

Tipología y fuentes de datos

PEC 4

UOC

Universitat Oberta
de Catalunya

[NOMBRE y APELLIDOS]

Fecha de entrega :
29 de diciembre de 2021

Tipología y fuentes de datos

PEC2

Esta PEC consiste en escribir el código necesario para ejecutar un *jupyter notebook* que acompaña a esta PEC y que contiene código python incompleto para que lo finalices adecuadamente según las especificaciones. Los fragmentos de código para completar están marcados con comentarios #TODO que deberán ser sustituidos por código válido python o consultas correctas, según corresponda.

El notebook está pensado para ejecutarse en el entorno Colab de Google (<https://colab.research.google.com/>) sin necesidad de ninguna instalación, pero puedes ejecutarlo localmente.

Parte 1 DBpedia [2 puntos]

Completando el código para acceder al punto SPARQL de la DBpedia para resolver el problema los problemas planteados con SPARQL

Parte 2 WikiData I [2.5 puntos].

Completar el código para consultar WikiData para resolver consultas acerca de programadores conocidos.

Parte 3 WikiData II [2 puntos].

Completar el código para resolver las cuestiones planteadas sobre capitales europeas y sus alcaldes.

Parte 3 Grafo de rutas aéreas [2.5 puntos].

En esta parte se van a tratar datos sobre aeropuertos y rutas aéreas como grafo para calcular diferentes métricas sobre ellos.

Parte 5 Introducción a los grafos de conocimiento [1 punto].

En esta parte se presenta cómo obtener las entidades y las relaciones presentes en un texto para poder obtener grafos de conocimiento que se pueden usar para representar el conocimiento.

Se pide:

- Entregar el notebook con el código y la ejecución.

Anexo

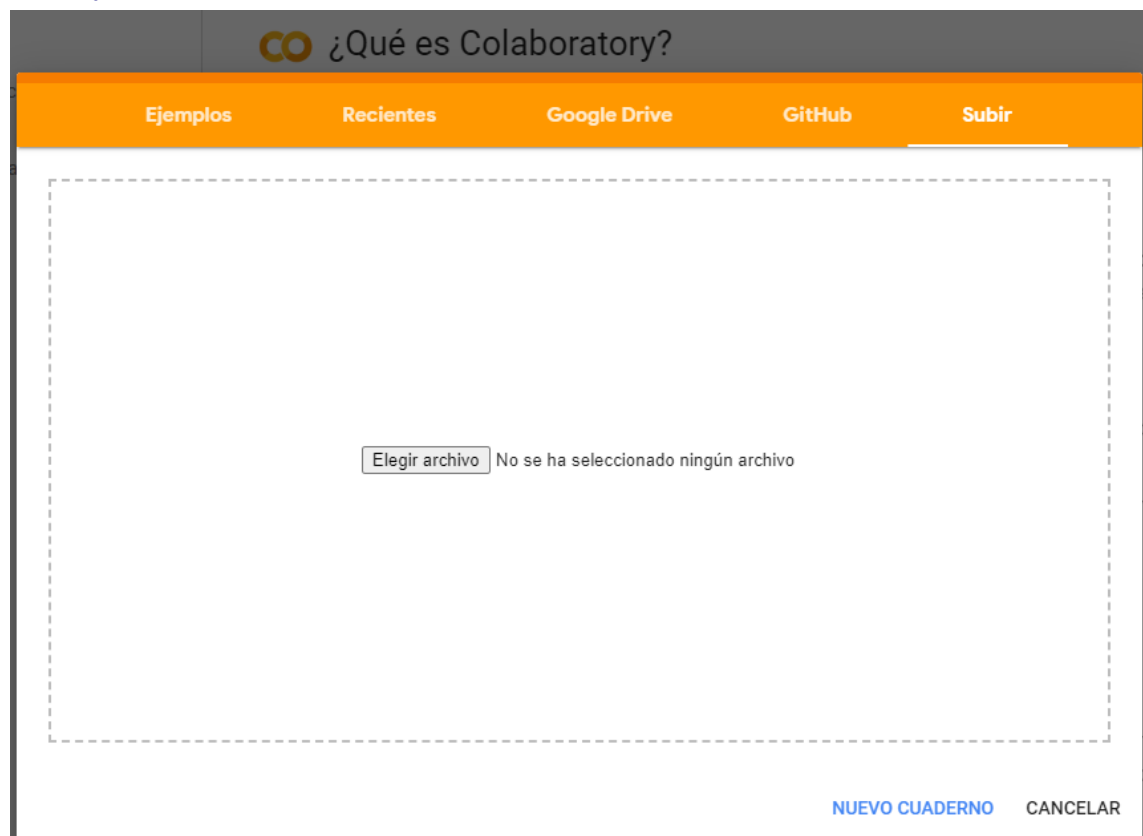
Descarga en tu equipo el notebook TyFdD-PEC4.ipynb. El notebook está pensado para que lo ejecutes en Google Colab.

Google Colab es un servicio en la nube que permite ejecutar Jupyter Notebooks accediendo con un navegador web. Tiene además las siguientes ventajas:

- Posibilidad de ejecución mediante GPUs
- Basado en jupyter notebook pudiendo crear y ejecutar libros en Python 2 o 3
- Tiene preinstaladas las librerías comunes usadas en ciencia de datos y la posibilidad de instalar otras.
- Enlaza con cuentas de Google Drive y desde github

Primero hay que entrar en sesión (login) con una cuenta de Google (la de la uoc debería funcionar).

Ahora ya se puede subir el notebook de la PEC a colab:



La ejecución del libro es exactamente igual que en cualquier jupyter notebook. Hay que pulsar Shift + Enter para que el código (python) se ejecute.

A partir de ahí debes completar el código python que falta que está marcado con **# TODO** y responder a las cuestiones planteadas escribiendo tanto las respuestas como el código con el que obtienes las respuestas.

Criterios de valoración

Cada uno de los apartados tiene un peso asignado en el total de la PEC. Se valorará, para cada apartado, la validez de la solución y la claridad de la argumentación.

Formato y fecha de entrega

Tenéis que enviar la PEC al buzón de Entrega y registro de EC disponible en el aula (apartado Evaluación). El formato del archivo que contiene vuestra solución puede ser .pdf, .odt, .doc y .docx. Para otras opciones, por favor, contactar previamente con vuestro profesor colaborador. El nombre del fichero debe contener el código de la asignatura, vuestro apellido y vuestro nombre, así como el número de actividad (PEC2). Por ejemplo *apellido1_nombre_tyfdd_pecX.pdf*. No olvidéis poner vuestro nombre y apellidos en el documento.

La fecha límite para entregar la PEC es **la indicada en la portada**.

Propiedad intelectual

Al presentar una práctica o PEC que haga uso de recursos ajenos, se tiene que presentar junto con ella un documento en que se detallen todos ellos, especificando el nombre de cada recurso, su autor, el lugar donde se obtuvo y su estatus legal: si la obra está protegida por el copyright o se acoge a alguna otra licencia de uso (Creative Commons, licencia GNU, GPL etc.). El estudiante tendrá que asegurarse que la licencia que sea no impide específicamente su uso en el marco de la práctica o PEC. En caso de no encontrar la información correspondiente tendrá que asumir que la obra está protegida por el copyright.

Será necesario, además, adjuntar los ficheros originales cuando las obras utilizadas sean digitales y su código fuente, si así corresponde.

