

UdeCataluña

Diplomado

Big Data Data Science con Python

Latinoamérica
Educación sin Fronteras

Diplomado en **Big Data - Data Science con Python**

Muchas organizaciones aún no conocen **el poder de los datos y el enorme potencial que tienen estos en la toma de decisiones**. Esta gran cantidad de información es una revolución dentro de las empresas por **su volumen, volatilidad y complejidad**. Gracias al sorprendente desarrollo de la computación, hoy disponemos de la capacidad de **procesar grandes volúmenes de información ofreciendo una nueva visión de la realidad** en términos de: patrones, tendencias o asociaciones; especialmente en lo relacionado con el comportamiento de los individuos y sus interacciones. Este conocimiento es el objeto de formación en nuestro Diplomado de Big Data y Data Science.

Actualmente los profesionales de Big Data son ampliamente solicitados; **están llamados a liderar la transformación del sector empresarial** aplicando el potencial de estas nuevas tecnologías.



Duración:
16 semanas



Modalidad: 100% Virtual
con clases en vivo +
acompañamiento académico



Dedicación:
8/10 horas por semana
Tiempo de trabajo del estudiante.

¡Ofrece soluciones Big Data!

Mediante este diplomado al estudiante se le proporcionarán **los fundamentos del análisis, la analítica empresarial (BA), inteligencia de negocios (BI) y la arquitectura de tecnología**, así como el diseño, desarrollo y gobernabilidad de soluciones de Big Data para empresas en diferentes campos de acción: el periodismo de datos, ciudades inteligentes, internet de las cosas, salud, administración, etc.

El diplomado tiene como fin alcanzar los siguientes **objetivos**:

01

Ofrecer al estudiante una aproximación global y práctica de Big Data, en la cual, se afiancen los conocimientos necesarios en el análisis y gestión de datos para su comprensión, así como las herramientas y estrategias que le permitan entender y rentabilizar el uso de Big Data.

02

Comprender las perspectivas a nivel económico y tecnológico en soluciones de Big Data.

03

Adquirir conocimientos sobre las diferentes soluciones de análisis para obtener beneficios en cualquier entorno empresarial.

Perfil del aspirante:

Profesionales en Estadística, Economía, Administración de Empresas, Matemáticas, Ingenierías o que cuenten con capacidades básicas en métodos cuantitativos y/o experiencia laboral en ámbitos donde el manejo de datos se convierte en información estratégica.

Conocimientos previos:

El aspirante al programa debe tener **conocimientos básicos en estadística** y conocimientos en el **uso y aplicación de métodos cuantitativos** que le permitan a través del uso de datos analizar, investigar y comprobar la información objeto de estudio.

Perfil ocupacional:

El profesional que se certifique en el diplomado podrá participar en proyectos de datos que involucren:

- Manejo de grandes volúmenes de datos.
- Minería de datos.
- Implementación y análisis de estadísticas empresariales.
- Inteligencia de negocios.
- Analítica empresarial.



Metodología

Formación 100% virtual, con enfoque constructivista que permite al estudiante ser protagonista de su aprendizaje. Nuestro esquema de Aula Invertida permite que el estudiante adquiera el conocimiento a su ritmo, para desarrollar aspectos prácticos bajo la orientación de los conferencistas expertos en las Clases Presenciales Virtuales. El proceso se complementa con talleres, foros, pruebas formativas, actividades de retroalimentación y asesorías para garantizar un completo aprendizaje.

Nuestra metodología incorpora estrategias de aprendizaje como:



Aula invertida

Comprende la revisión y análisis previo de materiales **por parte de los estudiantes**.



Encuentros sincrónicos

Se llevan a cabo entre docentes y estudiantes para **profundizar, aclarar y retroalimentar temas específicos**.



Talleres y trabajo colaborativo

Mediante actividades interactivas, casos de estudio y desarrollo de trabajos colaborativos, **se aplican los conceptos adquiridos**.



