

UNIVARIANTE VERTEILUNGEN II: QUANTILE

Martin-Luther-Universität Halle-Wittenberg

Institut für Soziologie

Übung Einführung in die deskriptive Statistik

Was machen wir heute?

- Wiederholung
- Lagemaße I:
 - Quantile
 - Urliste und Primärliste
 - Unklassierte Häufigkeitstabellen
 - Klassierte Häufigkeitstabellen

Aufgabe 1: Soziale Unterschiede

Eine Ungleichheitsforscherin interessiert sich dafür, ob soziale Unterschiede von den Befragten des ALLBUS 2014 als gerecht angesehen werden. Sie erhält folgende absolute Häufigkeiten:

	Häufigkeit
1 STIMME GAR NICHT ZU	727
2 STIMME EHER NICHT ZU	1593
3 STIMME EHER ZU	905
4 STIMME VOLL ZU	179
8 WEISS NICHT	48
9 KEINE ANGABE	19
Gesamt	3471

- Liegen fehlende Werte vor? Wenn ja, welche? Begründen Sie ihr Entscheidung. Welches Skalenniveau hat die Variable?
- Berechnen Sie die Gesamtprozente, gültigen Prozente sowie die kumulierten Prozente. Sind die kumulierten Prozente hier sinnvoll interpretierbar? Warum?
- Zeichnen Sie die empirische Verteilungsfunktion.

Aufgabe 1a: Lösung

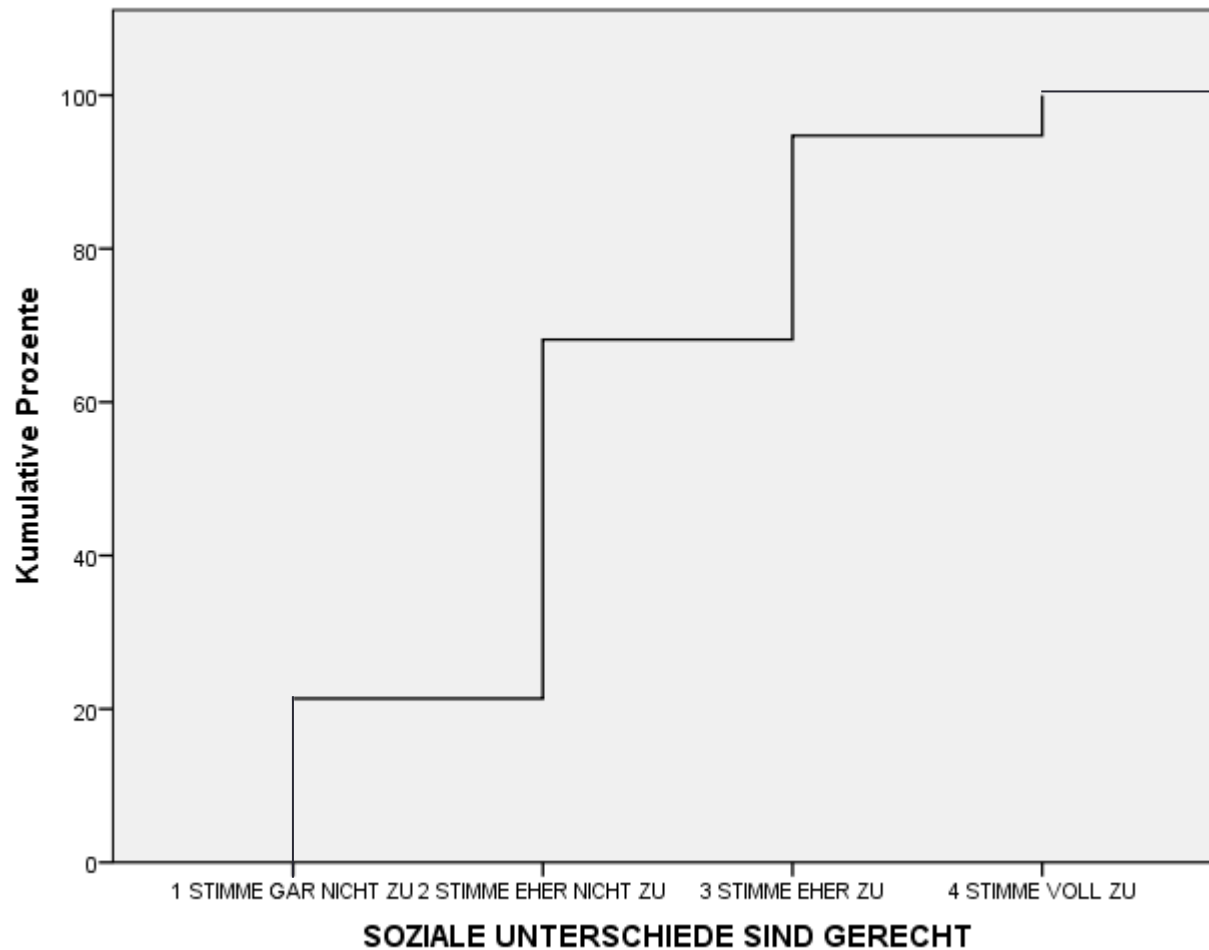
- fehlende Werte:
 - 8 weiß nicht
 - 9 keine Angabe
- Skalenniveau:
 - ordinal

