

Advanced Survey Statistics: Disclosure Control

Administratives und Outlook

Matthias Templ

Institut für Datenanalyse und Prozessdesign
School of Engineering
Zürcher Hochschule für Angewandte Wissenschaften

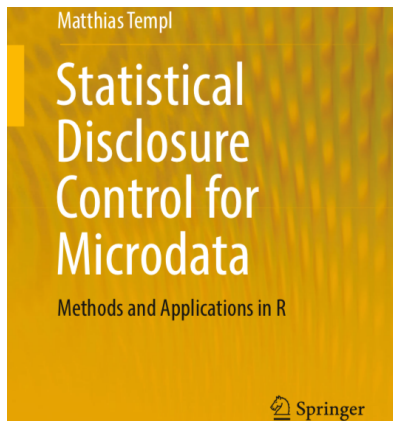
FU-Berlin, 2019

Zürcher Hochschule
für Angewandte Wissenschaften



Up to you.

- ▶ Erste Vorlesungsslides sind auf Deutsch vorbereitet, aber ich kann auf Englisch dazu sprechen
- ▶ Grundlage ist das Buch (Springer 2017, auf Englisch)



Zu meiner Person

- ▶ PhD in Techn. Math. an der Techn. Univ. Wien (TU Wien)
- ▶ Habilitation in Statistik TU Wien
- ▶ Arbeit im Methodenreferat der Statistik Austria
- ▶ Geoinformatik an der Palacky Univ. Olmütz
- ▶ Assoc.-Prof. am Institut für Statistics and Math. Methods in Economics, TU Wien (bis 2016)
- ▶ Dozent am Institut für Datenanalyse und Prozessdesign (IDP), Züricher Hochschule für Angewandte Wissenschaften (ZHAW)

Zu meiner Person

- ▶ PhD in Techn. Math. an der Techn. Univ. Wien (TU Wien)
- ▶ Habilitation in Statistik TU Wien
- ▶ Arbeit im Methodenreferat der Statistik Austria
- ▶ Geoinformatik an der Palacky Univ. Olmütz
- ▶ Assoc.-Prof. am Institut für Statistics and Math. Methods in Economics, TU Wien (bis 2016)
- ▶ Dozent am Institut für Datenanalyse und Prozessdesign (IDP), Züricher Hochschule für Angewandte Wissenschaften (ZHAW)

Hobbies:

- ▶ R. CRAN Task View Official Statistics and Survey Methodology (<https://cran.r-project.org/view=OfficialStatistics>)
- ▶ Editor Austrian Journal of Statistics (<https://www.ajs.or.at>)
- ▶ Kunst. Malen.
- ▶ Biologie und Klimaforschung
- ▶ Bier (nur wissenschaftlich! ;-))

- ▶ Was ist Euer Wissensstand bzgl. R?
- ▶ Wer hat welchen Background? (Statistik, Wirtschaftswissenschaften, Mathematik, Soziologie, ...?)
- ▶ Im welchen Semester?

Mai:

- ▶ 2 Vorlesungstage (9.5. und 10.5)
- ▶ Zuordnung von Aufgaben am Ende des zweiten Tages

Erste Juliwoche:

- ▶ 1.7. ganztags Vorlesung und 2.7. ev. nur halbtags
- ▶ Präsentationen der Aufgaben am 3.7. von 12:00 bis 18:00 Uhr

Ende Juli:

- ▶ Abgabe der Ausarbeitung der Aufgaben

Präsentation und Ausarbeitung

- ▶ in Gruppen zu je 3 Studierenden
- ▶ Präsentation der Aufgaben und (Teil)-Lösungen, ca. 20 Min. Präsentation und ca. 10 Min. Diskussion (genaue Länge wird noch fixiert)
- ▶ Nach Präsentation: Einarbeitung der Kommentare und Berücksichtigung des Juli-Stoffes. Abgabe Bericht Ende Juli. Ca. 5 Seiten ohne Tabellen, Grafiken und Referenzen. Max 10 Seiten mit Tabellen, Grafiken, Referenzen.

Note:

- ▶ 50% Präsentation, 50% Ausarbeitung

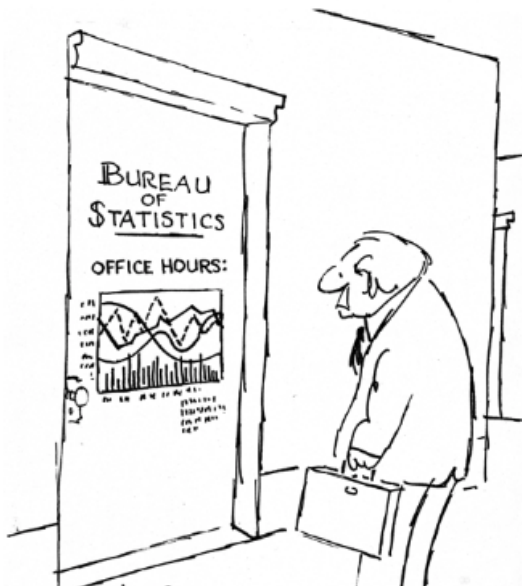
Wichtig:

- ▶ Wer hat was gemacht? Dies sollte klar in der Präsentation und Ausarbeitung vermerkt sein.

- ▶ Ihre aktive Teilnahme ist erwünscht.
- ▶ Ideal: Lehrgespräch statt Frontalunterricht
- ▶ Selbstständige Ausarbeitungen zwischendurch
- ▶ Grün markierte Kapitel im Buch sind speziell für diese Vorlesung wichtig.
- ▶ SDC ist manchmal ein „trockener“ Stoff. Trotzdem darf und soll in der Vorlesung auch gelacht werden.

Das techn. Verständnis der Methoden der Statistische Geheimhaltung (SDC) ist nur die halbe Miete. Die Anwendung ist datenspezifisch und erfordert Erfahrung mit Daten und erweitertes Verständnis von SDC.

Erreichbarkeit des Vortragenden (nur per Email)



- ▶ R und sdcMicro (get in touch, Buch Kapitel 1). Grundsätzlich aber sdcMicro und simPop Anwendungen nach Theorieteil
- ▶ Basics (Buch Kapitel 2) und Rechtliches (Folien)
- ▶ Disclosure Risk (Buch Kapitel 3)
- ▶ Anonymisierungsmethoden (Buch Kapitel 4)
- ▶ Utility (Buch Kapitel 5)

Falls Zeit bleibt:

- ▶ Optional: Einblick in synthetische Datengenerierung
- ▶ Optional: Case Studies
- ▶ Optional: Geheimhaltung in Tabellen
- ▶ Optional: Matching

Let's start

