

Learn Mixed Methods with the Titanic datasets

An exercise from www.mixedmethodstitanic.wordpress.com



As you may know, the Titanic hit an iceberg on April 14, 1912, at 11:40 p.m. and sunk completely about 2 hours and 40 minutes later. 2207 individuals were on board at the time of the collision and only 710 survived, despite the fact that the lifeboats could have saved 1178. The survival rates of individuals differed greatly according sex, classes, types of crew. Why is that?

Research question: What were the causes / mechanisms that led individuals on the *Titanic* to survive or perish?

Part 1: Theorizing

Find explanatory variables that may lead to higher or lower probability of surviving on the *Titanic*.



Tips: Specify the “causal story” of how exactly these explanatory variables might have influenced the response variable (survived/perished) and draw a sketch if this helps!

Part 2: Exploratory quantitative data analysis

In R or SPSS, analyse the quantitative dataset and answer the following questions using the variables class, gender, age, and survived/perished.



Tips: Look at the coding scheme in Textbox 1 and use the syntax in Textbox 2/3 in the Appendix to help you with crosstabs and graphs.

- a. Is there a bias in the testimonies? What is it and why?

- b. What is the proportion of individuals who survived and perished?

c. What is the composition of the population on the Titanic in general?

d. What is the distribution of the survivors according to gender, class, age, and type of crew? Did you find another interesting result?

Part 3: Exploratory qualitative data analysis

In MAXQDA, analyse the testimonies about what happened after the collision and answer the following questions using variables such as gender, class, filling rules, boat order, boat side, and authority.



Tips: Refer to the detailed MAXQDA procedure in Textbox 4 and the map of the lifeboats in the Appendix.

a. Compare the testimonies of 5 first-class and third-class women. What can you say?

b. Compare the filling rule „women and children first” for first-class and third-class women. What can you say?

c. Analyse the way authority was accepted during the process of filling the lifeboats. Was the crew important and why?

d. Analyse the rule „women and children first” during the process of filling the lifeboats. How was it applied?

e. In what sense did the two sides of the *Titanic* differ in applying this rule?

f. How did the way to the boat deck differ for first-, second-, and third-class passengers?

At this point, you should have noticed that your qualitative analysis shows the meaning and function of the quantitative variables in the “social game” that was played on the *Titanic*.

Part 4: Revisit the quantitative data

In R or SPSS, revisit the quantitative dataset in the light of your previous findings. This round should incorporate analyses that look at the specific time that individuals boarded a lifeboat and the side of the *Titanic* from where they boarded the lifeboat.



Tips: Use the syntax in Textbox 5/6 in the Appendix. Note that the “boatentertime” variable indicates 40.00 for 00:40 a.m. and 120.00 for 1:20 a.m.

a. Compare the boarding time across classes and types of crew. What can you say about the probability of being saved? Complete your answers with what you observed in the qualitative analysis.

b. Compare the distribution of sex and class according to boat side. What can you say about the filling rules? Complete your answers with what you observed in the qualitative analysis.

Part 5: Conclusion

Now that you have carried out a mixed methods study, some interesting points should be highlighted. Answer the following questions and do not hesitate to give your personal insight.

- a. Have you learned more by using both qualitative and quantitative data? Are our conclusions more valid than if we had had only one type of data? Why?

- b. Why do the datasets allow us to make meta-inferences so well?

- c. What are the limits to our analysis?

- d. How can we summarize the overall findings regarding our research question?

Congratulations, you have arrived at the end of the exercise!

Visit our website to learn more about this research:

www.mixedmethodstitanic.wordpress.com



Appendix

Textbox 1: Coding scheme (selection)

Level 1	Level 2	Level 3
Filling rules	Women & children first If no more women - fill randomly Couples first Fill with anybody Shortage of women Women hesitate/refuse to board	
Authority acceptance	Officers in charge People calm, follow orders People panic Men try to sneak into boats Men try to rush boats Officers re-establish order Officers shoot Fill from A deck	
Way to boat deck	Crowd size Arrival Time on Deck	Crowd_small Crowd_large Arrives before 0.40 Arrives 0.40 - 1.20 Arrives 1.20 - 2.05 Arrives after 2.05 or not (all boats gone)
	Way to boat deck	On first decks when impact occurs Informed through crew On other decks first Difficult way to boat deck Crowd obstructs passage Crew obstructs / unhelpful Reluctant to go to boat deck Finds gates locked Finds gates open
Age & strength		
Attribute_nationality		
Social_Network		Informed through social network Goes to deck with others Get on - encouraged by social network Social Network generic

```
*****  
Textbox 2: Exploratory quantitative analysis - SPSS  
*****
```

***** (1) Analysis of bias in testimonies: example gender**

crosstabs testimony by sex by lived
/cells count row
/statistics.

***** (2) Frequencies response variable**

Frequencies lived /bar chart.

***** (3) Frequencies explanatory variables**

Frequencies sex classcrew age_cat boatnumber boatside child country_5cat testimony / barchart.

Graph/histogram age.

***** (4) Some simple crosstabs**

crosstabs lived by classcrew sex
/cells count col
/statistics.

crosstabs lived by classcrew by sex
/cells count col
/statistics.

Graph
/bar(grouped)=mean(lived) by classcrew by sex.

```
#####  
Textbox 3: Exploratory data analysis - R  
#####
```

```
library(dplyr)  
library(tidyr)  
library(ggplot2)  
library(forcats)  
library(purrr)
```

```
# read in data (your own location file)  
Titanic_Mixed  
read.csv("~/Dropbox/2019_neu/02_Projects/P_Titanic/Titanic_Website/DATASET_QUAN/Titanic_Mixed.csv"  
)
```

```
# make it a tibble  
data <-tbl_df(Titanic_Mixed)
```

recode some variables

```
data <- data %>%  
  mutate(lived = recode(factor(lived), "0" = "perished", "1" = "survived")) %>%  
  mutate(sex = recode(factor(sex), "0" = "male", "1" = "female")) %>%
```

```

mutate (testimony = recode(factor(testimony), "0" = "no", "1" = "yes")) %>%
mutate (group = recode (factor(group), "0" = "Single", "1" = "Single w/servant",
"2" = "Couple", "3" = "Couple w/kids", "4" = "Couple w/servant",
"5" = "Single parent w/kids", "6" = "Family w/servant", "7" = "Family/friends",
"8" = "Crew groups", "9" = "Family/friends w/kids")) %>%
mutate (country_5cat = recode (factor(country_5cat), "0" = "England", "1" = "Ireland", "2" = "Sweden", "3" =
"USA",
"4" = "Others")) %>%
mutate (country_3cat = recode (factor(country_3cat), "0" = "England", "1" = "USA", "2" = "Others")) %>%
mutate (classcrew = recode (factor(classcrew), "1" = "1st class passenger", "2" = "2nd class passenger",
"3" = "3rd class passenger", "4" = "A la carte crew", "5" = "Deck crew",
"6" = "Engine crew", "7" = "Victualling crew")) %>%
mutate (classcrew1 = recode (factor(classcrew1), "1" = "1st class passenger", "2" = "2nd class passenger",
"3" = "3rd class passenger", "4" = "Crew")) %>%
mutate (age_cat = recode (factor(age_cat), "1" = "0-14", "2" = "15-30",
"3" = "31-40", "4" = "41-50", "5" = "51-60",
"6" = "61+", "99" = "NA")) %>%
mutate(age = as.numeric(age)) %>%
mutate(boatorder = as.numeric(boatorder)) %>%
mutate(boatside = recode(factor(boatside), "0" = "Starboard", "1" = "Port"))

glimpse(data)

```

(1) Analysis of bias in testimonies : example gender

```

options (digits = 3)

all <- data %>%
group_by(sex) %>%
summarize (All = n()) %>%
mutate (percent = All/sum(All) * 100) %>%
select (- sex)

survivors <- data %>%
filter(lived == "survived") %>%
group_by(sex) %>%
summarize (Survivors = n()) %>%
mutate (percent = Survivors/sum(Survivors)* 100) %>%
select (- sex)

testifiers <- data %>%
filter(lived == "survived" & testimony == "yes") %>%
group_by(sex) %>%
summarize (Testifiers = n()) %>%
mutate (percent = Testifiers/sum(Testifiers)*100)

```

```

bias_table <- cbind(testifiers, survivors, all)
bias_table

```

#(2) Frequencies response variable

```

data %>%
group_by(lived) %>%
count() %>%
ungroup() %>%
mutate(percent = n/sum(n))

```

(3) Frequencies explanatory variable

```

crosstabs <- function(df, Var1, Var2 ){

```

```

df %>%
  group_by (!! Var1, !! Var2) %>%
  summarise (n = n()) %>%
  mutate (perc = n/sum(n)* 100) %>%
  print()
}

explanatory_vars <- list("sex", "classcrew", "age_cat", "boatnumber",
  "boatside", "child", "country_5cat", "testimony")

explanatory_vars1 <- list("sex")

for (i in 1: length(explanatory_vars)){
  crosstabs(data, quo(lived), quo(eval(parse(text=paste0(explanatory_vars[[i]]))))))
}

glimpse(data)

# (4) Barchart

data %>%
  group_by(sex)%>%
  summarise (survival_rate = mean(lived == "survived"))%>%
  ggplot(aes(x = sex, y =survival_rate))+
  geom_bar(stat = "identity", fill = "darkblue")+
  labs(x = "", y = "Percentage survived", title = "Survival rate on the Titanic according to Gender")

# (4) Barchart

data$sex <- relevel(data$sex, ref = "female")

data %>%
  group_by(classcrew, sex,lived) %>%
  summarise(n = n()) %>%
  mutate (perc = n/sum(n)* 100) %>%
  filter (lived == "survived") %>%
  ggplot(aes(classcrew, perc, fill = sex))+
  geom_bar(stat = "identity", position = position_dodge(preserve = "single"))+
  labs(x = "Class & type of crew", y = "Percent",
    title = "Survival ratio according to class/crew & Sex")+
  theme(legend.title = element_blank())+
  scale_fill_manual(values=c("firebrick", "dodgerblue3"))+
  geom_text(aes(label = scales::percent(perc/100), y = perc + 2.2),
    position = position_dodge(width = 1), size = 3)

```

Textbox 4: Exploratory analysis in MAXQDA

(1) Compare first-class and third-class women

In MAXQDA, in the window with the list of codes, select the “Exercise” code. Right-click on it. A menu pops up. Select “Activate”.

Now go to tab mixed methods. Click activation by document variables. Activate first-class women, by putting [Gender] = F AND [Class] = 1 into the right-hand box. Click “Activate”. In the “View” menu, select “Selected Codings”. You should now see parts of the testimonies of three first-class women. Read through these testimonies.

Now change to see three third-class women. Click activation by document variables. Activate third-class women, by putting [Gender] = F AND [Class] = 3 into the right-hand box. Click OK. You should now see parts of the testimonies of three third-class women. Read through these testimonies. What is different in the accounts provided by first- and third-class women?

(2) Compare first-class and third-class women: codings about filling rule “women and children first”

Select the Exercise code. Right-click on the “Exercise” code. Select “Deactivate”. In the code “Filling rules” (experiencd), select the subcode “Women & Children first”. Right-click on it. Select “Activate”.

Now go to mixed methods tab. Select all first-class women as before. In the view menu, select “Selected Codings”. You should now see all coded testimonies where surviving first-class women say that the filling rule “women and children first” was used for their lifeboat. Do the same thing for third-class women, and compare.

(3) Make crosstabulation Authority acceptance * Boatorder

In MAXQDA, in the document window, activate all the documents (Boat 1 to Collapsible A).

In the code window, activate all codes under “Authority acceptance”.

Go to tab “Mixed Methods”. Go to Crosstabulation. Click on the Variable “Boat” in the left-hand window. Click below on “Insert all values into the table”. Click on the arrow. This puts all boats into the right-hand window. Now change the order of the boat (manually change the “Values” for each one and remove “boat 99”) until you have the following order (order in which lifeboats left the Titanic):

[Boat] = 7
[Boat] = 5
[Boat] = 3
[Boat] = 8
[Boat] = 1
[Boat] = 6
[Boat] = 16
[Boat] = 14
[Boat] = 12
[Boat] = 9
[Boat] = 11
[Boat] = 13
[Boat] = 15
[Boat] = 2
[Boat] = 10
[Boat] = 4
[Boat] = C
[Boat] = D
[Boat] = B
[Boat] = A

Click on OK. What generalizations can be made about authority acceptance during the filling of the lifeboats?

(4) Make crosstabulation Filling rules * Boatorder

In MAXQDA, in the document window, activate all the documents (Boat 1 to Collapsible A). These documents should now appear in red.

In the code window, activate all codes under Filling rules/Experienced. These codes should now appear in red.

Go to tab “Mixed Methods”. Go to Crosstabulation. Click on the Variable “Boat” in the left-hand window. Click below on “Insert all values into the table”. Click on the arrow. This puts all boats into the right-hand window. Now change the order of the boat (manually change the “Values” for each one and remove “boat 99”) until you have the following order (order in which lifeboats left the Titanic):

[Boat] = 7
[Boat] = 5
[Boat] = 3
[Boat] = 8
[Boat] = 1
[Boat] = 6
[Boat] = 16
[Boat] = 14
[Boat] = 12
[Boat] = 9
[Boat] = 11
[Boat] = 13
[Boat] = 15
[Boat] = 2
[Boat] = 10
[Boat] = 4
[Boat] = C
[Boat] = D
[Boat] = B
[Boat] = A

Click on OK. Did the application of filling rules change during the process of filling?

(5) Make crosstabulation Filling rules * Boatside

In MAXQDA, in the document window, activate all the documents (Boat 1 to Collapsible A). These documents should now appear in red.

