

Statistik für Digital Humanities

Organisatorisches / Pandemieplan

Dr. Jochen Tiepmar

Institut für Informatik
Computational Humanities
Universität Leipzig

12. Oktober 2020

[Letzte Aktualisierung: 11/10/2020, 09:10]

Modulinformationen

- Bachelor Digital Humanities
- Empfohlen ab 3. Semester
- Modulnummer:10-207-0003
- Arbeitsaufwand: 5 LP = 150 Arbeitsstunden
- Verwendbarkeit: SQ B.Sc.Informatik & B.Sc. DH, Kernmodul B.Sc. DH
- In Absprache mit Studiengangsverantwortlichen auch andere Verwendbarkeit möglich
- keine Teilnahmevoraussetzungen
- Vorlesung: montags 11:15 (asynchron)
- Übung: dienstags 11:15 (synchron)

Lernziele

Nach der aktiven Teilnahme am Modul *Statistik in den Digital Humanities* sind die Studierenden in der Lage

- statistische Grundbegriffe und Verfahren zu benennen und zu erklären
- ausgewählte Verfahren zu analysieren, zu beurteilen und diese selbstständig auf Problemstellungen im Kontext der Digital Humanities anzuwenden

Moodle Kurs

- Kursname: Statistik für Digital Humanities
- Kurskennzeichen: StatDH_WS2021
- Ergänzungsmaterialien und Übungsserien
- Weiterführende Informationen (Literatur etc.)
- (Freiwillige) Multiple Choice Tests zum Lösen während der Vorlesung
- Ergänzungsmaterialien neben der Vorlesung und den Übungsaufgaben sind nicht prüfungsrelevant

Literatur

Andy Field:
Discovering Statistics Using R (English Edition)
In der UBL vorrätig

Weitere Literaturempfehlungen

Peter Bruce:
Practical Statistics for Data Scientists: 50
Essential Concepts

Eileen Magnello, Borin Van Loon:
Statistics: A Graphical Guide

The R Project for Statistical Computing

- Frei verfügbare Softwareumgebung für statistische Analysen
- <https://www.r-project.org/>
- Selbstständige frühzeitige Installation und Einarbeitung empfehlenswert
- R selbst wird nicht gelehrt aber...
- Beispiele werden soweit sinnvoll mit R Skripten versehen
- Startprojekt im Moodle (Datensätze und Hallo Welt)
- Zur Prüfung steht R nicht zur Verfügung

Prüfungsleistung

- Details situationsbedingt unsicher
- Modulprüfung/Klausur (60 Minuten)
- Übungsschein = Prüfungsvorleistung
- **Unbedingt rechtzeitige Anmeldung / Abmeldung beachten**, auch bei Beendigung der Teilnahme (jtiepmar@informatik.uni-leipzig.de)
- Versäumtes Abmelden kann zu Fehlversuch der Klausur führen

Übungen

- 5 Übungsserien a 20 Punkten
- 50% der Punkte = Übungsschein/Klausurzulassung
- Übungsschein ist im Semester nicht wiederholbar
- Ausgabe nach der Vorlesung (Moodle)
- Abgabe spätestens am Abgabetag (montags) 24 Uhr über Moodle
- Gruppenarbeit (2-3 Studierende) möglich, **ist aber als solche zu markieren**
- 1 Abgabe pro Student_In
- Als Kopien interpretierte Abgaben werden mit 0 Punkten bewertet

Vorlesung

- Raumzeit: montags 11:15 asynchron
- Folien und Videos auf Uniserver
<http://www.informatik.uni-leipzig.de/~jtiepmar/lehre/statdh/ws2021/>
- Fragen usw. per Moodle-Forum oder Email

Seminare

- Raumzeit: dienstags 11:15 BigBlueButton synchron
- Besprechung der korrigierten Übungsaufgaben
- Lösungen der Aufgaben werden online gestellt
- Regulär 1 Termin pro Serie, Extratermine auf Anfrage

Syllabus

Wer zum 26.10 beginnt, kann den Übungstermin als Ausgleich nutzen

Vorlesung	Inhalt	Übungsblatt	Übung(Di)
12.10	Organisation		
19.10	Statistische Forschung		
26.10	Statistische Modelle		
02.11	Visualisierung	Ausgabe 1	
09.11	Annahmen Parametrischer Tests	Abgabe 1	
16.11	Kovarianz und Korrelation	Ausgabe 2	Ja
23.11	Regression	Abgabe 2	
30.11	Vergleich zweier Mittelwerte (t-Test)	Ausgabe 3	Ja
07.12	Vergleich mehrerer Mittelwerte (ANOVA)	Abgabe 3	
14.12	ANOVA II	Ausgabe 4	Ja
04.01	Nichtparametrische Testverfahren	Abgabe 4	
11.01	Kategorische Daten	Ausgabe 5	Ja
18.01	Faktoranalyse	Abgabe 5	
25.01	Logistische Regression		Ja
01.02	Prüfungsvorbereitung		Ja
Ferien	MANOVA		
Ferien	Mehrstufige Modelle (Multilevel Models)		