Aufgabe 1b: Lösung

SOZIALE UNTERSCHIEDE SIND GERECHT

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1 STIMME GAR NICHT ZU	727	20,9	21,4	21,4
	2 STIMME EHER NICHT ZU	1593	45,9	46,8	68,2
	3 STIMME EHER ZU	905	26,1	26,6	94,7
	4 STIMME VOLL ZU	179	5,2	5,3	100,0
	Gesamt	3404	98,1	100,0	
Fehlend	8 WEISS NICHT	48	1,4		
	9 KEINE ANGABE	19	,5		
	Gesamt	67	1,9		
Gesamt		3471	100,0		

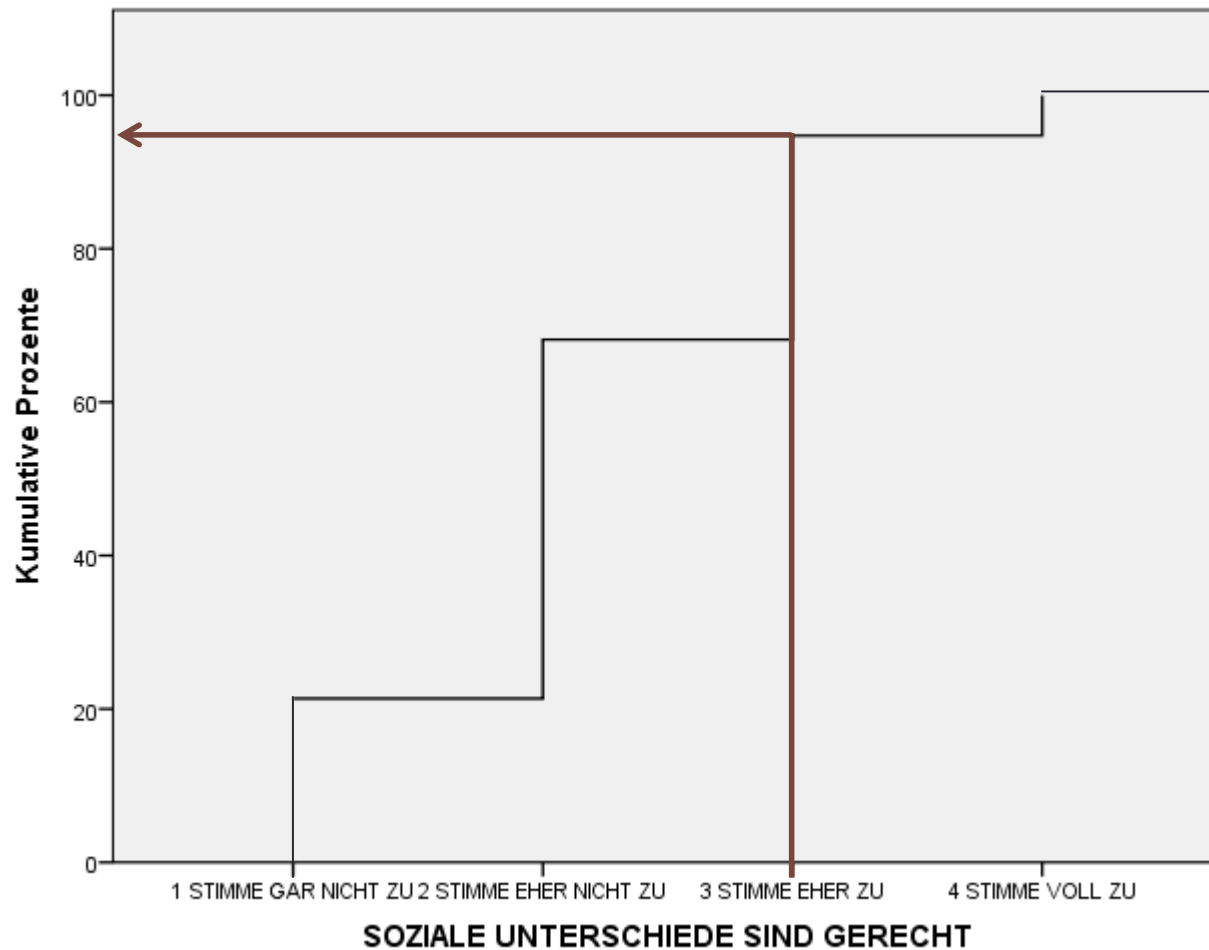
Aufgabe 1c: Lösung



Quiz: Wie viel Prozent der Befragten mit gültigen Werten, befinden sich unterhalb der Kategorie 2 „stimme eher nicht zu“?

Was bedeutet das inhaltlich?

Empirische Verteilungsfunktion: Fragestellung



Fragestellung:
Wie viel Prozent der Fälle
befinden sich unterhalb
einer Kategorie bzw.
eines Wertes?

