

TeamWork

Gruppe 13

03. Juni 2022

Oskar Friedrich Rabe
580240

Mauriz Weymann
528279



<https://gruppe13-sose2022.shinyapps.io/StatistikSoSe2022Gruppe13/>

Titanic

Gruppe 13

Navigation in Shiny

Unsere App bietet ein Auswahlmenü für bis zu 4 Features.

Für besonders aussagekräftige Kombinationen lassen sich unterschiedliche grafische Darstellungen anzeigen.

The screenshot shows the Titanic Shiny app interface. It includes several interactive elements: a slider for 'Choose number of features' set to 1; a dropdown menu for 'Survived vs' set to 'Fare'; radio buttons for 'Boxplot oder Histogram' with 'Boxplot' selected; a slider for 'Choose the maximum fare' set to 180; a slider for 'Choose size of bins' set to 10; and radio buttons for 'Relativ or absolut?' with 'relativ' selected.

Daten und Datenbereinigung

- "PassengerId" – Da PassengerId eine menschengeschaffene Variable ist und nichts über die Überlebenschance der Passagiere aussagt wurde die Spalte komplett entfernt
- "Name" – Der Name der Passagiere sagt ebenfalls nichts über die Überlebenschance aus. Er kann höchstens beding als Proxyvariable für Geschlecht genutzt werden, da es Geschlecht aber bereits gibt haben wir ihn ebenfalls entfernt
- "Age" – Bereinigung aller Werte zwischen 0 und 1 und in INT konvertiert um die Dezimalwerte mancher Datensätze zu entfernen
- "Survived" – Survived wurde in einen Faktor mit zwei Stufen konvertiert
- "Pclass" – Pclass wurde zu einem Faktor mit 3 Stufen konvertiert und in Class umbenannt
- "Sex" – Sex wurde nur zu einem Faktor mit 2 Stufen konvertiert
- "SibSp" – unverändert übernommen
- "Parch" – unverändert übernommen
- "Ticket" – nicht von Interesse
- "Fare" – gerundet
- "Cabin" – nur den Substring übernommen
- "Embarked" – faktorisiert und Namen der Häfen ausgeschrieben

```
> head(titanic_data)
  Survived Class Sex Age SibSp Parch Fare Cabin Embarked
1:   Died    3 male  22    1     0  7.25  <NA> Southampton
2: Survived    1 female 38    1     0 71.28    C  Cherbourg
3: Survived    3 female 26    0     0  7.92  <NA> Southampton
4: Survived    1 female 35    1     0 53.10    C  Southampton
5:   Died    3 male  35    0     0  8.05  <NA> Southampton
6:   Died    3 male  NA    0     0  8.46  <NA> Queenstown
```

Vermutung

Wir hatten die Vermutung, dass das Geschlecht und das Alter einen Einfluss auf die Überlebenschance hatten. Da die Chance im Eismeer zu überleben auch von der körperlichen Fitness, abhängen könnte, diese aber vermutlich mit zunehmendem Alter abnimmt.

Weiter schätzten wir die Überlebenschance besser in den oberen Klassen mit höheren Fare Kosten ein, da diese in der Titanic in den oberen Decks waren und somit näher an den Rettungsbooten.

Einen Zusammenhang zwischen einer besseren Überlebenschance und der Anzahl der Geschwister/Ehegatten und der Anzahl der Eltern/Kinder an Bord vermuteten wir zuerst nicht.

Beobachtung

- 62% aller Passagiere haben nicht überlebt
- Die Überlebenswahrscheinlichkeit sinkt von Klasse 1 nach 3
Klasse 1 hat die besten Chancen, 63% aus Klasse 1 haben überlebt
Der Zusammenhang zwischen Alter und Überleben ist mit einem Cramers V von 0.34 moderat
- Frauen hatten eine bessere Chance als Männer, 74% der Frauen haben überlebt, hingegen 84% der Männer nicht. Das Geschlecht hat dabei mit einem Cramers V von 0.54 den größten Einfluß auf die Chance zu überleben
- Kleinkinder bis ca 8 Jahre hatten bessere Chancen als der Rest und Personen zwischen ca 18 und 34 hatten besonders schlechte Chancen. Generell ist die Bedeutung des Alters moderat mit einem Cramers V von 0.35
- Personen mit nur einem Ehegatten oder Geschwisterteil an Bord haben zwar oft überlebt, eine Abhängigkeit besteht mit einem Cramers V von 0.2 nur leicht
- Personen mit Kindern oder Elternteilen an Board hatten geringfügig bessere Chancen als Personen ohne, doch auch hier besteht nur eine geringe Abhängigkeit, mit einem Cramers V von 0.18
- Der Hafen, an dem die Reise angetreten wurde, enthält mit einem Cramers V von 0.17 eine Information zur Wahrscheinlichkeit
- Der Fare-Preis hat auch nur einen geringen Zusammenhang mit der Überlebenschance, da es viele Fehler im Datensatz gibt
- Die Daten aus Cabin sind sehr ungenau, nur Daten aus der 1. Klasse sind vorhanden

Fazit

Die Gruppe mit der größten Überlebenswahrscheinlichkeit sind Frauen aus der 1. Klasse, wohingegen Männer aus der 3. Klasse die geringsten Chancen hatten.

