TU DORTMUND

Wintersemester 2024/2025

Fakultät Statistik

Prof. Dr. Jörg Rahnenführer

M.Sc. Marieke Stolte

Dr. Franziska Kappenberg

Dr. Julia Duda

Leitfaden zur Erstellung der Projektberichte für Fallstudien I

Layout und Umfang

Das endgültige Layout ist jedem Studierenden selbst überlassen, jedoch sollten folgende Regeln eingehalten werden:

- Blocksatz
- Ausreichend Korrekturrand (mindestens 2.5 cm) an allen vier Rändern freilassen
- Zeilenabstand 1.5-fach
- Nicht zu kleine Schriftgröße (z.B. 12pt für Times New Roman)
- Einheitliche Schriftart im gesamten Bericht
- Der "Kern" des Berichts (ohne Deckblatt, Inhaltsverzeichnis, Anhang und Literaturverzeichnis) darf maximal 12 Seiten umfassen, der Anhang maximal 3 Seiten.
- Seitenzahlen sollten vergeben werden. Die Nummerierung beginnt nach dem Inhaltsverzeichnis.

Hinweis

- Die Wahl des Textverarbeitungsprogramms ist freigestellt. Es wird jedoch LATEX empfohlen. Für das Zitieren sowie die Organisation des Literaturverzeichnisses wird BibLaTeX empfohlen.
- Achten Sie darauf, dass keine Wörter, Abbildungen oder Tabellen über den Rand hinaus geschrieben werden (siehe warnings in IATEX). Achten Sie bei der Verwendung von BibLaTeX auf sinnvolle Ergebnisse (z.B. Abkürzungen, Umlaute).
- Vermeiden Sie, dass einzelne Zeilen am Beginn oder Ende eines Absatzes auf einer anderen Seite stehen als die übrigen Zeilen (sog. "widow" bzw. "orphan lines").

Aufbau

Die nachfolgende Auflistung ist ein Grundgerüst für den Aufbau wissenschaftlicher Arbeiten, welches hier für die Verfassung der Berichte zwingend einzuhalten ist. Die zusätzlich aufgeführten Stichpunkte beschreiben den inhaltlichen Aufbau der einzelnen Kapitel. Zur Strukturierung der einzelnen Kapitel können Unterkapitel gebildet werden.

- Deckblatt
 - Veranstaltungsname
 - Thema des Projekts
 - Namen der Dozenten
 - Name des Verfassers und der übrigen Gruppenmitglieder sowie die Gruppennummer

- Datum

• Inhaltsverzeichnis

 Auflistung aller Kapitel (einschließlich Unterkapitel, falls vorhanden), des Anhangs und des Literaturverzeichnisses (nicht aber des Inhaltsverzeichnisses selbst), jeweils mit zugehöriger Nummer bzw. Buchstaben (vgl. Abschnitt Gliederung), und mit Angabe der Seitennummer, auf dem das entsprechende Kapitel beginnt

Hinweis

Das Inhaltsverzeichnis gehört auf die zweite Seite.

• 1 Einleitung

- Kurze allgemeine Einführung in das Thema
- Kurze Beschreibung von Inhalt und Ziel des Projekts
- Kurze Erläuterung der Vorgehensweise zur Problemlösung (inhaltlich und methodisch)
- Evtl. kurze Darstellung der zentralen Ergebnisse
- Überblick über die einzelnen Kapitel

Hinweis

Der Kapitelüberblick soll mit Inhalt gefüllt sein. Beispiel: Geben Sie an, welche statistischen Verfahren erläutert werden, anstatt lediglich zu schreiben, dass statistische Verfahren erläutert werden. Faustregel: Der Überblick darf nicht zu einem Bericht über ein anderes Projekt passen.

