

Diplomatura en Iniciación a la Programación y Análisis de Datos – UNIVERSIDAD DE BUENOS AIRES

a. **Denominación:** Diplomatura “Iniciación a la Programación y Análisis de Datos”

b.

c. **Objetivos**

El principal objetivo de la diplomatura es que los alumnos adquieran cierta familiaridad con la programación y el análisis de datos, entendiendo que es una herramienta poderosa y al alcance de todos para dar respuesta a diversos problemas.

Para eso, se buscará que los estudiantes obtengan una introducción al pensamiento computacional a través de un lenguaje de programación específico y fuertemente utilizado en el área disciplinar.

A su vez, se espera introducir a los estudiantes en el manejo de datos masivos en sus distintas etapas: obtención, procesamiento, visualización y análisis.

Es un objetivo central de esta diplomatura que los estudiantes sean capaces de aplicar los contenidos obtenidos a lo largo del curso para la resolución de una problemática real.

d. **Justificación**

La programación constituye una herramienta invaluable en el proceso de aprendizaje de muchas áreas del conocimiento. Una persona capaz de trabajar con algoritmos dispone de una formación que lo prepara para mucho más que escribir buenos programas, dispone de un instrumento de propósito general que será definitivo para su comprensión de muchísimos temas, y un potencial extraordinario para aplicarlo en diversas situaciones. La construcción de un programa requiere de una

profunda comprensión del problema y ayuda a ordenar el pensamiento.
Por otro lado, la capacidad de elaborar un programa informático resul

ser hoy una condición altamente valorada, y a veces imprescindible, para el desempeño profesional, laboral y académico.

Por otro lado, los datos están a nuestro alrededor, nos inundan y, en muchas ocasiones, nos ocultan la información que encierran. En los últimos tiempos, ha habido una importante explosión en herramientas tecnológicas y saberes científicos a través del acceso a grandes volúmenes de datos y del crecimiento de la capacidad de cómputo para procesarlos. El uso de estas herramientas permite resolver problemas complejos en una enorme diversidad de ámbitos tanto públicos como privados. Popularizar los conocimientos y la capacidad de usar la masividad de los datos que hay a nuestro alrededor para obtener información, resulta entonces una herramienta clave hacia una sociedad que brinde más oportunidades.

La diplomatura de “Iniciación a la Programación y Análisis de Datos” está pensada para estudiantes sin conocimientos previos de programación. Ha sido diseñada para brindar, a quienes lo completen, un primer encuentro con técnicas de programación y de análisis de datos. A lo largo de la diplomatura, sus estudiantes se familiarizarán con distintas herramientas del análisis de datos, incorporando metodologías típicas de visualización y descripción de la información, así como construcción de mapas y producción de aplicaciones interactivas. A la vez, se hará foco en la interpretación de gráficos comúnmente encontrados en la actualidad, haciendo foco en los errores más usuales de comprensión de los mismos y del uso del lenguaje técnico asociado a estos.

e. Pertinencia respecto a la unidad académica que la propone

Desde 2018 la Facultad de Ciencias Exactas y Naturales cuenta con la carrera de grado Licenciatura en Ciencia de Datos. Esta propone una fuerte formación científica en computación, matemática y datos. A su vez,

todas las carreras de la Facultad utilizan diversas herramientas informáticas.

La facultad lleva adelante desde 2018 “Exactas Programa”, destinado a estudiantes de la facultad de todas las carreras que quieran iniciarse en la programación. Estas experiencias permitieron explorar diversas estrategias para realizar un primer acercamiento a la programación y la ciencia de datos y conforman la base de esta diplomatura.

f. Estructura y modalidad de cursada

La Diplomatura está estructurada en dos (2) partes, cada una compuesta por cuatro (4) unidades.

Las clases de las unidades serán teórico-prácticas con exposición del equipo docente y trabajos grupales, apoyándose en casos de estudio y material bibliográfico pertinente. La duración de la Diplomatura es de 32 semanas, con una cursada de 4hs una vez a la semana, pudiendo complementarse virtual o asincrónicamente. Se busca que el uso de las nuevas tecnologías favorezcan el acompañamiento de las trayectorias del estudiantado desde un posicionamiento donde el espacio virtual genere aprendizaje colaborativo potenciando sus aprendizajes individuales desde la significatividad y la promoción de la autonomía.

g. Conteni

dos Parte 1

Unidad 1:

Python tutor: visualización de algunos conceptos instanciados en un lenguaje popular en Ciencias de Datos

Unidad 2:

Entornos de programación, introducción al modelado, programa, variables, listas, visualización de ejecuciones en python tutor. Ejecución de un programa. Funciones. Ciclos. Estructuras de Control.

