



## Memoria proyectos innovación 2018/2019

### MEMORIA A LA CONVOCATORIA PARA PROMOVER PROYECTOS Y ACCIONES DE INNOVACIÓN Y MEJORA EN LA UNIVERSIDAD DE MURCIA PARA EL CURSO 2018/2019

#### Nombre identificativo del Proyecto

Obtención y análisis de datos con R

#### Área de conocimiento, materias o ámbitos que abarca el proyecto:

Métodos Cuantitativos para la Economía y la Empresa

#### Datos del Coordinador

#### Datos del coordinador

Nombre: ISABEL Apellidos: PARRA FRUTOS Email: Teléfono: Tipo(PAS|PDI| Estudiante|Titulado): PDI Unidad: MÉTODOS CUANTITATIVOS PARA LA ECONOMÍA Y EMPRESA Centro: FACULTAD DE ECONOMÍA Y EMPRESA

#### Nombre del GID al que pertenece

Estadística en Ciencias Sociales

#### Relación completa de los miembros

	Apellidos y Nombre	NIF	Teléfono	E-mail	Departamento unidad	Facultad	Nombre del GID al que pertenece	Filiación
1	Arnaldos García, Fuensanta				Métodos Cuantitativos para la Economía y la Empresa	Economía y Empresa		
2	Díaz Delfa, M <sup>a</sup> Teresa				Métodos Cuantitativos para la Economía y la Empresa	Economía y Empresa		
3	Faura Martínez, Úrsula				Métodos Cuantitativos para la Economía y la Empresa	Economía y Empresa		
4	Molera Peris, Lourdes				Métodos Cuantitativos para la Economía y la Empresa	Economía y Empresa		
5	Pérez Castejón, Juan José				Métodos Cuantitativos para la Economía y la Empresa	Economía y Empresa		



## Memoria proyectos innovación 2018/2019

	Apellidos y Nombre	NIF	Teléfono	E-mail	Departamento unidad	Facultad	Nombre del GID al que pertenece	Filiación
6	Martínez Rodríguez, Francisco Javier			javierma@um.es	Ciencia Política, Antropología Social y Hacienda Pública	Economía y Empresa		

### Acción

#### CLASE INVERTIDA

#### a.- Descripción del proyecto. (Máximo 900 caracteres)

La cantidad de información de carácter socioeconómico disponible en Internet es cada vez mayor, y la variedad de formatos en las que se puede presentar es también considerable; además, muchos de ellos no son manejables directamente por el usuario medio. Saber cómo acceder a esos datos resulta cada vez más necesario. Para abordar esta temática, hemos elaborado una web didáctica con el software libre R que expone la forma de obtener datos de Internet con ejemplos de aplicación. Dicha web la podemos utilizar en nuestras asignaturas, aunque también puede ser usada por cualquier persona interesada de forma autónoma. Algunos de los aspectos más importantes tratados en el proyecto son: Exportación e importación a R de datos en diferentes formatos (Excel, SPSS...). Acceso a fuentes estadísticas: Banco Mundial, Reserva Federal,... Introducción a diferentes estructuras internas de almacenamiento

#### b.- Project description. (Máximo 900 caracteres)

The amount of socioeconomic information available online is huge, as is the variety of formats in which it is presented. Some of these are not easily manageable for the average user. Knowing how to access these data is increasingly necessary. To address this issue, we have created an educational website with the open source software R that explains how to get data from Internet through examples. This educational website can be used in any of the subjects taught in the degrees at the Faculty of Economics and Business. Moreover, it should also be of interest for people interested in learning for themselves how to get data. Some of the most important issues in the project are: Exporting and importing data in R with different formats (Excel, SPSS...). Access to official databases: World Bank, Federal Reserve Bank, Eurostat. Introduction to internal structures for saving information.

