EXPLORATIVE DATENANALYSE – LINUS STUHLMANN

Denkt nur den Teil «Deskriptive Statistik» ab, der Teil «Einführung in R» ist nicht enthalten.

EINFÜHRUNG

Statistik ist nützlich, um:

- Sinnvolle Entscheidungen zu treffen (Wissenschaft, Politik & Wirtschaft)
- Publikationen

Aufgaben der Statistik

- Planen (Versuchsplanung)
- Beschreiben (deskriptive Statistik)
- Schliessen (schliessende Statistik)
- Suchen (Explorative Statistik)

Definitionen

- Population
- Census bzw. Vollerhebung
- Stichprobe
- Repräsentativität

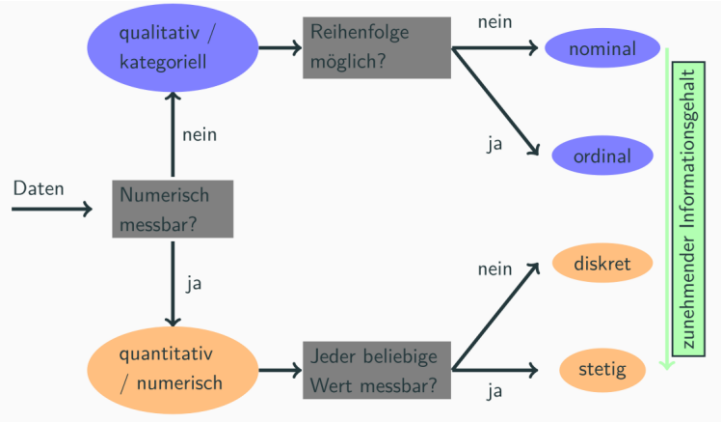
DATENTYPEN

Quantitativ → stetig & diskret

- Stetig → messbar, Fliesskommastelle
- Diskret → abzählbar, ganzzahlig

Qualitativ → ordinal, nominal

- Ordinal → Rangordnung
- Nominal → keine Rangordnung



ABSOLUTE HÄUFIGKEIT VS. RELATIVE HÄUFIGKEIT

- ***Absolute Häufigkeit*** ist die effektive Anzahl von Ereignissen in einer Population
  - $H_n(A)$ → 7
- ***Relative Häufigkeit*** ist die Anzahl von Ereignissen geteilt durch die Anzahl der Population
  - $\frac{H_n(A)}{n}$ → 7/21 = 0.33
  - Immer maximal 1 bzw. 100%
- **Modus**
  - Wert, der am häufigsten vorkommt in einem Datenset

DIAGRAMME

Je nach Datentyp gibt es verschiedene Darstellungsmöglichkeiten

Datentyp	Kategoriell (nominal & ordinal)	Metrisch (diskret & stetig)
Numerisch	Häufigkeitstabelle Modus	Lagemasse Streumasse
Grafisch	Kuchendiagramm Balkendiagramm	Histogramm Boxplot

Kuchendiagramm

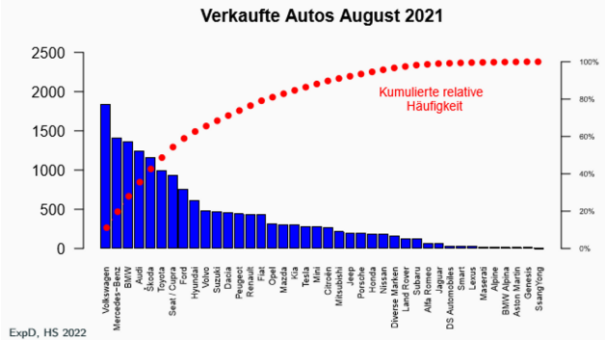
- Sinnvoll um einfache relative Häufigkeiten darzustellen.

Balkendiagramm

- Sinnvoll um absolute oder relative Häufigkeiten darzustellen.

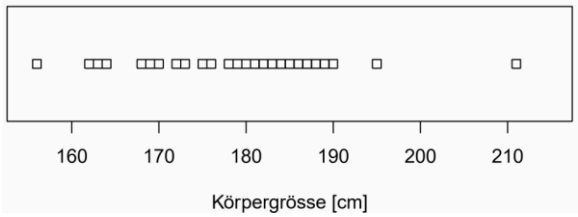
Pareto-Diagramm

- Sinnvoll um absolute und relative Häufigkeiten darzustellen.

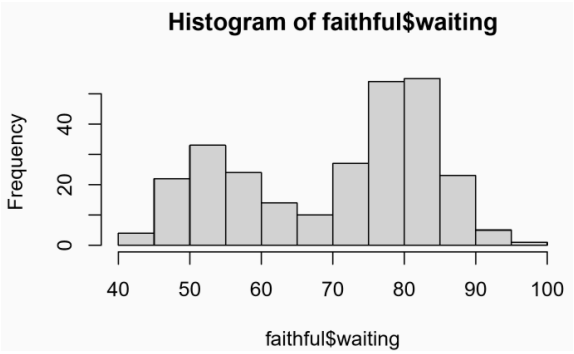


STRIPCHART

Es werden Daten auf einer Line aufgetragen. Jeder Datenpunkt ist durch eine kleine Box dargestellt.



