



Universidad
del Valle

Facultad de Ingeniería

Programa Académico de Estadística

Información básica

Curso	761017M - Prcesamiento de datos	Validable:	No
Profesores	Andrés Ochoa / Omar Rios	Habilitable:	No
e-mail	andres.ochoa@correounalvalle.edu.co omar.rios@correounalvalle.edu.co	Periodo:	febrero - junio 2026
		Créditos	3

Objetivos del curso

1. Dar al estudiante la posibilidad de utilizar herramientas computacionales que le permitan extraer información estadística a un conjunto de datos.
2. Que el estudiante adquiera los conocimientos fundamentales del manejo de herramientas computacionales para análisis estadístico y presentación de informes, tales como Excel, R, Python, Latex, Github, entre otros.
3. Adquirir la habilidad necesaria para trabajar con registros de información bajo diferentes formatos. Iniciar al estudiante en el proceso de conformación de informes descriptivos y resúmenes ejecutivos.

Contenido

1. Conceptos preliminares:
 - La metodología del análisis estadístico, validez y confiabilidad, Escalas de Medición.
 - Concepto de dato y base de datos.
 - Concepto de paquete estadístico.
 - Recodificación de una base de datos.
2. Herramientas de edición y presentación de informes:
 - Informes y presentaciones en Latex y Overleaf
 - Presentaciones a partir de Github y Git-Tea
3. Herramientas de Análisis de datos con Excel:
 - El Entorno Excel, Formulas y Funciones, Referencias Relativas, Operaciones con rangos, Ingreso de datos, herramientas de validación, ingreso de datos desde un formulario Excel.
 - Filtros y Autofiltros.
 - Funciones condicionales.
 - Generación de números aleatorios.
 - Módulo Análisis de Datos.
 - Gráficos con Excel.
 - Tablas y gráficos dinámicos.

4. Análisis de datos en R:
 - El entorno R, Objetos, Tipos de objetos, vinculación de librerías desde R, lectura de registros desde R, entorno gráfico de R, Estadística Básicas desde R.
 - Programación en R.
 - Otras utilidades en R: rmarkdown (flexdashboard), sweave, shiny.

5. Análisis de datos en Python.
 - Introducción a Python, carga de datos, creación y edición de objetos.
 - Estadísticas descriptivas, Herramientas gráficas: matplotlib y seaborn.

Texto

Notas de clase.

Evaluación

1. Trabajos en clase y tareas: 25%
2. Laboratorios (Informe y presentación): 75%

Bibliografía

- Géron, A. (2023). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (3rd ed.). O'Reilly Media.
- Grolemund, G. (2014). Hands-on programming with R: Write your own functions and simulations. O'Reilly Media.
- Maindonald, J., & Braun, W. J. (2010). Data analysis and graphics using R: An example-based approach (3rd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511841057>
- Matloff, N. (2011). The art of R programming: A tour of statistical software design. No Starch Press.
- Raschka, S., & Mirjalili, V. (2022). Python Machine Learning (4th ed.). Packt Publishing.
- VanderPlas, J. T. (2023). Python data science handbook: Essential tools for working with data (2.^a ed.). O'Reilly Media
- Wickham, H., & Grolemund, G. (2017). R for data science: Import, tidy, transform, visualize, and model data. O'Reilly Media. <https://r4ds.hadley.nz/>
- Wickham, H. (2019). Advanced R (2nd ed.). CRC Press. <https://adv-r.hadley.nz/>