• 2 Problemstellung

- Beschreibung des Datenmaterials
 - * Hintergrund und Art der Datenerhebung (geplanter Versuch, Beobachtungsstudie, Fragebogen usw.)
 - * Art und Umfang der Stichprobe (Vollerhebung, Schichtung usw.)
 - * Beschreibung aller Variablen (inhaltliche Bedeutung, Einheiten usw.)
 - * Diskussion der Datenqualität (fehlende Werte, Skalenniveau, Messgenauigkeit usw.)
- Beschreibung der Ziele des Projekts (inhaltliche und statistische Zielsetzung)

Hinweise

- Geben Sie immer eine Quelle für die Daten an.
- Erläutern Sie den Datenerhebungsprozess detailliert, sofern dieser Ihnen bekannt ist.
 Gehen Sie dabei ggf. auch auf durchgeführte Vorverarbeitungsschritte im Vergleich zu der von Ihnen angegeben Quelle ein.
- Nennen Sie alle im Datensatz vorhandenen Variablen. Variablen, die im Bericht verwendet werden, müssen genauer beschrieben werden. Variablen, die nachfolgend nicht weiter betrachtet werden, werden lediglich erwähnt. (Enthält ein Datensatz sehr viele weitere Variablen, so reicht es die Anzahl und ihre Themengebiete zu nennen.) Variablen dürfen im Vergleich zum Datensatz umbenannt sowie übersetzt werden. Zudem dürfen Abkürzungen eingeführt werden. Nichtssagende Abkürzungen wie X_1, X_2, \ldots sollten dabei vermieden werden.
- Kompliziertere Variablen müssen erläutert und interpretiert werden (z.B. Wie setzt sich ein vorliegender Score zusammen? Was bedeuten hohe bzw. niedrige Werte?)

Während die Einleitung lediglich eine kurze Beschreibung der inhaltlichen und methodischen Ziele enthält, erfolgt hier eine ausführliche Erläuterung dieser Ziele.

• 3 Statistische Methoden

- Beschreibung der verwendeten statistischen Verfahren, Modelle usw. inklusive ihrer Eigenschaften sowie der ihnen zugrundeliegenden Annahmen (hier sind auch mathematische Formeln nötig)
- Angabe der verwendeten Hilfsmittel (Software inklusive Versionsnummer, statistische Tafeln usw.)

Hinweise

- Es sollen alle zur Auswertung verwendeten Methoden (und nur diese) beschrieben werden
- Motivieren Sie jedes Verfahren: Geben Sie an, wozu es genutzt werden kann. Falls Sie mehrere Methoden beschreiben, die demselben Zweck dienen (z.B. Beschreibung der zentralen Lage), erläutern Sie die Unterschiede (z.B. Median ist robuster als arithmetisches Mittel Achtung: für diese Aussage muss auch Robustheit erklärt werden).
- Alle Methoden müssen sprachlich und mathematisch präzise erläutert werden.
- Erläutern Sie alle Bestandteile der von Ihnen angegebenen Formeln. Beispiel: Bei Angabe der Formel $\bar{x} := \frac{1}{n} \sum_{i=1}^{n} x_i$ muss erklärt werden, was x_1, \ldots, x_n sowie n inhaltlich bezeichnen. Darüber hinaus muss angegeben werden, was x_1, \ldots, x_n bzw. n erfüllen müssen (Skalenniveau, Wertebereich, etc.).
- Methoden, die Sie bereits in einem vorhergehenden Bericht beschrieben haben, dürfen Sie als bekannt voraussetzen.
- Achten Sie darauf, dass die von Ihnen in diesem Kapitel beschriebenen Methoden auch wirklich die sind, die Sie in der Auswertung verwenden: Lesen Sie daher aufmerksam die Hilfeseiten zu nicht selbst implementierter Software.
- Für jede statistische Methode müssen Sie eine Literaturquelle angeben, in welcher die Methode so, wie von ihnen angegeben, beschrieben wird. Ihre eigene Notation muss dabei nicht mit der der Quelle übereinstimmen (vgl. Abschnitt Referenzen).
- Erläutern Sie auch, wie die Ergebnisse der Methoden interpretiert werden. Z.B. Was sagen hohe bzw. niedrige Werte einer Kennzahl aus? Wann kann man anhand eines Plots davon ausgehen, dass Annahmen erfüllt bzw. nicht erfüllt sind?

