Bericht über die Hausarbeit

Sehr geehrter Leser,

Wir freuen uns, Ihnen unseren Bericht über unsere interaktive Visualisierung und die interessanten Ergebnisse hinsichtlich der Überlebenschancen der Passagiere auf der Titanic präsentieren zu können. In dieser Visualisierung haben wir Daten analysiert, die Informationen über verschiedene Faktoren enthalten, die sich auf die Überlebenschancen von Personen auswirken können. Wir haben versucht, diese Daten ansprechend und leicht verständlich darzustellen, um wertvolle Einblicke in den Datensatz zu ermöglichen.

Die Visualisierung basiert auf einem umfangreichen Datensatz, der Informationen zu Personen enthält, die an diesem Tag auf dem Schiff waren. Wir haben verschiedene visuelle Elemente verwendet, um die Daten zu präsentieren, darunter Balkendiagramme, Histogramme, Boxplots und Mosaikplots. Der Hauptzweck der Visualisierung war es, die Beziehungen zwischen verschiedenen Faktoren und den Überlebenschancen zu verdeutlichen und tiefergehende Einblicke zu ermöglichen.

Die Ergebnisse, die wir durch unsere Visualisierungen gefunden haben, sind äußerst spannend. Eine der interessantesten Erkenntnisse war, dass das Geschlecht einen signifikanten Einfluss auf die Überlebenschancen hatte. In unserer Visualisierung konnten wir deutlich zeigen, dass Frauen eine viel höhere Überlebensrate hatten als Männer. Basierend auf Bedingung 4 – relative Häufigkeit, erkennt man, dass Frauen eine 74,2% Chance hatten zu überleben, im Vergleich zu den Männern, die eine Überlebenschance von 18,9% hatten. Dieses Ergebnis kann auf verschiedene Gründe zurückgeführt werden, wie beispielsweise die Tatsache, dass Frauen bei der Evakuierung und Rettungsaktionen priorisiert wurden.

Ein weiterer interessanter Faktor, den wir analysiert haben, war das Alter der Personen. Hier zeigte sich, dass Kinder, Jugendliche und Erwachsene höheren Alters tendenziell eine höhere Überlebensrate hatten als Personen in anderen Altersgruppen. Diese Erkenntnis kann auf die erhöhte Aufmerksamkeit und Fürsorge zurückzuführen sein, die Kindern, Jugendlichen und Erwachsenen höheren Alters in solchen Situationen gewidmet werden.

Zusätzlich zu diesen beiden Faktoren haben wir auch andere Variablen untersucht, wie beispielsweise die Klassen der Passagiere. Dieser Faktor liefert ebenfalls interessante Einblicke und zeigt, dass Passagiere der ersten Klasse die höchsten Überlebenschancen hatten, Passagieren der zweiten Klassen die zweithöchsten und Passagiere der dritten Klasse die niedrigsten Überlebenschancen hatten. Das kann daran liegen, dass die Passagiere der ersten Klasse damals einen höheren Status hatten und somit mehr priorisiert wurden. Auch hier haben wir in den jeweiligen Passagierklassen untersucht, wie hoch die Wahrscheinlichkeit der Überlebenschance der Geschlechter ist. In den drei Passagierklassen erkennt man, dass Frauen eine höhere Überlebenschance im Vergleich zu den Männern hatten.

Als nächstes haben wir uns angeschaut, wie hoch die Überlebenschance von Eltern sind, die eine bestimmte Anzahl von Kindern hatten. In der Bedingung 3 kann man erkennen, dass Eltern mit einem, zwei und drei Kindern eine höhere Überlebenschance hatten, als Passagiere ohne Kinder. Interessanterweise kann man auch erkennen, dass ab vier Kindern, die Überlebenschance sehr niedrig, bzw. gar nicht vorhanden ist. Mögliche Gründe könnten sein, dass es vergleichsweise zu wenig Eltern gab, die 4 oder mehr Kinder hatten. Ein weiterer Grund könnte sein, dass Eltern mit 4 oder mehr Kinder, es viel schwieriger hatten, alle Kinder zu retten. Passagiere mit 0 Kindern hatten vergleichsweise eine niedrigere Wahrscheinlichkeit zu überleben. Knapp über 1/3 der Passagiere ohne Kinder haben überlebt, während knapp unter 2/3 nicht überlebt haben.

Des Weiteren haben wir uns angeschaut, wie die Überlebenschancen von Geschwister/Ehepartner standen. An dem Mosaikplot kann man erkennen, dass Passagiere, mit einem Geschwister/Ehepartner die höchste Überlebenschance hatten. Die größte Anzahl dieses Datensatz bildet die Gruppe, die keine Geschwister/Ehepartner hatten. Auch kann man an dieser Gruppe erkennen, dass knapp über nur 1/3 der Passagiere überlebt haben. Ein Grund könnte sein, dass alleinstehende Passagiere weniger bevorzugt wurden. Auf der anderen Seite könnte man behaupten, dass Passagiere mit einem Geschwister/Ehepartner, auch höchstwahrscheinlich Kinder hatten und dadurch auch priorisiert wurden. Interessant ist auch, dass desto mehr Geschwister an Bord waren, desto niedriger die Wahrscheinlichkeit der Überlebenschance ist. Einerseits kann es daran liegen, dass es generell weniger Datensätze über Geschwister/Ehepartner gibt und deswegen die Wahrscheinlichkeiten so extrem aussehen.

Insgesamt haben unsere interaktiven Visualisierungen eine Vielzahl von Informationen über die Überlebenschancen der Passagiere auf der Titanic dargestellt. Sie hat gezeigt, dass das Geschlecht, das Alter und auch die Anzahl der Familienmitglieder wichtige Faktoren sind, die sich auf die Überlebenswahrscheinlichkeit auswirken können. Die Visualisierung ermöglichte es den Benutzern, in die Daten einzutauchen und interessante Muster und Zusammenhänge zu entdecken.

Vielen Dank für Ihr Interesse an unserem Bericht. Ich hoffe, dass die beschriebene Visualisierung dazu beitragen kann, die Diskussion über Überlebenschancen und die damit verbundenen Faktoren weiter voranzutreiben.

Anmerkung: Die Prozente bei manchen Diagrammen sind umgekehrt. Wir haben dafür leider kein Fix gefunden.

Link: https://statistikprojekthtw.shinyapps.io/statistik\_projekt\_final/

Quellen:

https://statisticsglobe.com/adjust-space-between-ggplot2-axis-labels-and-plot-area-in-r

https://www.youtube.com/watch?v=mHRVFMQ54ZE

https://ggplot2.tidyverse.org/reference/index.html

<https://shiny.posit.co/>

Moodle: R-Befehle und R-Code aus der Vorlesung, sowie einzelne Teile aus den Aufgabenblätter