

Statistische Grafik

Dipl.-Ing. Hubert Schölnast, BSc
Stand: 15. Jänner 2022

Inhaltsverzeichnis

1	Säulendiagramm	3
1.1	Gruppieretes Säulendiagramm.....	4
1.1.1	Gruppieretes Säulendiagramm, 3D-Version.....	4
1.2	Gestapeltes Säulendiagramm	5
1.2.1	Gestapeltes Säulendiagramm (Anteile).....	5
2	Balkendiagramm	6
2.1	Streifendiagramm	7
3	Histogramm	7
3.1	Histogramme mit konstanter Intervallbreite	8
3.2	Histogramme mit variabler Intervallbreite	8
4	Kreisdiagramm (Tortendiagramm)	9
4.1	explodierte Kreisdiagramme	10
4.2	Halbkreisdiagramm.....	10
4.3	Ringdiagramm	11
4.4	Dreidimensionale Darstellung.....	11
4.5	Sunburst-Diagramm	12
5	Liniendiagramm	13

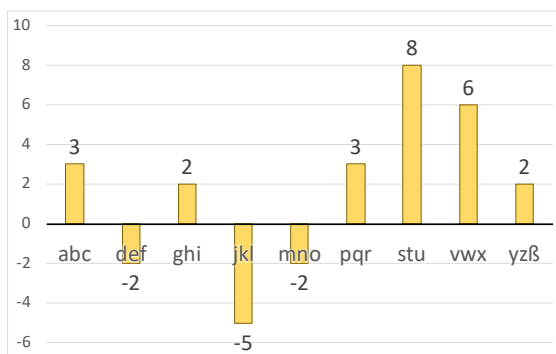
1 Säulendiagramm

Alternative Bezeichnung:

Stabdiagramm (insbesondere, wenn die Säulen sehr schmal sind)

Englisch:

Vertical bar chart



Allgemeines Säulendiagramm



Klimazaun des Wettermuseums Schreufa²

Bestattungsarten: durchschnittliche Kosten



*Kosten OHNE Abholung, Überführung, Kremation, Sarg, Trauerfeier, usw.

nominalskalierte X-Achse¹

Ein Säulendiagramm besteht aus Rechtecken gleicher Breite, die im gleichen Abstand entlang der X-Achse angeordnet sind. Die Höhen der Rechtecke sind proportional zu den verhältnisskalierten Werten, die dargestellt werden. Positive Werte werden nach oben aufgetragen, negative nach unten. Anders als beim Histogramm gibt es beim Säulendiagramm deutlich erkennbare Lücken zwischen den einzelnen Säulen.

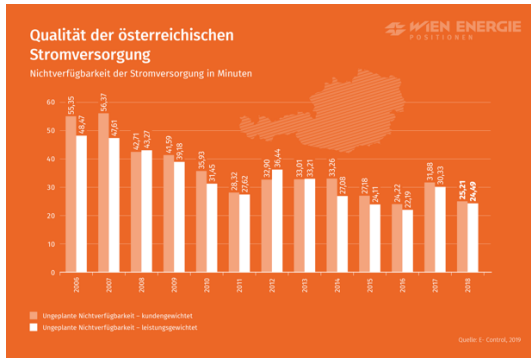
Die Breite der Rechtecke hat keine Bedeutung, auch nicht der Abstand zwischen den Kategorien.

Die verschiedenen Werte entlang der X-Achse entsprechen verschiedenen Kategorien, daher werden bei einem Säulendiagramm entlang der X-Achse nominalskalierte oder ordinalskalierte Werte aufgetragen. Bei Vorliegen einer Ordinalskala mit deutlich mehr als etwa einem Dutzend Werten empfiehlt es sich, statt eines Säulendiagramms ein Linien- oder Flächendiagramm zu verwenden.

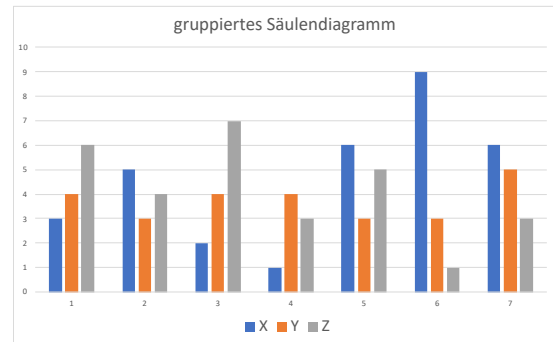
¹ <https://www.herold.at/blog/bestattungskosten/>

² https://de.wikipedia.org/wiki/Datei:Sonderpreis_Wettermuseum_Schreufa_Klimazaun.jpg

