1. In der Statistik gibt es vier Arten von Skalen:

I Qualitative Merkmalsart

1. Die **Nominalskala**
   Es gibt keine Möglichkeit der Anordnung. Zugeordnete Zahlen sind willkürlich. Es handelt sich um eine reine Beschreibung.
   Beispiele: Geschlecht, Beruf, Bundesland, Haarfarbe, Handy-Nummern, Farbe, Raucher-Nichtraucher.

2. Die **Ordinalskala**
   Die Ausprägungen können mit kleiner/größer, schlechter/besser, seltener/häufiger, … verglichen werden und so kann eine Rangordnung hergestellt werden.
   Beispiele: Bewertung einer Leistung durch Noten, Zufriedenheitsgrad, Schulabschlüsse, Platzierungen, Güteklassen.

II Quantitative Merkmalsart

3. Die **Intervallskala**
   Die Intervallskala weist gleichgroße Skalenabschnitte auf, wodurch Differenzen oder Summen zwischen den Daten sinnvoll interpretiert werden können. Es existiert kein **natürlicher** Nullpunkt. Intervallskalen können mit einer linearen Transformation $y = a \cdot x + b$ transformiert werden, ohne ihren Charakter zu verlieren.
   Beispiele: Temperatur in °C, nicht in Kelvin, da ein natürlicher Nullpunkt existiert, Längengrade der Erde (Ohne natürlichen Nullpunkt)

4. Die **Verhältnisskala**
   Bei der Verhältnisskala gibt es im Gegensatz zur Intervallskala einen absoluten Nullpunkt. Verhältnisskalen können mit einer Transformation $y = a \cdot x$ transformiert werden, ohne ihren Charakter zu verlieren.
   Beispiele: Temperatur in Kelvin, Alter in Jahren, Einkommen, Preis, Gewicht, Geschwindigkeit, Körpergröße.

Finden Sie weitere Beispiele?

2. Beispiel einer Nominalskala

Es wurden 54 Personen befragt, mit welchem Verkehrsmittel sie zur Arbeit kommen.
‚1'– zu Fuß, ‚2' – Fahrrad, ‚3' – öffentliches Verkehrsmittel, ‚4' – eigenes motorisiertes Fahrzeug, ‚5' –Mitfahrer motorisiertes Fahrzeug

j	x_j	h_j
1	1	6
2	2	11
3	3	13
4	4	19
5	5	5

Wieso ist das Merkmal ‚Verkehrsmittel' nominal skaliert, obwohl wir Zahlen verwenden?
Bestimmen Sie den Modus (häufigster, wahrscheinlichster Wert) und die relativen Häufigkeiten. Zeichnen Sie ein Balkendiagramm und ein Kreisdiagramm.

3. Beispiel einer Ordinalskala

Die Mathematiknoten von 50 Schülern wurden erhoben.

j	x_j	h_j
1	1	6
2	2	12
3	3	18
4	4	9
5	5	4
6	6	1

Wieso ist das Merkmal ‚Mathematiknote' ordinal skaliert?
Bestimmen Sie den Median, das 1. und 3. Quartil und das arithmetische Mittel $\bar{x}$.
Bestimmen Sie in einer Tabelle die relativen Häufigkeiten f_j, die kumulierten absoluten Häufigkeiten H_j und die kumulierten relativen Häufigkeiten F_j und zeichnen Sie die empirische Verteilungsfunktion F(x) über x. Bestimmen Sie hieraus nochmals den Median und die Quartile.

4. Beispiel einer Verhältnisskala

Aus einer Sendung wurden 80 Tomaten herausgegriffen und ihre Masse in Gramm bestimmt.

