

1. Información del equipo pedagógico y horarios de atención a estudiantes

Profesor magistral: Juan Carlos Tarapuez
Correo electrónico: j.tarapuezr@uniandes.edu.co
Horario de atención a estudiantes: Por definir.
Lugar de atención a estudiantes: Por definir.

Profesor magistral: Cristhian Julian Acosta Pardo
Correo electrónico: cj.acosta@uniandes.edu.co
Horario de atención a estudiantes: Jueves de 4:00 pm a 5:00 pm.
Lugar de atención a estudiantes: Vía [Zoom](#)

2. Descripción del curso

El curso Taller de R busca que los estudiantes adquieran una comprensión general de lenguaje de R y su utilidad para el análisis de datos, la investigación económica y otras aplicaciones académicas y profesionales. El énfasis será sobre la importación, organización y manejo de datos y la presentación de resultados, a través de herramientas que permitan una comunicación efectiva.

El curso no profundiza en conceptos estadísticos o econométricos; en cambio, busca que los estudiantes entiendan la estructura general de programación en R y la complementen con los conocimientos adquiridos en otras clases. Al final se espera que los estudiantes se sientan cómodos utilizando R para resolver problemas relacionados con análisis de datos; asimismo, se espera que identifiquen las diferentes aplicaciones de R y puedan continuar utilizándolo.

3. Resultados de aprendizaje

- ✓ Entender los archivos de ayuda sobre paquetes de R.
- ✓ Escribir funciones básicas.
- ✓ Escribir códigos para solucionar problemas de importación y limpieza de datos.
- ✓ Manipular datos para obtener información relevante.
- ✓ Presentar resultados estadísticos de forma adecuada.
- ✓ Elaborar documentos HTML y PDF.

4. Cronograma

Módulo 1: Introducción a R

Semana 1:

- ¿Qué es R? ¿Por qué R?
- Instalación de R y R Studio. Instalación y carga de paquetes.
- Paneles de R Studio y *scripts*.
- Personalización de R Studio.
- Conceptos básicos de sintaxis en R.
- Tipos de datos: entero, numérico, lógico, *character*. Missings: NaN, NA.
- Estructuras de datos: vectores, matrices, listas, *arrays*, *data frames*.
- Operaciones básicas.
- Funciones relevantes para la identificación y conversión de tipos de objetos y datos (coerción).
- ¿Dónde buscar ayuda?

Módulo 2: Funciones y *loops*

Semana 2:

- Asignar datos a objetos.
- Extraer información de objetos.
- Uso de vectores lógicos para extraer información de objetos.
- Comando "*function*".
- Uso de condicionales: *if*, *else*, *ifelse*.

Semana 3:

- Depuración de errores en funciones: *debug*, *undebug*, *traceback*.
- Inclusión de controles en las funciones: *warning*, *stop*.
- Loops: *while*, *for*, *repeat*.

Semana 4:

- Controles de flujo: *next*, *break*.
- Funciones *apply*, *sapply*, *tapply*, *lapply*.
- Instalar y cargar paquetes.
- Función *ddply*
-

Módulo 3: Bases de datos

Semana 5:

- Manejo básico *data frames*: selección de filas y/o columnas, selección condicionada.
- Creación de muestras y subconjuntos de *data frames*: *subset*, *sample_n*.
- Importación de datos: *dta*, *sav*, *csv*, *xlsx*.
- Inspección inicial de la información.

Semana 6:

- *Raw data vs tidy data*.
- Organización y limpieza de bases de datos con los paquetes *tidyr* y *dplyr*.
- Análisis descriptivo.
- Exportación de bases de datos: *dta*, *csv*, *xlsx*, *txt*.

Semana 7:

- Taller en clase #1: importación de base de datos *raw* y exportación de un archivo organizado (*tidy*).

Módulo 4: Visualización de datos

Semana 8:

- Gráficos del paquete base: histogramas, *scatter plot*, *bar plot*, *box plot*.
- Introducción a *ggplot2*.

Semana 9:

- Relación entre la organización de los datos y *ggplot2*: *dplyr* y *ggplot2*.
- Personalización de gráficos: ejes, colores, títulos.
- Exportación de gráficos.
- Otras herramientas para visualización de datos.

Módulo 5: Análisis espacial de datos

Semana 10:

- Análisis exploratorio de datos espaciales.
- Autocorrelación espacial (I de Moran).
- Métodos de corrección por autocorrelación espacial.

Semana 11:

- Taller en clase #2: ejercicio de visualización de datos y análisis espacial.

Módulo 6: Otras aplicaciones en R

Semana 12:

- Regresión lineal.
- Métodos de regresión para variables dicótomas.
- Aplicación en series de tiempo: modelos autorregresivos.

