IEEE VIS 2026 · Boston · TVCG paper

Generating Coherent Visualization Sequences for Multivariate Data by Causal Graph Traversal


Puripant Ruchikachorn (Joe)


Darius Coelho


Jun Wang


Kristina Striegnitz


Klaus Mueller

Chulalongkorn University · Stony Brook University · Union College

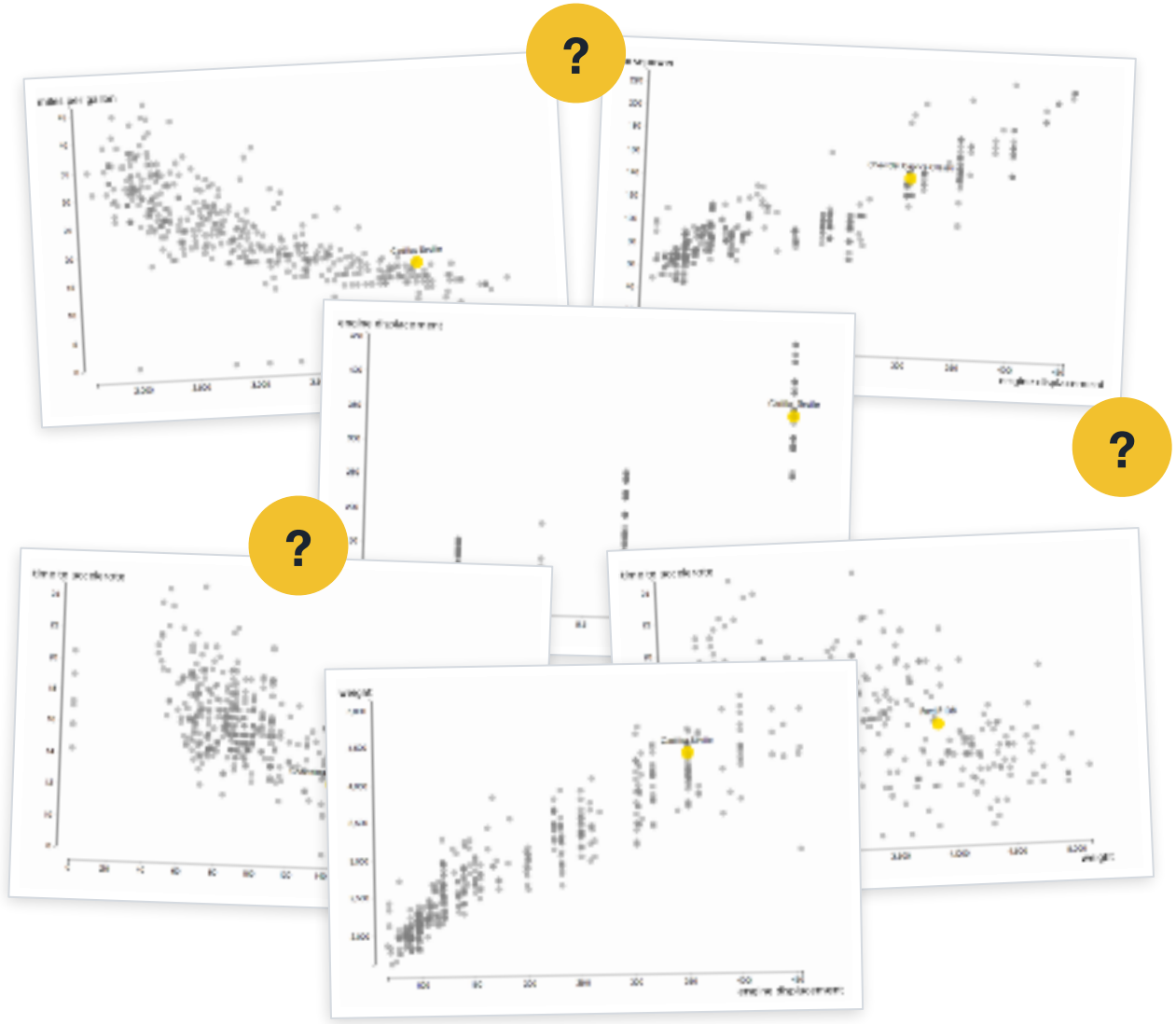
Many charts. In what sequence?

15
scatterplots from just 6 variables

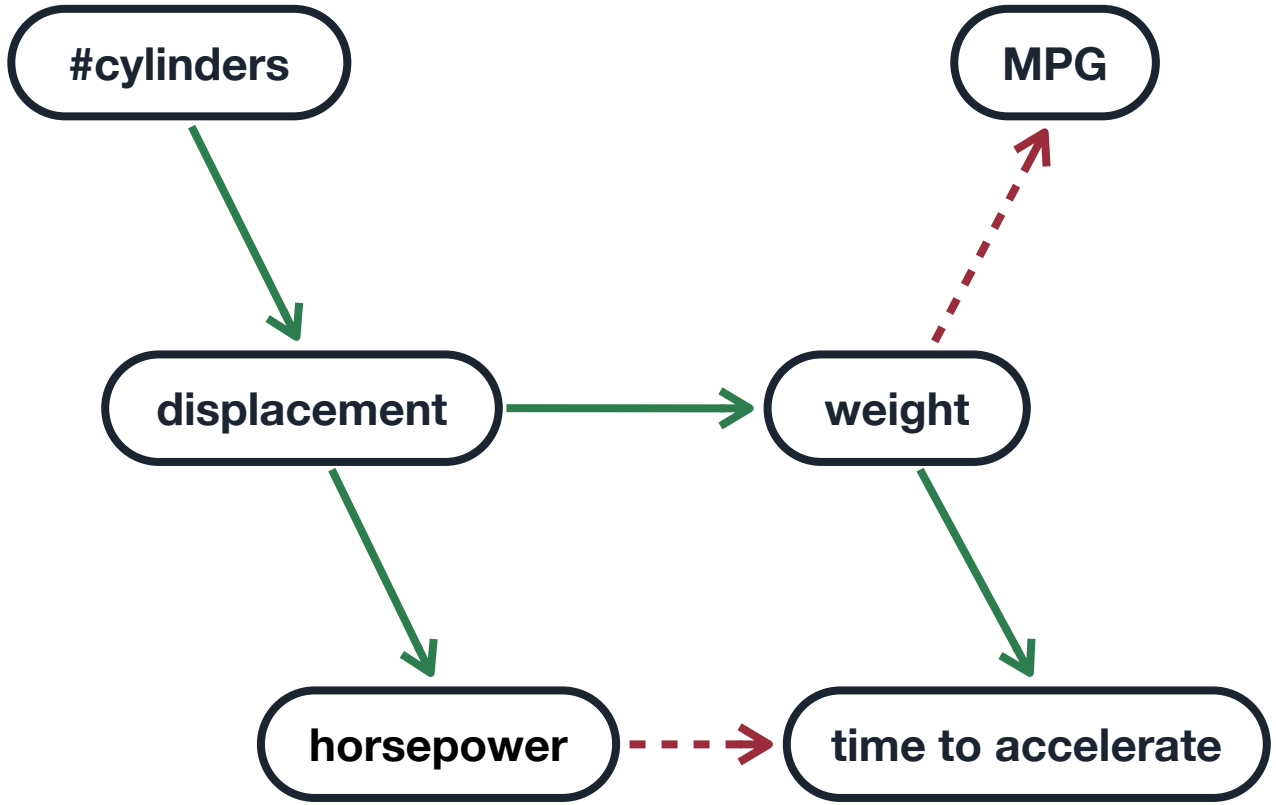
1.3 trillion
ways to order those 15 charts (15!)

720
axis orders for a 6-axis PCP (6!)

Orderings by correlation, similarity, or quality metrics make each chart informative, not the sequence coherent.



Many charts. Causal sequence.



positive negative Causal graph of the Cars dataset

CORRELATION

weight & MPG

r = –0.79 but the relation is symmetric.
which one drives the other?

CAUSALITY

weight → MPG

Each edge is directed and represents a cause-and-effect step of a sequence.

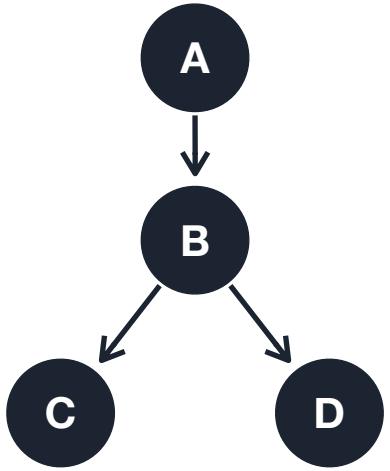
Many charts. Causal sequence.

1 Data

A	B	C	D

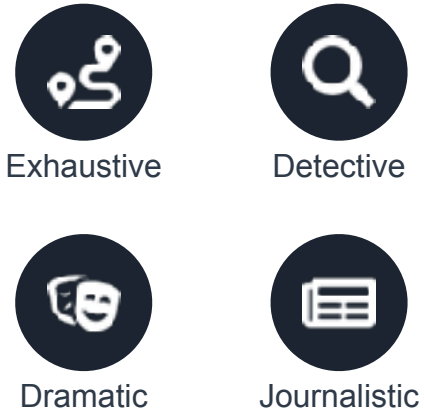
Numerical variables & light metadata, e.g., whether high or low values are desirable.