Übersicht I: statistische Kennwerte

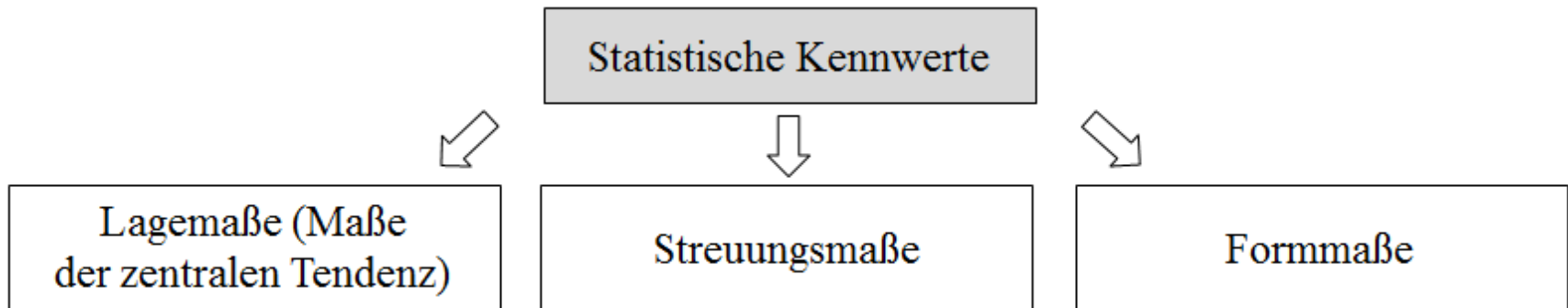


Bild: Kerstin Völkl / Christoph Korb

Übersicht II: Lagemaße

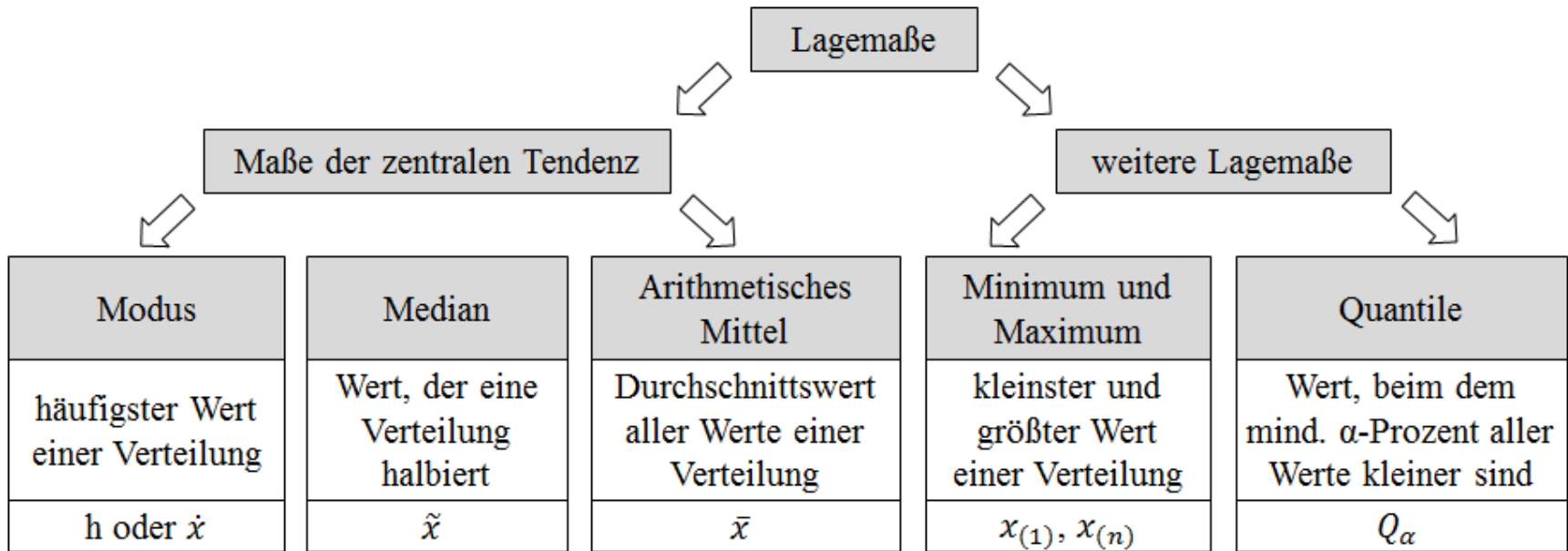


Bild: Kerstin Völkl / Christoph Korb

kommende Woche...

heute 😊

Quantile

- ab ordinalem Skalenniveau
- Quantil(wert):
 - „kleinste Ausprägung einer Variablen, für die gilt, dass ein (vorgegebener) kumulierter Anteil α (Alpha) der Verteilung kleiner oder gleich dieser Ausprägung ist“ (Kühnel / Krebs 2012: 71)
- Quantilanteil:
 - „kumulierter Anteil α (*Alpha*) einer Verteilung zu einem Quantilwert Q_α “ (Kühnel / Krebs 2012: 71)

Quantile II: Visualisierung der Fragestellung



Fragestellung:
Unterhalb welchem Wert
liegt mindestens ein Anteil
 α der gültigen Fälle?

