

logo_ult.pdf

QUITO: FERNANDO OVIEDO E8-65 Y JOSÉ BARBA

CONTACTOS: (02) 2 559-703 / (02) 2 602-560

MOVISTAR: 0998428362 / 0998890021

1. Quiénes somos ?

Source Stat Lab (SSL) es una empresa Ecuatoriana con sede en Quito especializada en fomentar el avance del conocimiento, misma que presta sus servicios de entrenamiento y capacitación en el lenguaje de programación R para profesionales en todas sus verticales (investigación, docencia, empresarial). SSL brinda capacitación, reportería y consultoría estadística/matemática con el uso de herramientas de software libre como: R, RStudio, R Analytic Flow & LaTeX a empresas globales y locales, así como instituciones públicas.

2. Historia

SSL nació en Septiembre 2014 con el propósito de dar servicio y asesoramiento en investigaciones aplicadas que requieran de estudios estadísticos/matemáticos asociados a las mismas.

3. Objetivos

Entre los objetivos de SSL se encuentran:

- Potenciar las actividades relacionadas con la Estadística que se llevan a cabo en varias universidades Ecuatorianas, así como ofrecer asesoramiento estadístico tanto a grupos de investigación como a particulares y empresas. Dicho asesoramiento se complementa con cursos de formación adaptados a las necesidades y requerimientos del usuario.
- Fomentar el uso del software estadístico R en actividades académicas, profesionales e investigación.

4. Cursos

Durante los últimos años, las nuevas tecnologías han permitido generar, almacenar y difundir grandes cantidades de información. Para poder extraer conocimiento y generar valor, hacen falta herramientas analíticas.

Dado que la estadística es la herramienta determinante para la toma de decisiones y la obtención de conocimiento, SSL ofrece los siguientes cursos:

- R Nivel Básico
- R Nivel Intermedio
- R Nivel Avanzado
- Gráficos con ggplot2
- Interfaces Web con Shiny
- Reportería Dinámica

R Nivel Básico

Descripción:

El curso de R Nivel Básico tiene como objetivo principal proporcionar al estudiante una visión general del entorno del lenguaje de programación R y sus aplicaciones. Al final del curso, el estudiante se encontrará en la capacidad de crear sus propias funciones, importar y exportar datos.

El enfoque del curso se lo hará sobre el ámbito estadístico, sin embargo, lo anterior no resta méritos para que el estudiante amplie sus posibilidades en diversas áreas. Por tratarse de un curso inicial el estudiante recibirá abundante información la cual le permita profundizar de forma autónoma en la utilización del programa.

Duración:

20 Hrs.

Requisitos:

Ninguno

Contenidos del curso:

1. Introducción
 - 1.1 Funcionamiento
 - 1.2 Ventajas
 - 1.3 Desventajas
2. Instalación y actualización
 - 2.1 Programa R
 - 2.2 Entorno de trabajo
 - 2.3 Instalación de paquetes
 - 2.4 Actualización de paquetes
 - 2.5 Actualización de R
 - 2.6 Obteniendo ayuda
3. IDE
 - 3.1 RStudio
 - 3.2 R Analytic Flow
 - 3.3 R Commander

4. Estructura de datos

4.1 Vectores

4.2 Matrices

4.3 Arrays

4.4 Data Frames

4.5 Listas

4.6 Factores & Tablas

5. Lectura y escritura de datos

5.1 Ingreso mediante el teclado

5.2 Lectura de archivos

5.3 Escritura de archivos

5.4 Accediendo a datos de la web

6. Funciones

6.1 Funciones de R

6.2 Estructuras de control

6.3 Creación de funciones

6.4 Edición de funciones

6.5 Almacenamiento de funciones

R nivel Intermedio

Descripción:

El curso R nivel Intermedio tiene dos propósitos principales:

1. Proporcionar al estudiante las principales herramientas de R utilizadas en la manipulación, tratamiento y depuración de la información de un conjunto de datos.
2. Mostrar al estudiante una forma de optimizar las líneas de código tanto en tiempo de ejecución como en apariencia (códigos más compactos).

Está dirigido a usuarios con un conocimiento básico de R. El curso se inicia analizando a detalle una amplia gama de R-funciones utilizadas en la práctica, mismas que permiten acelerar el tiempo de ejecución del código y a su vez reducir el tiempo en la generación de grandes códigos que R ya los tiene implementados de forma más eficiente.

Dentro de los temas a tratarse se encuentran las funciones **apply**, mismas que aplican una función a distintos objetos (vector, matriz, data frames, etc) evitando de esta forma la utilización de los típicos lazos iterativos **for**, **while**, **etc** que por lo general resultan ser poco eficientes.

