

SMI606: Week 5

Causality

Dr. Calum Webb

Sheffield Methods Institute, the University of Sheffield.

c.j.webb@sheffield.ac.uk



Sign In



Learning Objectives

What will I learn? How does this week fit into my course?

By the end of this week you will:

- Be able to describe what we mean by causality and how we can use statistical methods to explore probabilistic causation.
- Be able to explain three bases upon which causal inference can be made: robust dependence, consequential manipulation, and generative process.
- Be able to critically assess these foundations of causal inference.
- Be introduced to some tools for developing causal models (Directed Acyclic Graphs).
- Be able to identify which types of research design and data are appropriate for making causal inference.

Learning Objectives

What will I learn?

How does this week fit into my course?

- It is increasingly important for new social scientists to have a good working knowledge of causality and causal analysis.
- Being able to synthesize different forms of research into generative processes (e.g. DAGs) is increasingly a goal of postgraduate and postdoctoral research projects.
- This week's content will provide some of the fundamental building blocks for more advanced methods (e.g. Structural Equation Modelling, Longitudinal Modelling, A-B testing) you may need to learn in your career or may wish to cover in the rest of your studies.

What do we mean by causality?



'Causality'

Using the work and definitions developed by sociologist John H. Goldthorpe (2001).

Deterministic Causality

- In philosophy and theology, the idea of causality as a deterministic force: Given enough information about past events, every outcome can, in theory, be explained or predicted.

Hofer, 2003. *Causal Determinism*. Stanford Encyclopedia of Philosophy.

Probabilistic Causality

- Causes raise the probability of their effects, some of which may be screened off or superseded by other causal relations.

Hitchcock, 2018. *Probabilistic Causation*. Stanford Encyclopedia of Philosophy

[S]uppose that Billy and Suzy each throw a rock at a bottle, and that each has a certain probability of hitting and breaking it. As it happens, *Suzy's rock hits the bottle, and Billy's doesn't. As things actually happened, we would say that Suzy's throw caused the bottle to shatter, while Billy's didn't.* Nonetheless, Billy's throw increased the probability that the bottle would shatter ... Billy's throw had a tendency to shatter the bottle; it was a potential cause of the bottle shattering; it was the sort of thing that generally causes shattering; but it did not actually cause the bottle to shatter.

Hitchcock, 2018. Probabilistic Causation. Stanford Encyclopedia of Philosophy



[S]uppose that Billy and Suzy each throw a rock at a bottle, and that each has a certain probability of hitting and breaking it. As it happens, Suzy's rock hits the bottle, and Billy's doesn't. As things actually happened, we would say that Suzy's throw caused the bottle to shatter, while Billy's didn't. Nonetheless, *Billy's throw increased the probability that the bottle would shatter ... Billy's throw had a tendency to shatter the bottle; it was a potential cause of the bottle shattering; it was the sort of thing that generally causes shattering; but it did not actually cause the bottle to shatter.*

Hitchcock, 2018. Probabilistic Causation. Stanford Encyclopedia of Philosophy



Causation as "robust dependence": temporal causality



Causation as Robust Dependence

Causality can be argued to exist when we can observe a consistent, temporal causal effect.

"A variable, X, 'Granger causes' Y if, after taking into account all information apart from values of X, these values still add to one's ability to predict future values of Y."

Goldthorpe, 2001: 2

Causation as Robust Dependence

Causality can be argued to exist when we can observe a consistent, temporal causal effect.

"A variable, X, 'Granger causes' Y if, after taking into account all information apart from values of X, these values still add to one's ability to predict future values of Y."

Goldthorpe, 2001: 2

- Does knowing the value of X at a preceding time point improve our accuracy when predicting Y, after accounting for other competing explanations?

Causation as Robust Dependence

Causality can be argued to exist when we can observe a consistent, temporal causal effect.

"A variable, X, 'Granger causes' Y if, after taking into account all information apart from values of X, these values still add to one's ability to predict future values of Y."

Goldthorpe, 2001: 2

- Does knowing the value of X at a preceding time point improve our accuracy when predicting Y, after accounting for other competing explanations?
- Measurements of X and controls (Z) **must** precede measurements of Y.

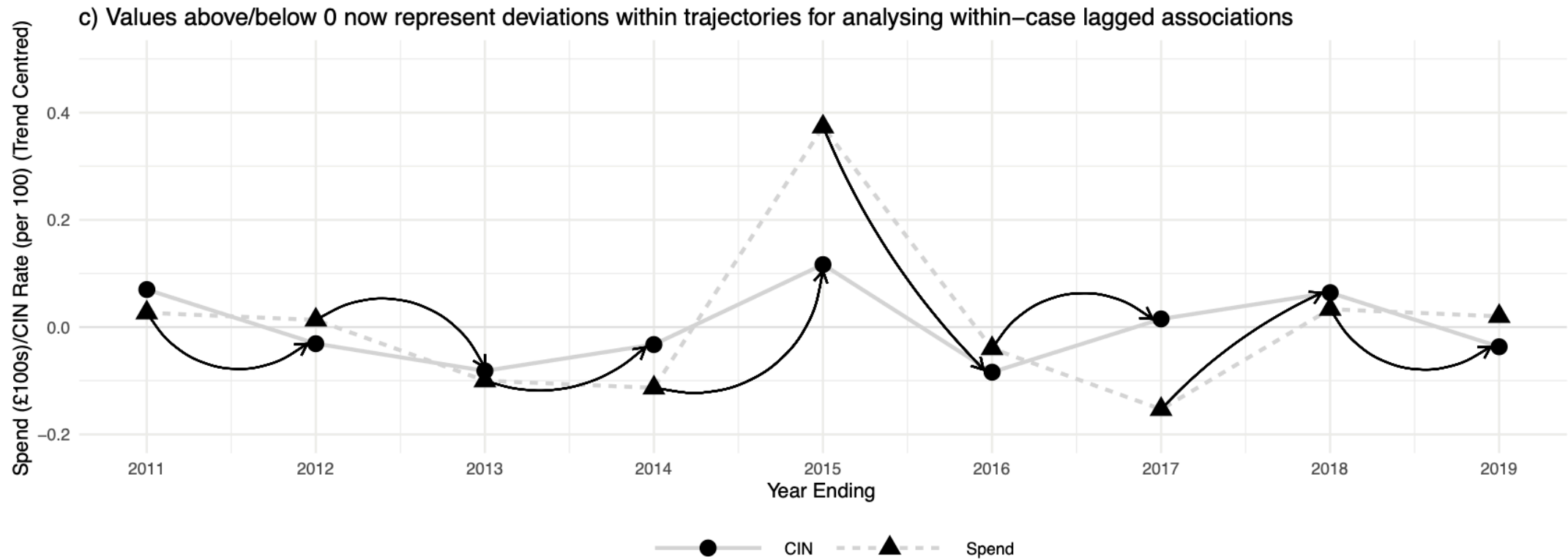
Causation as Robust Dependence

Causality can be argued to exist when we can observe a consistent, temporal causal effect.

