8 Glossar

Absolute Häufigkeit - relative Häufigkeit
Die absolute Häufigkeit gibt an, wie oft ein Ergebnis bei der mehrfachen Durchführung eines Zufallsexperiments auftritt.
Die relative Häufigkeit gibt den Anteil eines Ergebnisses an der Gesamtzahl der Versuchsdurchführungen an.

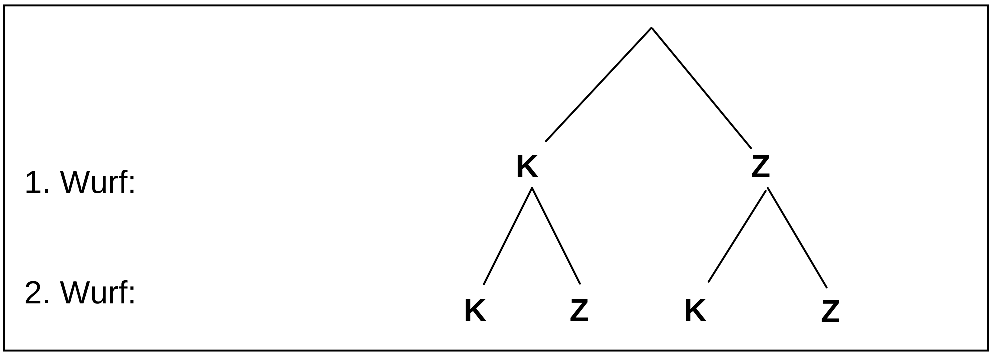
Beispiel:
Untersucht wird das Merkmal „Augenfarbe der Schülerin oder des Schülers“ einer Klasse. Mögliche Ausprägungen (Werte, Ergebnis) der Augenfarbe sind: blau, braun, grau, grün. D. h., jedem Schüler/jeder Schülerin wird genau eine Augenfarbe zugeordnet.
Haben nun 8 von 32 Schülerinnen bzw. Schülern blaue Augen, so ist

- 8 die absolute Häufigkeit
- und $\frac{8}{32} = \frac{1}{4}$ die relative Häufigkeit.

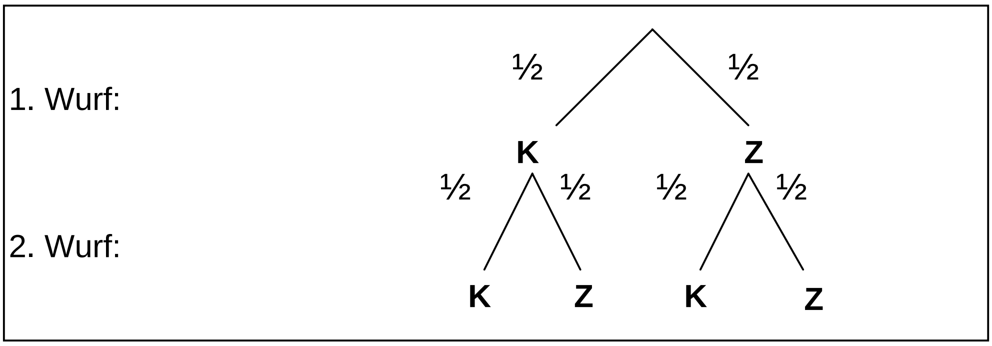
Balkendiagramm (siehe auch Säulendiagramm)
Wie Säulendiagramm, aber x- und y-Achse vertauscht, d. h., ein um 90° gedrehtes Balkendiagramm.

Baumdiagramm
Zur Darstellung unterschiedlicher Ergebnisse oder als Hilfe beim Zählen oder zur Darstellung mehrstufiger Zufallsversuche sind Baumdiagramme hilfreich.

1. Beispiel Abzählbaum:
Zweimaliges Werfen einer Münze. Dabei gibt es vier Ergebnisse für diesen zweistufigen Zufallsversuch: KK, KZ, ZK, ZZ.