2 Causal graph



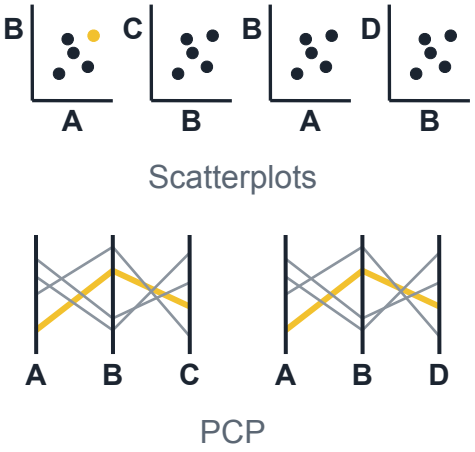
Relations are computed from causal analysis. Each edge has direction and strength. Analysts can edit the graph as they see fit.

3 Traversal



A traversal scheme picks which causal chains to show and in what order.

4 Sequence



One scatterplot per edge, or one PCP per chain, each with one highlighted example by rejection sampling.

Four Causal Graph Traversals for Four Sequences



Exhaustive

Depth-first, from root causes to leaf effects.

SHOWS

Every causal link



Detective

Backward, from a goal effect to its causes, like a whodunit.

SUGGESTS

The causes of a goal



Dramatic

A causal collider as a story conflict, after Freytag's pyramid.

RESOLVES

Which causal path dominates



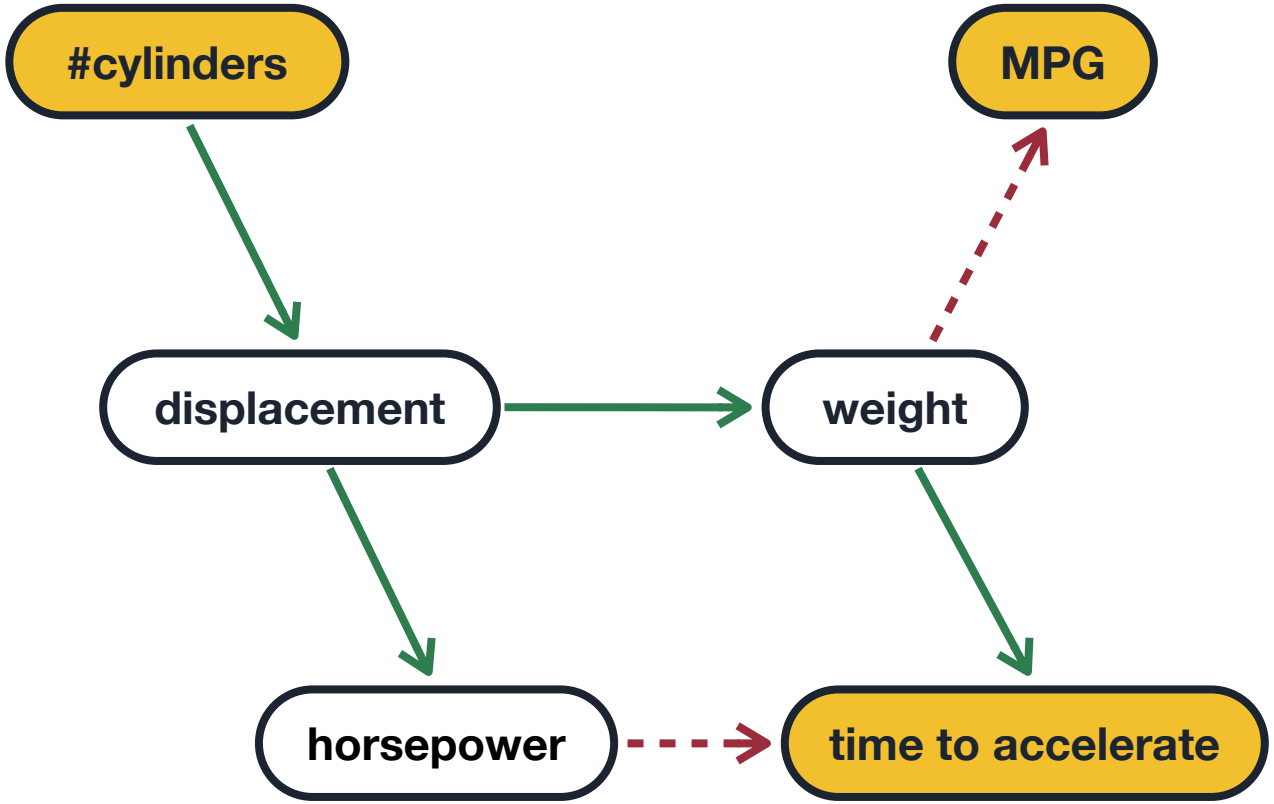
Journalistic

A balanced reporting on the causes of goal effects.