HISTOGRAMM

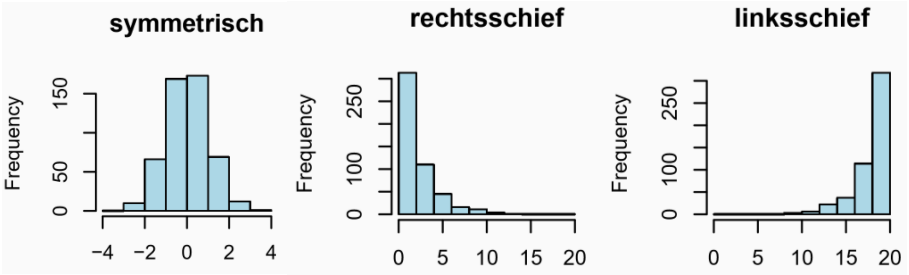


Faustregel für Klassenanzahl

k = $\sqrt{n}$

$k = 1 + \log_2(n)$ → Default in R

VERTEILUNGSFORMEN



Modalität → Anzahl Gipfel der Verteilung

- Unimodal → 1 Gipfel
- Bimodal → 2 Gipfel
- Multimodal → mehrere Gipfel

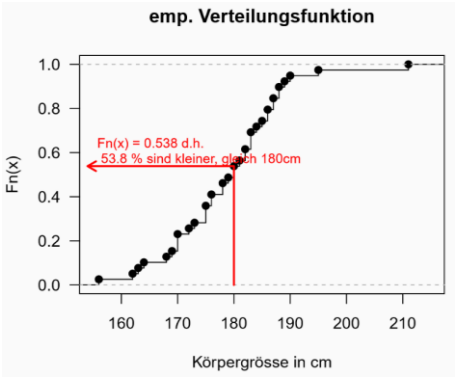
HISTOGRAMM MIT UNGLEICH GROSSEN KLASSEN

- Muss in relativen Häufigkeiten angegeben werden.
- Die aufsummierte Höhe von z.B. 3 Balken muss durch die Anzahl geteilt werden

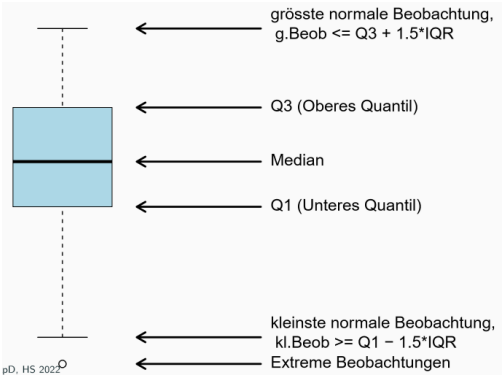
EMPIRISCHE VERTEILUNGSFUNKTION

Die empirische Verteilungsfunktion $F_n(x)$ zeigt die kumulierte relative Häufigkeit. Man kann z.B. durch das Diagramm ablesen, wie gross die Wahrscheinlichkeit ist, dass jemand aus dem Datensatz kleiner als 1,8 Meter ist. Ausrechnen kann man dies wie folgt: Anzahl Beobachtungen n = 100, Anzahl Beobachtungen unterhalb eines Wertes.

- 100/54 → 0.54 also 54% sind unter 180



BOXPLOT

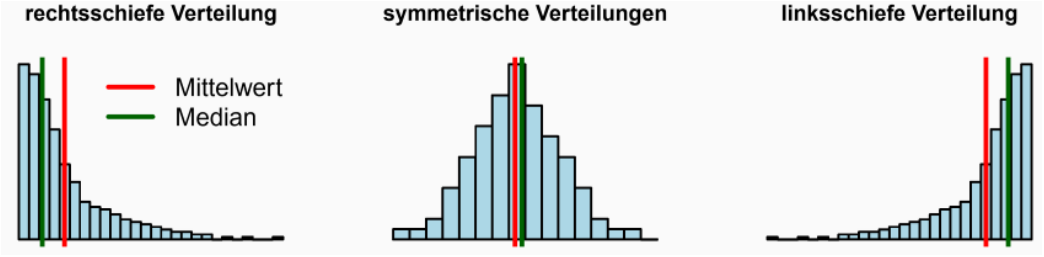


STATISTISCHE KENNZAHLEN

LAGEMASSE

Lagemasse beschreiben um welchen arithmetischen Mittelwert die Daten verteilt sind.

- **Arithmetisches Mittel** → Mittelwert
  - ist nicht robust → Mittelwert kann durch Ausreisser stark verfälscht werden
- **Median** → 50% der Daten liegen unter und 50% der Daten liegen über dem Median
  - bei geraden Beobachtungen ½ (n/2 + (n+1)/2)
- **Modus** → häufigster Wert



- Rechtschief → Mittelwert > Median
- Symmetrisch → Mittelwert ~ Median
- Linksschief → Mittelwert < Median

WEITERE LAGEMASSE

Geometrisches Mittel

$$g = \sqrt[n]{(x_1 * x_2 * x_3 * x_4 * \dots * x_n)} = \sqrt[n]{\prod_{i=1}^{n} x_i}$$

Harmonisches Mittel

$$h = \frac{n}{\frac{1}{x_1} + \dots + \frac{1}{x_n}}$$

Was ist die mittlere Geschwindigkeit, wenn man 100km mit 50km/h und dann 100km mit 100km/h zurücklegt?

$$\frac{2}{\frac{1}{50} + \frac{1}{100}} = 66.67 km/h$$

Quantile

Das Quantil legt fest, wie viele Werte unter und oberhalb eines Punktes liegen. So liegen z.B. unter dem 20. Quantil 20% der Werte unter und 80% der Werte oberhalb. Quantile gibt es [1;100]. Die wichtigsten Quantile sind dabei die folgenden:

- $Q_{25\%}$
- $Q_{50\%}$
- $Q_{75\%}$

Diese werden auch als «Quartile» bezeichnet (Q1, Q2 & Q3)

STREUMASSE

Streuungsmasse geben an, wie breit die Verteilung ist bzw. wie stark die Werte streuen. Die wichtigsten dabei sind:

Varianz

$$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_1 - \bar{x})^2$$

Standardabweichung

$$s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_1 - \bar{x})^2}$$

Die Varianz ist die quadrierte mittlere Abweichung der erhobenen Werte vom Mittelwert. Die Standardabweichung ist die Wurzel davon.

