



[Link zu den Folien](#)

Warum Statistik (allein) nichts hilft zur Entscheidungsfindung

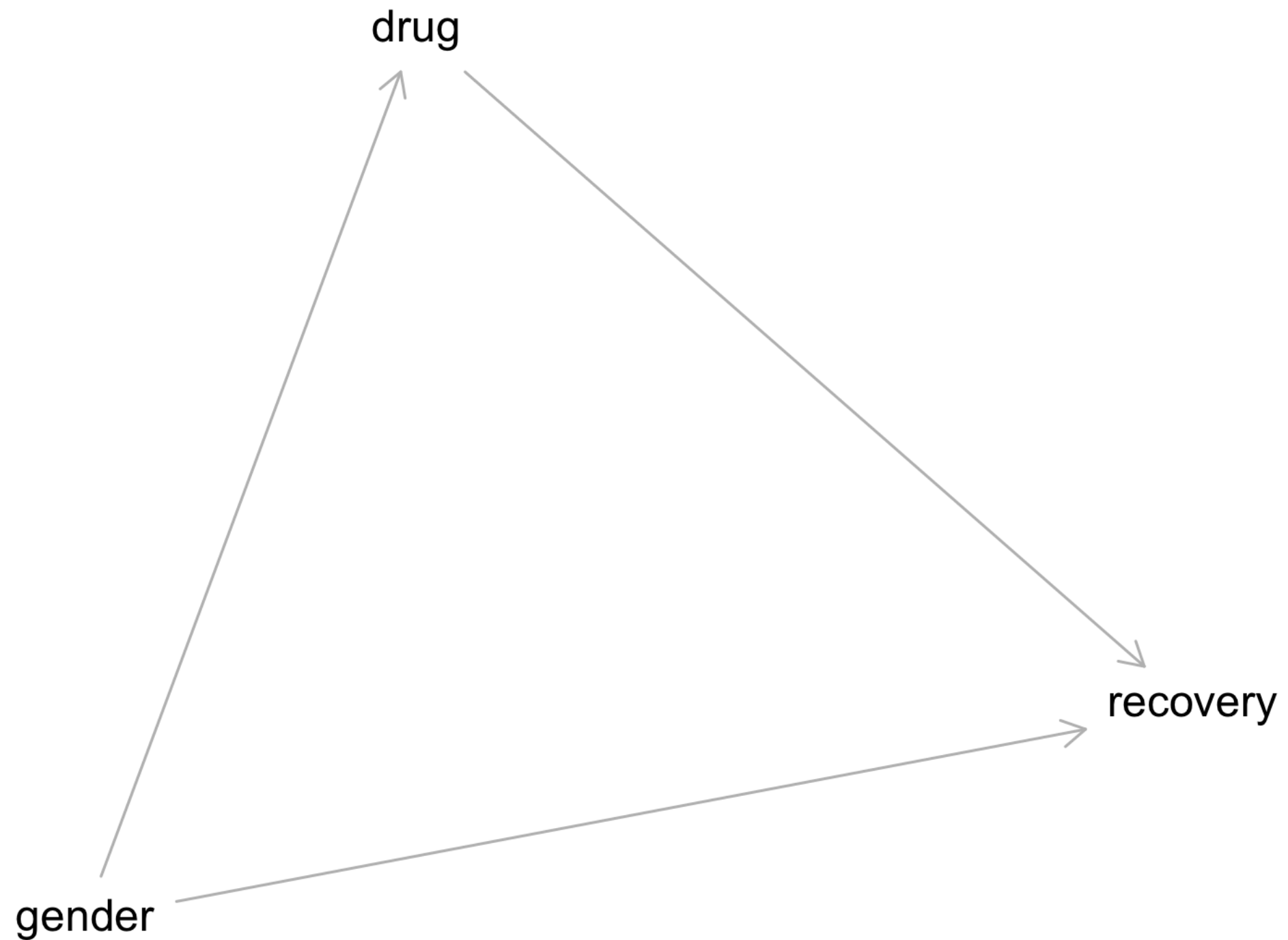
Einführung in die Kausalanalyse

Statistik, was soll ich tun?

Studie A: Was raten Sie dem Arzt?

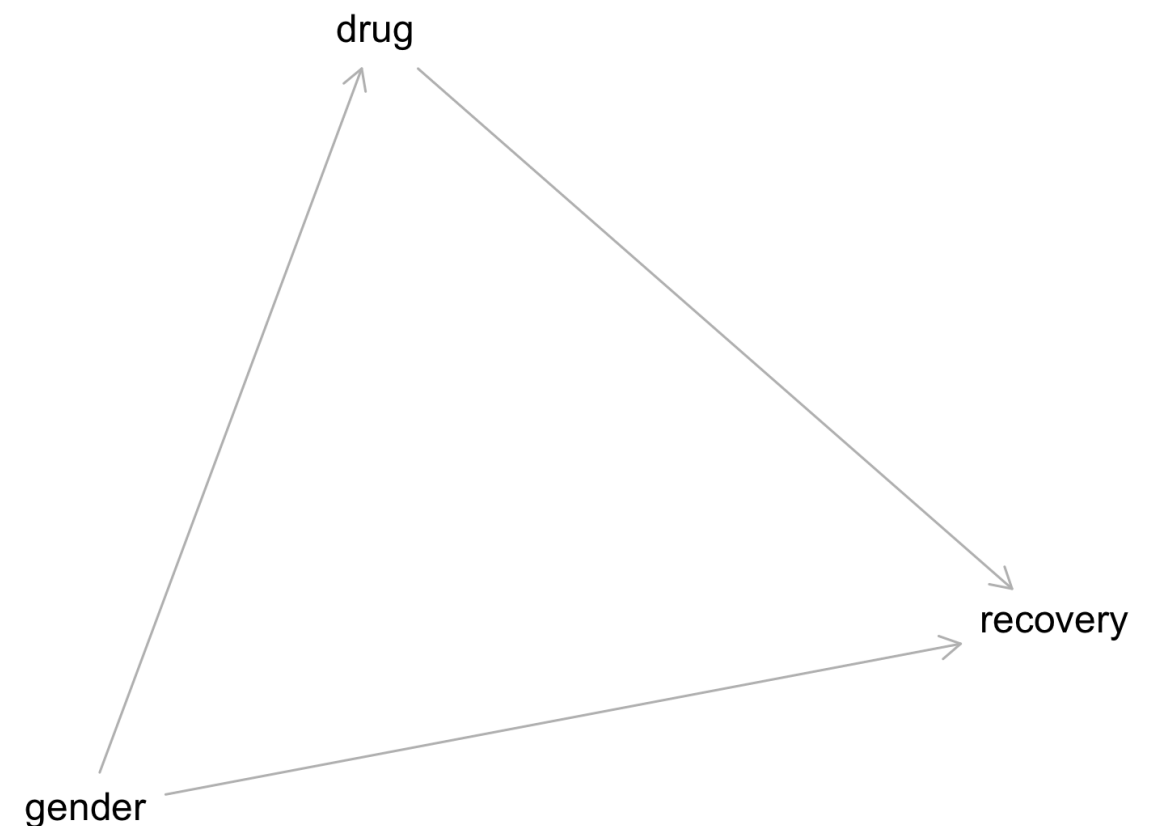
	Mit Medikament	Ohne Medikament
Männer	81/87 überlebt (93%)	234/270 überlebt (87%)
Frauen	192/263 überlebt (73%)	55/80 überlebt (69%)
Gesamt	273/350 überlebt (78%)	289/350 überlebt (83%)

Kausalmodell zur Studie A



Geschlecht „kosfundiert“ die Wirkung des Medikaments

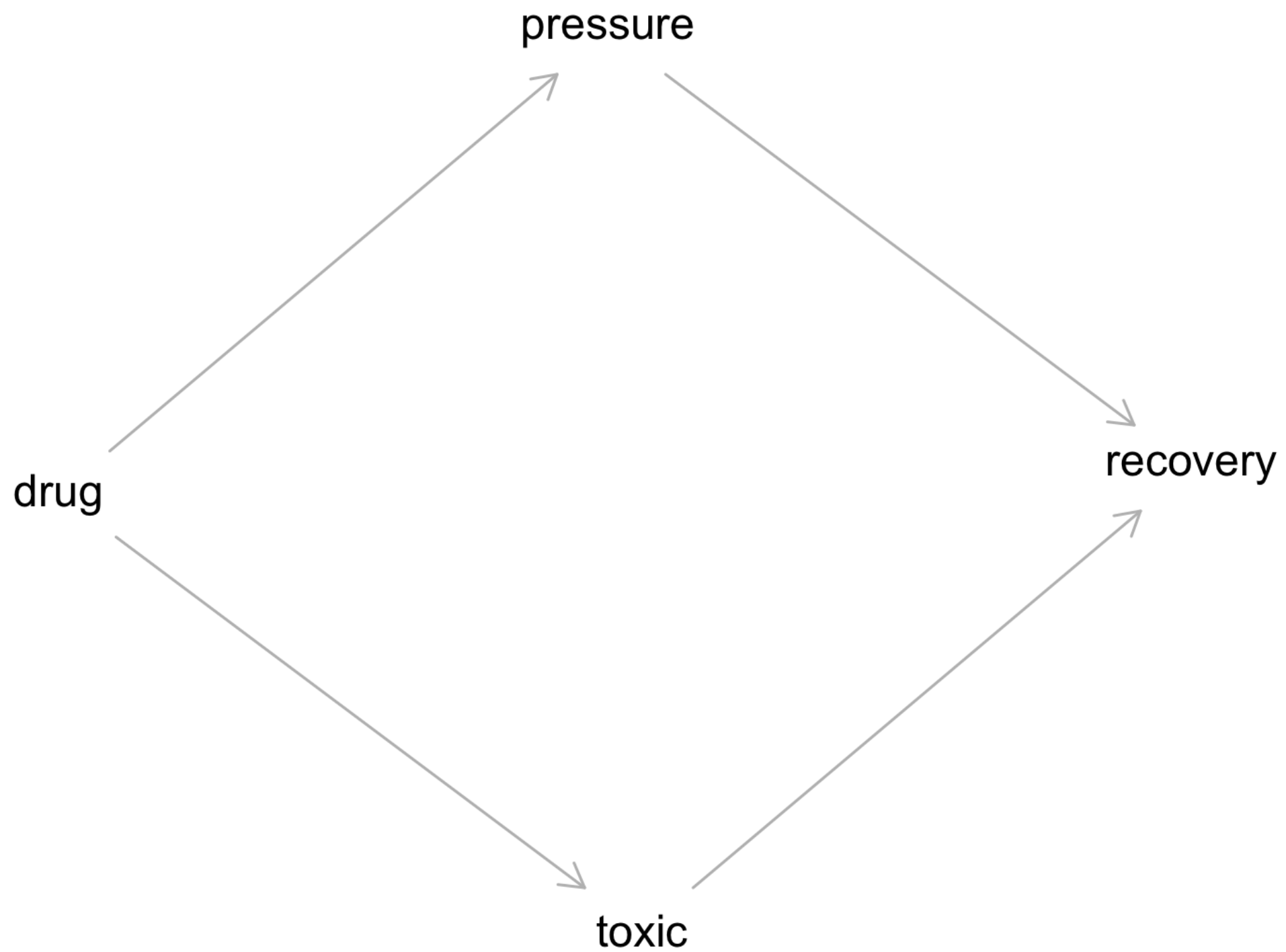
- Geschlecht hat einen Einfluss (+) auf Einnahme des Medikaments und auf Heilung (-).
- Das Medikament hat einen Einfluss (+) auf die Heilung.
- Betrachtet man die Gesamt-Daten zur Heilung, so ist der Effekt von Geschlecht und Medikament vermengt.



Studie B: Was raten Sie dem Arzt?

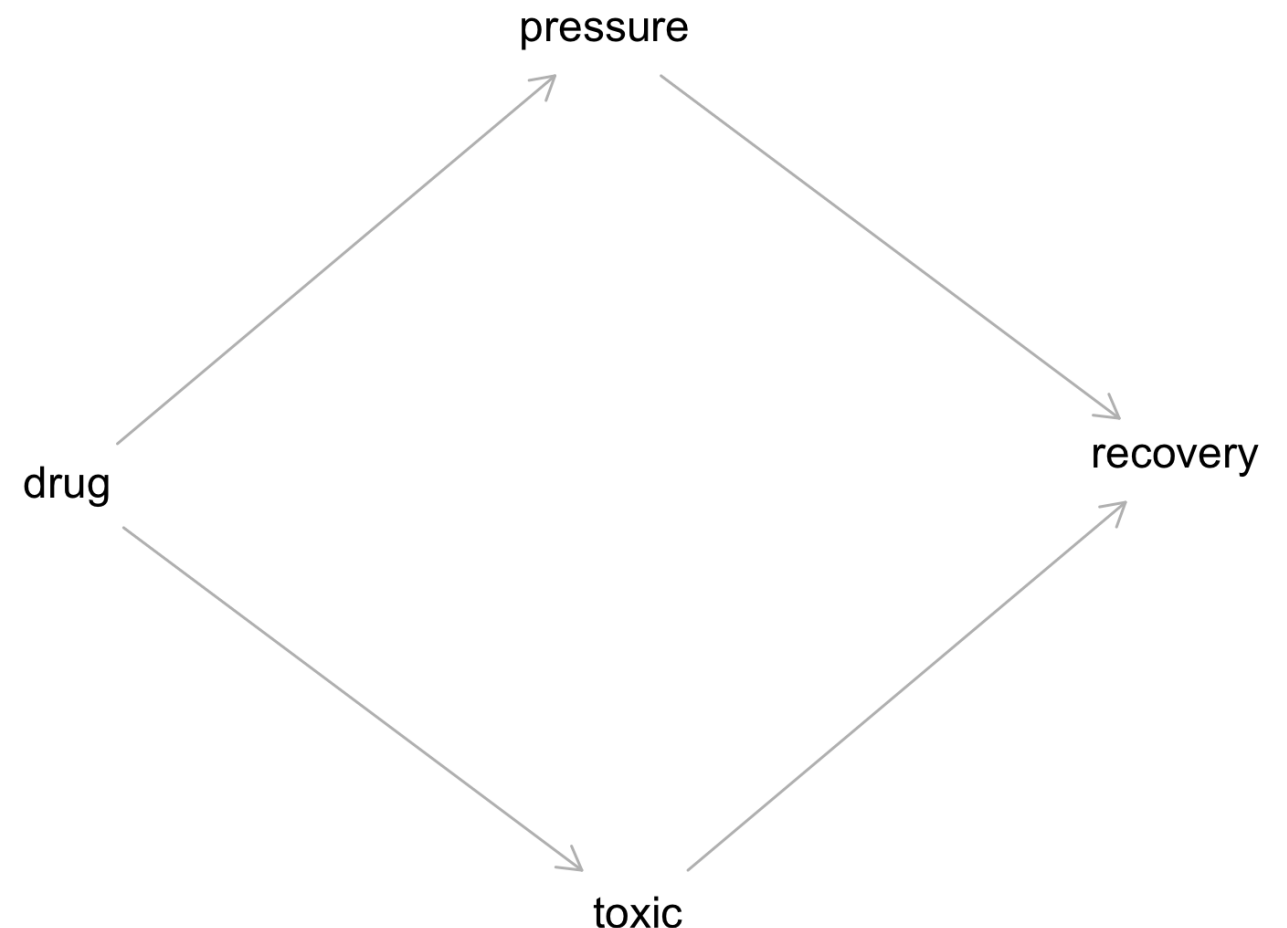
	Ohne Medikament	Mit Medikament
geringer Blutdruck	81/87 überlebt (93%)	234/270 überlebt (87%)
hoher Blutdruck	192/263 überlebt (73%)	55/80 überlebt (69%)
Gesamt	273/350 überlebt (78%)	289/350 überlebt (83%)

Kausalmodell zur Studie B



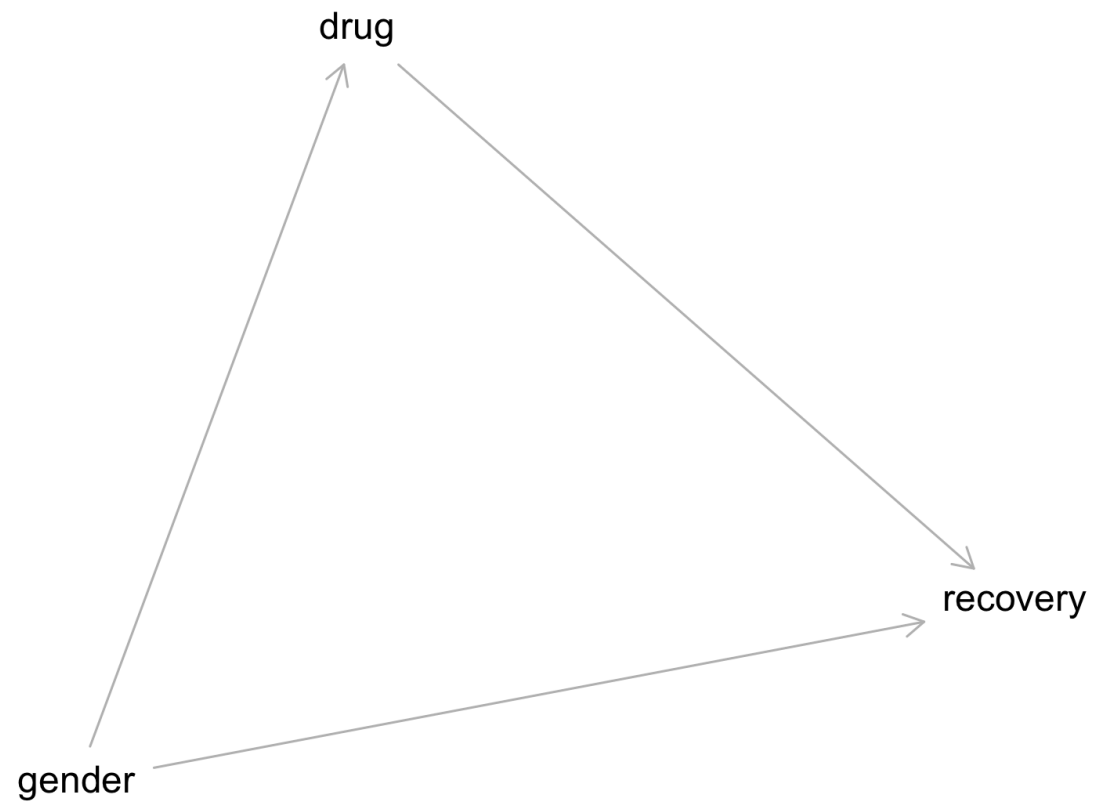
Geschlecht „kosfundiert“ die Wirkung des Medikaments

- Das Medikament senkt den Blutdruck und hat einen toxischen Effekt.
- Der verringerte Blutdruck bewirkt Heilung.
- Sucht man innerhalb der Leute mit gesenktem Blutdruck nach Effekten, findet man nur den toxischen Effekt: Gegeben diesen Blutdruck ist das Medikament schädlich aufgrund des toxischen Effekts. Der positive Effekt der Blutdruck-Senkung ist auf diese Art nicht zu sehen.

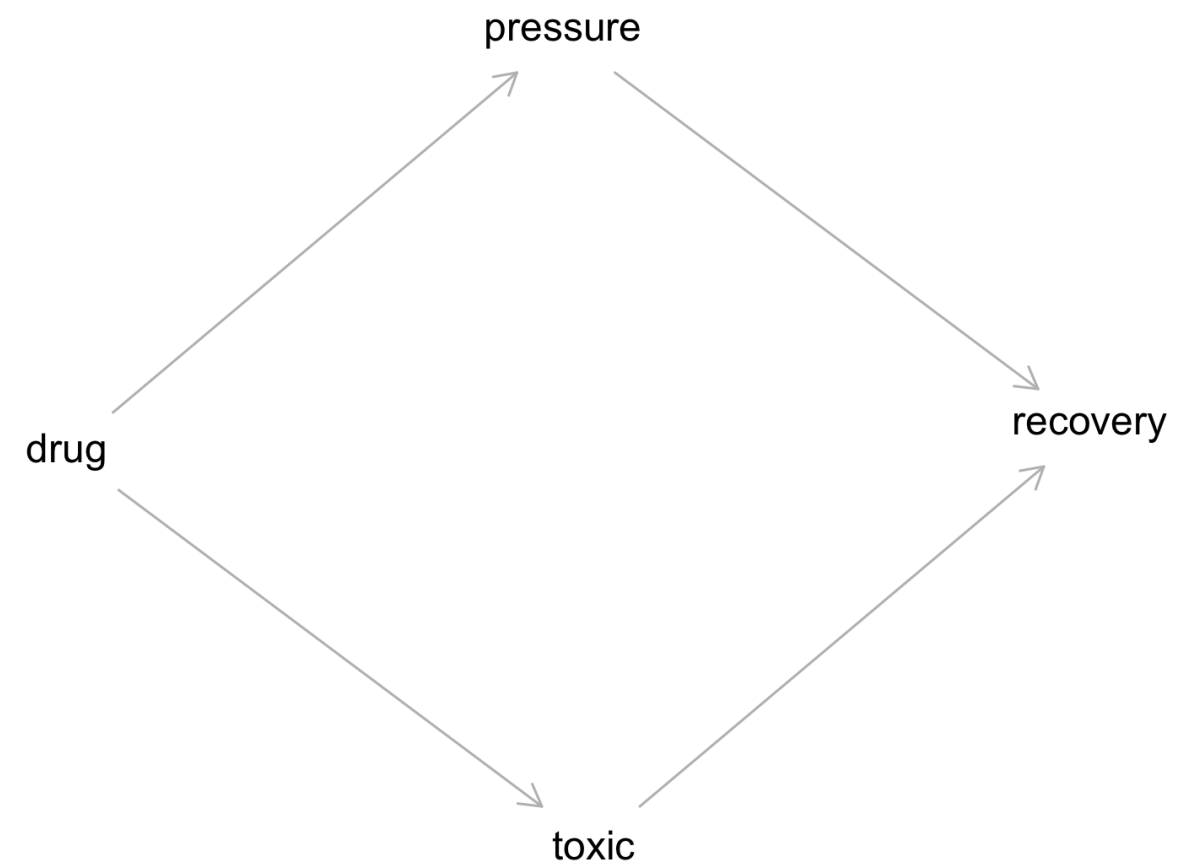


Das Kausalmodell entscheidet, ob Gesamt- oder Teildaten

Studie A



Studie B



Heiraten und Einkommen

„Studien zeigen, dass Einkommen und Heiraten (bzw. verheiratete sein) hoch korrelieren. Daher wird sich dein Einkommen erhöhen, wenn du heiratest.“

Eile mit Weile

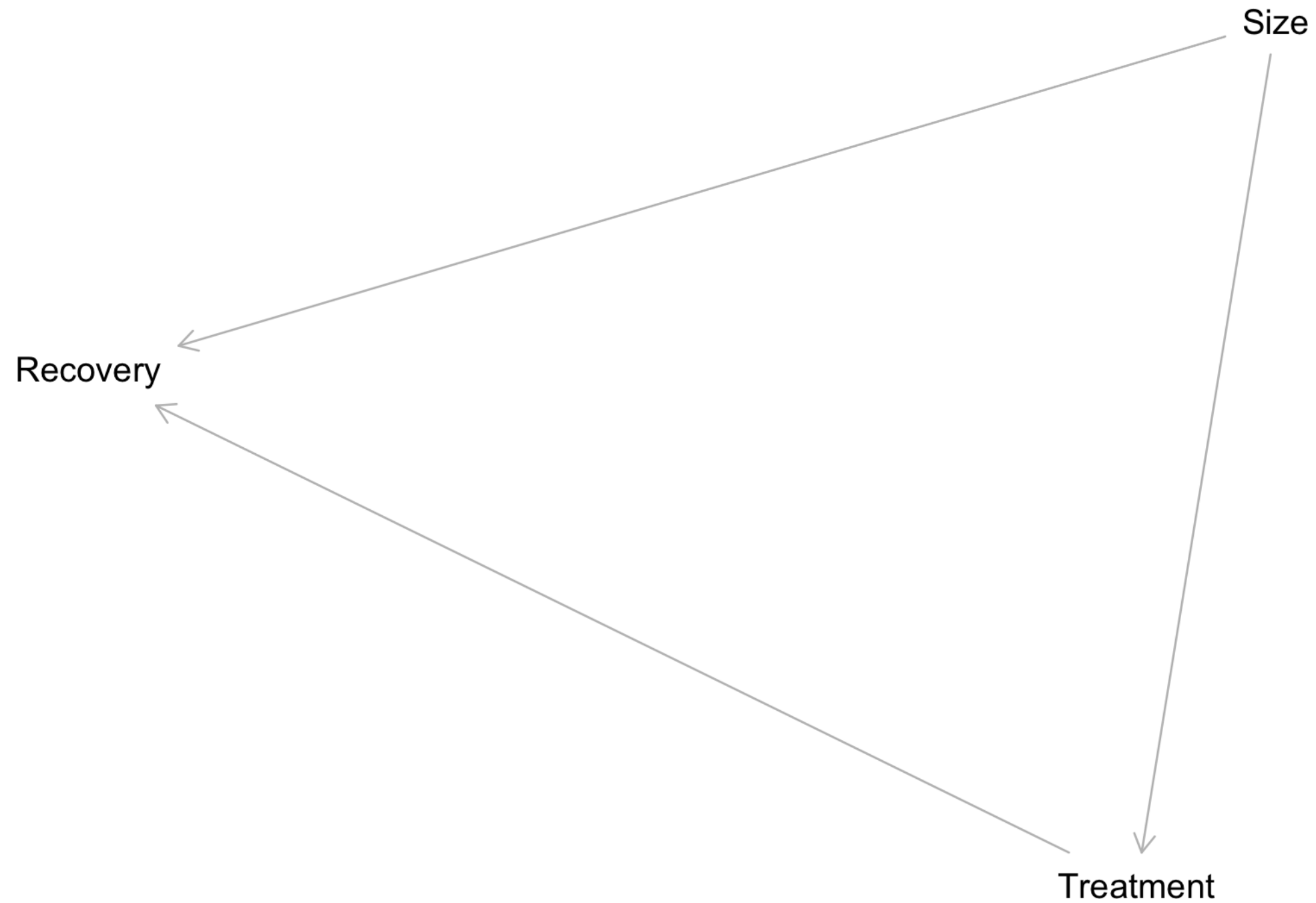
„Studien zeigen, dass Leute, die sich beeilen, zu spät zu ihrer Besprechung kommen. Daher lieber nicht beeilen, oder du kommst zu spät zu deiner Besprechung.“

Behandlung von Nierensteinen

Nehmen wir an, es gibt zwei Behandlungsvarianten bei Nierensteinen, Behandlung A und B. Ärzte tendieren zu Behandlung A bei großen Steinen (die einen schwereren Verlauf haben); bei kleineren Steinen tendieren die Ärzte zu Behandlung B.

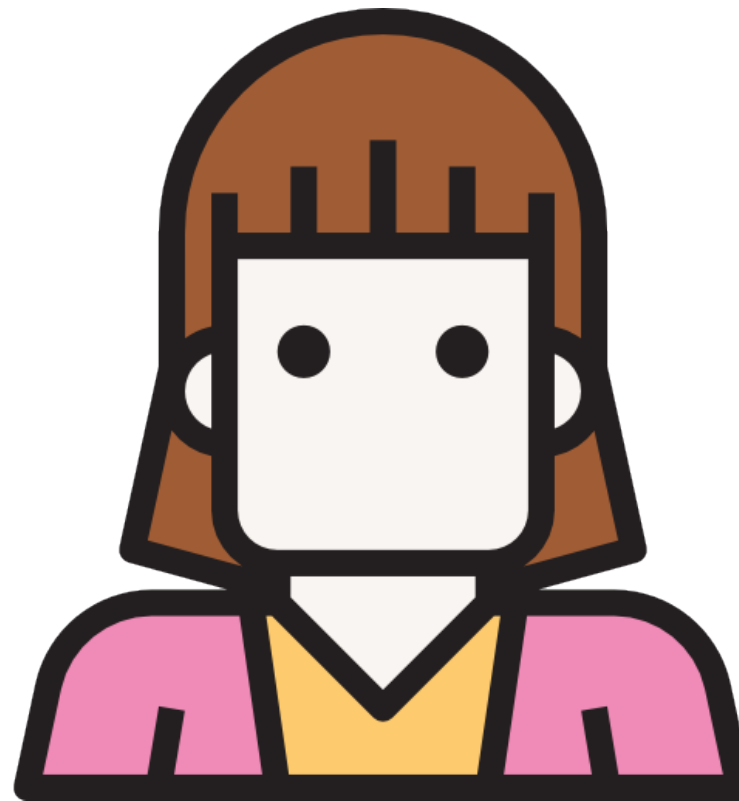
Sollte ein Patient, der nicht weiß, ob sein Nierenstein groß oder klein ist, die Wirksamkeit in der Gesamtpopulation (Gesamtdaten) oder in den stratifizierten Daten (Teildaten nach Steingröße) betrachten, um zu entscheiden, welche Behandlungsvariante er (oder sie) wählt?

Kausalmodell zur Nierenstein-Studie



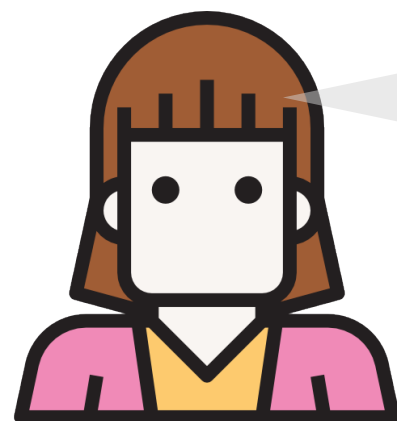
Angi, Don und Wolfi

Fallstudie: Angela hat einen neuen Job, als Marktforscherin



Angela M.,
Marktforscherin

Frisch nach Saratoga County gezogen, schicke Gegend



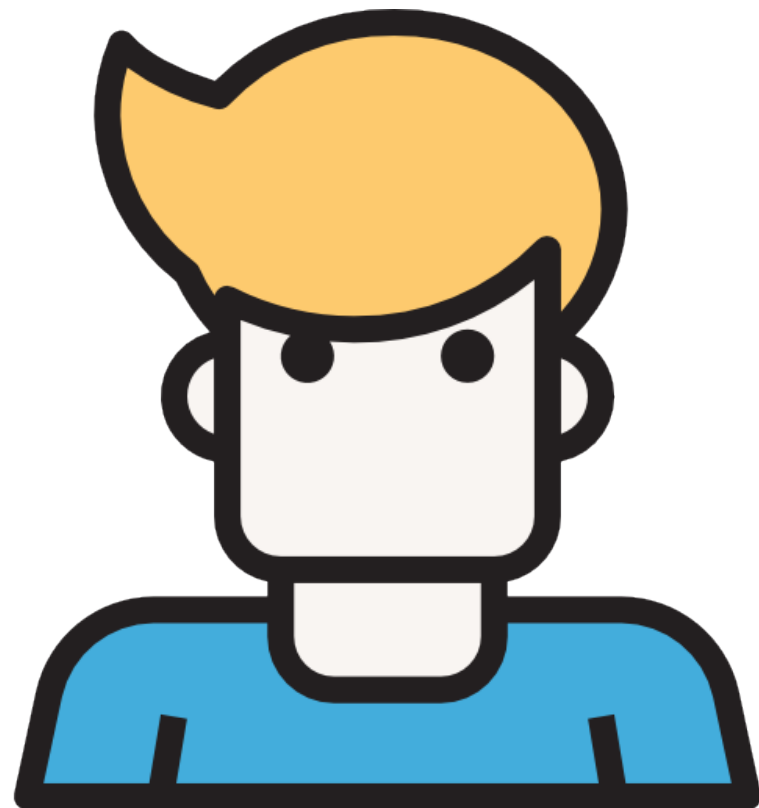
Yeah, I love my new job!

Angi



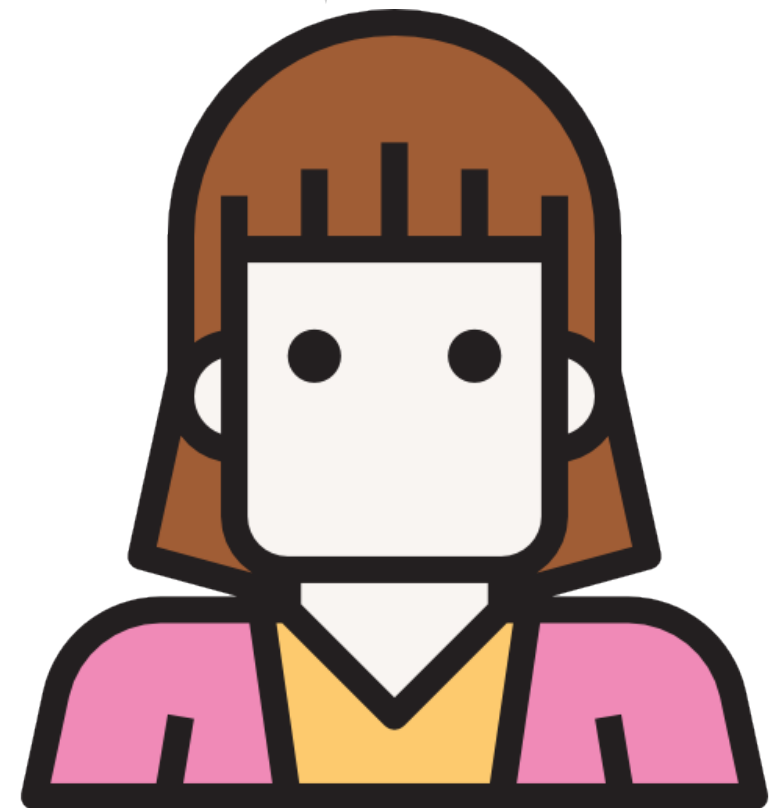
Neuer Job, erste Aufgabe: Verkaufswert von Immobilie vorhersagen

Wieviel ist mein Haus wert!
So groß, tolles Haus!



Don, Immobiliengigant

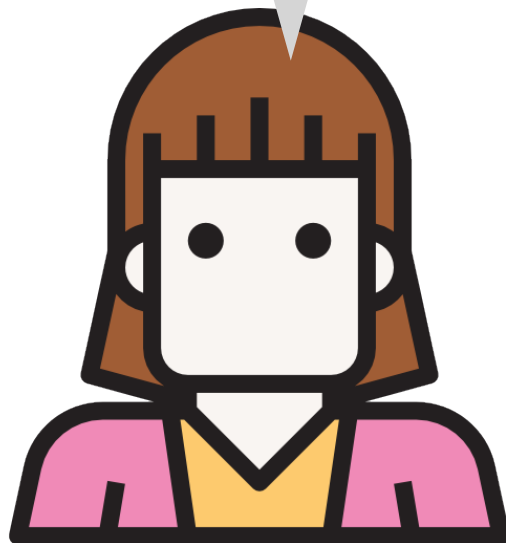
Find ich heraus.
Ich mach's
wissenschaftlich.



Angi

Angi schaut sich Immobilien-Daten zum Landkreis an

Ich liebe Daten! 🥰

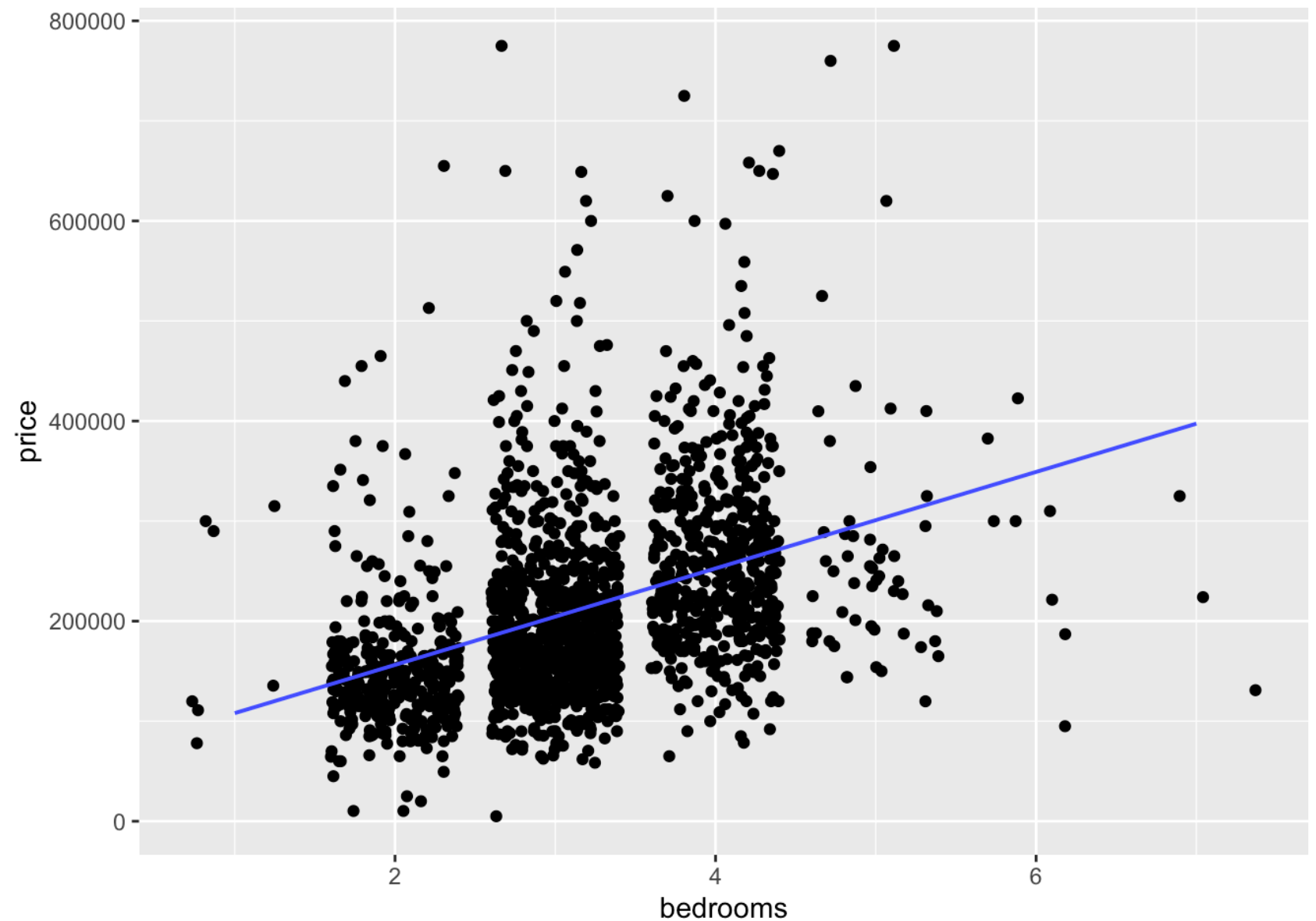
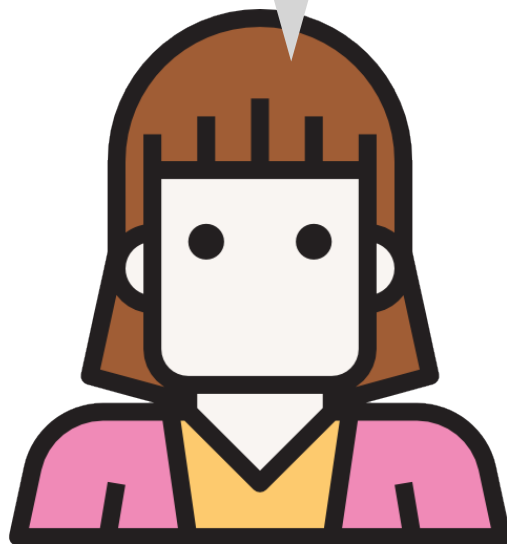


Angi

price	square-feet	age	a/c	fireplace	heating	...
132500	84.17	42	No	Yes	Electricity	
181115	181.44	0	No	No	Gas	
109000	180.60	133	No	Yes	Gas	
155000	180.60	13	No	Yes	Gas	
86060	78.04	0	Yes	No	Gas	
120000	107.02	31	No	Yes	Gas	
153000	255.67	33	No	Yes	Oil	
170000	154.40	23	No	Yes	Oil	
90000	151.62	36	No	No	Electricity	
122900	131.55	4	No	No	Gas	
...	

Model 1: Preis als Funktion der Anzahl der Zimmer

Hey Don!
Mehr Zimmer,
mehr Kohle!



Angi berichtet Don ihre Ergebnisse

```
library(mosaic)
```

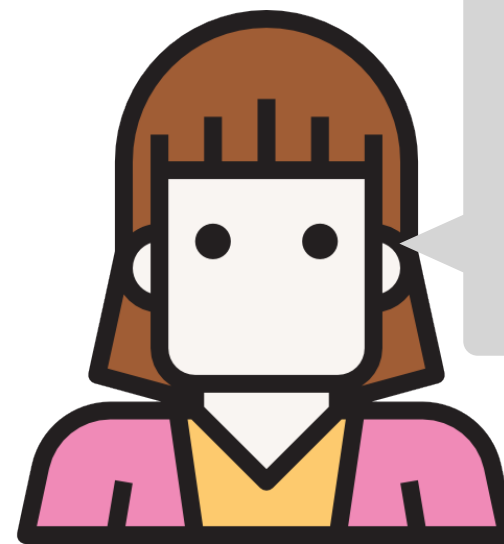
```
data("SaratogaHouses")
```

```
modell <- lm(price ~ bedrooms,  
data = SaratogaHouses)  
coef(modell)
```

```
## (Intercept)      bedrooms  
##      59862.96      48217.81
```

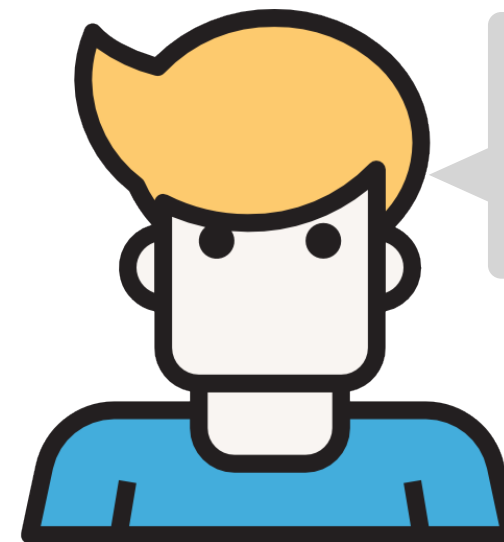
```
dons_house <- data.frame(bedrooms = 2)  
predict(modell, dons_house)
```

```
##           1  
## 156298.6
```



Angi

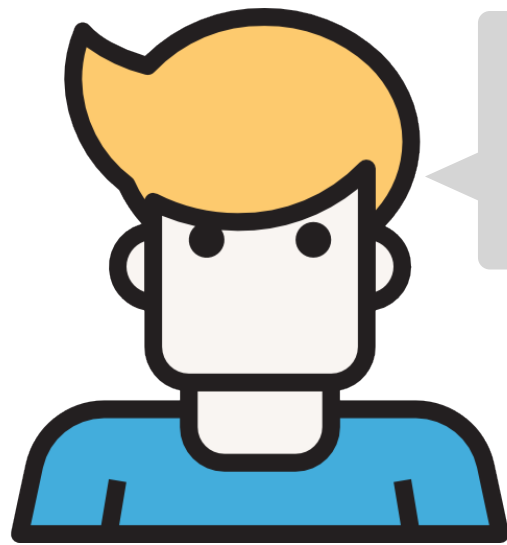
Jedes Zimmer ist knapp 50 Tausend wert. Dein Haus hat einen Wert von 150 Tausend.



Don

Zu wenig! 🤔 😡 🤯

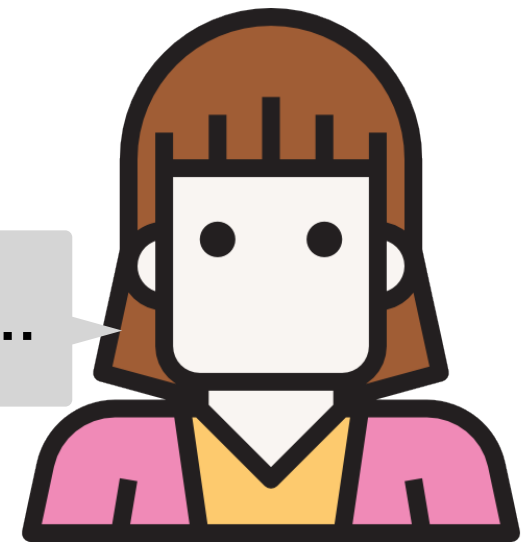
Don hat eine Idee: Jedes Zimmer in zwei Teile teilen!



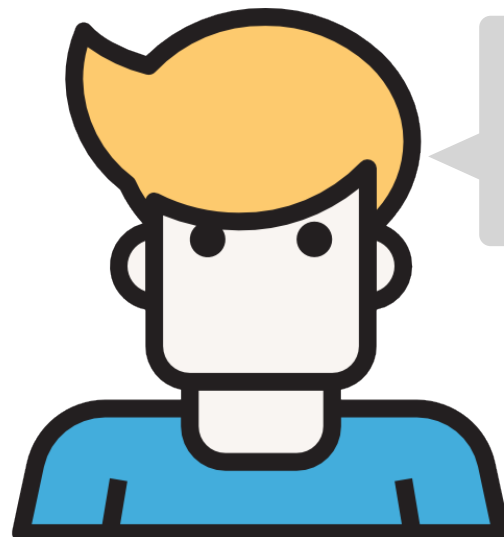
Don

Ich bau eine Mauer!
In jedes Zimmer!

Moment mal ...



Angi



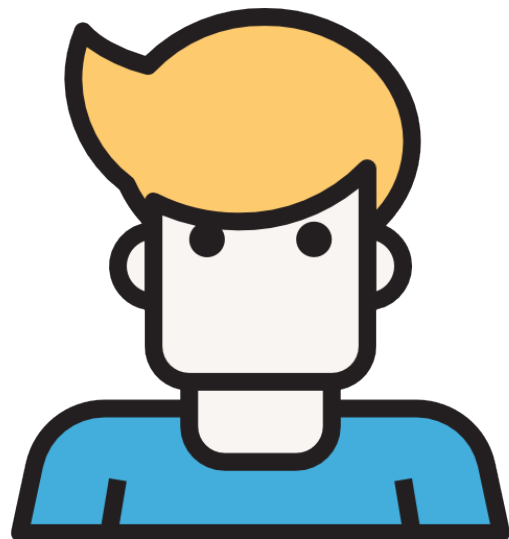
Rechne meinen Preis aus!
Jetzt!

Mit 4 (statt 2) Schlafzimmer steigt der Wert auf 250 Tausend

Laut Modell 1

```
dons_new_house <- data.frame(bedrooms = 4)  
predict(modell1, dons_new_house)
```

```
##           1  
## 252734.2
```

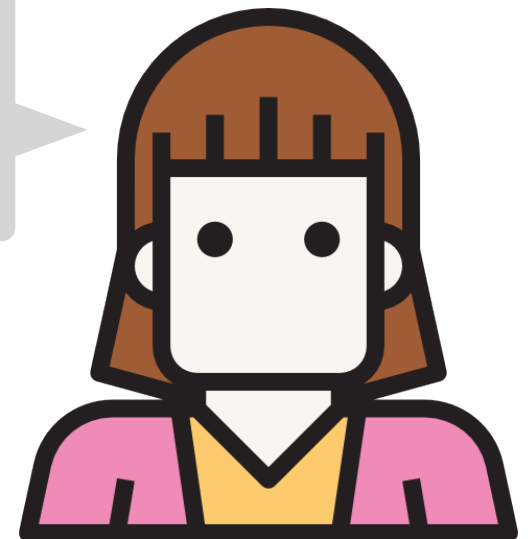


Don

Volltreffer! Jetzt verdien
ich 100 Tausend mehr!



Moment mal ...



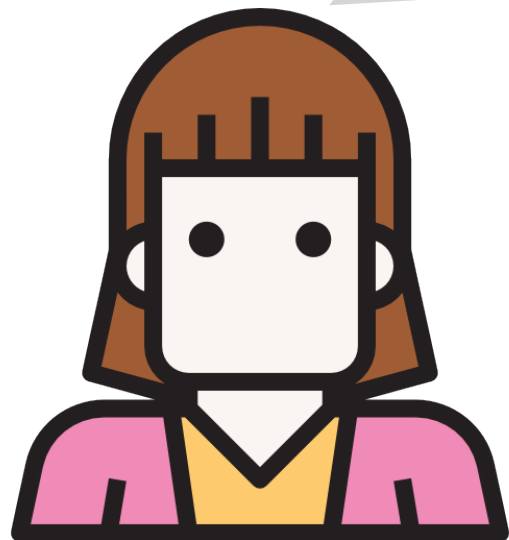
Angi

Modell 2: Preis als Funktion von Zimmerzahl und von Quadratmetern

```
model2 <- lm(price ~ bedrooms + livingArea, data = SaratogaHouses)
coef(model2)
```

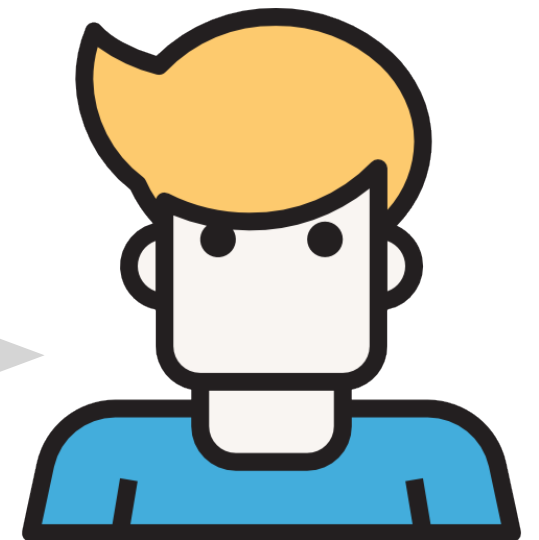
```
## (Intercept)    bedrooms  livingArea
##  36667.895    -14196.769      125.405
```

Die Zimmer halbieren
verringert den Wert,
Don!



Angi

Was, weniger
Geld?! Oh nein!

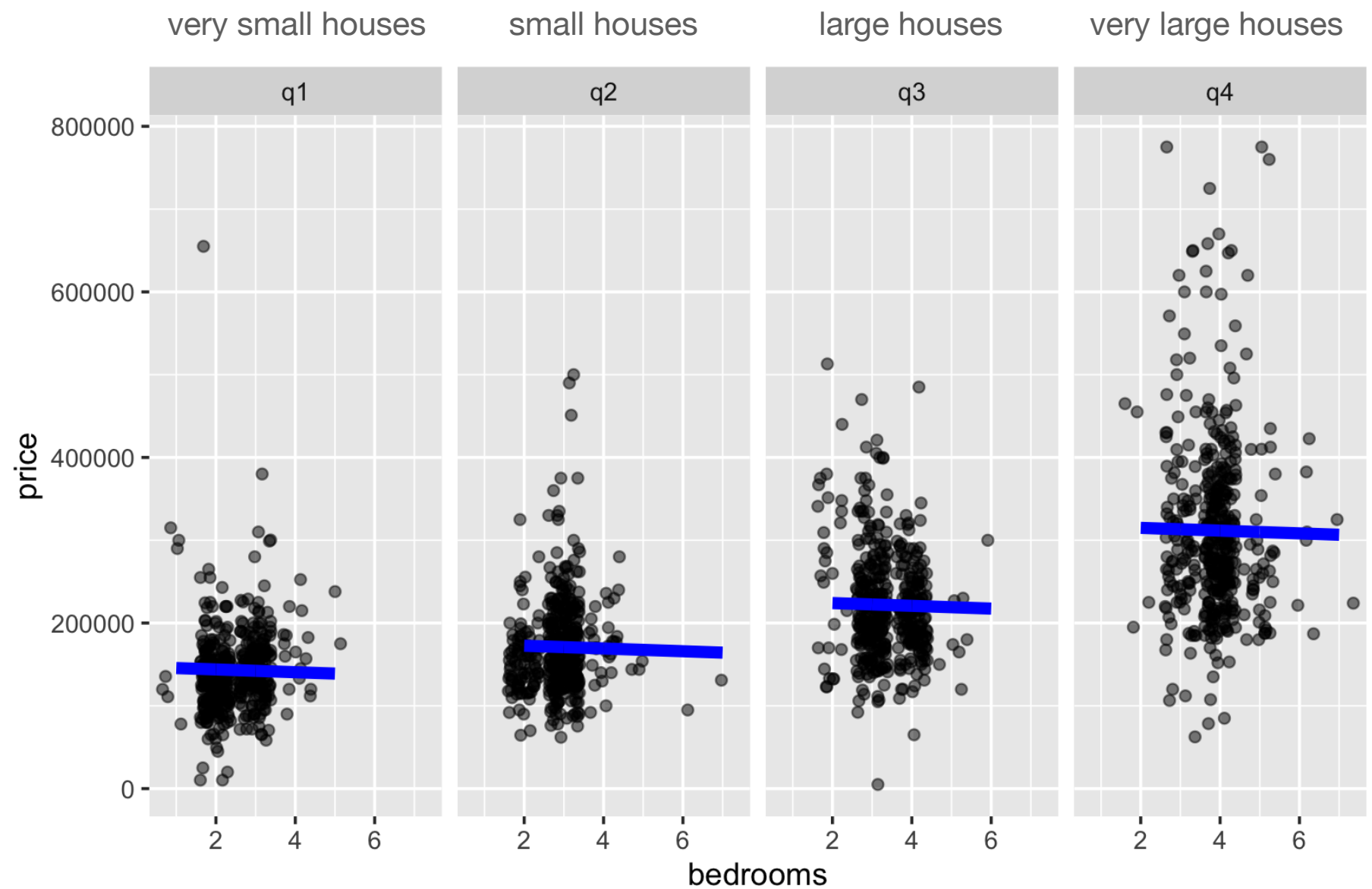
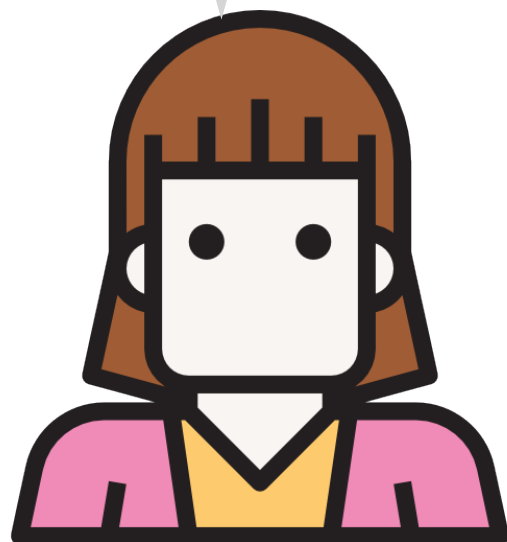


Don

Zimmerzahl ist **NEGATIV** mit dem Preis korreliert

... wenn man die Wohnfläche (Quadratmeter) kontrolliert

NE-GA-TIV!



Take-home message #1

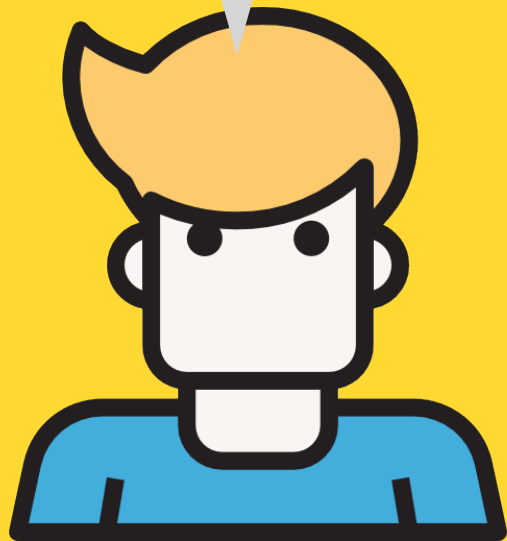
Prädiktoren hinzufügen oder wegnehmen hat oft einen Einfluss auf den Zusammenhang auf die anderen Prädiktoren im Modell.



Wolfi

Take-home message #2

Aber welchem Modell soll ich jetzt glauben?



Don

Keine Antwort von Statistik.



Wolfi

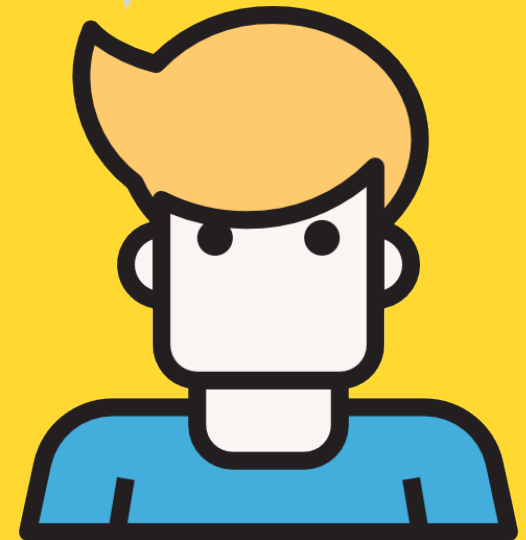
Take-home message #3

Beobachtungsstudien taugen oft nicht, um Entscheidungen zu treffen.



Wolfi

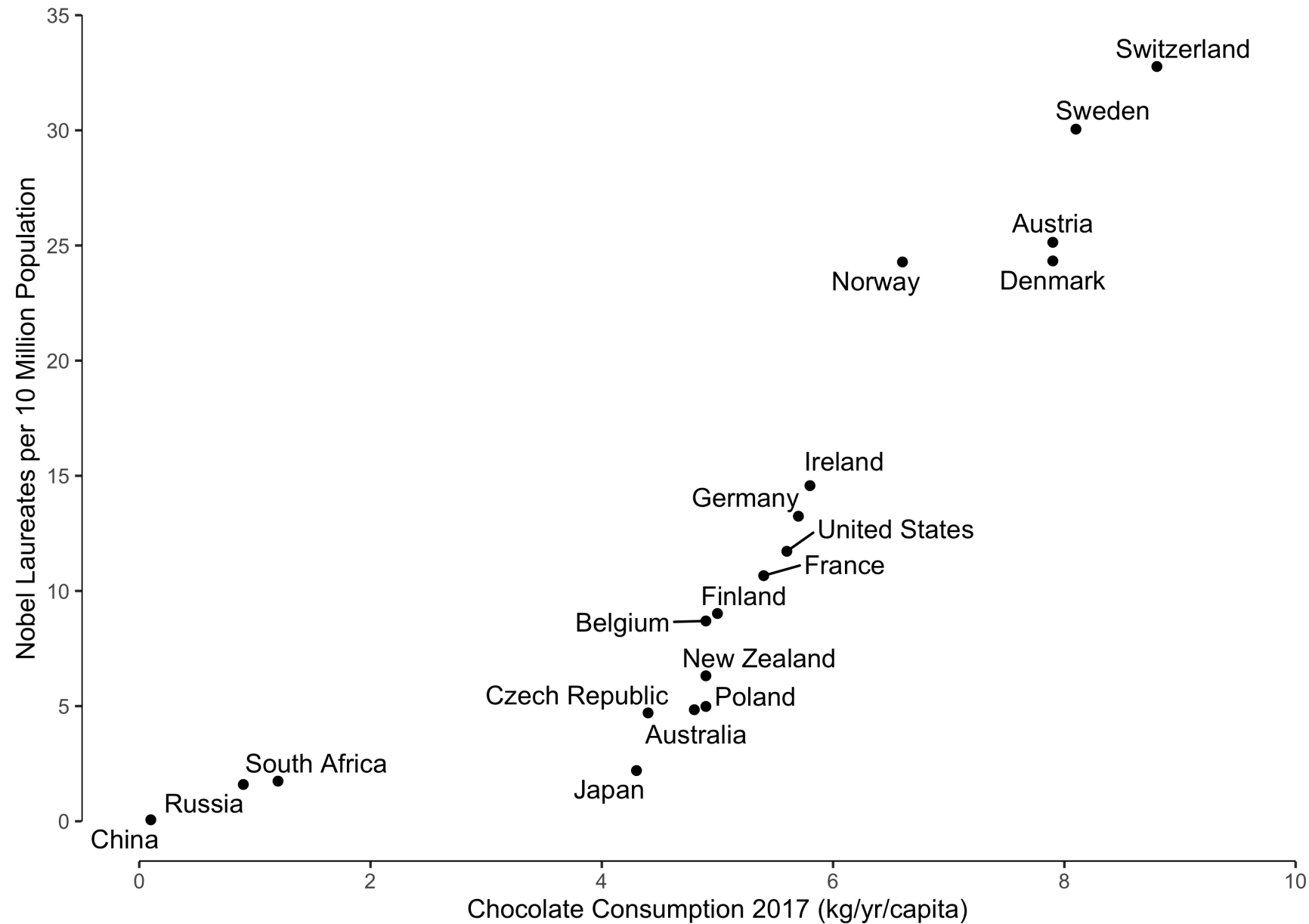
Hast du was gegen Wissenschaft?



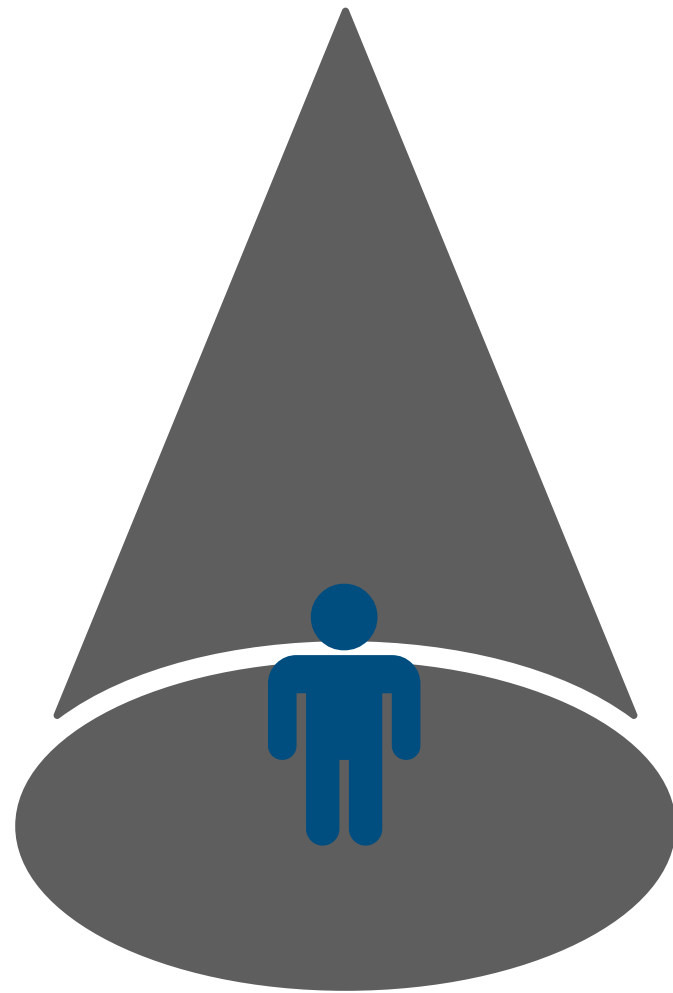
Don

Schoko futtern macht Nobelpreise?

Nobel Prizes and Chocolate Consumption



Statistische Zusammenhänge sind echt ... oder trügen

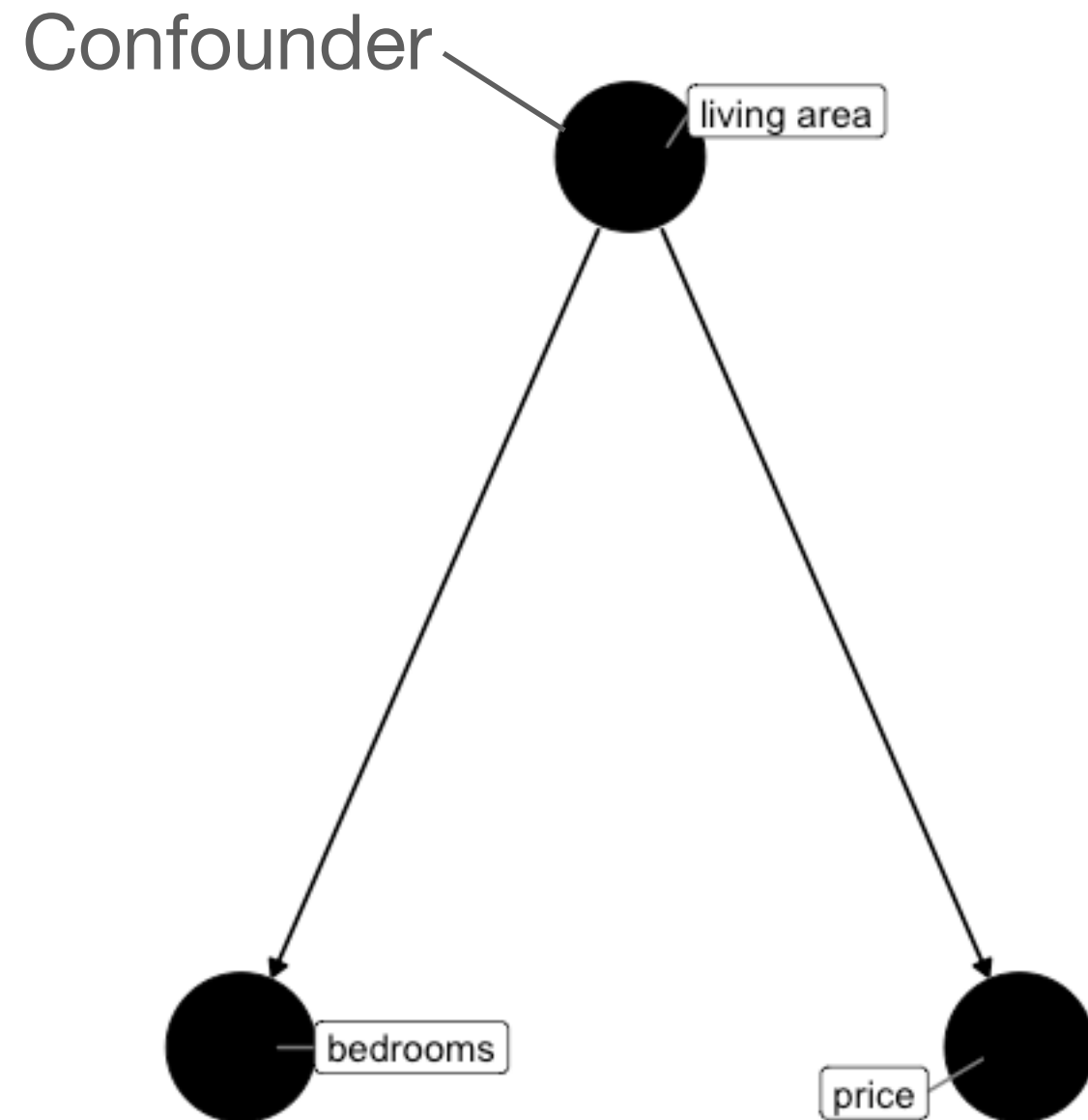


falsch: Scheinzusammenhang

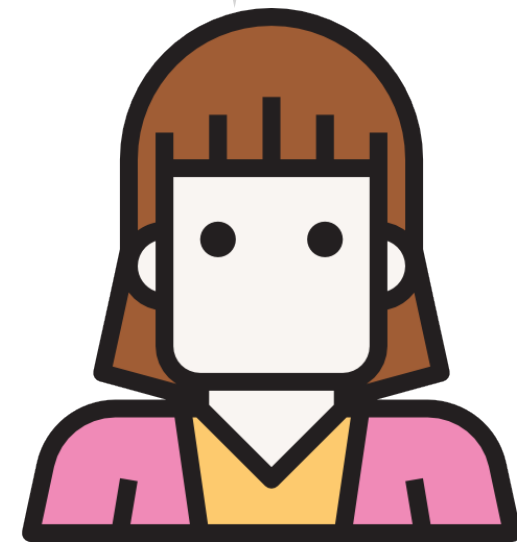


wahr: Kausalzusammenhang

Model 2: Wohnfläche als gemeinsame Ursache (Konfundierung)



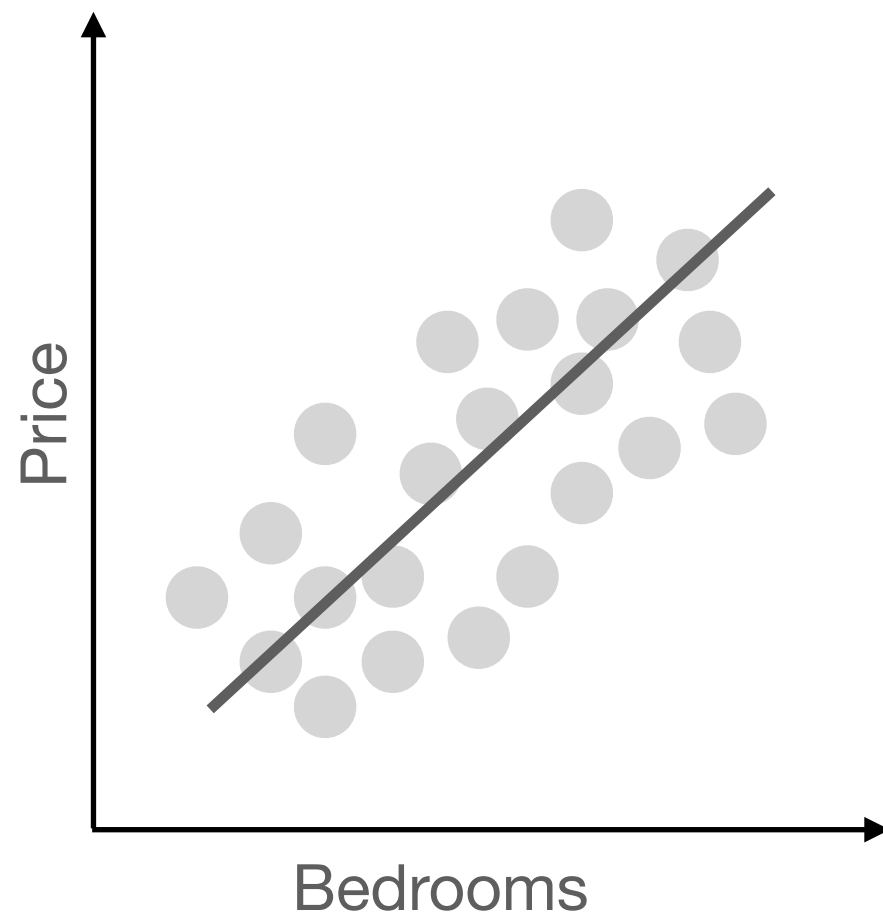
Model 2: Wir
adjustieren den
Preis.



Angi

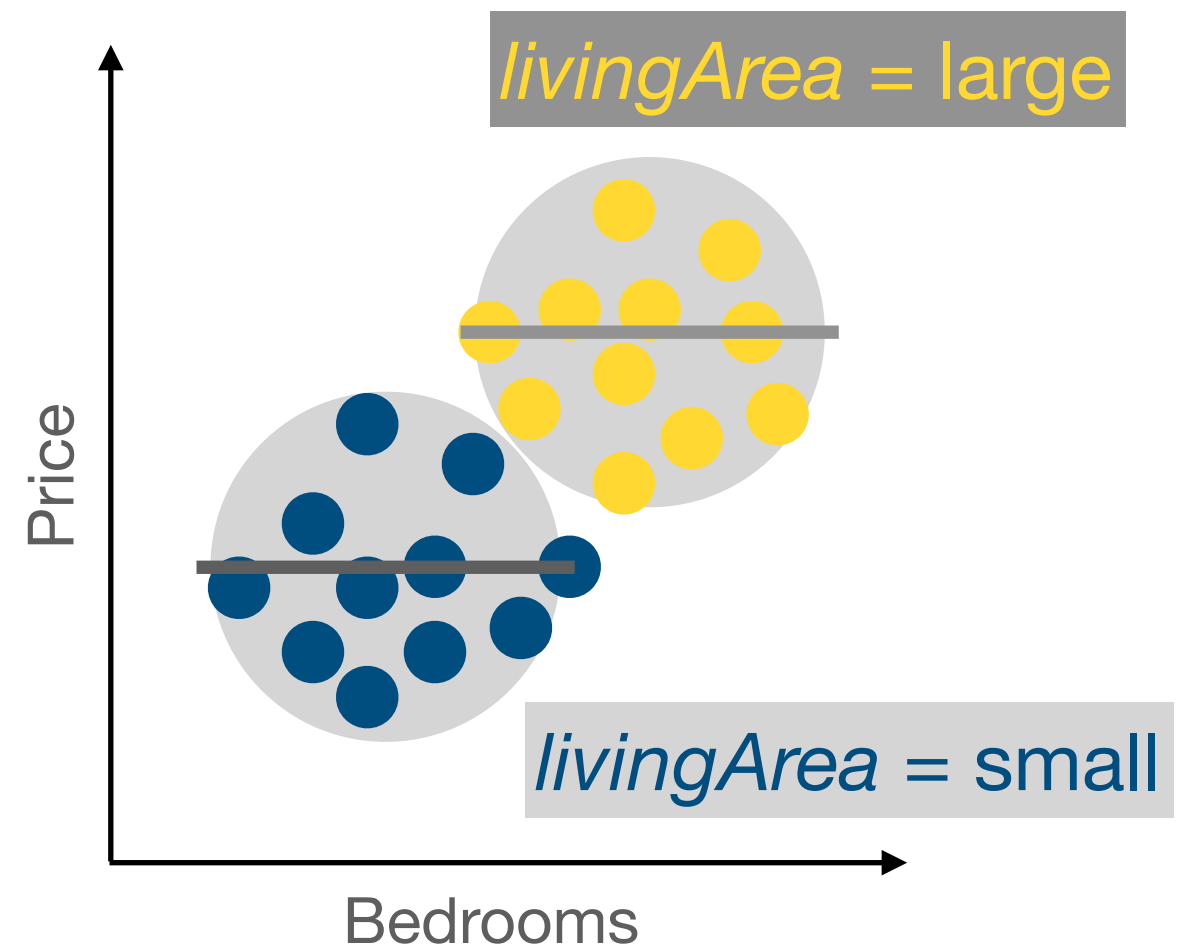
Konfundierer muss man kontrollieren

Model 1: Konfundierer *Wohnfläche*
nicht kontrolliert



Scheinkorrelation tritt auf

Model 1: Konfundierer *Wohnfläche*
ist kontrolliert



Scheinkorrelation verschwindet

Take-home message #4

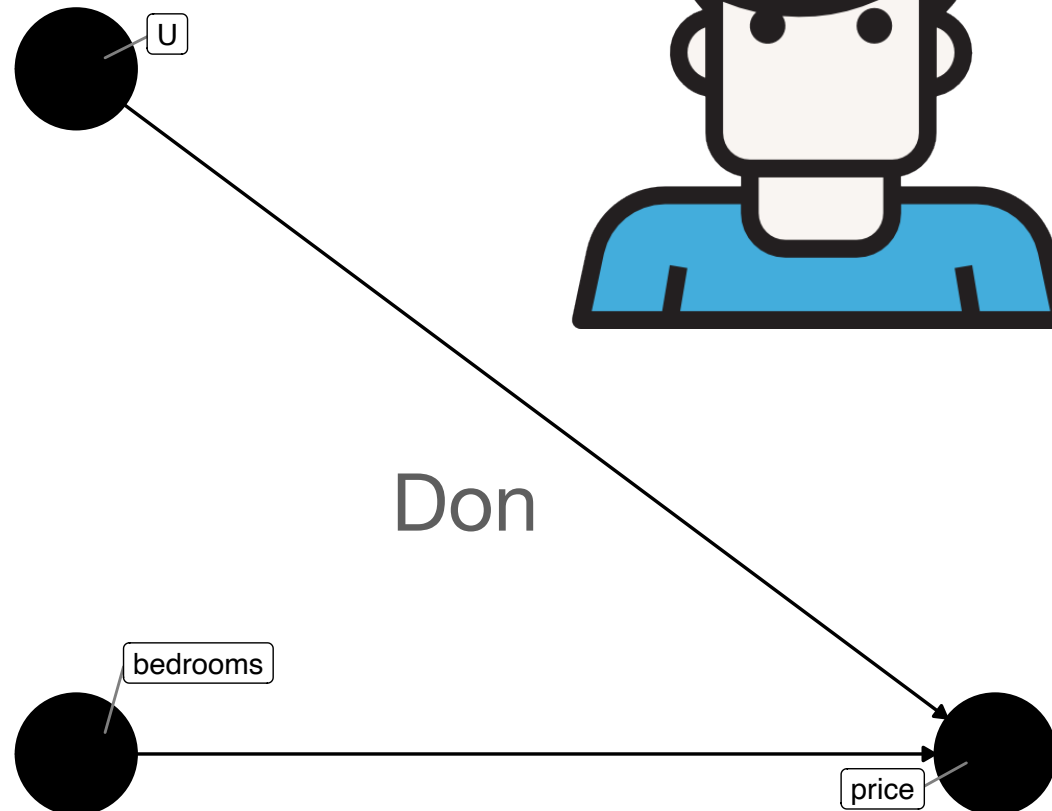
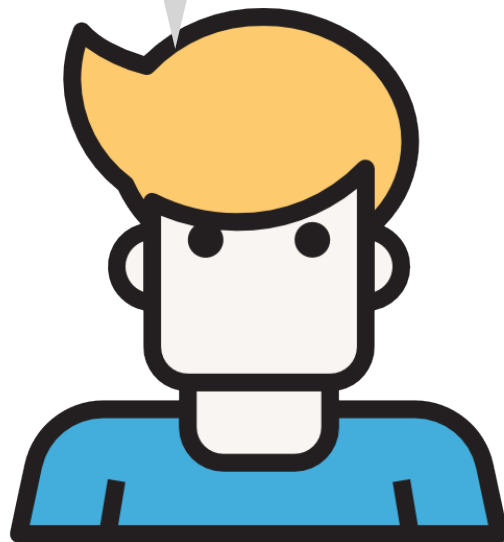


Wolfi

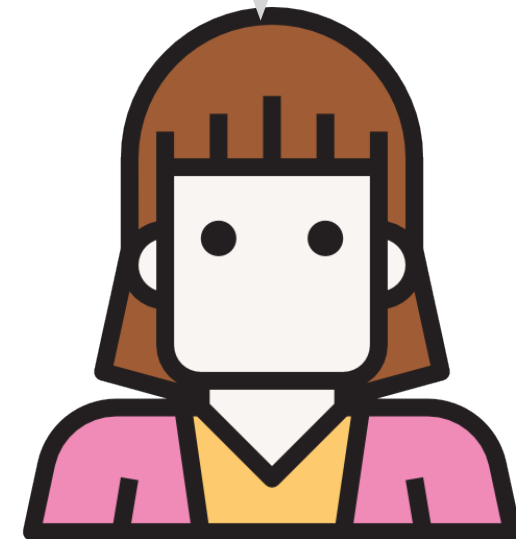
Nur mit einem Kausalmodell kann man
Scheinkorrelationen von echten, kausalen
Zusammenhängen unterscheiden

Modell 1 passt nicht zu den Daten

Modell 1 sagt, es gibt keine Konfundierung!



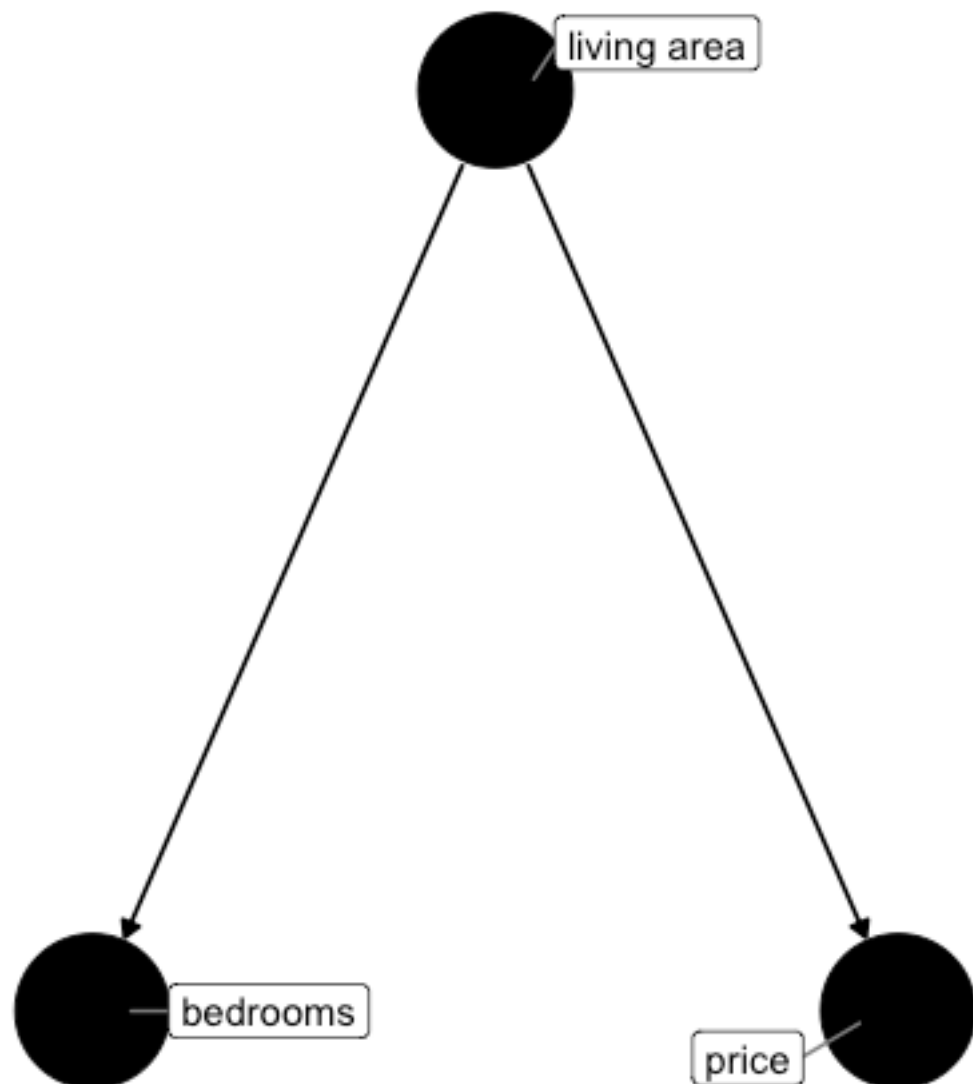
Modell 2 zeigt, dass es Konfundierung gibt. Dein Modell ist falsch, Don.



