

**Proceso selectivo para ingreso en la escala de Técnicos Superiores Especializados.
Cuarto ejercicio.**

**Programa Técnico de Datos: Minería, Gestión y Análisis de Big Data
en Sistemas Complejos**

- Cada candidato dispone de un ordenador personal para la realización de este ejercicio. Usuario: ifisc Password: ifisc
- El entorno de python anaconda está instalado.
- El ejercicio A se realizará desde el directorio /home/ifisc/ejercicioA y el B desde /home/ifisc/ejercicioB.
- Scripts, notebooks, programas, figuras, etc. que se hayan preparado durante la realización del ejercicio deben permanecer en el directorio al final del mismo.
- Se describirá por escrito en las hojas autocopiativas proporcionadas las tareas realizadas y las respuestas a las cuestiones concretas.

Ejercicio A: Obtención de datos de la Agencia Española de Meteorología (AEMET) usando su API, inserción en mongoDB y realización de consultas a la base de datos.

- Se usará la web opendata de la AEMET es <https://opendata.aemet.es>.
- Para acceder a los datos, usar la clave /home/ifisc/ejercicioA/aemet.key
- Documentación de la API: <https://opendata.aemet.es/centrodedescargas/AEMETApi>
- Lista de datasets disponibles: <https://opendata.aemet.es/dist/index.html?>. Se usará el dataset valores-climatologicos.
- Hay instalada la configuración estándar de mongoDB, a la cual se puede acceder desde shell o desde una aplicación con el path mongoddb://localhost:27017/

Descargar los valores diarios climatológicos de la estación B278, PALMA DE MALLORCA, AEROPUERTO en los años 2019 y 2020 e insertarlos en la base de datos mongoDB.

Finalmente hacer las consultas apropiadas a la base de datos para indicar:

1. El número de días en 2019 con temperatura máxima por encima de los 26° C así como dicho número para 2020.
2. El número de días con horas de sol (campo "sun" en los datos retornados) igual o inferior a 7, en el año 2019 así como en el 2020.

Ejercicio B: Obtención de datos la pagina web de publicaciones del IFISC y análisis de los mismos.

- Se usará la web de publicaciones del IFISC <https://ifisc.uib-csic.es/en/publications/>

Obtener el título, autores y revista o lugar de publicación de las publicaciones del IFISC en los años 2019 y 2020.

Descartar aquellas que no correspondan a publicaciones en revistas o libros si no a tesis doctorales o de máster (para las cuales el lugar de publicación es “PhD Thesis” o “Master Thesis”, en lugar de ser el nombre de una revista).

Construir la red de coautores donde cada uno de los autores (sea o no personal del IFISC) es un nodo y los enlaces entre nodos se establecen por publicaciones conjuntas durante 2019-20. Graficar dicha red de manera que el color o el grosor de los enlaces sea proporcional al número de publicaciones conjuntas y guardar dicha figura como pdf o como captura de pantalla.

Así mismo hacer un histograma con el número de publicaciones en cada revista, graficarlo y guardar la gráfica en pdf o captura de pantalla.

Indicar:

1. El autor con el mayor número de publicaciones conjuntas y cuantas son.
2. La revista con el mayor número de publicaciones y cuantas son.