

Technische Universität Dortmund
Fakultät Statistik
Kurs: Wissenschaftliches Arbeiten

Deskriptive Analyse des Titanic-Datensatzes

Gruppenmitglieder:
Paul Dickmann
Henning Hans
Katharina Hautzinger
Johannes Röhrig
Jannis Straub

Dozent: Steffen Maletz
Abgabedatum: 08.02.2026

Inhaltsverzeichnis

1 Einleitung	2
2 Beschreibung der Passagiergruppe	2
3 Zusammenhang zwischen Geschlecht und Überlebensrate	5
4 Zusammenhang zwischen Überlebensrate und Klassen	6
5 Zusammenhang zwischen Überleben und Alter	7
6 Zusammenhang zwischen Überleben und Fahrpreis	8
7 Zusammenhang zwischen Überleben, Geschlecht und Klasse	9
8 Fazit	10
9 Anhang	11

1 Einleitung

Ziel dieser Analyse ist es, herauszufinden, welche Faktoren die Überlebenschancen beim Untergang der Titanic, beeinflusst haben. Die Analyse basiert auf einem Datensatz über die Passagiere der Titanic aus dem Jahr 1912 zum Zeitpunkt des Schiffunglückes. Der Datensatz umfasst demographische Daten wie Geschlecht und Alter, sozio-ökonomische Merkmale wie Klasse und Ticketpreis sowie Daten über das Überleben. Zuerst wurde der Datensatz bereinigt. Anschließend erfolgte eine deskriptive Analyse, um Zusammenhänge und Unterschiede zwischen Überlebenden und Nicht-Überlebenden herauszufinden.

2 Beschreibung der Passagiergruppe

Zu Beginn der Analyse wurden die demographischen Merkmale der Passagiere auf der Titanic betrachtet. 64.8 % der Fahrgäste waren Männer, 35.2 % Frauen. Männer waren also deutlich in der Mehrheit. Das könnte daran liegen, dass die Crew auf dem Schiff wahrscheinlich vornehmlich aus Männern bestand.

Noch genauer kann man die Zusammensetzung der Passagiergruppe über die verschiedenen Anreden analysieren. 60.4 % wurden mit Mr angesprochen, waren also erwachsene Männer. 20.8 % waren unverheiratete Frauen (Anrede: Miss) und 14.4 % verheiratete Frauen (Anrede: Mrs). Außerdem bestanden die Passagiere zu 4.5 % auch aus kleinen Jungen (Anrede: Master). Weitergehend wurde die Altersverteilung der Passagiere beim Untergang analysiert (vgl. Abb. 1). Auf dem Schiff befanden sich Menschen im Alter von 0.42 bis 80 Jahren. Das arithmetische Mittel des Alters liegt bei 29.4 Jahren, während der Median und der Modalwert bei 30 Jahren liegen. Alle drei Werte liegen also ziemlich nah beieinander, was auf eine symmetrische Verteilung hindeutet. Das untere Quartil liegt bei 21 Jahren, das obere Quartil bei 35 Jahren. Die mittleren 50 % der Passagiere waren also zwischen 21 und 35 Jahre alt. Auch in Abb. 1 sieht man, dass die meisten Passagiere in der Klasse (25,30] einzuordnen sind. Die Standardabweichung liegt bei 13.25.