MAD (median absolute deviation)

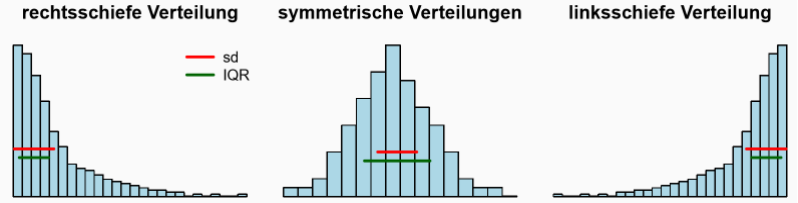
$$MAD_x = 1.4826 * median(|x_1 - \bar{x}|, |x_2 - \bar{x}| \dots, |x_n - \bar{x}|)$$

IQR (inter quartil range)

Ist der Interquartilsabstand, was bedeutet der Abstand zwischen Q1 und Q2. Darin liegen 50% der Werte.

$$IQR = Q_3 - Q_1$$

Wann IQR und wann Standardabweichung?



- Spannweite
    - Ist der Abstand zwischen Maxima und Minima der Daten.

BIVARIATE DARSTELLUNG

In der bivariaten deskriptiven Statistik geht es darum, zwei Variablen gleichzeitig darzustellen.

Wichtig:

Verschiedene Variablentypen benötigen unterschiedliche Handhabung.

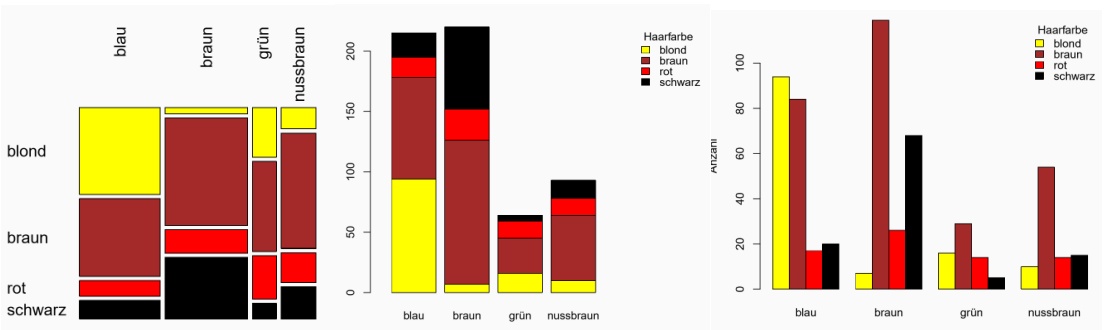
Es gibt die drei Fälle:

1. Kategoriell vs. kategoriell
2. Metrisch vs. kategoriell
3. Metrisch vs. metrisch

ZWEI KATEGORIELLE VARIABLEN

Folgende Diagramme machen Sinn:

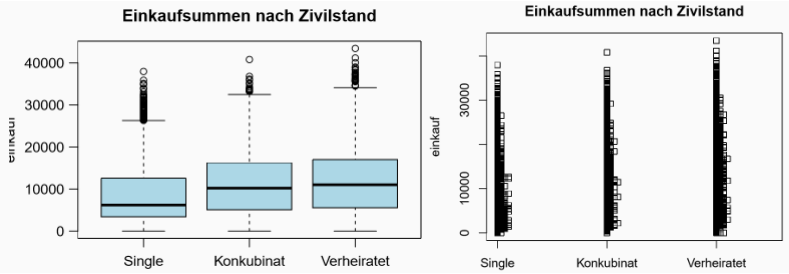
- Absolute Häufigkeiten
- Relative Häufigkeiten
- Mosaikplot (Bild 1)
- Gestapeltes Balkendiagramm (Bild 2)
- Gruppiertes Balkendiagramm (Bild 3)



KATEGORIELL VS. METRISCHE MERKMALE

Folgende Diagramme machen Sinn:

- Boxplot (x = kategoriell, y = metrisch) (Bild 1)
- Stripchart (Bild 2)

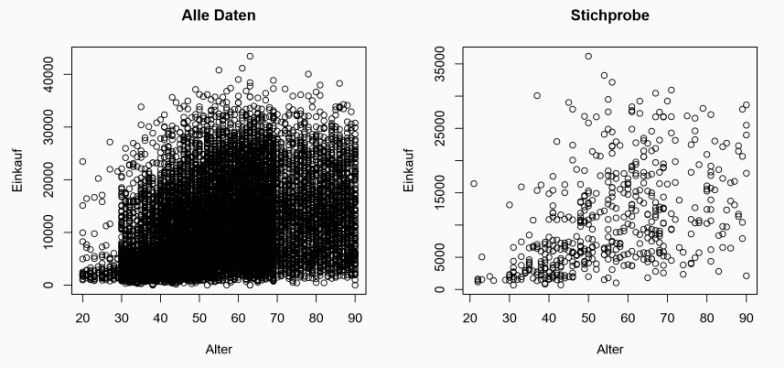


ZWEI METRISCHE MERKMALE

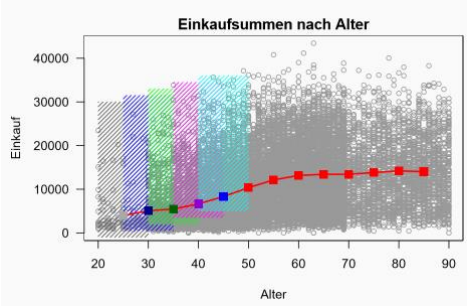
Folgende Diagramme sind sinnvoll:

- Scatterplot

Der Scatterplot ist sinnvoll bei zwei metrischen Merkmalen, allerdings können sich zu grosse Datensets als schwierig zu lesen erweisen, deshalb ist es ratsam eine Stichprobe zu ziehen.



