Universidad de Valladolid

Facultad de Ciencias

Trabajo Fin de Grado

Grado en Estadística

Tratamiento automático de encuestas con R

Autor:

D. Raúl Hernansanz Quevedo

Tutor:

D. Jesús M. Rodríguez

Agradecimientos

El llegar hasta aquí y terminar este trabajo no ha sido esfuerzo de uno solo y, aunque se merezcan mucho más, me gustaría al menos dar las gracias al resto de personas que lo han hecho posible.

Gracias a mi madre, Ana, por estar siempre ahí, por apoyarme en todo momento y por todo el cariño que me has dado durante estos cinco años, o bueno, durante estos 23 años.

Gracias a mis tíos, Gemma y Javier, que aunque estén un poco más lejos, siempre están ahí para apoyarme y ayudarme.

Gracias a esas dos personas que han hecho de esta etapa universitaria, una experiencia inolvidable. A Adrián por amenizar (casi todas) las clases y, con su sabiduría, solucionar hasta los peores problemas, y a Irene, por su paciencia, apoyo y por esa sonrisa que iluminaba los días más oscuros. Gracias una vez más y deciros que espero poder dároslas otra vez dentro de cinco, diez o quince años más.

Gracias también a mi tutor, Jesús M. Rodríguez por el tiempo y apoyo prestado durante la realización de este trabajo.

Por último, quiero agradecer también a esas personas que, a pesar de no estar, sé que me están dando ánimos como los que más. Papá, abuelos, va por vosotros.

Índice general

Índice de figuras **7**

1. Introducción **13**

1.1. Motivación 13

1.2. Objetivos 13

1.3. Estructura de la memoria 14

2. Conceptos previos **17**

2.1. Estructura de las encuestas 17

2.2. Librerías de R 20

2.2.1. Openxlsx 20

2.2.2. R2HTML 20

2.2.3. PNG 20

2.2.4. Officer 20

3. Paquetes de R para el tratamiento de encuestas **21**

3.1. Paquetes existentes para el tratamiento de encuestas 21

3.1.1. Survey 21

3.1.2. SQLSurvey 22

3.1.3. EnQuireR 22

3.1.4. Aplicación creada 22

4. Fuentes de microdatos **23**

4.1. En España 23

4.1.1. Estadística de Castilla y Leon 23

4.1.2. Instituto Nacional de Estadística 24

4.1.3. Consejo Superior de Investigación Científica 26
4.1.4. Datos por petición 27
4.2. En Europa 27
4.2.1. EUROSTAT, (Oficina Europea de Estadística) 27
4.3. Otros 28

5. Aplicación principal **29**
5.1. Explicación del funcionamiento de la aplicación 29
5.2. Módulos de la aplicación 31
5.2.1. Módulo de lectura de datos 31
5.2.2. Módulo de lectura del diccionario 32
5.2.3. Módulo de creación de tablas 32
5.2.4. Módulo de creación de gráficos de barras 33
5.2.5. Módulo de creación de gráficos de sectores 35
5.2.6. Módulo de modificación de nombres 36
5.2.7. Módulo de lectura de parámetros 37
5.2.8. Módulo de creación de informes 37
5.2.9. Módulo de lanzamiento de la aplicación 40
5.3. Ejemplo de funcionamiento 41

6. Pruebas de funcionamiento de la aplicación **43**

7. Conclusiones y trabajo futuro **51**
7.1. Conclusiones 51
7.2. Líneas de trabajo futuro 52

8. Bibliografía **53**

Anexos **55**

A. Manual de Instalación de R **57**

B. Manual de uso de la Interfaz **63**

C. Contenido del CD **71**

Índice de figuras

2.1. Ejemplo de microdatos 18
2.2. Estructura de microdatos 18

4.1. Web de estadística de Castilla y Leon 24
4.2. Web del Instituto Nacional de Estadística 25
4.3. Repositorio del Instituto Nacional de Estadística 25
4.4. Web del Instituto Nacional de Estadística 25
4.5. Web del Envejecimiento en red 26
4.6. Microdatos en la web del Envejecimiento en red 27
4.7. Microdatos en la web Eurostat 28

5.1. Módulo de creación de informes 1 37
5.2. Módulo de creación de informes 2 38
5.3. Módulo de creación de informes 3 39
5.4. Gráfico de barras verticales 41
5.5. Gráfico de barras horizontales 41
5.6. Gráfico de sectores 42
5.7. Tabla ejemplo 42

6.1. Fichero de texto 43
6.2. Estructura del diccionario 44
6.3. Informe en formato Word 45
6.4. Informe en formato Excel 46
6.5. Informe en formato Power Point 47
6.6. Informe en formato HTML 48
6.7. Informe en formato Imagen 49

A.1. Página web de R (parte 1) 57
A.2. Página web de R (parte 2) 58
A.3. Página web de R (parte 3) 58
A.4. Instalación de R (parte 1) 59
A.5. Instalación de R (parte 2) 59
A.6. Instalación de R (parte 3) 60
A.7. Instalación de R (parte 4) 60
A.8. Instalación de R (parte 5) 61
A.9. Instalación de R (parte 6) 61
A.10.Instalación de R (parte 7) 62

B.1. Contenidos de la carpeta de la aplicación 63
B.2. Interfaz de usuario (parte 1) 64
B.3. Interfaz de usuario (parte 2) 65
B.4. Ruta de R (parte 1) 65
B.5. Ruta de R (parte 2) 66
B.6. Ruta de R (parte 3) 66
B.7. Creación del diccionario 67
B.8. Creación de visualizaciones 68
B.9. Ruta de R (parte 4) 68
B.10.Ejemplo de variables 69
B.11.Ejemplo de parámetros 69

Abstract

En este Trabajo Fin de Grado se desarrollará una aplicación que creará documentos formados por distintas visualizaciones. Para esta tarea, la aplicación recibirá una serie de parámetros, entre los cuales se encuentran la lista de visualizaciones que el usuario desea, así como el formato de salida, creando con esto los diferentes informes. Además de esto, en este documento se hablará brevemente sobre las librerías que existen actualmente para el tratamiento de encuestas y sobre diferentes fuentes de datos desde las que cualquier persona puede descargarse resultados de encuestas.

Abstract

In this End of Degree Dissertation, an aplication that will create different types of documents formed by different visualizations will be developed. For this task, the aplication will receive some parameters, among which are the list of visualizations that the user wants, as well as the output format, creating with this the diferent reports. In addition to this, in this document we will talk about the libraries that currently exist for the treatment of surveys and about the different sources of data from which anyone can download survey results.

Capítulo 1

Introducción

Este proyecto titulado "Tratamiento automático de encuestas con R " ha sido desarrollado por Raúl Hernansanz Quevedo bajo la tutela de Jesús M. Rodríguez.

En este apartado se hablará de la motivación de este proyecto, de los objetivos propuestos para todo el trabajo y de la estructura de la memoria.

1.1. Motivación

Actualmente existen muchos profesionales que se dedican, entre otras cosas, a la realización de gráficos para mostrar diferentes datos de encuestas. Sin embargo, en la mayor parte de los casos, estos gráficos se realizan expresamente para una encuesta concreta y empleando Excel[1]. Esta forma de proceder no es para nada reutilizable, lo cual provoca una inversión de tiempo elevada para redactar cada informe.

En este trabajo se pretende desarrollar una solución o al menos, una primera versión de solución que permita realizar diversas visualizaciones para cualquier encuesta, simplemente conociendo el nombre de las variables que la componen.

1.2. Objetivos

Los objetivos a conseguir con este proyecto pueden resumirse en un objetivo principal, el cual a su vez puede ser desglosado en una serie de objetivos específicos, todos ellos necesarios para cumplir el principal.

- **Objetivo Principal:** Desarrollar una aplicación en R[3] que, recibiendo una serie de parámetros, realice diversos tipos de visualizaciones y las adjunte a un documento que puede ser de varios formatos para que el usuario pueda utilizarlo de base para crear informes. Esta aplicación también tendrá una interfaz de usuario de la que se habla en la memoria adjunta [2].

- **Objetivos Específicos:**

  1. Permitir al usuario crear gráficos de barras tanto verticales como horizontales y de sectores.
  2. Permitir al usuario crear tablas con hasta dos variables cruce.
  3. Exportar las visualizaciones a documentos en formato "Word", "Excel", "PowerPoint", "HTML" y "PNG"
  4. Leer archivos de datos de distintos formatos, como "CSV", "TXT", "TSV" y "XLSX".
  5. La aplicación debe ser modular, de manera que si el usuario desea añadir nuevas funcionalidades, le resulte sencillo.

1.3. Estructura de la memoria

Se describirán brevemente todos los apartados por los que estará formada la memoria.

1. **Introducción:** Se introducirá el trabajo, indicando los objetivos de los que consta el mismo.
2. **Conceptos previos:** Se hablará sobre la estructura que suelen presentar las encuestas, y sobre las librerías que se han utilizado para implementar la aplicación.
3. **Paquetes de R para el tratamiento de encuestas:** Mostrará algunos paquetes existentes actualmente destinados al tratamiento automático de encuestas.
4. **Fuentes de microdatos:** Se enumerarán diferentes fuentes de microdatos existentes tanto en España como en Europa.
5. **Aplicación principal:** En esta sección se explicará detalladamente como funciona la aplicación y los módulos que la forman.
6. **Pruebas de funcionamiento de la aplicación:** Se mostrará una prueba completa de la aplicación, adjuntando imágenes de los documentos que crea en diferentes formatos.
7. **Conclusiones y trabajo futuro:** Se exponen las conclusiones del proyecto, así como algunas posibles mejoras que pueden ser implementadas en el futuro.
8. **Bibliografía**
9. **Apéndice A. Manual de instalación de R:** A pesar de ser propiamente la parte más relacionada con el Trabajo Fin de Grado de Ingeniería Informática, se ha considerado interesante incluir también la interfaz para poder utilizar la aplicación de manera más sencilla. En esta sección se explica cómo instalar la herramienta R, para poder utilizar la aplicación.

10. **Apéndice B. Manual de uso de la interfaz:** Instrucciones para poder utilizar la aplicación.

11. **Apéndice C. Contenido del CD:** Contenido del CD

Capítulo 2

Conceptos previos

El objetivo de esta sección es hablar, de forma breve, del formato o estructura que van a seguir las encuestas a las que va a estar enfocada la aplicación. También se van a comentar las librerías que van a ser necesarias para dicha aplicación y su utilidad.

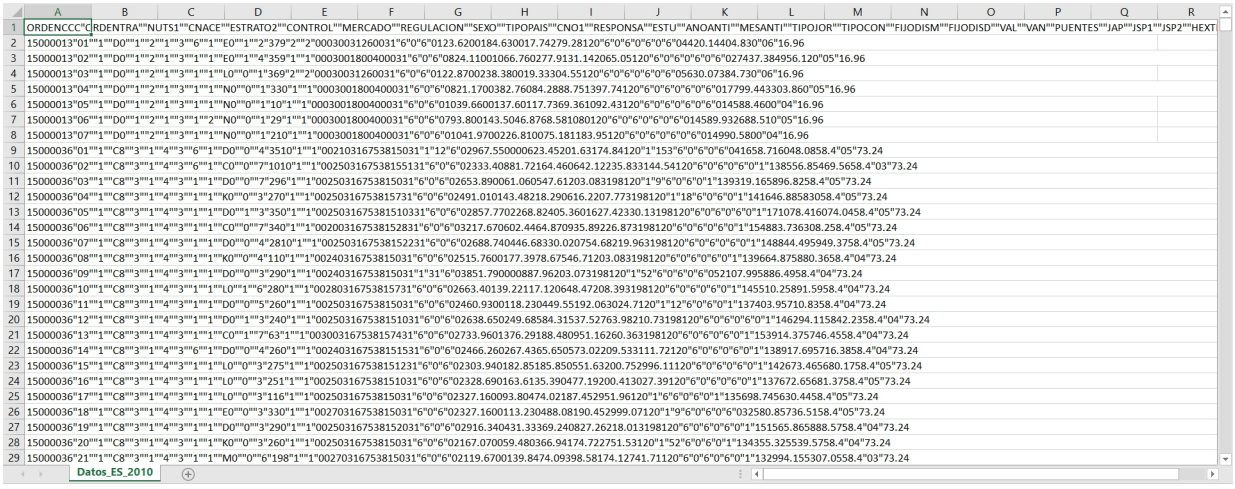
2.1. Estructura de las encuestas

Lo primero que se debe mencionar es que, esta aplicación, no está pensada para ser usada con conjuntos de datos de encuestas demasiado grandes pues la forma de lectura de esos mismos datos, así como la forma de trabajar con ellos, no ha sido optimizada para tales propósitos. No obstante, el ejemplo de encuesta que se va a utilizar en esta sección está formado por un total de 216.770 observaciones y 54 variables.

En esta sección se va a utilizar una encuesta sobre la estructura salarial, proveniente de la página web de Instituto Nacional de Estadística[4]. El archivo comprimido descargado de la página contiene un fichero de texto resumen, los propios microdatos en diferentes formatos y un fichero excel en el que se explica la estructura de cada variable, su longitud, una descripción y los códigos que la conforman así como el significado de cada uno. Para ilustrar la estructura, se utilizará el fichero con extensión csv y el de explicación.

En la figura 2.1 se muestra el fichero csv con los microdatos para esta encuesta. Como se puede ver, tiene una estructura muy sencilla donde la primera fila se trata de una cabecera que indicará el nombre de todas las variables, en este caso separados por tabuladores y el resto de filas son los datos en sí.

