

Welcome aboard the Titanic!

Program : the link between class and survival



A grayscale background image of the RMS Titanic at sea. In the top left corner, there is a small, colorized inset image of a woman in early 20th-century attire, including a large hat and a white jacket with a dark collar.

Dataset

parameters

passengers

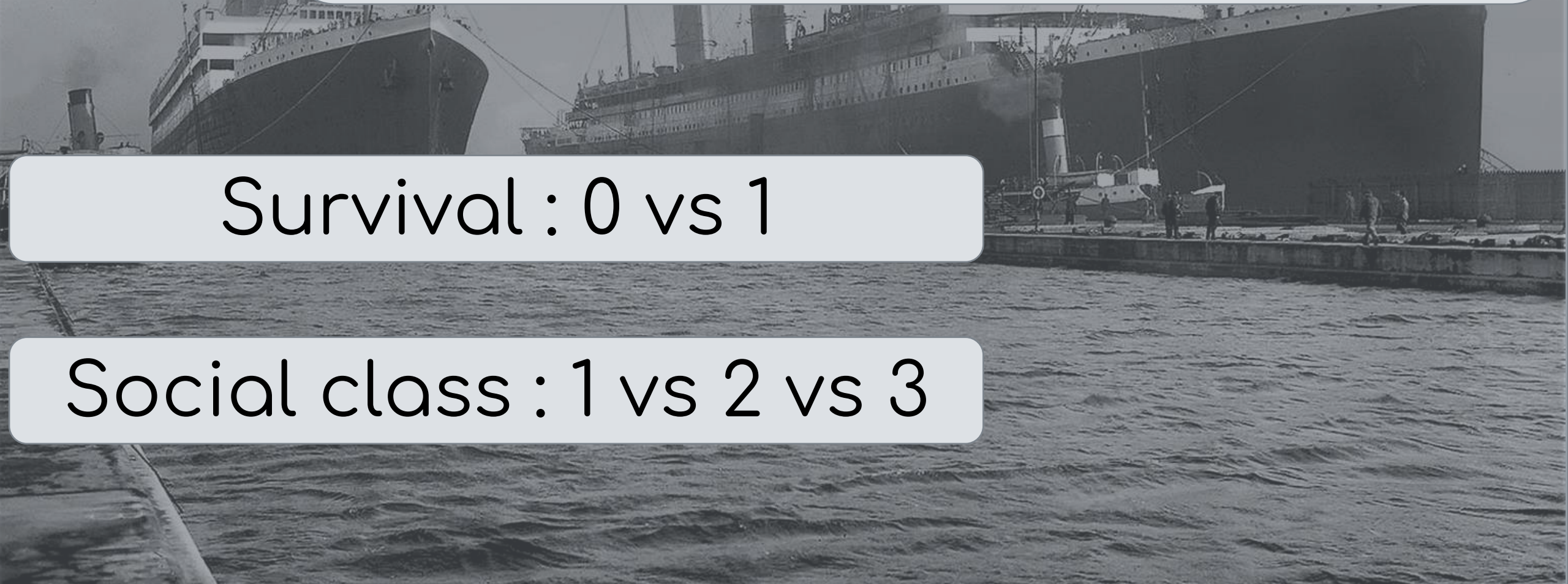
ID	survival	Social class
1	0	3
2	1	1
3	1	3
4	1	1



Is there a link between surviving the sinking of the Titanic and social class?

Survival : 0 vs 1

Social class : 1 vs 2 vs 3



Contingency table

```
titanic$Survived
titanic$Pclass    0    1
      1    80  136
      2    97   87
      3   372  119
```

2nd qualitative variable

```
(contingency_table=xtabs(~titanic$Pclass+ titanic$Survived))
```

