

Grundlagen
Häufigkeiten
Lagemaße
Streuung
Inferenzstatistik
Kreuztabellen
Gruppen-
unterschiede
Kovarianz/
Korrelation
Lineare
Regression

Herzlich willkommen zur
Vorlesung

Statistik

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-
unterschiede

Kovarianz/

Korrelation

Lineare

Regression

Vorlesung Statistik – für wen?

- BA Social Science/BA Sozialwissenschaften (Pflicht)
- Lehramt SoWi: Für einige Studiengänge Wahlmöglichkeit im Modul Didaktik und Methoden o.ä.
- M-PEB (gibt es wahrscheinlich nicht mehr): Pflicht; für „Teilnahmeschein“ Klausur
- Sonstige: ??
- **Nicht-Teilnehmende in LSF abmelden!!**

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen- unterschiede

Kovarianz/ Korrelation

Lineare

Regression

Leistungsnachweis

- Regelmäßige Bearbeitung von Tutoriumsaufgaben zur Klausurvorbereitung.
 - 3 KP: Mindestens 8 Aufgaben
 - 2 KP: Mindestens 5 Aufgaben(zwecks Klausurvorbereitung empfiehlt sich die Bearbeitung möglichst vieler Aufgaben!)

Ohne Bearbeitung der Tutoriumsaufgaben können die KP nicht erworben werden (dokumentierte Nachbereitung der Vorlesung)!

Die Bearbeitung der Tutoriumsaufgaben beginnt nächste Woche. Suchen Sie sich bis dahin eine/n Partner/in – die Aufgaben sollen in 2-er Teams bearbeitet werden. Bitte wählen Sie gut; spätere Partnerwechsel sind nicht möglich!

- Klausur am 25. 1. 2011

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-

unterschiede

Kovarianz/
Korrelation

Lineare

Regression

Mehr zu Tutoriumsaufgaben

- Es gibt **kein** regelmäßig stattfindendes mündliches Tutorium! Die Bearbeitung der Aufgaben erfolgt mit Hilfe von Moodle. Nur Studierende, die beim Bearbeiten erkennen lassen, dass sie große Schwierigkeiten haben, **müssen** dann an einem mündlichen Tutorium teilnehmen (dies gilt zumindest ab 3 KP).
- Sie müssen sich im Moodle-Kurs "Statistik 1" als Teilnehmer/in anmelden. Um sich allgemein bei Moodle anzumelden, nutzen Sie die Zugangsdaten, mit dem Sie auch Ihr Uni-Mail-Konto einsehen können. Die Anmeldung zu Moodle finden Sie unter <https://moodle.uni-siegen.de/> (Login ganz klein oben rechts). Für den Kurs selbst brauchen Sie keine gesonderten Zugangsdaten.

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-
unterschiede

Kovarianz/
Korrelation

Lineare

Regression

Statistik?

„ ...im materiellen Sinn die geordnete Menge von Informationen in Form empir. Zahlen („Statistiken“); im instrumentalen Sinn (Stat. Methoden) der Inbegriff der Verfahren, nach denen empir. Zahlen gewonnen, dargestellt, verarbeitet, analysiert und für Schlußfolgerungen, Prognosen und Entscheidungen verwendet werden.“

Der große Brockhaus Kompaktausgabe, Band 21,
18. Aufl. 1984, S. 29

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-
unterschiede

Kovarianz/
Korrelation

Lineare

Regression

...in dieser Vorlesung:

„ ...im materiellen Sinn die geordnete Menge von Informationen in Form empir. Zahlen („Statistiken“); **im instrumentalen Sinn (Stat. Methoden) der Inbegriff der Verfahren**, nach denen empir. Zahlen gewonnen, **dargestellt, verarbeitet, analysiert** und für **Schlußfolgerungen, Prognosen und Entscheidungen verwendet werden.**“

Der große Brockhaus Kompaktausgabe, Band 21,
18. Aufl. 1984, S. 29

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen- unterschiede

Kovarianz/ Korrelation

Lineare

Regression

Zahlen sind Daten!

„ ... **der Inbegriff der Verfahren**, nach denen empir. Zahlen gewonnen, **dargestellt, verarbeitet, analysiert** werden.“

- Die zu analysierenden Zahlen (oder Symbole) beziehen sich auf empirische Informationen → *Daten*.
- Es geht also nicht um abstrakte Zahlenspielerei – sondern darum, die in den Daten enthaltene Information bestmöglich zu nutzen.