Además de los datos, en la figura 2.2 podemos ver el archivo que indica la estructura de esta encuesta. Está formado por tres hojas, de las cuales solo se muestra la primera, que es donde podemos encontrar la mayor parte de la información de los datos de esta encuesta. En las otras dos hojas, es donde pueden verse, para cada variable, los códigos y descripciones que la componen, como puede ser el ejemplo de la variable SEXO que, si buscamos en la tabla 1, como está indicado en la tabla principal, podemos ver que está compuesta por dos códigos, 1 y 6, que representan a hombres y mujeres respectivamente.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	ORDENCCC"CRDENTRA""NUTS1""CNACE""ESTRATO2""CONTROL""MERCADO""REGULACION""SEXO""TIPOPAIS""CNO1""RESPONSA""ESTU""ANOANTI""MESANTI""TIPOJOR""TIPOCON""FIJODISM""FIJODISD""VAL""VAN""PUENTES""JAP""JSP1""JSP2""HEXT																	
2	15000013"01""1""D0""1""2""1""3""6""1""E0""1""2"379"2""2"00030031260031"6"0"6"0123.6200184.630017.74279.28120"6"0"6"0"6"0"6"04420.14404.830"06"16.96																	
3	15000013"02""1""D0""1""2""1""3""1""1""E0""1""4"359"1""1"00030018004000031"6"0"6"0824.11001066.760277.9131.142065.05120"6"0"6"0"6"0"6"027437.384956.120"05"16.96																	
4	15000013"03""1""D0""1""2""1""3""1""1""L0""0""1"369"2""2"00030031260031"6"0"6"0122.8700238.380019.33304.55120"6"0"6"0"6"0"6"05630.07384.730"06"16.96																	
5	15000013"04""1""D0""1""2""1""3""1""1""N0""0""1"330"1""1"00030018004000031"6"0"6"0821.1700382.76084.2888.751397.74120"6"0"6"0"6"0"6"017799.443303.860"05"16.96																	
6	15000013"05""1""D0""1""2""1""3""1""1""N0""0""1"10"1""1"00030018004000031"6"0"6"01039.6600137.60117.7369.361092.43120"6"0"6"0"6"0"6"014588.4600"04"16.96																	
7	15000013"06""1""D0""1""2""1""3""1""1""2""N0""0""1"29"1""1"00030018004000031"6"0"6"0793.800143.5046.8768.581080120"6"0"6"0"6"0"6"014589.932688.510"05"16.96																	
8	15000013"07""1""D0""1""2""1""3""1""1""N0""0""1"210"1""1"00030018004000031"6"0"6"01041.9700226.810075.181183.95120"6"0"6"0"6"0"6"014990.5800"04"16.96																	
9	15000036"01""1""C8""3""1""4""3""6""1""D0""0""4"3510"1""1"00210316753815031"1"12"6"02967.550000623.45201.63174.84120"1"153"6"0"6"0"6"041658.716048.0858.4"05"73.24																	
10	15000036"02""1""C8""3""1""4""3""6""1""C0""0""7"1010"1""1"00250316753815513 1"6"0"6"02333.40881.72164.460642.12235.833144.54120"6"0"6"0"6"0"1"138556.85469.5658.4"03"73.24																	
11	15000036"03""1""C8""3""1""4""3""1""1""D0""0""7"296"1""1"00250316753815031"6"0"6"02653.890061.060547.61203.083198120"1"9"6"0"6"0"1"139319.165896.8258.4"05"73.24																	
12	15000036"04""1""C8""3""1""4""3""1""1""K0""0""3"270"1""1"00250316753815731"6"0"6"02491.010143.48218.290616.2207.773198120"1"18"6"0"6"0"1"141646.88583058.4"05"73.24																	
13	15000036"05""1""C8""3""1""4""3""1""1""D0""1""3"350"1""1"00250316753815103 31"6"0"6"02857.7702268.82405.3601627.42330.13198120"6"0"6"0"6"0"1"171078.416074.0458.4"05"73.24																	
14	15000036"06""1""C8""3""1""4""3""1""1""C0""0""7"340"1""1"00200316753815283 1"6"0"6"03217.670602.4464.870935.89226.873198120"6"0"6"0"6"0"1"154883.736308.258.4"05"73.24																	
15	15000036"07""1""C8""3""1""4""3""1""1""D0""0""4"2810"1""1"00250316753815223 1"6"0"6"02688.740446.68330.020754.68219.963198120"6"0"6"0"6"0"1"148844.495949.3758.4"05"73.24																	
16	15000036"08""1""C8""3""1""4""3""1""1""K0""0""4"110"1""1"00240316753815031"6"0"6"02515.7600177.3978.67546.71203.083198120"6"0"6"0"6"0"1"139664.875880.3658.4"04"73.24																	
17	15000036"09""1""C8""3""1""4""3""1""1""D0""0""3"290"1""1"00240316753815031"1"31"6"03851.790000887.96203.073198120"1"52"6"0"6"0"6"052107.995886.4958.4"04"73.24																	
18	15000036"10""1""C8""3""1""4""3""1""1""L0""1""6"280"1""1"00280316753815731"6"0"6"02663.40139.22117.120648.47208.393198120"6"0"6"0"6"0"1"145510.25891.5958.4"04"73.24																	
19	15000036"11""1""C8""3""1""4""3""1""1""D0""0""5"260"1""1"00250316753815031"6"0"6"02460.9300118.230449.55192.063024.7120"1"12"6"0"6"0"1"137403.95710.8358.4"04"73.24																	
20	15000036"12""1""C8""3""1""4""3""1""1""D0""1""3"240"1""1"00250316753815103 1"6"0"6"02638.650249.68584.31537.52763.98210.73198120"6"0"6"0"6"0"1"146294.115842.2358.4"04"73.24																	
21	15000036"13""1""C8""3""1""4""3""1""1""C0""1""7"63"1""1"00300316753815743 1"6"0"6"02733.9601376.29188.480951.16260.363198120"6"0"6"0"6"0"1"153914.375746.4558.4"04"73.24																	
22	15000036"14""1""C8""3""1""4""3""6""1""D0""0""4"260"1""1"00240316753815153 1"6"0"6"02466.260267.4365.650573.02209.533111.72120"6"0"6"0"6"0"1"138917.695716.3858.4"04"73.24																	
23	15000036"15""1""C8""3""1""4""3""1""1""L0""0""3"275"1""1"00250316753815123 1"6"0"6"02303.940182.85185.850551.63200.752996.11120"6"0"6"0"6"0"1"142673.465680.1758.4"05"73.24																	
24	15000036"16""1""C8""3""1""4""3""1""1""L0""0""3"251"1""1"00250316753815103 1"6"0"6"02328.690163.6135.390477.19200.413027.39120"6"0"6"0"6"0"1"137672.65681.3758.4"05"73.24																	
25	15000036"17""1""C8""3""1""4""3""1""1""L0""0""3"116"1""1"00250316753815031"6"0"6"02327.160093.80474.02187.452951.96120"1"6"6"0"6"0"1"135698.745630.4458.4"05"73.24																	
26	15000036"18""1""C8""3""1""4""3""1""1""E0""0""3"330"1""1"00270316753815031"6"0"6"02327.1600113.230488.08190.452999.07120"1"9"6"0"6"0"6"032580.85736.5158.4"05"73.24																	
27	15000036"19""1""C8""3""1""4""3""1""1""D0""0""3"290"1""1"00250316753815231 1"6"0"6"02916.340431.33369.240827.26218.013198120"6"0"6"0"6"0"1"151565.865888.5758.4"04"73.24																	
28	15000036"20""1""C8""3""1""4""3""1""1""K0""0""3"260"1""1"00250316753815031"6"0"6"02167.070059.480366.94174.722751.53120"1"52"6"0"6"0"1"134355.325539.5758.4"04"73.24																	
29	15000036"21""1""C8""3""1""4""3""1""1""M0""0""6"198"1""1"00270316753815031"6"0"6"02119.6700139.8474.09398.58174.12741.71120"6"0"6"0"6"0"1"132994.155307.0558.4"03"73.24																	

Datos_ES_2010

Figura 2.1: Ejemplo de microdatos

Diseño de registro de la Encuesta Cuatrienal de Estructura Salarial (EES_ 2010)

Variable	Diccionario de la variable	Longitud	Tipo	Decimales	Posición	Orden	Diccionario ubicado en la hoja…	Descripción	Observaciones
ORDENCCC		8	A		1	1		NÚMERO DE ORDEN DEL CENTRO DE COTIZACIÓN	
ORDENTRA		2	A		9	2		NÚMERO DE ORDEN DEL TRABAJADOR	
NUTS1	TNUTS	1	A		11	3	Tablas1	NUTS1	VALORES DE 1 A 7
CNACE	TCNAE	2	A		12	4	Tablas2	CÓDIGO ACTIVIDAD ECONOMICA	Ramas de actividad
ESTRATO2	TTrabaj	1	A		14	5	Tablas1	TAMAÑO DE LA UNIDAD	2; DE 50 a 199; ESTRATO2= 3; 200 y más;
CONTROL	TContro	1	A		15	6	Tablas1	PROPIEDAD O CONTROL	1= PUBLICO -2=PRIVADO
MERCADO	TMercad	1	A		16	7	Tablas1	MERCADO	3=UNIÓN EUROPEA 4=MUNDIAL
REGULACION	TRegula	1	A		17	8	Tablas1	FORMA DE REGULACIÓN DE LAS RELACIONES LABORALES	AMBITO INFERIOR (AUTONÓMICO,
SEXO	TSexo	1	A		18	9	Tablas1	SEXO	1 (HOMBRE) Ó 6 (MUJER)
TIPOPAIS	TPais	1	A		19	10	Tablas1	NACIONALIDAD	
CNO1	TCNO	2	A		20	11	Tablas2	CODIGO DE OCUPACION	GRUPO PRINCIPAL CNO-11
RESPONSA	T1Sino	1	A		22	12	Tablas1	RESPONSABILIDAD EN ORGANIZACIÓN Y/O SUPERVISION	1(SI) Ó 0 (NO)
ESTU	TTitula	1	A		23	13	Tablas2	CODIGO DE LA TITULACION	Estudios
ANOANTI		2	N		24	14		AÑOS DE ANTIGÜEDAD	0 a 99
MESANTI		2	N		26	15		MESES DE ANTIGÜEDAD	0 a 12
TIPOJOR	TJornad	1	A		28	16	Tablas1	TIPO DE JORNADA	1=TIEMPO COMPLETO 2=TIEMPO PARCIAL
TIPOCON	TContra	1	A		29	17	Tablas1	DURACION DEL CONTRATO	DETERMINADA
FIJODISM		2	N		30	18		MESES DEL PERÍODO DE TRABAJO DEL TRABAJADOR FIJO DISCONTINUO	0 a 11
FIJODISD		2	N		32	19		DÍAS DEL PERÍODO DE TRABAJO DEL TRABAJADOR FIJO DISCONTINUO	0 a 31
VAL		2	N		34	20		DIAS DE VACACIONES ANUALES LABORABLES	0 a 99

Figura 2.2: Estructura de microdatos

Aunque no todos los datos que nos podamos encontrar van a estar tan detallados, y habrá veces en los que simplemente se pueda obtener el archivo de microdatos, sin explicación alguna. En estos casos, la mejor manera de informarse es contactar con la web desde la que han sido descargados y preguntar. A pesar de esto, hay algo que será común a casi todas las encuestas, el **factor de elevación**, esta variable estará incluida en las encuestas, generalmente como último campo y tendrá un nombre similar en todos los casos, en la encuesta mostrada de ejemplo, su nombre es "FACTOTAL". Esta variable se ha calculado en base al diseño de la encuesta y permite hacer los cálculos de forma sencilla, sin tener que trabajar con dicho diseño, ya que es el factor que hay que aplicar a cada observación para obtener lo que se quiera. Es muy importante tener en cuenta esta variable a la hora de realizar las operaciones porque, en caso de no utilizarla, no se ajustarán realmente a los resultados de la encuesta con la que se esté trabajando.

2.2. Librerías de R

Para la realización de la aplicación han sido necesarias una serie de librerías propias de R, las cuales van a ser detalladas en esta sección.

2.2.1. Openxlsx

Se trata de una librería [5] que simplifica la creación de archivos Excel y provee herramientas para leer o editar hojas de trabajo. En esta aplicación ha sido usada para eso mismo, la creación de los informes finales en formato Excel con las visualizaciones que sean necesarias.

2.2.2. R2HTML

Librería [6] que incluye diversos métodos para crear y escribir en un archivo HTML. Se ha utilizado para lo mismo que la anterior, crear los informes finales con las visualizaciones indicadas.

2.2.3. PNG

Esta librería [7], al igual que las dos anteriores, proporciona métodos para leer, crear o editar imágenes de bits guardadas con formato PNG. En la aplicación se utilizará para crear las visualizaciones que después se introducirán en el informe final o también para crear imágenes de las tablas solicitadas si el usuario solamente desea imágenes y no un informe.

2.2.4. Officer

Esta última librería [8] proporciona una forma de crear y editar archivos Word o Power Point de manera sencilla. Como en los otros casos, se utilizará para crear los informes en estos dos formatos.

Capítulo 3

Paquetes de R para el tratamiento de encuestas

A lo largo de esta sección se pretende hacer un breve inventario de paquetes ya existentes cuya función principal sea el tratamiento de encuestas automático. Además de esto se realizará un resumen de cada uno de ellos que hable tanto de para qué sirve, como de las características que posee. Finalmente también se expondrá a qué estará enfocada la aplicación que se está diseñando en este proyecto.

3.1. Paquetes existentes para el tratamiento de encuestas

3.1.1. Survey

El primer paquete que se va a mencionar es *survey* [9], este paquete implementa varias facilidades para analizar los datos obtenidos de encuestas complejas. No está diseñado para tratar encuestas de gran tamaño, para esto último existe una variante que se mencionará en posteriores apartados.

Algunas de las características que posee son las típicas de obtención de medias, cuantiles, ratios, modelos de regresión o tablas de contingencia. También se puede emplear para realizar muestreos multifase con o sin reemplazamiento ó para realizar gráficos.

Algo interesante de este paquete es que soporta el procesamiento en paralelo para ordenadores con varios núcleos y análisis multivariante, como puede ser el análisis en componentes principales.

3.1.2. SQLSurvey

Se trata de un paquete [10] muy similar al mencionado en la sección anterior, pero en este caso está diseñado para tratar encuestas de gran tamaño. Este paquete necesita de la instalación de uno adicional, *MonetDB* [11], se necesita de este paquete para poder realizar las conexiones con bases de datos SQL.

Las funciones que implementa este paquete son prácticamente las mismas que las del paquete *Survey* por lo que no se van a volver a enumerar todas ellas.

3.1.3. EnQuireR

El último paquete del que se va a hablar es *EnQuireR* [12], se trata un paquete que se centra en el análisis de datos categóricos y que además contiene varias herramientas para automatizar el proceso de una encuesta. Está diseñado para que cualquier persona sea capaz de utilizarlo aunque no posea gran conocimiento estadístico.

Incluye varios métodos de análisis tanto univariante como multivariante entre los que se incluyen.

- Análisis de correspondencias múltiple.
- Clustering.
- Análisis semántico.

Además de estos también implementa funciones para la realización de gráficos de diferentes tipos así como tablas y variedad de tests.

Algo muy interesante de este paquete es que posee métodos para automatizar la escritura de artículos pdf con las salidas de los análisis obtenidas.

3.1.4. Aplicación creada

En el caso de la aplicación que se está creando, se centra principalmente en la realización de gráficos o tablas de diferentes tipos. A pesar de que en este momento no realice ningún tipo de análisis estadístico, como la aplicación es modular, solo sería necesario crear un nuevo módulo que implemente los análisis que se deseen. En ese sentido, la aplicación que se está desarrollado es más fácilmente mantenible y actualizable.

Capítulo 4

Fuentes de microdatos

Existen infinidad de páginas web en las que se pueden encontrar ficheros de datos de encuestas que se han realizado en el pasado. La descarga de estos datos suele ser gratuita y brinda un mundo de posibilidades a cualquier persona que quiera hacer estudios o trabajos sobre ellos.

Esta sección estará dedicada a realizar un pequeño análisis sobre dichas páginas web, la estructura que tienen, cómo se pueden descargar o incluso qué tipos de encuestas hay. Se va a dividir la sección en dos partes, una dedicada a páginas web Españolas y la otra a páginas web Europeas.

Además de analizar las fuentes de datos existentes, se pretende estudiar, en el caso en el que sea posible, los posibles tipos de preguntas incluidas en algunas encuestas y la forma de codificar las respuestas, también se va a analizar los distintos tipos de informes que se crean actualmente en estas páginas.

4.1. En España

4.1.1. Estadística de Castilla y Leon

La primera página de la que se va a hablar se trata de la web de la Junta de Castilla y León de estadística [13], en ella, simplemente debemos poner el ratón sobre el botón marcado en azul de la figura 4.1 y accederemos a los diferentes tipos de encuestas.

Los tipos de encuestas principales que ofrece esta web son los listados a continuación. Además de estos, dentro de cada grupo, existen diferentes subgrupos entre los que se podrá elegir para que se asemeje más al objetivo que se busque.

- Demográficas.
- Laborales.

