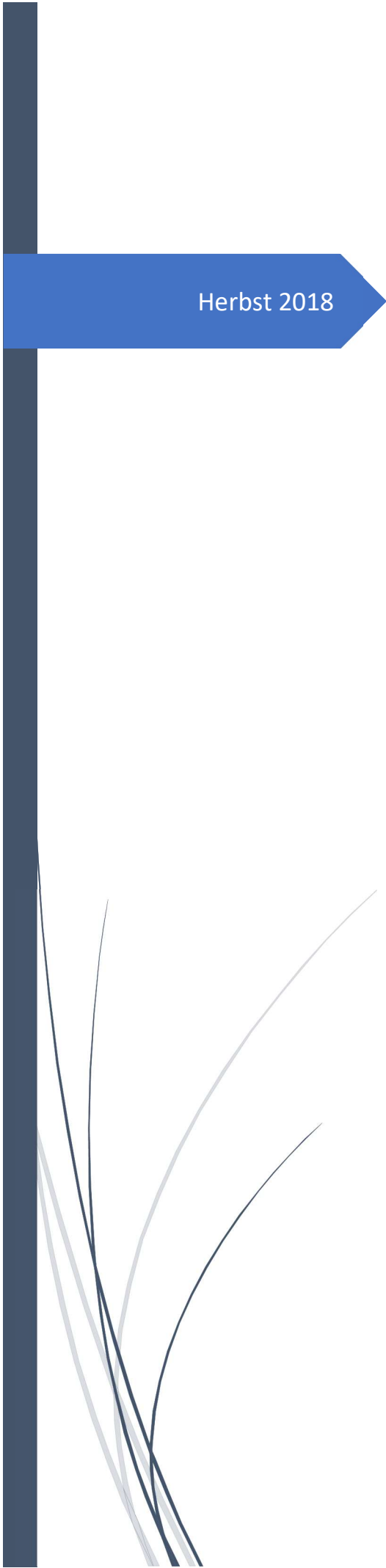


A dark blue vertical bar runs along the left edge of the slide. A blue arrow-shaped banner points to the right from this bar, containing the text 'Herbst 2018'. In the bottom-left corner, there are several thin, curved, light blue lines that sweep upwards and to the right.

Herbst 2018

Konzept der Veranstaltung „Methoden der Datenanalyse II“

Dr. Manuela Pötschke

UNIVERSITÄT KASSEL; ARBEITSGEBIET ANGEWANDTE STATISTIK

Vorwort

Viele Studierende, manche Kolleginnen und Kollegen, aber vor allem Menschen außerhalb der Universitäten denken, dass eine immer wiederkehrende Veranstaltung für die Dozentin keinerlei oder nur wenig Vorbereitungsarbeit mit sich bringt, weil die Inhalte ja von Jahr zu Jahr gleichblieben. Für die Konzeption der Lehrveranstaltungen zu den Methoden der Datenanalyse am Fachbereich 05 der Universität Kassel trifft das definitiv nicht zu. Vielmehr werden die Inhalte und vor allem die Didaktik der Veranstaltungen von Semester zu Semester weiterentwickelt. In die inhaltlichen Weiterentwicklungen fließen aktuelle Methodendiskussionen ein. Die didaktische Weiterentwicklung orientiert sich an Erkenntnissen der Lernforschung und an den Ergebnissen der kontinuierlich durchgeführten Forschung im Rahmen des Projektes „Statistik lernen“ (vgl. www.uni-kassel.de/go/statistik_lernen). Die Lernforschung zeigt z.B., dass das exemplarische Lernen an aktuellen Beispielen besser geeignet ist, auch abstrakte Konzepte zu verstehen, als der generelle Bezug auf formale statistische Modelle. Deshalb besteht eines der grundlegenden Prinzipien der Datenanalyseveranstaltungen im Zugang über Beispiele aus relevanten sozialwissenschaftlichen Surveys. Das erfordert dann für jede Veranstaltung die neue Aufbereitung passender Beispiele. Auch zeigt die neuere Lernforschung, dass der Einsatz multimedialer Mittel in der Lehre hilfreich sein kann. Dabei sind die Facetten der Wirksamkeit spezifischer Anwendungen aber noch nicht vollständig ausgeleuchtet. Hier hat die Veranstaltung auch experimentellen Charakter. Neuerungen werden ausprobiert und in den Kanon der zu verwendenden Methoden aufgenommen, wenn sie sich bewähren. Im anderen Fall werden sie im nächsten Durchgang wieder verworfen.