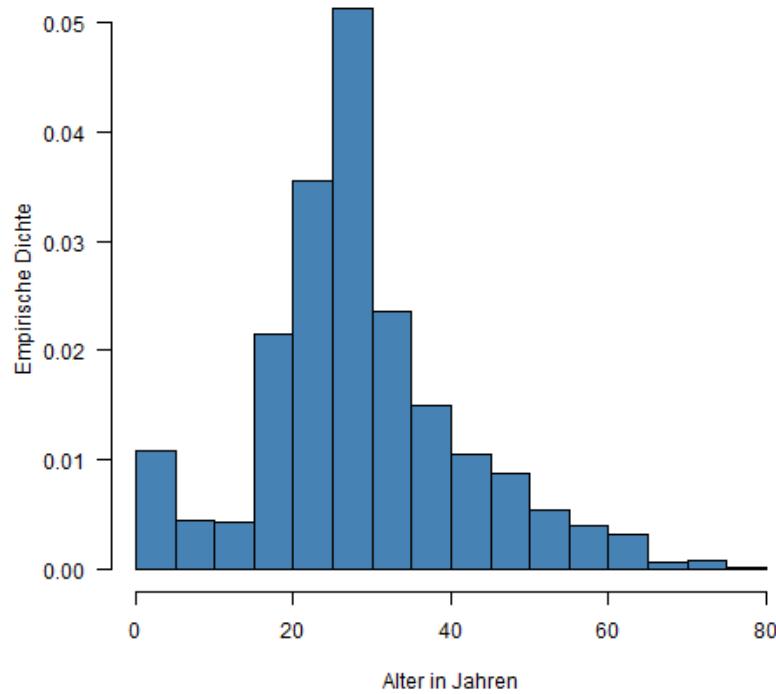


Abbildung 1: Histogramm zur Verteilung des Alters der Passagiere

Außerdem wurde auch die Verteilung der Passagiere auf die verschiedenen Buchungsklassen betrachtet (vgl. Abb. 2). Mit 55.1% hatten die meisten Kabinen der 3.Klasse gebucht. Die restlichen Passagiere verteilten sich ungefähr gleichmäßig auf die 1. und 2.Klasse: 24.2 % aller Passagiere waren in der 1.Klasse untergebracht und 20.7 % in der 2.Klasse.

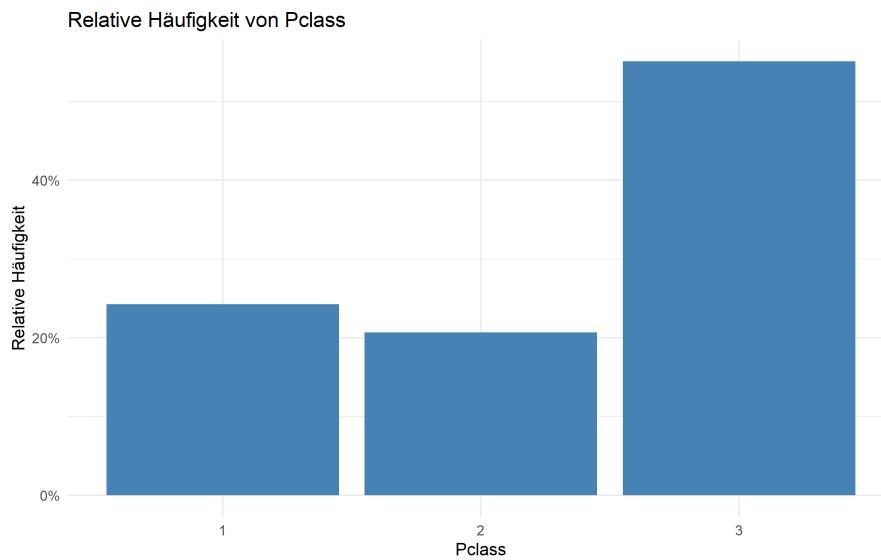


Abbildung 2: Balkendiagramm zur Verteilung der Buchungsklassen der Passagiere

Zuletzt wurde zur sozio-ökonomischen Beschreibung der Passagiergruppe die Variable Ticketpreis analysiert (vgl. Abb. 3). Die Ticketpreise erstreckten sich von 0 £ bis 512.3292 £ . Das arithmetische Mittel liegt bei 32.2 £ der Median bei 14.5 £ und der Modalwert bei 8.05 £ , was auf eine rechtsschiefe Verteilung hindeutet. Das untere Quartil liegt bei 7.8958 £ , das obere Quartil bei 31 £ . Die mittleren 50 % der Passagiere bezahlten also zwischen 7.8958 £ und 31 £ . In Abb. 3 sieht man, dass die meisten Passagiere in der Klasse [0,25] einzuordnen sind. Die Standardabweichung liegt bei 49.69 £ .