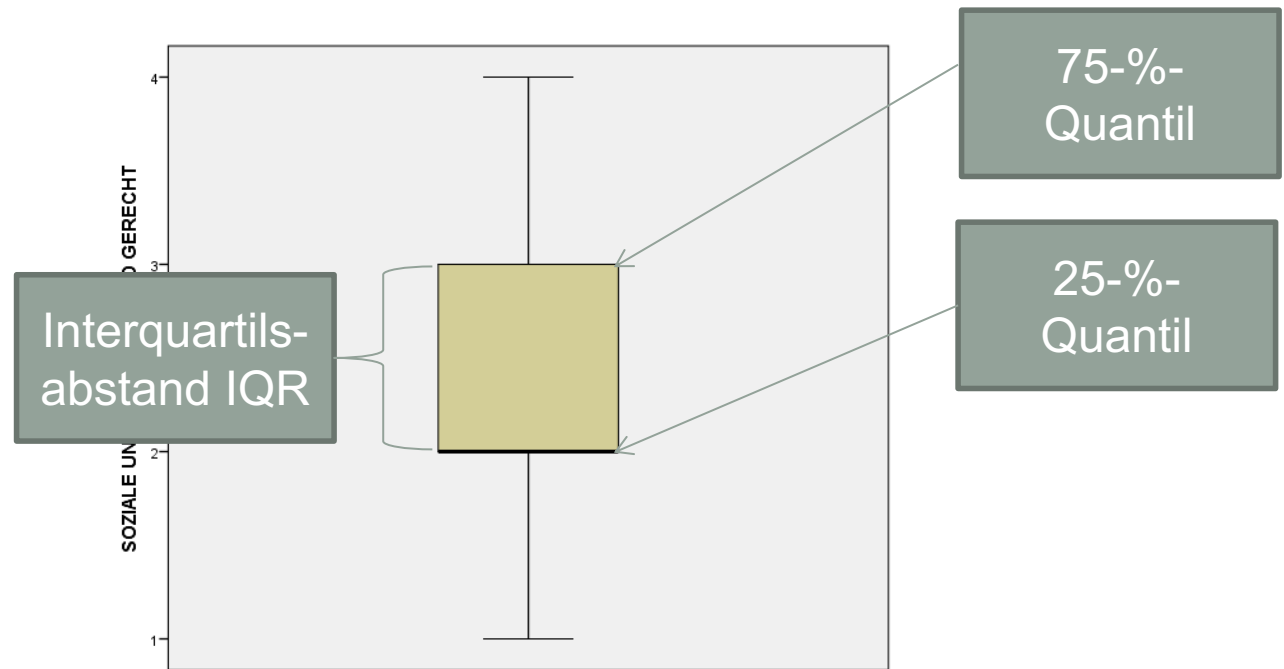
Quantile III: Spezielle Quantile

- Quartile
 - 1. Quartil = 25-%-Quantil (Q1)
 - 2. Quartil = 50-%-Quantil (Q2)
 - 3. Quartil = 75-%-Quantil (Q3)
 - 4. Quartil = 100-%-Quantil (Q4)
- Dezil:
 - 1. Dezil = 10-%-Quantil... (D1)
 - 9. Dezil = 90-%-Quantil... (D9)
- Perzentile:
 - 1. Perzentil = 1-%-Quantil
 - 12. Perzentil = 12-%-Quantil
 - 99. Perzentil = 99-%-Quantil

Lesehilfe: unterhalb des Quantilwerts für das 1. Quartil liegen mindestens 25 % der gültigen Fälle

Quantile IV: Bedeutung

- ermöglichen Aussagen über Verteilung
- Berechnung des Interquartilsabstands
- Box-Whisker-Plots



Quantile: ungeordnete Urliste

In einer (fiktiven) Befragung wurden 10 Leute gebeten anzugeben, wie viele Menschen in ihrem Haushalt leben. Dabei entstand die rechts stehende ungeordnete Urliste.

- Bestimmen Sie das 75-%-Quantil.
- Interpretieren Sie dieses inhaltlich.

	v351
1	3
2	2
3	3
4	1
5	1
6	4
7	2
8	1
9	2
10	2



Wie machen wir das?

Quantile: ungeordnete Urliste II

1. Urliste der Größe nach ordnen → Primärtafel erstellen
2. gesuchte Indexposition k in der geordneten Liste mit der Formel $k = n * \alpha$ berechnen

k : Index des Wertes in der geordneten Zahlenreihe

n : Gesamtzahl der gültigen Fälle

α : gesuchter Quantilsanteilstwert, also z.B. 0,75 für 75%

3. ggfs. erzielten Wert für k aufrunden
4. Wert für $x_{(k)}$ an Indexposition k abgelesen

Quantile: ungeordnete Urliste III

	v351
1	3
2	2
3	3
4	1
5	1
6	4
7	2
8	1
9	2
10	2

ungeordnete Urliste

Fälle aufsteigend sortieren

	v351
1	1
2	1
3	1
4	2
5	2
6	2
7	2
8	3
9	3
10	4

Primärtafel

Quantile: ungeordnete Urlisten IV

- Formel:
 - $k = n * \alpha$
- Berechnung:
 - $k = 10 * 0,75$
 - $k = 7,5$
 - aufrunden: $k = 8$
 - → in geordneter Liste Wert an Indexposition 8 ablesen

In einer (fiktiven) Befragung wurden 10 Leute gebeten anzugeben, wie viele Menschen in ihrem Haushalt leben. Dabei entstand die rechts stehende ungeordnete Urliste.

- a) Bestimmen Sie das 75%-Quantil.
- b) Interpretieren Sie dieses inhaltlich.

Quantile: ungeordnete Urlisten V

- Primärtafel

(k)	1	2	3	4	5	6	7	8	9	10
$x_{(k)}$	1	1	1	2	2	2	2	3	3	4

- Wert ablesen:

- $k = 8 \rightarrow x_{(k)} = 3$

- Interpretation:

- mindestens 75 Prozent der Haushalte haben drei oder weniger Mitglieder

Aufgabe 2: enge Freunde

In einer sozialwissenschaftlichen Befragung wurden 10 Personen gebeten anzugeben, wie viele enge Freunde sie haben. Dabei entstand folgende Urliste:

Person	1	2	3	4	5	6	7	8	9	10
Anzahl Freunde	3	1	2	10	3	3	4	3	2	3

- Bestimmen Sie das 25-%-Quantil und interpretieren Sie dieses.
- Bestimmen Sie das 50-%-Quantil und interpretieren Sie dieses.