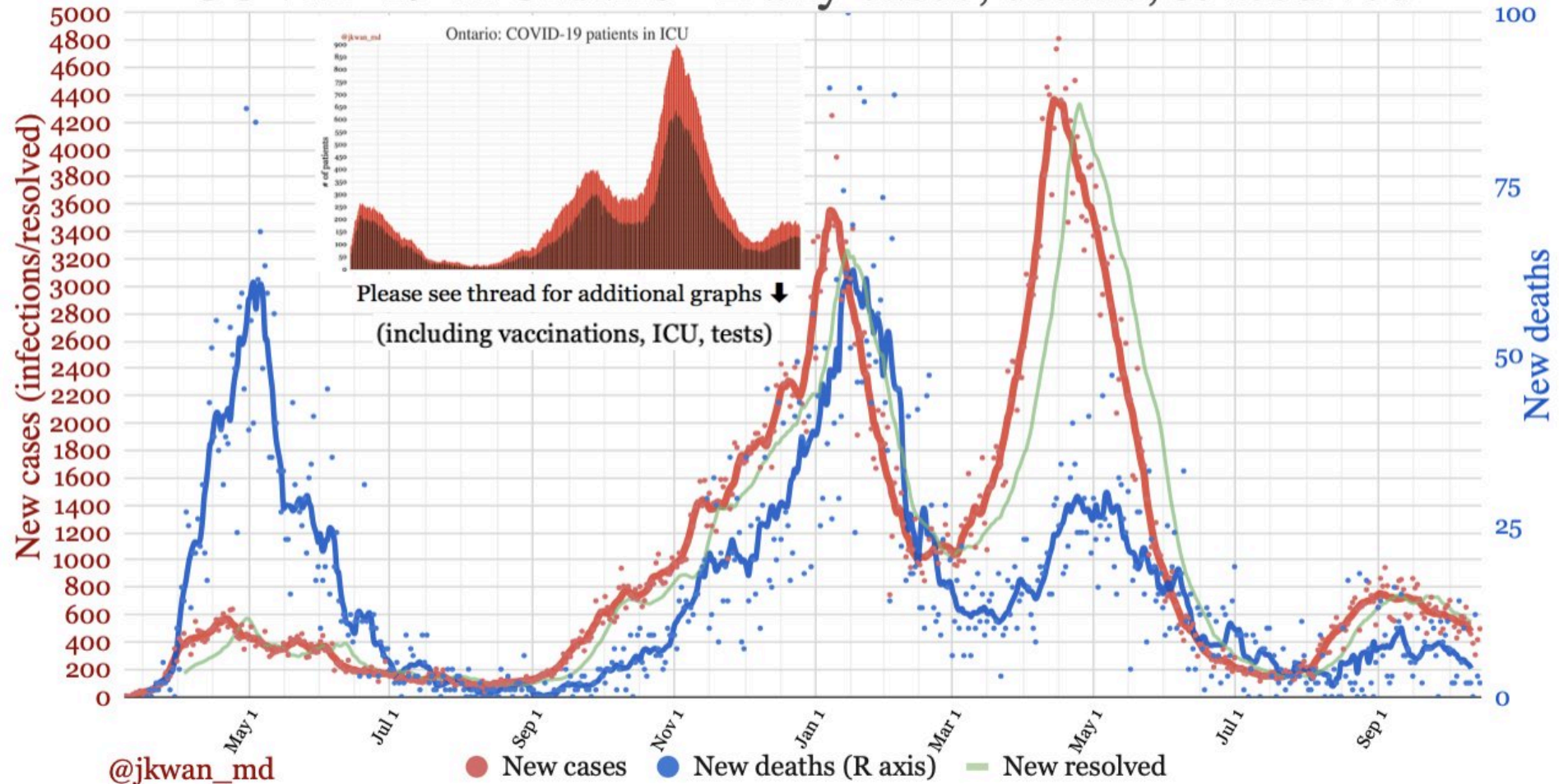
"A variable, X, 'Granger causes' Y if, after taking into account all information apart from values of X, these values still add to one's ability to predict future values of Y."

Goldthorpe, 2001: 2

- Does knowing the value of X at a preceding time point improve our accuracy when predicting Y, after accounting for other competing explanations?
- Measurements of X and controls (Z) **must** precede measurements of Y.
- Can be **deepened through 'Lazarfeldian elaboraton'**
 - Identify a 'signal' and 'response'
 - Iteratively introduce possible confounding factors and see if the pattern between 'signal' and 'response' stays the same (or continues to exist)



COVID-19 in Ontario - Daily cases, deaths, & resolved



Causation as Robust Dependence

Weaknesses of Granger causality

- **Easy for this approach to end up being atheoretical**, with causality claimed purely on the basis of statistical models.
 - "[S]ociologists have strongly criticized the supposition that statistical techniques can in themselves provide adequate causal explanations of social phenomena"

Causation as Robust Dependence

Weaknesses of Granger causality

- **Easy for this approach to end up being atheoretical**, with causality claimed purely on the basis of statistical models.
 - "[S]ociologists have strongly criticized the supposition that statistical techniques can in themselves provide adequate causal explanations of social phenomena"
- Is it possible to **fully incorporate all of the possible confounding variables**, or to adequately measure some causes.
 - "[A]vailable sociological theory may just not be strong enough to help produce models that can be treated as genuinely 'structural' — i.e. so parameterized that their co-efficients are sufficiently invariant and autonomous to sustain claims about the consequences of changes in the variables deemed to be 'exogenous'."

Causation as "consequential manipulation": experiments



Causation as Consequential Manipulation

Causation as some environmental factor that, when manipulated (or left alone), can be demonstrated to cause some change in an outcome variable of interest.

"Here, attention centres specifically on 'the consequences of performing particular acts' or, in other words, on establishing causation through experimental methods."

Goldthorpe, 2001: 4

- If I manipulate this thing, does the result change (probabilistically thinking)? If so, the thing I manipulated can be said to have caused a change in something else.

Causation as Consequential Manipulation

Causation as some environmental factor that, when manipulated (or left alone), can be demonstrated to cause some change in an outcome variable of interest.

"Here, attention centres specifically on 'the consequences of performing particular acts' or, in other words, on establishing causation through experimental methods."

