fu:stat thesis

Die Plattform **fu:stat thesis** wird von der Statistischen Beratungseinheit fu:stat betrieben, um Doktorandinnen und Doktoranden an der Freien Universität Berlin mit Rat und Tat beim Start des Empirieteils ihrer Promotion zu unterstützen. Dieses Wiki sammelt Informationen, Beispieldatensätze, Programmbeispiele, Videos und Animationen sowie Links auf Lehrbücher, die FU-Studierenden als E-Book zur Verfügung stehen. Die Software-Hinweise umfassen alle Statistik-Pakete, die hauptsächlich an der FU genutzt werden: JMP, R, SAS, SPSS und STATA.

Im Vordergrund stehen zunächst Fragestellungen in der Anfangsphase der empirischen Arbeiten. Fortgeschrittene Fragestellungen sollen weiterhin durch eine individuelle Beratung bei fu:stat geklärt werden.

Natürlich kann **fu:stat thesis** auch für die Anfertigung einer Masterarbeit genutzt werden. Allerdings sollte diese Plattform nicht als Einführungskurs für Statistikanfänger missverstanden werden.

Motivation

Immer mehr Dissertationen enthalten einen empirischen Teil. Die Analyse eines realen Datensatzes stellt den Verfasser bzw. die Verfasserin vor eine Aufgabe, für die sie häufig schlecht gerüstet sind: Die Einführungsveranstaltungen in Statistik liegen schon lange zurück und in der Frühphase des Studiums hat man eher daran gedacht, einen Pflicht- "Schein" zu erwerben, um sich dann der nächsten Vorlesung zu widmen.

Im Regelfall können die Promovierenden nur auf ein verschüttetes Statistikwissen auf elementarem Level zurückgreifen. Eigenes empirisches Arbeiten mit einem Datensatz ist vorher nie geübt worden. Datenmanagement, Umgang mit fehlenden Werten, Value Labels: Schulterzucken! Hinzu kommt Zeitdruck: Die Projektabgabe steht bevor oder der eigene Vertrag läuft bald aus. Schließlich noch die Ansprüche der Betreuenden, die auf die empirische Unerfahrenheit des Kandidaten kaum Rücksicht nehmen.

Allerdings verfügen die Promovierenden über eine weitaus höhere Arbeitsroutine als bei Studienbeginn. Auch ist die Arbeitsmotivation – die Anfertigung der Doktorarbeit – um ein Vielfaches höher als wenn es nur darum geht, eine Klausur zu bestehen.

In dieser schwierigen Startsituation möchte **fu:stat thesis** Hilfestellungen geben, die die Promovierenden in die Lage versetzen, den empirischen Teil der "Diss" anzugehen. Gerade in der Anfangsphase treten einige Standardsituationen auf, die sich gut eingrenzen lassen, so dass Hilfe suchende Studierende gut beraten werden können.

Es ist das Ziel von **fu:stat thesis**, Doktoranden in die Lage zu bringen, die empirischen Arbeiten einer Dissertation zu beginnen. Bei einem fortgeschrittenen Arbeitsstand werden auch komplexere Methoden und Modelle zum Einsatz kommen. Allerdings ist es unmöglich, hier auf begrenztem Raum einen Überblick zu geben. Eine empirische Beratung von einzelnen Dissertationsvorhaben sollte dann im Rahmen der Statistischen Beratungseinheit fu:stat durchgeführt werden. Umgekehrt fließen in einzelne Teilabschnitte viele Erfahrungen aus der Beratungstätigkeit von fu:stat ein. Vom formalen Niveau versuchen wir unsere Nutzer dort abzuholen, wo sie sind, was nicht ganz leicht ist.

fu:stat thesis befindet sich noch in der Erprobungsphase. Noch nicht alle Teile der Website sind öffentlich verfügbar und weitere Artikel werden in den kommenden Wochen erscheinen. Damit wir zu einem nützlichen Instrument der Graduiertenausbildung der FU werden, sind wir auf die Rückmeldungen unserer Nutzer angewiesen. Bitte senden Sie uns Ihre Rückmeldungen an:

thesis@stat.fu-berlin.de

fu:stat bietet regelmäßig Schulungen für Hochschulangehörige sowie für Unternehmen und weitere Institutionen an. Die Inhalte reichen von Statistikgrundlagen (Deskriptive, Testen, Schätzen, lineare Regressions) bis zu Methoden für Big Data. Es werden außerdem Kurse zu verschiedenen Software-Paketen gegeben. Auf Anfrage können wir auch gerne individuelle Inhouse-Schulungen bei Ihnen anbieten.

Unbekanntes Makro: 'vot

Im Einzelnen gibt **fu:stat thesis** Hilfestellungen in den folgenden Bereichen:

Das Verfassen einer empirischen Arbeit

- Ziele und Richtlinien
- To-Do's und Not-To-Do's in der Arbeitsorganisation

Datenmanagement

- Unterschiedliche Datenformate
- Variablenmanipulation
- Zusammenführen von Datensätzen
- Herausfiltern von Untergruppen

Das passende Statistik-Paket finden

- Das Datenformat

- Das Niveau des eigenen Statistikverständnisses
- Spezielle Erfordernisse der gewünschten Analyse
- Benutzerfreundlichkeit
- Lizenzgebühren
- Schulungen

Beschreibung der Statistik-Pakete

- JMP
- R (Nur mit FU-Login. Ohne Login besuchen Sie bitte die Seite unseres Kooperationspartners INWT-Statistics GmbH, auf der der Wikiartikel basiert.)
- SAS
- SPSS (Nur mit FU-Login. Ohne Login besuchen Sie bitte INWT-Statistics GmbH.)
- STATA (Nur mit FU-Login. Ohne Login besuchen Sie bitte INWT-Statistics GmbH.)
- Programmvergleiche (Nur mit FU-Login. Ohne Login besuchen Sie bitte INWT-Statistics GmbH.)

Explorative Analyse und Deskriptive Statistiken

- Erste deskriptive Auswertungen
- Graphische Übersichten

Standardauswertungen

- Das lineare Regressionsmodell
- Varianzanalyse (ANOVA)
- Statistische Tests für Kontingenztabellen (Chi-Quadrat Test, FET, McNemar)
- Multiples Testen
- Modelle für Cluster- und Paneldaten
- Logistische Regression
- Ordinale Regression

Modellauswahl und Modellgüte

- Wahl der Modellklasse: lineare Regression, Logit Modelle etc.
- Residuenplots
- Bestimmtheitsmaß R^2
- Modellselektion (AIC, BIC, Pseudo R^2, ...)

Die Gestaltung von Studien und ihre Auswertung

- Studiendesign
- Fragebogengestaltung: Fehler und Tipps zur Fehlervermeidung
- Von der Stichprobe zur Grundgesamtheit: Die Methodik der Stichprobentheorie

Umgang mit fehlenden Werten

- Vom Umgang mit fehlenden Werten - ein Überblick
- Imputationsverfahren

Beispieldatensätze und Beispielprogramme