

Sistemas y tecnologías para la gestión de la información

Actividad Guiada 1 - Extracción y almacenamiento de datos

> Objetivos

- Descarga de datos
- Explorar repositorios Open Data
- Lectura y escritura de JSON/XML
- Inserción y consulta en MongoDB

> Actividades entregables (Búsqueda de datos)

- Hacer una descarga de un conjunto de datos del Open Data de: Valencia, Madrid o Barcelona. Hablar sobre la fuente de los datos y la calidad de los datos. (Mínimo del conjunto de datos 100 registros o filas).

> Actividades entregables (Fichero e inserción)

- Generar un JSON de los datos en un fichero en local.
- Insertar en MongoDB todos los datos sin utilizar un bucle. Definir colección y código.

> Actividades entregables (Consultas)

- Mediante una consulta a Mongo obtener unos campos concretos de una consulta buscando en un atributo un valor concreto. El valor tiene que ser un valor numérico.
- Hacer una consulta de Mongo de forma que nos devuelva, en función de un atributo, el top 3 ordenado. ¿Existe el atributo “coordinates” en el conjunto de datos? ¿Cómo se consulta?
- Recorrer todos los registros de la colección y realizar una media en un atributo.

> Entregables

- Scripts y código utilizado para realizar las actividades. Enlace al repositorio del proyecto en GIT.
- Documento/Informe (formato pdf) con explicación de la actividad: definición del conjunto de datos y repositorio, estrategia de programación, resultados de las actividades y conclusiones.

Gracias