Andernfalls könnte ein gleitender Mittelwert sinnvoll sein, dieser setzt sich aus den Bereichen um diesen Wert zusammen.



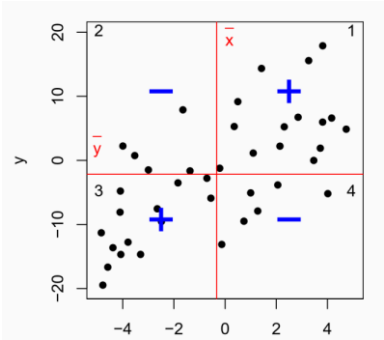
KOVARIANZ

Definition: Die Kovarianz informiert über die Richtung des Zusammenhangs, also ob er positiv oder negativ ist. Er misst allerdings nicht dessen Stärke. Die Kovarianz ist sehr anfällig auf Ausreisser. Die Kovarianz zweier identischen Variablen ist die Varianz.

$$cov(x,y) = \frac{1}{n-1}\sum_{i=1}^{n}[(x_i - \bar{x})(y_i - \bar{y})]$$

Bei einer positiven Kovarianz liegen die Datenpunkte vermehrt in den Feldern 1 & 3. Bei einer negativen Kovarianz liegen die Datenpunkte vermehrt in den Feldern 2 & 3. Wenn kein Zusammenhang besteht liegen gleich viele Punkte in allen vier Feldern.

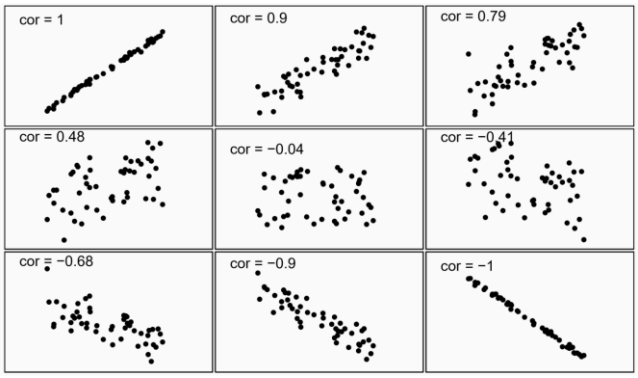
- Bei einem positiven Zusammenhang überwiegen die Datenpunkte mit positiven Vorzeichen Cov > 0
- Bei einem negativen Zusammenhang überwiegen die Datenpunkte mit negativen Vorzeichen Coc < 0



PEARSON-KORRELATION

Die Pearson-Korrelation misst die Stärke des *linearen* Zusammenhanges zwischen zwei Variablen. Sie ist dabei die standardisierte Kovarianz. Sie bewegt sich dabei im Intervall [-1; 1].

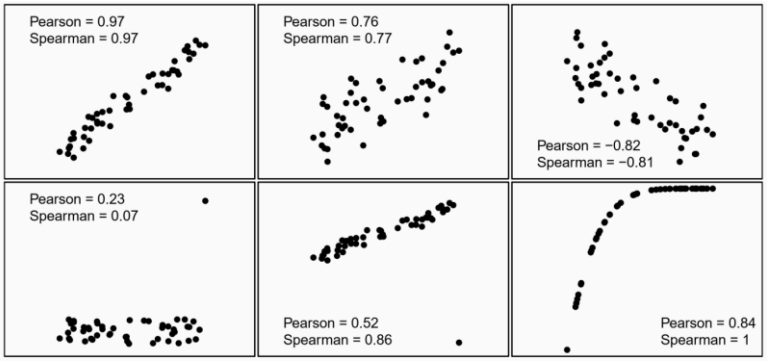
$$r_{xy} = \frac{Cov(x,y)}{s_x s_y}$$



SPEARMAN-KORRELATION

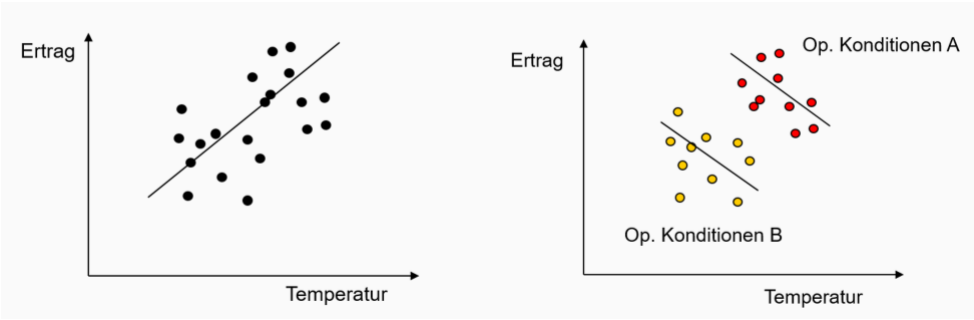
Die Spearman-Korrelation misst die Stärke des *monotonen* Zusammenhangs, d. h. wie nahe die Punkte um eine Kurve liegen, die von einer beliebigen, monotonen (steigen oder fallend) Funktion definiert sind.

