

Preregistration

My preregistration for the COS Preregistration Challenge

Lina Aragon-Baquero¹, Kristen Bill², Leah D'Aloisio³, Jack Goldman⁴, Briar Hunter^{5,6}

¹ University of Waterloo

² Wilfrid Laurier University

³ University of British Columbia

⁴ Institute of Forestry and Conservation, University of Toronto, Toronto, ON M5S 3E8,
Canada

⁵ Laurentian University

⁶ Toronto Zoo

05. October 2021

Study Information

Where sex and class determinant for survival aboard the Titanic?

My preregistration for the COS Preregistration Challenge

Description	It is well known that the survivorship upon the sinking of the Titanic was heavily influenced by sex and class of the individual. Studies reveal the role of sex and class was such a strong predictor of survival due to policies and social/physical barriers put in place that prioritized first class and women/children. However,
--------------------	--

to our knowledge, no studies have investigated the impact of other variables on survival. Therefore, using a multivariate model we will test not only for sex and class, but also effects of age, family members, ticket fare, home destination and boat number. Enter your response here.

Hypotheses	If both sex and class affected the survival aboard the Titanic, first-class passengers and women had a higher survival rate.
-------------------	--

Design Plan

Study type	This is a retrospective observation study based on data obtained from the historical event of the sinking of the Titanic.
-------------------	---

Blinding	Not applicable
-----------------	----------------

Study design	This study utilizes retrospective data from a one-time, natural event and thus data collection was constrained by availability of information gathered and documented from passengers and survivors at the time. On account of this being a one-time event with a defined number of people our design is a cross-sectional observation study.
---------------------	---

To understand the driving factors of survival aboard the titanic, we will use a mixed effect model approach in which sex and class will be fixed factors and age a random factor: $\text{Survival} \sim \text{Class} + \text{Sex} + (1|\text{Age})$.

Randomization	Not applicable.
----------------------	-----------------

Sampling Plan

Existing data	Registration prior to creation of data. As of the date of submission of this research plan for preregistration, the data have not yet been collected, created, or realized.
----------------------	--

Registration prior to any human observation of the data. As of the date of submission, the data exist but have not yet been quantified, constructed, observed, or reported by anyone - including individuals that are not associated with the proposed study. Examples include museum specimens that have not been measured and data that have been collected by non-human collectors and are inaccessible.

Registration prior to accessing the data. As of the date of submission, the data exist, but have not been accessed by you or your collaborators. Commonly, this includes data that has been collected by another researcher or institution.

Registration prior to analysis of the data. As of the date of submission, the data exist and you have accessed it, though no analysis has been conducted related to the research plan (including calculation of summary statistics). A common situation for this scenario when a large dataset exists that is used for many different studies over time, or when a data set is randomly split into a sample for exploratory analyses, and the other section of data is reserved for later confirmatory data analysis.

Registration following analysis of the data. As of the date of submission, you have accessed and analyzed some of the data relevant to the research plan. This includes preliminary analysis of variables, calculation of descriptive statistics, and observation of data distributions. Please see <https://cos.io/prereg> for more information.

**Explanation of
existing data**

Since this is a replication study, we have read prior papers analyzing the predictors of survival on the Titanic. The conclusions from these papers were that sex and class were the main influence on survival. However, no statistical tests were used to conclude these effects, rather they compared a table of those who survived and died. To determine the effect of class and sex, the authors also used information provided by survivors who described their experiences. We have not read any studies that statistically analyzed all determinants of survival, therefore we are unaware of any summary statistics.

Data collection procedures	This study utilized the publicly available titanic dataset found in R. The principal source for the data in this package is the Encyclopedia Titanica but a variety of researchers contributed to the datasets, including Eaton & Haas (1994) Titanic: Triumph and Tragedy, Patrick Stephens Ltd (CITATION: TitanicMETA) The study population used is not reproducible as it relied on the specific circumstances of the Titanic sinking, which we hope never happens again. Comparative analyses, however, can be performed on similar events such as the sinking of the Lusitania (Frey et al. 2010), but the exact conditions are unlikely to be repeated.
-----------------------------------	---

Sample size	Enter your response here.
--------------------	---------------------------

Sample size rationale	Enter your response here.
------------------------------	---------------------------

Stopping rule	Enter your response here.
----------------------	---------------------------

Variables

Manipulated variables	Enter your response here.
------------------------------	---------------------------

Measured variables	Enter your response here.
---------------------------	---------------------------

Indices	Enter your response here.
----------------	---------------------------

Analysis Plan

We will be using the predictor variables age, class, fare, gender, and boat ID to understand drivers of survival aboard the titanic.

```
library(knitr)
library(readr)
library(magrittr)
library(dplyr)
```

```
library(titanic)
#knitr::kable(head(titanic))
titanic<- read_csv(file="titanic.csv")
```

Statistical models To assess the drivers of survivability aboard the titanic, mixed effect models will be run with X fixed effects and Y random effects.

Transformations Enter your response here.

Inference criteria

Data exclusion Enter your response here.

Missing data Enter your response here.

Exploratory analyses (optional) Enter your response here.

Other

Other (Optional) Enter your response here.

References
