



## Curso de Análisis de datos en Python

# Syllabus

## Equipo Docente

### Alejandro Cataldo

Doctor en Ciencias de la Ingeniería. Pontificia Universidad Católica de Chile, 2018

Profesor asistente del Instituto de Ingeniería Matemática y Computacional de la Pontificia Universidad Católica de Chile. Su línea de investigación es la Investigación de Operaciones, Optimización y Programación Estocástica.

## Descripción del curso

En este curso los estudiantes aprenderán el uso adecuado y eficiente de Python como herramienta computacional para el análisis de datos, y su rol en la toma de decisiones dentro de una organización. Para ello se introducirán los conceptos básicos y generales sobre el manejo y análisis de datos, se explicará la importancia de una correcta visualización de los datos y de los resultados de los análisis que se realicen con estos, y se explicará en qué consisten los análisis descriptivos, predictivos, y prescriptivos, enseñando diferentes métodos para llevar a cabo estos análisis. El curso utilizará exposiciones a través de clases sincrónicas, análisis de casos, test basados en la interacción con bases de datos, contenidos entregados por las video clases, lecturas y reflexión individual a través de foros.

## Resultados de aprendizaje

- Describir el rol del uso de herramientas computacionales en el quehacer actual de las organizaciones, para desarrollar técnicas de ciencia de datos a través de Python.
- Aplicar conceptos y herramientas básicas de programación en Python, como lectura, escritura de archivos y edición masiva de datos, resolviendo problemas reales relacionados con el procesamiento y limpieza de datos.
- Distinguir el uso de la analítica descriptiva y prescriptiva para aplicar pronósticos en escenarios futuros inciertos y apoyar procesos de toma de decisiones.
- Aplicar algoritmos básicos de aprendizaje de máquina para el análisis descriptivo.
- Explicar la importancia de una visualización efectiva de los datos., que permiten evidenciar oportunidades y justificar decisiones en una organización.

## Estrategias metodológicas

- Exposiciones a través de clases sincrónicas
- Análisis de casos
- Test basados en la interacción con bases de datos
- Contenidos entregados por las video clases
- Lecturas y reflexión individual a través de foros.

## Estructura del Curso

### Módulo 1. Ciencia de Datos y Python

- Bienvenida e Introducción
- ¿Qué es la Ciencia de Datos?, ¿Por qué buscar evidencias en los datos?
- Información, datos, decisiones y programación, ¿cómo se integran estos conceptos?
- Secuencias de comandos en Python y Operaciones básicas

## Módulo 2. Manejo de Base de Datos y manejo de los datos

- Conceptos fundamentales de las Bases de Datos
- Conectores y librerías Python para trabajo con Bases de Datos relacionales
- Extracción directa de información desde un programa Python
- Librería de Python y el concepto de Dataframe

## Módulo 3. Visualización avanzada de Datos

- ¿Por qué es importante una visualización adecuada de los datos y resultados?
- ¿Cómo crear gráficos en Python?
- ¿Cómo crear una interface adecuada?
- Ejemplo práctico

## Módulo 4. Analítica descriptiva y predictiva

- ¿Qué es, por qué es importante y cómo se modela la incertidumbre?
- ¿Qué es la analítica descriptiva y predictiva?
- Cómo se utiliza Python en la construcción de modelos predictivos
- Decisiones bajo certidumbre, riesgo e incertidumbre

El curso tiene como requisitos de aprobación las siguientes instancias evaluativas:

Actividad	Evaluación
Tareas	40%
Controles	30%
Foros	10%
Trabajo Final	20%

## Información General

CURSO:	ANÁLISIS DE DATOS EN PYTHON
TRADUCCIÓN:	Data Analysis in Python
SIGLA:	MDS3010
CRÉDITOS:	5 UC
HORAS:	90 horas de dedicación total (24 directas y 66 indirectas)
REQUISITOS:	INF3103
RESTRICCIONES:	(Programa MDS) o (Programa MAN)
CONECTOR:	0
CARÁCTER:	Mínimo
TIPO:	Cátedra, Taller
CALIFICACIÓN:	Estándar
PALABRAS CLAVE:	Análisis de datos (Data analysis), Framework de análisis de datos
NIVEL FORMATIVO:	Magíster

## Bibliografía

### Mínima

- Kotu, V., & Deshpande, B., (2018). Data Science: Concept and Practice. Second Edition, Morgan Kaufmann.
- Severance, C. (2020). Python para Todos: Explorando la información con Python. Independently Published.
- Bertsimas, D., & Freund, R.M., (2008). Data. Models and Decisions. The Fundamentals of Management Science. Second Edition, Dynamic Ideas.
- Davenport, T., & Harris, J. (2017). Competing on analytics: The new science of winning. Harvard Business Press.

## Sugerida

- Winston, W. (2019). Microsoft Excel 2019 Data Analysis and Business Modeling: Data Analysis and Business Modeling. Pearson Education Limited.