



QUITO: FERNANDO OVIEDO E8-65 Y JOSÉ BARBA,

SECTOR: EL DORADO

CONTACTO: (02) 2 559 703 / (593) 0998890021

## 1. Quiénes somos ?

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

## 2. Historia

SSL nació en Septiembre 2014, mediante la iniciativa de un grupo de ingenieros matemáticos de la Escuela Politécnica Nacional, 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 y de 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, se requieren varias herramientas analíticas.

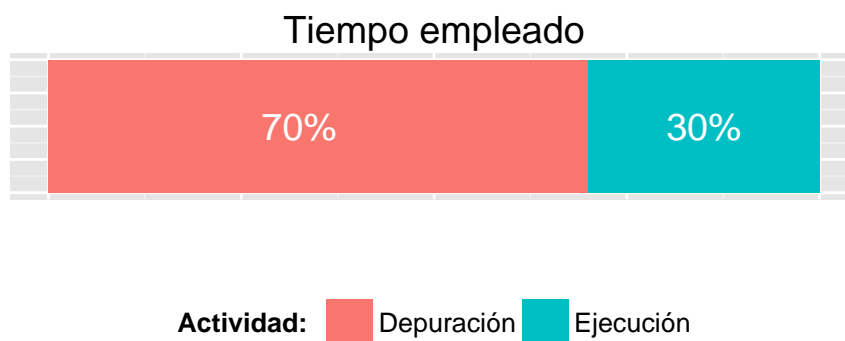
Dado que la estadística es la herramienta determinante para la toma de decisiones y la obtención de conocimiento en la mayoría de entidades del sector público y privado, SSL oferta al público en genral los siguientes cursos:

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

# R nivel Intermedio

## Descripción:

Generalmente cuando se aplica una determinada herramienta estadística o matemática, de muestreo, de control de calidad o cuando se estima un modelo de regresión, series temporales, etc. el 70 % del tiempo disponible es utilizado en la limpieza y depuración de la data disponible, por ejemplo en la detección y corrección de datos atípicos y faltantes, en la unión o consolidación de la data en un solo archivo cuando esta viene dada en distintas bases de datos, y únicamente el 30 % restante es destinado al desarrollo de la herramienta técnica o construcción del modelo.



El manejo intermedio de R busca corregir esta contradicción, pues lo ideal es destinar la mayor cantidad de tiempo disponible al análisis técnico de la data, la búsqueda de mejoras en las metodologías empleadas y la investigación de nuevas herramientas técnicas utilizadas en la actualidad. Es por esta razón, que 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 determinado. Luego de aprobar el curso el estudiante tendrá la capacidad de desarrollar códigos, que automaticen las tareas que comunmente las realiza manualmente al momento de trabajar con la información, capaces de adaptarse a distintas instancias que se puedan presentar.
2. Mostrar al estudiante la forma de como optimizar sus líneas de código tanto en tiempo de ejecución como en apariencia (códigos más compactos).

El curso R Nivel Intermedio está dirigido a usuarios con un conocimiento básico de R. Se analiza a detalle una enorme gama de R-funciones que resumen algoritmos muy empleados

en la práctica, las mismas que permiten acelerar el tiempo de ejecución de ciertas líneas de código y principalmente evitar malgastar el tiempo generando enormes códigos que en R ya se encuentran implementados y de una forma mucho más eficiente, por ejemplo uno de los temas a tratarse en este curso comprendiendo una herramienta muy empleada, la cual consiste en aplicar una misma función estadística a distintos objetos evitando usar los típicos lazos iterativos **for**, **while**, **etc** que por lo general resultan ser muy poco eficientes.

## **Duración:**

20 Hrs.

## **Requisitos:**

R nivel Básico.

## **Objetivos:**

- Proporcionar las principales herramientas de R, utilizadas en la manipulación, tratamiento y depuración de datos.
- Mostrar la forma de como optimizar de códigos tanto en tiempo de ejecución como en apariencia.

## **Contenidos:**

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 data continua
  - 1.6 Muestras aleatorias
2. Aplicación de funciones a cada elemento de un objeto
  - 2.1 Aplicación de funciones a una matriz o arreglo.
  - 2.2 Aplicación de funciones a un vector, lista, data frame.
  - 2.3 `mapply()`, versión multivariante de `sapply()`
  - 2.4 Aplicación de funciones a subgrupos de la data.
3. Funciones vectoriales

- 3.1 Funciones Vectoriales más utilizadas.
- 3.2 Creación de Funciones Vectoriales.
- 4. R-Funciones brillantes
  - 4.1 Funciones lógicas.
  - 4.1 Funciones matemáticas.
  - 4.1 Resúmenes escalares.
  - 4.1 Resúmenes Vectoriales.
  - 4.1 Buscando valores específicos en un objeto.
  - 4.1 Eliminación de registros duplicados.

## Bibliografía:

1. Teetor, P. (2011). *R cookbook*. O'Reilly Media, Inc. United States of America.
2. Matloff, N. (2011). *The art of R programming: tour of statistical software design*. No Starch Press, Inc. San Francisco.

## Métodos de evaluación.

La forma de evaluación se resume en la tabla siguiente, el estudiante aprobará el curso con una calificación superior o igual a 7/10 puntos.

Método de evaluación	Número	Puntaje	Porcentaje
Pruebas parciales	2	4	40 %
Deberes	2	2	20 %
Trabajo final	1	2	20 %
Examen final	1	2	20 %
Total	6	10	100 %

Tabla 1: Métodos de evaluación.