#### c.- Resultados. Se ha de especificar con claridad el proyecto realizado: plan de trabajo, resultados alcanzados, etc. (Máximo 3000 caracteres)

En una primera fase se realizó una puesta en común para decidir la forma de abordar los diferentes aspectos a tratar. Dada la gran cantidad de instituciones de las que se pueden extraer datos en diferentes formatos, se decidió realizar de manera individualizada una búsqueda de paquetes de R que permitieran (1) realizar el proceso de importación/exportación y (2) descargar datos de organismos oficiales (INE, Eurostat, Banco Mundial...). Además, se llevó a cabo una búsqueda de páginas webs con ejemplos de datos para leer en diferentes formatos. Se trabajó con una página colaborativa en la que se iba añadiendo la información recabada sobre los paquetes y las páginas webs de interés. En una segunda fase, se decidió con qué paquetes trabajar y cómo estructurar el contenido a desarrollar. Cada miembro del grupo se centró en uno o varios paquetes, aunque con apoyo del resto del grupo para revisar los materiales elaborados. En una tercera fase se puso en común el trabajo realizado, para depurarlo al máximo, y se decidió la manera de presentar todo el contenido de forma unificada, con objeto de que fuera sencillo encontrar la información que se busca en cada caso particular, sin tener que leer todo el material, con enlaces claros a los distintos puntos de interés. El resultado se ha plasmado en una website construida con RStudio, compuesta por varias páginas web en las que se tratan los aspectos teóricos necesarios y se desarrollan ejemplos prácticos de aplicación. La idea es enlazar esta website con otra website de análisis de datos con R elaborada en el marco de dos proyectos de innovación docente en



## Memoria proyectos innovación 2018/2019

los dos cursos anteriores. El objetivo final conjunto de todos estos proyectos es que nuestros alumnos dispongan de un material integrado que les permita abordar un ejercicio de análisis de datos de forma global, con un mismo paquete estadístico, y conseguir que esto resulte lo más sencillo posible. Además, un trabajo de este tipo puede tener distintos niveles de dificultad, ya que puede ser abordado desde los primeros cursos, en asignaturas de introducción a la estadística, hasta en asignaturas de cursos superiores, como en el trabajo Fin de Grado, en cuyo caso tendrá que ser más completo y elaborado. Todo esto obliga a diseñar cuidadosamente los materiales, de modo que puedan ser integrados en diferentes asignaturas de distintos grados y usados fácilmente de forma autónoma por los estudiantes. Avanzar en esta línea, a través de material diseñado a tal efecto, es una necesidad en el nuevo contexto del modelo enseñanza-aprendizaje en el que nos encontramos inmersos. Además, en este caso se trata de facilitar el trabajo en relación a cuestiones que se abordan en muchas asignaturas pero que, en la mayoría de las ocasiones, no se tratan expresamente en los contenidos de las mismas. Estos materiales serán la base para participar en las XI Jornadas de R que se celebrarán en Madrid y en algún congreso de innovación educativa que aún está por decidir.

### **d.- Summary. Se ha de especificar con claridad el proyecto realizado: plan de trabajo, resultados alcanzados, etc. (Máximo 3000 caracteres)**

