

Cronograma de la Tercera Escuela de Verano

Centro de Economía y Política Regional (CEPR) + GobLab UAI

4,5,6 y 7 de enero 2021

Modalidad Online (Zoom)

PROGRAMA

Lunes 4 de enero	09-12:00 hrs	Introducción al software R
	15-16:30 hrs	Taller de Generación de proyectos de Ciencia de Datos
Martes 5 de enero	09-12:00 hrs	Data cleaning and handling
	15-16:30 hrs	Understanding data
Miércoles 6 de enero	09-12:00 hrs	Data Visualization & EDA
Jueves 7 de enero	09-12:00 hrs	Machine Learning
	15-16:30 hrs	Webscraping con rvest

CONTENIDOS

Introducción al software: Entender la estructura de las fuentes de información y el cómo trabajarlas es un requisito esencial para poder introducirse a la ciencia de datos. En este tópico se estudiarán las diferentes estructuras de las bases de datos y el cómo alterarlas de forma eficiente para realizar ciencia de datos.

Taller de Generación de proyectos de Ciencia de Datos: Este taller presentará una metodología para explorar qué problemas públicos se pueden resolver con ciencia de datos. Se realizará un ejercicio práctico de análisis de factibilidad de un proyecto y se presentará una matriz de madurez de datos.

Data cleaning and handling: Este curso tiene como objetivo repasar ciertos conceptos básicos para el trabajo con bases de datos en "R". Además, se proporcionará una introducción a dplyr y data.table, herramientas usualmente utilizadas en "R" para estos fines.

Data Visualization & EDA: El uso de visualizaciones es una de las formas más eficaces para comunicar la información de las bases de datos. En este tópico se introducirá el uso de paquetes de visualización de datos en "R" como ggplot2, plotly, leaflet, entre otros.

Understanding Data: Este curso está orientado a discutir la manera correcta de describir y analizar datos. En un mundo donde la información y los datos son de fácil acceso, es importante aprender las habilidades para poder procesar esa información, realizar análisis preliminares, y presentarlos de forma que sean útiles para una potencial investigación. Discutiremos aspectos tanto de forma como de fondo del análisis de datos.



Introducción a técnicas de Machine Learning: El objetivo de este curso es que los estudiantes conozcan y apliquen los conceptos básicos del Machine Learning para generar modelos a partir de conjuntos de datos. Se utilizará el lenguaje de programación “R” y fuentes de datos de índole territorial para ejemplificar su uso.

Web scraping con rvest: Este curso tiene como objetivo introducir a los asistentes al web scraping, sus usos, alcances y limitaciones. Además, se proporcionará una introducción al rvest, una herramienta sencilla, útil y eficaz para introducirse al mundo del web scraping en “R”.