

# Willkommen zum 2. Teil des Projektkurs Data Science und Business Analytics

DR. ALEXANDER RIEBER

 [\\_ALEXRIEBER](#)

 [ALEXANDER.RIEBER@UNI-ULM.DE](mailto:ALEXANDER.RIEBER@UNI-ULM.DE)

# Mit wem haben Sie es zu tun?

## Ihr Dozent und Übungsleiter

### Dozent

Dr. Alexander Rieber

✉ [alexander.rieber@uni-ulm.de](mailto:alexander.rieber@uni-ulm.de)

Büro: Helmholtzstraße 18, Raum 1.22

### Übungsleiter

Julius Düker

✉ [\\_julius.dueker@uni-ulm.de](mailto:_julius.dueker@uni-ulm.de)

Büro: Helmholtzstraße 18, Raum 1.10

# Ihre Tutoren

## Tutorin

Lara Santak

✈ [lara.santak@uni-ulm.de](mailto:lara.santak@uni-ulm.de)

## Tutor

Nils Hönig

✈ [nils.hoenig@uni-ulm.de](mailto:nils.hoenig@uni-ulm.de)

## Tutor

Jonas Grüneberg

✈ [\\_jonas.grueneberg@uni-ulm.de](mailto:_jonas.grueneberg@uni-ulm.de)

## Tutor

Alexander Wilhelm

✈ [alexander.wilhelm@uni-ulm.de](mailto:alexander.wilhelm@uni-ulm.de)

# Vorlesungsdetails

# Aufbau der Veranstaltung

- + Interaktive Vorlesung mit Übungsaufgaben
  - + In den erste 6 Wochen
  - + Mittwochs von 12:15 Uhr - 13:45 Uhr in H14
  - + Freitags von 10:15 Uhr - 11:45 Uhr in H14
- + Anonymer Chat in der Veranstaltung (mit Umfragen)
- + RTutor Übungsaufgaben um Inhalte zu vertiefen
- + Ein gemeinsames Projekt mit dem Dozenten
- + Zwei selbstständige Projekte
- + Peer Review der selbstständigen Projekte
- + Probeklausur am 20.05.2022 in der Vorlesung! (Anwesenheitspflicht da Vorleistung)

# Wo finden Sie die Kursmaterialien?

Unserer Kommunikation mit Ihnen erfolgt über unsere Moodle Seite

Auf dieser Moodle Seite finden Sie:

# Wo finden Sie die Kursmaterialien?

Unserer Kommunikation mit Ihnen erfolgt über unsere Moodle Seite



Auf dieser Moodle Seite finden Sie:

+ Vorlesungsplaner per Excel 

# Wo finden Sie die Kursmaterialien?

Unserer Kommunikation mit Ihnen erfolgt über unsere Moodle Seite

Auf dieser Moodle Seite finden Sie:




- + Vorlesungsplaner per Excel 
- + Vorlesungsfolien 



# Wo finden Sie die Kursmaterialien?

Unserer Kommunikation mit Ihnen erfolgt über unsere Moodle Seite




Auf dieser Moodle Seite finden Sie:

- + Vorlesungsplaner per Excel 
- + Vorlesungsfolien 
- + Lehrvideos 

# Wo finden Sie die Kursmaterialien?

Unserer Kommunikation mit Ihnen erfolgt über unsere Moodle Seite





Auf dieser Moodle Seite finden Sie:

- + Vorlesungsplaner per Excel 
- + Vorlesungsfolien 
- + Lehrvideos 
- + RTutor Problem Sets -> RTutor

# Wo finden Sie die Kursmaterialien?

Unserer Kommunikation mit Ihnen erfolgt über unsere Moodle Seite







Auf dieser Moodle Seite finden Sie:

- + Vorlesungsplaner per Excel 
- + Vorlesungsfolien 
- + Lehrvideos 
- + RTutor Problem Sets -> RTutor
- + Aufgabenstellungen für die einzelnen Projekte 

# Wo finden Sie die Kursmaterialien?

Unserer Kommunikation mit Ihnen erfolgt über unsere Moodle Seite








Auf dieser Moodle Seite finden Sie:

- + Vorlesungsplaner per Excel 
- + Vorlesungsfolien 
- + Lehrvideos 
- + RTutor Problem Sets -> RTutor
- + Aufgabenstellungen für die einzelnen Projekte 
- + Forum für Fragen   !?

# Wo finden Sie die Kursmaterialien?

Unserer Kommunikation mit Ihnen erfolgt über unsere Moodle Seite








Auf dieser Moodle Seite finden Sie:

- + Vorlesungsplaner per Excel 
- + Vorlesungsfolien 
- + Lehrvideos 
- + RTutor Problem Sets -> RTutor
- + Aufgabenstellungen für die einzelnen Projekte 
- + Forum für Fragen   !?
- + Zoom Räume für die Kombi-Vorlesung 

# Wo finden Sie die Kursmaterialien?

Unserer Kommunikation mit Ihnen erfolgt über unsere Moodle Seite

Auf dieser Moodle Seite finden Sie:








- + Vorlesungsplaner per Excel 
- + Vorlesungsfolien 
- + Lehrvideos 
- + RTutor Problem Sets -> RTutor
- + Aufgabenstellungen für die einzelnen Projekte 
- + Forum für Fragen   !?
- + Zoom Räume für die Kombi-Vorlesung 

Bis zum 18. Mai werden Vorlesungen in Form von Präsenzvorlesungen, Lehrvideos, Tutorials und RTutor Problem Sets stattfinden. Weiterhin wird eine Projektarbeit innerhalb der Vorlesung mit ihnen gemeinsam erarbeitet. (Gibt 10% der Endnote)

# Wo finden Sie die Kursmaterialien?

Unserer Kommunikation mit Ihnen erfolgt über unsere Moodle Seite

Auf dieser Moodle Seite finden Sie:

- + Vorlesungsplaner per Excel 
- + Vorlesungsfolien 
- + Lehrvideos 
- + RTutor Problem Sets -> RTutor
- + Aufgabenstellungen für die einzelnen Projekte 
- + Forum für Fragen   !?
- + Zoom Räume für die Kombi-Vorlesung 

Bis zum 18. Mai werden Vorlesungen in Form von Präsenzvorlesungen, Lehrvideos, Tutorials und RTutor Problem Sets stattfinden. Weiterhin wird eine Projektarbeit innerhalb der Vorlesung mit ihnen gemeinsam erarbeitet. (Gibt 10% der Endnote)

Anschließend gibt es zwei selbstständige Projektarbeiten, welche in die Endnote einfließen werden.

