Estadistica descriptiva

# Analisis estadistico de datos

* **Analisis exploratorio**: objetivo resumir, representar y explicar los datos concretos de los que disponemos.
* **Análisis inferencial**: Inferir a partir de esos datos

**Objetivo final**: entender los datos lo mejor posible.

## Tipos de datos

* **Datos de tipo atributo**: Expresan una cualidad del individo. En R las listas de datos cuantitativos en vectores o factores si vamos a clasificar individuos.
* **Datos ordinales**: Similares a los cualitativos, se pueden ordenar de manera natural. Ejemplo: calificaciones en un control (En R se recomienda guardar como factor)
* **Datos cuantitativos**: edad, longitudes. En R guardaremos las listas de datos cuantivativos en vectores numéricos.

### Datos cualitativos

**Cualitativos**: corresponden a observaciones sobre cualidades de un objeto o individuo.

Suelen codificarse por medio de palabras, pero también se pueden usar números que jueguen papel de etiquetas. Es habitual representar No (falso, fracaso, ausente) con un 0 y si (verdadero, exito, presente) con un 1.

Los datos cuantitativos son aquellos que pueden ser iguales o diferentes, pero que no se admiten ningún otro tipo de comparación significativa.

Por lo tanto un mismo conjunto de datos puede ser cualitativo o de otro tipo segun el análisis que vayamos a hacer de él.

* **Variable cualitativa**: lista de observaciones de un tipo de datos cualitativos sobre un conjunto de objetos.
* **Niveles**: diferentes valores que pueden tomar estos datos. La variable sexo tiene dos niveles: M(Macho) y H(Hembra), o sinónimos.

En R se usa vectores y factores para representar variables cualitativas. Los factores nos servirán para agrupar las observaciones según los niveles de la variable. De esta manera podremos segmentar la población que representa la variable en grupos o subpoblaciones, asignando un grupo a cada nivel, y podremos comparar el comportamiento de otras variables sobre estos grupos.

#### Estudio de Frecuencias

Una variable cualitativa, para cada uno de sus niveles podemos contar cuantos datos hay en ese nivel(frecuencia absoluta) y que fración del toal representa.

* **frecuencia absoluta**: , del nivel , en esta variable cualitativa es el número de observaciones en las que toma el valor .
* **frecuencia relativa**: del nivel en esta variable cualitativa es la fraccion: