



Escuela de Software Libre

Herramientas fundamentales para todo Data Scientist



Organizan:



f Fepon EPN | MATH Consultores Ecuador

Con el Aval de:



ESCUELA
POLITÉCNICA
NACIONAL



Quito

Escuela Politécnica Nacional



Marzo-Mayo | 2018



info@mathconsultores.com
fepon@epn.edu.ec



(+593) 0958873211
(02) 2679-300 Ext.5114



Módulo 1

Introducción al Lenguaje de Programación R

LUGAR

La “Escuela de Software Libre”- Herramientas para todo Data Scientist se llevará a cabo en la Escuela Politécnica Nacional (EPN)

Aula: Esfot 22 (Tecnologos)

FECHAS

Horario entre semana: lunes a viernes de 18:00 a 21:00

- Fecha Inicio: lunes 5 de marzo
- Fecha Fin: miércoles 14 de marzo

CONTENIDOS

1. Introducción

- ✓ Funcionamiento.
- ✓ Instalación y actualización de librerías.
- ✓ Interfaz de usuario RStudio.

2. Estructuras de datos

- ✓ Vectores atómicos y Listas.
- ✓ Matrices y Arrays.
- ✓ Factores.
- ✓ Data Frames y Data Tables.
- ✓ Coerción.

3. Funciones

- ✓ Estructuras de control (loops y condicionales)
- ✓ Creación de funciones

- ✓ Ámbitos léxicos (Lexical Scoping)
- ✓ Ambientes (Environments)

4. Lectura y escritura de datos

- ✓ Lectura de archivos.
- ✓ Escritura de archivos.
- ✓ Gestión de directorios y archivos.
- ✓ Paquetes de lectura eficientes.

5. Tratamiento de datos

- ✓ Selección de variables y observaciones.
- ✓ Transformación de variables.
- ✓ Valores perdidos y duplicados.
- ✓ Recodificación.





6. Generación de datos

- ✓ Generación de secuencias
- ✓ Generación de números aleatorios
- ✓ Extracción de muestras
- ✓ Simulación

7. Gestión de R archivos y espacio de trabajo

- ✓ Cargar y listar objetos
- ✓ Función Load()

- ✓ Función Source()

8. Gráficos

- ✓ Sistemas y dispositivos gráficos.
- ✓ Gráficos de distribución.
- ✓ Gráficos de dispersión.
- ✓ Facetas (gráficos condicionales)

9. Estadística descriptiva

- ✓ Frecuencias.
- ✓ Estadísticos descriptivos.
- ✓ Regresión lineal.

```
chart_link = apf_create(p, filename="box/multiple")
chart_link
```

```
library(plotly)
set.seed(100)
d <- diamonds[sample(nrow(diamonds), 1000), ]
plot_ly(d, x = ~carat, y = ~price, color = ~carat,
        size = ~carat, text = ~paste("Clarity: ", clarity))
```

Average High and Low Temperatures in New York

