

**FAU – Fachbereich Wirtschafts- und Sozialwissenschaften**  
**Lehrstuhl für Statistik und Ökonometrie**  
**Statistik (RUW-2172/2170)**  
**Wintersemester 2020/21 – Syllabus**

**Bitte beachten:** Wegen der Corona-Pandemie findet dieses Modul im Wintersemester 2020/21 im Onlineformat statt. Für aktuelle Informationen bitte die Kommunikation über StudOn beachten!

**Vorlesung:** Aufzeichnungen der Vorlesung stehen in Form von kurzen Videoclips auf [StudOn](#) zur Verfügung.

**Übung:** Findet online über Zoom statt. Montag, 15:00 – 16:30; Dienstag, 9:45 – 11:15; Mittwoch, 11:30 – 13:00. Die Übungen starten in der ersten Vorlesungswoche. Zugangsdaten finden Sie auf [StudOn](#).

**Tutorium:** Findet online über Zoom statt. Donnerstag, 16:45 – 18:15; Freitag, 9:45 – 11:15.

Die Tutorien starten in der ersten Vorlesungswoche. Zugangsdaten finden Sie auf [StudOn](#).

**R-Übung:** Die Unterlagen für die R-Übung (inkl. Musterlösungen und kurzen Videoclips) finden Sie auf [StudOn](#).

**R-Sprechstunde:** Von studentischen Hilfskräften betreute Sprechstunde für Fragen zu den R-Aufgaben. Findet online über Zoom statt. Mittwoch, 9:45 – 11:15; Mittwoch, 13:15 – 14:45; Freitag, 13:15 – 14:45. Die R-Tutorien starten in der ersten Semesterwoche. Zugangsdaten finden Sie auf [StudOn](#).

**Dozent:**

Jonas Dovern

Raum: LG 4.169

E-Mail: [jonas.dovern@fau.de](mailto:jonas.dovern@fau.de)

Sprechstunde: nach Vereinbarung

**Koordinator für Übungen/Tutorien**

Daniel Perico

Raum: LG 4.172

E-Mail: [daniel.perico@fau.de](mailto:daniel.perico@fau.de)

Sprechstunde: nach Vereinbarung

**Koordinator für R-Übung/R-Tutorien**

Alexander Glas

Raum: LG 4.173

E-Mail: [alexander.glas@fau.de](mailto:alexander.glas@fau.de)

R-Sprechstunde: Donnerstag, 13:00 – 14:00

**Kurzbeschreibung:** Statistik und Wahrscheinlichkeitsrechnung verfolgen uns tagtäglich. Hilft ein bestimmtes Medikament gegen Kopfschmerzen? Wie lange warte ich im Schnitt auf die U-Bahn? Können Frauen besser Autofahren als Männer? Und natürlich sind statistische Verfahren auch nicht aus der Welt der BWL, VWL oder Soziologie wegzudenken. Zieht ein neues Webseiten-Design mehr Käufer an als das alte? Wie viele Kunden sollte eine Firma befragen, um belastbare Aussagen über die Interessen ihrer gesamten Kundschaft treffen zu können? Wie hoch ist das durchschnittliche Haushaltseinkommen in Deutschland? In

welchem Land sind die Einkommen besonders ungleich verteilt? Dauert es im Schnitt weniger als X Monate bis ein(e) Arbeitslose(r) wieder einen Job gefunden hat?

Dieser Kurs gibt eine Einführung in statistische Methoden, die dazu genutzt werden können, um i) Daten zu beschreiben, ii) Datenerhebungen zu planen, iii) von Stichproben auf ganze Populationen zu schließen, iv) Zufallsvorgänge zu analysieren und v) große Datensätze mittels Maschinellen Lernens zu analysieren.

Da mehr und mehr Berufsbilder in Wirtschaft und Forschung das praktische Arbeiten mit Datensätzen beinhalten, bietet der Kurs auch eine Einführung in die praktische Datenanalyse auf Basis der Statistiksoftware R.

