



PROMiDAT

IBEROAMERICANO

Programa Iberoamericano de
Formación en Minería de Datos

CD300

Fundamentos de Programación en R



(506) 4030.1205 - (506) 4030.1114



info@promidat.com



facebook.com/promidat



Twitter.com/promidat



www.promidat.com

Índice

Duración	2
Descripción	2
Objetivos	2
Metodología	2
Contenido	3
Evaluación	4
Bibliografía	4

Duración

Cuatro semanas.

Descripción



En este curso se presentarán desde cero los principales conceptos de programación y su aplicación en lenguaje R, enfocado al ámbito de la ciencia de datos.

Objetivos

En este curso el estudiante será capaz de:

1. Crear scripts de R.
2. Comprender el proceso de importación de datos en formato **csv** y **xlsx**.
3. Utilizar funciones estadísticas de R.
4. Crear soluciones a problemas personalizados creando sus propias funciones.
5. Utilizar Rstudio como herramienta para el desarrollo de análisis de datos.

Metodología

Basado en la teoría y en la aplicación directa de los conceptos aprendidos. Para esto se dispondrán de las siguientes herramientas:

1. Una vídeo conferencia semanal, las cuales quedarán grabadas en Zoom, para que los alumnos la puedan acceder en cualquier momento.
2. Trabajos prácticos semanales.
3. Foros para plantear dudas al tutor y compañeros.
4. Aula virtual en Moodle.

Luego de este curso el estudiante será capaz de:

Desarrollar scripts en lenguaje R que permiten un análisis de datos básico.

Contenido

1. Instalación y preparación del entorno de trabajo.

- a. Instalar RStudio.
- b. Instalar R.
- c. Instalaciones complementarias para Windows.
- d. Instalaciones complementarias para MacOS.

2. Introducción a los diagramas de flujo estructurados

- a. Esquema general de especificación.
- b. Modulación y bloque fundamental.
- c. Seguimiento de un diagrama de flujo.
- d. Estructuras fundamentales en un diagrama de flujo.
- e. El bloque de dos opciones (sí).
- ii. El bloque de opciones múltiples (case).
- iii. Condiciones compuestas.
- iv. El bloque de entrada condicionada (mientras).
- v. El bloque de entrada asegurada (hasta que).

3. Tipos de datos.

- a. Vectores Atómicos.
- b. Matrices.
- c. Listas.
- d. Tablas de datos.

4. Lectura de datos.

- a. Lectura de datos csv y xlsx.
- b. Tipos de datos y validación de archivos.
- c. Modificar tipos de datos de un archivo.
- d. Predicción mediante los modelos de Holt Winters.

5. Estructuras de control.

- a. IF - ELSE.

-
- b. Ciclos FOR.
 - c. Ciclos While.

6. Funciones.

- a. Estructura de una función.
- b. Buenas prácticas para el desarrollo de funciones.
- c. Manejo de errores.
- d. Documentación de una función.

Evaluación

El curso se evalúa con 4 tareas, una por semana, cada tarea tiene un valor de 25 puntos. La nota mínima de aprobación es de 70.

Bibliografía

- [1] J.J. Allaire et al. *Quarto*. DOI: 10.5281/ZENODO.5960048. Zenodo, Jan. 10, 2022. DOI: 10.5281/ZENODO.5960048. URL: <https://zenodo.org/record/5960048>.
- [2] John M. Chambers. *Software for data analysis: programming with R*. Statistics and computing. OCLC: ocn191243189. New York: Springer, 2008.
- [3] chocolatey. *Chocolatey Software Docs | Chocolatey - Software Management for Windows*. 2022. URL: <https://docs.chocolatey.org/en-us>.
- [4] Colin Gillespie and Robin Lovelace. *Efficient R programming: a practical guide to smarter programming*. First edition. OCLC: ocn964820738. Sebastopol, CA: O'Reilly Media, Inc, 2016.
- [5] Norman S. Matloff. *The art of R programming: tour of statistical software design*. OCLC: ocn711045702. San Francisco: No Starch Press, 2011.
- [6] Hadley Wickham. "Tidy Data". In: *Journal of Statistical Software* 59 (Sept. 12, 2014), pp. 1–23. DOI: 10.18637/jss.v059.i10. URL: <https://doi.org/10.18637/jss.v059.i10>.
- [7] Hadley Wickham and Garrett Grolemund. *R for data science: import, tidy, transform, visualize, and model data*. First edition. OCLC: ocn968213225. Sebastopol, CA: O'Reilly, 2016.