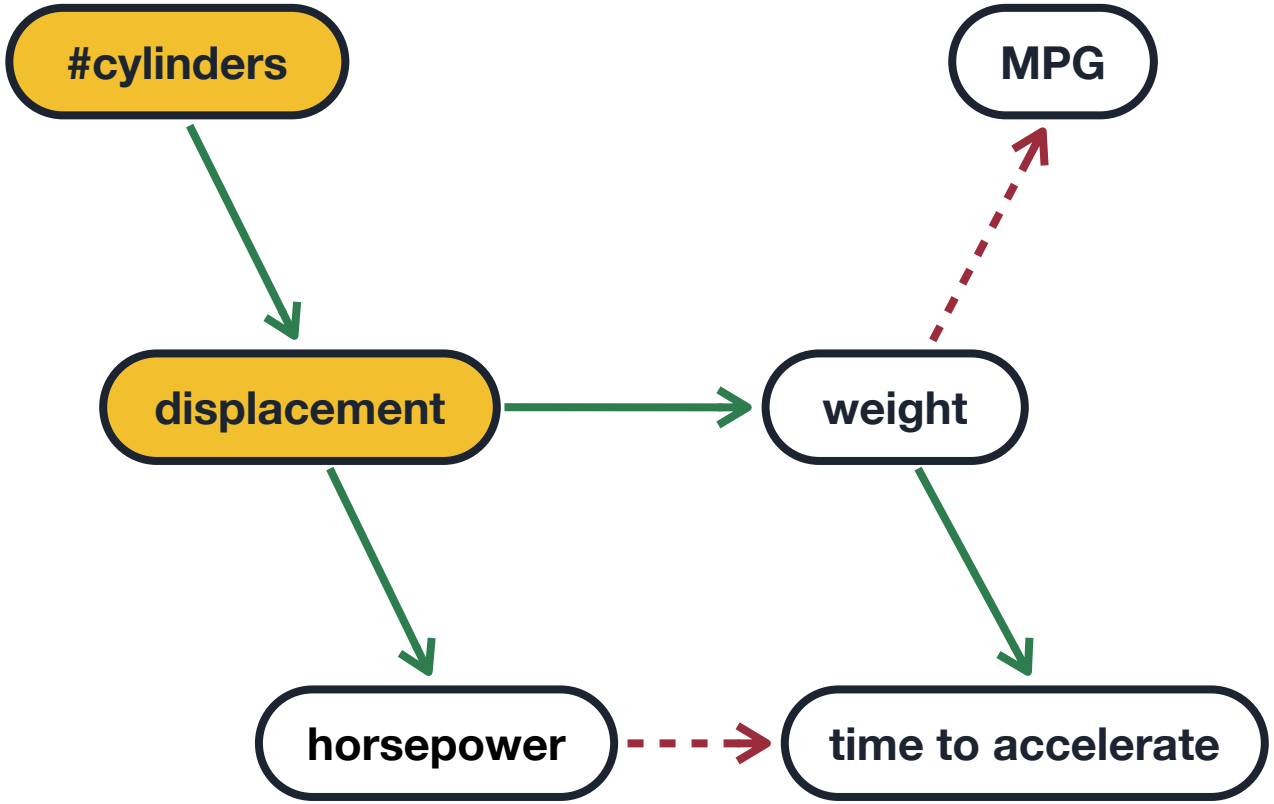
INVESTIGATES

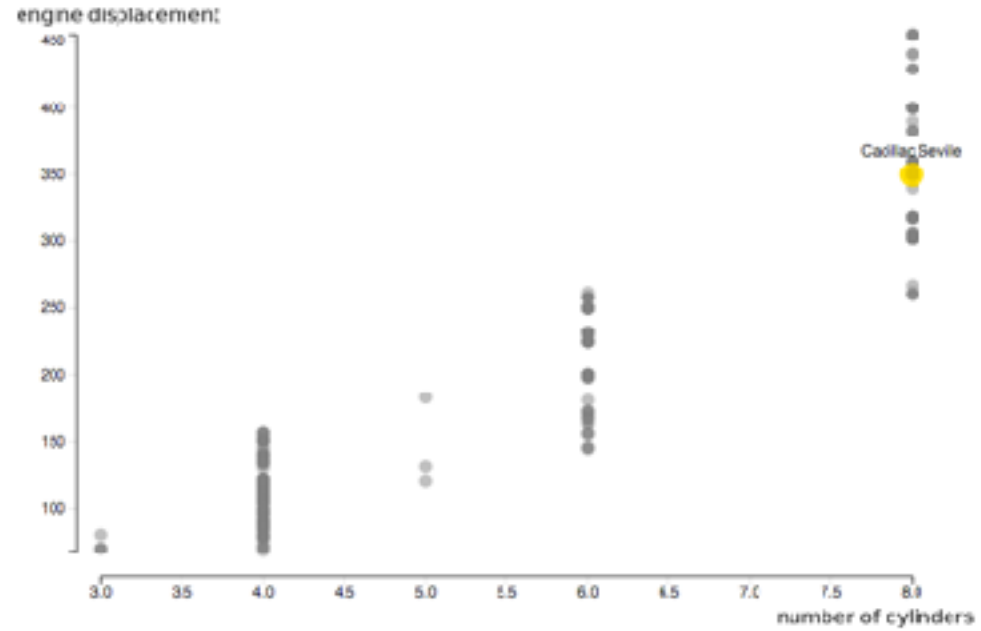
A trade-off between goals

Exhaustive Sequence: every causal link

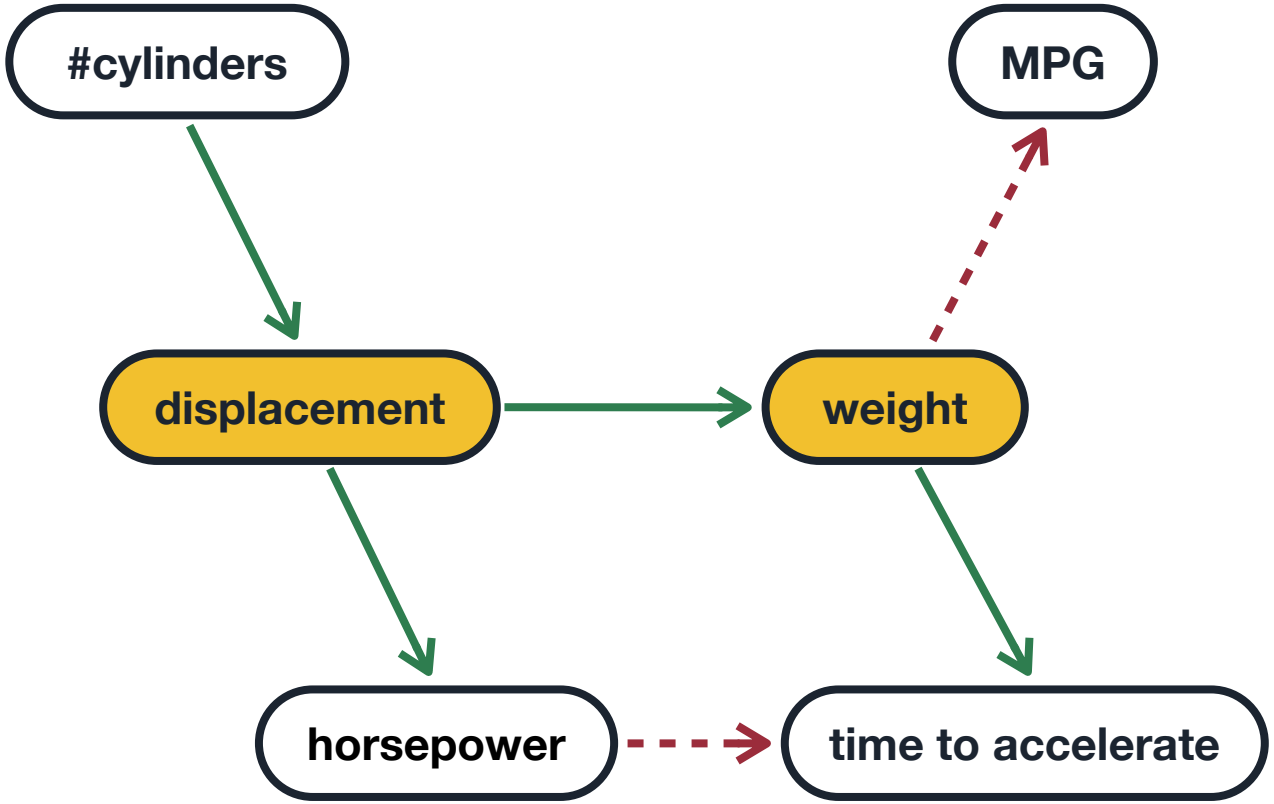


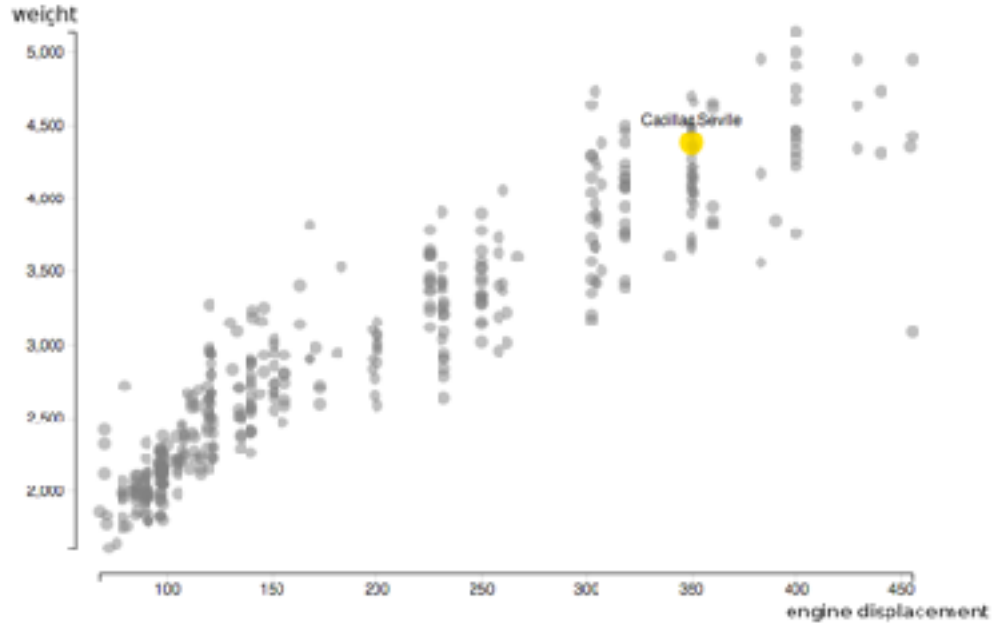
Exhaustive Sequence: every causal link



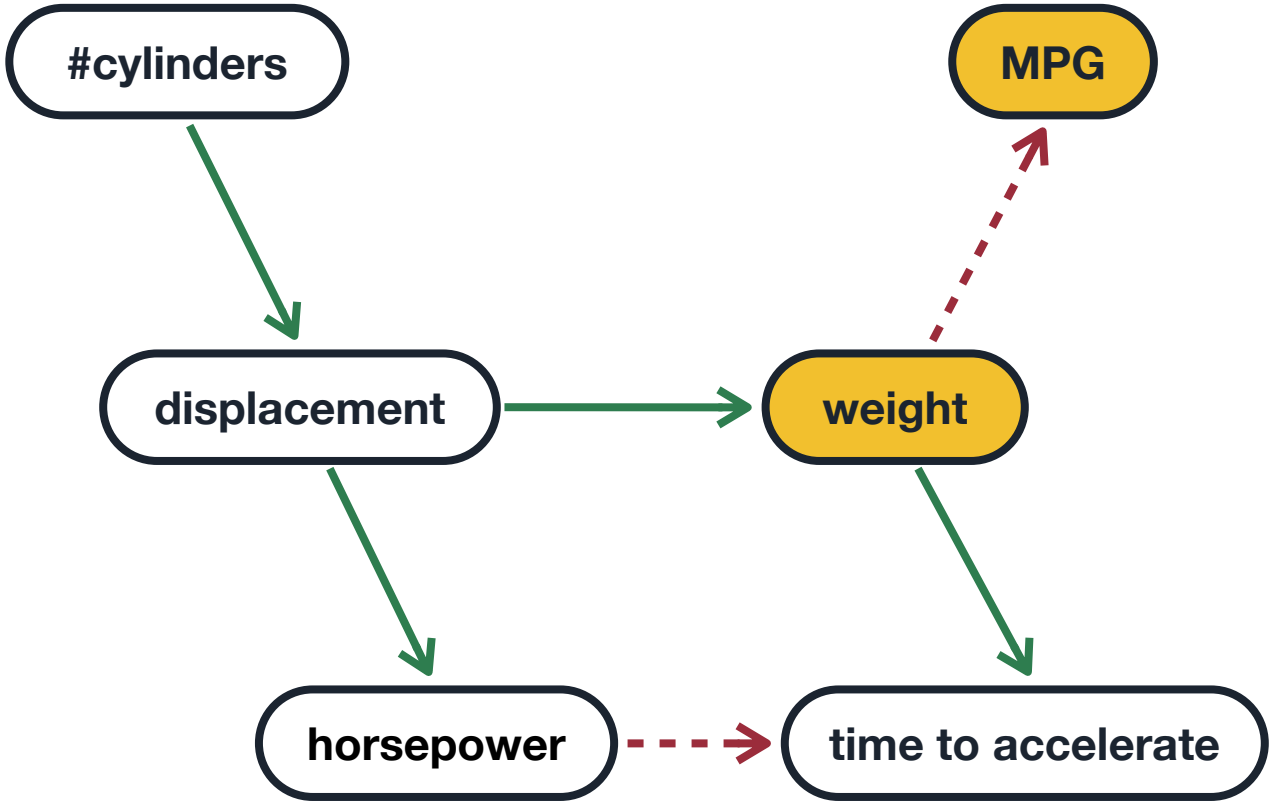


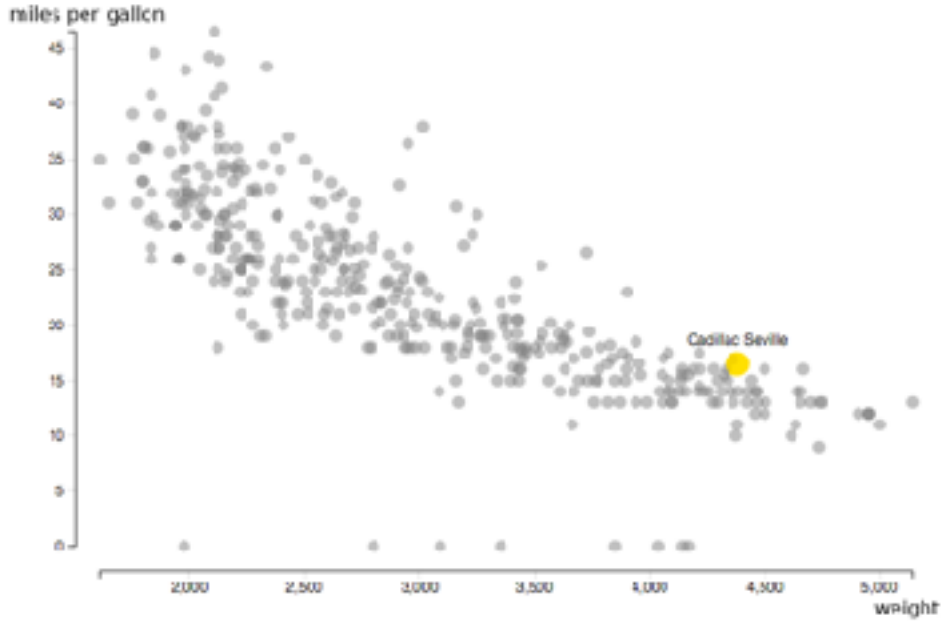
Exhaustive Sequence: every causal link



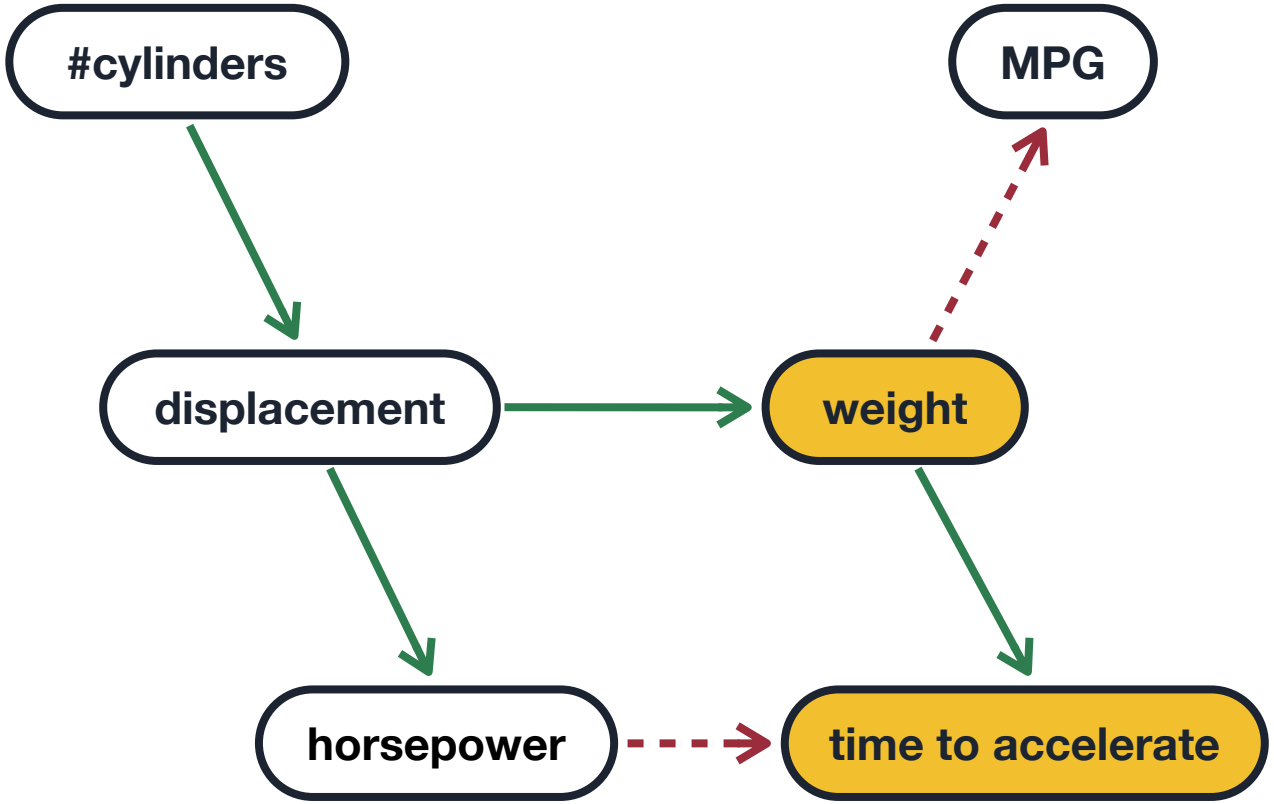


