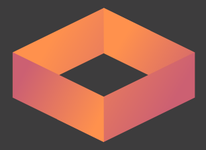


Group 1

Group Project Presentation

Causal Effect of Coughing on Car Accidents

University of Amsterdam - Causal Data Science

DISCUSSION POINTS

OVERVIEW

- Dataset
- Exploratory Data Analysis
- Estimands
- Causal Discovery
- Validation & Sensitivity Analysis
- Conclusion

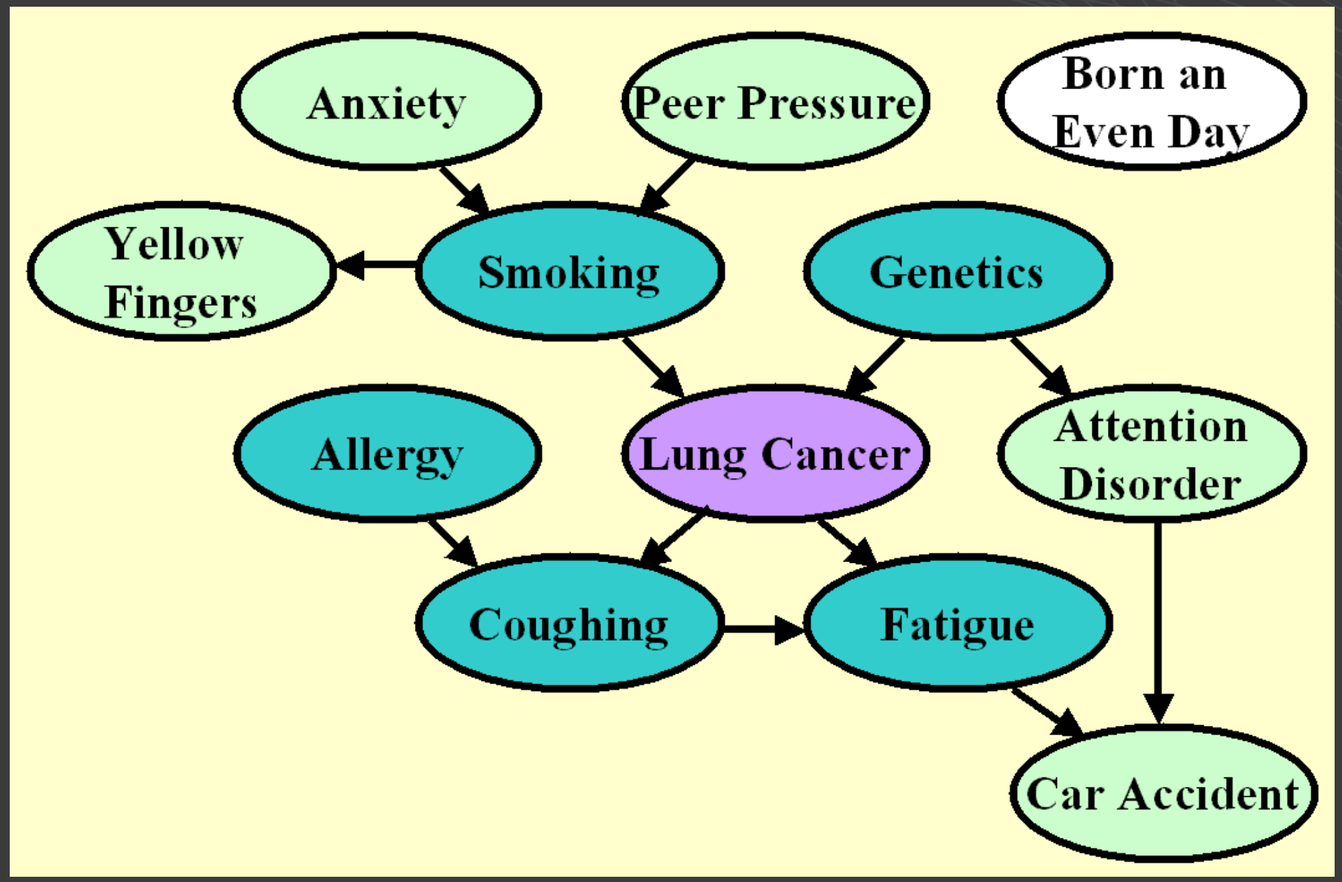
THE DATASET

DESCRIPTION

- LUCAS - LUng CAncer Simple set
  - LUCAS0 was used
- Artificially generated by causal Bayesian Network
- Contains only binary data
- The true causal graph is known
- Probabilities for graph are known as well

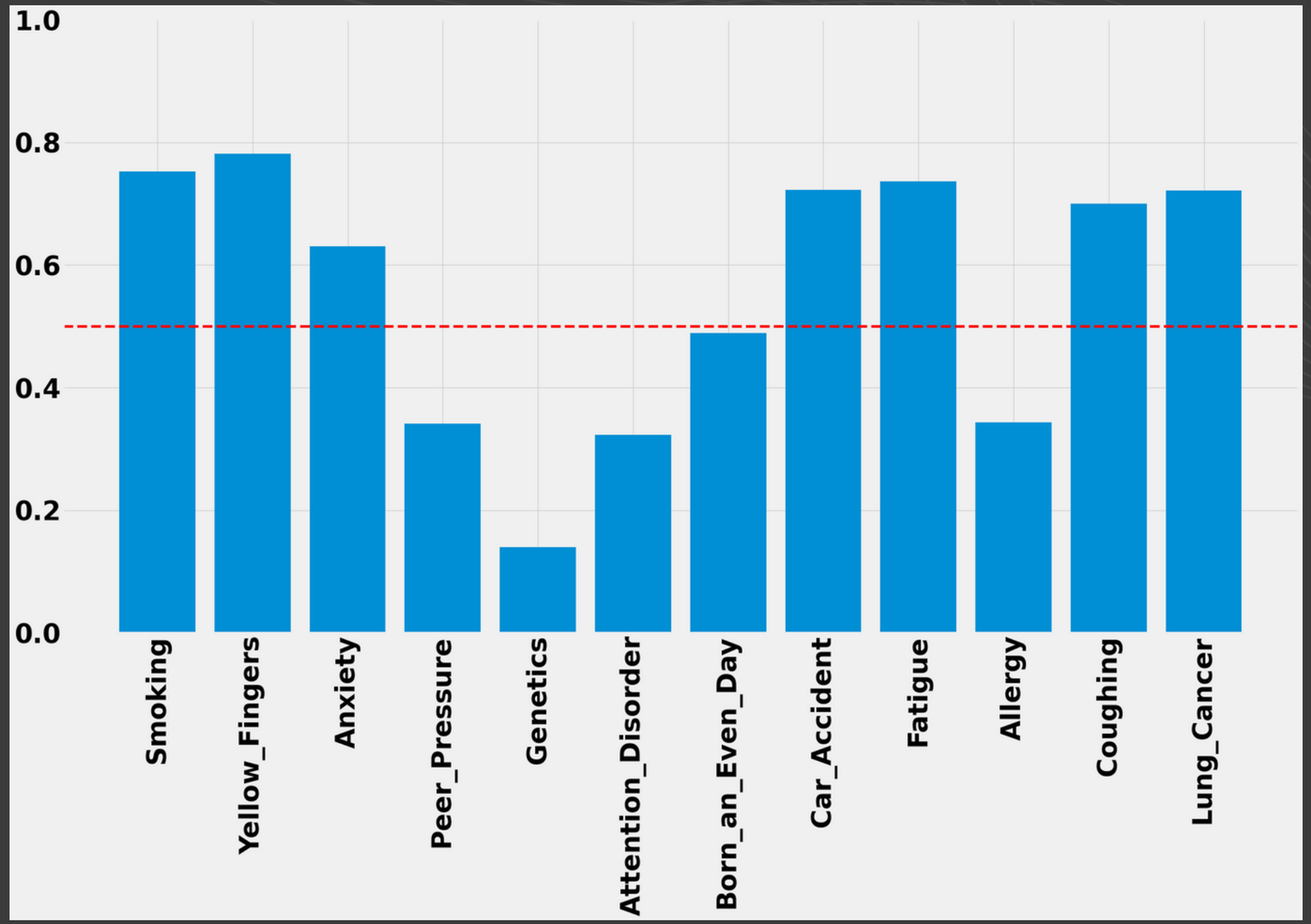
THE DATASET

CAUSAL GRAPH



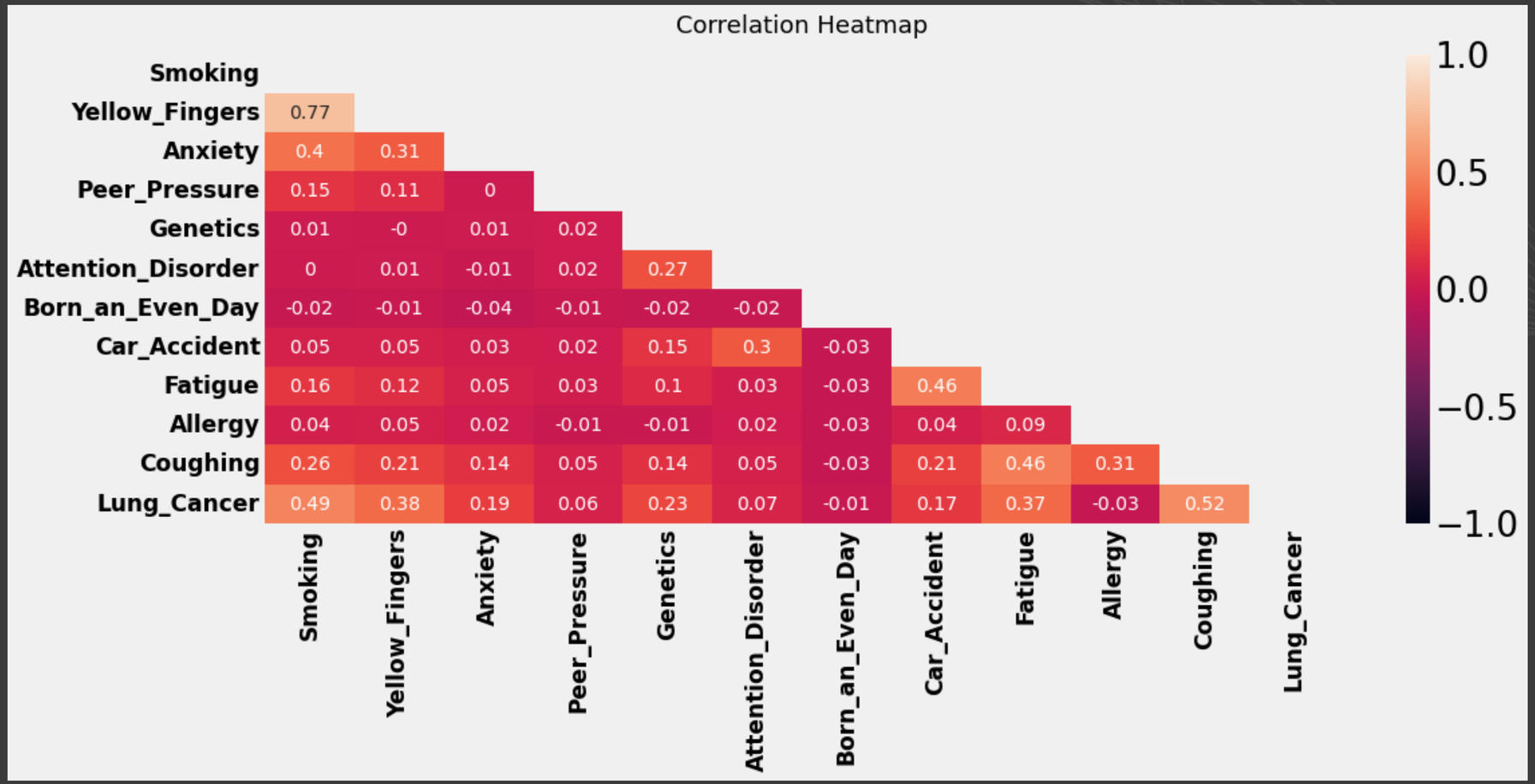
EXPLORATORY DATA ANALYSIS

MEAN VALUES



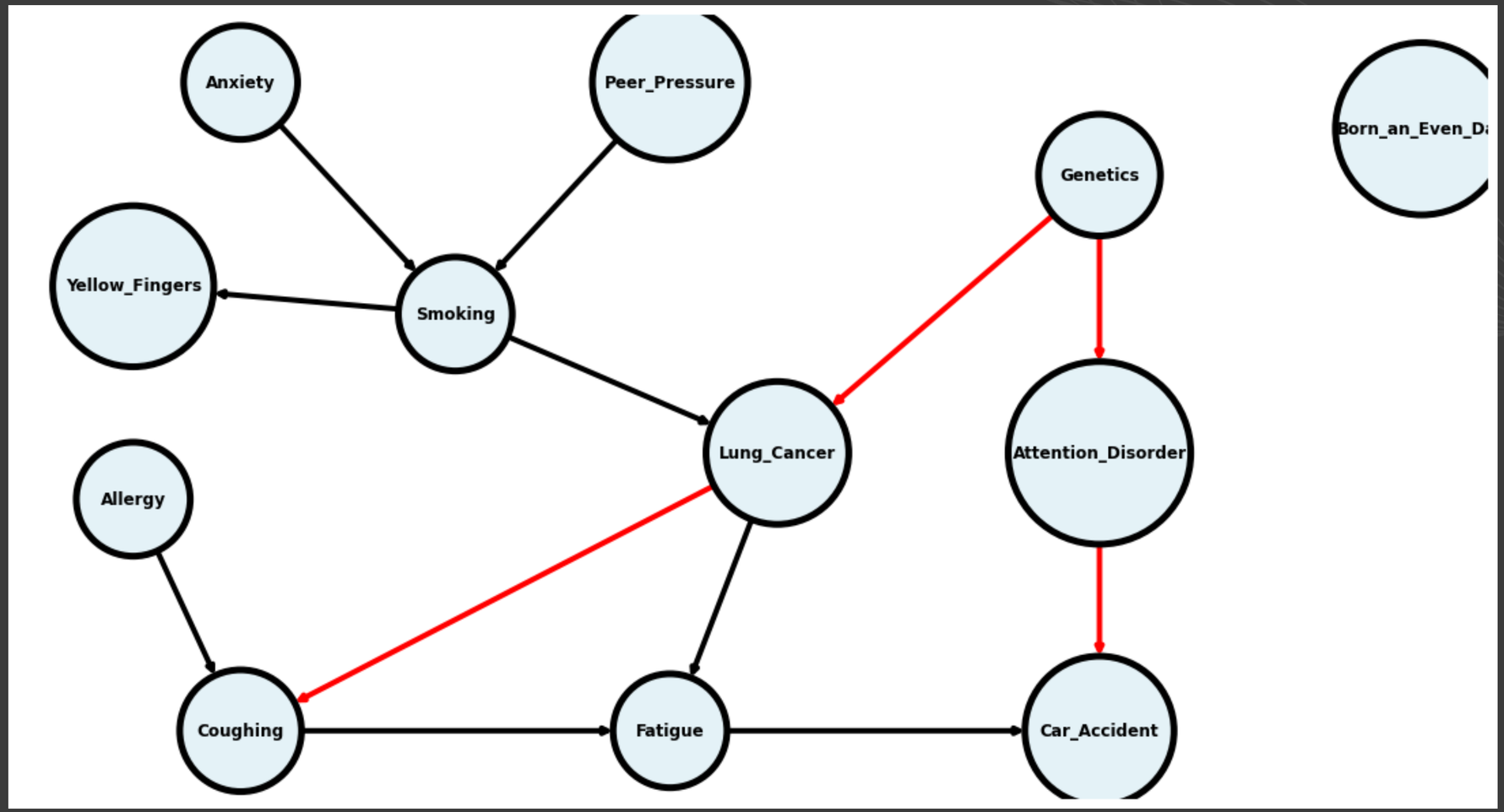
EXPLORATORY DATA ANALYSIS

CORRELATION VALUES



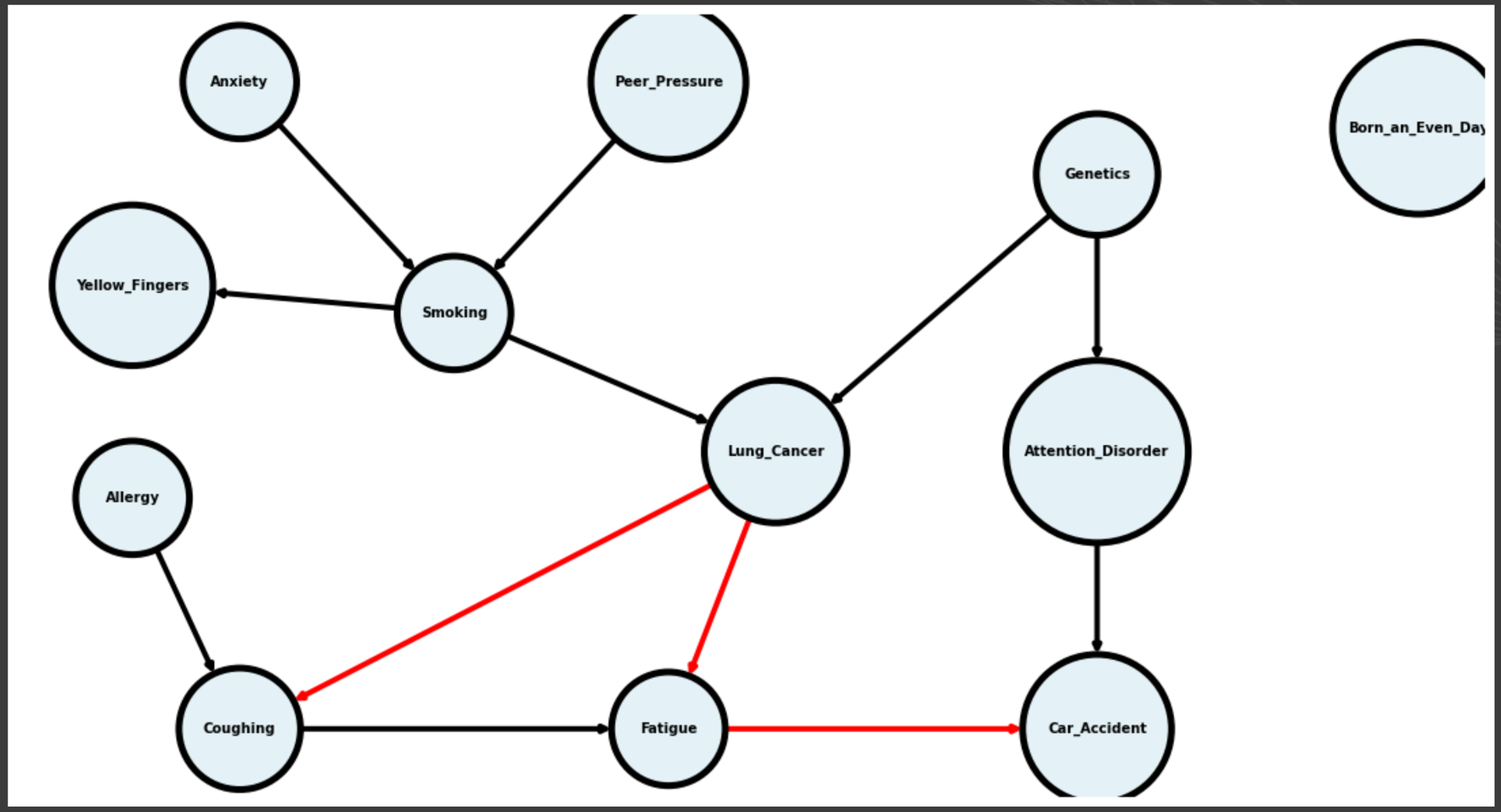
IDENTIFYING ESTIMANDS

BACKDOOR PATHS



IDENTIFYING ESTIMANDS

BACKDOOR PATHS



IDENTIFYING ESTIMANDS

- Frontdoor Paths are apparent
- Instrumental variables are apparent
- However, backdoor path is observed
- Therefore, those two approaches are not important for this work

ESTIMATING CAUSAL EFFECT

LOGISTIC REGRESSION

Car Accident ~ Coughing

Intercept: 0.34
Coefficients: 0.94

Car Accident ~ Coughing + Lung Cancer

Intercept: 0.19
Coefficients: 0.74 & 0.41
Accuracy: 72.3 %

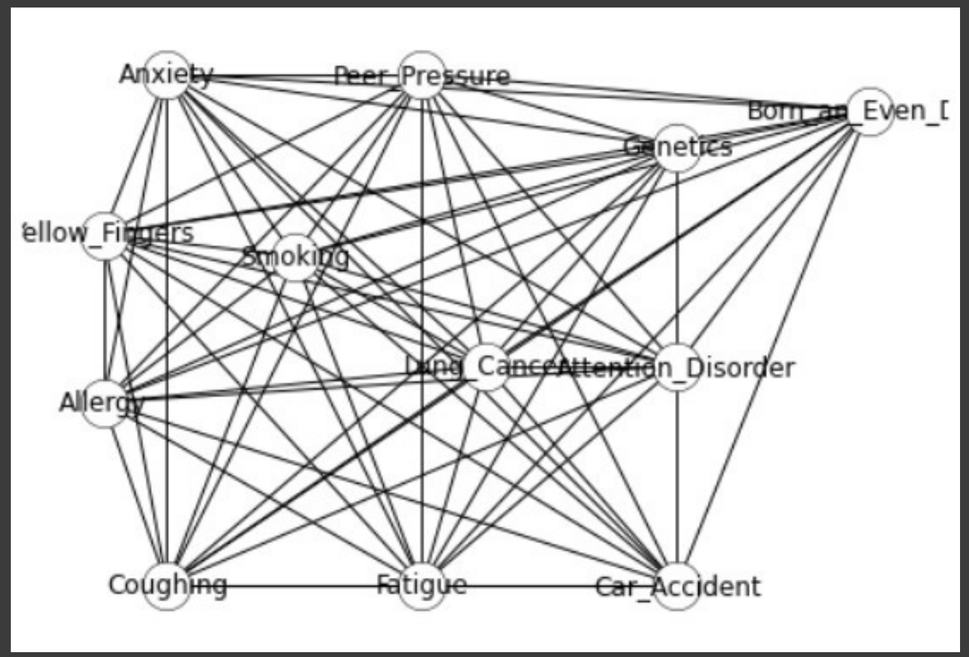
RESULTS

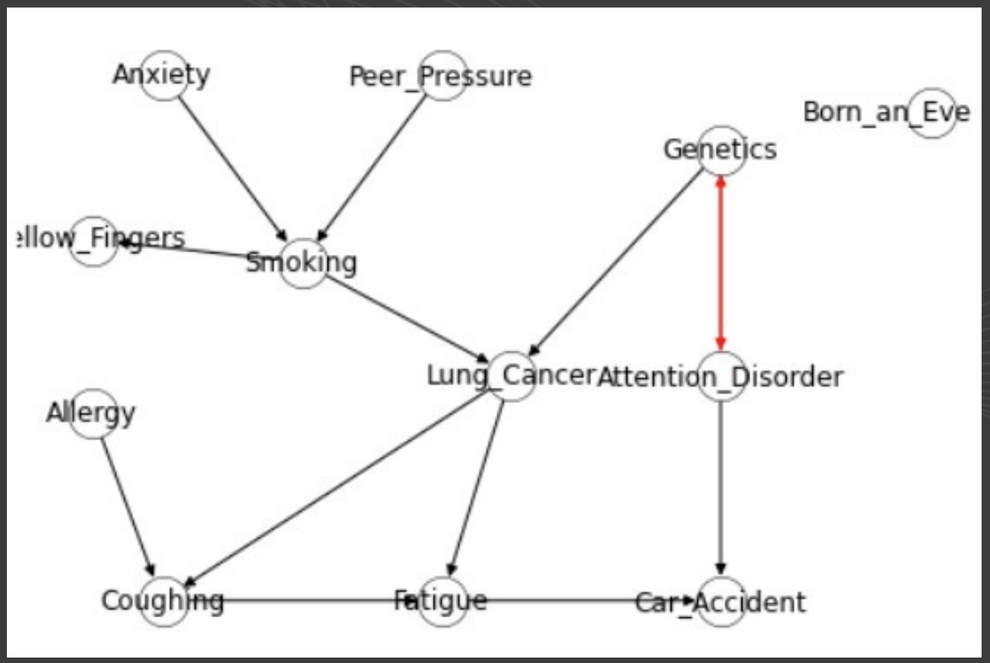
LOGISTIC REGRESSION

Coughing	Lung Cancer	Probability Car Accident
0	0	0.55
0	1	0.64
1	0	0.73
1	1	0.79

CAUSAL DISCOVERY

PC ALGORITHM



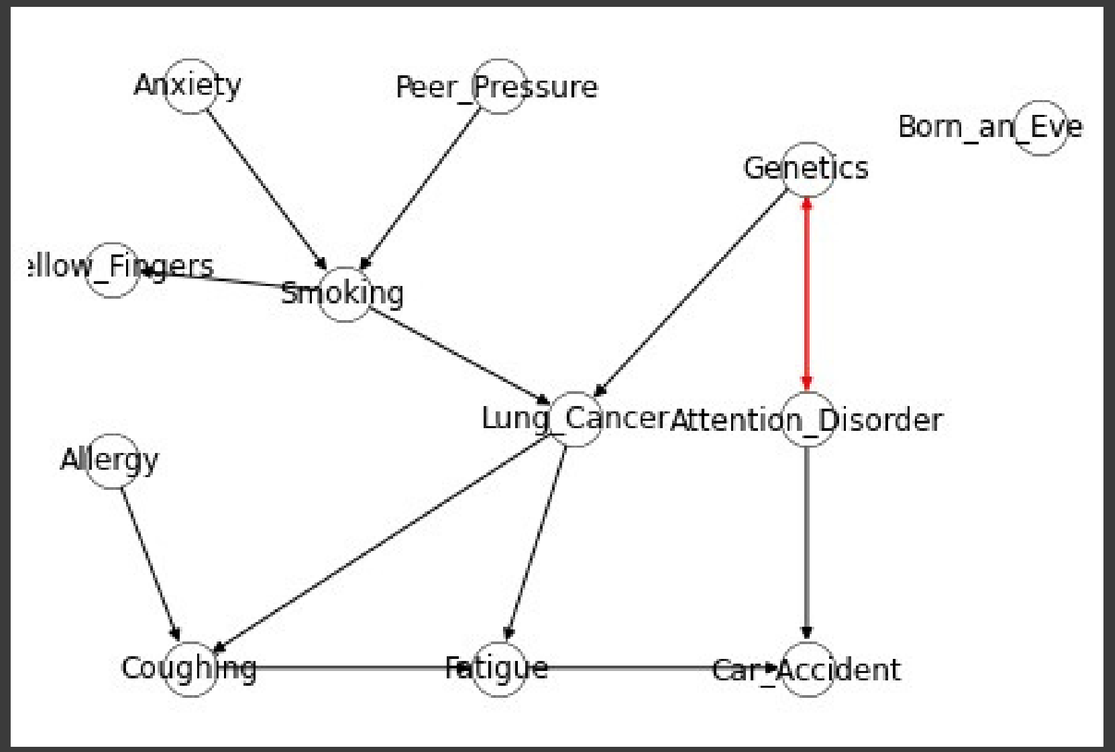


Edge Removal: Pingouin's partial correlation tests (α = 0.01)

Edge Direction: Using V-structures and independences found during edge removal

CAUSAL DISCOVERY

GES ALGORITHM



- GES algorithm is a score-based method (requiring the choice of a score function)
- Does not depend on conditional independence tests.
- GES always operates in the space of essential graphs
- Returns a valid essential graph (CPDAG)
  - Exactly same as resulting graph in PC algorithm
  - The edge between Genetics and Attention disorder is not directed

VALIDATION

SENSITIVITY ANALYSIS

Add a Random Common Cause	Add an Unobserved Common Cause	Use a subset of data
Estimated effect:0.1579 New effect:0.1578 p value:0.4599	Estimated effect: 0.15786 New effect: 0.1024	Estimated effect: 0.158 New effect: 0.1599 p value: 0.47

CONCLUSION

- Causal problem: Coughing → Car Accidents
- Valid adjustment set: {Lung Cancer}
- Estimated causal effect: ~0.16
- Coughing does increase the chance of having a car accident by around 16%
- PC and GES algorithm found the same causal graph

THANK YOU