



INFORMACIÓN GENERAL

Materia	Análisis de Datos con Lenguaje de Programación Estadísticos (Lenguaje R)
Titular / Cotitular	José Arturo Mora Soto / Alejandro García Fernández.
Fecha de elaboración	23-marzo-2015 / 09-abril-2015

INTRODUCCIÓN GENERAL DE LA MATERIA

El análisis de datos es fundamental para un profesional cuya responsabilidad es manejar grandes cantidades de información o *Big Data*. En este curso el alumno desarrollará las habilidades fundamentales de estadística, programación y visualización esenciales para cualquier científico de datos.

A lo largo del curso el alumno aprenderá a desarrollar programas en lenguajes de programación estadísticos, como el Lenguaje R, con la finalidad de desarrollar la capacidad de realizar programas para un análisis de datos efectivo. Además, el alumno aprenderá a configurar el entorno de desarrollo y a trasladar los elementos de control que ya conoce a un lenguaje estadístico. Finalmente, el alumno se expondrá a problemas prácticos de implementación como la importación y organización de datos, el desarrollo de comentarios al código y el importar librerías y crear librerías, entre otros.

Finalmente el alumno aprenderá a hacer representaciones visuales de datos usando librerías específicas, como *lattice* y *ggplot2* para lenguaje R, aplicando principios básicos de representación gráfica de datos para crear visualizaciones proviniendo de conjuntos de datos heterogéneos que permitan mejorar la toma de decisiones.

OBJETIVO GENERAL

El alumno será efectivo en el análisis de datos usando un lenguaje de programación estadístico.

OBJETIVOS PARTICULARES

1. Aprenderá a crear programas con un lenguaje de programación estadístico.
2. Conocerá los modelos estadísticos fundamentales en el diseño de experimentos.
3. Programará modelos matemáticos y estadísticos utilizando un lenguaje de programación estadístico.
4. Implementará visualizaciones de datos que den respuesta a un experimento o demanda de análisis de datos concreta.



TEMARIO

1. Desarrollo de programas con lenguajes estadísticos
 - 1.1. Introducción al lenguaje R
 - 1.2. Tipos de datos y objetos
 - 1.3. Operaciones de lectura y escritura
 - 1.4. Estructuras de control
 - 1.5. Funciones y ciclos
 - 1.6. Reglas de alcance de objetos en el lenguaje R
 - 1.7. Manejo de fechas y tiempo.
 - 1.8. Herramientas de depuración
 - 1.9. Simulación y medición del desempeño del código
 - 1.10. Técnicas de obtención y limpieza de datos
2. Herramientas estadísticas básicas
 - 2.1. Diseño de experimentos y rigor estadístico
 - 2.2. Variables y distribuciones
 - 2.3. Correlación y Medición
 - 2.4. Regresión simple y múltiple
 - 2.5. Comparación de grupos con Prueba-T y ANOVA
 - 2.6. Pruebas paramétricas y no paramétricas (U de Mann-Whitney, Kruskal-Wallis)
3. Visualización de datos
 - 3.1. Fundamentos de Visualización
 - 3.2. Principios de diseño
 - 3.3. Librerías de visualización de datos en el Lenguaje R
 - 3.4. Visualización dinámica de datos (d3.js, Shiny)

BIBLIOGRAFÍA

No.	Título	Autor	Editorial	Año
1	Head First Data Analysis: A learner's guide to big numbers, statistics, and good decisions	Michael Milton	O'Reilly Media	2009
2	Practical Data Science with R	Nina Zumel, John Mount	Manning Publications	2014
3	Visualize This: The FlowingData Guide to Design, Visualization, and Statistics	Nathan Yau	Wiley	2011
4	Software for Data Analysis, Programming with R	John Chambers	Springer	2009
5	Introductory Statistics with R	Peter Dalgaard	Springer	2004

EVALUACIÓN

ASPECTO A EVALUAR	PORCENTAJE
Proyecto de integrador de análisis de datos	25%
Proyecto intermedio de análisis de datos	15%
Tareas y Exámenes rápidos	60%



CENTRO DE
INVESTIGACIÓN EN
MATEMÁTICAS, A.C.

MAESTRÍA EN INGENIERÍA DE SOFTWARE

Plan de Estudios de 2015