Goldthorpe, 2001: 4

- If I manipulate this thing, does the result change (probabilistically thinking)? If so, the thing I manipulated can be said to have caused a change in something else.
- Something **must** be manipulated or manipulatable, and exposure to the change or difference must be either randomly chosen *or* be unrelated to the variables of interest (as in a random experiment).

Causation as Consequential Manipulation

Exercise

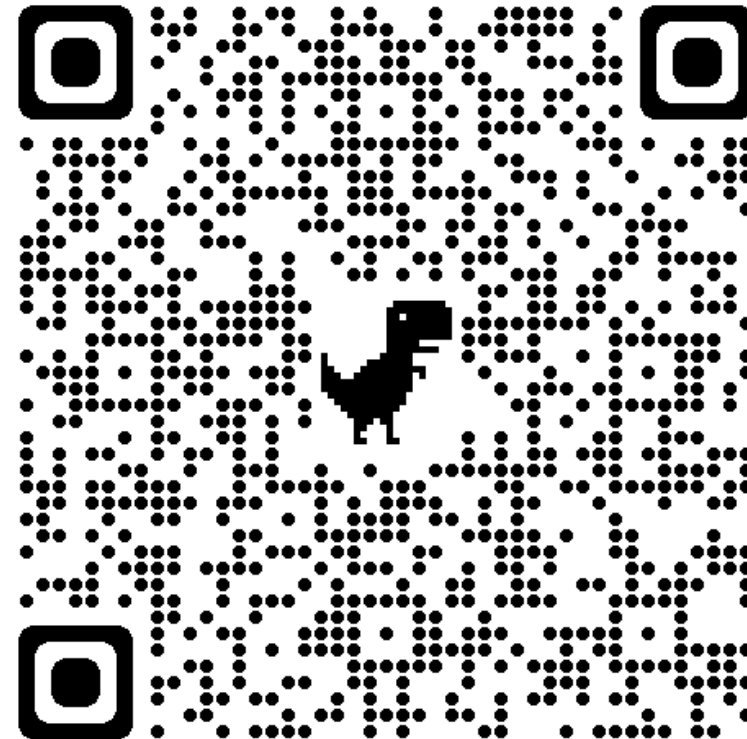
At the beginning of this session, chocolate buttons were randomly allocated behind the computer lab computers.

If you got chocolate buttons, you are the treatment group. If you did not get chocolate buttons, you are the control group.

- H_0 : Being given chocolate buttons will have no effect on students' enjoyment of learning quantitative methods.

Please now fill out the following form:

<https://forms.gle/sigw2jprMohiaBZM8>



Causation as Consequential Manipulation

How feasible are these kind of studies in your field? What are the potential weaknesses (having just been subjected to one)?

Wooclap Link

Event code: **AXXIZJ**

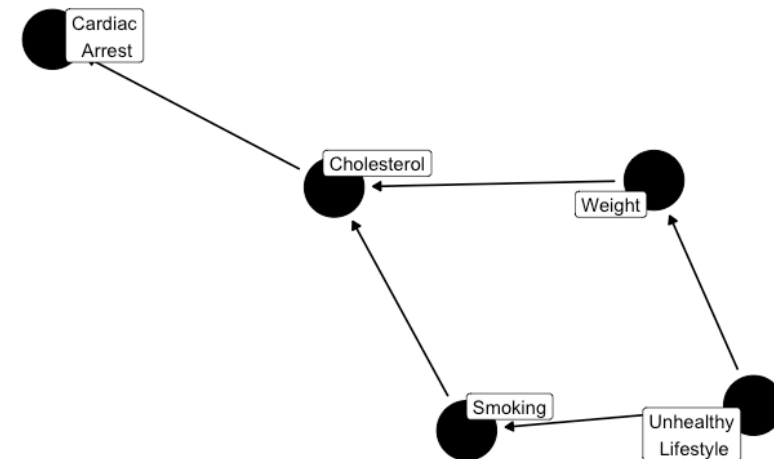


Causation as "generative process": DAGs and the New Science of Causality



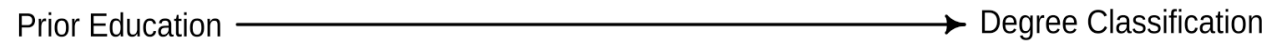
Causation as Generative Process

- **Start with a theoretical model** of a generative causal process based on a variety of evidence (qual + quants).
- **Develop the processes in this model**, ideally down to the 'methodological individual' — individual actions, characteristics, and events (if appropriate).
- Design hypotheses and research studies to **test the causal associations between each of the processes**. Some may be "stronger" causal evidence, others may be weaker by necessity.
- Aim for **falsification** of the processes and **adaptation** of the theoretical model, not "once-and-for-all" **validation**.



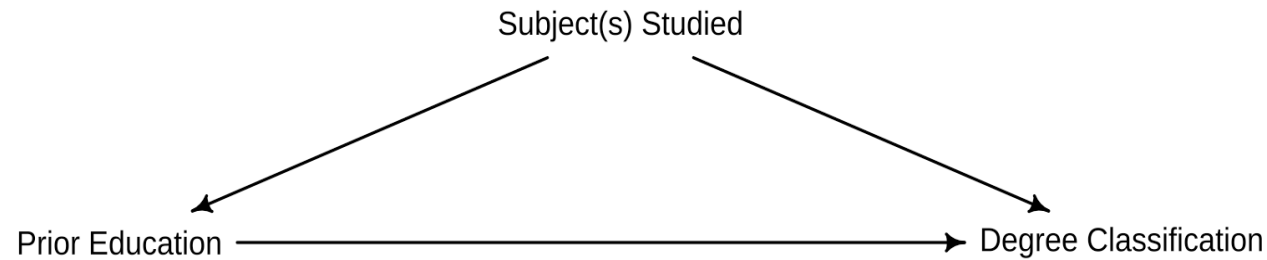
Directed Acyclic Graph example by [Malcolm Barrett](#)

DAGS: The Total Effect (Estimand)



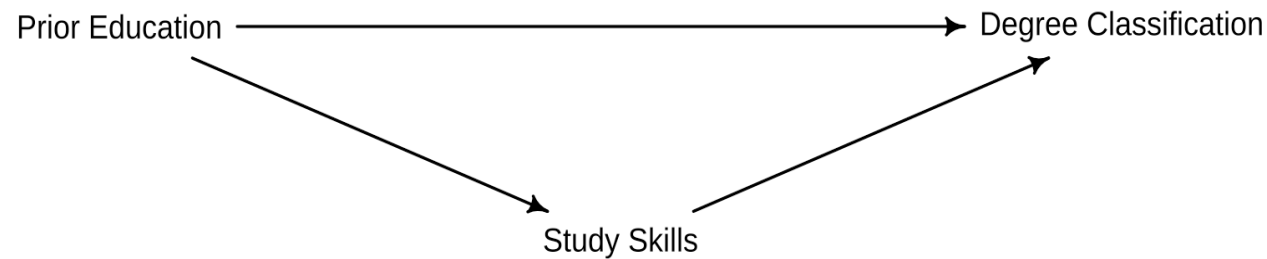
We start with the thing we want to estimate (our estimand), in this case the strength of the association between a person's secondary school outcomes and their university degree classification.