1.1 Gruppirtes Säulendiagramm



zwei Datenreihen im selben Diagramm³



allgemeines gruppiertes Säulendiagramm

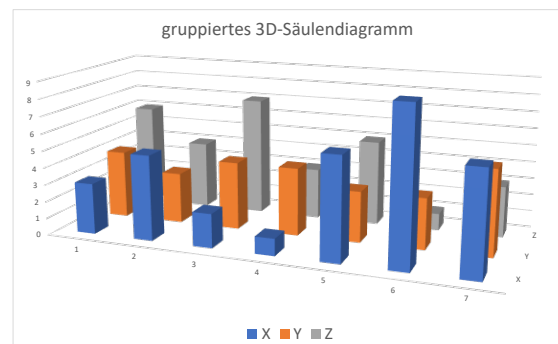
Bei gruppierten Säulendiagrammen werden auf der Y-Achse die Werte mehrerer Datenreihen über denselben X-Werten aufgetragen. Voraussetzung ist, dass die Datenreihen Größen derselben Qualität enthalten. Wenn z.B. in der ersten Datenreihe Geldbeträge dargestellt werden, sollten auch alle weiteren Datenreihen Geldbeträge zeigen.

In sehr seltenen Fällen wird davon abgewichen, dann muss aber einerseits im Begleittext ausdrücklich darauf aufmerksam gemacht werden, und zusätzlich sollten seitlich auch ebenso viele unterschiedliche Skalen angebracht werden, wie es unterschiedliche Qualitäten gibt. Grundsätzlich wird aber davon abgeraten, verschiedene Qualitäten gemeinsam in einem Säulendiagramm darzustellen.

Innerhalb einer Kategorie haben die Säulen nur einen kleinen oder gar keinen Abstand. Sogar Überlappungen sind möglich. Zwischen benachbarten Kategorien ist eine klare Lücke sichtbar.

1.1.1 Gruppirtes Säulendiagramm, 3D-Version

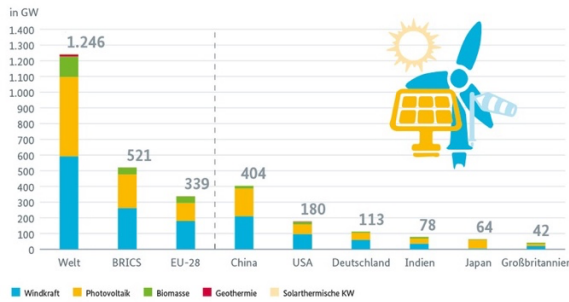
In der 3D-Variante können die Säulen auch hintereinander statt nebeneinander dargestellt werden. Grundsätzlich ist aber zu bedenken, dass es in 3D-Diagrammen immer zu perspektivischen Verzerrungen kommt und das gelegentlich Säulen durch andere Säulen verdeckt werden. Wenn man erreichen will, dass aus der Abbildung genaue Rückschlüsse auf die Daten gezogen werden können, sollte man stattdessen eine zweidimensionale Darstellung wählen.



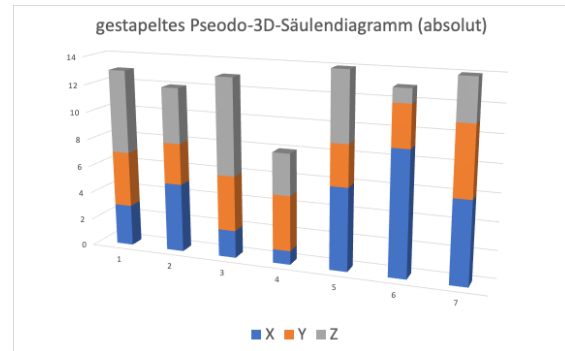
Die Säulen stehen hintereinander

³ Quelle: <https://positionen.wienenergie.at/beitraege/grafik-qualitaet-der-oesterreichischen-stromversorgung/>

1.2 Gestapeltes Säulendiagramm



gestapelte Werte⁴



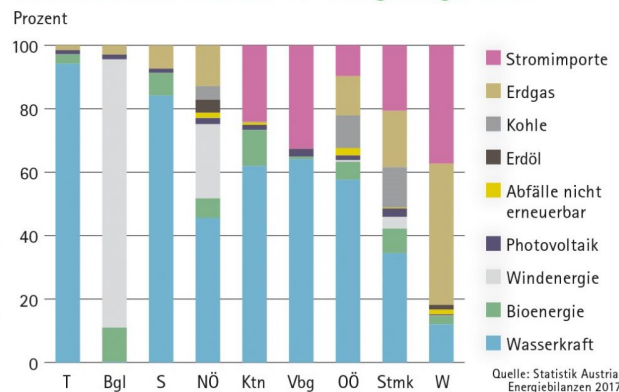
dreidimensionale Variante

Bei gestapelten Säulendiagrammen werden die Werte mehrerer Datenreihen nicht nebeneinander, sondern übereinander aufgetragen. Während bei gruppierten Säulendiagrammen die Forderung, nur Werte gleicher Qualität aufzutragen »nur« eine dringende Empfehlung war (also z.B. Längen nur gemeinsam mit anderen Längen oder Geldbeträge nur gemeinsam mit anderen Geldbeträgen), ist dies bei gestapelten Säulendiagrammen keine Empfehlung mehr, sondern verpflichtend. Denn hier ergeben die Einzelwerte in jeder Säule eine Summe, und das macht nur bei verhältnisskalierten Werten gleicher Qualität einen Sinn.

