
Vorlesung: Programmieren in Anwendungen
Übungsblatt 3

http://softech.informatik.uni-kl.de/Homepage/PIA_SS13

Installation von R

Auf der Website <http://www.r-project.org> erhalten Sie Information und Material zur Statistiksoftware R.

1. Laden Sie die neueste Version für Ihr Betriebssystem herunter und installieren Sie es.
2. Installieren Sie die Library ggplot2 inklusive aller abhängigen Pakete.

Geburtenstatistik

1. Laden Sie den vollen Datensatz aus dem Internet von <http://greenteapress.com/thinkstats/nsfg.html>.
2. Eine Übersicht der einzelnen Felder der Datensätze finden Sie unter <http://www.icpsr.umich.edu/nsfg6/>, z.B. Female Pregnancy File Codebook → Section B Pregnancy-Based Recodes.
3. Beantworten Sie an Hand der Daten die folgenden Fragen:
 - Wie schwer waren die Babys bei der Geburt?
 - Wie alt waren die Mütter zum Zeitpunkt der Geburt?
 - Nach wie vielen Schwangerschaftswochen wurde die Schwangerschaft festgestellt?
 - Wie lange wurden die Babys gestillt?
 - Wie änderten sich diese Kenngrößen bei späteren Schwangerschaften?
4. Visualisieren Sie die Informationen zum Familienstatus zu Beginn und am Ende der Schwangerschaft durch Balken- und Kreisdiagramme. Fügen Sie passende Achsenbeschriftungen, Titel, etc. hinzu.

Programmieren in R

1. Welchen Datentyp haben die Einträge des Vektors `c(7, 5+9i, 8.9)`?
2. Welchen Datentyp haben die Einträge des Vektors `c("9", "8.9")`?
3. Welchen Datentyp haben die Einträge des Vektors `c("9", FALSE, 9)`?
4. Erstellen Sie einen Vektor, der alle Monatsnamen enthält!
5. Erstellen Sie einen Vektor mit allen ungeraden Zahlen von 1 bis 1000! Wie viele Elemente enthält dieser Vektor? Addieren Sie zu jedem Element des Vektors den Wert 5!
6. Wie werden Rechenoperationen auf Vektoren durchgeführt, wenn die Vektoren unterschiedliche Länge haben?
7. Erzeugen Sie einen Vektor mit den Einträgen "A-1" bis "A-200"!
8. Wie unterscheiden sich die Funktionen `apply()`, `tapply()`, `sapply()` und `lapply()`?

Bevölkerungsstatistiken

Unter <http://opendata.service-bw.de> finden sich verschiedene Datensammlungen des Bundeslandes Baden-Württemberg.

1. Laden Sie den Datensatz von 2010 zur Bevölkerungsstatistik herunter.
2. Wie sind die Daten strukturiert?
3. Generieren Sie visuelle Übersichten zur Bevölkerungsverteilung (nach Geschlecht und Alter) von Stuttgart, Gundelfingen und Pfullendorf.