Aufgabe 2: Lösung

Person	1	2	3	4	5	6	7	8	9	10
Anzahl Freunde	3	1	2	10	3	3	4	3	2	3

Sortieren

(k)	1	2	3	4	5	6	7	8	9	10
$x_{(k)}$	1	2	2	3	3	3	3	3	4	10

Aufgabe 2a: Lösung

- 25-%-Quantil:

- $k = n * \alpha$

- $k = 10 * 0,25 = 2,5 \rightarrow \text{aufrunden } 3$

(k)	1	2	3	4	5	6	7	8	9	10
$x_{(k)}$	1	2	2	3	3	3	3	3	4	10

- $x_{(k)} = x_{(3)} = 2$

- Mindestens 25 Prozent der gültigen Fälle weisen höchstens einen Wert von zwei auf.

- Folglich haben mindestens 25 Prozent der Befragten mit gültigen Werten höchstens zwei enge Freunde.

Aufgabe 2b: Lösung

- 50-%-Quantil:

- $k = n * \alpha$

- $k = 10 * 0,5 = 5$

(k)	1	2	3	4	5	6	7	8	9	10
$x_{(k)}$	1	2	2	3	3	3	3	3	4	10

- $x_{(k)} = x_{(5)} = 3$

- Mindestens 50 Prozent der gültigen Fälle weisen höchstens einen Wert von drei auf.

- Folglich haben mindestens 50 Prozent der Befragten mit gültigen Werten angegeben, dass sie höchstens drei enge Freunde haben.

Primärtafel und unklassierte Häufigkeitstabelle

(k)	1	2	3	4	5	6	7	8	9	10
$x_{(k)}$	1	2	2	3	3	3	3	3	4	10



Anzahl enge Freunde

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	1	1	10,0	10,0	10,0
	2	2	20,0	20,0	30,0
	3	5	50,0	50,0	80,0
	4	1	10,0	10,0	90,0
	10	1	10,0	10,0	100,0
	Gesamt	10	100,0	100,0	

Wie können Quantile anhand der Häufigkeitstabelle bestimmt werden?

Quantile: unklassierte Häufigkeitstabelle

Anzahl enge Freunde

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig 1	1	10,0	10,0	10,0
2	2	20,0	20,0	30,0
3	5	50,0	50,0	80,0
4	1	10,0	10,0	90,0
10	1	10,0	10,0	100,0
Gesamt	10	100,0	100,0	

Quantilwert ist der Wert, bei dem bei kumulierten Prozente der vorgegebene Alpha-Wert erstmals erreicht oder überschritten wird...

Quiz: Bei welchem Wert liegt das 75%-Quantil? Was bedeutet es?

Aufgabe 3: Fachsemester

Anzahl enge Freunde

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig 1	1	10,0	10,0	10,0
2	2	20,0	20,0	30,0
3	5	50,0	50,0	80,0
4	1	10,0	10,0	90,0
10	1	10,0	10,0	100,0
Gesamt	10	100,0	100,0	

- Bestimmen Sie das 25-%-Quantil und interpretieren Sie dieses.
- Bestimmen Sie das zweite Quartil und interpretieren Sie dieses.
- Bestimmen Sie das 70. Perzentil und interpretieren Sie dieses.

Aufgabe 3a: Lösung

Anzahl enge Freunde

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig 1	1	10,0	10,0	10,0
2	2	20,0	20,0	30,0
3	5	50,0	50,0	80,0
4	1	10,0	10,0	90,0
10	1	10,0	10,0	100,0
Gesamt	10	100,0	100,0	

Mindestens 25 Prozent der Befragten mit gültigen Werten haben höchstens zwei enge Freunde.

Aufgabe 3b: Lösung

Anzahl enge Freunde

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig 1	1	10,0	10,0	10,0
2	2	20,0	20,0	30,0
3	5	50,0	50,0	80,0
4	1	10,0	10,0	90,0
10	1	10,0	10,0	100,0
Gesamt	10	100,0	100,0	

Mindestens 50 Prozent der Befragten haben mit gültigen Werten höchstens drei enge Freunde.

Aufgabe 3c: Lösung

Anzahl enge Freunde

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig 1	1	10,0	10,0	10,0
2	2	20,0	20,0	30,0
3	5	50,0	50,0	80,0
4	1	10,0	10,0	90,0
10	1	10,0	10,0	100,0
Gesamt	10	100,0	100,0	

Mindestens 70 Prozent der Befragten mit gültigen Werten haben höchstens 3 enge Freunde.

Probleme mit Häufigkeitstabellen für Variablen mit vielen Ausprägungen

ALTER: BEFRAGTE<R>

	Häufigkeit	Prozent	Gültige Prozent	Kumulative Prozente
Gültig 18	29	,8	,8	,8
19	54	1,6	1,6	2,4
20	42	1,2	1,2	3,6
21	56	1,6	1,6	5,2
22	41	1,2	1,2	6,4
23	49	1,4	1,4	7,8
24	46	1,3	1,3	9,1
--	--	--	--	---
88	4	,1	,1	99,8
89	2	,1	,1	99,9
90	3	,1	,1	99,9
91	2	,1	,1	100,0
Gesamtsumme	3468	99,9	100,0	
Fehlend KEINE ANGABE	3	,1		
Gesamtsumme	3471	100,0		

Quiz: Welches Problem haben wir an dieser Stelle?
Wie können wir das Problem beheben?

Klassenbildung

- Metrische Variablen sollten möglichst auch metrisch erhoben werden. (Es droht ansonsten der Verlust sinnvoller Informationen!)
- Klassierung der Daten bei der Erhebung nur sinnvoll, wenn dadurch Überforderung des Befragten vermieden wird oder man so Antwortverweigerung vermeiden kann (z. B. Einkommensvariable).
- Die nachträgliche Klassierung von Daten, um einen Überblick über die Daten zu bekommen oder für Grafiken ist in der Regel unproblematisch, da die Ausgangsdaten erhalten bleiben.