Die Werte werden vorher zu Rängen zusammengefasst, was verhindert, dass Ausreisser die Statistik verfälschen. Ausserdem können nicht lineare Zusammenhäng besser ermittelt werden.



INHOMOGENITÄTS-KORRELATION

Kann auftreten, wenn zusätzliche Grössen nicht miteinbezogen werden. Bspw.:



Zusammenhang wäre eigentlich fast Null, wenn die zwei Cluster für sich allein gesehen würden. In diesem Beispiel sind zwei verschiedene Gruppen zusammengenommen worden, was die Statistik verfälscht.

MULTIVARIATE DARSTELLUNGEN

SIMPSON-PARADOXON

An der Universität Barkley wurde eine Klage wegen Diskriminierung bei der Zulassung von Frauen eingereicht. Auslöser für diese Klage war eine Erhebung der Zulassungen an der Universität. Es wurden dabei deutlich mehr Männer als Frauen zugelassen.

Die Klage wurde abgeleht, da sich herausstellte, dass die Frauen sich mehr an Departementen mit einer höheren Durchfallquote beworben und Männer eher an Fakultäten mit einer hohen Aufnahmequote. Durch nicht Berücksichtigung von Gruppengrössen, in diesem Fall die Zulassungsraten an den einzelnen Departementen, kann ein falscher Eindruck entstehen. Für eine saubere Erhebung hätten die Daten der jeweiligen Fakultäten berücksichtigt werden müssen, um zu erkennen, dass die als ungerecht erscheinende Zulassungsrate auf die Anzahl der Bewerbungen pro Fakultät zurückzuführen ist.

- Tritt auf, wenn heterogene Gruppen aggregiert werden
- In diesem Fall beinhaltet die Statistik mindestens 3 Variablen
    - Zielvariable (Zulassung)
    - Beobachtete Variable (Geschlecht)
    - Störvariable (Departement)

Vereinfachtes Beispiel (benachteiligte Männer?)

	Frauen		Männer	
	Bewerberinnen	zugelassen	Bewerber	zugelassen
Fach 1	900	720 (80%)	200	180 (90%)
Fach 2	100	20 (20%)	800	240 (30%)
Summe	**1000**	740	**1000**	420

$$0.74 = 0.9 \cdot 0.8 + 0.1 \cdot 0.2, \qquad 0.42 = 0.2 \cdot 0.9 + 0.8 \cdot 0.3$$

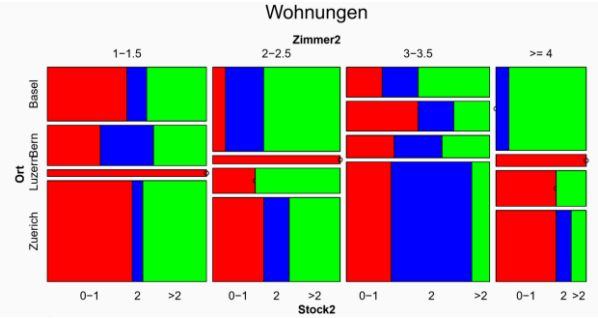
MULTIVARIATE DARSTELLUNGEN

Bei mehr als zwei Variablen hängt die Visualisierungsart mit den Datentypen der Variablen ab. Es muss dabei in folgende Fälle unterschieden werden:

- Mehrer kategorielle Variablen
- 1 quantitative Variable und mehrere kategorielle Variablen
- 2 quantitative und mehrere kategorielle Variablen
- Mehr als 2 quantitative Variablen

