

**UF1 - PROVA FINAL****Mòdul:** MP14 – Big Data**UF:** UF1 - Big Data**Professor:** Albert Guardiola**Data límit d'entrega:** 14/05/2023 23:59**Mètode d'entrega:** Per mitjà del Clickedu de l'assignatura. Les activitats entregades més enllà de la data límit només podran obtenir una nota de 5.**Instruccions:** Totes les tasques han de entregar-se en un sol document, de nom:*M14-UF1-PF-Nom\_Alumne***Resultats de l'aprenentatge:**

RA 1: Instal·la sistemes de dades massives (Big Data) garantint el seu funcionament i els requisits demanats.

RA 2: Programa pel Big data a partir de data sources.

RA3: Gestiona i explota sistemes de dades massives (Big Data) garantint la seva eficiència.

La Prueba Final de la UF1 del módulo M14 consiste en realizar un **proyecto individual de desarrollo de una aplicación web en el contexto de los datos masivos** (*Big Data*).

El proyecto debe:

- A: cumplir una serie de **requisitos técnicos**, que se indican en la rúbrica que se encuentra más abajo. Los requisitos técnicos se clasifican entre los de *obligado cumplimiento* y los *opcionales*.
- B: ir acompañado de un pequeño **dosier de proyecto** que documente aspectos relacionados con la justificación y documentación del proyecto. Los detalles de este dosier se indican en la rúbrica que se encuentra más abajo.

La parte más importante de este dosier será la *justificación del proyecto*: en ella se deben recoger los motivos, bien argumentados, por los que se ha escogido realizar el proyecto. Estos motivos pueden ser de índole personal o social, o ambas.

La entrega del proyecto pues, constará de dos partes, que se deberán entregar en un único fichero comprimido a través de Clickedu.

-A: **Enlace al repositorio del proyecto.**

-B: **PDF con el dosier de proyecto.**



El proyecto deberá exponerse, en las fechas indicadas (semanas del 15 y el 22 de mayo), ante el profesor. Cada alumno deberá demostrar:

- Que el proyecto responde a una finalidad bien clara.
- Que se han escogido datos de entrada relevantes para el proyecto.
- Que se tiene control sobre el código que implementa el proyecto.
- Que se tiene claro el alcance del proyecto y sus posibles correcciones, mejoras y ampliaciones ulteriores.

A continuación, se ofrece una rúbrica con los ítems obligatorios de los que debe constar el proyecto, así como algunas posibles ampliaciones opcionales.



### Rúbrica de evaluación:

	REQUISITOS TÉCNICOS (70%)	REQUISITOS DEL DOSIER DE PROYECTO (30%)
<b>OBLIGATORIOS</b> (max. 7 puntos)	<ul style="list-style-type: none"> <li>-Lenguaje de programación: <b>Python</b>.</li> <li>-Arquitectura de la aplicación: <b>servidor web</b> (Flask).</li> <li>-Uso de alguna <b>fuentes de datos abiertos</b> (<i>open data</i>).</li> <li>-Algún tipo de <b>análisis de los datos</b>, preferentemente a través de la librería <i>pandas</i>. El análisis de datos debe estar <u>convenientemente organizado en un pipeline</u> en el que se distinga: etapa de extracción, etapa de limpieza/acondicionamiento, etapa de análisis y etapa de visualización.</li> <li>-Uso coherente de las <b>estructuras de datos</b> básicas de Python: listas, diccionarios, maps... según convenga.</li> <li>-Paradigma funcional de programación (organización del código en <b>funciones</b>).</li> <li>-Uso de <b>base de datos</b> para la gestión de la persistencia de datos.</li> </ul>	<ul style="list-style-type: none"> <li>-<b>Justificación del proyecto</b>: explicación de las necesidades (personales y/o sociales) a las que viene a responder el proyecto.</li> <li>-<b>Explicación del uso que se hace de los datos masivos</b>: explicación de qué fuentes de datos masivos abiertos se utilizan, quién los proporciona, cómo se accede a ellas y de qué manera se analizan y presentan al usuario de la aplicación web.</li> <li>-<b>Referencia</b> a los materiales de terceros utilizados: origen de los fragmentos de código ajenos, del material multimedia incluido, de la información técnica consultada, etc.</li> <li>-Criterios formales de entrega: portada, índice, webografía. Corrección ortográfica i gramatical.</li> </ul>
<b>OPCIONALES</b> (max. 3 puntos)	<ul style="list-style-type: none"> <li>-Integración de gráficas <i>matplotlib</i> o <i>seaborn</i> en la aplicación.</li> <li>-Uso de algoritmos de aprendizaje automático para el análisis de datos.</li> <li>-Manejo de excepciones en el uso de recursos externos a la aplicación.</li> <li>-Uso de distintos módulos para organizar el código propio.</li> <li>-Alojamiento de la aplicación en algún servicio de <i>hosting</i> gratuito para Python (<i>pythonanywhere</i>, <i>render</i>...)</li> <li>-Interactividad con JavaScript: validación de formularios, programación de eventos...)</li> </ul>	<ul style="list-style-type: none"> <li>-Explicación detallada del código desarrollado.</li> <li>-Comentario sobre cómo se adaptaría la aplicación para un eventual procesamiento distribuido en <i>Spark</i>.</li> <li>-Lista de propuestas de mejora de la aplicación: corrección de errores (<i>fixes</i>), posibles nuevas funcionalidades, escalabilidad de la aplicación...</li> </ul>