

Proyecto 1: Titanic

Dado el conjunto de datos **Titanic** deberán realizar todo el proceso de *Data Wrangling* necesario para resolver las siguientes preguntas y tareas:

1. Encontrar el *patrón de supervivencia* de la catástrofe (**el mejor que ustedes puedan obtener**). (**Valor: 3 puntos**)
 - Por ejemplo: para sobrevivir, se tenía que ser niño de entre 3 meses y 5 años o mujer de entre 23 y 32 años; pudiéndose salvar también los hombres mayores de 50 años que embarcaron en Cherbourg.
 - Tienen que programar en Python la regla que encontraron, así como determinar la exactitud de la misma (el porcentaje de aciertos que su regla tiene con respecto a la supervivencia indicada en el conjunto de datos).
2. ¿Cuál es el pueblo de embarque más mortífero? (**Valor: 1 punto**)
3. ¿Cuál es la tarifa promedio pagada por los sobrevivientes y cuál la de los que no sobrevivieron? (**Valor: 1 punto**)
4. Realicen un gráfico que muestre lo mejor posible si la clase en la que viajaba el pasajero aseguraba su supervivencia. (**Valor: 1 punto**)
5. ¿Cuáles son los porcentajes de supervivencia y fallecimiento de niños, mujeres y hombres? (**Valor: 1 punto**)
6. ¿La edad influyó para **no** sobrevivir? (**Valor: 1 punto**)
7. Dados los datos personales de cada uno de ustedes (estudiantes de *Introducción a la Ciencia de Datos*) ¿hubiera sido más probable que se salvaran o que fallecieran? Si no desean utilizar sus datos verdaderos, elíjanlos al azar. Además, la edad estratifíquela en lustros (bloques de 5 años). (**Valor: 1 punto**)
8. ¿Cuál es el promedio de edad de los fallecidos que embarcaron en Queenstown? (**Valor: 1 punto**)

Recuerden que deben tomar una política con respecto a los datos faltantes e indicarla por escrito. Además, **deben** usar todos los gráficos necesarios que les apoyen para responder las preguntas y/o realizar las tareas solicitadas.