Duración:

20 Hrs.

Requisitos:

R nivel Básico.

Contenidos del curso:

1. Manipulación y depuración de bases de datos
 - 1.1 Unión de objetos y estructuras
 - 1.2 Unión de bases de datos por columnas comunes
 - 1.3 Valores perdidos y recodificación
 - 1.4 Filtrado y ordenamiento de bases de datos
 - 1.5 Discretización de variables
 - 1.6 Muestras aleatorias
2. Aplicación de una función a cada elemento de un objeto
 - 2.1 Aplicación de una función a una matriz o arreglo
 - 2.2 Aplicación de una función a una lista, data frame, vector

- 2.3 `mapply()` la versión multivariante de `sapply()`
- 2.4 Aplicación de una función a subgrupos de un vector
- 2.5 Aplicación de una función a subgrupos de un data frame
- 3. Funciones vectoriales
 - 3.1 Funciones vectoriales más utilizadas
 - 3.2 Creación de funciones vectoriales
- 4. R-Funciones brillantes
 - 4.1 Funciones lógicas
 - 4.2 Funciones matemáticas
 - 4.3 Resumen de escalares
 - 4.4 Resumen vectorial
 - 4.5 Buscando valores específicos en un objeto
 - 4.6 Eliminación de registros duplicados

R Nivel Avanzado

Descripción:

El curso de R Nivel Avanzado tiene una visión general de varios temas de investigación avanzada, tales como: ambientes, programación orientada a objetos, big data, etc. El curso se encuentra dirigido a los estudiantes que ya tengan experiencia y conocimiento del programa estadístico R y que se encuentren dispuestos a obtener una perspectiva de mayor profundidad respecto a la programación R con visión en convertirse en un desarrollador.

Duración:

20 Hrs.

Requisitos:

R Nivel Intermedio

Contenidos del curso:

1. Introducción
 - 1.1 Cómo trabaja R
2. Ambientes
 - 2.1 Ambientes básicos
 - 2.2 Recursividad
 - 2.3 Funciones
 - 2.4 Ambientes explícitos
3. Programación orientada a objetos
 - 3.1 Clases S3
 - 3.2 Clases S4
 - 3.3 Clases de referencia o R5
4. Programación funcional
 - 4.1 Funciones anónimas
 - 4.2 Clausuras
 - 4.3 Lista de funciones
5. Evaluación no estándar
 - 5.1 Captura de expresiones

5.2 Llamado desde otra función

5.3 Sustitución

6. Expresiones

6.1 Estructura

6.2 Names

6.3 Calls

6.4 Parsing & Deparsing

7. Big Data

7.1 Optimización de memoria

7.2 Carga de datos en disco

7.3 Manipulación de big data

Gráficos estadísticos con ggplot2

Descripción:

El objetivo del curso gráficos con **ggplot2** es dotar al estudiante de importantes herramientas útiles en la generación de gráficos estadísticos de alta calidad y complejidad a través del paquete ggplot2, mismo que en la actualidad es uno de los paquetes gráficos de R más utilizados, la principal ventaja que presenta a diferencia de otros dispositivos gráficos es la utilización de un lenguaje estándar que permite describir la forma de visualización de manera sencilla (gramática de gráficos), conocer dicha **gramática de los gráficos** permite al usuario optimizar sus gráficos estándar y la vez generar gráficos totalmente nuevos e innovadores.



Figura 1: Diagrama de cajas SPSS vs ggplot2.

Entre los gráficos estadísticos generados con **ggplot2** podemos enumerar los siguientes: histogramas, diagramas de densidad, de cajas y bigotes, de barras, de pastel, de dispersión, etc, adicionalmente, podemos crear **mapas geográficos** en los cuales es posible presentar resúmenes estadísticos por región, provincia, etc (Estadística espacial). En la Figura 1 se muestra un claro ejemplo de la diferencia entre la calidad de un gráfico **ggplot2** y de un gráfico realizado a través del programa estadístico clásico SPSS.

Duración:

20 Hrs.

Requisitos:

R nivel básico.

