representativos de Galicia. Los valores hallados están en concordancia: divergencias de Kullback-Leibler próximas a 0 dan correlaciones cerca de 1 (en valores absolutos).



# Capítulo 3

## Estudios Demoscópicos

La Real Academia Española define la demoscopia como el estudio de las opiniones, aficiones y comportamiento humanos mediante sondeos de opinión.

Si pretendemos saber que opinan los ciudadanos gallegos (principalmente sobre las elecciones) debemos centrar los estudios demoscópicos en los municipios testigo que calculamos en este capítulo, porque su opinión podemos extrapolarla a la de toda Galicia.

En el estudio de las mesas y municipios testigo, debemos ser conscientes de que esta capacidad es duradera a lo largo del tiempo. Aunque puede ser alterable, dependiendo del ciclo económico y social en el que nos encontremos. Todo ello, hace que unas mesas sean más pronosticadoras que otras.

En todo partido político existen varias teorías estadísticas que ayudan a buscar estrategias para alcanzar el éxito en las elecciones. Entre ellas encontramos la del *teorema del votante mediano*. Este teorema se fundamenta en el resultado de la decisión electoral que suele coincidir con lo que prefiere el ciudadano que en el momento de las elecciones es el dominante.

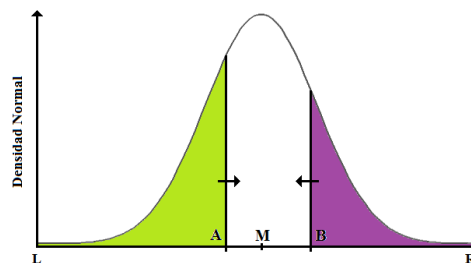


Figura 3.1: Distribución del votante mediano.

La Figura 3.1 es un modelo posible de la distribución de los votantes medianos. Si los partidos A y B pretenden ganarse al votante mediano (M), deben moverse hacia

el centro. Las áreas verde y violeta representan a los respectivos votantes que A y B esperan haber convencido.

De este modo, conociendo el carácter *middletown* de un municipio, si los partidos políticos consiguen obtener el voto del elector medio de dicho municipio, es muy probable que lo hayan conseguido también en el resto de la provincia, comunidad o nación (dependiendo de que elecciones se traten). Aumentando considerablemente las posibilidades de éxito.

Uno de los analistas de este teorema, Gordon Tulloch, en su libro “*Los motivos del voto*” sostiene que *en una democracia no es posible asegurar una representación ecuánime de las preferencias de los ciudadanos; necesariamente, algunos terminarán mejor representados que otros, incluyendo a muchos que pueden haber votado al candidato ganador. Es decir, una vez en el gobierno, el ganador no hará todo lo que todos sus votantes esperan, sino lo que le resulta posible o conveniente para poder gobernar en vistas de los equilibrios entre las distintas coaliciones... El argumento de Tullock es que si los candidatos buscan maximizar votos y no sólo insistir en sus propias propuestas, necesariamente deben moverse hacia el grueso de los votantes, los votantes “medios”, muchos indecisos e indefinidos hasta último momento. Son los votantes cuyas expectativas se ven finalmente mejor representadas (El teorema del votante medio, Revista Perfil).*

### Software propuesto para el estudio

El software libre R será nuestra herramienta de trabajo. R es un lenguaje de programación especialmente indicado para el análisis estadístico y gráficos creado inicialmente por Ross Ihaka y Robert Gentleman. Hoy en día, colaboran en su desarrollo miles de personas en todo el mundo. Al ser libre, los usuarios tienen total libertad para ejecutar, estudiar, distribuir, copiar, cambiar y mejorar el software. En la web <http://www.r-project.org/> no solo podemos acceder a los archivos necesarios para su instalación, sino que también tenemos acceso al resto de recursos del proyecto R. Está disponible para los sistemas operativos Windows, Macintosh, Unix y GNU/Linux. Este software es un proyecto GNU similar al lenguaje y entorno comercial S desarrollado por los laboratorios Bell. Algunas de las características que nos podemos encontrar en R son: almacenamiento y manipulación efectiva de datos; amplia colección de herramientas para el análisis de datos; descarga de librerías con implementaciones concretas de métodos estadísticos, funciones gráficas...

Para nuestro estudio hemos utilizado las librerías siguientes: `energy`, `pyramid`, `gdata`, `shiny`, `devtools`, `leaflet`, `RColorBrewer`, `scales`, `lattice` y `dplyr`.

## 3.1. Estudio demoscópico de Galicia

### 3.1.1. Contextualización

Para realizar el estudio sociodemográfico es importante conocer qué ocurría, tanto en el territorio nacional como internacional, sobre los aspectos socio-políticos y económicos en los años de las elecciones. Hagamos memoria:

- **Elecciones al parlamento de Galicia de 2005 (19/06/05)**

En el ámbito gallego, los candidatos a la presidencia de la Xunta de los partidos con representación en el parlamento son Manuel Fraga por el PP, Emilio Pérez Touriño por el PSOE y Anxo Quintana por el BNG. Estas elecciones se realizan en junio, coincidiendo con los exámenes universitarios, que mantienen a gran parte de los votantes más jóvenes ocupados. Además, son las primeras elecciones regionales después de la catástrofe del Prestige.

En España, el terrorismo ocupa los pensamientos de los ciudadanos (en el año anterior ocurren los atentados del 11-M). José Luís Zapatero (PSOE) es proclamado presidente del gobierno tras su victoria en las elecciones generales de 2004.

- **Elecciones al parlamento de Galicia de 2009 (01/03/09)**

En Galicia, cabe destacar el cambio del candidato del partido Popular a presidente de la Xunta: Manuel Fraga Iribarne es reemplazado por Alberto Núñez Feijóo.

En España, Zapatero revalidó su victoria como presidente del gobierno y mantiene su lucha contra el terrorismo como objetivo primordial.

Internacionalmente, Barack Obama es elegido presidente de Estados Unidos.

- **Elecciones al parlamento de Galicia de 2012 (21/10/12)**

En el ámbito gallego, las elecciones se adelantan seis meses. Los candidatos a presidente de la Xunta son Nuñez Feijóo por el PP, Pachi Vazquez por el PSOE, Francisco Jorquera por el BNG y Xosé Manuel Beiras por el AGE (Alternativa Galega de Esquerda).

En el año anterior, Mariano Rajoy (candidato por el PP) es proclamado presidente de España, en plena crisis económica.

### 3.1.2. Votos CERA

Debemos tener en cuenta que Galicia tiene una gran cantidad de emigrantes con derecho a voto, son los llamados votos CERA (ciudadanos residentes de forma permanente en el extranjero). Concretamente en el año 2012 son 397.382 votantes. Estos votos pueden repercutir en las estimaciones de buscar mesas testigo distorsionando la opinión de los ciudadanos que residen en la comunidad. Por todo ello, realizamos dos estudios:

**Estudio 1:** Teniendo en cuenta los votos CERA. Interesante para los partidos políticos, ya que les ayuda a conocer la opinión de todas las personas con derecho a voto. Los votos CERA están recogidos como una mesa electoral más de cada provincia.

**Estudio 2:** Omitiendo los votos CERA. De este modo observamos lo que realmente piensa la gente que vive en Galicia. Es el modelo adecuado para los estudios sociodemográficos, pero se manipulan los resultados finales de la comunidad.

Debido a la similitud de los resultados, analizamos con más detalle el **estudio 2**, ya que es el que mejor refleja lo que piensan los habitantes de Galicia. En el apéndice A mostraremos los resultados de ambos análisis al completo.

Elaboramos varios scripts, con el fin de ser más claros en la búsqueda de los resultados. En todos ellos cargamos siempre los datos que hemos preparado en la subsección 2.1.1:

```
datosS<-read.csv2(file="MesasGaliciaSinCERA.csv", sep=";", dec=".", header=T)
```

Muchas de las funciones empleadas se repiten en los scripts adaptadas a los datos que pretendemos analizar. Por lo que una vez explicadas, obviaremos su descripción en usos posteriores.

### 3.1.3. Selección de las mesas testigo

Para la selección de las mesas testigo elaboramos un script de R que calcula qué mesas tienen mayor similitud con los resultados de Galicia (`RankingMesasGalicia.R`). Contiene las siguientes funciones:

- `MesasGalicia`

Calcula el coeficiente de correlación lineal de Pearson (CCP), la divergencia de Kullback-Leibler (KL) y el test de correlación de Szekely (en la librería `energy`) para todas las mesas electorales de Galicia en las tres últimas elecciones. Y guarda las 1.300 mesas con mayor CCP ordenadas de mayor a menor en `MesasGalicia05`, `MesasGalicia09` y `MesasGalicia12`. Utilizamos para su selección el CCP, pero sería válido cualquiera de

los otros dos resultados ya que están en plena armonía, a mayor dependencia menor es la divergencia entre las variables. El código de R es el siguiente:

```
MesasGalicia<-function (mesas, totales, partidos, nummesas) {
  pearson<-c()
  KL<-c()
  independencia<-c()
  for (i in 1:nrow(mesas)) {
    x <- as.numeric(as.matrix(mesas[i,12:(11+partidos)]))
    y <-as.numeric(as.matrix(totales[1,12:(11+partidos)]))
    pearson<-c(pearson, abs(cor(x,y, method="pearson")))
    KL<-c(KL,sum(y*log10(y/x)))
    independencia<-c(independencia,dcov.test(y,x,R=199)$p.value)
  }
  mesas<-cbind(mesas, independencia)
  mesas<-cbind(mesas, KL)
  mesas<-cbind(mesas, pearson)
  seleccion <- sort(mesas[,ncol(mesas)], decreasing = TRUE)[1:nummesas]
  posicion<-which(mesas[,ncol(mesas)] %in% sort(mesas[,ncol(mesas)], decreasing = TRUE)
  [1:nummesas])
  mesas2<- mesas[posicion,]
  ordenmesas <-order(mesas2[,ncol(mesas2)], decreasing = TRUE)
  resultados<-mesas2[ordenmesas,]

  return(resultados)
}

# Resultados Galicia 2005:
mesas2005<-datosS[AÑO==2005 & (PROVINCIA=="A Coruña" | PROVINCIA== "Lugo" |
PROVINCIA== "Ourense" | PROVINCIA== "Pontevedra"),]
totales2005<-datosS[AÑO==2005 & PROVINCIA=="Galicia",]
MesasGalicia05 <- MesasGalicia(mesas = mesas2005, totales = totales2005, partidos=4,
nummesas=1300)

# Resultados Galicia 2009:
mesas2009<-datosS[AÑO==2009 & (PROVINCIA=="A Coruña" | PROVINCIA== "Lugo" |
PROVINCIA== "Ourense" | PROVINCIA== "Pontevedra"),]
totales2009<-datosS[AÑO==2009 & PROVINCIA=="Galicia",]
MesasGalicia09 <- MesasGalicia(mesas2009, totales2009, partidos=4,
nummesas=1300)

# Resultados Galicia 2012:
mesas2012<-datosS[AÑO==2012 & (PROVINCIA=="A Coruña" | PROVINCIA== "Lugo" |
PROVINCIA== "Ourense" | PROVINCIA== "Pontevedra"),]
totales2012<-datosS[AÑO==2012 & PROVINCIA=="Galicia",]
MesasGalicia12 <- MesasGalicia(mesas2012, totales2012, partidos=5,
nummesas=1300)
```

- `identificar_mesascomunes`

Esta función toma las mesas seleccionadas anteriormente y su CCP y comprueba si alguna se repite en las tres elecciones. Si una mesa se repite, significa que sus resultados se acercan mucho a los de Galicia en todas las elecciones y, por tanto, es candidata a ser *mesa testigo*. De las 1.300 mesas que disponemos en cada elección, 285 se encuentran entre las mejores.

```
identificar_mesascomunes <- function (mesas1, mesas2, nummesas) {
  coinciden <- matrix(0, nr=nummesas)
  for (i in 1:nummesas) {
    coinciden[which(apply(mesas1, 1, identical, mesas2[i,])), ] <- 1
  }

  return(coinciden)
}

a <- as.matrix(MesasGalicia05[ ,c(2:8, 19)])
b <- as.matrix(MesasGalicia09[ ,c(2:8, 19)])
c <- as.matrix(MesasGalicia12[ ,c(2:8, 19)])

# Mesas que coinciden en las elecciones de 2005 y de 2009:
mesas0509 <- identificar_mesascomunes(a,b, nummesas=1300)
sum(mesas0509)

# Mesas que coinciden en las elecciones de 2005 y de 2012:
mesas0512 <- identificar_mesascomunes(a, c, nummesas=1300)
sum(mesas0512)

# Mesas que coinciden en las tres elecciones:
mesastotales<-mesas0509 * mesas0512
sum(mesastotales)
```

- `imprime_mesas`

Tal y como su nombre indica, esta función imprime el nombre de las mesas electorales obtenidas en `identificar_mesascomunes`, mostrando además su CCP del año 2005 (en la siguiente función añadimos los coeficientes de los años 2009 y 2012).

```
imprime_mesas<- function (a, mesastotales, nummesas) {
  mesas <- matrix(ncol=8)
  for (i in 1:nummesas) {
    if (mesastotales[i,1]==1) {
      mesas <- rbind(mesas, a[i,])
    }
  }
  return(mesas[-1,])
}

nombremesas <- imprime_mesas(a, mesastotales, nummesas=1300)
colnames(nombremesas)[8] <- "pearson05"
```

