

Ejercicios de auto-comprobación

NumPy

Ejercicio 1.

- Crea un array de 45 enteros aleatorios entre -10 y 10.
- Rediménsionalo como matriz de 5x7
- Obtén los números positivos de esa matriz utilizando boolean indexing.
- Obtén con una función los datos no duplicados.
- Sobre los datos no duplicados, obtén la media.

Ejercicio 2.

- Crea dos arrays de 100 elementos con números aleatorios de distribuciones diferentes, pero en un rango de valores similar.
- Dibuja sus histogramas y compáralos.
- Busca en `scipy.stats` una función para hacer una comprobación (mediante un t-test) de si la media de los elementos en los dos arrays es diferente.

Pandas

Ejercicio 1.

- Carga en un DataFrame este dataset:
<https://archive.ics.uci.edu/ml/datasets/Automobile>
- Comprueba si tiene datos faltantes y trátalos de la manera que creas más conveniente.
- Obtén la longitud y anchuras medias por cada categoría ("body-style").
- Analiza el atributo peso ("curb-weight") y dibuja su histograma para apreciar si sigue alguna distribución concreta.
- Analiza cuál es el número de cilindros que se da con más frecuencia en el dataset.
- Obtén una lista ordenada de las diferentes marcas de coches.

Ejercicio 2.

- Carga en un DataFrame este dataset:
<https://archive.ics.uci.edu/ml/datasets/Census+Income>
- Estudia si hay variables como el nivel educativo o el país que influyen en que se tuviese un salario mayor o menor de 50K. Puedes simplemente obtener la cuenta de las veces que sucede una u otra cosa de los diferentes grupos.