1.2.1 Gestapeltes Säulendiagramm (Anteile)

Eine Sonderform des gestapelten Säulendiagramms liegt vor, wenn prozentuelle Anteile dargestellt werden. In diesem Fall ergeben alle Stapel dieselbe Höhe, nämlich 100%.

Stromverbrauch Anteile der Energieträger 2017



Alle Säulen sind gleich hoch⁵

⁴ Quelle: <https://www.bmwi.de/Redaktion/DE/Infografiken/Europa/global-installierte-leistung-der-erneuerbaren-energien-2018.html>

⁵ Quelle: <https://www.oekoenergieblog.at/2019/03/fakten-statt-fake-news/>

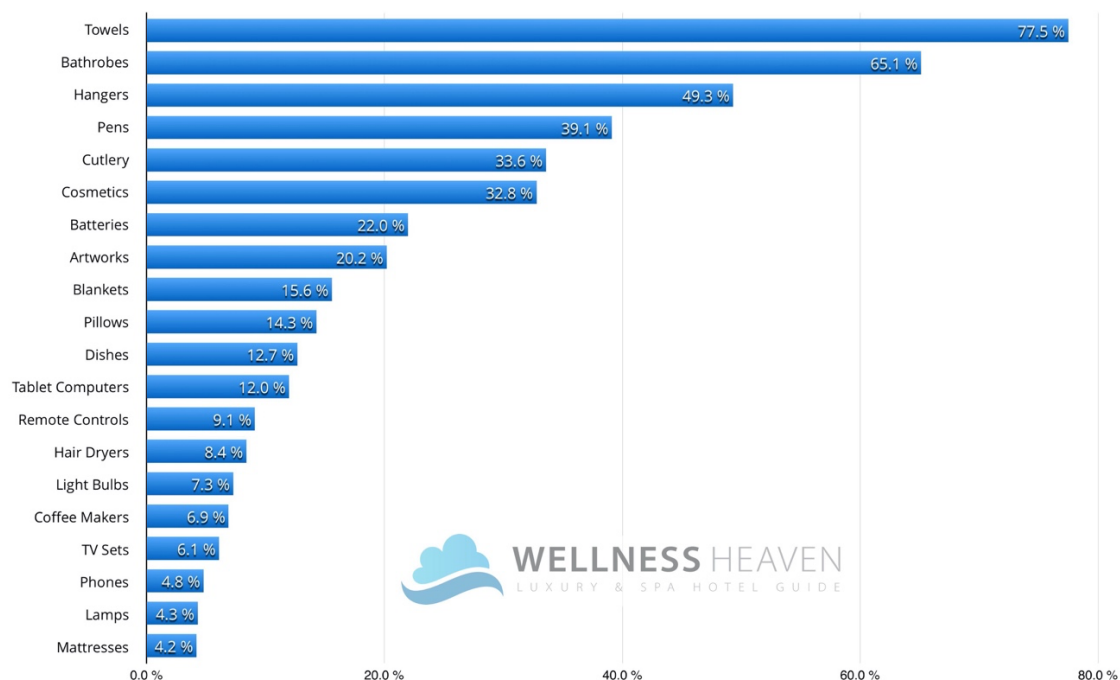
2 Balkendiagramm

Englisch:

Horizontal bar chart

Balkendiagramme sind Säulendiagramme mit vertauschten Achsen. Die Kategorien werden also nicht nebeneinander, sondern von untereinander angeordnet. Bei der Wahl, welchen Diagrammtyp man wählen soll, gilt folgende Faustregel:

- Wenn die Kategorien ordinalskaliert sind (d.h. wenn sie eine innere Ordnung aufweisen, wie z.B. Jahreszahlen, Monatsnamen usw.) sind Säulendiagramme vorzuziehen.
- Wenn die Kategorien nominalskaliert sind (also über keine innere Ordnung verfügen, wie z.B. Namen von Ländern, Sprachen usw.) und es Sinn macht, die Daten nach den Werten pro Kategorie zu sortieren, wie das z.B. in Ranglisten der Fall ist, sollte man bevorzugt über Balkendiagramme nachdenken.
- Auch wenn man deutlich mehr als 10 oder 15 nominalskalierte Kategorien hat, sind Balkendiagramme sinnvoll.
- In allen anderen Fällen kann man frei wählen



Rangliste mit 20 Kategorien (Am häufigsten gestohlene Gegenstände in Hotels)⁶

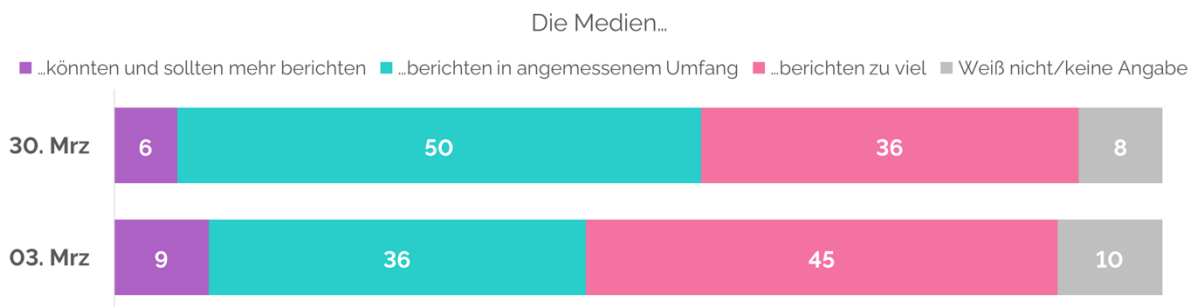
Ansonsten gilt für Balkendiagramme sinngemäß alles, was auch für Säulendiagramme gilt.

⁶ Quelle: <https://www.wellness-heaven.de/wellness/study-theft-in-hotels/>

2.1 Streifendiagramm

Das Streifendiagramm wird auch Blockdiagramm genannt. Es ist eigentlich nichts anderes als ein gestapeltes Anteils-Balkendiagramm, wobei mit »gestapelt« natürlich eine seitliche Aneinanderreihung gemeint ist. Den Namen Streifen- oder Blockdiagramm verwendet man vor allem dann, wenn das Diagramm aus nur einem Balken oder aus sehr wenigen Balken besteht.

Wie bewerten Sie die generelle mediale Berichterstattung über das Coronavirus und dessen Auswirkungen? (in %)



Befragungszeiträume und Stichprobe 03.03.2020 mit n=936 Befragten; 30.03.2020 mit n=6.892 Befragten
Die Ergebnisse sind repräsentativ für die deutsche Bevölkerung ab 18 Jahren.

YouGov

Streifendiagramm⁷

3 Histogramm⁸

Histogramme sind eine Sonderform der Säulendiagramme mit einem ganz bestimmten Zweck: Sie dienen der Darstellung von Häufigkeiten.

Der Wertebereich entlang der X-Achse wird in Klassen unterteilt. Im englischen Sprachgebrauch nennt man diese Klassen *bins* (Behälter, Gefäße, Eimer). In der grundlegenden Idee handelt sich um stetige Werte, deren Wertebereich in Intervalle unterteilt ist. Die Rechtecke, die im Diagramm dargestellt werden, sind so breit wie das Intervall das sie darstellen. Nachdem aber jedes dargestellte Intervall unmittelbar (ohne Lücke) an seine Nachbarintervalle angrenzt, grenzen in einem Histogramm auch die Rechtecke ohne Lücke direkt aneinander.

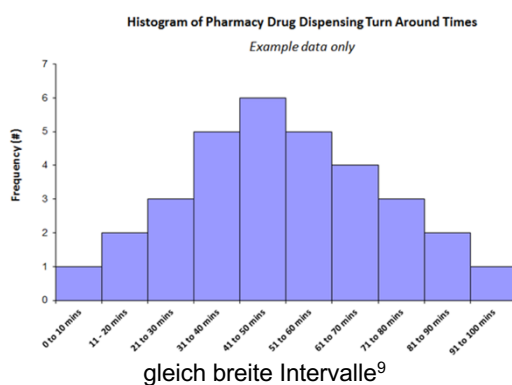
Man findet Histogramme aber auch bei diskreten Werten, insbesondere dann, wenn sich aufeinanderfolgende diskrete Werte zu Klassen zusammenfassen lassen. Es gibt aber auch Histogramme, bei denen jede Klasse aus nur einem diskreten Wert besteht. In diesem Fall wären Lücken zwischen den Rechtecken denkbar, man macht sie aber trotzdem nicht.

⁷ Quelle: <https://yougov.de/news/2020/05/11/covid-19-alle-yougov-ergebnisse-auf-einen-blick/>

⁸ Der Name »Histogramm« wurde 1891 vom englischen Mathematiker Karl Pearson eingeführt. Pearson stellte damit bestimmte historische Daten englischer Premierminister dar und machte aus »historical diagram« das Wort »histogram«. Vier Jahre danach (1895) verwendete er denselben Diagrammtyp zur Darstellung von Häufigkeiten, behielt aber den Namen bei.

Auf der Y-Achse eines Histogramms trägt man immer Zählwerte auf. Die Höhe des jeweiligen Rechtecks gibt also an, wie viele Elemente sich in der jeweiligen Klasse befinden. Der Zählwert kann entweder als absolute natürliche Zahl angegeben werden («Anzahl»), oder auch als prozentueller Anteil an der Gesamtmenge («Häufigkeit»). Die Form des Histogramms ist in beiden Fällen dieselbe.

