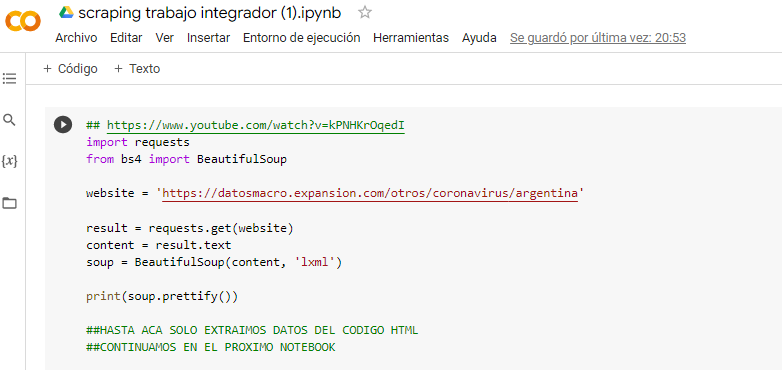
**PROYECTO WEB SCRAPING**

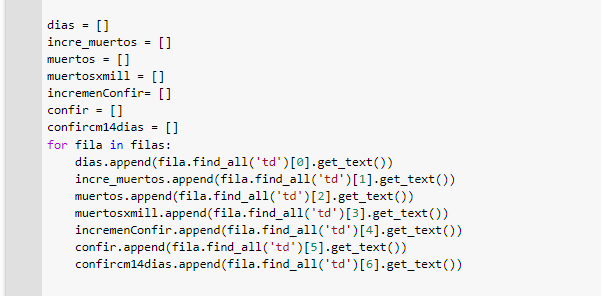
El actual desarrollo corresponde al PROYECTO INTEGRADOR I de la tecnicatura en ciencia de datos del ISPC.

En el presente proyecto se realiza el scraping de una página web para posteriormente generar un CSV con los datos obtenidos. El scraping los hicimos utilizando las librerías Beatifull Soup, request y pandas. La página web de donde se recolecta los datos es Datos Macros ('https://datosmacro.expansion.com/otros/coronavirus/argentina) y la información scrapeada corresponde a los datos de coronavirus de los últimos tres meses en argentina.

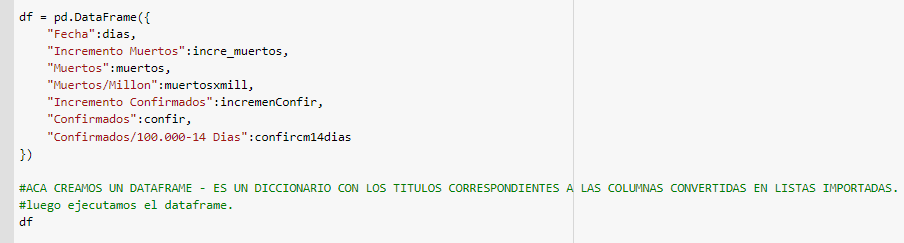


Como primer paso se importó la librería requests y Beautifull Soup. Requests es una biblioteca HTTP para el lenguaje de programación Python. El objetivo es hacer que las solicitudes HTTP sean más simples y BeautifullSoup es una biblioteca de Python para analizar documentos HTML. Esta biblioteca crea un árbol con todos los elementos del documento y puede ser utilizado para extraer información. Utilizando dichas librerías se crearon listas para obtener los valores que se encontraban en la pagina web.





En el segundo paso se utilizó pandas la cual es una librería de Python especializada en la manipulación y el análisis de datos. Ofrece estructuras de datos y operaciones para manipular tablas numéricas y series temporales, es como el Excel de Python. Con las listas creadas anteriormente se formó un diccionario con dichos datos y mediante dicha librería se creó un dataframe.





Posteriormente a la creación del dataframe con la función df.to\_csv se logra crear un archivo de texto descargable con los datos obtenidos.

