

Visualización:  
¿Que se esconde en mis datos?

# El plan

1. Lectura de datos
2. Limpieza y preprocesamiento
3. Visualización y formulación de hipótesis.

# Funciones, objetos y documentación útil

- `sklearn.decomposition.PCA`
- `sklearn.decomposition.KernelPCA`
- `sklearn.manifold.Isomap`
- `sklearn.manifold.TSNE`
- `sklearn.preprocessing`

# Dataset de Iris

1. Cargue los datos de Iris usando sklearn
2. Aplique la reducción de dimensionalidad usando PCA conservando las primeras 2 dimensiones.
3. Grafique los datos resultantes utilizando la clase de cada fila como color en función de la primera y segunda componente y de la segunda y tercera.
4. Responda: ¿Son linealmente separables las clases en las dimensiones reducidas?

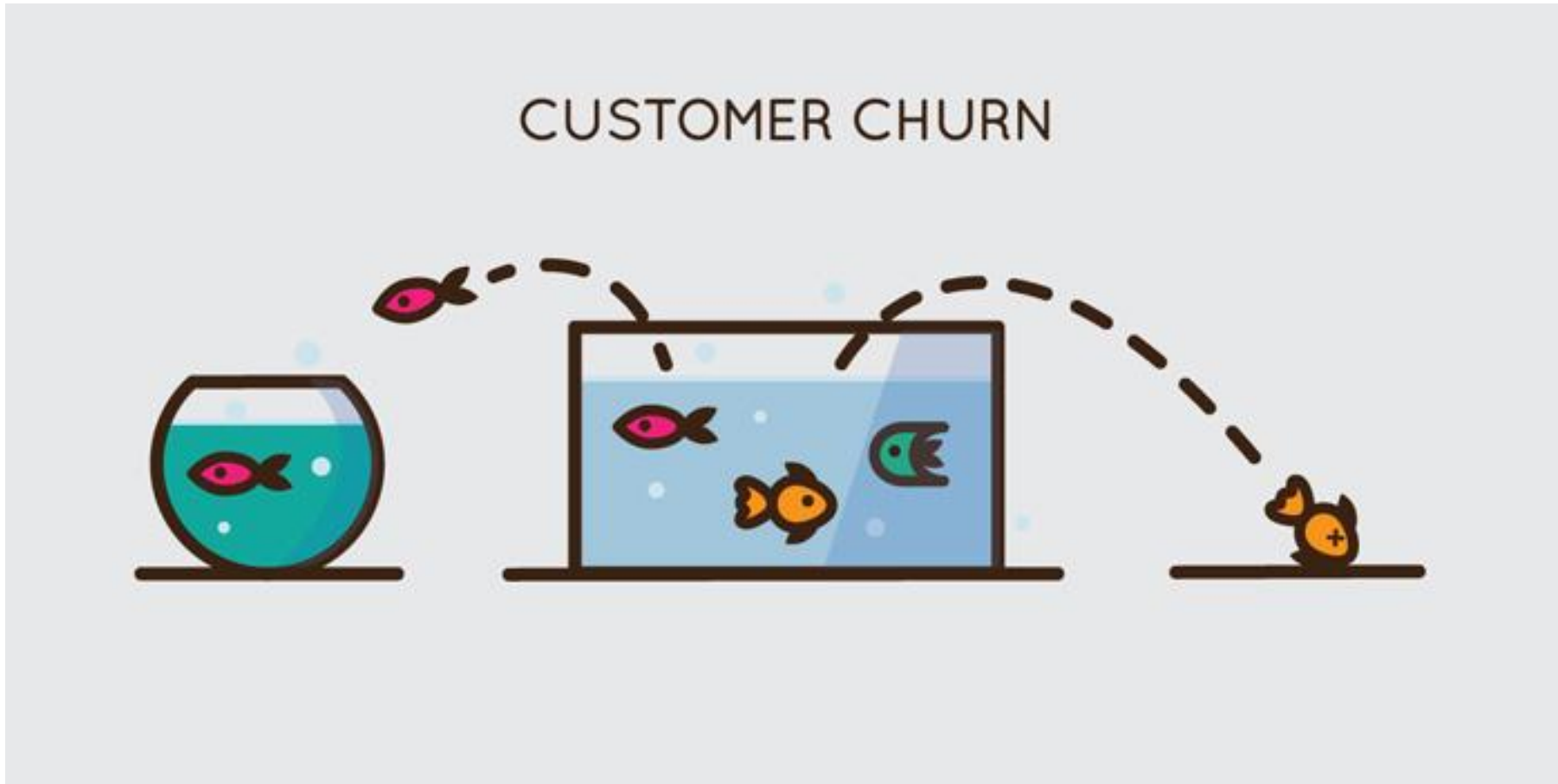
# Dataset de calidades de vinos tintos

1. Cargue los datos de calidad de vinos usando pandas, separe la columna de "quality" como la clase.
2. Aplique la reducción de dimensionalidad a 2D usando:
  - a. PCA
  - b. KernelPCA
  - c. Isomap
  - d. T-SNE

Experimente con diversos parámetros en el caso de Isomap, KernelPCA y T-SNE.

3. Grafique los datos resultantes utilizando la clase de cada fila como color.
4. Responda: ¿Son linealmente separables las clases en las dimensiones reducidas?
5. Normalice las columnas utilizando el StandardScaler de sklearn. Verifique las medias y desviaciones de los datos normalizados.
6. Repita los puntos 2, 3 y 4 utilizando los datos normalizados.

# Dataset de cancelación de clientes (churn)



# Datos de cancelación de clientes de una telefónica



# Datos de cancelación de clientes de una telefónica

- ¿Hace cuanto tiempo es cliente?
- Código de área + número de teléfono
- ¿Tiene un plan de llamadas internacional?
- ¿Tiene correo de voz?
- Número de mensajes en el correo de voz
- Número de minutos diarios
- Número de llamadas diarias
- Total de gasto diario
- ...
- ¿El cliente canceló su suscripción?



# Datos de cancelación de clientes de una telefónica

1. Cargue los datos de abandono de clientes usando pandas: elimine las columnas 'State','Area Code','Phone', separe la columna de "Churn?" cómo la clase.
2. Aplique la reducción de dimensionalidad a 2D usando:
  - a. PCA
  - b. KernelPCA
  - c. Isomap
  - d. T-SNE

Experimente con diversos parámetros en el caso de KernelPCA, Isomap y T-SNE.

3. Grafique los datos resultantes utilizando la clase de cada fila como color.
4. Responda: ¿Son linealmente separables las clases en las dimensiones reducidas?
5. Normalice las columnas utilizando el StandardScaler de sklearn. Verifique las medias y desviaciones de los datos normalizados
6. Repita los puntos 2, 3 y 4 utilizando los datos normalizados.