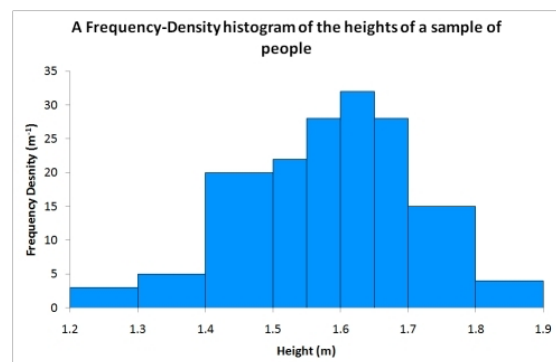
3.1 Histogramme mit konstanter Intervallbreite



Der Vorteil von Intervallen mit konstanter Breite ist, dass die Höhen und Flächen der Rechtecke direkt proportional zueinander sind. Anstatt die Höhe als Zählwert zu interpretieren, kann man genauso gut auch die Fläche verwenden. (Doppelt so große Fläche = doppelt so große Anzahl bzw. Häufigkeit.) Das ist nicht mehr möglich, wenn die Intervalle unterschiedliche Breiten haben.

3.2 Histogramme mit variabler Intervallbreite

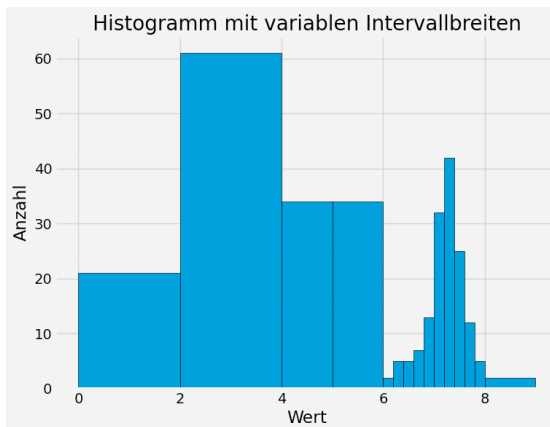
In manchen Fällen häufen sich die Messwerte an bestimmten Stellen, während andere Bereiche nur dünn besiedelt sind. Wenn man die Breiten zu schmal macht, hat man dann Intervalle, in denen es gar keine Werte gibt. Macht man sie zu breit, macht man damit vielleicht an den Häufungsstellen Details unkenntlich. Will man keinen der beide Effekte zulassen, kommt man um variable Intervallbreiten nicht herum.



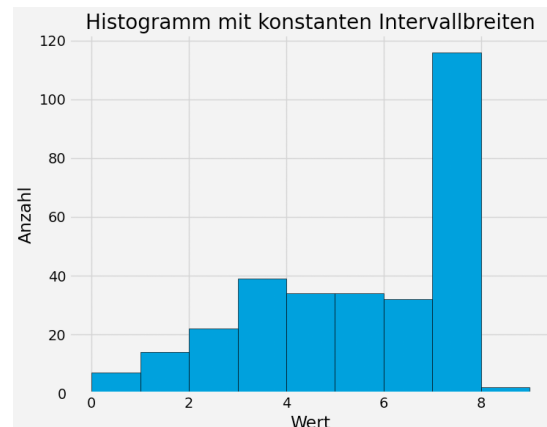
Der Nachteil liegt aber auf der Hand: Die Flächen der Rechtecke haben dann keine Bedeutung mehr. Insbesondere dürfen sie nicht als Anzahl oder Häufigkeit interpretiert werden, das wäre falsch. In einem Histogramm mit variabler Intervallbreite geben nur die Höhen der Rechtecke über die Häufigkeit oder Anzahl Auskunft.

⁹ Quelle: <https://www.cec.health.nsw.gov.au/Quality-Improvement-Academy/quality-improvement-tools/histogram>

¹⁰ Quelle: <https://www.dur.ac.uk/physics/students/labs/skills/data/histograms/>



unterschiedlich breite Intervalle



gleich breite Intervalle

Die beiden hier nebeneinander dargestellten Histogramme beruhen auf denselben Daten. Links sind die Werte zwischen 2 und 4 zu einem Rechteck der Höhe 60 zusammengefasst, rechts gibt es stattdessen 2 Intervalle, die 21 bzw. 39 Einheiten hoch sind. Dafür wurden die beiden rechten Rechtecke zwischen 6 und 8 in der linken Darstellung auf 10 Intervalle verteilt, die natürlich nicht nur schmaler, sondern auch niedriger sind.

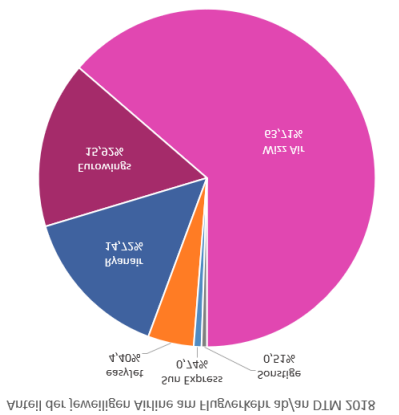
4 Kreisdiagramm (Tortendiagramm)

Alternative Bezeichnung:

Tortendiagramm, Kuchendiagramm (besonders dann, wenn der Kreis als dreidimensionaler Zylinder dargestellt wird)