In the code window, activate all codes under Filling rules/Experienced. These codes should now appear in red.

Go to tab “Mixed Methods”. Go to Crosstabulation. Click on the Variable “Boatside” in the left-hand window. Click below on “Insert all values into the table”. Click on the arrow. This puts all levels of "boatside" into the right-hand window. Now remove [Boatside] = 99

Click on OK. Use row percentages. Did the application of filling rules differ on port and starboard?

(6) Make crosstabulation Way to the boat deck * Classcrew

In MAXQDA, in the document window, activate all the documents (Boat 1 to Collapsible A), except “Other witnesses”. These documents should now appear in red.

In the code window, activate all codes under “Arrival time on deck” and “Way to boat deck – Experienced”. These codes should now appear in red.

Go to tab “Mixed Methods”. Go to Crosstabulation. Click on “Class” in the left-hand window. Click below on “Insert all values into the table”. Click on the arrow. This puts all classes into the right-hand window. Now change the order of the classes until you have the following order:

[Class] = 1
[Class] = 2
[Class] = 3
[Class] = Deck crew
[Class] = Restaurant crew
[Class] = Engine crew
[Class] = Victualling crew

Click on OK. Did different classes differ in terms of their experience when trying to get to the boat deck?

```
*****  
Textbox 5: Further quantitative analysis in SPSS  
*****
```

```
USE ALL.  
COMPUTE filter_$(boatentertime < 130).  
VARIABLE LABELS filter_$ 'boatentertime < 130 (FILTER)'.  
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.  
FORMATS filter_$(f1.0).  
FILTER BY filter_$.  
EXECUTE.
```

***** (1) Crosstabs: class and boarding time**

```
crosstabs boatentertime by classcrew  
/cells count row  
/statistics  
/barchart.
```

***** (2) Crosstabs : class, sex and boarding time**

```
crosstabs boatentertime by sex by boatside  
/cells count col  
/statistics.
```

```
#####  
Textbox 6: Further quantitative analysis in R  
#####
```

```
data$boatside <- relevel(data$boatside, ref = "Port")  
  
data %>%  
  filter (boatside != 99) %>%  
  group_by(sex, classcrew, boatside) %>%  
  summarize (n = n()) %>%  
  mutate(percent = n/sum(n)* 100)  
  
data %>%  
  filter (boatside != 99) %>%  
  group_by(sex, classcrew, boatside) %>%  
  summarize (n = n()) %>%  
  mutate(percent = n/sum(n)* 100) %>%  
  ggplot(aes(classcrew, percent, group = sex, fill = sex)) +  
  geom_col(position = position_dodge(preserve = "single")) +  
  facet_wrap(~ boatside )
```

(1) Calculate filling over time according to class / crew

```
data %>%  
  filter (boatorder != 99) %>%  
  group_by(sex, classcrew, boatorder) %>%  
  summarize (n = n()) %>%
```

```

mutate(percent = n/sum(n)* 100)

# (1bis) Plot lifeboat filling over time according to class / crew (Figure 1)

data %>%
  filter (boatorder != 99)%>%
  filter (boatside != 99)%>%
  group_by(boatside, classcrew1, boatorder) %>%
  summarize (n = n()) %>%
  mutate(percent = n/sum(n)* 100)%>%
  ggplot(aes(boatorder, percent, group = classcrew1, fill = classcrew1))+
  geom_col(position = position_dodge(preserve = "single"), width = 1.5)+
  facet_wrap(~ boatside )+
  labs(title = "Lifeboat filling over time according to class/crew and Port/Starboard", x = "Minutes after impact",
y = "Percent")+
  theme(legend.title = element_blank())

```

(2) Plot lifeboat filling according to gender and Port/Starboard (Figure 3)

```

data %>%
  select(boatside, sex1, boatentertime) %>%
  filter(boatside != "99") %>%
  filter(boatentertime < 130) %>%
  group_by (boatside, sex1, boatentertime) %>%
  summarize (n = n()) %>%
  mutate(cum_n = cumsum(n)) %>%
  ggplot(aes(x = boatentertime, y = cum_n, group = sex1, color = sex1))+
  geom_line()+
  facet_wrap(~ boatside)+
  labs(title = "Lifeboat filling according to gender and Port/Starboard", x = "Minutes after impact", y = "Entered
lifeboat")+
  scale_colour_discrete(labels = c("male", "female"))+
  theme(legend.title = element_blank())

```

Distribution of the boats on the boatdeck (Mersey's report, *Loss of the steamship « Titanic »*. Washington : August 12, 1912. ISBN 0-911962-14-X)



Lifeboat : 65 people x 14



Cutter (emergency boats) : 40 people x 2



Englehardt collapsible : 47 people x 4

