**Wirtschaftsstatistik**

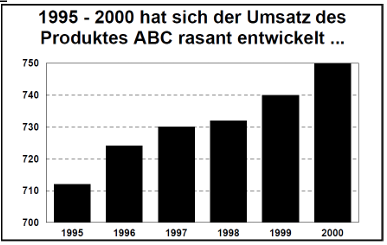
**Übungsblatt Modul 3**

**Häufigkeiten und Häufigkeitsverteilungen**

***Aufgabe 1***

*Nehmen Sie zum Schaubild (Abbildung) kritisch die Stellung.*

*Überlegen Sie was darf bei einer Tabelle und einem Diagramm auf keinen Fall fehlen?*



**Lösung:**

**Säulendiagramm**

* Es fehlt die Maßeinheit zu den Achsenbeschriftungen
* die Quellenangabe fehlt
* die absolute Umsatzentwicklung in 6 Jahren (von 1995 bis 2000) beträgt 38
* die relative Umsatzentwicklung in 6 Jahren (von 1995 bis 2000) beträgt 38 / 700 \* 100 = 5,4% = 0,9 % p. a. 🡆 Der Diagrammtitel übertreibt, die tatsächliche Umsatzentwicklung betrug in den 6 Jahren, auf welche sich die Auswertung bezieht, 5,4%, eine rasante Umsatzentwicklung kann man das nicht nennen
* Möglichkeit der Fehlinterpretation aufgrund der Täuschung durch die Skalierung der Ordinate (y-Achse), Die Skalierung der Ordinate täuscht einen höheren Anstieg vor, als das tatsächlich der Fall ist.  
  Skala der -Achse beginnt bei 700 und hat einen Abstand von 10 = 5 Schritte a 10  
  Dadurch wirkt die Umsatzentwicklung stärker, als sie tatsächlich ist.

***Aufgabe 2***

*Von 2006 bis 2007 ist der Umsatz eines Unternehmens um 50% gesunken.*

*Im darauffolgenden Jahr 2008 war das Unternehmen erfolgreicher: der Umsatz stieg von 2007 bis 2008 um 70%.*

*Welche der folgenden Aussagen ist richtig? (kurze Begründung durch Rechnung)*

*Der Umsatz ist im Zeitraum von 2006 bis 2008*

*a) um 10% gestiegen*

*b) um 15% gestiegen*

*c) um 15% gesunken*

**Lösung:**

Richtig ist Lösung c)

Der Umsatz ist um 15 % gesunken

**Wachstumsfaktor 2006/2007** = (100 – 50) / 100 = **0,50**

**Wachstumsfaktor 2007/2008** = (100 + 70) / 100 = **1,70**

**Wachstumsfaktor 2006/2008** = 0,50 \* 1,70 = **0,85**

**Das entspricht einer Wachstumsrate von -15%: (0,85 – 1) \* 100** = -15%

***Aufgabe 3***

*Im Rahmen einer Marktforschungsstudie wurden n = 12 Personen u.a. gefragt nach den drei Merkmalen Geschlecht G (w = weiblich, m = männlich), Alter A (Alter in Jahren) und Markenpräferenz M (A = Produkt A, B = Produkt B). Die Erhebung ergab die folgenden 12 Befragungsergebnisse (Beobachtungswertekombinationen):*

*(w, 37, A), (m, 65, A), (w, 26, A), (m, 37, B), (w, 21, B), (m, 29, A),*

*(w, 52, B), (m, 43, A), (w, 48, A), (m, 58, B), (w, 24, A), (m, 58, B).*

*Lesebeispiel:*

*Die 1. Person ist weiblich, 37 Jahre alt und bevorzugt Produkt A.*

*a)*

*Erstellen Sie die zwei folgenden zweidimensionalen Kreuztabellen:*

*Geschlecht x Markenpräferenz*

*klassiertes Alter x Markenpräferenz*

*(2 Altersklassen: 1. Klasse: bis unter 40 Jahre, 2. Klasse: 40 Jahre und älter)*