Figura 4.1: Web de estadística de Castilla y Leon

- Sociales.
- Económicas.
- Otros.

El problema principal de esta página es que solo se incluyen las visualizaciones que se han obtenido de cada conjunto de datos, y no te permite descargar los ficheros de datos propiamente dichos. Sin embargo, estos informes y tablas, pueden ser utilizados como ejemplo para saber qué clase de análisis realizan las empresas hoy en día y, con esto en cuenta, modificar o adaptar la aplicación desarrollada para este proyecto.

4.1.2. Instituto Nacional de Estadística

La siguiente página de la que se va a hablar es la del Instituto Nacional de Estadística. Para acceder a los microdatos que ofrece esta página, se puede proceder de dos formas diferentes.

- A través de la propia pagina web, mostrada en la figura 4.2.
- A través de una conexión FTP al repositorio de encuestas del INE, el cual se muestra en la figura 4.3.

En el primer caso, como se ve en la figura, se puede elegir entre diferentes temas, los cuales aparecen en la parte izquierda de la ventana. Dentro de cada tema, se puede elegir entre diferentes sub-temas para ajustar más la búsqueda. Se va a mostrar, por ejemplo, dentro del apartado “Mercado laboral”, el sub-tema “Actividad, ocupación y paro”.

En la figura 4.4, aparece el resultado de acceder a ese sub-tema. Una vez estamos ahí, solo hay que buscar lo que queremos y hacer click en la columna marcada de amarillo en la figura. Por último, solo sería necesario acceder al apartado de “Resultados” de la parte izquierda de la ventana, y podremos descargar los microdatos de la fecha que deseemos y en el formato que más nos convenga.

Figura 4.2: Web del Instituto Nacional de Estadística

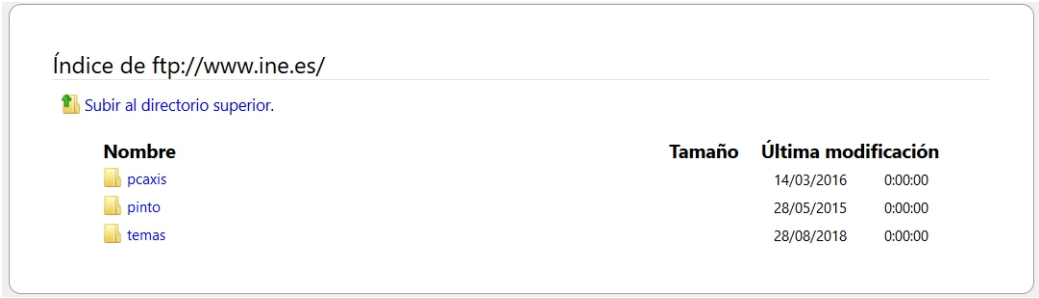


Figura 4.3: Repositorio del Instituto Nacional de Estadística

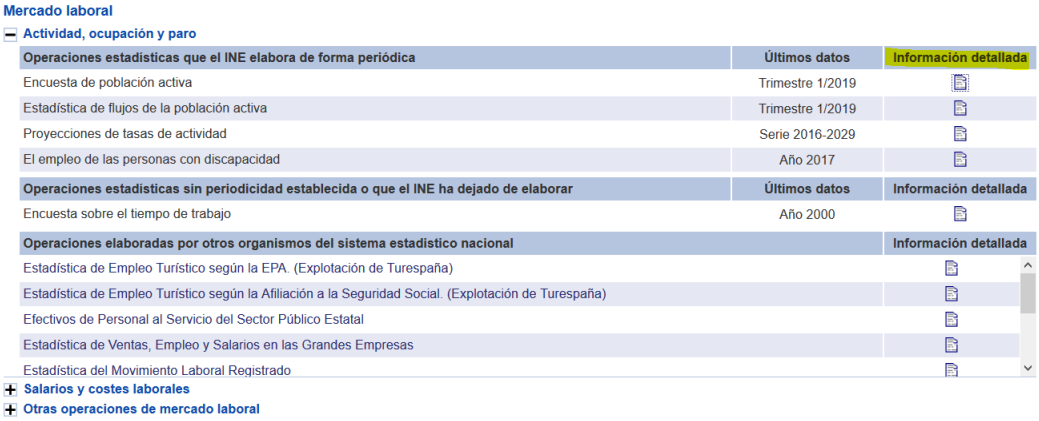


Figura 4.4: Web del Instituto Nacional de Estadística

Además de poder descargar los microdatos, la propia web dispone de las herramientas necesarias para obtener diferentes tablas y gráficos si el usuario así lo desea. Para ello, en vez de descargar los datos, hay que acceder al apartado y nos aparecerá la herramienta para indicar la estructura de las tablas y/o gráficos.

En cuanto a la segunda forma de acceder a los microdatos, la forma de descargarlos es más directa, a cambio de que es más complicado saber de qué son esos datos en la propia web. Para descargarlos, solamente hay que acceder al apartado "temas" mostrado en la figura y elegir el que queramos. Se descargará de esta forma un archivo comprimido, que contendrá los microdatos en diversos formatos, dependiendo del tema.

4.1.3. Consejo Superior de Investigación Científica

Otra de las páginas en las que se pueden encontrar algunos microdatos, aunque un poco anticuados, es el apartado de envejecimiento del Consejo Superior de Investigación Científica [14], mostrado en la figura 4.5

Figura 4.5: Web del Envejecimiento en red

La forma de acceder a los microdatos en esta web es sencilla, solamente deberemos pulsar en la pestaña llamada "Mapa de recursos" y nos aparecerá una nueva sección, justo al lado de esta última, denominada "Documentos e investigación". Tras acceder a ella, tendremos una serie de sub-secciones a las que podemos acceder, la que nos interesa es la de "Datos" . Una vez dentro de ésta, solo habrá que seleccionar la opción "Microdatos", tal y como se muestra en la figura 4.6. Ahí podremos seleccionar el tema que más nos convenga y descargar algunos datos en diferentes formatos.

Figura 4.6: Microdatos en la web del Envejecimiento en red

4.1.4. Datos por petición

Además de existir páginas que dan acceso de forma libre a sus microdatos, existen otra muchas que, para poder acceder a dichos datos, es necesario realizar una petición a los dueños. Por este motivo, simplemente van a ser mencionadas algunas, sin necesidad de entrar en detalles de cada una.

- Páginas web de Organismos Estadísticos por comunidades autónomas, algunas de ellas aparecen en la bibliografía de este mismo documento [15] [16] [17].

4.2. En Europa

La idea principal era hablar de fuentes de microdatos existentes en España, no obstante, existe una página no perteneciente a España de la que merece la pena hablar debido a la gran cantidad de datos que posee.

4.2.1. EUROSTAT, (Oficina Europea de Estadística)

Esta web [18], es una fuente increíble de datos de todo tipo, para acceder a ellos, existen diferentes formas, que se van a detallar a continuación.

1. A través de su sección de descarga masiva.
2. A través de su sección de datos estructurada.

3. A través de R, utilizando su librería propia.

Para las dos primeras, la forma de proceder es muy similar, solamente hay que poner el cursor sobre la pestaña de "Data", y seleccionar la opción "Database". De esta forma accederemos a los datos, estructurados en forma de árbol por el que podremos navegar hasta encontrar el tema que más nos convenga. La otra opción es, en esa misma ventana, pulsar sobre la opción "Bulk download" situada en la parte izquierda. La figura 4.7 muestra marcado en rojo la forma de acceder a los datos estructurados, y en azul la sección de descarga de datos masiva.

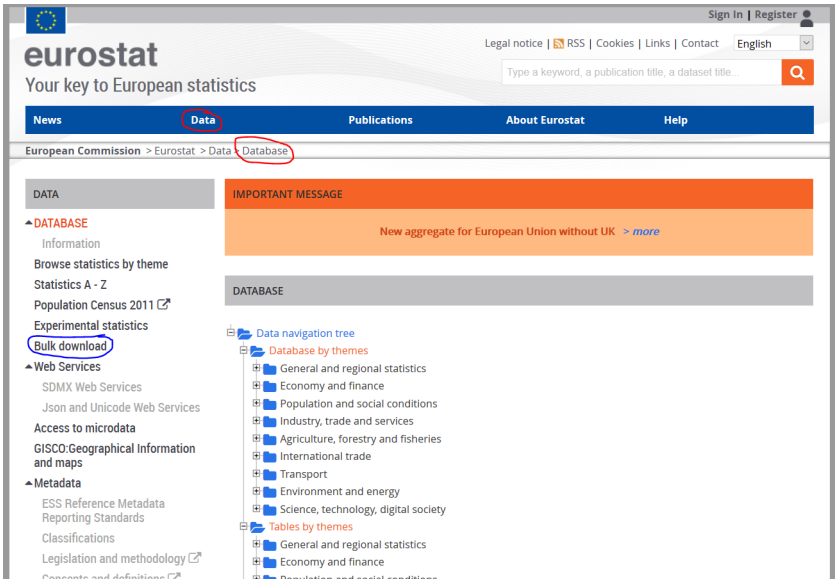


Figura 4.7: Microdatos en la web Eurostat

La última opción pasa por descargar los datos a través de la herramienta R, utilizando la librería propia "Eurostat" creada para ello. En las referencias de este documento se mostrará un enlace a un documento que indica la forma de hacerlo.

4.3. Otros

Por último, además de todas las posibles fuentes que se han indicado a lo largo de esta sección, se quiere mencionar una última, que no contiene microdatos, pero sí una lista de otras muchas páginas en las que sí que se puede acceder a ellos. Dicha web es la de la Universidad Autónoma de Barcelona [19]. En ella existe una pestaña denominada "Fuentes de datos" en la que podremos encontrar diferentes enlaces para acceder a diferentes páginas y descargar los microdatos.

Capítulo 5

Aplicación principal

5.1. Explicación del funcionamiento de la aplicación

A lo largo de toda esta sección se va explicar, de la forma más detallada posible, la estructura y el funcionamiento de la aplicación desarrollada, una vez esté todo explicado, se expondrá algún ejemplo que ilustre, de manera más práctica y visual, cómo actúa la aplicación en un determinado caso.

Lo primero será hablar un poco sobre el contexto. La intención que hay detrás de esta aplicación es tratar de automatizar el proceso que realizan algunas entidades de realizar una encuesta y, para esa encuesta exclusivamente, realizar una serie de gráficos y tablas que muestren los resultados, pues esto es una tarea bastante engorrosa y que se podría automatizar de varias formas.

En este caso, el lenguaje elegido es *R* y las visualizaciones a las que da soporte la aplicación son las siguientes

1. Gráficos de barras, tanto verticales como horizontales
2. Gráficos de sectores
3. Tablas de dos o tres variables

Todas estas visualizaciones están preparadas para representar diferentes tipos de medida, como puede ser la media, el total o incluso el porcentaje.

Además de esto, como el programa es modular, se puede ampliar con cualquier tipo de visualización añadiendo un nuevo módulo, o incluso modificar alguno ya existente para cambiar el estilo de un gráfico o tabla.

El funcionamiento global de la aplicación es bastante sencillo, debe recibir un archivo txt en el que estén indicados los siguientes campos.

- Las visualizaciones deseadas.

- Ruta donde se encuentra el archivo con los datos.
- Separador empleado en dicho archivo, puede ser cualquiera de los siguientes.
  - Punto y coma
  - Tabulador
  - Coma
- Formato en el que el usuario desea recibir las visualizaciones, puede ser cualquiera de los siguientes.
  - POWER POINT
  - WORD
  - EXCEL
  - HTML
  - IMÁGEN
- Ruta donde desea obtener el archivo con la salida.
- Nombre exacto del factor de elevación para la encuesta en cuestión.
- Nombre que se desea para el archivo de salida.
- Para cada visualización se deberán indicar.
  - Tipo de visualización, dentro de los indicados anteriormente.
  - Medida representada, también indicadas anteriormente.
  - Variable respuesta.
  - Variable de cruce uno.
  - Variable de cruce dos, esta última es opcional.

Además de esto, y de manera opcional, se introducirá un archivo llamado *Diccionario*, que pertenece únicamente a una encuesta concreta y contiene los siguientes campos

- Nombre de la variable: Se trata del nombre de la variable dentro del archivo de datos
- Código: Los diferentes códigos empleados en la encuesta
- Descripción: El significado de cada código

La utilidad de este "diccionario" es meramente ilustrativa, pues solo se utiliza para cambiar el nombre de las variables en las visualizaciones por su significado real dentro de la encuesta.

La aplicación recibe estos dos archivos que se crean por medio de una interfaz gráfica y, con ellos, realiza y adecúa las diferentes visualizaciones a las peticiones del usuario, depositando el documento solicitado en la ruta indicada. En la siguientes secciones se detallarán todos los módulos de los que dispone la aplicación en este momento, así como un ejemplo en el que se ilustrará todo.

5.2. Módulos de la aplicación

5.2.1. Módulo de lectura de datos

Como se puede suponer, el primer módulo creado para la aplicación se trata del que lee los datos, este módulo puede verse en el siguiente pseudocódigo.

```
Funcion lecturaDatos (ruta, separador=;)
        a1 ←obtener el formato a base de la ruta
        formato ←toupper(A1)
        Si (formato = CSV) entonces
                datos<- leer csv
                devolver (datos)
        Fin si
        Si (formato = TXT) entonces
                datos ←leer txt
                devolver (datos)
        Fin si
        Si (formato = TSV) entonces
                datos ←leer tsv
                devolver (datos)
        Fin si
        Si (formato = XLSX) entonces
                datos ←leer xlsx
                devolver (datos)
        Fin si
Fin función
```

Este módulo recibe como parámetros de entrada.

- Ruta: ruta en la que se encuentra el archivo de datos
- Separador: separador empleado en el archivo de datos, por defecto es la coma
- Formato: El formato en el que está el archivo de datos

En el caso de que el formato no sea proporcionado, el propio módulo detecta cuál es cortando la ruta por el punto y quedándose con lo que hay después del mismo.

Después, dependiendo del formato en el que esté, se lee de una forma u otra, ahora mismo los formatos soportados son los que se ven en el pseudocódigo mostrado.

5.2.2. Módulo de lectura del diccionario

El siguiente módulo se encarga de leer el archivo del diccionario, en el que caso de que exista. Tiene una estructura muy similar a la del módulo de lectura de datos, pero más simplificado, puede verse en el siguiente pseudocódigo.

```
Funcion lecturaDiccionario (rutad)
        diccionario ←leer diccionario
        devolver (diccionario)
Fin función
```

5.2.3. Módulo de creación de tablas

Una vez se han leído todos los datos, se comienza con la creación de las visualizaciones, el primer módulo para este propósito, es el encargado de crear las tablas solicitadas, éste módulo puede verse en el pseudocódigo mostrado a continuación.

```
Función crea_tablas (vars_cruce, var_calculo, factor, tipo, diccionario)
    Si (tipo = MEDIA) entonces
        tabla_beta ←calcula la tabla con las medias
        tabla ←cambia_nombres(tabla_beta, diccionario, vars_cruce)
    Fin si
    Si (tipo = TOTAL) entonces
        tabla_beta ←calcula la tabla con el total
        tabla ←cambia_nombres(tabla_beta, diccionario, vars_cruce)
    Fin si
    Si (tipo = PORCENTAJE) entonces
    tabla_beta ←calcula la tabla
        Si (longitud(vars_cruce)>1)entonces
            tabla2 ←calcula porcentajes por columnas (tabla_beta)
        Si no entonces
            tabla2 ←calcula porcentajes (tabla_beta)
        Fin si
        tabla ←cambia_nombres(tabla2, diccionario, vars_cruce)
    Fin si
Fin función
```

Este módulo recibe como parámetros de entrada.