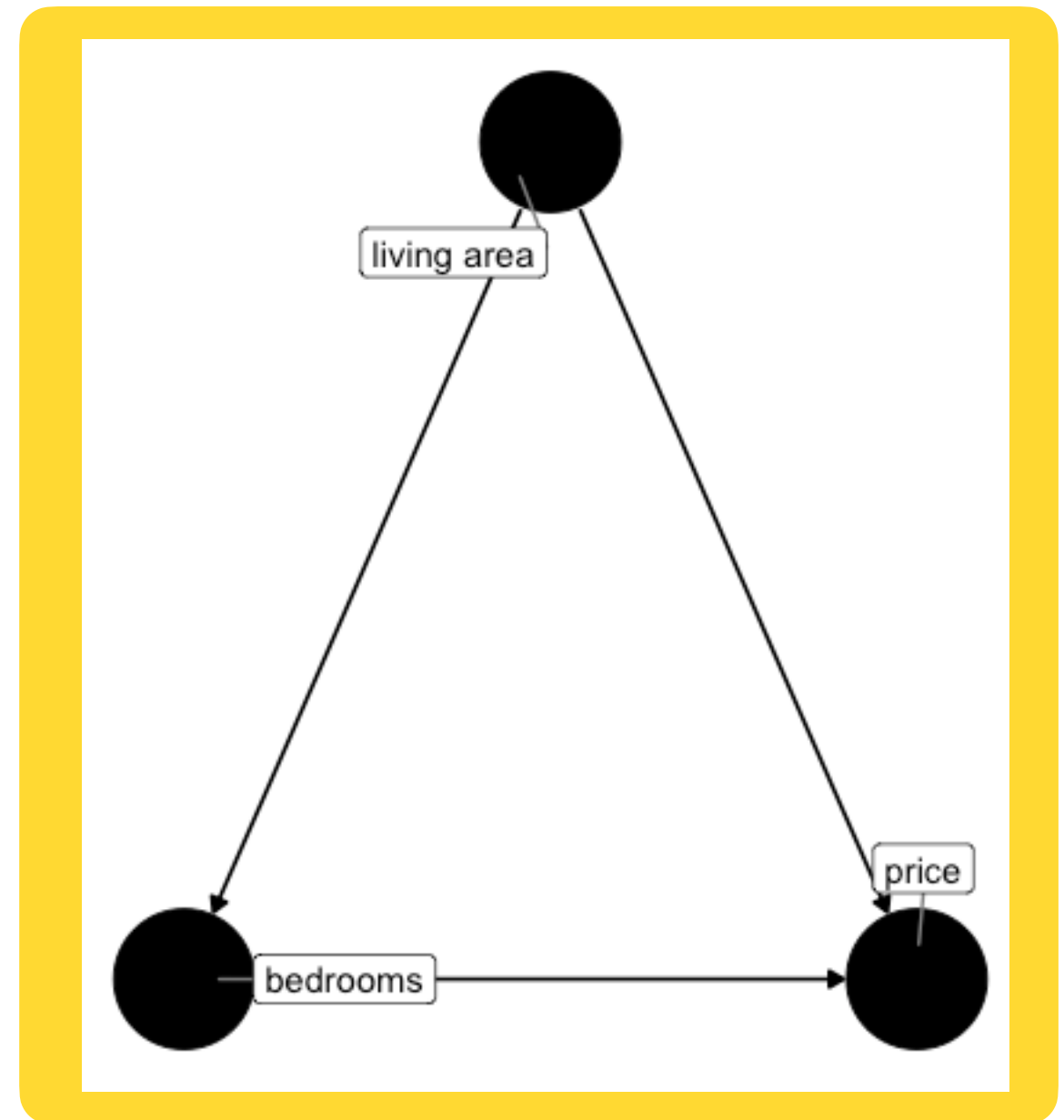
Angi

Angis Modell ist aber auch nicht richtig

Angis Modell (Modell 2)



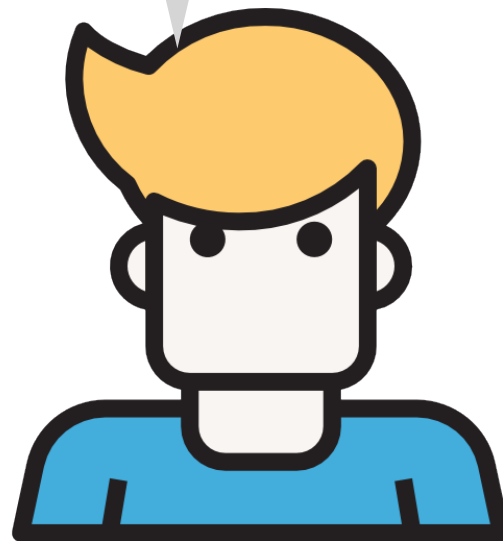
Wolfis model (Modell 3)



Wolffs Modell passt am besten auf die Daten.

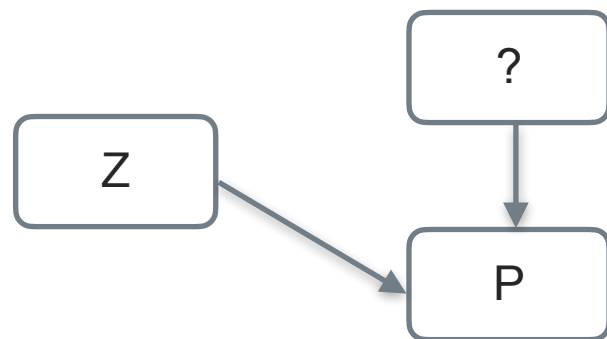
Überlegen Sie sich ein Kausalmodell für Dons Problem

Schaffst du nie,
du bist gefeuert!



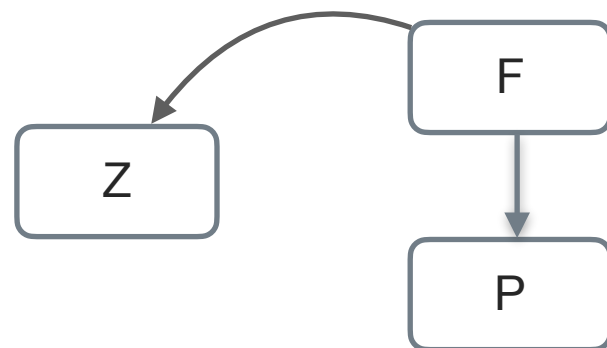
Jedes Kausalmodell impliziert (Un)Abhängigkeiten

M1



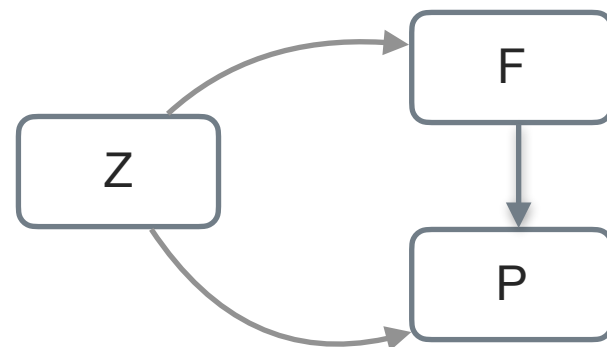
$$Z \perp ?$$

M2



$$Z \perp P \mid W$$

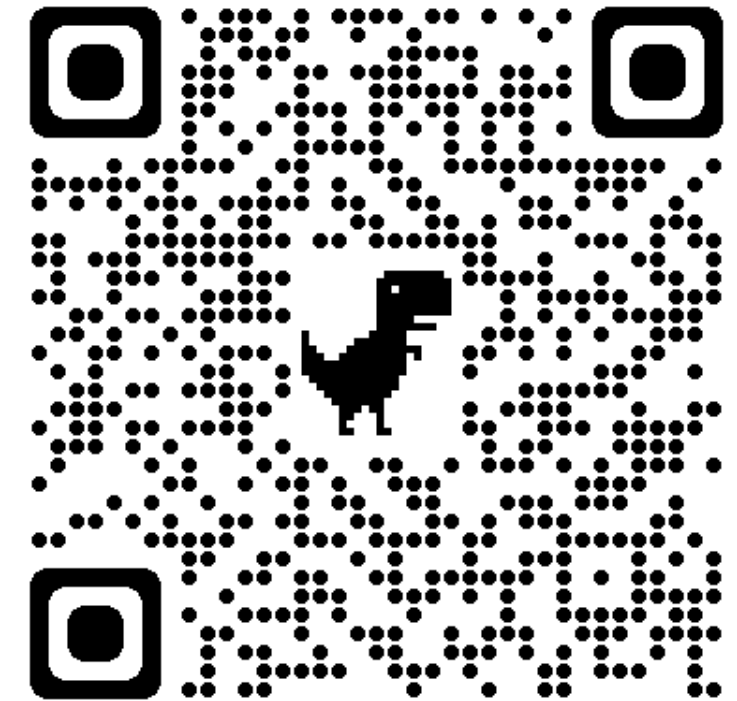
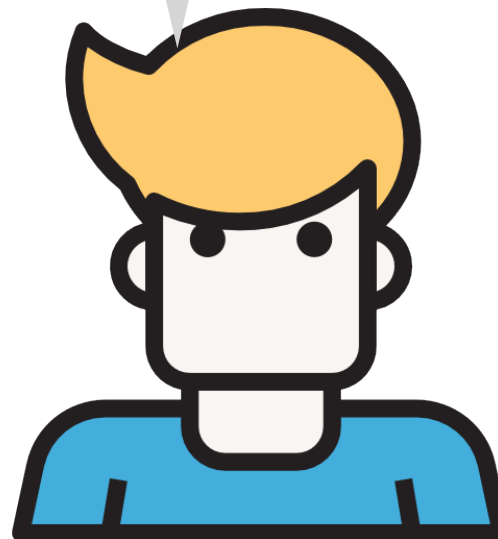
M3



$$\begin{array}{l} Z \not\perp P \\ W \not\perp P \\ W \not\perp Z \end{array}$$

Zeichnen Sie Ihr Kausalmodell bei dagitty

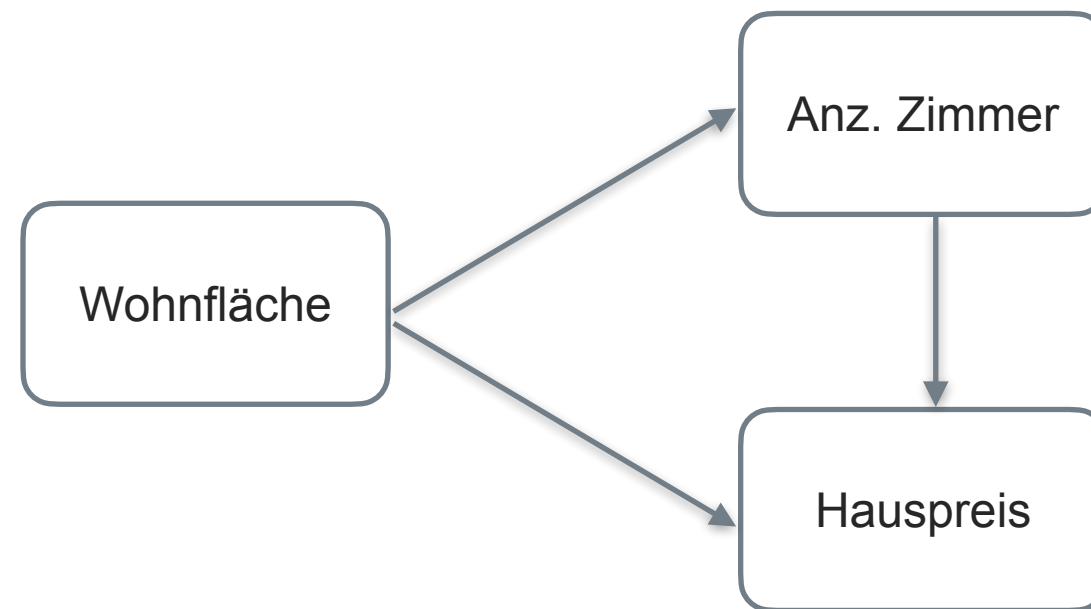
Schaffst du nie,
du bist gefeuert!



<http://dagitty.net/dags.html#>

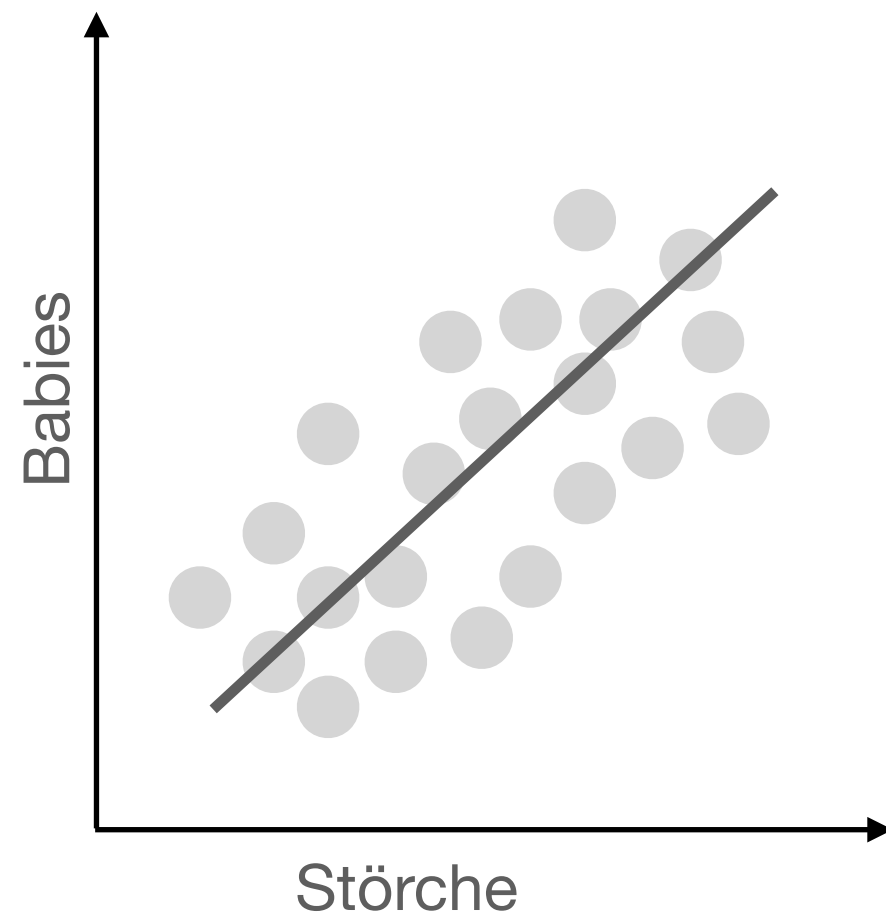
Zeit zum Üben

DAGs sind eine Art von Kausaldiagrammen

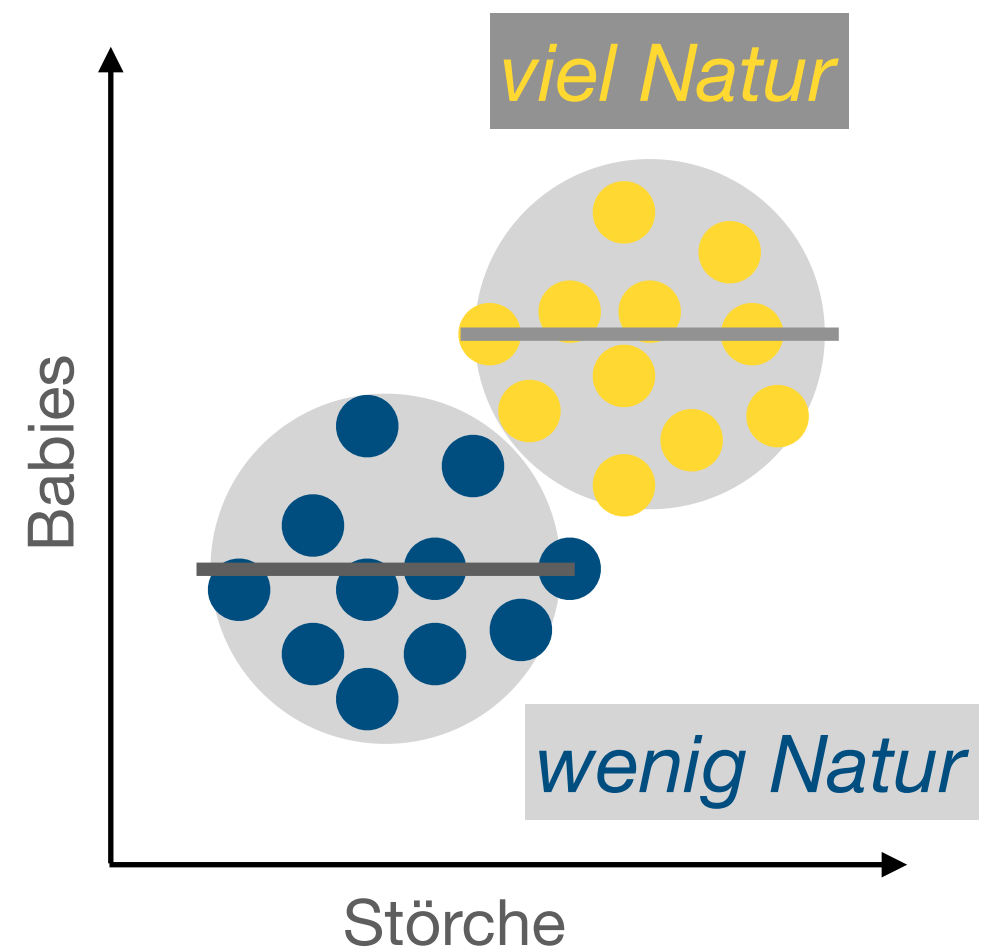


Von Störchen und Babies

Model 1: Konfundierer *Natur*
nicht kontrolliert

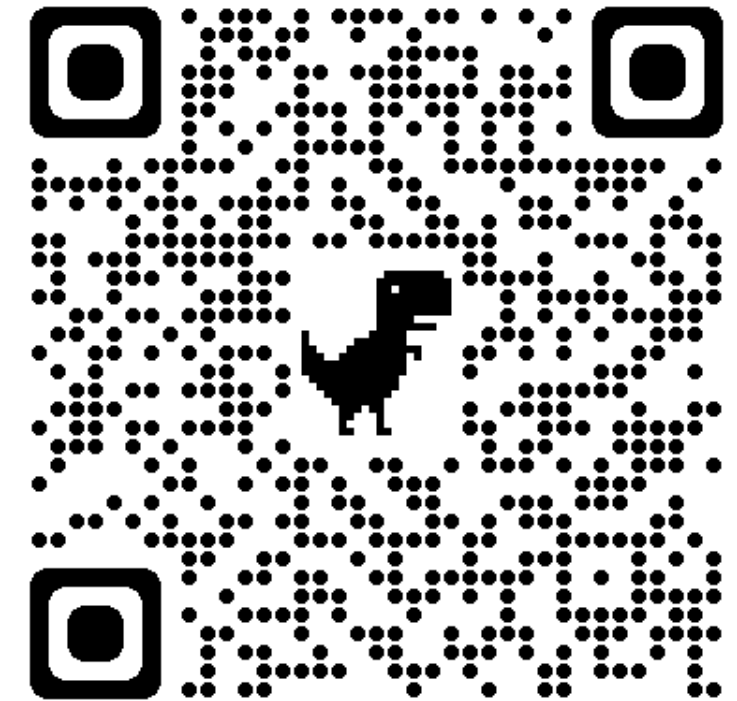
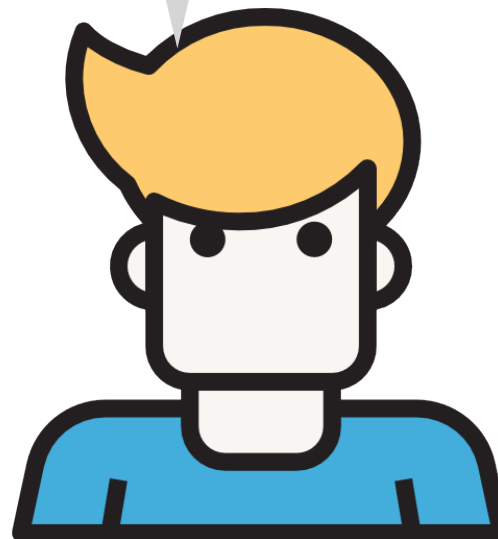


Model 1: Konfundierer *Natur*
ist kontrolliert



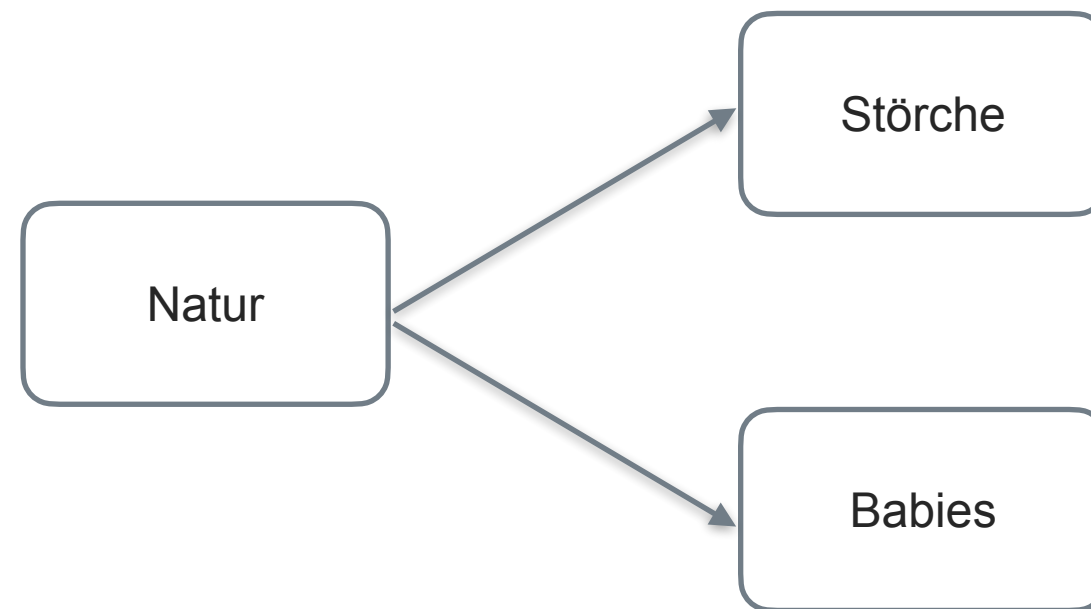
Zeichnen Sie Ihr Störche-Babies-Kausalmodell bei dagitty

Schaffst du nie,
du bist gefeuert!



<http://dagitty.net/dags.html#>

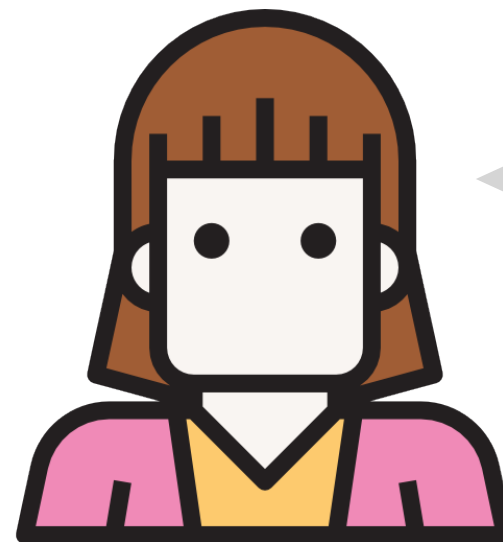
DAG für Störche-Babies konfundiert durch Natur



Von Intelligenz und Einkommen

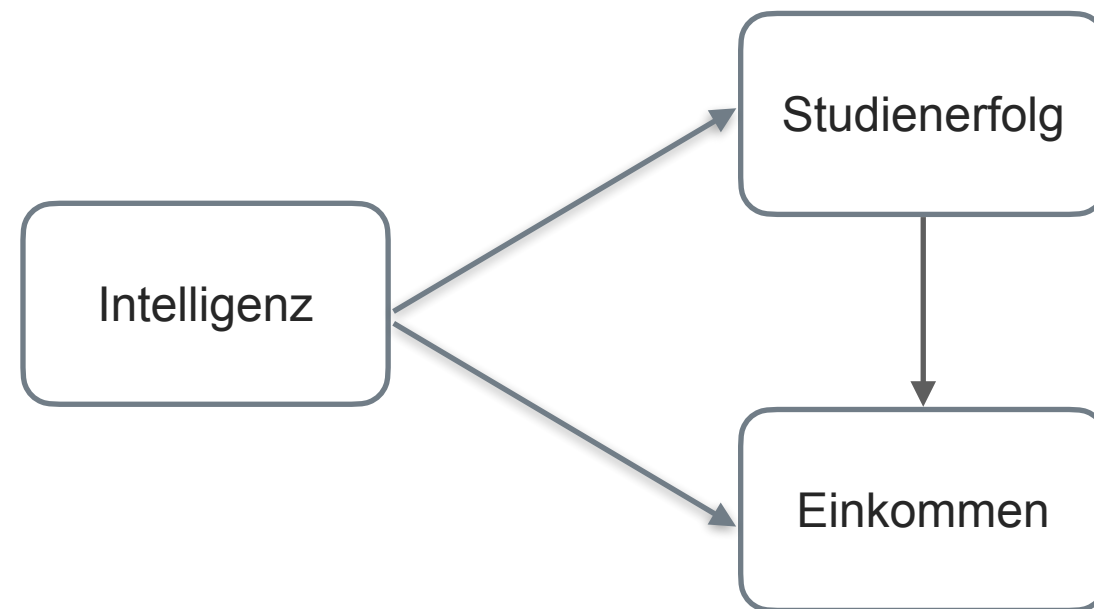
Intelligenz → Studienerfolg → Einkommen
Intelligenz → Einkommen

Wie sieht der DAG aus?



Wenn man am Effekt des Studienerfolgs interessiert ist:
Sollte man die Teil-Daten
(Studienerfolg hoch/gering)
betrachten?
Oder die Gesamt-Daten?

