

Title: "Marketing Preference Analysis: Children vs. Parents"

author: "Siamak Goudarzi"

2026-01-11

Project Overview

Ziel:

Welche Variablen werden verwendet?

Schritt 1:

Schritt 2:

Schritt 3:

Schritt 4:

Häufigkeitsanalyse und Kreuztabelle

Two-Way Contingency Table: Geschlecht, Präferenz

Deskriptive Analyse (Zeilenprozentsätze):

Two-Way Contingency Table: Geschlecht, Präferenz

Analyse der Testbedingungen (Voraussetzungen):

Interpretation der Ergebnisse:

Two-Way Contingency Table: Alter, Präferenz

Überprüfung der Voraussetzungen:

Deskriptive Analyse (Zeilenprozentsätze):

Inferenzstatistik (Chi-Quadrat-Test):

Abschließende Schlussfolgerung für den Kinderbereich:

Two-Way Contingency Table: Kind, Kaufpräferenz

Deskriptive Analyse (Zeilenprozentsätze):

Inferenzstatistik (Chi-Quadrat-Test):

Interpretation des Assoziationsmaßes (Yule's Q):

Schlussfolgerung:

Fazit:

Bar Plot: Geschlecht

Bar Plot: Alter

Bar Plot: Kind

Title: “Marketing Preference Analysis: Children vs. Parents”

author: “Siamak Goudarzi”

2026-01-11

Project Overview

This project explores the relationship between demographic factors (Gender and Age) and toy preferences (Design vs. Technology). The study employs Chi-Square independence tests and association measures to validate significance.

Ziel:

Teil 1 (Kinder):

Untersuchung, ob die Präferenz für Spielzeug (Technologie oder Design) vom Geschlecht oder Alter des Kindes abhängt.

Teil 2 (Eltern):

Untersuchung, ob die Kaufpräferenzen der Eltern vom Geschlecht ihres Kindes abhängen.

Welche Variablen werden verwendet?

In beiden Abschnitten handelt es sich um qualitative Variablen (nominal/ordinal):

Geschlecht (Mädchen/Junge) -> nominal

Präferenz (Technologie/Design) -> nominal

Altersgruppe (6–9, 10–13, 14+) -> ordinal

Schritt 1:

Daten organisieren (Kontingenztabelle) Zuerst müssen wir die Rohdaten in eine Häufigkeitstabelle (Kreuztabelle) umwandeln.

Schritt 2:

Grafische deskriptive Statistik: Zeichnen Sie gruppierte Balkendiagramme, um Unterschiede visuell zu beobachten.

Schritt 3:

Statistischer Hypothesentest (Chi-Quadrat-Test):

Die Beziehung zwischen zwei nominalen Variablen wird mithilfe des Chi-Quadrat-Unabhängigkeitstests untersucht, um festzustellen, ob der beobachtete Unterschied signifikant oder zufällig ist.

>>> Nullhypothese (H_0):

Es besteht kein Zusammenhang zwischen Geschlecht/Alter und Präferenz (sie sind unabhängig).

>>> Alternativhypothese (H_1):

Es besteht ein signifikanter Zusammenhang zwischen diesen Variablen.

Schritt 4:

Berechnung der Stärke des Zusammenhangs (Assoziationsmaß): Unter der Annahme, dass der Chi-Quadrat-Test signifikant ist, müssen wir die Stärke des Zusammenhangs mithilfe von Indikatoren messen.

```
> DS_Aufgabe1bKinder <-
+   readXL("Marketing.xlsx", rownames=FALSE,
+   header=TRUE, na="", sheet="Kinder", stringsAsFactors=TRUE)
```

Error:

! `path` does not exist: 'Marketing.xlsx'

```
> library(abind, pos=17)
```

Häufigkeitsanalyse und Kreuztabelle

Die erste Frage des Projekts lautet: Besteht ein Zusammenhang zwischen Geschlecht und Spielpräferenz?

Um dies zu klären, müssen wir eine Kreuztabelle erstellen.

Two-Way Contingency Table: Geschlecht, Präferenz

```
> local({
+   .Table <- xtabs(~Geschlecht+Präferenz, data=DS_Aufgabe1bKinder)
+   cat("\nFrequency table:\n")
+   print(.Table)
+   cat("\nRow percentages:\n")
```

```

+ print(rowPercents(.Table))
+ .Test <- chisq.test(.Table, correct=FALSE)
+ print(.Test)
+ library(psych)
+ cat("\nStärke des Zusammenhangs (Phi):\n")
+ print(phi(.Table))
+ })

```

Error:

! Objekt 'DS_Aufgabe1bKinder' nicht gefunden

Deskriptive Analyse (Zeilenprozentsätze):

Jungen:

Etwa 66,3 % wählten Design und 33,7 % Technologie.

Mädchen:

Etwa 50,5 % wählten Design und 49,5 % Technologie.