2. Beispiel Wahrscheinlichkeitsbaum:
Zweimaliges Werfen einer Münze, benutzen des Baumdiagramms für die Pfadregel.



$$P(\{KK\}) = P(\{KZ\}) = P(\{ZK\}) = P(\{ZZ\}) = \frac{1}{4}$$

Bedingte Wahrscheinlichkeit
Damit ist die Wahrscheinlichkeit für das Eintreffen eines Ereignisses gemeint, falls ein bestimmtes Ereignis (Bedingung) bereits eingetreten ist. Die bedingte Wahrscheinlichkeit bezieht sich deshalb auf eine Teilmenge der Ergebnismenge, die durch das Eintreffen einer Bedingung festgelegt wird.

> Beispiel: Beim Würfeln mit einem idealen Würfel ist die Wahrscheinlichkeit für eine 6 unter der Bedingung, dass eine gerade Zahl gewürfelt wurde, 1/3, denn 6 ist eine der drei gleichwahrscheinlichen Möglichkeiten 2, 4, 6. Die zusätzliche Bedingung verändert die Wahrscheinlichkeit, da sich die betrachtete Ergebnismenge ändert (siehe auch Vierfeldertafel).

Boxplot
Ein Boxplot ist ein Instrument der explorativen Datenanalyse[1]. Er ist ein Diagramm, das die Verteilung eines Ergebnisses grafisch veranschaulicht. Beim Boxplot genügt die Darstellung oder Angabe von fünf Kenngrößen, um insbesondere Aussagen zur Streuung und Symmetrie der Verteilung des betrachteten Ergebnisses zu machen. Gibt man mehrere Boxplots zu einer gemeinsamen Achse an, lassen sich Verteilungen auf einen Blick miteinander vergleichen. Das wird z. B. in der Medizin benutzt, wenn mehrere Gruppen von Patienten bezüglich ihrer Daten verglichen werden sollen.

Die fünf Kenngrößen, die in einem Boxplot gezeichnet werden, kann man leicht ermitteln, wenn man die Daten der Größe nach ordnet:

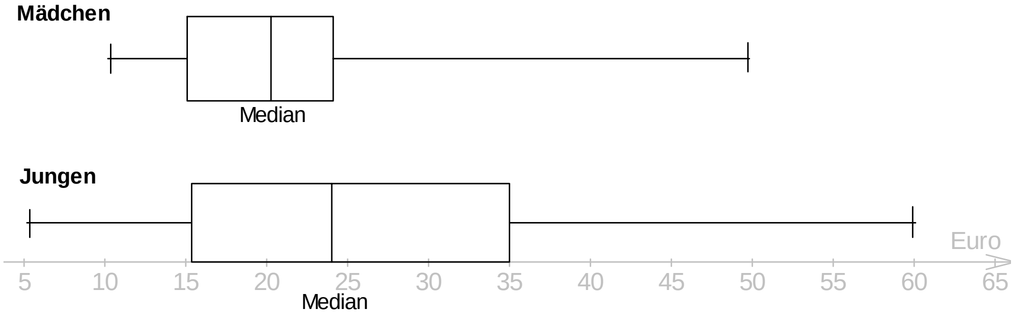
- Die erste Kenngröße ist der Median M (Wert in der Mitte der Datenreihe).
- Die zweite und dritte Kenngröße, das untere bzw. das obere Quartil (Q_1 bzw. Q_3), bilden die Werte, die die beiden durch den Median geteilten Hälften halbieren. Der Datensatz wird also insgesamt in vier Teile geteilt.

