

Georg-August-Universität Göttingen Modul B.Inf.1131: Data Science: Grundlagen <i>English title: Data Science: Basics</i>	6 C 4 SWS
<p>Lernziele/Kompetenzen: Das Modul vermittelt grundlegende Kompetenzen im Umgang mit Daten und ihrer Analyse. Es gliedert sich in vier Teilbereiche</p> <p>Konzepte. Nach erfolgreicher Teilnahme</p> <ul style="list-style-type: none"> • kennen Studierende verschiedene Datentypen und können sie mit deskriptiven Statistiken beschreiben • kennen Studierende verschiedene Arten der Datenerhebung (experimentelles Design) und können deren Vorteile und Risiken benennen • kennen Studierende verschiedene Formen von Voreingenommenheit (Bias) in den Daten und die resultierenden Risiken, und können neue Kontexte hinsichtlich Bias bewerten • kennen Studierende Probleme der Fairness in Datenverarbeitung und Erhebung und können neue Kontexte hinsichtlich Fairness bewerten. <p>Software Werkzeuge. Erfolgreiche Teilnahme befähigt Studierende zum</p> <ul style="list-style-type: none"> • benutzen einer Shell zur grundlegenden Datenvorverarbeitung • analysieren von Daten mit grundlegenden Softwarebibliotheken für Datenverarbeitung in Python (Pandas, Numpy, Scipy, Matplotlib, ...) • testen von Software und statischen Algorithmen auf Korrektheit <p>Statistische Werkzeuge. Erfolgreiche Teilnahme befähigt Studierende zum</p> <ul style="list-style-type: none"> • unterscheiden zwischen statistischer Inferenz und deskriptiver Statistik • beherrschen der Grundlagen statistischer Inferenz (Fehler, p-Wert, Trennschärfe, Null-Hypothese, Konfidenzintervalle, ...) und vorhersagen welche Parameter diese beeinflussen • durchführen einfacher statistischer Tests mit Bootstrap- und Permutationstests • anwenden grundlegender Methoden des überwachten und unüberwachten Maschinellen Lernen (Klassifikation, Regression, Clustering). <p>Stil. Erfolgreiche Teilnahme befähigt Studierende zum</p> <ul style="list-style-type: none"> • anwenden guter Praktiken von Visualisierung von Daten • verfassen aussagekräftiger Projektberichte • strukturieren von reproduzierbaren Daten- und Softwareprojekten • strukturieren von Software für Wiederverwendbarkeit • anwenden von Prinzipien guter Codestrukturierung und -praktiken • anwenden grundlegende Formen des Projekt- und Team-Managements 	<p>Arbeitsaufwand: Präsenzzeit: 56 Stunden Selbststudium: 124 Stunden</p>
Lehrveranstaltung: Data Science: Grundlagen (Vorlesung, Übung)	4 SWS
Prüfung: Klausurähnliche Hausarbeit (Bearbeitungszeit: 1 Woche) Prüfungsanforderungen:	6 C