Klassenbildung: Anforderungen

- Klassen sollten...
 - überschneidungsfrei (disjunkt) sein
 - keine Lücken aufweisen
 - möglichst gleich groß sein (Ausnahme: erste / letzte Kategorie) oder inhaltlich begründet sein

Klassierte Häufigkeitstabellen

ALTER: BEFRAGTE<R>, KATEGORISIERT

		Häufigkeit	Prozent	Gültige Prozent	Kumulative Prozente
Gültig	1 - 18-29 JAHRE	583	16,8	16,8	16,8
	2 - 30-44 JAHRE	752	21,7	21,7	38,5
	3 - 45-59 JAHRE	1083	31,2	31,2	69,7
	4 - 60-74 JAHRE	749	21,6	21,6	91,3
	5 - 75-89 JAHRE	296	8,5	8,5	99,9
	6 - UEBER 89 JAHRE	5	,1	,1	100,0
	Gesamtsumme	3468	99,9	100,0	
Fehlend	9 - KEINE ANGABE	3	,1		
Gesamtsumme		3471	100,0		

Quiz: Welche Vor- und Nachteile sind mit dieser Darstellungsform verbunden?

Quantile: klassierte Häufigkeitstabellen

ALTER: BEFRAGTE<R>, KATEGORISIERT

	Häufigkeit	Prozent	Gültige Prozent	Kumulative Prozente
Gültig				
1 - 18-29 JAHRE	583	16,8	16,8	16,8
2 - 30-44 JAHRE	752	21,7	21,7	38,5
3 - 45-59 JAHRE	1083	31,2	31,2	69,7
4 - 60-74 JAHRE	749	21,6	21,6	91,3
5 - 75-89 JAHRE	296	8,5	8,5	99,9
6 - UEBER 89 JAHRE	5	,1	,1	100,0
Gesamtsumme	3468	99,9	100,0	
Fehlend				
9 - KEINE ANGABE	3	,1		
Gesamtsumme	3471	100,0		

Hinweis: der älteste Befragte war 91 Jahre alt...

Quiz: Welches Problem haben wir, wenn wir hier Quantile mit unserem bisherigen Weg bestimmen wollen?

Quantile: klassierte Häufigkeitstabellen II

- Lösung: Interpolationsformel

- $Q_\alpha = u_k + \frac{\alpha - cp_{k-1}}{p_k} * (o_k - u_k)$

- k : Klasse, in der der gesuchte Anteil erstmals überschritten wird

- u_k : exakte untere Grenze der Klasse k

- o_k : exakte obere Grenze der Klasse k

- α : vorgegebener Quantilanteil

- cp_{k-1} : kumulierter Anteilswert der Klasse $k-1$

- p_k : gültiger Anteilswert der Klasse k

Quantile: klassierte Häufigkeitstabellen III

Code	Wertelabel	u_k	o_k	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
1	18-29 JAHRE	17,5	29,5	583	16,8	16,8	16,8
2	30-44 JAHRE	29,5	44,5	752	21,7	21,7	38,5
3	45-59 JAHRE	44,5	59,5	1083	31,2	31,2	69,7
4	60-74 JAHRE	59,5	74,5	749	21,6	21,6	91,3
5	75-89 JAHRE	74,5	89,5	296	8,5	8,5	99,9
6	90-91 JAHRE	89,5	91,5	5	0,1	0,1	100,0
	Gesamt			3468	99,9	100%	
9	KEINE ANGABE			3	0,1		
	Gesamt			3471	100,0		

exakte Klassengrenzen

Quantile: klassierte Häufigkeitstabellen IV

Code	Wertelabel	u_k	o_k	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
1	18-29 JAHRE	17,5	29,5	583	16,8	16,8	16,8
2	30-44 JAHRE	29,5	44,5	752	21,7	21,7	38,5
3	45-59 JAHRE	44,5	59,5	1083	31,2	31,2	69,7
4	60-74 JAHRE	59,5	74,5	749	21,6	21,6	91,3
5	75-89 JAHRE	74,5	89,5	296	8,5	8,5	99,9
6	90-91 JAHRE	89,5	91,5	5	0,1	0,1	100,0
	Gesamt			3468	99,9	100%	
9	KEINE ANGABE			3	0,1		
	Gesamt			3471	100,0		

Aufgabe: Bestimmen Sie das 25-%-Quantil.
1. Schritt: Kategorie bestimmen

Quantile: klassierte Häufigkeitstabellen V

Code	Wertelabel	u_k	o_k	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
1	18-29 JAHRE	17,5	29,5	583	16,8	16,8	16,8
2	30-44 JAHRE	29,5	44,5	752	21,7	21,7	38,5
3	45-59 JAHRE	44,5	59,5	1083	31,2	31,2	69,7
4	60-74 JAHRE	59,5	74,5	749	21,6	21,6	91,3
5	75-89 JAHRE	74,5	89,5	296	8,5	8,5	99,9
6	90-91 JAHRE	89,5	91,5	5	0,1	0,1	100,0

$$k = 2$$

$$Q_\alpha = u_k + \frac{\alpha - cp_{k-1}}{p_k} * (o_k - u_k)$$

Achtung, die Formel arbeitet mit dem Anteilswert (zwischen 0 und 1 liegend)!

Quiz: Was ist für die einzelnen Werte in die Formel einzusetzen?

