

Titanic

First draft: 27 November 2016

1	What happened	1
2	The E-T data	3
3	Comparison between R and E-T data sets	6
4	Children	7
5	Lifeboats	7
6	Crew	8
7	Third class	8
8	Models and probabilities	8

This handout draws mostly from material at the website

<https://www.encyclopedia-titanica.org>

and from a 2002 class handout (using Titanic data to illustrate ways to analyse counts) for Stat 230, written by John Hartigan.

Don't believe everything you see on the WWW, even if it involves **nice animation** or claims to have **1912 Original Video Footage**. (For the second example, where was the cameraman standing at 2:18?

1 What happened

S:story

Just over a hundred years ago, on her maiden voyage the ocean liner Titanic hit an iceberg in the North Atlantic (at about **this location** according to the website <http://matei.org/>) and sank within three hours, with great loss of life.

According to <https://www.britannica.com/topic/Titanic>: the collision occurred at 11:30_{PM} on 14 April 1912; the ship carried enough lifeboats

for only about half the number of people on board; the water was very cold; the first rescue ship, Carpathia, arrived at the scene at 3:30_{AM}, more than an hour after the Titanic had sunk.

The disaster became the subject of two government inquiries, one in Britain (British Board of Trade) and one in the United States (US Senate). Transcripts of testimony and final reports are available online at <http://www.titanicinquiry.org>.

The deaths of passengers (1st, 2nd, and 3rd class tickets, sex, age) and crew were distributed rather unevenly across those on board. **R** comes with a small data set in the form of a four-way array, which is easier to understand when viewed as a flat table:

##			Survived	No	Yes
##	Class	Sex	Age		
##	1st	Male	Child	0	5
##			Adult	118	57
##		Female	Child	0	1
##			Adult	4	140
##	2nd	Male	Child	0	11
##			Adult	154	14
##		Female	Child	0	13
##			Adult	13	80
##	3rd	Male	Child	35	13
##			Adult	387	75
##		Female	Child	17	14
##			Adult	89	76
##	Crew	Male	Child	0	0
##			Adult	670	192
##		Female	Child	0	0
##			Adult	3	20

Clearly women and children in 1st and 2nd class suffered relatively light casualties, perhaps due to the noble sacrifice made by their male fellow travellers—or so one version of the myth goes. (Actually the transcripts from the inquiries make clear that the traditional ‘women and children first’ procedure was followed when the lifeboats were being loaded.)

You should be aware that even the numbers in the tables are the subject of much disagreement. It is not completely clear,

<https://www.encyclopedia-titanica.org/titanic-statistics.html>,

who was and who was not actually on the ship at what times.

The Encyclopedia Titanica site has a lot of detailed information that might be useful for a closer look.

2 The E-T data

S:ETdata

I have placed the data extracted from the table at

<https://www.encyclopedia-titanica.org/titanic-passengers-and-crew/>

on ClassesV2 in the tab-delimited file `Titanic_data.txt`, which corresponds to my data frame `people` (often abbreviated to `pp` to save on typing):

```
names(people); table(people$status); table(people$prefix)

## [1] "url"      "Name"      "Age"      "Class.Dept" "Ticket"
## [6] "Joined"    "Job"      "Boat:Body" "status"     "prefix"
##
## biography      cross survivor      victim
##           248           1           711           1496
##
##           Captain   Col. Colonel   Don.   Doa   Dr       Fr   Major
##           3         2           1       3       1       10      6     3
## Master   Miss   Mlle   Mme.   Mr   Mrs   Ms   Rev   Rev.
##           64   278     2     1   1815  214   1     1     1
##   Revd   Sig.   Sir   Sr.   Sra.
##           1    42     1     2     3
```

The column `url` (showing the file linked to each name) comes from the links on the web page., The column `status` is extracted from `url`. The column `prefix` comes from the `Name`. Presumably one can infer sex from the `prefix`, with an occasional clarification from the detailed information pointed to by the `url`.

Remark. Note the anachronistic ‘Ms.’ for Mary McGovern. Perhaps the spellings of names has been altered in some cases. See the discussion at <https://www.encyclopedia-titanica.org/titanic-statistics.html> for more about the difficulties with names. Also Doa is actually Doña. Be aware that webpages sometimes use special characters (such as the \mathcal{L} sign) that disappear when processed by \LaTeX . And don’t be sexist: even in Edwardian times there were [female doctors](#).

The column `Boat:Body` contains the grisly information about the fate of each individual: either a number for the lifeboat they entered, a number for a body (presumably identified by some coroner), or a blank (mostly for bodies never recovered, although there are blanks for one cross-channel passenger, 248 biographies, and 84 survivors). Do not assume that everyone who entered a lifeboat also survived: you should follow the links from

`people[grep("A",people$B),]`. We need to figure out what `Boat:Body` really means.

You should go to the E-T website and click on individual names to read some heart wrenching stories, which might be helpful in understanding some of the numbers. For example, the entry for [Mrs. Bessie Waldo Allison](#) gives an explanation for why she was one of the few women in 1st class who died and why her baby son was the only member of the family to survive. You will also learn from the ticket information that there were others travelling with the family, such as [Miss Alice Cleaver](#), the nursemaid for the baby son. From that link:

The next day, Alice Cleaver and Sarah Daniels realized that they, along with Trevor and the cook—Mildred Brown, were the only survivors of their party.

There is no Mildred Brown on the passenger list, but the link to the 2nd class passenger [Miss Amelia Mary Brown](#) explains:

Amelia, known as Millie, was hired by the wealthy Allison family of Montral and she and their chauffeur George Swane were booked as second class passengers on Titanic whilst the Allisons would travel in first class. She boarded the Titanic at Southampton on 10 April 1912 (ticket number 248733 which cost £13 and she shared a cabin with Amelia Lomore, Elizabeth Nye and Selina Cook.

On the night of the sinking Amelia was reluctant to rise from her bed despite the warnings from her roommates and George Swane. She only rose when Selina Cook exclaimed that she was surely the only person on the entire ship still in bed.

```
allisons.maybe <-
  c(grep("113781",people$Ticket),
    grep("Mrs::Selena",people$Name),
    grep("Miss::Amelia",people$Name),
    grep("SWANE",people$Name),
    grep("NYE",people$Name))
people[allisons.maybe,c(2:5,7:8)]
```

##		Name	Age	Class.Dept
## 35	ALLISON::Mr::Hudson	Joshua Creighton	30	1st Class Passenger
## 36	ALLISON::Mrs::Bessie	Waldo	25	1st Class Passenger
## 37	ALLISON::Miss::Helen	Lorraine	2	1st Class Passenger
## 38	ALLISON::Master::Hudson	Trevor	1	1st Class Passenger

```

## 446      CLEAVER::Miss::Alice Catherine  22 1st Class PassengerServant
## 562      DANIELS::Miss::Sarah Rebecca  36 1st Class PassengerServant
## 484              COOK::Mrs::Selena    22      2nd Class Passenger
## 306      BROWN::Miss::Amelia Mary    18 2nd Class PassengerServant
## 2181             SWANE::Mr::George    19 2nd Class PassengerServant
## 1667             NYE::Mrs::Elizabeth  29      2nd Class Passenger
##
##              Ticket                    Job Boat:Body
## 35  113781:GBP151 16s      Businessman      [135]
## 36  113781:GBP151 16s
## 37  113781:GBP151 16s
## 38  113781:GBP151 16s                                11
## 446 113781:GBP151 16s      Nursemaid      11
## 562 113781:GBP151 16s      Personal Maid    8
## 484  14266:GBP10 10s
## 306   248733:GBP13  Cook (Personal)    11
## 2181  248734:GBP13      Chauffeur      [294]
## 1667  29395:GBP10 10s                                11

```

Hmmm! When I followed the links it seemed that all except Nye were really members of the Allison party, but only the nursemaids got to travel first class. From the Daniels page:

Sarah later came into the employment of the Hudson Allison family of Montral, Québec and was enlisted as a maid to the family. She boarded the Titanic at Southampton on 10 April 1912 as a first class passenger travelling with the Allisons, Hudson and his wife Bess and their two infant children Loraine and Trevor and their children's nurse Alice Cleaver (ticket number 113781 which cost £151, 16s). A further two servants, Amelia Brown and George Swane, travelled in second class.

Real data gets complicated. Who knows whether any of these connections were relevant to the survival of the individuals. It appears that all except baby Trevor of the Allisons perished but all except George Swane of the entourage survived.

The E-T table also shows where people joined the ship, and gave some clues about persons who might have left the ship:

```

options(width=90)
table(people$Class,people$Joined)

##
##              Belfast Southampton Cherbourg Queenstown
## 1st Class Passenger              4         156         123         3
## 1st Class PassengerCross Channel Passenger    0          22          0          0
## 1st Class PassengerH&W Guarantee Group       4           0           0           0
## 1st Class PassengerServant                 0          18          20           0
## 2nd Class Passenger                       0         233          25           7

```

##	2nd Class Passenger	Cross Channel Passenger	0	9	0	0
##	2nd Class Passenger	H&W Guarantee Group	6	0	0	0
##	2nd Class Passenger	Musician	0	8	0	0
##	2nd Class Passenger	Servant	0	4	1	0
##	3rd Class Passenger		0	494	102	113
##	Deck Crew		48	41	0	0
##	Deck Crew	Titanic Officers	8	2	0	0
##	Engineering Crew		221	282	0	0
##	Restaurant Staff		0	69	0	0
##	Victualling Crew		124	304	0	0
##	Victualling Crew	Postal Clerk	0	5	0	0

Not surprisingly, many of the crew were on the ship before it left Belfast, where the Titanic was built. Titanic was launched on May 31, 1911. Britannica: “It then began the fitting-out phase, as machinery was loaded into the ship and interior work began.” There must have been many crew members working on the ship before (and after?) it began its maiden voyage. The actual maiden voyage started at Southampton (south coast of England), stopping at Cherbourg in France to pick up further passengers (and drop off cross-channel passengers?), then stopping at **Queenstown** in Ireland before setting out to New York.