Englisch:

pie chart, circle chart



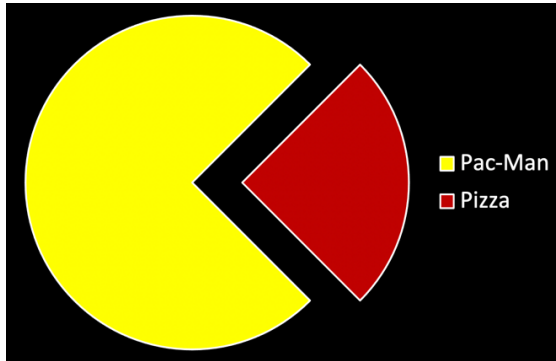
Kreisdiagramm¹¹

Kreisdiagramme sind immer Anteilsdiagramme. Wenn man ein gestapeltes Anteils-Säulendiagramm mit nur einer Säule oder ein Streifendiagramm mit nur einem Balken zeichnen kann, kann man stattdessen auch ein Kreisdiagramm zeichnen.

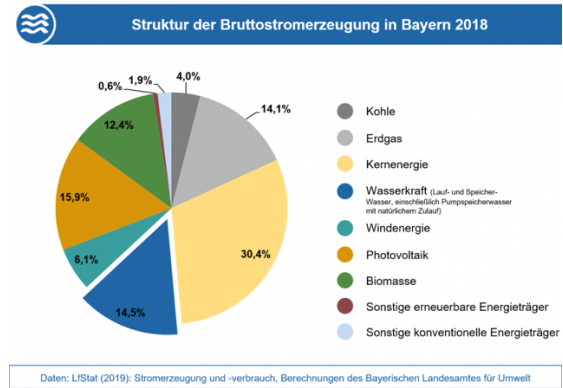
100% entsprechen einem Winkel von 360°, die entsprechenden Anteile sind mit diesem Umrechnungsfaktor in einen Winkel umzurechnen.

¹¹ Quelle: https://de.wikipedia.org/wiki/Datei:Anteile_der_Airlines_am_Flugverkehr_in_DTM_2018.png

4.1 explodierte¹² Kreisdiagramme



Kreisdiagramm mit explodiertem Segment

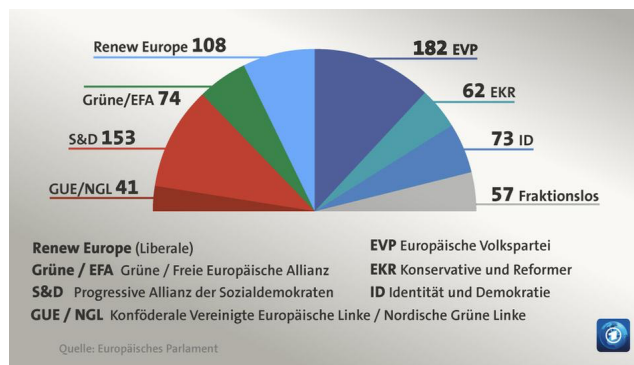


Kreisdiagramm mit explodiertem Segment¹³

Um einzelne Segmente eines Kreisdiagramms hervorzuheben, kann man sie ein wenig aus der Kreismitte herausziehen. Diesen Vorgang nennt man Explosion.

4.2 Halbkreisdiagramm

Anstelle die 100% auf 360° aufzuteilen kann man stattdessen auch einen Winkel von 180° verwenden. Das Ergebnis nennt man Halbkreisdiagramm.



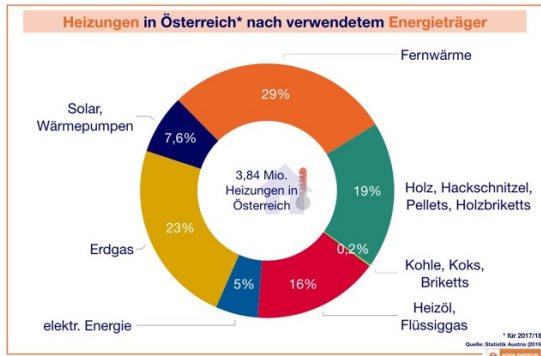
Halbkreisdiagramm¹⁴

¹² In einer Explosionszeichnung werden die Bestandteile eines Konglomerats so dargestellt, als wären sie durch eine Explosion auseinanderbewegt worden.

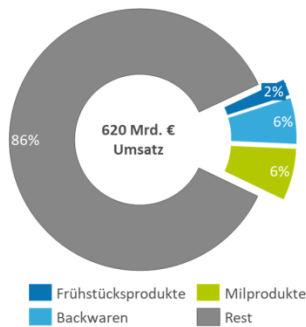
¹³ Quelle: https://www.energieatlas.bayern.de/thema_wasser/daten.html

¹⁴ Quelle: https://www.tagesschau.de/multimedia/bilder/eu-parlament-143~_origin-2da487b9-c3ae-4263-a12e-0dff651300d9.html

4.3 Ringdiagramm



Ringdiagramm¹⁵



Ringdiagramm mit drei explodierten Segmenten¹⁶

NR-Wahl 2019 – Sitze im Nationalrat

Vorläufiges Endergebnis mit Briefwahl (ohne Wahlkarten)