  Auf diese Weise liegen zwischen dem unteren Quartil und dem Median 25 % aller Werte genauso wie zwischen dem Median und dem oberen Quartil. Diese Werte bilden den Kasten (Box). Handelt es sich bei der Liste um eine gerade Anzahl von Daten, nimmt man den benachbarten Wert.
- Die vierte und fünfte Kenngröße bilden der kleinste (Minimum M_i) bzw. der größte Wert (Maximum M_a) des Datensatzes. Sie werden durch „Antennen“ (whiskers) an der Box dargestellt.

  Beispiel:
  „Wie viel Euro gibst du pro Monat für das Telefonieren mit dem Handy aus?"
  Geordnete Datenreihe für die Antwort von 33 Jungen:

  5 (=M_i), 6, 6, 10, 12, 14, 15, 16, 16(=Q_1),17, 17, 18, 20, 20, 22 , 22, 24(=M), 24, 24, 25, 25, 28, 32, 33, 35(=Q_3), 35, 38, 40, 40, 42, 48, 50, 60 (=M_a).

  Die Befragung von Mädchen wird als Boxplot in dasselbe Diagramm eingetragen:



[1] Explorative Datenanalyse (EDA): Tukey (Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley), New York
Siehe auch Deutsche Gesellschaft für Psychologie
http://www.dgps.de/fachgruppen/methoden/mpr-online/issue1/art3/node8.html

Ohne genauere Kenntnis über die Einzeldaten der Befragung bei den Mädchen wird deutlich:

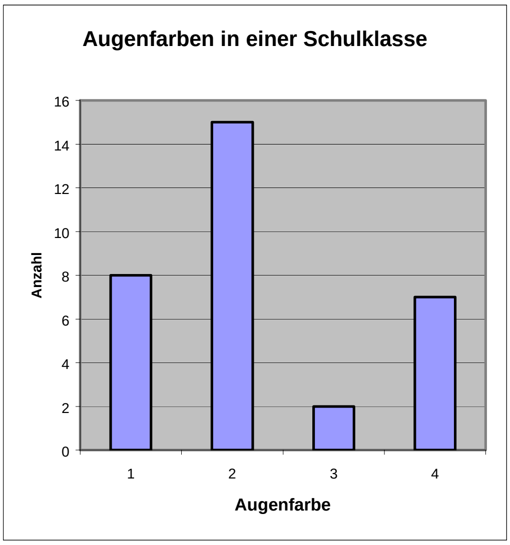
> 50 % der Mädchen geben zwischen 15 € und 24 € monatlich für ihr Handy aus. Der geringste Betrag liegt bei 10 €, der höchste bei 50 €. Verglichen mit den Ausgaben der Jungen geben 50 % der Mädchen so viel Geld aus wie 25 % der Jungen, also sind die Mädchen im Mittel sparsamer.

Diagramm
Die Ergebnisse einer Datenerhebung werden grafisch dargestellt.
Folgende Diagrammarten sind im Glossar aufgeführt: Balkendiagramm, Boxplot, Kreisdiagramm, Liniendiagramm, Piktogramm, Säulendiagramm.

Ergebnis und Ereignis
Ein Zufallsexperiment (z. B. das Werfen eines Würfels) hat einen Ausgang: das Ergebnis. Im Fall des Würfels sind die Augenzahlen mögliche Ergebnisse. Diese fasst man zur Ergebnismenge zusammen. Oft interessiert man sich aber dafür, ob eines von mehreren Ergebnissen eingetroffen ist. Man spricht davon, dass ein Ereignis eingetreten ist. Im Fall des Würfels ist das Ereignis „weder eine Eins noch eine Sechs zu würfeln" eingetreten, wenn ein Element der Menge {2, 3, 4, 5} das Ergebnis des Zufallsexperiments ist.

Gesetz der großen Zahl
Mit zunehmender Zahl der Durchführungen eines Zufallsexperiments verändern sich die relativen Häufigkeiten eines Ergebnisses immer weniger. Sie liegen dann in der Nähe der Wahrscheinlichkeit eines Ereignisses.

Grafische Darstellungen und Skalierung
Wenn man etwas „messen“ möchte, muss eine Skala festgelegt werden, auf der die Ergebnisse abgetragen werden können.



Im Beispiel der Augenfarbe kann man dieser auch Zahlenwerte zuordnen, so z. B. 1 für blau, 2 für braun usw. Dann können die Werte auch auf einer Skala angegeben werden und z. B. die absoluten Häufigkeiten auch als Säulendiagramm entsprechend dargestellt werden.

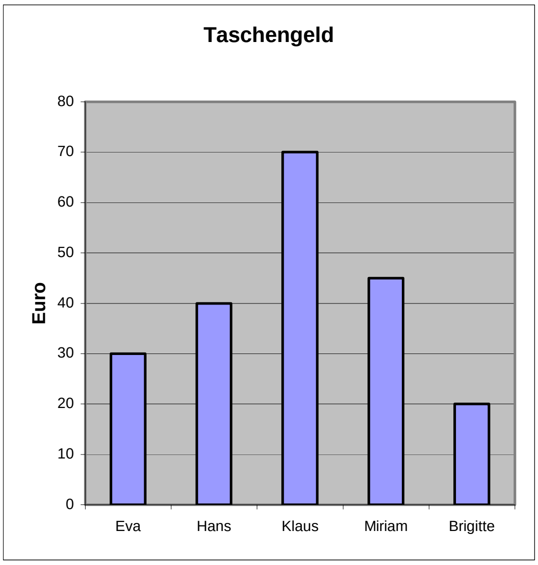
Entscheidend ist dabei, wie man mit den Skalenwerten umgehen kann. Bei der Augenfarbe handelt es sich um eine sog. Nominalskala. Man kann diese weder sortieren, noch kann man mit diesen Zahlen rechnen oder etwa Mittelwerte bilden.
Die nächste Art von Skalen ist die Ordinalskala. Das bekannteste Beispiel von Merkmalen, die auf dieser Skala abgebildet werden, sind die Schulnoten. Sie lassen sich anordnen, denn eine Eins ist besser als eine Zwei und diese wiederum besser als eine Drei usw. So kann man den Wert in der Mitte der geordneten Reihe angeben (**Median**) und den am häufigsten vorkommenden (**Modalwert**). Weil aber gleiche Differenzen inhaltlich unterschiedlich sind (der Unterschied zwischen 2 und 3 ist ein anderer als zwischen 4 und 5) kann man nicht sinnvoll mit diesen Noten rechnen und insbesondere keine Mittelwerte bilden. Dies ist erst bei metrischen Skalen möglich.

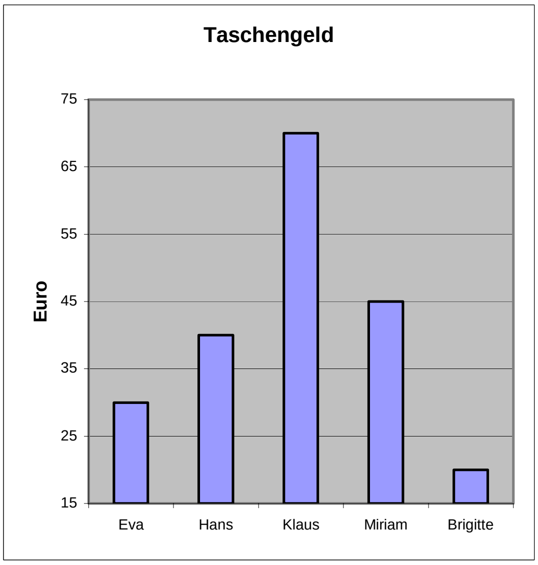
Nachfolgende Übersicht stellt unter anderem eine Beziehung zwischen verschiedenen Skalen und den dafür angemessenen grafischen Darstellungen her.