• 4 Statistische Auswertung

- Gegebenenfalls Überprüfung der zu Grunde liegenden Annahmen der verwendeten statistischen Verfahren
- Ausführliche Darstellung der Ergebnisse, aufbereitet mit Hilfe von Tabellen und Abbildungen
- Interpretation der Ergebnisse im Hinblick auf die Problemstellung

Hinweise

 Das erste Unterkapitel sollte (mit Ausnahme des ersten Berichts) stets eine kurze deskriptive Beschreibung des vorliegenden Datenmaterials enthalten. Alle Variablen müssen univariat betrachtet werden. Ggf. sind auch bivariate Betrachtungen sinnvoll. Die in den nachfolgenden Unterkapiteln gewonnenen Erkenntnisse sollen auch

- in Bezug zu den Ergebnissen der Deskription gesetzt werden.
- Beschreiben Sie nachvollziehbar, wie Sie von den Daten zu den Ergebnissen gekommen sind, z.B. was mit fehlenden Werten geschehen ist.
- Ein "roter Faden" fördert die Verständlichkeit ihrer Ergebnisse. Erläutern Sie daher, wozu die jeweiligen Analysen dienen und stellen Sie Zusammenhänge zwischen den Analysen her (Beispiel: "Diese Analyse dient zur Überprüfung der Annahmen für …"). Geben Sie zu Beginn eines Unterkapitels bzw. Abschnitts kurz an, was Sie in diesem untersuchen werden. Führen Sie dann die Untersuchung durch. Formulieren Sie abschließend ein kurzes Fazit hinsichtlich der zu untersuchenden Fragestellung.
- Interpretation meint objektive Schlussfolgerungen anhand der Daten. Subjektive Anmerkungen gehören in die Diskussion im Kapitel "Zusammenfassung".

• 5 Zusammenfassung

- Kurze Wiederholung der Fragestellung und des Datenmaterials des Projekts
- Kurze Darstellung der wichtigsten Ergebnisse
- Diskussion der Ergebnisse (mögliche Schlussfolgerungen, Warnung vor Fehlinterpretationen usw.)
- Ausblick (offene Fragen, Hinweis auf mögliche weitere Untersuchungen usw.)

Hinweise

- Die Einleitung und die Zusammenfassung sollen eigenständig lesbar sein. Das heißt insbesondere, dass kein Wissen aus den anderen Abschnitten vorausgesetzt werden darf.
- Beantworten Sie alle Fragen. Auch nicht signifikante Testergebnisse sind Ergebnisse.
- Geben Sie an, mit welchen statistischen Verfahren Sie Ihre Ergebnisse erhalten haben. Erwähnen Sie bei statistischen Tests auch das Signifikanzniveau.
- Ein kritisches Hinterfragen der Ergebnisse kann angebracht sein (z.B. multiples Testniveau). Sie sollten jedoch nicht sämtliche Ergebnisse infrage stellen.

• Literaturverzeichnis

- Auflistung aller zur Auswertung herangezogenen Zeitschriftenartikel, Bücher (dazu gehören z.B. auch Software-Manuals) und Web-Seiten
- Die Literaturquellen werden alphabetisch nach dem Nachnamen des Erstautors sortiert.

Hinweise

- Beispiel für eine Lehrbuchangabe:
 Schumacher, M. und Schulgen, G. (2008): Methodik klinischer Studien, 3. Auflage,
 Springer, Berlin.
- Beispiel für einen Zeitschriftenartikel:
 Byth, K., Cox, D. R. und Forder, P. (2006): Assessing the relationship between symptoms of allergic rhinoconjunctivitis and pollen counts, Australian and New Zealand Journal of Statistics 48(4), 417–428.
- Bei Web-Seiten muss das Datum der Abfrage mit angegeben werden. Beispiel:
 Wikipedia (2024): Statistik. URL: https://de.wikipedia.org/wiki/Statistik
 (besucht am 01.10.2024)

- Im Literaturverzeichnis werden alle Autoren aufgeführt (und nicht mit et al. abgekürzt).
- Alle im Literaturverzeichnis aufgeführten Arbeiten müssen im Text zitiert werden und umgekehrt.

Anhang

- Weitere wichtige Tabellen und Abbildungen

Hinweise

- Bezugnahme auf Elemente des Anhangs erfolgt immer unter Angabe der Seitenzahl, auf welcher die Tabelle bzw. Abbildung zu finden ist
- Ob Tabellen bzw. Abbildungen im Kapitel "Auswertung" oder im Anhang platziert werden, sollte nach Verständlichkeit und Übersichtlichkeit entschieden werden. Die Auswertung soll gelesen werden können, ohne dass es nötig ist, mehrere Seiten zur entsprechenden Grafik zu blättern.

Gliederung

- Längere Kapitel sollten in Unterkapitel mit aussagekräftigen Überschriften unterteilt werden. Zusätzlich sollte der Text durch Absätze und ggf. Abschnittsüberschriften strukturiert werden. Zwischen zwei Absätzen soll deutlich mehr Abstand gelassen werden, als zwischen Zeilen innerhalb eines Absatzes.
- Alle Kapitel (und ggf. die Unterkapitel) des Kerns erhalten eine Nummer, wobei eine mehr als dreistufige Gliederung unübersichtlich und daher unüblich ist. Empfohlen wird eine höchstens zweistufige Gliederung.
- Bei einer Gliederung im Anhang (die nicht zwingend erforderlich ist, z.B. wenn der Anhang kurz ist) werden Buchstaben verwendet.
- Inhalts- und Literaturverzeichnis erhalten weder eine Nummer noch einen Buchstaben.
- Die Kapitel "Einleitung" bis "Zusammenfassung" werden mit den Zahlen 1 bis 5 nummeriert. Unterkapitel dieser Kapitel werden mit 1.1, 1.2, ..., 2.1, 2.2, etc. nummeriert. Die Kapitel "Literaturverzeichnis" und "Anhang" erhalten keine Nummerierung. Falls das Kapitel "Anhang" in Unterkapitel unterteilt wird, so werden diese Unterkapitel mit A, B, C, etc. durchnummeriert.
- Das aus der Gliederung entstehende Inhaltsverzeichnis sollte durch die Unterkapitel bereits einen inhaltlichen Überblick geben. Faustregel: Der Überblick darf nicht zu einem Bericht über ein anderes Projekt passen. Allerdings sollte der Überblick auch nicht zu detailliert sein.

Abbildungen und Tabellen

- Alle Abbildungen und Tabellen müssen eine aussagekräftige Unterschrift erhalten (Beispiel für eine Tabellenunterschrift: "Lage- und Streuungsmaße für die Variable 'Alter' getrennt nach Geschlecht"). Redundante Überschriften in Abbildungen sollten vermieden werden.
- Abbildungen und Tabellen werden durchnummeriert, z.B.: "Abbildung 1: Titel oder Erklärung der Abbildung" bzw. "Tabelle 1: Titel oder Inhalt der Tabelle"