Grafik: © APA



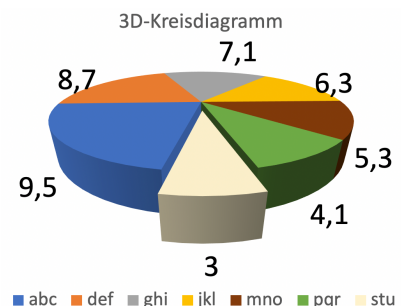
Halbringdiagramm¹⁷

Lässt man die Mitte eines Kreisdiagramms frei, erhält man ein Ringdiagramm. Der Platz in der Mitte eignet sich gut für zusätzliche Beschriftungen.

Je größer der Innendurchmesser im Verhältnis zum Außendurchmesser ist, desto mehr ähnelt der Ring einem gebogenen Streifendiagramm.

4.4 Dreidimensionale Darstellung

Bei Kreisdiagrammen fällt es ohnehin schon schwer, die Größen verschiedener Anteile miteinander zu vergleichen. Wenn das Diagramm dann auch noch als dreidimensionales Objekt dargestellt wird, machen es die perspektivischen Verzerrungen nahezu unmöglich die Größen richtig einzuschätzen.



Perspektivisch verzerrtes Tortendiagramm

¹⁵ Quelle: <https://positionen.wienenergie.at/beitraege/grafik-heizungen-nach-energietraegern/>

¹⁶ Quelle: <https://www.atrimeo.com/news/marktanalyse-anteile-der-fruehstueckprodukte-am-europaeischen-lebensmittelmarkt.html>

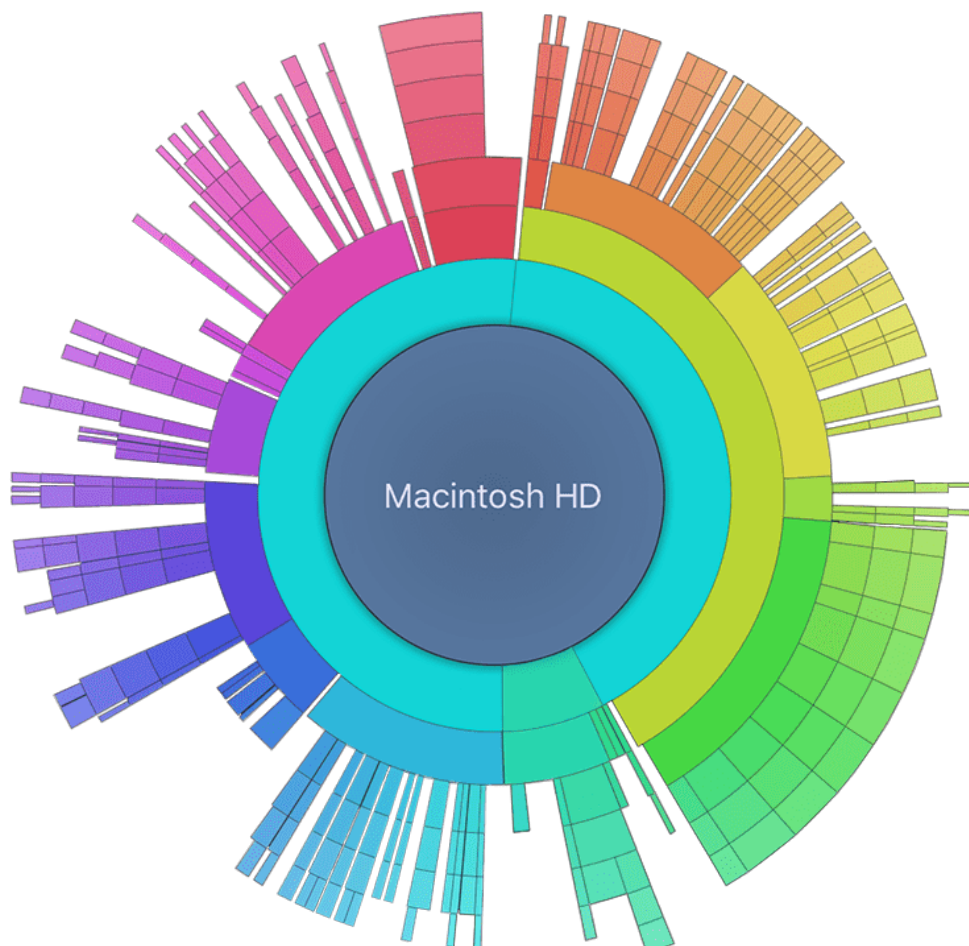
¹⁷ Quelle: <https://www.vienna.at/ergebnis-der-nationalratswahl-2019-so-hat-oesterreich-gewaehlt/6368037>

4.5 Sunburst-Diagramm

Englisch: sunburst chart

Dieser Diagrammtyp besteht aus konzentrischen Ringdiagrammen mit nach außen hin immer stärker fragmentierten Ringen. Dieser Typ wird manchmal zur Darstellung von hierarchischen Größenverhältnissen verwendet, wie man sie z.B. in Dateisystemen findet. Die folgende Abbildung zeigt die Größen verschiedener Verzeichnisse in einem Dateisystem an. Je weiter oben im Verzeichnisbaum ein Verzeichnis liegt, desto weiter innen wird es dargestellt. Große Verzeichnisse erstrecken sich über große Winkel.