Name	Beispiele	Diagrammarten	erlaubte Relation	zulässige statistische Operationen
Nominalskala	Geschlecht, Lieblingsfarbe, Telefonnummer	Piktogramme, Balken- und Säulendiagramme, Kreisdiagramme	Gleichheit und Ungleichheit	Berechnung von absoluten, relativen, prozentualen Häufigkeiten, Modalwert
Ordinalskala	Tabellenplätze im Sport, Schulnoten	zusätzlich: Boxplots	zusätzlich: Größer- und Kleinerrelation, Rangordnung von Objekten	Häufigkeitsverteilungen Modalwert, Median, Berechnung von Quantilen
Metrische Skalen Intervallskala	Kalenderdaten, Temperaturskalen in Celsius, Punktezahlen bei Leistungstests	zusätzlich: Boxplots und Perzentilbänder, die um das arithmetische Mittel angeordnet sind	Größe von Unterschieden bei freier Wahl der Maßeinheit und des Nullpunktes	zusätzlich: arithmetische Mittelwerte, Standardabweichung
Verhältnisskala	Längen, Gewicht, Alter, Temperaturskala in Kelvin	Diagrammarten: alle	Zusätzlich: Proportionen, Existenz eines absoluten Nullpunktes	Geometrischer Mittelwert, Variationskoeffizient

Grafische Darstellungen und Täuschungen

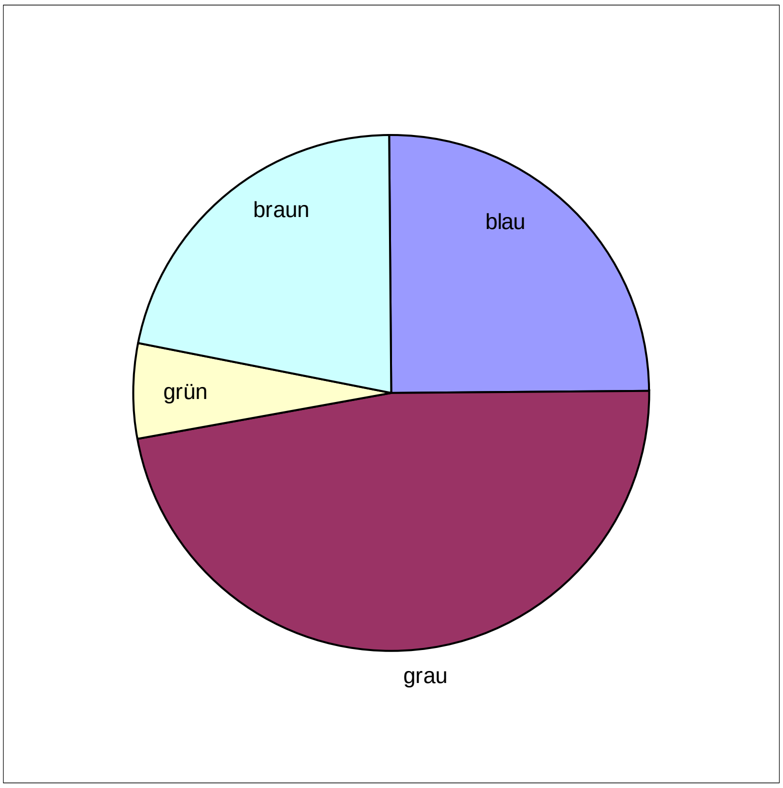
Häufig werden Grafiken so gestaltet, dass sie einen falschen Eindruck vermitteln. Dies ist z. B. der Fall, wenn von der y-Achse nur ein Ausschnitt abgebildet wird.





Falsche Eindrücke entstehen auch dadurch, dass räumliche Darstellungen benutzt werden, bei denen alle Längen entsprechend vergrößert werden (vgl. Beispiel „Wir fördern Familien“ aus der Agenda 2010, Seite 3).