- addPearson

Como adelantamos en la función `imprime_mesas`, añadimos los coeficientes de correlación de Pearson de los años 2009 y 2012. Y una nueva columna con la suma de estos tres coeficientes que ordenamos de forma decreciente. Cuanto mayor sea este sumatorio, que oscila entre 0 y 3, mayor es la dependencia de la mesa con respecto a Galicia (en el conjunto de los resultados de las tres elecciones).

```
addPearson <- function(mesas1, mesas2) {
  pearsonY <- c()
  for (i in 1:nrow(mesas1)) {
    pearsonY <- c(pearsonY, b[which(apply(mesas2[, 1:7], 1, identical,
      mesas1[i, 1:7])), 8])
  }

  # Añadir columna pearson
  mesas1 <- cbind(mesas1, pearsonY)

  return(mesas1)
}

# Añadir 2009
mesaspearson <- addPearson(nombremesas, b)
colnames(mesaspearson)[9] <- "pearson09"

# Añadir 2012
mesaspearson<- addPearson(mesaspearson, c)
colnames(mesaspearson)[10] <- "pearson12"

# Añadir columna de sumas
sumapearson <- c()
for(i in 1:nrow(mesaspearson)) {
  sumapearson <- c(sumapearson, sum(as.numeric(mesaspearson[i, 8:10])))
}

mesaspearson <- cbind(mesaspearson, sumapearson)

# Ordenar de mayor a menor según la columna de sumapearson
ordenmesaspearson <-order(mesaspearson[,ncol(mesaspearson)], decreasing = TRUE)
resultadosmesaspearson<-mesaspearson[ordenmesaspearson,]

resultadosmesaspearson
```

**Resultado de mesas testigo para Galicia**

En el cuadro 3.1 mostramos las 15 mejores mesas testigo. La primera columna refleja la provincia de donde proviene la mesa; en la segunda columna, tenemos el nombre de la mesa electoral formado por: código municipio, municipio, distrito, sección y mesa; y la tercera columna, es el sumatorio de los CCP de las tres elecciones.

	<b>Provincia</b>	<b>Mesa electoral</b>	$\sum_{i=1}^3 \text{CCP}$
1	A Coruña	42 Lousame 1 1 B	2,9978719
2	Lugo	13 Cervo 1 2 A	2,9972689
3	A Coruña	78 Santiago de Compostela 2 6 C	2,9938972
4	Lugo	65 Vilalba 1 7 U	2,9925219
5	Pontevedra	27 Meaño 2 2 U	2,9921801
6	Ourense	85 Verín 1 1 B	2,9920981
7	A Coruña	57 Noia 2 4 A	2,991708
8	A Coruña	57 Noia 2 6 B	2,9916811
9	A Coruña	43 Malpica de Bergantiños 1 2 B	2,9914323
10	Lugo	51 Ribadeo 1 2 A	2,9912435
11	Pontevedra	51 Sanxenxo 1 6 U	2,9912402
12	A Coruña	2 Ames 3 1 B	2,9911627
13	A Coruña	57 Noia 1 1 B	2,9909612
14	Pontevedra	3 Baiona 2 1 A	2,9908296
15	Lugo	22 Guitiriz 4 2 B	2,9905138

Cuadro 3.1: Mejores mesas testigo para Galicia.

Como podemos observar en el Cuadro 3.1 todas las mesas superan el umbral de 2.99 en el sumatorio de los CCP de las tres elecciones. Esto implica que en todas las elecciones estas mesas tienen una alta dependencia con Galicia. Dentro de las 15 mesas, tres de ellas pertenecen al mismo municipio, Noia. ¿Será Noia la ciudad testigo de la comunidad gallega? ¿Y de su provincia? Lo averiguaremos en las siguientes subsecciones.



### 3.1.4. Selección del municipio testigo

Con el script Galicia.R evaluamos qué municipio es el que mejor refleja los resultados electorales de Galicia. Aplicamos las siguientes funciones:

- MesasGalicia
- identificar\_mesascomunes
- imprime\_mesas

Las funciones `identificar_mesascomunes` e `imprime_mesas` sufren una pequeña modificación, omitimos las columnas reservadas al CCP ya que ahora nos interesa saber cuantas mesas hay pertenecientes a cada municipio.

- `porcentajes_municipios`

Con esta función calculamos el número y el porcentaje de mesas con mayor CCP que tiene cada municipio. De este modo, los municipios con mayor porcentaje son los candidatos a ser la ciudad testigo de Galicia.

```
porcentajes_municipios <- function(nombremesas, datos) {

  municandidatos <- unique(nombremesas[, 4])

  porcentajesMuni <- matrix(ncol = 4)[-1, ]

  for (i in 1:length(municandidatos)) {
    # Nombre del municipio
    nombremunicipio <- municandidatos[i]

    # Vamos a calcular el número total de mesas de cada municipio:
    municipios <- datos[AÑO == 2005 & MUNICIPIO == nombremunicipio, ]
    x <- dim(municipios)[1] # número de mesas del municipio

    # Vamos a calcular el número de mesas de cada municipio:
    seleccion <- which(nombremesas[, 4] == nombremunicipio)
    y <- length(seleccion)
    porcentaje <- round(y*100/x, 2)
    muniactual <- c(nombremunicipio, x, y, porcentaje)
    porcentajesMuni <- rbind(porcentajesMuni, muniactual)
  }

  rownames(porcentajesMuni) <- NULL
  porcentajesMuni <- porcentajesMuni[order(as.numeric(porcentajesMuni[, 4]),
    decreasing=TRUE), ]
  colnames(porcentajesMuni) <- c("Municipio", "Nº mesas", "Nº mesas coinciden",
    "Porcentaje")
  porcentajesMuni <- as.data.frame(porcentajesMuni)
```

```

    return(porcentajesMuni)
}
porcentajes_municipios(nombremesas, datosS)

```

### Resultado de municipios testigo para Galicia

Finalmente, obtenemos los posibles candidatos a municipio testigo. En el cuadro 3.2 mostramos los 15 municipios con mayor porcentaje de mesas con coeficiente de correlación de Pearson alto y el censo electoral de las últimas elecciones.

	Municipio	Nº mesas	Porcentaje	Censo
1	Rábade	2 de 2	100 %	1.329
2	<b>Baiona</b>	8 de 12	66,67 %	9.888
3	<b>Nigrán</b>	9 de 18	50 %	14.466
4	Páramo (O)	2 de 4	50 %	1.397
5	Cabanas	2 de 4	50 %	2.841
6	Paderne de Allariz	1 de 2	50 %	1.372
7	<b>Marín</b>	14 de 29	48,28 %	20.455
8	<b>Noia</b>	9 de 19	47,37 %	12.405
9	Curtis	3 de 7	42,86 %	3.593
10	Ribadumia	3 de 7	42,86 %	4.122
11	Meis	3 de 7	42,86 %	4.186
12	<b>Guarda (A)</b>	5 de 12	41,67 %	8.551
13	Lourenzá	2 de 5	40 %	2.075
14	<b>Ribadeo</b>	5 de 13	38,46 %	8.179
15	<b>Betanzos</b>	7 de 19	36,84 %	11.057

Cuadro 3.2: Posibles municipios testigo de Galicia.

El municipio testigo no solo debe acercarse a los resultados electorales de Galicia, sino que sus características generales deben ser lo más similares posibles a las características de la comunidad gallega.

Además, debemos buscar alguna manera de filtrar los datos, puesto que, en nuestra forma de selección, los municipios pequeños tienen más probabilidades de ser candidatos a municipios testigo que los más poblados.

Para ello, calculamos el tamaño de la muestra. Esto es, fijamos el tamaño mínimo del censo electoral que debe tener el municipio elegido. Lo calculamos a un nivel de confianza  $3\sigma$  (es decir, el 97,7% de probabilidad de éxito en la estimación) y con un

error muestral del 2,5%. De este modo, haciendo los cálculos con los datos de las últimas elecciones, para Galicia su municipio testigo debe superar el **7.122** individuos con derecho a voto (explicación y cálculos más detallados en el apartado 4.1.1).

En el cuadro 3.2 los municipios que están en negrita (Baiona, Nigrán, Marín, Noia, A Guarda, Ribadeo y Betanzos) son los que tienen un censo superior a 7.122. Todos ellos poseen un gran porcentaje de mesas con CCP en las tres elecciones, aunque Baiona destaca entre ellas con 8 de sus 12 mesas.

Por todo ello, tomamos a Baiona como candidato a ser municipio testigo de Galicia.

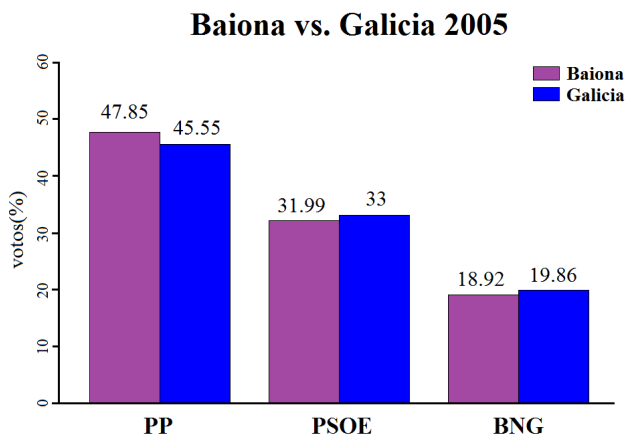
### 3.1.5. Baiona, ¿ciudad testigo de Galicia?

Comprobemos la similitud entre los resultados obtenidos por los distintos partidos en Baiona y los generales de Galicia. Recordemos que en este estudio hemos omitido los votos CERA.

#### • Elecciones al parlamento de Galicia de 2005

En el 2005, Baiona tenía 9.398 electores, de los cuales votaron 6.396. Esto presenta el 68,06% de participación. Los resultados electorales dan 2.999 votos al PP (47,85%), 2.005 votos al PSOE (31,99%) y 1.186 votos al BNG (18,92%). Los votos restantes fueron para los partidos que no alcanzaron el umbral de voto exigido para entrar en el reparto de parlamentarios, para votos en blanco o nulos.

Trasfiriendo estos datos para toda Galicia, los electores eran 2.311.794, de los que votaron 1.574.350, lo que representa el 68,10% de participación. Los resultados generales dan 704.198 votos al PP (45,55%), 510.171 votos al PSOE (33%) y 307.021 votos al BNG (19,86%).



Comparativa de los resultados electorales 2005

Asombrosamente el nivel de participación de Galicia y Baiona es prácticamente el mismo. Los porcentajes a las mayores fuerzas políticas del momento también son muy similares como podemos observar en la gráfica de la izquierda.

Calculando el coeficiente de correlación de Pearson (CCP), todas las mesas tienen una dependencia muy alta con Galicia, 11 de las 12 mesas superan el umbral del 0,9 y las mesas que están coloreadas tienen una tasa superior a 0,99 (cuadro 3.3). Destaca la mesa que está compuesta por la población de Santa María de Afuera, **12B**, que se encuentra rozando la perfección. Por lo contrario, la mesa de la zona de Baredo (Santa María), **13U** es la que posee el valor más bajo, 0,81955.

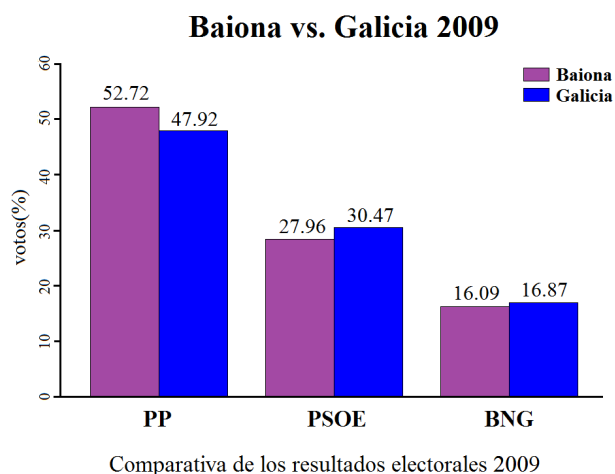
MESA	CENSO	T.VOTOS	PP	PSOE	BNG	OTROS	KL	CCP
1 1 U	956	675	0,50	0,32	0,18	0,01	0,0023907	0,9947439
1 2 A	794	517	0,41	0,36	0,22	0,00	0,0067315	0,9819579
<b>1 2 B</b>	771	527	0,46	0,33	0,20	0,02	0,0000245	0,9999661
1 3 U	888	568	0,54	0,15	0,30	0,01	0,0476427	0,8195586
1 4 A	671	458	0,49	0,31	0,18	0,02	0,0018164	0,9914273
1 4 B	702	489	0,48	0,31	0,19	0,02	0,0012262	0,9947882
2 1 A	683	456	0,43	0,34	0,22	0,01	0,0027481	0,9948935
2 1 B	773	487	0,48	0,36	0,15	0,01	0,0052592	0,9883899
2 2 U	956	667	0,58	0,29	0,11	0,01	0,0189050	0,9512711
2 3 A	567	397	0,42	0,39	0,17	0,01	0,0034007	0,9772525
<b>2 3 B</b>	600	437	0,50	0,32	0,17	0,01	0,0023670	0,9923722
2 4 U	1.037	718	0,41	0,39	0,20	0,01	0,0046948	0,9751051
Galicia	2.311.794	1.574.350	0,46	0,33	0,20	0,02		

Cuadro 3.3: Resultados de 2005 con distancias de Kullback-Leibler (KL) y coeficientes de correlación de Pearson (CCP).

