



## PRÁCTICA 2. MANEJO DE FICHEROS. INTERACCIÓN CON OTROS PROGRAMAS: EXCEL Y SPSS

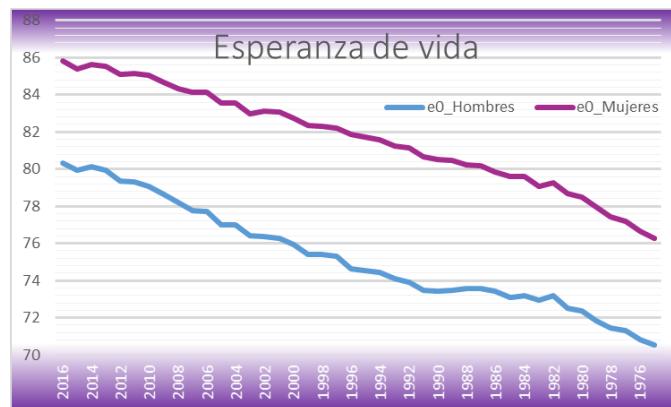
### Ejercicio 1

En la web del Instituto Nacional de Estadística, <http://www.ine.es/>, entramos en IneBase/Demografía y Población/Indicadores Demográficos Básicos. Descarguemos los últimos indicadores demográficos (Edad media a la maternidad, Indicador Coyuntural de Fecundidad Total y Españolas/Extranjeras, Esperanza de vida al nacimiento y a los 65 años, Razón de masculinidad al nacimiento).

Descargue y prepare su fichero (si tiene dificultad, puede descargarlo de la web en [http://www.ugr.es/~mdhuete/Informatica/Datos\\_demográficos.xls](http://www.ugr.es/~mdhuete/Informatica/Datos_demográficos.xls)).

Año	Edad_media	ICF_Total	ICF_Españolas	ICF_Extranjeras	e0_Hombres	e0_Mujeres	e65_Hombres	e65_Mujeres	Razón_masc_nac
2016	31,997702	1,337999	1,279723	1,717761	80,310001	85,836114	19,143034	23,050313	105,763103
2015	31,899652	1,333638	1,283599	1,662833	79,923839	85,408152	18,793134	22,645315	106,196256
2014	31,778309	1,320230	1,272676	1,621409	80,121264	85,641977	19,057086	22,923396	106,790848

Realice gráficos que permitan mostrar la evolución temporal de cada uno de los indicadores. Por ejemplo:



- Guarde una copia de los datos y prepárelo para llevar a R-Commander (ordene dentro de excel los datos, según el año de observación y suprima los años con datos faltantes).
- Importe sus datos a un fichero .Rdata.
- Guarde los datos en el Menú Datos/Conjunto de Datos Activo como Demo.RData.
- En el menú Gráficas, realizaremos un gráfico de líneas.



## Ejercicio 2

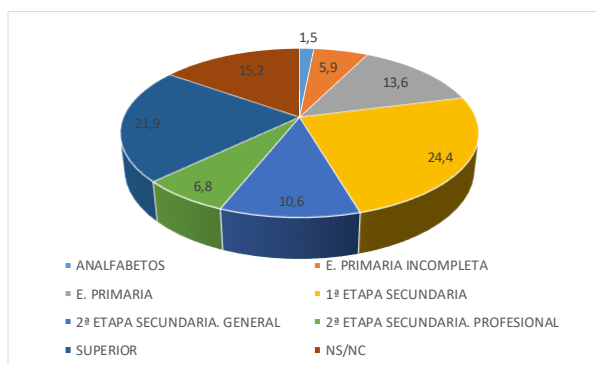
Descargue el siguiente fichero que contiene una muestra aleatoria del 5% de los datos recogidos en la Encuesta de Población Activa (1er trimestre del año 2016). Esta encuesta está elaborada por el INE todos los trimestres del año para conocer el estado de España en cuanto a actividad y paro. Consta de más de 160.000 registros.

[http://www.ugr.es/~mdhuete/Informatica/Muestra5\\_EPA.xls](http://www.ugr.es/~mdhuete/Informatica/Muestra5_EPA.xls)

1.- En una hoja nueva, vamos a crear algunas tablas de frecuencias:

- Obtengamos la tabla de frecuencias para el “Nivel de Estudios” de España. Necesita usar las órdenes `CONTAR.SI(Muestra_EPA!$G$2:$G$8002;B4)` y `CONTAR.BLANCO(Muestra_EPA!$G$2:$G$8002)` para los registros que están vacíos. Calcule también las frecuencias relativas en porcentaje y realice un diagrama de sectores con ellas.