- Auch im Anhang müssen Abbildungen und Tabellen nummeriert werden und eine Unterschrift enthalten.
- Auf jede Abbildung und Tabelle muss im Text mindestens einmal Bezug genommen werden (das gilt auch für diejenigen im Anhang). Dabei wird die entsprechende Nummer der Abbildung oder Tabelle verwendet (Beispiel: "In Tabelle 1 ist der Zusammenhang zwischen …").
- Eine Abbildung und eine Tabelle erscheinen typischerweise auf der Seite oder der Folgeseite des ersten Ansprechens im Text, aber niemals vorher.
- Abbildungen und Tabellen sollen selbsterklärend sein. Dazu gehört, dass alle Achsen beschriftet sind. Vergessen Sie hierbei nicht die Einheiten. Zusätzliche relevante Informationen können z.B. in einer Legende dargestellt werden. Informationen, die zum Verständnis nötig sind, aber nicht innerhalb der Abbildung bzw. Tabelle platziert werden können, können im Untertitel mitgeteilt werden.
- Die Größe der Beschriftung der Abbildungen sollte etwa der Schriftgröße des Berichts entsprechen.
- Falls Farben verwendet werden, sollten gut zu unterscheidende Farben gewählt werden.
 Wird der Bericht später in schwarz-weiß gedruckt, empfehlen sich beispielsweise Graustufen.
- Format von Tabellen in der Regel: Zahlen rechtsbündig, Text linksbündig, ansonsten Inhalt auf bestimmte Weise zentriert, falls es die Lesbarkeit erhöht. Alle Zahlen in einer Spalte müssen dieselbe Anzahl an Nachkommastellen aufweisen (ggf. Nullen ergänzen). Sofern sinnvoll, sollte in allen Spalten einer Tabelle dieselbe Anzahl an Nachkommastellen verwendet werden. Variablen, die nur ganzzahlige Werte annehmen, werden ohne Nachkommastellen dargestellt.
- Stellen Sie nicht Variablen mit verschiedenen Einheiten in einem Plot gemeinsam dar.
- Alle Abbildungen und Tabellen (auch die im Anhang) müssen im Text erläutert werden. Dazu sollte zunächst kurz erwähnt werden, was in einer Abbildung bzw. Tabelle dargestellt ist. Beispiel: "In Tabelle 1 sind die Ergebnisse der Bundestagswahl 2017 getrennt nach Bundesländern dargestellt." Anschließend muss erklärt werden, welche Erkenntnisse aus der Abbildung bzw. Tabelle gewonnen werden können. Beispiel: "Partei X erzielt ihr bestes Ergebnis in Bundesland Y." Man darf nicht davon ausgehen, dass eine Abbildung bzw. Tabelle doch offensichtlich etwas zeigen würde, alle Erkenntnisse müssen in Textform festgehalten werden.
- Verwenden Sie Vektorgrafiken (z.B. .pdf).

Referenzen

• Innerhalb des Berichts erfolgen Referenzen auf eine Literaturangabe mittels Autor (Jahr) oder (Autor, Jahr). Bei zwei Autoren sind beide anzugeben, bei mehr als zwei Autoren ist die Kurzform Autor 1 et al. (Jahr) bzw. (Autor 1 et al., Jahr) zu wählen. Sind mehrere Arbeiten eines Autors aus demselben Jahr vorhanden, werden sie zusätzlich durch Buchstaben nach den Jahreszahlen unterschieden: (2005a), (2005b) usw.

Beispiel für die oben genannten Publikationen:

- Schumacher und Schulgen (2008) bzw. (Schumacher und Schulgen, 2008)

- Byth et al. (2006) bzw. (Byth et al., 2006)
- Bei Büchern muss das entsprechende Kapitel bzw. die entsprechende Seite mitangegeben werden. Beispiel: "[...] Die disjunktive Normalform für die Systemzuverlässigkeitsfunktion ist aus Hartung et al. (2005, S. 756) entnommen."
- Für alle Informationen, die nicht zum Allgemeinwissen gehören und die nicht der Projektbeschreibung entnommen wurden, muss eine Quelle angegeben werden.
- Auch für die verwendete Software ist eine Referenz anzugeben, wobei die Herausgeber häufig angeben, wie man auf die Software verweisen kann. Bei Verwendung der Software R müssen auch alle verwendeten Pakete zitiert werden, welche nicht zum Basisumfang von R gehören.
- Vorlesungsskripte, Wikipedia, etc. sind keine zulässigen Quellen für statistische Methoden.
- Quellen dürfen nicht wortwörtlich übernommen werden, sondern die Ideen müssen in eigenen Worten formuliert werden. Die eigene Formulierung darf nicht zu nah an der der Originalquelle sein. Bei Formeln ist es nötig, dass nicht die Notation jeder Quelle exakt übernommen wird, sondern dass im Bericht eine konsistente Notation vorliegt.
- Falls verfügbar, wählen Sie Quellen, welche (im Hochschulnetz) online abrufbar sind.