#### • Elecciones al parlamento de Galicia de 2009

En Baiona, aumentó el número de electores con respecto a las elecciones anteriores, 9.766, de los cuales fueron a las urnas 6.930. Esto presenta el 70,96% de participación, aumenta casi tres puntos porcentuales en comparación con la del año 2005. Los resultados electorales dan 3.572 votos al PP (52,72%), 1.892 votos al PSOE (27,93%) y 1.090 votos al BNG (16,09%).

En Galicia, había un total de 2.311.584, de los que votaron 1.630.174, lo que supone el 70,52% de participación. Los resultados generales dan 256.518 votos al PP (47,92%), 483.776 votos al PSOE (30,47%) y 267.814 votos al BNG (16,86%). Los votos restantes fueron para votos en blanco, votos nulos o para los partidos que no alcanzaron votos suficientes para entrar en el reparto de parlamentarios.



Las variaciones, tanto en la tasa de participación como en los votos en Galicia, son paralelas a los registrados en Baiona. Desciende la participación de los electores; el PP aumenta su popularidad; el BNG y el PSOE la disminuyen, sobretodo este último.

El coeficiente de correlación de Pearson (CCP), no es inferior en ninguna mesa de 0,9 y las mesas que están coloreadas superan el 0,99 (cuadro 3.4). Aumenta el número de mesas, ya que la mesa 24U se desglosa en dos (24A y 24B).

MESA	CENSO	T.VOTOS	PP	PSOE	BNG	OTROS	KL	CCP
1 1 U	944	664	0,48	0,30	0,18	0,04	0,0001769	0,9995909
1 2 A	797	531	0,48	0,30	0,19	0,03	0,0035529	0,9948048
1 2 B	795	574	0,53	0,28	0,16	0,03	0,0037389	0,9938206
1 3 U	928	647	0,64	0,20	0,15	0,01	0,0336132	0,9499359
1 4 A	676	486	0,53	0,30	0,14	0,04	0,0031953	0,9951455
1 4 B	753	532	0,54	0,25	0,16	0,05	0,0038196	0,9806497
2 1 A	711	485	0,48	0,32	0,16	0,04	0,0008457	0,9981610
2 1 B	763	519	0,52	0,32	0,13	0,02	0,0070586	0,9974994
2 2 U	949	722	0,61	0,22	0,15	0,02	0,0188487	0,9597335
2 3 A	584	413	0,49	0,32	0,15	0,04	0,0011108	0,9984198
2 3 B	648	476	0,54	0,29	0,13	0,04	0,0045968	0,9916171
2 4 A	563	405	0,49	0,30	0,18	0,04	0,0003561	0,9991518
2 4 B	655	476	0,45	0,27	0,24	0,04	0,0067407	0,9678866
Galicia	2.311.584	1.630.174	47,92	30,47	16,871	4,74		

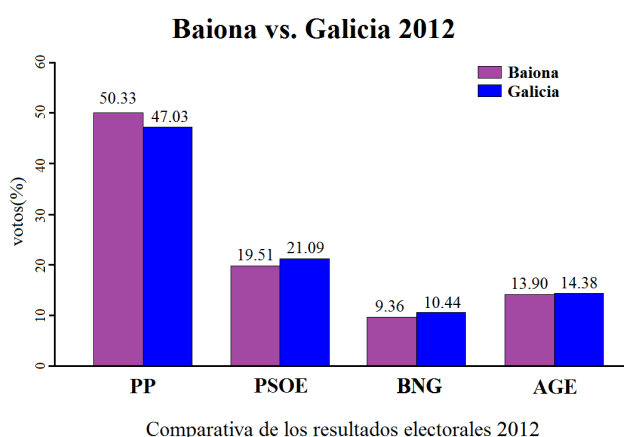
Cuadro 3.4: Resultados de 2009 con distancias de Kullback-Leibler (KL) y coeficientes de correlación de Pearson (CCP).

#### • Elecciones al parlamento de Galicia de 2012

El censo electoral tuvo un ligero aumento con respecto a elecciones anteriores en Baiona que quedó en 9.888, de los cuales votaron 6.201, lo que supone una baja participación (62,71 %). Los resultados electorales dan 2.951 votos al PP (50,33 %),

1.144 votos al PSOE (19,51 %) y 549 votos al BNG (9,36 %). Aparece un nuevo partido que alcanza el umbral para optar a representación parlamentaria, Alternativa Galega de Esquerda (AGE), que obtiene 815 votos (13,90 %).

En Galicia, el censo electoral en el 2012 es de 2.300.335, de los que ejercieron su derecho a voto 1.468.425, lo que supone el 63,85 % de participación. La abstención es muy elevada, situándose en un 36,15 %. Los resultados generales dan 654.907 votos al PP (47,03 %), 293.714 votos al PSOE (21,09 %), 145.424 votos al BNG (10,44 %) y 200.211 votos al AGE (14,38 %).



Al igual que en las elecciones anteriores, las variaciones tanto en la tasa de participación como en los votos en Galicia son paralelos a los registrados en Baiona. Los votos a los partidos en el municipio vuelven a ser proporcionales a los de la comunidad.

Una vez calculado el coeficiente de correlación de Pearson (CCP), tenemos de nuevo una alta correlación entre todas las mesas y los resultados finales, ningún coeficiente es inferior a 0,9. Las mesas que están coloreadas superan el 0,99 (cuadro 3.5).

MESA	CENSO	T.VOTOS	PP	PSOE	BNG	OTROS	AGE	KL	CCP
1 1 U	911	565	0,47	0,20	0,08	0,05	0,19	0,0053526	0,9841515
1 2 A	809	468	0,44	0,25	0,11	0,08	0,12	0,0023655	0,9891334
1 2 B	816	512	0,49	0,18	0,11	0,07	0,15	0,0017537	0,9929199
1 3 U	908	567	0,60	0,15	0,07	0,07	0,10	0,0175328	0,9813598
<b>1 4 A</b>	686	428	0,51	0,19	0,10	0,05	0,15	0,0026041	0,9975609
1 4 B	750	456	0,46	0,20	0,08	0,06	0,19	0,0039968	0,9867976
2 1 A	721	420	0,47	0,23	0,08	0,07	0,15	0,0023257	0,9944325
2 1 B	765	468	0,48	0,26	0,08	0,06	0,12	0,0043317	0,9888279
2 2 U	974	669	0,59	0,18	0,08	0,06	0,09	0,0142623	0,9908710
2 3 A	640	406	0,47	0,18	0,09	0,11	0,14	0,0056422	0,9852939
2 3 B	651	402	0,56	0,16	0,06	0,08	0,14	0,0128766	0,9812279
2 4 A	582	395	0,49	0,19	0,12	0,07	0,13	0,0012142	0,9958778
2 4 B	675	445	0,46	0,18	0,18	0,06	0,12	0,0097215	0,9664887
Galicia	2.300.335	1.468.425	47,03	21,09	10,44	14,38			

Cuadro 3.5: Resultados de 2012 con distancias de Kullback-Leibler (KL) y coeficientes de correlación de Pearson (CCP).

Aparte de los resultados estadísticos tan brillantes, debemos de tener en cuenta las características sociales propias de municipio de Baiona.

Este municipio está situado al suroeste de la provincia de Pontevedra, regogado en una pequeña bahía de la ría de Vigo. Tiene una extensión de  $34,26 \text{ Km}^2$ , con 11.933 habitantes (censo de población y viviendas, 2011 del INE), con respecto a Galicia se puede considerar una ciudad media en cuanto al tamaño de su población.

Tiene una gran importancia histórica, ya que en marzo de 1493, Martín Alonso Pinzón arribó a las costas de Baiona tras su viaje a América convirtiendo a esta villa en la primera de Europa que supo la noticia del descubrimiento del Nuevo Mundo. Por ello el primer fin de semana de marzo se celebra en el casco histórico de Baiona una fiesta medieval llamada “La Arribada”.

El turismo es su principal fuente de ingresos. Debido a ello, su comercio es de los más importantes de la zona que, sobre todo, ofrece productos del mar. No debemos olvidar el entorno natural incomparable, donde destacan sus costas y el parque natural de las Islas Cíes.

Baiona posee un clima suave, que facilita el cultivo de viñedos y huertos. Su ría es el lugar idóneo para la práctica de los deportes náuticos, al abrigo de los vientos y del mar. Su puerto deportivo está considerado uno de los mejores de Galicia.

Centrándonos en la población, en la Figura 3.2 vemos representado la estructura de la población por sexo y edad de Baiona y de Galicia.

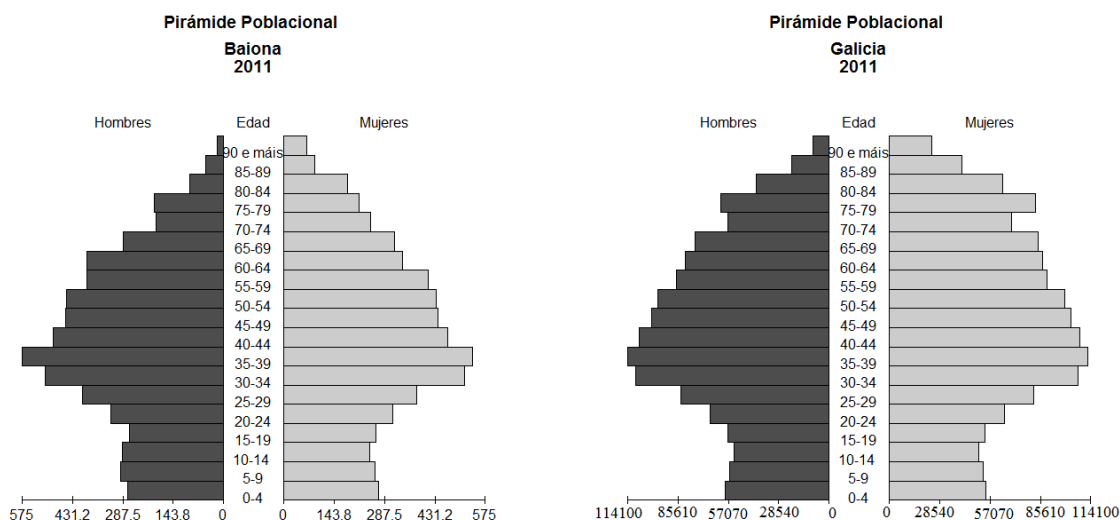


Figura 3.2: Pirámides poblacionales de Baiona y Galicia (2011).

Observamos que desde los 0 a los 40 años las pirámides poblacionales son muy

similares. Pero a partir de esa edad varía, reflejando que Baiona es un municipio más joven que la media de Galicia, hay menos personas mayores, sobre todo lo observamos en las mujeres.

Otros datos de interés son: el ingreso medio mensual por persona y la tasa de paro (INE, 2012 y 2014 respectivamente). En Galicia el ingreso medio de un habitante es de 734€/mes, muy similar al de Baiona, que es de 737€/mes. En cambio, la tasa de paro en Galicia es del 23,2% mientras que en Baiona es más elevado (38,2%).

Todo apunta a que Baiona es una de las ciudades estadísticamente perfectas de Galicia, aún así completemos el análisis para comprobar si sus resultados electorales al **parlamento europeo de 2014** son extrapolables a los de toda la comunidad. En los siguientes cuadros se recogen los datos de dichas elecciones:

	<b>Galicia</b>		<b>Baiona</b>	
T. Votantes	1.035.487	45,45 %	4.626	45,82 %
Abstención	1.242.655	54,55 %	5.469	54,18 %
Votos nulos	26.714	2,58 %	111	2,51 %
Votos en blanco	29.927	2,97 %	153	3,34 %

Cuadro 3.6: Votos al parlamento europeo 2014 de Baiona y Galicia.

Podemos observar en el cuadro 3.6, que los porcentajes de los votantes de Baiona son muy parecidos a los porcentajes de Galicia. En el cuadro 3.7 tenemos los porcentajes y proporciones a los partidos políticos, los valores del cálculo de la divergencia de Kullback-Leibler (0,0040946) y coeficiente de correlación de Pearson (0,9887496). Los valores hallados muestran que hay muy poca distancia entre las votaciones, de modo que su correlación es alta.

<b>Partidos</b>	<b>Galicia</b>		<b>Baiona</b>	
PP	35,16 %	0,3516	38,29 %	0,3829
PSOE	21,73 %	0,2173	18,64 %	0,1864
AGE	10,52 %	0,1052	10,68 %	0,1068
PODEMOS	8,34 %	0,0834	8,53 %	0,0853
BNG	7,90 %	0,0790	6,54 %	0,0654
UPyD	3,47 %	0,0347	3,70 %	0,0370
OTROS	9,88 %	0,0988	10,20 %	0,1020
			KL	<b>0,0040946</b>
			CCP	<b>0,9887496</b>

Cuadro 3.7: Resultados al parlamento europeo 2014 de Baiona y Galicia con divergencia de Kullback-Leibler (KL) y coeficiente de correlación de Pearson (CCP).



Como conclusión general, debido a los resultados de los cálculos (tanto para las elecciones al parlamento de Galicia como para las elecciones al parlamento Europeo) y a las características propias del municipio, podemos decir que Baiona es una de las ciudades estadísticamente perfectas de Galicia. (Todos los cálculos y gráficas están recogidos en el script `Baiona.R`)

Es posible que exista más de una ciudad que refleje lo que piensa la comunidad. Salvador Naya, en el ensayo “*Betanzos, magic town*”, muestra la posibilidad de que Betanzos sea también una las ciudades testigo de Galicia.

El análisis para buscar el municipio testigo de Galicia, lo reiteramos para cada provincia, con el fin de encontrar los municipios que reflejen con exactitud los resultados de la provincia a la que pertenecen. Todos los scripts realizados se encuentran en el CD adjunto con esta memoria.

## 3.2. Municipios parecidos a A Coruña

En el script `Coruna.R`, evaluamos qué municipio es el que mejor refleja los resultados electorales de A Coruña. Aplicamos las siguientes funciones vistas anteriormente:

- `MesasCoruna`
- `identificar_mesascomunes`
- `imprime_mesas`
- `porcentajes_municipios`

La función `MesasCoruna` es una variación de `MesasGalicia`. Cambiamos la selección de los datos, es decir, en vez de tener todas las mesas de Galicia, tomamos solamente las de Coruña y comparamos sus resultados con los generales de Coruña. Y en vez de guardar las 1.300 mesas, guardamos las 300 mesas con mayor CCP.

Tal y como hemos hecho para Galicia, en las provincias también filtramos los datos adaptando el tamaño censal que debe tener el municipio testigo. Por tanto, a un nivel de confianza  $3\sigma$  y con un error muestral del 2,5% el censo que debe superar el municipio coruñés es de 6.598 electores.

En el cuadro 3.8 mostramos los primeros seis posibles candidatos a ciudad testigo de A Coruña. Los municipios que resaltan en negrita (Noia, Sada y Betanzos) son los que tienen un censo superior a 6.598. Y como consecuencia, son los municipios candidatos a ser ciudad testigo de A Coruña.

	Municipio	Nº mesas	Porcentaje	Censo
1	Cabanas	3 de 4	75 %	2.841
2	Bergondo	2 de 10	20 %	5.613
3	<b>Noia</b>	3 de 19	15,79 %	12.405
4	<b>Sada</b>	2 de 14	14,29 %	12.135
5	Lousame	1 de 8	12,5 %	3.218
6	<b>Betanzos</b>	2 de 19	10,53 %	11.057

Cuadro 3.8: Posibles municipios testigo de A Coruña.

### 3.3. Municipios parecidos a Lugo

En el script `Lugo.R`, evaluamos qué municipio es el que mejor refleja los resultados electorales de Lugo siguiendo el mismo procedimiento que para A Coruña. Sus funciones son:

- `MesasLugo`
- `identificar_mesascomunes`
- `imprime_mesas`
- `porcentajes_municipios`

La lista con los primeros doce posibles candidatos a municipio testigo de Lugo es:

	Municipio	Nº mesas	Porcentaje	Censo
1	Rábade	2 de 2	100 %	1.329
2	Guntín	6 de 7	85,71 %	2.727
3	Páramo (O)	3 de 4	75 %	1.397
4	Lourenzá	3 de 5	60 %	2.075
5	Castro de Rei	4 de 7	57,14 %	4.611
6	Bóveda	2 de 4	50 %	1.426
7	Cervo	5 de 11	45,45 %	3.916
8	Valadouro (O)	2 de 5	40 %	1.869
9	Cospeito	3 de 8	37,5 %	4.512
10	Baralla	2 de 6	33,33 %	2.519
11	<b>Viveiro</b>	7 de 21	33,33 %	12.907
12	Becerreá	2 de 6	33,33 %	2.751

Cuadro 3.9: Posibles municipios testigo de Lugo.

El tamaño censal apropiado para la provincia de Lugo es de 6.994 electores. Si observamos el cuadro 3.9, tenemos que entre los doce posibles candidatos, Viveiro, con un censo electoral de 12.907, es el único candidato a ciudad testigo de Lugo.

### 3.4. Municipios parecidos a Ourense

En el script `Ourense.R`, evaluamos qué municipio es el que mejor refleja los resultados electorales de Ourense. Aplicamos las siguientes funciones vistas anteriormente:

- `MesasOurense`
- `identificar_mesascomunes`
- `imprime_mesas`
- `porcentajes_municipios`

En `MesasOurense` cambiamos la selección de los datos, tomando solamente los de Ourense. Y guardamos las 300 mesas con mayor CCP.

Los primeros diez posibles candidatos a municipio testigo de Ourense son los siguientes:

	<b>Municipio</b>	<b>Nº mesas</b>	<b>Porcentaje</b>	<b>Censo</b>
1	Melón	3 de 3	100 %	1.323
2	Paderne de Allariz	2 de 2	100 %	1.372
3	Gudiña (A)	3 de 3	100 %	1.204
4	Beade	1 de 1	100 %	438
5	Montederramo	3 de 3	100 %	831
6	Larouco	1 de 1	100 %	507
7	Peroxa (A)	5 de 6	83,3 %	1.905
8	Pereiro de Aguiar (O)	6 de 8	75 %	5.261
9	<b>Verín</b>	16 de 22	<b>72,73 %</b>	10.770
10	Taboadela	2 de 3	66,67 %	1.432

Cuadro 3.10: Posibles municipios testigo de Ourense.

El tamaño censal fijado para la provincia de Ourense es de 7.267 electores. Observando el cuadro 3.10, Verín es el candidato a ser municipio testigo de Ourense.

### 3.5. Municipios parecidos a Pontevedra

En el script `Pontevedra.R`, al igual que para el resto de provincias, evaluamos qué municipio es el que mejor refleja los resultados electorales de Pontevedra. Las funciones empleadas son:

- `MesasPontevedra`
- `identificar_mesascomunes`
- `imprime_mesas`
- `porcentajes_municipios`

En `MesasPontevedra` cambiamos la selección de los datos, tomando solamente los de Pontevedra. Y guardamos las 300 mesas con mayor CCP.

La lista con los primeros cinco posibles candidatos a municipio testigo de Pontevedra:

	<b>Municipio</b>	<b>Nº mesas</b>	<b>Porcentaje</b>	<b>Censo</b>
1	<b>Baiona</b>	6 de 12	50 %	9.888
2	Portas	1 de 4	25 %	2.621
3	<b>Marín</b>	5 de 29	17,24 %	20.455
4	Ribadumia	1 de 7	14,29 %	4.122
5	Rosal (O)	1 de 7	14,29 %	5.447
6	Meis	1 de 7	14,29 %	4.186

Cuadro 3.11: Posibles municipios de Pontevedra.

Una vez más filtramos los datos fijando el tamaño muestral del censo electoral que debe superar el municipio testigo. Para Pontevedra, este umbral es de 7.269 electores.

En el cuadro 3.11 tenemos los primeros seis posibles candidatos a ciudad testigo de Pontevedra. Los municipios que poseen un censo superior a 7.269 son Baiona y Marín. Consecuentemente, son los municipios candidatos a ser ciudad testigo de Pontevedra. Parece que Baiona es la ciudad que refleja con mayor exactitud los resultados tanto de la provincia a la que pertenece como los de la comunidad.

# Capítulo 4

## Investigación Social

La investigación social se define como el proceso que permite obtener nuevos conocimientos en el campo de la realidad social (investigación pura) o que permite estudiar una situación social para diagnosticar necesidades y problemas a los efectos de aplicar los conocimientos con finalidades prácticas (investigación aplicada).

Probablemente, la encuesta es la estrategia de investigación más conocida y practicada por los investigadores sociales.

Sus antecedentes se remontan a los mismos orígenes de la indagación social empírica. En concreto, a los estudios realizados por John Sinclair (*Informe estadístico de Escocia, 1791-1825*), James Kay Shuttleworth (*Las condiciones morales y físicas de la vida de los obreros de la industria textil de Manchester, 1832*) y Charles Booth (*Vida y trabajo de los habitantes de Londres, 1889-1891*).

Pero es a partir de la Segunda Guerra Mundial cuando la encuesta se convierte en la estrategia predominante en la investigación social. A ello contribuyeron, desde la vertiente privada, los estudios de mercado llevados a cabo en EEUU por Gallup, Crossley y Ropper; y, desde la vertiente universitaria, los estudios de Lazarsfeld (sobre todo el que llevó con Berelson: *La elección del pueblo, 1944*) y de Stouffer et al. (*El soldado americano, 1949*), principalmente.

### 4.1. Un ejemplo práctico: Predicción electoral mediante la encuesta

En esta sección, vamos a proponer una forma de como predecir en el mismo día de las elecciones al parlamento gallego sus resultados.

Construimos una estrategia de investigación basada en las declaraciones verbales

de una población concreta, es decir, realizaremos una encuesta. Para los objetivos de esta investigación, la modalidad de encuesta más adecuada es mediante la entrevista personal.

#### 4.1.1. Selección del tamaño muestral

En primer lugar, calculamos el tamaño de la muestra. Partimos de los datos recogidos de las últimas elecciones (2012):

ELECCIONES 2012					
Censo Electoral	PP	PSOE	AGE	BNG	OTROS
2.300.335	654.907	293.714	200.211	145.424	98.290

Cuadro 4.1: Datos de las Elecciones al Parlamento de Galicia 2012.

Como la varianza poblacional es desconocida, tomamos el producto de las probabilidades  $P$  (de aparición del suceso) y  $Q$  (no ocurrencia del suceso).

$$P_i = \frac{\text{Partido}}{\text{Censo}} \cdot 100 \quad Q_i = 100 - P_i \quad (4.1)$$

Tenemos así, la siguiente tabla:

	PP	PSOE	AGE	BNG	OTROS
$P_i$	28,47	12,77	8,70	6,32	4,27
$Q_i$	71,53	87,23	91,30	93,68	95,731

Cuadro 4.2: Cálculo de las probabilidades  $P$  y  $Q$  de los datos de las Elecciones 2012.

Ya estamos en condiciones de calcular el tamaño muestral ( $n$ ). Lo realizaremos para dos niveles de confianza y errores de muestreo distintos. Posteriormente se elegirá el que mejor se ajuste a nuestras necesidades.

$$N = 2.300.335 \quad (\text{censo electoral 2012})$$

Como  $N > 100.000$  unidades, entonces  $n = \frac{z^2 \hat{P}^2 (1 - \hat{P})}{\epsilon^2}$ .

Por lo tanto,

$$n = \frac{z^2 \sum_{i=1}^5 P_i Q_i}{\epsilon^2} \quad (4.2)$$

#### 4.1. UN EJEMPLO PRÁCTICO: PREDICCIÓN ELECTORAL MEDIANTE LA ENCUESTA 41

donde  $z$  son las unidades de desviación típica correspondientes al nivel de confianza elegido y  $\epsilon$  es el error muestral.

##### TAMAÑO MUESTRAL: $n$

Error muestral	Nivel de confianza $2\sigma$ (95,5 %)	Nivel de confianza $3\sigma$ (99,7 %)
2.5 %	3.165,13 $\simeq$ 3.166	7.121,55 $\simeq$ 7.122
2 %	4.945,52 $\simeq$ 4.946	11.127,42 $\simeq$ 11.128
1 %	19.782,08 $\simeq$ 19.783	44.509,68 $\simeq$ 44.510

Cuadro 4.3: Posibles tamaños de la muestra.

Se quiere que la encuesta sea efectiva y a su vez lo menos costosa posible. Además, sabemos que en la investigación social, es frecuente que los errores de muestreo se encuentren comprendidos entre el 2 % y 2,5 %. Observando el cuadro 4.3, vemos que hay grandes diferencias sobre el tamaño de la muestra dependiendo del error y nivel de confianza que se aplique. Para nuestro caso, el tamaño muestral más adecuado, es de **3.166** encuestados a un nivel de confianza  $2\sigma$  y con un error muestral del 2,5 %.

Ya se tiene el tamaño muestral, 3.166 personas. Pero, ¿a qué votantes preguntar?

#### 4.1.2. Selección de los votantes a encuestar

A estas alturas, no es difícil averiguar en donde debemos ir a hacer una encuesta. En el capítulo 3 hemos analizado con detalle las mesas y municipios de Galicia. Siendo más concretos, en la subsección 3.1.3 hallamos las mesas testigo para la comunidad realizando el coeficiente de correlación de Pearson (CCP), la divergencia de Kullback-Leibler (KL) y el test de correlación de Szekely para las tres elecciones. Seguidamente, escogíamos las mesas que disponían de un alto CCP (en todas las elecciones) y las ordenábamos de mayor a menor correlación.

Por lo tanto, disponemos de una lista con 285 mesas ordenadas de mayor a menor dependencia. Según del personal que dispongamos para realizar la encuesta, elegimos el número de mesas. Consideramos que una buena cifra es de 46 mesas.

A continuación, en los siguientes cuadros separados por provincias, mostramos las 46 mesas resultantes. En la primera columna, se encuentra el nombre de la mesa electoral formado por: código municipio, municipio, distrito, sección y mesa. Y en la segunda columna, se exponen en donde se sitúa cada mesa (datos obtenidos del boletín oficial de cada provincia, BOP).

