When A/B Experimentation is not possible.

Alternative Methodologies for causality estimation.

Jennifer Vlasiu | 29 June 2022

Session Objectives:

- Introduction to the concept of A/B Experimentation, and why it matters
- Challenges and situations when A/B Experimentation may not be possible
- Basic understanding of what is meant by estimating a causal effect
- Introduction to a few alternative Methodologies for causality estimation
- Demo: Walkthrough Jupyter and R on how to implement causal impact

Session Objectives:

- Introduction to the concept of A/B Experimentation, and why it matters
- Challenges and situations when A/B Experimentation may not be possible
- Basic understanding of what is meant by estimating a causal effect
- Introduction to a few alternative Methodologies for causality estimation
- Demo: Walkthrough Jupyter and R on how to implement causal impact

What is A/B Experimentation

A/B Experimentation is the gold standard for estimating causal effects is a truly randomized experiment, evenly distributed treatment and control cohorts.

A/B Experimentation is where each participant is randomly assigned a group.

- **Treatment** – all the participants in the group that **do** receive an intervention
- **Control** – all the participants in the group that **do NOT** receive an intervention

Session Objectives:

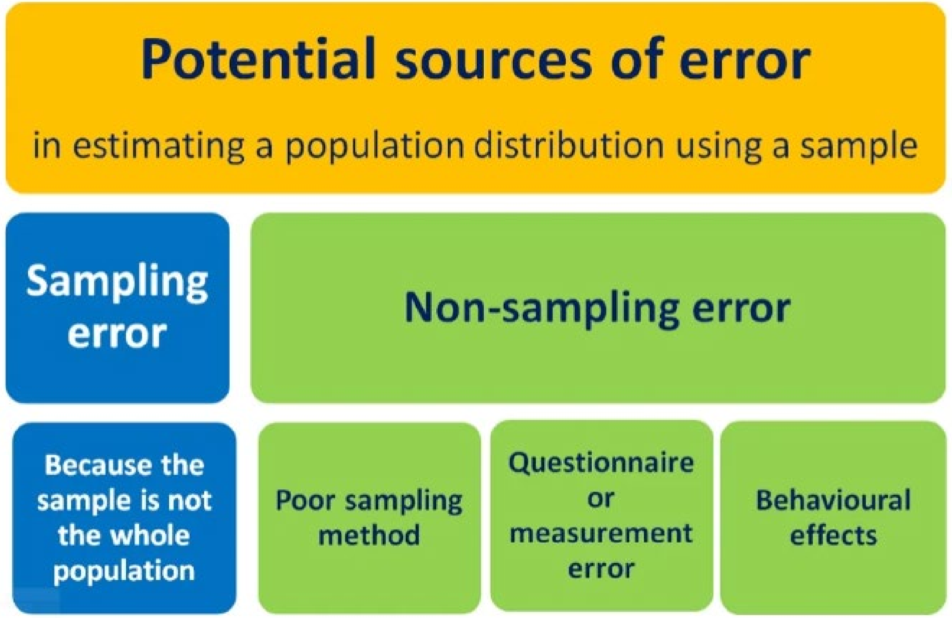
- Introduction to the concept of A/B Experimentation, and why it matters
- Challenges and situations when A/B Experimentation may not be possible
- Basic understanding of what is meant by estimating a causal effect
- Introduction to a few alternative Methodologies for causality estimation
- Demo: Walkthrough Jupyter and R on how to implement causal impact

When might the "Gold Standard" A/B Experimentation not be possible?

A/B Experimentation is the gold standard for estimating causal effects is a truly randomized experiment, evenly distributed treatment and control cohorts.

- Too Expensive
- Too Difficult
- Unethical
- Assembly/Instrumentation not done

In these situations we want to obtain a causal effect, in the absence of experimentation



Source: https://learnandteachstatistics.wordpress.com/2014/09/04/sampling-and-non-sampling-error/

The Challenge of Inference in the Social Sciences (slideshare.net)

Session Objectives:

- Introduction to the concept of A/B Experimentation, and why it matters
- Challenges and situations when A/B Experimentation may not be possible
- Basic understanding of what is meant by estimating a causal effect
- Introduction to a few alternative Methodologies for causality estimation
- Demo: Walkthrough Jupyter and R on how to implement causal impact

What is Causal Inference?