Kreisdiagramm



Laplace-Experimente
Sind alle Ergebnisse eines Zufallsexperiments gleich wahrscheinlich, spricht man von einem Laplace-Versuch oder einem Laplace-Experiment.
Es gilt dann für die Wahrscheinlichkeit eines Ereignisses:

$$P(\mathit{Ereignis}) = \frac{\mathit{Anzahl\ der\ zum\ Ereignis\ gehörenden\ Ergebnisse}}{\mathit{Anzahl\ aller\ Ergebnisse}}$$

Spielgeräte, die zu Laplace-Experimenten führen:
- Glücksräder
- verschiedene Spielwürfel
- Münze
- Urne

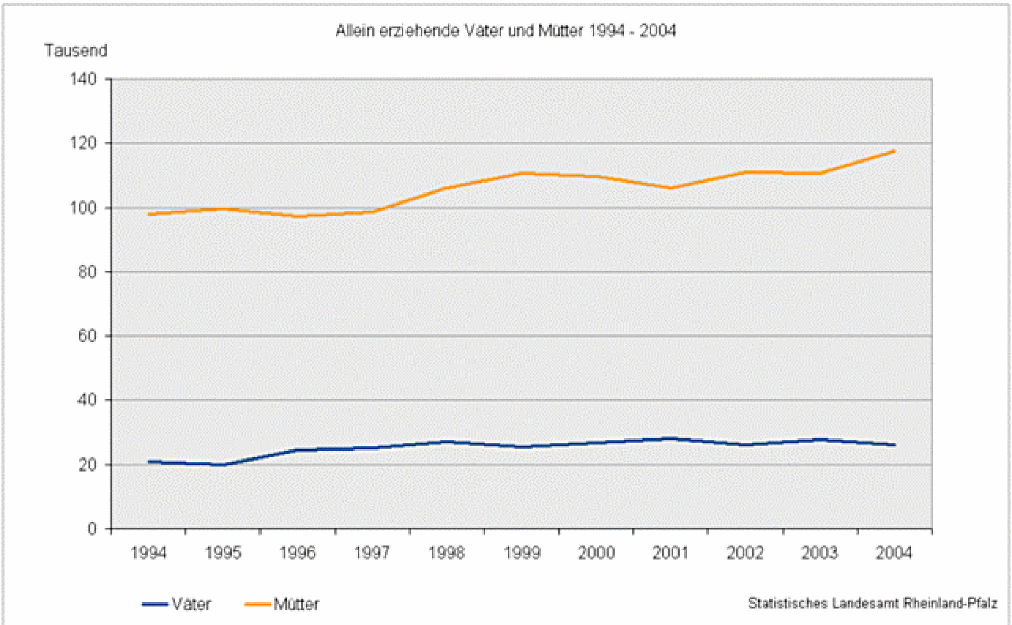
Nicht-Laplace-Experimente

Spielgeräte, die nicht zu Laplace-Experimenten führen:

- Reißnagel
- unregelmäßiger Würfel (Riemer- oder Ludos-Würfel)
- Schweinchen (aus dem Spiel „Schweinerei“)

Liniendiagramm

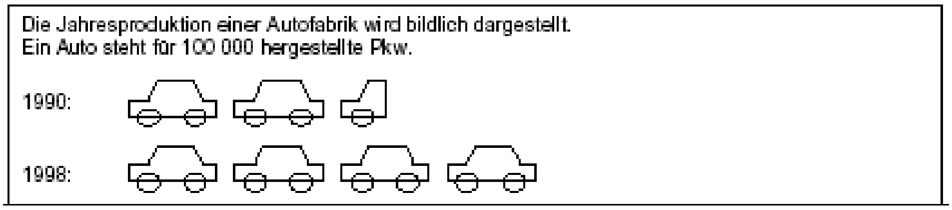
Liniendiagramme werden verwendet, um Trends in Datensätzen zu verdeutlichen.



