

# Recapitulando

## ¿En qué consiste Data Manipulation?

Se trata del proceso que nos permitirá comprender y asegurar la calidad de nuestros datos. De esta manera podremos extraer los mejores resultados de nuestro análisis. Es uno de los pasos más importantes de Data Analytics ya que la calidad de los datos serán determinantes para la calidad de nuestros resultados.

El proceso consiste en tres etapas denominadas ETL. En un primer paso extraemos los datos necesarios para nuestro análisis, luego haremos las transformaciones requeridas para el procesamiento, por ejemplo eliminar valores extremos, y por último cargamos los datos listos para analizarlos en un data warehouse.

---

## Y no te olvides, siempre que vayas a comenzar un proceso de Data Analytics, preguntate:

- ¿Qué datos necesito para mi análisis? ¿Cómo puedo consultarlos?
- ¿Qué tipos de datos son? ¿Qué aporte hacen a nuestro análisis?
- ¿Es necesario hacer alguna transformación en los datos para poder interpretarlos?
- ¿Hay valores que deba excluir para el análisis?
- ¿Las transformaciones se ejecutaron correctamente?
- ¿El data warehouse está listo para aplicar un modelo de datos?