*b)*

*Versuchen Sie eine dreidimensionale Kreuztabelle zu erstellen für die drei Merkmale:*

*Geschlecht x klassiertes Alter x Markenpräferenz*

*c)*

*Welche der drei Merkmale kann man als unabhängige bzw. abhängige Merkmale betrachten?*

*d)*

*Sie erheben bei einer Grundgesamtheit Daten für 3 Merkmale.*

*Was liefert mehr Information: die dreidimensionale Häufigkeitsverteilung oder alle drei möglichen ein- und zweidimensionalen Häufigkeitsverteilungen zusammen?*

**a)**

**Erstellen Sie die zwei folgenden zweidimensionalen Kreuztabellen:**

* **Geschlecht x Markenpräferenz**
* **klassiertes Alter x Markenpräferenz**

**(2 Altersklassen: 1. Klasse: bis unter 40 Jahre, 2. Klasse: 40 Jahre und älter)**

**zweidimensionalen Kreuztabelle: Geschlecht x Markenpräferenz**

relative Spaltenhäufigkeit

Zellenwert / Spaltensumme

z. B. 3,00 / 6,00 = 0,50

Die Summe aller relativen Spaltenhäufigkeiten = 1

z. B. 0,50 + 0,50 = 1

relative Zeilenhäufigkeit

Zellenwert / Zeilensumme

z. B. 3,00 / 7,00 = 0,43

Die Summe aller relativen Zeilenhäufigkeiten = 1

z. B. 0,43 + 0,57 = 1

relative Häufigkeit der Merkmalsausprägungskombination

Zellenwert / Gesamtanzahl aller Ausprägungen

z. B. 3,00 / 12,00 = 0,25

Die Summe aller relativer Häufigkeiten zur Merkmalsausprägungs-kombination= 1

z. B. 0,25 + 0, 33 + 0,25 + 0,17

|  |  |  |  |
| --- | --- | --- | --- |
| Merkmal Markenpräferenz M | Merkmal  Geschlecht G | | |
| m | w | **Summe** |
| **Produkt A**  relative Spaltenhäufigkeit  relative Zeilenhäufigkeit  relative Häufigkeit der Merkmalsausprägungskombination | **3,00**  0,50  0,43  0,25 | **4,00**  0,67  0,57  0,33 | **7,00**  0,58 |
| **Produkt B**  relative Spaltenhäufigkeit  relative Zeilenhäufigkeit  relative Häufigkeit der Merkmalsausprägungskombination | **3,00**  0,50  0,60  0,25 | **2,00**  0,33  0,40  0,17 | **5,00**  0,42 |
| **Summe** | **6,00**  0,50 | **6,00**  0,50 | **12,00**  1,00 |

**zweidimensionalen Kreuztabelle: klassiertes Alter x Markenpräferenz**

2 Altersklassen:

* Klasse: bis unter 40 Jahre
* Klasse: 40 Jahre und älter)

|  |  |  |  |
| --- | --- | --- | --- |
| Merkmal Markenpräferenz M | Merkmal  Klassiertes Alter A | | |
| b. u. 40 | 40+ | Summe |
| **Produkt A**  relative Spaltenhäufigkeit  relative Zeilenhäufigkeit  relative Häufigkeit der Merkmalsausprägungskombination | **4,00**  0,67  0,57  0,33 | **3,00**  0,50  0,43  0,25 | **7,00**  0,58 |
| **Produkt B**  relative Spaltenhäufigkeit  relative Zeilenhäufigkeit  relative Häufigkeit der Merkmalsausprägungskombination | **2,00**  0,33  0,40  0,17 | **3,00**  0,50  0,60  0,25 | **5,00**  0,42 |
| Summe | **6,00**  0,50 | **6,00**  0,50 | **12,00**  1,00 |

