



Technische Hochschule  
Ingolstadt

# Statistik

Themengebiet: Deskriptive Statistik  
VE 02: Tabellarische und grafische Darstellung von Daten

Wintersemester 2024/25

Dozent: Prof. Dr. Sören Gröttrup

Vorlage von Prof. Dr. Max Krüger

Fakultät Informatik, Technische Hochschule Ingolstadt (THI)

Der vorliegende Foliensatz ist ausschließlich für den persönlichen, vorlesungsinternen Gebrauch im Rahmen der Vorlesungen zur Mathematik und Statistik für anwendungsorientierte Informatik an der Fakultät Informatik der Technischen Hochschule Ingolstadt (THI) bestimmt.

Der Foliensatz wird kontinuierlich korrigiert, aktualisiert und erweitert.

### ***– Urheberrechtlich geschütztes Material –***

Die Weitergabe an Dritte sowie Veröffentlichungen in jeglicher Form (insb. Hochladen ins Internet, Social Media, Videoplattformen, etc.) sind u.a. aus urheberrechtlichen Gründen in keinem Fall gestattet.

## Thema: Tabellarische und grafische Darstellung von Daten

- Absolute und relative Häufigkeiten
- Tabellendarstellungen
- Grafische Darstellungen von Daten
- Problematische Darstellungen

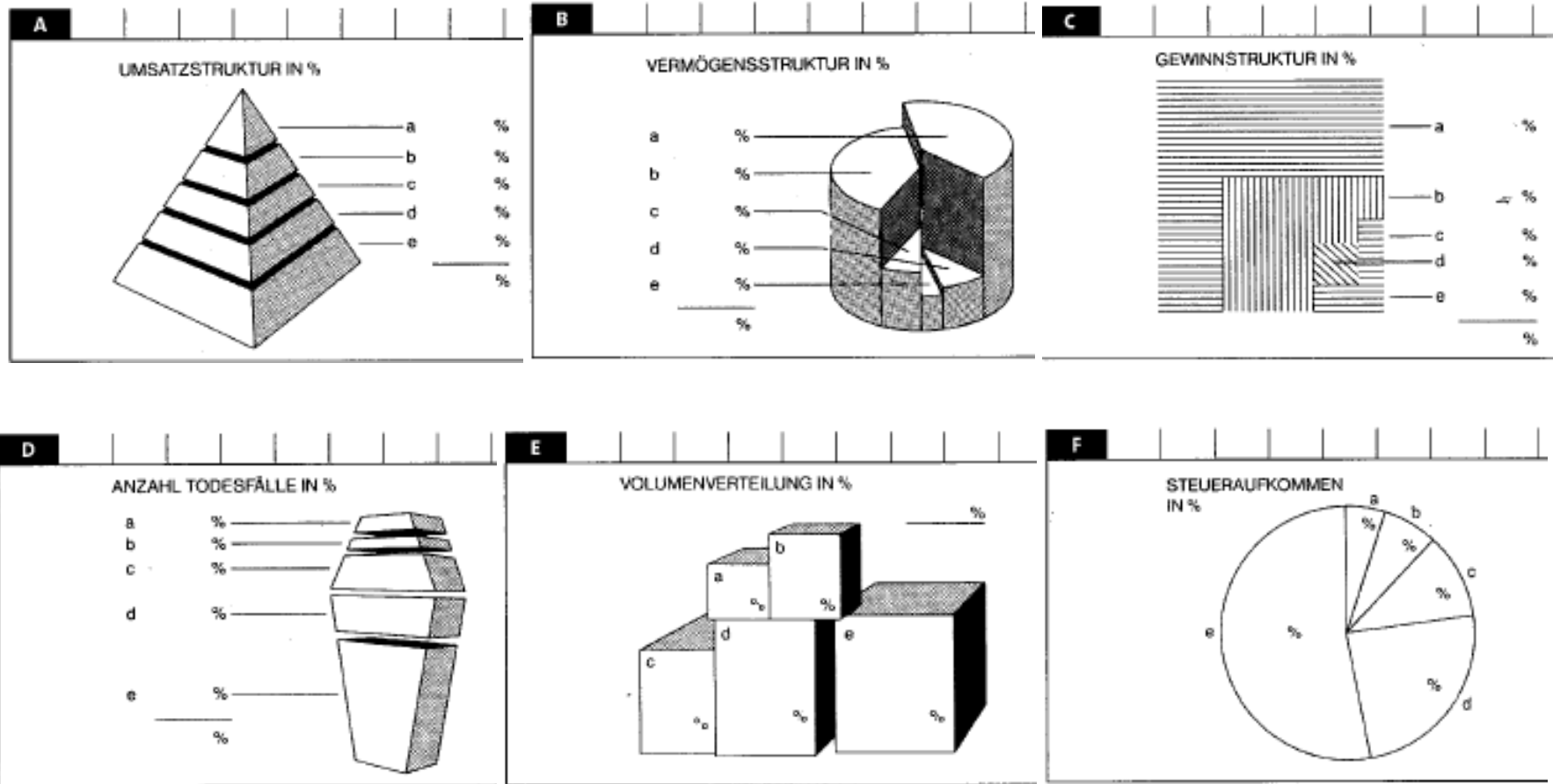
## VE 02: Tabellarische und grafische Darstellung von Daten

- 2.1 Experiment: Ablesen von Prozentdarstellungen
- 2.2 Häufigkeiten und Tabellendarstellungen
- 2.3 Stab-, Säulen-, Kurven-, Balken- und Kreisdiagramme
- 2.4 Strichdiagramme und Histogramme
- 2.5 Piktogramme, Polaritätsdiagramme, Netzdiagramme
- 2.6 Box-Plot-Diagramme und empirische Verteilungsfunktion
- 2.7 Problematische grafischen Darstellungen

# Experiment: Ablesen von Prozentdarstellungen



In jedem dieser sechs Schaubilder setzen Sie bitte nach Ihrem ersten Eindruck ohne weitere Überlegung die Prozentsätze für die einzelnen Anteile ein. (Zeitansatz: 5 min)



Aufgaben- und Bildquelle: Gene Zelazny: *Wie aus Zahlen Bilder werden*. (6. Auflage). Wiesbaden: Gabler Verlag, 2005. (ISBN-10:3-8349-0074-5)

# Ablezen der Prozentdarstellungen



| <b>A</b>              | Ihre Schätzung | Lösung | Delta-Betrag |
|-----------------------|----------------|--------|--------------|
| a                     | 10 %           |        |              |
| b                     | 10 %           |        |              |
| c                     | 15 %           |        |              |
| d                     | 25 %           |        |              |
| e                     | 40 %           |        |              |
| Summe der Delta-Werte |                |        |              |

| <b>B</b>              | Ihre Schätzung | Lösung | Delta-Betrag |
|-----------------------|----------------|--------|--------------|
| a                     |                |        |              |
| b                     |                |        |              |
| c                     |                |        |              |
| d                     |                |        |              |
| e                     |                |        |              |
| Summe der Delta-Werte |                |        |              |

| <b>C</b>              | Ihre Schätzung | Lösung | Delta-Betrag |
|-----------------------|----------------|--------|--------------|
| a                     |                |        |              |
| b                     |                |        |              |
| c                     |                |        |              |
| d                     |                |        |              |
| e                     |                |        |              |
| Summe der Delta-Werte |                |        |              |

| <b>D</b>              | Ihre Schätzung | Lösung | Delta-Betrag |
|-----------------------|----------------|--------|--------------|
| a                     |                |        |              |
| b                     |                |        |              |
| c                     |                |        |              |
| d                     |                |        |              |
| e                     |                |        |              |
| Summe der Delta-Werte |                |        |              |

| <b>E</b>              | Ihre Schätzung | Lösung | Delta-Betrag |
|-----------------------|----------------|--------|--------------|
| a                     |                |        |              |
| b                     |                |        |              |
| c                     |                |        |              |
| d                     |                |        |              |
| e                     |                |        |              |
| Summe der Delta-Werte |                |        |              |

| <b>F</b>              | Ihre Schätzung | Lösung | Delta-Betrag |
|-----------------------|----------------|--------|--------------|
| a                     |                |        |              |
| b                     |                |        |              |
| c                     |                |        |              |
| d                     |                |        |              |
| e                     |                |        |              |
| Summe der Delta-Werte |                |        |              |