DAGS: Confounders and the Direct Effect



We know that there will be certain confounders in our data: things that influence both the predictor/cause (Prior Education) and the outcome/effect (Degree Classification). We need to control for these things to get an unbiased estimate of the effect. For example, here, we might want to only include people who studied the same subjects.

DAGS: Pipes, Direct and Indirect Effects



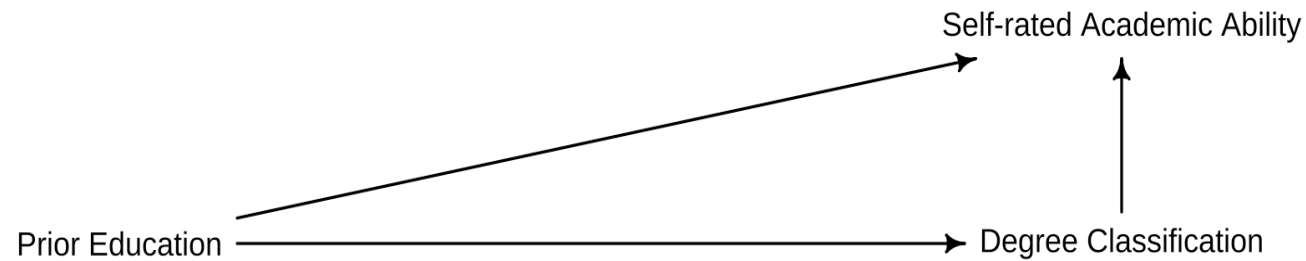
"Pipes" occur when the effects of the 'cause' operates through some mechanism(s), for example, the effect of prior education on degree outcome is partially mediated through prior education teaching more effective study skills. You may or may not want to condition effects based on mediators.

DAGS: "Blocked" Pipes: Post-treatment Bias



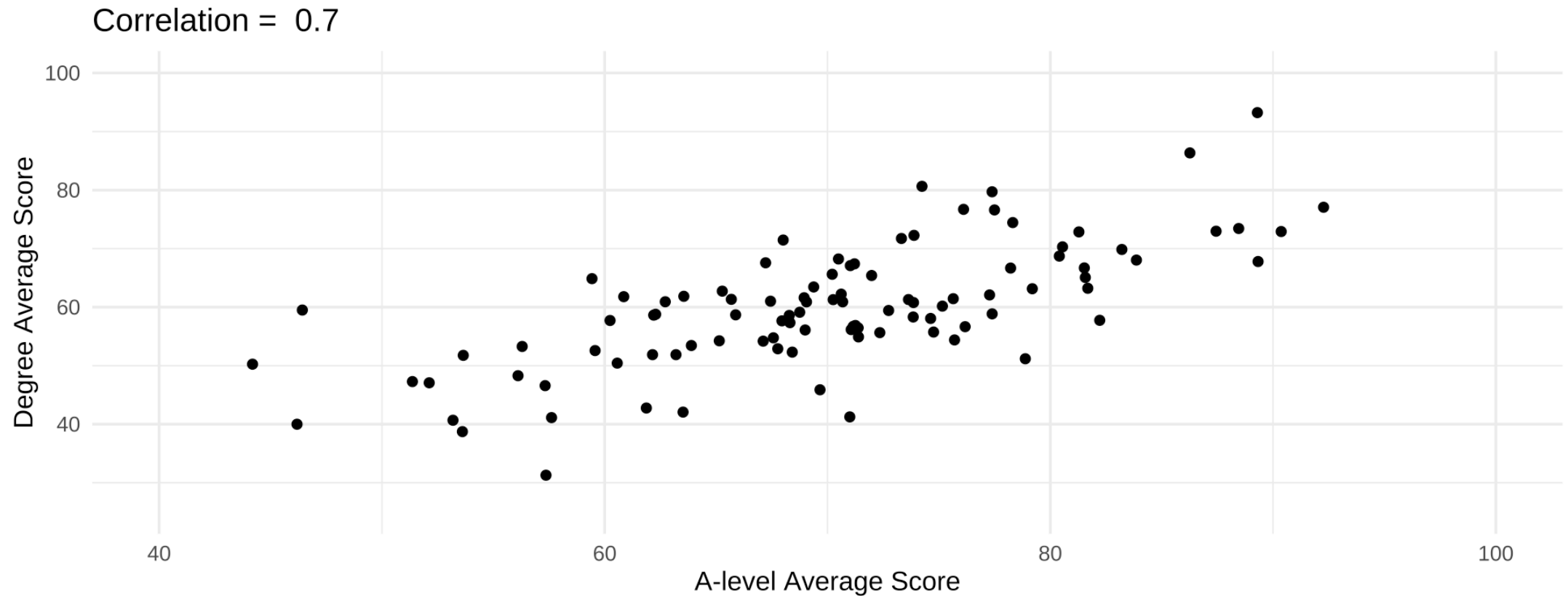
We can also imagine some of these pipes being 'blockers' rather than mediators — test scores will obviously dictate degree classification and may fully mediate prior education, so conditioning on test scores as well as degree classification would incorrectly negate any effect of prior education.

DAGS: Collider Bias

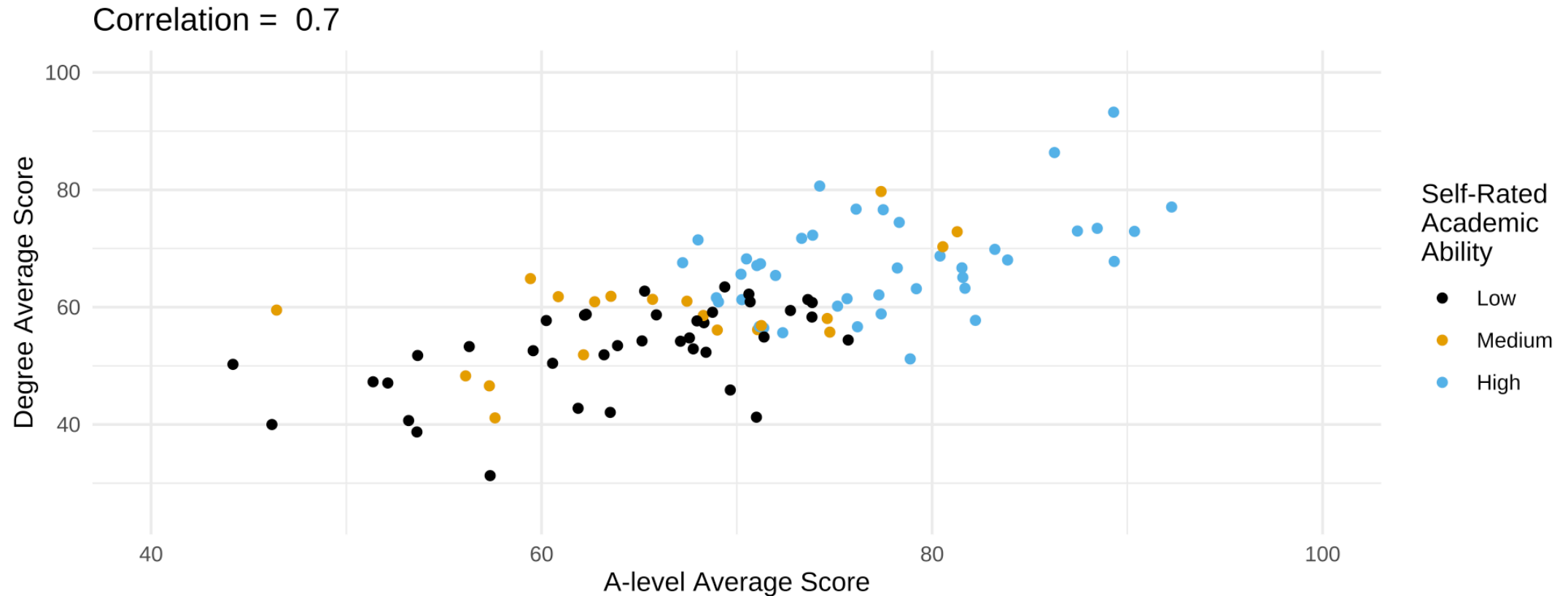


Collider bias is an often overlooked source of bias that occurs where the relationship between a cause and an outcome is conditioned on something that is a common consequence of both; e.g. here past educational performance will likely impact a person's self-rated academic ability.

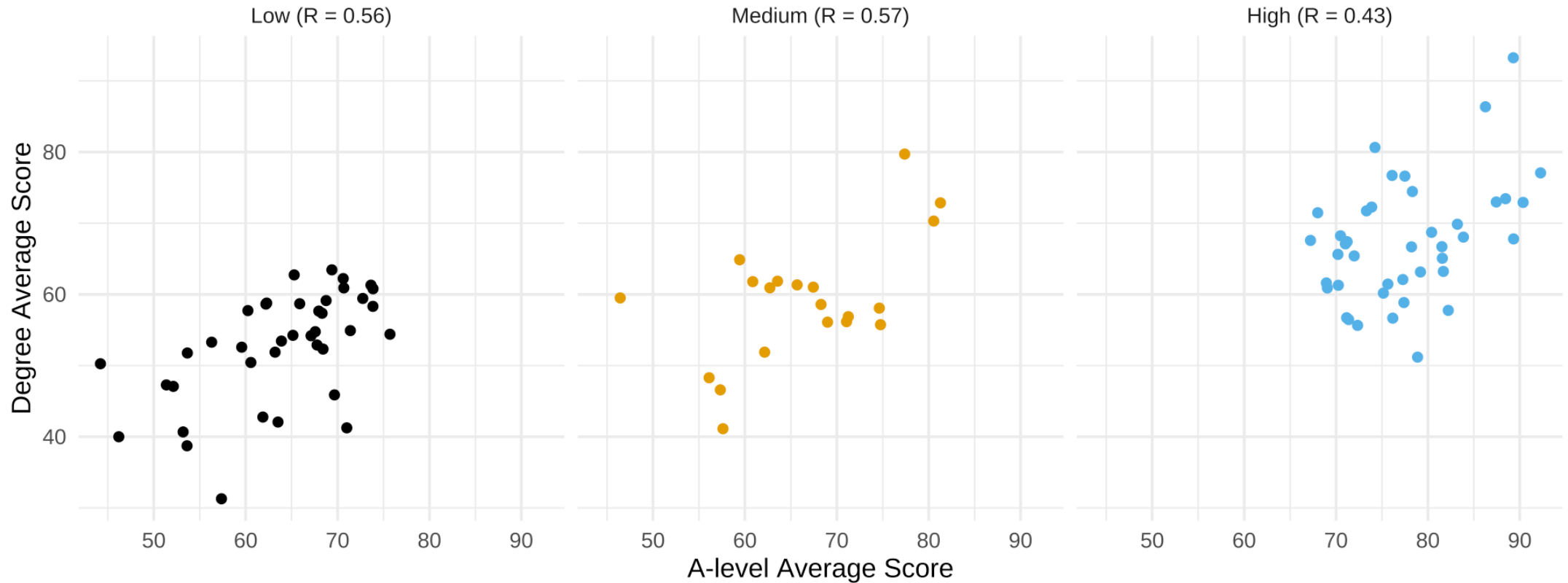
Collider Bias



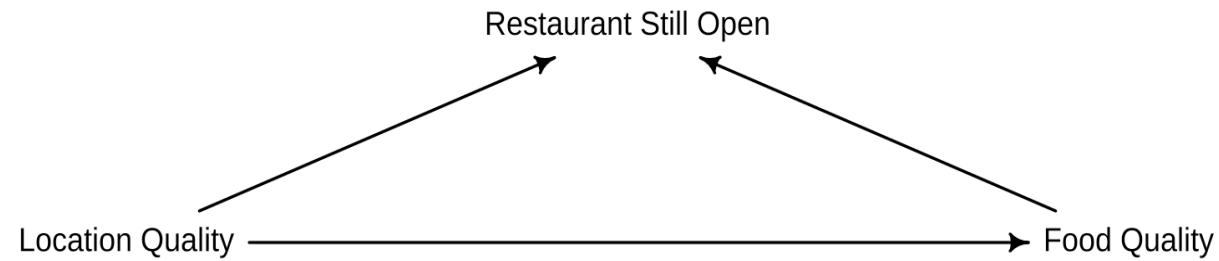
Collider Bias



Collider Bias



Collider Bias

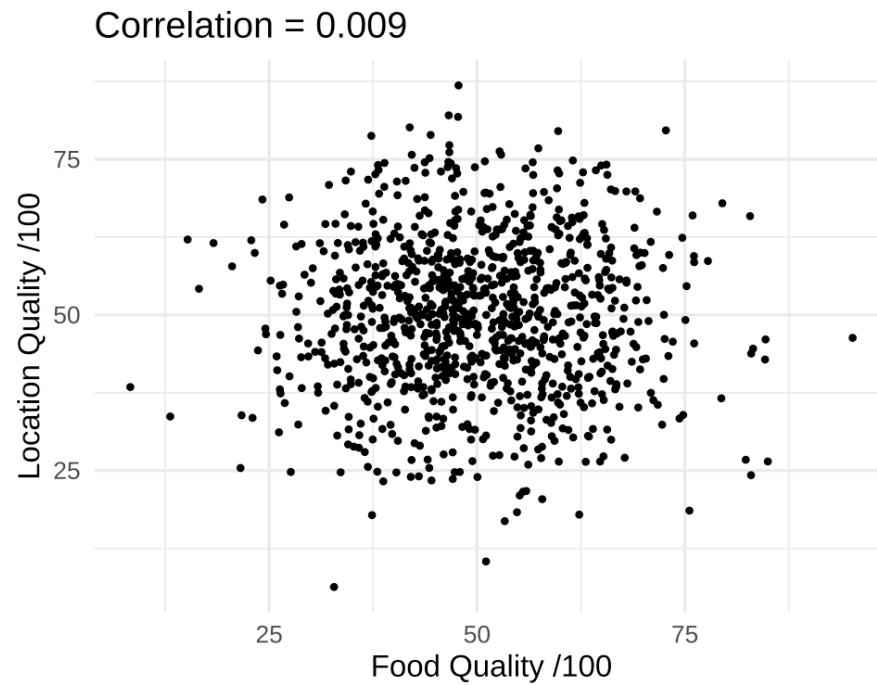


What can happen when we have a collider that determines selection?

Do restaurants in better locations have better quality food?

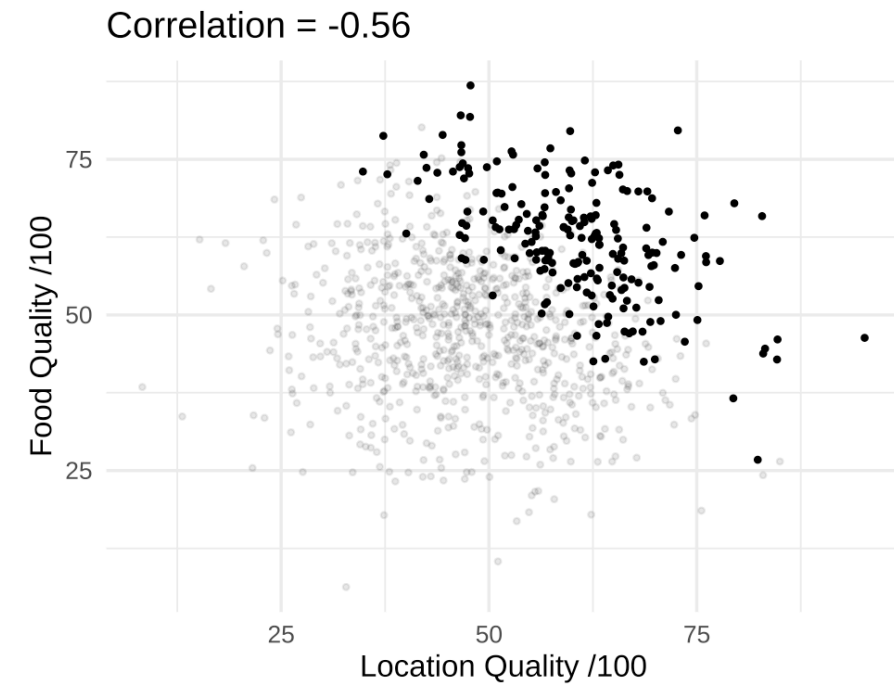
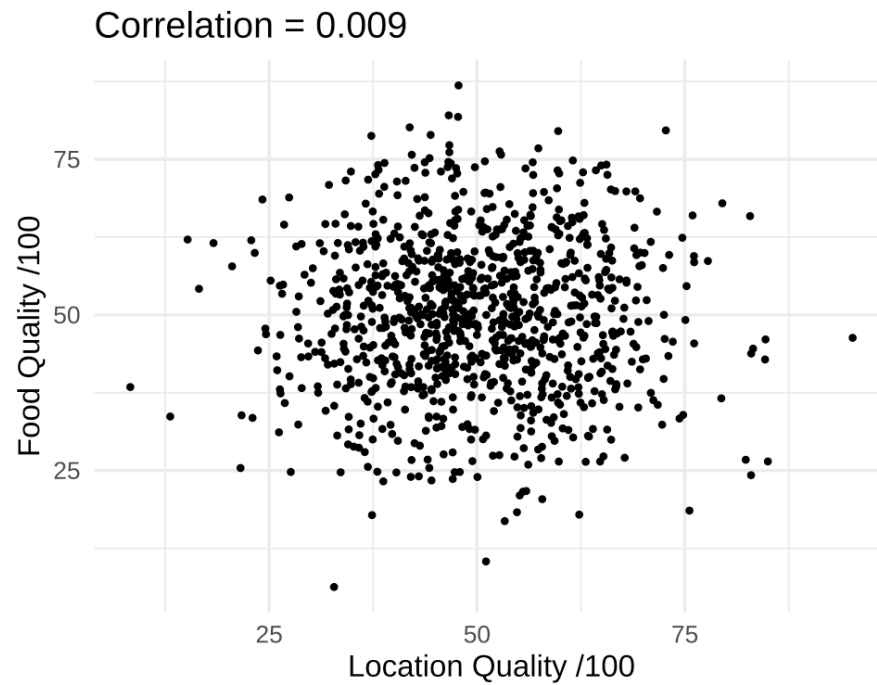


Collider Bias



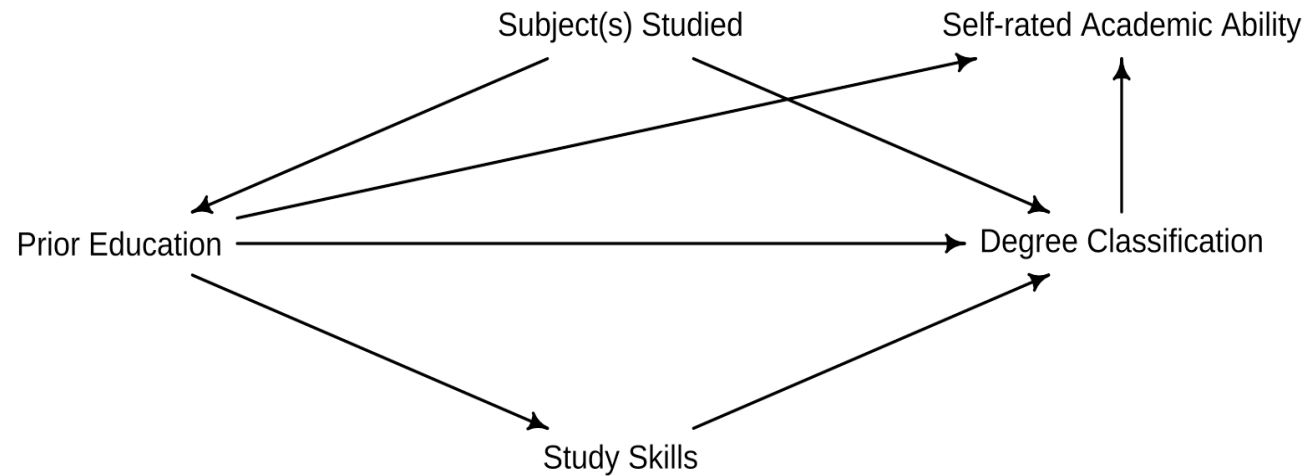
Before including whether a restaurant is open or closed we find no relationship.

Collider Bias



But when only including currently open restaurants...

DAGS: Putting it all together



You can follow the rules of 'do-calculus' to create adjustment sets to find out what kinds of confounding variables you should condition on to identify the total and/or direct effects; the model and findings can then be developed, and their causal claims challenged.

DAGS: Dagitty and Further Reading

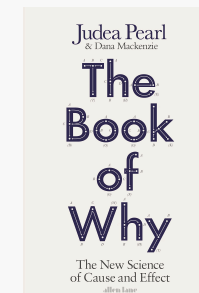
```
dagitty::adjustmentSets(dag, exposure = "Prior Education", outcome = "Degree Classification", effect = "total")
```

```
## { Subject(s) Studied }
```

```
dagitty::adjustmentSets(dag, exposure = "Prior Education", outcome = "Degree Classification", effect = "direct")
```

```
## { Study Skills, Subject(s) Studied }
```

- Peters, J, et al. "Elements of Causal Inference". [Open Access Version Available](#) from MIT Press.
- Pearl, J. et al. "Causal Inference in Statistics: A Primer", or "The Book of Why".
- McElreath, R. "Statistical Rethinking: A Bayesian course with examples in R and Stan. 2nd Edition"



Which kinds of research design support causal analyses?



Research Designs & Statistical Causal Analyses

- **Survey**
 - Cross-sectional
 - Longitudinal
- **Experiment**
 - Randomised Controlled Trials
 - Survey Experiments
 - Natural Experiments
- **Observation**
 - Administrative Data
 - Scraped/"Big" Data
 - Observational or Sensor Data

Research Designs & Statistical Causal Analyses

- Survey

- Cross-sectional
- Longitudinal

- Experiment

- Randomised Controlled Trials
- Survey Experiments
- Natural Experiments

- Observation

- Administrative Data
- Scraped/"Big" Data
- Observational or Sensor Data

Cross-sectional Survey

Face-to-face or postal questionnaire-based data collection.

- Usually a very large sample size (>200,000) and representative of the population.
- However, **usually not possible to make causal inferences from** based on either robust dependence or consequential manipulation.
- UNLESS, the cross-sectional survey either
 - Included data on something where an argument for a natural experiment could be made.
 - Included a survey experiment design in it.

Example:

- Labour Force Survey

Research Designs & Statistical Causal Analyses

- Survey

- Cross-sectional
- Longitudinal

- Experiment

- Randomised Controlled Trials
- Survey Experiments
- Natural Experiments

- Observation

- Administrative Data
- Scraped/"Big" Data
- Observational or Sensor Data

Longitudinal Survey

Face-to-face or postal questionnaire-based data collection.

- Usually a smaller but often substantial (~5000-15,000) representative sample, **but** longitudinal surveys follow the same people for multiple years.
- However, **usually possible to make causal inferences from** on the basis of **robust dependence**.
- But may also include:
 - Included data on something where an argument for a natural experiment could be made.
 - Included a survey experiment design in it.

Example:

- UK HLS/Understanding Society

Research Designs & Statistical Causal Analyses

- **Survey**

- Cross-sectional
- Longitudinal

- **Experiment**

- **Randomised Controlled Trials**
- Survey Experiments
- Natural Experiments

- **Observation**

- Administrative Data
- Scraped/"Big" Data
- Observational or Sensor Data

Randomised Controlled Trial

A trial of some treatment(s) on some outcome where participants are randomly assigned to either treatment or control groups.

- Usually a small sample, not often representative.
- Specifically designed to **allow for causal inference based on consequential manipulation**.

Example:

- Adena, Alizade & Bohner, 2019. Quality certifications for nonprofits, charitable giving, and donor's trust

Research Designs & Statistical Causal Analyses

- **Survey**

- Cross-sectional
- Longitudinal

- **Experiment**

- Randomised Controlled Trials
- **Survey Experiments**
- Natural Experiments

- **Observation**

- Administrative Data
- Scraped/"Big" Data
- Observational or Sensor Data

Survey Experiment

A trial of some treatment(s) on some outcome where participants are randomly assigned to either treatment or control groups, however, this is tested through randomising questions and/or question wording or prompts in a survey.

- Usually a small sample, not often representative but can be when incorporated into a large longitudinal or cross-sectional survey as a module.
- Specifically designed to **allow for causal inference based on consequential manipulation**.

Example:

- [Zapryanova, 2016. Eurosceptic Cues and Citizen Attitudes](#)

Research Designs & Statistical Causal Analyses

- **Survey**

- Cross-sectional
- Longitudinal

- **Experiment**

- Randomised Controlled Trials
- Survey Experiments
- **Natural Experiments**

- **Observation**

- Administrative Data
- Scraped/"Big" Data
- Observational or Sensor Data

Natural Experiment

A naturally occurring scenario where people are randomised into a suitable treatment or control group by circumstances outside of the researcher's control, and as such, can be compared to assess causal effects by consequential manipulation. Data can come from administrative records, surveys, or both.

- Usually a small sample, not often representative.
- An organic scenario that **allows for causal inference based on consequential manipulation.**

Example:

- [The Oregon Health Insurance Experiment](#)

"[i]n early 2008, Oregon opened a waiting list for its Medicaid program for low-income adults that had previously been closed to new enrollment. Approximately 90,000 people signed up for the available 10,000 openings. The state drew names from this waiting list by lottery to fill the openings."

Research Designs & Statistical Causal Analyses

- **Survey**

- Cross-sectional
- Longitudinal

- **Experiment**

- Randomised Controlled Trials
- Survey Experiments
- Natural Experiments

- **Observation**

- **Administrative Data**
- Scraped/"Big" Data
- Observational or Sensor Data

Administrative Data

Data that is primarily collected for an administrative purpose which can be re-purposed for social science research. Commonly comes from government departments, schools, and welfare, tax, or healthcare administration.

- Often an enormous sample, usually the majority of the population and often close to 100%.
- Often collected over a person's lifespan and updated, therefore **allowing for causal inference by robust dependence** (or by consequent manipulation in the case of natural experiments, e.g. Universal Credit rollout).

Example:

- Continuous Recording of Social Housing Lettings and Sales (CORE)

Research Designs & Statistical Causal Analyses

- **Survey**

- Cross-sectional
- Longitudinal

- **Experiment**

- Randomised Controlled Trials
- Survey Experiments
- Natural Experiments

- **Observation**

- Administrative Data
- **Scraped/"Big" Data**
- Observational or Sensor Data

Scraped Data

Data that is largely unstructured but publicly available on the internet or in some other form (like a pdf). This data is then 'scraped' and sorted into a dataset using its underlying web structure (e.g. its HTML code).

- Sometimes not much of a meaningful sample, e.g. unlikely to be a good sample of people, might be a good sample of interactions on a platform (e.g. Tweets on Twitter) if selected randomly.
- Often longitudinal in nature but sometimes temporality is lost if not scraped repeatedly (e.g. prices of items on Amazon, tweets that get deleted)
- **Some potential for causal inference by robust dependence or consequent manipulation.**

Example:

- Scraping Twitter data in R

Research Designs & Statistical Causal Analyses

- **Survey**

- Cross-sectional
- Longitudinal

- **Experiment**

- Randomised Controlled Trials
- Survey Experiments
- Natural Experiments

- **Observation**

- Administrative Data
- Scraped/"Big" Data
- **Observational or Sensor Data**

Observational or Sensor Data

Traditionally, observational data is collected by researchers watching and recording events that happen. Increasingly, collected by sensors (e.g. geotracking of devices, wearable medical devices).

- Not frequently used in university social science, but much data like this is held by large corporations (e.g. Google, Apple, Fitbit, Garmin).
- Often has a temporal component, so **can feasibly be used for establishing robust dependence**. Could form the basis for a natural experiment. **Researcher can manipulate environment** in some way and then try and observe cause through **consequential manipulation**, but is this ethical?

Example:

- Weekly People-Movement GPS Patterns in Spain during the pandemic
- [Pokemon Go](#) / [Strava](#) / [Garmin](#) / [Fitbit](#)

Recap: The difference between inference (inferential statistics) and causality (causal inference)



Inference

Can we generalise from our sample to the population of interest?

Is our data a sample or is it the entire population?

- If the data is of the entire population of interest inference (p-values) **are usually not meaningful** because we have nothing larger than our sample we can generalise to.

Inference

Can we generalise from our sample to the population of interest?

Is our data a sample or is it the entire population?

- If the data is of the entire population of interest inference (p-values) **are usually not meaningful** because we have nothing larger than our sample we can generalise to.

Is our sample a random sample or otherwise *representative* of our population?

- If our sample *is not* representative of our population then **we cannot safely make inferences about the whole population from it**, and should not use p-values, *unless* our data is from an experimental design where we are making inferences based on comparing control and treatment group.

If our data *is* a sample *and* that sample is representative of the wider population of interest, or a result of an experiment, we can use hypothesis testing (p-values).

Inference

Can we generalise from our sample to the population of interest?

Is our data a sample or is it the entire population?

- If the data is of the entire population of interest inference (p-values) **are usually not meaningful** because we have nothing larger than our sample we can generalise to.

Is our sample a random sample or otherwise *representative* of our population?

- If our sample *is not* representative of our population then **we cannot safely make inferences about the whole population from it**, and should not use p-values, *unless* our data is from an experimental design where we are making inferences based on comparing control and treatment group.

If our data *is* a sample *and* that sample is representative of the wider population of interest, or a result of an experiment, we can use hypothesis testing (p-values).

The answers are not in the data!

Causality

Can we say anything about whether changes in one variable *causes* changes in another?

Does our study/analysis look at how *preceding* changes in one variable are associated with *subsequent* changes in another variable? (Robust Dependence/Granger Causality)

- For example, lagged panel models, cross-lagged panel models. If no, we **cannot infer anything about causality.**

Causality

Can we say anything about whether changes in one variable *causes* changes in another?

Does our study/analysis look at how *preceding* changes in one variable are associated with *subsequent* changes in another variable? (Robust Dependence/Granger Causality)

- For example, lagged panel models, cross-lagged panel models. If no, we **cannot infer anything about causality**.

Does our study use an experimental design or leverage a natural experiment? (Consequential Manipulation)

- If it does, then we **can** make causal inferences about the **effect of the treatment on the outcome**.

Causality

Can we say anything about whether changes in one variable *causes* changes in another?

Does our study/analysis look at how *preceding* changes in one variable are associated with *subsequent* changes in another variable? (Robust Dependence/Granger Causality)

- For example, lagged panel models, cross-lagged panel models. If no, we **cannot infer anything about causality**.

Does our study use an experimental design or leverage a natural experiment? (Consequential Manipulation)

- If it does, then we **can** make causal inferences about the **effect of the treatment on the outcome**.

Does our study present a theoretically-sound causal model and condition appropriately on possible confounders/colliders and other 'backdoor paths'? (Generative Process)

- If it does, then we can **accept the evidence as causal but encourage falsification and development of our model**.

Causality

Can we say anything about whether changes in one variable *causes* changes in another?

Does our study/analysis look at how *preceding* changes in one variable are associated with *subsequent* changes in another variable? (Robust Dependence/Granger Causality)

- For example, lagged panel models, cross-lagged panel models. If no, we **cannot infer anything about causality**.

Does our study use an experimental design or leverage a natural experiment? (Consequential Manipulation)

- If it does, then we **can** make causal inferences about the **effect of the treatment on the outcome**.

Does our study present a theoretically-sound causal model and condition appropriately on possible confounders/colliders and other 'backdoor paths'? (Generative Process)

- If it does, then we can **accept the evidence as causal but encourage falsification and development of our model**.

The answers are not in the data!

R Exercise

- This week, you will be practicing running the inferential statistics tests we covered last week using data from a **mixture** of studies (some where causation can be established, some where it cannot).
- Download the **week-5-exercise.zip** file on the Blackboard page and open the Rproject file. Open the **week-5-exercise.Rmd**.
- Go through the exercises. There is a challenge at the end with some open data experiments!