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-
unterschiede

Kovarianz/
Korrelation

Lineare

Regression

Statistik ist ...

1. Schwer?

?????

Außer Grundrechenarten keine Kenntnisse nötig!

2. Für Analyse von Menschen und Gesellschaften
unangemessen?

?????

Wäre nur richtig, wenn behauptet würde, dass
Analyse **ausschließlich** auf Statistik beruhen soll
→ Statistik ist **Hilfsmittel**

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-
unterschiede

Kovarianz/
Korrelation

Lineare

Regression

Statistik ist ...

1. Wichtig!

Statistik durchzieht unser Leben. Ein kritisches Verständnis von Statistik (was kann sie, was kann sie nicht? Wie muss sie korrekt angewendet werden?) ist deshalb von großer Bedeutung.

2. beruflichen Erfolg versprechend!

Alle Untersuchungen zeigen, dass Absolventen mit guten Methodenkenntnissen keine Job-Probleme haben (im Gegenteil).

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-
unterschiede

Kovarianz/
Korrelation

Lineare

Regression

Ein Tipp von Gero von Randow

„Ein paar Bemerkungen zum Umgang mit Formeln. Zugegeben, Formeln sind die Geheimwaffe einer internationalen Verschwörung gegen Ihr Selbstbewußtsein. Aber am besten tun Sie so, als würde Ihnen das nichts ausmachen, das verwirrt den Gegner.“

Gero von Randow. Das Ziegenproblem. Denken in Wahrscheinlichkeiten. Rowohlt 1992, S. 17.

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen- unterschiede

Kovarianz/ Korrelation

Lineare Regression

Hilfsmittel

Lehrbücher: Siehe Vorlesungsplan

Tutorien:

- Die Aufgaben helfen zur Klausurvorbereitung; sie sind Klausurangaben ähnlich, meist etwas komplexer, nicht jedoch schwerer.
- Es finden außerdem bei Bedarf mündliche Tutorien statt (genaue Daten werden noch bekannt gegeben).
- → Das „Supertutorium“ ist kein Tutorium für Studierende, sondern die regelmäßige Besprechung des Prof's mit den TutorInnen!

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen- unterschiede

Kovarianz/ Korrelation

Lineare

Regression

Was macht (man mit) Statistik? I

→ Daten anschauen; siehe ausgeteiltes Blatt.

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-
unterschiede

Kovarianz/
Korrelation

Lineare

Regression

Was macht (man mit) Statistik? II

→ Daten visualisieren (oder sonst zusammenfassen)

| | | |
|-----|--|---------------|
| 7* | | 2 |
| 8* | | 0 |
| 9* | | |
| 10* | | |
| 11* | | 46 |
| 12* | | 0003445778899 |
| 13* | | 11236688 |
| 14* | | 0179 |

Daten zu Studiendauer; Quelle: Ausgeteilte Daten

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-
unterschiede

Kovarianz/
Korrelation

Lineare
Regression

Was macht (man mit) Statistik? III

Kennzahlen für die „zentrale Tendenz“ und die „Streuung“ (oder Variabilität) der Datenwerte errechnen.

Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen- unterschiede

Kovarianz/ Korrelation

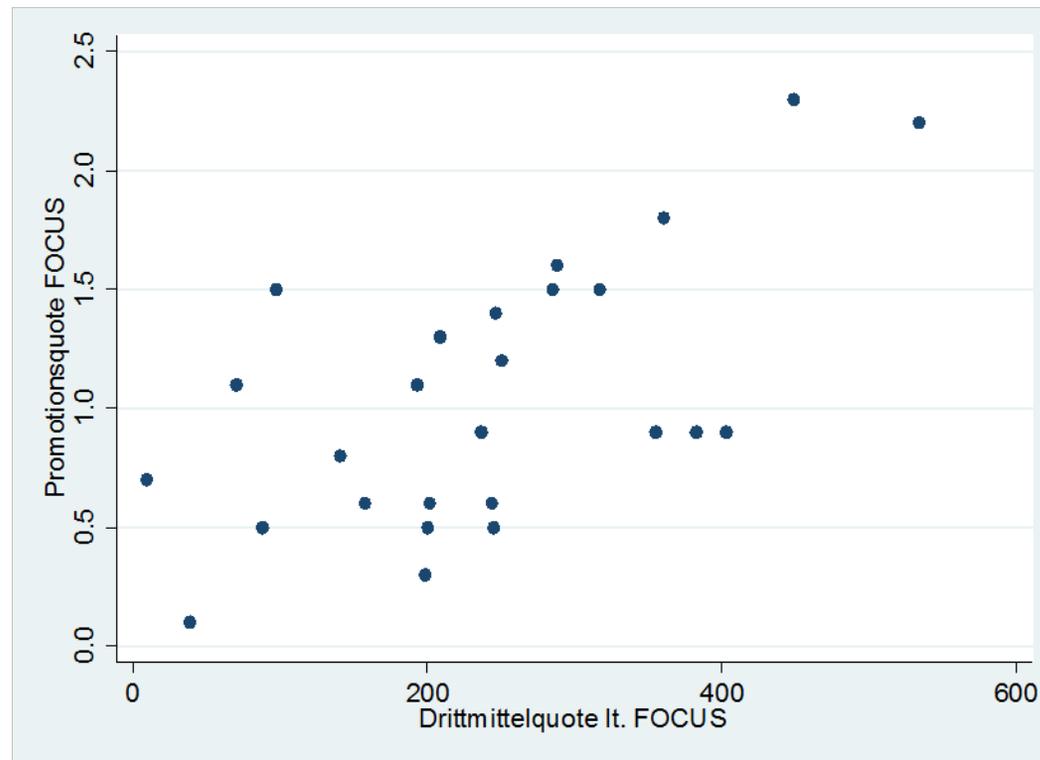
Lineare Regression

Lineare

Regression

Was macht (man mit) Statistik?

Zusammenhänge untersuchen



Grundlagen

Was ist Statistik?

Organisatorisches

Deskriptiv vs.

Inferenz

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen- unterschiede

Kovarianz/ Korrelation

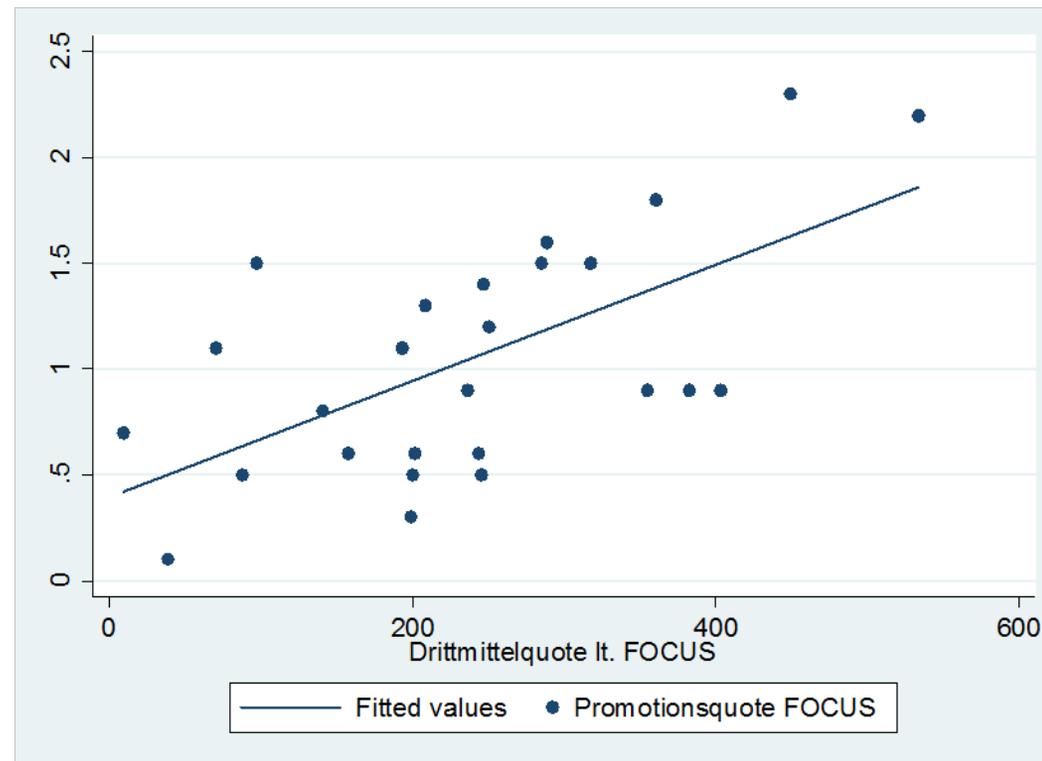
Lineare Regression

Lineare

Regression

Was macht (man mit) Statistik?