- Vars_cruce: Array con todas las variables de cruce implicadas en la creación de la tabla
- Var_calculo: La variable respuesta que se quiere representar con la tabla
- Factor: El factor de elevación asociado a la encuesta

- Tipo: El tipo de medida que se quiere representar
- Diccionario: Contiene el diccionario creado con anterioridad

El módulo se divide en dos partes bastante diferenciadas, en la primera parte, en función del tipo de medida a representar, se emplea la función de *R* llamada "tapply" para crear la tabla con el cruce de las variables, la forma de calcular esta tabla cambia dependiendo del tipo de medida a representar, después se llama a la función "cambia_nombres" (de la cual se hablará más adelante), esta función modifica el nombre de las columnas y filas de la tabla empleando el diccionario creado.

5.2.4. Módulo de creación de gráficos de barras

El siguiente módulo de creación de visualizaciones se encarga de crear los gráficos de barras en el caso de que sean solicitados, el funcionamiento de este módulo se puede ver en el siguiente pseudocódigo.

Este módulo recibe como parámetros de entrada.

- Vars_cruce: Array con todas las variables de cruce implicadas en la creación del gráfico
- Var_calculo: La variable respuesta que se quiere representar con el gráfico
- Factor: El factor de elevación asociado a la encuesta
- Tipo: El tipo de medida que se quiere representar
- Diccionario: Contiene el diccionario creado con anterioridad
- H: indica si el gráfico de barras es horizontal o vertical
- ÍndiceG: indica cuántos gráficos hay ya creados, con objetivo de asignar bien el nombre del archivo de destino.

En este caso, la forma de actuar también dependerá fundamentalmente del tipo de medida a representar, sin embargo, la forma de proceder será prácticamente idéntica, solo cambiando la manera de crear la tabla con los datos que se representarán en el gráfico.

Lo primero que se hace es obtener el nombre que tendrá la imagen creada, concatenando el coletilla "auxG" con el índice correspondientes, después hay que crear la tabla con los datos y modificar el nombre de las filas y columnas tal y como se ha hecho en el módulo explicado anteriormente, una vez hecho esto se crea el gráfico de barras con esta tabla y, dependiendo de la cantidad de variables de cruce que existan, se crea la leyenda asociada al gráfico y, además se indica, en cada barra, el número exacto que se está representando.

```
Función crea_barras (vars_cruce, var_calculo, factor, tipo, h, diccionario, indiceG)
        nombreA ← crear el nombre que tendrá la imagen del grafico
        Si (tipo = MEDIA) entonces
                tabla_beta ← calcula la tabla con las medias
                tabla ← cambia_nombres(tabla, diccionario, vars_cruce)
                n ← ncol (tabla)
                png (nombreA)
                a1 ← crea el gráfico de barras
                Si (longitud(vars_cruce)>1) entonces
                        crear la leyenda
                Fin si
                añadir texto a las barras
                cierra el creador de imágenes
        Fin si
        Si (tipo = TOTAL) entonces
                tabla_beta ← calcula la tabla con los totales
                tabla ← cambia_nombres(tabla, diccionario, vars_cruce)
                n ← ncol (tabla)
                png (nombreA)
                a1 ← crea el gráfico de barras
                Si (longitud(vars_cruce)>1) entonces
                        crear la leyenda
                Fin si
                añadir texto a las barras
                cierra el creador de imágenes
        Fin si
        Si (tipo = PORCENTAJE) entonces
                tabla_beta ← calcula la tabla con los totales
                tabla ← cambia_nombres(tabla, diccionario, vars_cruce)
                Si (longitud(vars_cruce)>1)entonces
                        tabla2 ← calcula porcentajes por columnas (tabla_beta)
                Si no entonces
                        tabla2 ← calcula porcentajes (tabla_beta)

                n ← ncol (tabla)
                png (nombreA)
                a1 ← crea el gráfico de barras
                Si (longitud(vars_cruce)>1) entonces
                        crear la leyenda
                Fin si
                añadir texto a las barras
                cierra el creador de imágenes
        Fin si
Fin función
```

5.2.5. Módulo de creación de gráficos de sectores

Otro módulo de creación de visualizaciones se encarga de crear los gráficos de sectores si el usuario lo solicita, este módulo se puede ver en el pseudocódigo de la siguiente página.

```
Función crea_pie (vars_cruce, var_calculo, factor, tipo, h, diccionario, indiceG)
        nombreA ← crear el nombre que tendrá la imagen del grafico
        Si (tipo = MEDIA) entonces
                tabla_beta ← calcula la tabla con las medias
                tabla ← cambia_nombres(tabla, diccionario, vars_cruce)
                n ← ncol (tabla)
                png (nombreA)
                a1 ← crea el gráfico de sectores
                crear la leyenda
                cierra el creador de imágenes
        Fin si
        Si (tipo = PORCENTAJE) entonces
                tabla_beta ← calcula la tabla con los totales
                tabla ← cambia_nombres(tabla, diccionario, vars_cruce)
                tabla2 ← calcula porcentajes (tabla_beta)
                n ← ncol (tabla)
                png (nombreA)
                a1 ← crea el gráfico de sectores
                crear la leyenda
                cierra el creador de imágenes
        Fin si
        Si (tipo = TOTAL) entonces
                tabla_beta ← calcula la tabla con los totales
                tabla ← cambia_nombres(tabla, diccionario, vars_cruce)
                n ← ncol (tabla)
                png (nombreA)
                a1 ← crea el gráfico de sectores
                crear la leyenda
                cierra el creador de imágenes
        Fin si
Fin función
```

Este módulo recibe como parámetros de entrada.

- Vars_cruce: Array con todas las variables de cruce implicadas en la creación del gráfico
- Var_calculo: La variable respuesta que se quiere representar con el gráfico
- Factor: El factor de elevación asociado a la encuesta
- Tipo: El tipo de medida que se quiere representar
- Diccionario: Contiene el diccionario creado con anterioridad

- ÍndiceG: indica cuántos gráficos hay ya creados, con objetivo de asignar bien el nombre del archivo de destino.

Una vez más, la manera de proceder dependerá del tipo de medida que se quiere representar pero, al igual que en el caso anterior, los cálculos a realizar serán casi idénticos, solo cambiando la forma de calcular los elementos que formarán parte de la tabla con la que se representará el gráfico.

En primer lugar, se obtendrá el nombre que tendrá la imagen creada, concatenando el coletilla "auxG" con el índice que corresponda, después se creará la tabla con los correspondientes datos y se modificarán los nombres de columnas y filas con la función *Cambia_nombres*, la cual se mostrará a continuación. Después se creará el gráfico de sectores junto con su leyenda y se exportará al formato deseado.

5.2.6. Módulo de modificación de nombres

Este módulo es una función auxiliar que sirve para modificar, en función del diccionario proporcionado, los nombres que reciben las columnas y las filas de una determinada tabla que es recibida como parámetro. Dicha función puede verse a continuación en el siguiente pseudocódigo.

```
Función cambia_nombres (tabla, diccionario, vars_cruce)
    Si (longitud(vars_cruce)>1) entonces
        row ←registros del diccionario que coincidan con las filas de la tabla
        actualizar los nombres de las filas de la tabla

        col ←registros del diccionario que coincidan con las columnas de la tabla
        actualizar los nombres de las columnas de la tabla
        devuelve (tabla)
    Si no entonces
        row ←registros del diccionario que coincidan con las filas de la tabla
        actualizar los nombres de las filas de la tabla
        devuelve (tabla)
    Fin si
Fin función
```

Este módulo recibe como parámetros de entrada.

- Tabla: Tabla a la que se desea modificar los nombres.
- Diccionario: Diccionario asociado a la encuesta.
- Vars_cruce: Array que contiene las variables que son representadas en la tabla.

En este caso la forma de proceder esta ligada a la cantidad de variables que existen en el cruce, ya que si solo existe 1 (además de la variable respuesta), solo habría que modificar los nombres de las filas de la tabla y, si existen más, habría que modificar tanto el nombre de las filas como el de las columnas.

5.2.7. Módulo de lectura de parámetros

La función de este módulo es leer el archivo de parámetros creado por la interfaz gráfica, es un módulo muy sencillo, como puede verse en el siguiente pseudocódigo.

```
Funcion lectura_parametros ()
        parametros ← leer archivo de parametros
        Devolver (parametros)
Fin función
```

El módulo simplemente lee el archivo de parámetros, ubicado siempre en el mismo directorio, y lo devuelve.

5.2.8. Módulo de creación de informes

Este módulo es el encargado de crear todos los tipos de informes, dependiendo del formato de salida indicado por el usuario, actuará de una forma u otra. Debido a la complejidad del mismo, se ha decidido mostrar el código completo en las figuras 5.1, 5.2 y 5.3.

```
creacionInforme<-function(formatoS,nombreD,rutaD,indiceG,indiceT,tablas){
  if(formatoS=="PowerPoint"){
    doc<-read_pptx()
    if(indiceG>0){
      for(i in 1:indiceG-1){

        nombre<-paste("auxG",i,sep="_")
        nombreF<-paste(nombre,"png",sep=".")
        ruta<-paste("./resultados",nombreF,sep="/")
        add_slide(doc,layout = "Title and Content", master="Office Theme")
        doc <- on_slide(doc, index = i+1)
        img.file <- file.path(ruta )
        ph_with(x = doc, external_img(img.file, 1, 1),location = ph_location_type(type = "body"))
      }
    }
    if(indiceT>0){
      for(i in 1:indiceT-1){
        add_slide(doc,layout = "Title and Content", master="Office Theme")
        doc <- on_slide(doc, index = i+1)
        doc<-ph_with_table(doc,as.data.frame(tablas[i]),first_column=TRUE,last_row=TRUE,
                           last_column=TRUE,location = ph_location_type(type = "body"))
      }
    }
    nombreSalida<-paste(nombreD,"pptx",sep=".")
    final<-paste(rutaD,nombreSalida,sep="/")
    print(doc,final)
  }
```

Figura 5.1: Módulo de creación de informes 1

```
if(formatoS == "Word"){
  doc<-read_docx()
  if(indiceG>0){
    for(i in 1:indiceG-1){
      nombre<-paste("auxG",i,sep="_")
      nombreF<-paste(nombre,"png",sep=".")
      ruta<-paste("./resultados",nombreF,sep="/")
      doc<-body_add_img(doc,ruta,width=5,height=5)
    }
  }
  if(indiceT>0){
    for(i in 1:indiceT-1){
      doc<-body_add_table(doc,as.data.frame(tablas[i]),first_column=TRUE)
    }
  }
  nombreSalida<-paste(nombreD,"docx",sep=".")
  final<-paste(rutaD,nombreSalida,sep="/")
  print(doc,final)
}
if(formatoS=="Excel"){
  wb <- createWorkbook()
  if(indiceG>0){
    for(i in 1:indiceG-1){
      nombre<-paste("auxG",i,sep="_")
      nombreF<-paste(nombre,"png",sep=".")
      ruta<-paste("./resultados",nombreF,sep="/")
      addWorksheet(wb, nombre)
      insertImage(wb, nombre, ruta,width = 6, height = 3, startRow = 1,
                  startCol = 1, units = "in", dpi = 300)}
  }
  if(indiceT>0){
    for(i in 1:indiceT-1){
      name<-paste("Tabla",i,sep="_")
      addWorksheet(wb, name)
      writeData(wb, name, tablas[i])}}
  nombreSalida<-paste(nombreD,"xlsx",sep=".")
  final<-paste(rutaD,nombreSalida,sep="/")
  saveWorkbook(wb, file = final, overwrite = TRUE)
}
```

Figura 5.2: Módulo de creación de informes 2

```
  if(formatoS=="Imagen"){
    if(indiceT>0){
      for(i in 1:indiceT-1){
        nombreSalida<-paste(nombreD,"png",sep=".")
        final<-paste(rutaD,nombreSalida,sep="/")
        png(final, height = 50*nrow(tablas[i]), width = 200*ncol(tablas[i]))
        grid.table(tablas[i])
        dev.off()
      }
    }
  }

  if(formatoS=="HTML"){
    rutaF<-paste(rutaD,"resultadosHTML",sep="/")
    dir.create(rutaF,showWarnings = FALSE)
    HTMLStart(outdir=rutaF,filename=nombreD, extension="html")
    if(indiceG>0){
      for(i in 1:indiceG-1){
        nombre<-paste("auxG",0,sep="_")
        nombreF<-paste(nombre,"png",sep=".")
        ruta<-paste("./resultados",nombreF,sep="/")
        img<-readPNG(ruta)
        plot(1:2, type='n',ann=FALSE,axes = FALSE)
        rasterImage(img, 0.95, 0.95, 2.05, 2.05, interpolate=FALSE)
        HTMLplot()
      }
    }

    if(indiceT>0){
      for(i in 1:indiceT-1){
        HTML(tablas[i])
      }
    }
    HTMLStop()
  }
}
```

Figura 5.3: Módulo de creación de informes 3

Este módulo recibe como parámetros de entrada.

- FormatoS: Formato en el que el usuario desea obtener el informe.
- NombreD: Nombre que el usuario desea dar al informe.
- RutaD: Ubicación en la que el usuario desea guardar el informe.
- ÍndiceG: Cantidad de gráficos creados a añadir en el informe.
- ÍndiceT: Cantidad de tablas creadas a añadir en el informe.
- Tablas: Lista de tablas creadas.

La forma de proceder para cada tipo de formato es bastante similar, generalmente se crea el documento en memoria y, una vez creado, se realizan dos bucles en los que se van añadiendo primero los gráficos y después las tablas. Para añadir cada gráfico o tabla se obtiene el nombre de dicha visualización mediante una concatenación de la coletilla "auxG" precedida de la ruta, la cual nunca cambia, y seguida del formato. Una vez añadido todo, se indica el nombre del archivo de destino, y se crea finalmente el documento.

En el caso del formato Power Point, se debe añadir una transparencia nueva antes de insertar cada gráfico o tabla y, además, indicar el estilo de dicha transparencia. Esto también pasa en el formato Excel.

Cuando el formato de salida se trate de imagen, como los gráficos ya están en ese formato, solo se deberán crear imágenes de las tablas.

Por último, si el formato es HTML, a la hora de insertar los gráficos, primero hay que leer la imagen ya creada, convertirla a plot de R y después insertarla.

5.2.9. Módulo de lanzamiento de la aplicación

Se trata del módulo principal de la aplicación, el cual ejecutará el resto. Para comenzar, obtiene todos los parámetros necesarios así como el diccionario de los ficheros de texto generados por la interfaz.

Después de esto comienza un bucle para crear todas las visualizaciones y las tablas solicitadas. Tras crearlas, se lanza una llamada al módulo de creación de informes y se finaliza la ejecución.

5.3. Ejemplo de funcionamiento

En esta sección se van a exponer, mediante imágenes, algunos ejemplos de gráficos que crea la aplicación y, en la próxima sección se mostrará un ejemplo completo de funcionamiento con todos los posibles formatos de informes que puede generar la aplicación.