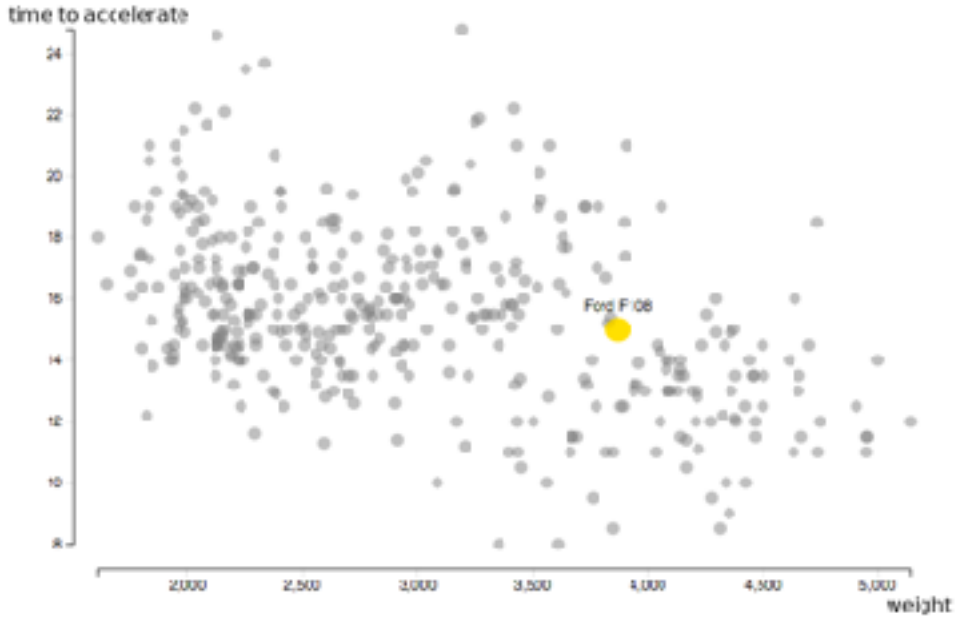
Exhaustive Sequence: every causal link



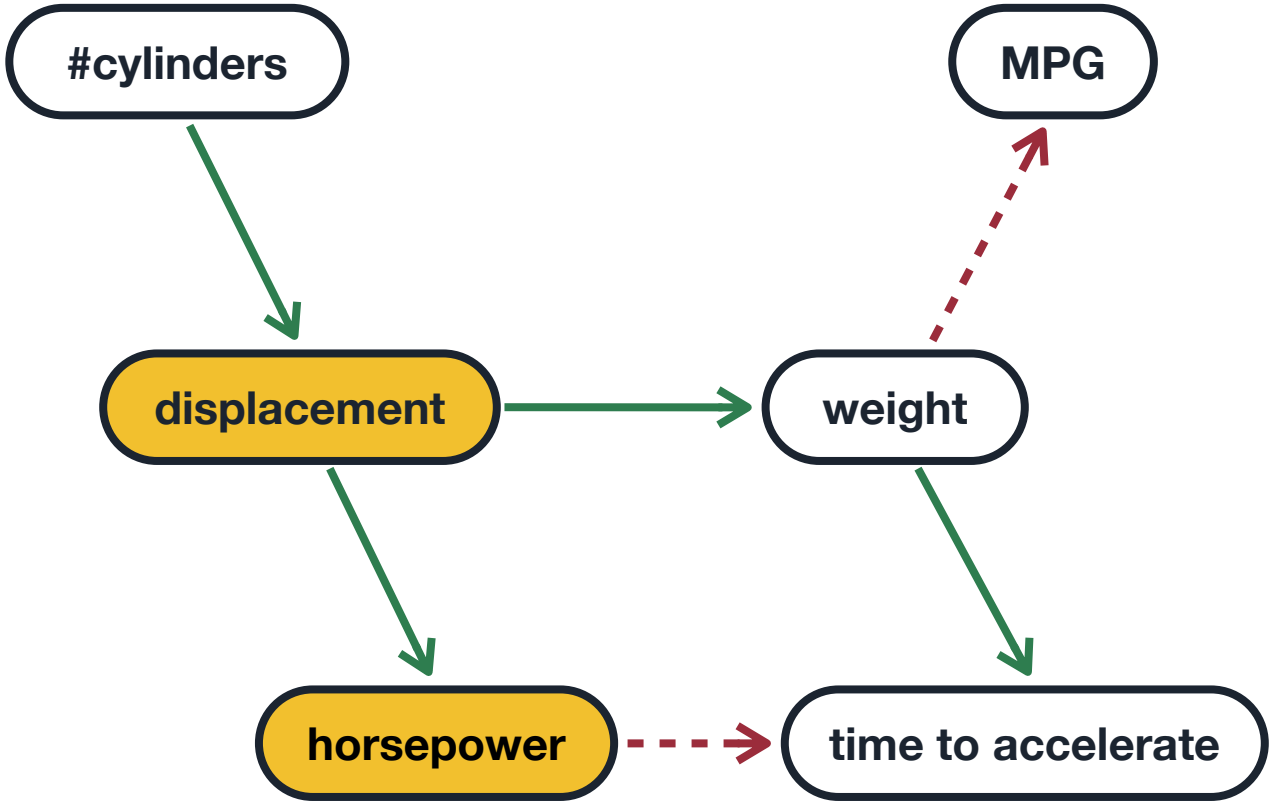


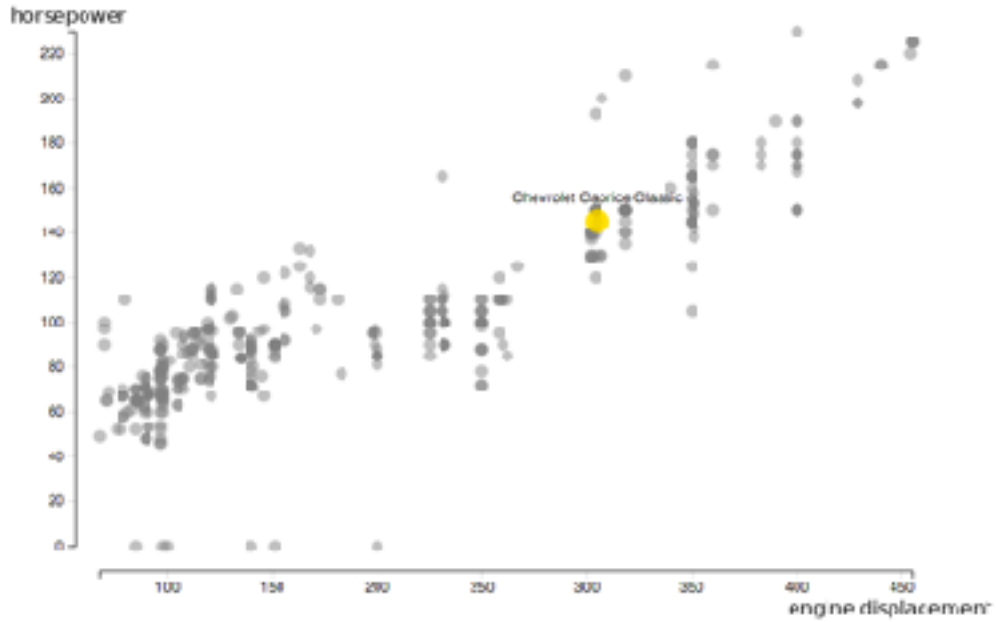
Exhaustive Sequence: every causal link



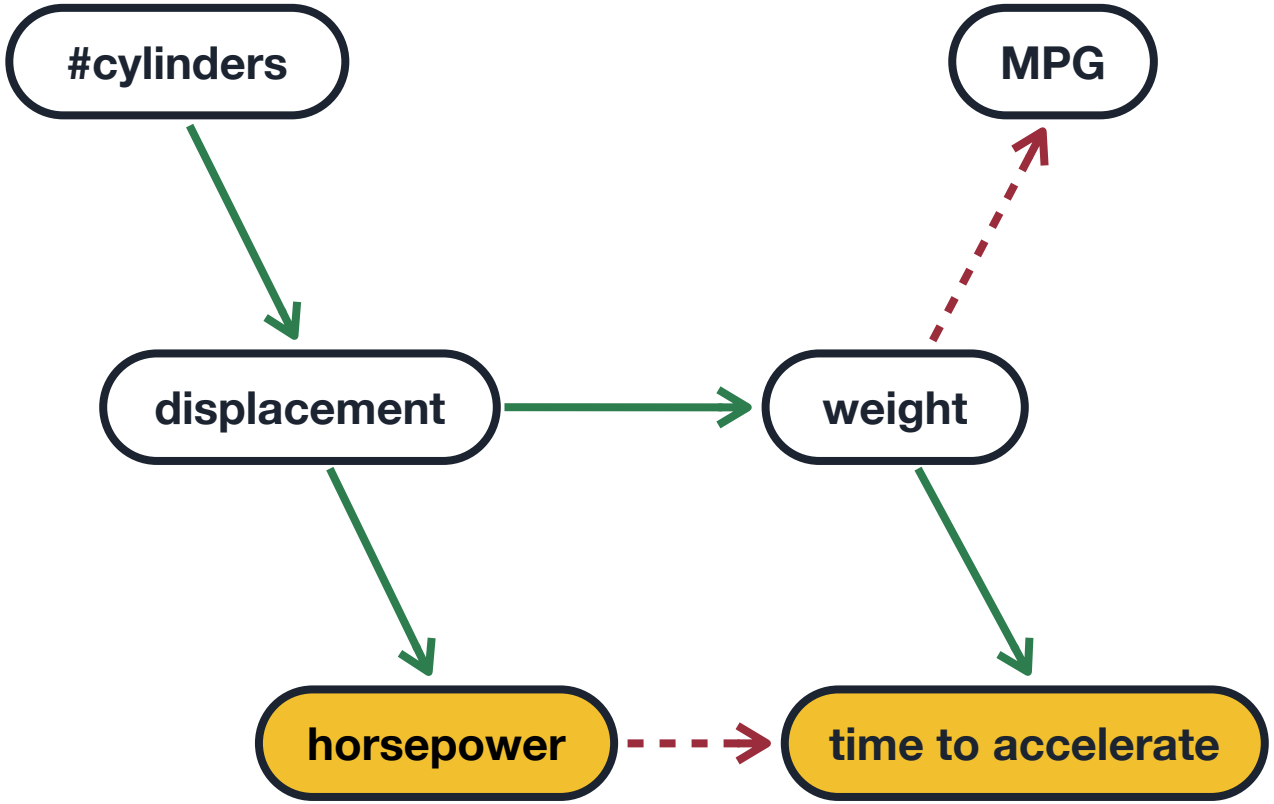


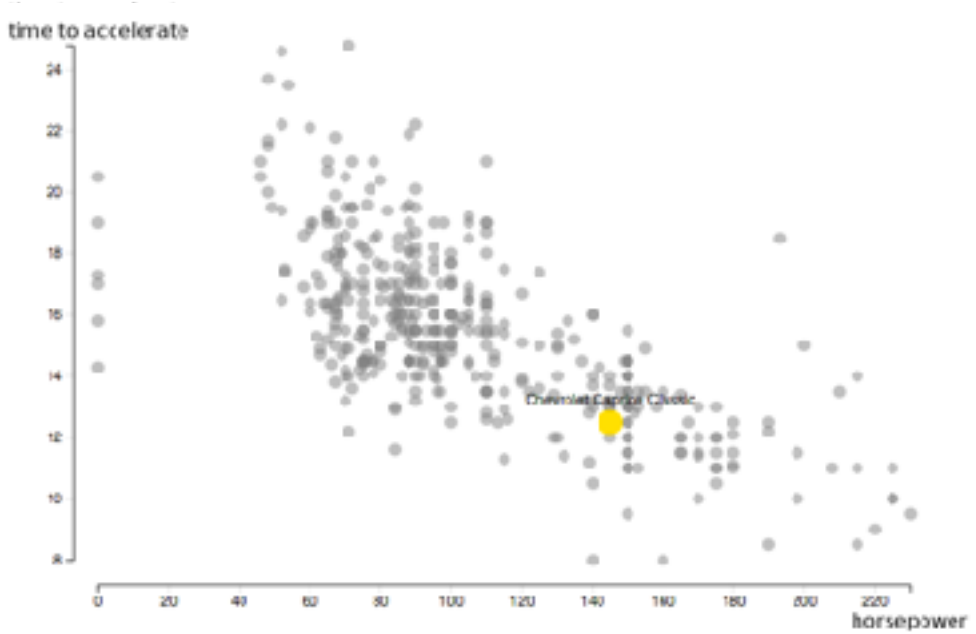
Exhaustive Sequence: every causal link



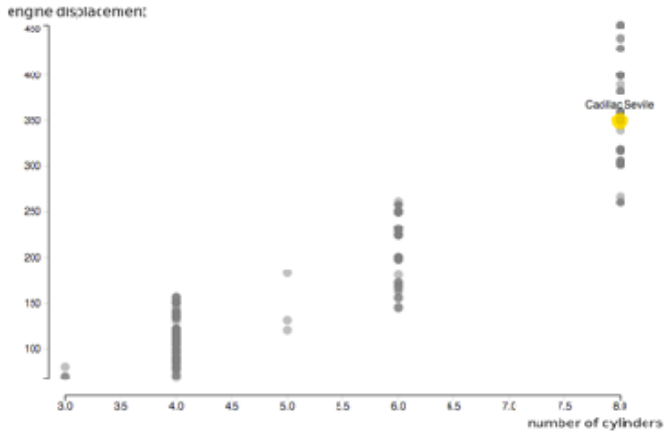


Exhaustive Sequence: every causal link

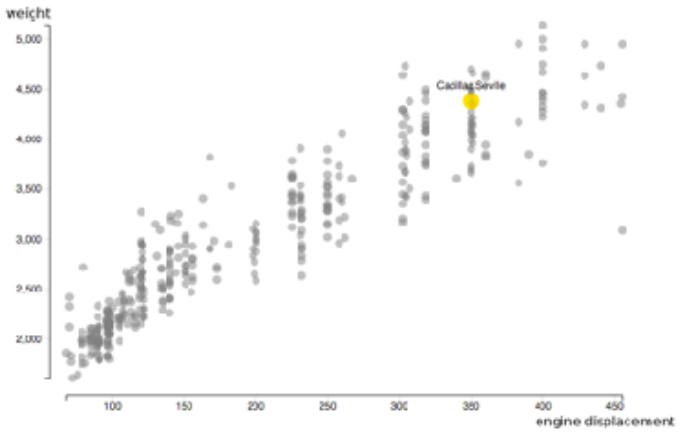




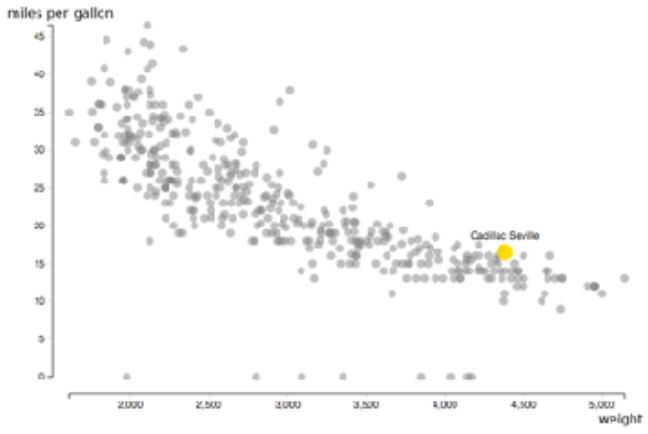
Exhaustive Sequence: every causal link



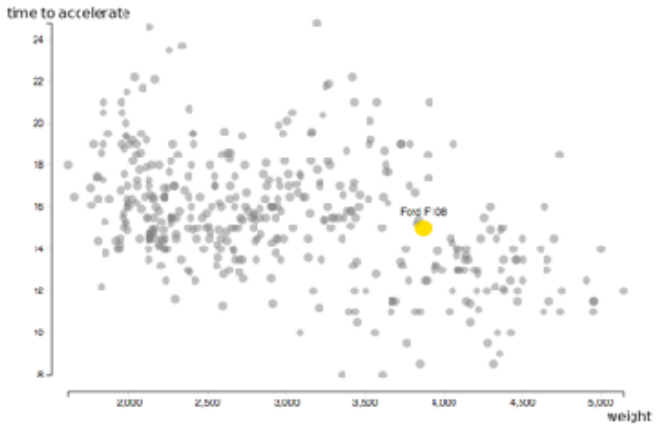
1 #cylinders → displacement



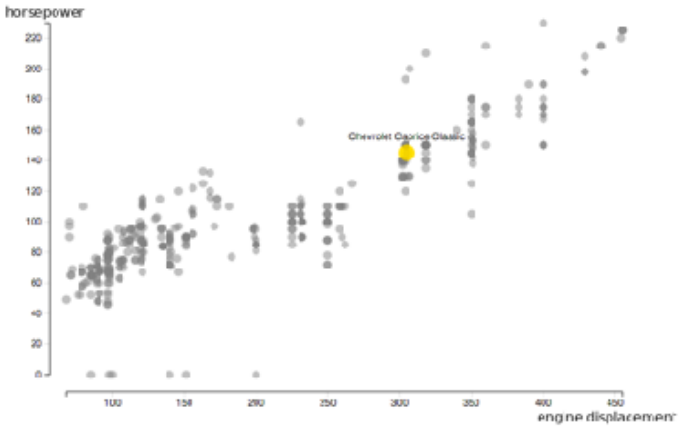
2 displacement → weight



3 weight → MPG

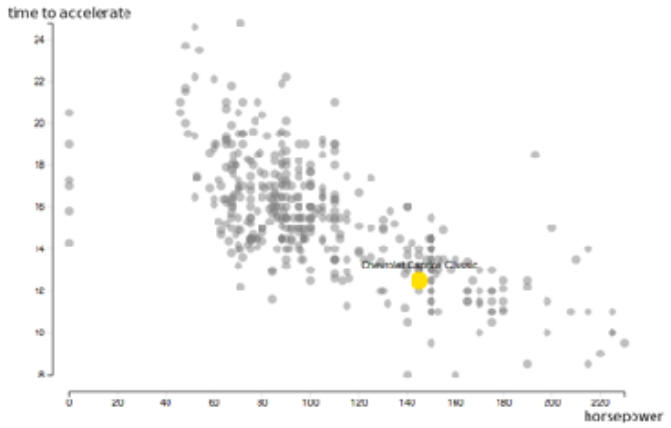


4 weight → time to accelerate



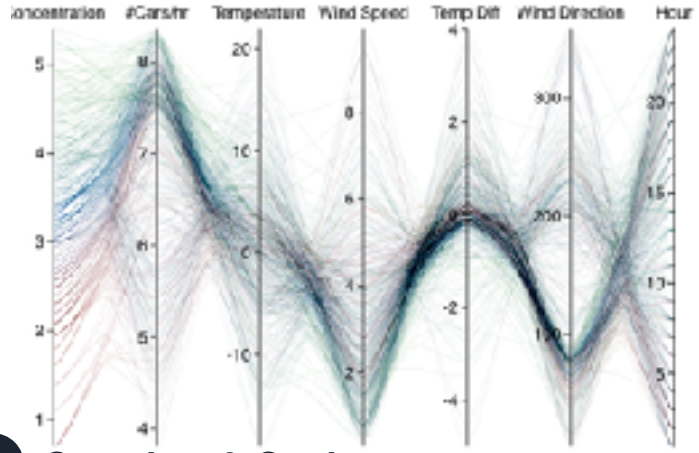
5 displacement → horsepower

↺ after a flashback to frame 1



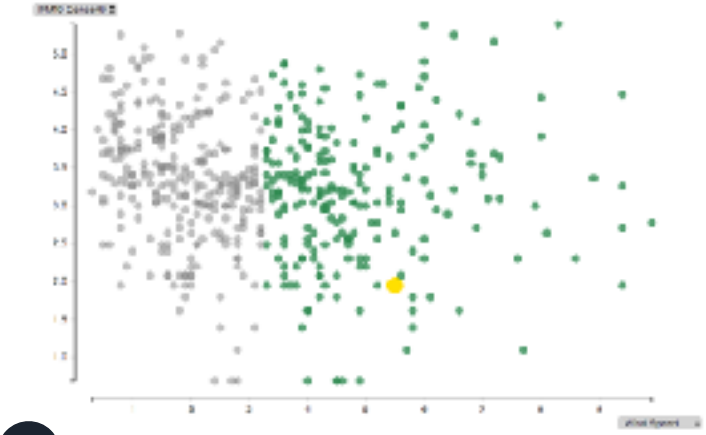
6 horsepower → time to accelerate

Detective Sequence: from effect back to causes



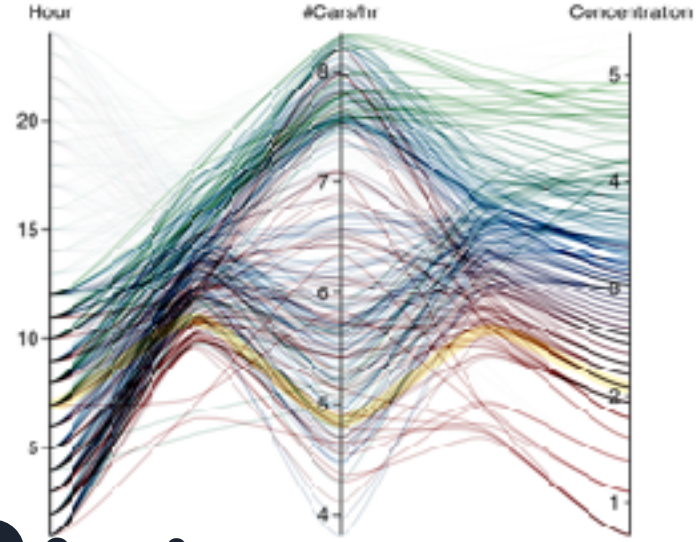
1 Overview & Goal

Low PM10 concentration



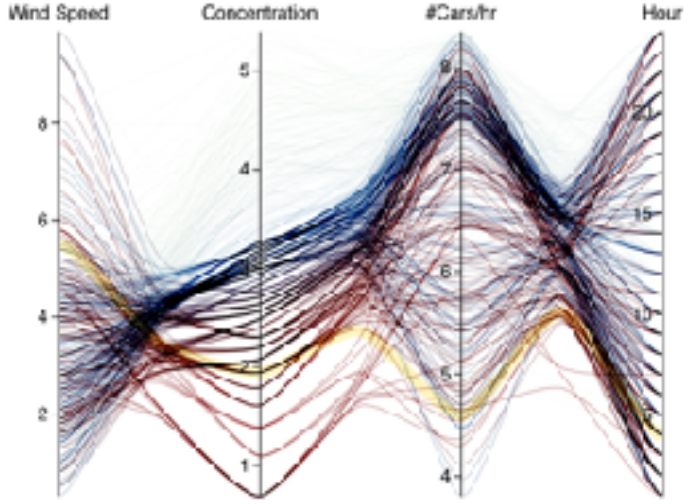
2 Cause 1

Strong wind reduces PM10



3 Cause 2

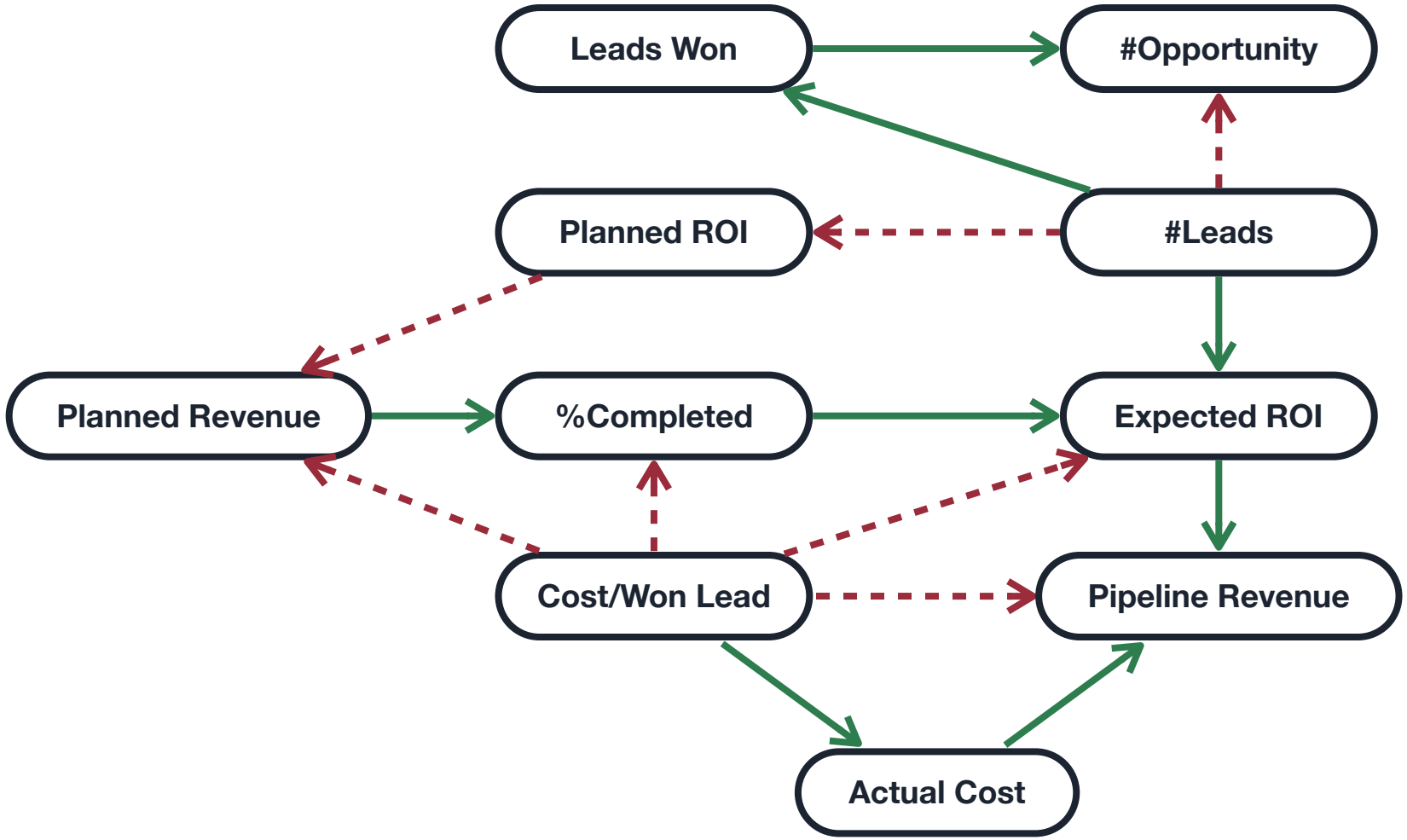
Early hour, less cars per hour, less PM10



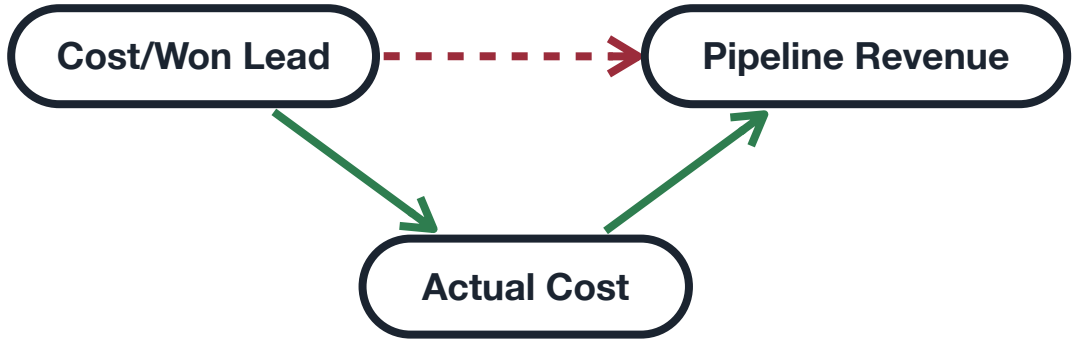
