**Explicación Examen Interciclo**

**Freddy L. Abad L.**

Inicialmente, se procedió a implementar los procesos para recuperar el url, los nombres de los países y posteriormente concatenar el url predeterminado con el resultado del Paso 1. Como resultado se obtuvo un archivo Excel con los urls.

El paso 2 consistió en acceder a los proyectos de 3 países, el proceso fue acceder al archivo filtrado con los 3 urls de los países y acceder a la información de los proyectos, posteriormente se obtuvo el nombre, presupuesto, institución y resumen de los proyectos.

Del RESULTADO DE ESTE FUE UN EXCEL CON LA INFORMACION ANTES MENCIONADA

Posteriormente, se procedió a analizar cada resumen, tokenizar, filtrar palabras, ngrams, para obtener una tabla de frecuencias, y finalmente normalizar y clusterizar. Si bien no se implementó el análisis de los 23 resúmenes, ya que se procedió a archivos por cada resumen.

Dado esto se puede implementar kmeans para obtener la clusterizacion deseada.