Quantile: klassierte Häufigkeitstabellen VI

Code	Wertelabel	u_k	o_k	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
1	18-29 JAHRE	17,5	29,5	583	16,8	16,8	16,8
2	30-44 JAHRE	29,5	44,5	752	21,7	21,7	38,5
3	45-59 JAHRE	44,5	59,5	1083	31,2	31,2	69,7
4	60-74 JAHRE	59,5	74,5	749	21,6	21,6	91,3
5	75-89 JAHRE	74,5	89,5	296	8,5	8,5	99,9
6	90-91 JAHRE	89,5	96,5	5	0,1	0,1	100,0

$$Q_\alpha = u_k + \frac{\alpha - cp_{k-1}}{p_k} * (o_k - u_k)$$

$$Q_{0,25} = 29,5 + \frac{0,25-0,168}{0,217} * (44,5 - 29,5) = 35,17 \text{ Jahre}$$

Der Wert für das interpolierte 25-%-Quantil für die vorliegende gruppierte Häufigkeitsverteilung liegt bei etwa 35 Jahren.

Aufgabe 4: Quantile für klassiertes Alter

Code	Wertelabel	u_k	o_k	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
1	18-29 JAHRE			583	16,8	16,8	16,8
2	30-44 JAHRE			752	21,7	21,7	38,5
3	45-59 JAHRE			1083	31,2	31,2	69,7
4	60-74 JAHRE			749	21,6	21,6	91,3
5	75-89 JAHRE			296	8,5	8,5	99,9
6	90-91 JAHRE			5	0,1	0,1	100,0

Bestimmen Sie das 75-%-Quantil!
Interpretieren Sie Ihr Ergebnis.

Aufgabe 4: Lösung

Code	Wertelabel	u_k	o_k	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
1	18-29 JAHRE	17,5	29,5	583	16,8	16,8	16,8
2	30-44 JAHRE	29,5	44,5	752	21,7	21,7	38,5
3	45-59 JAHRE	44,5	59,5	1083	31,2	31,2	69,7
4	60-74 JAHRE	59,5	74,5	749	21,6	21,6	91,3
5	75-89 JAHRE	74,5	89,5	296	8,5	8,5	99,9
6	90-91 JAHRE	89,5	91,5	5	0,1	0,1	100,0

$$Q_\alpha = u_k + \frac{\alpha - cp_{k-1}}{p_k} * (o_k - u_k)$$

$$Q_{0,75} = 59,5 + \frac{0,75 - 0,697}{0,216} * (74,5 - 59,5) = 63,18 \text{ Jahre}$$

Literaturhinweis

- Kerstin Völkl / Christoph Korb (2018): Deskriptive Statistik. Eine Einführung für Politikwissenschaftlerinnen und Politikwissenschaftler. S. 60-70.

Übungsaufgabe 1: Interesse an Spielfilmen

Im ALLBUS 2014 wurden die Befragten gebeten anzugeben, wie sehr sie sich für Spielfilme interessieren. Mithilfe der Daten entstand folgende Häufigkeitstabelle:

		Häufigkeit	Prozent	Gültige Prozent	Kumulative Prozente
Gültig	1 - SEHR STARK	459	13,2		
	2 - STARK	1325	38,2		
	3 - MITTEL	1160	33,4		
	4 - WENIG	339	9,8		
	5 - UEBERHAUPT NICHT	88	2,5		
	Gesamtsumme	3371	97,1		
Fehlend	0 - TRIFFT NICHT ZU	98	2,8		
	9 -KEINE ANGABE	2	,1		
	Gesamtsumme	100	2,9		
Gesamtsumme		3471	100,0		

- Ergänzen Sie die gültigen Prozente und kumulierten Prozente!
- Bestimmen Sie das 50-%-Quantil und 3. Quartil. Interpretieren Sie Ihr Ergebnis inhaltlich!

Übungsaufgabe 1a: Lösung

		Häufigkeit	Prozent	Gültige Prozent	Kumulative Prozente
Gültig	1 - SEHR STARK	459	13,2	13,6	13,6
	2 - STARK	1325	38,2	39,3	52,9
	3 - MITTEL	1160	33,4	34,4	87,3
	4 - WENIG	339	9,8	10,1	97,4
	5 - UEBERHAUPT NICHT	88	2,5	2,6	100,0
	Gesamtsumme	3371	97,1	100,0	
Fehlend	0 - TRIFFT NICHT ZU	98	2,8		
	9 - KEINE ANGABE	2	,1		
	Gesamtsumme	100	2,9		
Gesamtsumme		3471	100,0		

Übungsaufgabe 1b: Lösung II

		Häufigkeit	Prozent	Gültige Prozent	Kumulative Prozente
Gültig	1 - SEHR STARK	459	13,2	13,6	13,6
	2 - STARK	1325	38,2	39,3	52,9
	3 - MITTEL	1160	33,4	34,4	87,3
	4 - WENIG	339	9,8	10,1	97,4
	5 - UEBERHAUPT NICHT	88	2,5	2,6	100,0
	Gesamtsumme	3371	97,1	100,0	
Fehlend	0 - TRIFFT NICHT ZU	98	2,8		
	9 - KEINE ANGABE	2	,1		
	Gesamtsumme	100	2,9		
Gesamtsumme		3471	100,0		

$$Q_{\alpha} = Q_{0,50} = 2.$$

Mindestens 50 % der Befragten mit gültigen Werten haben höchstens einen Wert von 2 erzielt. Das bedeutet hier, dass mindestens 50 % der Befragten ein starkes oder noch stärkeres Interesse an Spielfilmen haben.