Studierende, die den Kurs erfolgreich abschließen, können i) Datensätze mittels geeigneter Maßzahlen beschreiben und geeignete Formen zur Visualisierung von Daten auswählen, ii) grundlegende Stichprobenverfahren erläutern, iii) Parameter von Verteilungen schätzen und verschiedene Arten von Hypothesentests anwenden, iv) induktive Methoden als Grundlage des Arbeitens in empirischen Wissenschaften einsetzen und statistische Ergebnisse kritisch überprüfen, v) die Funktionsweise ausgewählter Verfahren des Maschinellen Lernens erklären und vi) die behandelten Methoden praktisch in der Statistiksoftware R implementieren.

**Prüfung:** Das Modul wird mit einer schriftlichen Klausur (120 Minuten) am Ende des Semesters geprüft. Die Klausur besteht zu rund 2/3 aus einem Multiple-Choice-Teil und zu rund 1/3 aus Aufgaben, die frei beantwortet werden müssen. Für Studierende, die die Kursversion RUW-2172 belegen (im Prinzip alle außer Studierende des Wirtschaftsingenieurwesens), beinhaltet die Klausur auch Fragen zur praktischen Umsetzung statistischer Analysen in R. Ein Termin zum Nachschreiben der Klausur wird am Ende des Sommersemesters angeboten.

**Klausurtermin:** Wird später vom Prüfungsamt bekanntgegeben.

## **Themen des Kurses**

### Deskriptive Statistik:

Wofür braucht man Statistik?

Welche Arten von Daten gibt es?

Wie kann man die Verteilung von Daten darstellen?

Wie misst man den „Durchschnitt“ einer Variablen?

Welche Maße geben sonst noch Auskunft über die Verteilung einer Variablen?

Wie misst man Konzentration in Märkten bzw. Ungleichheit (z.B. der Einkommen)?

Wie kann man Datensätze mit mehreren Variablen auswerten?

Wie misst man einen (linearen) Zusammenhang zwischen zwei Variablen?

Wie visualisiert man Daten am besten?

### Stichprobenverfahren:

Wie erhebt man eigentlich eine Datenstichprobe?

### Wahrscheinlichkeitsrechnung:

Was ist eine Wahrscheinlichkeit?

Wie kann man bedingte Wahrscheinlichkeiten berechnen?

Wie viele Möglichkeiten gibt es bestimmte Elemente zu kombinieren?  
Was ist eine Zufallsvariable und wie beschreibt man sie mathematisch?  
Was gibt es für wichtige Verteilungsfunktionen für diskrete Daten?  
Was gibt es für wichtige Verteilungsfunktionen für stetige Daten?

#### Grundlagen des Maschinellen Lernens:

Was ist Maschinelles Lernen?  
Wie kann man Beobachtungseinheiten datengetrieben in Gruppen aufteilen?  
Wie kann man Zufallsvariablen auf Basis anderer Information voraussagen?

#### Schließende Statistik:

Was kann man aus einer zufälligen Stichprobe schließen?  
Wie kann man einen unbekanntem Verteilungsparameter schätzen?  
Wie kann man Schätzunsicherheit darstellen?  
Wie kann man testen, ob ein Parameter einen bestimmten Wert hat?  
Wie kann man testen, ob zwei Stichproben ähnlich sind?  
Wie kann man testen, welche Verteilung am besten zu den Daten passt?  
Wie testet man, ob es einen Zusammenhang zwischen zwei Variablen gibt?

#### Die Geschichte der Statistik/Wahrscheinlichkeitsrechnung

##### **Lehrbücher:**

- Schlittgen, Rainer (2012), Einführung in die Statistik – Analyse und Modellierung von Daten (12. Auflage), Oldenbourg Wissenschaftsverlag GmbH, München. [Hauptlehrbuch; als elektronische Version in der FAU-Bibliothek verfügbar.]
- James, Gareth, Daniela Witten, Trevor Hastie und Robert Tibshirani (2013), An Introduction to Statistical Learning, Springer, Heidelberg. [Einige Exemplare in der FAU-Bibliothek verfügbar; als elektronische Version unter <http://faculty.marshall.usc.edu/gareth-james/ISL/ISLR%20Seventh%20Printing.pdf> verfügbar.]
- Kauermann, Göran und Helmut Küchenhoff (2010), Stichproben – Methoden und praktische Umsetzung in R, Springer, Heidelberg. [Als elektronische Version in der FAU-Bibliothek verfügbar.]