# Interaktive Vorlesung

- + Interaktive Vorlesung mit integrierten Übungsaufgaben
- + Geblockte Vorlesungseinheiten zu Beginn des Semesters
- + Synchrone Lehre, aufgezeichnet und gestreamed via Zoom, dient der praktischen Anwendung aus den Lehrvideos
  - + **Anders als im letzten Semester** werden diese Vorlesungseinheiten hauptsächlich theoretische Inhalte präsentieren
  - + Case-Study wird in der Einheit zur linearen Regression näher beleuchten
  - + Kausale Effekte werden anhand eines Experiments und einer Instrumentalvariablenschätzung aufgezeigt
- + RTutor Problem Sets und Tutorials sollen zum "learning-by-doing" anregen
- + Vorlesungs- und Übungsunterlagen werden zu Beginn der jeweiligen Vorlesungswoche hochgeladen



Was Sie in dieser Vorlesung lernen

Was haben Sie im ersten Teil des Projektkurses gelernt?

# Was haben Sie im ersten Teil des Projektkurses gelernt?

- (Große) Datenmengen einlesen, bearbeiten und aus verschiedenen Datenquellen zusammenführen

# Was haben Sie im ersten Teil des Projektkurses gelernt?

- ✚ (Große) Datenmengen einlesen, bearbeiten und aus verschiedenen Datenquellen zusammenführen
- ✚ Mit diversen Problemen rund um die Datenaufbereitung umgehen

# Was haben Sie im ersten Teil des Projektkurses gelernt?

- ✚ (Große) Datenmengen einlesen, bearbeiten und aus verschiedenen Datenquellen zusammenführen
- ✚ Mit diversen Problemen rund um die Datenaufbereitung umgehen
- ✚ Zusammenarbeit im Team + Versionierung über Github

# Was haben Sie im ersten Teil des Projektkurses gelernt?

- ✚ (Große) Datenmengen einlesen, bearbeiten und aus verschiedenen Datenquellen zusammenführen
- ✚ Mit diversen Problemen rund um die Datenaufbereitung umgehen
- ✚ Zusammenarbeit im Team + Versionierung über Github
- ✚ Regeln zur Erstellung geeigneter Grafiken und Tabellen um ihre Analyse zu untermauern

# Was haben Sie im ersten Teil des Projektkurses gelernt?

- ✚ (Große) Datenmengen einlesen, bearbeiten und aus verschiedenen Datenquellen zusammenführen
- ✚ Mit diversen Problemen rund um die Datenaufbereitung umgehen
- ✚ Zusammenarbeit im Team + Versionierung über Github
- ✚ Regeln zur Erstellung geeigneter Grafiken und Tabellen um ihre Analyse zu untermauern
- ✚ Transparente und reproduzierbare Analysen mit RMarkdown erstellen

# Was haben Sie im ersten Teil des Projektkurses gelernt?

- ✚ (Große) Datenmengen einlesen, bearbeiten und aus verschiedenen Datenquellen zusammenführen
- ✚ Mit diversen Problemen rund um die Datenaufbereitung umgehen
- ✚ Zusammenarbeit im Team + Versionierung über Github
- ✚ Regeln zur Erstellung geeigneter Grafiken und Tabellen um ihre Analyse zu untermauern
- ✚ Transparente und reproduzierbare Analysen mit RMarkdown erstellen
- ✚ Selbstständig nach Lösungen für Probleme bei empirischen Problemen zu suchen



# Was haben Sie im ersten Teil des Projektkurses gelernt?

- ✚ (Große) Datenmengen einlesen, bearbeiten und aus verschiedenen Datenquellen zusammenführen
- ✚ Mit diversen Problemen rund um die Datenaufbereitung umgehen
- ✚ Zusammenarbeit im Team + Versionierung über Github
- ✚ Regeln zur Erstellung geeigneter Grafiken und Tabellen um ihre Analyse zu untermauern
- ✚ Transparente und reproduzierbare Analysen mit RMarkdown erstellen
- ✚ Selbstständig nach Lösungen für Probleme bei empirischen Problemen zu suchen
- ✚ (Kritische) Bewertung anderer Gruppen über einen Review Report

# Was haben Sie im ersten Teil des Projektkurses gelernt?

- + (Große) Datenmengen einlesen, bearbeiten und aus verschiedenen Datenquellen zusammenführen
- + Mit diversen Problemen rund um die Datenaufbereitung umgehen
- + Zusammenarbeit im Team + Versionierung über Github
- + Regeln zur Erstellung geeigneter Grafiken und Tabellen um ihre Analyse zu untermauern
- + Transparente und reproduzierbare Analysen mit RMarkdown erstellen
- + Selbstständig nach Lösungen für Probleme bei empirischen Problemen zu suchen
- + (Kritische) Bewertung anderer Gruppen über einen Review Report

Das letzte Semester war eher technisch (programmierlastig) ausgelegt und auf deskriptive Analysen beschränkt.

Was werden Sie im zweiten Teil des  
Projektkurses lernen?