# Überprüfung abgelesenen Prozentdarstellungen



|          |                |        |              |
|----------|----------------|--------|--------------|
| <b>A</b> | Ihre Schätzung | Lösung | Delta-Betrag |
|----------|----------------|--------|--------------|

|          |                |        |              |
|----------|----------------|--------|--------------|
| <b>B</b> | Ihre Schätzung | Lösung | Delta-Betrag |
|----------|----------------|--------|--------------|

|          |                |        |              |
|----------|----------------|--------|--------------|
| <b>C</b> | Ihre Schätzung | Lösung | Delta-Betrag |
|----------|----------------|--------|--------------|

*Lösungswerte in der Vorlesung ...*

|                       |      |
|-----------------------|------|
| Summe der Delta-Werte | 26 % |
|-----------------------|------|

|                       |      |
|-----------------------|------|
| Summe der Delta-Werte | 14 % |
|-----------------------|------|

|                       |      |
|-----------------------|------|
| Summe der Delta-Werte | 10 % |
|-----------------------|------|

|          |                |        |              |
|----------|----------------|--------|--------------|
| <b>D</b> | Ihre Schätzung | Lösung | Delta-Betrag |
|----------|----------------|--------|--------------|

|          |                |        |              |
|----------|----------------|--------|--------------|
| <b>E</b> | Ihre Schätzung | Lösung | Delta-Betrag |
|----------|----------------|--------|--------------|

|          |                |        |              |
|----------|----------------|--------|--------------|
| <b>F</b> | Ihre Schätzung | Lösung | Delta-Betrag |
|----------|----------------|--------|--------------|

*Lösungswerte in der Vorlesung ...*

|                       |      |
|-----------------------|------|
| Summe der Delta-Werte | 26 % |
|-----------------------|------|

|                       |      |
|-----------------------|------|
| Summe der Delta-Werte | 10 % |
|-----------------------|------|

|                       |     |
|-----------------------|-----|
| Summe der Delta-Werte | 2 % |
|-----------------------|-----|

## VE 02: Tabellarische und grafische Darstellung von Daten

2.1 Experiment: Ablesen von Prozentdarstellungen

2.2 Häufigkeiten und Tabellendarstellungen

2.3 Stab-, Säulen-, Kurven-, Balken- und Kreisdiagramme

2.4 Strichdiagramme und Histogramme

2.5 Piktogramme, Polaritätsdiagramme, Netzdiagramme

2.6 Box-Plot-Diagramme und empirische Verteilungsfunktion

2.7 Problematische grafischen Darstellungen



## Absolute Häufigkeit einer Merkmalsausprägung

Für eine Stichprobe  $(x_1, \dots, x_n)$  heißt die Anzahl  $n_j$  der Beobachtungswerte mit der Merkmalsausprägung  $u_j$  die **absolute Häufigkeit der Merkmalsausprägung**  $u_j$ .

## Eigenschaft

Sei  $m$  die Anzahl unterschiedlicher Merkmalsausprägungen. Dann gilt

$$\sum_{j=1}^m n_j = n$$

## Relative Häufigkeit einer Merkmalsausprägung

Für eine Stichprobe  $(x_1, \dots, x_n)$  mit der absoluten Häufigkeit  $n_j$  der Merkmalsausprägung  $u_j$  heißt der relative (prozentuale) Anteil

$$h_j = \frac{n_j}{n}$$

der Merkmalsausprägung  $u_j$  am Stichprobengesamtumfang  $n$  die **relative Häufigkeit der Merkmalsausprägung  $u_j$** .

## Eigenschaft

Sei  $m$  die Anzahl unterschiedlicher Merkmalsausprägungen. Dann gilt:

- $0 \leq h_j \leq 1$  für alle  $j \in \{1, \dots, m\}$
- Es ist

$$\sum_{j=1}^m h_j = 1$$

| Zensur | Strichliste | absolute<br>Häufigkeit | relative<br>Häufigkeit | prozentualer<br>Anteil |
|--------|-------------|------------------------|------------------------|------------------------|
| 1      |             | 0                      | 0,0                    | 0 %                    |
| 2      | II          | 2                      | 0,08                   | 8 %                    |
| 3      | III I       | 6                      | 0,24                   | 24 %                   |
| 4      | III IIII    | 9                      | 0,36                   | 36 %                   |
| 5      | III         | 5                      | 0,2                    | 20 %                   |
| 6      | III         | 3                      | 0,12                   | 12 %                   |
|        |             | n=25                   | Summe = 1,00           | Summe = 100%           |

## **Aufgabe:**

Diskutieren Sie die Vor- und Nachteile von Tabellendarstellungen im Vergleich zu grafischen Darstellungen von Daten.

## VE 02: Tabellarische und grafische Darstellung von Daten

2.1 Experiment: Ablesen von Prozentdarstellungen

2.2 Häufigkeiten und Tabellendarstellungen

2.3 Stab-, Säulen-, Kurven-, Balken- und Kreisdiagramme

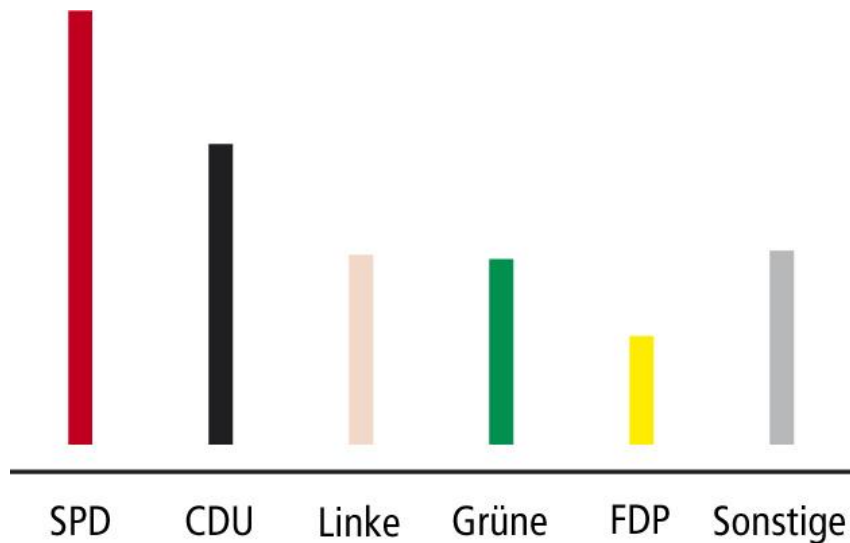
2.4 Strichdiagramme und Histogramme

2.5 Piktogramme, Polaritätsdiagramme, Netzdiagramme

2.6 Box-Plot-Diagramme und empirische Verteilungsfunktion

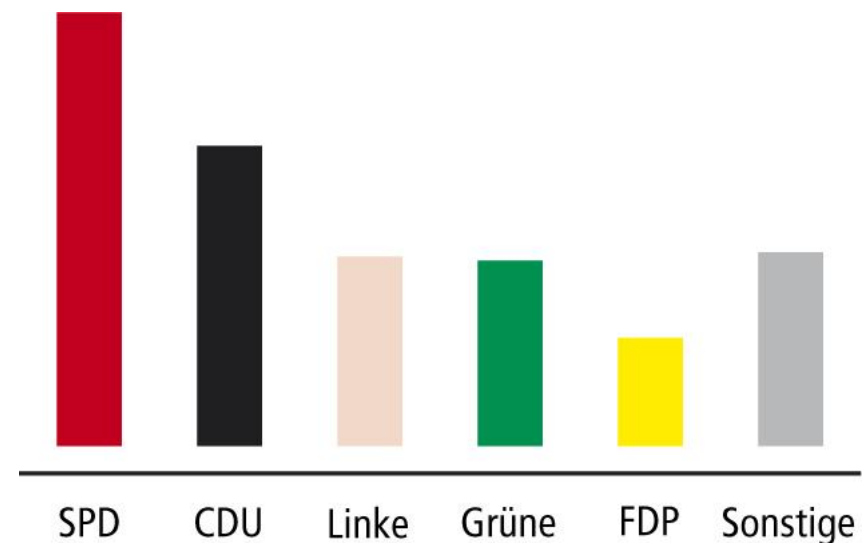
2.7 Problematische grafischen Darstellungen

Stabdiagramm zur  
Berliner Landtagswahl 2006:



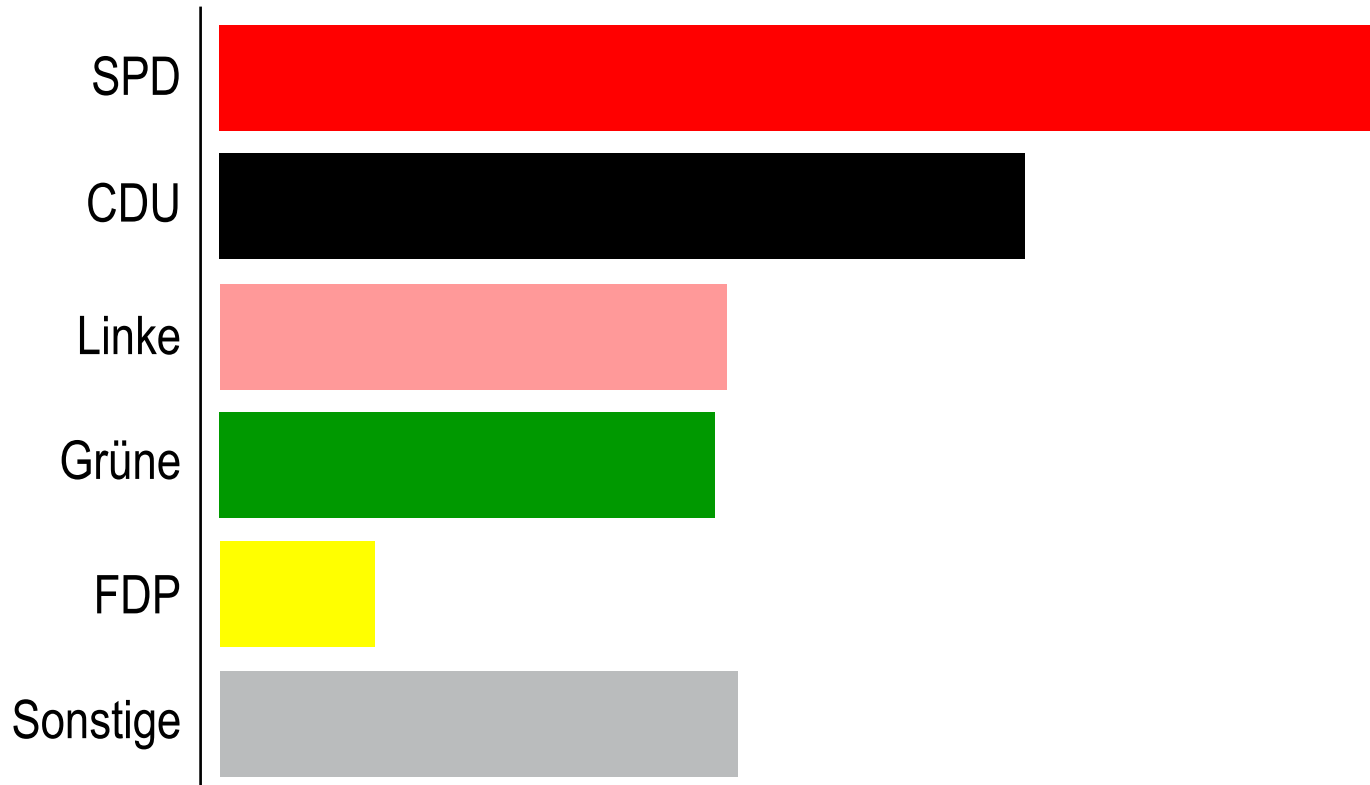
Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008

Säulendiagramm zur  
Berliner Landtagswahl 2006:



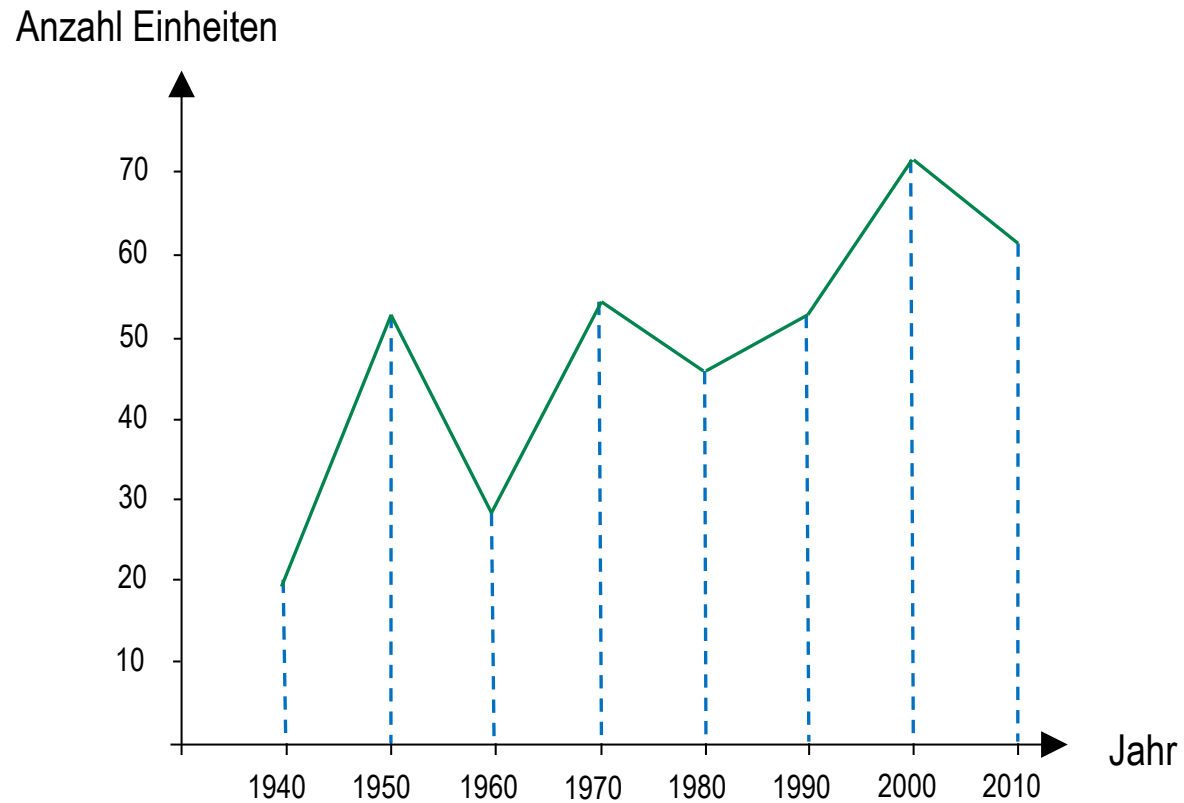
Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008

Balkendiagramm zur Berliner Landtagswahl 2006:

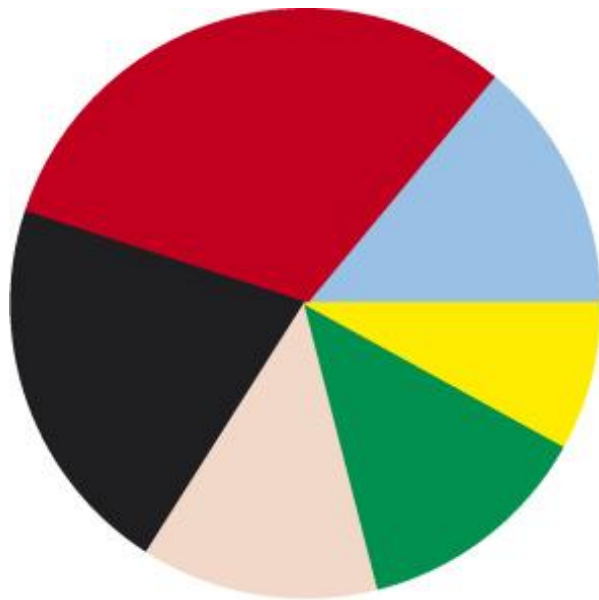


Quelle: in Anlehnung an

Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008







Berlin



Friedrichshain

- Sonstige
- SPD
- CDU
- Linke
- Grüne
- FDP

Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008

Anmerkung: (vgl. [5,p.5])

Beim Vergleich verschiedener Grundgesamtheiten ist die Kreisfläche proportional zur Gesamthäufigkeit  $n$ , d.h. der Radius ist proportional zu  $\sqrt{n}$ .