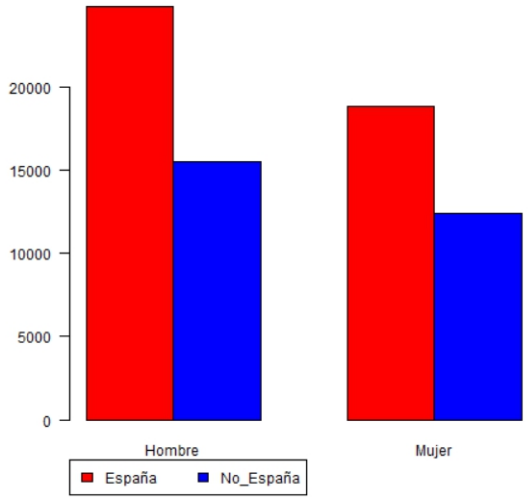


Figura 5.4: Gráfico de barras verticales

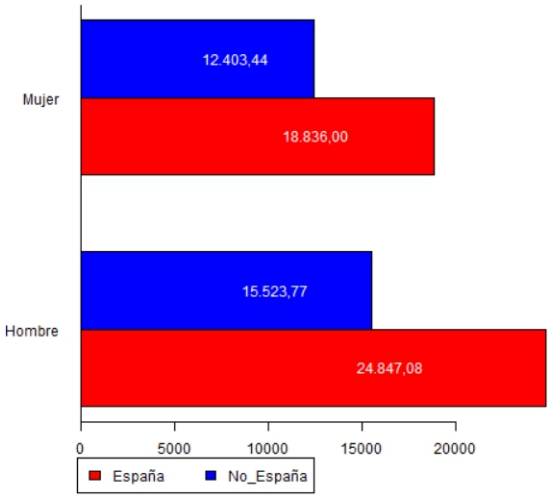


Figura 5.5: Gráfico de barras horizontales

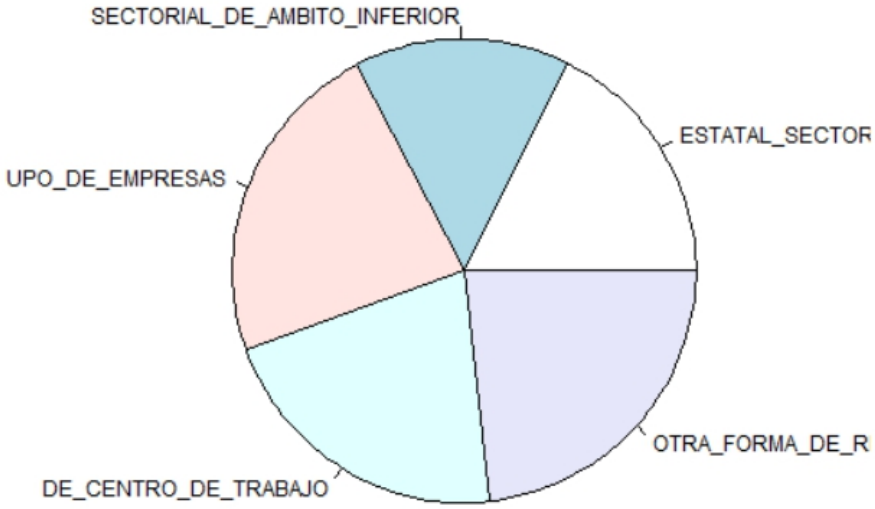


Figura 5.6: Gráfico de sectores

	España	No_España
Hombre	61.55	38.45
Mujer	60.30	39.70

Figura 5.7: Tabla ejemplo

Capítulo 6

Pruebas de funcionamiento de la aplicación

A continuación se van a realizar una serie de pruebas de funcionamiento de la aplicación desarrollada. En ellas, se va a emplear el mismo archivo de texto, para probar los diferentes formatos de informes que soporta la aplicación.

El fichero de texto que se va a utilizar para todas las pruebas se muestra en la figura 6.1

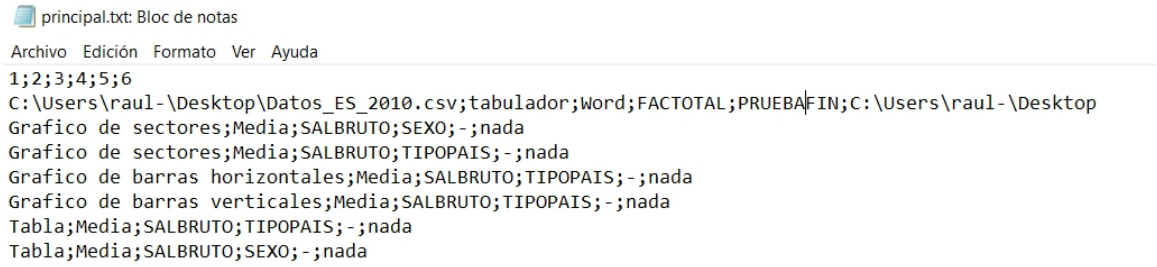


Figura 6.1: Fichero de texto

Se puede ver que, en la primera linea del archivo, se encontrarán los siguientes elementos, separados por punto y coma.

- Ruta del archivo de datos.
- Separador.
- Formato de destino.
- Nombre del factor de elevación que emplea la encuesta con la que se está trabajando.
- Nombre que se le dará al archivo de destino.

- Ruta de destino.

En el resto del fichero, estarán todas las visualizaciones solicitadas, con sus características, también separadas por punto y coma. En este caso se van a crear las siguientes.

- Gráfico de sectores con la variable sexo.
- Gráfico de sectores con la variable tipo de país.
- Gráfico de barras horizontales con la variable tipo de país.
- Gráfico de barras verticales con la variable tipo de país
- Tabla con la variable tipo de país.
- Tabla con la variable sexo.

En cuanto al diccionario, puede verse, en la figura 6.2, que sigue una estructura muy similar al ya expuesto.

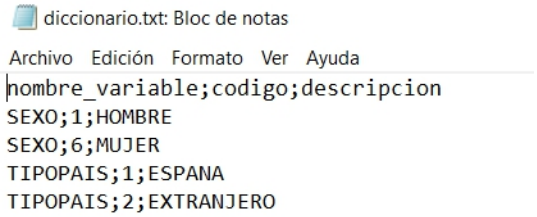


Figura 6.2: Estructura del diccionario

En cada linea de este fichero, se encontrarán, en el orden que aparece en la cabecera, el nombre de la variable, el código, y el significado de ese código.

Cabe destacar que, estos dos archivos, se crearán automáticamente a través de una interfaz gráfica también diseñada e implementada para este proyecto y de la cual se habla en la memoria adjunta.

Una vez se tienen estos dos ficheros, solamente quedaría ejecutar el programa en R y él ya se encargará de crear las visualizaciones solicitadas de manera automática, utilizando el diccionario para renombrar los ejes de los gráficos y los nombres de filas y columnas de las tablas. En el caso de que no exista diccionario, se mantendrá el código leído del archivo de datos.

Se irán mostrando los diferentes resultados para cada formato de salida en las figuras 6.3, 6.4, 6.5, 6.6 y 6.7.

Puede verse como cada visualización creada, independientemente del formato de salida, se coloca en una hoja distinta. Una vez ahí, el usuario puede redimensionarla a su gusto o añadir los comentarios que considere oportunos.