# Was werden Sie im zweiten Teil des Projektkurses lernen?

✚ Wahrscheinlichkeitsrechnung (kurz)

# Was werden Sie im zweiten Teil des Projektkurses lernen?

- + Wahrscheinlichkeitsrechnung (kurz)
- + (Multivariate) Lineare Regression

# Was werden Sie im zweiten Teil des Projektkurses lernen?

- + Wahrscheinlichkeitsrechnung (kurz)
- + (Multivariate) Lineare Regression
- + Interpretation von Schätzern

# Was werden Sie im zweiten Teil des Projektkurses lernen?

- + Wahrscheinlichkeitsrechnung (kurz)
- + (Multivariate) Lineare Regression
- + Interpretation von Schätzern
- + Kausale Zusammenhänge aufdecken
  - + In Experimenten
  - + In Beobachtungsdaten

# Was werden Sie im zweiten Teil des Projektkurses lernen?

- + Wahrscheinlichkeitsrechnung (kurz)
- + (Multivariate) Lineare Regression
- + Interpretation von Schätzern
- + Kausale Zusammenhänge aufdecken
  - + In Experimenten
  - + In Beobachtungsdaten
- + Konzept der Kausalität in Beobachtungsdaten mittels directed acyclical graphs (DAG)



# Was werden Sie im zweiten Teil des Projektkurses lernen?

- + Wahrscheinlichkeitsrechnung (kurz)
- + (Multivariate) Lineare Regression
- + Interpretation von Schätzern
- + Kausale Zusammenhänge aufdecken
  - + In Experimenten
  - + In Beobachtungsdaten
- + Konzept der Kausalität in Beobachtungsdaten mittels directed acyclical graphs (DAG)
- + Instrumentalvariablenschätzung

# Was werden Sie im zweiten Teil des Projektkurses lernen?

- + Wahrscheinlichkeitsrechnung (kurz)
- + (Multivariate) Lineare Regression
- + Interpretation von Schätzern
- + Kausale Zusammenhänge aufdecken
  - + In Experimenten
  - + In Beobachtungsdaten
- + Konzept der Kausalität in Beobachtungsdaten mittels directed acyclical graphs (DAG)
- + Instrumentalvariablenschätzung

Dieses Semester werden Sie ihre ökonometrischen Kenntnisse vertiefen!

# Ziele für dieses Semester

Nach diesem Semester werden Sie in der Lage sein:

# Ziele für dieses Semester

Nach diesem Semester werden Sie in der Lage sein:

- ✚ Regressionstabellen zu lesen und die jeweiligen Schätzer zu interpretieren

# Ziele für dieses Semester

Nach diesem Semester werden Sie in der Lage sein:

- + Regressionstabellen zu lesen und die jeweiligen Schätzer zu interpretieren
- + Die (statistische) Signifikanz von Schätzern zu beurteilen

# Ziele für dieses Semester

Nach diesem Semester werden Sie in der Lage sein:

- + Regressionstabellen zu lesen und die jeweiligen Schätzer zu interpretieren
- + Die (statistische) Signifikanz von Schätzern zu beurteilen
- + Hypothesen abzuleiten und eigene Fragenstellungen mittels statistischer Analysen zu beantworten

# Ziele für dieses Semester

Nach diesem Semester werden Sie in der Lage sein:

- + Regressionstabellen zu lesen und die jeweiligen Schätzer zu interpretieren
- + Die (statistische) Signifikanz von Schätzern zu beurteilen
- + Hypothesen abzuleiten und eigene Fragenstellungen mittels statistischer Analysen zu beantworten
- + Zu verstehen, wann Sie ihre Ergebnisse kausal interpretieren können

# Übungsaufgaben mit RTutor



# Wie Sie die Vorlesungsinhalte vertiefen

- + Wöchentliche RTutor Problem Sets, welche die Unterrichtseinheit aufgreifen und vertiefen
- + *Individuelle Abgabe der RTutor Problem Sets als **Vorleistung***
  - + Insgesamt 3 RTutor Problem Sets müssen bearbeitet werden
  - + In jedem RTutor Problem Set müssen Sie mind. 80% der Gesamtpunkte erreichen

# Wie Sie die Vorlesungsinhalte vertiefen

- + Wöchentliche `RTutor` Problem Sets, welche die Unterrichtseinheit aufgreifen und vertiefen
- + *Individuelle Abgabe der `RTutor` Problem Sets als **Vorleistung***
  - + Insgesamt 3 `RTutor` Problem Sets müssen bearbeitet werden
  - + In jedem `RTutor` Problem Set müssen Sie mind. 80% der Gesamtpunkte erreichen

Die `RTutor` Problem Sets behandeln vorlesungsbegleitend folgende Inhalte:

- + Regressionsanalysen in R
- + Kausalität in Experimenten
- + Kausalität in Beobachtungsdaten

# Teamzusammensetzung

# Teamzusammensetzung

Im letzten Semester gab es einige Stimmen, die eine andere Gruppenzuteilung angeregt hatten.

Daher gilt in diesem Semester: Sie dürfen sich ihre Teammitglieder selbst aussuchen!

# Teamzusammensetzung

Im letzten Semester gab es einige Stimmen, die eine andere Gruppenzuteilung angeregt hatten.

Daher gilt in diesem Semester: Sie dürfen sich ihre Teammitglieder selbst aussuchen!

Was benötigen wir hierfür?

- + Auf Moodle haben wir eine Gruppenwahl erstellt. Dort können Sie sich in Gruppen von drei Personen selbstständig zusammenfinden.
  - + Gruppenwahl bis spätestens zum 08.05.2022
- + Die Gruppen können sich anschließend auf einen Termin für das Tutorium bewerben (mehrere Termine stehen zur Auswahl)

# Teamzusammensetzung

Im letzten Semester gab es einige Stimmen, die eine andere Gruppenzuteilung angeregt hatten.

Daher gilt in diesem Semester: Sie dürfen sich ihre Teammitglieder selbst aussuchen!

Was benötigen wir hierfür?

- + Auf Moodle haben wir eine Gruppenwahl erstellt. Dort können Sie sich in Gruppen von drei Personen selbstständig zusammenfinden.
  - + Gruppenwahl bis spätestens zum 08.05.2022
- + Die Gruppen können sich anschließend auf einen Termin für das Tutorium bewerben (mehrere Termine stehen zur Auswahl)

Voraussetzung zur Teilnahme am 2. Teil des Projektkurses in diesem SoSe:

- + Erster Teil des Projektkurses im WiSe 21/22 bestanden

Unterstützungsangebote

# Wo bekommen Sie Hilfe?

- + Auf Moodle gibt es wieder ein Diskussionsforum, scheuen Sie sich nicht dort ihre Fragen zu stellen!
  - + Bitte benutzen Sie dieses Forum, bevor Sie eine Mail an den Tutor, Übungsleiter oder Dozent stellen
  - + Wir haben einige Beispiele aufgearbeitet, wie Sie Fragen stellen können, auf die Sie schnell antworten erhalten.  
[Das Dokument finden Sie hier](#)

Insbesondere während der Vorlesungsphase ist das Moodle Forum sehr wichtig um allgemeine Fragen zu klären!



