

Memoria sobre el análisis de fallecimientos por accidente aéreo

Inicie el trabajo buscando bases de datos sobre accidentes aéreos, encontré un archivo en Kaggle el cual se convertiría en mi principal Dataframe, durante la marcha descubrí que tenía datos sin rellenar o con símbolos indicando que se desconocía el valor, por lo que corregí esto rellenando los "Nan" con "unknown", por suerte en el caso de fallecimientos, venía como entero. Eliminé unas cuantas columnas debido a que no tendría uso alguno en esta ocasión.

Teniendo el data set como estaba no era suficiente, tuve que realizar unas pocas limpiezas, por lo que primero pensé que podría necesitar y que quería utilizar para ello, así cree nuevos data sets con las columnas necesarias para su gráfica específica. Me habría gustado hacer algún tipo de análisis sobre el número de pasajeros y relacionarlo al número de fallecimientos, pero me encontré con que dentro de la rama pasajeros venían strings y no supe como separar los enteros de sus correspondientes strings, de igual manera me encontré con el problema de que en la fecha, unas cuantas de ellas venían con "???" en vez de un valor, intenté corregirlo, pero este me daba otro tipo de errores así que acabe editando dicho "???" por "1", lo cual influye en lo mínimo a los cálculos finales.

Muchas gracias por leer todas las dificultades por las que pase para terminar este proyecto espero que sea de su agrado, si me preguntan. ¿lo podría haber mejorado?, siempre hay una mejora que hacerle a un proyecto y que me habría gustado añadirle, pero debido al tiempo y motivos personales, así como esta me hace sentir satisfecho, podría haber usado graficas más complejas o limpiar y agrupar de otra manera los datos para poder tener un análisis más profundo como contar las veces que un mismo modelo de avión o motor ha fallado. Pero en mi opinión no se relacionaban mucho con mi tema por lo que las mantuve simples y entendible..