An initial brainstorming session determined the issues we wanted to cover in the project. Socioeconomic information is provided in a wide variety of formats. Moreover, a large number of institutions supplies valuable data using diverse formats. Hence, we decided to seek out and explain the use of R packages to (1) aid the exporting/importing process and (2) download data from official sources (INE, Eurostat, World Bank, etc.). We also looked for specific and interesting examples of how to download other information. We agreed to use an educational website to sum up the information about the different R packages and the institutional websites that could be interesting for our students. Next, we chose the R packages to be used and the structure to follow when explaining each them. Each member of the group focused on one or more R packages and one or more institutional websites, although with the support of the whole group when reviewing. Thirdly, we shared out the work to correct any errors found, and we decided on how to present all the information in an homogeneous way to help with the search for particular information without having to go through all the material. We also highlighted useful links to the different points of interest. The final result is an educational website obtained using RStudio, which includes several web pages with guidelines for downloading data with R through examples. This website will be linked to that of the multivariate data analysis with R built in the educational innovation projects undertaken in the two previous academic years. The ultimate goal of all these educational innovation projects is to provide our students with a complete framework to solve a data analysis problem in a simple, overall manner. This type of analysis can have different levels of difficulty, as it can be undertaken in the first years of the degree, in subjects like ¿Introduction to Statistics¿ or in more complex subjects as the ¿Final Degree Project¿ (TFG) where the analysis will be more complete. We have kept in mind these different levels of difficulty when creating the materials to allow their inclusion in subjects of several degrees and foster autonomous student learning. Advancing in this line, through specifically designed materials, is a necessity in the new context of the teaching-learning model in which we are immersed. Furthermore, it is a question of facilitating work that relates to the issues addressed in many subjects but which, in most cases, are not expressly dealt with in the contents. We firmly believe that this is the course to follow; creating materials for autonomous study is a necessity in today¿s teaching and learning. We must also facilitate data analysis in a variety of subjects for all the degrees taught in our Faculty. These materials will be the base for participating in the ¿XI Jornadas de usuarios de R¿ in Madrid in November 2019 and also in some educational innovation conferences, yet to be determined.

### **e.- Problemas detectados. Se ha de especificar los problemas o limitaciones detectados durante la elaboración del proyecto. (Máximo 3000 caracteres)**

Uno de los principales problemas con los que nos hemos encontrado ha sido la gran cantidad de posibilidades para poder importar/exportar desde el software R, con diferentes paquetes o funciones según el formato del fichero que se quiera leer o escribir, lo que nos ha obligado a decidir qué alternativas serían las más interesantes para los alumnos de Ciencias Sociales a los que, en principio, va dirigido este material.

### **f.- Problems detected. Se ha de especificar los problemas o limitaciones detectados durante la**



## Memoria proyectos innovación 2018/2019

### elaboración del proyecto. (Máximo 3000 caracteres)

One of the main problems we have found is that there are so many ways to import/export data in R, with different packages and commands, depending on the format of the file we want to read or write. This has meant that we have had to opt for a format that we believe best addresses the needs of the students for whom this material has been created, i.e., students in Social Sciences.

### g.- Recomendaciones futuras. (Máximo 3000 caracteres)

Estos materiales elaborados para favorecer el trabajo autónomo del alumnado aún no han sido incorporados en las tareas de nuestras asignaturas, por lo que es necesario testar cómo responden los estudiantes para detectar las fortalezas y debilidades del trabajo desarrollado así como posibles mejoras que podrían aumentar su utilidad. El próximo curso se van a confeccionar actividades en alguna asignatura de primero, para que los alumnos trabajen con la importación/exportación de datos. También sería deseable poder contar con alumnos internos que pudieran explotar más a fondo este trabajo y que nos ofrecieran un feedback de las virtudes y defectos encontrados. Además, sería particularmente interesante trabajar con el website en el contexto del Trabajo Fin de Grado, que es donde se requiere un mayor trabajo autónomo por parte del alumno.

### h.- Future recommendations. (Máximo 3000 caracteres)

These materials designed to encourage the autonomous work of students have not yet been incorporated into the tasks of our subjects, so it is necessary to test how students respond to detect the strengths and weaknesses of the work developed as well as possible improvements that could increase its usefulness. The next course some activities will be made in some subject of first, so that the students work with the import / export of data. It would also be desirable to be able to count on internal students who could exploit this work more thoroughly and offer us a feedback on the virtues and defects found. In addition, it would be particularly interesting to work with the website in the context of the Final Degree Project, which is where more independent work by the student is required.