**WICHTIG:**

* **Summenzeile und Summenspalte nicht vergessen!**

**b)**

**Versuchen Sie eine dreidimensionale Kreuztabelle zu erstellen für die drei Merkmale:**

**Geschlecht x klassiertes Alter x Markenpräferenz**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Merkmal Markenpräferenz M | Geschlecht m | | Geschlecht w | | Summe |
| klass. Alter  b. u. 40 | klass. Alter  40+ | klass. Alter  b. u. 40 | klass. Alter  40+ |  |
| **Produkt A**  relative Spaltenhäufigkeit  relative Zeilenhäufigkeit  relative Häufigkeit der Merkmalsausprägungskombination | **1**  0,50  0,14  0,08 | **2**  0,50  0,29  0,17 | **3**  0,75  0,43  0,25 | **1**  0,50  0,14  0,08 | **7,00**  0,58 |
| **Produkt B**  relative Spaltenhäufigkeit  relative Zeilenhäufigkeit  relative Häufigkeit der Merkmalsausprägungskombination | **1**  0,50  0,20  0,08 | **2**  0,50  0,40  0,17 | **1**  0,25  0,20  0,08 | **1**  0,50  0,20  0,08 | **5,00**  0,42 |
| Summe | **2,00**  0,17 | **4,00**  0,33 | **4,00**  0,33 | **2,00**  0,17 | **12,00**  1,00 |

**c)**

**Welche der drei Merkmale kann man als unabhängige bzw. abhängige Merkmale betrachten?**

Die **Merkmale Alter und Geschlecht sind unabhängige Merkmale (Regressor), da diese Merkmale die Markenpräferenz (Wahl des Produktes A oder B) beeinflussen** können.

Die Markenpräferenz (Produktwahl) ist somit das vom Alter bzw. Geschlecht abhängige Merkmal. Das **Merkmal Markenpräferenz ist also die abhängige Variable** (Regressand).

Es wird beobachtet, ob dieses Merkmal durch andere (sog. unabhängige) Merkmale beeinflusst wird.

Die Markenwahl (abhängiges Merkmal) kann die unabhängigen Merkmale Alter und Geschlecht nicht beeinflussen.

Es ist sinnvoll zu untersuchen, ob die Markenwahl bei Frauen und Männern unterschiedlich oder gleich ist bzw. ob die Markenwahl bei einer Produktgruppe altersabhängig ist. Denn es ist möglich, dass Frauen und Männer ein unterschiedliches Kaufverhalten haben. Genauso ist möglich bzw. sogar wahrscheinlich, dass jüngere Konsumenten andere Marken wählen als ältere Konsumenten

Regression:

Aufteilung einer Variablen in einen systematischen und einen zufälligen Teil zur angenäherten Beschreibung einer Variablen als Funktion anderer.

**Merke:**

**Unabhängige Merkmale (Variablen) (hier Alter und Geschlecht) beeinflussen die abhängigen Variablen (hier die Markenpräferenz).**

**Abhängige Merkmale werden von unabhängigen Merkmalen beeinflusst.**

**d)**

**Sie erheben bei einer Grundgesamtheit Daten für 3 Merkmale.**

**Was liefert mehr Information: die dreidimensionale Häufigkeitsverteilung oder alle drei möglichen ein- und zweidimensionalen Häufigkeitsverteilungen zusammen?**

In der dreidimensionalen Häufigkeitsverteilung werden die Informationen zur Markenpräferenz dediziert nach Geschlecht und Alter (männlich bis unter 40 Jahre, männlich 40+, weiblich bis unter 40 Jahre, weiblich 40+) aufgezeigt. Die ein- und zweidimensionale Häufigkeitsverteilung enthält diese detaillierte Aufschlüsselung nicht. Die ein- und zweidimensionale Häufigkeitsverteilung ist mit einem Informationsverlust behaftet.

