

Universidad Nacional de Quilmes

Secretaría de Extensión Universitaria

Ciclo: Análisis de datos para ciencias sociales.

Nombre del curso: Introducción a la analítica de datos

Docente/s a cargo: Germán Silva (14hs), Matías Conde (14hs) Marcela Grinszpun (6hs), Emilio Cabello (6hs),

Carga horaria total estimada: 40 horas

Destinado a: abiertos a la comunidad en general.

Cantidad de estudiantes: Cupo mínimo: 15 inscriptos. Cupo máximo 50 estudiantes.

Modalidad de dictado: Virtual sincrónica (2h teórico + 1 práctico) + 1hs asincrónico

Conocimientos previos requeridos: Metodología y técnicas de Investigación social. Análisis de datos cuantitativos en ciencias sociales, Estadística Básica.

Fundamentación:

El presente curso tiene como propósito la presentación de la teoría y herramientas metodológicas disponibles en la actualidad para uso en el análisis de datos entendido como la recopilación, transformación y organización de los datos para sacar conclusiones, hacer análisis descriptivos, predictivos e impulsar una toma de decisiones fundamentada en datos.

Creemos que la utilización de estas herramientas técnicas no debería implementarse sin una discusión epistemológica respecto de los marcos en los cuales se encuadran sus posibilidades.

El presente curso posee como finalidad proveer a los cursantes las herramientas necesarias para poder analizar e interpretar de manera pertinente todo dato que se encuentre estructurado, lo que resulta de vital trascendencia en la investigación científica aplicada en las ciencias sociales, como así también cada vez más en el mercado laboral en el estado actual.

Históricamente ha sido de gran relevancia el conocimiento pertinente al manejo de bases de datos y la exploración de sus contenidos, pero en el contexto actual de explosión en el almacenamiento masivo de datos, y su habitual liberación de bases por parte de organismos o filtraciones, se vuelve por demás indispensable el manejo de las herramientas que nos permitan acceder analíticamente a la información que allí se encuentra.

En la actualidad, existen distintas aplicaciones de software para la sistematización de la información, preparación, limpieza, normalización, análisis y visualización que utilizaremos para la ejercitación y la aplicación de los conceptos.

Objetivos:

Los objetivos generales del presente ciclo serán que los cursantes:

- Adquieran destrezas analíticas básicas como: limpieza, análisis y visualización de datos, y herramientas claves del proceso -hojas de cálculo, SQL, programación en R, PowerBI).
- Logren organizar los datos y a realizar análisis descriptivos mediante hojas de cálculo, SQL y programación en R.
- Aprendan a visualizar y presentar los resultados en tableros, presentaciones y plataformas de visualización.
- Transiten por una experiencia inmersiva de las prácticas y los procesos que transitan los analistas de datos actualmente.

Contenidos:

MODULO 1: Fundamentos del Análisis de Datos

De manera introductoria se verán de modo sucinto las principales etapas por las que atraviesa hoy en día el analista de datos, desde la obtención, la limpieza y transformación de los datos, el procesamiento y la presentación de los análisis realizados. Se presentarán una amplia variedad de términos y conceptos relevantes para el rol de un analista de datos, sobre el ciclo de vida de los datos y el proceso de análisis de datos.

Contenidos

- Introducción al Análisis de datos
- Introducción a Bases de Datos
- Introducción a los Datos Estructurados
- Introducción a Herramientas de BI

Herramientas: Excel, SPSS, SQL, OpenRefine, PowerBI

MÓDULO 2: Fundamentos de Estadística

El curso presenta los principales conceptos de la Estadística Descriptiva, enfatizando aspectos de especial importancia en el procesamiento y análisis de datos: modos usuales de recolección de datos y su registro/sistematización; tipos de variables, medición y valores; medidas de tendencia central y de dispersión en el análisis de conjuntos de datos; entre otros.

Los contenidos se abordan teniendo en cuenta los alcances de las capacidades a desarrollar y los problemas propios del campo profesional.

Contenidos

- Tipos de fuentes de datos disponibles.
- Evaluación de la confiabilidad y la adecuación de los datos disponibles a los procesamiento estadísticos requeridos para el análisis.
- Universo, población y muestra.
- Resumiendo los Datos: Medidas de tendencia central, dispersión y posición.
- Gráficos estadísticos.

Herramientas: Excel, SPSS.

MÓDULO 3: Visualización de Datos

La capacidad de visualizar información de manera precisa es una habilidad clave en el trabajo de un analista de datos. Este curso brinda herramientas teóricas y prácticas para que los gráficos, tableros e informes de resultados desarrollados no contengan errores conceptuales y luzcan profesionales.