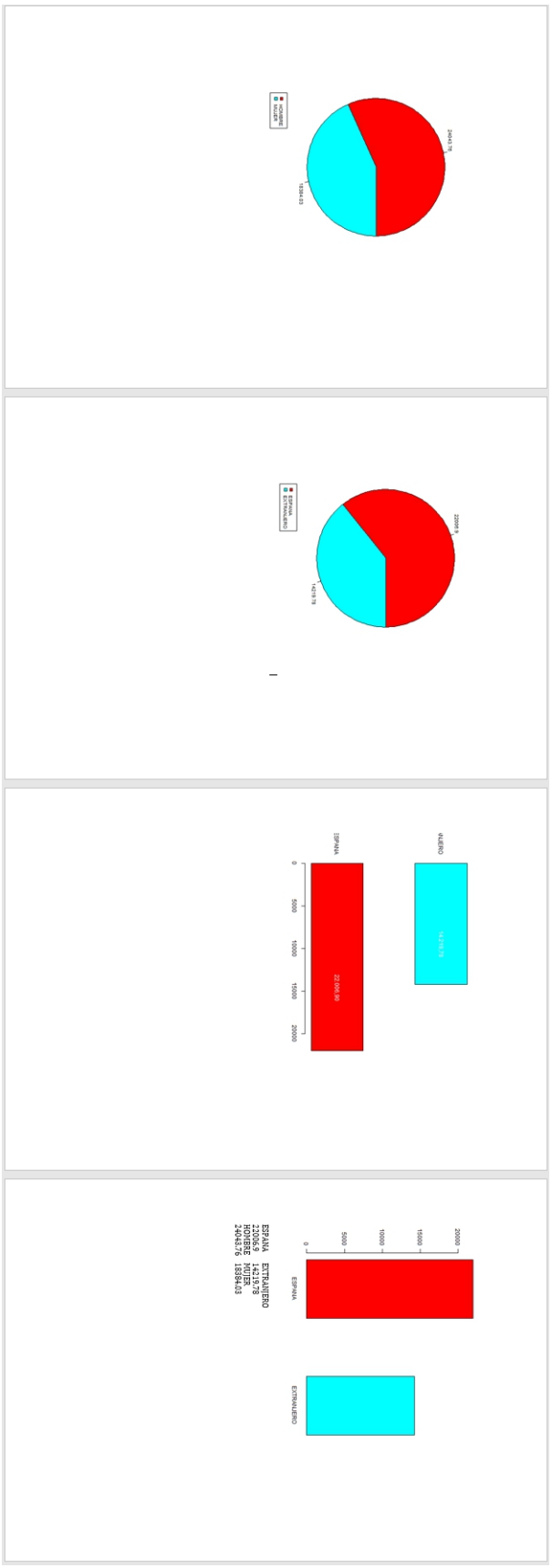


Figura 6.3: Informe en formato Word

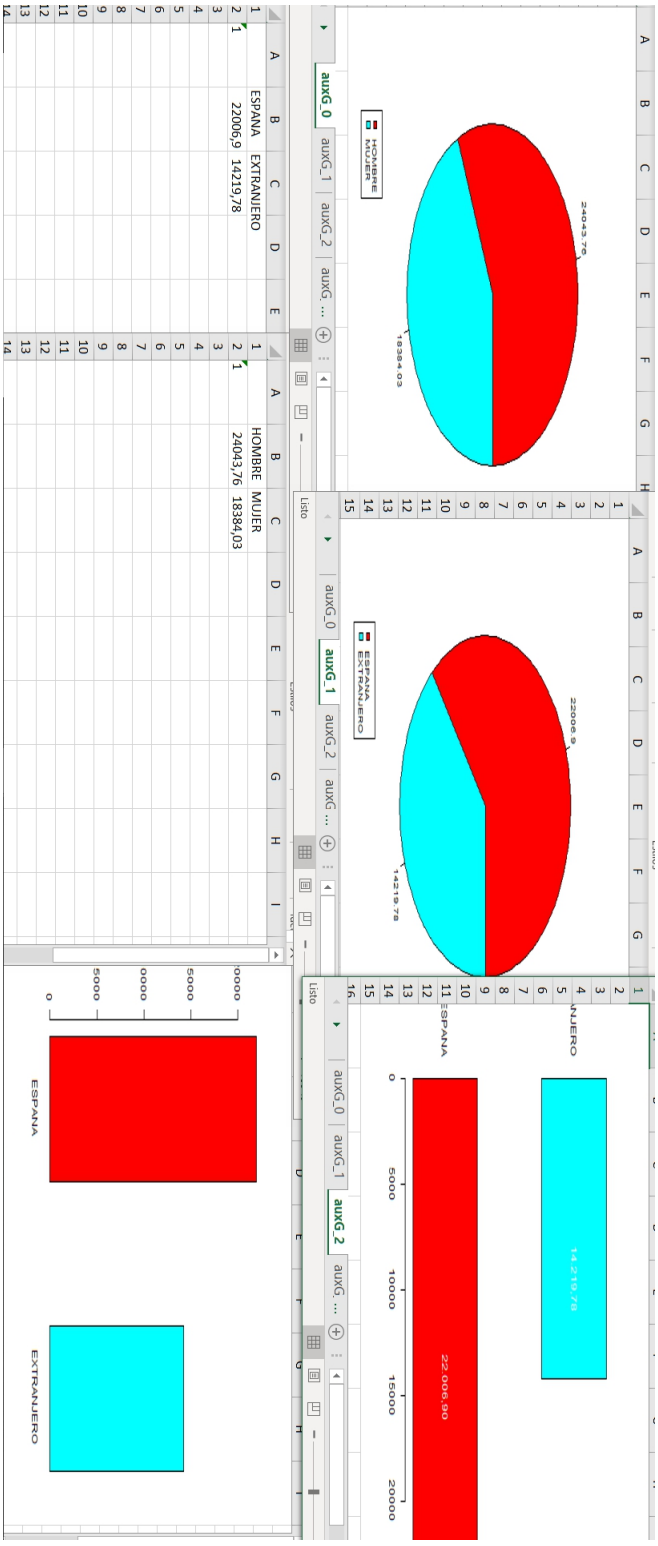


Figura 6.4: Informe en formato Excel

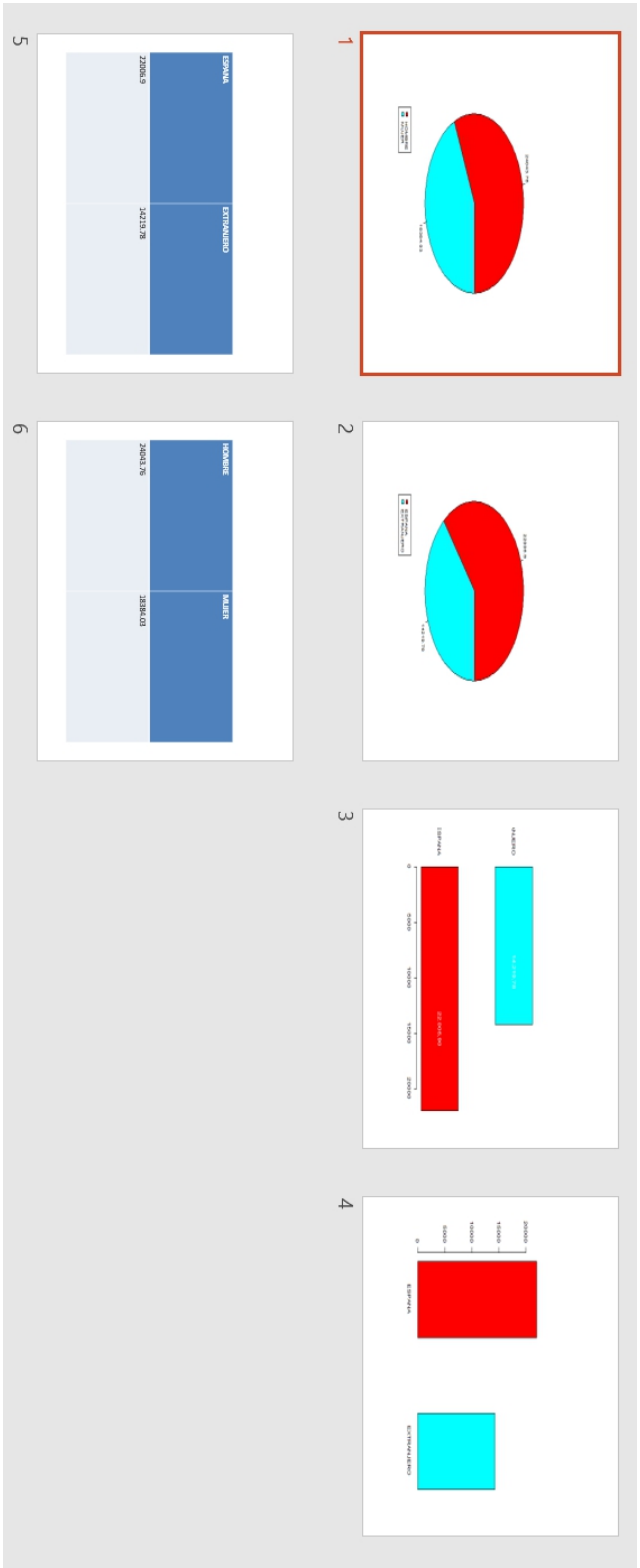


Figura 6.5: Informe en formato Power Point

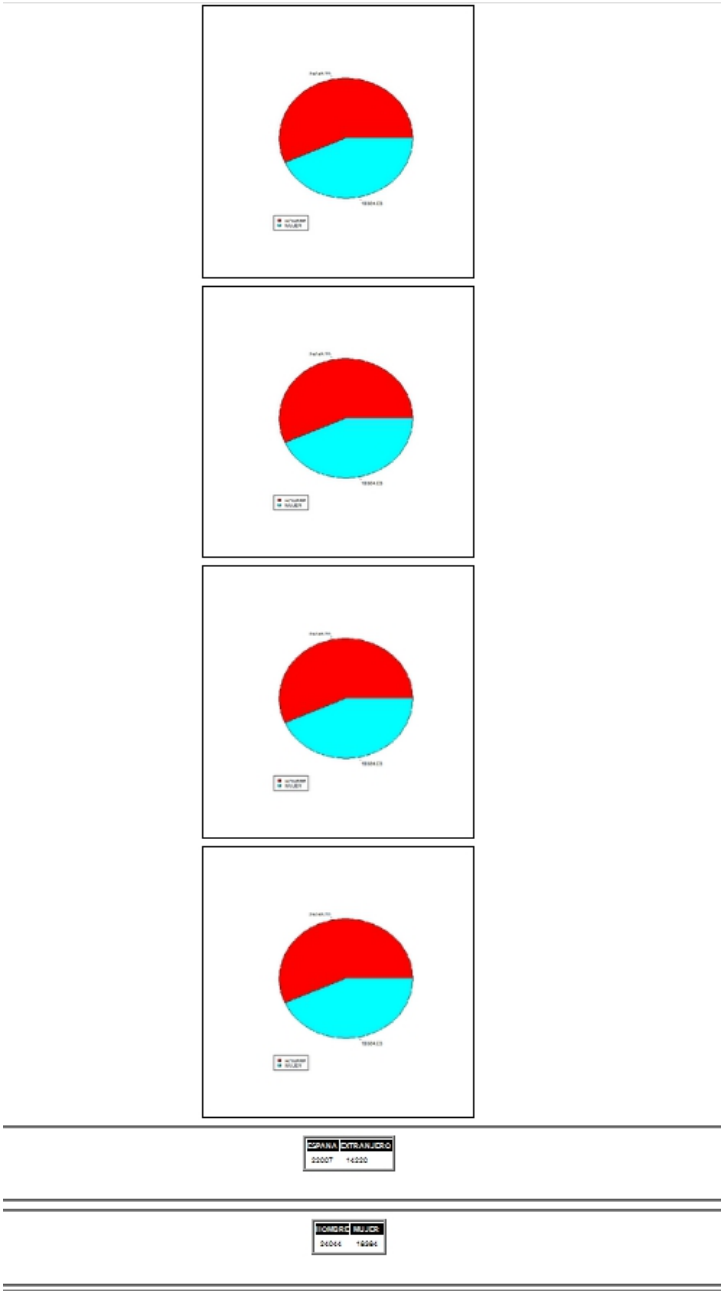


Figura 6.6: Informe en formato HTML

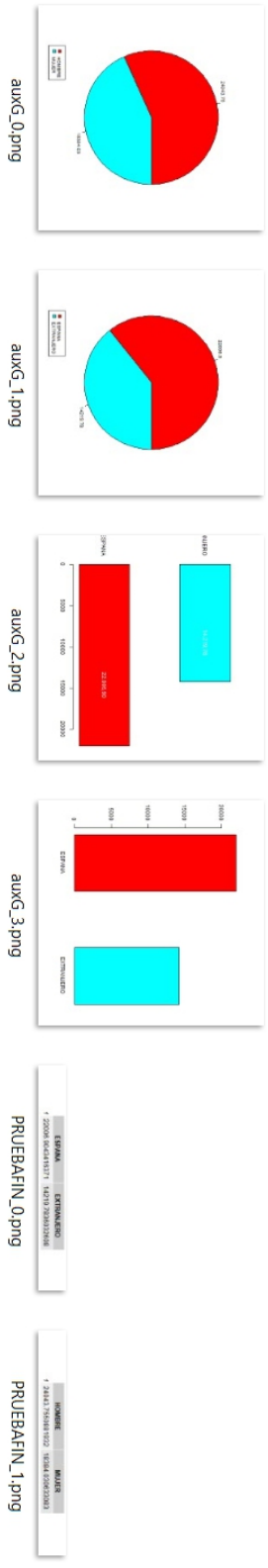


Figura 6.7: Informe en formato Imagen

Capítulo 7

Conclusiones y trabajo futuro

7.1. Conclusiones

Al finalizar este trabajo fin de grado, se puede concluir que se han alcanzado los objetivos planteados en su inicio. Se ha desarrollado una aplicación que genera informes partiendo de datos de encuestas.

Dicha aplicación, recibiendo una serie de parámetros que el usuario habrá indicado previamente a través de una interfaz, podrá crear todos los tipos de visualizaciones planteados inicialmente y, además de ello, podrá generar también documentos en los que adjuntará todas las visualizaciones creadas. Esos documentos podrán ser de cualquiera de los formatos también planteados al comienzo de la creación de la aplicación.

Toda la implementación de la aplicación ha sido desarrollado de forma modular, con la intención de que el usuario pueda modificar uno de ellos sin afectar al resto, o incluso que pueda añadir nuevos módulos sin que la funcionalidad hasta ese momento desarrollada, se vea afectada.

Con este proyecto de final de grado, se han sentado las bases de una aplicación que puede ir mejorando hasta convertirse en una herramienta muy potente que pueda utilizar cualquier profesional en su trabajo.

7.2. Líneas de trabajo futuro

Se ha desarrollado toda la aplicación de manera que fuera lo más modular posible, ya que de esta manera, es muy sencillo implementar nueva funcionalidad sin modificar la ya existente. A continuación se enumeran algunas ideas que podrían llevarse a cabo en el futuro.

- Optimizar las funciones ya desarrolladas para que sean más eficientes.
- Añadir módulos que permitan desarrollar análisis estadísticos tales como regresiones lineales, clustering, etc.
- Crear una interfaz de usuario, basándose en la ya desarrollada, pero utilizando una librería propia de R, como puede ser Shiny [20] de manera que la comunicación entre interfaz y aplicación sea más sencilla y fluida. Además, a través de esta librería se podrían crear aplicaciones web interactivas que resulten más atractivas y funcionales que las normales.
- Ampliar la cantidad de formatos de entrada o de salida que soporta actualmente la aplicación.

Capítulo 8

Bibliografía

[1] Microsoft Office, diversas herramientas de procesado de textos, etc. Fecha de último acceso, 8 de Julio de 2019. Página web https://products.office.com/es-es/home

[2] Raúl Hernansanz Quevedo, 10 de junio de 2019, "Desarrollo de una aplicación para el tratamiento automático de encuestas", memoria del Trabajo de Fin de Grado de Ingeniería Informática.

[3] The R Project for Statistical Computing, entorno de programación, fecha de último acceso, 9 de Julio de 2019. Página web https://www.r-project.org/

[4] Instituto Nacional de Estadística, fecha de último acceso, 10 de Julio de 2019. Página web https://www.ine.es/

[5] Paquete de R, información al respecto en la documentación del CRAN, fecha de último acceso, 10 de Julio de 2019. Página web https://cran.r-project.org/web/packages/openxlsx/index.html

[6] Paquete de R, información al respecto en la documentación del CRAN, fecha de último acceso, 10 de Julio de 2019. Página web https://cran.r-project.org/web/packages/R2HTML/index.html

[7] Paquete de R, información al respecto en la documentación del CRAN, fecha de último acceso, 10 de Julio de 2019. Página web https://cran.r-project.org/web/packages/png/index.html

[8] Paquete de R, información al respecto en la documentación del CRAN, fecha de último acceso, 10 de Julio de 2019. Página web https://cran.r-project.org/web/packages/officer/index.html

[9] Paquete de R, información al respecto en la documentación del CRAN, fecha de último acceso, 10 de Julio de 2019. Página web https://cran.r-project.org/web/packages/survey/index.html

[10] Paquete de R, fecha de último acceso, 10 de Julio de 2019. Página web http://sqlsurvey.r-forge.r-project.org/

[11] Paquete de R, fecha de último acceso, 10 de Julio de 2019. Página web https://www.monetdb.org/Documentation/UserGuide/MonetDB-R

[12] Paquete de R, fecha de último acceso, 10 de Julio de 2019. Página web http://enquirer.free.fr/

[13] Página web de estadística de Castilla y Leon, fecha de último acceso, 10 de Julio de 2019. Página web https://estadistica.jcyl.es/web/jcyl/Estadistica/es/Plantilla100/1246989275272/_/_/_

[14] Página web sobre el envejecimiento de la población del Consejo Superior de Investigaciones Científicas, versión antigua. Fecha de último acceso, 10 de Julio de 2019. Enlace web http://envejecimiento.csic.es/

[15] Página web del Centro de Investigaciones Sociológicas, fecha de último acceso, 10 de Julio de 2019. Enlace web http://www.cis.es/cis/opencms/ES/index.html

[16] Página web del Instituto Vasco de Estadística, fecha de último acceso, 10 de Julio de 2019. Enlace web http://www.eustat.eus/indice.html

[17] Página web del Instituto Gallego de Estadística, fecha de último acceso, 10 de Julio de 2019. Enlace web https://www.ige.eu/web/index.jsp?paxina=001&idioma=gl

[18] Página web del Instituto Europeo de Estadística, fecha de último acceso, 10 de Julio de 2019. Enlace web https://ec.europa.eu/eurostat

[19] Página web de la Universidad Autónoma de Barcelona, fecha de último acceso, 10 de Julio de 2019. Enlace web http://pagines.uab.cat/plopez/content/inicio

[20] Paquete de R para construir aplicaciones web, último acceso el 23 de Junio de 2019. Página web https://shiny.rstudio.com/

Anexos

Apéndice A

Manual de Instalación de R

En esta sección se van a explicar, de forma detallada, todos los pasos necesarios para realizar la instalación de la herramienta R, necesaria para poder utilizar la aplicación desarrollada en este proyecto.

Antes de comenzar con las instrucciones para instalar R, cabe mencionar que, aunque dicha herramienta funciona en diversos sistemas operativos, como **Windows**, **Linux** o **Mac**, la aplicación desarrollada está pensada para Windows, por lo que este manual será para instalar R en este sistema operativo.

El primer paso será hacer doble click sobre el acceso directo llamado R que está incluido en la carpeta de la aplicación. Esto nos llevará a la página principal de R, donde deberemos presionar en el enlace marcado en rojo en la figura A.1, una vez hecho esto, tendremos que entrar en el siguiente enlace, también marcado en rojo en la figura A.2 y, finalmente, en el enlace de descarga de la figura A.3.

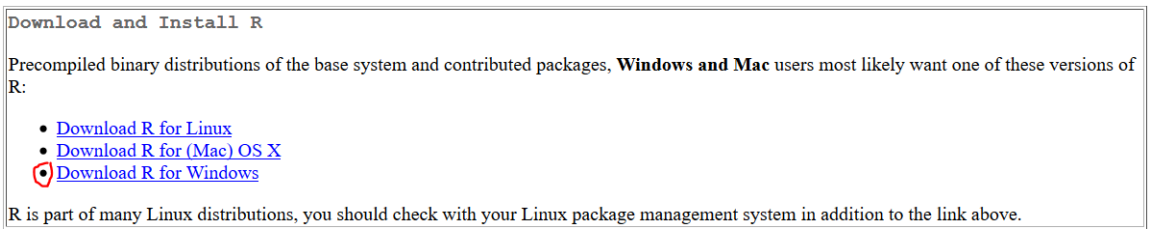


Figura A.1: Página web de R (parte 1)

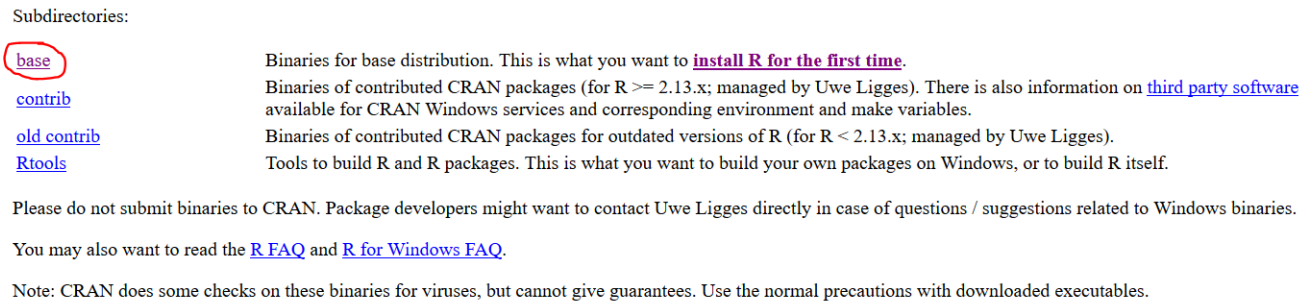


Figura A.2: Página web de R (parte 2)

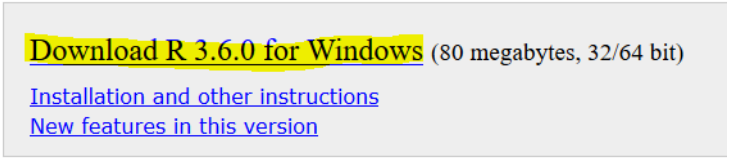


Figura A.3: Página web de R (parte 3)

Seguidamente habrá que ejecutar el archivo que hemos descargado y seguir las instrucciones que se describirán a continuación.

Para comenzar seleccionaremos el idioma en el que deseamos que esté la instalación, en mi caso será Español, una vez elegido, pulsaremos en **Aceptar**. Aparecerá entonces el acuerdo de licencia de la herramienta y tendremos que pulsar en **Siguiente**.

