# EJERCIOS DE MACHINE LEARNING CON PYTHON

# Ej01: Visualizar un gráfico de barras apiladas

Con el fichero drinks.csv crear una gráfica de barras apiladas. En el eje X expresamos los continentes y en el y las porciones de vino, cerveza y licor.

# Ej02: Crear un boxplot y graficar una regresión

Cargar la base de datos drinks.csv y graficar un boxplot que represente la cantidad de porciones de vino consumidas por continente.

A continuación, graficar un regresión entre las porciones de vino que consume un país y las porciones de cerveza.

# Ej03: Explorar un dataframe

Cargar la base de datos cars.csv (no tiene cabecera), a continuación agrupar los coches por marca y sacar el consumo de carburante en ciudad de cada modelo y su precio.

# Ej04: Explorar un dataframe 2

Con la base del ejercicio 3 crear una nueva columna en la que se agrupen los precios de los coches en tres categorías de precio (Bajo: 1 – 5000 €, Medio: 5001 – 10000 €, Alto: 10001 - 45000 €).

A continuación, crear un gráfico de barras con la cantidad de coches por cada categoría de precio.

Recuerda que los precios están en dólares.

# Ej05: Exploración de datos

Con los datos del csv autos cumplir los siguiente objetivos:

1 - Crea un diagrama de dispersión para ver si existe correlación entre el Precio y el Tamaño del Motor.

2 - Realizar con el método de Correlación Pearson si existe correlación o fortaleza entre RPM y el Precio.

3 - Realizar un Test para establecer si existe diferencia sustancial entre el precio Medio de las Marcas de Autos Audi y alfa-romero.

# Ej06: Generar un modelo lineal simple y múltiple

Con la base de datos cars.csv:

* Crear un modelo lineal simple: Precio – Ancho de motor.
* Crear un modelo lineal múltiple: Precio – Peso de motor, consumo ciudad, tamaño de motor.

# Ej07: Generar un modelo múltiple

Al igual que en el ejercicio anterior, crear un modelo múltiple con las variables que desees.

# Ej08: Generar modelo polinómico simple

Generar un modelo polinómico simple con la base de datos cars.

# Ej09: Generar modelos polinómicos simples

Generar un modelos polinómicos simples con la base de datos drinks. Generar modelos que expliquen el consumo de cerveza, vino y licor en función del consumo del alcohol puro.

# Ej10: Modelo polinómico simple

Generar un modelo polinómico múltiple en el que las variables cerveza, vino y licor expliquen el consumo de alcohol puro.

# Ej11: Modelo lineal múltiple con entrenamiento

Crear un modelo múltiple para explicar el precio de la base de datos cars.

Utilizar un entrenamiento del 85%.

Conseguir un coeficiente R2 mayor a 0.79.

# Ej12: Modelo polinómico simple con entrenamiento

Generar un modelo polinómico simple de tercer orden para explicar el precio de los coches en función de la cantidad de caballos.

Usar un 70% de los datos para el entrenamiento.

# Ej13: Utilizar el método Ridge

Generar un modelo polinómico múltiple que explique el precio de los coches. Con este modelo aplicar el método Ridge y comparar distintos valores de alfa con R2.

# Ej14: Modelar con el fichero del PIB de China

Con la base de datos del PIB de China modelar un modelo de regresión lineal y un modelo de regresión polinómico. Graficar los modelos y los valores predictivos con los reales.

# Ej15: ANOVA

Con el fichero NASDAP realizar un test ANOVA para comparar los valores de cierre de las empresas AAME y AACC.

# Ej16: Modelar la base de datos Nasdaq

Con la base de datos Nasdaq modelar la variable stock volume.

Seleccionar dos variables explicativas y componer un modelo arista.

# Ej17: K vecinos

Emplear el modelo de K vecinos cercanos para explicar la variable cuscat de la base de datos telecust.

Encontrar el valor de K que genere un mejor modelo.

# Ej18: Árbol de decisión

Crear un árbol de decisiones con el fichero drug200.csv. Utilizar las varibales edad, colesterol y sexo como variables independientes y el tipo de medicamento como dependiente.

Generar un gráfico con el esquema conseguido.

# Ej19: Recapitulación.

01 -> Modelo de regresión múltiple: Explicar en un modelo de regresión múltiple la variable Alcohol Puro del fichero drinks.csv. Crear 2 gráficos de cajas que muestren los valores reales de consumo de alcohol por continente y otro que represente los valores predichos.

02 -> Crear un modelo de regresión polinomial múltiple con el fichero autos.csv. Utilizar las variables Caballos, Tamaño del motor, Consumo en Autopista como variables independientes para predecir el precio.

03 -> Crear un modelo arista para dividir los datos, la variable dependiente sería el precio y las independiente los Caballos, el Tamaño del Motor, el Consumo en Autopista, el Tipo de Combustible y la Cantidad de Puertas.

Graficar los valores reales frente a los predichos.

04 -> Crear un algoritmo de K-vecinos cercanos con el fichero telecust. Emplear todas las variables independientes. Averiguar cual es el valor de K más adecuado.

05 -> Crear un árbol de clasificación con el fichero drug200 y visualizar el esquema.