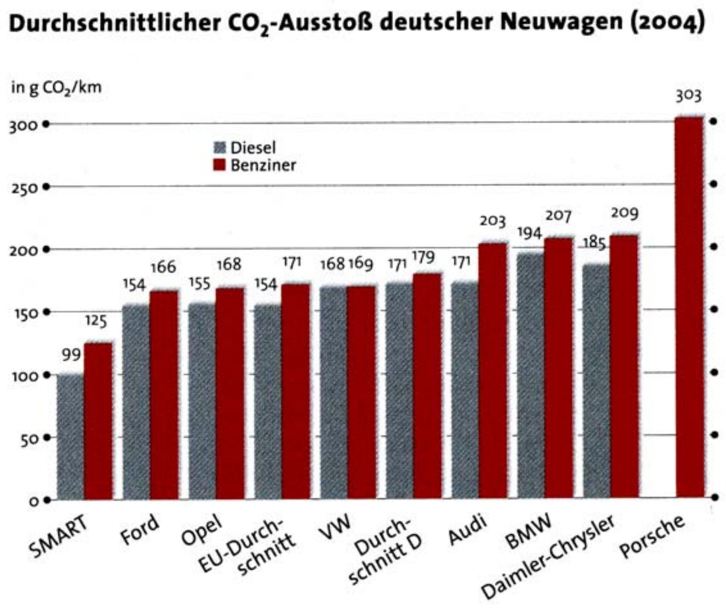
Median

Wert, der in der Mitte einer geordneten Datenreihe liegt.

Piktogramm

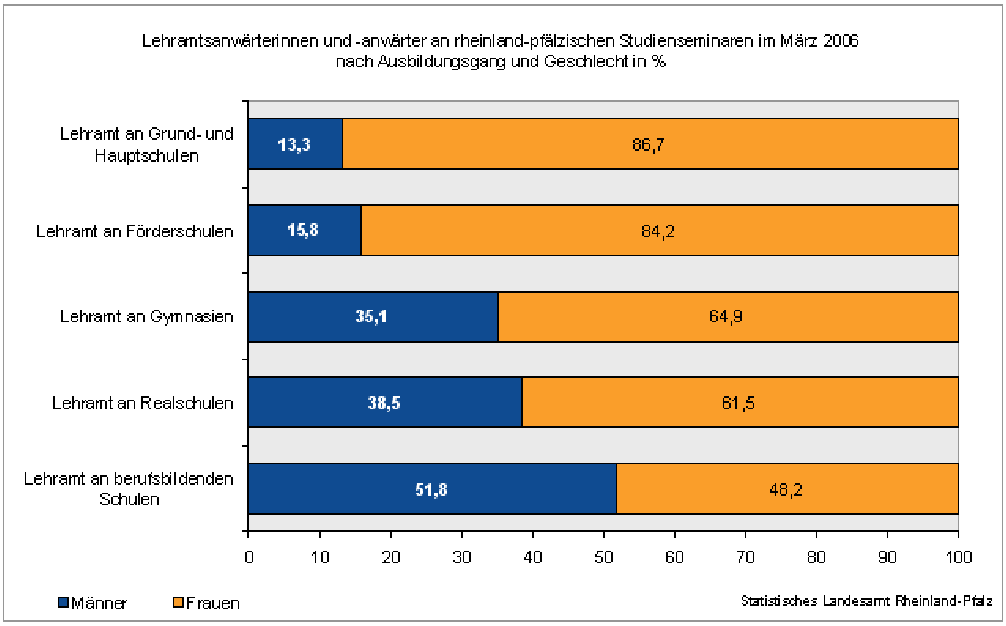


Säulendiagramm (siehe auch Balkendiagramm):



Streifendiagramm

Der Balken wird so unterteilt, dass die Länge eines Streifens die absolute oder relative Häufigkeit eines Merkmals angibt.



Es gibt auch Säulendiagramme, bei denen die Säulen als Streifendiagramme ausgebildet sind.