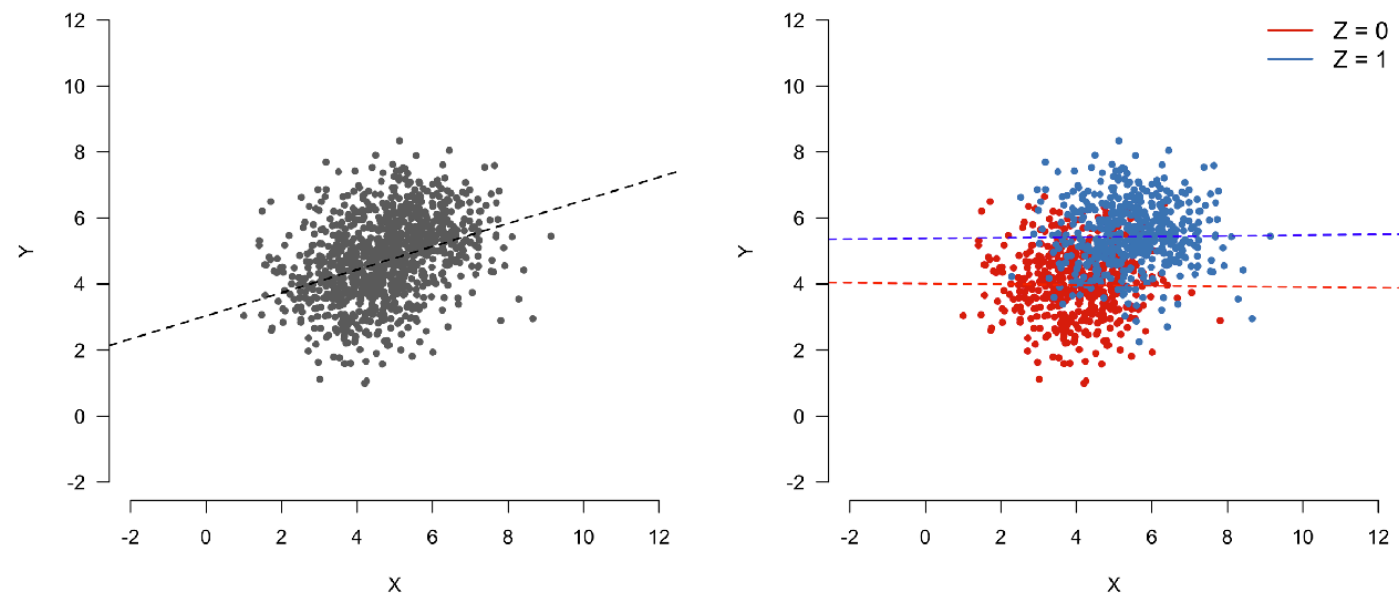
DAG für Intelligenz, Studienerfolg und Einkommen



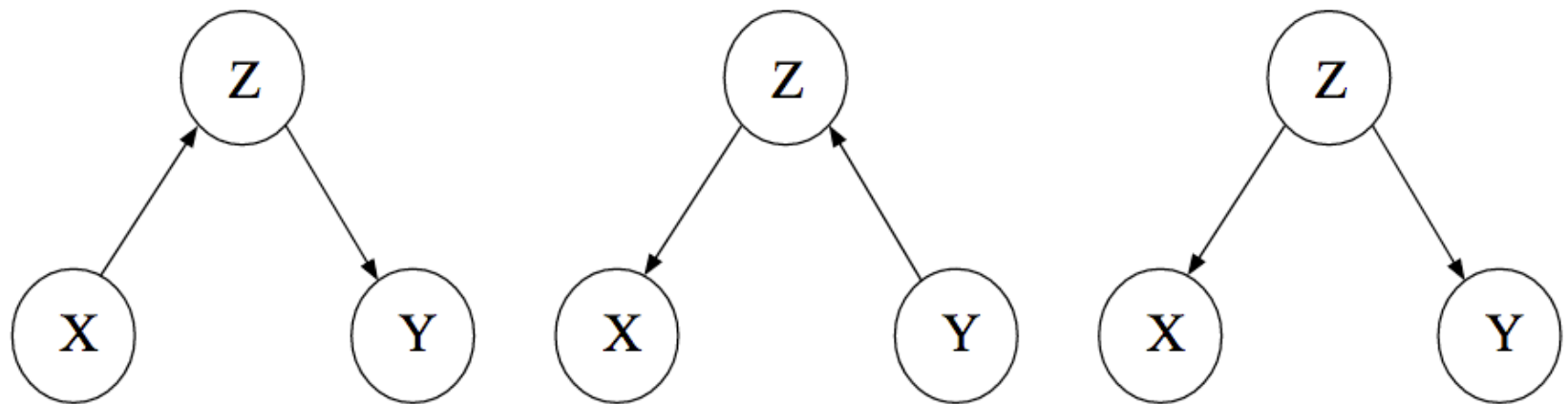
Ist man am direkten Effekt von Studienerfolg interessiert, so sollte man Intelligenz blockieren (z.B. Intelligenz stratifizieren).

Zu einer Datenlage passen leider meist viele DAGs

Zu dieser
Datenlage ...

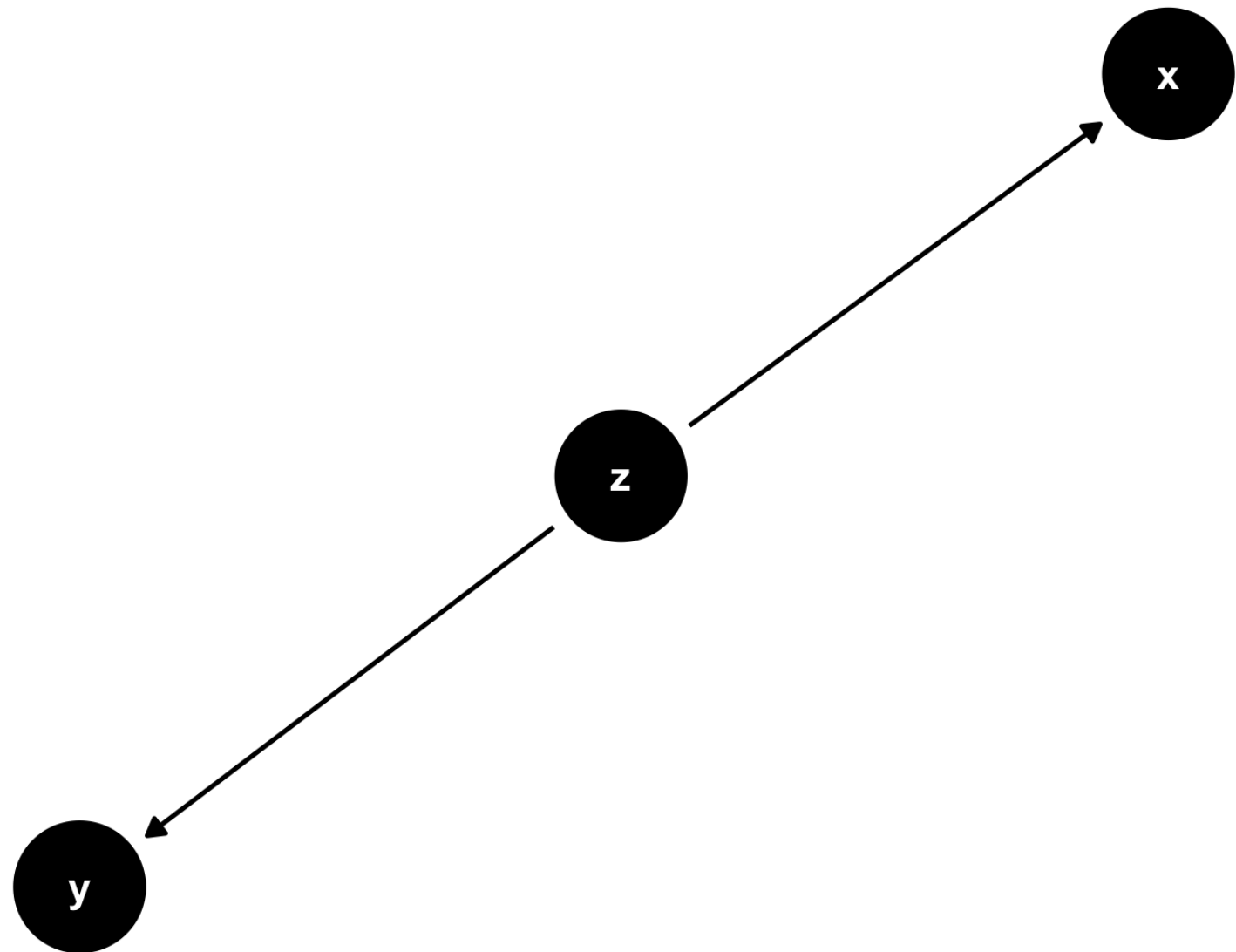
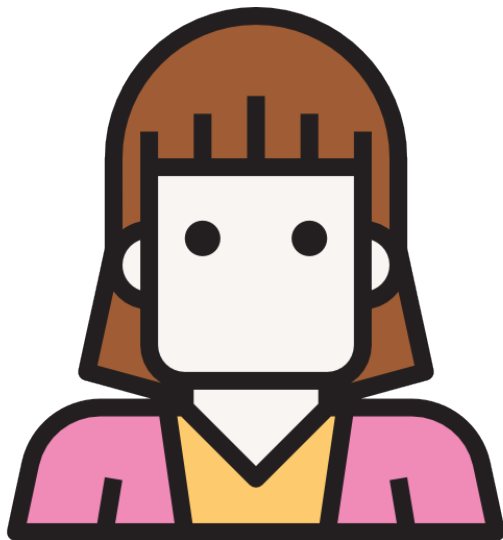


passen diese
DAGs:



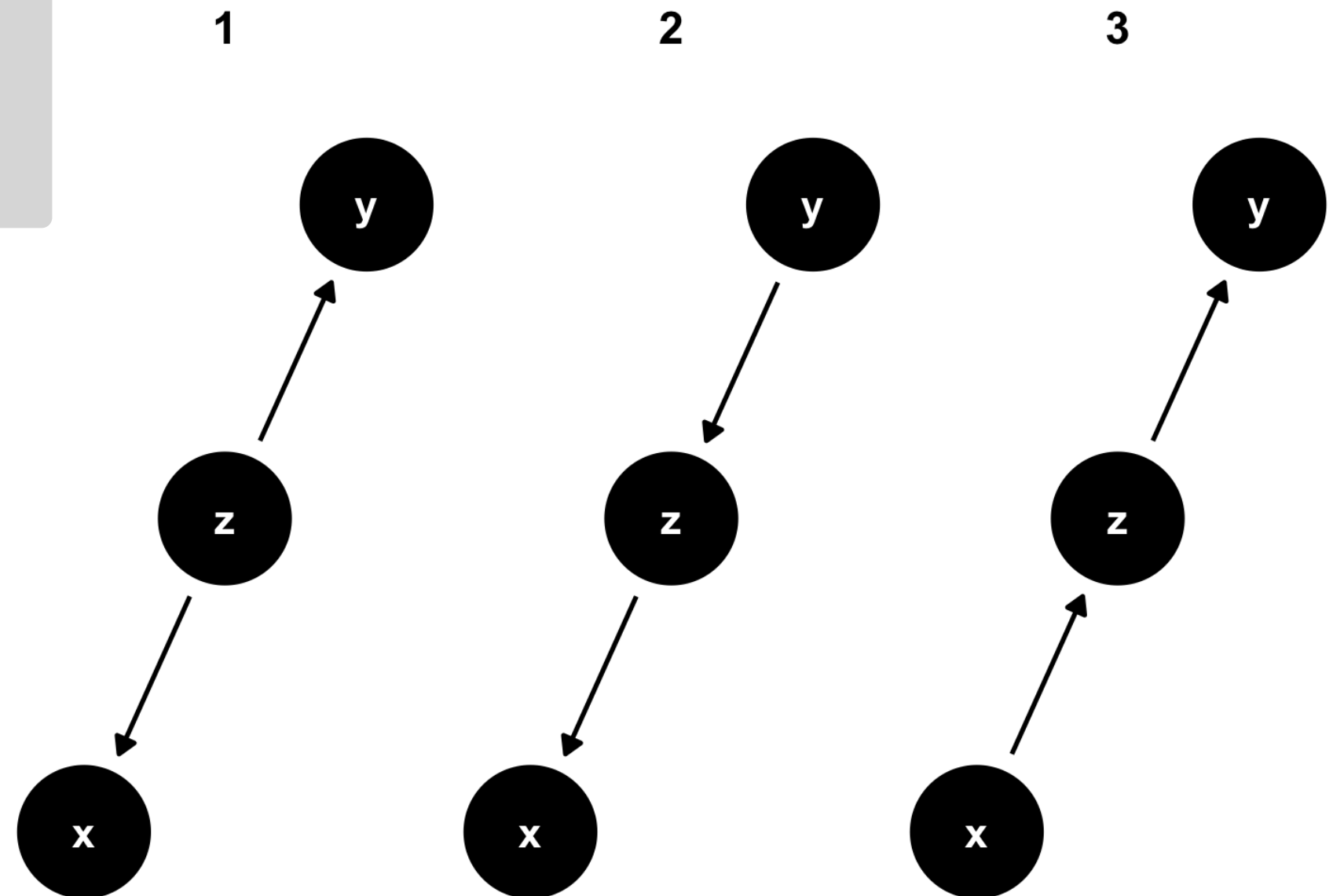
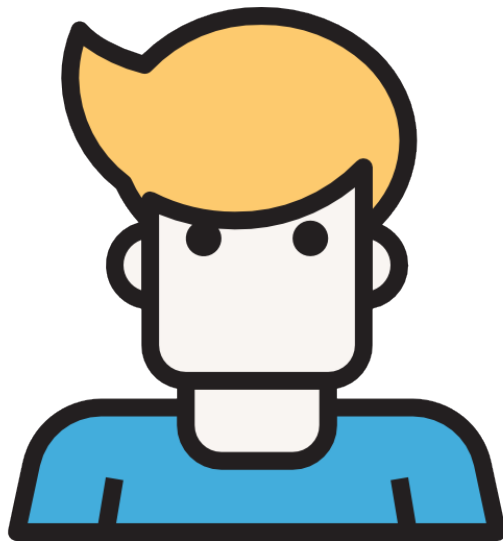
Welche DAGs führen zu den gleichen Daten?

Es gibt noch andere DAGs, die die gleichen Daten produzieren, wie dieser DAG. Aber welche?



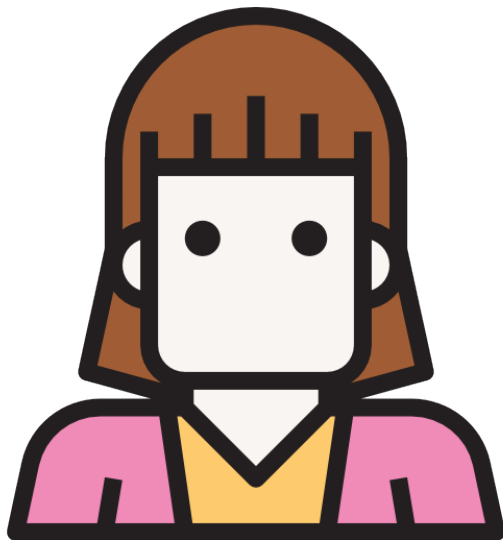
Mehrere DAGS erzeugen die gleiche Datenstruktur

Daten kann man nicht trauen?!



Konfundierung, Montezumas Rache bei Beobachtungsstudien

Da gibt's viele
Beispiele.



X	Y	Z
?	?	?
...		
...		

Konfundierer blockieren

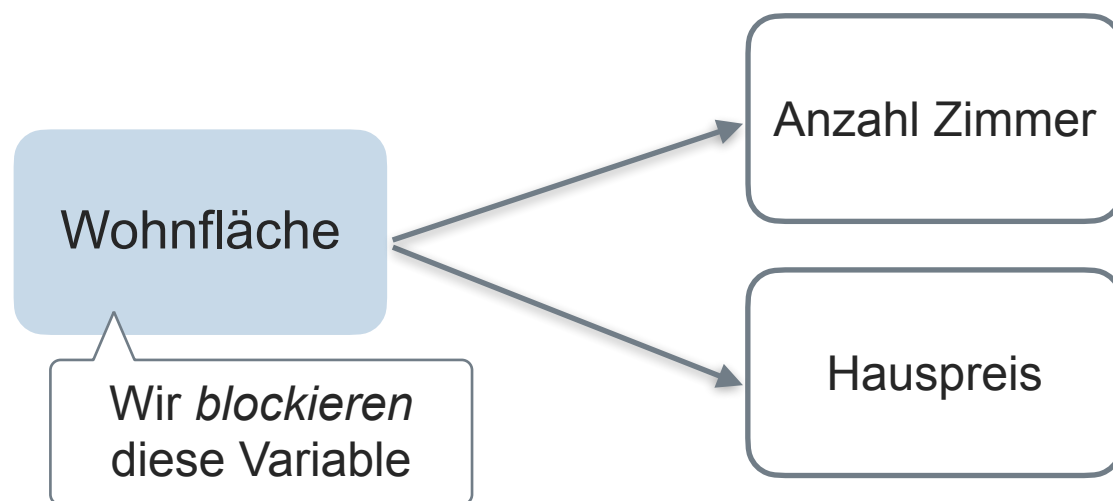
Was tun bei Konfundierung? Blockieren!

Blockieren

1. Aufnahme der Konfundierungsvariable in ein Regressionsmodell
2. Getrennte Analyse von Subgruppen
3. Durchführung eines randomisierten Experiments
4. ...

- Durch Blockieren wird der „Durchfluss“ von statistischer Abhängigkeit blockiert (unterbrochen).
- Um Kausaleffekte zu entdecken, müssen wir nicht-kausale Pfade blockieren, so dass nur „echte“, d.h. kausale Pfade übrig bleiben.

Gibt es einen kausalen Effekt von Zimmer auf Hauspreis?



- Durch das Blockieren von Wohnfläche kann keine statistische Abhängigkeit (z. B. Korrelation) mehr von *Zimmer* über Wohnfläche nach *Hauspreis* fließen
- Der Pfad von *Wohnfläche* nach *Hauspreis* ist ebenfalls geschlossen
- Laut dem Modell darf keine (statist.) Abhängigkeit zwischen *Zimmer* und *Hauspreis* übrig bleiben, wenn man *Wohnfläche* blockiert.

Was tun bei Konfundierung? Blockieren!

Blockieren

1. Aufnahme der Konfundierungsvariable in ein Regressionsmodell
2. Getrennte Analyse von Subgruppen
3. Durchführung eines randomisierten Experiments
4. ...

1. Aufnahme der Konfundierungsvariable in ein Regressionsmodell:

```
my.adj.model <- lm(price ~ bedrooms + livingArea,  
                   data = SaratogaHouses)
```

```
my.adj.model
```

Coefficients:

(Intercept)	bedrooms	livingArea
36667.9	-14196.8	125.4

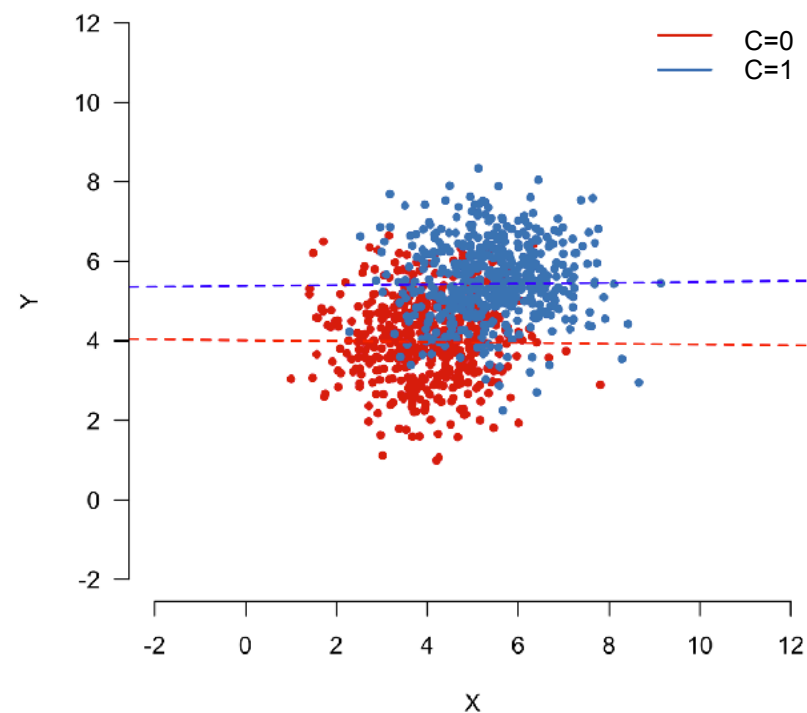
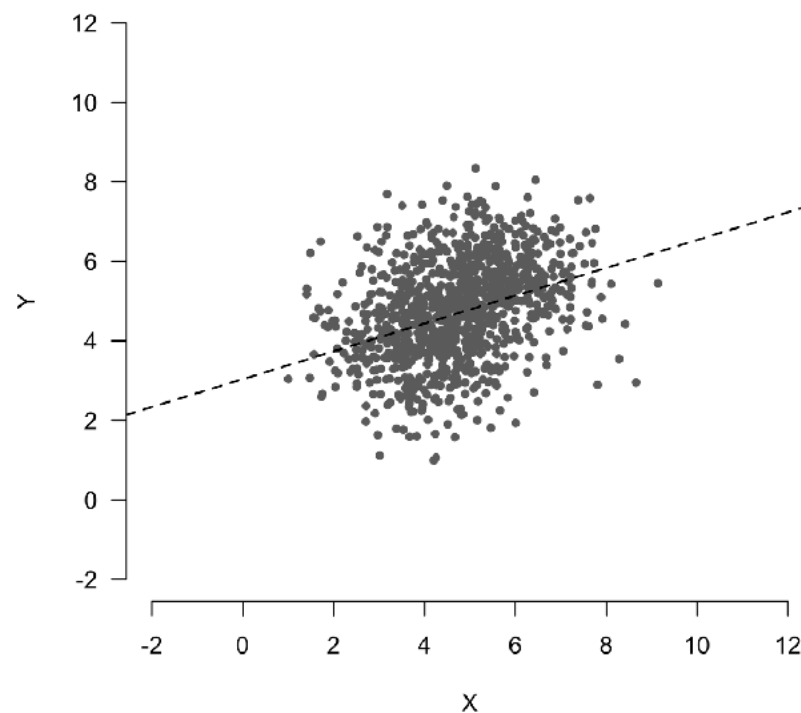
Tatsächlich hat die Variable *Anzahl der Zimmer* (bedrooms) einen Effekt, allerdings wirkt sie negativ (nicht positiv) auf den *Hauspreis* (price)!

Was tun bei Konfundierung? Blockieren!

Blockieren

1. Aufnahme der Konfundierungsvariable in ein Regressionsmodell
- 2. Getrennte Analyse von Subgruppen**
3. Durchführung eines randomisierten Experiments
4. ...

2. Getrennte Analyse von Subgruppen (Konstanthalten, Stratifizieren)

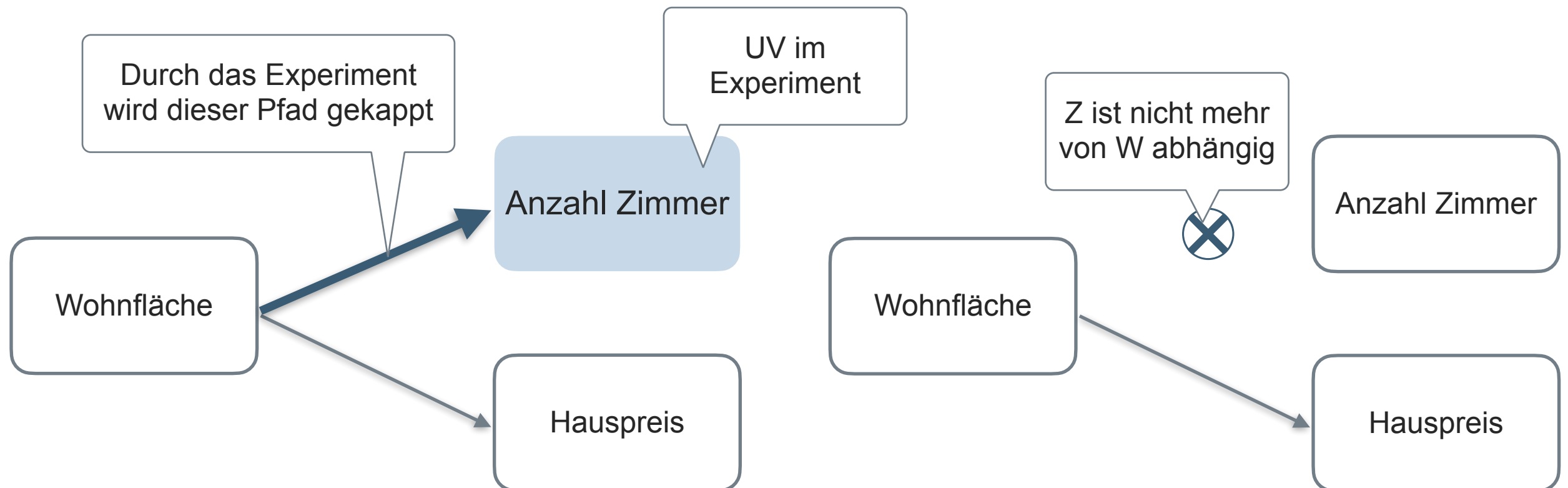


Was tun bei Konfundierung? Blockieren!

Blockieren

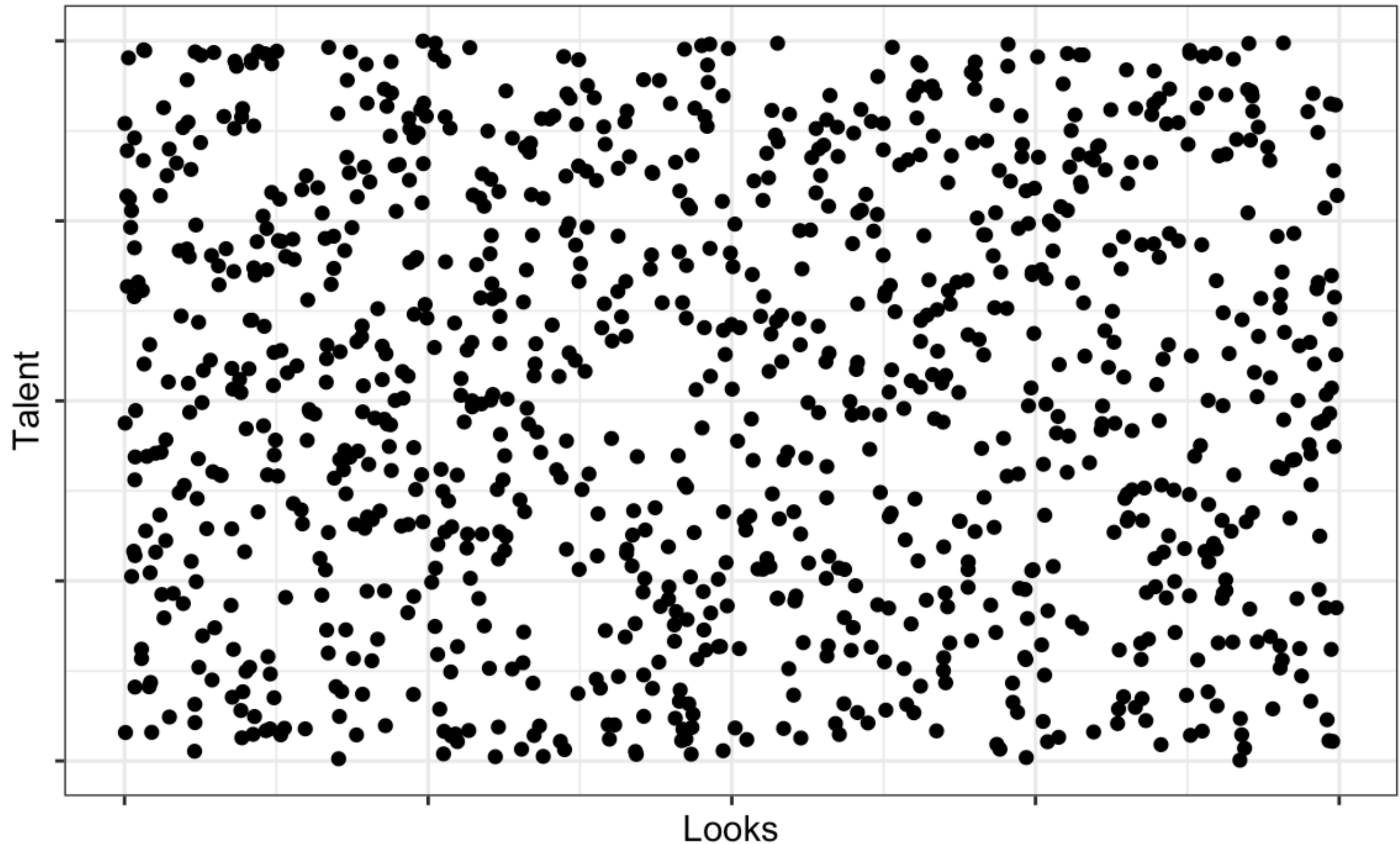
1. Aufnahme der Konfundierungsvariable in ein Regressionsmodell
2. Getrennte Analyse von Subgruppen
- 3. Durchführung eines randomisierten Experiments**
4.

3. Durchführung eines randomisierten Experiments

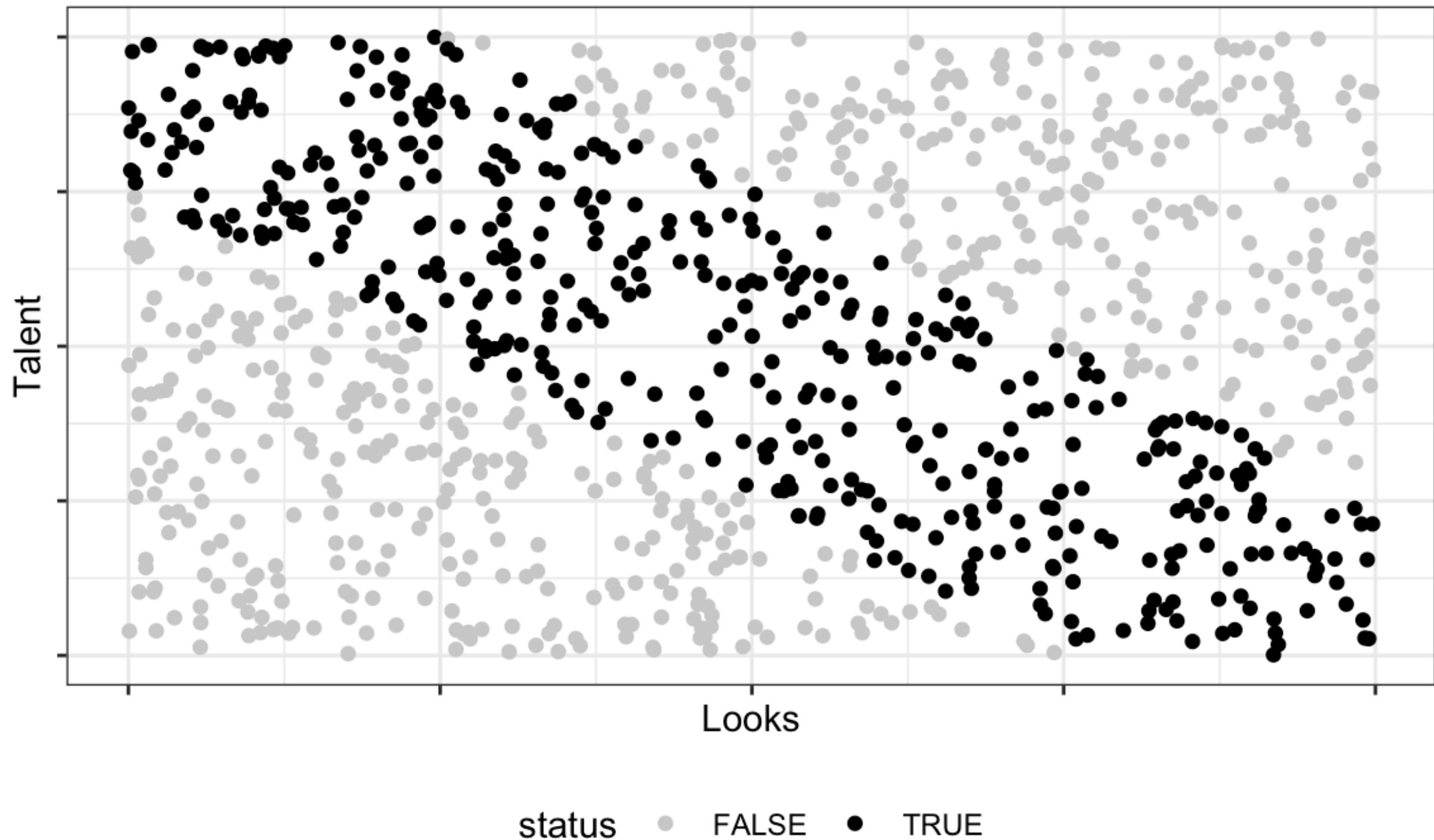


Gott ist gerecht:
Schönheit und Intelligenz
sind unabhängig (?)

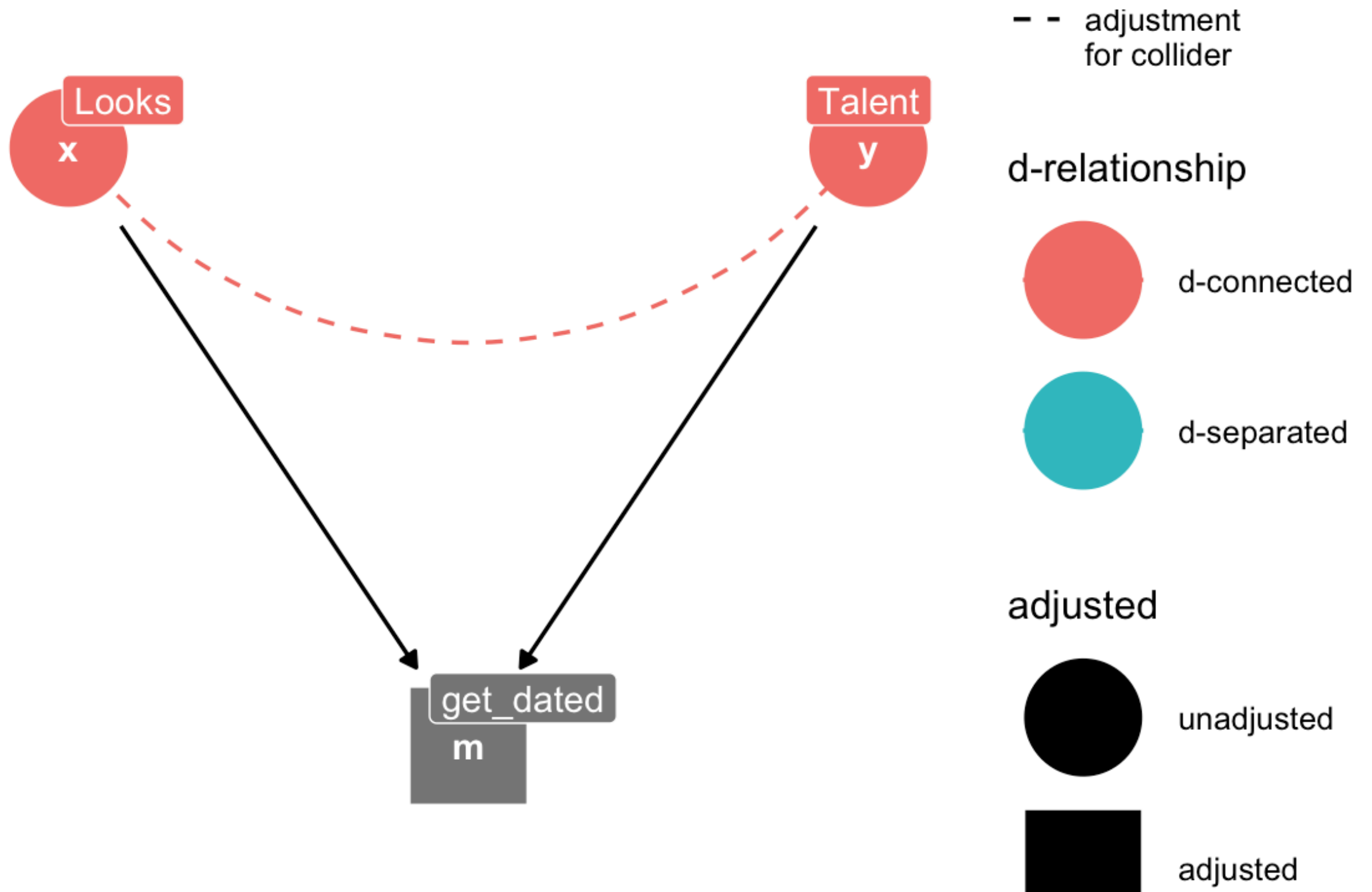
In Wahrheit kein Zusammenhang von Intelligenz und Schönheit



Ihre Dates sind aber entweder schön oder schlau (?)

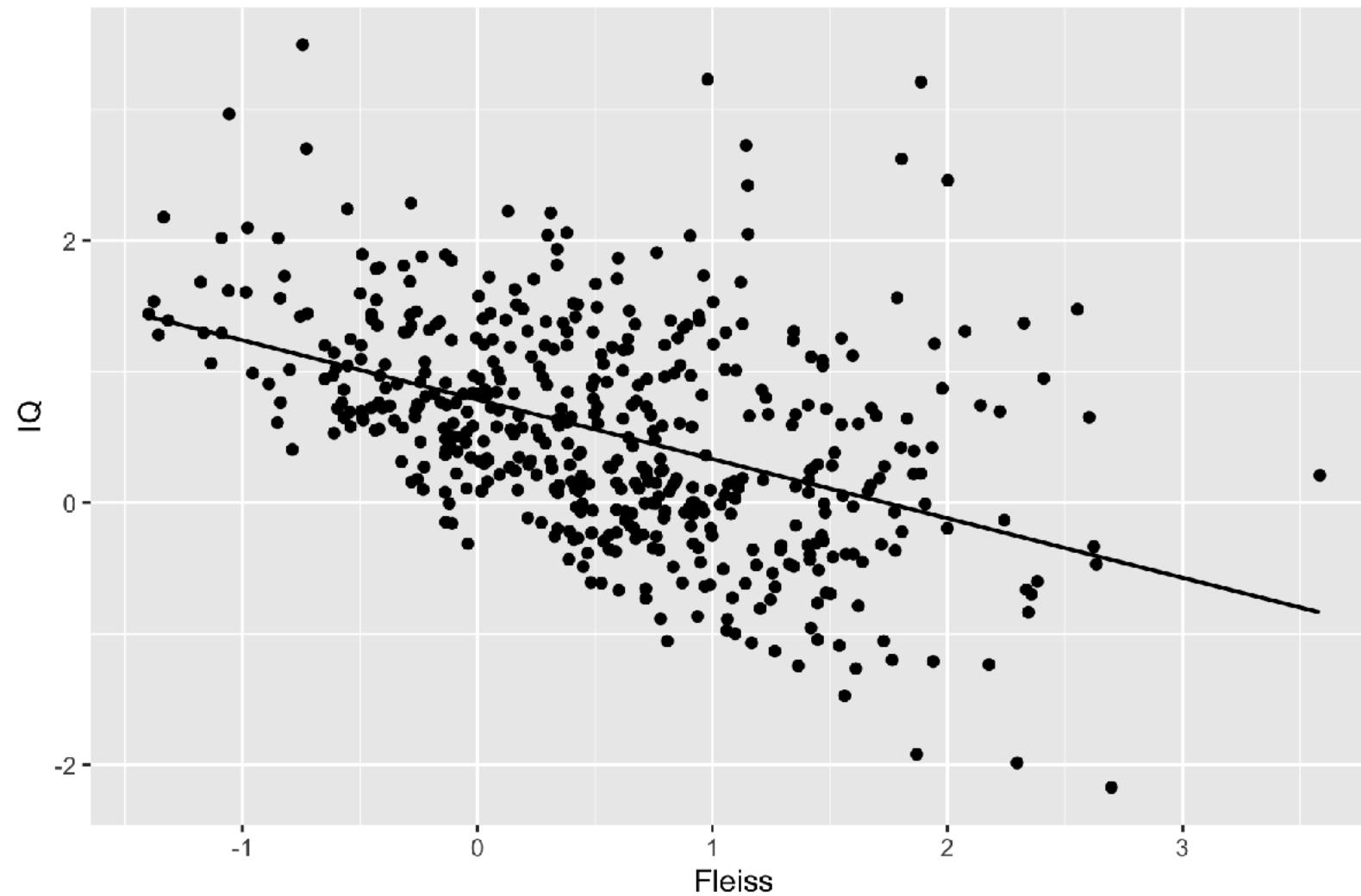


DAG zeigt bedingte Abhängigkeit



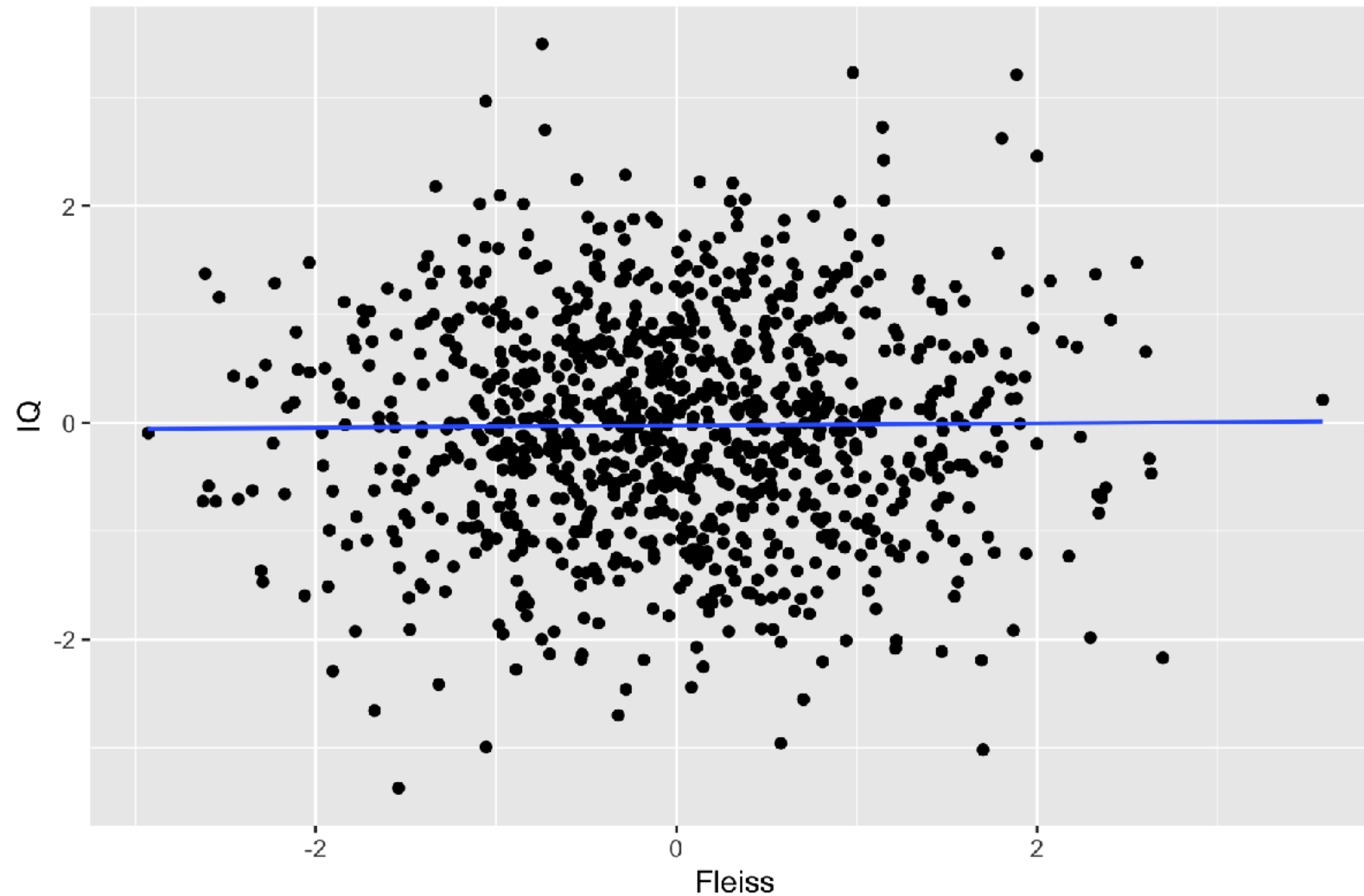
Kollisions- Verzerrung

Die Schlaunen sind faul, sagt eine wissenschaftliche Studie



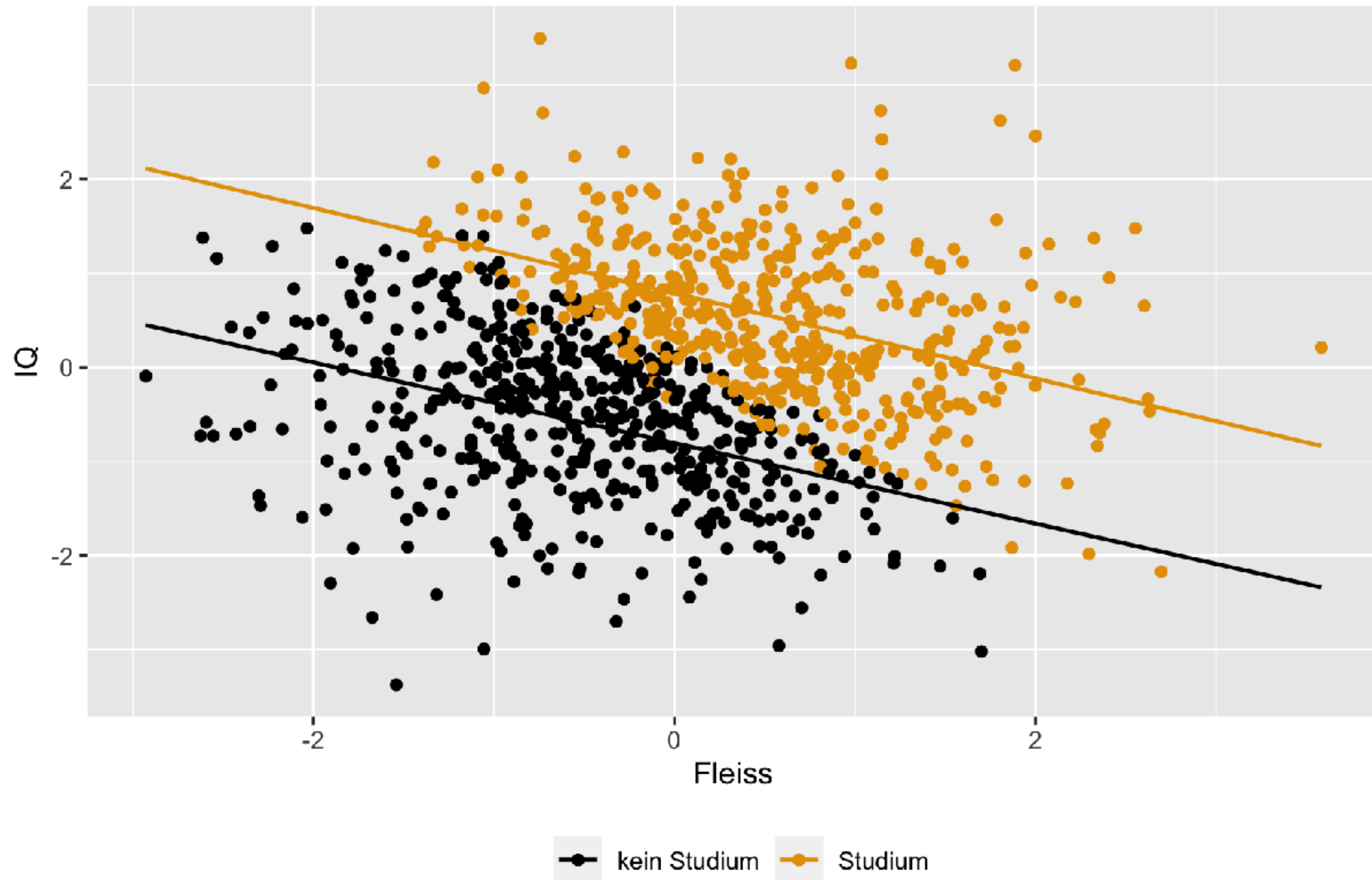
n=492 Studis

Replikation (Studie 2) fand keinen Zusammenhang



n=1000 Personen (Studis und Nicht-Studis)

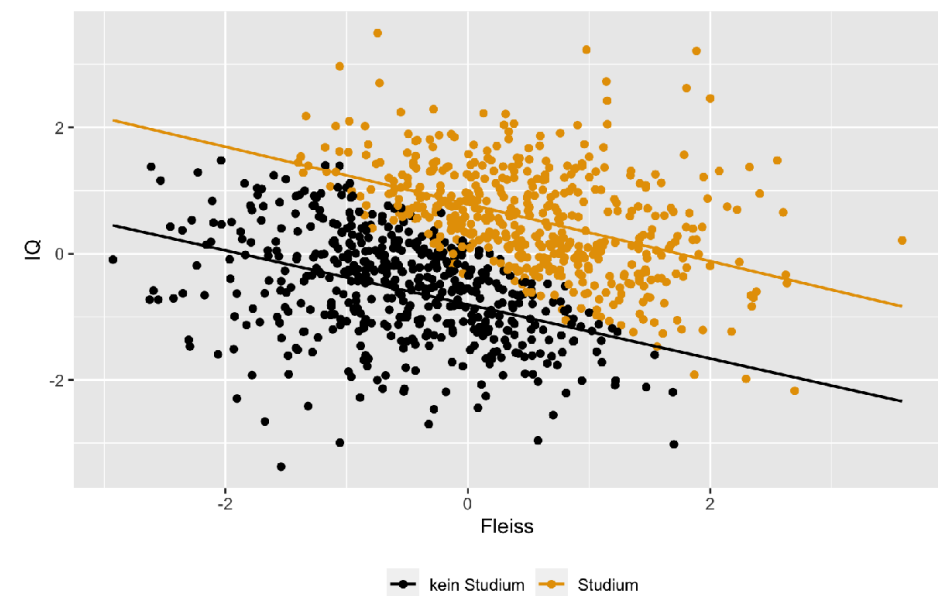
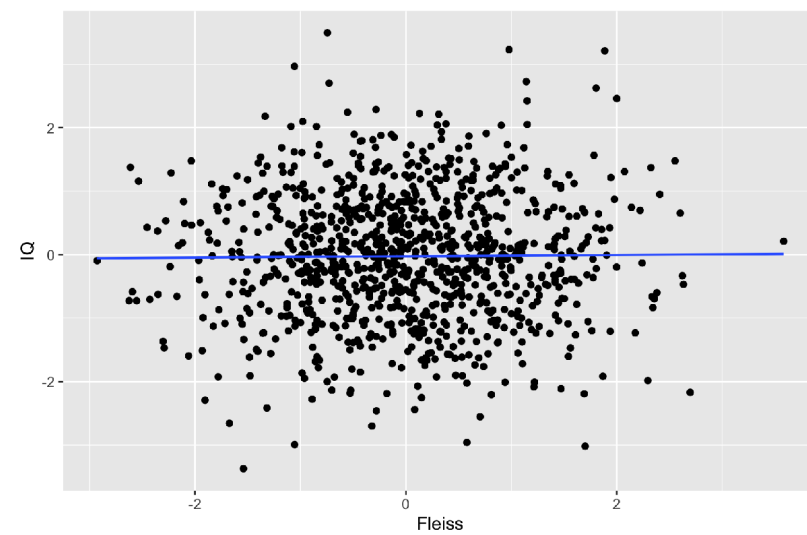
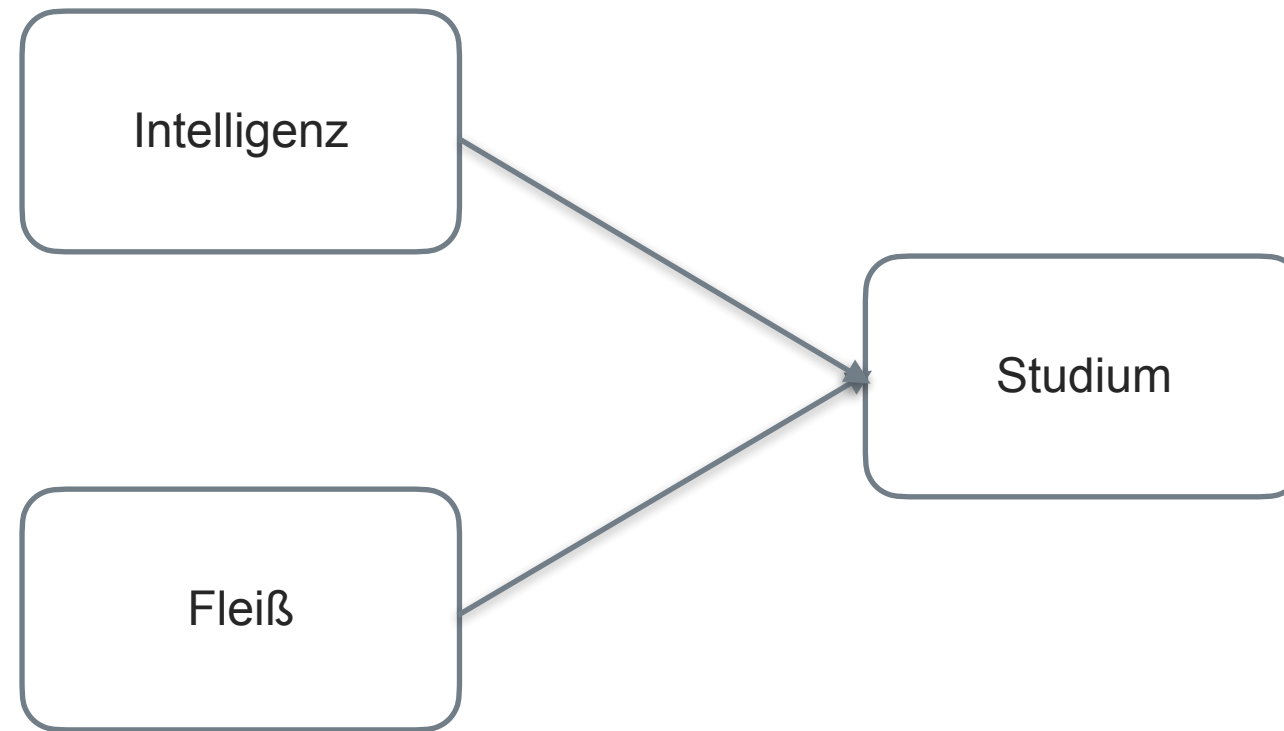
Subgruppen (Studie 2) zeigen negativen Zusammenhang



Woher kommt der (scheinbare) negative Zusammenhang?