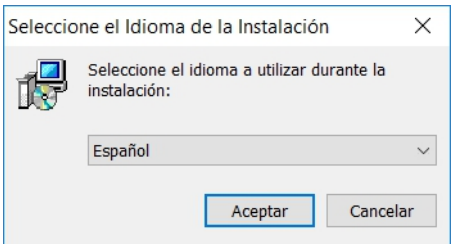


Figura A.4: Instalación de R (parte 1)

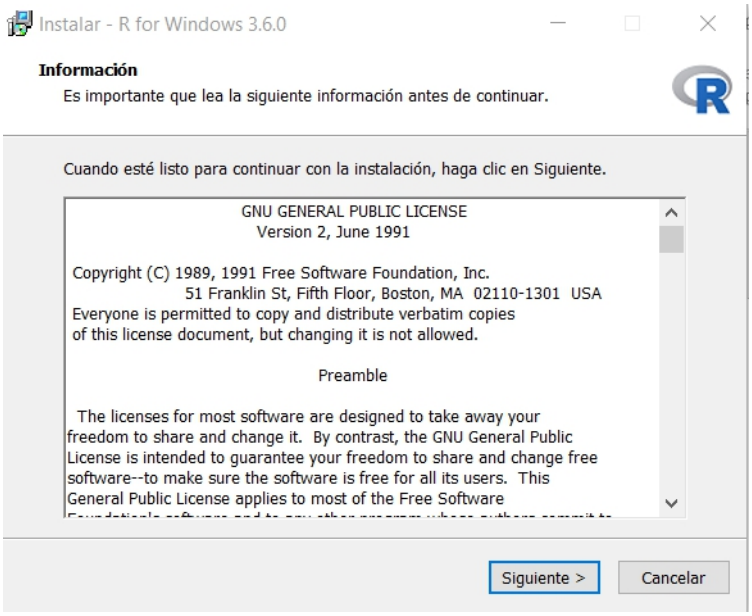


Figura A.5: Instalación de R (parte 2)

A continuación deberemos indicar la ruta en la que deseamos instalar R, esta ruta habrá que tenerla presente puesto que la aplicación desarrollada hará uso de ella, pero eso ya se explicará más adelante en esta misma sección.

Tras introducir la ruta que más nos interese, presionaremos en **Siguiente**. En la siguiente ventana, habrá que indicar los componentes que deseamos instalar. En este caso recomiendo instalar todos, para evitar posibles problemas. Con esto presente, pulsaremos en **Siguiente** tal y como aparece en la figura A.7.

Ahora pincharemos en **Siguiente** en las dos ventanas que siguen a la de los componentes, es decir, figuras A.8 y A.9 hasta que el instalador pida seleccionar las tareas

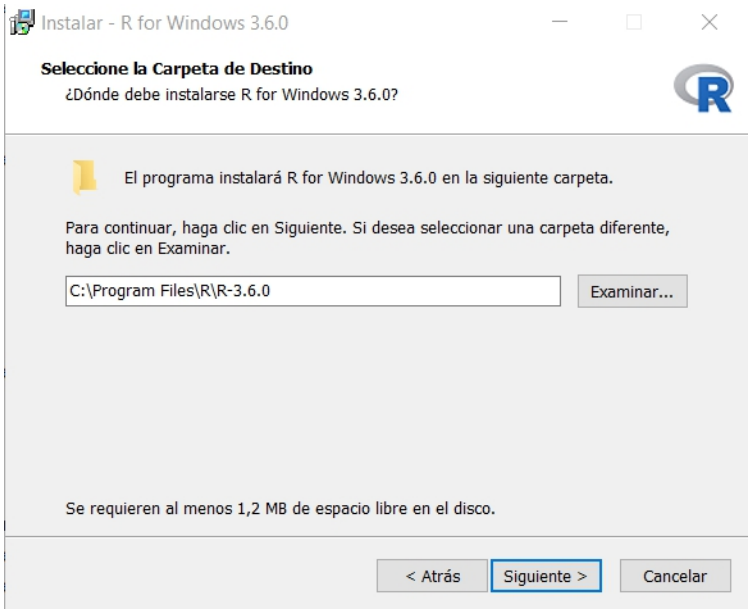


Figura A.6: Instalación de R (parte 3)

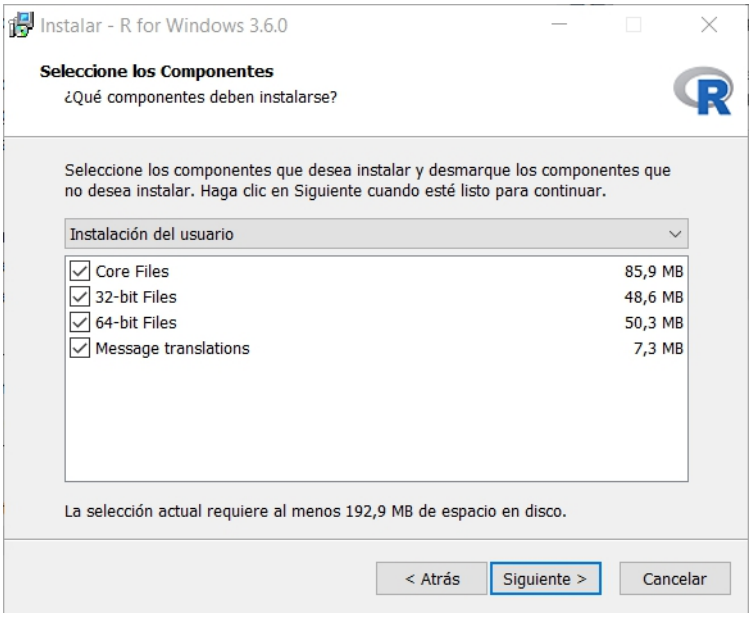


Figura A.7: Instalación de R (parte 4)

adicionales, como en la figura A.10. En este caso, además de las que aparecen marcadas por defecto, también seleccionaré la de crear un acceso directo en el escritorio. Tras esto, pulsaremos **Siguiente** y comenzará el proceso de instalación. Cuando termine, pulsaremos **Finalizar** y habrá terminado la instalación.

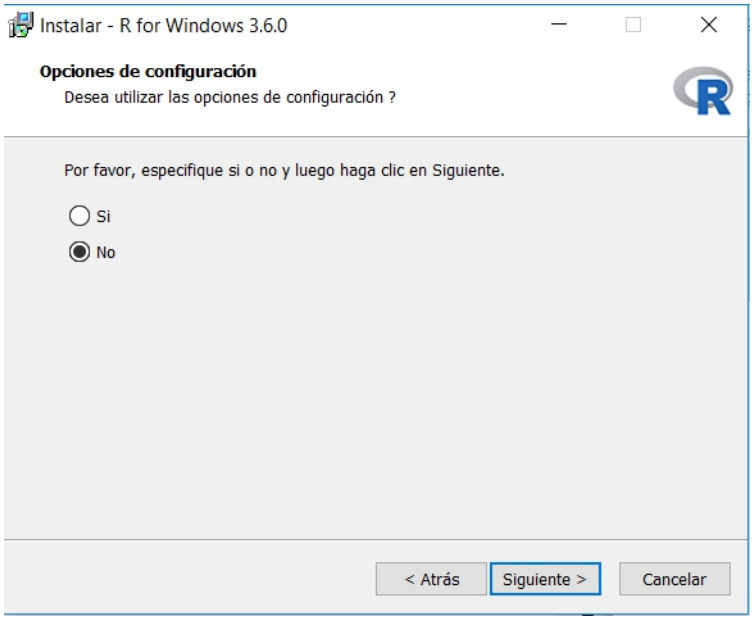


Figura A.8: Instalación de R (parte 5)

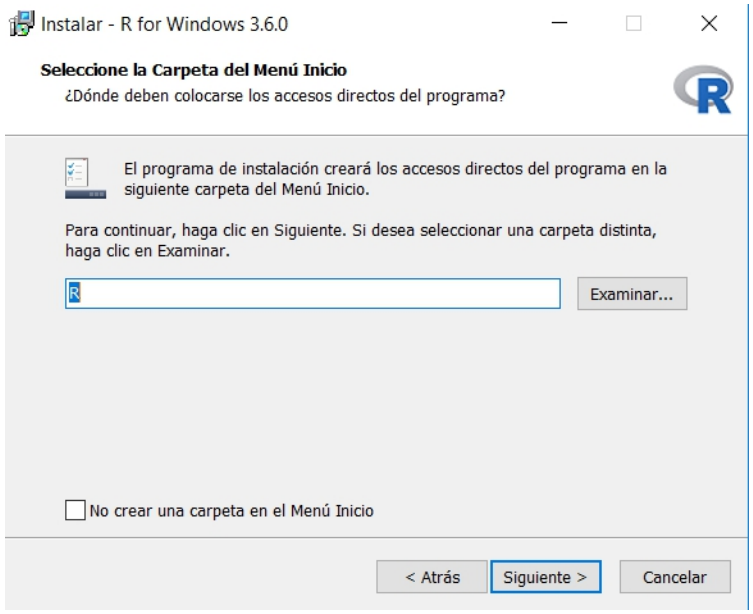


Figura A.9: Instalación de R (parte 6)

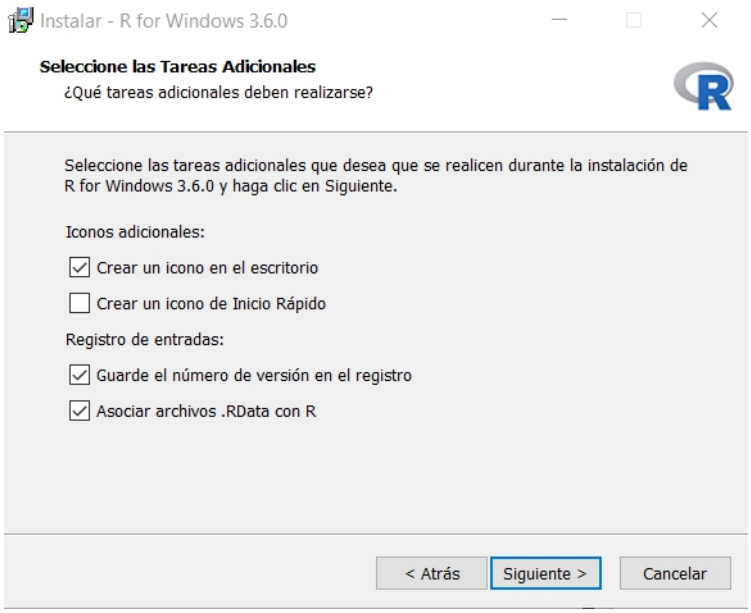


Figura A.10: Instalación de R (parte 7)

Apéndice B

Manual de uso de la Interfaz

Esta sección estará dedicada a explicar cómo debe ser usada la aplicación desarrollada para este proyecto. Como es una aplicación hecha con java, dicho programa deberá estar instalado para que pueda funcionar.

Teniendo en cuenta que se debe haber instalado R primeramente, se supondrá que ya se ha extraído el archivo comprimido en el que se encuentra la aplicación. En la carpeta principal, mostrada en la figura B.1 se pueden encontrar los siguientes archivos y directorios.

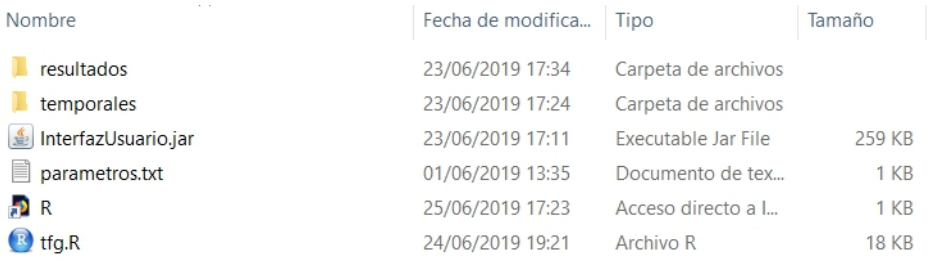


Figura B.1: Contenidos de la carpeta de la aplicación

- **tfg.R:** Archivo R que contiene el código fuente de la aplicación. Este archivo solo deberá ser abierto en el caso en el que se desee añadir o modificar algún módulo de la misma.

- **R:** Acceso directo a la web de descarga de la herramienta R.

- **parametros.txt:** Archivo de texto en el que el usuario podrá indicar algunos parámetros que utilizará la interfaz. Se debe tener en cuenta que no solo bastará añadir parámetros a este archivo, sino que será necesario modificar el código fuente de la aplicación R para contemplar los nuevos parámetros.

- **InterfazUsuario.jar:** Archivo java ejecutable que lanzará la interfaz de la aplicación.
- **temporales:** Directorio en el que se guardarán los archivos de texto que creará la interfaz de usuario y que son necesarios para que la aplicación R funcione. No es necesario que el usuario acceda a este directorio.
- **resultados:** Directorio en el que se guardarán las imágenes de los gráficos que cree el usuario, estas imágenes se incluirán más tarde en los informes indicados.

Para iniciar la aplicación se deberá hacer doble click sobre el archivo jar llamado InterfazUsuario, apareciendo de esta forma la ventana inicial de la aplicación, la cual se muestra en la figura B.2.

Figura B.2: Interfaz de usuario (parte 1)

Lo primero será hacer click en cualquier parte de la imagen para comenzar a utilizar la aplicación. De esta forma se abrira la ventana principal de la aplicación, mostrada en la figura B.3, la forma de utilizar la interfaz es muy sencilla, el primer paso sera introducir la ruta donde se encuentra el archivo de RScript, para ello pulsaremos en el botón con los 3 puntos situado en la parte derecha de ese mismo campo y buscaremos la ruta en la que tenemos instalado R; una vez en ella, accederemos a la carpeta **bin** marcada en la figura B.4.

Tras esto, habrá que entrar a la carpeta **x64**, también marcada en la figura B.5. Por último seleccionaremos el archivo **Rscript.exe** indicado en la figura B.6.

Una vez seleccionado el ejecutable de R, introduciremos el resto de parámetros comenzando por la ubicación del archivo de datos, el cual lo podremos introducir de la misma forma que la ruta de R. Después habrá que indicar el separador que emplea dicho archivo y el formato en el que deseamos el informe final.

