

Fortgeschrittene Datenanalyse mit R

Friedrich Leisch

1. Übung, 9.5.2006

1. Legen Sie ein Verzeichnis für eine persönliche R Bibliothek und ein Arbeitsverzeichnis für diese LVA an. Erstellen Sie eine R-Verknüpfung am Desktop („R Icon“), das R im neuen Arbeitsverzeichnis startet und zusätzlich zu den systemweiten Bibliotheken auch Ihre private Bibliothek kennt.
Tests: `.libPaths()` und `getwd()` am R Prompt.
Installieren Sie ein Paket Ihrer Wahl von CRAN in die private Bibliothek (Default falls an erster stelle in `.libPaths()`).
2. Legen Sie in Excel irgendeinen Datensatz im “Wide“-Format an, importieren Sie die Daten nach R und formen Sie den Datensatz ins „Long“-Format um.
3. Lösen Sie das Problem der Umformung von Namen aus dem Format „**Vorname Nachname**“ ins Format „**Nachname, Vorname**“ mittels regulärer Ausdrücke.
4. Ersetzen Sie alle Umlaute in den Namen durch normale Zeichen, d.h., aus „ä“ wird „ae“ etc.
5. Wiederholen Sie die Analyse des CI-TCP-Dumps, und „spielen“ Sie ein wenig mit den regulären Ausdrücken.
6. Erstellen Sie eine Tabelle der CRAN-Klienten mit den meisten TCP/IP-Paketen.
7. Fassen Sie für die drei Klienten mit den meisten Paketen diese zeitlich geeignet zusammen (etwa pro Sekunde) und zeichnen Sie den Zeitverlauf der Verbindung.