**A Coruña:**

<b>MESA ELECTORAL</b>	<b>LOCAL ELECTORAL</b>
2 Ames 3 1 B	Antiga Escola de Firmistáns
5 Arteixo 3 1 B	Escola de Educación Infantil de Lagoa
9 Betanzos 1 7 A	Local Municipal Rúa Saavedra Meneses
22 Cedeira 1 2 A	Colexio Ed. Infantil e Primaria Nicolás del Río
42 Lousame 1 1 B	Colexio Público Integrado de Cernadas de Castro
43 Malpica de Bergantiños 1 2 B	Ludoteca Municipal
56 Negreira 1 3 B	Colexio de Ed. Infantil e Primaria “O Coto”
57 Noia 1 1 B	Local de Ensaio
57 Noia 2 1 A	Casa so Concello (Salón de Actos)
57 Noia 2 4 A	Instituto Social da Mariña
57 Noia 2 6 B	Casa da Xuventude
59 Ordes 1 1 A	Casa da Cultura
61 Ortigueira 3 1 B	Antiga Escola de San Claudio
65 Padrón 3 1 U	Centro Social da Igrexa
78 Santiago de Compostela 2 6 C	Centro Sociocultural da Peregrina
78 Santiago de Compostela 2 8 A	Facultade de Filoloxía
78 Santiago de Compostela 2 12 U	Centro Sociocultural da Gracia

Cuadro 4.4: Nombre de las mesas electorales a las que realizar la encuesta.

**Lugo:**

<b>MESA ELECTORAL</b>	<b>LOCAL ELECTORAL</b>
10 Castro de Rei 1 2 U	Escuela de Outeiro
13 Cervo 1 2 A	CEIP de Galdín
15 Cospeito 1 3 B	Escuela de Xermar
22 Guitiriz 4 2 B	Escuela Formación Profesional
28 Lugo 6 1 B	Escola de “O Veral”
43 Páramo (O) 1 1 B	Centro Social Páramo-Vilarmosteiro
51 Ribadeo 1 1 B	Oficina de Información y Turismo
51 Ribadeo 1 2 A	Cine Teatro
65 Vilalba 1 7 U	Casa Consistorial (Entrada JM <sup>a</sup> Chao)
901 Baralla 1 1 B	Casa de Cultura

Cuadro 4.5: Nombre de las mesas electorales a las que realizar la encuesta.



#### 4.1. UN EJEMPLO PRÁCTICO: PREDICCIÓN ELECTORAL MEDIANTE LA ENCUESTA 43

##### Ourense:

MESA ELECTORAL	LOCAL ELECTORAL
58 Pereiro de Aguiar (O) 1 2 B	Hogar Social de Vilariño
85 Verín 1 1 B	Colegio Amaro Refojo

Cuadro 4.6: Nombre de las mesas electorales a las que realizar la encuesta.

##### Pontevedra:

MESA ELECTORAL	LOCAL ELECTORAL
3 Baiona 1 2 B	Centro Público Integrado de Covaterreña
3 Baiona 1 4 A	Centro Público Integrado de Covaterreña
3 Baiona 2 1 A	Colexio Educación Primaria de Sabarís
26 Marín 1 1 B	Biblioteca Pública Municipal
26 Marín 1 11 B	Residencia Tercera Edad
26 Marín 3 1 A	Casa de Cultura de Seixo
27 Meaño 2 2 U	Colegio de Educación Infantil y Primaria de Coirón
45 Redondela 1 3 B	CEIP de Santo Paio de Abaixo
46 Ribadumia 2 1 B	Centro Socio-Cultural de Couto de Arriba
51 Sanxenxo 1 1 B	Colexio de Educación Infantil e Primaria Florida
51 Sanxenxo 1 6 U	Centro Privado Abrente
55 Tui 1 1 B	Instituto de Educación Secundaria San Paio
57 Vigo 3 3 A	Pavillón do Berbes
57 Vigo 3 21 A	Escuela de Música
57 Vigo 3 15 B	Casa del Ayuntamiento
57 Vigo 6 24 B	IES República Uruguay
57 Vigo 7 11 B	Estación de Autobuses

Cuadro 4.7: Nombre de las mesas electorales a las que realizar la encuesta.

Como conclusión, tenemos que las encuestas se deberían realizar a 69 votantes de cada una de las 46 mesas que se muestran en los cuadros anteriores abarcando una muestra de 3.174 individuos.

Dependiendo del presupuesto del que se disponga, podríamos aumentar el número de mesas a las que encuestar, o incluso variar el tamaño y error muestral.

### 4.1.3. Ventajas e inconvenientes de la encuesta

Un papel importante en el desarrollo de la encuesta es el del entrevistador. Es el que formula las preguntas y quien anota las respuestas de cada uno de los individuos seleccionados de la muestra. Por lo tanto, el éxito de la investigación descansa bastante en la actuación del entrevistador.

El entrevistador, no sólo debe tener la capacidad de establecer empatía con el entrevistado y conseguir su cooperación, sino que también debe de estar adecuadamente preparado para:

- a) Formular correctamente las preguntas del cuestionario.
- b) Asegurar la adecuación de las respuestas y su correspondiente anotación.
- c) Tomar decisiones en el campo, sin la asistencia de un supervisor.

En nuestra encuesta, la intervención del entrevistador se convierte en decisiva, ya que puede facilitar el trabajo de campo, pero también puede obstaculizarlo.

Como toda estrategia de investigación social, la encuesta personal tiene ventajas e inconvenientes. Se muestran en el cuadro siguiente:

<b>VENTAJAS</b>	<b>INCONVENIENTES</b>
- Permite conseguir un mayor porcentaje de respuestas.	- Encarece los costes del estudio, en tiempo y dinero.
- Favorece el tratamiento de temas complejos.	- La presencia del entrevistador puede provocar efectos reactivos en las respuestas de los entrevistados.
- Se obtienen respuestas de mayor calidad y espontaneidad.	- Se generan dificultades para acceder a domicilios particulares y a determinados grupos de población.
- Se puede obtener información complementaria del entrevistado ajena al cuestionario.	

Cuadro 4.8: Ventajas e inconvenientes de la encuesta personal.

# Capítulo 5

## Web Interactiva: Shiny

Con el objeto de que los interesados en este tipo de estudios puedan visualizar información, investigar y manipular datos, hemos creado una web interactiva a través del paquete `shiny`.

Shiny es un paquete desarrollado por el grupo `Rstudio` que se dio a conocer a finales del año 2012. Con él, podemos realizar interfaces web, para la creación de informes interactivos y visualizaciones utilizando R. Estas aplicaciones nos permiten especificar, de forma interactiva, parámetros de entrada utilizando los controles deslizantes, menús desplegables y campos de texto. Además permite incorporar cualquier número de salidas como gráficos, tablas y resúmenes.

Poseer conocimientos de desarrollo web como HTML, CSS y JavaScript, resulta muy útil, ya que son elementos importantes a la hora de crear nuestra web con shiny. A continuación los describimos brevemente:

- **HTML**: es el lenguaje con el que se crean las páginas web. Fundamentalmente se trata de un conjunto de etiquetas que sirven para definir el texto y otros elementos que compondrán una página web.
- **CSS**: este lenguaje sirve para organizar la presentación y aspecto de una página web. Una vez creados los contenidos en HTML, se utiliza el lenguaje CSS para definir el aspecto de cada elemento: color, tamaño y tipo de letra del texto, separación horizontal y vertical entre elementos, posición de cada elemento dentro de la página...
- **JavaScript**: es un lenguaje que puede interactuar con el código HTML, permitiendo a los programadores web utilizar contenido dinámico. Hoy en día es el motor de las aplicaciones más conocidas en el ámbito de Internet: Google, Facebook, Twitter, Outlook...

A la hora de compartir la web interactiva con el resto del mundo existen varias opciones:

a) Implementarla en uno de los servidores que ofrecen en la web de Shiny. Disponen dos modalidades:

i) Subir la aplicación a un servidor que se encuentra en período de pruebas. Se necesita solicitar permiso y es gratuita.

ii) Implementarla en una versión del servidor más profesional. Esta modalidad es de pago y su precio mínimo es de 9.995 dólares al año (el precio varía según la capacidad que se necesite).

b) Descargando e instalando el servidor de Shiny.

Nosotros no hemos subido, por el momento, nuestra aplicación a la nube. La ejecutamos directamente desde `Rstudio`.

## 5.1. Descripción de la web

Nuestra web consta de tres pestañas: *Mapa Interactivo*, *Explorador de Datos* y *Resultados*. En las dos primeras se muestran los datos más relevantes de las elecciones al parlamento de Galicia del año 2012 y, en la última, se añaden los datos relativos a las elecciones de 2005 y 2009.

A continuación, desglosamos cada pestaña.

### • Mapa Interactivo

En la pestaña del *Mapa Interactivo*, pretendemos visualizar y ubicar los resultados electorales del año 2012 por municipios. Al mismo tiempo podemos compararlos con los resultados generales obtenidos en cada provincia y en toda la comunidad.

Disponemos de un panel móvil que se compone de:

- Un menú desplegable: se puede seleccionar la provincia a la que pertenecen los municipios que queremos visualizar en el mapa.

- Dos gráficos de barras: muestran los porcentajes de los votos obtenidos por los principales partidos políticos. El primero varía en función de lo que escojamos en el menú desplegable, el segundo es fijo y muestra los porcentajes que se obtuvieron en toda Galicia.

En la figura 5.1 observamos el panel móvil con las opciones de su menú desplegable y los gráficos de barras correspondientes al porcentaje de votos obtenidos por los principales partidos políticos.



Figura 5.1: Panel móvil.

El mapa lo creamos a partir de *mapbox* (similar a *google maps*). En él, podemos visualizar los 315 municipios de Galicia.

En la figura 5.2, en la imagen de la izquierda, mostramos todos los municipios de Galicia. Cada color representa una provincia y el tamaño de los círculos que rodean a cada municipio son proporcionales al censo electoral que poseen.

Si seleccionamos (“clicamos”) en un municipio, obtenemos una ventana emergente con la siguiente información sobre él:

- Coeficiente de correlación de Pearson.
- Nombre, provincia y código postal.
- Censo electoral.
- El porcentaje de los votos recibidos de las principales fuerzas políticas.

En la figura 5.2, en la imagen de la derecha, mostramos la provincia de A Coruña previamente seleccionada en el menú desplegable del panel móvil y además visualizamos la información de un municipio elegido al azar.

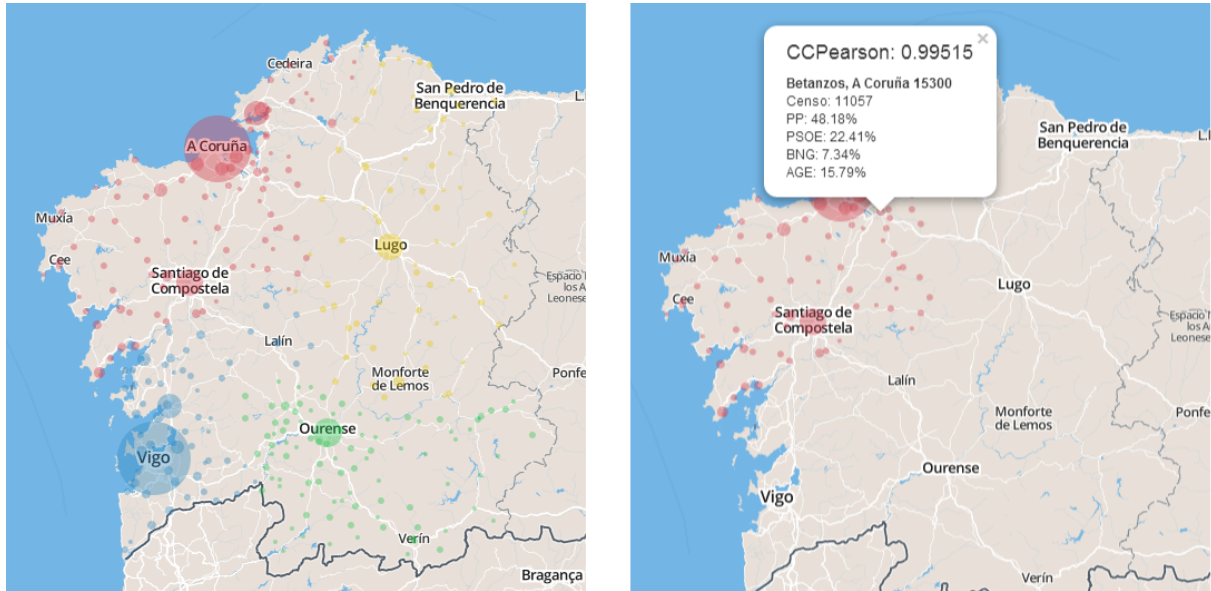


Figura 5.2: Mapa.

Como resultado de la combinación entre el panel móvil y el mapa, obtenemos la pestaña *Mapa Interactivo* tal y como se muestra en la figura 5.3.