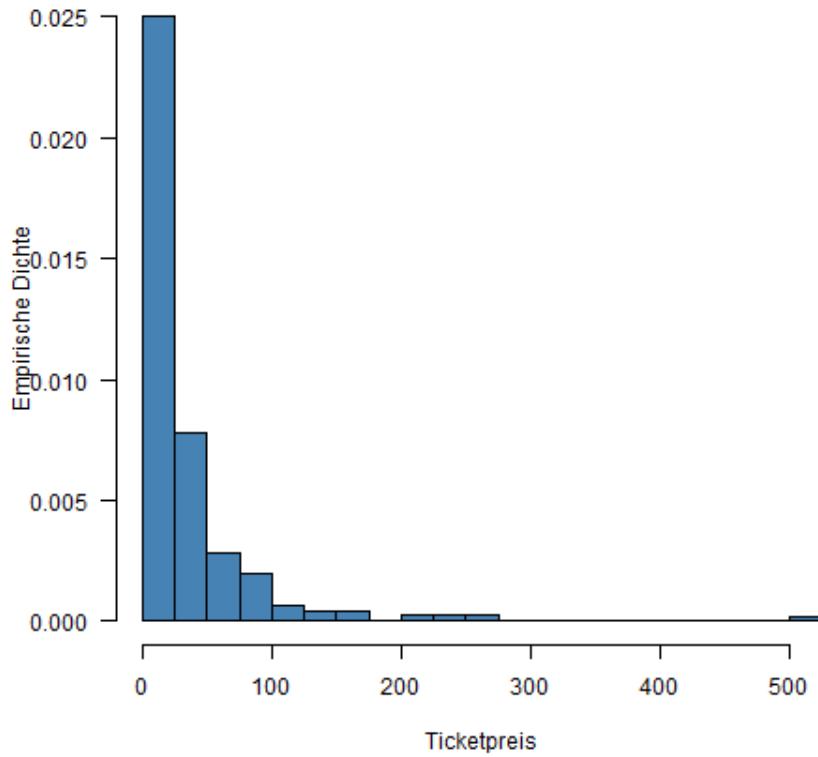


Abbildung 3: Balkendiagramm zur Verteilung der Ticketpreise der Passagiere

Zur Überlebensverteilung lässt sich sagen, dass 38.4 % der Passagiere überlebt haben, während die restlichen 61.6 % beim Untergang der Titanic gestorben sind. Die Überlebenden waren also in der Minderheit.

Insgesamt lässt sich zur Zusammensetzung der Passagiere auf der Titanic sagen, dass mehr Männer als Frauen auf der Titanic waren und sich alle im Alter zwischen 0 und 80 Jahren befanden, wobei die Hälfte zwischen 21 und 35 Jahren war. Außerdem gehörten die meisten der Passagiere zur 3.Klasse, wobei die Hälfte aller Passagiere zwischen 7.8958 £ und 31 £ für ihr Ticket bezahlten. Außerdem ist die Mehrheit der Passagiere beim Untergang gestorben.

3 Zusammenhang zwischen Geschlecht und Überlebensrate

Zunächst wurde der Zusammenhang der Variablen „Sex“ (male/female) und „Survived“ (Yes/No) untersucht. Die Kontingenztabelle zeigt, dass insgesamt mehr Männer (577) als Frauen (314) an Bord waren, jedoch mehr Frauen (233) als Männer (109) überlebt haben. Die relativen Häufigkeiten verdeutlichen dies: Etwa 65 % der Passagiere waren Männer, von denen nur ca. 19 % überlebten, während bei den Frauen etwa 74 % überlebten.

Auch die spaltenweisen relativen Häufigkeiten bestätigen den Zusammenhang: Rund 85 % der Nicht-Überlebenden waren Männer, während etwa 68 % der Überlebenden Frauen waren. Der Eindruck eines Zusammenhangs zwischen Geschlecht und Überlebensrate wird zusätzlich durch das korrigierte Kontingenzmaß nach Pearson gestützt, das mit ca. 0.67 einen starken Zusammenhang anzeigt.

Das gestapelte Balkendiagramm (vgl. Abb. 4) bestätigt diese Ergebnisse ebenfalls: Obwohl mehr Männer an Bord waren, ist der Anteil der Überlebenden bei Frauen deutlich höher, wodurch der zuvor festgestellte Zusammenhang visuell verdeutlicht wird.

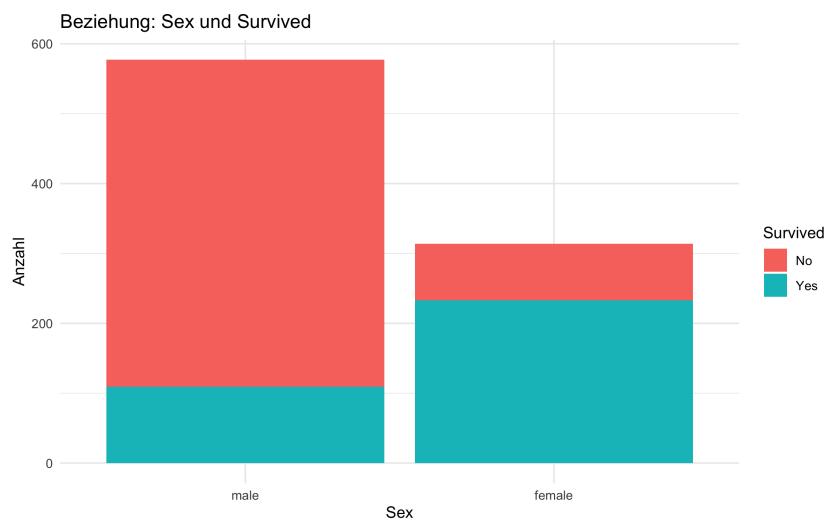


Abbildung 4: Überlebensanteil nach Geschlecht

Hinsichtlich der Interpretation dieser Ergebnisse lässt sich festhalten, dass dieser unterschiedliche Überlebensanteil wahrscheinlich mit dem Prinzip „Frauen und Kinder zuerst“ begründet ist. Demnach wurden bei der Rettung der Passagiere der Titanic, konkret bezogen auf die Besetzung der Rettungsboote, Frauen und Kinder bevorzugt. Dadurch wurden Männer als letztes gerettet, beziehungsweise waren somit die meisten der Passagiere, welche nicht gerettet werden konnten, männlich. Zudem halfen Männer eher dabei, die Rettungsboote zu besetzen, blieben also länger auf der Titanic zurück, weshalb die Überlebenschancen geringer waren. Dass jedoch nicht ausschließlich Frauen überlebt haben, liegt also vor allem daran, dass viele gerettete Kinder männlich waren.

4 Zusammenhang zwischen Überlebensrate und Klassen

Ein weiterer untersuchter Zusammenhang betrifft die Variablen „Pclass“ (1, 2, 3) und „Survived“ (Yes/No). Die Kontingenztabelle zeigt, dass die meisten Passagiere der 3. Klasse angehörten und nicht überlebt haben (372). Auffällig ist zudem, dass nur in der 1. Klasse die Mehrheit überlebt hat (136).

Betrachtet man die relativen Häufigkeiten, zeigt sich ein deutlicher Unterschied zwischen den Klassen: In der 1. Klasse überlebten etwa 63 % der Passagiere, in der 2. Klasse ca. 47 % und in der 3. Klasse nur ca. 24 %. Gleichzeitig fällt auf, dass der größte Anteil der Überlebenden aus der 1. Klasse stammt (ca. 40 %), während der größte Anteil der Nicht-Überlebenden aus der 3. Klasse kommt (ca. 68 %).