1st qualitative variable

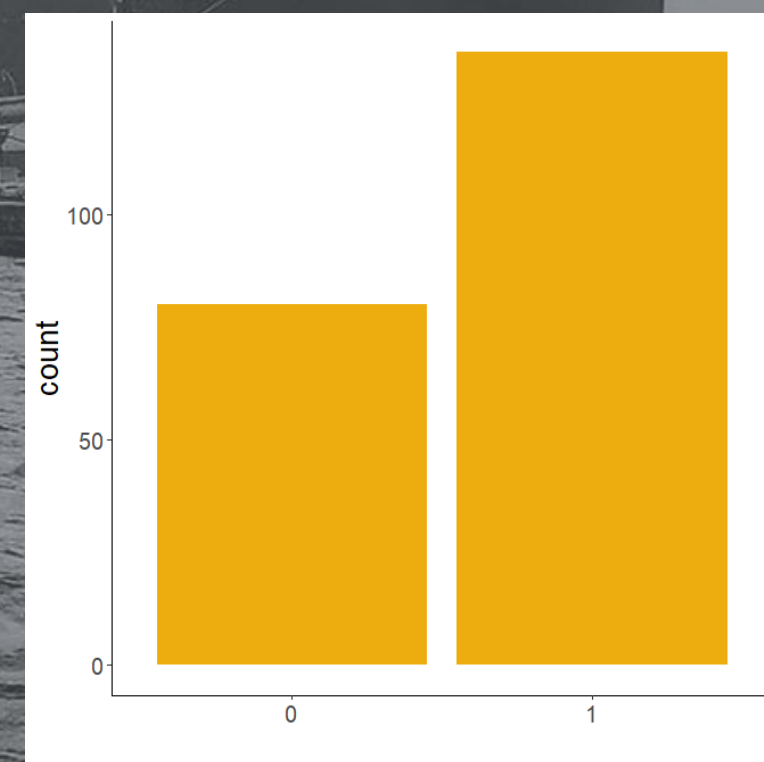
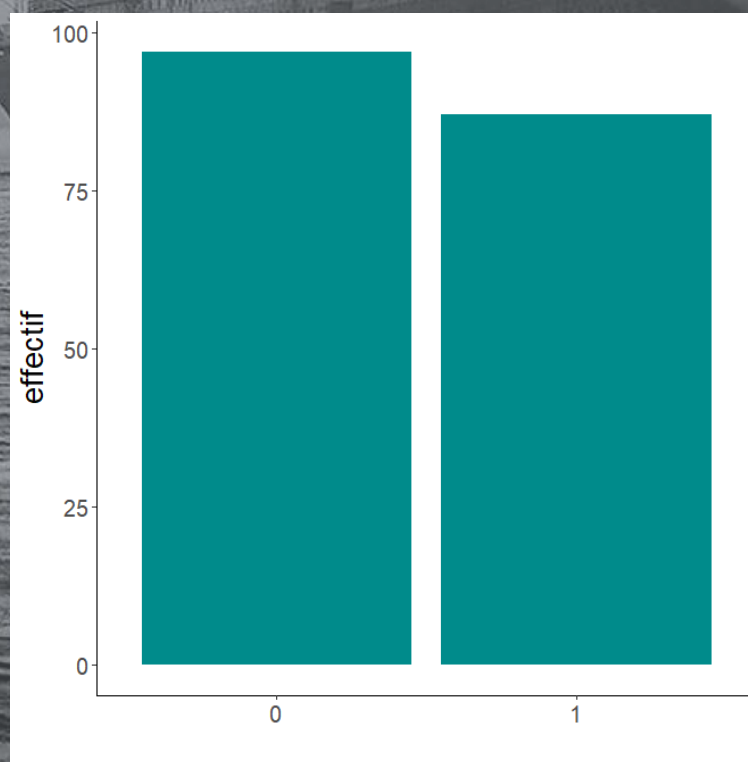
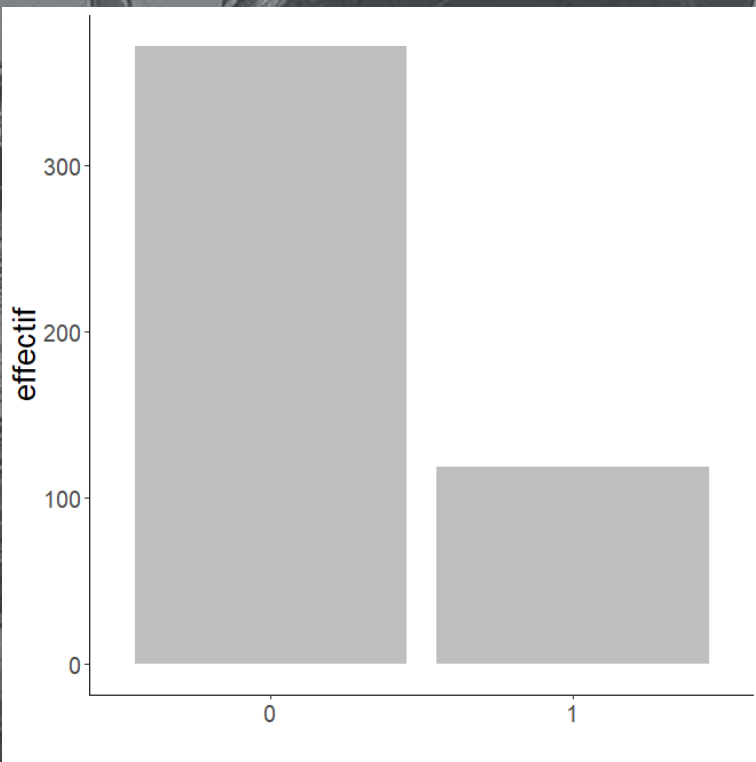
Do I tell the same story across 'lines'?

```
titanic$Survived
titanic$Pclass    0    1
1      80  136
2      97   87
3     372  119
```