>>> Erste Beobachtung:

Jungen scheinen deutlich mehr Interesse an Design zu haben als Mädchen, während sich die Mädchen fast gleichmäßig auf die beiden Optionen verteilen.

>>> Inferenzstatistik (Chi-Quadrat-Test):

Um zu beurteilen, ob dieser Unterschied in der Grundgesamtheit signifikant ist oder nur zufällig auftritt, betrachten wir die folgenden Werte:

$\chi^2 = 4,7316$: Der Wert der Teststatistik.

p-Wert = 0,02961: Dies ist der wichtigste Wert für die Entscheidungsfindung.

>>> Interpretation des p-Werts:

Da der Wert $p = 0,0296$ kleiner als das Standardfehler niveau ($\alpha = 0,05$) ist, verwerfen wir die Nullhypothese (H_0).

>>> Schlussfolgerung:

Es besteht ein signifikanter Zusammenhang zwischen „Geschlecht“ und „Spielpräferenz“.

Anders ausgedrückt: Das Geschlecht beeinflusst die Wahl des Spieltyps (Technologie oder Design).

Two-Way Contingency Table: Geschlecht, Präferenz

```
> local({
+   .Table <- xtabs(~Geschlecht+Präferenz, data=DS_Aufgabe1bKinder)
+   cat("\nFrequency table:\n")
+   print(.Table)
+   cat("\nRow percentages:\n")
+   print(rowPercents(.Table))
+   .Test <- chisq.test(.Table, correct=FALSE)
+   print(.Test)
+   cat("\nExpected counts:\n")
+   print(.Test$expected)
+ })
```

Error:
! Objekt 'DS_Aufgabe1bKinder' nicht gefunden

Analyse der Testbedingungen (Voraussetzungen):

In der Ausgabe der erwarteten Häufigkeiten sind alle vier Zahlen größer als 5 (die niedrigste Zahl ist 38,29).
Gemäß Handout 2.6, Seite 4, bedeutet dies, dass der Chi-Quadrat-Test gültig ist und der Fisher-Test nicht erforderlich war.

Interpretation der Ergebnisse:

>>> deskriptiv:

Jungen neigen mit 66,3 % eher zum Design, während Mädchen nahezu gleichmäßig (50:50) auf beide Optionen verteilt sind.

>>> Inferenzstatistisch:

Der p-Wert von 0,029 liegt unter dem Signifikanzniveau von 0,05.

>>> Interpretation der Effektstärke (Phi):

Der Phi-Koeffizient liegt bei ca. 0.16. Dies deutet nach Cohen (1988) auf einen schwachen Zusammenhang

zwischen dem Geschlecht und der Spielzeugpräferenz hin.

>>> Ergebnis:

Die Nullhypothese wird verworfen. Dies bedeutet, dass das Geschlecht einen signifikanten Einfluss auf die Präferenz der Kinder hat.

Two-Way Contingency Table: Alter, Präferenz

```
> local({
+   .Table <- xtabs(~Alter+Präferenz, data=DS_Aufgabe1bKinder)
+   cat("\nFrequency table:\n")
+   print(.Table)
+   cat("\nRow percentages:\n")
+   print(rowPercents(.Table))
+   .Test <- chisq.test(.Table, correct=FALSE)
+   print(.Test)
+   cat("\nExpected counts:\n")
+   print(.Test$expected)
+ })
```

Error:

! Objekt 'DS_Aufgabe1bKinder' nicht gefunden

Überprüfung der Voraussetzungen:

In der Tabelle der erwarteten Häufigkeiten sind alle erwarteten Werte größer als 5 (der niedrigste Wert ist 24,14).

Daher ist der Chi-Quadrat-Test gemäß Handout 2.6, Seite 4, vollständig gültig.

Deskriptive Analyse (Zeilenprozentsätze):

Die prozentualen Anteile zeigen deutlich den Trend sich ändernder Interessen im Laufe des Lebens:

6- bis 9-Jährige: Die Mehrheit (54,7 %) interessiert sich für Technik.

10- bis 13-Jährige: Nun hat sich das Interesse verschoben, und die Mehrheit (57,1 %) interessiert sich für Design.

14+-Jährige: Die Neigung zum Design nimmt stark zu (74,1 %).

Inferenzstatistik (Chi-Quadrat-Test):

p-Wert = 0,005341.

Da dieser Wert deutlich kleiner ist als das 5%-Fehlerniveau (0,05), wird die Nullhypothese (H_0) mit hoher Wahrscheinlichkeit verworfen.

Abschließende Schlussfolgerung für den Kinderbereich:

Neben dem Geschlecht hat auch das Alter einen sehr bedeutenden Einfluss auf die Spielvorlieben von Kindern.

