|  |  |  |  |
| --- | --- | --- | --- |
| (Modul-)Titel | | RSDS-Modulbez. | |
| Data Science mit R (Applied Data Science with R) | | RSDS\_ADR | |
| (Modul-)Verantwortliche/r | | Fakultät | |
| Prof. Dr. Hößl | | BW | |
| Lehrende/r / Dozierende/r | | Angebotsfrequenz | |
| Prof. Dr. Hößl | | Wechselnder Fächerkatalog, die Veröffentlichung der im jeweiligen Semester angebotenen Module erfolgt online. | |
| Lehrform | | Unterrichtssprache | |
| Seminaristischer Unterricht | | Deutsch | |
| Art der Prüfung | | Voraussetzungen | |
| Empirische Projektarbeit  Kurzpräsentation der zu untersuchenden Forschungsfrage ist Zulassungsvoraussetzung zur Projektarbeit | | Grundkenntnisse statistischer Verfahren hilfreich | |
| Teilnehmerzahl | Modultyp | | Arbeitsaufwand (evtl. SWS und ECTS) |
| 10 BW  7 weitere Interessierte | WPF (BW) FW | | 4 SWS / 5 ECTS |
| Zielfakultäten/ -studiengänge: | Für Bachelor | | Für Master |
| Geöffnet für folgende Fakultäten:  A, ANK, B, EI, IM, M, S (7) | Studienabschnitt ab 2. Studienabschnitt | |  |
| Inhalt (Kurzbeschreibung) | | | |
| Die Inhalte werden – soweit möglich - im Kontext von Praxisbeispielen vermittelt.  Grundlagen der Empirie und Datenanalyseprozess  Einführung in R und die Paketgruppe tidyverse  Datenanalyseprozess mittels R und tidyverse  Einsatz ausgewählter statistischer, ökonometrischer Methoden und Verfahren des Machine Learning  Darstellung von Ergebnissen mittels statischer und interaktiver Grafiken  Automatisierte Generierung von Reports mittels RMarkdow | | | |
| Lernziel | | | |
| **Qualifikationsziele**  Die nachfolgenden Qualifikationsziele werden in verschiedene Dimensionen unterteilt. Jede Dimension entspricht dabei einer angestrebten Kompetenzstufe. Folgende Kompetenzstufen werden unterteilt:  Niveaustufe 1 (Kennen): oberflächliches Verstehen einfacher Strukturen bzw. Abfrage erworbenen Wissens  Niveaustufe 2 (Können): oberflächliches Verstehen mehrerer Strukturen bis zu tieferem Verständnis von Beziehungen zwischen Strukturen bzw. Gelerntes übertragen, zerlegen, kombinieren und einsetzen  Niveaustufe 3 (Verstehen und Anwenden): tieferes Verständnis von Beziehungen zwischen Strukturen bis zur Abstraktion und Erweiterung auf andere Strukturen bzw. Wissen hinterfragen und/oder bewerten, Zusammenhänge und Auswirkungen erläutern  Die jeweilige Dimensionszuordnung der Qualifikationsziele wird durch die Ergänzung der jeweiligen Ziffer (1,2 oder 3) in der Kompetenzbeschreibung dargestellt.  Im Einzelnen haben die Studierenden nach Abschluss des Moduls auf Basis wissenschaftlicher Methoden die folgenden Lernziele erreicht:  Fachkompetenz  Die Studierenden sind dazu befähigt, empirische Fragestellungen mittels geeigneter Methoden zu überprüfen (3), Schlussfolgerungen aus den Ergebnissen abzuleiten, (3) diese geeignet zu kommunizieren (2) und die Vorgehensweise kritisch zu reflektieren (3).  