

LICENCIATURA EN  
**CIENCIA  
DE DATOS**

**PROGRAMA**

Big data y la Salud  
Dr. Sebastián Simondi

## 1. Cantidad de horas semanales y totales

Horas semanales: 6 hs. (seis horas)

Horas totales: 96 hs. (noventa y seis)

## 2. Nombres de las/los integrantes del equipo docente

Dr. Sebastián Simondi

## 3. Fundamentación

Desde hace unos años los avances tecnológicos en diferentes sectores y en especial en el de la Salud han hecho posible la recolección de una gran cantidad de datos relacionados con los pacientes, sus patologías y los centros sanitarios donde son tratados. Éstos datos se pueden obtenerse de diferentes fuentes tales como informes epidemiológicos, historias clínicas electrónicas, dispositivos de telemedicina e incluso de celulares. Procesar, curar y analizar toda esa información puede ayudar a la toma de decisiones por parte de los profesionales de la ciencia de la salud y de los gestores de centros sanitarios públicos y privados. Es en este contexto que surge la necesidad de aplicar técnicas denominadas de Big Data que permiten inferir una capa de inteligencia, en la que resulta de especial relevancia la utilización de modelos predictivos que ayuden por ejemplo a la detección temprana de enfermedades, a reducir los tiempos de investigación médica, o bien, avanzar hacia una medicina personalizada utilizando los datos de pacientes.

La propuesta de este taller es tomar grandes bases de datos reales y comenzar con el análisis de las mismas a partir de los conocimientos adquiridos por los estudiantes en materias que han cursado de la carrera y según el interés particular de cada uno. Una vez comenzado el análisis se reforzará y complementará sus conocimientos con nuevas técnicas de Big Data.

## 4. Programa sintético

Este taller de Big Data y Salud se propone integrar todas las herramientas de programación, inferencia estadística y matemática que los estudiantes han adquirido en diversas materias de la carrera y complementarlas con nuevas técnicas de Bigdata para abordar bases de datos reales. Comenzaremos con la exploración de las bases y trabajaremos en el análisis, visualización y presentación en diferentes gráficos para organizar, describir y comprender cada tipo de datos presentado. Continuaremos con un análisis estadístico que nos permita identificar patrones y tendencias que nos permita proporcionar información objetiva y cuantitativa para la toma de decisiones. En esta etapa utilizaremos distintos softwares de alto rendimiento tales como R y Python. Para finalizar estudiaremos diversas estrategias para comunicar los resultados obtenidos tales como tablas, gráficos y otros medios audiovisuales para la información sea más accesible y fácil de entender.

## 5. Objetivos

- Integrar y aplicar los conocimientos adquiridos a lo largo de la carrera por los estudiantes a grandes bases de datos provenientes de la Ciencia de la Salud.
- Generar presentaciones dinámicas y accesibles de los resultados obtenidos, como así también informes técnicos específicos.

## 6. Propósitos de la enseñanza

El propósito principal de este Taller, por un lado, es brindar a los estudiantes un espacio en donde aplicar todos los conocimientos acumulado a lo largo de la carrera a bases de datos específicas y reales. Además, adquirir nuevas herramientas a medida que surjan nuevas necesidades a partir de los casos de estudios. Por otro lado, brindaremos especial interés a la comunicación efectiva de los resultados obtenidos, tanto a un público técnico como a uno no especializados.

## 7. Contenidos

### **Unidad 1: El concepto de Big Data en las Ciencia de la Salud.**

El concepto de Bigdata: antecedentes y actualidad. Evolución histórica del concepto. Diferencias entre los conceptos de Data Mining, Business Intelligence, Data Lake, Big Data. Aplicaciones de Big Data en la Salud. Desafíos y oportunidades.

### **Unidad 2: Visualización y comunicación de la información.**

Visualización de la información de las bases de datos en diferentes softwares de alto rendimiento, tales como R y Python. Geolocalización de datos y presentación de información espacial. Comunicación efectiva de los resultados obtenidos en esta primera etapa de exploración.

### **Unidad 3: Análisis Estadístico.**

Aplicación de técnicas de asociación y de análisis multivariado al análisis de los datos proveniente de la Ciencia de la Salud. Conceptos generales. Ordenamientos. Clasificación e Inferencia. Sesgos y errores frecuentes. Adecuada interpretación y comunicación de los resultados.

### **Unidad 4: Análisis de casos de estudio.**

Aplicación a Casos de Estudio de los temas desarrollados en las unidades anteriores.

Calderón, C. L. P. (2016). Big data in health in Spain: now is the time for a national strategy. *Gaceta sanitaria*, 30(1), 63-65.

Cuenca, G. M., & Oliván, J. A. S. (2018). Del CMBD al Big Data en salud: un sistema de información hospitalaria para el siglo XXI. *Scire: representación y organización del conocimiento*, 77-89.

Malvicino, F. (2017). Big Data aplicada al sector salud en Argentina. Definiendo una agenda de trabajo. Documento de Trabajo. Centro Interdisciplinario de Estudios en Ciencia Tecnología e Innovación (CIECTI-MinCyT). Buenos Aires.

Lobo, J. F. C. (2020). Big data: utilidad en ciencias de la salud. *GICOS: Revista del Grupo de Investigaciones en Comunidad y Salud*, 5(1), 5-5.

Lv, Z., & Qiao, L. (2020). Analysis of healthcare big data. *Future Generation Computer Systems*, 109, 103-110.

Raghupathi, W., Raghupathi, V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst* 2, 3 (2014). <https://doi.org/10.1186/2047-2501-2-3>

Sesmero, J. M. M., & de Toledo, H. Aplicación del Big Data en Salud: aplicaciones, oportunidades y sesgos.

## 9. Metodología

La cursada del Taller será en forma virtual. Cada encuentro semanal comenzará con exposiciones teóricas-prácticas de los diferentes temas por parte del docente, en donde se ejemplificará los contenidos en bases de datos concretas. Luego, en una segunda de la clase, cada estudiante compartirá los avances realizado en sus casas, mostrando las aplicaciones de los conceptos abordados en las clases anteriores. Todo esto, además de servir como evaluación continua, tendrá como objetivo principal el armado del trabajo final donde deberá integrarse todos los contenidos abordados en el Taller.

## 10. Uso del campus virtual e integración de TIC en la propuesta pedagógica

El uso del campus virtual será esencial para el dictado de este Taller, ya que el mismo se realizará a distancia. En él se subirá el material bibliográfico, se realizarán las consultas a través del foro, se publicarán los links de los encuentros y se subirán las grabaciones de las clases. Además, será el medio oficial de comunicación

## 11. Evaluación

La evaluación se realizará de manera continua, a partir de los avances presentado por cada estudiante a lo largo del cursado de la materia. Por otra parte, la aprobación del Taller se obtendrá con la realización de una presentación para un público no especializado y un informe técnico de los resultados obtenidos a partir de los avances obtenidos en el desarrollo de cada una de las unidades temáticas.