

# Programación Técnica y Científica

## Práctica evaluable 1 (hasta 2,5 puntos)

El objetivo de esta práctica es trabajar con los diferentes tipo de datos que ofrece python así como los módulos csv, numpy, bs4, matplotlib y otros que se necesiten para el manejo de datos procedentes de páginas webs y ficheros .csv. Los datos de entrada son suministrados en ficheros .csv y .html, habrá que leerlos correctamente y calcular valores estadísticos para generar tablas html de salida y ciertos gráficos para representar la información obtenida. No se puede usar el módulo “pandas” (prohibido el tipo de dato dataframe). Se debe usar el tipo de dato diccionario preferentemente y los numpy array. La práctica tiene que funcionar correctamente en la máquina virtual suministrada por el profesor que tiene Ubuntu 20.04, Python 3.8.3, Anaconda 4.8.3. y Spyder 4.1.4.

Como información de entrada vamos a trabajar con valores de población obtenidos de la página web del Instituto Nacional de Estadística (<https://www.ine.es/>).

En Prado tendremos este guión de la práctica, varios ejemplos de ayuda y los ficheros con los datos de entrada:

1. Fichero **poblacionProvinciasHM2010-17.csv** contiene datos de la población por sexo y provincia desde el año 2010 a 2017.
2. Página web **comunidadesAutonomas.html** contiene el listado de las comunidades autónomas con sus códigos.
3. Página web **comunidadAutonoma-Provincia.html** contiene el listado de provincias de cada comunidad autónoma con sus códigos.

Se pide crear un conjunto de programas en python para obtener los siguientes resultados **R1** a **R6**. En todos los casos los valores numéricos de población tienen que tener el separador de miles. Todos los valores numéricos que tengan decimales deben estar redondeados a 2 cifras decimales. En las páginas webs generadas la codificación debe ser utf8 y los caracteres tienen que verse correctamente en el sistema operativo de la máquina virtual (Ubuntu 20.04).

**R1.** Calcular la variación de la población por provincias desde el año 2011 a 2017 en términos absolutos y relativos generando la página web 1 (que debe llamarse **variacionProvincias.html**) que contenga una tabla parecida a la que se puede observar en el siguiente ejemplo:

Variación anual en la población por provincias  
Unidades: Personas, Porcentaje

	Variación absoluta							Variación relativa						
	2017	2016	2015	2014	2013	2012	2011	2017	2016	2015	2014	2013	2012	2011
Total Nacional	15.124 <sup>1</sup>	-67.374 <sup>1</sup>	-146.959 <sup>1</sup>	-358.442 <sup>1</sup>	-135.538 <sup>1</sup>	74.828 <sup>1</sup>	169.462 <sup>1</sup>	0.03 <sup>1</sup>	-0.14 <sup>1</sup>	-0.31 <sup>1</sup>	-0.76 <sup>1</sup>	-0.29 <sup>1</sup>	0.16 <sup>1</sup>	0.36 <sup>1</sup>
02 Albacete	-2.086 <sup>1</sup>	-2.462 <sup>1</sup>	-2.407 <sup>1</sup>	-3.020 <sup>1</sup>	-2.830 <sup>1</sup>	519 <sup>1</sup>	636 <sup>1</sup>	-0.53 <sup>1</sup>	-0.62 <sup>1</sup>	-0.61 <sup>1</sup>	-0.75 <sup>1</sup>	-0.70 <sup>1</sup>	0.13 <sup>1</sup>	0.16 <sup>1</sup>
03 Alicante/Alacant	-11.127 <sup>1</sup>	-18.588 <sup>1</sup>	-13.391 <sup>1</sup>	-77.204 <sup>1</sup>	1.732 <sup>1</sup>	9.783 <sup>1</sup>	7.842 <sup>1</sup>	-0.61 <sup>1</sup>	-1.00 <sup>1</sup>	-0.72 <sup>1</sup>	-3.97 <sup>1</sup>	0.09 <sup>1</sup>	0.51 <sup>1</sup>	0.41 <sup>1</sup>
04 Almería	2.375 <sup>1</sup>	3.086 <sup>1</sup>	-477 <sup>1</sup>	2.359 <sup>1</sup>	-4.890 <sup>1</sup>	1.400 <sup>1</sup>	7.259 <sup>1</sup>	0.34 <sup>1</sup>	0.44 <sup>1</sup>	-0.07 <sup>1</sup>	0.34 <sup>1</sup>	-0.69 <sup>1</sup>	0.20 <sup>1</sup>	1.04 <sup>1</sup>
01 Aragón/Alava	2.448 <sup>1</sup>	478 <sup>1</sup>	1.716 <sup>1</sup>	515 <sup>1</sup>	-1.140 <sup>1</sup>	3.330 <sup>1</sup>	1.875 <sup>1</sup>	0.76 <sup>1</sup>	0.15 <sup>1</sup>	0.53 <sup>1</sup>	0.16 <sup>1</sup>	-0.35 <sup>1</sup>	1.04 <sup>1</sup>	0.59 <sup>1</sup>
33 Asturias	-7.648 <sup>1</sup>	-8.621 <sup>1</sup>	-10.527 <sup>1</sup>	-6.409 <sup>1</sup>	-9.195 <sup>1</sup>	-4.127 <sup>1</sup>	-2.854 <sup>1</sup>	-0.73 <sup>1</sup>	-0.82 <sup>1</sup>	-0.99 <sup>1</sup>	-0.60 <sup>1</sup>	-0.85 <sup>1</sup>	-0.38 <sup>1</sup>	-0.26 <sup>1</sup>
05 Ávila	-1.814 <sup>1</sup>	-2.411 <sup>1</sup>	-2.090 <sup>1</sup>	-1.810 <sup>1</sup>	-2.440 <sup>1</sup>	-1.439 <sup>1</sup>	808 <sup>1</sup>	-1.12 <sup>1</sup>	-1.46 <sup>1</sup>	-1.25 <sup>1</sup>	-1.07 <sup>1</sup>	-1.42 <sup>1</sup>	-0.83 <sup>1</sup>	0.47 <sup>1</sup>
06 Badajoz	-4.229 <sup>1</sup>	-2.617 <sup>1</sup>	-4.199 <sup>1</sup>	-2.800 <sup>1</sup>	-804 <sup>1</sup>	612 <sup>1</sup>	1.784 <sup>1</sup>	-0.62 <sup>1</sup>	-0.38 <sup>1</sup>	-0.61 <sup>1</sup>	-0.40 <sup>1</sup>	-0.12 <sup>1</sup>	0.09 <sup>1</sup>	0.26 <sup>1</sup>
07 Baleares, Illes	8.779 <sup>1</sup>	2.741 <sup>1</sup>	1.037 <sup>1</sup>	-8.232 <sup>1</sup>	-7.765 <sup>1</sup>	6.325 <sup>1</sup>	7.065 <sup>1</sup>	0.79 <sup>1</sup>	0.25 <sup>1</sup>	0.09 <sup>1</sup>	-0.74 <sup>1</sup>	-0.69 <sup>1</sup>	0.57 <sup>1</sup>	0.64 <sup>1</sup>
08 Barcelona	33.357 <sup>1</sup>	19.758 <sup>1</sup>	138 <sup>1</sup>	-17.141 <sup>1</sup>	-11.125 <sup>1</sup>	22.951 <sup>1</sup>	17.952 <sup>1</sup>	0.60 <sup>1</sup>	0.34 <sup>1</sup>	0.00 <sup>1</sup>	-0.31 <sup>1</sup>	-0.20 <sup>1</sup>	0.42 <sup>1</sup>	0.33 <sup>1</sup>
48 Bizkaia	726 <sup>1</sup>	-1.199 <sup>1</sup>	-3.130 <sup>1</sup>	-4.542 <sup>1</sup>	-1.992 <sup>1</sup>	2.667 <sup>1</sup>	2.048 <sup>1</sup>	0.06 <sup>1</sup>	-0.10 <sup>1</sup>	-0.27 <sup>1</sup>	-0.39 <sup>1</sup>	-0.17 <sup>1</sup>	0.23 <sup>1</sup>	0.18 <sup>1</sup>
09 Burgos	-2.824 <sup>1</sup>	-3.007 <sup>1</sup>	-2.898 <sup>1</sup>	-4.348 <sup>1</sup>	-3.722 <sup>1</sup>	-687 <sup>1</sup>	831 <sup>1</sup>	-0.78 <sup>1</sup>	-0.83 <sup>1</sup>	-0.79 <sup>1</sup>	-1.17 <sup>1</sup>	-0.99 <sup>1</sup>	-0.18 <sup>1</sup>	0.22 <sup>1</sup>
10 Cáceres	-3.629 <sup>1</sup>	-2.602 <sup>1</sup>	-2.436 <sup>1</sup>	-1.572 <sup>1</sup>	-3.322 <sup>1</sup>	-1.849 <sup>1</sup>	363 <sup>1</sup>	-0.90 <sup>1</sup>	-0.64 <sup>1</sup>	-0.60 <sup>1</sup>	-0.38 <sup>1</sup>	-0.80 <sup>1</sup>	-0.45 <sup>1</sup>	0.09 <sup>1</sup>
11 Cádiz	-454 <sup>1</sup>	-395 <sup>1</sup>	109 <sup>1</sup>	1.683 <sup>1</sup>	-6.872 <sup>1</sup>	1.645 <sup>1</sup>	6.780 <sup>1</sup>	-0.04 <sup>1</sup>	-0.03 <sup>1</sup>	0.01 <sup>1</sup>	0.14 <sup>1</sup>	-0.54 <sup>1</sup>	0.13 <sup>1</sup>	0.55 <sup>1</sup>
39 Cantabria	-1.911 <sup>1</sup>	-2.973 <sup>1</sup>	-3.477 <sup>1</sup>	-3.232 <sup>1</sup>	-1.973 <sup>1</sup>	740 <sup>1</sup>	871 <sup>1</sup>	-0.33 <sup>1</sup>	-0.51 <sup>1</sup>	-0.59 <sup>1</sup>	-0.55 <sup>1</sup>	-0.33 <sup>1</sup>	0.12 <sup>1</sup>	0.15 <sup>1</sup>

Las fórmulas a aplicar son:

variación absoluta 2017=población 2017 – población 2016

variación relativa 2017=(variación absoluta 2017 / población 2016) \* 100

Para una mejor comprensión se puede observar un ejemplo similar al fichero solicitado en la página web con título: **variacionProvincias2011-17.html** también disponible en Prado.

Vamos a conservar el mismo orden de listado que aparece en el fichero **poblacionProvinciasHM2010-17.csv** para mantener la coherencia con los listados del INE.

**R2.** Usando el listado de comunidades autónomas que podemos obtener del fichero **comunidadesAutonomas.html**, así como de las provincias de cada comunidad autónoma que podemos obtener de **comunidadAutonoma-Provincia.html** y los datos de **poblacionProvinciasHM2010-17.csv**, hay que generar una **página web 2** (fichero **poblacionComAutonomas.html**) con una tabla con los valores de población de cada comunidad autónoma en cada año de 2010 a 2017, indicando también los valores desagregados por sexos (de manera semejante a como aparece en el fichero **poblacionProvinciasHM2010-17.csv**)

Conservar el orden que presenta **comunidadesAutonomas.html**

**R3.** Usando Matplotlib, para las 10 comunidades con más población media de 2010 a 2017, generar un gráfico de columnas que indique la población de hombres y mujeres en el año 2017, salvar el gráfico a fichero e incorporarlo **a la página web 2** del punto R2.

**R4.** Generar una **página web 3** (fichero **variacionComAutonomas.html**) con una tabla con la variación de población por comunidades autónomas desde el año 2011 a 2017, indicando variación absoluta, relativa y desagregando dicha información por sexos, es decir, variación absoluta (hombres, mujeres) y relativa (hombres, mujeres). Para los cálculos, hay que actuar de manera semejante que en el apartado **R1**.