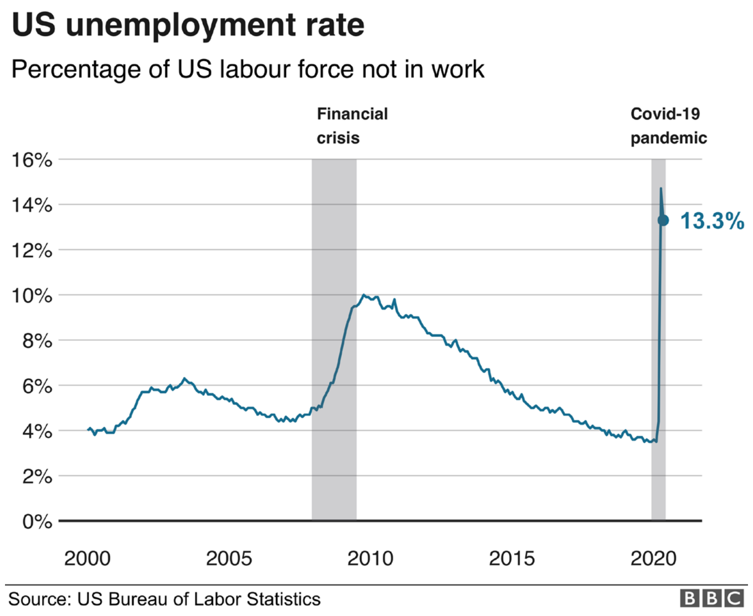
Causal inference is the branch of statistics that is concerned about effects; the consequences of our actions.

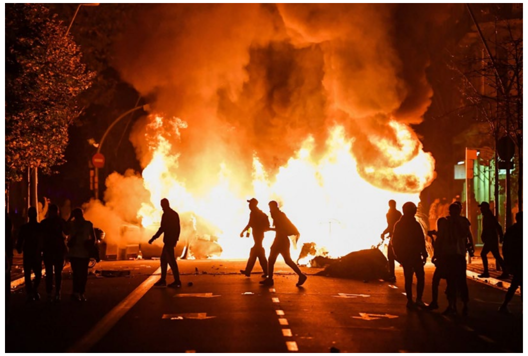
Identifying one causal law can often be **more important** than identifying multiple correlation patterns that you might find

Understanding causal effects is the **bedrock** of big data and **of modern data science.**

What is Causal Inference?

Causal inference seeks to understand the effect of some variable(s) on some other variables(s)





Example: The causal effect of **unemployment** on the **probability that a riot** will occur

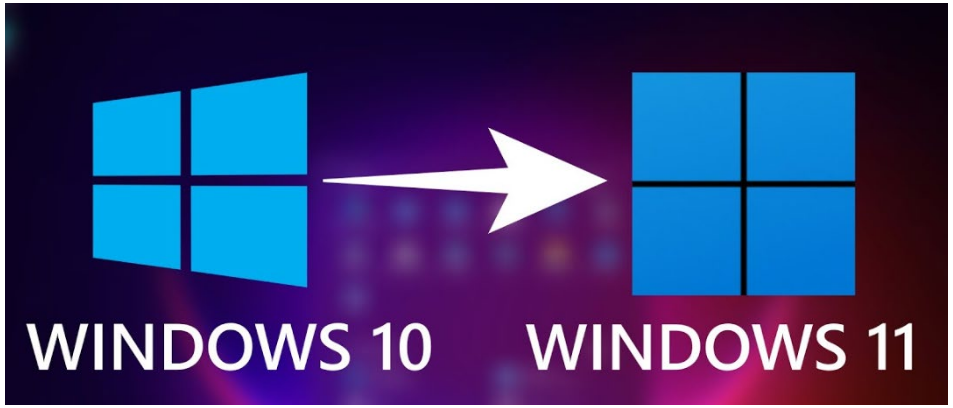
Example: The causal effect of **a riot** on **government spending**

[US unemployment sees surprise improvement in May - BBC News](#)
[State Unemployment Map, January 2016 (businessinsider.com)](#)

What is “Encouragement” Design?