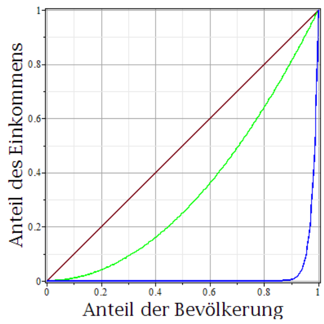
Von .... bis unter ...	h_j	w_j	x_j	$h_j \cdot x_j$	f_j	f_j^*	F_j
0 – 30	4						
30 – 50	10						
50 – 60	14						
60 – 65	25						
65 – 75	15						
75 – 100	12						
$\sum$	80	100	–		1	–	–

Vervollständigen Sie die Tabelle durch die Klassenbreiten w_j, die Klassenmitten x_j, die geschätzten Massen pro Klasse $h_j \cdot x_j$, die relativen Häufigkeiten f_j, die relativen Dichten f_j^* und die kumulierten relativen Häufigkeiten $F_j = \sum_{k=1}^{j} f_k$.

Zeichnen Sie ein Histogramm: $f^*(x)$ über x und konstruieren Sie den Modus D aus diesem Histogramm (Siehe Skript Seite 10).
Zeichnen Sie die klassierte Verteilungsfunktion: $F^*(x)$ über x.

Bestimmen Sie den Mittelwert $\bar{x} = \frac{1}{n}\sum_{j=1}^{m} h_j \cdot x_j$.

5. Interpretieren Sie die drei Lorenz-Kurven.



6. Gegeben sind die 25 Daten 1, 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 8.
Bestimmen Sie den Modus, die drei Quartile und das arithmetische Mittel und zeichnen Sie einen Box-Plot.
Erstellen Sie eine Tabelle wie auf Seite 9 des Skripts.
Zeichnen Sie die Diagramme für die relative Häufigkeit und die kumulierte relative Häufigkeit.
Bestimmen Sie die Varianz s^2 und die Standardabweichung s.

7. Beweisen Sie für den Spezialfall von zwei positiven Messwerten x_1 und x_2, dass $\bar{x}_{harm} \le \bar{x}_{geom} \le \bar{x}_{arith}$ gilt.
Das Gleichheitszeichen gilt im Fall $x_1 = x_2$.

8. Ein Kapital K_0 wird für n Jahre, $n \ge 1$, angelegt. Im j-ten Jahr, $j = 1, \dots, n$, wird es mit p_j % verzinst.
Bestimmen Sie eine Formel für den mittleren jährlichen Zinssatz $\bar{p}$ %.

Zeigen Sie, dass $\bar{p} = \left(\bar{q}_{geom} - 1\right) \cdot 100$ beträgt, wobei $q_j = 1 + \frac{p_j}{100}$ für $j = 1, \dots, n$ gilt.

Beispiel: Es sei $n = 4$ mit 2%, 3%, 4% und 5%. Bestimmen Sie $\bar{p}$.

9. Ein Schwimmbecken wird von n Rohren, $n \geq 1$, befüllt. Das j-te Rohr würde allein das Becken in der Zeit t_j in Stunden füllen. In welcher Zeit t füllen alle n Rohre zusammen das Becken?

Zeigen Sie, dass $t = \frac{1}{n} \cdot \bar{t}_{harm}$.

Beispiel: Es sei $n = 3$ mit $t_1 = 20$, $t_2 = 25$, $t_3 = 30$ jeweils in Stunden. Bestimmen Sie $\bar{t}_{harm}$ und t.

10. Zusammensetzen von drei Messreihen

1. Messreihe	1,1	1,2	–	–	–
2. Messreihe	0,9	1,0	0,8	–	–
3. Messreihe	1,0	0,8	1,1	1,0	1,1

Bestimmen Sie die drei Mittelwerte $\bar{x}_1$, $\bar{x}_2$ und $\bar{x}_3$ und die drei Varianzen s_1^2, s_2^2 und s_3^2.

Bestimmen Sie $\bar{x}$ und s^2 für die drei Messreihen zusammen mit Hilfe der Formeln $\bar{x} = \frac{1}{n} \cdot \sum_{j=1}^{m} n_j \cdot \bar{x}_j$ und

$$s^2 = s_{int}^2 + s_{ext}^2 \quad \text{mit} \quad s_{int}^2 = \frac{1}{n} \cdot \sum_{j=1}^{m} n_j \cdot s_j^2 \quad \text{und} \quad s_{ext}^2 = \frac{1}{n} \cdot \sum_{j=1}^{m} n_j \cdot \left(\bar{x}_j - \bar{x}\right)^2 .$$

Bestimmen Sie nochmals $\bar{x}$ und s^2 für die drei Messreihen mit Hilfe der 10 Daten – als eine Messreihe aufgefasst.