Sie kennen die einzelnen Schritte des Datenanalyseprozesses (1) und können diese für empirische Fragestellungen anhand der Analysesoftware R umsetzen (3). Sie sind sie in der Lage, Daten zu importieren, bereinigen und aufzubereiten (2), etwaige Datenprobleme zu identifizieren und diese - falls möglich - zu beheben (3). Sie kennen ausgewählte statistische/ökonometrische Verfahren sowie Methoden des Machine Learnings (1) und können diese zielgerichtet einsetzen (2). Die Ergebnisse ihrer Analysen können Sie für Dritte verständlich und nachvollziehbar aufbereiten und darstellen (2).  Studierende sind dazu befähigt, die vermittelten Inhalte im Rahmen von Studien-, Bachelor- und Masterarbeiten sowie bei weiteren empirischen Fragestellungen eigenständig anzuwenden (3).  Sozialkompetenz  Die Studierenden sind zum Selbststudium und zum vertieften eigenen Zeitmanagement befähigt (2). Sie sind in der Lage, gemeinschaftlich Fragestellungen anzugehen und hierfür adäquate Lösungen zu erarbeiten (2).  Methodenkompetenz  Die Studierenden sind dazu befähigt, mit den Analyseverfahren kompetent umzugehen (2) und ihre analytischen Fähigkeiten problemadäquat mittels R umzusetzen (3).  Persönliche Kompetenz  Die Studierenden sind in der Lage, Probleme und Fragestellungen angemessen empirisch zu untersuchen (2). Sie kennen die mit statistischen Auswertungen verbundenen Fallstricke (1) und können die Qualität von Untersuchungen und die dabei verwendeten Verfahren kritisch beurteilen (2). | | | |
| Zusätzliche Informationen: | | | |
| **Besonderes:**  Teilnehmerzahl ist begrenzt auf **20 Studierende**. Davon werden bis zu **10 Plätze im Rahmen der RSDS** für Studierende anderer Fakultäten reserviert.  Anmeldung erforderlich. Die Details zum Anmeldeverfahren für FW-Fächer sind für BW-Studierende in Moodle bzw. für Studierende anderer Fakultäten auf der Website der RSDS abrufbar. Eine Teilnahme ohne vorherige Anmeldung ist grundsätzlich nicht möglich.  Es wird ein eigener Rechner benötigt, auf dem die Programme R, RStudio und Rtools, sowie weitere R-Pakete, installiert sind.  Für einige Lerneinheiten wird die Lernplattform DataCamp verwendet. Um diese zu nutzen, müssen Sie sich unter Verwendung Ihrer OTH-Adresse einen Account bei DataCamp anlegen. Für die Dauer der Veranstaltung erhalten Sie kostenlosen Zugriff auf alle Inhalte und können kostenlos Kurse, Skill Tracks und Zertifizierungen anschließen.  Die Vorlesungszeiten sind dem Stundenplan zu entnehmen.  **Literatur:**  Pflichtliteratur  Ismay, C. / Kim, A. Y.: An Introduction to Statistical and Data Sciences with R, 2018  Wickham, Hadley / Grolemound, Garret: R for Data Science, O’Reilley Media, 2016  Xie, Yihui / Allaire, J. J. / Grolemound, Garret: R Markdown: The Definitive Guide, 2018  Zusätzlich empfohlene Literatur  Heuman, C. / Shomaker, S.: Introduction to Statistics and Data Analysis, Springer International Publishing, 2016  Kleiber, C. / Zeileis, A.: Applied Econometrics with R, Springer Science and Business Media, 2008  Schlittgen, R.: Angewandte Zeitreihenanalyse mit R, 3. Auflage, De Gruyter Oldenbourg, 2015  Shumway, R. H. / Stoffer, D. S.: Time Series Analysis and Its Applications, Springer Science and Business Media, 2006  Wickham, Hadley: Advanced R, Chapman and Hall, 2017  Wickham, Hadley: ggplot2, Springer, 2009  Wollschläger, D.: Grundlagen der Datenanalyse mit R, 3. Auflage, Springer Spektrum, 2014 | | | |