

Diplomado Ciencia de los Datos e Internet de las Cosas

Módulo I. Programación R con R Studio.

Objetivo

Construir programas en lenguaje de programación R en entorno R Studio para analizar, interpretar y comunicar datos con enfoque estadístico y probabilístico presentados mediante documentos *markdown* y publicados en la nube con el servicio *RPubs*.

Dirigido a:

- Estudiantes: Reforzar aprendizajes de probabilidad y estadística
- Académicos: Desarrollar habilidades en el uso de herramientas estadística, probabilidad y lenguaje de programación como alternativa de enseñanza.
- Investigadores. Alternativa para difundir trabajos con análisis estadístico de datos.
- Empresa: Aprovechar las ventajas de utilizar R en el ámbito estadístico y probabilístico para el desarrollo de análisis de datos.
- A la comunidad. Personas interesadas en R y R Studio

Requisitos:

- Laptop y/o equipo de cómputo
- Instalar R y R Studio,
- Conectividad a internet
- Conocimientos generales de programación y HTML
- Conocimientos generales de probabilidad y estadística
- Enlaces google meet

Documentación

Temas:

- Reconocimiento de entorno de trabajo
- Programación R
- Estructuras de datos en R
- Funciones del usuario en R
- Visualización de datos
- Archivos markdown y publicación rpubs
- Importar y exportar datos
- Librería dplyr
- Simulación de datos
- Gráficos avanzados ggplot2() y plotly()
- Strings and Dates en R
- Datos semiestructurados
- R shiny

Material del módulo

<https://github.com/rpizarro/CIIT.-Diplomado-en-Ciencia-de-los-Datos-e-IoT>

Github.com/rpizarro/Diplomado Ciencia de los Datos e IoT. Módulo I. Programación R con R Studio

Sesión 01. Ciencia de los datos, entorno R y R Studio

- Instalar R y R Studio
- Entorno R Studio
- Programación R en consola y construcción de scripts
- Tipos de datos, variables, función c()

Sesión 02. Programación R

- Números, operaciones matemáticas y funciones matemáticas, strings, booleanos
- Operadores de asignación `<-` y `->` `=`; operadores matemáticos; operadores lógicos
- Condicionales, Ciclos

Sesión 03. Estructuras de datos en R, funciones

- Creación y llamada de funciones
- Estructuras de datos: vectores, listas, matrices, arreglos, data frames, factores
- Visualización de datos, `plot()`, `bar()`, ...

Sesión 04. Importar y Exportar datos

- Importar y exportar datos `read.csv()`, `write.csv()`
- Importar datos Excel
- Importar datos texto
- Exportar con `write.csv()`

Sesión 05. Estadística descriptiva en R, comunicación de datos con archivos markdown y publicar en R Pubs

- Variable de interés
- Muestra con `sample()`
- Medidas de tendencia central `mean()`, `median()`, `quantile()`, `table()`,
- Medidas de dispersión `var()`, `sd()`
- Función `summary()`
- Frecuencias e histograma, funciones `which()`, `subset()` e `hist()`

- Tablas de frecuencias y su uso; librería `fdth()`

Sesión 06 Documentos markdown, R Cloud y publicar en Rpubs

- Muestra con `sample()`
- Crear archivo markdown HTML
- Crear archivos bloc de notas
- Acceder a R Studio Cloud
- Crear documentos PDF
- Publicar en Rpubs

Sesión 07 Probabilidad y distribuciones de probabilidad

- Muestra con `sample()`
- Combinaciones y Permutaciones
- Probabilidad
- Operaciones de conjuntos
- Variables aleatorias discretas y continuas
- Distribución binomial
- Distribución Normal

Sesión 08 Simulación y visualización de datos

- Librería `dplyr`
- Simulación de datos
- Gráficos avanzados `ggplot2()` y `plotly()`

Sesión 09.

- Strings and Dates en R
- Datos semiestructurados

Sesión 10

- R Shiny