

### **Módulo 1. Big Data (4 horas)**

- ❖ ¿Qué es Big Data y por qué ha influido tanto en la última década?
- ❖ 5 V del Big Data.
- ❖ Herramientas para el procesamiento de grandes volúmenes de datos.
- ❖ Importancia de las visualizaciones en la presentación de conclusiones y hallazgos de análisis de datos.
- ❖ Computación en la nube como factor de éxito en el Big Data.

### **Módulo 2. Repaso de conceptos Python (6 horas)**

- ❖ Uso de Colaboratory de Google con python.
- ❖ Tipos de datos.
- ❖ Entrada y salida de información por medio de consola.
- ❖ Bloques condicionales.
- ❖ Bloques de interacción.
- ❖ Instalación de librerías externas.

### **Módulo 3. Librería Pandas (16 horas)**

- ❖ Generalidades, origen de la librería y principales usos.
- ❖ Tipos de datos soportados por Pandas.
- ❖ Transformación de tipos de datos.
- ❖ Carga de datos de diferentes orígenes.
- ❖ Selección de filas y columnas.
- ❖ Filtrado de datos.
- ❖ Filtrado condicional de datos.
- ❖ Tratamiento de datos faltantes.
- ❖ Agrupaciones condicionales de datos.
- ❖ Herramientas básicas de visualización de Pandas.
- ❖ Combinación lógica de diferentes fuentes de datos en una sola fuente.
- ❖ Manejo de campos de tipo fecha.
- ❖ Manejo básico de series temporales.
- ❖ Repaso estadística básica con Pandas.

#### **Módulo 4. Librería plotly (6 horas)**

- ❖ Generalidades, origen de la librería y principales usos.
- ❖ Visualizaciones de puntos.
- ❖ Visualizaciones de líneas.
- ❖ Visualizaciones de barras e histogramas.
- ❖ Visualizaciones circulares.
- ❖ Visualizaciones especiales en 3D y mapas de calor.

#### **Módulo 5. Proyectos propuestos (4 horas)**

- ❖ Análisis exploratorio de datos.
- ❖ Preparación de datos para algoritmos de machine learning.
- ❖ Detección de anomalías.