

TRABAJO FINAL ESTADÍSTICA Y PROBABILIDAD (TITANIC).

Presentado por: Juan Camilo Calle, Sofía Correa, Juan David López, Daniel Felipe Salazar, Valentina Sánchez

Presentado a: Beatriz Martínez González

**Risaralda, Universidad Tecnológica de Pereira
Pereira, Colombia**



RESUMEN

Este documento tiene como finalidad arrojar conclusiones a partir de una base de datos en la cual se puede identificar la relación entre dos o más variables de forma que cada relación con su respectiva gráfica permita observar y examinar situaciones concretas para obtener conclusiones acertadas. Para este caso, el espacio muestral serán los pasajeros del Titanic y se estudiará cuál es la relación entre los sobrevivientes y la clase o precio de su tiquete, asimismo observar si existe una relación entre su sexo con los datos anteriores. La metodología del trabajo es observacional de tipo correlacional, con diseño no experimental; los datos se extraen del sitio *data.world* y las gráficas aquí contenidas están hechas en lenguaje de Python.

PALABRAS CLAVE: Análisis estadístico, Clase, Correlación de Pearson, sobrevivientes, no sobrevivientes, base de datos, Titanic, botes salvavidas, género, precio tiquete.

ABSTRACT

The purpose of this document is to draw conclusions from a database in where it is possible to identify the relationship between two or more variables, so each relationship with its respective graph allows observing and examining specific situations to obtain correct conclusions. In this case, the sample space will be the passengers of the Titanic and the relationship between the survivors and the class or price of their ticket will be studied. There is a relationship between your gender and the above data. The methodology of the work is observational of a correlational type, with non-experimental design; the data are extracted from the data.world page and the graphs contained here are made in Python language.

KEYWORDS: Statistical analysis, Class, Pearson's correlation, survivors, non-survivors, database, Titanic, lifeboats, gender, ticket price.

INTRODUCCIÓN

Este documento está elaborado con datos reales del suceso histórico que se dio a cabo el 15 de abril de 1912, que fue el hundimiento del barco más grande del

mundo durante esta época. El RMS Titanic fue un transatlántico que zarpó en Southampton, Inglaterra con rumbo a New York, EE. UU., pero a las 23:40 del 14 de abril, cuatro días después de zarpar, y a

unos 600 km al sur de Terranova, el Titanic chocó con un iceberg; al momento de la colisión el navío contaba con 2208 personas a bordo, de las cuales 1309 cuentan con un registro y es esta la base de datos en la que se basa el documento [1].

Es válido aclarar que la razón por la cual la base de datos no contiene el mismo número de personas como pasajeros en el Titanic, es debido a que en la época realizar un censo a una población tan grande se hacía difícil puesto que la tecnología era menor, sin mencionar que los pasajeros ascienden y descienden en cada parada.

Se tomarán los datos de los pasajeros concerniente a sexo, clase del tiquete, el bote al cual subió y si sobrevivió al naufragio.

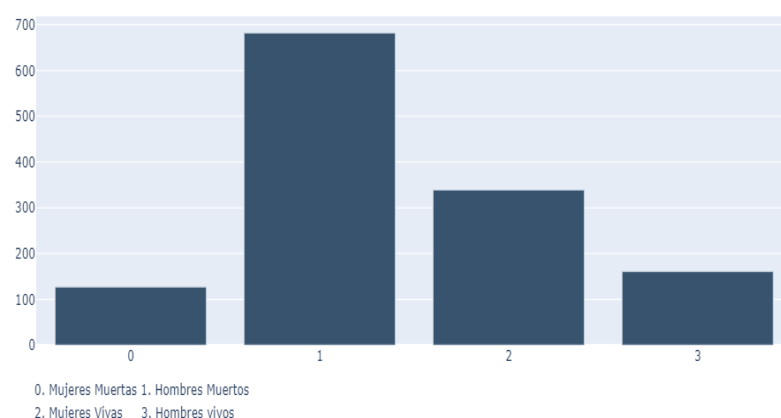
Durante el desarrollo se pretende llegar a una conclusión respecto a qué determinó que una persona sobreviviera al naufragio, si fueron: el género, la clase del tiquete y/o su ubicación en el navío.

DESARROLLO

En primera instancia es importante dimensionar la información que se tiene de la base de datos desde una perspectiva más gráfica. Para obtener las gráficas aquí contenidas, fue necesario filtrar la base de datos de forma que se pudiera extraer la información de mayor interés; se hizo una limpieza bastante general, que permitiera un buen estudio de la situación.

