

Information Mining - Wintersemester 2014

Dipl.-Inform. Vu Tran, LF 139

vtran@is.inf.uni-due.de

Übungsblatt 8

Aufgabe 16: R und Rkward

R^1 ist ein Werkzeug für statistische Analysen und Grafiken. Eine grafische Benutzeroberfläche für R ist $Rkward^2$. Installiere R und $Rkward$ und mache dich mit den Anwendungen vertraut. Installationshinweise und ein ausführliches Handbuch sind auf den angegebenen Webseiten zu finden.

Aufgabe 17: Beschreibende Statistik mit R

Berechne aus den Beispieldaten³ die folgenden Werte:

- (a) Median des Alters
- (b) Arithmetisches Mittel des Jahreseinkommens
- (c) Anzahl der Verheirateten

Zeichne anschließend aus der Spalte mit dem Alter der Kunden ein Histogramm mit 6 Intervallen.

Aufgabe 18: Signifikanztest und Korrelation mit R

Benutze die Beispieldaten aus Aufgabe 17.

- (a) Gibt es einen signifikanten Unterschied zwischen dem Einkommen der Leute, die geantwortet haben und denen, die nicht geantwortet haben? Benutze dafür den t-Test.
- (b) Gibt es eine Korrelation zwischen dem Alter und dem Einkommen? Führe eine Korrelationsanalyse nach Spearman durch.

(Bitte umblättern)

¹<http://www.r-project.org/>

²http://sourceforge.net/apps/mediawiki/rkward/index.php?title=Main_Page

³http://www.is.inf.uni-due.de/courses/im_ws14/uebung/data_a17.csv

Aufgabe 19: Lineare Regression mit R

Die Beispieldaten⁴ enthalten die Anzahl der Studienanfänger der letzten Jahre.

- (a) Erzeuge ein Streudiagramm aus den Daten
- (b) Ermittle die Regressionsgerade und zeichne sie ins Streudiagramm ein
- (c) Die lineare Regression kann zur Prognose benutzt werden. Wieviele Studienanfänger werden im Jahr 2014 erwartet?
- (d) Gibt es eine Korrelation zwischen dem Jahr und der Anzahl der Studienanfänger? Führe eine Korrelationsanalyse nach Spearman durch.

⁴http://www.is.inf.uni-due.de/courses/im_ws14/uebung/data_r.csv