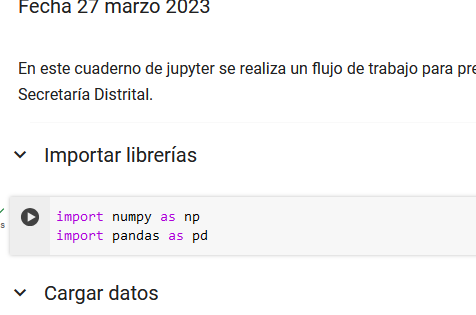
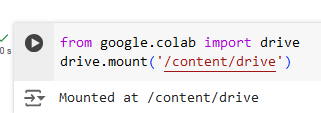
Roger Alberto Arias Triana - Taller 4

Se importan las librerías:



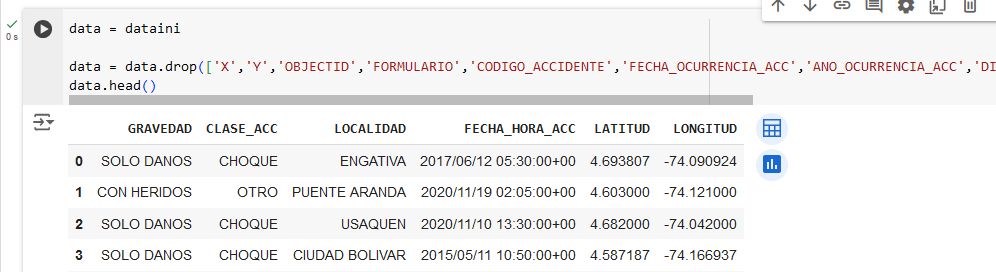
Se da el acceso a drive, aquí se debe subir el archivo de siniestros compartido por el profe



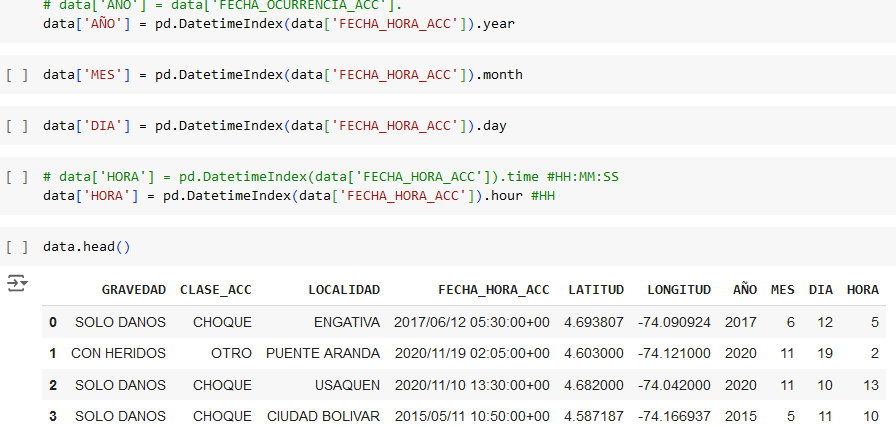
Una vez subido el archivo se lee con Pandas y se hace un head para ver los primeros registros



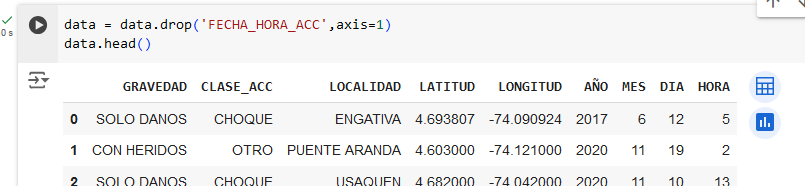
Se eliminan campos que no se necesitan con la función drop y se visualizan los primeros registros con head



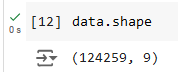
Se trabaja con el datos time stamp con las funciones de fecha para separar en campos diferentes y se visualiza la información



