

TALLER - FUNDAMENTOS EN ANÁLISIS DE INFORMACIÓN Y MINERÍA DE DATOS EN AGRICULTURA ESPECÍFICA POR SITIO

Como parte de la colaboración entre ICDF, INTA y CIAT en el marco del Memorando de entendimiento para *“Desarrollar investigación en prácticas de adaptación al cambio climático en el cultivo de arroz”*, llevaremos a cabo el taller *“Fundamentos en análisis de información y minería de datos en agricultura específica por sitio”*. El propósito es fortalecer las habilidades de investigadores nacionales en la implementación de métodos modernos de análisis de datos, los cuales podrán ser implementados como soporte en la toma de decisiones en agricultura. En esta capacitación se incluirá la búsqueda e integración de distintas fuentes de datos disponibles con variables de clima y suelo.

Tema Central:

Análisis de datos comerciales en agricultura con un enfoque BIG DATA, a través del uso de metodologías de exploración de datos, aprendizaje de máquina (Machine Learning) y visualización.

Fechas y lugar:

Del 20 al 22 de marzo, Ciudad Managua, Nicaragua, oficinas del INTA.

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 de datos, como análisis o procesamiento de datos.

Cupo máximo del taller:

15 Personas

Antecedentes:

Este taller se ha realizado en varios países de Latinoamérica como Uruguay, Honduras, Nicaragua y Colombia, y han contado con la participación de representantes del sector público y privado. Entre todas las instituciones con las que hemos trabajado cabe la pena mencionar a INTA (Instituto Nacional de Técnica) en Argentina, INIA (Instituto Nacional de Investigación Agropecuario) en Uruguay, FLAR (Fondo Latinoamericano para Arroz de Riego), COPECO (Comisión Permanente de Contingencias) en Honduras, FEDEARROZ (Federación Nacional de Arroceros) y FENALCE (Federación Nacional de Cerealistas y Leguminosas) 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/>.
 - 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

FECHA: 20 de marzo

Hora	Tema
8:00 – 9:30	Bienvenida y apertura del taller.
9:30 – 10:00	Presentación, Agricultura Digital y Datos.
10:00 – 10:40	Introducción a análisis de datos en agricultura. <ul style="list-style-type: none"> • Rápido repaso de conceptos básicos de análisis de datos aplicado a agricultura.
10:40 – 12:00	Introducción al software R. <ul style="list-style-type: none"> • Revisión del funcionamiento de R en los equipos de los participantes. • Manipulación de objetos en R • Lectura de datos. • Medidas descriptivas.
12:00 – 1:30	Almuerzo.
1:30 – 3:00	Continuación introducción a R. <ul style="list-style-type: none"> • Gráficos descriptivos en R (ggplot2)
3:00 – 4:30	Caso práctico <ul style="list-style-type: none"> • Descripción del objetivo del caso práctico. • Análisis descriptivo (Promedios, máximos, mínimos, desviaciones estándar, histogramas y boxplot)

2^{do} día

FECHA: 21 de marzo

Hora	Tema
8:00 – 8:30	Revisión caso práctico, introducción a R
8:30–9:30	Análisis de varianza, regresión lineal múltiple (repaso y aplicación).
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. <ul style="list-style-type: none"> Fuentes de información secundaria (worldclime, CRU, estaciones meteorológicas,...)
10:40 – 12:00	Caso práctico ejercicio con conjunto de datos de ejemplo. <ul style="list-style-type: none"> Modelo de regresión lineal multiple, datos faltantes, análisis de varianza.
12:00 – 1:30	Almuerzo
1:30 – 2:20	Presentación, Consideraciones para la alineación de bases de datos y procesamiento.
2:20 – 3:00	Presentación, minería de datos en agricultura. (Algunos métodos aplicados)
3:00 – 4:30	Caso práctico. <ul style="list-style-type: none"> Alineación de bases de datos y minería de datos.

3^{er} día

FECHA: 22 de marzo

Hora	Tema
8:00 – 8:30	Revisión del caso práctico de alineación de bases de datos y minería de datos
8:30 – 9:00	Otros repositorios para análisis complementarios
9:00 – 10:30	Script de REGRESSION METHODS
10:30 – 11:30	Presentación, Visualización Tableau
11:30 – 12:00	Casó práctico. Implementación del script de análisis con un conjunto de datos de ejemplo.
12:00 – 1:30	Almuerzo
1:30 – 3:20	Continuación casó práctico. Implementación del script de análisis con el conjunto de datos de ejemplo.
3:20 – 4:10	Presentación y discusión de resultados.
4:00-4:30	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>