Zusammenhänge modellieren



Grundlagen

Was ist Statistik?

Organisatorisches

**Deskriptiv vs.
Inferenz**

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-

unterschiede

Kovarianz/
Korrelation

Lineare

Regression

Was macht Statistik?

1. Deskriptive Statistik

Daten „beschreiben“ → Daten und ihre Zusammenhänge durch Kennziffern (z.B. Mittelwerte), Graphiken und statistische Modelle auf das Wesentliche reduzieren.

Hinweis: In dieser Vorlesung werden einige Elemente der sog. „explorativen Datenanalyse“ der deskriptiven Statistik zugeordnet.

Grundlagen

Was ist Statistik?

Organisatorisches

**Deskriptiv vs.
Inferenz**

Urliste und

Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-

unterschiede

Kovarianz/
Korrelation

Lineare

Regression

Was macht Statistik?

2. Inferenzstatistik oder schließende Statistik

Wenn Stichprobe vorliegt – was sagen die Daten über die Grundgesamtheit? Mit welcher (Un-)Sicherheit ist unser Schluss auf die Grundgesamtheit behaftet?

Mit Inferenzstatistik beschäftigen wir uns in dieser Vorlesung nur kurz – aber intensiv!

Grundlagen

Was ist Statistik?
Organisatorisches
Deskriptiv vs.
Inferenz

**Urliste und
Datenmatrix**

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-
unterschiede

Kovarianz/
Korrelation

Lineare

Regression

Ein wenig Terminologie

- **Urliste**: Ausdruck v.a. statistischer Fachwissenschaftler für die Auflistung der Messwerte.
- **Datenmatrix**: Die Anordnung der Daten vor allem (aber nicht nur) in EDV-Programmen für die weitere Analyse.
- Die Messwerte (Daten) beziehen sich auf (ein oder mehrere) **Merkmale**.
- Die Merkmale weisen in aller Regel mehrere **Ausprägungen** auf (die Ausprägungen/Werte unterscheiden sich von einer Untersuchungseinheit zur nächsten). Dann spricht man auch von **Variablen**.
- Die Untersuchungseinheiten heißen auch **Merkmals-träger**.

Grundlagen

Was ist Statistik?
Organisatorisches
Deskriptiv vs.
Inferenz
Urliste und
Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-
unterschiede

Kovarianz/
Korrelation

Lineare

Regression

Messniveaus (Skalenniveaus)

- 1. Nominalskala:** Unterschiedliche Werte bedeuten „etwas Anderes“, sonst nichts.
- 2. Ordinalskala:** Werte bedeuten Rangordnung (erste, zweite, dritte), sagen aber nichts über Abstände.
- 3. Intervallskala:** Abstände sind inhaltlich interpretierbar (Unterschied zwischen 30°C und 20°C ist identisch mit Unterschied zwischen 20°C und 10°C).
- 4. Verhältnisskala (Ratioskala):** Verhältnisse zwischen Messwerten sind interpretierbar (Temperatur in Kelvin!).

Grundlagen

Was ist Statistik?
Organisatorisches
Deskriptiv vs.
Inferenz
Urliste und
Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen- unterschiede

Kovarianz/ Korrelation

Lineare

Regression

Messniveaus in der Praxis

„Mir fällt es momentan schwer, einen Sinn im Studium zu entdecken.“

1 = trifft voll und ganz zu

2

3

4

5 = trifft überhaupt nicht zu

Intervallskala? Ratioskala? Ordinalskala?

Summenwert aus fünf solcher Items?

Grundlagen

Was ist Statistik?
Organisatorisches
Deskriptiv vs.
Inferenz
Urliste und
Datenmatrix

Messniveaus

Häufigkeiten

Lagemaße

Streuung

Inferenzstatistik

Kreuztabellen

Gruppen-
unterschiede

Kovarianz/
Korrelation

Lineare

Regression

Einige weitere Begriffe

- Intervall- und Ratioskala werden auch als metrische Skalen bezeichnet, entsprechend spricht man von **metrischen Variablen** (oder metrischen Daten).
 1. Nominalskalierte und ordinale Variablen (Daten) werden auch als **kategoriale Variablen** (Daten) bezeichnet;
 2. Wir unterscheiden ferner zwischen
 - **Diskreten Merkmalen**
Ausprägungen sind abzählbar.
 - **Stetigen Merkmalen**
Im Prinzip sind alle Werte (evtl. innerhalb eines Intervalls) möglich.