Es útil mostrar la relación que hay entre el número de mujeres y hombres que sobrevivieron y los que no, tal que sea más sencillo visualizar la situación en cuestión. Es de notar que el mayor número de no sobrevivientes son hombres y el mayor porcentaje de sobrevivientes corresponde a las mujeres. Al analizar esta información se puede decir que la preferencia al embarcar los botes salvavidas la obtuvieron las mujeres, ya que tuvieron una diferencia del 35,6% por encima de los hombres, lo que corresponde a 178 mujeres.

Sobrevivientes vs Sexo

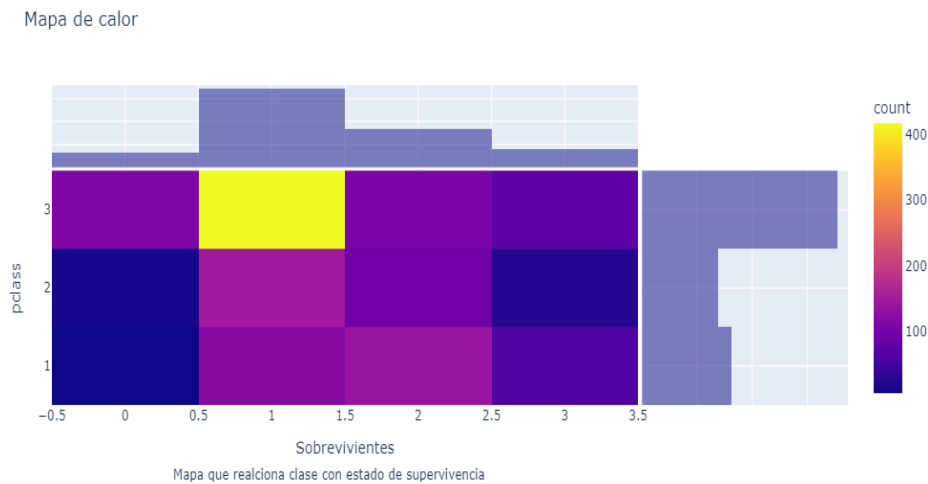


Gráfica 1 Sobrevivientes vs Sexo

Fuente: propia.

De la misma manera, al hablar de los sobrevivientes, se pueden introducir las clases; las cuales fueron primera clase, segunda clase y tercera clase. En la Gráfica 2 se presenta una relación más visual para los sobrevivientes y su respectiva clase

dentro del barco, esta corresponde a un mapa de calor en donde la mayor concentración corresponde a los varones no sobrevivientes.



Gráfica 2 Clases vs Estado Supervivencia

Fuente: propia.

En la clase 3, fueron los hombres los que más cantidad de muertos aportaron al desastre del Titanic, se dice entonces, que la tercera clase estaba compuesta en su mayoría por personas de clase más baja que se sabe, se encontraban en los lugares más

retirados de la cubierta según se observa en la figura 1. En base a esto, es fácil deducir que a las personas de la clase tres se les dificultó ascender hasta los botes salvavidas, más que al resto de los pasajeros.

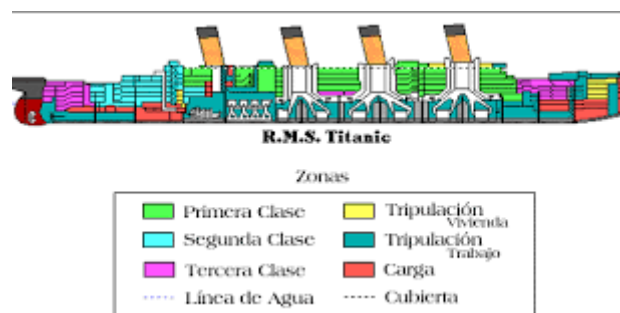


Figura 1 Distribución del RMS Titanic

Fuente: Blogger [2]

Al momento del desastre, uno de los factores fundamentales para la supervivencia en el Titanic fueron los botes salvavidas. El transatlántico contaba con 14 botes salvavidas, 2 botes de emergencia y 4 botes plegables; los primeros tenían capacidad para 65 pasajeros, los segundos para 40 y los botes plegables para 47 personas [3]. La embarcación no estaba preparada para una catástrofe de tal magnitud, de hecho, los botes podían albergar en total 1178 personas, de los 1309 que pagaron por su ticket, esto sin mencionar que en estos números no se tiene