**Die dreidimensionale Häufigkeitsverteilung liefert also mehr Informationen als alle ein- und zweidimensionale Häufigkeitsverteilung zusammen.**

**Nur aus der dreidimensionalen Häufigkeitsverteilung kann man erfahren, welchen Zusammenhang es zwischen allen drei Merkmalen gibt.**

***Aufgabe 4***

*Ein Merkmal X mit m Merkmalsausprägungen (x1, …, xm) wird bei n Untersuchungseinheiten gemessen.*

*Beantworten Sie für die entsprechende Häufigkeitsverteilung die folgenden Fragen:*



absolute Häufigkeit: Anzahl des Auftretens einer bestimmten Merkmalsausprägung

** auf den Spicker!**

**bedeutet: Summe der absoluten Häufigkeiten = Anzahl n der Stichproben (Summe der Einzelhäufigkeiten)**



relative Häufigkeit: Verhältnis der absoluten Häufigkeit zur Summe der Einzelhäufigkeiten (aller Häufigkeiten)

** auf den Spicker!**

**bedeutet: Summe der relativen Häufigkeiten = 1 oder 100% (je nach Aufgabenstellung**



absolute kumulierte Häufigkeit

**H(xm) = h(x1) + h(x2) + … h(xn) = n auf den Spicker!**

**bedeutet: kumulierte absolute Häufigkeiten = absolute Häufigkeit1 + absolute Häufigkeit2 + … = Anzahl der Stichproben**

relative kumulierte Häufigkeit

**F(xm) = f(x1) + f(x2) + … f(xn) = 1 (100%) auf den Spicker!**

**bedeutet: kumulierte relative Häufigkeiten = relative Häufigkeit1 + relative Häufigkeit2 + … = 1**

***Aufgabe 5***

*Vervollständigen Sie die folgende Häufigkeitstabelle (relative Häufigkeiten sind ganze Zahlen!)*

*Die kursiv und grün formatierte Zahlen sind die Vorgaben in der Aufgabenstellung.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| i | Merkmalsausprägung  xi | h(xi) | f(xi) (%) | H(xi) | F(xi) (%) |
| 1 | *3* | 240 | *15%* | 240 | 15% |
| 2 | *4* | 400 | 25% | 640 | *40%* |
| 3 | *5* | 320 | 20% | *960* | 60% |
| 4 | *6* | 640 | 40% | 1.600 | 100% |
| Σ |  | *1.600* | **100%-** | - | - |

h(xi) absolute Häufigkeit

f(xi) (%) relative Häufigkeit hi / n

**sinnvoll nur für Rangmerkmale absolute und metrische Merkmale**

H(xi) absolute Summenhäufigkeit (absolute kumulierte Häufigkeit)

F(xi) (%) relative Summenhäufigkeit (relative kumulierte Häufigkeit)

relative Häufigkeit fi %

Klasse Nr.

i

1 240 / 1600 \* 100 = 15

2 400 / 1600 \* 100 = 25

3 320 / 1600 \* 100 = 20

4 640 / 1600 \* 100 = 40

absolute Summenhäufigkeit Hi

1 240 = 240

2 240 + 400 = 640

3 240 + 400 + 320 = 960

4 240 + 400 + 320 + 640 = 1600

relative Summenhäufigkeit Fi

1 15 = 15

2 15 + 25 = 40

3 15 + 25 + 20 = 60

4 15 + 25 + 20 + 40 = 100

***Aufgabe 6***

*Vervollständigen Sie die folgende klassierte Häufigkeitstabelle (relative Häufigkeiten mit einer Nachkommastelle!)*

*Die kursiv und grün formatierte Zahlen sind die Vorgaben in der Aufgabenstellung.*

**WICHTIG:**

**Auf richtigen Wert hinter dem Komma runden!**

**Bei relativen Werten, die als %-Werte dargestellt werden sollen, immer das %-Zeichen hinter dem Wert schreiben, auch wenn in der Spalte % angegeben ist.**