Diese Ergebnisse deuten darauf hin, dass die Überlebensrate mit der Klasse zunahm: Je höher die Klasse, desto größer die Wahrscheinlichkeit zu überleben. Das korrigierte Kontingenzmaß nach Pearson beträgt ca. 0.455 und weist damit auf einen mittleren Zusammenhang hin. Der Wert ist moderat, da sich die Überlebensraten der 2. und 3. Klasse zwar unterscheiden, aber nicht so stark wie die der 1. Klasse.

Das gestapelte Balkendiagramm (vgl. Abb. 5) visualisiert diese Eindrücke: Während die 3. Klasse die größte Passagierzahl aufweist, ist der Anteil der Überlebenden in der 1. Klasse deutlich am höchsten und in der 3. Klasse am niedrigsten.

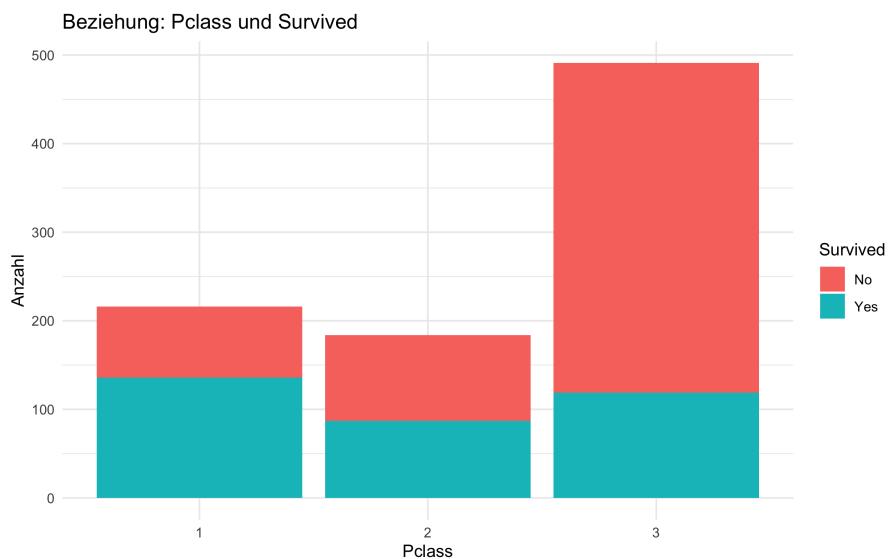


Abbildung 5: Überlebensanteil nach Klassen

Bezüglich der Interpretation dieser Ergebnisse lässt sich festhalten, dass der unterschiedliche Überlebensanteil innerhalb der Klassen vor allem in der Lage der Klassen innerhalb des Schiffes begründet liegt. Die 1. Klasse lag oben im Schiff, die 2. Klasse mittig, und die 3. Klasse weiter unten im Schiff. Ausgehend davon, dass die Passagiere vom oberen Teil des Schiffes aus in die

Rettungsboote evakuiert wurden, war also die Entfernung der 1.Klasse zu den Rettungsbooten am geringsten und von der 3.Klasse am höchsten. Diese unterschiedlichen Ausgangslagen begründen also weshalb der Überlebensanteil der niedrigeren Klasse geringer waren. Zudem konnte aufgrund der kürzeren Distanz und der geringeren Anzahl an Passagieren der 1.Klasse, diese schneller von der Schiffscrew erreicht und informiert werden, was schließlich in einer schnelleren Evakuierung von einem Großteil der 1.Klasse mündete.

5 Zusammenhang zwischen Überleben und Alter

Die deskriptiven Statistiken zeigen nur geringe Unterschiede: Nicht-Überlebende waren im Mittel etwas älter ($M = 30.2$; $SD = 12.7$) als Überlebende ($M = 28.1$; $SD = 14.1$). Die punktbiseriale Korrelation beträgt $r = -0.099$ und weist auf einen sehr schwachen negativen Zusammenhang hin.

