

1. ESCOGER DATOS → #observaciones, # atributos numéricos
- KAGGLE
  - UCI datasets
  - Citihub
  - web ...
  - DATOS PROPIOS
- etiqueta de clase

## 2. EXPLORACIÓN CON PANDAS

- FILTRAR
- ESTADÍSTICA
- PRESENCIA DE NANS
- IMPUTACIÓN
- VISUALIZACIÓN

## 3. REDUCCIÓN DE DIMENSIONALIDAD (PCA)

#componentes necesarios, proyección

## 4. CLUSTERING

- K-means
- Agglomerative
- GMM
- GMM + BIC → how many clusters?
- EVALUACIÓN (Adjusted Rand Score)

## 5. MODELOS DE REGRESIÓN

- UNIVARIADA
  - MULTIVARIADA
  - ÁRBORES DE DECISIÓN
  - EVALUACIÓN RESIDUOS
    - test de Normalidad (q-q plot)
- } SELECCIÓN  
ORDEN  
POLINÓMICO