4 Suggestions

increase wind speed or decrease #cars/hr & hour to reduce PM10 concentration

Dramatic Sequence: a causal conflict

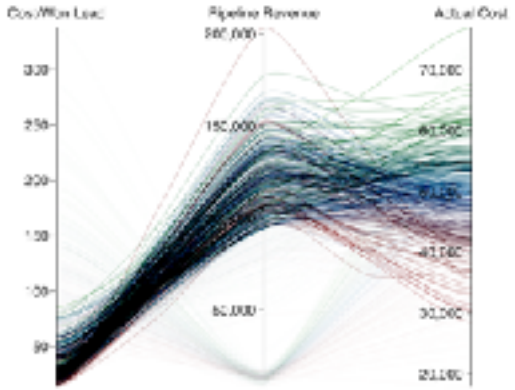


Dramatic Sequence: a causal conflict



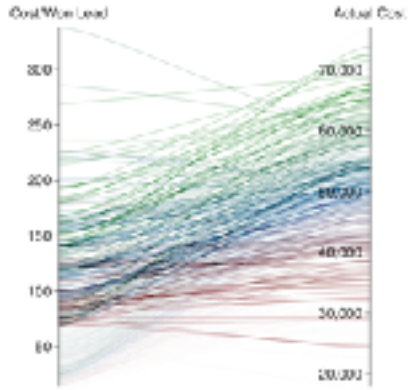
Causal conflict: a collider whose two causes are linked, and whose direct and indirect paths disagree.

Cost/Won Lead raises Pipeline Revenue through Actual Cost, but lowers it directly.