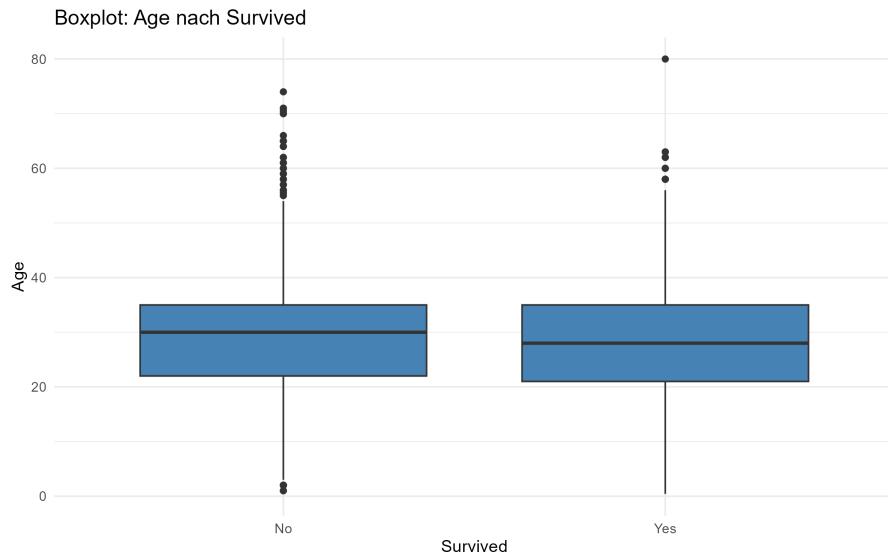


Abbildung 6: Boxplot des Alters nach Überlebensstatus

Der Boxplot (vgl. Abb. 6) zeigt stark überlappende Altersverteilungen mit ähnlichen IQRs. Beide Gruppen sind leicht rechtsschief, wobei Ausreißer im höheren Altersbereich bei Nicht-Überlebenden häufiger auftreten.

Insgesamt zeigen sich nur geringfügige Lageunterschiede, sodass das Alter keinen klaren Trennfaktor zwischen den Gruppen darstellt.

6 Zusammenhang zwischen Überleben und Fahrpreis

Zwischen Fahrpreis und Überleben zeigt sich ein deutlich stärkerer Zusammenhang. Überlebende zahlten im Mittel einen höheren Fahrpreis ($M = 48.4$; $SD = 66.6$) als Nicht-Überlebende ($M = 22.1$; $SD = 31.4$). Die punktbiseriale Korrelation von $r = 0.327$ weist auf einen moderaten positiven Zusammenhang hin, sodass mit steigenden Fahrpreisen auch die Überlebenswahrscheinlichkeit zunimmt.

