

EXAME PARCIAL DE MARZO DE SISTEMAS DE BIG DATA. CURSO: 2023/24**INSTRUCCIÓNs:**

- Debes **entregar un arquivo notebook** que faga tódalas accións.
- No código debes empregar a notación recomendada por **PEP8**.
- **Podes empregar:** os apuntes de clase (incluíndo aula virtual) e o código que teñas entregado e mesmo o github da asignatura, tamén a URL dos datos, **pero NADA máis**.
- Debes empregar como ferramentas: miniconda, jupyterlab, vscode e as dependencias aprendidas.

Empregando selenium, scrapea aos fuxitivos máis buscados de Europa da páxina **(7 Puntos)**:

<https://eumostwanted.eu/es>

Pista: Quizás debas empregar un selector por URL (tag a).

De cada fuxitivo (de cada páxina) recolle: Nome do sospeitoso, país que o busca, data de nacemento, nacionalidade, etnia, idiomas falados.

Carga os datos en pandas e salva a formato parquet. (1 Punto)

Indica mediante consultas de PANDAS:

1. Cantos fuxitivos busca cada país. (1 Punto)
2. A idade media, a mediana da idade e os valores extremos. (1 Punto)