Class : 3

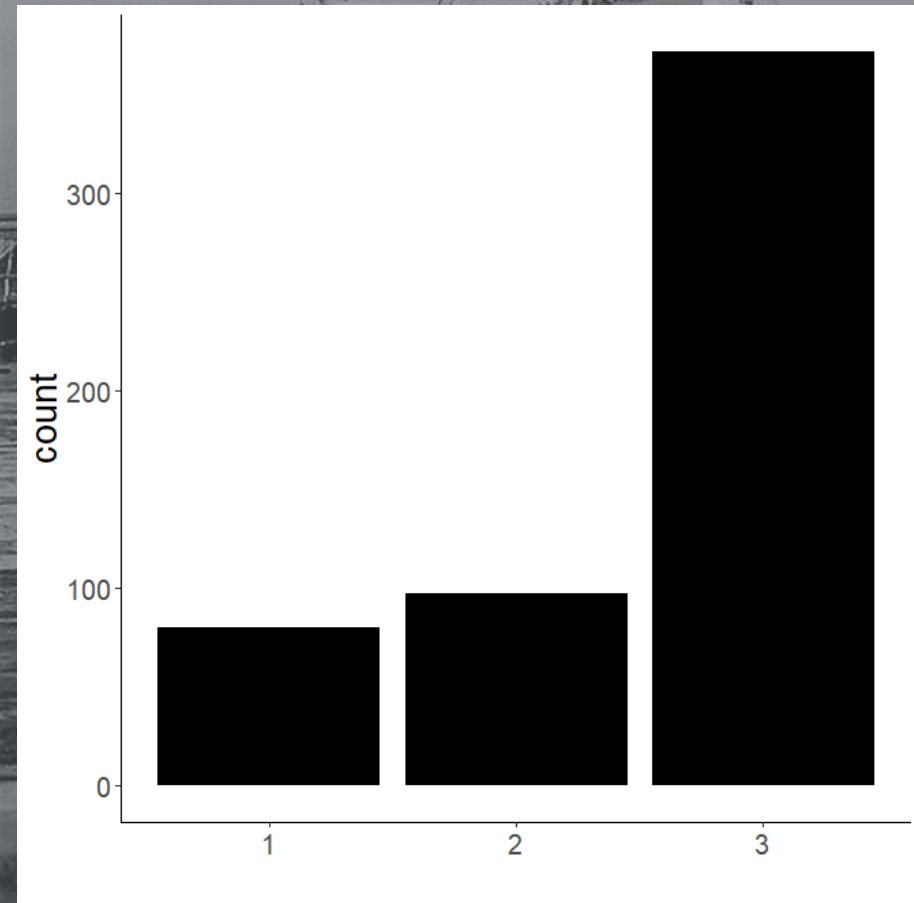
Class : 2

Class : 3

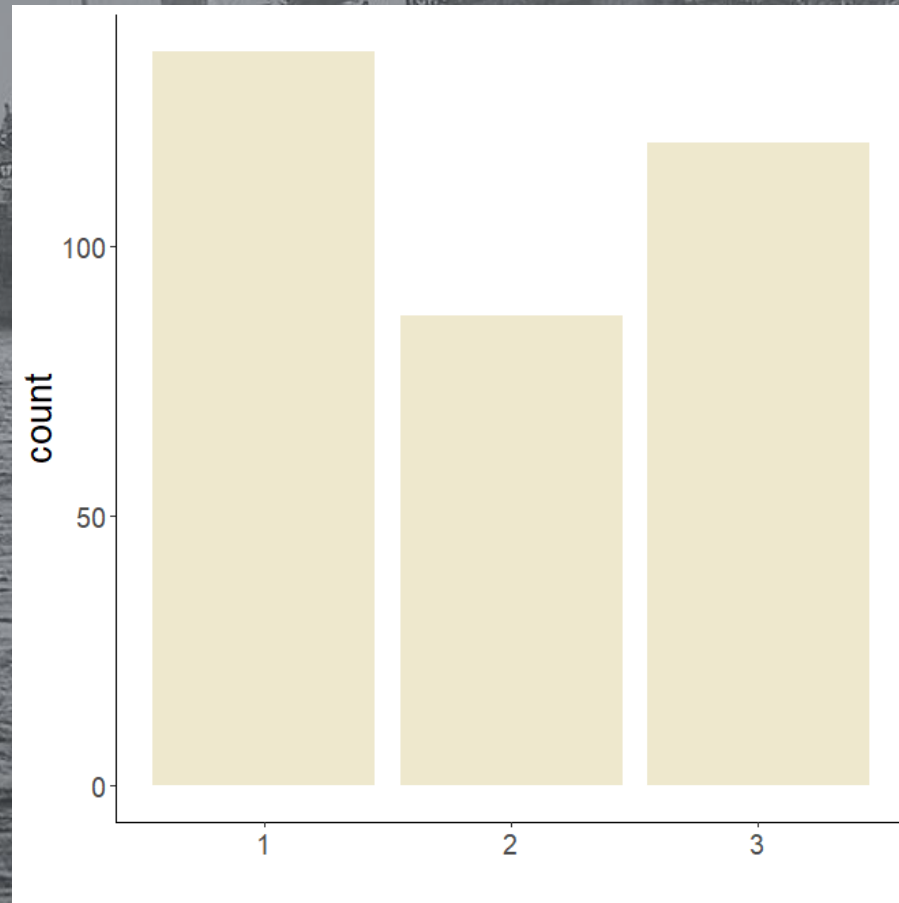


Do I tell the same story across 'columns'?


Survived: 0



Survived: 1



titanic\$Survived		
titanic\$Pclass	0	1
1	80	136
2	97	87
3	372	119



Link between 2 qualitative variables

Theoretical
counts > 5

χ^2

Theoretical
counts < 5

Fisher test

hypothesis testing

H0 (Null Hypothesis): There is no association between the variables (they are independent).

H1 (Alternative Hypothesis): There is an association between the variables (they are not independent).



How to get theoretical counts ?

We run a χ^2 test as if we had the right to.

We extract theoretical values.

We interpret the χ^2 test or we run fisher test

Theoretical count extraction

```
titanic$Survived
titanic$Pclass           0           1
1 133.0909 82.90909
2 113.3737 70.62626
3 302.5354 188.46465
```



```
test=chisq.test(contingency_table)
test$expected
```

Decision ?

Chi² interpretation

Pearson's Chi-squared test

```
data:  contingency_table  
X-squared = 102.89, df = 2,  
p-value < 2.2e-16
```



```
1 | test=chisq.test(contingency_table)  
   | test$expected  
   | test
```

Alternative : Fisher test

Fisher's Exact Test for Count Data

```
data: contingency_table  
p-value < 2.2e-16  
alternative hypothesis: two.sided
```



```
1 |  
fisher.test(contingency_table)
```


In your field



A hand is holding a brown cardboard takeout box. The box has a handle cutout at the top. The text "Take home message for chi²" is printed on the upper part of the box, and "Link between 2 qualitative variables" is printed on the lower part. The background is a blurred image of a person's arm in a white shirt.

Take home
message for χ^2

Link between 2
qualitative variables