

## Fundamentación

El análisis de datos se ha convertido en una disciplina esencial en un mundo donde la generación de datos es cada vez más masiva y omnipresente. Las organizaciones, desde las grandes corporaciones hasta las pequeñas empresas, requieren habilidades para transformar datos en información valiosa que les permita tomar decisiones estratégicas, innovar en productos y servicios, y resolver problemas complejos.

Este curso tiene como objetivo dotar a los estudiantes de las herramientas y metodologías necesarias para interpretar, procesar y presentar datos de manera efectiva. A lo largo del curso, se abordarán desde los conceptos fundamentales del análisis de datos hasta aplicaciones más complejas como el **Machine Learning** y el análisis de **series temporales**. Además, se hará especial énfasis en la **visualización de datos** y en la creación de narrativas efectivas basadas en datos (**Storytelling**), lo que permitirá a los estudiantes no solo interpretar los datos, sino también comunicar sus hallazgos de forma clara y persuasiva.

El curso está diseñado de forma modular, comenzando con los fundamentos y avanzando hacia aplicaciones prácticas. Los estudiantes aprenderán a manejar y analizar datos utilizando herramientas y lenguajes como **Python**, **SQL** y **Power BI**, mientras que el trabajo práctico y el proyecto final les permitirán aplicar lo aprendido en un caso real.

Al final del curso, los estudiantes estarán preparados para enfrentar desafíos en el análisis de datos en diversos sectores, tales como **finanzas**, **salud**, **transporte**, **marketing**, entre otros. Este conocimiento es clave en el actual entorno profesional, donde los datos impulsan decisiones clave y estrategias empresariales.

## Contenido del Curso

### Módulo 1: Introducción al Análisis de Datos

- **Clase 1: Introducción al Mundo de los Datos**
  - Historia y evolución del análisis de datos.
  - Importancia y aplicaciones en diversos campos.
  - Tipos de datos y fuentes de datos.
- **Clase 2: Herramientas y Procesos del Análisis de Datos**
  - Herramientas (Python, SQL, Power BI).
  - Ciclo de vida.
  - Ética.

### Módulo 2: Repaso Integrado de Python y SQL

- **Clase 1: Programación Orientada a Objetos en Python**
  - Fundamentos de POO.
  - Principios de POO.
- **Clase 2: Consultas Avanzadas en SQL**
  - JOINS avanzados.
  - Subconsultas.
  - Vistas.
  - Funciones personalizadas.
- **Clase 3: Laboratorio Práctico**
  - Objetivo del laboratorio.
  - Descripción del caso.
  - Requisitos del sistema.

- Implementación paso a paso.

### Módulo 3: Manipulación de Datos con Pandas

- **Clase 1: Introducción a EDA, Series y DataFrames**
  - Exploratory Data Analysis (EDA).
  - Creación y manipulación de Series y DataFrames.
- **Clase 2: Operaciones con DataFrames, Indexado y Filtrado**
  - Operaciones básicas y avanzadas.
  - Técnicas de indexado y filtrado.
- **Clase 3: Manejo de Fechas y Datos Ausentes**
  - Manejo de fechas y tiempos con Pandas.
  - Tratamiento de datos faltantes.
- **Clase 4: Limpieza de Datos y Muestras Aleatorias**
  - Técnicas de limpieza de datos.
  - Obtención de muestras aleatorias.
- **Clase 5: Combinación de Conjuntos**
  - Fusión y combinación de datasets.
  - Laboratorio práctico con Pandas.

### Módulo 4: Estadística Descriptiva

- **Clase 1: Medidas de Resumen para Variables Cuantitativas**
  - Media, mediana, moda.
  - Varianza y desvío estándar.
- **Clase 2: Correlación y Causalidad, Correlación en SQL**
  - Diferencias entre correlación y causalidad.
  - Aplicación de correlación en bases de datos SQL.

### Módulo 5: Presentación de Datos con Python

- **Clase 1: Creación de Figuras y Estilos**
  - Figuras y paneles.
  - Personalización de estilos.
- **Clase 2: Gráficos Avanzados**
  - Leyendas, etiquetas y títulos.
  - Tipos de gráficos avanzados.
- **Clase 3: Presentación del Análisis**
  - Presentación efectiva de datos y análisis.
  - Laboratorio práctico de visualización.

### Módulo 6: Visualización de Datos con Power BI

- **Clase 1: Introducción a Power BI**
  - Navegación y uso de Power BI.
- **Clase 2: Limpieza y Transformación de Datos**
  - Técnicas de limpieza y preparación de datos en Power BI.
- **Clase 3: Informes y Dashboards**
  - Creación de informes y paneles interactivos.
  - Uso de funciones DAX.
  - Laboratorio práctico creando un dashboard.

## Módulo 7: Storytelling con Datos

- **Clase 1: Fundamentos del Storytelling**
  - Qué es el Storytelling aplicado a datos.
  - Principios de comunicación efectiva.
- **Clase 2: Diseño de Visualizaciones Impactantes**
  - Uso del color, jerarquía y enfoque visual.
- **Clase 3: Construcción de Historias Basadas en Datos**
  - Técnicas para estructurar narrativas con datos.
  - Laboratorio práctico contando una historia.

## Módulo 8: Machine Learning Básico

- **Clase 1: Introducción a Machine Learning**
  - Tipos de aprendizaje: supervisado y no supervisado.
  - Procesos clave: limpieza, modelado y validación.
- **Clase 2: Modelos Predictivos**
  - Regresión lineal y logística.
  - Clasificación básica.

## Módulo 9: Proyecto Final

- **Propuesta:**
  - El estudiante seleccionará un área temática de su interés (como finanzas, salud, transporte, retail, etc.) para desarrollar un proyecto aplicado que integre todos los conocimientos adquiridos..
- **Componentes:**
  - **Limpieza y preparación de datos:** tratamiento de inconsistencias, datos faltantes y transformación de variables.
  - **Visualización avanzada:** creación de dashboards o gráficos significativos con Python o Power BI.
  - **Modelado opcional:** aplicación de modelos básicos de Machine Learning o predicciones de series temporales, según el enfoque del proyecto.
- **Presentación:**
  - El proyecto será presentado a través de una **narrativa de storytelling basada en datos**, utilizando Power BI o Python, con foco en la interpretación de resultados y en la comunicación clara de insights.

## Bibliografía

1. McKinney, W. (2017). *Python for Data Analysis: Data wrangling with pandas, numpy, and IPython*. O'Reilly Media.
2. Tanimura, C. (2021). *SQL for Data Analysis: Advanced techniques for transforming data into insights*. O'Reilly Media.
3. VanderPlas, J. (2016). *Python Data Science Handbook: Essential tools for working with data*. O'Reilly Media.
4. Knaflitz, C. N. (2015). *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Wiley.
5. Müller, A. C., & Guido, S. (2016). *Introduction to Machine Learning with Python: A guide for data scientists*. O'Reilly Media.
6. Sweigart, A. (2019). *Automate the Boring Stuff with Python: Practical programming for total beginners*. No Starch Press.

7. Wheelan, C. (2013). *Naked Statistics: Stripping the dread from the data*. W.W. Norton & Company.