Simulation

Bei der Simulation werden Wahrscheinlichkeiten durch Zufallsexperimente näherungsweise bestimmt. Häufig werden dazu Urnen- oder Würfelexperimente sowie Glücksräder benutzt. Simulationen können auch mit dem Computer ausgeführt werden.

Spannweite

Differenz zwischen dem größten und kleinsten Wert einer Datenreihe.

Wahrscheinlichkeit
Bei den Laplace-Experimenten lässt sich die Wahrscheinlichkeit ermitteln, indem man überlegt, welche Ergebnisse für das Eintreten des gesuchte Ereignisses günstig sind. Dann ergibt der Quotient aus der Anzahl der für das Ereignis günstigen Ergebnisse und der Anzahl aller Ergebnisse die gesuchte Wahrscheinlichkeit.
Damit liegen die Zahlenwerte für Wahrscheinlichkeiten im Intervall [0;1].

Beispiel:
Gesuchtes Ereignis ist die Augenzahl beim Würfeln, die durch 2 oder 3 teilbar ist:
Das Ereignis enthält die Augenzahlen 2, 3, 4, 6:

$$E = \{2, 3, 4, 6\};\ P(E) = \frac{Anzahl\ der\ günstigen\ Ergebnisse}{Anzahl\ der\ möglichen\ Ergebnisse} = \frac{4}{6} = \frac{2}{3}$$

Bei denjenigen Zufallsexperimenten, bei denen die einzelnen Ergebnisse nicht gleich wahrscheinlich sind, kann die Wahrscheinlichkeit über die relative Häufigkeit annähernd bestimmt werden: Bei einer genügend großen Anzahl von Versuchen stabilisiert sich die relative Häufigkeit auf einen bestimmten Wert. Dieser gilt als Schätzwert für die Wahrscheinlichkeit (vgl. Gesetz der großen Zahl).

Zeitreihe
Die Ausprägung eines Merkmals wird über längere Zeit immer wieder überprüft

Beispiel:
Entwicklung der Teilnehmerzahlen in den deutschen Digitalnetzen
(Stand Juni 2003, Quelle http://www.handy-seiten.de/14-Statistiken/14-statistiken.html)

Zufallsexperimente
Versuche, bei denen verschiedene Ergebnisse möglich sind. Es lässt sich aber nicht vorhersagen, welches Ergebnis letztendlich eintritt (vgl. Laplace-Experiment).

9 Anhang

Folgende Lehrerinnen und Lehrer haben an der Fortbildungsveranstaltung des ILF „Stochastik im neuen Lehrplan“ vom 19.9. - 20.9.2006 teilgenommen.

Petra Anlauf-Lenz	Kopernikus Realschule Ludwigshafen
Michael Erlenbach	Realschule Linz
Sandra Fernow	Regionale Schule Bobenheim-Roxheim
Ulrike Fischer	Overberg-Hauptschule Koblenz
Kerstin Heß	Realschule Schifferstadt
Alexa Hirschfeld	Kopernikus Realschule Ludwigshafen
Stefanie Holländer	Regionale Schule Bobenheim-Roxheim
Angelika Hormann	Overberg-Hauptschule Koblenz
Nina Jahnke	Integrierte Gesamtschule Koblenz
Michael Kurth	Göttenbach-Gymnasium Idar-Oberstein
Patrick Loosen	Regionale Schule Boppard
Markus Mallmann	Realschule Emmelshausen
Thomas Ohlhorst	Albert-Einstein-Gymnasium Frankenthal
Jörg Pfeiffer	Overberg-Hauptschule Koblenz
Joachim Schmitt	Hermann-Staudinger-Realschule Konz
Ingrid Sohn	Kopernikus Realschule Ludwigshafen
Wolfgang Steinbring	Grund- und Hauptschule Höhn
Rainer Vogt	Regionale Schule Boppard