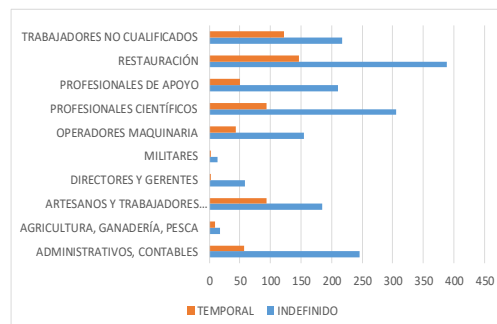
Tabla de frecuencias NIVEL ESTUDIOS	ni	fi (%)
ANALFABETOS	124	1,5
E. PRIMARIA INCOMPLETA	469	5,9
E. PRIMARIA	1089	13,6
1ª ETAPA SECUNDARIA	1953	24,4
2ª ETAPA SECUNDARIA. GENERAL	850	10,6
2ª ETAPA SECUNDARIA. PROFESIONAL	545	6,8
SUPERIOR	1755	21,9
NS/NC	1216	15,2
<b>Sumas</b>	<b>8001</b>	<b>100</b>



- Obtenga la tabla de frecuencias para el tipo de contrato usando también la función `CONTAR.SI`.
- Halle la distribución bidimensional de la ocupación según el tipo de contrato. En este caso, necesita usar la orden `CONTAR.SI.CONJUNTO(Muestra_EPA!$K$2:$K$8002; $B$32; Muestra_EPA!$L$2:$L$8002;C$31)`.

Realice otra tabla con las frecuencias relativas y gráficas.

Bidimensional OCUPACIÓN/TIPO CONTRATO	INDEFINIDO	TEMPORAL
ADMINISTRATIVOS, CONTABLES	246	57
AGRICULTURA, GANADERÍA, PESCA	17	10
ARTESANOS Y TRABAJADORES CUALIFICADOS	184	94
DIRECTORES Y GERENTES	58	2
MILITARES	14	2
OPERADORES MAQUINARIA	155	43
PROFESIONALES CIENTÍFICOS	306	94
PROFESIONALES DE APOYO	211	50
RESTAURACIÓN	389	146
TRABAJADORES NO CUALIFICADOS	217	122



- Calcule los estadísticos más importantes del “Número de meses en la empresa”. Debe usar las funciones Promedio, Percentil, VarP.

Estadísticos Número de Meses en la Empresa			
TOTALES			
Media=	151		
Mediana=	118		
P80=	283,6		
Varianza=	18823,6		
CV=	0,90753503		

- Ahora seleccione mediante un filtro, solamente a las Mujeres y copie los datos a una hoja nueva. Vuelva a calcular los estadísticos. Realice lo mismo para los varones. Compare los datos obtenidos en ambos casos.

#### MUJER VARÓN

321	14
26	0
192	310
429	409
120	2
129	146
192	357
236	612
318	72
183	63
251	252
35	215
75	132
48	216
96	487
13	396
384	252
252	8

#### Estadísticos Número de Meses en la Empresa

TOTALES	
Media=	151
Mediana=	118
P80=	283,6
Varianza=	18823,6
CV=	0,90753503

#### MUJERES

Media=	
Mediana=	
P80=	
Varianza=	
CV=	

#### HOMBRES

Media=	
Mediana=	
P80=	
Varianza=	
CV=	

- Seleccione las celdas no vacías del número de meses en la empresa (mediante un filtro), al igual que el Sexo y cópielas a una hoja nueva. Sustituya “Varón” por 1 y “Mujer” con 0. Estime el número de meses en la empresa en función del Sexo y comente el ajuste (use las funciones PENDIENTE y INTERSECCION.EJE) para hallar la pendiente y ordenada en el origen respectivamente.

	A	B	C	D	E	F	G
1	MESES EN EMPRESA	SEXO			AJUSTE DE REGRESIÓN LINEAL SIMPLE		
2							
3					Pendiente:		
4					Ordenada:		
5							

- Realice el último ejercicio del Seminario de Estadística Descriptiva, usando la hoja de cálculo, para calcular todas las medidas estadísticas a través de las marginales, como sigue:



				Estimamos ahorro de una familia que ingresa 830 euros y calculamos el coeficiente de determinación					
X: Ingresos mensuales (x100€)				(usamos marcas de clase de los intervalos)					
Y: Ahorro (x100€)									
	X/Y	0,1	0,6	2	5	ni*	xi*ni	xi^2*ni	
	1	2	0	0	0				
	5	0	0	1	0				
	10	0	0	4	1				
	16	0	0	2	5				
	n*j								
	yj*nj								
	yj^2*nj								
	Cov								