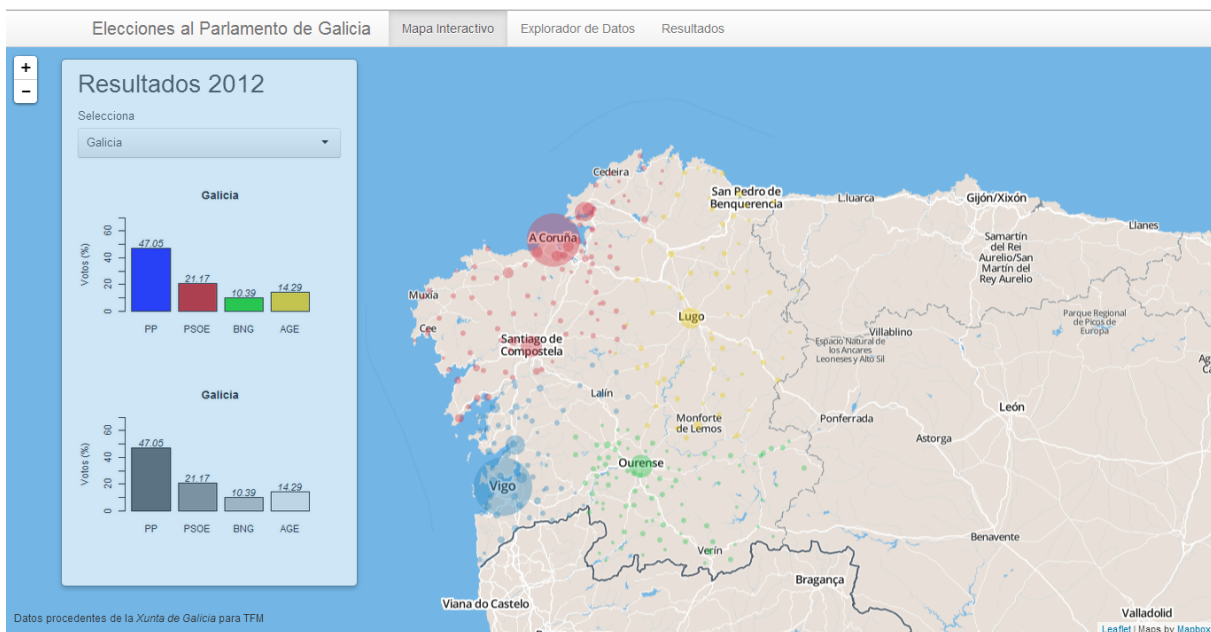


Figura 5.3: Mapa Interactivo.

### • Explorador de Datos

Tal y como indica su nombre, es un espacio destinado a informar sobre los municipios.

De cada municipio disponemos de: la provincia a la que pertenecen, el código postal, el censo electoral, el porcentaje de los votos obtenidos por los partidos políticos (PP, PSOE, BNG y AGE), el coeficiente de correlación de Pearson, las coordenadas geográficas (latitud y longitud) y un icono de localización que nos lleva directamente al mapa de la pestaña anterior, indicando dónde se ubica el municipio elegido.

Al lado de cada columna encontramos unas flechas que nos permiten ordenar los datos de mayor a menor, o viceversa, en función de cualquiera de ellas.

Sobre las columnas, disponemos de un menú desplegable, donde podemos seleccionar las provincias que deseamos consultar. Automáticamente emerge otro menú en el que podemos buscar un municipio por el código postal. En la figura 5.4 observamos los menús. En el ejemplo, hemos seleccionado las provincias de Lugo y Pontevedra.

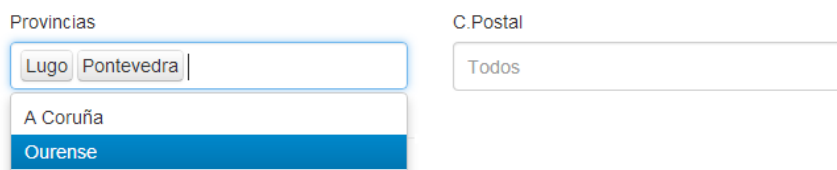


Figura 5.4: Menús desplegables.

En la figura 5.5, tenemos marcados tres elementos. En primero limita el número de municipios que aparecen en la tabla (lo tenemos limitado a 10). El segundo, *Search* busca cualquier dato de la tabla, independientemente de la columna en la que se encuentre. Y el tercer elemento, haciendo “clic” sobre el icono de localización (*crosshair*), sitúa en el mapa el municipio deseado mostrando a su vez información sobre él, tal y como veíamos en la figura 5.2.

Municipio	Provincia	CP	Censo	PP	PSOE	BNG	AGE	CCPearson	Latitud	Longitud	Mapa
Baiona	Pontevedra	36300	9888	50.33	19.51	9.36	13.9	0.99819	42.12015	-8.852367	

Figura 5.5: Otras opciones del Explorador de Datos.

Inicialmente, la pestaña *Explorador de Datos* tiene el aspecto que se expone en la figura 5.6.

Municipio	Provincia	CP	Censo	PP	PSOE	BNG	AGE	CCPearson	Latitud	Longitud	Mapa
Abegondo	A Coruña	15318	4960	56.84	14.70	8.71	14.94	0.98375	43.21912	-8.297905	+
Ames	A Coruña	15864	21911	41.07	16.62	8.65	24.31	0.92085	42.90380	-8.658368	+
Aranga	A Coruña	15317	1907	66.50	17.41	6.47	7.05	0.98990	43.23410	-8.015973	+
Ares	A Coruña	15624	4940	42.41	24.66	13.47	14.78	0.97885	43.42789	-8.243909	+
Arteixo	A Coruña	15142	24035	42.01	18.97	11.47	15.95	0.99338	43.30387	-8.510297	+
Arzúa	A Coruña	15810	5505	54.95	16.11	12.67	12.67	0.98431	42.92741	-8.165074	+
Baña(A)	A Coruña	15863	3600	58.88	21.15	7.77	6.69	0.98887	42.96196	-8.757251	+
Bergondo	A Coruña	15165	5613	50.59	17.81	5.14	19.83	0.97436	43.30897	-8.232594	+
Betanzos	A Coruña	15300	11057	48.18	22.41	7.34	15.79	0.99515	43.27843	-8.213123	+
Boimorto	A Coruña	15818	2004	55.72	26.47	11.11	4.17	0.97154	43.00080	-8.133835	+

Figura 5.6: Explorador de Datos.

## • Resultados

En este último apartado, disponemos de los principales resultados a las elecciones al parlamento de Galicia de los años 2005, 2009 y 2012.

En el menú desplegable de la izquierda, seleccionamos el año del que queremos información.



Figura 5.7: Menú desplegable.

Una vez seleccionado el año que queremos consultar, se nos expone lo siguiente:



- Una tabla con los votos y los escaños conseguidos por las principales fuerzas políticas en cada provincia y en la comunidad.

Resultados	PP.escaños	PP.votos	PSOE.escaños	PSOE.votos	BNG.escaños	BNG.votos	AGE.escaños	AGE.votos	OTROS
A Coruña	13	263167	5	108886	2	55716	4	96278	38201
Lugo	9	98284	4	43370	1	16622	1	19639	7973
Ourense	8	90026	4	43340	1	15492	1	14245	15961
Pontevedra	11	209804	5	101988	3	58197	3	70666	37545
Galicia	41	661281	18	297584	7	146027	9	200828	99680

Figura 5.8: Votos y escaños de los principales partidos políticos (2012).

- Dos diagramas de sectores que muestran los votos y escaños de Galicia.

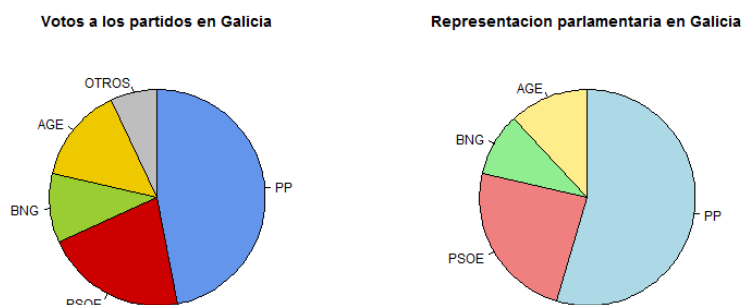


Figura 5.9: Representación de los votos y los escaños de Galicia (2012).

- Una segunda tabla y un diagrama de barras con los porcentajes de los votos obtenidos por los distintos partidos tanto en Galicia como en cada una de las provincias.

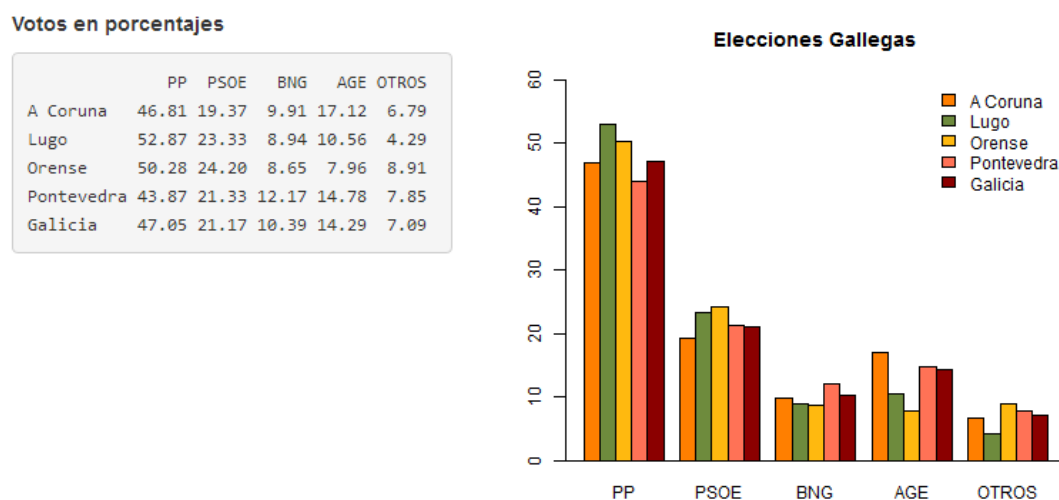


Figura 5.10: Tabla y diagrama de barras de los porcentajes de los votos de Galicia y por provincias (2012).

Finalmente, el conjunto de la pestaña *Resultados* es:

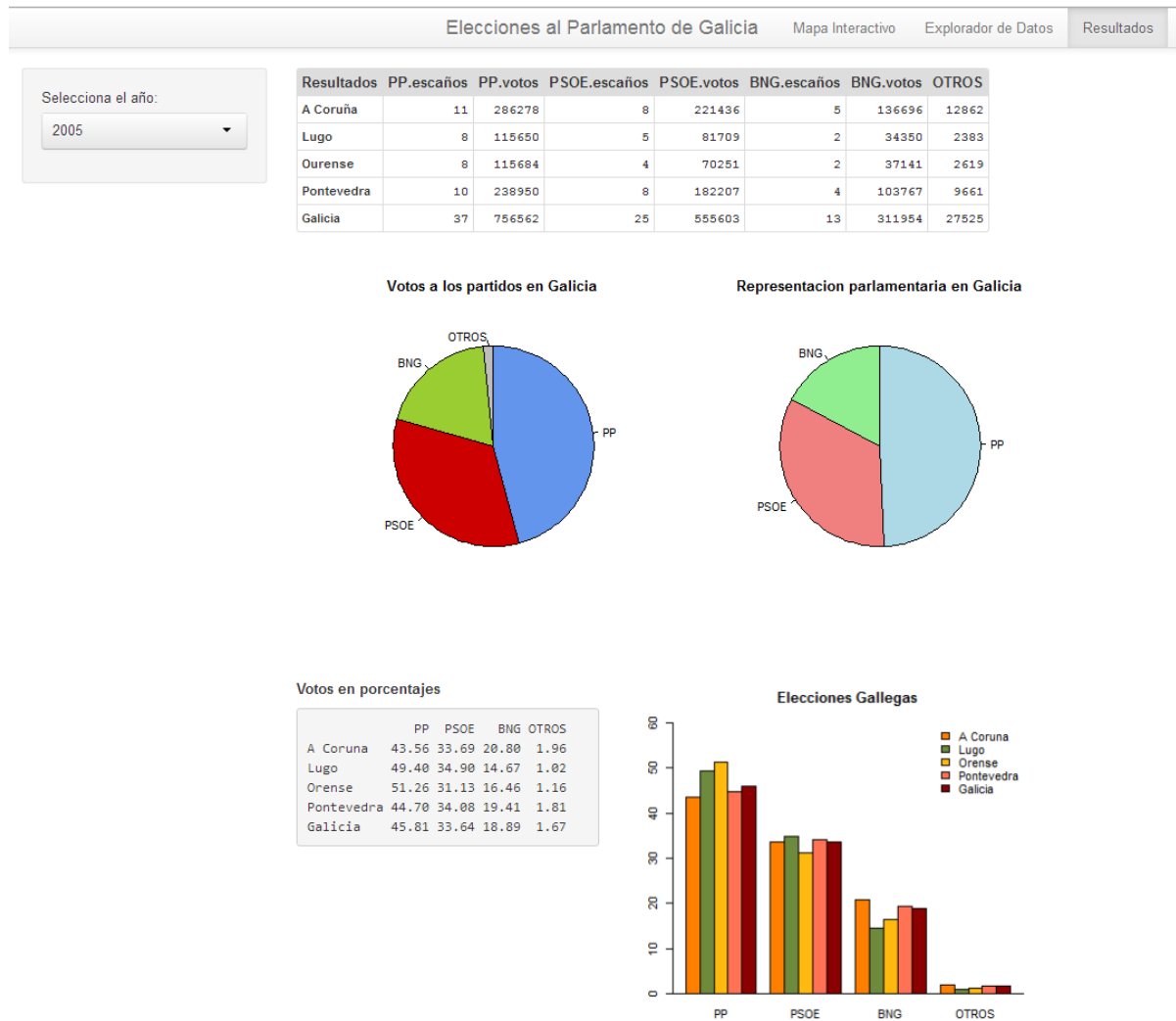


Figura 5.11: Resultados.