Contenidos del curso:

1. Introducción e instalación del paquete
2. La gramática de ggplot2
 - 2.1 Generación del primer gráfico ggplot2
3. Gráficos de distribución
 - 3.1 Histogramas
 - 3.2 Curvas de densidad
 - 3.3 Diagrama de caja-bigotes
 - 3.4 Múltiples gráficos de distribución
4. Gráficos de líneas
 - 4.1 Gráfico de líneas básicas
 - 4.2 Gráfico de líneas múltiples
 - 4.3 Apariencias de líneas (color, estilo, forma, etc).
5. Gráficos de dispersión
 - 5.1 Gráfico de dispersión básico
 - 5.2 Agrupación puntos con color, forma, etc.
 - 5.3 Adición de líneas de ajuste de modelos de regresión
6. Mapas geográficos
 - 6.1 Gráfico de mapa geográfico básico
 - 6.2 Adición de color, tipo de línea, etc. para cada provincia
 - 6.3 Adición de resúmenes estadísticos por región, provincia, etc.
 - 6.3 Adición de escalas de color
7. Formato del gráfico
 - 7.1 Modificación de escala, color, letra, etc. de ejes
 - 7.2 Modificación de posición, color, letra, etc. de legendas
 - 7.3 Modificación de márgenes, color de fondo

8. Color en gráficos

8.1 Escala discreta de colores

8.2 Escala continua de colores

9. Facetas

9.1 Generación de gráficos por subgrupo de la data (variable discreta)

9.2 Generación de gráficos por dos subgrupos de la data (dos variables discretas)

Interfaces de Usuario con Shiny

Descripción:

Implementar un algoritmo en R permite disminuir considerablemente el tiempo de ejecución de una determinada tarea a diferencia de realizarla manualmente. El usuario de R siempre debe tener en mente que sus algoritmos serán utilizados por otros usuarios, entre ellos se debe considerar un pequeño porcentaje de usuarios que no poseen conocimiento suficiente del lenguaje R, es por esta razón que es necesario desarrollar interfaces gráficas mediante las cuales el usuario pueda ejecutar el algoritmo a través de botones y controladores (widgets), sin la necesidad de visualizar las líneas de código.



Figura 2: Simulación de variables aleatorias.

El curso *Interfaces de Usuario con Shiny* tiene como finalidad proveer al estudiante un conjunto de herramientas utilizadas en la generación de interfaces gráficas a través del paquete Shiny. Una vez que el estudiante logre comprender su estructura y funcionamiento resultará sencillo e intuitivo su uso. En la Figura 1 mostramos un ejemplo de una

interfaz que simula un número de observaciones determinado de algunas distribuciones de probabilidad clásicas.

Duración:

20 Hrs.

Requisitos:

R nivel básico.

Contenidos del curso:

1. Introducción a Shiny
 - 1.1 Estructura de una aplicación
 - 1.2 Ejecución de una aplicación
 - 1.3 Ejemplos
2. Diseño de la interfaz de usuario
 - 2.1 Construcción de la interfaz de usuario
 - 2.2 Funciones y contenidos HTML
 - 2.3 Formato de texto
 - 2.4 Adición de imágenes y videos
3. Añadir widgets de control (botones)
 - 3.1 Widgets de control básicos
 - 3.2 Adición de widgets de control
4. Objetos de salida reactiva
 - 4.1 Adición de un R objeto a la interfaz de usuario
 - 4.2 Construcción de un objeto reactivo
 - Textos reactivos
 - Gráficos reactivos
 - Tablas reactivas, etc.
5. Carga y descarga de archivos
 - 5.1 Carga de archivos y direcciones de archivos
 - 5.2 Descarga de archivos de resultados
6. Publicación de aplicaciones Shiny
 - 6.1 Ejecución de Interfaz mediante RStudio
 - 6.2 Ejecución de Interfaz mediante un Navegador Web

Reportería Dinámica

Descripción:

El curso de Reportería Dinámica tiene como finalidad integrar R con LaTeX, dichas herramientas permitirán generar reportes que se actualicen automáticamente a medida que los datos cambien. Esto facilita la elaboración de reportes evitando recurrir al famoso copy-paste de todos los resultados y gráficos de un programa a otro.

Al final del curso, el estudiante se encontrará en la capacidad de crear sus propios informes dinámicos en varios formatos digitales, mismos que eliminan la posibilidad de cometer errores de transcripción.

Duración:

20 Hrs.

Requisitos:

R Nivel Básico

Contenidos del curso:

1. Introducción
 - 1.1 Antecedente
 - 1.2 Sweave
2. Reportería
 - 2.1 Buenas & malas prácticas
 - 2.2 Restricciones
3. Editores
 - 3.1 RStudio
 - 3.2 LyX
 - 3.3 Otros editores
4. Mi primer reporte
 - 4.1 Configuraciones básicas
 - 4.2 Knitr
 - 4.3 LaTeX
 - 4.4 Markdown

5. Dinamizando la reportería

5.1 Syntaxis

5.2 Chunks

5.3 R Scripts

5.4 Tablas

5.5 Gráficos

5.6 Publicación en la web

6. Cache

6.1 Implementación

6.2 Actualización