Semana 13:

- Introducción a *webscrapping*: *rvest* (nociones básicas)
- Introducción a *text mining*: *tm*, *tidytext*. (nociones básicas)

Módulo 7: RMarkdown y Shiny

Semana 14:

- Introducción a RMarkdown.
- Elaboración de documentos PDF.
- Elaboración de documentos HTML, utilizando Shiny.

Semana 15:

- Análisis de componentes principales.
- Análisis de cluster.

Semana 16:

- Taller final en clase: importación y limpieza de datos, análisis descriptivo, visualización de datos y presentación de conclusiones utilizando RMarkdown.

5. Referencias

- R Development Core Team. *R-Manuals*. <https://cran.r-project.org/manuals.html>.
- Wiley, M, Wiley, J. *Advanced R Data Programming and the cloud*. Apress. 2016
- Wickham, H. *Advanced R. The R Series*. CRC Press. 2015.
- S.P. Paul. *Introductory Time Series with R*.
- Chang Winston. *R graphics Cookbook*. O'Reilly.
- Matthias Kohl. *Introduction to statistical data analysis with R*.
- Shiny. <https://shiny.rstudio.com>.

6. Metodología

El curso se realizará de forma 100% presencial. Las sesiones estarán divididas en dos partes; durante la primera parte el profesor introducirá técnicas, comandos y conceptos relacionados con la utilización de R y utilizará ejemplos que lo ilustren; durante la segunda parte los estudiantes deberán trabajar individualmente en un ejercicio, el cual deberán entregar hasta más tardar el sábado a la media noche. Se calificarán todas las tareas que realicen, aunque no se tendrá en cuenta una de ellas en la calificación de tareas del semestre.

Adicionalmente, en el semestre se realizarán tres (3) talleres en clase sobre temas que hayan sido cubiertos; estos talleres deben ser desarrollados por los estudiantes durante la clase, sin embargo, tendrán hasta la medianoche del sábado de la semana siguiente para entregarlos.

El curso tiene siete (7) módulos, los cuales buscan iniciar al estudiante en el uso de R y presentar algunas aplicaciones, que motiven su continuo aprendizaje y utilización de R durante el curso y después de haberlo terminado.

7. Evaluaciones

El principal criterio para la evaluación es los tres talleres que se realizarán en clase. La calificación final será determinada de la siguiente manera:

Taller en clase #1: 30%

Taller en clase #2: 25%

Taller final en clase: 30%

Ejercicios en clase: 15%

Los talleres en clase se evaluarán teniendo en cuenta: 1) Organización del código/*script*, 2) Que el código genere un resultado correcto y 3) Número de errores en el código. Los talleres deben ser presentados a través de Bloque Neón. Para aprobar el curso cada estudiante debe alcanzar una nota igual o superior a 3.00. Las calificaciones definitivas de las materias serán numéricas de uno punto cinco (1,50) a cinco punto cero (5,00), en unidades, décimas y centésimas.

Justificación de ausencia: Según el artículo 45 del Reglamento General de Estudiantes de Pregrado, los estudiantes tendrán ocho días hábiles para presentar una excusa válida

y, de ser aceptada, el profesor programará el supletorio en las dos semanas siguientes. Tenga en cuenta que en el art. 45 se especifican un mínimo de excusas válidas, pero también se especifica que “[e]l profesor podrá tener en cuenta otras circunstancias que a su criterio puedan justificar la ausencia del estudiante”.

Reclamos: Según los artículos 64, 65 y 66 del Reglamento general de estudiantes de pregrado, el estudiante tendrá cuatro días hábiles después de la entrega de la evaluación calificada para presentar un reclamo. El profesor responderá al reclamo en los cinco días hábiles siguientes. Si el estudiante considera que la respuesta no concuerda con los criterios de evaluación podrá solicitar un segundo calificador al Consejo de la Facultad en los cuatro días hábiles posteriores a la recepción de la decisión del profesor.

Fraude: Se considera fraude cuando, en una actividad, dos o más estudiantes presentan el mismo código, la secuencia de comandos es idéntica y no es posible identificar el proceso de aprendizaje de cada uno. En caso de detectar un fraude se seguirán los lineamientos establecidos en el artículo 116 del [Reglamento General de Estudiantes de Pregrado](#).

8. Asistencia

Según los artículos 43 y 44 del Reglamento General de Estudiantes de Pregrado (RGEPr), los estudiantes pueden no asistir hasta al 20% de las sesiones. Si la inasistencia supera dicho umbral el profesor seleccionará la nota de ejercicio de clase más alta para ser reemplazada por 0.

9. Políticas generales de los cursos de Economía y fechas importantes

Los estudiantes deben consultar [este enlace](#), donde se encuentran las reglas sobre asistencia a clase, excusas válidas, fraude académico y faltas disciplinarias, reclamos, políticas de bienestar y fechas importantes del semestre.