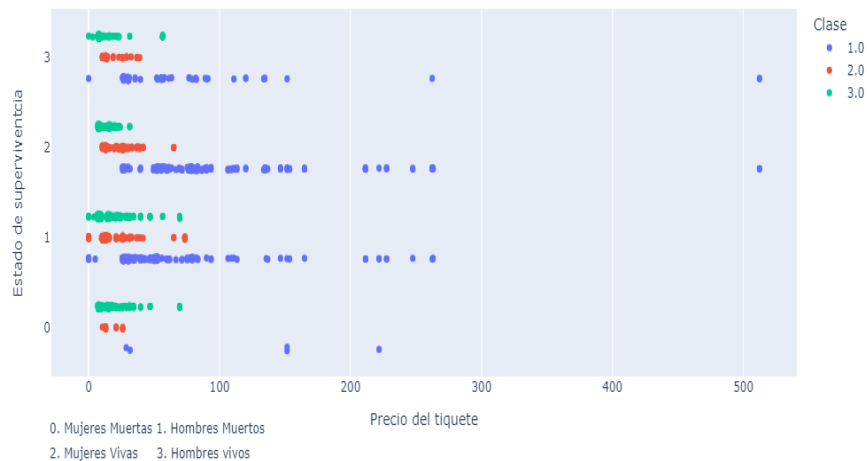
en cuenta la tripulación del barco, que eran alrededor de 899 individuos, de los cuales no existe registro alguno en la base de datos aquí citada [1].

Como se sabe, los pasajeros que pagaron más por su ticket, por lo general, pertenecían a las clases superiores. A excepción de algunos que pagaron menos por su ticket o incluso nada, y hacían parte de la primera clase. Se puede suponer, por esto, que su abordaje era gracias a alguna cortesía u obsequio.

Sirva analizar si al interior del barco existían preferencias por las clases

superiores o las personas más adineradas, en especial, al momento de abandonar el barco durante el siniestro. Sin lugar a duda es necesario relacionar qué tantas personas sobrevivieron de acuerdo con el precio que

pagaron o a la clase que pertenecían, ya que ayuda a saber si existen favoritismos, pleitesía o algún tipo de clasismo o si, por el contrario, el trato era igual para todas las personas.



Gráfica 3 Estado supervivencia vs Precio tiquete.

Fuente: Propia.

pclass	Sobrevivientes	
1.0	0	5
	1	118
	2	139
	3	61
2.0	0	12
	1	146
	2	94
	3	25
3.0	0	110
	1	418
	2	106
	3	75

Tabla 1 Estado supervivencia por clases.

Fuente: Propia.

De relacionar el precio del tiquete con el estado de supervivencia de cada persona por sexo y clase, se observa que en la población femenina de la tercera clase fue la que presentó mayor número de muertes en comparación con la de la primera clase; como se puede confirmar en la *tabla 1*, entonces, se tiene que en la tercera clase murieron el 13,60% del total de mujeres a bordo, mientras que, en la primera clase, solo el 0,62%. Es decir, más de la mitad del

total de las mujeres de la tercera clase, no lograron sobrevivir.

De forma análoga para los hombres, las muertes no presentan una diferencia tan drástica como en las mujeres con respecto a sus respectivas clases. Los hombres no sobrevivientes de la primera clase se le atribuye un 13,99% del total de la población masculina, mientras que, para la tercera clase la cantidad de muertes fue de 49,58%, lo que supone que una sexta parte del total de hombres en la tercera clase, sobrevivieron. De la misma forma es apropiado decir que la menor cantidad de sobrevivientes hombres corresponde a la segunda clase, mientras que en la tercera clase el estado de supervivencia estuvo en relación 6:1.

Es evidente que la prelación fue para las mujeres, pero dentro de este grupo, las afortunadas fueron las de la primera clase; quienes tuvieron una mayor probabilidad de sobrevivir.

Una práctica habitual en los naufragios con un número insuficiente de botes salvavidas, es embarcar mujeres y niños primero; esta

podría ser una razón válida para asegurar el por qué existió un mayor número de sobrevivientes femeninas que masculinos. Para corroborar la distribución de los botes al momento de la evacuación, es de utilidad tener todas las variables; Horas, fechas, procedimientos, clases, nombres, ordenes,

entre otras. ¿Será posible, entonces, conocer el orden de descenso de los botes salvavidas? Ya que debido a esto era más fácil acceder a determinados botes dependiendo del lugar donde se encontrase el pasajero.

