## DIPLOMATURA UNIVERSITARIA EN CIENCIA DE DATOS Y PYTHON

En esta diplomatura, aprenderemos cómo usar Python para ciencia de datos. Aprenderemos cómo recopilar datos, limpiar datos, hacer visualizaciones y construir modelos descriptivos y predictivos usando Python.

El objetivo principal es desarrollar habilidades analíticas y de programación para análisis de datos. Para lograrlo, vamos a resolver situaciones problemáticas aplicadas y proyectos en los cuales se pondrán en práctica todos los conceptos de programación utilizados en ciencia de datos.

La herramienta de aplicación será el lenguaje Python con sus principales librerías utilizadas en ciencia de datos, como Pandas, Numpy y Scikit Learn. Se las usará para realizar tareas del tipo que quienes trabajan con datos realizan a diario (como ser limpieza, visualización y recopilación de datos y creación de modelos).

Nuestro horizonte es que en la era de la información, los datos nos rodean por todas partes y en volúmenes inmensos. Dentro de estos datos hay respuestas a preguntas de diferentes sectores sociales (política, negocios, ciencia, etc.). Seguro que teniendo acceso a un gran conjunto de datos y pudiéndolos analizar, seríamos capaces de encontrar las respuestas a muchos problemas.

FUNDAMENTACION

Python es un lenguaje multiplataforma que puede ser programado mediante Programación Orientada a Objetos. Su sintaxis es simple, distinguiéndose de otros lenguajes como Java, C, PHP y JavaScript en la ausencia de paréntesis y la no incorporación de punto y coma al final de cada instrucción, entre otros aspectos. Es un lenguaje muy completo, con soporte para bases de datos relacionales, no relacionales u orientados a objetos, capaz de incorporar ORM´s, con frameworks de desarrollo muy complejos.

En este trayecto formativo sentaremos las bases del lenguaje aprendiendo a trabajar con sus diferentes tipos de objetos y bases de datos.

Se espera que los participantes incorporen criterios para diseñar y utilizar repositorios de datos.

OBJETIVOS GENERALES

Formar profesionales en Ciencia de Datos que:

* + Puedan adaptarse a diferentes contextos de disciplinas, servicios y trabajo
  + Comprendan y puedan aplicar estructuras básicas de la programación, en particular orientada al análisis de bases de datos, usando Phyton como lenguaje de programación.
  + Puedan diseñar un repositorio de datos.
  + Puedan reconocer la pertinencia de herramientas del análisis estadístico

¿Qué se espera aprender?

* Procedimientos básicos en un proyecto que implique herramientas de ciencia de datos.
* Cuadernos Python y Jupyter
* Cómo manipular y analizar conjuntos de datos.
* Realizar análisis estadísticos básicos e implementar métodos de aprendizaje automático.
* Limpieza, depuración y análisis de conjuntos de datos

DESTINATARIOS

Esta diplomatura está dirigida a todos los interesados en ingresar al mundo de la ciencia de datos. Es conveniente una experiencia previa con cualquier lenguaje de programación (Java, C, Pascal, Fortran, C ++, R, Python, PHP, etc.).

DURACIÓN Y MODALIDAD

6 meses de cursado. 3h semanales.

Cantidad de horas totales: 140 hs. reloj Teórico-Prácticas, distribuidas en actividades presenciales y asincrónicas, más un Proyecto Final Integrador.

Modalidad de cursado mixta, mediante plataforma educativa Moodle y en la UTN Concordia Salta 277.

LUGAR

El ámbito de coordinación y supervisión, preparación de materiales didácticos, recursos tecnológicos, construcción y gestión de las aulas virtuales y de las plataformas de simulación on-line, será en la Facultad Regional Concordia, al igual que las clases presenciales.

CANTIDAD MÁXIMA DE ESTUDIANTES

El cupo máximo por cada Cohorte será de 50 (cincuenta) participantes. Podrá habilitarse varias cohortes simultáneamente, dependiendo de la disponibilidad de docentes y de los recursos necesarios.

REQUISITOS DE ADMISIÓN E INSCRIPCIÓN

Escuela secundaria terminada. Es recomendable conocimientos previos de programación.

PLAN DE ESTUDIOS

Diseño curricular para dos cuatrimestres enfocado en Ciencia de Datos con Python:

Módulo 1: **Introducción a la programación con Python**

Introducción a Python y su sintaxis básica

Estructuras de control de flujo

Funciones y módulos en Python

Programación orientada a objetos en Python

Módulo 2: **Fundamentos de estadística y probabilidad**.

Conceptos básicos de estadística: descriptiva e inferencial. Medidas de centralización y dispersión. Representaciones gráficas.

Probabilidad: conceptos básicos, distribuciones de probabilidad. Simulación de experimentos aleatorios. Introducción a los Métodos estadísticos.

Modelización estadística: regresión lineal, regresión logística. Ejemplos de modelización.

Módulo 3: **Manipulación de datos con Python**

Manejo de datos en Python: estructuras de datos, archivos. Introducción a SQL.

Librerías para manipulación de datos: NumPy, Pandas.

Exploración y visualización de datos, limpieza, normalización, transformación con Python: Matplotlib, Seaborn.

Introducción a Tableau, QlikSense, Splunk, QlikView, Power BI

Módulo 4: **Introducción al aprendizaje automático**

Técnicas de pre procesamiento de datos.

Introducción a técnicas de aprendizaje supervisado y no supervisado con scikit learn.

Módulo 5: **Proyecto de ciencia de datos con Python**

Desarrollo de un proyecto de ciencia de datos utilizando Python y las técnicas aprendidas en los módulos anteriores. El proyecto puede ser seleccionado por los estudiantes o propuesto por el docente.

Este diseño curricular está enfocado en proporcionar una base sólida en programación con Python, estadística y manipulación de datos, así como en técnicas de aprendizaje automático y minería de datos. El proyecto final busca que los estudiantes apliquen los conocimientos adquiridos en una tarea práctica y real, permitiéndoles desarrollar habilidades en resolución de problemas y trabajo en equipo.

EVALUACIÓN Y APROBACIÓN

Para aprobar cada módulo, se deberá cumplimentar con los siguientes requisitos:

75% de asistencia a los encuentros sincrónicos.

Contestación del 100% de los cuestionarios y demás actividades propuestas por cada docente durante el desarrollo del módulo, con un puntaje mínimo de 6(seis) en cada uno.

Presentación y defensa de un proyecto integrador del módulo.

DIRECCIÓN Y COORDINACIÓN

Director de la carrera:

Coordinación: Ing. Carlos Blanc

CERTIFICADOS

Se entregará el certificado de “Diplomado en Ciencia de Datos con Python”, a aquellos participantes que tengan aprobados los 5 módulos, cumplan con la asistencia y presenten y defiendan su proyecto.