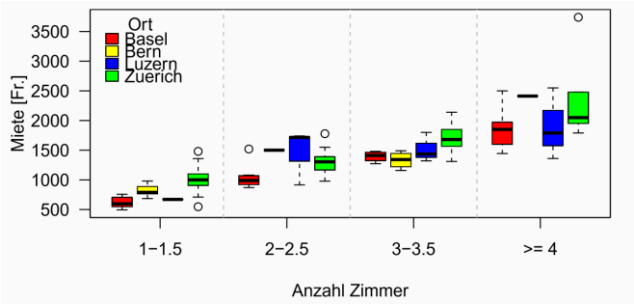
MEHRERE KATEGORIELLE VARIABLEN

Mehrfacher Mosaicplot

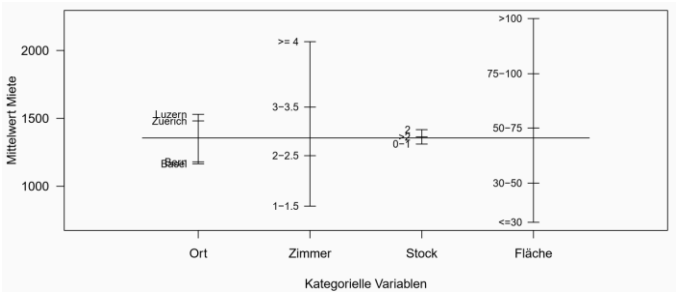


1 QUANTITATIVE UND MEHRERE KATEGORIELLE VARIABLEN

Mehrfacher Boxplot

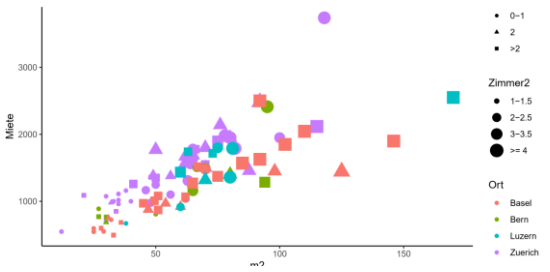


Faktorplot



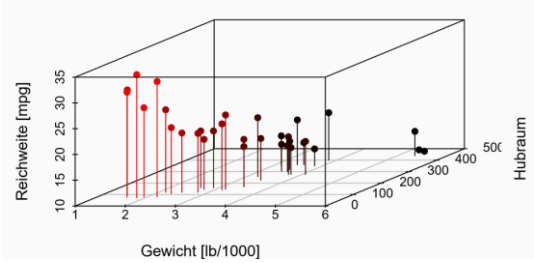
2 QUANTITATIVE UND MEHRERE KATEGORIELLE VARIABLEN

Streudiagramm mit Grössen, Formen und Farben

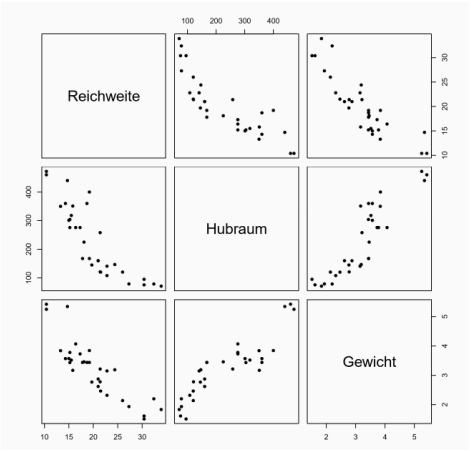


MEHR ALS 2 QUANTITATIVE VARIABLEN

Bei 3 metrischen und einer kategoriellen Variable, sind 3D-Plots möglich



Bei mehr als zwei metrischen Werten eigen sich ausserdem die Streudiagramms-Matrix oder eine Korrelations-Heatmap.



TRANSFORMATIONEN

Daten müssen in manchen Fällen transformiert werden, um für die Analyse eine bessere Form zu haben. Durch Transformationen können sich Kennzahlen und/oder die Verteilung von Variablen verändern.

Gründe für Datentransformation:

- Zusammenfassen von Beobachtungen in Klassen
- Umrechnen von Einheiten
- Informative Darstellung
- Daten standardisieren
- Änderung von unvorteilhaften Formen der Verteilung

NOMINALE DATEN

- Transformation **ohne** Informationsverlust
  - Umbenennung (männlich/weiblich → 0/1)
- Transformation **mit** Informationsverlust
  - Klassen mit wenigen Werten zu «sonstige» zusammenfassen

ORDINALE DATEN

- Transformation **ohne** Informationsverlust
  - Ausprägungen zu Zahlen zusammenfassen
- Transformation **mit** Informationsverlust
  - Zusammenfassen von Klassen

QUANTITATIVE DATEN

Lineare Transformation

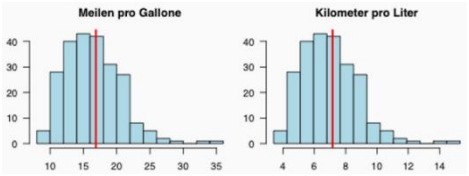
Beispiel lineare Transformation

- Umrechnung von Gallone pro Meile in Liter pro Kilometer
- 1 mpg = (1.6093 km)/ (3.7854 l) = 0.425143 km/l
- Transformationsfunktion: $f(x) = 0.425143 * x$

- Änderung der Masseinheit (von ft in cm)
  - $x \rightarrow f(x) = ax$
- Änderung von Messwerten (Waage mit verstelltem Nullpunkt)
  - $x \rightarrow f(x) = x + b$
- Verschiebung der Skala und des Mess-Nullpunktes (Fahrenheit in Celsius)
  - $x \rightarrow f(x) = ax + b$
  - $C = (F - 32) * \frac{5}{9}$: wobei $32\ F° = Nullpunkt$

Wirkung von linearen Transformationen

Durch eine lineare Transformation verändert sich die Verteilung und Form von Histogrammen nicht, nur die Achsenbeschriftung bzw. die Lagemasse.



- Mittelwert mpg: 16.86 → Mittelwert km/l 7.167
  - $\bar{y} = a * \bar{x} + b$
    - $7.167 = 0.425143 * 16.86$
- Standardabweichung mpg: 4.256 → Standardabweichung km/l: 1.809
  - $sd_y = |a| * sd_x$
    - $1.809 = |0.425143| * 4.251$

Verallgemeinert:

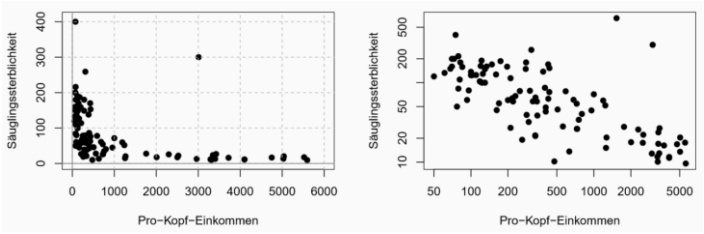
- $Lage_y = a * Lage_x + b$
- $Streuung_y = |a| * Streuung_x$

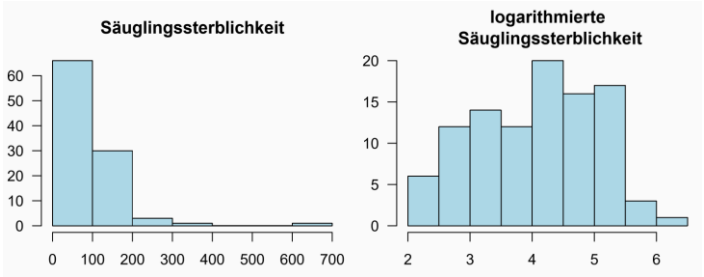
Wobei:

- Lage: arith. Mittel, Median, Quantile, Modus
- Streuung: Standardabweichung, MAD, IQR

