

### **Módulo 1. Repaso de conceptos básicos de python (6 horas)**

- ❖ Instalación de entorno Anaconda.
- ❖ Creación de entornos virtuales y el uso de los mismos.
- ❖ Programación en notebooks.
- ❖ Programación de scripts clases y métodos (conceptos básicos de programación orientada a objetos).
- ❖ Tipos de datos en python.

### **Módulo 2. Manipulación de datos con pandas (20 horas)**

- ❖ Pandas como librería para manipulación de datos.
- ❖ Lectura de archivos desde diferentes fuentes, csv, xls, SQL, txt, json y lectura desde url.
- ❖ Manipulación básica para poder entender los datos que se leyeron en el apartado anterior.
- ❖ Visualizaciones básicas para entender el comportamiento estadístico básico de los datos.
- ❖ Reemplazo de valores faltantes o fuera de rango: entender el concepto de outlier.
- ❖ Creación de conjuntos y subconjuntos condicionados.
- ❖ Agrupaciones de datos y operaciones con datos agrupados.
- ❖ Agrupaciones de varias fuentes de datos para creación de una única fuente consolidada (en esta parte se verán las operaciones de conjuntos aplicadas a la manipulación de datos).
- ❖ Correlación estadística.

### **Módulo 3. Aprendizaje supervisado (7 horas)**

- ❖ Regresión lineal.
- ❖ Regresión lineal múltiple.
- ❖ Validación de la regresión lineal.
- ❖ Regresión logística.
- ❖ Árboles de decisión.

**Módulo 4. Aprendizaje no supervisado (7 horas)**

- ❖ Clustering.
- ❖ Análisis de componentes principales.
- ❖ Kmeans.

**Duración: 40 Horas**