Übungsaufgabe 1b: Lösung II

	Häufigkeit	Prozent	Gültige Prozent	Kumulative Prozente
Gültig	1 - SEHR STARK	459	13,2	13,6
	2 - STARK	1325	38,2	52,9
	3 - MITTEL	1160	33,4	87,3
	4 - WENIG	339	9,8	97,4
	5 - UEBERHAUPT NICHT	88	2,5	100,0
	Gesamtsumme	3371	97,1	100,0
Fehlend	0 - TRIFFT NICHT ZU	98	2,8	
	9 - KEINE ANGABE	2	,1	
	Gesamtsumme	100	2,9	
Gesamtsumme		3471	100,0	

$Q_3 = Q_{0,75} = 3$. Mindestens 75 % der Befragten mit gültigen Werten haben höchstens einen Wert von 3 erzielt. Das bedeutet hier, dass mindestens 75 % der Befragten ein mittleres oder noch stärkeres Interesse an Spielfilmen haben.

Übungsaufgabe 2: Alter ALLBUS 2012

Ein Wissenschaftler interessiert sich dafür, wie alt mindestens 75 Prozent der Befragten des ALLBUS 2012 höchstens waren. Leider hat er nur folgende klassierte Häufigkeitstabelle zur Bestimmung zur Verfügung:

	Häufigkeit
Gültig	
18-29 JAHRE	598
30-44 JAHRE	743
45-59 JAHRE	1039
60-74 JAHRE	789
75-89 JAHRE	289
UEBER 89 JAHRE	15
Gesamt	3473
Fehlend	
KEINE ANGABE	7
Gesamt	3480

- Bestimmen Sie den einen Wert für das 75-%-Quantil mithilfe der entsprechenden Interpolationsformel und interpretieren Sie dieses.
- Welches Problem ist mit der Bestimmung über klassierte Daten verbunden? Was könnte man machen, um das Problem zu umgehen?

Übungsaufgabe 2a: Lösung

Kategorie j	Altersgruppe	Exakte Grenzen $u_j - < o_j$	Klassen- mitte m_j	Häufigkeit n_j	Klassenmitte * Häufigkeit $m_j * n_j$
1	18-29 Jahre	17,5-<29,5	23,5	598	14053
2	30-44 Jahre	29,5-<44,5	37	743	27491
3	45-59 Jahre	44,5-<59,5	52	1039	54028
4	60-74 Jahre	59,5-<74,5	67	789	52863
5	75-89 Jahre	74,5-<89,5	82	289	23698
6	89-96 Jahre	89,5-<96,5	93	15	1395
	Summe			3473	173528
	Durchschnitt				49,9648719

Übungsaufgabe 3: Töpferkurs

An der halleschen Volkshochschule nehmen verschiedene Leute an einem Töpferkurs teil. Für das Alter der Personen wurde hierbei folgende Werte im Anmeldebogen angegeben: 22, 23, 28, 31, 35, 37, 41, 42, 46, 47, 51, 53, 55, 57, 62 Jahre.

Bestimmen Sie für die Verteilung die empirischen Quantilwerte für folgende Prozentwerte: 25%, 50%, 75% und 90%.

Übungsaufgabe 3: Lösung

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$x_{(k)}$	22	23	28	31	35	37	41	42	46	47	51	53	55	57	62

- $k = n * \alpha$
- $Q_{0,25}: k = 15 * 0,25 = 3,75 \rightarrow k = 4, x_{(4)} = 31 \text{ Jahre}$
- $Q_{0,50}: k = 15 * 0,50 = 7,50 \rightarrow k = 8, x_{(8)} = 42 \text{ Jahre}$
- $Q_{0,75}: k = 15 * 0,75 = 11,25 \rightarrow k = 12, x_{(12)} = 53 \text{ Jahre}$
- $Q_{0,90}: k = 15 * 0,90 = 13,50 \rightarrow k = 14, x_{(14)} = 57 \text{ Jahre}$

Übungsaufgabe 4

Ein Soziologe hat in einem Einrichtungshaus die Besucher nach ihrem Alter befragt und folgende Urliste dabei erhalten:

60, 61, 58, 37, 41, 45, 24, 33, 32, 52, 54, 48, 26, 37.

Helfen Sie ihm dabei aus der Urliste eine Primärtafel aufzustellen und berechnen Sie anschließend das erste, zweite und dritte Quartil sowie das 10-%-Quantil und das 90-%-Quantil.

Übungsaufgabe 4: Lösung

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$x_{(k)}$	24	26	32	33	37	37	41	45	48	52	54	58	60	61

- $k = n * \alpha$
- $Q_{0,25}: k = 14 * 0,25 = 3,5 \rightarrow \text{aufrunden} \rightarrow k = 4, x_{(4)} = 33 \text{ Jahre}$
- $Q_{0,50}: k = 14 * 0,50 = 7,0 \rightarrow k = 7, x_{(7)} = 41 \text{ Jahre}$
- $Q_{0,75}: k = 14 * 0,75 = 10,5 \rightarrow \text{aufrunden} \rightarrow k = 11, x_{(11)} = 54 \text{ Jahre}$
- $Q_{0,10}: k = 14 * 0,10 = 1,4 \rightarrow \text{aufrunden} \rightarrow k = 2, x_{(2)} = 26 \text{ Jahre}$
- $Q_{0,90}: k = 14 * 0,90 = 12,6 \rightarrow \text{aufrunden} \rightarrow k = 13, x_{(13)} = 60 \text{ Jahre}$