# Wo bekommen Sie Hilfe?

- + Auf Moodle gibt es wieder ein Diskussionsforum, scheuen Sie sich nicht dort ihre Fragen zu stellen!
  - + Bitte benutzen Sie dieses Forum, bevor Sie eine Mail an den Tutor, Übungsleiter oder Dozent stellen
  - + Wir haben einige Beispiele aufgearbeitet, wie Sie Fragen stellen können, auf die Sie schnell antworten erhalten.  
[Das Dokument finden Sie hier](#)

Insbesondere während der Vorlesungsphase ist das Moodle Forum sehr wichtig um allgemeine Fragen zu klären!

- + Es finden parallel zu den Projektarbeiten Tutorien statt, wobei jede Gruppe einen Termin pro Woche **heraussuchen kann**

# Tutorium

- + Ab der 6. Vorlesungswoche gibt es ein vorlesungsbegleitendes Tutorium.
  - + Die Termine werden auf der Homepage bekannt gegeben
  - + Jede Gruppe kann ihren präferierten Termin angeben
- + Die Tutorien werden wöchentlich als Live-Tutorien stattfinden, hierfür haben wir Gruppenarbeitsräume reserviert
- + Die Tutoren und die Tutorin beantworten auch ihre Fragen auf Moodle
  - + Bitte stellen Sie ihre Fragen im Forum, wenn ihr Problem allgemeiner Natur ist

# Tutorium

- + Ab der 6. Vorlesungswoche gibt es ein vorlesungsbegleitendes Tutorium.
  - + Die Termine werden auf der Homepage bekannt gegeben
  - + Jede Gruppe kann ihren präferierten Termin angeben
- + Die Tutorien werden wöchentlich als Live-Tutorien stattfinden, hierfür haben wir Gruppenarbeitsräume reserviert
- + Die Tutoren und die Tutorin beantworten auch ihre Fragen auf Moodle
  - + Bitte stellen Sie ihre Fragen im Forum, wenn ihr Problem allgemeiner Natur ist

Das Tutorium soll hauptsächlich dazu dienen Sie bei ihren individuellen Projektausarbeitungen zu unterstützen!

Mit welchen Daten bekommen *Sie* es  
zu tun?

# Case-Study: Statistische Analyse der Verschuldung in Deutschland

# Case-Study: Statistische Analyse der Verschuldung in Deutschland

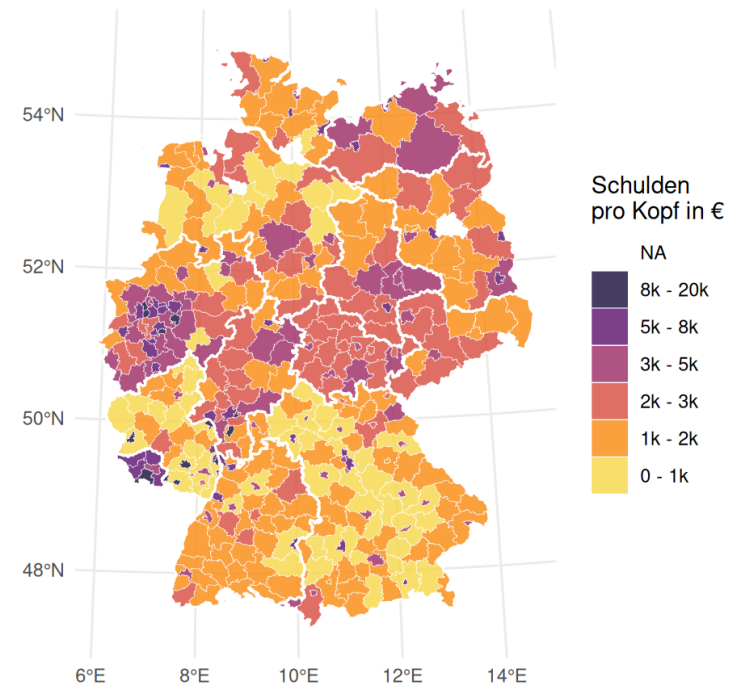
In der Case-Study widmen Sie sich der Frage:

Gibt es einen Zusammenhang zwischen den Verschuldung eines Landkreises und dessen Arbeitslosenquote?

**Fokus in diesem Semester:** Statistische Analyse mittels linearer Regression

Zur Beantwortung dieser Frage werten Sie u.a. Informationen vom statistischen Bundesamt auf Landkreisebene aus.

Wie verschuldet sind die deutschen Landkreise?  
Öffentliche Schulden pro Kopf in 2017



# Projekt 4: Bildungsrendite

# Projekt 4: Bildungsrendite

In diesem Projekt widmen Sie sich der Frage:

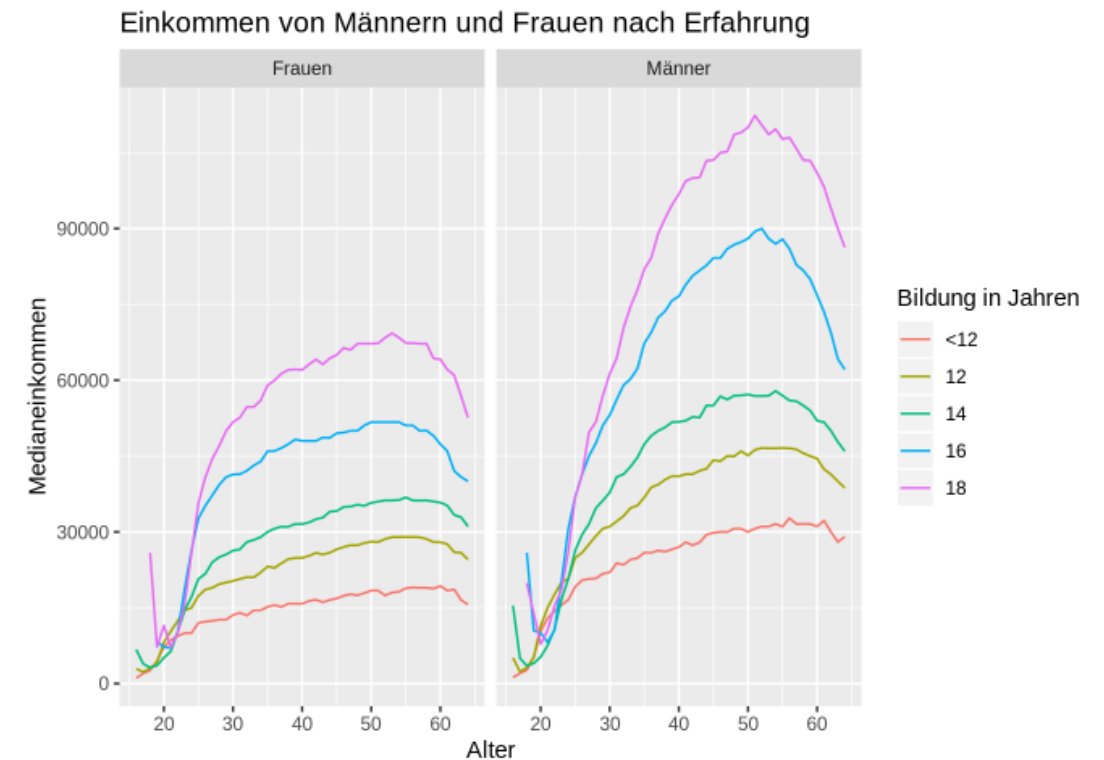
**Welche Auswirkung hat Bildung auf das spätere Einkommen?**

Frage nicht trivial, da es mehrere sogenannte Confounder gibt:

- + Hintergrund, Motivation, Fähigkeiten
- + **Lösung:** Verwenden von Instrumentalvariablenschätzern

Weiterhin:

- + Nutzung eines API um auf die amerikanischen ACS Daten zuzugreifen





# Projekt 5: Bewertungen von Hotels auf Booking

# Projekt 5: Bewertungen von Hotels auf Booking

In diesem Projekt widmen Sie sich der Frage:

Kommen manche Hotels besser durch die Corona Pandemie als andere? Woran liegt das?

- + Zur Beantwortung dieser Frage **scrapen Sie selbstständig** die Webseiten von Booking.com zu bestimmten Hotels und werten die Reviews der Hotelbesucher statistisch aus.

The screenshot shows the Booking.com homepage. At the top, there's a navigation bar with the Booking.com logo, currency (EUR), language (German), and links for 'Ihre Unterkunft anmelden', 'Registrieren', and 'Anmelden'. Below this is a menu with icons for 'Aufenthalte', 'Flüge', 'Flug + Hotel', 'Mietwagen', 'Attraktionen', and 'Flughafentaxis'. The main heading is 'Finden Sie das perfekte Hotel bei Booking.com', followed by the tagline 'Alles von 5-Sterne-Luxushotels bis hin zu günstigen Zimmern'. A search bar contains the text 'Wohin reisen Sie?', dates 'Mi, 20 Apr. — Sa, 23 Apr.', and guest information '2 Erwachsene · 0 Kinder · 1 Zimmer'. A 'Suche' button is on the right. Below the search bar is a checkbox 'Ich reise geschäftlich'. A warning box states: 'Bitte informieren Sie sich über Reiseeinschränkungen. Möglicherweise sind Reisen nur für bestimmte Zwecke erlaubt und insbesondere touristische Reisen sind unter Umständen nicht gestattet. Mehr erfahren'. Below this is a breadcrumb trail 'Startseite > Alle Hotels'. The section 'Last-minute-Hotels in der Nähe' is highlighted, with the text 'Finden Sie ein tolles Angebot für ein Hotel für heute Nacht oder für eine bevorstehende Reise'. Two buttons are present: 'Hotels für heute Nacht' and 'Hotels für dieses Wochenende'. The first hotel listed is 'Gasthof Hotel Zum Hirsch\*\*\*S' in Kirchen, with a rating of 9.0 and a price starting at €85 for today. The second hotel is 'Köhlers Krone' in Ehingen, with a rating of 8.9 and a price starting at €62 for today, marked as a 'Super Preis-Leistungs-Verhältnis'. The third hotel is 'Hotel Adler' in Ehingen, with a rating of 8.6. The fourth hotel is 'Best Western Plus BierKulturHotel Schwanen' in Ehingen, with a rating of 8.5.

# Projekt 6: Effekte von besserer Krankenversicherung

# Projekt 6: Effekte von besserer Krankenversicherung

In diesem Projekt widmen Sie sich der Frage:

Welche Auswirkung auf die Gesundheit hat eine bessere Krankenversicherung (mehr Deckung)? Oder führt diese nur zu mehr Arztbesuchen?

Zur Beantwortung dieser Frage werten Sie einen **bereitgestellten Datensatz** aus einem Feldexperiment in Oregon aus. Dort wurden zufällig ausgewählten Einwohnern unterschiedliche (staatliche) Krankenversicherungen bereitgestellt.