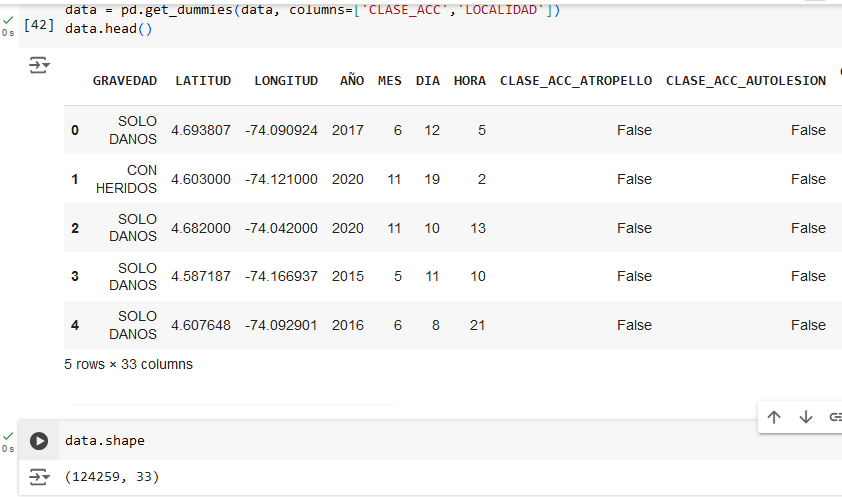
Se elimina el campo ‘fecha\_hora\_acc’



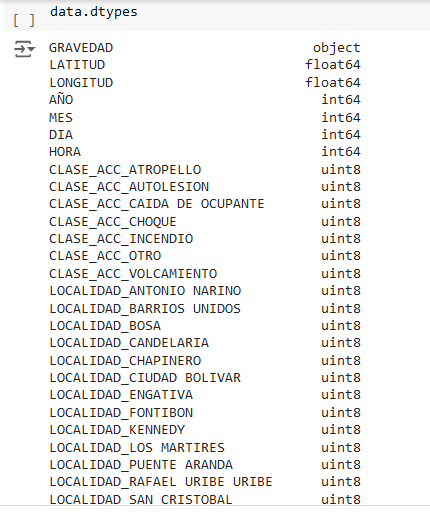
Se trae el numero de filas y columnas del dataset



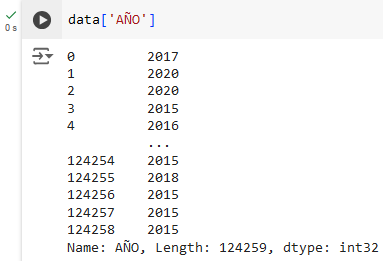
Se extraen los valores del campo calaseacc en campos diferentes discriminados por valor con la función get\_dummies, se muestra nuevamente la dimensión del dataset y se observa que el número de columnas ha aumentado a 33



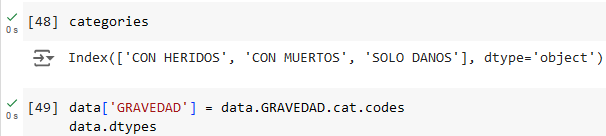
Luego se describen los tipos de campos del dataset utilizando la función dtypes:



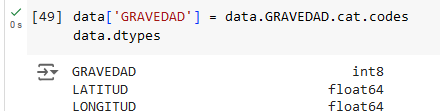
Luego se muestran los valores de determinado campo, se imprime el valor del campo como si se tratara de un arreglo:



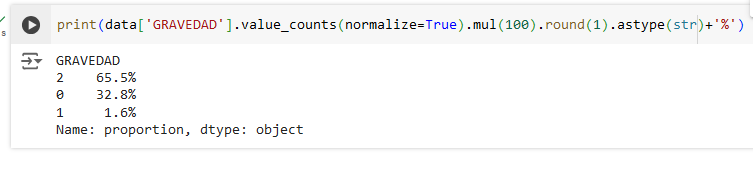
Se asigna un valor númerico al campo gravedad y se muestra:



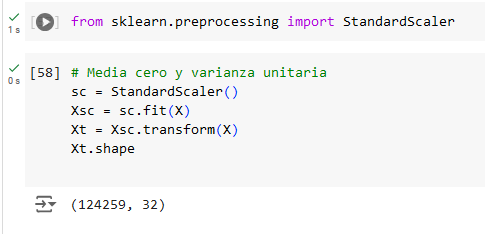
Nuevamente se describen los campos y se observa que ‘gravedad’ ahora es numérico



Se muetsran los valores porcentuales de los registros de gravedad:



Para finalizar se haya la varianza unitaria con la función StandarScaler:



Se finaliza con el uso de RandomForest:

