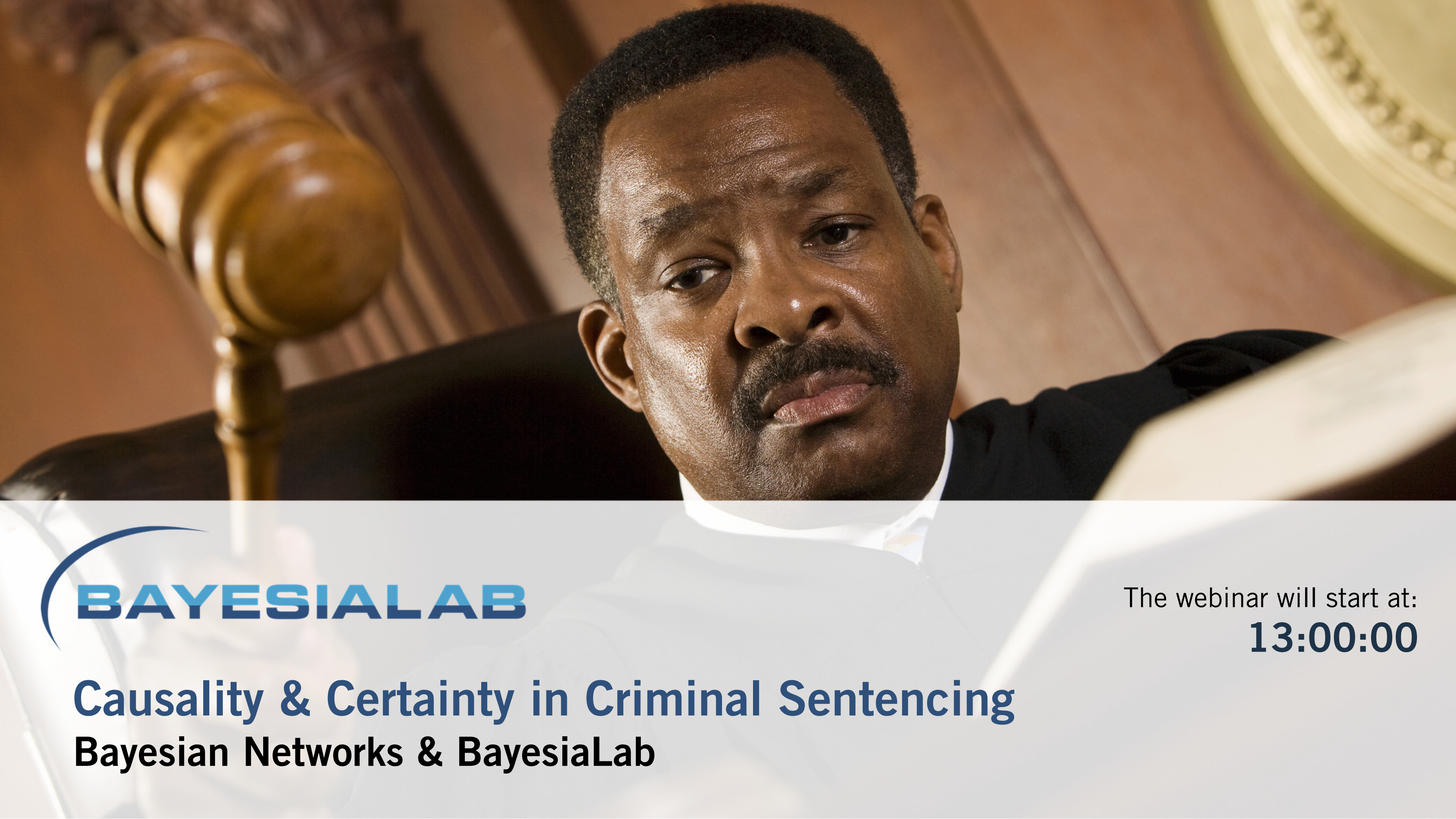


The webinar will start at:
13:00:00

Causality & Certainty in Criminal Sentencing

Bayesian Networks & BayesiaLab

Today's Program

Introduction

- Our Team & Our Company

Motivation & Background

- Disparity in Criminal Sentencing
- Federal Sentencing Guidelines
- Legislation & Supreme Court Rulings

20 min.

Analyzing Federal Sentencing Data

- Certainty: Information Theory & Entropy
- Fairness: Causality

40 min.

Your BayesiaLab Team Today

stefan.conrady@bayesia.us

stacey.blodgett@bayesia.us

clare.gora@bayesia.us

Disambiguation

BAYESIA
Your Decision Partner

Our Company

BAYESIALAB

Our Product

The Paradigm

BAYESIAN NETWORKS*

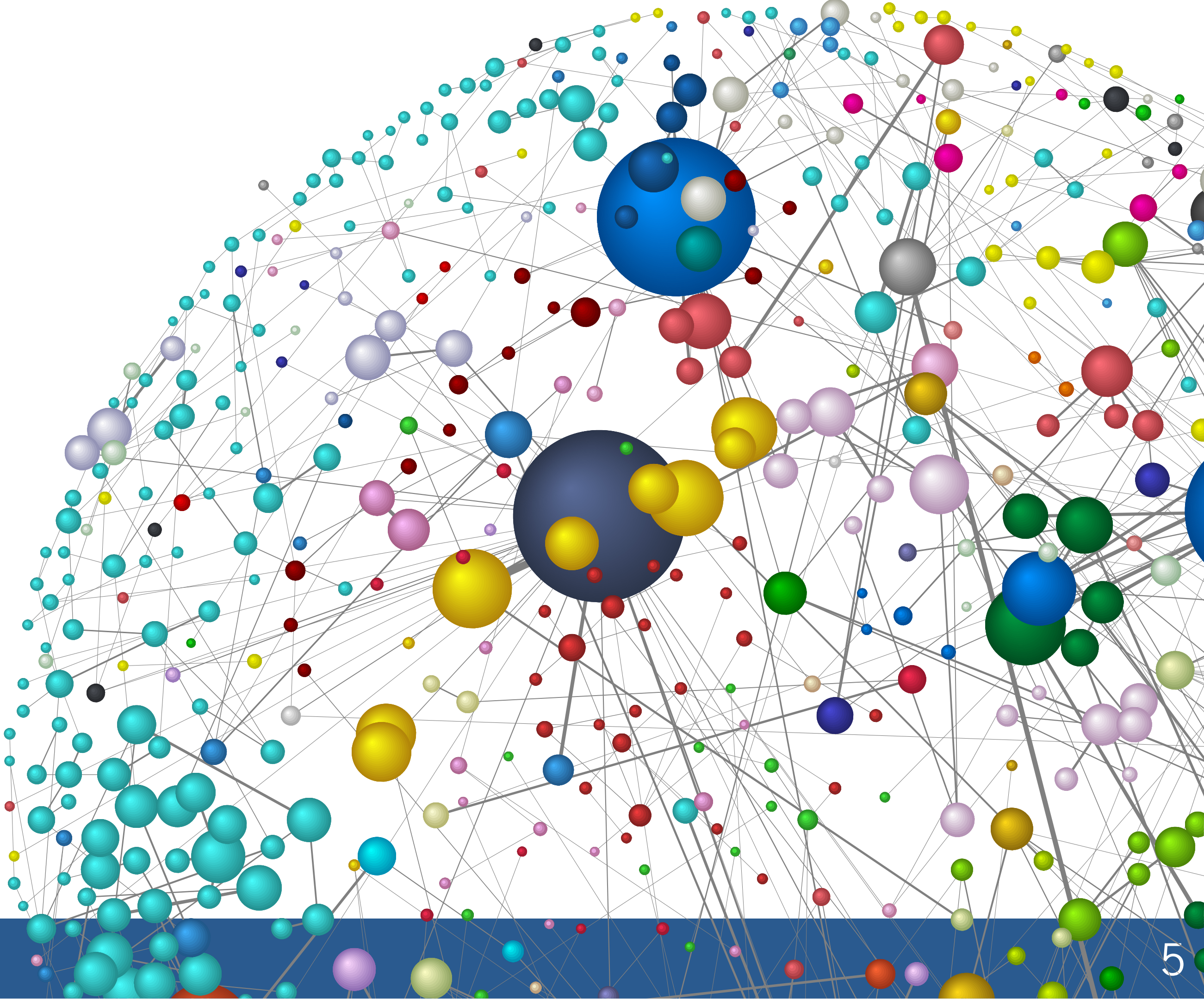
Judea Pearl
Cognitive Systems Laboratory
Computer Science Department
University of California, Los Angeles, CA 90024
judea@cs.ucla.edu

Bayesian networks were developed in the late 1970's to model distributed processing in reading comprehension, where both semantical expectations and perceptual evidence must be combined to form a coherent interpretation. The ability to coordinate bi-directional inferences filled a void in expert systems technology of the early 1980's, and Bayesian networks have emerged as a general representation scheme for uncertain knowledge [Pearl, 1988, Heckerman *et al.*, 1995, Jensen, 1996, Castillo *et al.*, 1997].

Bayesian networks are directed acyclic graphs (DAGs) in which the nodes represent vari-

BAYESIA
Your Decision Partner

Co-founded in 2001
by Dr. Lionel Jouffe &
Dr. Paul Munteanu



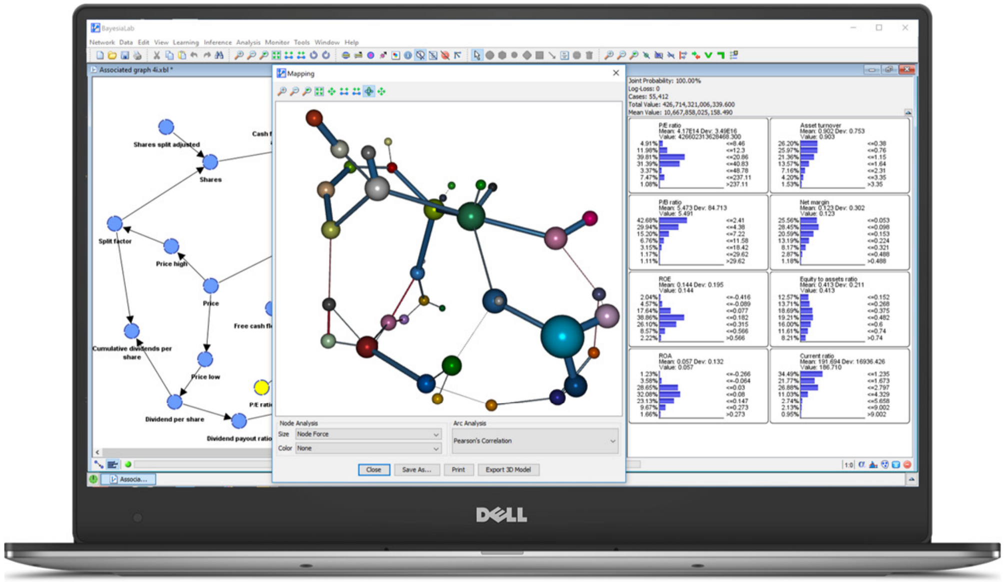
BAYESIALAB

Bayesian Networks & BayesiaLab

A Practical Introduction for Researchers

- Free download:
  www.bayesia.com/book
- Hardcopy available on Amazon:
  http://amzn.com/0996533303

BAYESIALAB

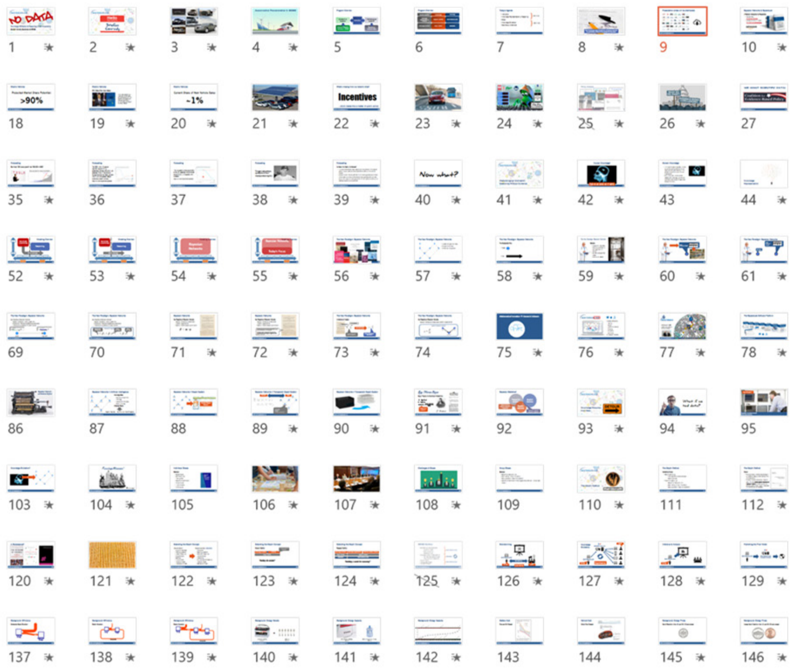


A desktop software for:

- encoding
- learning
- editing
- performing inference
- analyzing
- simulating
- optimizing

with Bayesian networks.

Webinar Slides, Data, and Recording Available

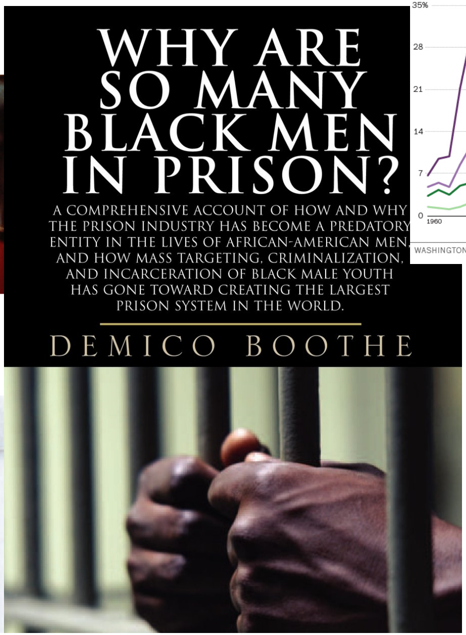


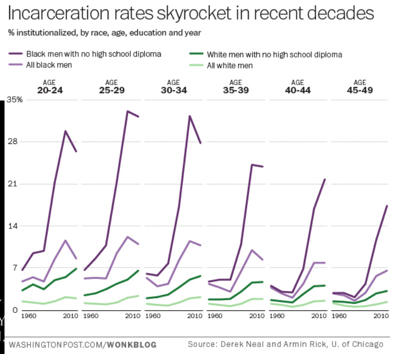
Caveat

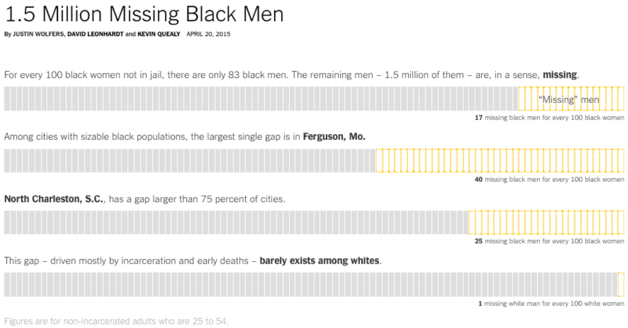
- In this webinar, we focus exclusively on methodological questions and do not endeavor to arrive a substantive findings.
- While we work with real data, any effect size estimates are for illustration purposes only and should be considered fictional.

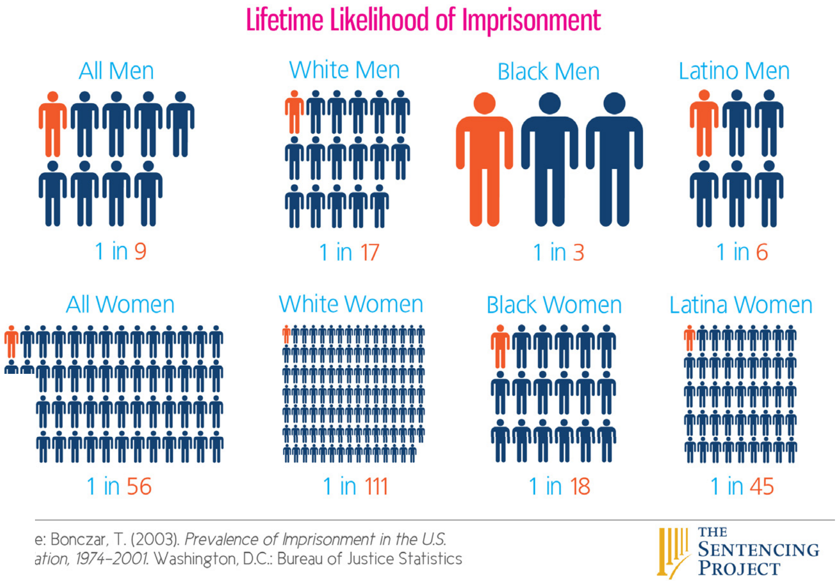
Background & Motivation

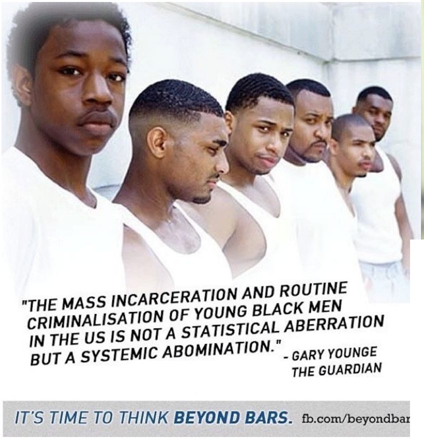

BLACK VOICES 10/04/2013 03:24 pm ET | Updated Oct 04, 2013
1 In 3 Black Males Will Go To Prison In Their Lifetime, Report Warns
By Saki Knafo
GETTY

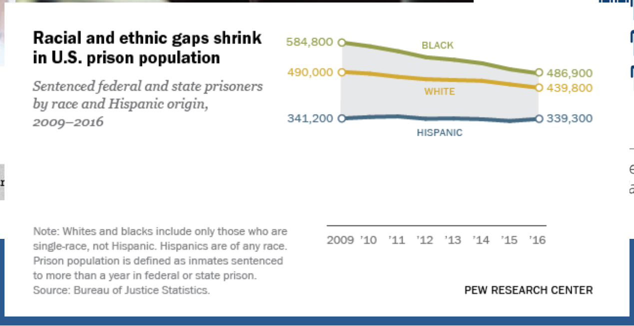

WHY ARE SO MANY BLACK MEN IN PRISON?
A COMPREHENSIVE ACCOUNT OF HOW AND WHY THE PRISON INDUSTRY HAS BECOME A PREDATORY ENTITY IN THE LIVES OF AFRICAN-AMERICAN MEN AND HOW MASS TARGETING, CRIMINALIZATION, AND INCARCERATION OF BLACK MALE YOUTH HAS GONE TOWARD CREATING THE LARGEST PRISON SYSTEM IN THE WORLD.
DEMICO BOOTHE


Incarceration rates skyrocket in recent decades
% institutionalized, by race, age, education and year
Black men with no high school diploma
All black men
White men with no high school diploma
All white men
AGE 20-24
AGE 25-29
AGE 30-34
AGE 35-39
AGE 40-44
AGE 45-49
35%
28
21
14
7
0
1960
2010
WASHINGTONPOST.COM/WONKBLOG
Source: Derek Neal and Armin Rick, U. of Chicago


1.5 Million Missing Black Men
By JUSTIN WOLFERS, DAVID LEONHARDT and KEVIN QUEALY APRIL 20, 2015
For every 100 black women not in jail, there are only 83 black men. The remaining men – 1.5 million of them – are, in a sense, missing.
"Missing" men
17 missing black men for every 100 black women
Among cities with sizable black populations, the largest single gap is in Ferguson, Mo.
40 missing black men for every 100 black women
North Charleston, S.C., has a gap larger than 75 percent of cities.
25 missing black men for every 100 black women
This gap – driven mostly by incarceration and early deaths – barely exists among whites.
1 missing white man for every 100 white women
Figures are for non-incarcerated adults who are 25 to 54.


Lifetime Likelihood of Imprisonment
All Men
1 in 9
White Men
1 in 17
Black Men
1 in 3
Latino Men
1 in 6
All Women
1 in 56
White Women
1 in 111
Black Women
1 in 18
Latina Women
1 in 45
e: Bonczar, T. (2003). Prevalence of Imprisonment in the U.S.
ation, 1974-2001. Washington, D.C.: Bureau of Justice Statistics
THE SENTENCING PROJECT


"THE MASS INCARCERATION AND ROUTINE CRIMINALISATION OF YOUNG BLACK MEN IN THE US IS NOT A STATISTICAL ABERRATION BUT A SYSTEMIC ABOMINATION." - GARY YOUNGE THE GUARDIAN
IT'S TIME TO THINK BEYOND BARS. fb.com/beyondbar

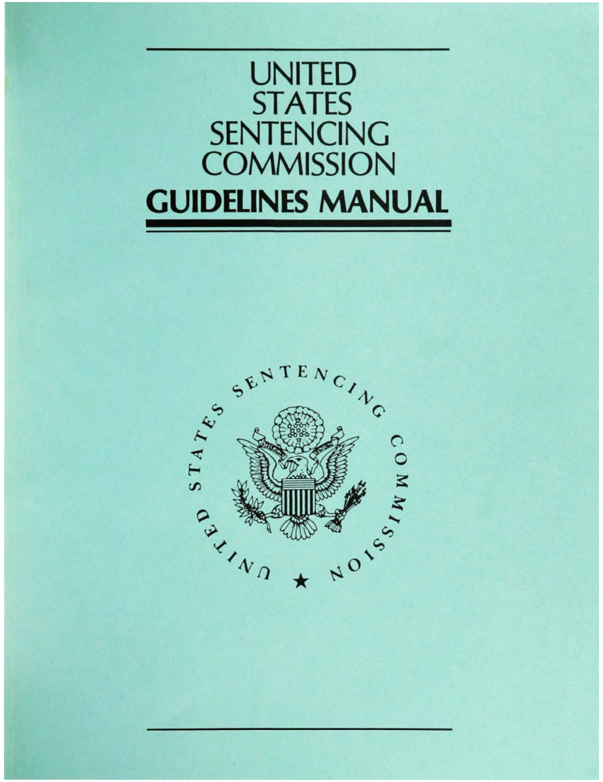

Racial and ethnic gaps shrink in U.S. prison population
Sentenced federal and state prisoners by race and Hispanic origin, 2009–2016
584,800
BLACK
486,900
490,000
WHITE
439,800
341,200
HISPANIC
339,300
2009 '10 '11 '12 '13 '14 '15 '16
Note: Whites and blacks include only those who are single-race, not Hispanic. Hispanics are of any race. Prison population is defined as inmates sentenced to more than a year in federal or state prison.
Source: Bureau of Justice Statistics.
PEW RESEARCH CENTER


Too much discretion?

Background

Federal Sentencing Guidelines

- The Guidelines are the product of the United States Sentencing Commission, which was created by the Sentencing Reform Act of 1984. The Guidelines' primary goal was to alleviate sentencing disparities that research had indicated were prevalent in the existing sentencing system, and the guidelines reform was specifically intended to provide for **determinate sentencing**.



Background

Federal Sentencing Guidelines

Offense Seriousness

➔ Base Level Offense

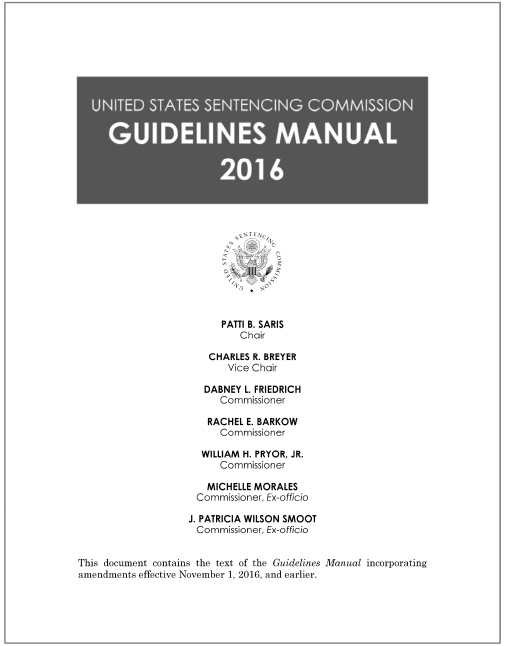
± Specific Offense Characteristics

± Adjustments

+ Criminal History

= Final Offense Level

➔ Sentence Range


UNITED STATES SENTENCING COMMISSION
GUIDELINES MANUAL
2016

PATTI B. SARIS
Chair

CHARLES R. BREYER
Vice Chair

DABNEY L. FRIEDRICH
Commissioner

RACHEL E. BARKOW
Commissioner

WILLIAM H. PRYOR, JR.
Commissioner

MICHELLE MORALES
Commissioner, *Ex-officio*

J. PATRICIA WILSON SMOOT
Commissioner, *Ex-officio*

This document contains the text of the *Guidelines Manual* incorporating amendments effective November 1, 2016, and earlier.

PART D — OFFENSES INVOLVING DRUGS AND NARCO-TERRORISM

Historical Note	Effective November 1, 1987. Amended effective November 1, 2007 (amendment 711).

1. UNLAWFUL MANUFACTURING, IMPORTING, EXPORTING, TRAFFICKING, OR POSSESSION; CONTINUING CRIMINAL ENTERPRISE

§2D1.1. Unlawful Manufacturing, Importing, Exporting, or Trafficking (Including Possession with Intent to Commit These Offenses); Attempt or Conspiracy

(a) Base Offense Level (Apply the greatest):

(1) **43**, if the defendant is convicted under 21 U.S.C. § 841(b)(1)(A), (b)(1)(B), or (b)(1)(C), or 21 U.S.C. § 960(b)(1), (b)(2), or (b)(3), and the offense of conviction establishes that death or serious bodily injury resulted from the use of the substance and that the defendant committed the offense after one or more prior convictions for a similar offense; or

(2) **38**, if the defendant is convicted under 21 U.S.C. § 841(b)(1)(A), (b)(1)(B), or (b)(1)(C), or 21 U.S.C. § 960(b)(1), (b)(2), or (b)(3), and the offense of conviction establishes that death or serious bodily injury resulted from the use of the substance; or

(3) **30**, if the defendant is convicted under 21 U.S.C. § 841(b)(1)(E) or 21 U.S.C. § 960(b)(5), and the offense of conviction establishes that death or serious bodily injury resulted from the use of the substance and that the defendant committed the offense after one or more prior convictions for a similar offense; or

(4) **26**, if the defendant is convicted under 21 U.S.C. § 841(b)(1)(E) or 21 U.S.C. § 960(b)(5), and the offense of conviction establishes that death or serious bodily injury resulted from the use of the substance; or

(5) the offense level specified in the Drug Quantity Table set forth in subsection (c), except that if (A) the defendant receives an adjustment under §3B1.2 (Mitigating Role); and (B) the base offense level under subsection (c) is (i) level **32**, decrease by **2** levels; (ii) level **34** or level **36**, decrease by **3** levels; or (iii) level **38**, decrease by **4** levels. If the resulting offense level

SENTENCING TABLE
(in months of imprisonment)

	Offense Level	Criminal History Category (Criminal History Points) I (0 or 1)	II (2 or 3)	III (4, 5, 6)	IV (7, 8, 9)	V (10, 11, 12)	VI (13 or more)
Zone A	1	0–6	0–6	0–6	0–6	0–6	0–6
	2	0–6	0–6	0–6	0–6	0–6	1–7
	3	0–6	0–6	0–6	0–6	2–8	3–9
	4	0–6	0–6	0–6	2–8	4–10	6–12
	5	0–6	0–6	1–7	4–10	6–12	9–15
	6	0–6	1–7	2–8	6–12	9–15	12–18
	7	0–6	2–8	4–10	8–14	12–18	15–21
	8	0–6	4–10	6–12	10–16	15–21	18–24
Zone B	9	4–10	6–12	8–14	12–18	18–24	21–27
	10	6–12	8–14	10–16	15–21	21–27	24–30
	11	8–14	10–16	12–18	18–24	24–30	27–33
Zone C	12	10–16	12–18	15–21	21–27	27–33	30–37
	13	12–18	15–21	18–24	24–30	30–37	33–41
Zone D	14	15–21	18–24	21–27	27–33	33–41	37–46
	15	18–24	21–27	24–30	30–37	37–46	41–51
	16	21–27	24–30	27–33	33–41	41–51	46–57
	17	24–30	27–33	30–37	37–46	46–57	51–63
	18	27–33	30–37	33–41	41–51	51–63	57–71
	19	30–37	33–41	37–46	46–57	57–71	63–78
	20	33–41	37–46	41–51	51–63	63–78	70–87
	21	37–46	41–51	46–57	57–71	70–87	77–96
	22	41–51	46–57	51–63	63–78	77–96	84–105
	23	46–57	51–63	57–71	70–87	84–105	92–115
	24	51–63	57–71	63–78	77–96	92–115	100–125
	25	57–71	63–78	70–87	84–105	100–125	110–137
	26	63–78	70–87	78–97	92–115	110–137	120–150
	27	70–87	78–97	87–108	100–125	120–150	130–162
	28	78–97	87–108	97–121	110–137	130–162	140–175
	29	87–108	97–121	108–135	121–151	140–175	151–188
	30	97–121	108–135	121–151	135–168	151–188	168–210
	31	108–135	121–151	135–168	151–188	168–210	188–235
	32	121–151	135–168	151–188	168–210	188–235	210–262
	33	135–168	151–188	168–210	188–235	210–262	235–293
	34	151–188	168–210	188–235	210–262	235–293	262–327
	35	168–210	188–235	210–262	235–293	262–327	292–365
	36	188–235	210–262	235–293	262–327	292–365	324–405
	37	210–262	235–293	262–327	292–365	324–405	360–life
	38	235–293	262–327	292–365	324–405	360–life	360–life
	39	262–327	292–365	324–405	360–life	360–life	360–life
	40	292–365	324–405	360–life	360–life	360–life	360–life
	41	324–405	360–life	360–life	360–life	360–life	360–life
	42	360–life	360–life	360–life	360–life	360–life	360–life
	43	life	life	life	life	life	life

Background

Guidelines are "advisory" not mandatory

(Bench Opinion) OCTOBER TERM, 2004 1

Syllabus

NOTE: Where it is feasible, a syllabus (headnote) will be released, as is being done in connection with this case, at the time the opinion is issued. The syllabus constitutes no part of the opinion of the Court but has been prepared by the Reporter of Decisions for the convenience of the reader. See *United States* v. *Detroit Timber & Lumber Co.*, 200 U. S. 321, 337.

SUPREME COURT OF THE UNITED STATES

Syllabus

UNITED STATES *v.* BOOKER

CERTIORARI TO THE UNITED STATES COURT OF APPEALS FOR THE SEVENTH CIRCUIT

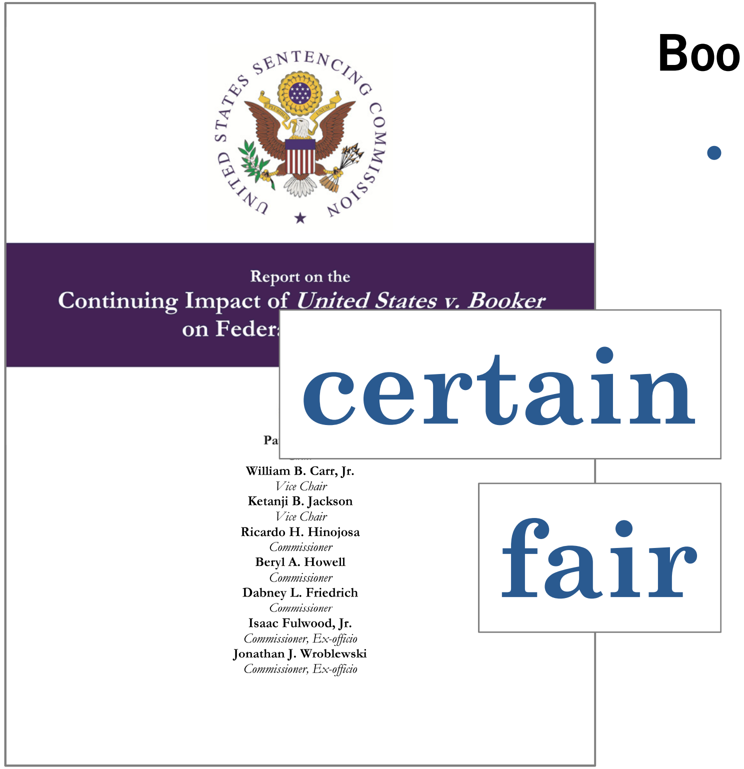
No. 04–104. Argued October 4, 2004—Decided January 12, 2005*

Under the Federal Sentencing Guidelines, the sentence authorized by the jury verdict in respondent Booker's drug case was 210-to-262

Key Events	Congress & USSC	Supreme Court	Congress	Supreme Court	Supreme Court
	Sentencing Guidelines 1987	Koon 1996	PROTECT Act 2003	Booker 2005	Gall 2007

Judiciary Discretion

Research Questions



Booker Report

- The United States Sentencing Commission ("the Commission") submits this report to Congress[1] on the impact of United States v. Booker[2] on federal sentencing in order to assist Congress in its efforts to ensure **certain and fair sentencing** that avoids unwarranted sentencing disparities while maintaining sufficient flexibility,[3] as envisioned in the Sentencing Reform Act of 1984 ("SRA").

Research Questions

- How can we translate “certain” and “fair” into measurable quantities that can be estimated from data?

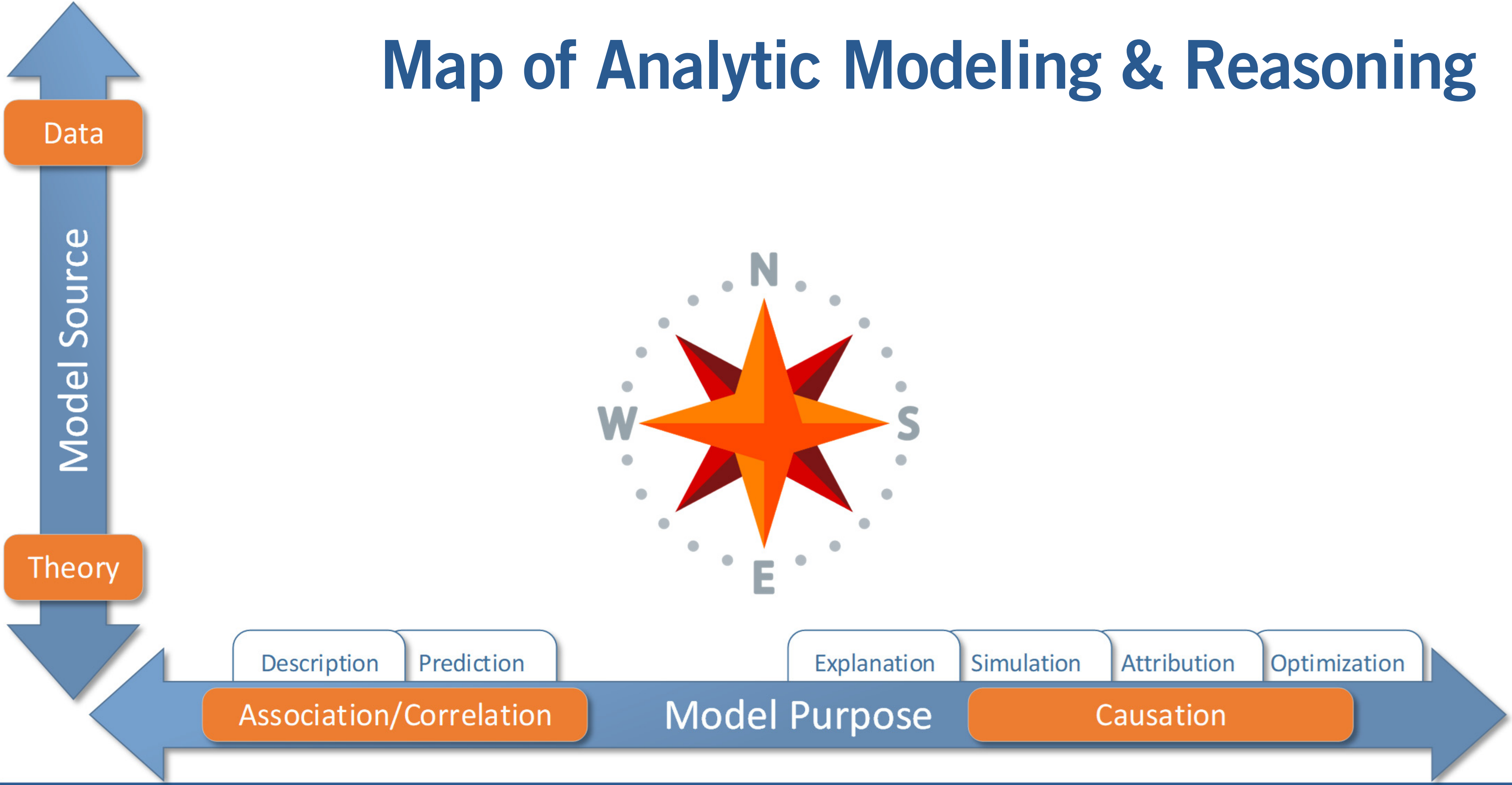
“Certain”

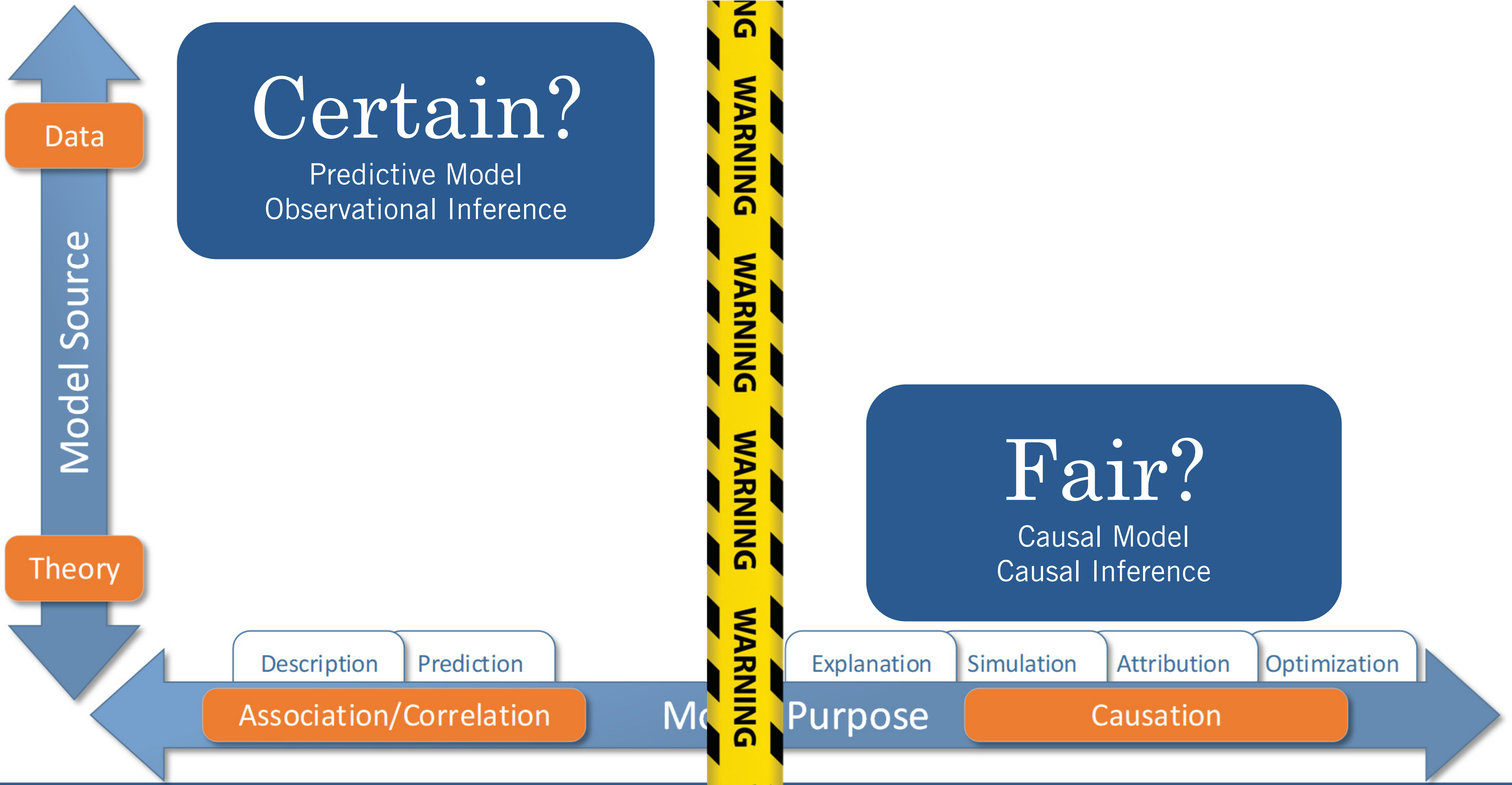
- We quantify uncertainty by calculating the entropy of the distributions of sentences.

“Fair”

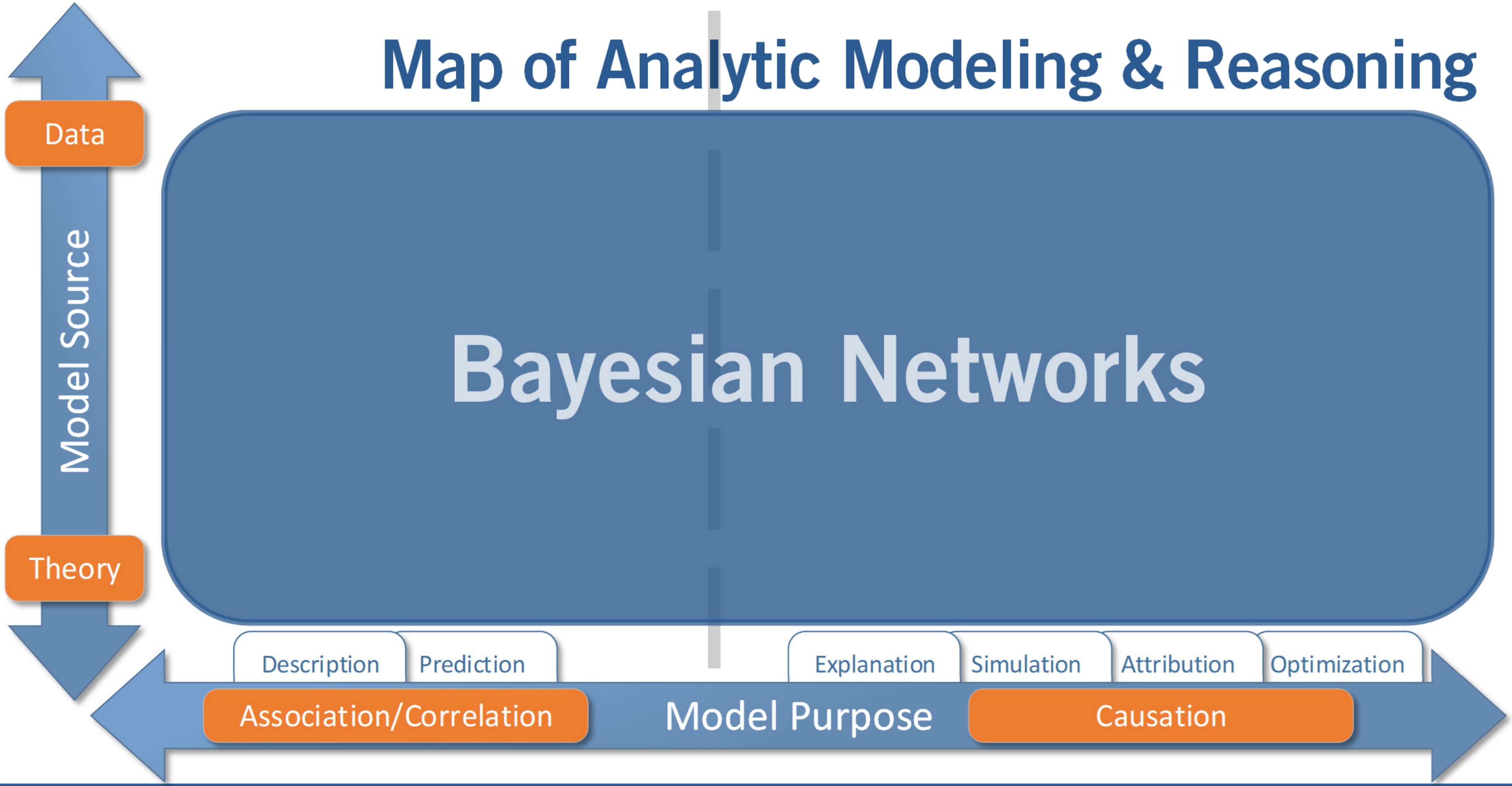
- We consider “fair” as the absence of extra-legal causes in sentencing. Thus, we need to estimate the potential effects of extra-legal causes.

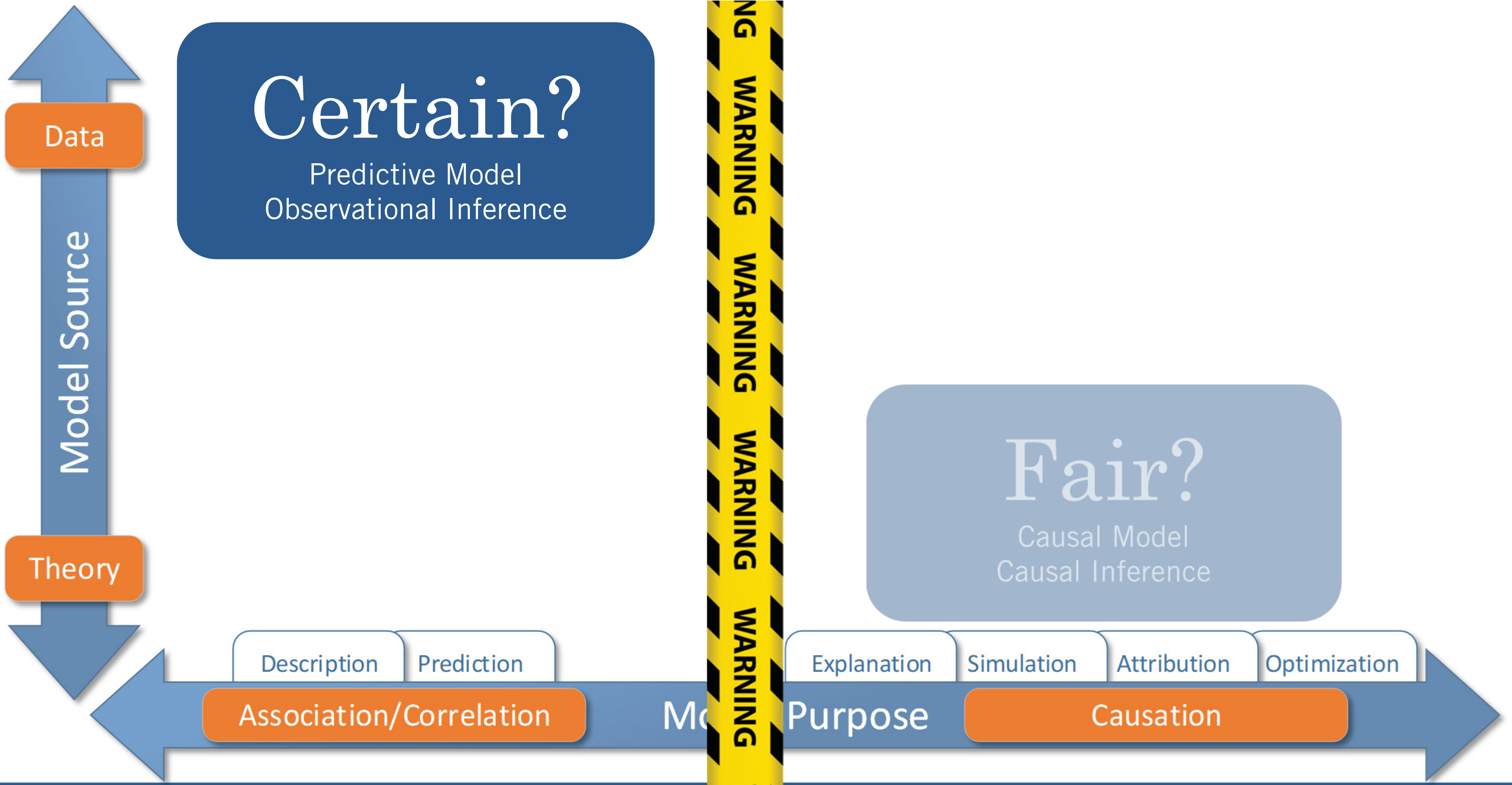
Map of Analytic Modeling & Reasoning




Certain?
Predictive Model
Observational Inference
Fair?
Causal Model
Causal Inference
Data
Model Source
Theory
Description
Prediction
Explanation
Simulation
Attribution
Optimization
Association/Correlation
Purpose
Causation
WARNING
WARNING
WARNING
WARNING

Map of Analytic Modeling & Reasoning




Certain?
Predictive Model
Observational Inference
Data
Model Source
Theory
Fair?
Causal Model
Causal Inference
Description
Prediction
Explanation
Simulation
Attribution
Optimization
Association/Correlation
Purpose
Causation
WARNING

UNITED STATES
SENTENCING COMMISSION

GUIDELINES | RESEARCH | POLICYMAKING | EDUCATION | ABOUT | BY TOPIC

Individual Offender Datafiles

The Commission's individual datafiles provide information on the sentences imposed in cases involving individuals. In these datafiles the individual is the unit of analysis. These datafiles do **not** contain information from the Commission's Organizational, Resentencing, or Appeals Databases.

Datafiles (downloads are SAS and SPSS compatible)

- Fiscal Year 2017
- Fiscal Year 2016
- Fiscal Year 2015
- Fiscal Year 2014
- Fiscal Year 2013
- Fiscal Year 2012
- Fiscal Year 2011
- Fiscal Year 2010
- Fiscal Year 2009
- Fiscal Year 2008
- Fiscal Year 2007
- Fiscal Year 2006
- Fiscal Year 2005
- Fiscal Year 2004
- Fiscal Year 2003
- Fiscal Year 2002

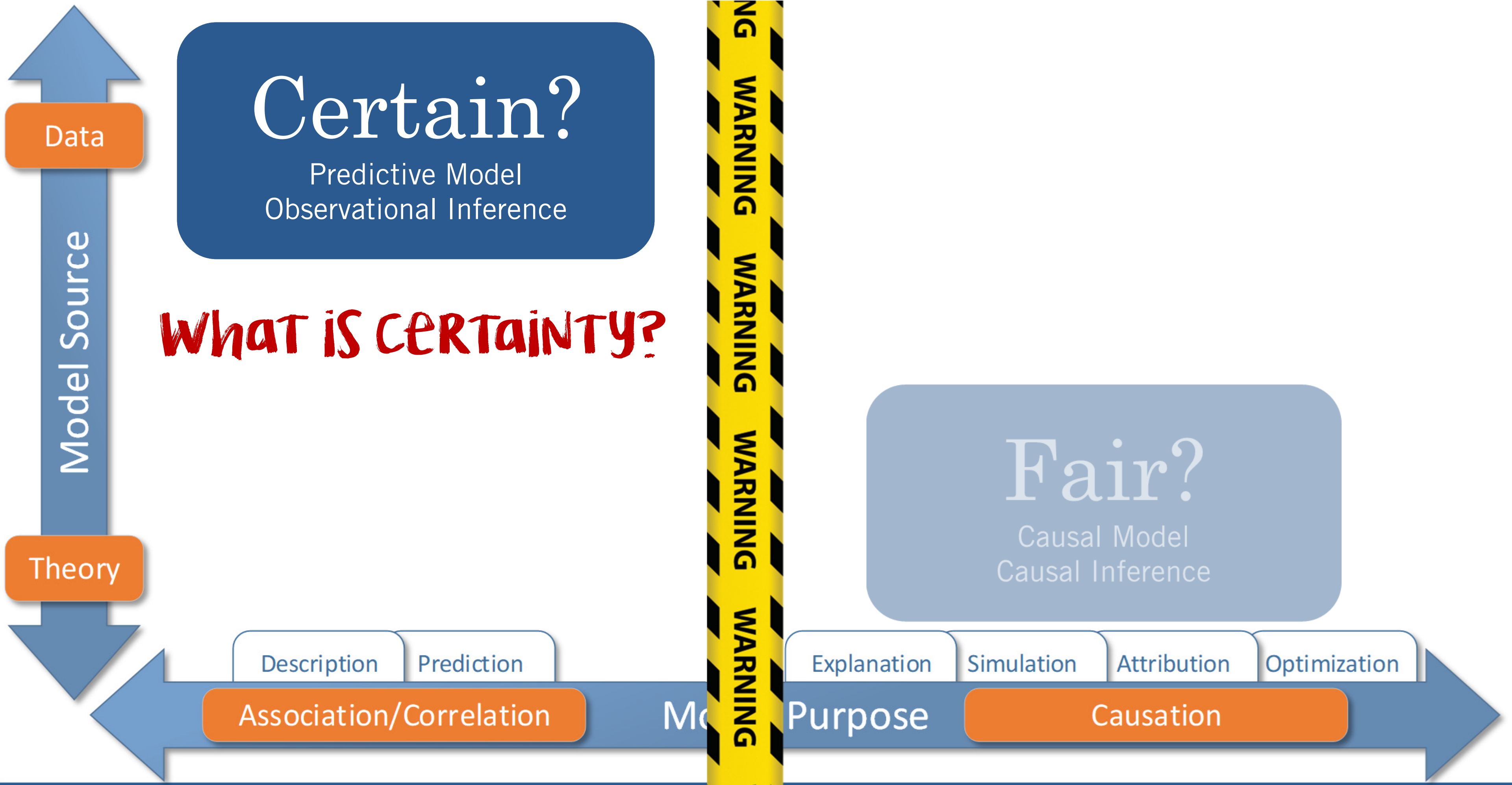
Study Dataset: USSC Booker Report

Federal Criminal Sentences

319,172 Observations 1996-2011

- accap: Armed Career Criminal Status
- age: Age
- aggdum: Aggravated Role Adjustment
- booker2: Sentence vs. Guideline Range
- career: Career Offender Impact on Final Offense Level
- caroffap: Career Offender Status Applied
- circdist: Circuit District
- mitdum: Mitigating Role Application
- moncirc: Judicial Circuit
- monsex: Gender
- newcit: Citizenship
- newcnvtn: Plea or Trial
- neweduc: Education
- newrace: Race
- offtype2: Primary Offense Type **(Drug Trafficking Only)**
- period: Time Period **(Koon/PROTECT/Booker/Gall Only)**
- primary: Primary Drug Type
- quarter: Fiscal Year Quarter
- sa: Sentence vs. Guideline Range (Expanded)
- safevalve: Safety Valve Provision
- sentimp: Sentence Type
- totchpts: Criminal History Points Applied
- weapon: SOC Weapon Enhancement
- xcrhissr: Defendant's Final Criminal History Category (I-VI)
- glmin: Trumped Guideline Minimum
- fy: Fiscal Year
- usscidn: USSC ID
- sensplt0: Trumped Total Prison Sentence
- loss_2b: Dollar Amount of Loss
- gdl: Chapter Two Guideline

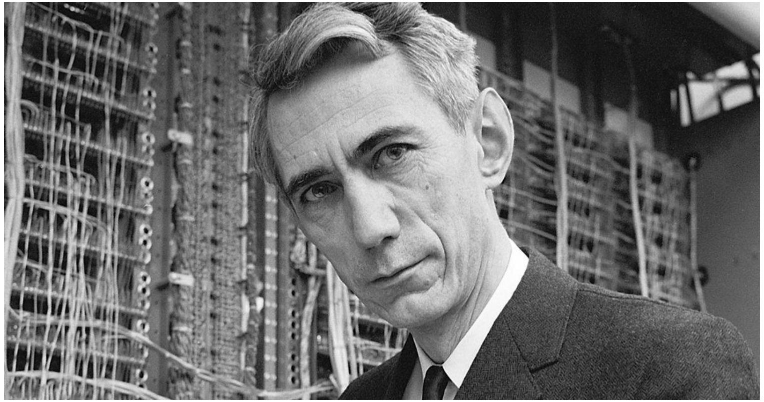
https://www.ussc.gov/sites/default/files/zip/BookerReport2012.zip


Certain?
Predictive Model
Observational Inference
What is certainty?
Data
Model Source
Theory
Description
Prediction
Association/Correlation
Mo
WARNING
Purpose
Explanation
Simulation
Attribution
Optimization
Causation
Fair?
Causal Model
Causal Inference

Information Theory

"Information is the resolution of uncertainty."

Claude Shannon, 1948

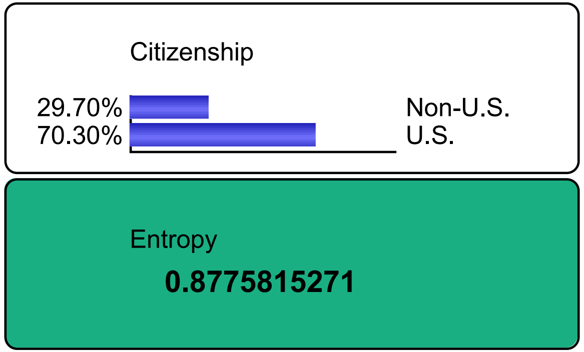


Claude Shannon (1916-2001)

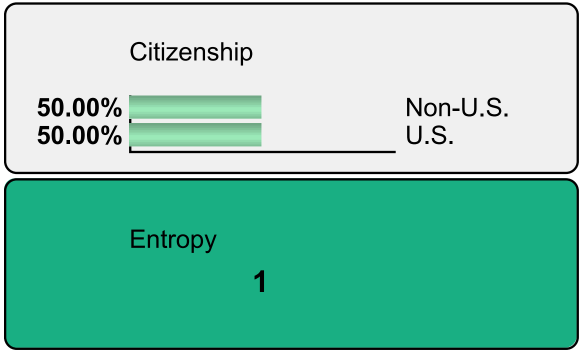
Information Theory

Entropy, a measure of "uncertainty"

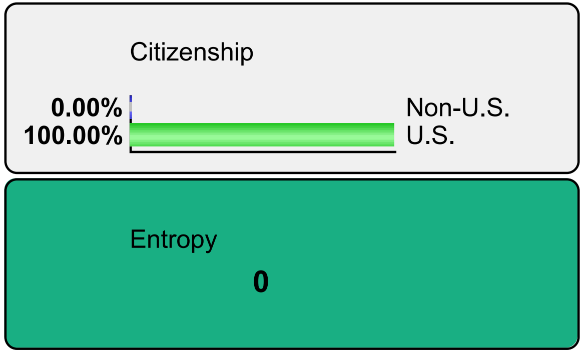
$$H(X) = -\sum_{x \in X} p(x) \times \log_2 p(x)$$



Marginal Entropy

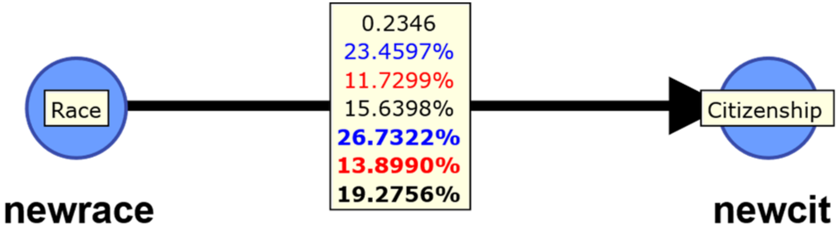


Maximum Entropy



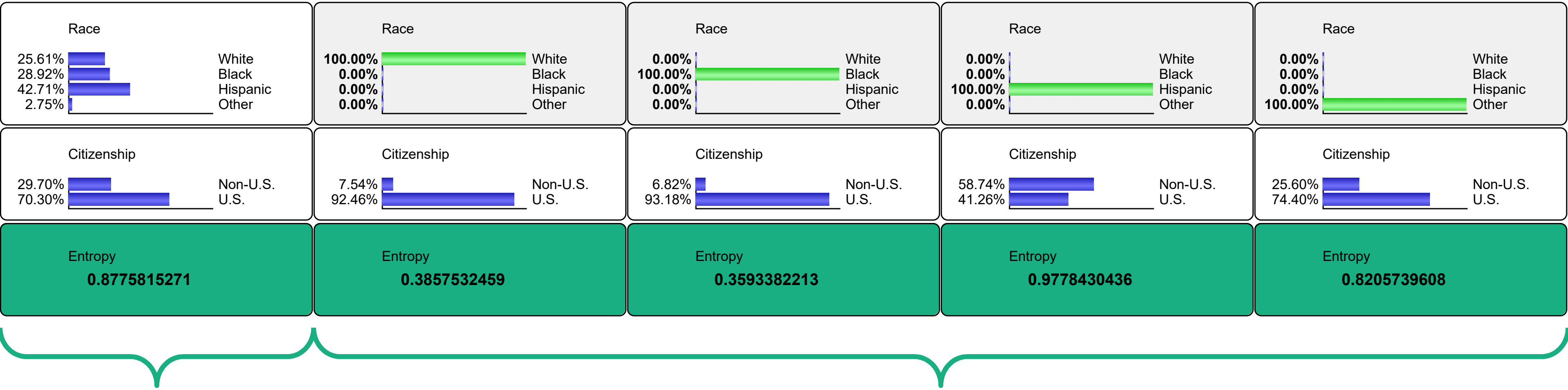
Minimum Entropy

Information Theory



Conditional Entropy

$$H(\text{Citizenship}|\text{Race})$$

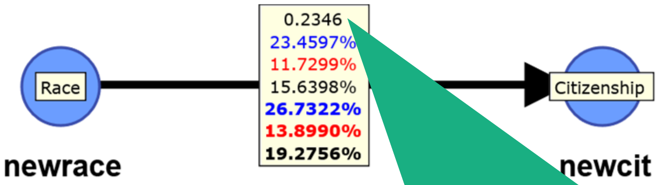


$H(\text{Citizenship})=0.8776$ $\qquad$ $H(\text{Citizenship}|\text{Race})=0.6429$

Information Theory

Mutual Information

$$\underbrace{\mathrm{I(Citizenship,\ Race)}}_{\substack{0.2346\\ \text{Mutual Information}}}=\underbrace{\mathrm{H(Citizenship)}}_{\substack{0.8776\\ \text{Marginal Entropy}}}-\underbrace{\mathrm{H(Citizenship|Race)}}_{\substack{0.6429\\ \text{Conditional Entropy}}}$$



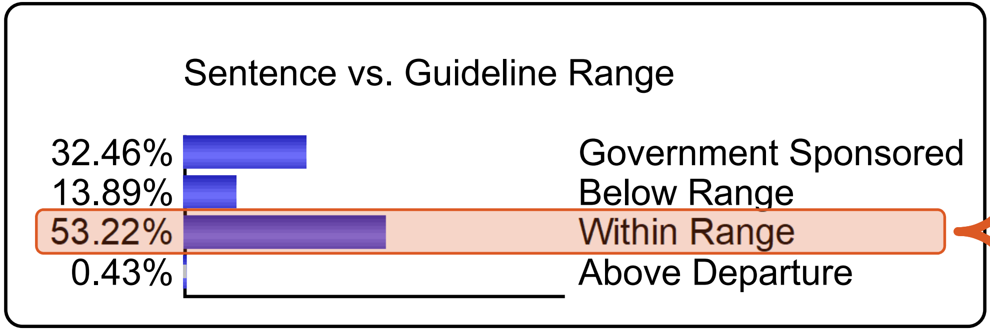
This is the amount of information Citizenship and Race have in common.

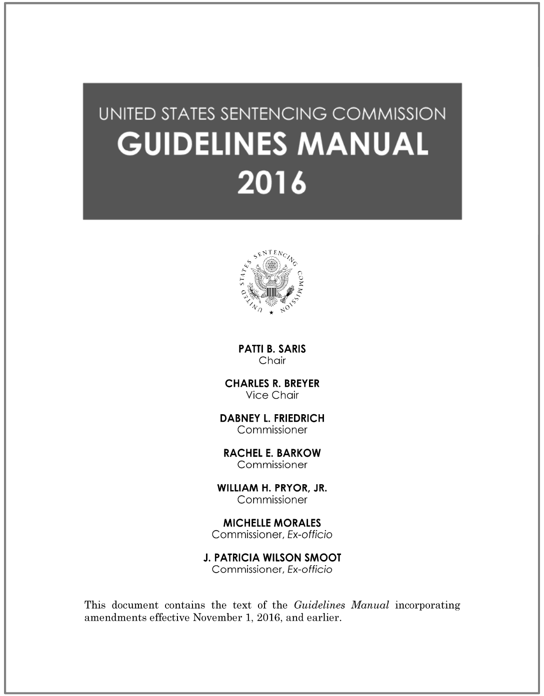
Mutual Information is a symmetrical metric.

Certain?

Departures from Guideline (booker2)

- What is the entropy over time?




UNITED STATES SENTENCING COMMISSION
GUIDELINES MANUAL
2016

PATTI B. SARIS
Chair

CHARLES R. BREYER
Vice Chair

DABNEY L. FRIEDRICH
Commissioner

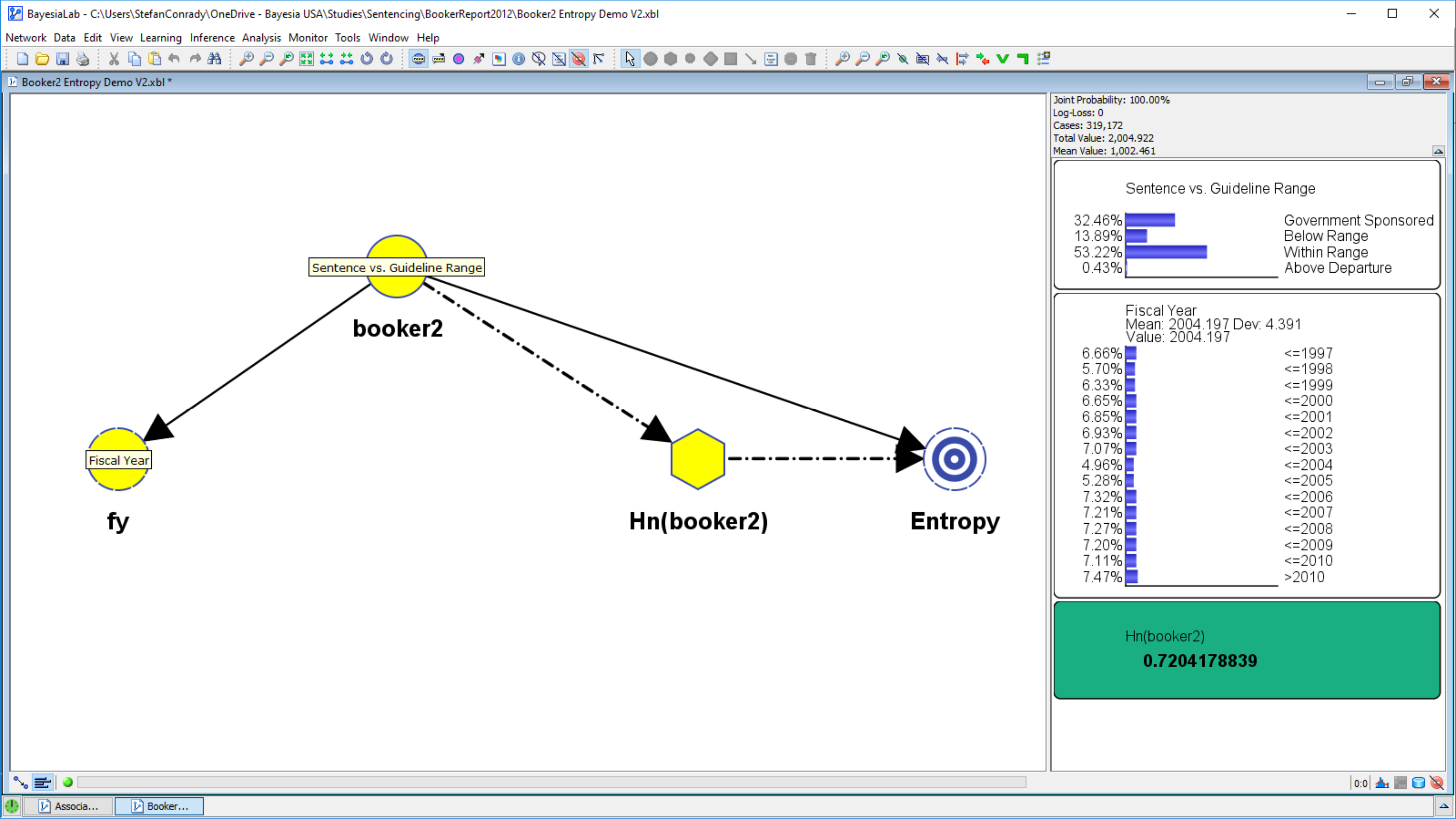
RACHEL E. BARKOW
Commissioner

WILLIAM H. PRYOR, JR.
Commissioner

MICHELLE MORALES
Commissioner, *Ex-officio*

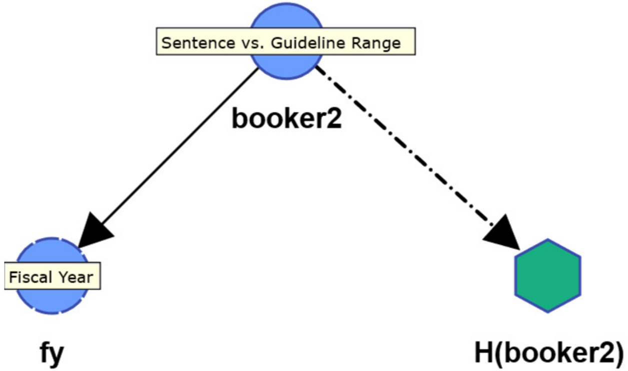
J. PATRICIA WILSON SMOOT
Commissioner, *Ex-officio*

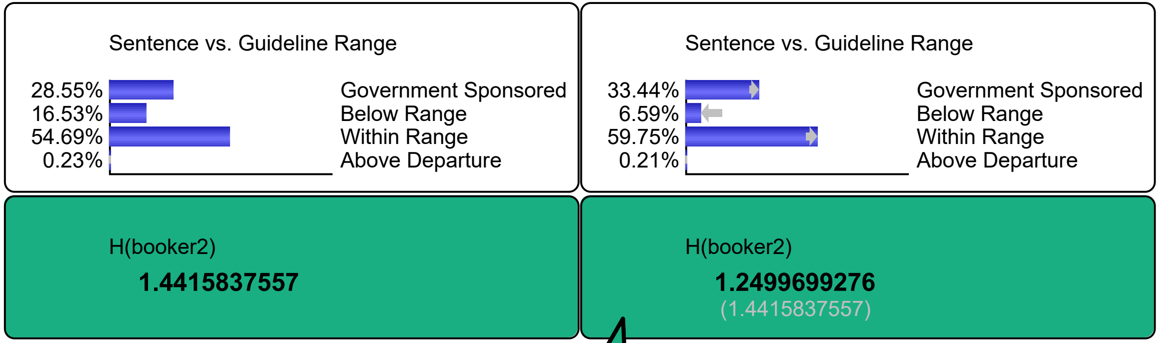
This document contains the text of the *Guidelines Manual* incorporating amendments effective November 1, 2016, and earlier.


BayesiaLab - C:\Users\StefanConrady\OneDrive - Bayesia USA\Studies\Sentencing\BookerReport2012\Booker2 Entropy Demo V2.xbl
Network Data Edit View Learning Inference Analysis Monitor Tools Window Help
Booker2 Entropy Demo V2.xbl *
Sentence vs. Guideline Range
booker2
Fiscal Year
fy
Hn(booker2)
Entropy
Joint Probability: 100.00%
Log-Loss: 0
Cases: 319,172
Total Value: 2,004.922
Mean Value: 1,002.461
Sentence vs. Guideline Range
32.46% Government Sponsored
13.89% Below Range
53.22% Within Range
0.43% Above Departure
Fiscal Year
Mean: 2004.197 Dev: 4.391
Value: 2004.197
6.66% <=1997
5.70% <=1998
6.33% <=1999
6.65% <=2000
6.85% <=2001
6.93% <=2002
7.07% <=2003
4.96% <=2004
5.28% <=2005
7.32% <=2006
7.21% <=2007
7.27% <=2008
7.20% <=2009
7.11% <=2010
7.47% >2010
Hn(booker2)
0.7204178839
Associa...
Booker...

Certain?

Measuring Uncertainty as a Function of Time Period





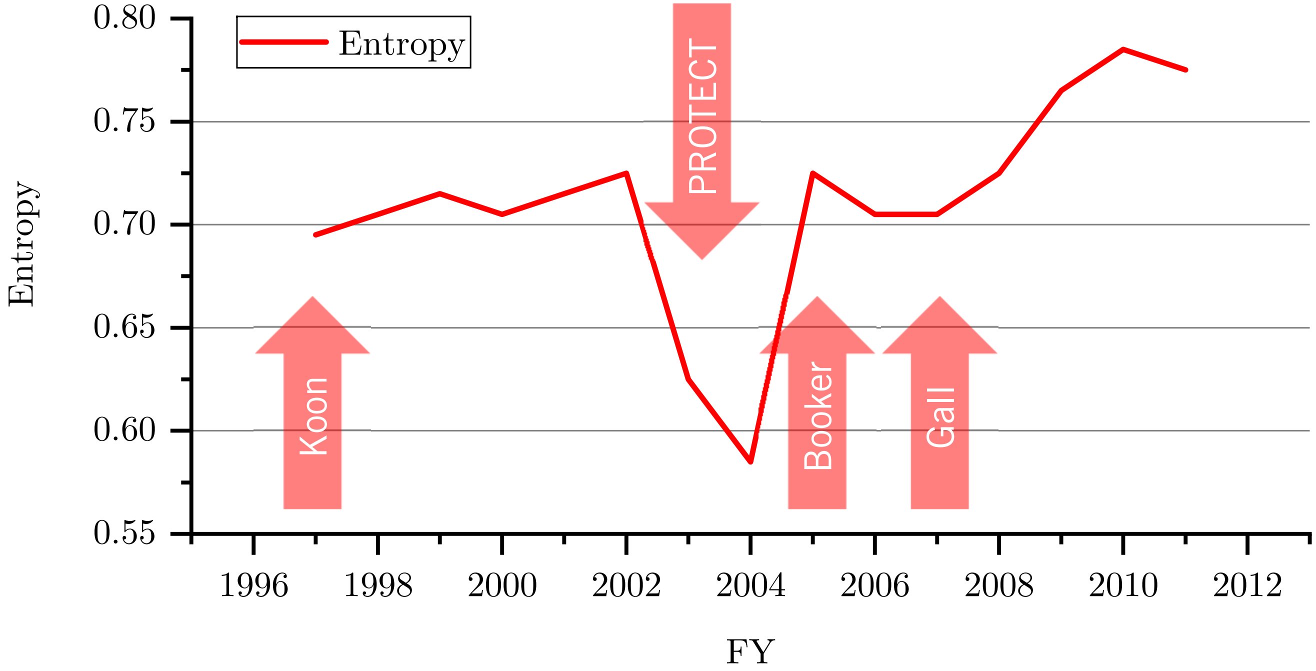
2002

2003
PROTECT Act

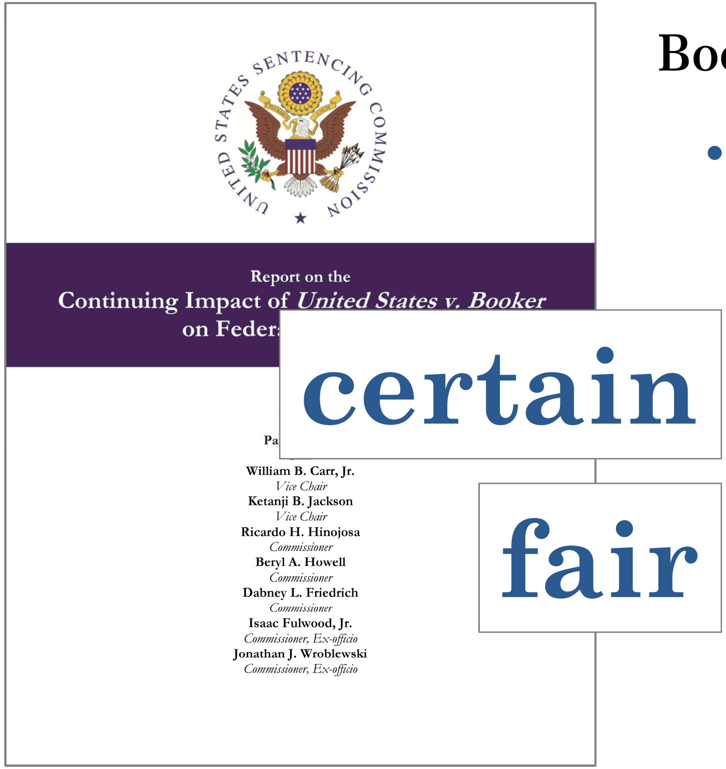
Lower Entropy
➔ Less Uncertainty

Conditional Entropy

Normalized Entropy of Sentence vs. Guideline Range over Time

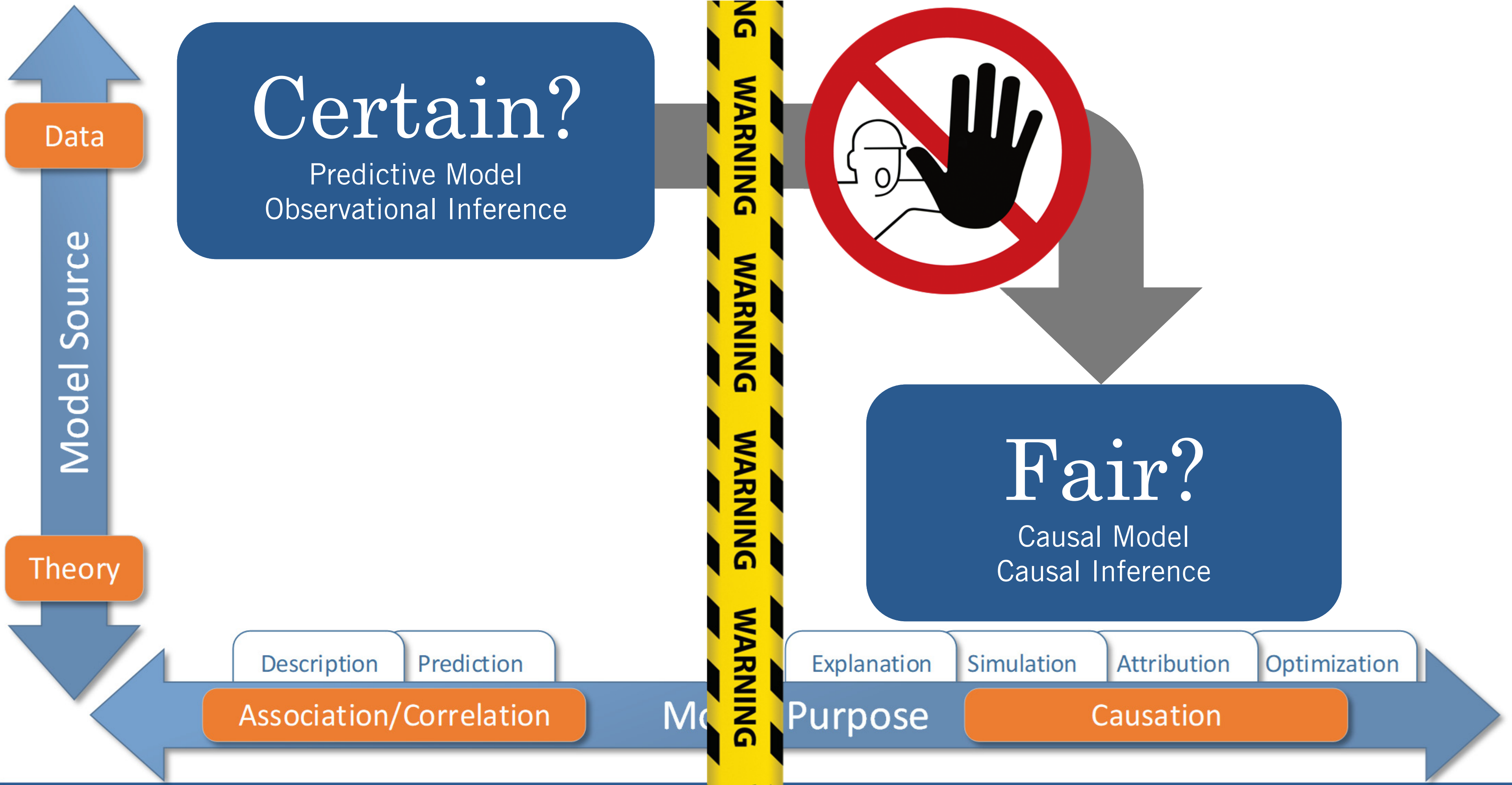


Research Questions



Booker Report

- The United States Sentencing Commission ("the Commission") submits this report to Congress[1] on the impact of United States v. Booker[2] on federal sentencing in order to assist Congress in its efforts to ensure **certain and fair sentencing** that avoids unwarranted sentencing disparities while maintaining sufficient flexibility,[3] as envisioned in the Sentencing Reform Act of 1984 ("SRA").

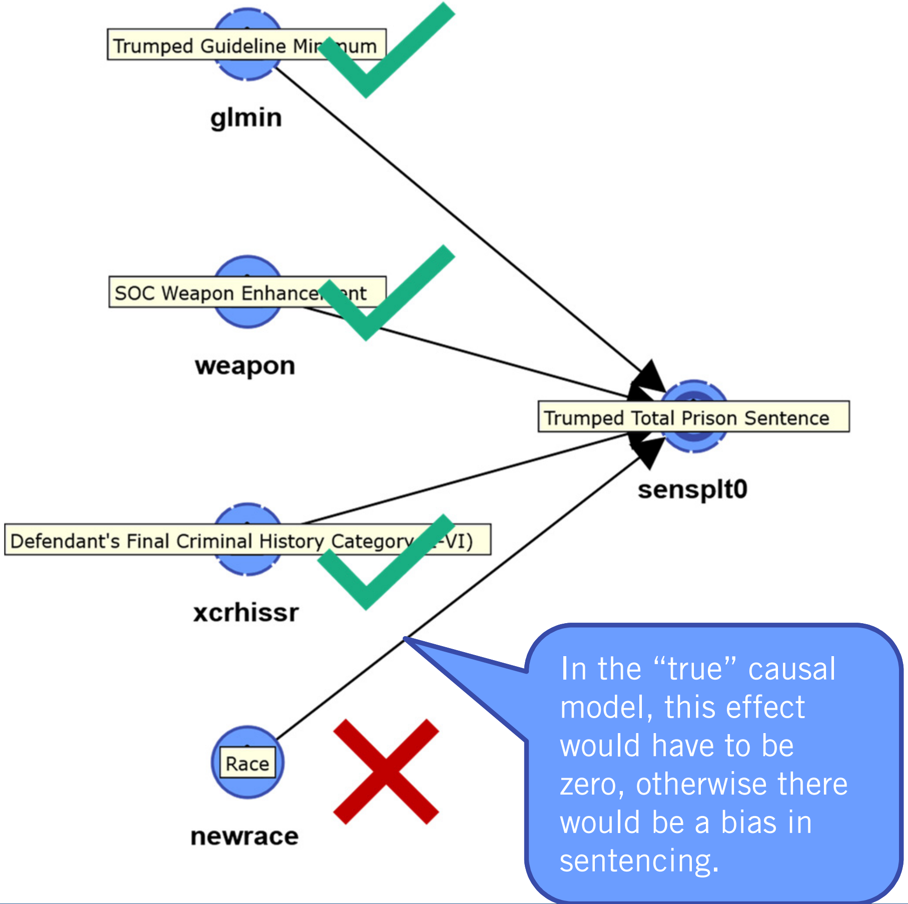


Certain?
Predictive Model
Observational Inference
Fair?
Causal Model
Causal Inference
Data
Model Source
Theory
Description
Prediction
Explanation
Simulation
Attribution
Optimization
Association/Correlation
Mo
Purpose
Causation
WARNING
WARNING
WARNING
WARNING

Causality?

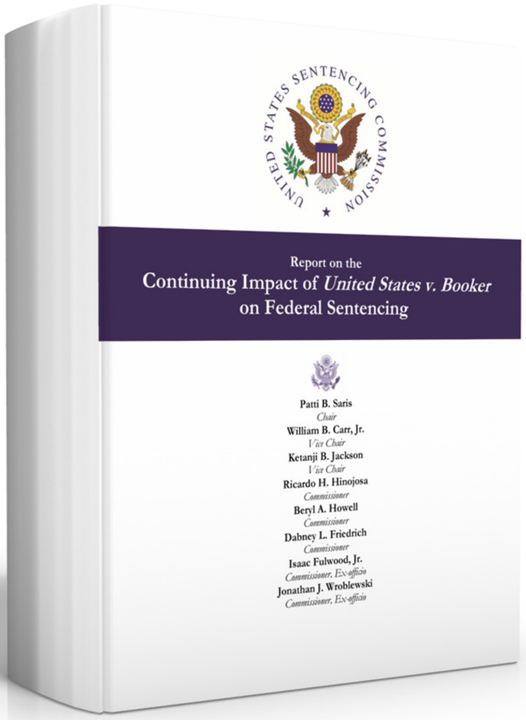
Why is this a causal question?

A possible
causal model



Race must not "cause" sentence!

Causality?

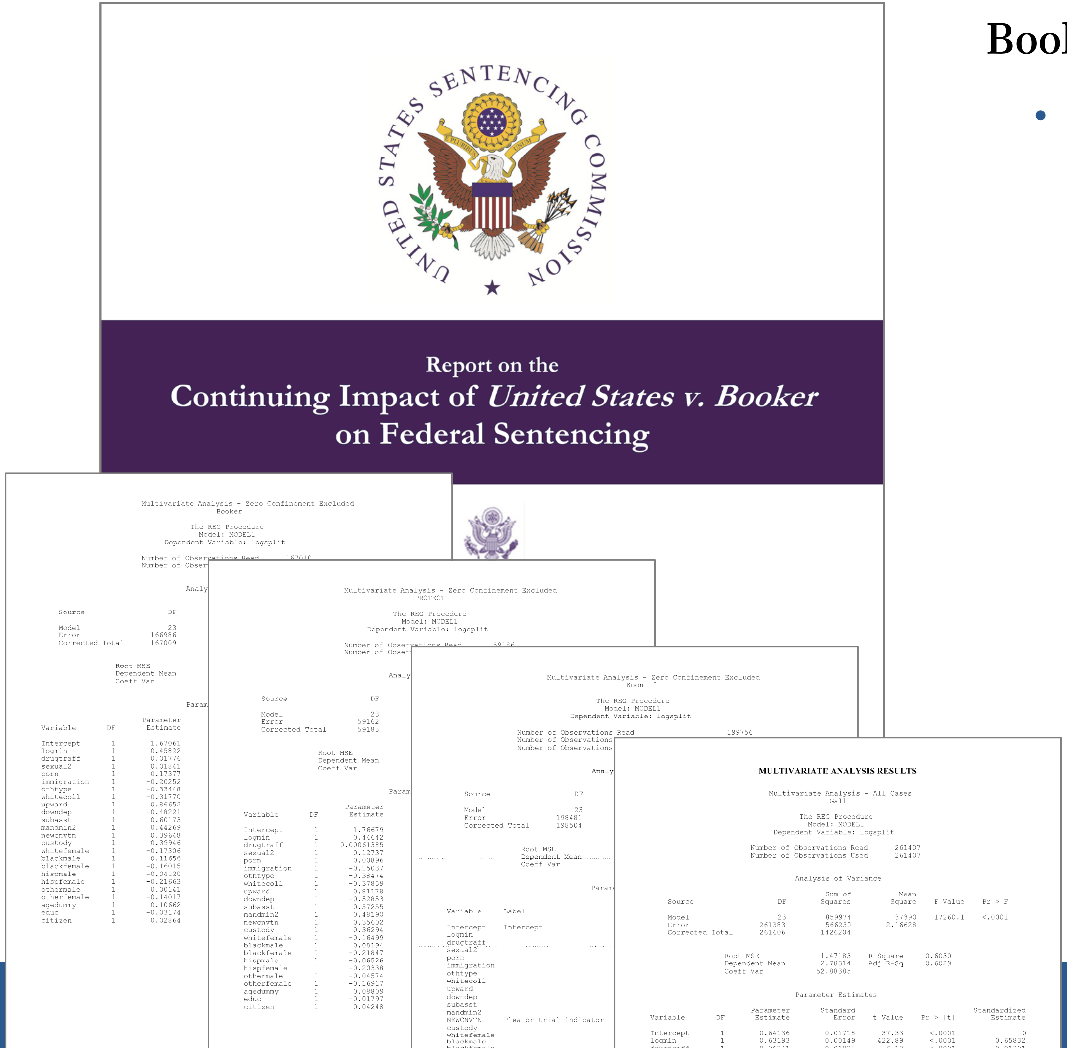


Booker Report:

- 1430 pages
- 1 footnote on "causality"

[19] Correlation and causation are different concepts. A variable that is correlated with another may not be caused by it.

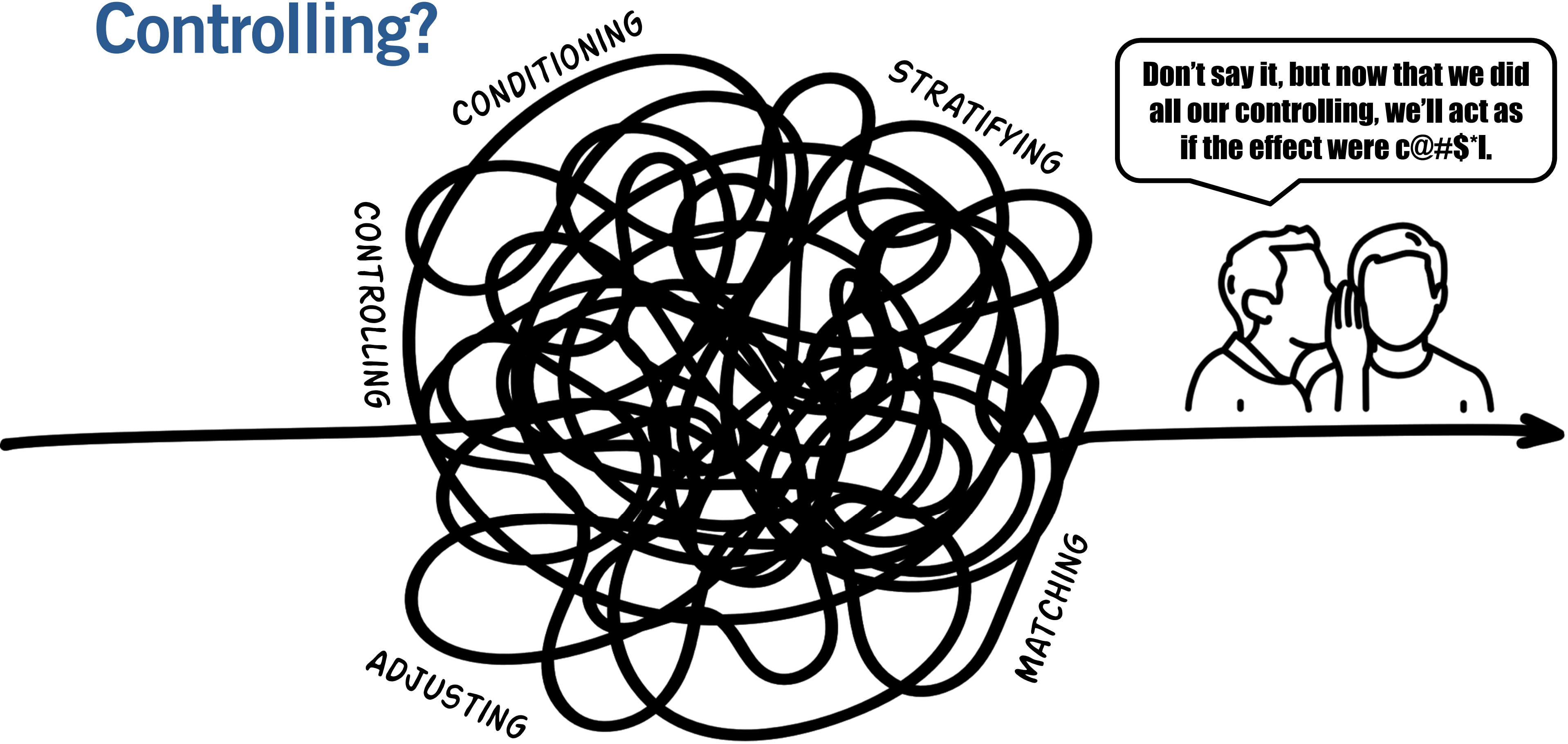
> [I]n interpreting the results of a multiple regression analysis, it is important to distinguish between correlation and causality. Two variables are correlated when the events associated with the variables occur more frequently together than one would expect by chance . . . . A correlation between two variables does not imply that one event causes the second. Therefore, in making causal inferences, it is important to avoid spurious correlation. Spurious correlation arises when two variables are closely related but bear no causal relationship because both are caused by a third, unexamined variable . . . . Causality cannot be inferred by data analysis alone; rather, one must infer that a causal relationship exists on the basis of an underlying causal theory that explains the relationship between the two variables. Even when an appropriate theory has been identified, causality can never be inferred directly. One must look for empirical evidence that there is a causal relationship. Conversely, the fact that two variables are correlated does not guarantee the existence of a relationship; it could be that the model – a characterization of the underlying theory – does not reflect the correct interplay among the explanatory variables.



Booker Report

- "The principal benefit of multivariate regression analysis is that it controls for the effect of each factor in the analysis by comparing offenders who are similar to one another in relevant ways. For example, controlling for the presumptive sentence (guideline minimum) and offense type means that Black male offenders convicted of firearms offenses and who faced a guideline minimum of 46 months of imprisonment are compared to White male offenders convicted of firearms offenses who faced a guideline minimum of 46 months of imprisonment. By controlling for such factors and comparing similarly situated offenders to each other, multivariate regression analysis seeks to answer the question: if two offenders are similar in certain ways, what other factors might be associated with those two offenders receiving different sentences? In addition, multivariate regression analysis measures the extent of the difference in outcomes."

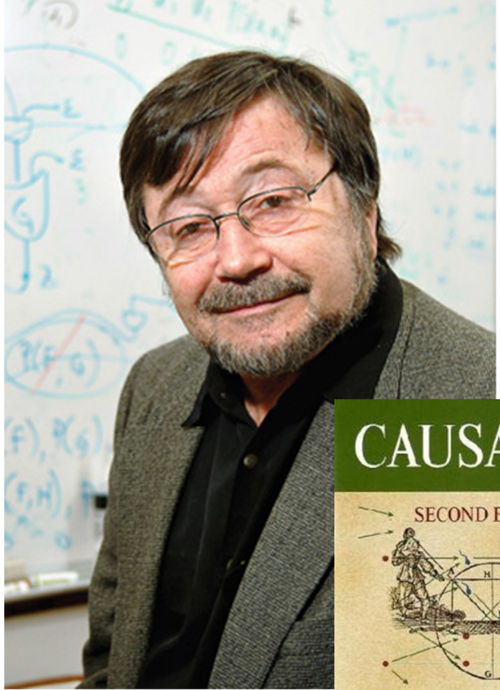
Controlling?

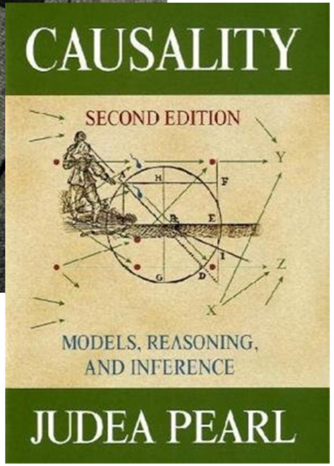


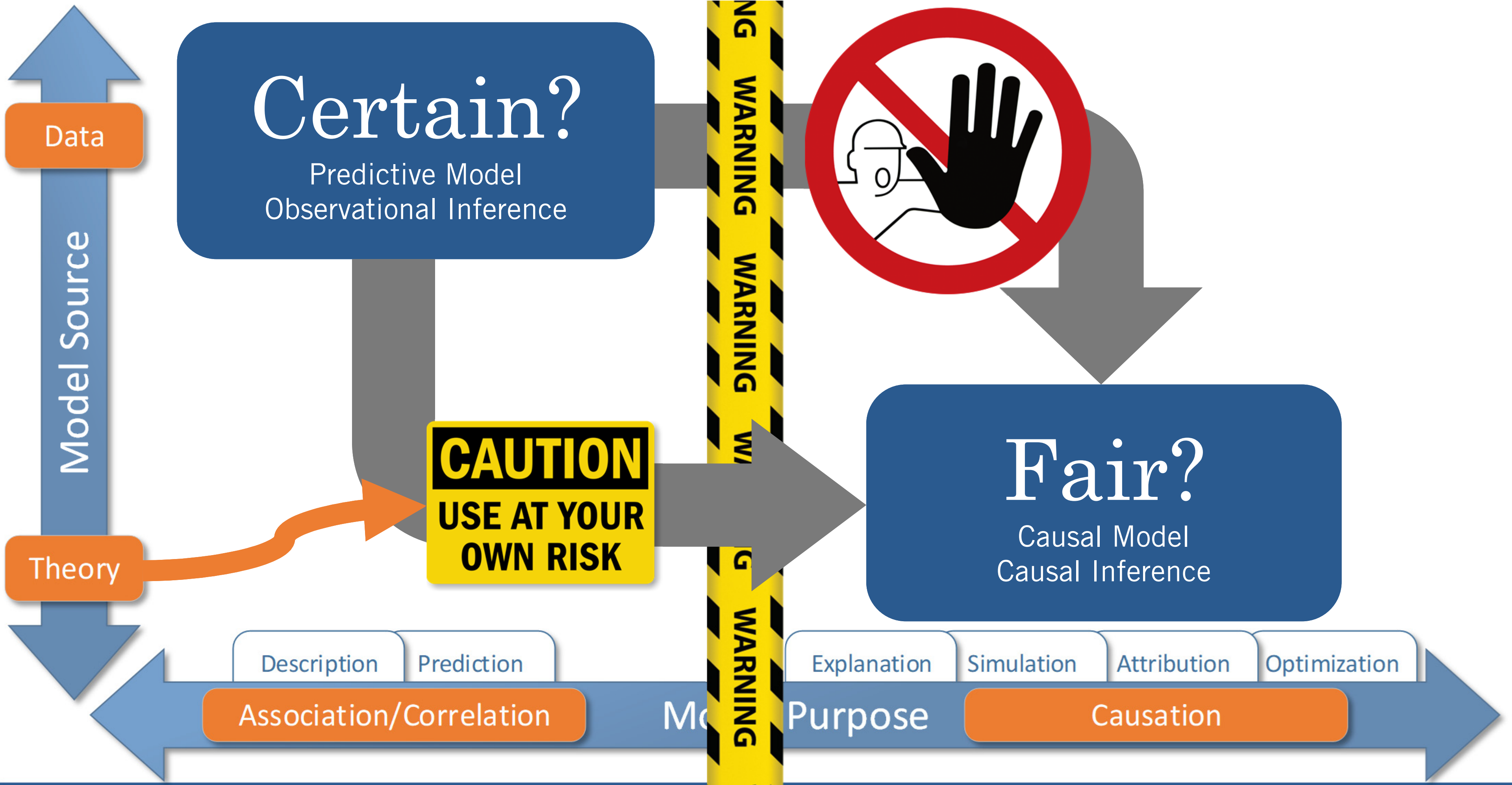
Controlling?

“The prevailing attitude is that adding more covariates can cause no harm and can absolve one from thinking about causal relationships…”

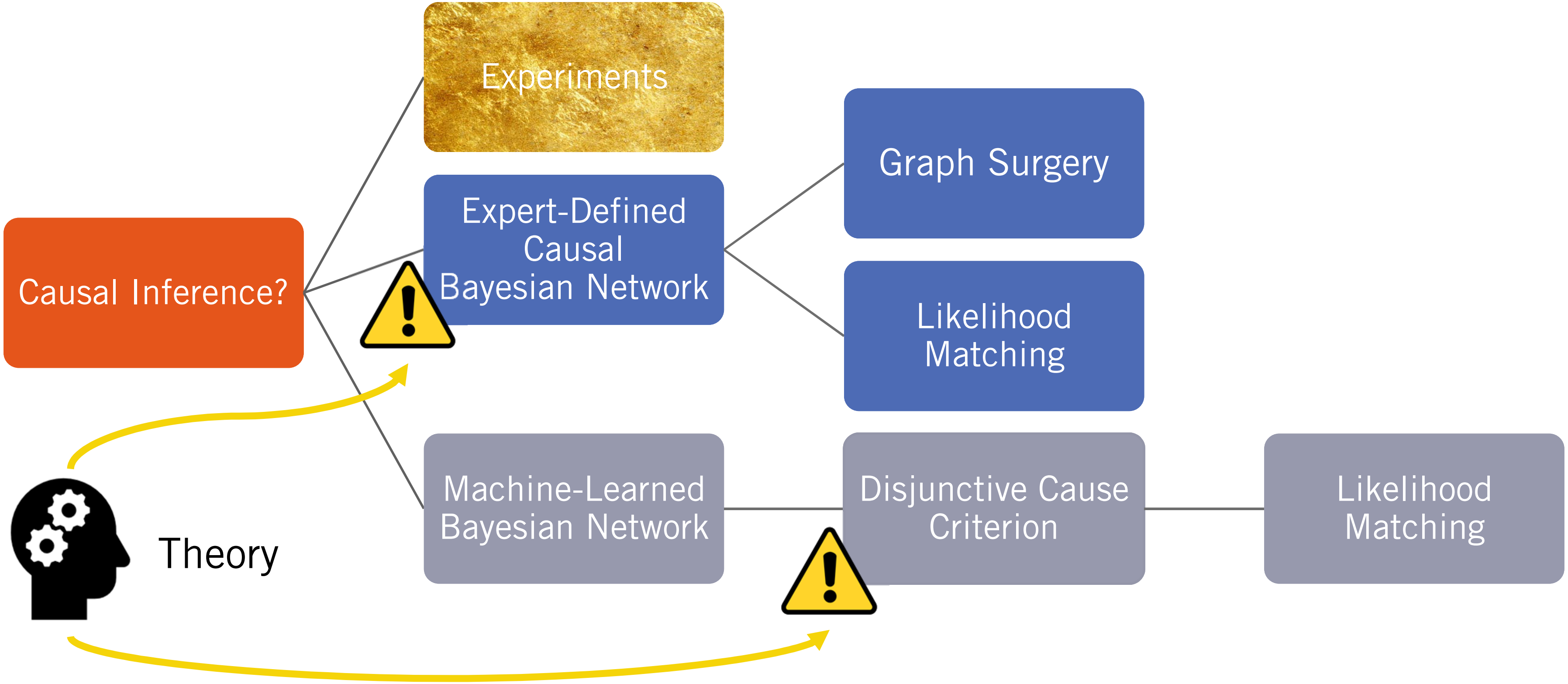
Judea Pearl in Causality (2009)






Data
Model Source
Theory
Certain?
Predictive Model
Observational Inference
WARNING
CAUTION
USE AT YOUR OWN RISK
Fair?
Causal Model
Causal Inference
Description
Prediction
Explanation
Simulation
Attribution
Optimization
Association/Correlation
Purpose
Causation

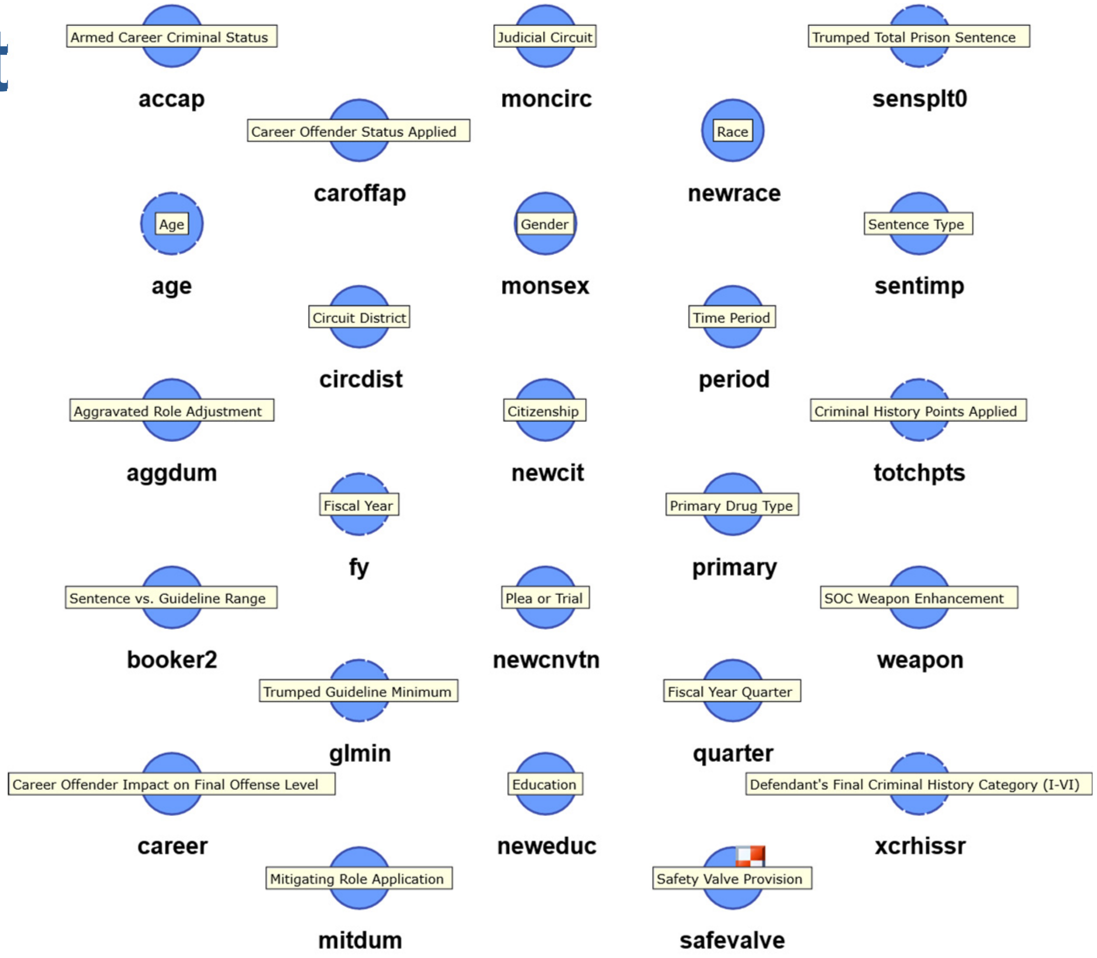
Causality?


Experiments
Causal Inference?
Expert-Defined Causal Bayesian Network
Graph Surgery
Likelihood Matching
Machine-Learned Bayesian Network
Disjunctive Cause Criterion
Likelihood Matching
Theory

Model Development

Encoding a Causal Model

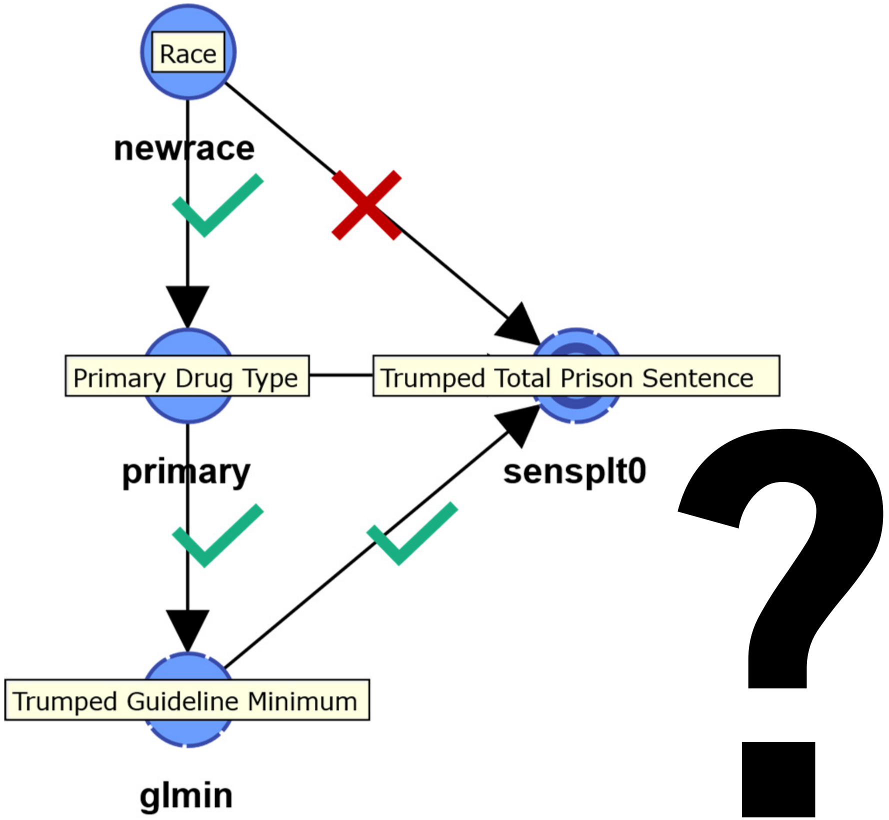
- What causes the *Total Prison Sentence*?
- Everything?



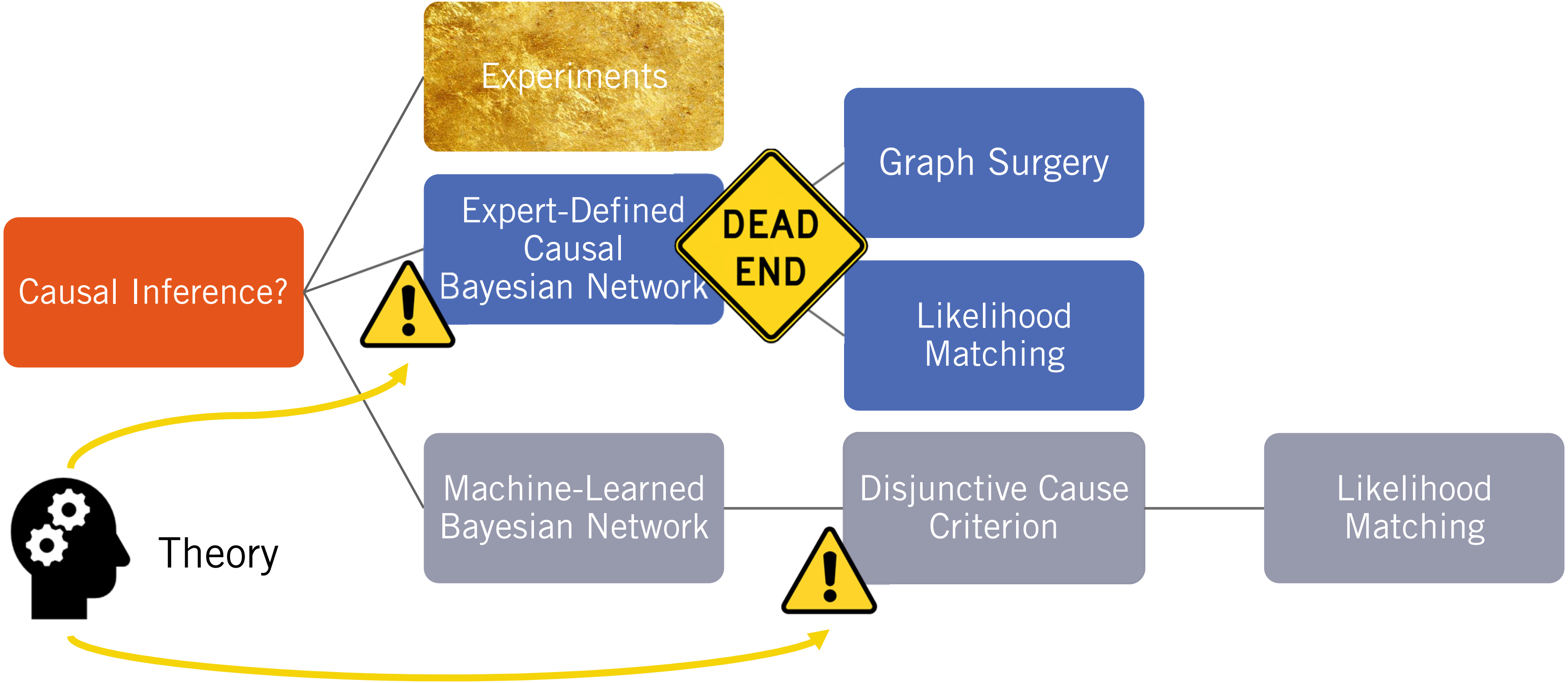
Model Development

Encoding a Causal Model

- How do the variables cause each other?
- There are way too many assumption we would have to make!



Causality?



Experiments
Expert-Defined Causal Bayesian Network
DEAD END
Graph Surgery
Likelihood Matching
Causal Inference?
Machine-Learned Bayesian Network
Disjunctive Cause Criterion
Likelihood Matching
Theory

Disjunctive Cause Criterion

Published in final edited form as:
Biometrics. 2011 December ; 67(4): 1406–1413. doi:10.1111/j.1541-0420.2011.01619.x.

A new criterion for confounder selection

Tyler J. VanderWeele and
Departments of Epidemiology and Biostatistics, Harvard School of Public Health 677 Huntington Avenue, Boston, MA 02115, Phone: 617-432-7855; Fax: 617-4321884

Ilya Shpitser
Department of Epidemiology, Harvard School of Public Health 677 Huntington Avenue, Boston, MA 02115

Tyler J. VanderWeele: tvanderw@hsph.harvard.edu

Abstract

We propose a new criterion for confounder selection when the underlying causal structure is unknown and only limited knowledge is available. We assume all covariates being considered are pretreatment variables and that for each covariate it is known (i) whether the covariate is a cause of treatment, and (ii) whether the covariate is a cause of the outcome. The causal relationships the covariates have with one another is assumed unknown. We propose that control be made for any covariate that is either a cause of treatment or of the outcome or both. We show that irrespective of the actual underlying causal structure, if any subset of the observed covariates suffices to control

Disjunctive Cause Criterion

“Confounder” in BayesiaLab

VanderWeele and Shpitser (2011)

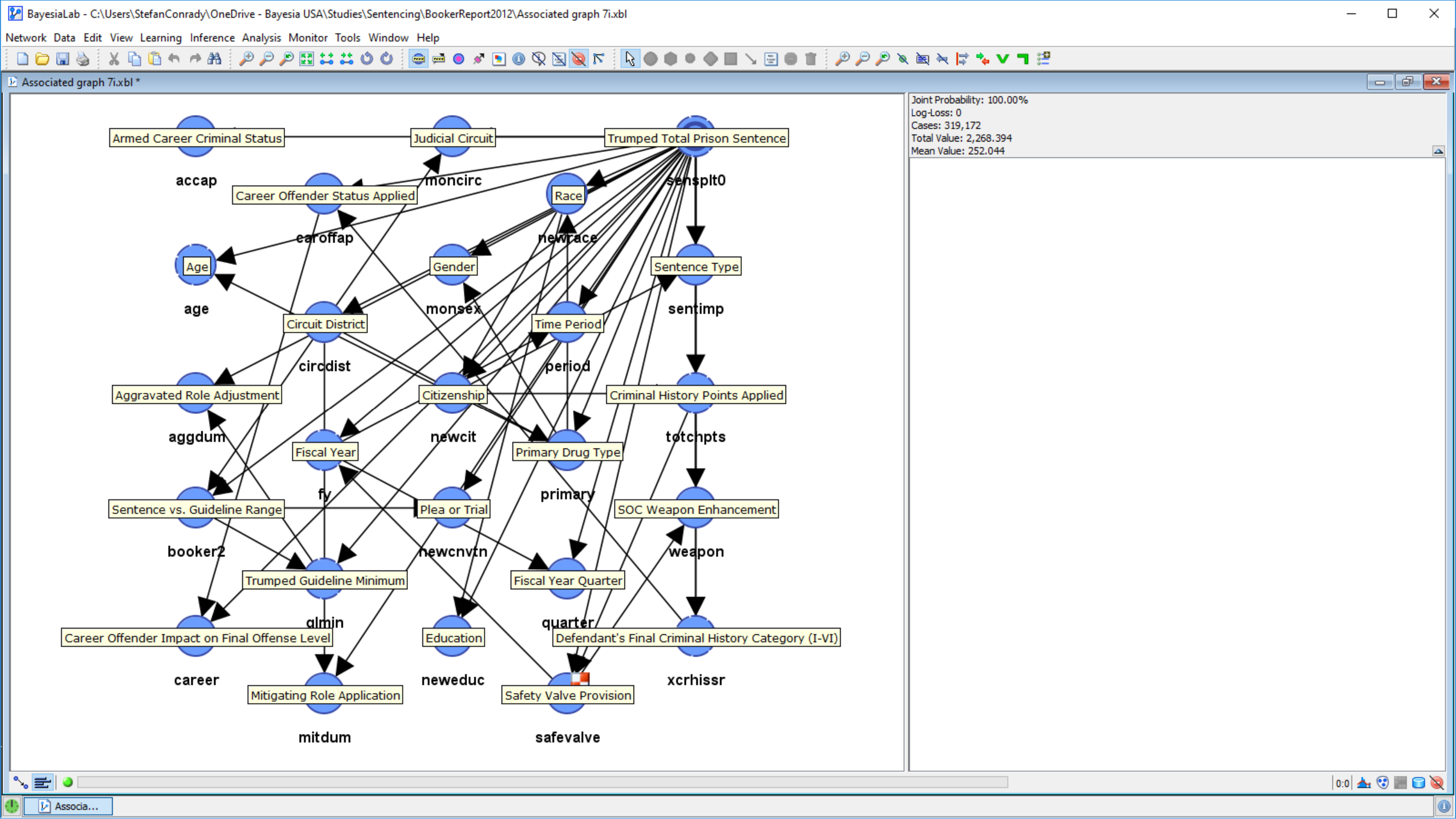
- “We propose that control be made for any [**pre-treatment**] **covariate** that is either a cause of **treatment** or of the **outcome** or both.”

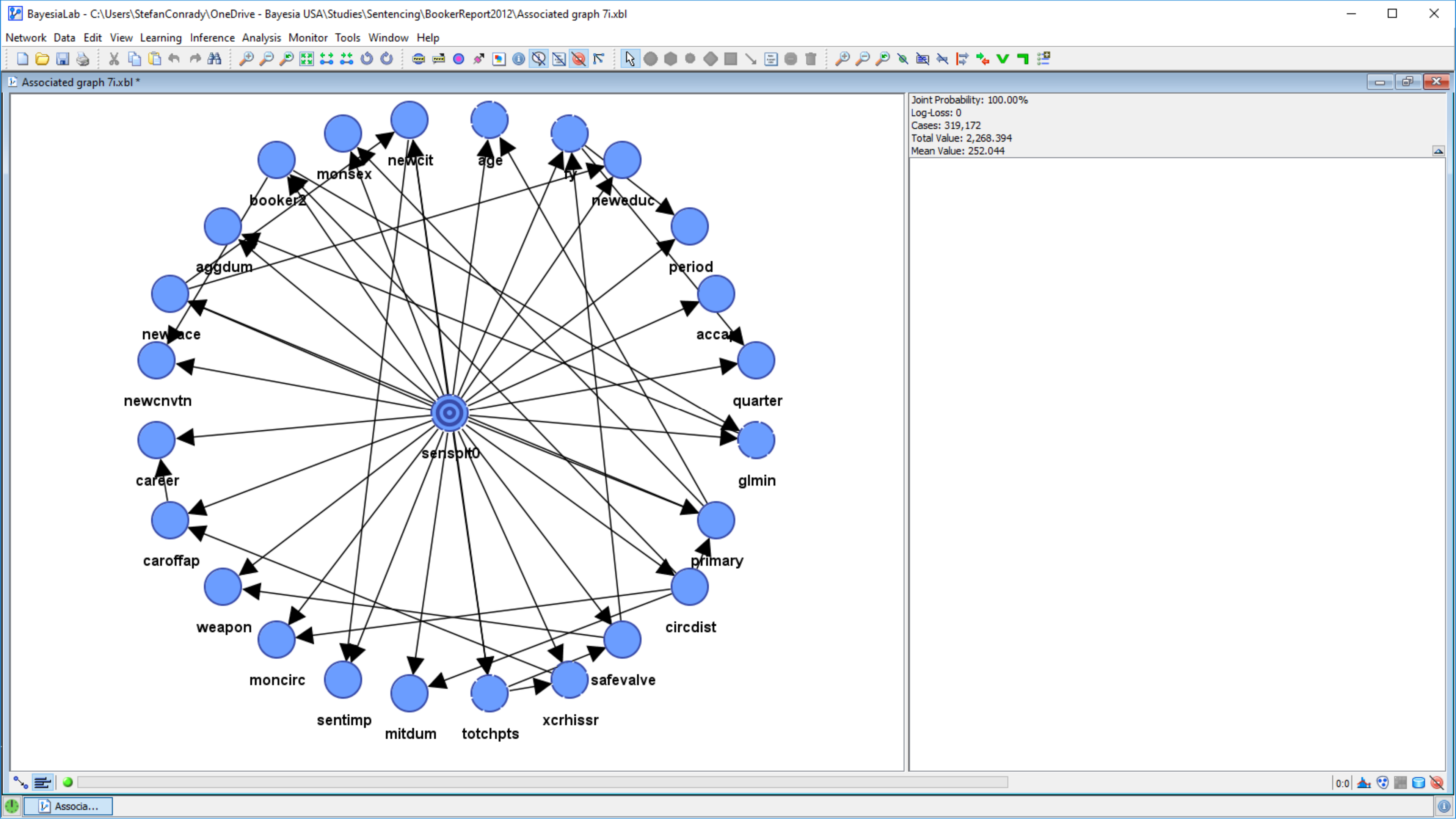
Race

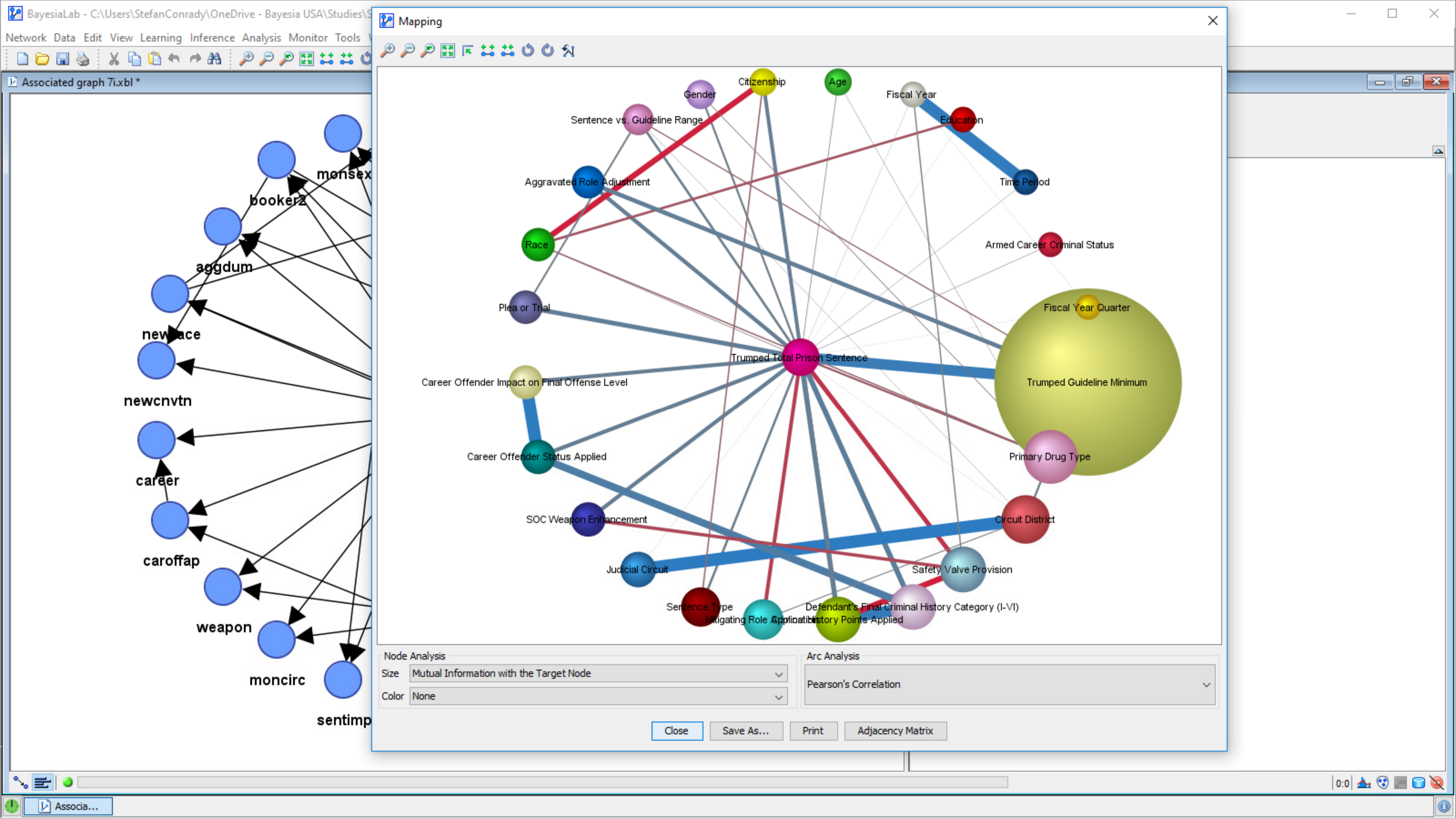
Sentence

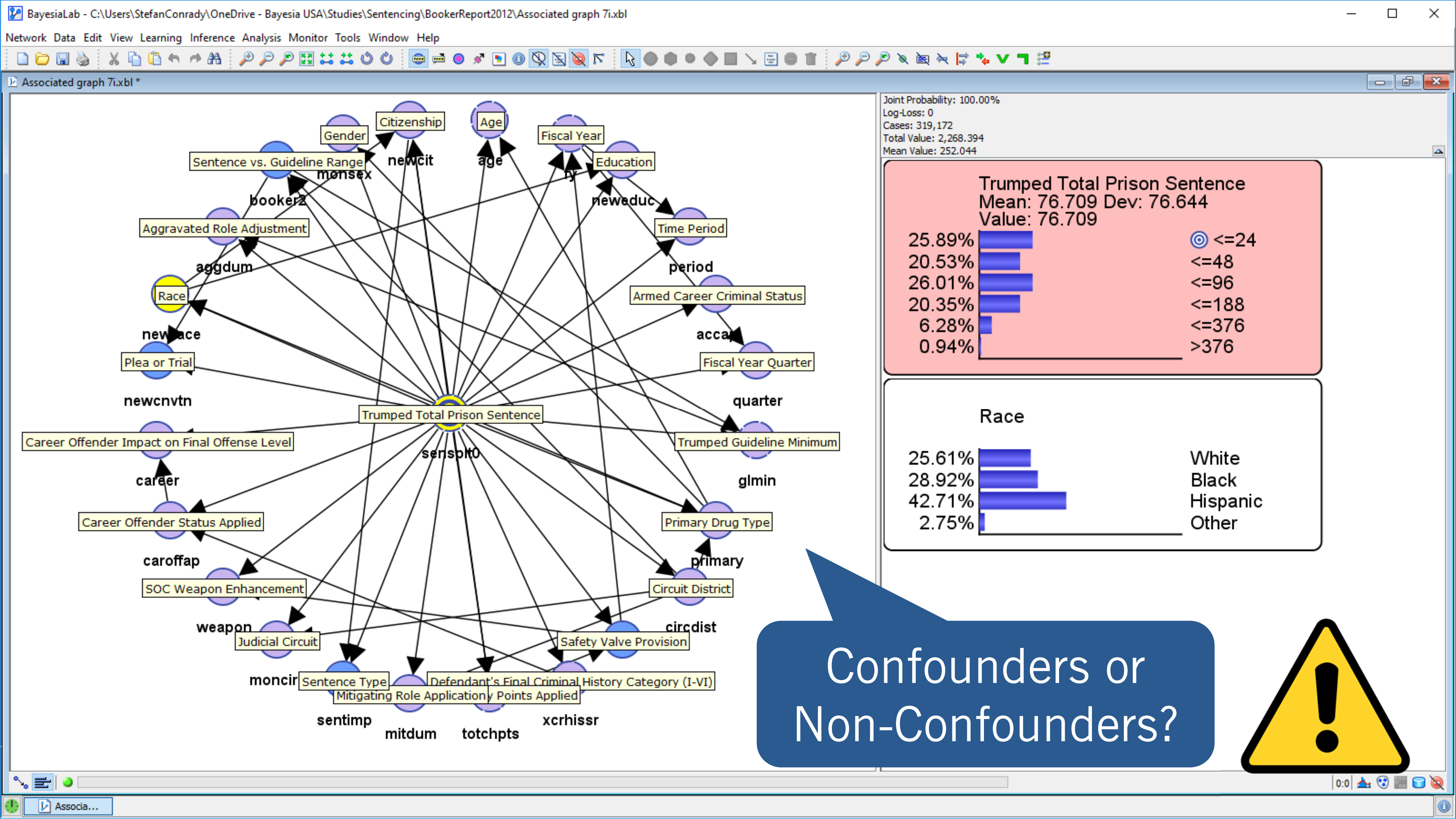
Implementation in BayesiaLab:
Likelihood Matching on Confounders in
Direct Effects Analysis → Causal Effect

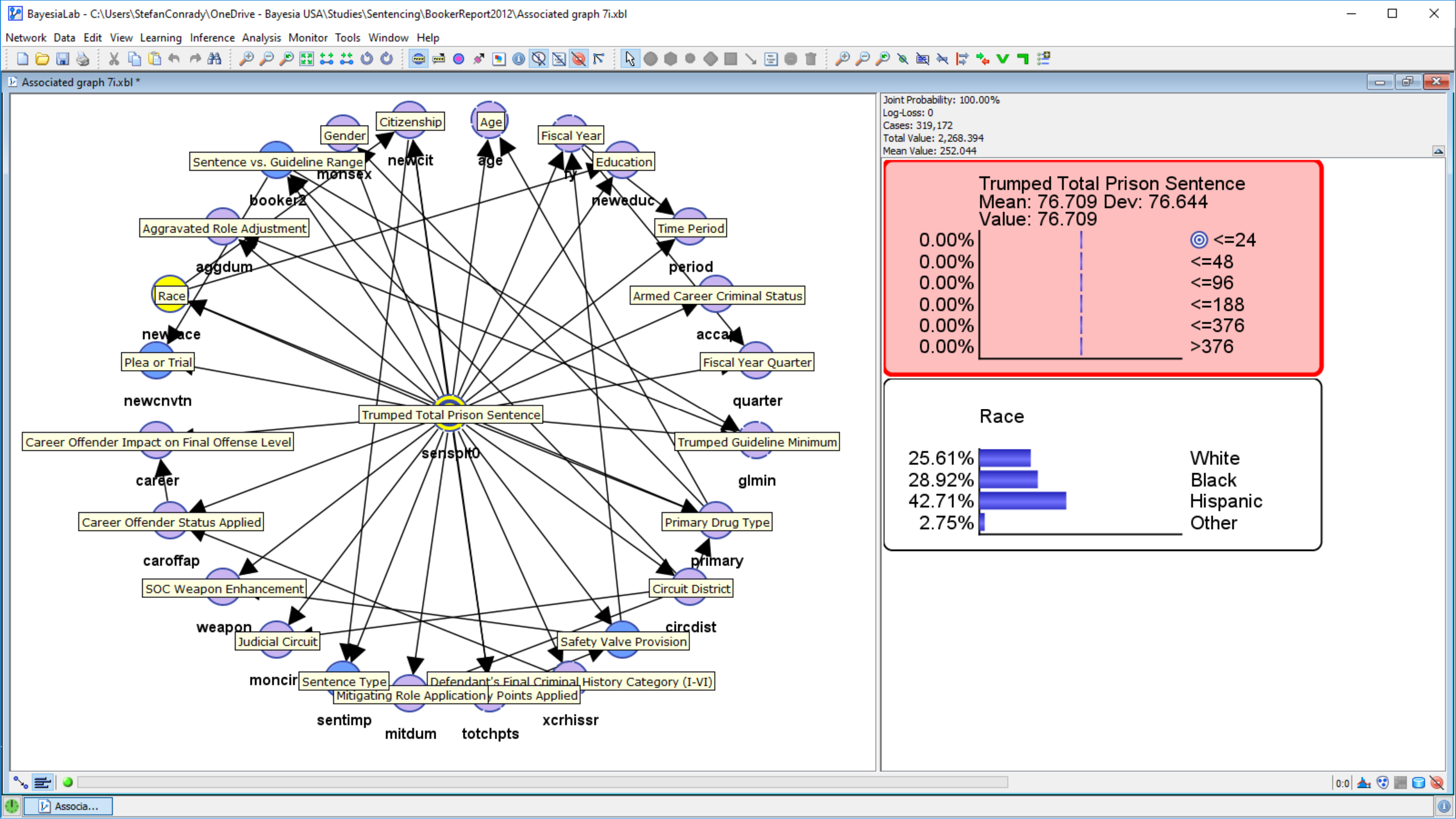
IMPORTANT ASSUMPTION:
NO UNOBSERVED CONFOUNDERS

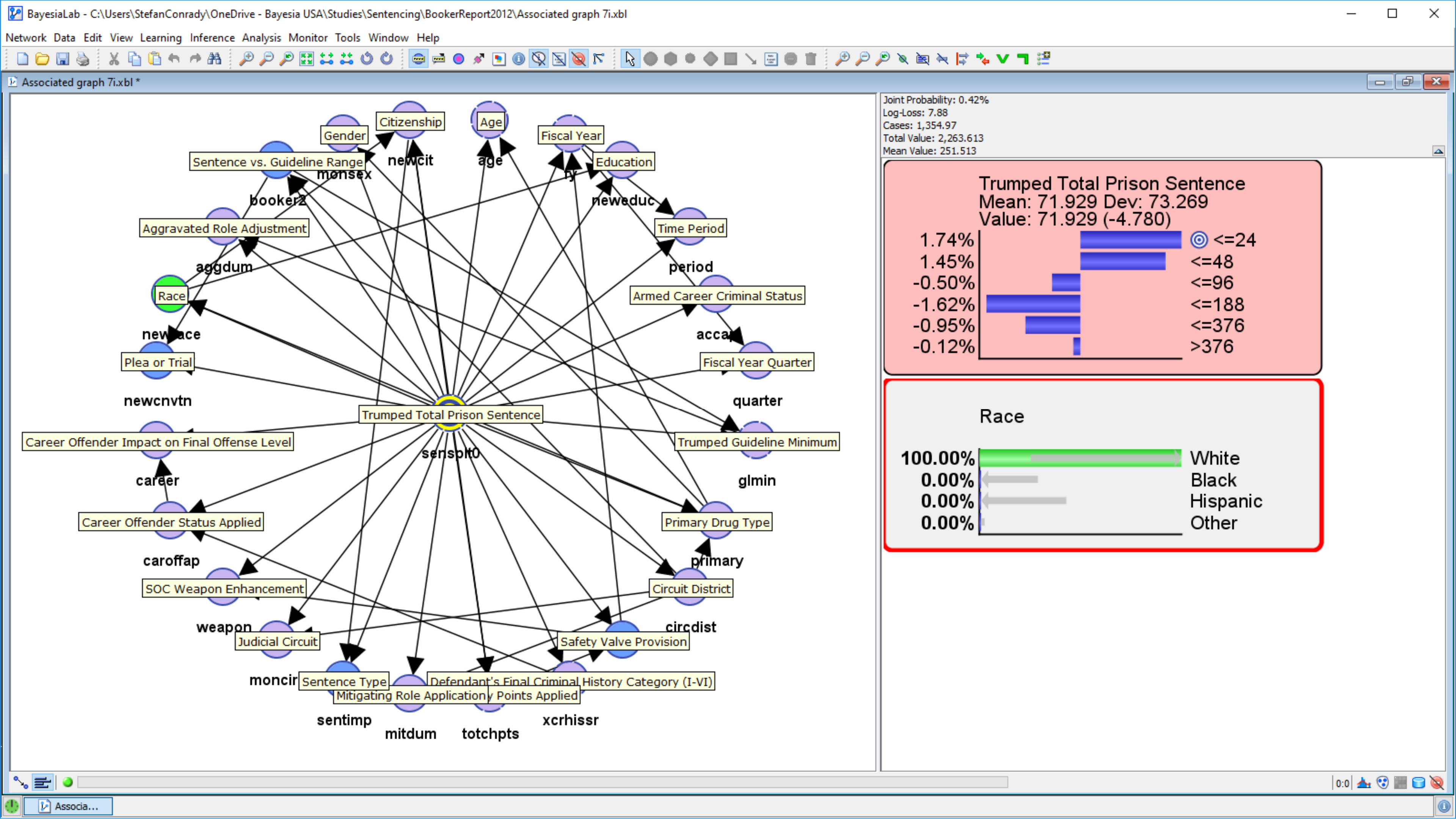

BayesiaLab - C:\Users\StefanConrady\OneDrive - Bayesia USA\Studies\Sentencing\BookerReport2012\Associated graph 7i.xbl
Network Data Edit View Learning Inference Analysis Monitor Tools Window Help
Associated graph 7i.xbl *
Armed Career Criminal Status
accap
Judicial Circuit
moncirc
Trumped Total Prison Sentence
senspl t0
Career Offender Status Applied
caroffap
Race
newrace
Age
age
Gender
monsex
Sentence Type
sentimp
Circuit District
circdist
Time Period
period
Aggravated Role Adjustment
aggdum
Citizenship
newcit
Criminal History Points Applied
totchpts
Fiscal Year
fy
Primary Drug Type
primary
Sentence vs. Guideline Range
booker2
Plea or Trial
newcnvtn
SOC Weapon Enhancement
weapon
Trumped Guideline Minimum
glmin
Fiscal Year Quarter
quarter
Career Offender Impact on Final Offense Level
career
Education
neweduc
Defendant's Final Criminal History Category (I-VI)
xcrhissr
Mitigating Role Application
mitdum
Safety Valve Provision
safevalve
Joint Probability: 100.00%
Log-Loss: 0
Cases: 319,172
Total Value: 2,268.394
Mean Value: 252.044
Associa...

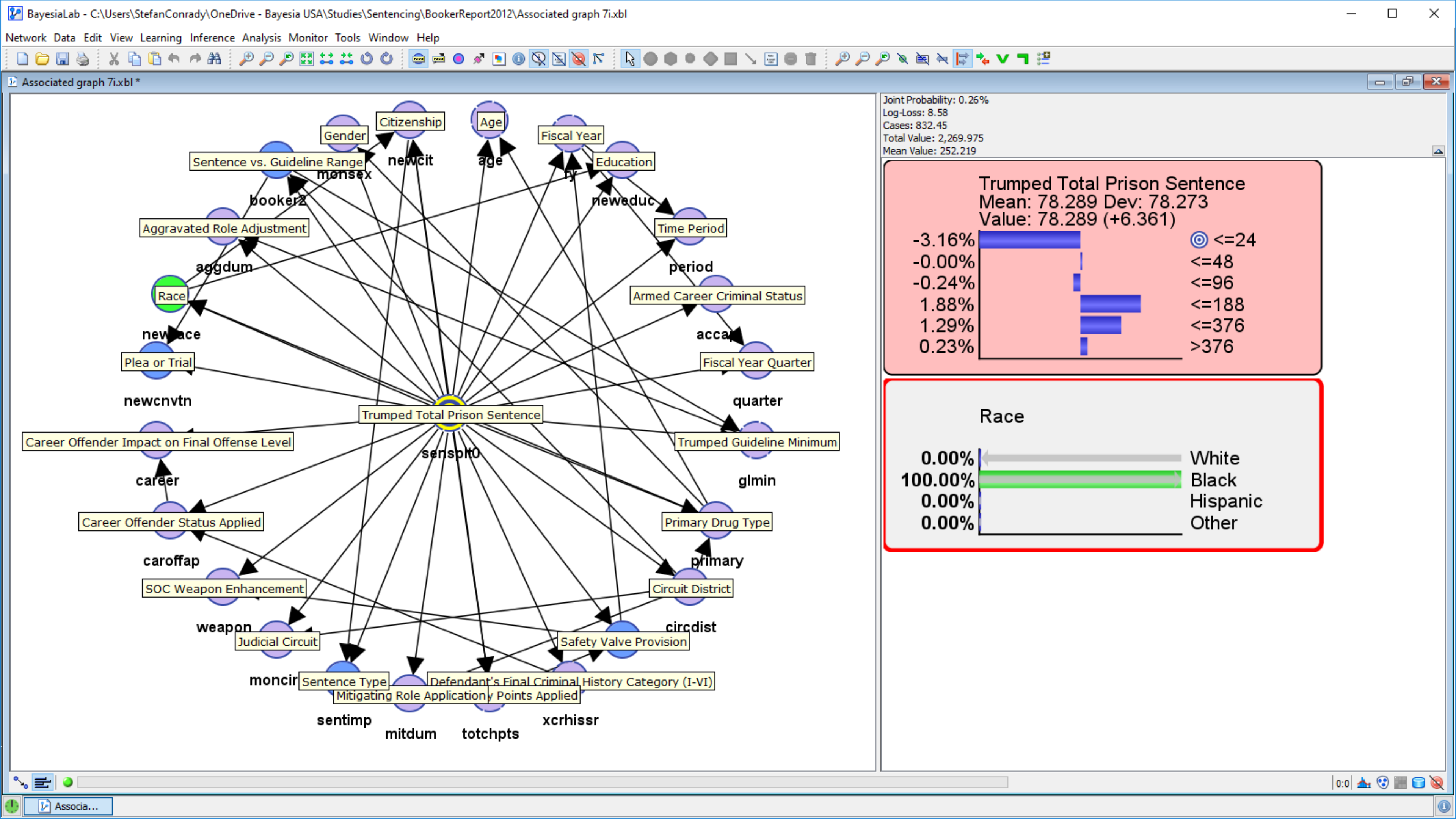

BayesiaLab - C:\Users\StefanConrady\OneDrive - Bayesia USA\Studies\Sentencing\BookerReport2012\Associated graph 7i.xbl
Network Data Edit View Learning Inference Analysis Monitor Tools Window Help
Associated graph 7i.xbl *
Joint Probability: 100.00%
Log-Loss: 0
Cases: 319,172
Total Value: 2,268.394
Mean Value: 252.044
monsex
newcit
age
neweduc
booker2
aggdum
period
newcnvtn
quarter
glmin
career
caroffap
primary
weapon
circdist
moncirc
safevalve
sentimp
mitdum
totchpts
xcrhissr
Associa...


BayesiaLab - C:\Users\StefanConrady\OneDrive - Bayesia USA\Studies\S
Network Data Edit View Learning Inference Analysis Monitor Tools
Associated graph 7i.xbl *
monsex
booker2
aggdum
newcnvtn
career
caroffap
weapon
moncirc
sentimp
Mapping
Citizenship
Age
Gender
Fiscal Year
Sentence vs. Guideline Range
Education
Aggravated Role Adjustment
Time Period
Race
Armed Career Criminal Status
Plea or Trial
Fiscal Year Quarter
Trumped Total Prison Sentence
Trumped Guideline Minimum
Career Offender Impact on Final Offense Level
Career Offender Status Applied
Primary Drug Type
SOC Weapon Enhancement
Circuit District
Judicial Circuit
Safety Valve Provision
Sentence Type
Defendant's Final Criminal History Category (I-VI)
Node Analysis
Size
Mutual Information with the Target Node
Color
None
Arc Analysis
Pearson's Correlation
Close
Save As...
Print
Adjacency Matrix
Associa...

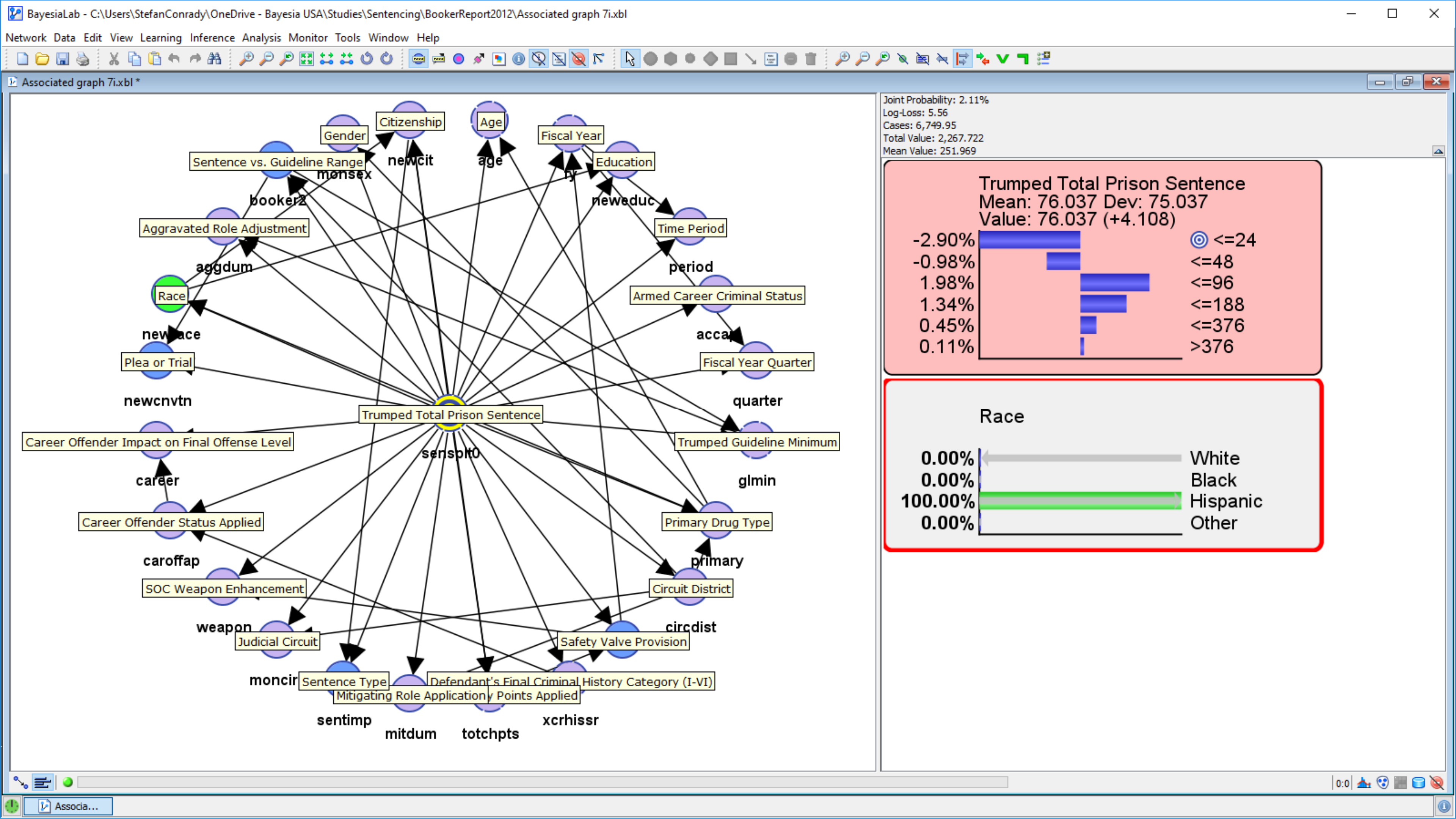

BayesiaLab - C:\Users\StefanConrady\OneDrive - Bayesia USA\Studies\Sentencing\BookerReport2012\Associated graph 7i.xbl
Network Data Edit View Learning Inference Analysis Monitor Tools Window Help
Associated graph 7i.xbl *
Gender
Citizenship
Age
Fiscal Year
Education
Sentence vs. Guideline Range
monsex
newcit
age
fy
booker2
neweduc
Aggravated Role Adjustment
Time Period
aggdum
period
Race
Armed Career Criminal Status
newrace
acca
Plea or Trial
Fiscal Year Quarter
newcnvtn
quarter
Trumped Total Prison Sentence
Career Offender Impact on Final Offense Level
Trumped Guideline Minimum
sensplt0
career
glmin
Career Offender Status Applied
Primary Drug Type
caroffap
primary
SOC Weapon Enhancement
Circuit District
weapon
circdist
Judicial Circuit
Safety Valve Provision
moncir
Sentence Type
Defendant's Final Criminal History Category (I-VI)
Mitigating Role Application
Points Applied
sentimp
mitdum
totchpts
xcrhissr
Joint Probability: 100.00%
Log-Loss: 0
Cases: 319,172
Total Value: 2,268.394
Mean Value: 252.044
Trumped Total Prison Sentence
Mean: 76.709 Dev: 76.644
Value: 76.709
25.89% <=24
20.53% <=48
26.01% <=96
20.35% <=188
6.28% <=376
0.94% >376
Race
25.61% White
28.92% Black
42.71% Hispanic
2.75% Other
Confounders or
Non-Confounders?
Associa...

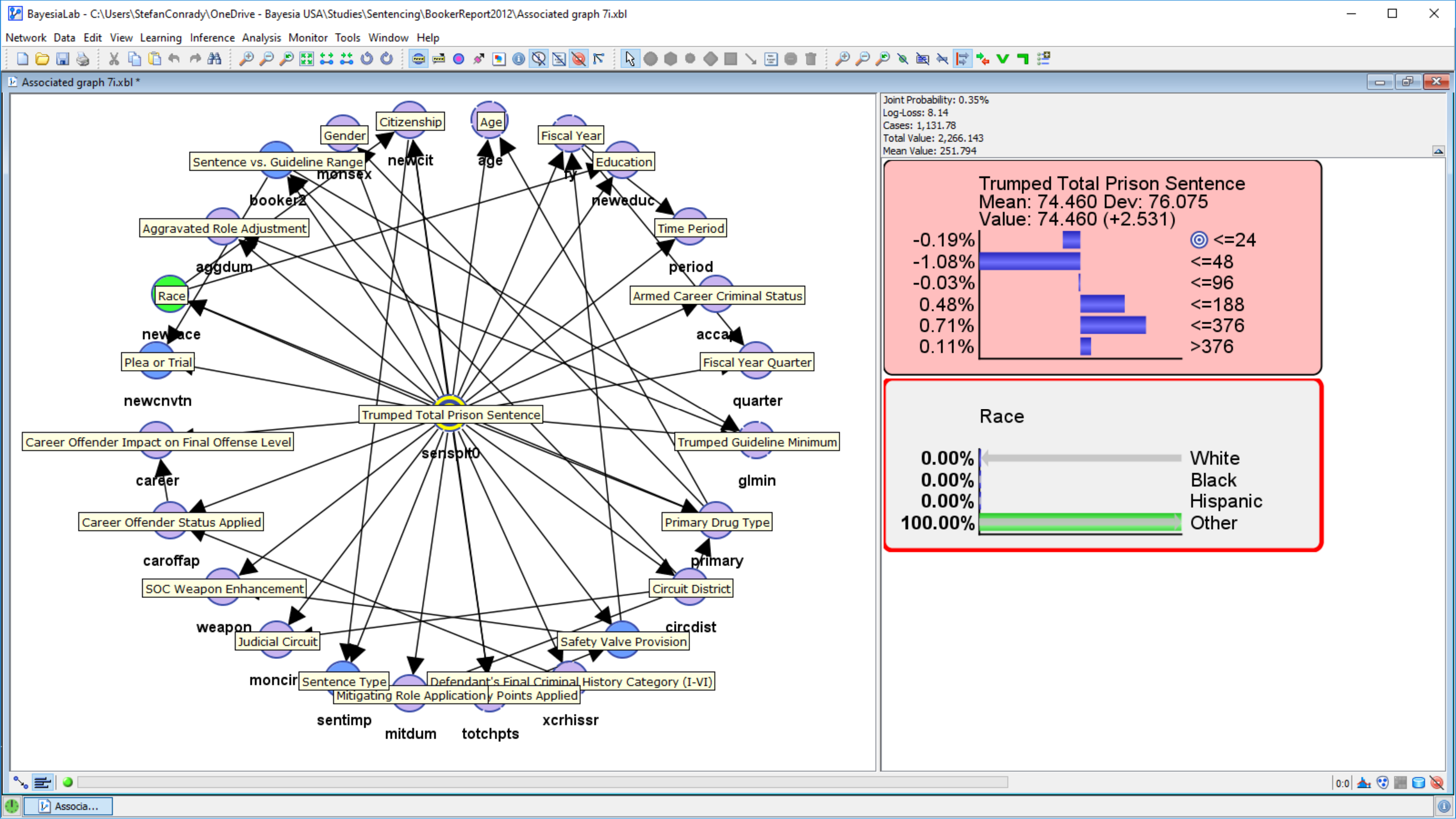

BayesiaLab - C:\Users\StefanConrady\OneDrive - Bayesia USA\Studies\Sentencing\BookerReport2012\Associated graph 7i.xbl
Network Data Edit View Learning Inference Analysis Monitor Tools Window Help
Associated graph 7i.xbl *
Citizenship
newcit
Age
age
Gender
monsex
Fiscal Year
fy
Education
neweduc
Sentence vs. Guideline Range
booker2
Aggravated Role Adjustment
aggdum
Time Period
period
Race
newrace
Armed Career Criminal Status
acca
Plea or Trial
newcnvtn
Fiscal Year Quarter
quarter
Trumped Total Prison Sentence
senspltt0
Career Offender Impact on Final Offense Level
career
Trumped Guideline Minimum
glmin
Career Offender Status Applied
caroffap
Primary Drug Type
primary
SOC Weapon Enhancement
weapon
Circuit District
circdist
Judicial Circuit
moncir
Safety Valve Provision
Sentence Type
sentimp
Mitigating Role Application
mitdum
Defendant's Final Criminal History Category (I-VI)
xcrhissr
Points Applied
totchpts
Joint Probability: 100.00%
Log-Loss: 0
Cases: 319,172
Total Value: 2,268.394
Mean Value: 252.044
Trumped Total Prison Sentence
Mean: 76.709 Dev: 76.644
Value: 76.709
0.00% <=24
0.00% <=48
0.00% <=96
0.00% <=188
0.00% <=376
0.00% >376
Race
25.61% White
28.92% Black
42.71% Hispanic
2.75% Other
Associa...


BayesiaLab - C:\Users\StefanConrady\OneDrive - Bayesia USA\Studies\Sentencing\BookerReport2012\Associated graph 7i.xbl
Network Data Edit View Learning Inference Analysis Monitor Tools Window Help
Associated graph 7i.xbl *
Gender
Citizenship
Age
Fiscal Year
Sentence vs. Guideline Range
newcit
age
Education
monsex
fy
booker2
neweduc
Aggravated Role Adjustment
Time Period
aggdum
period
Race
Armed Career Criminal Status
newrace
acca
Plea or Trial
Fiscal Year Quarter
newcnvtn
quarter
Trumped Total Prison Sentence
Career Offender Impact on Final Offense Level
Trumped Guideline Minimum
sensplt0
career
glmin
Career Offender Status Applied
Primary Drug Type
caroffap
primary
SOC Weapon Enhancement
Circuit District
weapon
circdist
Judicial Circuit
Safety Valve Provision
moncir
Sentence Type
Defendant's Final Criminal History Category (I-VI)
Mitigating Role Application
Points Applied
sentimp
xcrhissr
mitdum
totchpts
Joint Probability: 0.42%
Log-Loss: 7.88
Cases: 1,354.97
Total Value: 2,263.613
Mean Value: 251.513
Trumped Total Prison Sentence
Mean: 71.929 Dev: 73.269
Value: 71.929 (-4.780)
1.74% <=24
1.45% <=48
-0.50% <=96
-1.62% <=188
-0.95% <=376
-0.12% >376
Race
100.00% White
0.00% Black
0.00% Hispanic
0.00% Other
Associa...
0:0



BayesiaLab - C:\Users\StefanConrady\OneDrive - Bayesia USA\Studies\Sentencing\BookerReport2012\Associated graph 7i.xbl
Network Data Edit View Learning Inference Analysis Monitor Tools Window Help
Associated graph 7i.xbl *
Gender
monsex
Citizenship
newcit
Age
age
Fiscal Year
fy
Education
neweduc
Sentence vs. Guideline Range
booker2
Aggravated Role Adjustment
aggdum
Time Period
period
Race
newrace
Armed Career Criminal Status
acca
Plea or Trial
newcnvtn
Fiscal Year Quarter
quarter
Trumped Total Prison Sentence
senspltO
Career Offender Impact on Final Offense Level
career
Trumped Guideline Minimum
glmin
Career Offender Status Applied
caroffap
Primary Drug Type
primary
SOC Weapon Enhancement
weapon
Circuit District
circdist
Judicial Circuit
moncir
Safety Valve Provision
Sentence Type
sentimp
Defendant's Final Criminal History Category (I-VI)
xcrhissr
Mitigating Role Application
mitdum
Points Applied
totchpts
Joint Probability: 0.26%
Log-Loss: 8.58
Cases: 832.45
Total Value: 2,269.975
Mean Value: 252.219
Trumped Total Prison Sentence
Mean: 78.289 Dev: 78.273
Value: 78.289 (+6.361)
-3.16% <=24
-0.00% <=48
-0.24% <=96
1.88% <=188
1.29% <=376
0.23% >376
Race
0.00% White
100.00% Black
0.00% Hispanic
0.00% Other
Associa...

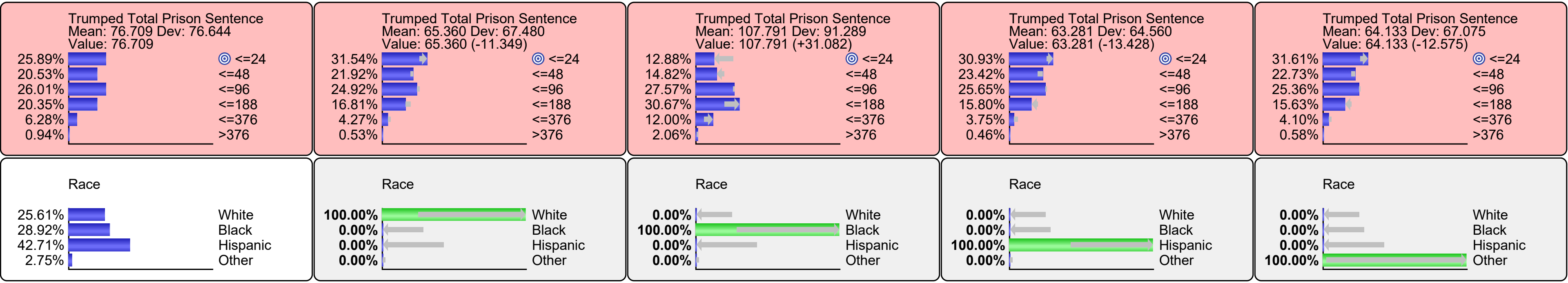

BayesiaLab - C:\Users\StefanConrady\OneDrive - Bayesia USA\Studies\Sentencing\BookerReport2012\Associated graph 7i.xbl
Network Data Edit View Learning Inference Analysis Monitor Tools Window Help
Associated graph 7i.xbl *
Citizenship
newcit
Age
age
Gender
monsex
Fiscal Year
fy
Education
neweduc
Sentence vs. Guideline Range
booker2
Time Period
period
Aggravated Role Adjustment
aggdum
Armed Career Criminal Status
acca
Race
newrace
Fiscal Year Quarter
quarter
Plea or Trial
newcnvtn
Trumped Total Prison Sentence
senspltO
Career Offender Impact on Final Offense Level
career
Trumped Guideline Minimum
glmin
Career Offender Status Applied
caroffap
Primary Drug Type
primary
SOC Weapon Enhancement
weapon
Circuit District
circdist
Judicial Circuit
moncir
Safety Valve Provision
Sentence Type
sentimp
Mitigating Role Application
mitdum
Points Applied
totchpts
Defendant's Final Criminal History Category (I-VI)
xcrhissr
Joint Probability: 2.11%
Log-Loss: 5.56
Cases: 6,749.95
Total Value: 2,267.722
Mean Value: 251.969
Trumped Total Prison Sentence
Mean: 76.037 Dev: 75.037
Value: 76.037 (+4.108)
-2.90% <=24
-0.98% <=48
1.98% <=96
1.34% <=188
0.45% <=376
0.11% >376
Race
0.00% White
0.00% Black
100.00% Hispanic
0.00% Other
Associa...


BayesiaLab - C:\Users\StefanConrady\OneDrive - Bayesia USA\Studies\Sentencing\BookerReport2012\Associated graph 7i.xbl
Network Data Edit View Learning Inference Analysis Monitor Tools Window Help
Associated graph 7i.xbl *
Citizenship
newcit
Gender
monsex
Age
age
Fiscal Year
fy
Education
neweduc
Sentence vs. Guideline Range
booker2
Aggravated Role Adjustment
aggdum
Time Period
period
Race
newrace
Armed Career Criminal Status
acca
Plea or Trial
newcnvtn
Fiscal Year Quarter
quarter
Trumped Total Prison Sentence
senspltO
Career Offender Impact on Final Offense Level
career
Trumped Guideline Minimum
glmin
Career Offender Status Applied
caroffap
Primary Drug Type
primary
SOC Weapon Enhancement
weapon
Circuit District
circdist
Judicial Circuit
moncir
Safety Valve Provision
Sentence Type
sentimp
Mitigating Role Application
mitdum
Defendant's Final Criminal History Category (I-VI)
xcrhissr
Points Applied
totchpts
Joint Probability: 0.35%
Log-Loss: 8.14
Cases: 1,131.78
Total Value: 2,266.143
Mean Value: 251.794
Trumped Total Prison Sentence
Mean: 74.460 Dev: 76.075
Value: 74.460 (+2.531)
-0.19% <=24
-1.08% <=48
-0.03% <=96
0.48% <=188
0.71% <=376
0.11% >376
Race
0.00% White
0.00% Black
0.00% Hispanic
100.00% Other
Associa...

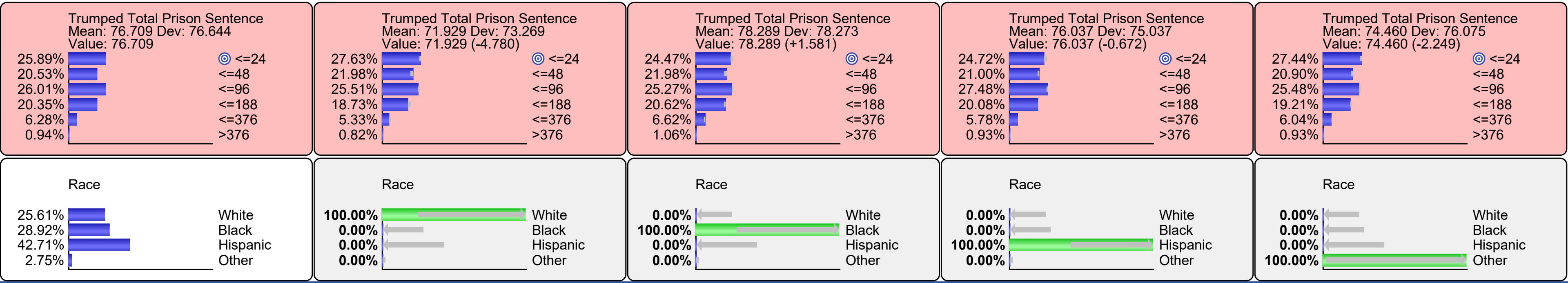
Effects Analysis

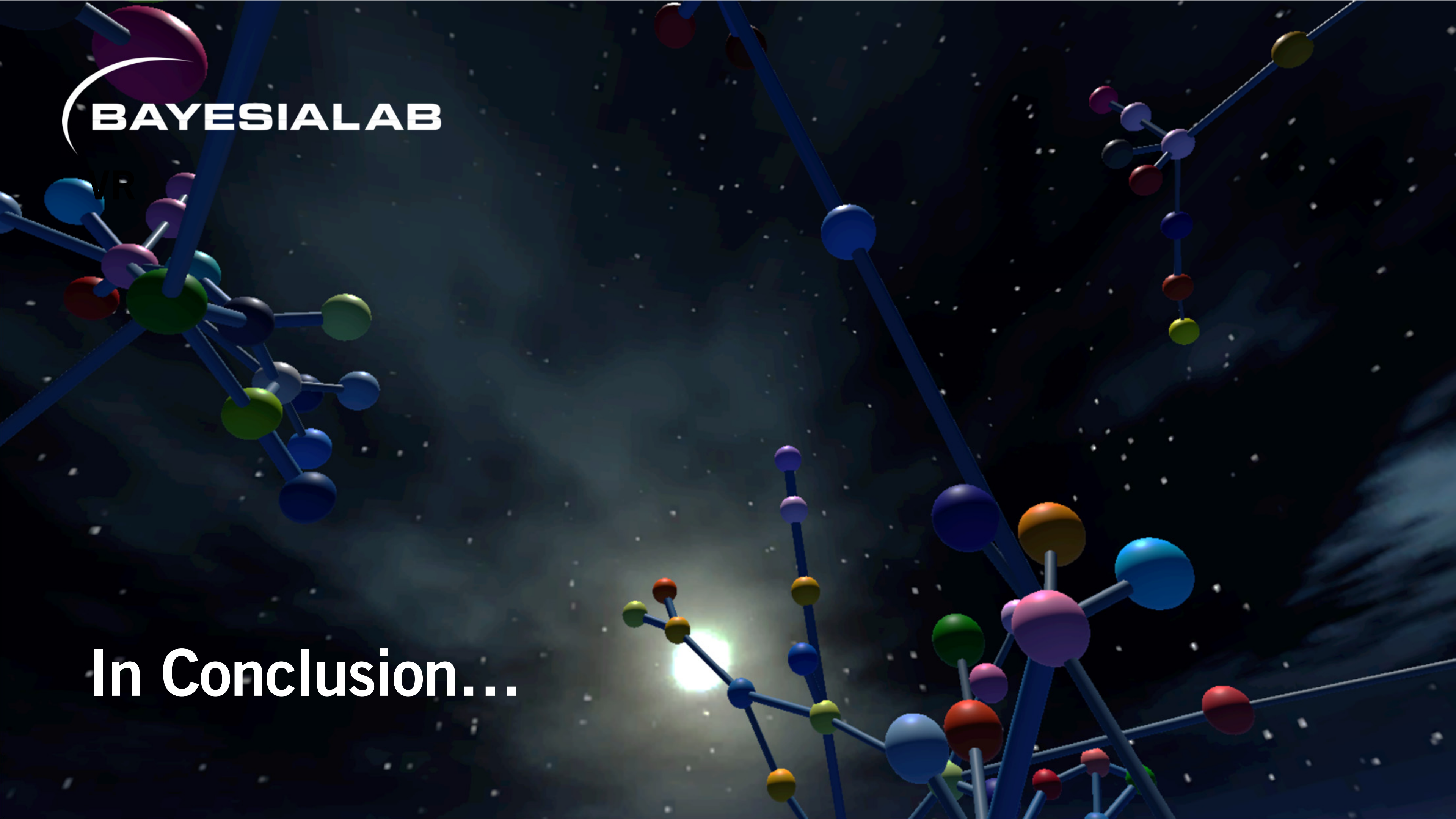
IMPORTANT ASSUMPTION:
NO UNOBSERVED CONFOUNDERS

Observational Inference



Causal Inference





In Conclusion…

Upcoming Events

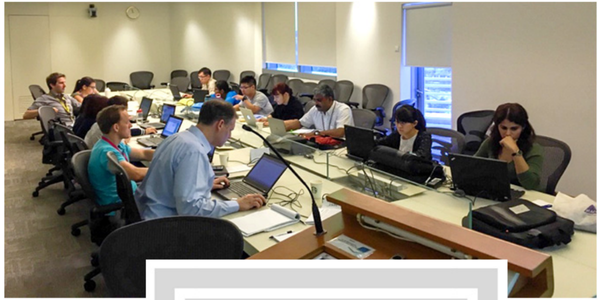
Webinars & Seminars:

- May 18 Webinar: Geographic Optimization Mapping with BayesiaLab
- June1 Webinar: Topic t.b.d.
- June 19 Seminar in Chicago: Knowledge Discovery in Financial Data

Register here: bayesia.com/events

BayesiaLab Courses Around the World in 2018

- May 16–18
  Seattle, WA
- June 26–28
  Boston, MA
- July 23–25
  San Francisco, CA
- August 29–31
  London, UK
- September 26–28
  New Delhi, India
- October 29–31
  Chicago, IL
- December 4–6
  New York, NY



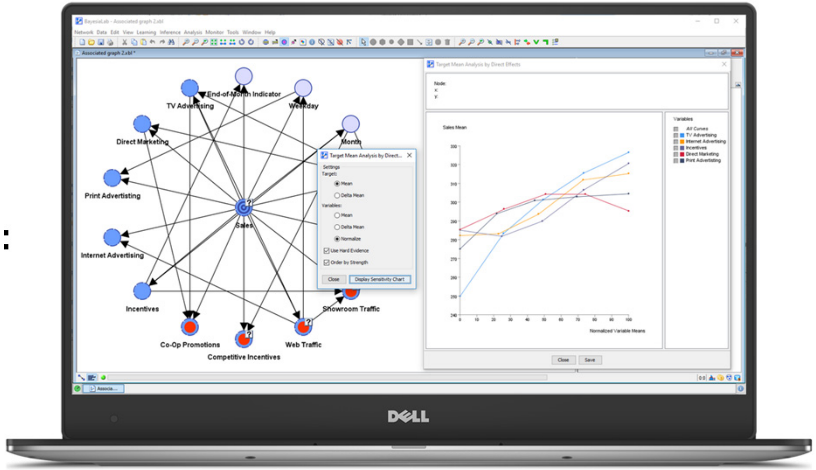
BAYESIALAB
Introductory Course

Learn More & Register: bayesia.com/events

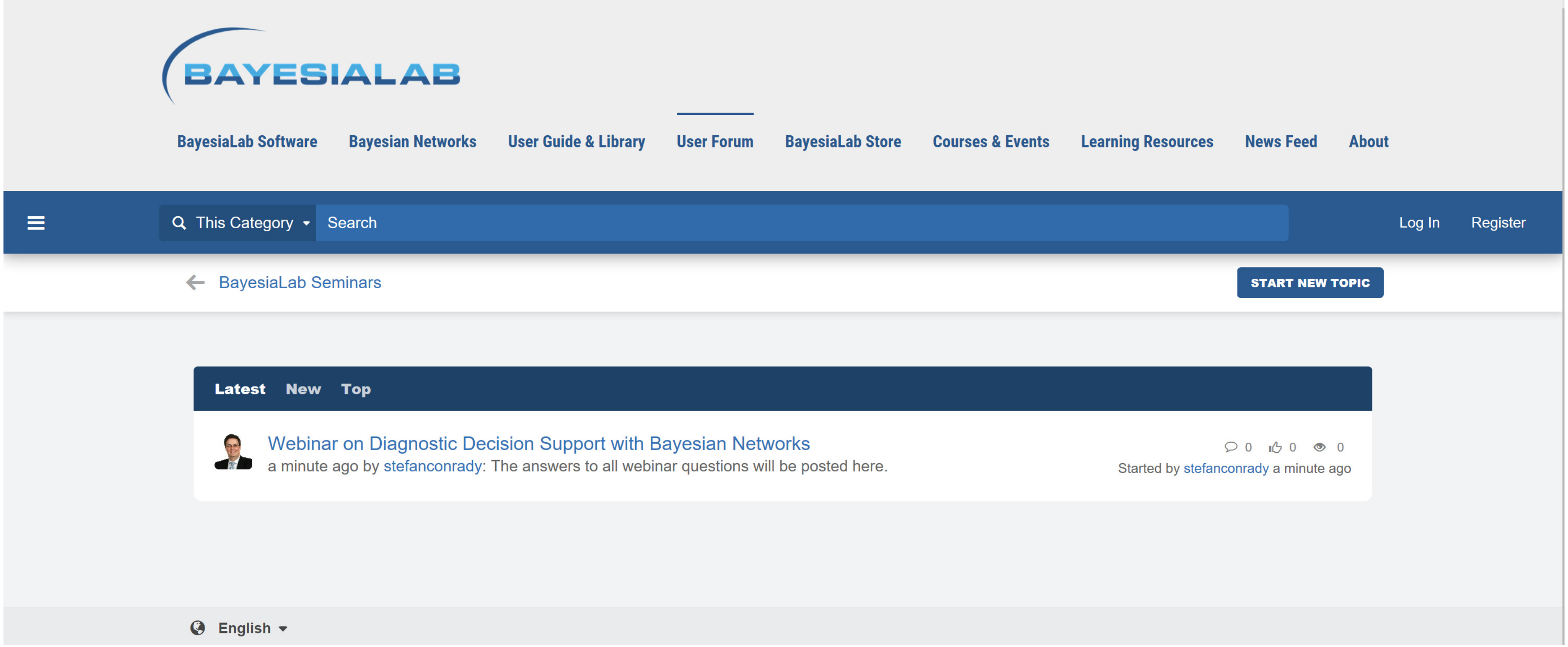
BayesiaLab Trial

Try BayesiaLab Today!

- Download Demo Version:
  www.bayesialab.com/trial-download
- Apply for Unrestricted Evaluation Version:
  www.bayesialab.com/evaluation



User Forum: bayesia.com/community




6th Annual BayesiaLab Conference in Chicago
November 1–2, 2018

Thank You!

stefan.conrady@bayesia.us

BayesianNetwork

linkedin.com/in/stefanconrady

facebook.com/bayesia