

### Aufgabe 1

Im Rahmen der Evaluierung der Programme eines Softwarepaketes werden folgende Merkmale erfasst:

- (1) Laufzeit
- (2) Anzahl der Änderungen seit dem letzten Release
- (3) ID des Programms (durchnumeriert nach alphabetischer Sortierung des Programmnamens)
- (4) Komplexität (bewertet auf einer 4-Punkte-Skala: 0=einfach, 1=mittel, 2=komplex, 3=hochkomplex)
- (5) Alter des Programms (in Tagen seit Erstspeicherung)
- (6) Verwendete Programmiersprache (1=C/C++, 2=Java, 3=Python, 4=andere)

Klassifizieren sie diese Merkmale als nominal/ordinal/intervallskaliert/absolutskaliert bzw. stetig/diskret.

### Aufgabe 2

Bei welchen Merkmalstypen (nominal/ordinal/intervallskaliert/absolutskaliert) ergibt Ihrer Meinung nach die Berechnung des arithmetischen Mittels Sinn bzw. keinen Sinn?

### Aufgabe 3

11 Studierende werden danach befragt, wie viele Stunden sie in der letzten Woche mit Literaturrecherche in der Bibliothek verbracht haben. Die relativen Häufigkeiten (auf zwei Nachkommastellen gerundet) stellen sich wie folgt dar:

Stunden ( $y_i$ )	relative Häufigkeit ( $f_i$ )
1	0.18
2	0.09
3	0.27
4	0.45

Erstellen Sie eine mögliche Urliste, auf der diese Häufigkeitstabelle beruhen könnte.

### Aufgabe 4

Eine Studie zur Untersuchung der Internetnutzung sei an  $n$  Personen durchgeführt worden. 35% der Befragten gehörten zur Altersklasse von unter 20 Jahren, 40% zur Altersklasse von 20 bis unter 50 Jahren und 25% zur Alterklasse von 50 Jahren oder darüber. In diesen Altersklassen (aufsteigend sortiert) werde das Internet von 80%, 70% bzw. 30% regelmäßig genutzt. Wie hoch ist (in Abhängigkeit von  $n$ ) die absolute Häufigkeit der regelmäßigen Internetnutzer in allen Altersklassen zusammengefasst? Wie hoch ist die zugehörige relative Häufigkeit? Unter welchen Voraussetzungen vermittelt Ihnen das Ergebnis einen guten Eindruck von der Häufigkeit der regelmäßigen Internetnutzer in der Gesamtbevölkerung?

### Aufgabe 5

631 Jugendliche im Alter zwischen 12 und 19 Jahren wurden befragt, in welchem Alter sie sich zum ersten Mal verliebt haben (Merkmal  $X$ ). Die Ergebnisse sind in der nachfolgenden Tabelle aufgeführt.

Alter in Jahren	Absolute Häufigkeit ( $h_i$ )
1.0	3
3.0	5
4.0	16
5.0	20
6.0	32
7.0	32
8.0	58
9.0	36
10.0	70
11.0	61
11.5	1
12.0	116
13.0	72
13.5	1
14.0	59
15.0	28
16.0	11
17.0	4
18.0	3
19.0	3

Wie viele Ausprägungen hat das Merkmal  $X$ ? Führen Sie für den Datensatz folgende Untersuchungen durch:

- Zeichnen Sie die empirische Verteilungsfunktion im Bereich von 0 bis 10 Jahren.
- Beantworten Sie folgende Fragen:
  - (1) Wie viele der Befragten (absolut und relativ) waren beim ersten Verliebtsein jünger als 7 bzw. älter als 13 Jahre?
  - (2) In welchem Alter waren 60% der Befragten schon mindestens einmal verliebt?

Angenommen, die obige Tabelle sei eine “bereinigte Darstellung” der Umfrage, in dem Sinn, daß ursprünglich 648 Jugendliche im Alter zwischen 12 und 19 Jahren befragt worden waren, von denen aber 17 die Antwort “Ich war bis jetzt noch nie verliebt” gegeben haben. Wie ändern sich Ihre Ergebnisse zu den obigen 2 Fragen, wenn sie diese 17 Personen in die Auswertung mit einbeziehen?