

Kursnummer: DLBDBSC01	Kursname: Statistical Computing	Gesamtstunden: 150 h ECTS Punkte: 5 ECTS
Kurstyp: Pflicht Kursangebot : WS, SS Course Duration : Minimaldauer 1 Semester		Zugangsvoraussetzungen: keine
Kurskoordinator(en) / Dozenten / Lektoren: Siehe aktuelle Liste der Tutoren im Learning Management System		Bezüge zu anderen Modulen: Siehe Modulbeschreibung
<p>Beschreibung des Kurses:</p> <p>Das Statistical Computing verbindet die Konzepte und Methoden der Statistik mit den Werkzeugen aus der Informatik. Das Ergebnis sind unter anderem Statistik-Programme und -Programmiersprachen, die viele nützliche Funktionen zur Analyse von digital verfügbaren Datenquellen bieten. In diesem Kurs wird den Studierenden die Programmiersprache R vermittelt, um damit anschließend statistische Verfahren (z. B. Regressionsanalyse, Varianzanalyse) anwenden zu können. Im Rahmen einer Fallstudie sollen die erworbenen Kompetenzen eingesetzt werden, um Zusammenhänge aus komplexen Datenquellen zu gewinnen und graphisch darzustellen.</p> <p>Kursziele:</p> <p>Nach erfolgreichem Abschluss des Kurses sind die Studierenden in der Lage,</p> <ul style="list-style-type: none"> • den Begriff Statistical Computing einzuordnen und abzugrenzen. • sich eine PC-Arbeitsumgebung zu schaffen, mit der Aufgaben aus dem Themengebiet Statistical Computing bearbeitet werden können. • einfache Programme mit der Programmiersprache R zu schreiben. • mit R Daten zu importieren und zu exportieren. • mit R verschiedene statistische Verfahren anzuwenden, von der deskriptiven Statistik über die Inferenzstatistik bis hin zur Varianz- und Regressionsanalyse. <p>Lehrmethoden:</p> <p>Die Lehrmaterialien enthalten Skripte, Video-Vorlesungen, Übungen, Podcasts, (Online-) Tutorien und Fallstudien. Sie sind so strukturiert, dass Studierende sie in freier Ortswahl und zeitlich unabhängig bearbeiten können.</p> <p>Inhalte des Kurses:</p> <p>1. Einstieg in das Statistical Computing</p> <p>1.1 Begriffsbestimmung und Abgrenzung</p> <p>1.2 Statistik-Programm vs. Statistik-Programmiersprache</p> <p>1.3 Einrichtung der Arbeitsumgebung</p> <p>2. Grundlagen der Programmierung mit R</p> <p>2.1 R als Taschenrechner</p> <p>2.2 Zuweisungen und Variablen</p> <p>2.3 Vektoren und Matrizen</p> <p>2.4 Logik</p>		

2.5 Funktionen

2.6 Datentypen und Datenstrukturen

3. Auf Daten zugreifen

3.1 Daten eingeben

3.2 Import und Export von externen Dateien

3.3 Datenmanagement in R

4. Deskriptive Statistik

4.1 Univariate deskriptive Statistik

4.2 Bivariate deskriptive Statistik

5. Inferenzstatistik

5.1 Verteilungen

5.2 Stichproben

5.3 t-Tests

6. Varianzanalyse

6.1 Grundlagen und Abgrenzung zum t-Test

6.2 Einfaktorielle Varianzanalyse

6.3 Zweifaktorielle Varianzanalyse

7. Regressionsanalyse

7.1 Korrelation

7.2 Lineare Regression

7.3 Weitere Modelle und Verfahren

Literatur:

- Ligges, U. (2008): Programmieren mit R. 3. Auflage, Springer, Heidelberg.
- Luhmann, M. (2015): R für Einsteiger. Einführung in die Statistiksoftware für die Sozialwissenschaften. Beltz, Weinheim, Basel.
- Toomey, D. (2017): Jupyter for Data Science. Exploratory analysis, statistical modeling, machine learning, and data visualization with Jupyter. Packt Publishing, Birmingham, UK.
- Vanderplas, J. (2017): Data Science mit Python. Das Handbuch für den Einsatz von Ipython, Jupyter, NumPy, Pandas, Matplotlib, Scikit-Learn. mitp, Frechen.
- Wollschläger, D. (2015): Grundlagen der Datenanalyse mit R. Eine anwendungsorientierte Einführung. Springer Spektrum, Berlin Heidelberg.

Prüfungszugangsvoraussetzung:

- Kursabhängig: Begleitende Online-Lernkontrolle (max. 15 Minuten je Lektion, bestanden / nicht bestanden)

Prüfungsleistung:

Schriftliche Ausarbeitung: Fallstudie

Zeitaufwand Studierende (in Std.): 150

Selbststudium (in Std.): 110

Selbstüberprüfung (in Std.): 20

Tutorien (in Std.): 20