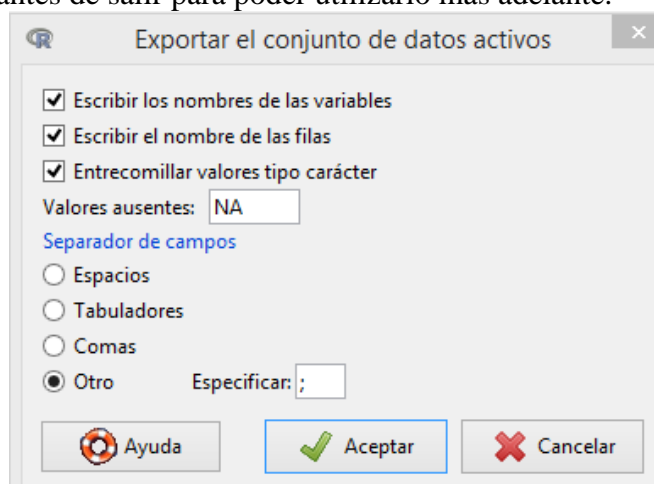
Para calcular la covarianza, necesita fijar las filas o columnas para poder pegar la fórmula.  
 Por ejemplo, para n11 necesita escribir “E7\*E\$6\*\$D7”.



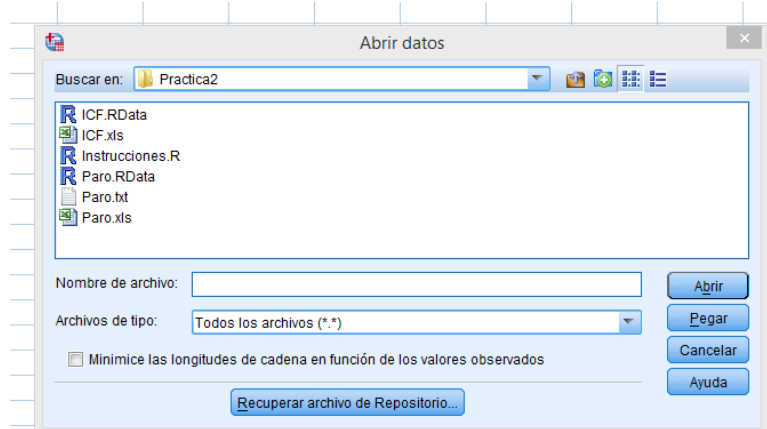
### Ejercicio 3

Vamos a ver ahora cómo es posible realizar ejercicios con un paquete específico de Estadística (en particular, SPSS). Dependiendo de la versión disponible en los ordenadores, las siguientes imágenes pueden variar levemente.

- a) En el Menú Datos/conjunto de datos activo, vamos a exportar nuestros datos .RData. Elegimos separaciones con “;” y lo guardamos en la carpeta Practica2 como Paro.txt. Puede cerrar R, para lo que sigue. Si lo desea, guarde el archivo de instrucciones antes de salir para poder utilizarlo más adelante.



- b) Abra el fichero Paro.txt y sustituya el punto decimal por una coma. Entramos en SPSS. En el Menú Archivo/Abrir datos, buscamos el fichero Paro.txt. Importamos los datos.



- c) Comprobamos ahora que nuestro fichero está correctamente importado. En caso contrario, modificamos lo necesario:

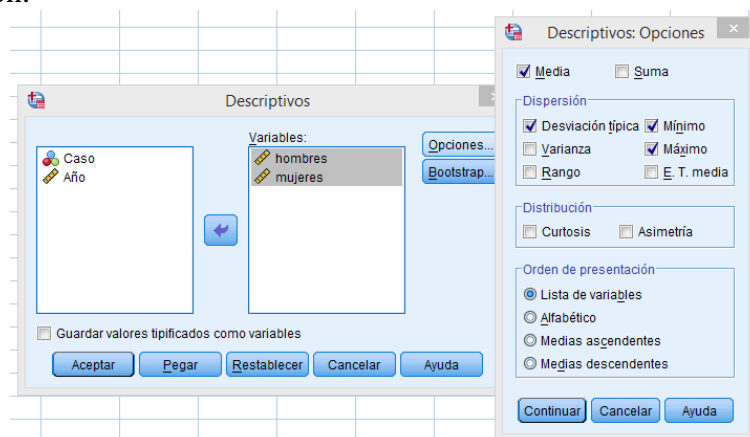


	Caso	Año	hombres	mujeres
1	1	2006	30	33
2	2	2007	21	28
3	3	2008	38	26
4	4	2009	47	26
5	5	2010	58	37
6	6	2011	68	32
7	7	2012	71	43
8	8	2013	76	49

- d) Construimos un gráfico de líneas en el Menú Gráficos. Si entra en el editor de gráficos de la ventana de resultados, puede modificar el mismo pulsando doble clic en la parte que desea modificar.



- e) En el menú Analizar/Estadísticos Descriptivos/Descriptivos, calcule el número medio de demandantes de empleo para hombres y mujeres en todo el periodo de observación.



## Ejercicio 4



Busque ahora la población de su municipio en SIMA y construya en Excel un fichero que contenga los parados de su municipio (por sexo) y la población de mujeres y hombres. Llévelo a R-comander y construya una tasa de paro en cada año de observación. Grafique dicha tasa con R-Commander, con SPSS y con Excel.