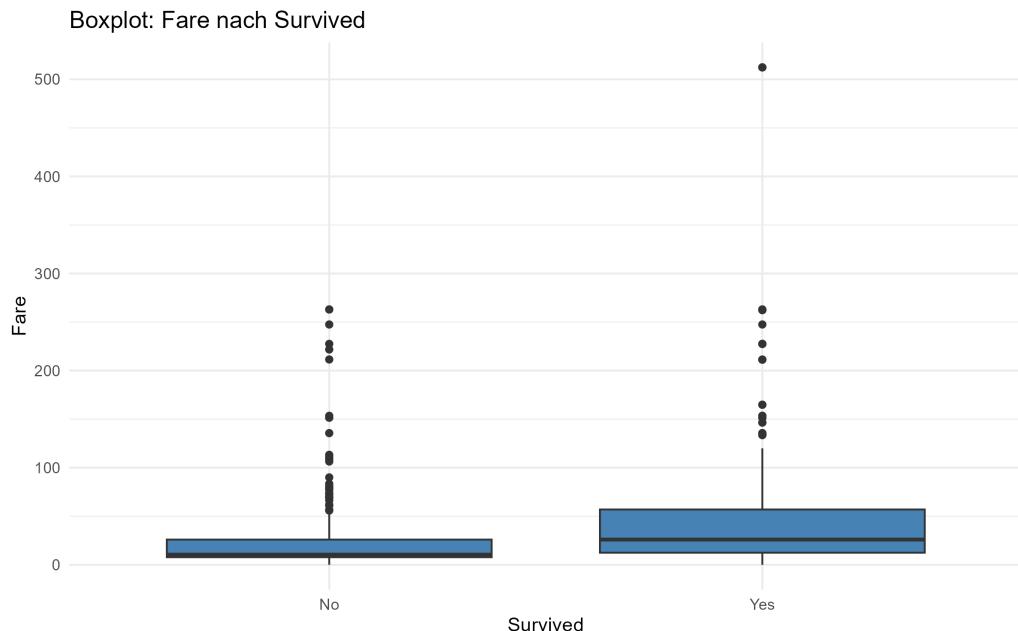


Abbildung 7: Boxplot des Fahrpreises nach Überlebensstatus

Der Boxplot (vgl. Abb. 7) verdeutlicht diese Beziehung durch eine klare Verschiebung der Verteilung zu höheren Fahrpreisen bei den Überlebenden. Ihr Median liegt deutlich über dem der Nicht-Überlebenden, zudem ist der Interquartilsabstand größer, was auf eine stärkere Streuung der mittleren Werte hindeutet. In der Gruppe der Nicht-Überlebenden treten zahlreiche hohe Ausreißer auf, die zu einer ausgeprägten Rechtsschiefe der Verteilung führen.

Insgesamt unterstützen sowohl die deskriptiven Kennwerte als auch die grafische Darstellung die Annahme, dass höhere Fahrpreise - und damit indirekt eine höhere soziale Stellung - mit einer höheren Überlebenswahrscheinlichkeit verbunden waren.