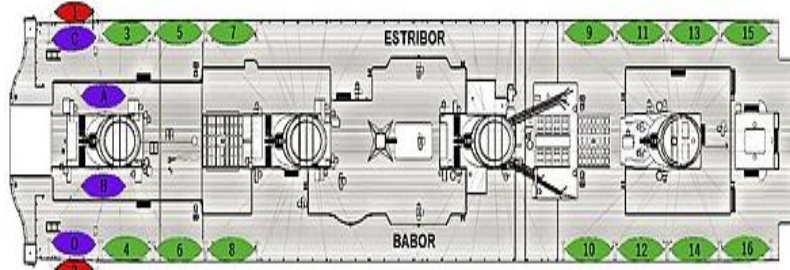


Figura 2 Ubicación de botes salvavidas

Fuente: Wikipedia. [4]

El capitán Smith dio la orden para bajar los botes salvavidas, sus dos oficiales decidieron empezar desde la mitad de la cubierta hacia la proa. Debido a un malentendido de ejecución, el oficial Lightoller, embarcó en los botes sólo a mujeres y niños; mientras que el oficial Murdoch, les daba preferencia a las mujeres y niños, pero también les daba la oportunidad a los hombres de abordarlos. Esta es una de las razones por las cuales los hombres representan muchos muertos incluso los de la primera clase, ya que estos botes podían llevar hasta 65 pasajeros, pero al no dejar pasar hombres, algunos se iban muy vacíos. Otra creencia es que despachaban a los botes con poca gente para que no hubiese riesgo de que se agrietaran por el peso antes de llegar abajo. De una forma u otra el panorama inicial era que aún con una buena organización al evacuar las personas a bordo, iba a ser imposible salvarlos a todos.

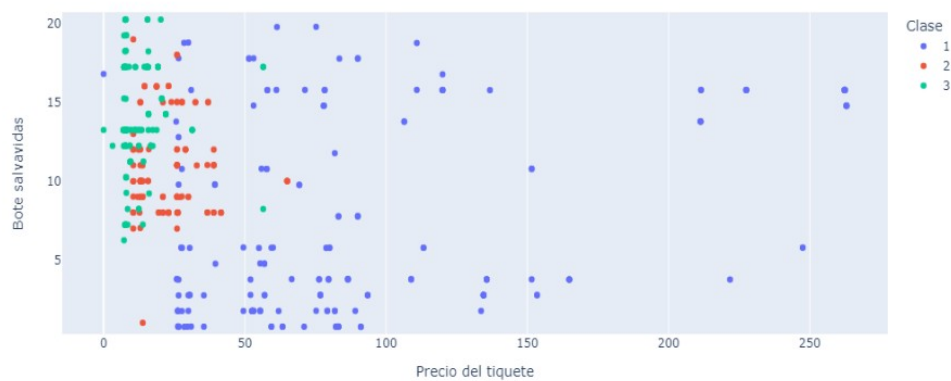
El orden de salida de los botes salvavidas fue el siguiente, por ambos costados empezaron desde la mitad del barco hacia la proa; lo que según la *figura 1* indica que se encontraban más personas de la primera y de la tercera clase. [5]

Con el fin de obtener un mejor análisis y una mejor grafica gracias al efectuar el siguiente outlayer, se crea una tabla con los valores de orden de salida a un valor numérico continuo.

Orden	Nº de Bote	Orden	Nº de Bote
1	7	11	11
2	5	12	13
3	3	13	15
4	8	14	2
5	1	15	10
6	6	16	4
7	16	17	C
8	14	18	D
9	12	19	B
10	9	20	A

Tabla 2 Orden de salida botes salvavidas

Fuente: Propia.



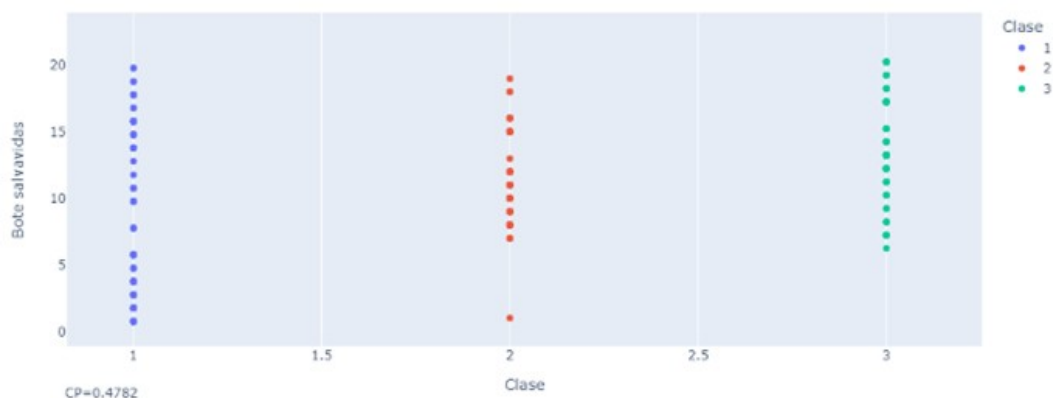
Gráfica 4. Orden Cronológico botes vs Precio del tickete.