## 5.2. Ejecución de la web

Los archivos necesarios para la ejecución de esta web desde RStudio o R, se encuentran en el CD adjunto dentro de la carpeta *web*. Previamente, es necesario instalar los siguientes paquetes: `shiny`, `leaflet`, `RColorBrewer`, `scales`, `lattice` y `dplyr`.

Puede que haya problemas al instalar el paquete `leaflet` puesto que no se encuentra en CRAN. Y por tanto, debemos compilar:

```
install.packages('devtools')
devtools::install_github('leaflet-shiny', 'jcheng5')
```

El paquete `devtools` es un conjunto de herramientas de desarrollo de R que se utiliza para instalar paquetes que no están en CRAN.

Ahora solo queda abrir el archivo `server.R` de la carpeta *web* y realizar lo siguiente:

- Si tenemos la última versión de RStudio, disponemos de un botón que ejecuta directamente nuestra aplicación web (`Run App`).

- Si no tenemos la última versión de RStudio o estamos utilizando R, compilamos en la consola las dos líneas siguientes:

```
library(shiny)
runApp("../web")
```

Desde este momento, ya podemos manejar la página web.



# Apéndice A

## Resultados

A continuación, mostramos los resultados obtenidos de los dos análisis, con y sin votos CERA, del estudio demoscópico para las elecciones al parlamento de Galicia en las convocatorias de 2005, de 2009 y de 2012.

- **Análisis Galicia: mesas electorales**

Con votos CERA		
Provincia	Mesa electoral	$\sum_{i=1}^3 \text{CCP}$
1	A Coruña 78 Santiago de Compostela 2 6 C	2.9960962
2	A Coruña 42 Lousame 1 1 B	2.9950379
3	Lugo 13 Cervo 1 2 A	2.9944452
4	Lugo 51 Ribadeo 1 2 A	2.9941261
5	A Coruña 57 Noia 2 1 B	2.9920098
6	Pontevedra 46 Ribadumia 1 1 B	2.9912943
7	Ourense 85 Verín 1 1 B	2.9911489
8	Pontevedra 26 Marín 1 1 B	2.9909609
9	Pontevedra 57 Vigo 7 11 B	2.9909172
10	Pontevedra 3 Baiona 2 1 A	2.9906418
11	Lugo 43 O Páramo 1 1 B	2.9900368
12	Pontevedra 26 Marín 1 11 B	2.9899774
13	Lugo 65 Vilalba 1 7 U	2.9897545
14	A Coruña 9 Betanzos 1 7 A	2.9896224
15	A Coruña 61 Ortigueira 3 1 B	2.9895962

Sin votos CERA		
Provincia	Mesa electoral	$\sum_{i=1}^3 \text{CCP}$
1	A Coruña 42 Lousame 1 1 B	2.9978719
2	Lugo 13 Cervo 1 2 A	2.9972689
3	A Coruña 78 Santiago de Compostela 2 6 C	2.9938972
4	Lugo 65 Vilalba 1 7 U	2.9925219
5	Pontevedra 27 Meaño 2 2 U	2.9921801
6	Ourense 85 Verín 1 1 B	2.9920981
7	A Coruña 57 Noia 2 4 A	2.991708
8	A Coruña 57 Noia 2 6 B	2.9916811
9	A Coruña 43 Malpica de Bergantiños 1 2 B	2.9914323
10	Lugo 51 Ribadeo 1 2 A	2.9912435
11	Pontevedra 51 Sanxenxo 1 6 U	2.9912402
12	A Coruña 2 Ames 3 1 B	2.9911627
13	A Coruña 57 Noia 1 1 B	2.9909612
14	Pontevedra 3 Baiona 2 1 A	2.9908296
15	Lugo 22 Guitiriz 4 2 B	2.9905138

• Análisis Galicia: municipios

Con votos CERA			
	Municipio	Nº mesas	Porcentaje
1	Rábade	2 de 2	100 %
2	Baiona	8 de 12	66.67 %
3	Ribadumia	4 de 7	57.14 %
4	Guarda (A)	6 de 12	50 %
5	Páramo (O)	2 de 4	50 %
6	Cabanas	2 de 4	50 %
7	Paderne de Allariz	1 de 2	50 %
8	Noia	9 de 19	47.37 %
9	Marín	13 de 29	44.83 %
10	Rúa (A)	3 de 7	42.86 %
11	Meis	3 de 7	42.86 %
12	Nigrán	7 de 18	38.89 %
13	Ribadeo	5 de 13	38.46 %
14	Laxe	3 de 8	37.5 %
15	Negreira	5 de 15	33.33 %
16	Meira	1 de 3	33.33 %
17	Melón	1 de 3	33.33 %
18	Gomesende	1 de 3	33.33 %
19	Gudiña (A)	1 de 3	33.33 %
20	Mos	6 de 19	31.58 %
21	Betanzos	6 de 19	31.58 %
22	Cee	3 de 10	30 %
23	Bergondo	3 de 10	30 %
24	Curtis	2 de 7	28.57 %
25	Viveiro	6 de 21	28.57 %
26	Rosal (O)	2 de 7	28.57 %
27	Cervo	3 de 11	27.27 %
28	Verín	6 de 22	27.27 %
29	Lousame	2 de 8	25 %
30	Cariño	2 de 8	25 %
31	Meaño	2 de 8	25 %
32	Oroso	2 de 8	25 %
33	Portas	1 de 4	25 %
34	Ordes	4 de 18	22.22 %
35	Taboada	2 de 9	22.22 %
36	Oza dos Ríos	1 de 5	20 %
37	Monforte de Lemos	6 de 30	20 %
38	Lourenzá	1 de 5	20 %
39	Silleda	3 de 16	18.75 %
40	Lugo	21 de 114	18.42 %
41	Padrón	2 de 11	18.18 %
42	Ponteceso	2 de 11	18.18 %
43	Baralla	1 de 6	16.67 %
44	Viana do Bolo	1 de 6	16.67 %
45	Porto do Son	2 de 12	16.67 %
46	Pastoriza (A)	1 de 6	16.67 %
47	Agolada	1 de 6	16.67 %
48	Foz	2 de 12	16.67 %
49	Corgo (O)	1 de 6	16.67 %
50	Paderne	1 de 6	16.67 %
51	Boqueixón	1 de 6	16.67 %
52	Outes	2 de 12	16.67 %
53	Vilardevós	1 de 6	16.67 %
54	Malpica de Bergantiños	2 de 14	14.29 %
55	Castro de Rei	1 de 7	14.29 %
56	Boborás	1 de 7	14.29 %
57	Redondela	5 de 37	13.51 %
58	Ortigueira	2 de 15	13.33 %
59	Cospeito	1 de 8	12.5 %
60	Moraña	1 de 8	12.5 %

Sin votos CERA			
	Municipio	Nº mesas	Porcentaje
1	Rábade	2 de 2	100 %
2	Baiona	8 de 12	66.67 %
3	Nigrán	9 de 18	50 %
4	Páramo (O)	2 de 4	50 %
5	Cabanas	2 de 4	50 %
6	Paderne de Allariz	1 de 2	50 %
7	Marín	14 de 29	48.28 %
8	Noia	9 de 19	47.37 %
9	Curtis	3 de 7	42.86 %
10	Ribadumia	3 de 7	42.86 %
11	Meis	3 de 7	42.86 %
12	Guarda (A)	5 de 12	41.67 %
13	Lourenzá	2 de 5	40 %
14	Ribadeo	5 de 13	38.46 %
15	Betanzos	7 de 19	36.84 %
16	Meira	1 de 3	33.33 %
17	Melón	1 de 3	33.33 %
18	Gomesende	1 de 3	33.33 %
19	Gudiña (A)	1 de 3	33.33 %
20	Verín	7 de 22	31.82 %
21	Bergondo	3 de 10	30 %
22	Viveiro	6 de 21	28.57 %
23	Rosal (O)	2 de 7	28.57 %
24	Ordes	5 de 18	27.78 %
25	Padrón	3 de 11	27.27 %
26	Negreira	4 de 15	26.67 %
27	Lousame	2 de 8	25 %
28	Oroso	2 de 8	25 %
29	Meaño	2 de 8	25 %
30	Portas	1 de 4	25 %
31	Laxe	2 de 8	25 %
32	Outes	3 de 12	25 %
33	Trasmiras	1 de 4	25 %
34	Taboada	2 de 9	22.22 %
35	Vedra	2 de 9	22.22 %
36	Malpica de Bergantiños	3 de 14	21.43 %
37	Mos	4 de 19	21.05 %
38	Monforte de Lemos	6 de 30	20 %
39	Oza dos Ríos	1 de 5	20 %
40	Cambados	3 de 16	18.75 %
41	Cervo	2 de 11	18.18 %
42	Ponteceso	2 de 11	18.18 %
43	Lugo	19 de 114	16.67 %
44	Muros	3 de 18	16.67 %
45	Baralla	1 de 6	16.67 %
46	Viana do Bolo	1 de 6	16.67 %
47	Porto do Son	2 de 12	16.67 %
48	Pastoriza (A)	1 de 6	16.67 %
49	Agolada	1 de 6	16.67 %
50	Barro	1 de 6	16.67 %
51	Ares	1 de 6	16.67 %
52	Aranga	1 de 6	16.67 %
53	Boqueixón	1 de 6	16.67 %
54	Corgo (O)	1 de 6	16.67 %
55	Vilardevós	1 de 6	16.67 %
56	Sanxenxo	3 de 19	15.79 %
57	Rúa (A)	1 de 7	14.29 %
58	Castro de Rei	1 de 7	14.29 %
59	Incio (O)	1 de 7	14.29 %
60	Boborás	1 de 7	14.29 %

## • Municipios parecidos a A Coruña

Con votos CERA			
	Municipio	Nº mesas	Porcentaje
1	Cabanas	3 de 4	75 %
2	Bergondo	3 de 10	30 %
3	Noia	3 de 19	15.79 %
4	Lousame	1 de 8	12.5 %
5	Betanzos	2 de 19	10.53 %
6	Cee	1 de 10	10 %
7	Pontedeume	1 de 12	8.33 %
8	Porto do Son	1 de 12	8.33 %
9	Sada	1 de 14	7.14 %
10	Malpica de Bergantiños	1 de 14	7.14 %
11	Ames	2 de 29	6.9 %
12	Ordes	1 de 18	5.56 %
13	Santiago de Compostela	6 de 113	5.31 %
14	Boiro	1 de 22	4.55 %
15	Oleiros	1 de 39	2.56 %
16	Coruña (A)	1 de 327	0.31 %

Sin votos CERA			
	Municipio	Nº mesas	Porcentaje
1	Cabanas	3 de 4	75 %
2	Bergondo	2 de 10	20 %
3	Noia	3 de 19	15.79 %
4	Sada	2 de 14	14.29 %
5	Lousame	1 de 8	12.5 %
6	Betanzos	2 de 19	10.53 %
7	Pontedeume	1 de 12	8.33 %
8	Porto do Son	1 de 12	8.33 %
9	Malpica de Bergantiños	1 de 14	7.14 %
10	Ames	2 de 29	6.9 %
11	Negreira	1 de 15	6.67 %
12	Ordes	1 de 18	5.56 %
13	Boiro	1 de 22	4.55 %
14	Santiago de Compostela	5 de 113	4.42 %
15	Oleiros	1 de 39	2.56 %
16	Coruña (A)	1 de 327	0.31 %

## • Municipios parecidos a Lugo

Con votos CERA			
	Municipio	Nº mesas	Porcentaje
1	Rábade	2 de 2	100 %
2	Páramo (O)	3 de 4	75 %
3	Guntín	5 de 7	71.43 %
4	Castro de Rei	4 de 7	57.14 %
5	Bóveda	2 de 4	50 %
6	Ourol	2 de 5	40 %
7	Lourenzá	2 de 5	40 %
8	Valadouro (O)	2 de 5	40 %
9	Palas de Rei	3 de 8	37.5 %
10	Cospeito	3 de 8	37.5 %
11	Cervo	4 de 11	36.36 %
12	Viveiro	7 de 21	33.33 %
13	Foz	4 de 12	33.33 %
14	Guitiriz	4 de 12	33.33 %
15	Becerreá	2 de 6	33.33 %
16	Meira	1 de 3	33.33 %
17	Ribadeo	4 de 13	30.77 %
18	Lugo	35 de 114	30.7 %
19	Sarria	6 de 20	30 %
20	Friol	3 de 10	30 %
21	Paradela	1 de 4	25 %
22	Baleira	1 de 4	25 %
23	Saviñao (O)	3 de 12	25 %
24	Taboada	2 de 9	22.22 %
25	Mondoñedo	2 de 9	22.22 %
26	Samos	1 de 5	20 %
27	Monterroso	1 de 5	20 %
28	Baralla	1 de 6	16.67 %
29	Pastoriza (A)	1 de 6	16.67 %
30	Monforte de Lemos	5 de 30	16.67 %
31	Incio (O)	1 de 7	14.29 %
32	Vilalba	3 de 24	12.5 %
33	Fonsagrada (A)	1 de 12	8.33 %

