

ACCIÓN FORMATIVA PARA DOCTORADO EN CIENCIAS ECONÓMICAS Y EMPRESARIALES
UNIVERSIDAD DE ALMERÍA

Título: Herramientas informáticas para el análisis de datos y presentación de resultados (Parte 1).

Docente:

Román Salmerón Gómez (romansg@ugr.es), profesor titular de Universidad del Departamento de Métodos Cuantitativos para la Economía y la Empresa de la Universidad de Granada (UGR). Desde el curso 2013-14 hasta la actualidad imparte la asignatura “Herramientas Informáticas para la Investigación” del máster universitario en Técnicas Cuantitativas en Gestión Empresarial de la UGR. Los contenidos de dicha asignatura versan sobre el sistema de composición tipográfica LaTeX y el entorno de programación R.

Contenidos y estructura:

- **BLOQUE 1:** Entorno de programación R.
 - **Sesión 1 (4 horas):** Introducción al entorno de programación R. Descarga, instalación y primeros pasos en R/RStudio: archivos R Script, objetos, manipulación de objetos, cálculos sencillos, leer datos de un archivo, guardar datos, generar datos de forma aleatoria, listas y hojas de datos, algunos gráficos simples. Librerías (dplyr).
 - **Sesión 2 (4 horas):** Ejecución condicional (comando if) y ciclos (comandos for, repeat y while). Escritura de nuevas funciones. Ejemplos prácticos.
 - **Sesión 3 (4 horas):** Creación de paquetes en R. Publicación en R-Force, Github y CRAN.
- **BLOQUE 2:** Análisis de datos con el entorno de programación R.
 - **Sesión 4 (4 horas):** Análisis descriptivo de un conjunto de datos e inferencia en una y dos poblaciones. Aleatoriedad y normalidad.
 - **Sesión 5 (4 horas):** Modelo de regresión lineal simple/múltiple. Modelos de datos de panel.
 - **Sesión 6 (4 horas):** Modelos de variable dependiente binaria (logit/probit).

Cronología (24 horas en 6 sesiones de 4 horas):

- **Semana 1:**
 - **Día 06/03/2024:** mañana sesión 1 de 10:00h a 14:00h, tarde sesión 2 de 16:30h a 20:30h.
 - **Día 07/03/2024:** mañana sesión 3 de 9:30h a 13:30h.
- **Semana 2:**
 - **Día 13/03/2024:** mañana sesión 4 de 10:00h a 14:00h, tarde sesión 5 de 16:30h a 20:30h.
 - **Día 14/03/2022:** mañana sesión 6 de 9:30h a 13:30h.

Contenidos:

• Sesión 1:

- Enlace a video tutorial que explica la descarga e instalación de R en Windows: <https://drive.google.com/file/d/14HNeLUCW8yIDuJSK6totTehC9e2AaszK/view?usp=sharing>.
- Enlace a video tutorial que explica la descarga e instalación de RStudio en Windows: <https://drive.google.com/file/d/14HNeLUCW8yIDuJSK6totTehC9e2AaszK/view?usp=sharing>.
- Algunos objetos básicos: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/01_objetos_basicos/. Aproximadamente 30 minutos.
- Vectores: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/02_vectores/. Aproximadamente 30 minutos.
- Matrices: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/03_matrices/. Aproximadamente 30 minutos.
- Lectura de datos: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/04_lectura_datos/. Aproximadamente 30 minutos.
- Datos aleatorios y guardar datos: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/05_datos_aleatorios/. Aproximadamente 30 minutos.
- Librería *dplyr*: documentación del paquete en la web oficial CRAN <https://cran.r-project.org/web/packages/dplyr/index.html> y algunos ejemplos de interés en <https://rsanchezs.gitbooks.io/rprogramming/content/chapter9/dplyr.html> y <https://rpubs.com/joser/dplyr/>. Aproximadamente 30 minutos.

• Sesión 2:

- Ejecución repetitiva de comandos: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/06_bucles/. Ejercicio propuesto: crear un bucle que me permita jugar al juego de las siete y media. Aproximadamente 1 hora y 30 minutos.
- Creación de funciones: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/07_funciones/. Ejercicio propuesto: crear una función que me permita hacer estadística descriptiva unidimensional. Aproximadamente 2 horas.

• Sesión 3: Descargar <https://github.com/rnoremlas/documentos/tree/7ce0ae17a22153f4f06382edfc5a04e455490e49/2023.06.04%20Curso%20Almer%C3%ADa>.

- Hace falta instalar MikTeX completo: descarga (<https://drive.google.com/file/d/16XCdMgco-mSHtBgi4qa-IK7yFDMVLZxE/view?usp=sharing>) e instalación de programas (<https://drive.google.com/file/d/16U7cQM8-qOosNRXPQF1gWvGflymn6L9u/view?usp=sharing>) relacionados con LaTeX.

• Sesión 4:

- Análisis descriptivo de un conjunto de datos:
 - Estadística descriptiva unidimensional: algunas representaciones gráficas https://rnoremlas.quarto.pub/un_rincon_para_r/posts/08_descriptiva/ y principales características unidimensionales https://rnoremlas.quarto.pub/un_rincon_para_r/posts/09_descriptiva/. Aproximadamente 1 hora.
 - Estadística descriptiva bidimensional: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/10_descriptiva/. Aproximadamente 1 hora.
- Probabilidad: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/11_probabilidad/. Aproximadamente 20 minutos.
- Inferencia en una y dos poblaciones: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/12_inferencia/. Aproximadamente 1 hora.
- Aleatoriedad y normalidad: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/12_inferencia/. Aproximadamente 20 minutos.

• Sesión 5:

- Modelo de regresión lineal simple: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/13_regresion/. Aproximadamente 30 minutos.
- Modelo de regresión lineal múltiple: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/14_regresion/. Aproximadamente 1 hora y 30 minutos.

- Apoyo teórico: introducción a la Econometría (<https://hdl.handle.net/10481/89703>), modelo lineal general (<https://hdl.handle.net/10481/89706>), multicolinealidad (<https://hdl.handle.net/10481/89709>), heterocedasticidad (<https://hdl.handle.net/10481/89716>), autocorrelación (<https://hdl.handle.net/10481/89717>), formulario (<https://hdl.handle.net/10481/89720>).
 - Modelos de datos de panel: https://rnoremlas.quarto.pub/un_rincon_para_r/posts/21_datos_panel/. Aproximadamente 2 horas.
 - Apoyo teórico: https://www.ugr.es/%7Eromansg/material/WebEco/04-Eco2/Transparencias/04_Panel.pdf.
- **Sesión 6:**
 - Modelos de variable dependiente binaria (logit/probit): https://rnoremlas.quarto.pub/un_rincon_para_r/posts/20_logit_probit/.
 - Apoyo teórico: https://www.ugr.es/%7Eromansg/material/WebEco/04-Eco2/Transparencias/03_Discreta.pdf.

Material de apoyo:

- Fundamentos de ciencia de datos con R: <https://cdr-book.github.io/>.
- GitHub de Román Salmerón: <https://github.com/rnoremlas/documentos.git>.
- Web de Román Salmerón: <https://www.ugr.es/~romansg/web/docencia.html>.
- Perfil en DIGIBUG de Román Salmerón: <https://digibug.ugr.es/browse?authority=2681&type=author>.