Actividades de retroalimentación

El estudiante verifica su aprendizaje y desarrolla actividades que le permiten **construir de manera autónoma su propio aprendizaje**.

Certificado

Los estudiantes que finalicen el programa con un cumplimiento igual o superior al 80% de las actividades programadas se les otorgará, por parte de **UdeCataluña**, el certificado en:

Diplomado en Big Data - Data Science con Python



Módulo 1

Introducción a Data Science y Big Data

Objetivo:

Conocer las ventajas del aprovechamiento de datos en el ambiente empresarial, el concepto de Data Science y Big Data, sus características, combinaciones y crecimiento. Comprender las generalidades de la metodología CRISP-DM.

- Datos, información y conocimiento en la toma de decisiones.
- Introducción y conceptualización de Big Data.
- Las Vs de Big Data.
- Roles y habilidades técnicas en el análisis de datos.
- Diferencias entre Big Data y Data Science.
- Machine Learning e Inteligencia Artificial.
- Taller práctico - Recolección de datos.

Módulo 3

Análisis Descriptivo y Exploratorio

Objetivo:

Presentar algunas técnicas de análisis descriptivo y exploratorio de datos relevantes dentro del proceso de entendimiento de la información. Además, trabajar algunas técnicas de visualización para la presentación de resultados del análisis de grandes conjuntos de datos.

- Manejo de data frames con Pandas.
- Medidas de tendencia central, de posición o localización.
- Introducción a la visualización de datos con Matplotlib y Seaborn.
- Análisis de tendencias y correlación.
- Introducción a Power BI Pro.
- Taller práctico – Exploración y visualización de datos.

Plan de Estudios

Módulo 2

Introducción a la Programación con Python

Objetivo:

Dominar los conceptos básicos de programación necesarios para el análisis de datos empleando Python.

- Introducción a Google Colaboratory.
- Variables y tipos de objetos.
- Operadores lógicos y de comparación.
- Condicionales y funciones.
- Bucles o iteradores.
- Estructuras de datos.
- Carga de archivos.
- Introducción a Markdown.
- Retos prácticos de programación.

Módulo 4

Limpieza y Transformación con Python

Objetivo:

Comprender la importancia de la limpieza y transformación de datos dentro del proceso CRISP-DM, además de poner en práctica diversas metodologías para la limpieza y transformación con datos reales.

- Detección y corrección de errores, valores faltantes y valores atípicos.
- Transformación de variables: normalizar, escalar, numerizar y/o discretizar.

Módulo 5

Machine Learning con Scikit-Learn

Objetivo:

Profundizar en la presentación, tanto teórica como práctica, de técnicas de modelamiento derivadas de la estadística o de Machine Learning empleando la librería especializada scikit-learn.

- Concepto de Machine Learning y técnicas de aprendizaje computacional.
- Tipos de aprendizaje computacional: aprendizaje supervisado y no supervisado.
- Tipos de problemas computacionales: regresión, clasificación y agrupamiento.
- Algoritmos de Machine Learning.
- Vecinos más cercanos.
- Árboles de decisión y bosque aleatorio.
- Regresión lineal y logística.
- Redes neuronales artificiales.
- Taller práctico – Aprendizaje computacional y evaluación de modelos.

UdeCataluña se reserva el derecho de apertura o aplazamiento de los cursos, seminarios, diplomados en caso de no contar con el número mínimo de inscritos. El grupo docente estará sujeto a cambios según disponibilidad de su agenda al igual que el cronograma de actividades académicas.

Módulo 6

Arquitectura y Herramientas Big Data - Spark y Mongo DB

Objetivo:

Examinar las diferentes herramientas Big Data que logren sacar el máximo valor y aspectos fundamentales en el análisis de datos procedentes de diversas fuentes.

- Arquitectura relacional – Introducción a SQL.
- Taller práctico de SQL.
- Ecosistema Big Data.
- Almacenamiento Big Data – Tecnologías NO-SQL.
- Taller práctico de MongoDB.
- Procesamiento Big Data - Apache Spark.
- Taller práctico de Spark.