**Bei ganzzahligen Werten (wie z. B. 50,0%) immer die geforderten Nachkommastellen (also hier ein 0) schreiben!**

**Wenn es nicht auf 100% aufgeht (z. B. 100,1) dann auch die 1 hinter dem Komma schreiben!**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Klasse Nr.  i | Umsatzklasse (Mio €) | Anzahl Filialen hi | Anteil fi  (%) | Hi | Fi (%) |
| 1 | 0 b.u. 40 | *50* | 16,7% | 50 | 16,7% |
| 2 | 40 b.u. 100 | *150* | 50,0% | 200 | 66,7% |
| 3 | 100 b.u. 200 | *50* | 16,7% | 250 | 83,4% |
| 4 | 200 b.u. 400 | *50* | 16,7% | 300 | 100,1% |
| Σ |  | *300* | 100,1% |  |  |

hi absolute Häufigkeit

fi relative Häufigkeit hi / n \* 100

**sinnvoll nur für Rangmerkmale absolute und metrische Merkmale**

Hi absolute Summenhäufigkeit (absolute kumulierte Häufigkeit)

H(x1) = h(x1)

H(x2) = h(x1)+h(x2

H(x3) = h(x1)+h(x2)+h(x3)

**H(xi) = h(x1)+h(x2)+ ... +h(xi) = n**

Die absolute Summenhäufigkeit der letzten Zeile = der Summe aller Häufigkeiten ( = Anzahl n der Stichproben), i. Bsp. 300

Fi relative Summenhäufigkeit (relative kumulierte Häufigkeit)

F(x1) = f(x1)

F(x2) = f(x1)+f(x2)

F(x3) = f(x1)+f(x2)+f(x3)

**F(xi) = f(x1)+f(x2)+ ... +f(xi) = 1 (100%)**

Die relative Summenhäufigkeit der letzten Zeile = 100 oder 1 (je nachdem, ob Dezimalwerte oder Prozentwerte für die relative Häufigkeit verwendet wurden)

relative Häufigkeit fi

Klasse Nr.

i

1 50 / 300 \* 100 = 16,7

2 150 / 300 \* 100 = 50

3 50 / 300 \* 100 = 16,7

4 50 / 300 \* 100 = 16,7

absolute Summenhäufigkeit Hi

1 50 = 50

2 16,7 + 50 = 66,7

3 66,7 + 16,7 = 83,4

4 83,4 + 16,7 = 100,1 gerundet 100

relative Summenhäufigkeit Fi

1 16,7 = 16,7

2 30 + 35 = 65

3 30 + 35 + 20 = 85

4 30 + 35 + 20 + 15 = 100

***Aufgabe 7***

*Im Rahmen einer Bürgerbefragung wurden n = 1.200 Personen u.a. gefragt nach ihrem Alter (A) und ihrer Meinung zur Ausweitung der Fußgängerzone in der Innenstadt (F).*

*Das Merkmal Alter wurde wie folgt klassiert:*

*„bis unter 40 Jahre“ und „40 Jahre und älter“.*

*Auf die Frage nach der Ausweitung der Fußgängerzone konnte nur mit „ja“ oder „nein“ geantwortet werden.*

*Die Befragungsergebnisse wurden ausgezählt.*

*Es ergaben sich die in der folgenden Kreuztabelle aufgeführten absoluten Häufigkeiten für die Merkmalsausprägungskombinationen.*

**WICHTIG:**

**Auf richtigen Wert hinter dem Komma runden!**

**Bei relativen Werten, die als %-Werte dargestellt werden sollen, immer das %-Zeichen hinter dem Wert schreiben, auch wenn in der Spalte % angegeben ist.**