Encouragement Design is a term used when we are **unable to force** all experimental subjects to take the (randomly) assigned treatment or control groups. It includes:

1. Can’t force users in treatment control to take the treatment



Example: **Software Upgrades**. One can only encourage the users to take the treatment by making “Download Upgrade” as an option

Example: **Marketing Campaign via Notification**. User may have turned off notifications, hence they won’t see the campaign

2. Can’t force users in the Control group to NOT take the treatment

Example: **Software Upgrades**. User could download the upgrade directly from the internet even when “Download Upgrade” option is not available

[turn off notifications chrome - Bing images]

Session Objectives:

- Introduction to the concept of A/B Experimentation, and why it matters
- Challenges and situations when A/B Experimentation may not be possible
- Basic understanding of what is meant by estimating a causal effect
- Introduction to a few alternative Methodologies for causality estimation
- Demo: Walkthrough Jupyter and R on how to implement causal impact

Alternative Methodologies for Causality Estimation

As with all non-experimental approaches to causal inference, valid conclusions require strong assumptions. Try at least two methods, with different assumptions to give confidence

	Google Google Causal Impact Bayesian Time Series	Uber CausalML	Microsoft Microsoft EconML	Microsoft Microsoft "Do Why" Package	IBM IBM Causal Library
About	• **Used to estimate the causal effect of a designed intervention** on a time series (i.e. daily clicks generated from ad campaign		• Estimating causal effects with ML • Helps to estimate individualized casual responses from observational or experimental data	• Use people who think about prediction problems, to also consider Causation problems • End-to-end library for causal inference, wrapper for different libraries (that hold some estimation code) that can be called from Do Why • Outer layer, common API that helps you reason.	Used to estimate the causal effect of an intervention on **some outcome from real-world, non-experimental observational data**
Underlying Statistical Technique	• **Bayesian Time-Series Model** to estimate how response metric may have evolved after the intervention and if the intervention had not occurred. The delta is then deemed the net impact			• Wrapper for causal-inference estimation approaches	• Wrapper for causal-inference estimation approaches
Model Assumptions	• Assumes a graph, not explicitly stated			• Need to explicitly understand causal effect graph (can use DAGitty.net • Like an API	• Package encompasses suite of causal methods using a sklearn-like fit-predict API
When to Use	• When randomized experiment is not available. • When seeking to understand causal effects; what would have happened in the absence of the action whose effects we're interested in		• Recommendation A/B Testing; Interpret experiments with imperfect compliance	• Breaks down causality into 4 steps, to evaluate causal models	• If looking for flexible causal modelling, with high tweaking availability.
	CausalImpact (google.github.io) This package in R implements a Bayesian approach to causal impact estimation in time	GitHub - uber/causalml: Uplift modeling and causal inference with machine learning algorithms	GitHub - microsoft/EconML: ALICE (Automated Learning and Intelligence for Causation and	Causality and Machine Learning – Microsoft Research	GitHub - IBM/causallib: A Python package for modular causal inference analysis and model

High Level Implementation Steps

Step 1:
Setup Knowledge Graph
(DAGitty - drawing and analyzing causal diagrams (DAGs))

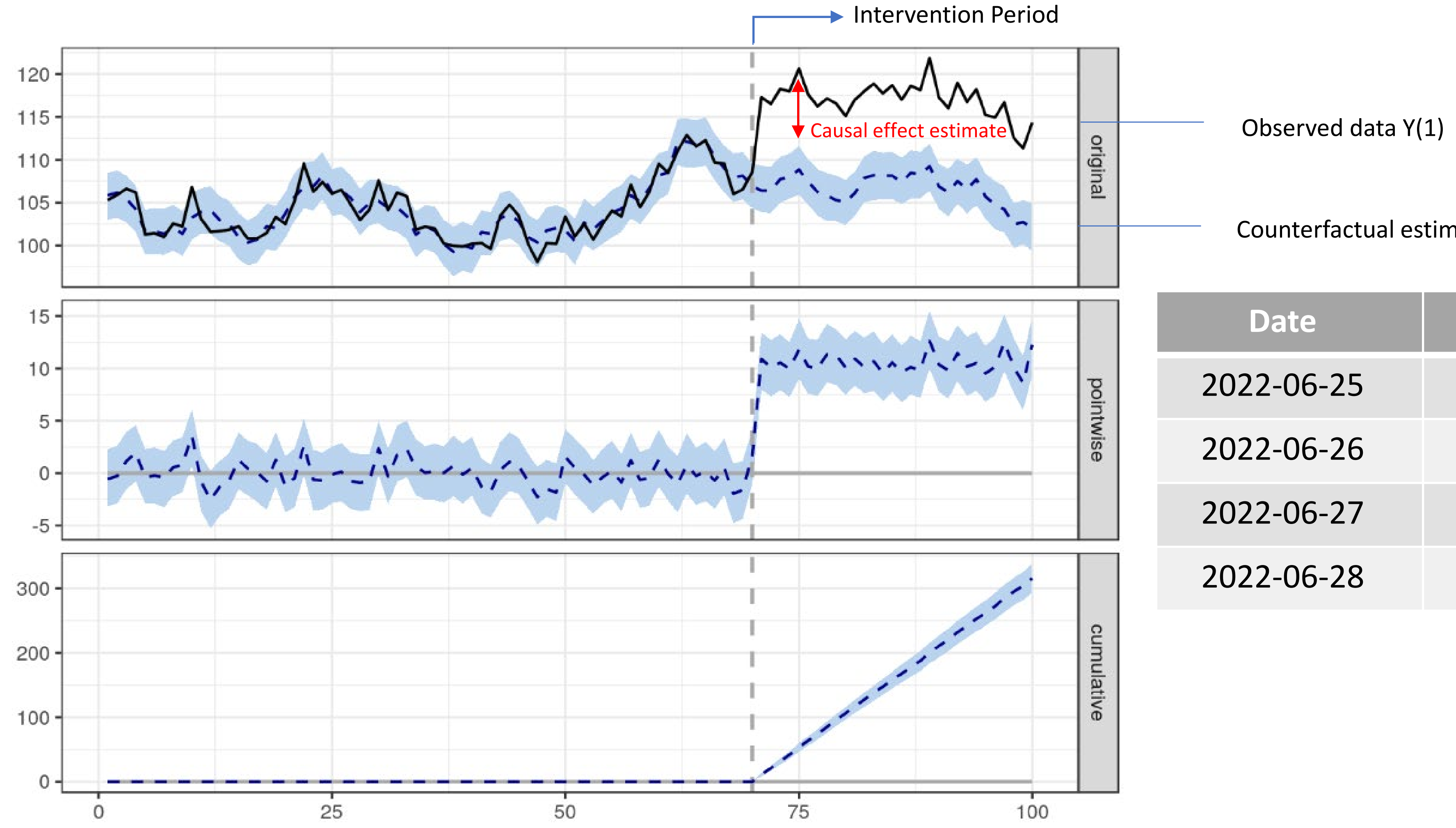
Step 2:
Validate via refute tests if knowledge graph holds

Step 3:
Select and implement causal impact packages from Google, Microsoft, Uber

Step 4:
Implement and validate via wrapper, such as Microsoft's Do Why Package

Google Causal Impact in R

- Uses Bayesian structural time-series; statistical technique that intakes 3 variables
- Looks at observed data y1, and to understand causal effects, we estimate the counterfactual; what would have happened in the absence of the action whose effects we're interested in
- Since don't really have an experiment, there isn't a true "control" in experiment sense. If want to be able to estimate something that looks like a control, can do so via **synthetic control**
- **Difference between counterfactual and the actual observed data is the causal effect**



Date	Pre-Period	Post-Period
2022-06-25	105	115
2022-06-26	106	120
2022-06-27	107	115
2022-06-28	106	113

Session Objectives:

- Introduction to the concept of A/B Experimentation, and why it matters
- Challenges and situations when A/B Experimentation may not be possible
- Basic understanding of what is meant by estimating a causal effect
- Introduction to a few alternative Methodologies for causality estimation
- Demo: Walkthrough Jupyter and R on how to implement causal impact

Inferring effect of an event using Causal Impact – DEMO

References/Helpful Links

- Inferring the effect of an event using CausalImpact by Kay Brodersen – YouTube
- The Challenge of Inference in the Social Sciences (slideshare.net)
- GitHub - IBM/causallib: A Python package for modular causal inference analysis and model evaluations