Insgesamt bedeutet das, dass ausgehend von einem Grundstock an konzeptionellen Überlegungen jede Veranstaltung wieder neu überdacht und angepasst wird. Das vorliegende Paper ist dem Grundstock gewidmet und stellt die basalen Überlegungen zur Veranstaltung „Methoden der Datenanalyse II“ vor.

Inhaltsverzeichnis

Formale Einordnung der Veranstaltung „Methoden der Datenanalyse II“	1
Grundlegende didaktische Prinzipien der Veranstaltung	1
Inhaltliches Konzept der Veranstaltung.....	3

Formale Einordnung der Veranstaltung

„Methoden der Datenanalyse II“

Die Veranstaltung richtet sich an Studierende im Hauptfach Soziologie und an Studierende im Hauptfach Politikwissenschaft mit dem Nebenfach Soziologie. Die empfehlenden Studienpläne sehen die Veranstaltung im 3. (Hauptfach Soz.) bzw. 5. Semester (Nebenfach Soz.) vor.

Die Veranstaltung umfasst eine Plenumssitzung á 4 SWS und ein Tutorium á 2 SWS. Die Plenumssitzung beinhaltet Inputsequenzen und Anwendungssequenzen.

Die zwei zentralen Lernziele der Veranstaltung sind:

- Studierenden können die Ergebnisse aus empirischen Studien anderer nachvollziehen, verstehen und kritisch beurteilen.
- Studierende können für eigene empirische Fragen angemessene Analyseverfahren begründet auswählen und die softwarebasierten Schätzergebnisse korrekt interpretieren.

Um die Ziele zu erreichen, liegen der Veranstaltung grundlegende didaktische Prinzipien und eine neu konzipierte inhaltliche Struktur zugrunde.

Grundlegende didaktische Prinzipien der Veranstaltung

Komplexe Methoden der Datenanalyse werden auf der Basis von Beispielanwendungen eingeführt.

Der Bezug auf praktische Beispiele erfolgt in verschiedener Hinsicht. Ausgangspunkt der Einführungen sind Anwendungen der Verfahren, die in Fachartikeln verwendet wurden. Dadurch ist der Zugang konkret und inhaltlich nachvollziehbar, da die Ergebnisse der statistischen Analysen auch verbalisiert vorliegen. Danach schließen sich weitere Beispiele zur Einführung an, die dann den Modelloutput stärker in den Fokus nehmen. Diese Beispiele basieren auf einem zentralen Survey für die Sozialforschung, dem ALLBUS. Der ALLBUS hat mehrere Vorteile für Studierende. Der Datensatz selbst und zugehörige Dokumentationen sind leicht zugänglich (jeder Studierende kann sich die Daten, den Fragebogen und Methodenmaterial nach einer Anmeldung auf www.gesis.org herunterladen). Der Datensatz umfasst Informationen zu aktuellen Themen von allgemeinem Interesse und er ist dabei nicht zu komplex. In den Tutorien können die Studierenden dann eigenen Fragestellungen auf der Basis des ALLBUS oder weiterer Surveys nachgehen und damit stärker ihre eigenen Interessen einbringen. Das wird möglich, weil die Tutorien systematisch in die Nutzung von Softwareanwendungen zur Datenanalyse einführen. Im Mittelpunkt stehen dabei (in den Semestern in unterschiedlicher Gewichtung) die Programme SPSS, R und STATA.

Der Zugang zu spezifischen Verfahren erfolgt über eine Herausforderung oder Aufgabe, vor der die Datenanalyse steht.

Datenanalyse steht insgesamt vor zwei zentralen Aufgaben. Statistische Modelle und aussagenlogische Zugänge können erstens der Erklärung sozialer Phänomene dienen. Besondere Bedeutung erlangt in diesem Zusammenhang die Frage nach der Kausalität sozialer Gegebenheiten. Zweitens können Modelle dazu genutzt werden, Strukturen zu entdecken oder zu überprüfen. Zur Lösung beider Aufgaben stehen zahlreiche Ansätze zur Verfügung. Eingeführt wird in die grundsätzlichen, so dass

die Studierenden nach der Veranstaltung auch selbst einen Zugang zu weiteren, nicht explizit behandelten Modellen finden können.

Die inhaltliche Gliederung innerhalb der beiden Richtungen orientiert sich am Ablauf der Interpretationsschritte, die in einer praktischen Datenanalyse zu gehen sind. Dadurch gewinnen die Studierenden nicht nur Sicherheit in der Anwendung der einzelnen Verfahren, sondern auch in der Strukturierung eines empirischen Forschungsvorhabens. Für die Abfassung empirischer Berichte, Hausarbeiten und Abschlussarbeiten ist das eine essentielle Kompetenz (vgl. das Paper „Hinweise für Forschungsberichte und empirische Abschlussarbeiten“ in diesem Material).

Neuere Erkenntnisse der Lernforschung werden in die Konzeption aufgenommen und die Lehre entsprechend gestaltet.

Der Schwerpunkt liegt dabei auf „wünschenswerten Erschwernissen“ im Lernprozess, weil empirische Studien deren Wirkungen auf einen nachhaltigen Lernerfolg zeigen konnten.¹ Zu den vorgesehenen Erschwernissen zählen Aufgaben zur Generierung eigener Lösungen zu statistischen Problemen, der Zugang zu den Gegenständen über verschachteltes Lernen und die Konstruktion von Testgelegenheiten. Das verschachtelte Lernen schlägt sich dabei auch in der Präsentation der Themen nieder. Über die beiden zur Verfügung stehenden Beamer werden abwechselnd Verfahren besprochen, so dass der Fokus vor allem auf den gemeinsamen Aspekten liegt und darauf aufbauend, die Spezifika jedes Verfahrens hervorgehoben werden kann.

Testgelegenheiten bieten sich durch die Bereitstellung zweier umfangreicher Onlinetests, die mehrfach und parallel zur Veranstaltung eigenständig bearbeitet werden können, durch die gemeinsame Besprechung der Lösungen einer Probeklausur zu Beginn des letzten Drittels der Veranstaltung und durch Aufgabenblätter, die zum Teil im Tutorium bearbeitet werden können.

Darüber hinaus waren bisher begleitende Übung von Veranstaltungsinhalten nach dem Prinzip des verteilten Lernens über eine Kooperation mit der Kollegin Barzagar-Nazari in die Veranstaltung eingebunden. Die Übungen werden solange angeboten, wie die Kooperation besteht.

Der Heterogenität in den Voraussetzungen wird mit einem Set von Lernunterstützungen begegnet.

Dazu zählen E-Learning-Angebote, Literaturempfehlungen, Handlungsanleitungen und Testgelegenheiten. Die Bereitstellung der Unterstützungen erfolgt über die Homepage „von Studis für Studis“ (www.uni-kassel.de/go/statistik_lernen) und über die Lernplattform Moodle.

Den E-Learning-Angeboten kommt hier eine besondere Funktion zu. Für „Methoden der Datenanalyse II“ wurden bisher Videofilme gedreht und Web Based Trainings gestaltet. Sie verfolgen jeweils spezifische Ziele. Die Filme stellen statistische Konzepte mit dem Verweis auf Alltagssituationen der Studierenden dar oder weisen einen eher motivationsfördernden Charakter auf. Die aktuellsten Entwicklungen sind WBTs. Diese sind als eigenständige Lehreinheiten zu verstehen, die jeweils ein Thema in den Blick nehmen. Neben der interaktiven Erarbeitung spielen die Übungen hier eine besondere Rolle.

¹ Vgl. Lipowsky, F., Richter, T., Borromeo Ferri, R., Ebersbach, M. & Hänze, M. (2015). Wünschenswerte Erschwernisse beim Lernen. *Schulpädagogik Heute*, 11.

Eine besondere Verbindung stellen Quiz über die digitale Anwendung ARSnova (vgl. <https://arsnova.thm.de/mobile>) dar. Zu Beginn einer Veranstaltung eingesetzt können Fragen zu den Inhalten vergangener Sitzungen an das Bisherige anknüpfen und zur Wissensaktivierung beitragen. In der Veranstaltung kann die Anwendung genutzt werden, um den Lernfortschritt zu dokumentieren und auf fehlende Links in den kognitiven Verknüpfungen der Studierenden hinzuweisen. Zum Ende wird die Anwendung vor allem als Instrument der Ergebnissicherung eingesetzt.²

Die Lernfortschritte werden durch eine begleitende Lernforschung dokumentiert und analysiert.