**Bei ganzzahligen Werten (wie z. B. 50,0%) immer die geforderten Nachkommastellen (also hier ein 0) schreiben!**

**Wenn es nicht auf 100% aufgeht (z. B. 100,1) dann auch die 1 hinter dem Komma schreiben!**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | F / A ⇒  ⇓ | 40-  bis unter 40 Jahre alt | 40+  40 Jahre und älter | Σ |
| Ausweitung der  Fußgängerzone | ja | *396*  (2) 60,0%  (3) 66,0%  (4) 33,0% | *204*  (2) 37,8%  (3) 34,0%  (4) 17,0% | (1)  600  50,0% |
| nein | *264*  (2) 40,0%  (3) 44,0%  (4) 22,0% | *336*  (2) 62,2%  (3) 56,0%  (4) 28,0% | (1)  600  50,0% |
|  | Σ | 660  (1) 55% | 540  (1) 45% | *1.200*  *100%* |

a) Tragen Sie in die obige Kreuztabelle ein:

(1) die beiden Randverteilungen (absolut und relativ in % mit 1 Nachkommastelle),

(2) die relativen Spaltenhäufigkeiten in % (1 Nachkommastelle),

(3) die relativen Zeilenhäufigkeiten in % (1 Nachkommastelle),

(4) die relativen Häufigkeiten der Merkmalsausprägungskombinationen in % (1 Nachkommastelle).

b)

Beantworten Sie mit Hilfe der obigen Kreuztabelle die folgenden Fragen:

(1) Wie viel % der Befragten, die die Ausweitung der Fußgängerzone befürworten, sind 40 Jahre und älter? ...34..................%.

🡆 Wert aus Zelle „40 Jahre und älter“/“ja“ : durch Zeilensumme aus Zeile „Ja“ \* 100  
204 / 600 \* 100 = 34%

(2) Wie viel % der jüngeren Befragten (bis unter 40 Jahre), befürworten die Ausweitung der Fußgängerzone? .........60............%.

🡆 Wert aus Zelle „bis unter 40 Jahre“/“ja“ : durch Spaltensumme aus Spalte „bis unter 40 Jahre“ \* 100  
396 / 660 \* 100 = 60%

(3) Wie viel % der Befragten sind bis unter 40 Jahre alt und befürworten die Ausweitung der Fußgängerzone? ..........33...........%.

🡆 Wert aus Zelle „bis unter 40 Jahre“/“ja“ : Gesamtanzahl der Einzelausprägungen \* 100

396 / 1200 \* 100 = 60%

(4) Wie viel % der Befragten, die 40 Jahre und älter sind, sind gegen die Ausweitung der Fußgängerzone? .....62,2................%.

🡆 Wert aus Zelle „40 Jahre und älter“/“nein“ : durch Spaltensumme aus Spalte „40 Jahre und älter“ \* 100  
336 / 540 \* 100 = 62,2%

***Aufgabe 8***

*Im Rahmen einer Marktforschungsstudie wurden n = 2.000 Personen u. a. gefragt nach den drei Merkmalen*

*Geschlecht G,*

*Alter A (klassiert in 2 Klassen: "b.u. 40 Jahre" und "40 Jahre und älter") und*

*Markenpräferenz M (A = Produkt A, B = Produkt B).*

*Die Befragungsergebnisse sind in der folgenden dreidimensionalen Häufigkeitstabelle zusammengefasst:*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | weiblich | | männlich | |  |
| M / G A | bis unter 40 Jahre alt | 40 Jahre und älter | bis unter 40 Jahre alt | 40 Jahre und älter | Σ |
| Produkt A | 400 | 150 | 200 | 450 | 1200 |
| Produkt B | 100 | 450 | 200 | 50 | 800 |
| Σ | 500 | 600 | 400 | 500 | 2000 |

Beantworten Sie auf der Basis der obigen Tabelle die folgenden Fragen (mit 1 Nachkommastelle):

a) Wie viel % der jüngeren Frauen (bis unter 40 Jahren) bevorzugen Produkt B?

20%

100 / 500 \* 100 = 20%

b) Wie viel % der älteren Männer (40 Jahre und älter) bevorzugen Produkt A?

90%

450 / 500 \* 100 = 90%

c) Wie viel % der Frauen bevorzugen Produkt A?

