Herramientas y tecnologías usadas para el trabajo.

Las herramientas que se usaran para este reto que nos en parte fueron recomendadas son las siguientes:

* Python:
  + Nosotros nos recomendaron este lenguaje de programación ya que este lenguaje es el que contiene más librerías para hacer nuestro trabajo y nos facilita mucho el trabajo al momento de hacer nuestro análisis, además la curva de aprendizaje de esta misma es muy alta y nos permite la realización de diversos modelos en pocas líneas de código.
  + Se utilizará especialmente Jupyter Notebook, esta es una interfaz amigable para mostrar contenido como en documento y mostrar gráficas y tablas en cada conjunto de código.
  + Posibles librerías para usar:
    - Numpy
    - Pandas
    - Sklearn y sus derivados (Modelos de aprendizaje)
    - Matplotlib
    - Seaborn
    - Statsmodels
* C++ / OpenCv:
  + Se va a utilizar los archivos proporcionados por nuestro socio formador de Nebraska para el análisis de imágenes de ríos, de las cuales nosotros tenemos que usar la información tomada y hacer otro análisis mediante la herramienta anteriormente mencionada.
* AWS (cloud):
  + Teniendo en cuenta nuestro crédito otorgado por nuestra institución el Instituto Tecnológico y de Estudios Superiores de Monterrey tenemos 100 dólares por persona las cuales somos 5 en el equipo por lo cual individualmente tenemos 500 dólares de crédito los cuales en un futuro lo podríamos usar para usar una herramienta proporcionada por AWS la cual nos permite conectarnos con Google Colab pero darle una potencia en específico que nos ayude a que nuestro procesamiento tenga más poder y se mucho más rápido ya que en un futuro podemos hacer análisis de imágenes y esto si necesitaría de un poder un poco más grande.

Modelo de almacenamiento de los datos

En este caso tanto el espacio de trabajo como el almacenamiento de los datos están localizados en una carpeta en un drive compartido con todos nuestros miembros del equipo, teniendo en cuenta como estamos trabajando con un Google Colab este obligatoriamente tiene que ser almacenado en esta nube y además nuestros datos nos fueron otorgados mediante un CSV, esto implica que si en algún momento del reto nuestro socio formador nos diera más datos o generara más datos este nos los tendría que dar en un archivo del mismo formato el cual tendría que volver a subirse a este mismo espacio compartido.

Enfoque de Big Data

Con respecto al enfoque de Big Data podemos decir que nuestro proyecto podría tener este enfoque pero analizando muy bien la situación en la que estamos llegamos a la conclusión de que no lo es ya que para que sea esto tendría que ser una cantidad muy pesada de datos difíciles de procesar y capturar, esto implica que nuestro socio formador nos diera una nube donde este hosteado todos nuestros datos donde nosotros le hagamos un request para jalarlos, además de esto nuestro socio formador tendría que estar constantemente tomando datos y analizándolos y nosotros en tiempo real estarlos analizando.