Mathematische Formeln

- Mathematische Formeln und Symbole (ob eingerückt oder im Fließtext) müssen vom restlichen Schriftbild abgehoben werden (mathfont in LATEX).
- Einmal gewählte Bezeichnungen sollen im gesamten Bericht beibehalten werden.
- Wichtig ist, dass unterschiedliche Größen klar voneinander abgegrenzt werden: Großbuchstaben für Zufallsvariablen, Kleinbuchstaben für Beobachtungen.
- Auch mathematische Formeln können nummeriert werden (mittels "(Nummer)"), so dass im Text auf sie Bezug genommen werden kann.
- Größere Formeln sollten abgesetzt werden.
- Namen dürfen nicht im mathfont geschrieben werden (z.B. $\exp(x^2)$ statt $\exp(x^2)$ oder $\min\{0, x\}$ statt $\min\{0, x\}$.
- Bestandteile von Formeln werden auch im Text im mathfont geschrieben. Beispiele: $\sum_{i=1}^{n} x_i$ ist die Summe von n Werten (nicht n Werten), t-Test, p-Wert.
- Zu viele Indizes wirken unübersichtlich.

Formulierungen

- Der Bericht (abgesehen von der Zusammenfassung) wird im Präsens geschrieben. Die Zusammenfassung wird in einer konsistenten Vergangenheitsform geschrieben. Wichtig: Erkenntnisse, die weiterhin gelten, werden auch dort im Präsens verfasst (z.B. Zwischen X und Y besteht eine empirische Korrelation von 0.3).
- Sprachlich sollten präzise und sachliche/wissenschaftliche Formulierungen gewählt werden. Zu vermeiden sind: Erlebnisstil, Umgangssprache, Füllwörter, zu viele Wortwiederholungen und lange Schachtelsätze. Der Text muss flüssig lesbar sein.
- Verwenden Sie eine passive Ausdrucksweise (z.B. "es wird" statt "wir haben"). Vermei-

- den Sie dabei Konstruktionen mit "soll" (z.B. "wird untersucht" statt "soll untersucht werden").
- Beachten Sie die Sprachlogik. Beispiel: "daher" oder "deshalb" geben an, dass es einen kausalen Zusammenhang gibt.

Weitere Hinweise

- Zitate sind als solche kenntlich zu machen (Einrücken des Textes, Anführungszeichen, exakte Angabe des Buches/Artikels und der Seitenzahl) und dürfen nicht aus längeren Textpassagen bestehen. Das Abschreiben von Textpassagen ist nicht erlaubt!
- An vielen Stellen ist die genaue Formatierung nicht eindeutig festgelegt. Wichtig ist hierbei, dass eine einmal gewählte Form konsistent im gesamten Bericht angewendet wird.
- Ein guter Bericht zeichnet sich dadurch aus, dass er inhaltlich und formal korrekt ist sowie dadurch, dass die vorliegende Fragestellung unter Verwendung angemessener Verfahren sinnvoll beantwortet wird. Es geht nicht darum, alles über den vorliegenden Datensatz aufzuschreiben. Vielmehr geht es darum, die zentralen Erkenntnisse, welche hinsichtlich der Fragestellung gewonnen wurden, anschaulich, präzise und korrekt darzustellen.
- Bevor Sie Ihren Bericht abgeben, sollten Sie diesen mindestens einmal vollständig Korrektur lesen. Es kann hilfreich sein, die Berichte zusätzlich innerhalb der Gruppe gegenseitig Korrektur zu lesen.