## VE 02: Tabellarische und grafische Darstellung von Daten

2.1 Experiment: Ablesen von Prozentdarstellungen

2.2 Häufigkeiten und Tabellendarstellungen

2.3 Stab-, Säulen-, Kurven-, Balken- und Kreisdiagramme

2.4 Strichdiagramme und Histogramme

2.5 Piktogramme, Polaritätsdiagramme, Netzdiagramme

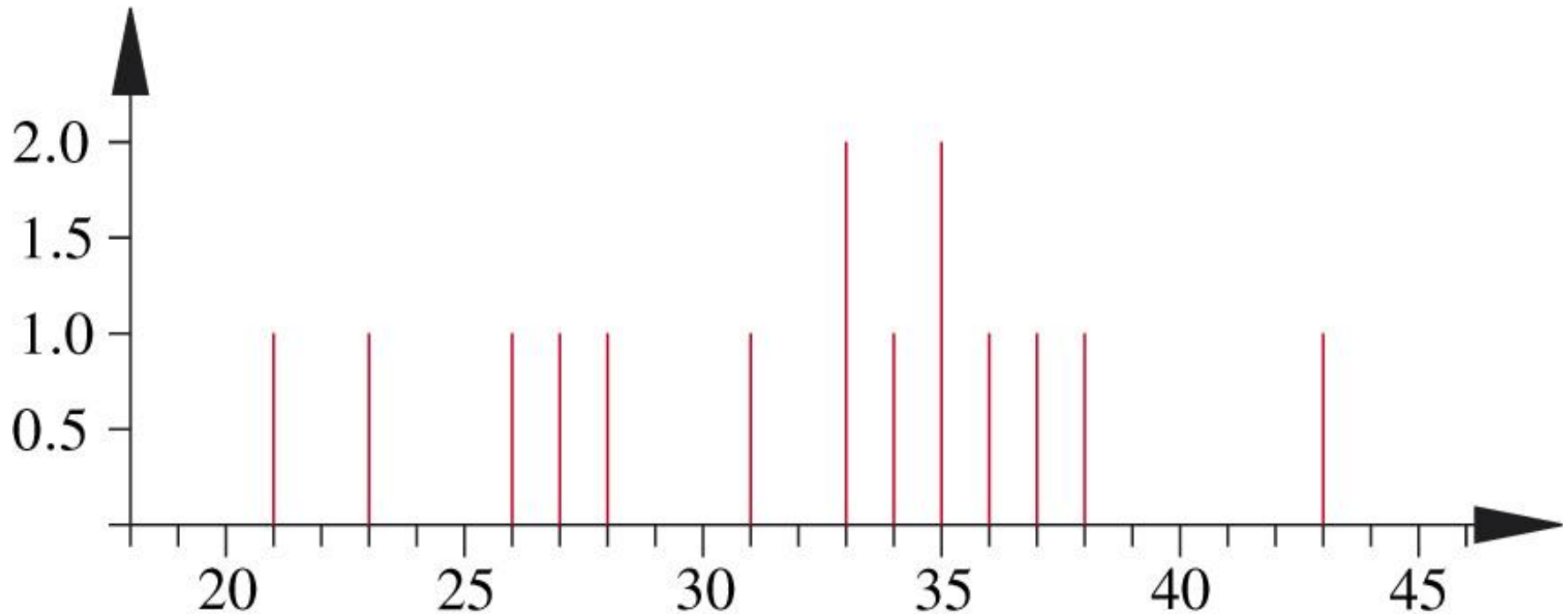
2.6 Box-Plot-Diagramme und empirische Verteilungsfunktion

2.7 Problematische grafischen Darstellungen

# Strichdiagramme bei diskreten Ausprägungen der Werte

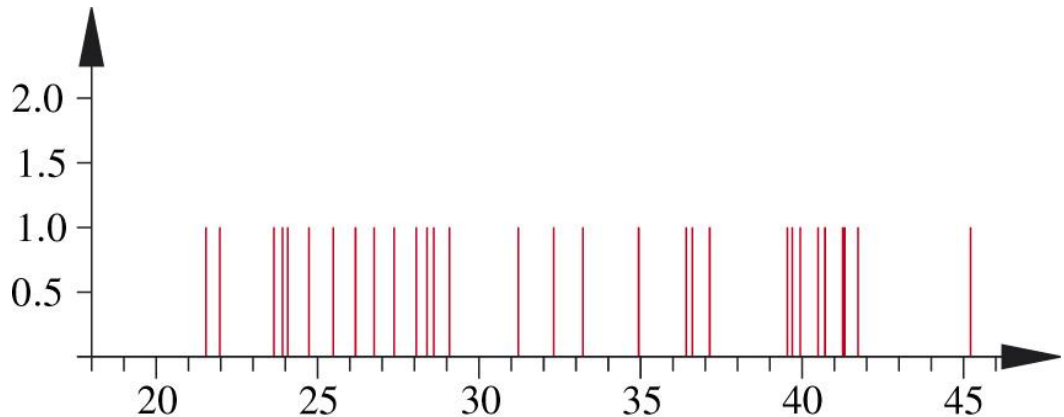


Strichdiagramm mit 15 gemessenen Stickstoff-Monoxid-Werten:



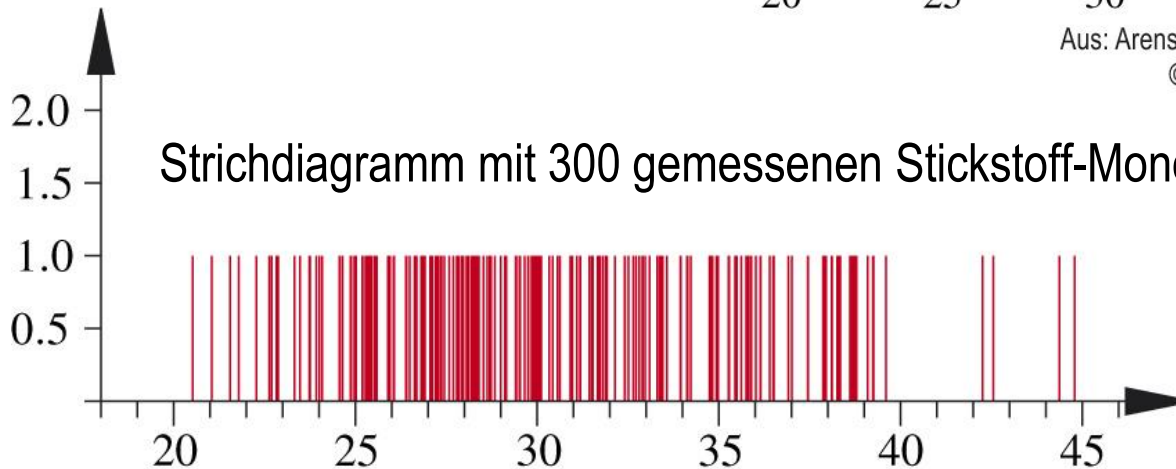
Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008

Strichdiagramm mit 30 gemessenen Stickstoff-Monoxid-Werten:



Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008

Strichdiagramm mit 300 gemessenen Stickstoff-Monoxid-Werten:



Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008

## Histogramm

Ein Histogramm ist die flächentreue Darstellung der Häufigkeitsverteilung eines gruppierten stetigen Merkmals.

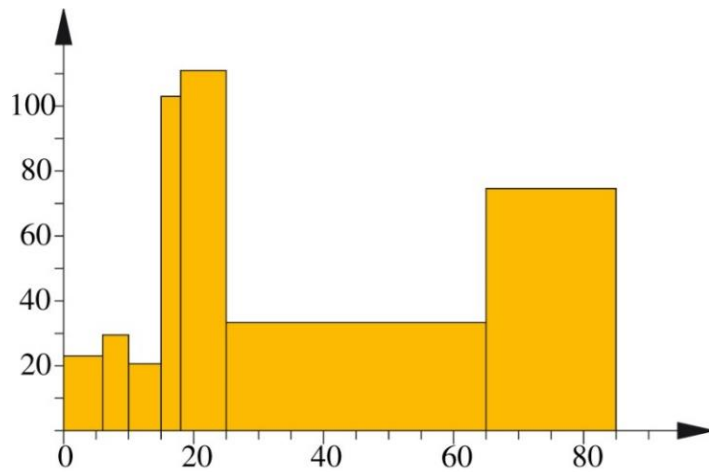
- Die stetigen Merkmalsausprägungen werden in endlich viele Gruppenintervalle (Klassen) zusammengefasst.
- Die Fläche der Säule über einer Klasse entspricht der (absoluten bzw. relativen) Klassenbesetzung:

$$a_j \cdot b_j = n_j \quad \text{bzw.} \quad a_j \cdot b_j = h_j,$$

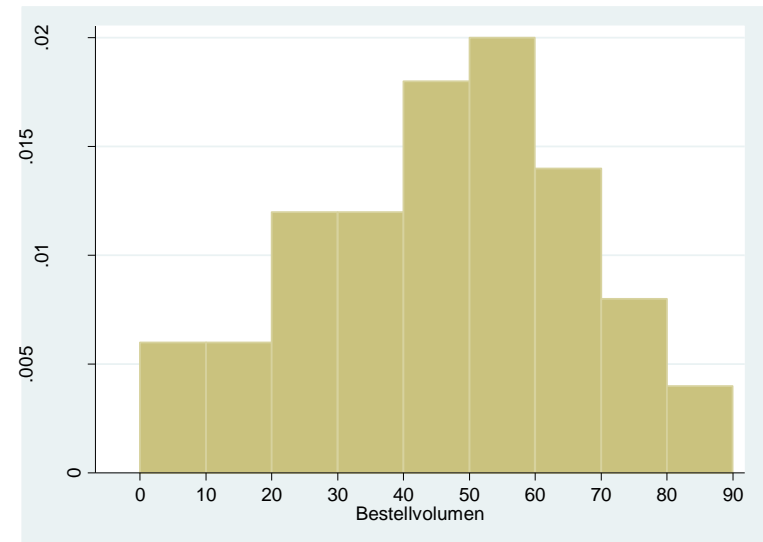
wobei

- $n_j$  = Anzahl Elemente der  $j$ -ten Klasse bzw.  $h_j$  = Anzahl Elemente der  $j$ -ten Klasse
- $b_j$  = Breite des  $j$ -ten Klassenintervalls
- $a_j$  = Höhe des  $j$ -ten Balkens (Häufigkeitsdichte).

## Beispiele



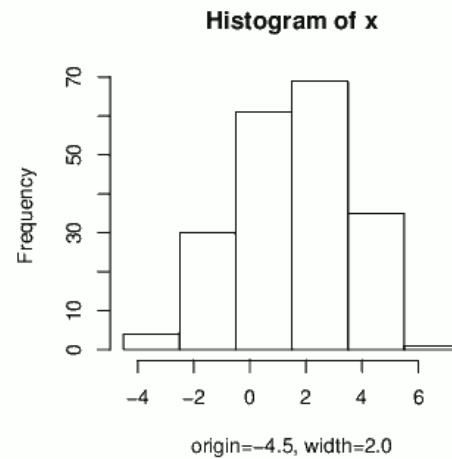
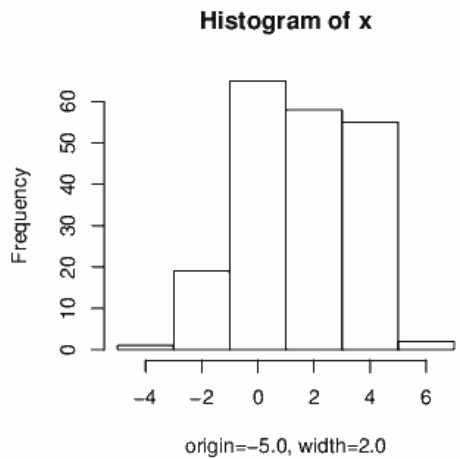
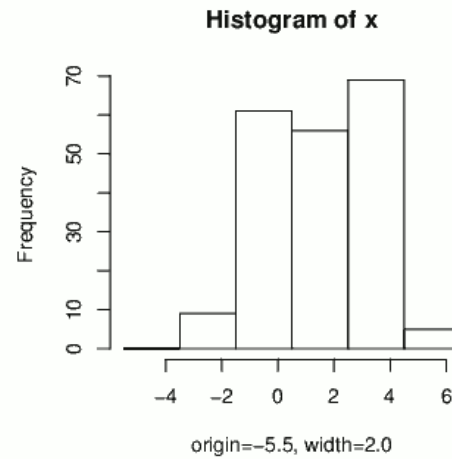
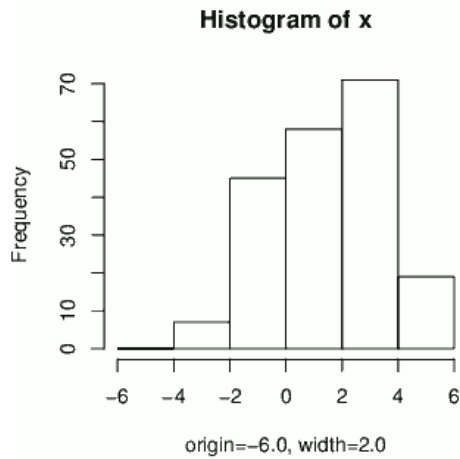
Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008



## Anmerkungen

- Bei **gleichbreiten** Klassen sind die Höhen der Balken zu den Klassenhäufigkeiten proportional
- Bei **ungleichen** Klassenbreiten sind die Höhen der Balken nicht mehr zu den Klassenhäufigkeiten proportional
- Konsequenz ungleicher Klassenbreiten für Balkenhöhe: Sind zwei Klassen gleich stark besetzt aber eine Klasse doppelt so breit, wie die andere, dann muss die Höhe der Säule der breiteren Klasse die Hälfte der Höhe der schmaleren Klasse betragen
- Ränder können nicht offen sein. Es muss ein minimaler und maximaler Wert gewählt werden.

# Randpunktwahl bei Histogrammen



Quelle: o.V., <http://upload.wikimedia.org/wikipedia/commons/thumb/8/8c/Histogram4.png/300px-Histogram4.png> (abgerufen: 08.10.2013)



## VE 02: Tabellarische und grafische Darstellung von Daten

2.1 Experiment: Ablesen von Prozentdarstellungen

2.2 Häufigkeiten und Tabellendarstellungen

2.3 Stab-, Säulen-, Kurven-, Balken- und Kreisdiagramme

2.4 Strichdiagramme und Histogramme

2.5 Piktogramme, Polaritätsdiagramme, Netzdiagramme

2.6 Box-Plot-Diagramme und empirische Verteilungsfunktion

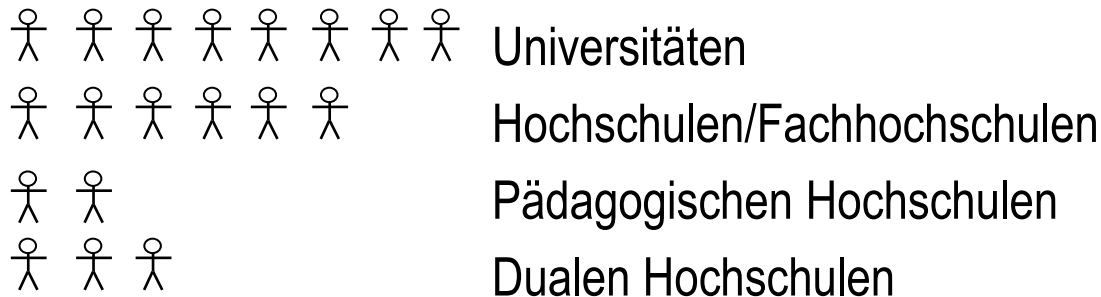
2.7 Problematische grafischen Darstellungen

## Piktogramme

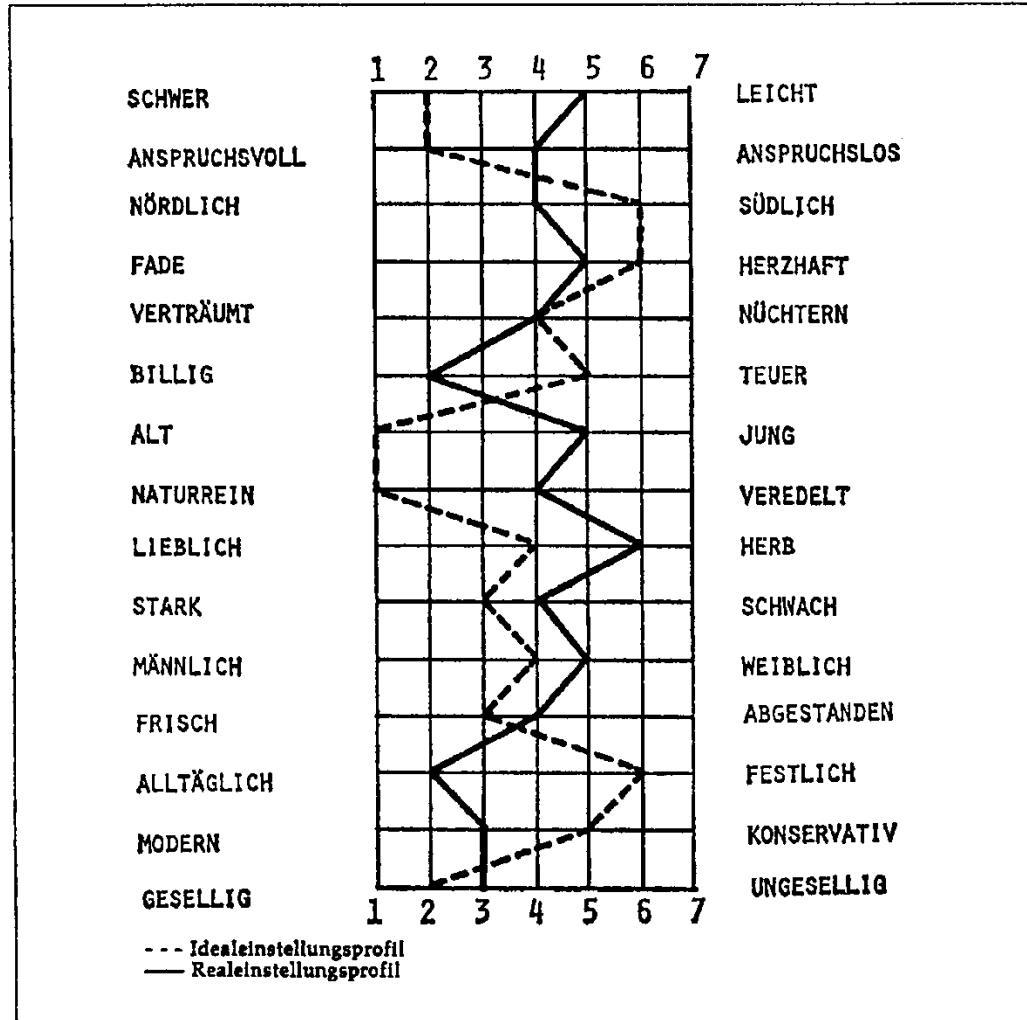
Diagramme mit einer Anordnung von gleichartigen Symbolen, wobei jedes Symbol für eine festgelegte Anzahl von Einheiten steht.

### Beispiel:

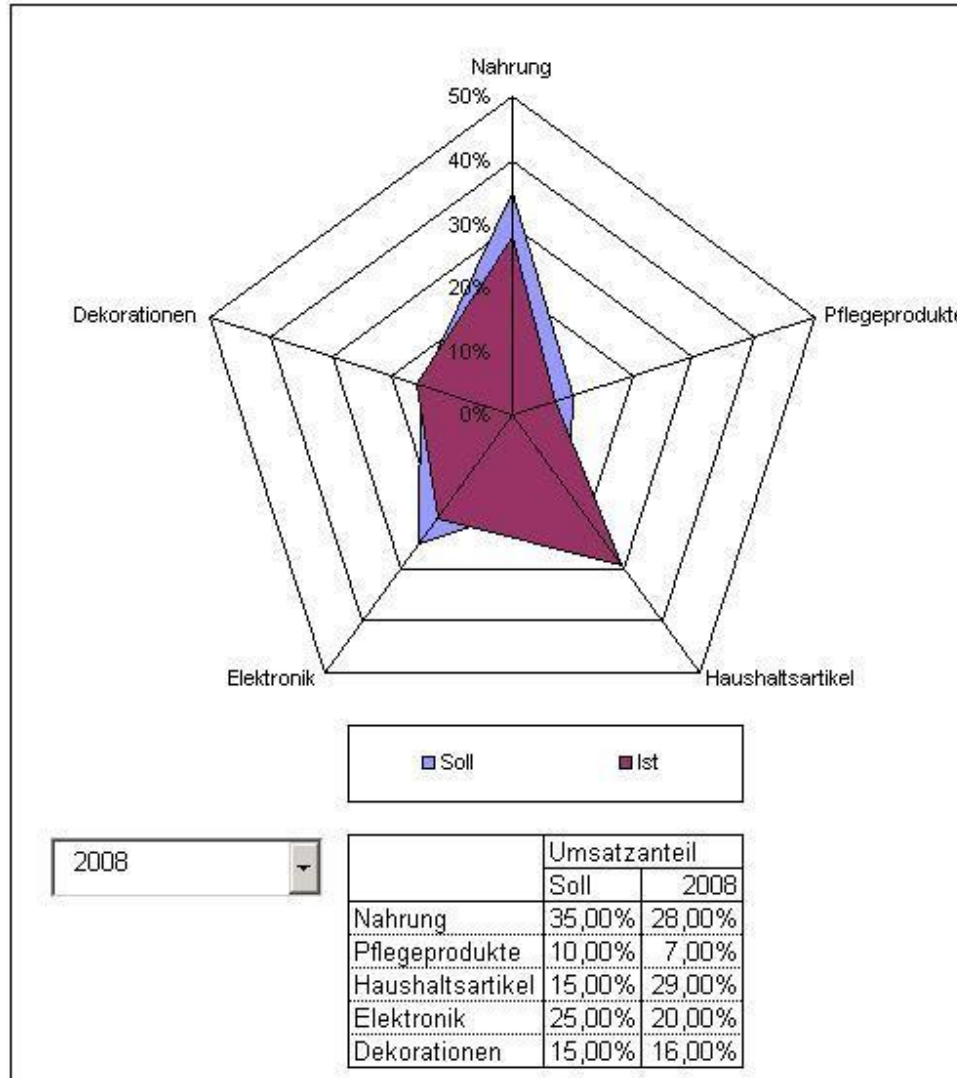
Studierende an ...



 = 10.000 Studierende



Bildquelle: Heribert Meffert: *Marketing* (7. Auflage). Wiesbaden: Gabler Verlag, 1986. (ISBN-10:3-409-69014-X)



Quelle: reimus.NET GmbH, Alexander Wildt, <http://www.controllingportal.de/Fachinfo/Excel-Tipps/Das-Erstellen-von-Netzdiagrammen.html> (abgerufen: 12.10.2013)

## VE 02: Tabellarische und grafische Darstellung von Daten

2.1 Experiment: Ablesen von Prozentdarstellungen

2.2 Häufigkeiten und Tabellendarstellungen

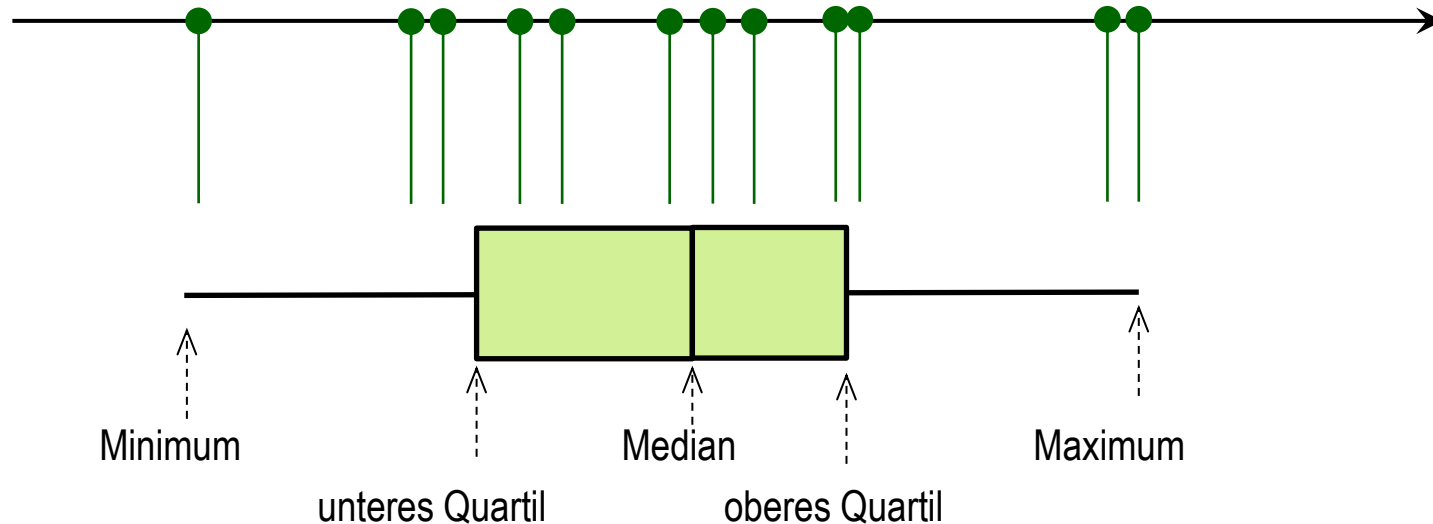
2.3 Stab-, Säulen-, Kurven-, Balken- und Kreisdiagramme