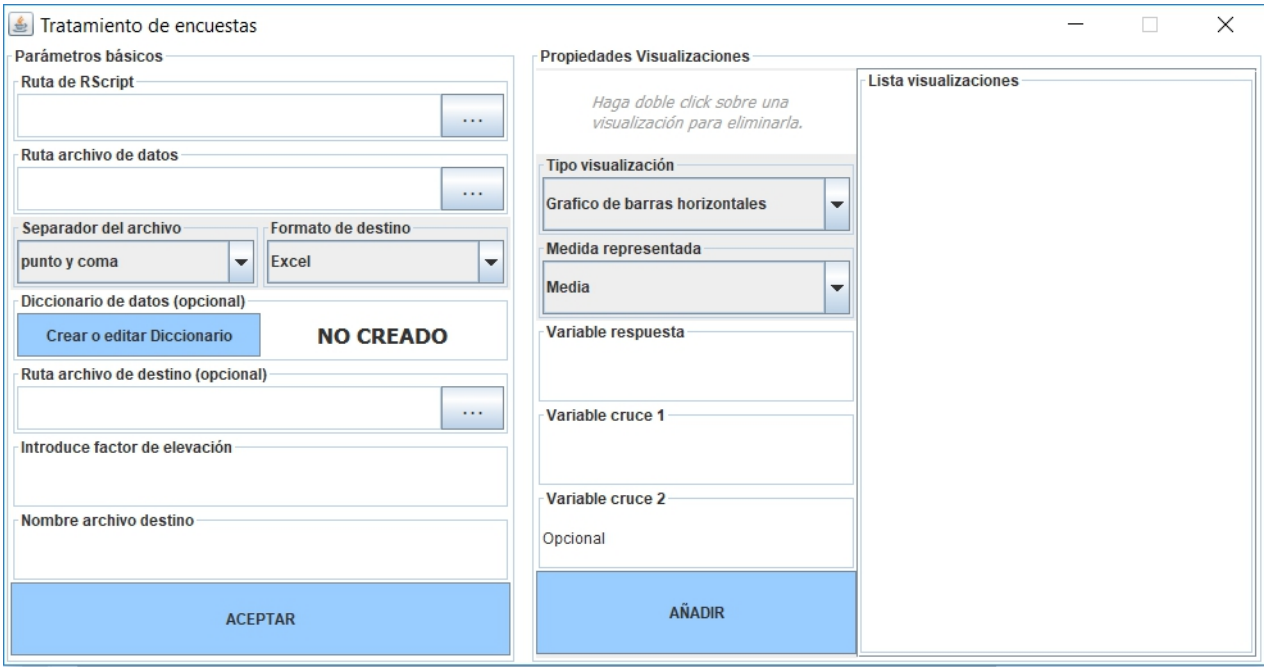


Figura B.3: Interfaz de usuario (parte 2)

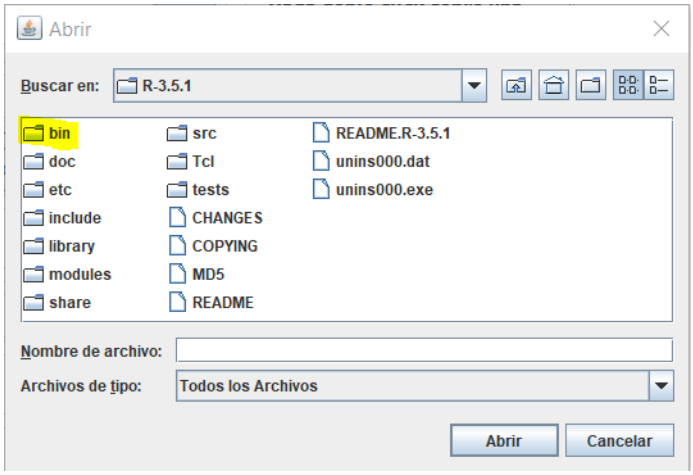


Figura B.4: Ruta de R (parte 1)

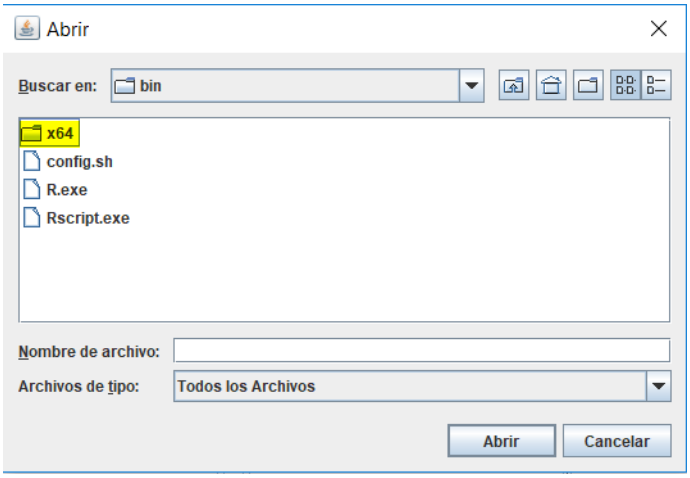


Figura B.5: Ruta de R (parte 2)

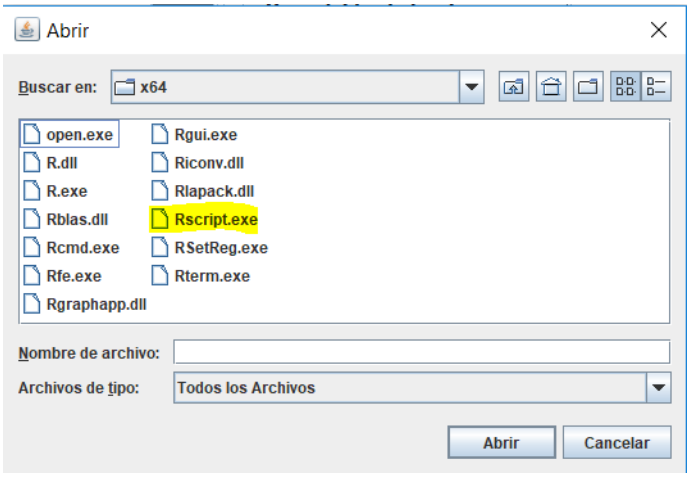


Figura B.6: Ruta de R (parte 3)

Tras esto, podremos crear un diccionario, cuya utilidad ya ha sido explicada en diferentes puntos de este mismo documento. Para hacerlo, lo primero será pulsar en el botón "Crear o editar Diccionario", mostrándose así la ventana de la figura B.7, una vez ahí, podremos crear o eliminar variables a nuestro gusto. Para crearlas, solo debemos introducir su nombre, código y descripción y pulsar sobre "Añadir", de esta forma se mostrará la variable creada en el campo de texto inferior, para ilustrar esto, se ha creado la variable "ejemplo". En el caso de que se desee eliminar dicha variable, solo habrá que hacer doble click sobre cualquier parte de ella en el panel de texto.

Figura B.7: Creación del diccionario

Hay que tener en cuenta que cuando se esté creando el diccionario, los cambios no se guardarán si el usuario no presiona el botón **Aceptar**. Cuando el diccionario esté creado o en el caso en el que no se desee crear, habrá que introducir la ruta donde queremos los resultados, si esta ruta no se indica, por defecto estará en la misma ubicación que la carpeta de la aplicación.

Lo siguiente será escribir el nombre **exacto** del factor de elevación que se encuentra en la encuesta con la que vamos a trabajar. El último parámetro a indicar, será el nombre que el usuario desea dar al informe resultado.

Cuando todos los parámetros estén indicados, habrá que introducir las visualizaciones que se deseen. Esto se hará en la parte derecha de la ventana principal, mostrada en la figura B.8, donde habrá que introducir el tipo de visualización deseada, la medida que queremos representar, la variable respuesta y, al menos, la primera variable de cruce. **Todas** las variables introducidas deben estar escritas de manera **exacta** a como están escritas en la encuesta con la que estemos trabajando, incluyendo si están o no en mayúsculas.

Una vez se ha rellenado todo, pulsaremos "Añadir" y el resultado de la visualización aparecerá en el panel de texto derecho. Para este ejemplo, se ha solicitado un Gráfico de barras horizontales, que represente la media de la variable *REJEMPLO* frente a la variable *CEJEMPLO1*, como no se ha introducido una segunda variable de cruce, ese campo aparece con un guión. La forma de eliminar una visualización es la misma que la

utilizada con las variables del diccionario, bastará con hacer doble click en cualquier parte de dicha visualización en el panel de texto.

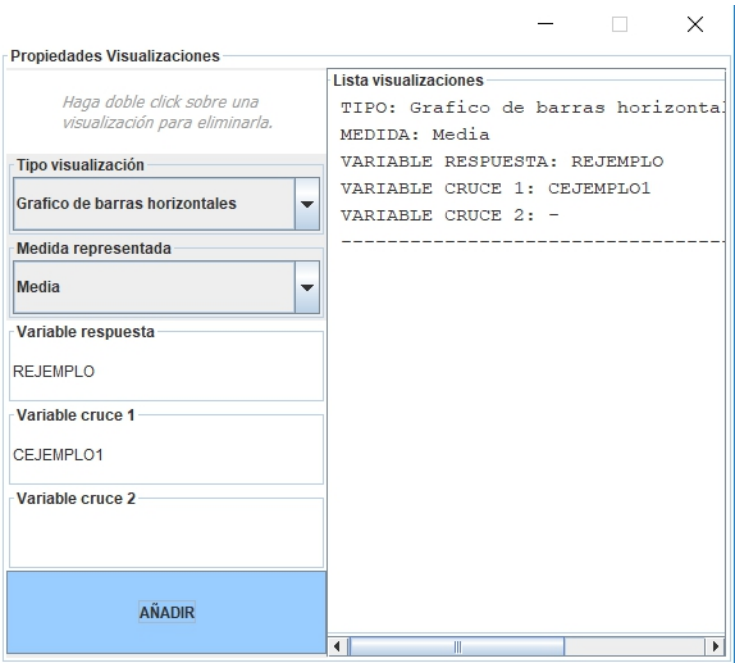


Figura B.8: Creación de visualizaciones

Finalmente, cuando todo esté introducido, pulsaremos el botón **Aceptar** y, tras un poco de tiempo, se creará el informe en la ubicación indicada. En el caso de que no se cree, se deberán comprobar las variables introducidas por si hubiera algún error.

Algo a tener en cuenta es que, si cuando presionamos **Aceptar**, el archivo que deseamos ya existe en la ubicación indicada, aparecerá una ventana de alerta, mostrada en la figura B.9, en este caso existirán dos opciones, la primera es reemplazar el archivo pulsando **Reemplazar** y la segunda es cancelar la operación pulsando **Cancelar**, si hacemos esto último, volveremos a la ventana de la interfaz.

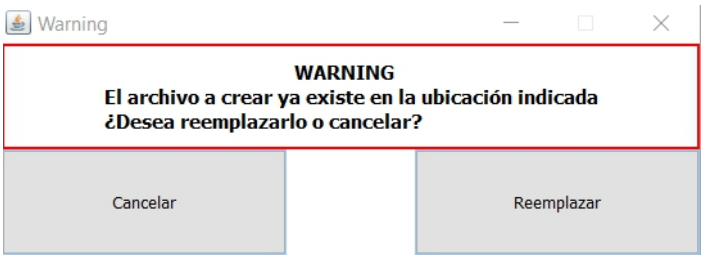


Figura B.9: Ruta de R (parte 4)

Una **observación importante** es que, a la hora de depositar la carpeta que contiene todos los elementos de la aplicación, se **debe** hacer en un lugar con una ruta absoluta no muy larga, y sin caracteres raros, debido a la codificación de los ficheros de texto generados y de limitaciones de lectura en rutas demasiado extensas.

Una vez explicado el funcionamiento general de la aplicación, se van indicar algunas cosas que se pueden realizar para probarla. Empleando el archivo de Excel incluido entre el contenido del CD, se tendría que introducir, como factor de elevación, la variable **"FACTOTAL"** y, como separador del archivo, el **tabulador**. En la figura B.10 se muestran algunas variables que se pueden introducir como variables de cruce para las visualizaciones. Como variable respuesta, se introducirá **"SALBRUTO"**. Para mostrar un ejemplo, en la imagen B.11 se han introducidos todos los parámetros necesarios para crear un gráfico de barras horizontales.

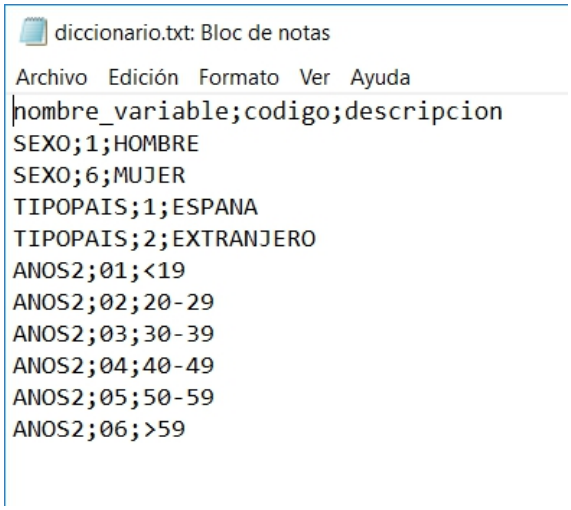


Figura B.10: Ejemplo de variables

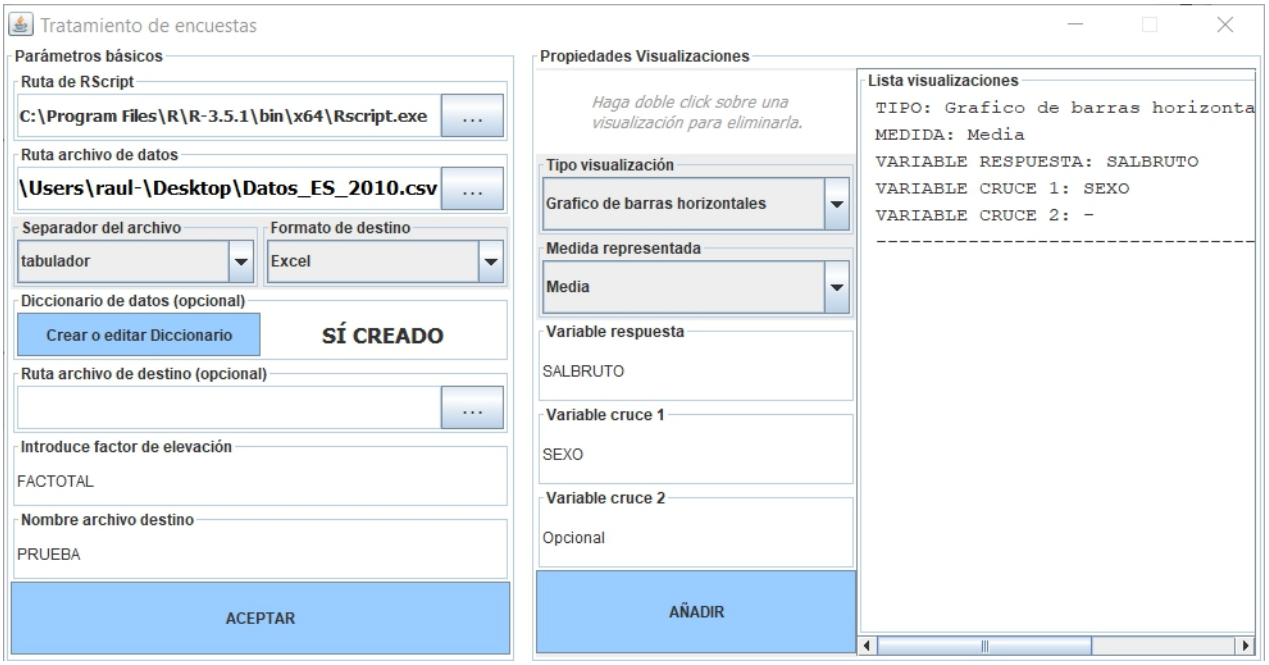


Figura B.11: Ejemplo de parámetros

Apéndice C

Contenido del CD

```
/
└── Tratamiento automatico de encuestas con R.zip
    ├── Tratamiento automatico de
    │   encuestas con R.pdf ..... Memoria del Trabajo de Fin de
    │                             Grado en Estadística.
    └── Programa
        ├── resultados ..... Carpeta donde se almacenarán las
        │                    imágenes de las visualizaciones
        │                    de forma temporal.
        ├── Datos_ES_2010.csv ..... Fichero Excel con datos
        │                           de ejemplo para probar la
        │                           aplicación.
        ├── temporales ..... Carpeta donde se crearán los
        │                    archivos de texto que leerá la
        │                    aplicación de R.
        ├── InterfazUsuario.jar ..... Ejecutable que lanza la
        │                             aplicación.
        ├── parametros.txt ..... Documento de texto en el que
        │                        se pueden introducir diversos
        │                        parámetros.
        ├── R ..... Acceso directo a la página de
        │           descarga de R.
        └── tfg.R ..... Código fuente de la aplicación
                        R.
```