50,0%

550 / 1100 \* 100 = 50%

d) Wie viel % der A-Käufer sind Frauen?

45,8%

550 / 1200 \* 100 = 45,8%

e) Wie viel % der Befragten sind Frauen, bis unter 40 Jahre alt und A-Käuferinnen?

20,0%

400 / 2000 \* 100 = 20,0%

f) Wie viel % der Frauen, die 40 Jahre und älter sind, bevorzugen Produkt A?

25,0%

150 / 600 \* 100 = 25,0%

g) Wie viel % der Befragten sind Männer?

45,0%

900 / 2000 \* 100 = 45,0%

h) Welchen Wert hat f(B | b.u. 40 J.) 🡆 bedeutet Anteil der u. 40 zu Produkt B

33,3%

300 / 900 \* 100% = 33,3%

300 ist Summe aus Zeile „Produkt B“ und den beiden Spalten „bis unter 40“ = 100 + 200 = 300

900 ist die Spaltensumme aus beiden Spalten „bis unter 40“ = 500 + 400 = 900

i) Wie viel Prozent der A-Käufer, die 40 Jahre und älter sind, sind Frauen?

25,0%

150 / 600 \* 100 = 25,0% 600 ist die summe aus 150 + 450

j) Wie viel Prozent der Befragten sind A-Käufer, 40 Jahre und älter und Männer?

22,5%

450 / 2000 \* 100 = 22,5%

***Aufgabe 9***

*Für die 20 Unternehmen einer Branche wurden im Jahr 2005 die folgenden Umsätze (in Mio. €) ermittelt:*

*35, 150, 190, 20, 8, 74, 44, 89, 25, 12, 17, 5, 22, 10, 13, 150, 47, 65, 49, 55*

*Erstellen Sie für diese statistische Reihe in der folgenden Tabelle eine klassierte Häufigkeitsverteilung für die gegebenen Klassen*

*(mit absoluten und relativen (in % mit 1 Nachkommastelle) Häufigkeiten und Summenhäufigkeiten).*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Klasse Nr.  i | Umsatzklasse (Mio €) | Anzahl Filialen hi | Anteil fi  (%) | Hi | Fi (%) |
| 1 | 0 b.u. 20 | 6 | 30,0% | 6 | 30,0% |
| 2 | 20 b.u. 50 | 7 | 35,0% | 13 | 65,0% |
| 3 | 50 b.u. 100 | 4 | 20,0% | 17 | 85,0% |
| 4 | 100 b.u. 200 | 3 | 15,0% | 20 | 100,0% |
| Σ |  | 20 | 100,0% | - | - |

hi absolute Häufigkeit

fi relative Häufigkeit hi / n \* 100

**sinnvoll nur für Rangmerkmale absolute und metrische Merkmale**

Hi absolute Summenhäufigkeit (absolute kumulierte Häufigkeit)

H(x1) = h(x1)

H(x2) = h(x1)+h(x2

H(x3) = h(x1)+h(x2)+h(x3)

**H(xi) = h(x1)+h(x2)+ ... +h(xi) = n**

Fi relative Summenhäufigkeit (relative kumulierte Häufigkeit)

F(x1) = f(x1)

F(x2) = f(x1)+f(x2)

F(x3) = f(x1)+f(x2)+f(x3)

**F(xi) = f(x1)+f(x2)+ ... +f(xi) = 1 (100%)**

relative Häufigkeit fi

Klasse Nr.

i

1 6 / 20 \* 100 = 30

2 7 / 20 \* 100 = 35

3 4 / 20 \* 100 = 20

4 3 / 20 \* 100 = 15

absolute Summenhäufigkeit Hi

1 6 = 6

2 6 + 7 = 13

3 6 + 7 + 4 = 17

4 6 + 7 + 4 + 3 = 20

relative Summenhäufigkeit Fi

1 30 = 30

2 30 + 35 = 65

3 30 + 35 + 20 = 85

4 30 + 35 + 20 + 15 = 100