**R5.** Usando Matplotlib, para las 10 comunidades elegidas en el punto R3 generar un gráfico de líneas que refleje la evolución de la población total de cada comunidad autónoma desde el año 2010 a 2017, salvar el gráfico a fichero e incorporarlo **a la página web 3** del punto R4.

**R6.** Usando como entrada la página web 1 generada en el apartado 1 llamada **variacionProvincias.html** y el fichero proporcionado **variacionProvincias2011-17.html** hay que implementar un programa que compare los datos de variación de población de 2011 a 2017 (absoluta y relativa) de ambos ficheros para comprobar que son los mismos valores en cada caso. Después usar el fichero **comunidadesAutonomasBis.html** para generar una versión Bis de las páginas web 2 (ver R2 y R3) y web 3 (ver R4 y R5) con sus respectivos gráficos debiendo llamarse **poblacionComAutonomasBis.html** y **variacionComAutonomasBis.html** respectivamente.

### **Criterios de corrección**

La práctica se va a puntuar según los apartados entregados hasta 2,5 puntos. Cada apartado correcto cuenta la siguiente puntuación, R1(0.5), R2(0.5), R3(0.25), R4(0.5), R5(0.25). R6(0.5)

Se tendrán en cuenta los siguientes criterios.

- Obtener los resultados especificados en cada caso y que funciona en la máquina virtual.
- Calidad del software: no repetir código, modularización correcta en funciones, documentar adecuadamente, generalidad de la solución...
- Para conseguir la máxima puntuación hay que realizar todos los apartados de la práctica y usar los tipos de dato diccionario y numpy array.
- Comprobar que se entregan los scripts de python utilizados y las 5 páginas webs.

### **Entrega**

Se debe crear una carpeta llamada “poblacion” que contenga los ficheros de entrada (en una carpeta llamada “entradas”) y los scripts creados por el estudiante R1.py a R6.py. Las funciones comunes se crean en un fichero llamado “funciones.py”. Debe existir un script llamado “**main.py**” que al ejecutarse lance todos los scripts R1.py a R6.py. Las salidas deben ir a “poblacion/resultados” y las imágenes a “poblacion/imagenes”. Hay que hacer una memoria PDF de una página donde aparezcan el nombre

completo del estudiante y se conteste a las siguientes cuestiones:

1. Indica cuantos apartados de la práctica has resuelto.
2. ¿Has organizado tu código en funciones y diferentes scripts?
3. ¿Has utilizado el tipo de dato diccionario?
4. ¿Has utilizado el tipo de dato numpy.array?

Todo se comprime y se entrega en Prado. Importante: antes de realizar la entrega, hay que comprobar que funciona todo en la máquina virtual, que también tenéis en Prado, sin necesidad de instalar ningún módulo de python adicional a los ya existentes. Para mayor claridad, la estructura de los directorios y nombres de los ficheros se muestra a continuación.

```
\---poblacion
|  funciones.py
|  main.py
|  memoria.pdf
|  R1.py
|  R2.py
|  R3.py
|  R4.py
|  R5.py
|  R6.py
|
+---entradas
|  |  comunidadAutonoma-Provincia.html
|  |  comunidadesAutonomas.html
|  |  comunidadesAutonomasBis.html
|  |  poblacionProvinciasHM2010-17.csv
|  |  variacionProvincias2011-17.html
|  |
|  +---comunidadAutonoma-Provincia_files
|  |  |
|  +---comunidadesAutonomas_files
|  |
|  \---variacionProvincias2011-17_files
|
+---imagenes
|  R3.png
|  R3_bis.png
|  R5.png
|  R5_bis.png
|  R5_biss.png
|
\---resultados
    poblacionComAutonomas.html
    poblacionComAutonomasBis.html
    variacionComAutonomas.html
    variacionComAutonomasBis.html
    variacionProvincias.html
```

**Plazo: hasta el 29 noviembre de 2020.**