7 Zusammenhang zwischen Überleben, Geschlecht und Klasse

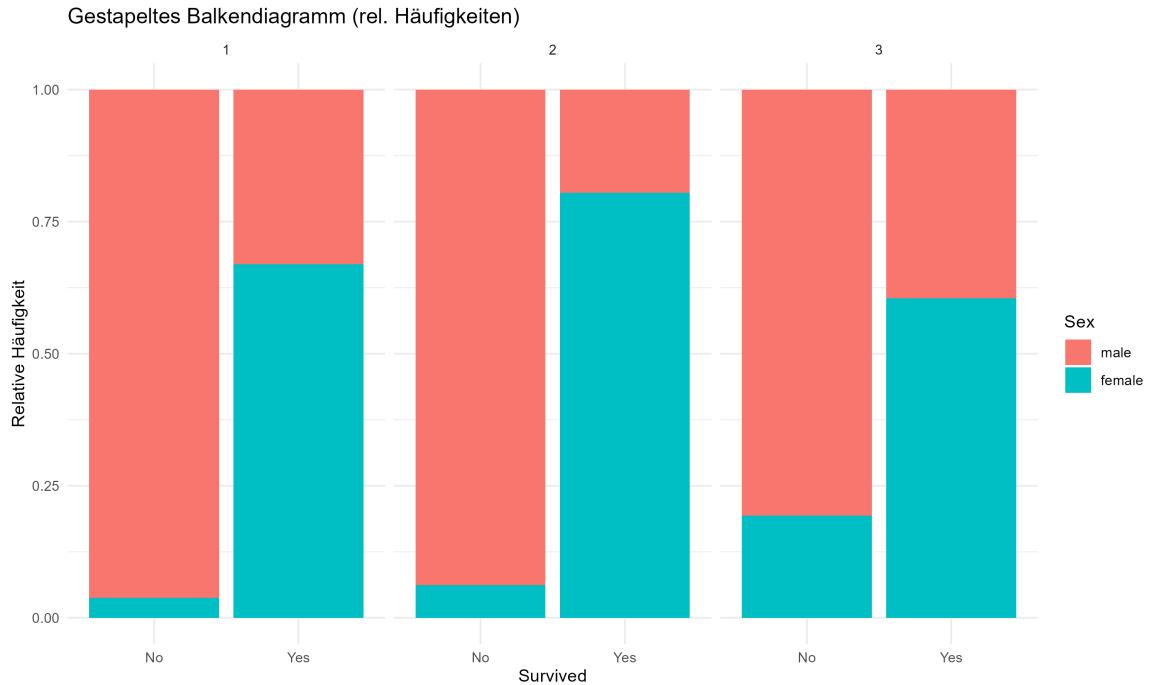


Abbildung 8: Überlebensstatus nach Geschlecht und Klasse

Das gestapelte Balkendiagramm (vgl. Abb. 8) zeigt deutliche Unterschiede im Überlebensstatus nach Geschlecht und Klasse. Über alle Klassen hinweg ist der Anteil der Überlebenden bei Frauen deutlich höher als bei Männern. Besonders ausgeprägt ist dieser Unterschied in der 1. und 2. Klasse: In der 1. Klasse überleben rund 65 % der Frauen, in der 2. Klasse knapp 80 %. Bei Männern dominieren in diesen Klassen die Nicht-Überlebenden.

Ein möglicher Grund dafür ist die bessere Zugänglichkeit zu Rettungsbooten und die Priorisierung von Frauen (und Kindern) bei der Evakuierung, wodurch Frauen der 1. Klasse schneller Zugang zu Rettungsmitteln hatten.

In der 3. Klasse zeigt sich ein anderes Muster: relativ betrachtet überleben dort die meisten Männer, während der Anteil überlebender Frauen am geringsten ist. Da in der 3. Klasse zudem knapp 55 % der Passagiere reisten (vgl. Abb. 2) – also die größte Gruppe insgesamt – beeinflusst dieses Muster die Gesamtstatistik stark. Schlechtere Unterbringung und geringerer Zugang zu Rettungsbooten könnten dazu geführt haben, dass Frauen der 3. Klasse besonders benachteiligt waren.

Insgesamt verdeutlicht die Darstellung, dass sowohl Geschlecht als auch soziale Klasse die Überlebenswahrscheinlichkeit beeinflussen, wobei Frauen und Passagiere höherer Klassen tendenziell bessere Chancen hatten.

8 Fazit

Die Analyseergebnisse lassen sich in mehreren zentralen Punkten zusammenfassen. Die Passagiere der Titanic waren überwiegend männlich, wobei das durchschnittliche Alter bei etwa 30 Jahren lag. Hinsichtlich der Klassenzugehörigkeit zeigte sich, dass der Großteil der Reisenden der 3. Klasse angehörte, während die 1. Klasse als teuerste Kategorie die geringste Anzahl an Passagieren umfasste.

Die Untersuchung der Überlebensraten – insbesondere in Bezug auf Geschlecht und Klassenzugehörigkeit – ergab deutliche Unterschiede. Von den insgesamt rund 60 % der Überlebenden waren die Mehrheit Frauen. Neben dem Geschlecht bestand zudem ein klarer Zusammenhang zwischen der Überlebenswahrscheinlichkeit und der gebuchten Klasse. Ausschlaggebend hierfür waren vor allem unterschiedliche Priorisierungen bei der Rettung sowie die räumliche Lage der einzelnen Klassen an Bord. Die vergleichsweise hohe Überlebensrate der Passagiere der 1. Klasse deutet zudem darauf hin, dass höhere Ticketpreise mit einer erhöhten Überlebenswahrscheinlichkeit einhergingen.

9 Anhang

Aufteilung des Berichts

Katharina Hautzinger:

1. Einleitung
2. Beschreibung der Passagiergruppe

Johannes Röhrig:

3. Zusammenhang zwischen Geschlecht und Überlebensrate
4. Zusammenhang zwischen Überlebensrate und Klassen
8. Fazit

Jannis Straub:

5. Zusammenhang zwischen Überleben und Alter
6. Zusammenhang zwischen Überleben und Fahrpreis
7. Zusammenhang zwischen Überleben, Geschlecht und Klasse