



# Data Science para profesionistas: Ciencia de Datos

Tomar decisiones, generar reportes, crear visualizaciones, responder preguntas y resolver problemas sobre bases de datos con los lenguajes R y Python, haciendo uso de diferentes librerías, herramientas y paquetes para el análisis de datos como Pandas, Numpy, Matplotlib, entre otras. Aplicar modelos de regresión, clasificación y predicción en un proyecto de Machine Learning.



Curso de  
**8 meses**



Acceso a  
**Plataforma BEDU**



Adéntrate al mundo del análisis de datos; conoce los fundamentos de la programación para crear tus tableros, hacer informes y presentar resultados.

### Conocimientos / Habilidades recomendadas

- Pensamiento lógico
- Manejo de computadoras

### Requerimientos técnicos

- Windows 7 64bits / dual core / 4gb RAM
- Mac OS X Yosemite / dual core / 4gb RAM

### Programas y tecnologías que aprenderás



\*Los programas a aprender pueden variar según la línea de estudio que se decida tomar.

# Temario

## Introductorio

### Nivel uno

#### Temario

1. Introducción a estadística para Data Análisis
2. Excel para Análisis de Datos
3. Python desde cero
4. SQL para principiantes

### **Módulo I** | Introducción a Bases de Datos Nivel dos

**Objetivo del módulo.** Generar un conjunto de datos a partir de la manipulación de bases de datos relacionales y no relacionales para su posterior análisis, con el fin de seguir buenas prácticas para diseñar y ejecutar poderosas y eficientes consultas SQL y NoSQL.

#### Temario

1. Fundamentos de SQL
2. Agrupaciones y subconsultas
3. Joins y Vistas
4. Configuración de Base de Datos locales
5. Fundamentos de MongoDB
6. Consultas en MongoDB
7. Agregaciones
8. Query competition

### **Módulo II** | Programación y Estadística con R Nivel tres

**Objetivo del módulo.** Recopilar información valiosa de una base de datos para alcanzar una mayor comprensión y entendimiento para la toma de decisiones. Utilizar bases de datos obtenidas mediante MySQL y Mongo con el objetivo de realizar limpieza, manipulación y análisis en R. Crearás visualizaciones, predicciones y modelaje matemático.

#### Temario

1. Introducción a R y Software
2. Programación y manipulación de datos en R
3. Análisis Exploratorio de Datos (AED o EDA) con R
4. Algunas distribuciones, Teorema del Límite Central y Contraste de Hipótesis
5. Regresión lineal y Clasificación
6. Series de tiempo
7. RStudio Cloud y conexiones con BDs
8. Dashboards con Shiny, GUI

### **Módulo III** | Procesamiento de Datos con Python Nivel tres

**Objetivo del módulo.** Crear rutinas y scripts utilizando Python para automatizar la obtención, limpieza, procesamiento y transformación de datos provenientes de diferentes fuentes y formatos como: archivos locales, APIs y Bases de Datos externas con el fin de tener data en condiciones óptimas para su futuro análisis y visualización.

#### Temario

1. Fundamentos de Python
2. Estructuras de Datos y Funciones
3. Programación funcional, operadores lógicos y funciones lambda
4. Pandas y Análisis Exploratorio de Datos
5. Funciones Vectorizadas y Limpieza de Datos
6. APIs, automatización y concatenación de DataFrames
7. Transformación, filtración y reordenamiento de datos
8. Bases de datos, merge y agrupaciones

### **Módulo IV** | Análisis de Datos con Python Nivel cuatro

**Objetivo del módulo.** Utilizar Python y sus librerías para realizar análisis robustos de datos, aplicando modelos estadísticos y matemáticos que permitan encontrar patrones y relaciones en los datos con el fin de generar visualizaciones de análisis univariados, bivariados y multivariados con Seaborn y Matplotlib y aplicar modelos de regresión, clasificación y predicción.

#### Temario

1. Estimados de Locación y Variabilidad
2. Introducción a la visualización de datos: Distribuciones
3. Exploración de Variables Categóricas y Análisis Multivariable
4. Correlaciones y Regresión Linear Simple
5. Distribuciones Muestrales y Técnicas de Evaluación de Modelos
6. Visualización de Datos Avanzada
7. Pruebas A/B y Procesamiento de Lenguaje Natural
8. Introducción a Machine Learning: Clasificación No Supervisada y Supervisada

### **Módulo V** | Machine Learning Nivel cuatro

**Objetivo del módulo.** Identificar, seleccionar, construir e implementar algoritmos de Machine Learning con Python, con el objetivo de identificar si un problema amerita utilizar Machine Learning, seleccionar el mejor algoritmo, construirlo, implementarlo y probar su eficacia.

#### Temario

1. Introducción a Machine Learning
2. Separación, validación y evaluación para algoritmos de ML
3. Algoritmos no supervisados
4. Problemas de regresión
5. Árboles de decisión
6. Clasificación y redes neuronales artificiales
7. Arena de clasificadores
8. Fundamentos de procesamiento de imágenes





# ¡Prepárate para retar tu potencial!

 /BeduOrg

 @Bedu\_Org

 @Bedu\_org

 Bedu