Im Rahmen der Veranstaltung werden die Studierenden zweimal befragt: zu Beginn des Semesters nach ihren Vorkenntnissen, Einstellungen und Erwartungen bezogen auf das Themenfeld der Datenanalyse und die Veranstaltung. Am Ende des Semesters werden sie gebeten, den Lernprozess und ihre Klausurvorbereitung zu reflektieren und einzuschätzen, welche Unterstützungsmaßnahmen ihnen wie geholfen haben. Ergänzt werden die Befragungsdaten durch spezifische und allgemeine Ergebnisse aus der Klausur.

Inhaltliches Konzept der Veranstaltung

Üblicherweise sind Veranstaltungen im Bereich der Datenanalyse linear aufgebaut.³ Nach einer sehr ausführlichen Einführung in das lineare Regressionsmodell schließt sich zumeist die Behandlung weiterer Verfahren an, die dann deutlich kürzer in Abgrenzung oder mit Bezug zum linearen Modell behandelt werden. Studierenden fällt es mit diesem Zugang häufig schwer, grundsätzliche Konzepte der Statistik, wie beispielsweise die Vorstellung von statistischen Modellen zu verstehen. Das Vorgehen fördert eher die Vorstellung, dass die Verfahren jedes für sich immer etwas völlig Neues beinhalten und unverbunden nebeneinanderstehen.

Ausgangspunkt des erneuerten inhaltlichen Konzeptes sind zwei zentrale Aufgaben der angewandten Datenanalyse in den Sozialwissenschaften: die Erklärung sozialer Phänomene und die systematische Strukturierung empirischer Beobachtungen. Dazu werden im Bereich der Methoden zur Datenanalyse II typischerweise Regressions- und Kovarianzverfahren eingeführt.

Auf der Basis des problembezogenen Zugangs werden die einzelnen Verfahren (lineare Regression, logistische Regression, Mehrebenenanalyse, Analyse latenter Konstrukte, Qualitative Comparative Analysis, Faktorenanalyse, Clusteranalyse) hier jedoch nicht nacheinander abgehandelt (wie es der Standard in der klassischen Lehre und in den meisten Lehrbüchern ist). Sondern sie werden zum Teil parallel eingeführt.

² ARSnova bietet weitere Anwendungen, die in der Veranstaltung bisher jedoch nicht genutzt werden. Die Nutzung von Feedbackfragen und Einschätzungen der Studierenden wird im nächsten Semester erprobt und bei positiver Evaluation in das allgemeine Konzept aufgenommen. Der Grad der Interaktivität wird dadurch deutlich erhöht. Die Anforderungen an die Flexibilität der Dozentin steigt enorm an.

³ Vgl. dazu die Empfehlungen der Methodensektionen der Deutschen Vereinigung für Politikwissenschaft (DVPW) und der Deutschen Gesellschaft für Soziologie (DGS) zur grundlegenden Ausbildung im Bereich empirischer Datenanalyse.

Das findet sich in der inhaltlichen Struktur der Veranstaltung wieder. Die einzelnen Themen sind dann über die konkreten Verfahren hinweg:

- Datenanalyse II im Forschungsprozess
- Erklärungsmodelle und Strukturierungsmodelle
- Interpretation von Signifikanzen in Modellen zur Datenanalyse
- Interpretation von Modellkoeffizienten
- Interpretation der Güte von Modellen
- Beurteilung von Modellvoraussetzungen
- Umgang mit wenigen Fällen in Populationen und Teilstichproben
- Umgang mit fehlenden Werten

Dieser Ansatz erlaubt zum einen den Zugriff auf datenanalytisches Denken insgesamt und stellt zum zweiten die gemeinsamen Probleme und Lösungen verschiedener Erklärungsmodelle in den Mittelpunkt. Dadurch können prinzipielle Themen wie Missing Data Analysis oder Gewichtung logisch integriert und nachfolgend deutlicher herausgearbeitet werden, worin die Unterschiede in den Anwendungsvoraussetzungen und Schätzungen in den Verfahren bestehen.