11. Aufgabe aus einem SAT-Test (Scholastic Assessment Test) aus den USA

a. In einer Klasse mit 30 Schülern wird ein durchschnittliches Testergebnis von 80 Punkten erzielt. Das durchschnittliche Testergebnis der 20 Mädchen beträgt 85. Bestimmen Sie das durchschnittliche Testergebnis der Jungen.

b. In einer Klasse mit p Schülern wird ein durchschnittliches Testergebnis von 70 Punkten erzielt, in einer anderen Klasse mit n Schülern beträgt das durchschnittliche Testergebnis 92 Punkte. Das durchschnittliche Testergebnis beider Klassen zusammen beträgt 86 Punkte. Bestimmen Sie das Verhältnis p / n.

12. Nochmals zu den 80 Tomaten aus Aufgabe 4

Von .... bis unter ...	w_j	h_j	f_j	$f_j \cdot w_j^2$	x_j	$f_j \cdot x_j$	$(x_j - \bar{x})^2$	$f_j \cdot (x_j - \bar{x})^2$
0 – 30		4						
30 – 50		10						
50 – 60		14						
60 – 65		25						
65 – 75		15						
75 – 100		12						
$\sum$	100	80	1		–		–	

Füllen Sie die Tabelle aus und bestimmen Sie näherungsweise den Mittelwert und die Varianz.

13. Zehn Firmen wurden nach der Höhe ihres Umsatzes in Mio. Euro befragt. Es ergab sich die Tabelle:

j	Umsatz x_j	absolute Häufigkeit h_j	relative Häufigkeit f_j	kumulierte relative Häufigk. F_j	relative Merkmalssumme MS. $\ell_j = \frac{h_j \cdot x_j}{\sum_{k=1}^{4} h_k \cdot x_k}$	kumulierte rel. MS. $L_j = \sum_{k=1}^{j} \ell_k$
1	5	2	0,2	0,2	0,05	0,05
2	10	1	0,1	0,3	0,05	0,10
3	15	2				
4	30	5				
$\sum$		10	1	–	1	–

Tragen Sie L_j über F_j auf und zeichnen Sie die Lorenzkurve.

Dabei ist F_j der Prozentsatz von der Befragten, deren Umsatz bis zu x_j beträgt.

L_j ist der Prozentsatz des Gesamtumsatzes $\sum_{k=1}^{4} h_k \cdot x_j = 200$ bis zum Umsatz x_j.

Geben Sie Beispiele für die Zuordnungen $L_j \to F_j$ und umgekehrt.

(Bestimmen Sie den Gini-Koeffizienten G.)

14. 100 Leute wurden nach der Höhe ihres wöchentlichen Einkommens befragt. Es ergab sich die Tabelle:

j	Einkommen von ... bis unter ...	absolute Häufigkeit h_j	Klassenbreite w_j	Klassenmitte m_j	relative Häufigkeit f_j	Relative Dichte f_j^*	kumulierte relative Häufigk. F_j	relative Merkmalssumme MS. $\ell_j = \frac{h_j \cdot m_j}{\sum_{k=1}^{m} h_k \cdot m_k}$	kumulierte rel. MS. $L_j = \sum_{k=1}^{j} \ell_k$
1	0 – 200	15	200	100	0,15	0,00075	0,15	0,040	0,040
2	200 – 400	50	200	300	0,50	0,0025	0,65	0,395	0,435
3	400 – 600	20							
4	600 – 800	10							
5	800 – 1000	5							
$\sum$		100	1000	–	1	–	–	1	–

Tragen Sie L_j über F_j auf und zeichnen Sie die Lorenzkurve.

Dabei ist F_j der Prozentsatz von der Befragten, deren Einkommen bis zur Klasse j gehört.

L_j ist der Prozentsatz des Gesamteinkommens $\sum_{k=1}^{m} h_k \cdot m_k$ bis zum Einkommen der Klasse j.

Geben Sie Beispiele für die Zuordnungen $F_j \to L_j$ und umgekehrt.