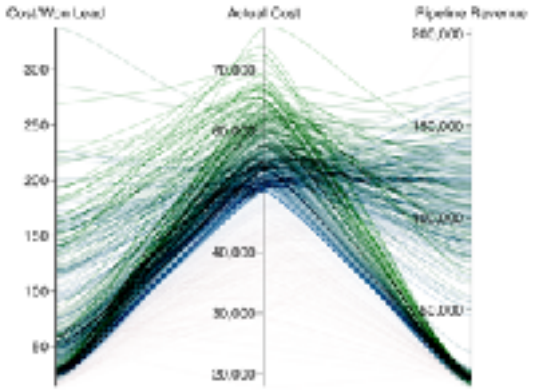
1

all three variables



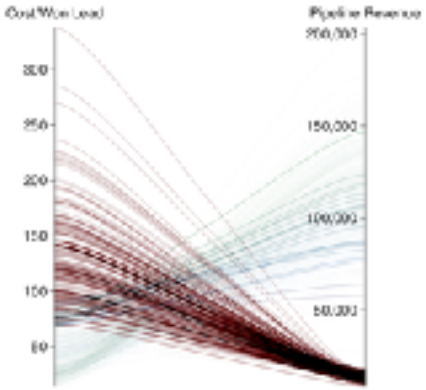
2

high Cost/Won Lead
→ high Cost



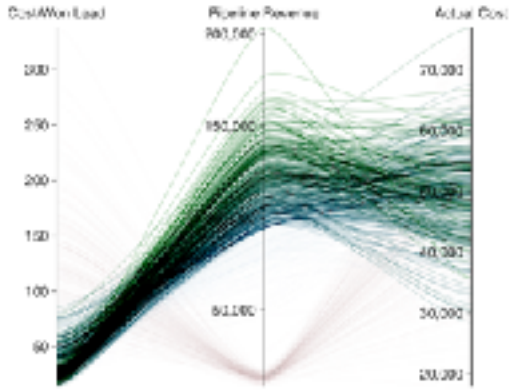
3

high Cost
→ high Revenue



4

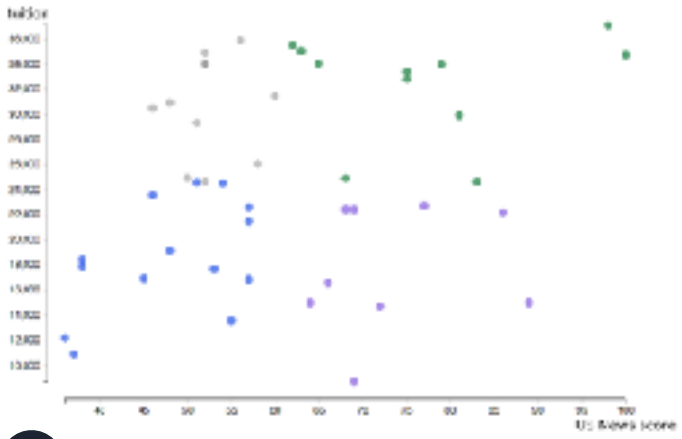
high Cost/Won Lead
→ low Revenue



5

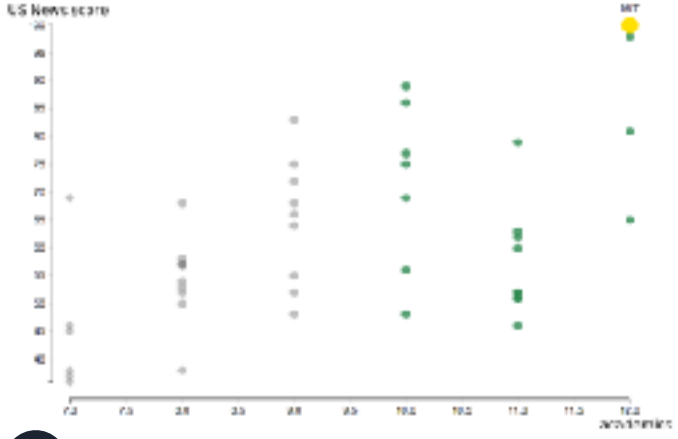
Cost/Won Lead
dominates

Journalistic Sequence: balancing a trade-off



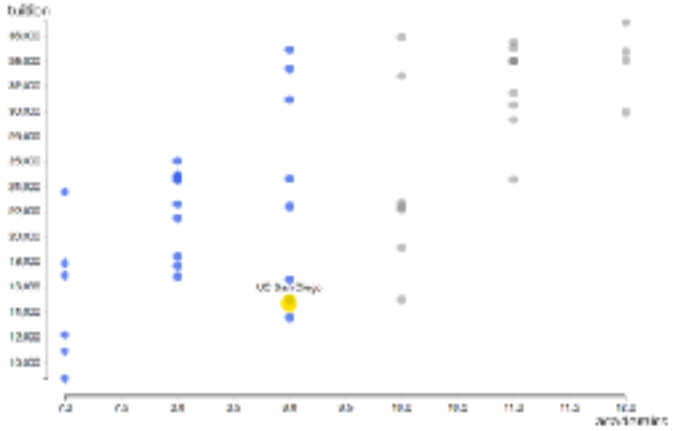
1 Goals

high US News score, low tuition



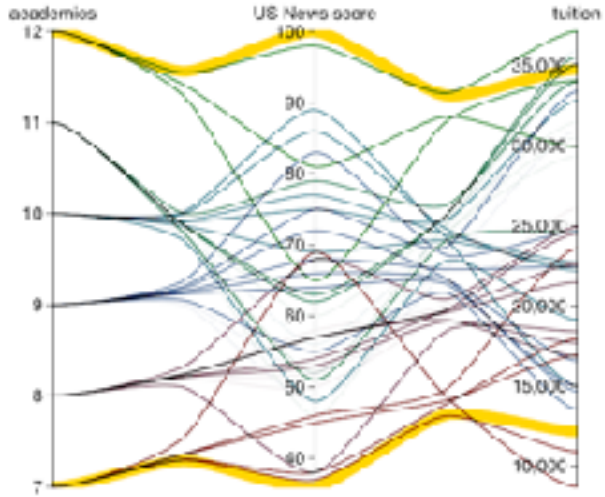
2 High academic score

Caused by high academics



3 Low tuition

Caused by low academics



4 Trade-off

The goals are incompatible and need to compromise.

Crowdsourced Study: do causal sequences help?

50 Mechanical Turk participants,
41 with valid responses

1. **Tutorial**
   Scatterplot or PCP; no causal language
2. **4 Datasets, 4 Sequences**
   Cars, Sales, College, PM10: one sequence each
3. **Write**
   Describe the relationships you see
4. **Score**
   Two coders rate every causal clause

SEQUENCE TYPES

Exhaustive
causal: depth-first

Detective
causal: goal-first

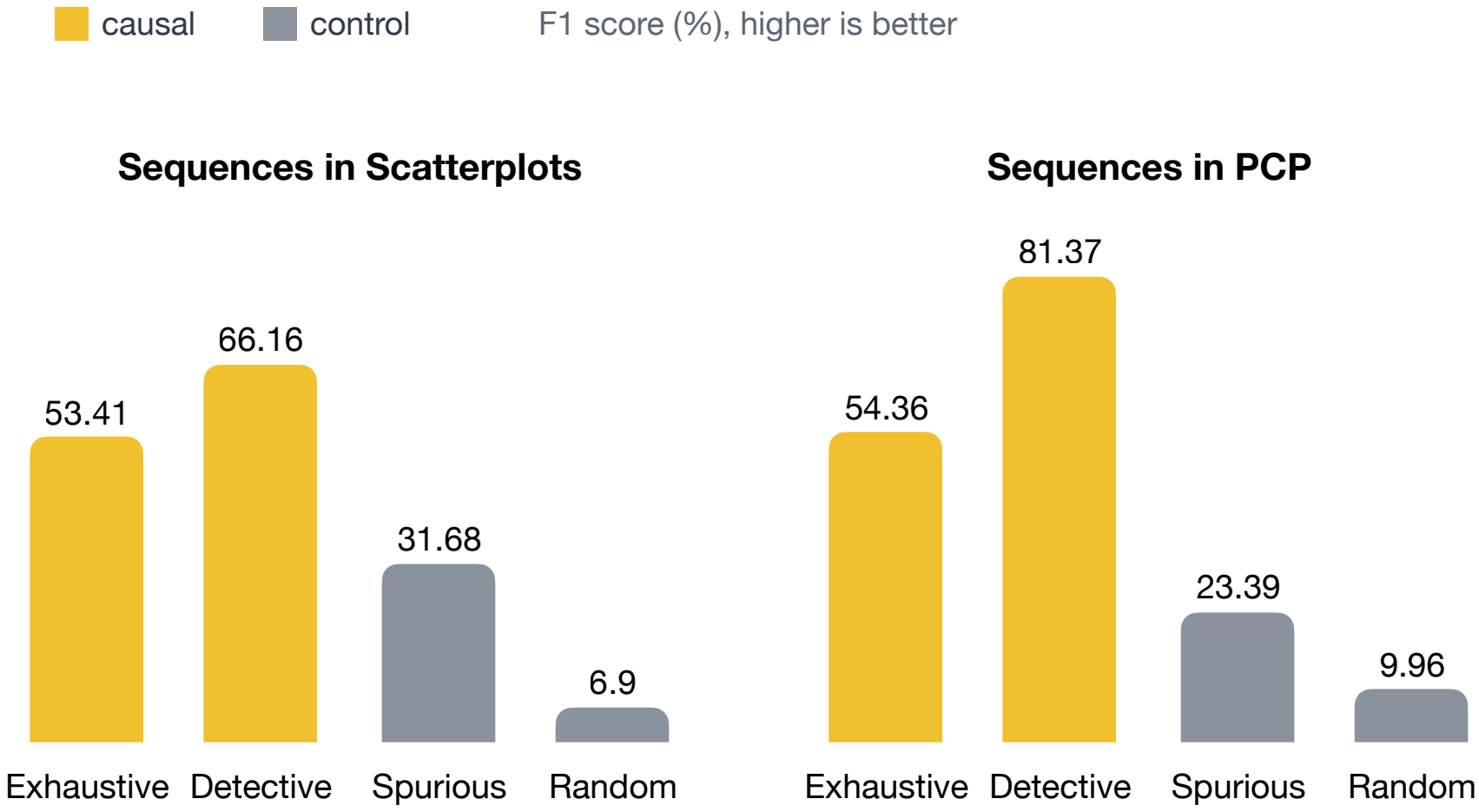
Spurious
correlated pairs,
random order

Random
random pairs

CLAUSE SCORE (for precision, recall & F1)

1	0.5	0
right pair, right direction	right pair, wrong direction	wrong pair

Crowdsourced Study: causal sequences beat both controls



81.37

F1 for detective + PCP, the best combination, driven by recall (90.74%)

$p < .01$

Sequence type matters (one-way ANOVA)

Tukey HSD: every causal vs. control pair differs. However, exhaustive vs. detective does not.

Scatterplot $F(3, 76) = 24.85$; PCP $F(3, 80) = 21.47$. About 5 valid responses per dataset × sequence × chart type.

Crowdsourced Study: causal sequences feel familiar

> *It raises the kinds of questions a journalist or policy analyst would want to explore further.*

Data storyteller (S4), on the journalistic sequence

4 data storytellers and 1 journalism professor
20–30 minute sessions

WHAT WORKED

- Sequences were understandable
- Highlighted examples grounded the story
- 4 of 5 had never seen a PCP, yet understood it
- Journalistic sequence felt like an investigative story

WHAT THEY ASKED FOR

- Less clutter with grouping or highlighting top/bottom 10
- More chart types
- Several datasets in one story
- Injecting domain knowledge
- Adjustable length

Limitations & Future Work

Limitations

Small data
Under a dozen dimensions and a few hundred points

Small study
About 5 responses per condition

Scope
Numerical variables only; single-level conflicts only

Precomputed graphs
Not yet connected to a live causal-inference engine

Future Work

More charts and data
Additional chart types; Not only larger in size but also different types: time-varying and categorical data

Larger studies
More, and more diverse, participants

Words with charts
Natural-language text generated by LLM

New sequences
Other narrative strategies

IEEE VIS 2026 · Boston · TVCG paper

Generating Coherent Visualization Sequences for Multivariate Data by Causal Graph Traversal

More recall than correlation-based or random sequences

Exhaustive, detective, dramatic & journalistic

Each chart sequence follows the causal chain


Puripant Ruchikachorn (Joe)


Darius Coelho


Jun Wang


Kristina Striegnitz


Klaus Mueller

Chulalongkorn University · Stony Brook University · Union College