Die Lücken entstehen durch Dateien. Im Diagramm werden nur Verzeichnisse dargestellt. Enthält ein Verzeichnis Unterverzeichnisse und Dateien, werden im nächsten Ring nur die Unterverzeichnisse dargestellt.



sunburst chart¹⁸

¹⁸ Quelle: <https://disk-usage-mac-app.peatix.com>

5 Liniendiagramm

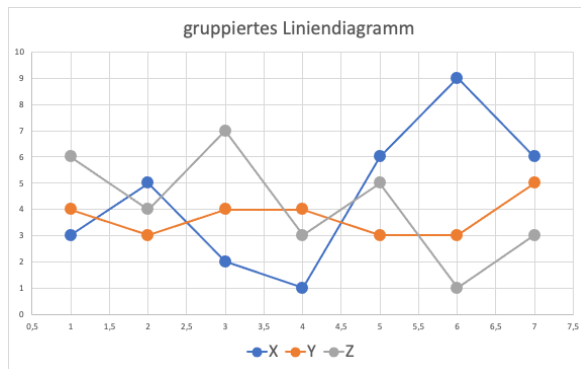
Alternative Bezeichnung:

Streckenzug, Polygonzug (wenn die Linien aus geraden Strecken bestehen)

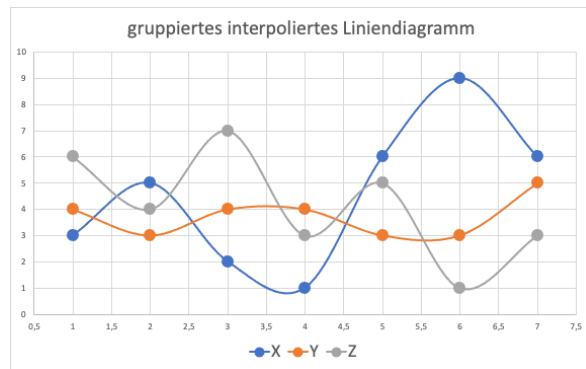
Kurvendiagramm (vor allem dann, wenn gekrümmte Linien gezeichnet werden)

Englisch:

line chart, line plot, line graph, curve plot



Streckenzug



Kurvendiagramm

In den beiden hier dargestellten Diagrammen sind die eigentlichen Datenpunkte durch kleine Kreise dargestellt. Punkte, die zum selben Merkmal gehören, werden in derselben Farbe gezeichnet und durch Linien miteinander verbunden. Die Punkte geben somit dieselbe Information wieder wie die Säulen in einem Säulendiagramm. Ein Liniendiagramm wird aber einem Säulendiagramm vorgezogen, wenn die Werte auf der X-Achse verhältnis- oder intervallskaliert sind, und wenn es sich um ein stetiges Merkmal handelt.

Beispiel:

Die drei Kurven, die in den Abbildungen dargestellt sind, könnten die Temperaturwerte sein, die an drei verschiedenen Orten zu bestimmten Uhrzeiten gemessen wurden. Die Uhrzeiten bilden die X-Achse, die Temperaturen die Y-Achse. Gemessen wurde nur zu jeder vollen Stunde. An den drei Orten herrschte aber auch dazwischen irgendwelche Temperaturen, die aber nicht gemessen wurden.

Die Linien sollen anzeigen, dass es auch zwischen den tatsächlichen Messungen Werte gab, die ungefähr jene Werte haben könnten, die dem Kurvenzug entsprechen. Ob man einen eckigen Streckenzug oder eine geglättete Kurve mit runden Übergängen wählt, hängt einerseits von den Möglichkeiten der Software ab, die man dafür verwendet, andererseits kann es auch von der Art des dargestellten Zusammenhangs abhängen. Sicherlich ist diese Entscheidung auch eine Frage des optischen Designs.

Gegenbeispiel:

Es macht keinen Sinn, die Säulengrafik mit den Bestattungskosten von der ersten Seite als Liniendiagramm darzustellen, weil entlang der X-Achse diskrete nominalskalierte Kategorien

aufgetragen sind. Es gibt keine Zwischenwerte zwischen See- und Feuerbestattung, daher macht es auch keinen Sinn, hier eine Verbindungslinie zu zeichnen.

Anwendungsgebiete

Liniendiagramme kann man immer einsetzen, wenn auf der X-Achse stetige metrische Werte aufgetragen werden. Besonders bietet sich dieser Grafik-Typ für Zeitreihen an. Wenn auf der X-Achse diskrete Kategorien aufzutragen sind, sollte man stattdessen ein Säulendiagramm verwenden.