Sin votos CERA			
	Municipio	Nº mesas	Porcentaje
1	Rábade	2 de 2	100 %
2	Guntín	6 de 7	85.71 %
3	Páramo (O)	3 de 4	75 %
4	Lourenzá	3 de 5	60 %
5	Castro de Rei	4 de 7	57.14 %
6	Bóveda	2 de 4	50 %
7	Cervo	5 de 11	45.45 %
8	Valadouro (O)	2 de 5	40 %
9	Cospeito	3 de 8	37.5 %
10	Baralla	2 de 6	33.33 %
11	Viveiro	7 de 21	33.33 %
12	Becerreá	2 de 6	33.33 %
13	Meira	1 de 3	33.33 %
14	Ribadeo	4 de 13	30.77 %
15	Sarria	6 de 20	30 %
16	Friol	3 de 10	30 %
17	Lugo	32 de 114	28.07 %
18	Monforte de Lemos	8 de 30	26.67 %
19	Palas de Rei	2 de 8	25 %
20	Guitiriz	3 de 12	25 %
21	Paradela	1 de 4	25 %
22	Baleira	1 de 4	25 %
23	Saviñao (O)	3 de 12	25 %
24	Taboada	2 de 9	22.22 %
25	Mondoñedo	2 de 9	22.22 %
26	Ourol	1 de 5	20 %
27	Foz	2 de 12	16.67 %
28	Vilalba	4 de 24	16.67 %
29	Pastoriza (A)	1 de 6	16.67 %
30	Antas de Ulla	1 de 6	16.67 %
31	Incio (O)	1 de 7	14.29 %
32	Pantón	1 de 7	14.29 %
33	Fonsagrada (A)	1 de 12	8.33 %

• Municipios parecidos a Ourense

Con votos CERA			
	Municipio	Nº mesas	Porcentaje
1	Melón	3 de 3	100 %
2	Paderne de Allariz	2 de 2	100 %
3	Gudiña (A)	3 de 3	100 %
4	Beade	1 de 1	100 %
5	Montederramo	3 de 3	100 %
6	Larouco	1 de 1	100 %
7	Peroxa (A)	5 de 6	83.33 %
8	Verín	15 de 22	68.18 %
9	Taboadela	2 de 3	66.67 %
10	Gomesende	2 de 3	66.67 %
11	Pereiro de Aguiar (O)	5 de 8	62.5 %
12	Leiro	2 de 4	50 %
13	Mezquita (A)	1 de 2	50 %
14	Padrenda	2 de 4	50 %
15	Oímbra	2 de 4	50 %
16	Rubiá	2 de 4	50 %
17	Arnoia (A)	1 de 2	50 %
18	San Xoán de Río	1 de 2	50 %
19	Ribadavia	5 de 11	45.45 %
20	Celanova	4 de 9	44.44 %
21	Toén	2 de 5	40 %
22	Carballiño (O)	7 de 20	35 %
23	Vilardevós	2 de 6	33.33 %
24	Vilamartín de Valdeorras	2 de 6	33.33 %
25	Verea	1 de 3	33.33 %
26	Barco de Valdeorras (O)	6 de 19	31.58 %
27	Xinzo de Limia	5 de 17	29.41 %
28	Rúa (A)	2 de 7	28.57 %
29	Xunqueira de Ambía	1 de 4	25 %
30	Trasmiras	1 de 4	25 %
31	Cenlle	1 de 4	25 %
32	Carballeda de Valdeorras	1 de 4	25 %
33	Lobeira	1 de 4	25 %
34	San Cibrao das Viñas	1 de 4	25 %
35	Barbadás	2 de 10	20 %
36	Piñor	1 de 5	20 %
37	Viana do Bolo	1 de 6	16.67 %
38	Lobios	1 de 6	16.67 %
39	Ourense	21 de 144	14.58 %
40	Boborás	1 de 7	14.29 %
41	Nogueira de Ramuín	1 de 7	14.29 %

Sin votos CERA			
	Municipio	Nº mesas	Porcentaje
1	Melón	3 de 3	100 %
2	Paderne de Allariz	2 de 2	100 %
3	Gudiña (A)	3 de 3	100 %
4	Beade	1 de 1	100 %
5	Montederramo	3 de 3	100 %
6	Larouco	1 de 1	100 %
7	Peroxa (A)	5 de 6	83.33 %
8	Pereiro de Aguiar (O)	6 de 8	75 %
9	Verín	16 de 22	72.73 %
10	Taboadela	2 de 3	66.67 %
11	Gomesende	2 de 3	66.67 %
12	Celanova	5 de 9	55.56 %
13	Leiro	2 de 4	50 %
14	Mezquita (A)	1 de 2	50 %
15	Oímbra	2 de 4	50 %
16	Padrenda	2 de 4	50 %
17	Arnoia (A)	1 de 2	50 %
18	Rubiá	2 de 4	50 %
19	San Xoán de Río	1 de 2	50 %
20	Ribadavia	5 de 11	45.45 %
21	Toén	2 de 5	40 %
22	Vilardevós	2 de 6	33.33 %
23	Verea	1 de 3	33.33 %
24	Carballiño (O)	6 de 20	30 %
25	Xinzo de Limia	5 de 17	29.41 %
27	Xunqueira de Ambía	1 de 4	25 %
28	Trasmiras	1 de 4	25 %
29	Cenlle	1 de 4	25 %
30	Carballeda de Valdeorras	1 de 4	25 %
31	Lobeira	1 de 4	25 %
32	San Cibrao das Viñas	1 de 4	25 %
33	Cualedro	1 de 4	25 %
34	Barco de Valdeorras (O)	4 de 19	21.05 %
35	Barbadás	2 de 10	20 %
36	Piñor	1 de 5	20 %
37	Bande	1 de 5	20 %
38	Viana do Bolo	1 de 6	16.67 %
39	Lobios	1 de 6	16.67 %
40	Vilamartín de Valdeorras	1 de 6	16.67 %
41	Ourense	21 de 144	14.58 %
42	Boborás	1 de 7	14.29 %
43	Nogueira de Ramuín	1 de 7	14.29 %



• Municipios parecidos a Pontevedra

Con votos CERA			
	Municipio	Nº mesas	Porcentaje
1	Baiona	6 de 12	50 %
2	Portas	1 de 4	25 %
3	Marín	5 de 29	17.24 %
4	Mos	3 de 19	15.79 %
5	Ribadumia	1 de 7	14.29 %
6	Rosal (O)	1 de 7	14.29 %
7	Meaño	1 de 8	12.5 %
8	Moraña	1 de 8	12.5 %
9	Nigrán	2 de 18	11.11 %
10	Cuntis	1 de 9	11.11 %
11	Sanxenxo	2 de 19	10.53 %
12	Ponteareas	2 de 29	6.9 %
13	Gondomar	1 de 16	6.25 %
14	Silleda	1 de 16	6.25 %
15	Cambados	1 de 16	6.25 %
16	Redondela	2 de 37	5.41 %
17	Tui	1 de 20	5 %
18	Pontevedra	3 de 100	3 %
19	Estrada (A)	1 de 42	2.38 %
20	Vigo	7 de 364	1.92 %

Sin votos CERA			
	Municipio	Nº mesas	Porcentaje
1	Baiona	6 de 12	50 %
2	Portas	1 de 4	25 %
3	Marín	5 de 29	17.24 %
4	Ribadumia	1 de 7	14.29 %
5	Rosal (O)	1 de 7	14.29 %
6	Meis	1 de 7	14.29 %
7	Moraña	1 de 8	12.5 %
8	Meaño	1 de 8	12.5 %
9	Cuntis	1 de 9	11.11 %
10	Mos	2 de 19	10.53 %
11	Sanxenxo	2 de 19	10.53 %
12	Neves (As)	1 de 10	10 %
13	Redondela	3 de 37	8.11 %
14	Ponteareas	2 de 29	6.9 %
15	Gondomar	1 de 16	6.25 %
16	Nigrán	1 de 18	5.56 %
17	Poio	1 de 19	5.26 %
18	Tui	1 de 20	5 %
19	Porriño (O)	1 de 20	5 %
20	Estrada (A)	1 de 42	2.38 %
21	Vigo	8 de 364	2.2 %
22	Pontevedra	1 de 100	1 %



# Bibliografía

- [1] COTELO, CARMEN y NAYA, SALVADOR, *Betanzos, magic town*, Biblos, 2013.
- [2] CEA D ANCONA, M. ÁNGELES, *Metodología cuantitativa: estrategias y técnicas de investigación social*, Síntesis, 2001.
- [3] SÁNCHEZ CARRIÓN, J.JAVIER, *Manual de análisis de datos*, Alianza Editorial, 1995.
- [4] G. J. SZEKELY y M. L. RIZZO, *Energy statistics: A class of statistics based on distances*, *Journal of Statistical Planning and Inference* (2013).
- [5] BAYES, THOMAS (1763), *An Essay towards solving a Problem in the Doctrine of Chances*. *Philosophical Transactions of the Royal Society of London* 53: pp. 370-418.
- [6] WEISSTEIN, ERIC W., *Correlation Coefficient*, MathWorld. Wolfram Research.
- [7] BEELEY, CHRIS, *Web Application Development with R Using Shiny*, 2013.
- [8] PARDO LLORENTE, LEANDRO, *Teoría de la Información Estadística*. Departamento de Estadística e I. O. de la Facultad de Matemáticas (Universidad Complutense de Madrid).
- [9] BERNARDO, J. MIGUEL, *Una alternativa a la ley d'Hondt*. *El País* (02/03/2004)
- [10] *El teorema del votante medio*. *Revista Perfil* (26/01/2013)
- [11] HOTELLING, HAROLD (1929), *Stability in Competition*. *The Economic Journal* 39: pp. 41-57.
- [12] SINCLAIR, JOHN, *Informe estadístico de Escocia, 1791-1825*.
- [13] KAY SHUTTLEWORTH, JAMES, *Las condiciones morales y físicas de la vida de los obreros de la industria textil de Manchester, 1832*.
- [14] BOOTH, CHARLES, *Vida y trabajo de los habitantes de Londres, 1889-1891*.

- [15] LAZARSELD y BERELSON, *La elección del pueblo, 1944*.
- [16] STOUFFER, *El soldado americano, 1949*.
- [17] MATEOS, ARACELI y PENADÉS, ALBERTO, *Las encuestas electorales en la prensa escrita (2008-2011). Errores, sesgos y transparencia*, Universidad de Salamanca.
- [18] TENA-SÁNCHEZ, JORDI, *El Impacto Electoral de los Atentados del 11M en Madrid. Mecanismos explicativos*. Department of Sociology, Universitat Autònoma de Barcelona.
- [19] Abert@s Portal Open Data da Xunta de Galicia: <http://abertos.xunta.es/catalogo/administracion-publica/-/dataset/0261/elecciones-parlamento-galicia-resultados-2005> [Consulta: viernes, 27 de junio de 2014]
- [20] Abert@s Portal Open Data da Xunta de Galicia: <http://abertos.xunta.es/catalogo/administracion-publica/-/dataset/0260/elecciones-parlamento-galicia-resultados-2009> [Consulta: viernes, 27 de junio de 2014]
- [21] Abert@s Portal Open Data da Xunta de Galicia: <http://abertos.xunta.es/catalogo/administracion-publica/-/dataset/0048/elecciones-parlamento-galicia-resultados-2012> [Consulta: viernes, 27 de junio de 2014]
- [22] Ministerio del Interior: <http://www.infoelectoral.mir.es/index.htm> [Consulta: viernes, 27 de junio de 2014]
- [23] Ministerio del Interior: <http://www.resultados2014.interior.es/99PE/DPE11999CM.htm?d=200> [Consulta: viernes, 27 de junio de 2014]
- [24] Iniciativa Debate: <http://iniciativadebate.org/2013/04/10/diferencia-entre-voto-en-blanco-nulo-y-abstencion/> [Consulta: viernes, 27 de junio de 2014]
- [25] Lista de municipios y calles españolas con su código postal: <http://agj.es/> [Consulta: viernes, 27 de junio de 2014]
- [26] Instituto Galego de Estatística: <http://www.ige.eu/web/index.jsp?paxina=001&idioma=gl> [Consulta: viernes, 27 de junio de 2014]
- [27] Instituto Nacional de Estadística: <http://www.ine.es/> [Consulta: viernes, 27 de junio de 2014]

- [28] Paquete energy: <http://www.sciencedirect.com/science/article/pii/S0378375813000633> [Consulta: viernes, 27 de junio de 2014]
- [29] Paquete pyramid: <http://minato.sip21c.org/swtips/R.html#PYRAMID> [Consulta: viernes, 27 de junio de 2014]
- [30] Paquete gdata: <http://cran.r-project.org/web/packages/gdata/index.html> [Consulta: viernes, 27 de junio de 2014]
- [31] Paquete shiny: <http://shiny.rstudio.com/> [Consulta: viernes, 27 de junio de 2014]
- [32] Paquete devtools: <http://www.rstudio.com/products/rpackages/devtools/> [Consulta: viernes, 27 de junio de 2014]
- [33] Paquete RColorBrewer: <http://colorbrewer2.org/> [Consulta: viernes, 27 de junio de 2014]
- [34] Paquete lattice: <http://cran.r-project.org/web/packages/lattice/lattice.pdf> [Consulta: viernes, 27 de junio de 2014]
- [35] Paquete scales: <https://github.com/hadley/scales> [Consulta: viernes, 27 de junio de 2014]
- [36] Paquete leaflet: <https://github.com/jcheng5/leaflet-shiny> [Consulta: viernes, 27 de junio de 2014]
- [37] Paquete dplyr: <https://github.com/hadley/dplyr> [Consulta: viernes, 27 de junio de 2014]