Apart from the fact that some passengers who were supposed to board at some of those ports failed to do so, how do we determine who was actually on the ship as it crossed the Atlantic? Is it enough to rely on the **status**?

The **R** data set **Titanic** gives aggregated information about 2201 people, whereas **people** contains information for 2456 people. My guess is that the $711 + 1496 = 2207$ with status ‘survivor’ or ‘victim’ were the ones on the ship when the collision occurred.

```
on.board <- grep("victim|survivor",people$status,perl=T)
table(people$status[on.board])
```

```
##
## survivor   victim
##         711    1496
```

3 Comparison between R and E-T data sets

S:compare

It simplifies the analysis to turn the 4-way array **Titanic** into a data frame:
Tidf <- **data.frame(Titanic)**.

```
Tidf <- data.frame(Titanic)
```

Are the data sets mostly in agreement?

4 Children

S:kids

The E-T site offers many specialized listings, such as “Children on the Titanic”. I have extracted the data from that list into a text file.

```
children <- read.table("children.txt",header=T,sep="\t",quote="")
is.kid <- match(children$url,people$url)
rbind(freq=table(pp$Age[is.kid]),cum=cumsum(table(pp$Age[is.kid])))

##      1  2  3  4  5  6  7  8  9 10 11 12 13 14 18 32 34 38 39 45 46 61
## freq 11 13  8 12  8  4  8 10  9  7  7  4  9  8  1  1  2  1  2  1  1  1
## cum  11 24 32 44 52 56 64 74 83 90 97 101 110 118 119 120 122 123 125 126 127 128
```

What is going on? Has the meaning of “child” changed since 1912? Which persons were actually counted as children for Titanic? Who actually benefitted from ‘women and children first’?

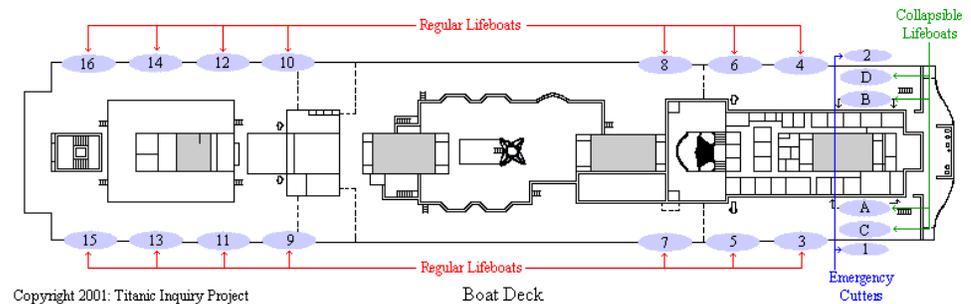
This problem needs to be fixed before we engage in any formal analysis.

```
kar <- intersect(is.kid,on.board) # kids atrisk
table(people$Class[kar],people$status[kar])
```

```
##
##              survivor victim
## 1st Class Passenger      6      1
## 2nd Class Passenger     26      1
## 3rd Class Passenger     33     57
```

5 Lifeboats

S:lifeboats



In what order were they launched? Who was on each lifeboat? What was the theoretical capacity of each boat?

6 Crew

S:crew

Why did they suffer heavy casualties?

```
is.crew <- grep("Crew|Guarantee|Musician|Staff", pp$Class, perl=T)
car <- intersect(is.crew ,on.board) # crew at risk
table(people$Class[car],people$status[car])

##
##
## survivor victim
## 1st Class PassengerH&W Guarantee Group 0 3
## 2nd Class PassengerH&W Guarantee Group 0 6
## 2nd Class PassengerMusician 0 8
## Deck Crew 39 19
## Deck CrewTitanic Officers 4 4
## Engineering Crew 71 253
## Restaurant Staff 3 66
## Victualling Crew 94 332
## Victualling CrewPostal Clerk 0 5
```

Maybe something along the following lines would be informative:

```
fireman <- read.table("Occupation/Fireman.txt",header=T,sep="\t",quote="",e
far <- intersect(match(fireman$url,pp$url,nomatch=0),on.board) # fireman at
table(pp$status[far])

##
## survivor victim
## 46 127
```

7 Third class

S:third

8 Models and probabilities

S:probs

What increased or decreased the chances of survival on the Titanic? That might seem a strange question. We know (more or less) who died and who survived. Are we thinking of some hypothetical collection of universes where many Titanics set out across their Atlantics and some don't? Or are we looking for some simple explanation for why some people died and

others survived? Are we planning to place blame on those whose negligence might have caused deaths? Are we (many years after the fact) thinking about matters of **liability and insurance**? Or are we but noble (and humble) statisticians seeking to understand the mysteries of the universe?

With the E-T data we have an opportunity to learn in more detail about who died and who survived.