	<i>N</i>	Percent with any visits			Number of visits		
		Mean value in control group	Effect of Medicaid coverage	<i>P</i> value	Mean value in control group	Effect of Medicaid coverage	<i>P</i> value
<i>Overall</i>							
All visits	24,646	34.5	7.0 (2.4)	0.003	1.022 (2.632)	0.408 (0.116)	<0.001
<i>By emergency-department use in the prerandomization period</i>							
No visits	16,930	22.5	6.7 (2.9)	0.019	0.418 (1.103)	0.261 (0.084)	0.002
One visit	3881	47.2	9.2 (6.0)	0.127	1.115 (1.898)	0.652 (0.254)	0.010
Two+ visits	3835	72.2	7.1 (5.6)	0.206	3.484 (5.171)	0.380 (0.648)	0.557
Five+ visits	957	89.4	0.7 (8.3)	0.932	6.948 (7.635)	2.486 (2.079)	0.232
Two+ outpatient visits	3402	73.2	9.6 (6.0)	0.111	3.658 (5.375)	0.560 (0.742)	0.450

# Was Sie in den Projekten lernen

✚ Analysen auf **echten Daten**

# Was Sie in den Projekten lernen

- + Analysen auf **echten Daten**
- + Nutzung eines API, Webscraping, Analyse von experimentellen Daten

# Was Sie in den Projekten lernen

- + Analysen auf **echten Daten**
- + Nutzung eines API, Webscraping, Analyse von experimentellen Daten
- + Deskriptive Analysen mittels Grafiken und Tabellen

# Was Sie in den Projekten lernen

- + Analysen auf **echten Daten**
- + Nutzung eines API, Webscraping, Analyse von experimentellen Daten
- + Deskriptive Analysen mittels Grafiken und Tabellen
- + Regressionsanalysen (univariat und multivariat) und Interpretation von Schätzern



# Was Sie in den Projekten lernen

- + Analysen auf **echten Daten**
- + Nutzung eines API, Webscraping, Analyse von experimentellen Daten
- + Deskriptive Analysen mittels Grafiken und Tabellen
- + Regressionsanalysen (univariat und multivariat) und Interpretation von Schätzern
- + Kausale Zusammenhänge verstehen und diese aufdecken

# Was Sie durch Review Reports lernen

- ✚ Rekapitulation des Projekts

# Was Sie durch Review Reports lernen

- + Rekapitulation des Projekts
- + Kritische Auseinandersetzung mit der Arbeit von Mitstudierenden

# Was Sie durch Review Reports lernen

- + Rekapitulation des Projekts
- + Kritische Auseinandersetzung mit der Arbeit von Mitstudierenden
- + Erkennen was die andere Gruppe gut/schlecht gemacht hat und daraus für ihre eigene Arbeit Erkenntnisse gewinnen

# Was Sie durch Review Reports lernen

- + Rekapitulation des Projekts
- + Kritische Auseinandersetzung mit der Arbeit von Mitstudierenden
- + Erkennen was die andere Gruppe gut/schlecht gemacht hat und daraus für ihre eigene Arbeit Erkenntnisse gewinnen

Die Review Reports für jedes Projekt sind *individuelle* Abgaben!

Wie setzt sich die Note zusammen?

# Notengebung

Nach den Vorlesungseinheiten und der Einführung in R gibt es drei Projekte:

# Notengebung

Nach den Vorlesungseinheiten und der Einführung in R gibt es drei Projekte:

- + Erstes Projekt: Zusammen mit dem Dozenten: 10 Punkte = 10% der Note
  - + Nur Basiscode (spezielle für die Grafiken) wird geliefert und Sie sollten diesen optimieren
  - + Interpretation der Regression und Instrumentalvariablenschätzer wird in der Vorlesung ausführlich besprochen



# Notengebung

Nach den Vorlesungseinheiten und der Einführung in R gibt es drei Projekte:

- + Erstes Projekt: Zusammen mit dem Dozenten: 10 Punkte = 10% der Note
  - + Nur Basiscode (spezielle für die Grafiken) wird geliefert und Sie sollten diesen optimieren
  - + Interpretation der Regression und Instrumentalvariablenschätzer wird in der Vorlesung ausführlich besprochen
- + Zweites Projekt: 30 Punkte = 30% der Note

# Notengebung

Nach den Vorlesungseinheiten und der Einführung in R gibt es drei Projekte:

- + Erstes Projekt: Zusammen mit dem Dozenten: 10 Punkte = 10% der Note
  - + Nur Basiscode (spezielle für die Grafiken) wird geliefert und Sie sollten diesen optimieren
  - + Interpretation der Regression und Instrumentalvariablenschätzer wird in der Vorlesung ausführlich besprochen
- + Zweites Projekt: 30 Punkte = 30% der Note
- + Drittes Projekt: 30 Punkte = 30% der Note

# Notengebung

Nach den Vorlesungseinheiten und der Einführung in R gibt es drei Projekte:

- + Erstes Projekt: Zusammen mit dem Dozenten: 10 Punkte = 10% der Note
  - + Nur Basiscode (spezielle für die Grafiken) wird geliefert und Sie sollten diesen optimieren
  - + Interpretation der Regression und Instrumentalvariablenschätzer wird in der Vorlesung ausführlich besprochen
- + Zweites Projekt: 30 Punkte = 30% der Note
- + Drittes Projekt: 30 Punkte = 30% der Note
- + Multiple-Choice Abschlussprüfung: 30 Punkte = 30% der Note

# Notengebung

Nach den Vorlesungseinheiten und der Einführung in R gibt es drei Projekte:

- + Erstes Projekt: Zusammen mit dem Dozenten: 10 Punkte = 10% der Note
  - + Nur Basiscode (spezielle für die Grafiken) wird geliefert und Sie sollten diesen optimieren
  - + Interpretation der Regression und Instrumentalvariablenschätzer wird in der Vorlesung ausführlich besprochen
- + Zweites Projekt: 30 Punkte = 30% der Note
- + Drittes Projekt: 30 Punkte = 30% der Note
- + Multiple-Choice Abschlussprüfung: 30 Punkte = 30% der Note

Es werden immer **Gruppen von drei Personen** geformt, welche die **Projekte zusammen** abgeben. Natürlich dürfen alle Teilnehmer Projekte gerne miteinander diskutieren.

Jedoch muss jede Gruppe eine *individuelle Ausarbeitung* abgeben. Wenn sich die Lösung einzelner Abschnitte der Projektes (oder das komplette Projekt) zu stark ähnelt, wird dies mit 0 Punkten für das Projekt geahndet (für alle beteiligten Gruppen).

