

Fundamentos en análisis de información y minería de datos en agricultura específica por sitio

Fechas y lugar:

Del 28 al 30 de octubre del 2019, Ciudad Lima - Perú.

Instructores: El taller será presentado por el Estadístico y Magister en Computación Hugo Andrés Dorado y el profesional en Matemáticas Juan Camilo Rivera, ambos investigadores del Centro Internacional de Agricultura Tropical (CIAT - Colombia), con varios años de experiencia trabajando en minería de datos aplicada a la agricultura.

Requisitos del participante: Conocimiento básico de estadística, Excel y preferiblemente programación. Además, este curso es recomendable para las personas que trabajen en temas relacionados con el manejo, procesamiento o análisis de datos.

Antecedentes: Este taller se ha realizado en 4 países de Latinoamérica: Uruguay, Honduras, Nicaragua y Colombia, y ha contado con la participación de representantes del sector público y privado, tales como: INTA (Instituto Nacional de Técnica) en Argentina, INIA (Instituto Nacional de Investigación Agropecuaria) en Uruguay, FLAR (Fondo Latinoamericano para Arroz de Riego), COPECO (Comisión Permanente de Contingencias) en Honduras, FEDEARROZ (Federación Nacional de Arroceros de Colombia) y FENALCE (Federación Nacional de Cerealistas y Leguminosas de Colombia), entre otros. Adicionalmente el contenido de este taller se ha incorporado en Diplomados y en programas académicos de carreras del sector agropecuario en Colombia.

El objetivo de cada taller es brindar las herramientas en análisis de datos mediante el software de uso libre R para que cada asistente, al finalizar el curso, las aplique en sus bases de datos.

Requerimientos: Computador portátil y preferiblemente con los siguiente software instalados:

- Software estadístico R. Se puede descargar en <https://cran.r-project.org/> ó <https://cran.dcc.uchile.cl/>.
- R Studio. Es una interfaz para R. Se puede descargar en (<https://www.rstudio.com/products/rstudio/download/>).
- Tableau. Es un software para visualizar datos. Se puede descargar en <https://www.tableau.com/es-es/products/desktop/download>) **ATENCIÓN. Porfavor descargarlo en la semana del taller ya que la versión gratuita tiene un duración de 12 días.**
- Slack. Es el canal de comunicación. Hay dos opciones descargar el software o usar la versión, recomendamos esta última. En este link <https://slack.com/> pueden crear su usuario y contraseña.

Contactos: Cualquier inquietud, por favor escribir a: h.a.dorado@cgiar.org ó J.C.Rivera@cgiar.org

PROGRAMA.

1^{er} día, 28 de octubre

Hora	Tema
8:00 – 9:00	Instalación de software y paquetes.
9:00 – 10:00	Presentaciones: Agricultura Digital y Datos
10:00 – 10:20	Receso
10:20 – 10:40	Presentación de la base de datos de ejemplo
10:40 – 12:00	Introducción al software R.
	• Crear un proyecto
	• Manipulación de objetos
	• Funciones básicas.
	• Lectura de datos.
12:00 – 1:00	• Gráficos base (Dispersión, histograma, curvas)
	Almuerzo.
1:00 – 2:00	Análisis estadístico descriptivo (Repaso de conceptos) (Promedio, mediana, rango, desviaciones estándar, histogramas, boxplot,...)
2:00 – 3:00	Gráficos descriptivos en R (ggplot2)
3:00 – 3:20	Receso
3:20 – 4:30	• Ejercicio práctico (Gráficos con ggplot2)

2^{do} día, 29 de octubre

Hora	Tema
8:00 – 8:30	Revisión ejercicio práctico, introducción a R
8:30 – 9:30	Limpieza de base de datos (errores comunes)
9:30 – 10:00	Presentación, procesamiento y decisiones sobre los datos faltantes.
10:00 – 10:40	Presentación, obtención de datos de clima, suelo y procesamiento de información.
	• Fuentes de información secundaria (Worldclim, CRU, estaciones meteorológicas)
10:40 – 11:00	Receso
11:00 – 12:00	Presentación, alineación de bases de datos y procesamiento.
	• Ejercicio práctico
12:00 – 1:00	Almuerzo
1:00 – 3:20	Introducción a minería de datos en agricultura. (Algunos métodos aplicados)
3:20 – 3:40	Receso
3:40 – 4:30	• Ejercicio práctico (Minería de datos)

3^{er} día, 30 de octubre

Hora	Tema
8:00 – 8:30	Revisión del caso práctico de alineación de bases de datos y minería de datos
8:30 – 9:00	Análisis de varianza, regresión lineal múltiple (repaso y aplicación).
9:00 – 10:30	Presentación, Visualización con Tableau
	<ul style="list-style-type: none"> Ejercicio práctico (Tableau)
10:30 – 10:40	Receso
10:40 – 12:00	<ul style="list-style-type: none"> Continuación ejercicio práctico (Tableau)
12:00 – 1:00	Almuerzo
1:00 – 3:30	Desarrollo de un caso práctico con bases de datos propias o de ejemplo (Taller en grupos).
3:30 – 4:10	Presentación y discusión de resultados.
4:10 – 4:30	CIERRE

Links de interés:

- <https://conceptosclaros.com/instalar-r-primeros-pasos/>
- http://cran.r-project.org/doc/contrib/rdebuts_es.pdf
- <https://github.com/bigdataciat>