



Ciencias Sociales
Universidad de la República
URUGUAY

Programa de Curso Edición 2024

Análisis exploratorio

Especialización en Economía, opción Ciencia de Datos

Docente/s y correo electrónico: *Guillermo Lezama,*
guillermo.lezama@cienciassociales.edu.uy

Créditos: 4

Régimen de cursado: *Presencial / Híbrido*

Carga y distribución de horas estimada: *20 horas presenciales y 40 horas de trabajo domiciliario.*

Conocimientos previos recomendados: Python básico, Jupyter Notebook, Estadística Básica

Contenidos y organización del curso:

Parte 1: Estadísticas Descriptivas y Preparación de Datos

- Introducción y principios del EDA.
- Estadísticas descriptivas y visualización de datos.
- Técnicas de limpieza y preparación de datos.
- Uso de SQL para análisis de datos.

Parte 2: Manejo y Transformación de Variables

- Creación, modificación y eliminación de variables.
- Feature Engineering.
- Análisis exploratorio gráfico.
- Comparación de grupos para identificar disparidades.

Parte 3: Técnicas Avanzadas de EDA

- Análisis de correlación y estadístico exploratorio.
- Reducción de dimensionalidad para la exploración.
- Clustering para describir datos.
- EDA aplicado a análisis de series temporales y datos de texto.
- Generación y prueba de hipótesis basadas en EDA.

Método de enseñanza: Clases teóricas con aplicaciones prácticas. Ejercicios domiciliarios en cada clase.

Sistema de evaluación:

- **Proyecto final entregable (60%):** Aplicar técnicas de EDA en un dataset a determinar y presentar los hallazgos.
- **Trabajos de clase (40%):** Ejercicios prácticos realizados durante el curso.

Bibliografía Recomendada

No se seguirá un libro específico, pero se utilizará alguno de estos libros para cada unidad.

- "Python for Data Analysis" by Wes McKinney.
- "Python Data Science Handbook" by Jake VanderPlas.
- "Learning SQL" by Alan Beaulieu.
- "Practical Statistics for Data Scientists" by Peter Bruce y Andrew Bruce.
- "Introduction to Time Series Forecasting With Python", Jason Brownlee.
- "The Effect: An Introduction to Research Design and Causality" by Nick Huntington-Klein, Chapters 3 and 4.