2.4 Strichdiagramme und Histogramme

2.5 Piktogramme, Polaritätsdiagramme, Netzdiagramme

2.6 Box-Plot-Diagramme und empirische Verteilungsfunktion

2.7 Problematische grafischen Darstellungen



Erläuterungen: (vgl. [1,pp.44-46])

Minimum: unterer Extremwert

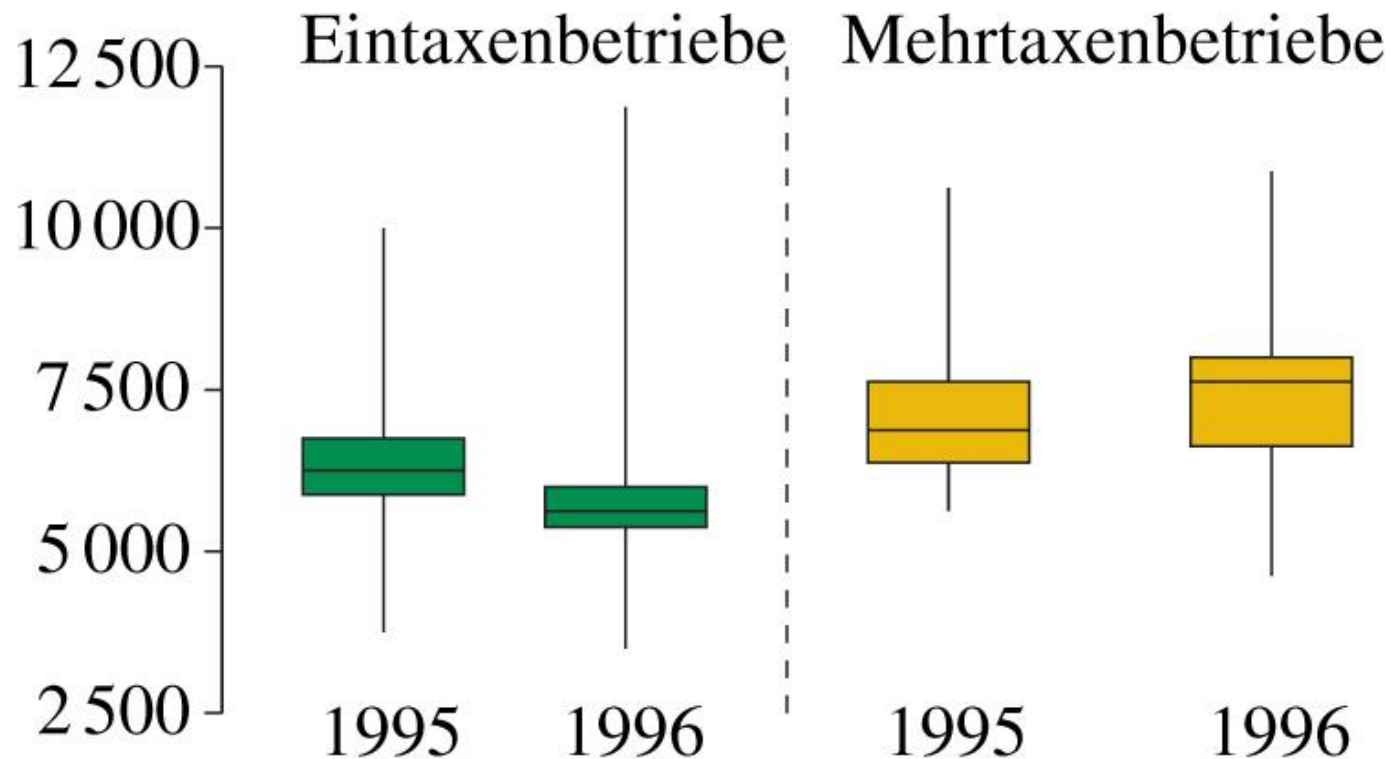
Maximum: oberer Extremwert

Median: Zentralwert (= mittlerer Wert)

unteres Quartil: Zentralwert der unteren Hälfte

Oberes Quartil: Zentralwert der oberen Hälfte

Monatlicher Umsatz pro Taxi in einer deutschen Großstadt:



Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008

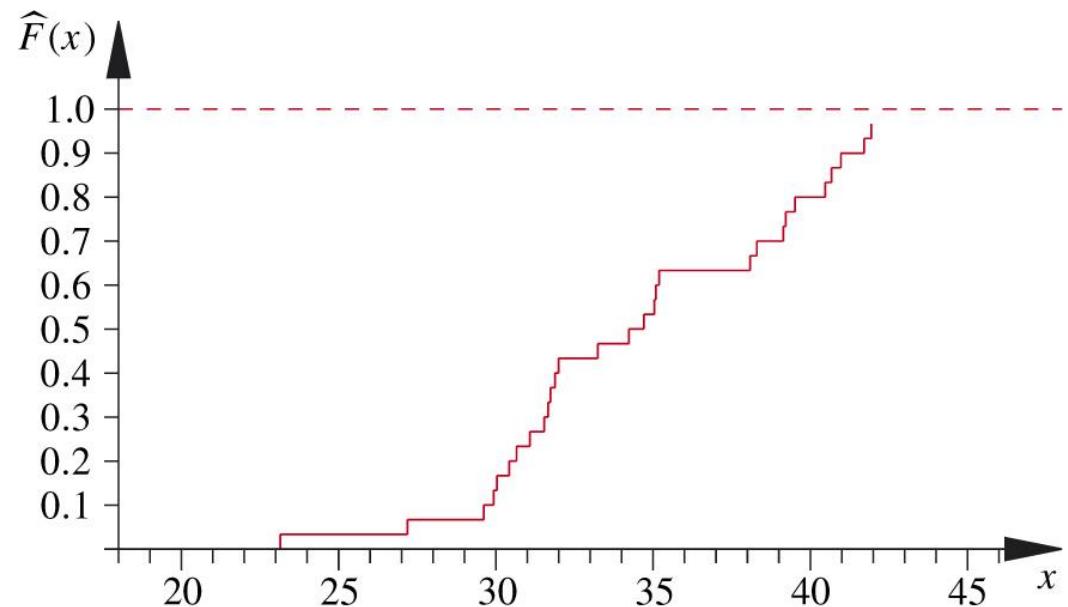
## (Empirische) Verteilungsfunktion (gem. [5,p.7])

Der Wert  $\hat{F}(x)$  der empirischen Verteilungsfunktion [einer Stichprobe  $(x_1, \dots, x_n)$ ] ist die relative Häufigkeit der Elemente, deren Ausprägung kleiner oder gleich  $x$  sind, sprich

$$\hat{F}(x) = \frac{1}{n} |\{i \mid x_i \leq x\}|$$

### Beispiel:

Empirische Verteilungsfunktion  
mit 30 Beobachtungen



Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008



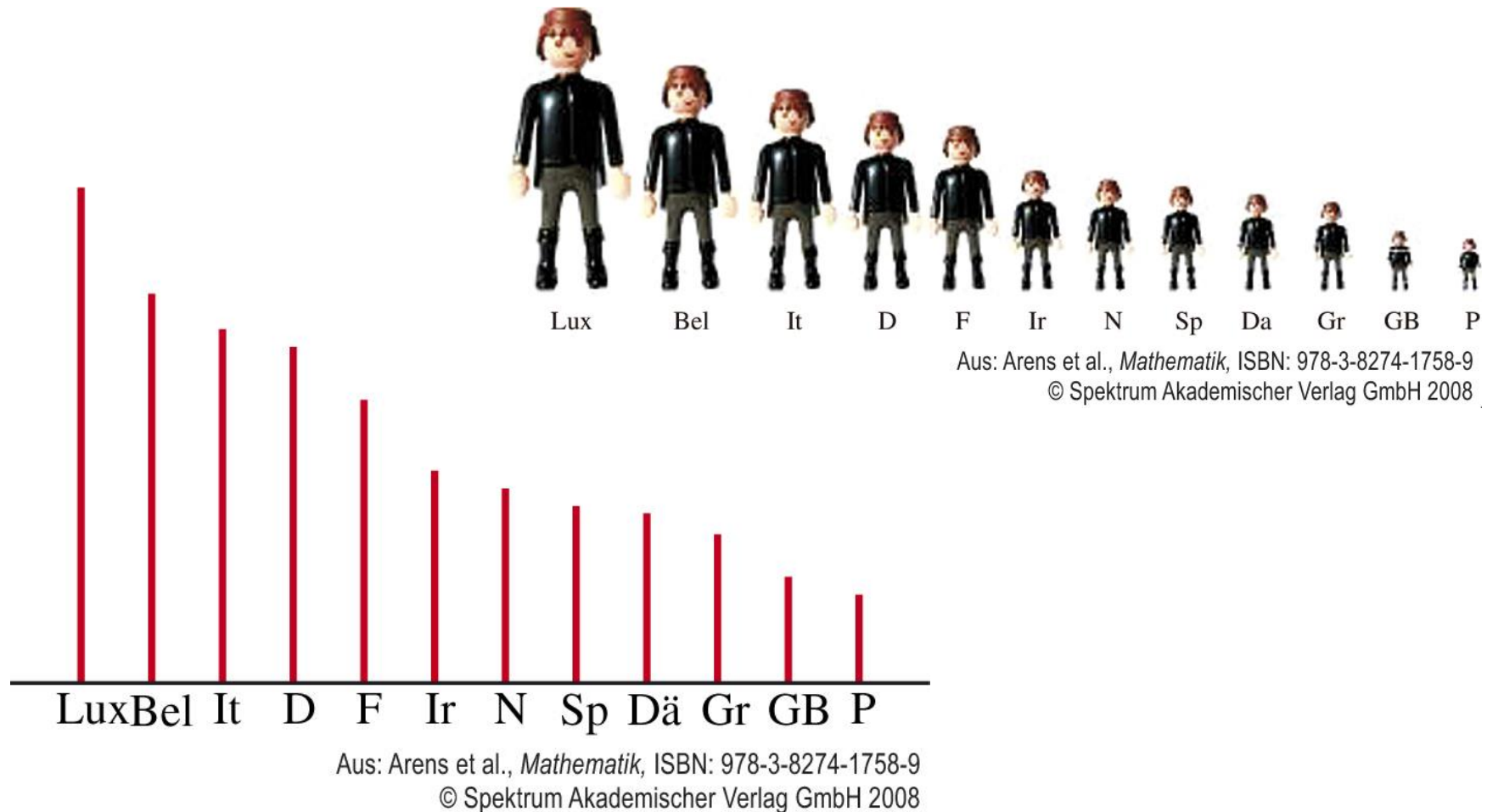
## VE 02: Tabellarische und grafische Darstellung von Daten

- 2.1 Experiment: Ablesen von Prozentdarstellungen
- 2.2 Häufigkeiten und Tabellendarstellungen
- 2.3 Stab-, Säulen-, Kurven-, Balken- und Kreisdiagramme
- 2.4 Strichdiagramme und Histogramme
- 2.5 Piktogramme, Polaritätsdiagramme, Netzdiagramme
- 2.6 Box-Plot-Diagramme und empirische Verteilungsfunktion
- 2.7 Problematische grafischen Darstellungen

# Problematische Darstellungen (1): Irreführende Volumina



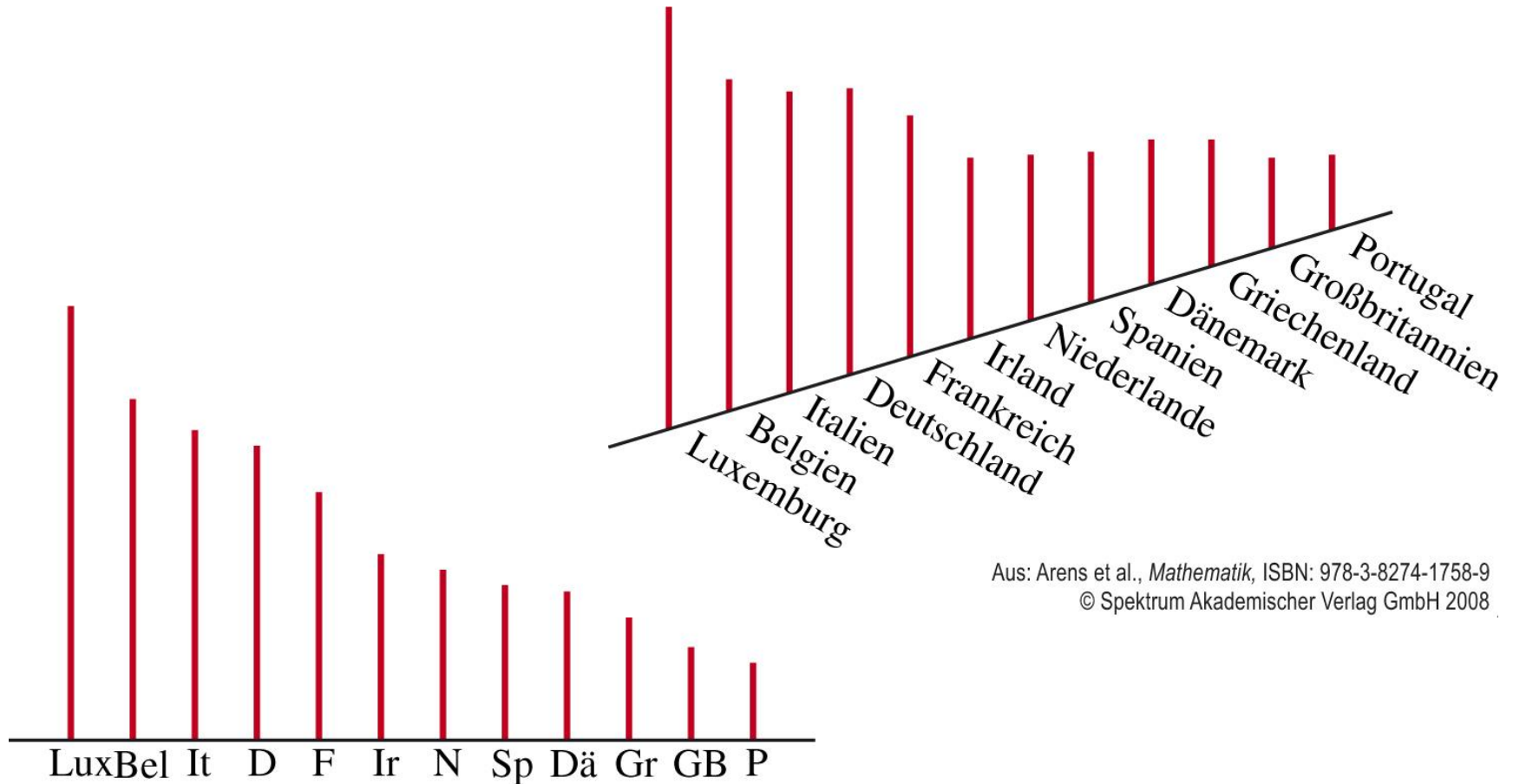
Subventionen der EU pro Beschäftigten im Jahr 1994:



# Problematische Darstellungen (2): Perspektivische Darstellung



Subventionen der EU pro Beschäftigten im Jahr 1994:



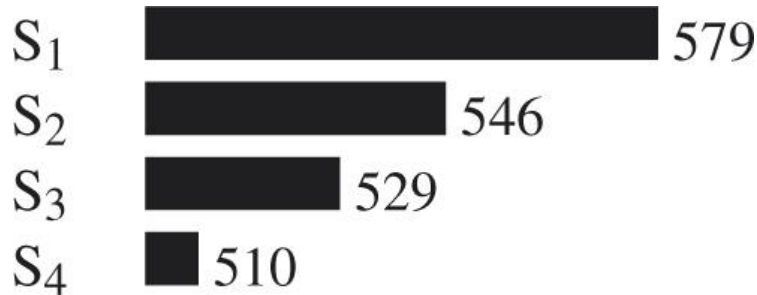
Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008

Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008

# Problematische Darstellungen (3): Nullpunktunterdrückung



Vergleich von PISA-Testergebnisse:



Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008



Aus: Arens et al., *Mathematik*, ISBN: 978-3-8274-1758-9  
© Spektrum Akademischer Verlag GmbH 2008

# Alternative zur Nullpunktunterdrückung



Vergleich von PISA-Testergebnisse:

