

Name:	Studiengang: <input type="checkbox"/> B.A. <input type="checkbox"/> MA.
Vorname:	In FlexNow angemeldet: <input type="checkbox"/> Ja <input type="checkbox"/> Nein
Matrikelnummer:	
Studienfächer:	Fachsemester:

Allgemeine Hinweise:

1. Überprüfen Sie bitte, ob Sie alle Seiten der Klausurangabe vollständig erhalten haben (Gesamtzahl: **9**)
2. **Bearbeitungszeit: 90 Minuten**, maximal erreichbare **Punktzahl: 82**. Die jeweils erreichbare Punktzahl ist bei jeder Frage angegeben. Bitte teilen Sie Ihre Arbeitszeit entsprechend ein.
3. Denken Sie daran, die Daten oben einzutragen, **bevor** Sie mit der Bearbeitung beginnen.
4. Treffen Sie bitte die Auswahl Ihrer Antworten bei Multiple-Choice-Fragen **direkt** auf dieser Klausurangabe.
5. Verwenden Sie für die Beantwortung der Freitext-Fragen ebenfalls diese Klausurangabe. Sie können jederzeit auch die Rückseiten beschreiben, falls der Platz auf der Vorderseite nicht ausreichen sollte. Bitte geben Sie in jedem Fall an, auf welche Frage sich die Lösung jeweils bezieht.
6. Benutzen Sie keine Bleistifte, keine rot schreibenden Stifte und kein TippEx, o.ä.
7. Zugelassene Hilfsmittel: **KEINE**
8. Mobiltelefone sowie Computer am Arbeitsplatz - auch ausgeschaltet - sind **nicht zugelassen**.
9. Geben Sie keine mehrdeutigen (oder mehrere) Lösungen an. In solchen Fällen wird stets die Lösung mit der geringeren Punktzahl gewertet. Eine richtige und eine falsche Lösung ergeben also null Punkte.
10. Wenden Sie sich bei Unklarheiten in den Aufgabenstellungen immer an die Aufsichtsführenden. Hinweise und Hilfestellungen werden dann, falls erforderlich, offiziell für alle Teilnehmer durchgegeben.

Teil I) BENUTZERSTUDIEN (MAX. 41 PUNKTE INSGESAMT)

Aufgabe I.1)

a) Beurteilen Sie für jede der folgenden Aussagen zu (Fragebogen-basierten) Umfragen, ob diese wahr oder falsch ist: (5 Punkte)

Wahr		Falsch
<input type="checkbox"/>	Umfragen ermöglichen ein tiefes Verständnis einer bestimmten Situation.	<input type="checkbox"/>
<input type="checkbox"/>	Umfragen sind eine gute Methode, neue, wenig erforschte Themen zu explorieren.	<input type="checkbox"/>
<input type="checkbox"/>	Umfragen können benutzt werden, um Erklärungen für bestimmte Verhaltensweisen zu finden.	<input type="checkbox"/>
<input type="checkbox"/>	Mit Umfragen erhobene Daten sind objektiv und naturalistisch.	<input type="checkbox"/>
<input type="checkbox"/>	Im Vergleich zu Interviews erlauben Umfragen die Sammlung einer größeren Datenmenge in kürzerer Zeit.	<input type="checkbox"/>

b) Beurteilen Sie für jede der folgenden Aussagen, ob diese wahr oder falsch ist: (5 Punkte)

Das Ziehen einer Stichprobe als geschichtete Zufallsstichprobe (stratified random sampling)...

Wahr		Falsch
<input type="checkbox"/>	... ist der einfachen Zufallsstichprobe (standard random sampling) immer vorzuziehen	<input type="checkbox"/>
<input type="checkbox"/>	... ermöglicht umfangreichere Stichproben.	<input type="checkbox"/>
<input type="checkbox"/>	... ergibt eine Stichprobe, die gleichmäßig über bestimmte interessante Variablen verteilt ist.	<input type="checkbox"/>
<input type="checkbox"/>	... erlaubt die Anwendung mehrerer Schichtvariablen beim Erstellen der Stichprobe.	<input type="checkbox"/>

c) Sie machen ein Praktikum bei einer bekannten Firma in der Gegend. Ihre Aufgabe ist es, zu untersuchen, wie Mitarbeiter das Intranet der Firma zum Wissenserwerb und Erfahrungsaustausch benutzen. Sie entscheiden sich für eine Umfrage, weil Sie glauben, dass die Methode am besten geeignet ist, dieses Ziel zu erreichen. Nennen Sie drei Schichtvariablen (stratifying criteria), die Sie benutzen können, um eine geeignete geschichtete Zufallsstichprobe (stratified random sampling) für diese Fragestellung zu ziehen. (3 Punkte)

d) Die Firma hat 5.021 Mitarbeiter, die Intranetkonten haben. Sie ziehen eine Stichprobe mit einer Größe von 500 und schicken diesen Mitarbeitern einen Link zu einer Onlineumfrage. Sie erhalten 391 ausgefüllte Fragebögen. Wie groß ist Ihre Stichprobenbasis (sampling frame)? Umkreisen Sie die richtige Antwort: (2 Punkte)

- 391
- 5.021
- 5.900
- 500
- 391/500 (78.2%)

Aufgabe I.2)

a) Für welche der folgenden Daten wäre Grounded Theory eine passende Analyseverfahren? (5 Punkte)

Passend		Unpassend
<input type="checkbox"/>	Interviewtranskripte	<input type="checkbox"/>
<input type="checkbox"/>	Beobachternotizen	<input type="checkbox"/>
<input type="checkbox"/>	Umfragen mit Likertskalen	<input type="checkbox"/>
<input type="checkbox"/>	Anfragen einer Websuche	<input type="checkbox"/>
<input type="checkbox"/>	Schrittzählerdaten	<input type="checkbox"/>

b) Erläutern Sie den Unterschied zwischen induktiver und deduktiver Kodierung im Rahmen der Grounded Theory. (2 Punkte)

Aufgabe I.3)

a) Nennen Sie zwei Gründe, die für die Evaluation eines Informationssystems mit Hilfe einer Laborstudie sprechen. (2 Punkte)

b) Im Kurs haben wir eine Reihe potentieller Quellen für das Problem der Verzerrung (bias) in Laborstudien besprochen. Zwei dieser Quellen waren Ermüdung (fatigue) (d.h. dass Aufgaben am Ende eines Tests aufgrund von Müdigkeit weniger gut erledigt werden) und teilnehmerspezifische Charakteristika (participant characteristics) (d.h. dass Teilnehmer an einer Studie z.B. bestimmte Fähigkeiten oder Erfahrungen etc. haben, die das Ergebnis beeinflussen).

Wie können die Auswirkungen dieser Verzerrungsquellen eingedämmt bzw. kontrolliert werden? Nennen Sie je einen Aspekt für Müdigkeit und einen für teilnehmerspezifische Charakteristika. (4 Punkte)

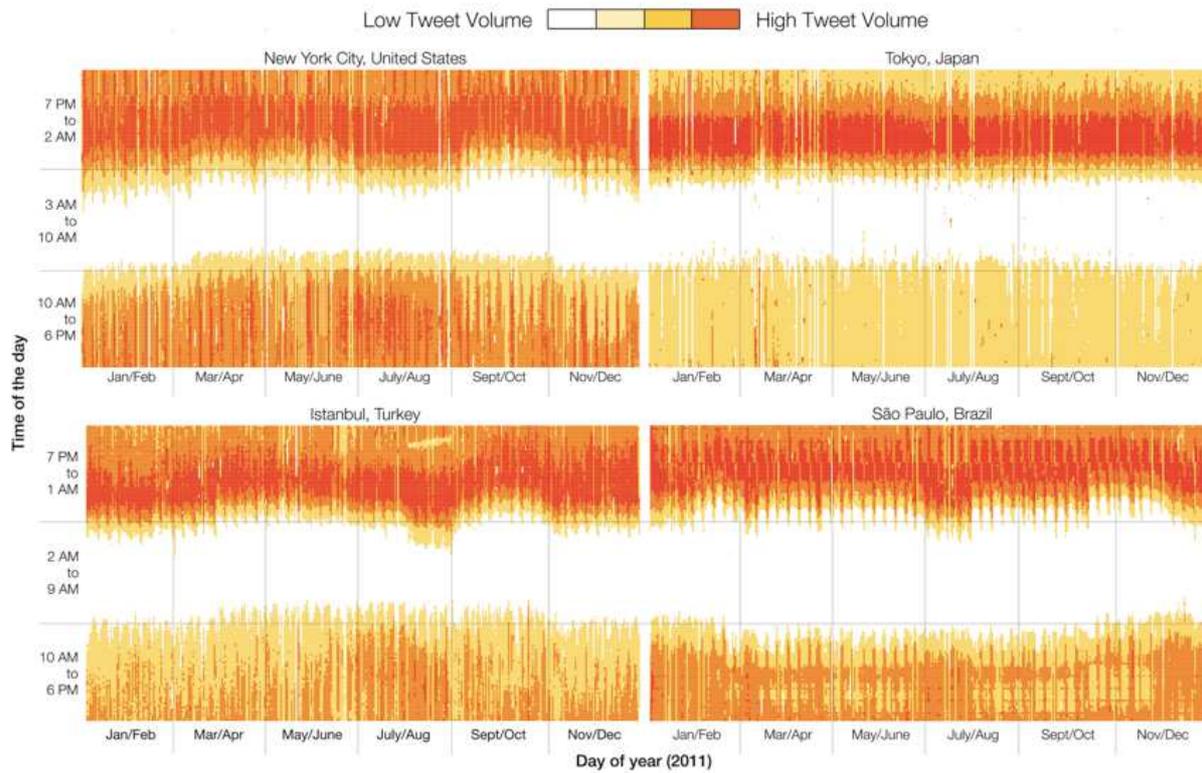
c) Nennen Sie drei Vor- und drei Nachteile des Experimentaldesigns *within groups* und *between groups*. (6 Punkte)

d) In welcher Situation ist das eine Design dem anderen vorzuziehen? Begründen Sie Ihre Einschätzung. (2 Punkte)

Aufgabe I.4)

Die nachstehende Graphik stellt die Menge an Tweets dar, die zu verschiedenen Tageszeiten in vier Städten veröffentlicht wurden: New York, Tokio, Istanbul und Sao Paolo. Die Farbintensität spiegelt die Menge der Tweets wider, d.h. je dunkler, desto mehr Tweets. Die gesamte Graphik basiert auf naturalistischen Log-Daten von Twitter.

Nennen Sie zwei Dinge, die Ihnen am Verhalten der Twitternutzer in verschiedenen Städten auffallen. (5 Punkte)



TEIL II STATISTISCHE DATENAUSWERTUNG (MAX. 41 PUNKTE INSGESAMT)

Aufgabe II.1) Skalierung (5 Punkte)

Beurteilen Sie für jede der folgenden Aussagen, ob diese wahr oder falsch ist.

Wahr		Falsch
<input type="checkbox"/>	Die Überführung von intervallskalierten Daten in nominalskalierte Daten ist mit Informationsverlust verbunden.	<input type="checkbox"/>
<input type="checkbox"/>	Intervallskalierte Likert-Skalen zeichnen sich dadurch aus, dass benachbarte Skalenpunkte verschieden weit voneinander entfernt sind.	<input type="checkbox"/>
<input type="checkbox"/>	Ist ein Merkmal auf einer Nominalskala messbar, so ist dieses Merkmal auch auf einer Ordinalskala messbar.	<input type="checkbox"/>
<input type="checkbox"/>	Zur Berechnung des Medians ist mindestens Ordinalskalenniveau nötig.	<input type="checkbox"/>
<input type="checkbox"/>	Der Modalwert kann ausschließlich auf Verhältnisskalenniveau bestimmt werden.	<input type="checkbox"/>

Aufgabe II.2) Korrelationsrechnung (4 Punkte)

Gegeben seien zwei auf metrischem Skalenniveau gemessene Merkmale X u. Y. Im Kurs haben wir zwei Korrelationskoeffizienten besprochen: den Korrelationskoeffizienten von Pearson r und Spearmans Rho r_s .

Beurteilen Sie für jede der folgenden Aussagen, ob diese wahr oder falsch ist.

Wahr		Falsch
<input type="checkbox"/>	Alle nachstehenden Aussagen sind falsch.	<input type="checkbox"/>
<input type="checkbox"/>	Aus $r = -1$ folgt immer $r_s = -1$.	<input type="checkbox"/>
<input type="checkbox"/>	Aus $r_s > 0$ folgt immer $r > 0$.	<input type="checkbox"/>
<input type="checkbox"/>	Ist $r = r_s = 0$ dann liegt kein Zusammenhang zwischen X und Y vor.	<input type="checkbox"/>

Aufgabe II.3) Lineare Regression (6 Punkte)

Eine informationswissenschaftliche Studie zeigt, dass es umso länger dauert, eine Datei auf der eigenen Festplatte wiederzufinden, je älter diese ist. Im Rahmen dieser Untersuchung wurden das Alter A einer Datei in Tagen und die nötige Zeit Z , diese wieder zu finden, in Sekunden bestimmt. Aus diesen Untersuchungsdaten ergab sich das folgende lineare Regressionsmodell:

$$\hat{Z} = 40 + 0,997 * A$$

Beurteilen Sie für jede der folgenden Aussagen, ob diese wahr oder falsch ist.

Wahr**Falsch**

- | | | |
|--------------------------|---|--------------------------|
| <input type="checkbox"/> | Der Koeffizient 40 gibt die Steigung der Regressionsgeraden an. | <input type="checkbox"/> |
| <input type="checkbox"/> | Eine gerade erstellte Datei wird in 40 Sekunden wiedergefunden. | <input type="checkbox"/> |

Der Wert des Koeffizienten 0,997 wurde mit Hilfe der nachstehenden Beziehung geschätzt.

$$\hat{\beta} = \frac{\sum_{i=1}^n (a_i - \bar{a})(z_i - \bar{z})}{\sum_{i=1}^n (z_i - \bar{z})^2}$$

Ist die folgende Aussage wahr oder falsch?

- | | | |
|--------------------------|--|--------------------------|
| <input type="checkbox"/> | Die arithmetischen Mittel der Realisierungen der beiden Zufallsvariablen waren ungefähr gleich groß. | <input type="checkbox"/> |
|--------------------------|--|--------------------------|

Aufgabe II.4) Schätzer u. Verteilungen (6 Punkte)

Beurteilen Sie für jede der folgenden Aussagen, ob diese wahr oder falsch ist.

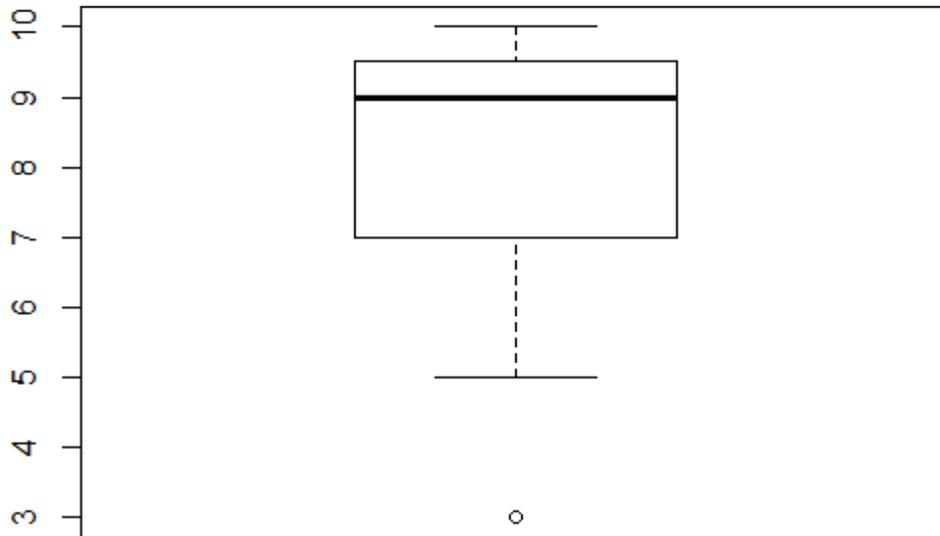
Wahr**Falsch**

- | | | |
|--------------------------|--|--------------------------|
| <input type="checkbox"/> | Normalverteilungen werden durch Mittelwert und Varianz definiert. | <input type="checkbox"/> |
| <input type="checkbox"/> | Je größer der Bias eines Parameterschätzers desto besser nähert dieser den wahren Wert des Parameters in der Grundgesamtheit an. | <input type="checkbox"/> |
| <input type="checkbox"/> | Die Stichprobenvarianz unterschätzt die Populationsvarianz systematisch. | <input type="checkbox"/> |

Aufgabe II.5) 5-Punkte-Zusammenfassung (8 Punkte)

Welche Aussagen können Sie auf Grund des dargestellten Boxplots über die Verteilung der Daten machen bezogen auf

- die Links- oder Rechtssteilheit der Verteilung?
- das Minimum?
- das Maximum?
- den Median?
- den Wert 7?



Aufgabe II.6) t-Test (12 Punkte)

Nennen Sie 3 von 4 Voraussetzungen, die für die sinnvolle Anwendung eines t-Tests für 2 unabhängige Stichproben gegeben sein müssen.

- a) Erläutern Sie jede der von Ihnen genannten Voraussetzungen kurz bezüglich der Umsetzung.
- b) Welche Gegenmaßnahmen können ergriffen werden, wenn eine der von Ihnen genannten Voraussetzungen verletzt wurde?
- c) Was besagt das Signifikanzniveau bei statistischen Tests allgemein?
- d) Was schließen Sie daraus, wenn ein zweiseitiger t-Test ein signifikantes Ergebnis hat?