**Aufbau des Kurses:** Die Veranstaltung bietet Ihnen vielfältige Möglichkeiten die Grundlagen der Statistik zu lernen. Wöchentlich können Sie zum Lernen die Vorlesungsaufzeichnungen zu 2 Themenbereichen, 1 (Online-)Übung, 1 (Online-)Tutorium, und das R-Angebot, das über StudOn bereitgestellt wird, nutzen. Um einen optimalen Lernerfolg zu erzielen, empfehlen wir Ihnen jeweils i) die relevante Literatur vor den Vorlesungen anzuschauen, ii) die Übungsaufgaben eigenständig vor den Übungen und Tutorien zu rechnen und, falls Sie Probleme mit einer Aufgabe hatten, *nach* den Übungen/Tutorien nochmal zu probieren (ohne die Lösung daneben liegen zu haben) sowie iii) die Implementation des Stoffs in R eigenständig zu üben. Wichtig ist, dass Sie ab der ersten Woche mitarbeiten, weil viele Elemente des Kurses aufeinander aufbauen und Sie sonst schnell den Anschluss verlieren.

Zusätzlich zu den Veranstaltungen gibt es auf StudOn zu jedem Themenkomplex ein Online-Quiz, das Sie dazu nutzen können Ihren Lernfortschritt zu überprüfen. Machen Sie diese Quizzes kontinuierlich während des Semesters, um von diesem Angebot maximal zu profitieren.

### **Informationen zu R-Übung und R-Sprechstunde**

In der R-Übung wird gezeigt, wie mit Hilfe des Statistikprogrammes R der aus Vorlesung und Übung bekannte Stoff auf reale Datensätze angewendet werden kann. Zusätzlich zu den in den Übungen besprochenen Aufgaben (dieses Semester in Form von Videos) gibt es Übungsblätter mit Zusatzaufgaben, die der freiwilligen eigenständigen Kontrolle Ihres Lernfortschritts dienen. Bei Problemen mit diesen Aufgaben können Sie Hilfe im Rahmen der R-Sprechstunde bekommen. Sort können auch Fragen zu R im Allgemeinen geklärt werden und es wird Hilfe bei der Vor- und Nachbereitung der Aufgaben aus der R-Übung angeboten. Eine Vermittlung von zusätzlichen Inhalten in Form von Frontalunterricht findet in den Sprechstunden nicht statt. (Bitte beachten Sie, dass diese Veranstaltungen für Studierende des Wirtschaftsingenieurwesens nicht verpflichtend sind; die Studierenden können aber natürlich gerne freiwillig an dem Lehrangebot teilnehmen.)

### **Informationen zu R/R-Studio**

R ist eine freie Programmiersprache, die vor allem für statistische Analysen und zur Visualisierung von Daten genutzt wird. R-Studio ist eine integrierte Entwicklungsumgebung und grafische Benutzeroberfläche für R (ein „Editor“).

Um die Grundlagen des Programms zu lernen, müssen Sie in den ersten beiden Semesterwochen den von uns entworfenen [Onlinekurs „Basiskurs R/RStudio“](#) auf StudOn bearbeiten. Dort wird auch erklärt, wie Sie R und R-Studio auf Ihrem Computer installieren können.

Das Wichtigste, um die Datenanalyse mit R (oder jedem anderen Software-Paket) zu lernen, ist: Üben, üben, üben!!! Haben Sie keine Angst vor Fehlermeldungen! Googeln Sie nach Lösungen (es hatten immer schon andere das gleiche Problem)! Probieren Sie sich an den Aufgaben der Übungsblätter ohne direkt in die Lösung zu schauen! Programmieren Sie Beispiele (ggf. leicht modifiziert) nach!