Módulo 7

Herramientas Complementarias - Reproducibilidad y Agilismo

Objetivo:

Desarrollar habilidades complementarias para desarrollar proyectos de análisis y aprovechamiento de datos.

- Reproducibilidad y Técnicas de Implementación.
- Pilares del agilismo y metodología SCRUM.



Valores agregados

- ✓ Incluye uso de librerías como Numpy, Pandas, Matplotlib, Seaborn, Scikit-Learn.
- ✓ Orientación de un proyecto de analítica avanzada a través de la metodología CRISP-DM.
- ✓ Domina los conceptos básicos de programación necesarios para el análisis de datos empleando Python.
- ✓ Uso de Herramientas Big Data como Pyspark y SparkSQL en Databricks.
- ✓ Incluye uso de Inteligencia Artificial Generativa como herramienta de productividad en la explotación de Big Data y caso de uso de Prompt Engineering en Ciencia de Datos
- ✓ Incluye material complementario en lenguaje R (videos explicativos, talleres guiados, Notebooks).
- ✓ Incluye Acreditación de Competencias mediante Insignia Digital (emite UdeCataluña y Acreditta).
- ✓ Incluye curso complementario autoformativo (40 horas certificadas) en Fundamentos de Programación con Python.
- ✓ Incluye curso complementario autoformativo (40 horas certificadas) a elección según la oferta académica de la universidad. **(oferta cursos complementarios, ¡click aquí!)**





Conferencistas

Julián Rojas Díaz

Profesional de la Pontificia Universidad Javeriana. Experto en Ciencias de datos con certificaciones internacionales entre ellas: Data Engineer de la University of Chicago, Data Science Professional Certificate y Data Engineering Professional Certificate de IBM.

Actualmente Data Scientist & Pricing associate en Melonn y Data Scientist en PREDIK Data-Driven.

Amplia experiencia en Data analysis, Machine Learning, Data visualization, Storytelling, Cloud services, Analytics Engineering y en herramientas como: Python, SQL, DBT, PostgreSQL, Excel, Tableau, Power BI, Hive, AWS (Athena, S3, Sagemaker, Glue), ApacheSpark, Presto, Shell. Docente universitario.

Harold Uribe Romero

Profesional de la Universidad Nacional de Colombia. Científico y Analista de Datos con +8 años de experiencia. Experto en Power BI y Tableau y SQL.

Amplia experiencia en Machine Learning; Desarrollo e implementación de modelos predictivos utilizando Python (Pandas, Seaborn, NumPy, scikit-learn) gestionado su despliegue efectivo a través de Flask y Docker.

Conocimientos avanzados en plataformas como AWS y GCP, reducir el churn, optimizar KPIs y proporcionar insights que permitan a las empresas y negocios tomar decisiones informadas y basadas en datos. Actualmente, Business Intelligence (BI) and Data Coordinator en Liberty Latin America. Docente universitario.

Mejora tus competencias con UdeCataluña



Manejas tu tiempo

Accede a los contenidos y clases grabadas cuando desees.



Desde cualquier dispositivo

Descarga el material de estudios a tu PC, tablet o teléfono móvil.



Expertos temáticos

Aprende de docentes certificados y con amplia experiencia.

UdeCataluña

Bogotá, Colombia.

Cra 18 B No. 106 A - 15
+57 601 919 6734

Miami: +1 (786) 292 42 31

info@udecatalunya.edu.co

www.udecatalunya.edu.co

 @unicatalunya_

 UdeCataluña

 /uceduccion

 /UniversidadCatalunya

Los recursos obtenidos de este programa nos permiten ofrecer programas de pregrado sin costo para el estudiante (Beca del 100%).

www.ucatalunya.edu.co/pregrado/ingenieria-informatica
www.ucatalunya.edu.co/pregrado/administracion-de-empresas.

UdeCataluña es una Institución de Educación Superior, vigilada por el MEN. Resolución No 21329 de 2016.