Unidad 3:

Tipos de variables, lógica trivariada, operaciones disponibles con cada tipo de variables.

Unidad 4:

Manejo de archivos. Presentación de bibliotecas. Resolución de diferentes problemas para afianzar los conceptos presentados en las unidades 2 y 3. Nuevas estructuras.

Parte 2

Unidad 1:

Introducción a la ciencia de datos. Consideraciones prácticas y éticas. Distintos tipos de datos en ciencia de datos. Datos estructurados y no estructurados. Análisis exploratorio de datos numéricos. Procesamiento y preparación de datos numéricos. Lectura de archivos de datos.

Unidad 2

Estadística descriptiva. Medias, medianas, cuartiles, desvío estándar, error estándar.

Rudimentos de probabilidad y manejo de distribuciones.

Estandarización, normalización, detección y remoción de extremos, datos faltantes. Visualización de datos. Visualizaciones básicas y cuándo usarlos: gráfico de dispersión, gráfico de barras, histogramas, diagramas de cajas, violinplots, gráfico de torta y otros. Identificación de valores atípicos.

Unidad 3:

Manejo de información geográfica: formatos, sistemas de coordenadas, proyecciones.

Detección de patrones espaciales y temporales mediante la visualización de datos geográficos.

Unidad 4:

Obtención, preprocesado y limpieza de datos no numéricos. Expresiones regulares. Uso de interfaz de programación de aplicaciones (APIs), obtención de llaves y solicitudes. Manejo del portal datos.gob.ar. Scrapeo de páginas web. Visualización: Nubes de palabras

h. Carga horaria total: 128 Horas

i. Destinatarios y requisitos de ingreso

Esta diplomatura está destinada a cualquier persona con interés en introducirse la programación y el análisis de datos. Se requiere contar con Secundario completo

j. Cronograma de actividades:

Parte 1

Unidad	descripción	horas por clase	cantidad de clases	horas totales
1	Conceptos fundacionales de la programación	4	4	16
2	Programas y funciones	4	4	16

3	Tipos de variables y operaciones	4	4	16
4	Conceptos de modularización	4	4	16

Parte 2

Unidad	descripción	horas por clase	cantidad de clases	horas totales
1	Introducción a la ciencia de datos	4	4	16
2	estadística descriptiva	4	4	16
3	Manejo de información geográfica	4	4	16
4	Datos no numéricos	4	4	16

k. Curriculum vitae del Coordinador, quien deberá ser docente de la Universidad.

- **Esteban Mocskos**

Licenciado en Cs de la computación y Doctor por la Universidad de Buenos Aires, área

Cs de la computación. Profesor adjunto en el Departamento de Computación de la Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires. Investigador independiente en CONICET.

Director del Laboratorio Interdisciplinario de Computación de Alto Rendimiento (LICAR) con actividades en el Depto de Computación, Facultad de Ciencias Exactas y Naturales, UBA y Centro de

Simulación Computacional p/Aplic Tecnológicas (CSC-CONICET).
Coordinador de “Exactas Programa”

• **Lucía Pedraza**

Lic. en Cs. Matemáticas. FCEyN-UBA. Doctoranda en Cs. Físicas, período 2018 - 2023 (IFIBA, UBA-CONICET). JTP en el Departamento de Matemática de la FCEyN en el 2022, ayudante de 1ra en el 2018 y ayudante de 2da en el período 2013-2017, Ayudante de 2da en el Departamento de Matemática del CBC en el período 2014-2021. Coordinadora del Taller de Datos Populares. FCEyN (2021- 2022).

• **Diego Garbervetsky**

Licenciado en Cs de la computación y Doctor por la Universidad de Buenos Aires, área Cs de la computación. Es profesor asociado con dedicación exclusiva en Exactas/UBA e investigador Independiente del CONICET. Actualmente es el Director del Instituto UBA/CONICET de Investigación en Ciencias de la Computación (ICC) y consejero superior de la UBA. Fue Director del Departamento de Computación y Consejero por el claustro de Profesores en Exactas/UBA. Posee mas de 50 publicaciones en revistas y conferencias internacionales de renombre.

I. Nómina de colaboradores/docentes

- José Crespo
- Mariano Camilo González Lebrero
- Yamila Bauer
- Juliana Revés
- Tomás Cicchini

m. n. Modo de Evaluación y condiciones de aprobación:

Para aprobar la diplomatura se requerirá un 75% de asistencia, la entrega y aprobación de trabajos grupales/individuales (dependiendo de la cantidad de asistentes) donde se dé cuenta de los conocimientos adquiridos en cada parte. En caso de no estar aprobado se otorgará una instancia de compensación, donde deberá reelaborar el mismo con las orientaciones dadas en el tiempo y forma estipulada por dicho equipo.