Mit zunehmendem Alter verlagert sich die Präferenz von „Technologie“ hin zu „Design“.

```
> DS_Aufgabe1bEltern <-
+   readXL("Marketing.xlsx", rownames=FALSE,
+   header=TRUE, na="", sheet="Eltern", stringsAsFactors=TRUE)
```

Error:

! `path` does not exist: 'Marketing.xlsx'

Two-Way Contingency Table: Kind, Kaufpräferenz

```
> local({
+   .Table <- xtabs(~Kind+Kaufpräferenz, data=DS_Aufgabe1bEltern)
+   cat("\nFrequency table:\n")
+   print(.Table)
+   cat("\nRow percentages:\n")
+   print(rowPercents(.Table))
+   .Test <- chisq.test(.Table, correct=FALSE)
+   print(.Test)
+   a <- .Table[1,1]; b <- .Table[1,2]
+   c <- .Table[2,1]; d <- .Table[2,2]
+   q_val <- (a*d - b*c) / (a*d + b*c)
+   cat("\nJules Q (Assoziationsmaß):\n")
+   print(q_val)
+ })
```

```
Error:  
! Objekt 'DS_Aufgabe1bEltern' nicht gefunden
```

Deskriptive Analyse (Zeilenprozentsätze):

>>> Eltern von Jungen:

70 % bevorzugen Technik, nur 30 % interessieren sich für Design.

>>> Eltern von Mädchen:

Hier ist das Verhältnis genau umgekehrt: 55 % bevorzugen Design, 45 % interessieren sich für Technik.

Inferenzstatistik (Chi-Quadrat-Test):

p-Wert = 0,01661 Dieser Wert ist kleiner als 0,05.

Interpretation des Assoziationsmaßes (Yule's Q):

Mit einem Yule's Q von 0.48 liegt ein moderater bis starker Zusammenhang vor. Dies bestätigt, dass das

Geschlecht des Kindes die Kaufentscheidung der Eltern deutlich beeinflusst.

Schlussfolgerung:

Die Nullhypothese (H_0) wird verworfen. Dies bedeutet, dass ein signifikanter Zusammenhang zwischen dem Geschlecht des Kindes und den Einkaufspräferenzen der Eltern besteht.

Anders ausgedrückt: Eltern treffen unterschiedliche Einkaufsentscheidungen, je nachdem, ob ihr Kind ein Mädchen oder ein Junge ist.

Fazit:

>>> Präferenzen der Kinder nach Geschlecht:

Der Unterschied ist signifikant ($p = 0,029$). Jungen sind designorientierter.

>>> Präferenzen der Kinder nach Alter:

Der Unterschied ist hochsignifikant ($p = 0,005$). Mit zunehmendem Alter steigt die Designorientierung deutlich an.

>>> Kaufpräferenzen der Eltern:

Der Unterschied ist signifikant ($p = 0,016$). Eltern kaufen für Jungen hauptsächlich Technikartikel und für Mädchen eher Designartikel.

Bar Plot: Geschlecht

```
> with(DS_Aufgabe1bKinder, Barplot(Geschlecht, by=Präferenz, style="divided",
+   legend.pos="above", xlab="Geschlecht", ylab="Frequency", label.bars=TRUE))
```

Error:

! Objekt 'DS_Aufgabe1bKinder' nicht gefunden

Bei den Jungen ist die Design-Gruppe deutlich größer als die Technik-Gruppe (ca. 66 % gegenüber 34 %).

Bei den Mädchen hingegen sind die beiden Gruppen nahezu gleich groß (50:50-Verteilung).

Bar Plot: Alter

```
> with(DS_Aufgabe1bKinder, Barplot(Alter, by=Präferenz, style="divided",
+   legend.pos="above", xlab="Alter", ylab="Frequency", label.bars=TRUE))
```

Error:

! Objekt 'DS_Aufgabe1bKinder' nicht gefunden

>>> 6- bis 9-Jährige:

Die Spalte „Technik“ ist höher. Jüngere Kinder interessieren sich mehr für Technologie.

>>> 10- bis 13-Jährige:

Die Spalten sind umgekehrt, und die Spalte „Design“ ist etwas höher.

>>> Ab 14 Jahren:

Die Spalte „Design“ ist deutlich höher und hat einen klaren Vorsprung (74 %).

Bar Plot: Kind

```
> with(DS_Aufgabe1bEltern, Barplot(Kind, by=Kaufpräferenz, style="parallel",
+   legend.pos="above", xlab="Geschlecht des Kindes", ylab="Prozent",
+   main="Kaufpräferenz der Eltern", scale="percent", label.bars=TRUE))
```

Error:

! Objekt 'DS_Aufgabe1bEltern' nicht gefunden

>>> Jungen:

Die Spalte für Technik ist deutlich länger als die für Design. Dies zeigt, dass die überwiegende Mehrheit der Eltern (70 %) technologische Angebote für ihre Söhne bevorzugt.

>>> Mädchen:

Die Spalte für Design ist länger (55 %). Dies bedeutet, dass Eltern bei Mädchen eher Wert auf Design legen.