# Klausur

**Vorleistung 1:** Die **RTutor Problem Sets** müssen **individuell** bearbeitet und eingereicht werden um für die Projekte und Klausur zugelassen zu werden

- + Letztes RTutor Problem Set ist am 16. Mai fällig
- + Mindestens 80% der Punkte **pro Problem Set** müssen erreicht werden um für die Projekte zugelassen zu werden
- + Mindestens 30% der Punkte in der Probeklausur in KW 20

# Klausur

**Vorleistung 1:** Die **RTutor Problem Sets** müssen **individuell** bearbeitet und eingereicht werden um für die Projekte und Klausur zugelassen zu werden

- + Letztes RTutor Problem Set ist am 16. Mai fällig
- + Mindestens 80% der Punkte **pro Problem Set** müssen erreicht werden um für die Projekte zugelassen zu werden
- + Mindestens 30% der Punkte in der Probeklausur in KW 20

**Vorleistung 2:** Die **Peer Reviews** müssen **individuell** bearbeitet und eingereicht werden um für die Klausur zugelassen zu werden

- + Letzter Review Report ist am 19. Juli fällig
- + Studenten ranken die Review Reports nach Nützlichkeit:
  - + Bei *mindestens* einem Projekt muss der Review Report *mindestens* auf Platz 2 gerankt werden.
  - + Wenn alle Reports als "nützlich" eingestuft wurden, entscheidet der Dozent über die Zulassung zu Klausur

# Klausur

**Vorleistung 1:** Die **RTutor Problem Sets** müssen **individuell** bearbeitet und eingereicht werden um für die Projekte und Klausur zugelassen zu werden

- + Letztes RTutor Problem Set ist am 16. Mai fällig
- + Mindestens 80% der Punkte **pro Problem Set** müssen erreicht werden um für die Projekte zugelassen zu werden
- + Mindestens 30% der Punkte in der Probeklausur in KW 20

**Vorleistung 2:** Die **Peer Reviews** müssen **individuell** bearbeitet und eingereicht werden um für die Klausur zugelassen zu werden

- + Letzter Review Report ist am 19. Juli fällig
- + Studenten ranken die Review Reports nach Nützlichkeit:
  - + Bei *mindestens* einem Projekt muss der Review Report *mindestens* auf Platz 2 gerankt werden.
  - + Wenn alle Reports als "nützlich" eingestuft wurden, entscheidet der Dozent über die Zulassung zu Klausur

Die multiple choice Abschlussprüfung beinhaltet Fragen zu den drei Teilprojekten, den Vorlesungsinhalten, der Case-Study und den RTutor Problem Sets.

# Teilen von Code

- + Viel ist im Web verfügbar und darf auch gerne verwendet werden
  - + Wenn Sie Code aus dem Internet verwenden, dann müssen Sie die Quelle entsprechend kennzeichnen!
  - + Falls Sie die Quelle nicht zitieren wird dies als Plagiat gewertet und wird mit einer 5.0 für das gesamte Projekt geahndet
- + `RTutor` Problem Sets sind selbstständig auszuführen. Code darf nicht mit anderen geteilt werden!



# Teilen von Code

- + Viel ist im Web verfügbar und darf auch gerne verwendet werden
  - + Wenn Sie Code aus dem Internet verwenden, dann müssen Sie die Quelle entsprechend kennzeichnen!
  - + Falls Sie die Quelle nicht zitieren wird dies als Plagiat gewertet und wird mit einer 5.0 für das gesamte Projekt geahndet
- + RTutor Problem Sets sind selbstständig auszuführen. Code darf nicht mit anderen geteilt werden!

Projekte dürfen innerhalb der Gruppe und auch gerne mit anderen Gruppen diskutiert werden, jedoch **darf kein Code** an andere Gruppen **weitergegeben werden!**

# Notengebung der Projekte

- ✚ Im zweiten Projekt muss neben der schriftlichen Ausarbeitung (70% der Note für das Projekt) auch ein Screencast (30% der Note für das Projekt) gehalten werden
  - ✚ Der Screencast sollte nicht länger als 5 Minuten gehen

# Notengebung der Projekte

- ✚ Im zweiten Projekt muss neben der schriftlichen Ausarbeitung (70% der Note für das Projekt) auch ein Screencast (30% der Note für das Projekt) gehalten werden
  - ✚ Der Screencast sollte nicht länger als 5 Minuten gehen
- ✚ Im dritten Projekt muss neben der schriftlichen Ausarbeitung (70% der Note für das Projekt) auch ein Screencast (30% der Note für das Projekt) gehalten werden
  - ✚ Der Screencast sollte nicht länger als 5 Minuten gehen

# Notengebung der Projekte

- ✚ Im zweiten Projekt muss neben der schriftlichen Ausarbeitung (70% der Note für das Projekt) auch ein Screencast (30% der Note für das Projekt) gehalten werden
  - ✚ Der Screencast sollte nicht länger als 5 Minuten gehen
- ✚ Im dritten Projekt muss neben der schriftlichen Ausarbeitung (70% der Note für das Projekt) auch ein Screencast (30% der Note für das Projekt) gehalten werden
  - ✚ Der Screencast sollte nicht länger als 5 Minuten gehen

Bitte bereiten Sie den Screencast frühzeitig vor, er gibt 30% der Note!

# Notenschlüssel

Der Notenschlüssel für die Zusammensetzung der Endnote wurde auf Moodle hochgeladen.

Dieser Notenschlüssel ist unsere Referenzgröße, wir behalten uns jedoch vor davon abzuweichen!

# Dokumentation und Coding Standard

Sie sollten mit ihren Gruppenpartnern zusammenarbeiten, dafür gilt es einige Prinzipien zu beachten:

- + Benutzen Sie immer einfache Textdateien um miteinander zusammen zu arbeiten
  - + Dateien, welche Sie miteinander bearbeiten sollten mit einem Texteditor zu lesen sein (Notepad ++ / vim / eclipse ...)
  - + Hier eignet sich das in der Vorlesung vorgestellte RMarkdown bestens
- + Strukturieren Sie ihren Code (Kommentare)
- + Begrenzen Sie ihren Code (max. 80 Zeichen pro Zeile)
- + Rücken Sie einzelne Bausteine ein (Alles was zu einer Funktion gehört sollte mit vier Leerzeichen eingerückt werden)

# Zeitmanagement

- + Definieren Sie Meilensteine
  - + Bis wann muss was von wem erstellt worden sein?

# Zeitmanagement

- + Definieren Sie Meilensteine
  - + Bis wann muss was von wem erstellt worden sein?
- + Bleiben Sie in Kontakt mit ihrem/ihrer Gruppenpartner/in, bspw. über Github oder Moodle, um über ihren Projektfortschritt zu sprechen



# Zeitmanagement

- + Definieren Sie Meilensteine
  - + Bis wann muss was von wem erstellt worden sein?
- + Bleiben Sie in Kontakt mit ihrem/ihrer Gruppenpartner/in, bspw. über Github oder Moodle, um über ihren Projektfortschritt zu sprechen
- + Stellen Sie unbedingt Fragen im Forum auf Moodle!
  - + Hier können Fragen oft sehr schnell beantwortet werden und meist haben mehrere Gruppen die gleiche Frage
  - + Wenn Sie die Antwort auf eine Frage wissen, dann scheuen Sie sich nicht diese in Moodle zu posten!

# Zeitmanagement

- + Definieren Sie Meilensteine
  - + Bis wann muss was von wem erstellt worden sein?
- + Bleiben Sie in Kontakt mit ihrem/ihrer Gruppenpartner/in, bspw. über Github oder Moodle, um über ihren Projektfortschritt zu sprechen
- + Stellen Sie unbedingt Fragen im Forum auf Moodle!
  - + Hier können Fragen oft sehr schnell beantwortet werden und meist haben mehrere Gruppen die gleiche Frage
  - + Wenn Sie die Antwort auf eine Frage wissen, dann scheuen Sie sich nicht diese in Moodle zu posten!
- + Nutzen Sie die Möglichkeit des Tutoriums und sprechen Sie dort Schwierigkeiten direkt an

# Zeitmanagement

- + Definieren Sie Meilensteine
  - + Bis wann muss was von wem erstellt worden sein?
- + Bleiben Sie in Kontakt mit ihrem/ihrer Gruppenpartner/in, bspw. über Github oder Moodle, um über ihren Projektfortschritt zu sprechen
- + Stellen Sie unbedingt Fragen im Forum auf Moodle!
  - + Hier können Fragen oft sehr schnell beantwortet werden und meist haben mehrere Gruppen die gleiche Frage
  - + Wenn Sie die Antwort auf eine Frage wissen, dann scheuen Sie sich nicht diese in Moodle zu posten!
- + Nutzen Sie die Möglichkeit des Tutoriums und sprechen Sie dort Schwierigkeiten direkt an
- + Kommen Sie bei tiefergehenden Fragen frühzeitig auf den Dozenten zu

# Zeitmanagement

- + Definieren Sie Meilensteine
  - + Bis wann muss was von wem erstellt worden sein?
- + Bleiben Sie in Kontakt mit ihrem/ihrer Gruppenpartner/in, bspw. über Github oder Moodle, um über ihren Projektfortschritt zu sprechen
- + Stellen Sie unbedingt Fragen im Forum auf Moodle!
  - + Hier können Fragen oft sehr schnell beantwortet werden und meist haben mehrere Gruppen die gleiche Frage
  - + Wenn Sie die Antwort auf eine Frage wissen, dann scheuen Sie sich nicht diese in Moodle zu posten!
- + Nutzen Sie die Möglichkeit des Tutoriums und sprechen Sie dort Schwierigkeiten direkt an
- + Kommen Sie bei tiefergehenden Fragen frühzeitig auf den Dozenten zu
- + Planen Sie genügend Zeit für die Erstellung und das Halten des Screencast ein
  - + Der Screencast gibt 30% der Projektnote -> Nicht auf die leichte Schulter nehmen!

# Vorbereitung bis zur Vorlesung am Freitag

- + Schauen Sie sich die Lehrvideos an:
  - + Stichprobe vs. Population
  - + Stichprobenvarianz und Standardfehler
  - + Konfidenzintervall

# Github für die Zusammenarbeit

Wir werden weiterhin unsere [private Organisation auf Github](#) nutzen.

Falls Sie ihre Git-Kenntnisse auffrischen wollen können Sie jederzeit auf den [Moodle-Kurs des letzten Semesters](#) zugreifen.  
Die Unterlagen bleiben dort weiterhin für Sie verfügbar!

Recap des 1. Teils der Vorlesung

# Recap zu R und dem tidyverse

Welche Pakete verwenden Sie um .csv/.dta/.xlsx Dateien einzulesen?



# Recap zu R und dem tidyverse

Welche Pakete verwenden Sie um .csv/.dta/.xlsx Dateien einzulesen?

Was sagt dieses Symbol  $\%>\%$ ?

# Recap zu R und dem tidyverse

Welche Pakete verwenden Sie um .csv/.dta/.xlsx Dateien einzulesen?

Was sagt dieses Symbol  $\%>\%$ ?

Wie führen Sie mehrere Datensätze zusammen?

# Recap zu Github

Warum sollte ihr Commit in Github Desktop immer beschrieben werden?

# Recap zu Github

Warum sollte ihr Commit in Github Desktop immer beschrieben werden?

Warum sollten Sie häufig committen?

# Recap zur Visualisierung

Warum sollte ein Balkendiagramm immer bei Null beginnen? Gilt diese Regel für alle Diagramme?

# Recap zur Visualisierung

Warum sollte ein Balkendiagramm immer bei Null beginnen? Gilt diese Regel für alle Diagramme?

Warum sollten Sie ihre Tabellen nicht als Tibble Output präsentieren?

# Recap der Case-Study

Welches Bild ergibt sich für die Verschuldung in Deutschland? Gibt es regionale Unterschiede?

# Recap der Case-Study

Welches Bild ergibt sich für die Verschuldung in Deutschland? Gibt es regionale Unterschiede?

Sind diese regionalen Unterschiede auch bei anderen Kenngrößen sichtbar?



# Recap Theoretischer Inhalte

Wann können Sie aus einer Stichprobe valide Aussagen zur Grundgesamtheit ziehen?

# Recap Theoretischer Inhalte

Wann können Sie aus einer Stichprobe valide Aussagen zur Grundgesamtheit ziehen?

Welchen Effekt hat die Stichprobengröße auf ihre Stichprobenvarianz?