



PROGRAMA DE CURSO
Programación para Estadística Computacional I
XS-2210

Características del curso

Horas:	4 práctica
Créditos:	4
Requisitos:	MA1021; XS1130
Correquisitos:	
Ciclo:	I-2024
Clasificación:	propio
Modalidad:	Presencial
Horario:	Martes 5:00 pm a 8:50 pm
Consulta:	Miércoles 5:00 pm a 7:00 pm
Profesor:	Lic. Michael Sánchez Soto

Descripción

Este curso está dirigido a estudiantes de la carrera del Bachillerato en Estadística para que aprendan a programar en un lenguaje estadístico, así como el uso de control de versiones para el registro de cambios en el código y el trabajo colaborativo. El curso está orientado a adquirir destrezas en conceptos básicos de programación, manipulación de archivos de datos, realizar análisis básicos de estadística descriptiva y visualizarlos adecuadamente. Se trata de un curso que hace un uso intensivo del computador para habituar al estudiantado al ambiente de programación en R realizando trabajos prácticos grupales.

Objetivo general

Desarrollar destrezas en programación para manipular datos, realizar análisis básicos de estadística descriptiva y visualizarlos adecuadamente haciendo uso de control de versiones de código.



Objetivos específicos

Al finalizar el curso el/la estudiante estará en la capacidad de:

1. Utilizar software colaborativo para facilitar la programación en R.
2. Aplicar los conceptos básicos de programación para utilizarlos en el lenguaje R.
3. Aplicar funciones de las principales librerías de R en situaciones prácticas.

Contenidos

1. Aspectos básicos de la programación en R:
 - a) Conceptos básicos de lenguajes de programación.
 - b) Representación de algoritmos: diagramas de flujo.
 - c) R y RStudio, instalación y usos básicos.
 - d) Configuraciones globales de RStudio.
 - e) Uso básico de archivos de extensión .R y .Rmd y su relación con otros lenguajes de programación.
 - f) Tipos de variables.
 - g) Operadores de asignación, aritméticos, lógicos, relacionales.
 - h) Estructuras de datos: vectores, matrices, arreglos, listas, conjuntos de datos.
 - i) Indexación y sustitución.
 - j) Estructuras de control iterativas (*for* y *while*).
 - k) Estructuras de control condicionales (*if*, *else*, *ifelse*).
 - l) Familia de funciones *apply*.
 - m) Diferencias entre mensajes, advertencias y errores.
2. Otros procedimientos para programar con R:
 - a) Uso de librerías y funciones de R y el conflicto entre librerías.
 - b) Creación de funciones personalizadas en R.
 - c) Lectura y escritura de datos: .csv, .xlsx, .sav, .txt.
 - d) Uso de la librería *tidyverse*.
 - e) Procesamiento y manipulación de un *data.frame* con *dplyr*: selección, filtrado, agrupamiento, cálculo y modificación de variables.



- f) Procesamiento y manipulación de un data.frame con tidy: manipulación de formatos largo y ancho en un conjunto de datos.
 - g) Procesamiento y manipulación de datos temporales con lubridate.
 - h) Operadores de uso común: %>%, %in% y %nin%
 - i) Unión de datos con las funciones left_join(), right_join(), inner_join(), anti_join(), full_join() de la librería dplyr y con la función do.call().
3. Visualización de datos:
- a) Construcción de gráficos con funciones base de R.
 - b) Construcción de gráficos por capas con ggplot2.
4. Análisis de datos:
- a) Medidas de posición y variabilidad en R.
 - b) Medidas de asociación y correlación en R.
 - c) Distribuciones de probabilidad en R: forma de sus funciones (p_, r_, q_).
 - d) Generación de variables aleatorias.
5. Versiones de control con Git:
- a) Control de cambios con Git.
 - b) Flujos de trabajo con Git.
 - c) Colaboración con Git.
 - d) Diferencias entre Git y Github.
6. Productos de datos: Generación de reportes de extensión .html, .docx y .pdf.

Metodología

Este es un curso fundamentalmente práctico. Los contenidos del curso serán abordados en clase (laboratorio) y cada docente desarrollará actividades formativas para apoyar el proceso de enseñanza-aprendizaje. Para abordar los contenidos del curso el personal docente facilitará espacios para la puesta en práctica de los conocimientos adquiridos, por lo que se aplicará una metodología de un aprendizaje activo, donde cada estudiante debe realizar las actividades propuestas por su profesor o profesora, así como asistir y participar en los espacios de trabajo facilitados. El proceso para apoyar el aprendizaje del estudiantado incluye pruebas cortas y trabajos prácticos.



Evaluación

Rubro	Valor
Quices (4)	20
Tareas programadas (2)	30
Exámenes (2)	40
Proyecto	10

Cronograma

Semana	Fechas	Contenido	Asignaciones
1	12 marzo	Conceptos básicos de computación, algoritmos y lenguajes de programación	
2	19 marzo		Quiz 1
3	26 marzo		
4	2 abril	Introducción a R	Tarea 1 - Asignación
5	9 abril		Quiz 2
6	16 abril		
7	23 abril		Examen 1
8	30 abril	R Intermedio	Tarea 1 - Entrega
9	7 mayo		
10	14 mayo		Tarea 2 - Asignación
11	21 mayo		Quiz 3
12	28 mayo	GIT	Proyecto -Asignación
13	4 junio	Análisis y Visualización de datos	
14	11 junio		Tarea 2 - Entrega
			Quiz 4
15	18 junio	Productos de datos	Proyecto - Entrega
16	25 junio		Examen 2
17	1 julio		Ampliación

Bibliografía

- Aguilar Fernández, E. & Zamora Araya, A. (2020). Introducción a la estadística descriptiva con R. <https://tinyurl.com/2oo9bmwk>



- Pace, L. (2012). Beginning R: An Introduction to Statistical Programming.
<https://tinyurl.com/2qj7tyaz>
- Venables, W.N. & Smith, D. M. (2010). An Introduction to R
<https://tinyurl.com/y9u5nxsy>
- Wickham, H. & Golemund, G. (2018). R para Ciencia de Datos.
<https://es.r4ds.hadley.nz/>

Reglamentación importante adicional

1. Reglamento de Régimen Académico Estudiantil:
https://www.cu.ucr.ac.cr/normativ/regimen_academico_estudiantil.pdf
2. Reglamento de Orden y Disciplina de los Estudiantes de la Universidad de Costa Rica:
https://www.cu.ucr.ac.cr/normativ/orden_y_disciplina.pdf
3. Reglamento de la Universidad de Costa Rica contra el Hostigamiento Sexual:
https://www.cu.ucr.ac.cr/normativ/hostigamiento_sexual.pdf