Streng monotone Transformation

- $x \to f(x), f\ monoton$
- Beispiel streng monotone Transformation
  - Logarithmieren der Achse enspricht dem Transformieren der x & y Variable mit log()
  - Der Logarithmus zieht die Werte nahe 0 auseinander und schiebt grosse Werte zusammen.
  - Auf der Log-Skala sieht man nun einen linearen Zusammenhang





x	$\log_{10}(x)$
1	0
2	0.3
5	0.7
10	1

*2 +0.3
*2 +0.3

STICHPROBEN

Stichproben vs. Population

Population:

- Parameter direkt bestimmbar
- Keine Informationsunsicherheit

Stichprobe:

- Informationsunsicherheit
- Schätzwerte für Parameter

Eine Stichprobe ist repräsentativ, wenn die Auswahl die typischen Merkmale der Grundgesamtheit getreu ihrer relativen Häufigkeit abbildet.

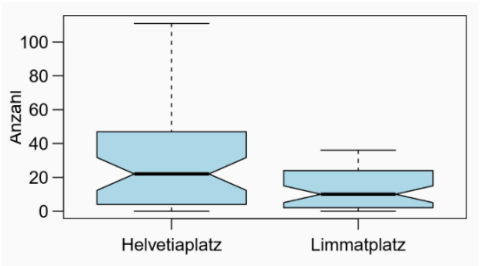
- Je grösser die Stichprobe (n) desto kleiner ist die Standardabweichung
- $\sqrt{n} * s_x$ bleibt hingegen konstant, d.h. die Standardabweichung ist umgekehrt proportional zur Wurzel aus n.
    - Möchte man die Standardabweichung halbieren, muss man die Stichprobengrösse **vervierfachen**

SIGNIFIKANTEN UNTERSCHIED PRÜFEN

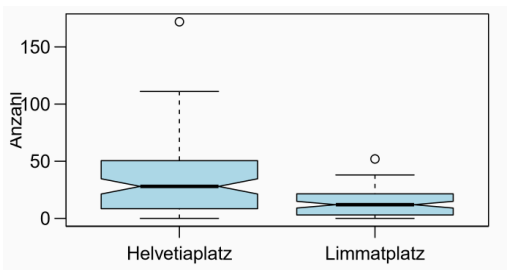
Durch das einzeichenen von Kerben in einem Boxplot ($Median \pm \frac{1.58*IQR}{\sqrt{n}}$) lässt sich der Intervall einer Stichprobe angeben, indem der wahre Median «ziemlich sicher» liegt.

Wenn geprüft werden soll, ob ein signifikanter Unterschied zwischen zwei Datensätzen besteht, kann dies durch die Kerben in Boxplots geprüft werden. Die Kerben werden wie folgt eingezeichnet: $Median \pm \frac{1.58*IQR}{\sqrt{n}}$. Überschneiden sich die Kerben nicht, so liegt ein signifikanter Unterschied zwischen den Gruppen vor, der nicht nur auf Zufall begründet ist. Dadurch kann die Grösse der Stichprobe n ermittelt werden.

links: nicht signifikanter Unterschied



rechts: signifikanter Unterschied



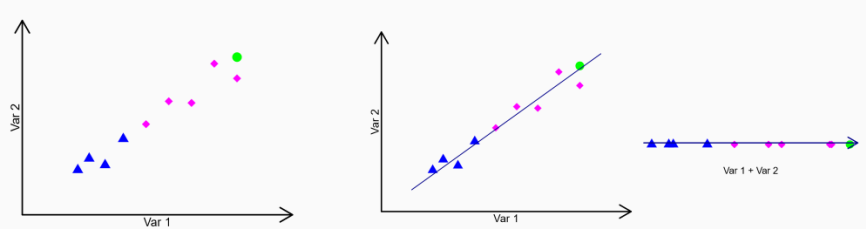
PCA PRINCIPAL COMPONENT ANALYSIS

DIMENSIONSREDUKTION

Hochdimensionale Daten lassen sich nicht mehr so einfach visualisieren. Also ist es wünschenswert, die Dimensionen zu reduzieren. Dies wird auch für Modelle (z.B. im Machine Learning) benötigt. Es kann helfen, irrelevante Daten und Rauschen aus den Daten zu entfernen.

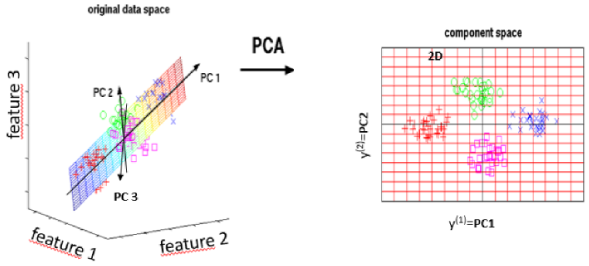
- Die Lösung dieser Probleme ist, Hauptkomponentenanalyse (PCA).

Beispiel für eine Reduktion von 2-dimensionalen Daten auf 1-dimensionale Daten:



Es wird eine Achse so durch die Daten gelegt, um die Daten am besten zu beschreiben. Am besten heisst in diesem Fall:

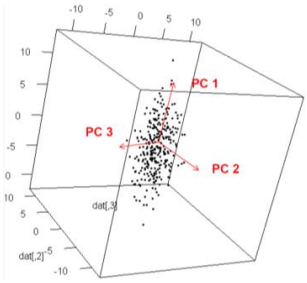
- Drehung der Achse, so dass minimal Informationen verloren gehen
- Dies ist mathematisch äquivalent, die Richtung zu finden die die höchste Varianz aufweisst
- die Projektionslinie mit der minimalen Summe der quadratischen Abstände der Daten



Es wird probiert, möglichst viel Information in den Daten zu erhalten.