### i.- Memoria final detallada del proyecto de innovación docente. (Extensión máxima de 4000 caracteres)

Este proyecto se enmarca en la línea de anteriores proyectos de innovación docente, cuyo objetivo general es acercar a los estudiantes una herramienta de análisis estadístico tan potente y versátil como es R. En esta convocatoria hemos elaborado un website que reúne, en forma de páginas webs organizadas, la ayuda necesaria para aprender a descargar datos de diferentes organismos con información de carácter socioeconómico, así como importar ficheros de datos en una gran variedad de formatos. Actualmente hay una gran disponibilidad de datos en distintas fuentes (Banco Mundial, INE, Eurostat, Reserva Federal, etc.) y con diversos formatos, algunos de los cuales no son de manejo habitual por parte del usuario medio. R también puede acceder y descargar la información almacenada en múltiples bases de datos, así como leer o importar directamente ficheros en distintos formatos o desde el portapapeles. Para ello, basta con conocer el paquete que permita realizar esta operación, instalarlo, cargarlo y utilizar la función que importará los datos. Se han elaborado un conjunto de páginas web (organizadas en un website) sobre cómo importar/exportar datos para cada formato de fichero (csv, Excel, SPSS, STATA, SAS, EViews, json, PC-Axis, Google Sheets) así como desde el portapapeles. También hemos incluido una gran variedad de bases de datos, pero que más adelante se podrían ampliar para dar cobertura a lo más utilizado por los estudiantes de los grados de ADE, Economía, Marketing y Sociología, como: 1. INEbase es la base de datos que proporciona el Instituto Nacional de Estadística (INE). Se pueden descargar en R datos de INEbase a través de varios paquetes como son `¿pxR¿` (importar ficheros tipo Pc-Axis), `¿MicroDatosEs¿` (importar los microdatos de algunas encuestas como el censo y la EPA) e `¿INEbaseR¿` (proporciona acceso directo a la API del INE). 2. Eurostat es el portal de datos abiertos de la Unión Europea. Es posible conectarse a su API directamente desde R con el paquete `¿eurostat¿` o descargarse ficheros en distintos formatos para después importarlos a R con un paquete genérico como `¿rio¿`. 3. El Banco Mundial elabora una base de datos temporales comparables entre países que abarcan una gran variedad de indicadores de desarrollo mundial: educación, medio ambiente, género, salud, población, pobreza, tecnología. Además, proporciona la posibilidad de descargar microdatos, dando acceso a los datos recopilados a través de encuestas por muestreo de hogares, establecimientos comerciales u otras instalaciones. Los datos se pueden descargar en distintos formatos (excel, csv, txt con tabuladores) para después importarlos



## Memoria proyectos innovación 2018/2019

a R, pero también hay disponibles varios paquetes de R que permiten el acceso a la API del Banco Mundial como `¿wbstats¿` y `¿WDI¿`. 4. Portal de datos económicos de la Reserva Federal. Ofrece la posibilidad de descargar los datos en dos formatos (.xls y .csv) y el gráfico en tres (.PNG, .pptx y .pdf). Explicamos como descargar directamente en R sus datos con el paquete `¿quantmod¿`. 5. En Oanda se dispone de información sobre precios de monedas (dólar, euro, bitcoin¿), materias primas (maíz, soja, azúcar, trigo, petróleo crudo brent, cobre, gas natural¿) y metales preciosos (oro, plata, platino, paladio). Se da un ejemplo de cómo importar datos desde Oanda con el paquete `¿quantmod¿`. 6. Yahoo Finanzas. Es posible descargar la información financiera incluida en esta web mediante el paquete `¿quantmod¿`. Toda esta información se ha reunido y esquematizado en un website con ejemplos de aplicación para que el estudiante sea capaz de encontrar lo que necesita fácilmente. Con este website se acerca a los estudiantes la posibilidad de usar datos reales lo que puede suponer una motivación mayor. De forma temporal el material elaborado está a disposición de la comisión en [https://www.um.es/estadempresa/visualizacion\\_provisional\\_18\\_19/](https://www.um.es/estadempresa/visualizacion_provisional_18_19/)

**El/la solicitante firma la presente memoria final.**

SI