Fuente: Propia.

Gran Bretaña, más aún Inglaterra, se ha caracterizado por ser una nación clasista [6] y para la época en la que se hundió el Titanic, este fenómeno social se vivía a flor de piel. Esto fue un factor determinante para la supervivencia de unos u otros pasajeros; pues ¿De qué otra forma los primeros seis botes irían cargados solo con pasajeros de la primera clase, teniendo en cuenta que el segundo barco descendió media hora después del primero y una hora y media después de la colisión con el iceberg? [4]. No es de extrañar que los

oficiales pudieran distinguir rápidamente a las damas de la primera clase por su acento refinado, y embarcarlas en los primeros botes que descendieron.

Como es evidente hasta el momento, existe una estrecha relación entre el haber podido pagar más por el tickete y haber obtenido un puesto en los primeros botes salvavidas. Si se agrupan los precios de los tickets por categorías y se correlacionan con el orden de salida de los botes salvavidas, se obtiene la siguiente gráfica.



Gráfica 5 Correlación Clases vs Orden de descenso

Fuente: Propia.

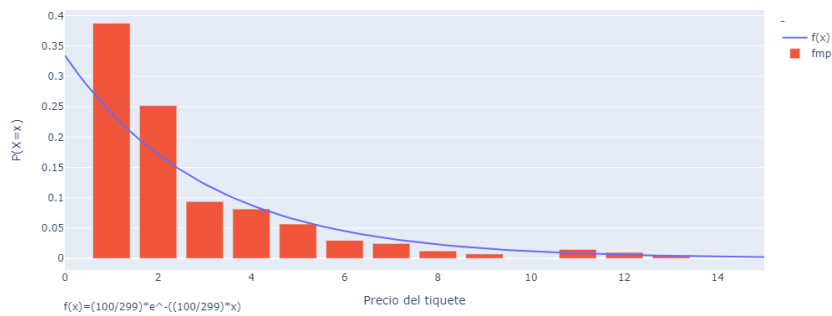
Para conseguir el coeficiente de correlación de Pearson y, otras de las gráficas aquí contenidas, fue necesario hacer una depuración más exhaustiva, de forma que se pudieran obviar ciertos datos que no

aportaban nada a la investigación. Un ejemplo de este fue una familia que pagó 512 dólares por su boleto, este precio se sale del rango de estudio del documento por eso, al igual que las personas que tienen doble o

triple asignación de botes salvavidas no fueron tenidas en cuenta. También se hizo una tabla de guía para el cambio de variable correspondiente al orden de salida de los botes y el nombre que recibieron desde un principio. Existe un bienestar al encontrar el valor de la correlación para las variables en cuestión, que, aunque bajo, muestra la existencia de una relación entre estos datos y en consecuencia confirma nuestras suposiciones iniciales.

RESULTADOS

- El 61,802% de los pasajeros murieron durante la catástrofe, esto sin mencionar que en realidad fue el 73,279% de las personas a bordo las que fallecieron en este incidente, pues en la base de datos sólo hay registro de quienes pagaron por su ticket y no de quienes trabajaron en el navío.
- El porcentaje de hombres que abordaron el transatlántico era mayor que el de las mujeres (por cada 1 mujer había 1,8 hombres) pero las cifras se invirtieron al hablar de supervivencia. Sólo el 19,09% de los hombres, sobrevivieron contra el 72,75% de las mujeres que también sobrevivieron.
- Como se puede observar en la gráfica 5, el coeficiente de Pearson corresponde a 0.4785 y se confirma que existe una correlación positiva; por lo que, si una aumenta, la otra por consiguiente debe aumentar. Esto quiere decir que ser una clase más alta, está estrechamente relacionado con obtener uno de los primeros botes.
- Se identificó un alto nivel de elitismo, en donde los sobrevivientes en su mayoría y tanto para hombre como para mujeres, pertenecían a la primera clase.
- Por último, es necesario mencionar la media muestral que se obtuvo durante la investigación, ésta corresponde a los siguientes valores: Para el caso que se mencionó sobre la clase y precio de los tickets, se tiene 1.874044 y 50.266729, respectivamente. De forma que la clase es una muestra de las personas que iban a bordo y a qué categoría pertenecía de las tres posibles, y el precio del ticket es una variable muestral mucho más extensa pues había muchos precios que dependían de múltiples factores.
- Asimismo, la varianza muestral se obtuvo para la variable de “clase” con un valor de 0.719252, esta hace alusión a la desviación típica al cuadrado y en caso de necesitarlo se usaría para conseguir la covarianza. Seguido a esto, con la variable “precio del ticket” que corresponde a una variable de mayor envergadura, se obtuvo una varianza de 3374.970762.