Se repasan fundamentos de diseño como la aplicación de colores y la adecuación al branding de empresas e instituciones, al mismo tiempo que se trabaja sobre la correcta selección de visualizaciones según la información disponible y el destino del análisis.

Contenidos

- Elección de gráficos de acuerdo a la información disponible.
- Desarrollo de informes para distintos niveles: de los informes operativos a los tableros estratégicos.
- Fundamentos de diseño para no diseñadores: principios que deben seguir las visualizaciones y los reportes.
- Creación de visualizaciones con herramientas de inteligencia de negocios (BI).
- Visualizaciones con herramientas online.
- Storytelling: cómo crear una narrativa que nos permita contar historias a partir de los datos. Presentación de Resultados.

Herramientas: Tableau, Power Bi, Google DataStudio, Datawrapper y Flourish

Propuesta Didáctica:

El desarrollo del curso contempla instancias teóricas y prácticas. Las clases se darán de forma sincrónica de manera virtual por videoconferencia con una frecuencia semanal donde se desarrollarán los conceptos teóricos y aplicaciones. También se realizarán actividades de forma asincrónica a través del campus - foros, ejercitación, cuestionarios, lecciones. Una vez por semana, se realizarán también encuentros de consulta para la realización de las actividades.

La modalidad de evaluación tendrá en cuenta tanto los procesos de los estudiantes durante el curso a partir de trabajos prácticos que se irán realizando, como la aprobación de un trabajo práctico final integrador.

Al finalizar el curso, los estudiantes desarrollarán un trabajo incluyendo los distintos conocimientos adquiridos a lo largo de los diferentes módulos. A partir de datos abiertos provistos por la universidad, los alumnos deberán presentar un informe de resultados, la base de datos trabajada y la documentación correspondiente a la misma.

Criterios de Evaluación:

Se tendrá en cuenta el grado de participación en los encuentros sincrónicos pautados, la comprensión de los contenidos vistos en cada unidad y su correcta aplicación en la actividad final.

Requisitos de Aprobación:

La aprobación del curso requerirá una participación de los encuentros sincrónicos no inferior al 75% y la aprobación de la actividad final y la resolución de las actividades de aprendizaje semanales que el profesor proponga.

Bibliografía Obligatoria:

- Cruscianelli, S. - ¿Qué es el periodismo de datos?. En Cuadernos de Periodistas.
<http://www.cuadernosdeperiodistas.com/media/2013/12/106-1241.pdf> .
- Huff, D. (2015). Cómo mentir con estadísticas. Barcelona: Crítica.
- López-Roldán, P.; Fachelli, S. (2015), Metodología de la Investigación Social Cuantitativa. Bellaterra (Cerdanyola del Vallès): Dipòsit Digital de Documents, Universitat Autònoma de Barcelona. 1ª edición. Edición Digital
- Sosa Escudero, W., (2019) "Big data: Breve manual para conocer la ciencia de datos que ya invadió nuestras vidas". Editorial Siglo XXI.
- Sosa Escudero, W., (2014, 2021) "¿Qué es y que no es estadística?. Editorial Siglo XXI.
- Vittert, L. (2022) "50 principios de la ciencia de datos". BLUME

Bibliografía de Consulta:

- Aguirre, Niño, Simonetti (2005). Estadística Aplicada en las Ciencias Sociales y Humanas. Universidad Nacional de Misiones. Editorial Universitaria. Posadas.
- Archenti, Piovani, Marradi, (2007). Metodología de las Ciencias Sociales. Emecé. Buenos Aires.
- Cole, N., (2021) "Storytelling con datos : visualización de datos para profesionales de los negocios", Editorial Anaya Multimedia
- Corbalán, F. y Sanz, G. (2011) "La Conquista del Azar". EDITEC.
-data1.pdf
- Figueroa, Norberto. ¿Qué hay detrás de Big Data?. En
- García Ferrando, M. (1989). Socioestadística. Introducción a la estadística en Sociología. Alianza Editorial. Madrid.
- Grant, R. (2019). Data visualization: Charts, maps and interactive graphics.
- Grima, P., "La certeza absoluta y otras ficciones" (2011), EDITEC.
<https://articulosit.files.wordpress.com/2013/03/que-hay-detrc3a1s-de-big>
- Huff, D. (2015). Cómo mentir con estadísticas. Barcelona: Crítica.
- Samaja, J. (1999). Epistemología y metodología: Elementos para una teoría de la investigación científica. Buenos Aires: EUDEBA. Parte 1.
- VVAA "El Atlas de la Educación" (2019). Capital Intelectual.