Schritte der Dimensionsreduktion:

1. Verschieben des Koordinatensystems in den Schwerpunkt der Daten.
2. Rotation des ursprünglichen Koordinatensystems der Hauptkomponenten, sodass die Varianz entlang der ersten Hauptkomponente am grössten ist. (die quadrierte Abweichung der Daten am kleinsten)
3. Der grösste Teil der restlichen Varianz soll entlang der zweiten Hauptkomponente liegen.



Die Hauptkomponentenanalyse wurde als Rotation definiert, sodass die Varianz entlang der ersten Hauptkomponente am grössten ist. Für diese Rotation wird eine Multiplikation der Datenmatrix X mit einer Rotationsmatrix A durchgeführt.

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \times \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pp} \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1p} \\ z_{21} & z_{22} & \cdots & z_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{np} \end{bmatrix}$$

Daraus folgt eine Linearkombination für die neuen Werte der rotierten Datenmatrix.

$$\begin{aligned} z_{.1} &= a_{11}x_{.1} + a_{21}x_{.2} + \ldots + a_{p1}x_{.p} \\ z_{.2} &= a_{12}x_{.1} + a_{22}x_{.2} + \ldots + a_{p2}x_{.p} \\ &\vdots \\ z_{.p} &= a_{1p}x_{.1} + a_{2p}x_{.2} + \ldots + a_{pp}x_{.p} \end{aligned}$$

Die Rotationsmatrix ist eine Kovarianzmatrix und wird wie folgt berechnet:

$$Cov(X) = \begin{pmatrix} var(x_1) & cov(x_1, x_2) & \ldots & cov(x_1, x_p) \\ cov(x_2, x_1) & var(x_2) & \ldots & cov(x_2, x_p) \\ \vdots & \vdots & & \vdots \\ cov(x_p, x_1) & cov(x_p, x_2) & \ldots & var(x_p) \end{pmatrix}$$

- Wenn einige Variablen korreliert sind, dann haben Elemente ausserhalb der Diagonale Werte ungleich 0. → Das heisst einige Variablen enthalten teilweise redundante Informationen.
- Wenn Elemente ausserhalb der Diagonale Werte = 0 haben, dann sind alle Variablen unkorreliert.
- Bei der PCA werden original korrelierte Variablen in neue unkorrelierte Variablen transformiert → Hauptkomponenten

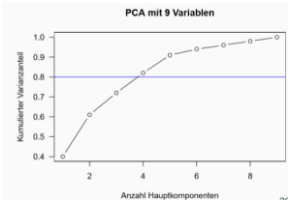
Ergebnis der PCA

Es werden original korrelierte Variablen in neue unkorrelierte Variablen transformiert (Hauptkomponenten).

$$Cov(Z) = \begin{pmatrix} var(z_1) & 0 & \ldots & 0 \\ 0 & var(y_2) & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & var(z_p) \end{pmatrix}$$

Wahl der Anzahl bestehenbleibender Dimensionen

- Anzahl von Dimensionen k sollte so gewählt werden, dass etwa 80% der totalen Varianz durch die Hauptkomponenten erklärt wird.



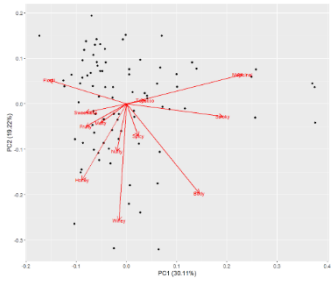
ZUSAMMENFASSUNG PCA

Ziel einer PCA:

Die Hauptkomponenten der Variablen zu finden, die den grössten Einfluss auf die Daten haben, um Dimensionen zu reduzieren. Häufig machen wenige Hauptkomponenten einen Grossteil der Varianz in Daten aus. Es wird eine Drehung im p-dimensionalen Raum durchgeführt, um die Hauptkomponenten so zu konstruieren, dass sie untereinander nicht korreliert sind und den grössten Teil der Totalvarianz ausmachen. Beispielsweise kann es Sinn machen einen 10-dimensionalen Datensatz auf 3 Hauptkomponenten zu reduzieren. Wenn die Daten verschiedene Einheiten aufweisen, sollten sie skaliert werden. Durch die PCA gehen keine Informationen verloren, allerdings werden die Hauptkomponenten nach ihrem Informationsgehalt nach geordnet und darauf gehofft, dass die ersten Hauptkomponenten einen Grossteil der Informationen ausmachen und diejenigen mit wenig Information zu streichen und dadurch die Dimensionen zu reduzieren. Die Faustregel dabei ist 80% der Information sollte erhalten bleiben.

INTERPRETATION BIPLOT

- kurze Pfeile
  - schlecht in der Projektion abgebildete Variablen
- Pfeile die nicht orthogonal zu Komponentenachsen
  - tragen wenig zu Komponenten bei, da in andere Dimensionen zeigend
- Winkel zwischen den Pfeilen
  - zeigen Korrelation untereinander



VORGEHEN BEI EINER PCA

1. Vorliegende Daten plotten und prüfen, ob sie stark streuen. Wenn sie Skalen unterschiedlich, standardisieren.
   a. Boxplot für Skalierung
   b. Pairsplot für Korrelation
2. Prcomp() durchführen mit scale = T, wenn standardisiert werden muss
3. Darstellung der ersten 2 PCs
4. Überprüfung wieviele Hauptkomponenten 80% der Varianz abbilden
5. Biplot zur Interpretation nutzen