Gráfica 6. Distribución

Fuente: Propia.

- El gráfico de barras corresponde a la función de masa de probabilidad en donde el precio del tiquete está dividido en rangos de 20, para este caso se hizo indispensable calcular la esperanza, de forma que al obtener el inverso de esta, por consiguiente se pueda extraer λ que tiene un valor de $\frac{100}{299}$ como se puede confirmar en la gráfica 6. A continuación se puede examinar la línea en azul que corresponde en últimas a la distribución que como es evidente es una distribución exponencial. Lo que quiere decir que la distribución sea exponencial es que en nuestro espacio muestral hubo más personas que pagaron menos por su tiquete que otras, lo que es lo mismo, había más personas con menos recursos entre los pasajeros del Titanic.

DISCUSIÓN Y CONCLUSIONES

Acerca de la preparación que tuvo el RMS Titanic se puede decir abiertamente que no tuvieron las medidas necesarias para una eventualidad de tal magnitud, lo que supuso una mayor cantidad de muertos que hubiesen podido evitar al haber tenido una

mayor organización y cantidad de botes salvavidas.

Si bien, las mujeres tuvieron prioridad, es claro que hubo un sesgo en el abordaje de los botes salvavidas, en el que tanto la ubicación en el barco como la clase influyeron en la probabilidad de sobrevivir, entonces, más que ser mujer o niño era necesario ser adinerado, o en su defecto pertenecer a la clase superior dentro del barco.

De forma similar se pudo concluir que no solo influía el ser adinerado para abordar un bote salvavidas, sino también qué tan rápido podía conseguir uno, es decir, entre mayor fuera su clase, aumenta la probabilidad de que el bote que reciba descienda en último lugar.

Esta correlación obtuvo un coeficiente de 0.4782 esto implica de forma directa que, al ser de las primeras clases, obtienen un bote más rápido y en consecuencia existe una mayor probabilidad de sobrevivir.

Dicho lo anterior, se hace cada vez más claro que un fenómeno social puede influir de manera negativa, incluso en una situación de vida o muerte. El clasismo llevó a disminuir las probabilidades de supervivencia de muchas personas a bordo, porque de qué otra forma se explica que seis de los primeros botes salieron únicamente con mujeres de la primera clase y cargados a la mitad.

Finalmente se puede evidenciar que lo que determinó principalmente la supervivencia

de las personas a bordo fue su posición burocrática, el pertenecer a una clase más alta y haber pagado más por su tiquete, lo que brindaba una mayor probabilidad de obtener un bote salvavidas y sobrevivir.

REFERENCIAS

- [1] @nrippner, «data.world,» 2016. [En línea]. Available:
<https://data.world/nrippner/titanic-disaster-dataset>. [Último acceso: 2 Noviembre 2021].
- [2] T. P. o. t. Opera, «Blogger,» 24 Enero 2009. [En línea]. Available:
<http://usstitanic.blogspot.com/2009/01/planos-del-rms-titanic.html#axzz7Djgidq3q>. [Último acceso: 26 Noviembre 2021].
- [3] E. M. d. Pozuelo, «La mayor maquina jamas construida,» *La Vanguardia*, pp. 8-9, 09 Abril 2012.
- [4] p. Wikipedia, «Wikipedia,» 02 Noviembre 2021. [En línea]. Available:
https://es.wikipedia.org/wiki/Botes_salvavidas_del_RMS_Titanic. [Último acceso: 30 Noviembre 2021].
- [5] J. Fellowes y B. Lavery, *a night to remeber*, London, WC2R ORL: Penguin Books, 2012.
- [6] L. ventoso, «Inglaterra, todavia un mundo de "ariba y abajo",» *ABC Internacional*, 27 Diciembre 2015.