

**Lic. en Sociología**  
**Universidad Nacional del Chaco Austral**  
**Estadísticas Aplicadas a las Ciencias Sociales**

**CIENCIA DE DATOS Y SOFTWARE LIBRE PARA CIENCIAS SOCIALES**

**PLAN DE TRABAJO**

**ADSCRIPTO:** MERCADO, Isaac Pablo Rubén

**CÁTEDRA:** Estadísticas Aplicadas a las Ciencias Sociales

**CARRERA:** Licenciatura en Sociología

**PERIODO:** 2do Cuatrimestre

**DOCENTE RESPONSABLE DE LA ADSCRIPCIÓN:** Mg. Facundo Matías Kalin e Ing. Sabrina Laglaive

**LUGAR DE TRABAJO:** Cátedra de Estadísticas Aplicadas a las Ciencias Sociales – UNCAus Virtual.

**FUNDAMENTACIÓN:**

La propuesta consiste en dar a conocer a los y las estudiantes de la carrera de Licenciatura en Sociología de la Universidad Nacional del Chaco Austral herramientas gratuitas de software libre y ciencias de datos para trabajar mediante técnicas de recolección y análisis de datos sociales.

El uso de las nuevas tecnologías de la información y comunicación así como la exigencia de hacer públicos los datos que se producen en el estado nacional y provincial generan un nuevo escenario caracterizado por la masividad de datos presentes, un volumen que años atrás parecían impensados, es a partir de aquí, que esta información es valiosa para las ciencias sociales ya que las mismas cuentan con una gran cantidad de registros de variables que hacen posible nuevos y más profundos análisis.

En este sentido se hace indispensable el uso de nuevas herramientas metodológicas, estadísticas y técnicas de las ciencias de datos capaces de procesar la información,

proporcionando de esta manera a los y las estudiantes un nuevo panorama para realizar investigación científica.

El objetivo principal del curso es realizar una introducción al lenguaje de programación Python, el cual es una de los principales lenguajes para realizar análisis de datos, con un enfoque orientado principalmente a perfiles provenientes de las Ciencias Sociales.

Python es un lenguaje de programación open source y libre, es uno de los más usados en la actualidad y escogido por la industria para desarrollar modelos. Es un lenguaje sencillo, con una curva de aprendizaje lo suficientemente plana para atraer nuevos usuarios de las más diversas formaciones, especialmente aquellos que buscan trabajar con grandes volúmenes de datos dentro del ámbito de las Ciencias Sociales.

De tal modo, la propuesta busca que los estudiantes adquieran una comprensión general del lenguaje de programación y su utilidad para el análisis de datos sociales.

En el curso se centrará en realizar una introducción a algunos conceptos teóricos para ubicarnos en el tema para luego aprender la sintaxis básica de programación, el manejo de datos sociales y su visualización. Así, el énfasis será sobre la importación, organización, manejo de datos sociales y la presentación de los resultados.

La propuesta consiste en brindar dos clases, una teórica y otra práctica, de alrededor de 1 hora cada una, de manera totalmente virtual.

### **OBJETIVO GENERAL**

- Introducir a los estudiantes de la carrera de Lic. en Sociología en conceptos y herramientas básicas para el análisis y visualización de datos sociales.

### **OBJETIVO ESPECÍFICO**

- Lograr que los y las estudiantes hagan uso de la información disponible en internet como datos abiertos para construir conocimientos sobre problemáticas sociales.
- Qué reconozcan la importancia de la estadística como una herramienta de investigación.
- Qué aprendan los procedimientos básicos de limpieza, descripción y análisis de datos cuantitativos a partir del uso de software libres.

- Qué logren familiarizarse con los procedimientos más adecuados a utilizar para el tratamiento de los datos según los objetivos de investigación y los tipos de variables con las cuales se trabaje.
- Que aprendan a realizar un análisis exploratorio de datos (EDA).
- Que aprendan a visualizar de manera eficiente la información.

## **CONTENIDO**

- Introducción al concepto de ciencia de datos. Historia. Diferencia entre Data Mining, Data Lake, Big Data, Business Intelligence. Debates actuales. Introducción al aprendizaje supervisado y no-supervisado. Introducción a la utilización de Ciencias de Datos en las Ciencias Sociales.
- Software Libre y lenguajes de programación. Qué es un lenguaje de programación. Beneficios de aplicación de herramientas computacionales para las ciencias sociales. ¿R o python?. Similitudes y diferencias.
- Introducción a python. ¿Cómo instalar y utilizar python?. Jupyter Notebook y Google Colab. Variables y tipos. Operadores de comparación. Tipos de datos: diccionarios, listas, tuplas, cadenas. Control de Flujo. Ciclos. Funciones. Clases y Objetos.
- Introducción a los módulos y librerías de python: pandas, numpy, matplotlib y seaborn. Principales funciones y métodos de cada paquete. Aplicaciones prácticas.
- Búsqueda de la información. Fuentes abiertas. Base de datos estructuras y no estructuradas.
- Construcción de análisis estadístico descriptivo univariable y análisis exploratorio de datos mediante la utilización de pandas.
- Visualización de la información mediante matplotlib y seaborn.
- Introducción a las herramientas para extracción de información de páginas web (web scrapping).
- Caso de análisis sobre un dataset presentado.

## **CRONOGRAMA**

- La propuesta consiste en dos clases, una de contenido teórico en la cual se abordarán los principales conceptos y otra de contenido práctico en la cual se aplicará en un caso práctico los conceptos tratados.
- El día y el horario es a convenir entre los y las estudiantes.
- Las clases se brindarán mediante una plataforma de reuniones virtuales a convenir (zoom, meet, etc).

## **REQUISITOS:**

- Soporte técnico para tomar clases virtuales (computadora con acceso de internet).
- No se requiere experiencia previa en programación o matemáticas avanzadas.
- Es recomendable estar familiarizado con algunos conceptos básicos de estadísticas descriptivas, aunque no es indispensable.

## **BIBLIOGRAFÍA**

- POLDRACK, R. (2022): Statistical Thinking for the 21st Century. Universidad de Stanford. Disponible en: <https://statstheinking21.org/>
- VÁZQUEZ BRUST, A. (2020): Ciencia de Datos para Gente Sociable. Una introducción a la exploración, análisis y visualización de datos. Disponible en: [https://bitsandbricks.github.io/ciencia\\_de\\_datos\\_gente\\_sociable/](https://bitsandbricks.github.io/ciencia_de_datos_gente_sociable/)
- BECERRA, G. (2018): Big Data como objeto de estudio y método para la investigación empírica en sociología y psicología social. STS, Simposio Argentino sobre Tecnología y Sociedad. Santiago del Estero.
- GUTIÉRREZ-PUEBLA, J.; GARCÍA-PALOMARES, J.C.; SALAS-OLMEDO, M.H. (2016): Big (Geo) Data en Ciencias Sociales: Retos y Oportunidades. Revista de Estudios Andaluces, vol. 33, 1-23. Disponible en: [https://institucional.us.es/revistas/andaluces/33/ART1\\_GUTIERREZ%20PUEBLA.p  
df](https://institucional.us.es/revistas/andaluces/33/ART1_GUTIERREZ%20PUEBLA.pdf)

- PALIOTTA, A.P. (2018): Nuevas profesiones y técnicas de *web data minning* en Argentina: El caso del *data scientist*. Revista del Centro de Estudios de Sociología del Trabajo. Universidad de Buenos Aires. Facultad de Ciencias Económicas.