- Nehmen wir an, Studenti kann nur werden, wer entweder intelligent oder (auch) fleißig ist
- Intelligenz und Fleiß sind unabhängig, hat man also eine intelligente Person (aus der Allgemeinbevölkerung) weiß man nichts über ihren Fleiß
- Weiß man aber, dass die Person ei Studenti ist, ändert unser Wissen: Ist dis Studenti faul, wissen wir sofort, dass dis Studenti intelligent ist.
- Fazit: Zwei unabhängige Variablen werden abhängig, wenn man auf eine dritte Variable bedingt, die die gemeinsame Wirkung (Studium) der beiden anderen Variablen ist

Scheinkorrelation von Intelligenz und Fleiß: Kollision

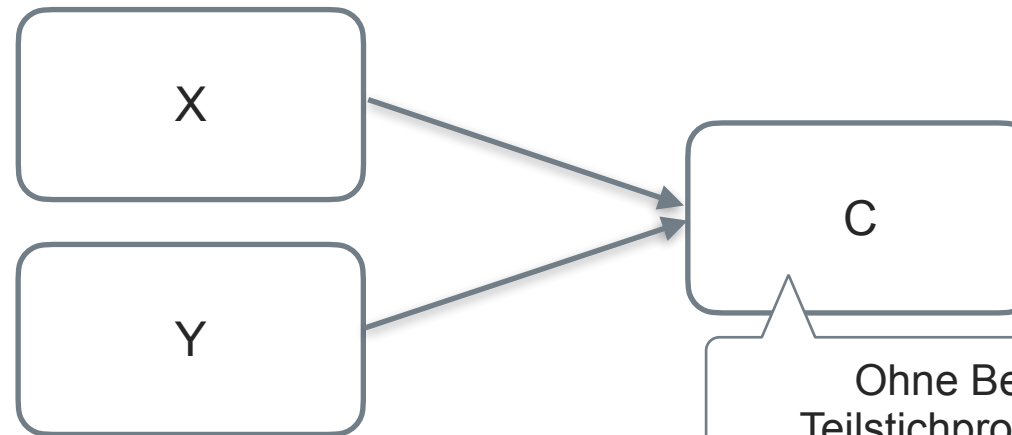


Warum entsteht die Kollisions-Verzerrung?

- $Z = X + Y$, wobei X und Y unabhängig sind
- Wenn ich Ihnen sage, $X = 3$, lernen Sie nichts über Y , da die beiden Variablen unabhängig sind
- Aber: Wenn ich Ihnen zuerst sage, $Z = 10$, und dann sage, $X = 3$, wissen Sie sofort, was Y ist ($Y = 7$).
- Also: X und Y sind abhängig – *gegeben* Z .

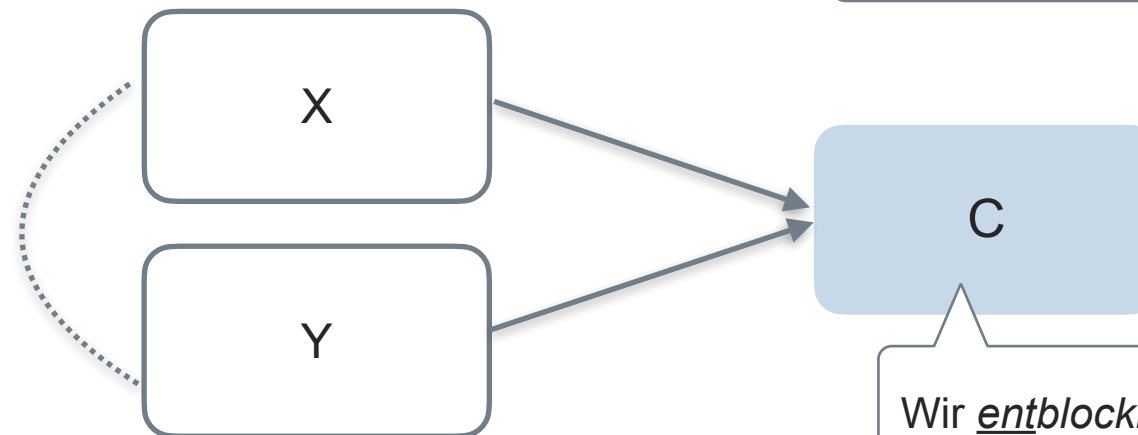
Merkmale einer Verzerrung durch Kollision

Kausalstruktur



Ohne Beschränkung auf eine Teilstichprobe ist der Pfad X - C - Y blockiert

Verzerrung durch Kollision tritt auf



Wir entblockieren den Pfad X - C - Y durch die Beschränkung auf eine Subgruppe

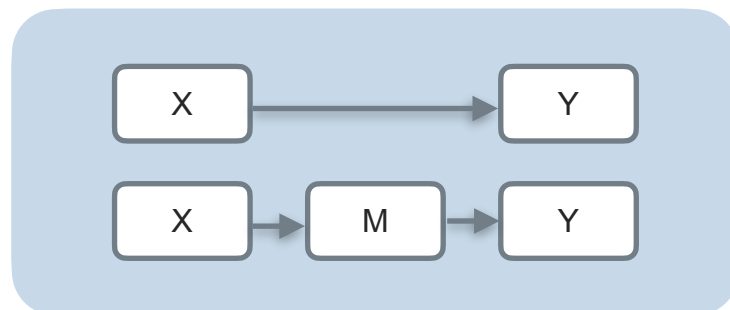
Atome der Kausalität

Drei Arten von Kausal-Elementen

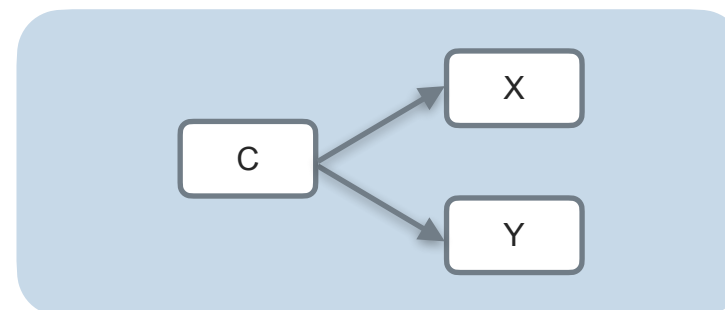
Nur
Kausalzusammenhänge

Achtung,
Scheinzusammenhänge

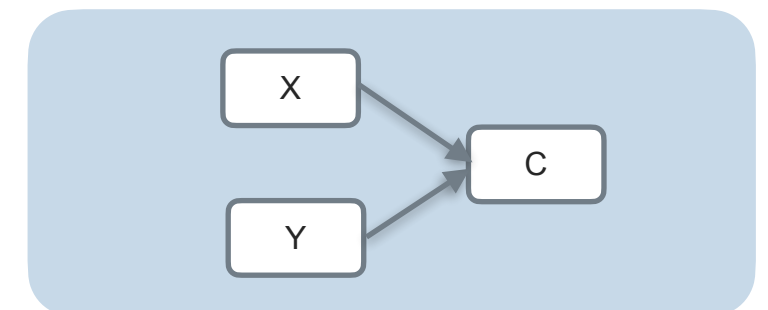
Mediation



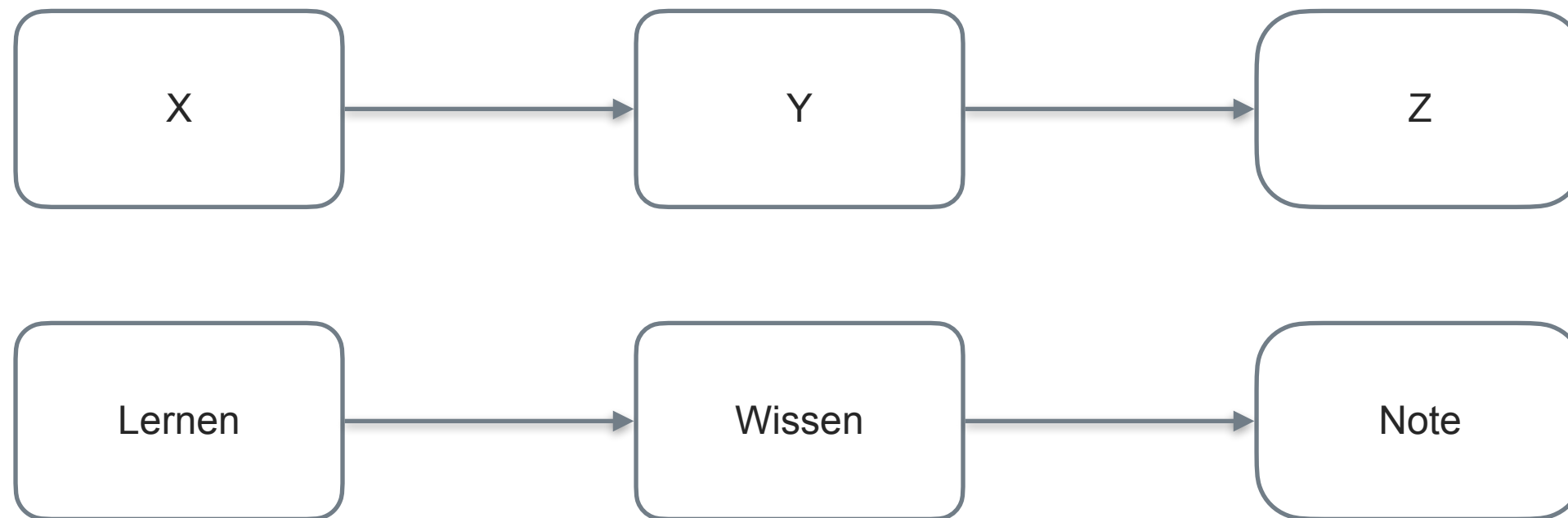
Konfundierung



Auswahlverzerrung



Mediatoren sind Wirkketten

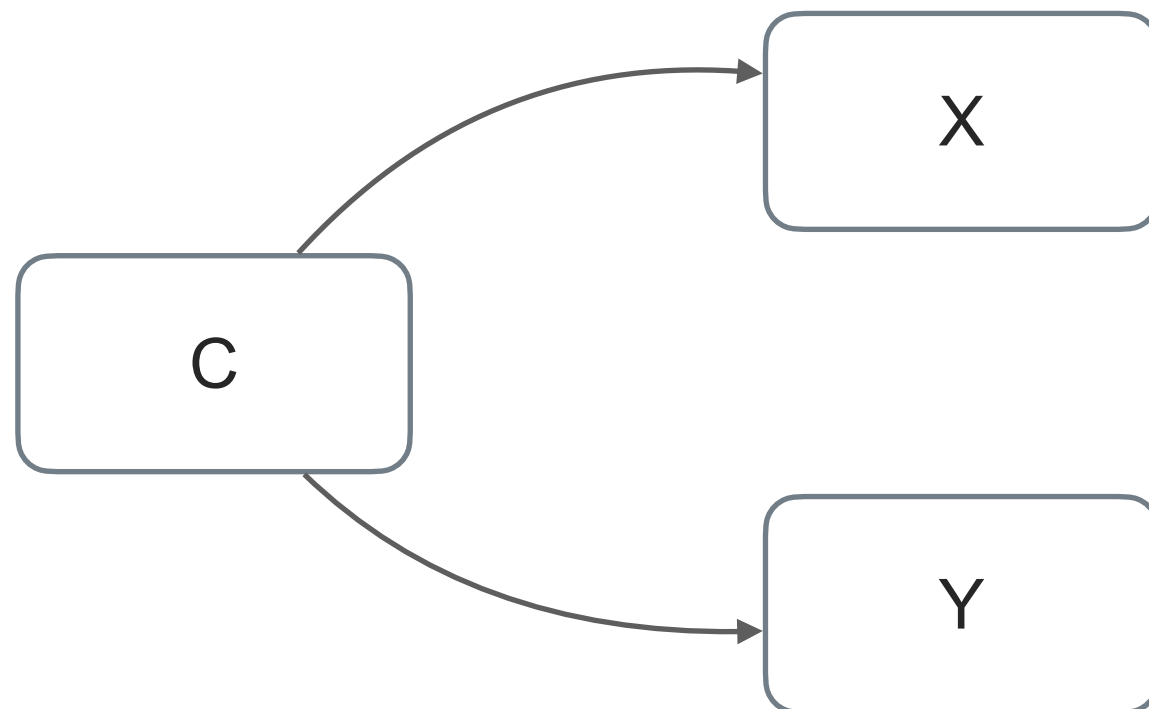


- Z und Y sind stochastisch abhängig $Z \not\perp Y$
- Y und X sind (stochastisch) abhängig $Y \not\perp X$
- Z und X sind abhängig $Z \not\perp X$
- Z und X sind unabhängig, gegeben Y $Z \perp X | Y$

Nennen Sie Beispiele von Wirkketten

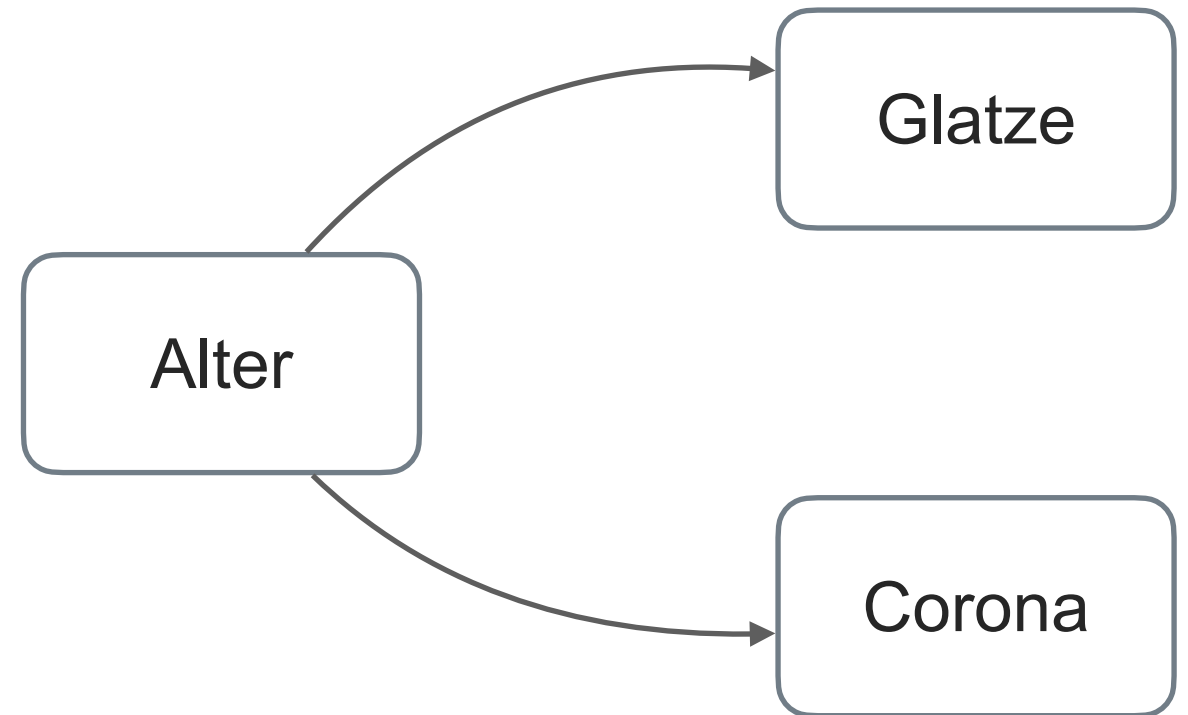
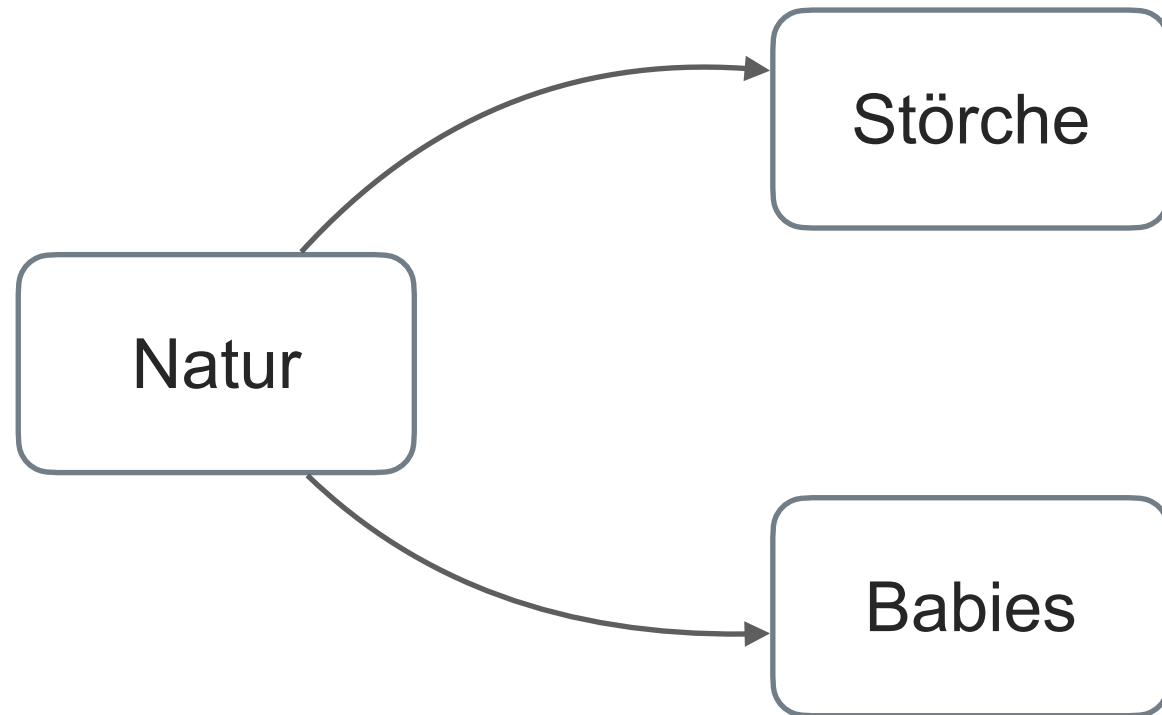


Ein Konfundierer (eine „Gabel“) ist eine gemeinsame Ursache

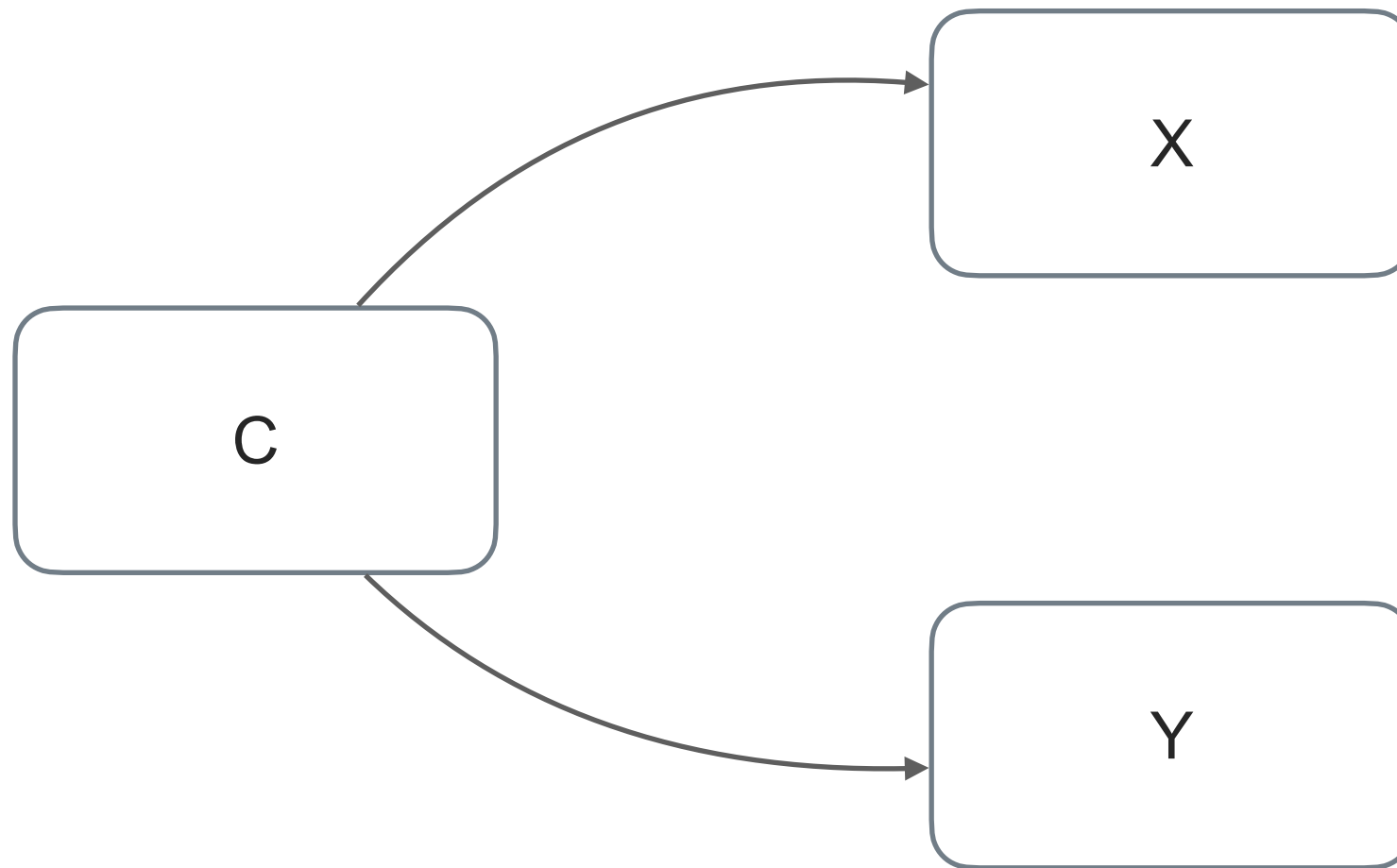


- C und X sind abhängig
- C und Y sind abhängig
- X und Y sind abhängig
- X und Y sind unabhängig, gegeben C

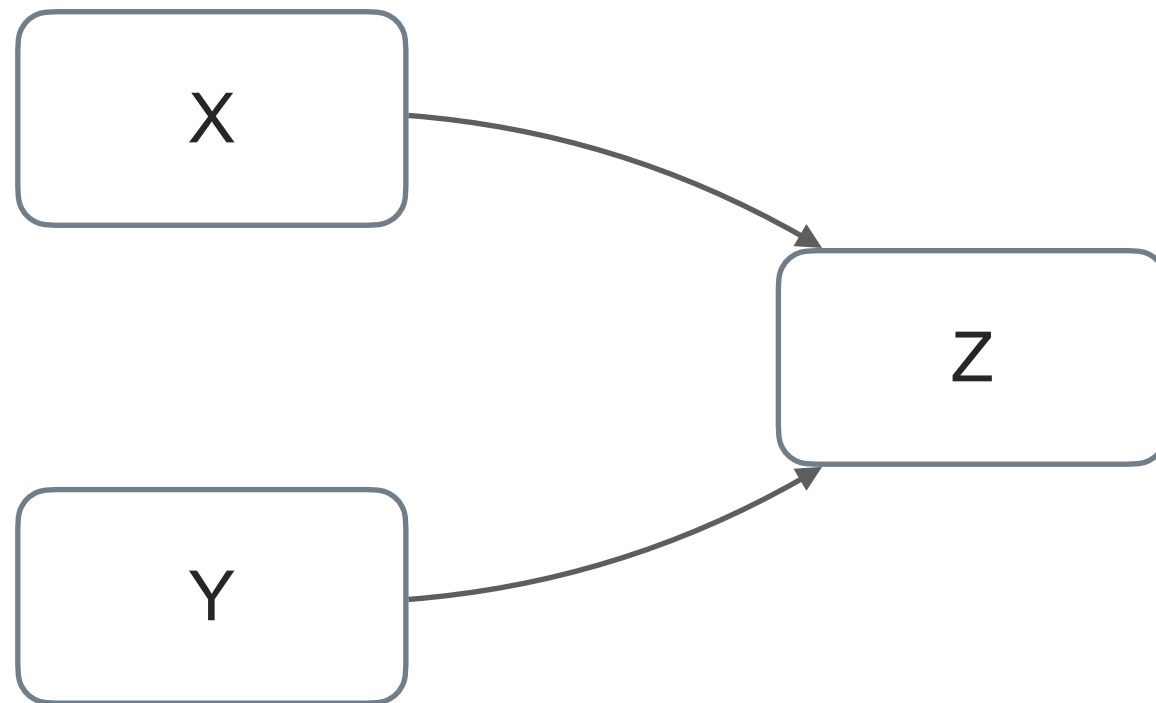
Beispiele für Konfundierungen



Nennen Sie Beispiele für Konfundierungen



Kollisionen sind gemeinsame Wirkungen



- X und Z sind abhängig
- Y und Z sind abhängig
- X und Y sind unabhängig
- X und Y sind abhängig, gegeben Z

Wann man eine Variable
nicht kontrollieren sollte

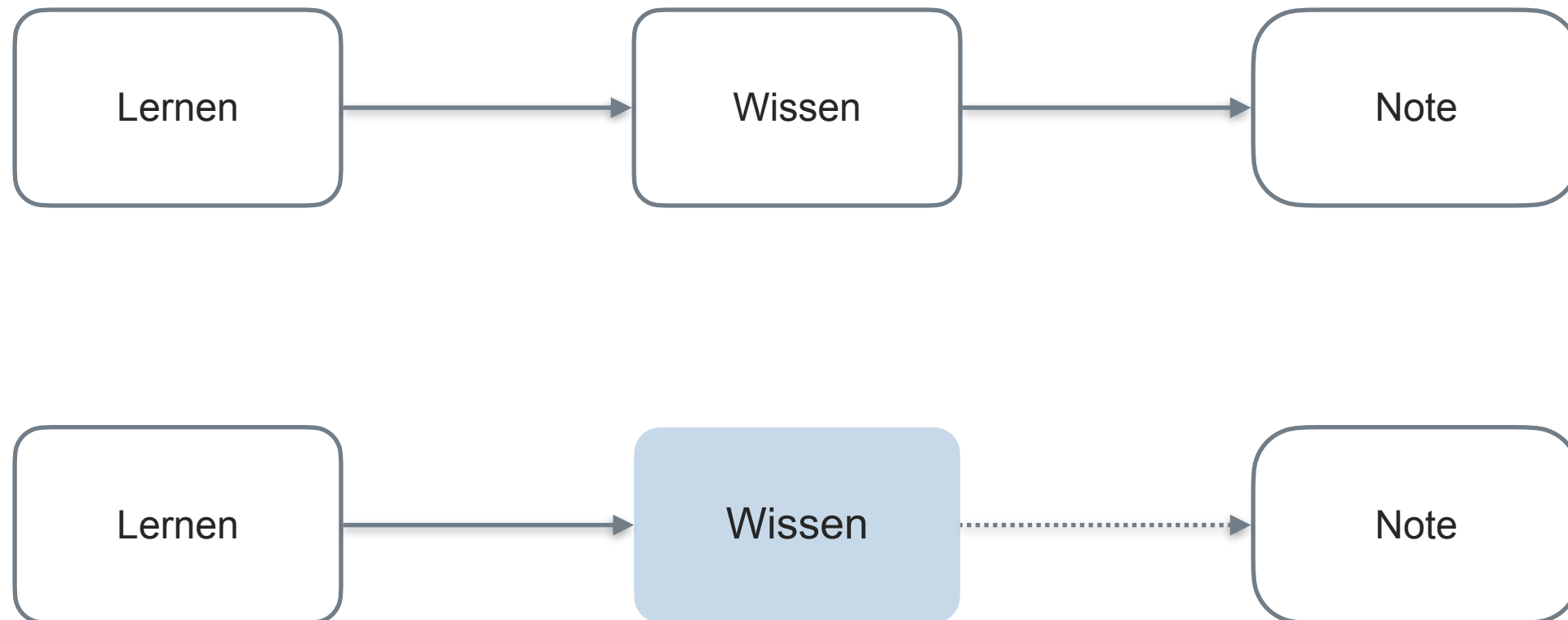
Guter Einstigsartikel

- Lesen Sie den Abschnitt „Controlling for mediators: removing the association of interest“ im Artikel von Rohrer, 2019!

<https://journals.sagepub.com/doi/full/10.1177/2515245917745629>

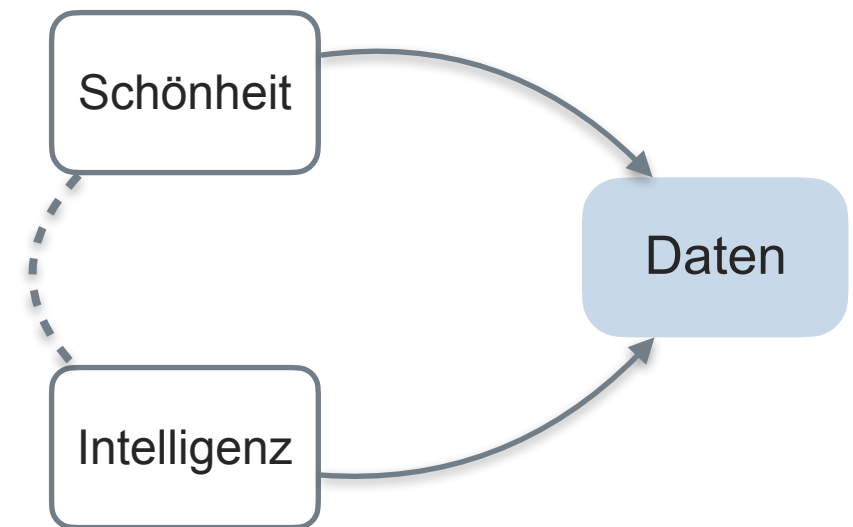
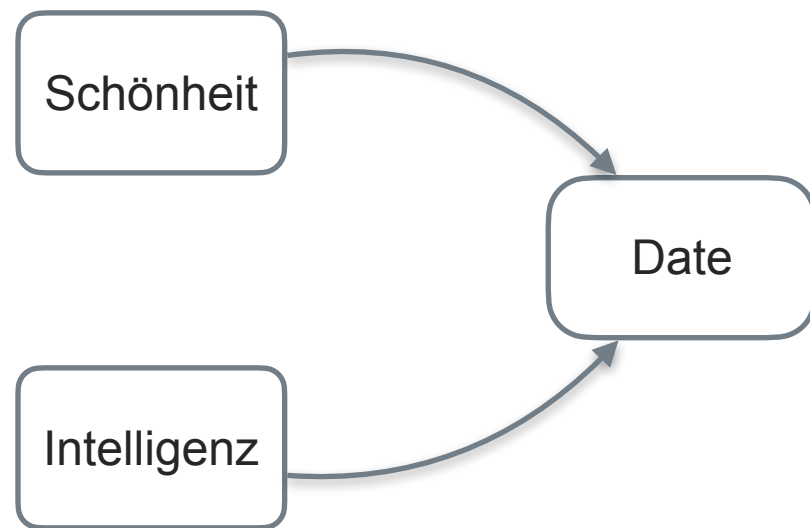
(Als PDF oder HTML frei verfügbar.)

Du sollst keine Mediatoren kontrollieren



Kontrolle von Mediatoren blockt den Fluss der Assoziation.

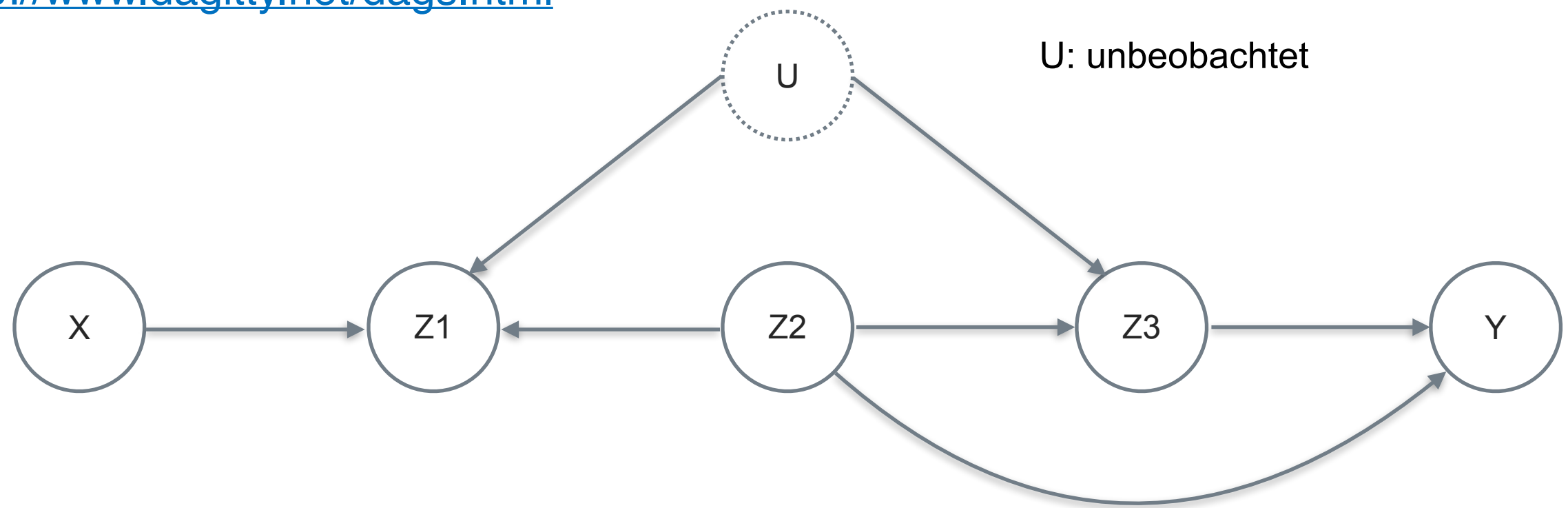
Du sollst keine Kollisionsvariablen kontrollieren



Zeit zu üben

Zeichnen Sie diesen DAG in [dagitty.net](http://www.dagitty.net)

<http://www.dagitty.net/dags.html>



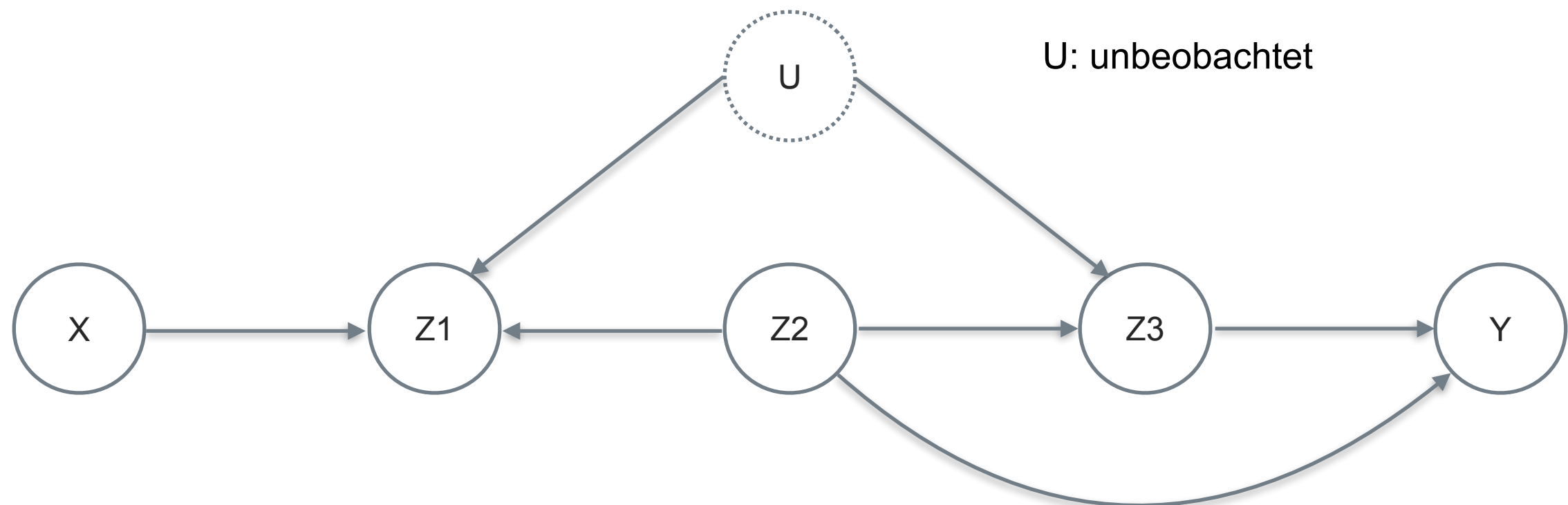
Bezeichnen Sie dabei die Variablen wie folgt:

- X: UV (exposure)
- Y: AV (outcome)
- U: unmeasured

Welche Pfade muss man blockieren?

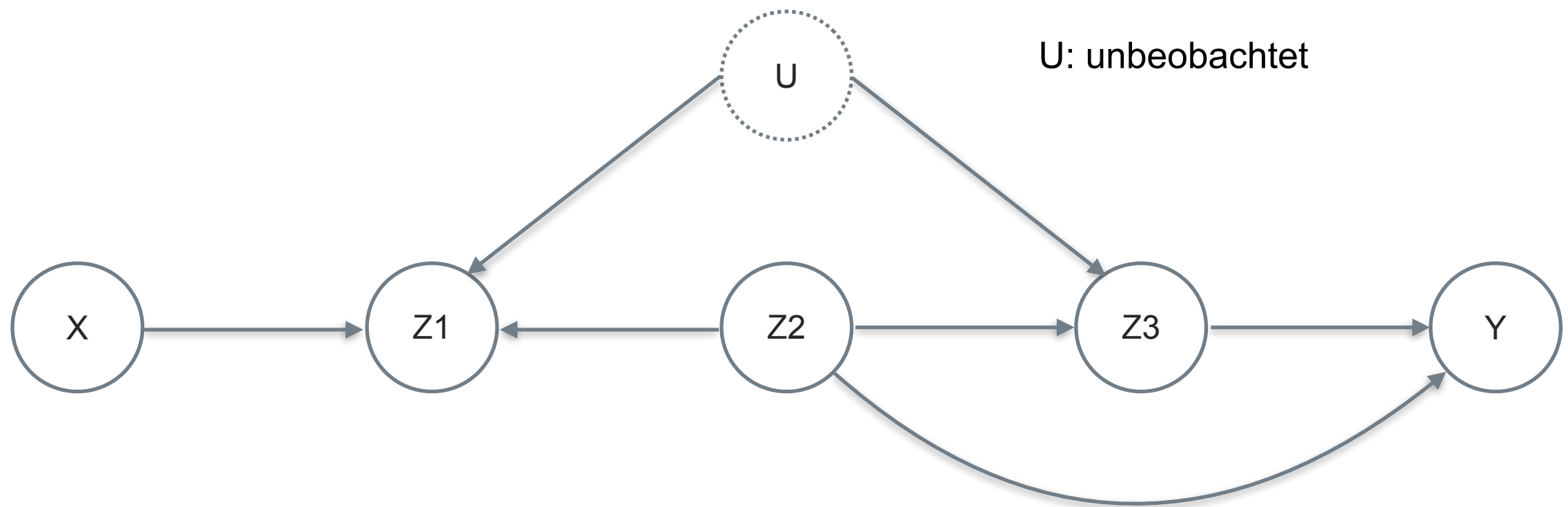
Welche Pfade muss man blockieren, um den kausalen Effekt von X auf Y zu bemessen?

(S. Elwert, 2013, S. 253).



Welche Pfade muss man blockieren?

LÖSUNG



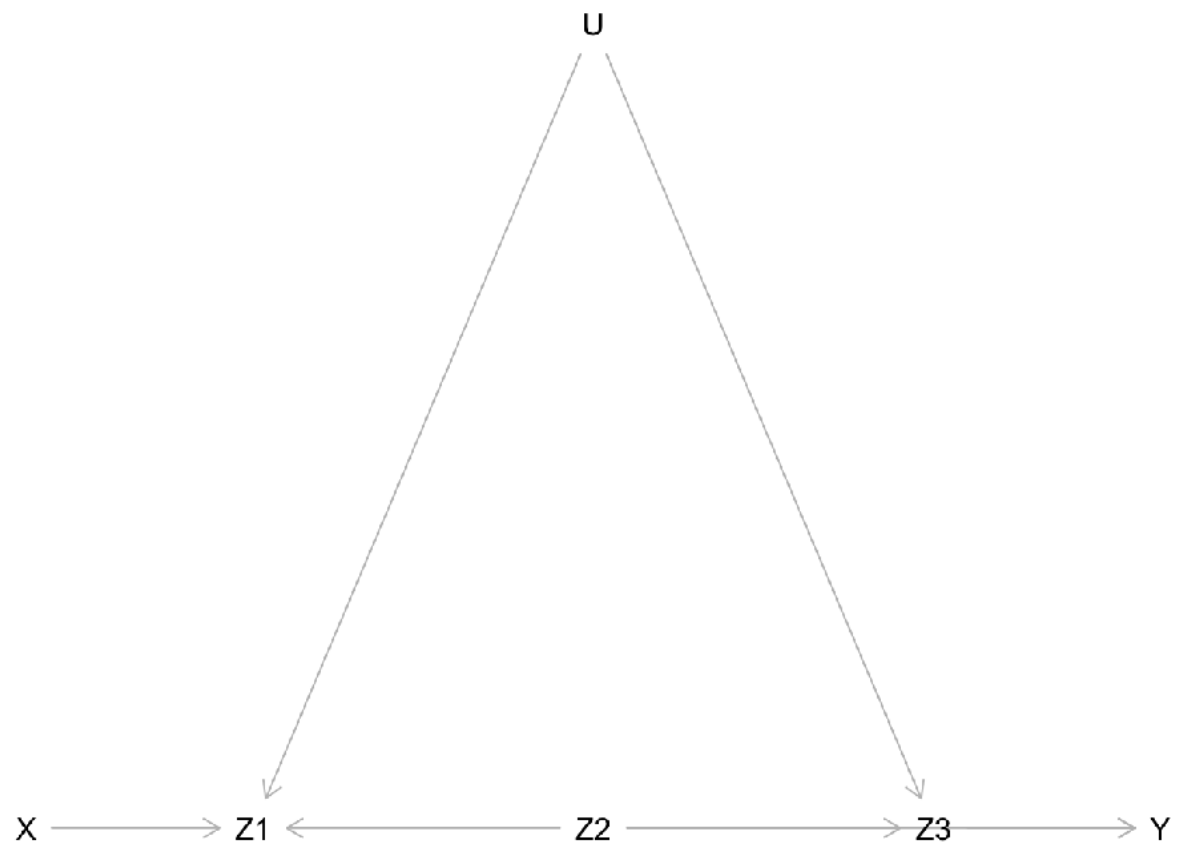
- X und Y sind unabhängig, da alle Pfade die Kollision bei Z1 enthalten. Sofern *nicht* auf eine *Kollisionsvariable* bedingt wird, ist der Pfad *blockiert*.
- Bedingt man auf Z2 oder Z3, ändert sich nichts, da sich kein Pfad zwischen X und Y öffnet.
- Bedingt man (nur) auf Z1, so öffnen sich drei Pfade von X zu Y: $X - Z1 - Z2 - Z3 - Y$, $X - Z1 - U - Z3 - Y$, $X - Z1 - Z2 - Y$, so dass X und Y statistisch abhängig werden.
- Bedient man zusätzlich zu Z1 auf Z3, so schließen sich die Pfade $X - Z1 - Z2 - Z3 - Y$ und $X - Z1 - U - Z3 - Y$, aber ein Pfad bleibt offen: $X - Z1 - Z2 - Y$. Außerdem öffnet sich ein weiterer Pfad: $X - Z1 - U - Z3 - Z2 - Y$. Durch Bedingen auf Z2 würden sich die Pfade wieder schließen.

Erstellen Sie den DAG in R!

```
library(dagitty)
```

```
dag1 <- dagitty('
dag {
  U [latent, pos = "2.5, 0"]
  X [exposure, pos = "0, 1"]
  Y [outcome, pos = "5, 1"]
  Z1 [pos = "1, 1"]
  Z2 [pos = "2.5, 1"]
  Z3 [pos = "4, 1"]
  U -> Z1
  U -> Z3
  X -> Z1
  Z2 -> Y
  Z2 -> Z1
  Z2 -> Z3
  Z3 -> Y
}' )
```

```
plot(dag1)
```



Die Wissenschaft liebt
Kausalität (zu sehr)

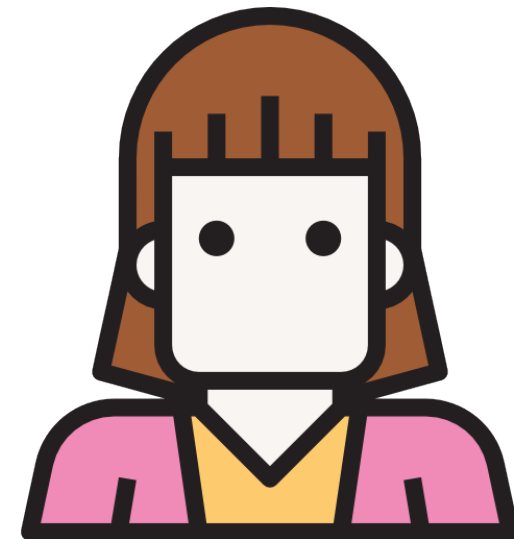
Es gibt drei Arten von wissenschaftlichen Studien, sagt Wolfi

Das Gute, das
Böse und das
Hässliche



Wolfi

Wolfi, reiß dich
bitte zusammen!



Angi

Es gibt drei Arten von wissenschaftlichen Studien,

Deskription

„Welche
Konsumertypen
gibt es?“

Vorhersage

„Sagen
Facebook-Likes
die Persönlichkeit
vorher?“

Kausalität

„Erhöht
Achtsamkeits-
training die
Konzentration?“

Wissenschaft ist meist an kausalen Fragen interessiert

Deskription

Vorhersage

Kausalität

Kausalität ist cool.

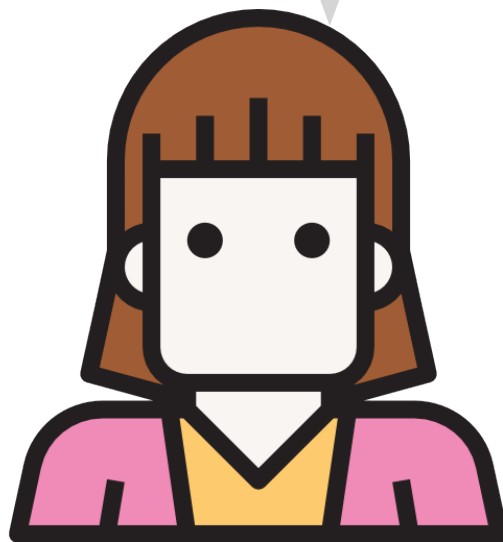


Wolfi

Journal of Applied Psychology mag auch Kausalforschung

Ausgabe 4 in 2020

10 von 12 Studien
der Ausgabe
verwenden Kausal-
Sprache.



Angi

Title	quotes (abstract)	causal language?
The generation and function of moral emotions in teams: An integrative review.	„influence on individual team members' moral emotions“	yes
On melting pots and salad bowls: A meta-analysis of the effects of identity-blind and identity-conscious diversity ideologies.	„improve intergroup relations“ „the effects of identity-blind ideologies“	yes
Political affiliation and employment screening decisions: The role of similarity and identification processes.	„to examine the effects of“	yes
A dynamic account of self-efficacy in entrepreneurship.	„self-efficacy energizes action because“	yes
Coworker support and its relationship to allostasis during a workday: A diary study on trajectories of heart rate variability during work.	„We examined the effect of“	yes
A theoretical assessment of dismissal rates and unit performance, with empirical evidence.	"utility analysis suggests that increasing dismissal rates can improve performance“	yes
Motivation to lead: A meta-analysis and distal-proximal model of motivation and leadership.	„the three MTL types partially explained the relationship“	no
Putting leaders in a bad mood: The affective costs of helping followers with personal problems.	„how such helping acts may impact leaders“ „leaders with high (vs. low) managerial experience were less affected by“	yes
When goals are known: The effects of audience relative status on goal commitment and performance.	„investigating how the perceived relative status of a goal audience influences goal commitment“	yes
Selecting response anchors with equal intervals for summated rating scales.		no
It hurts me too! (or not?): Exploring the negative implications for abusive bosses.	„we propose that perpetrated abuse impacts these supervisor outcomes“	yes
How can employers benefit most from developmental job experiences? The needs-supplies fit perspective.	„developmental job experiences (DJE) lead to positive work-related outcomes“	yes

Fazit

Take-home Message der Take-home Messages



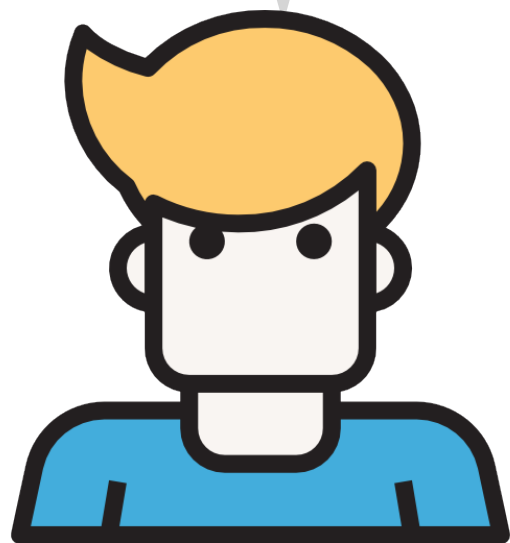
Wolfi

Glaub keiner Beobachtungsstudie.

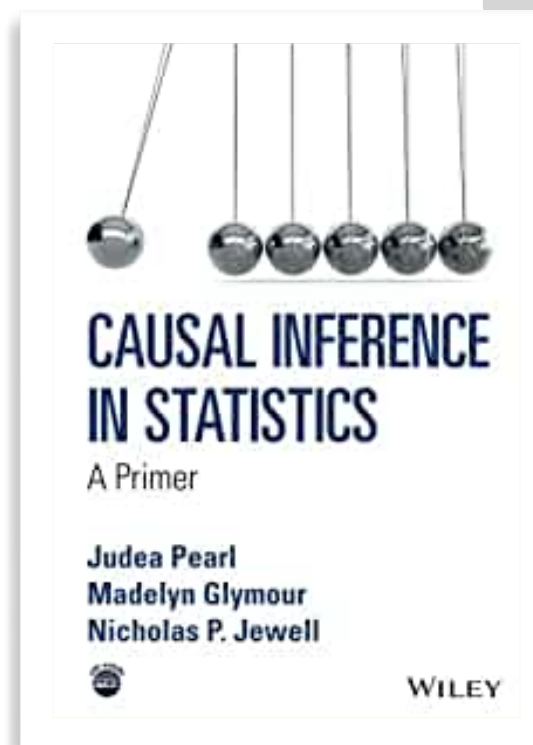
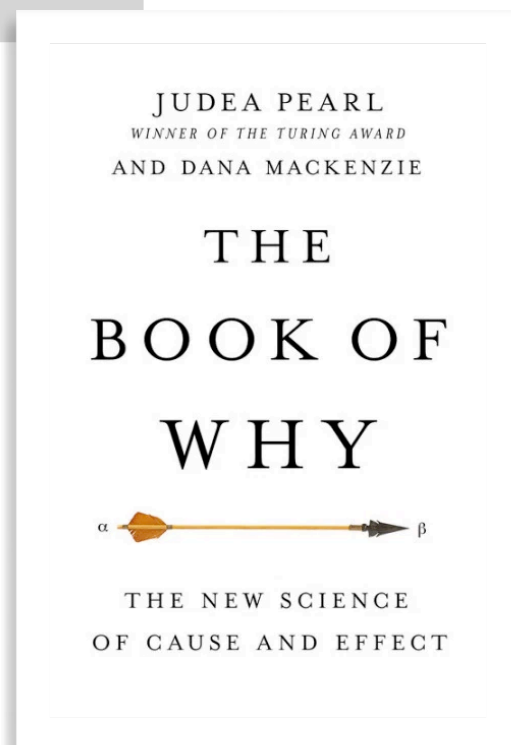
Es sei denn, sie präsentiert ein gutes
Kausalmmodell.

Kausalinferenz ist ein Juwel für die Statistik

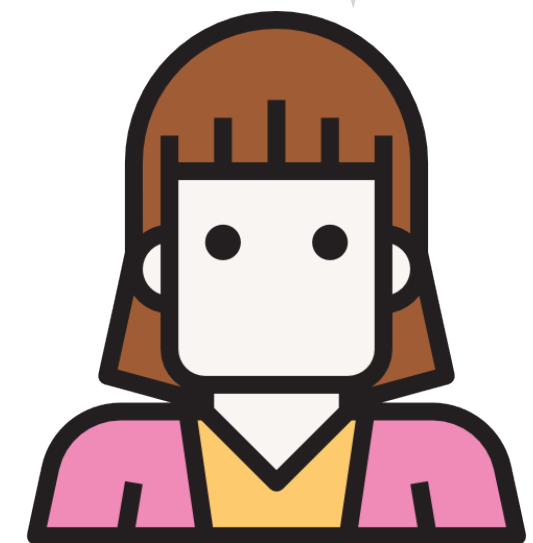
Das erzählt doch bloß ein verrückter Prof!



Don



Hm, aber Judea Pearl hat dafür den Turing Award erhalten.



Angi

Einstiegsliteratur



- Elwert, F. (2013). Graphical causal models. In S. Morgan (Ed.), *Handbook of causal analysis for social research* (pp. 245–273). Springer. https://www.researchgate.net/publication/278717528_Graphical_Causal_Models
- Lübke, K., Gehrke, M., Horst, J., & Szepannek, G. (2020). Why We Should Teach Causal Inference: Examples in Linear Regression with Simulated Data. *Journal of Statistics Education*, 1–17. <https://doi.org/10.1080/10691898.2020.1752859>
- Rohrer, J. M. (2018). Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data. *Advances in Methods and Practices in Psychological Science*, 1(1), 27–42. <https://doi.org/10.1177/2515245917745629>

- Corvetti, C. (2006). Saratoga Houses. <https://rdr.io/cran/mosaicData/>
- Dablander, F. (2020). An Introduction to Causal Inference [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/b3fkx>
- Dedering, U. (2010). Map of the USA [Map]. https://en.wikipedia.org/wiki/Saratoga_Springs,_New_York#/media/File:Usa_edcp_relief_location_map.png
- Elwert, F. (2013). Graphical causal models. In S. Morgan (Ed.), Handbook of causal analysis for social research (pp. 245–273). Springer. https://www.researchgate.net/publication/278717528_Graphical_Causal_Models
- Hernán, M. A., Hsu, J., & Healy, B. (2019). A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks. *Chance*, 32(1), 42–49. <https://doi.org/10.1080/09332480.2019.1579578>
- item2101. (2020). Avatar Icon Pack [Icon]. www.flaticon.com. <https://www.flaticon.com/packs/avatar-14?k=1587995971688>
- Lübke, K. (2020, February). Introduction to Causal Inference. Dozententage der FOM, Essen.
- Lübke, K., Gehrke, M., Horst, J., & Szepannek, G. (2020). Why We Should Teach Causal Inference: Examples in Linear Regression with Simulated Data. *Journal of Statistics Education*, 1–17. <https://doi.org/10.1080/10691898.2020.1752859>
- Pearl, J. (2009). *Causality*. Cambridge university press.
- Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal inference in statistics: A primer*. Wiley.
- Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect* (First edition). Basic Books.
- Rohrer, J. M. (2018). Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data. *Advances in Methods and Practices in Psychological Science*, 1(1), 27–42. <https://doi.org/10.1177/2515245917745629>
- Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, 25(3), 289–310. <https://doi.org/10.1214/10-STS330>
- UpstateNYer. (2009). Saratoga County, New York, USA,. https://en.wikipedia.org/wiki/Saratoga_Springs,_New_York#/media/File:Downtown_Saratoga_Springs.jpg