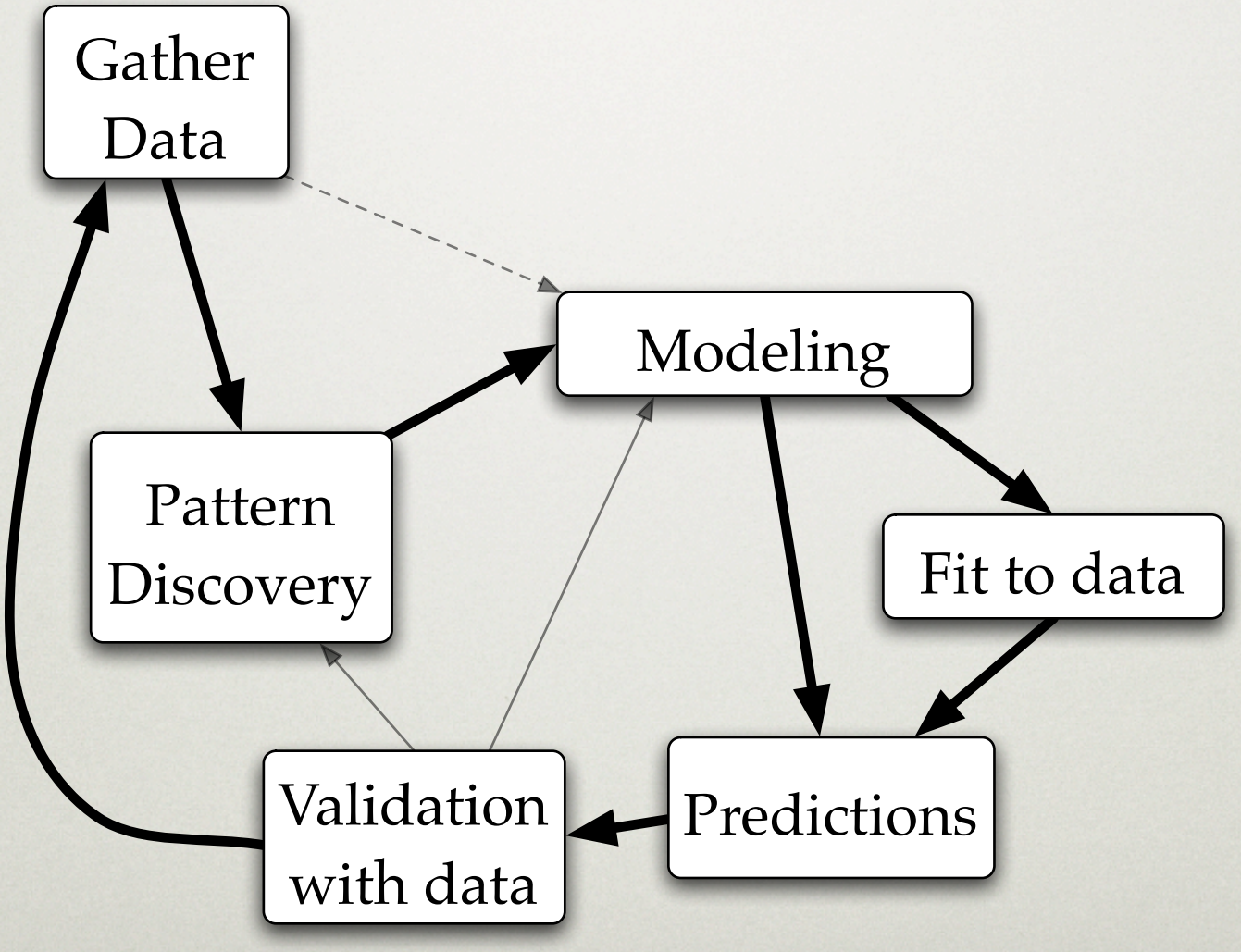
Analyzing and Modeling Global Terrorism

Aaron Clauset
Santa Fe Institute

On Modeling


Gather Data
Modeling
Pattern Discovery
Fit to data
Validation with data
Predictions

Studying Terrorism

What is "terrorism"?

- definitions differ, change over time
- at least as old as ancient Greeks, Persians

Many subjective, conflicting claims, beliefs

- causes, trends, strategies, etc.

Involves human decisions and technology

- political, capricious behavior
- technologically enhanced violence

Studying Terrorism

What we might like to know:

- is terrorism changing?
- are there any interesting patterns?
- underlying similarities, differences worldwide?
- any chance of modeling?
- any chance of prediction, intervention?

Defining "Terrorism"

Definition is important:

- determines which violent acts we study

Terrorism seems different from other violence

- non-state actors
- targets civilians
- instills fear
- for political purposes

Defining "Terrorism"

Definition is important:

- determines which violent acts we study

Terrorism seems different from other violence

- non-state actors
- targets civilians
- instills fear
- for political purposes

} a reasonably narrow definition

Defining "Terrorism"

What constitutes an "event" of terrorism?

- single target
- single location
- single day

} a reasonably narrow definition

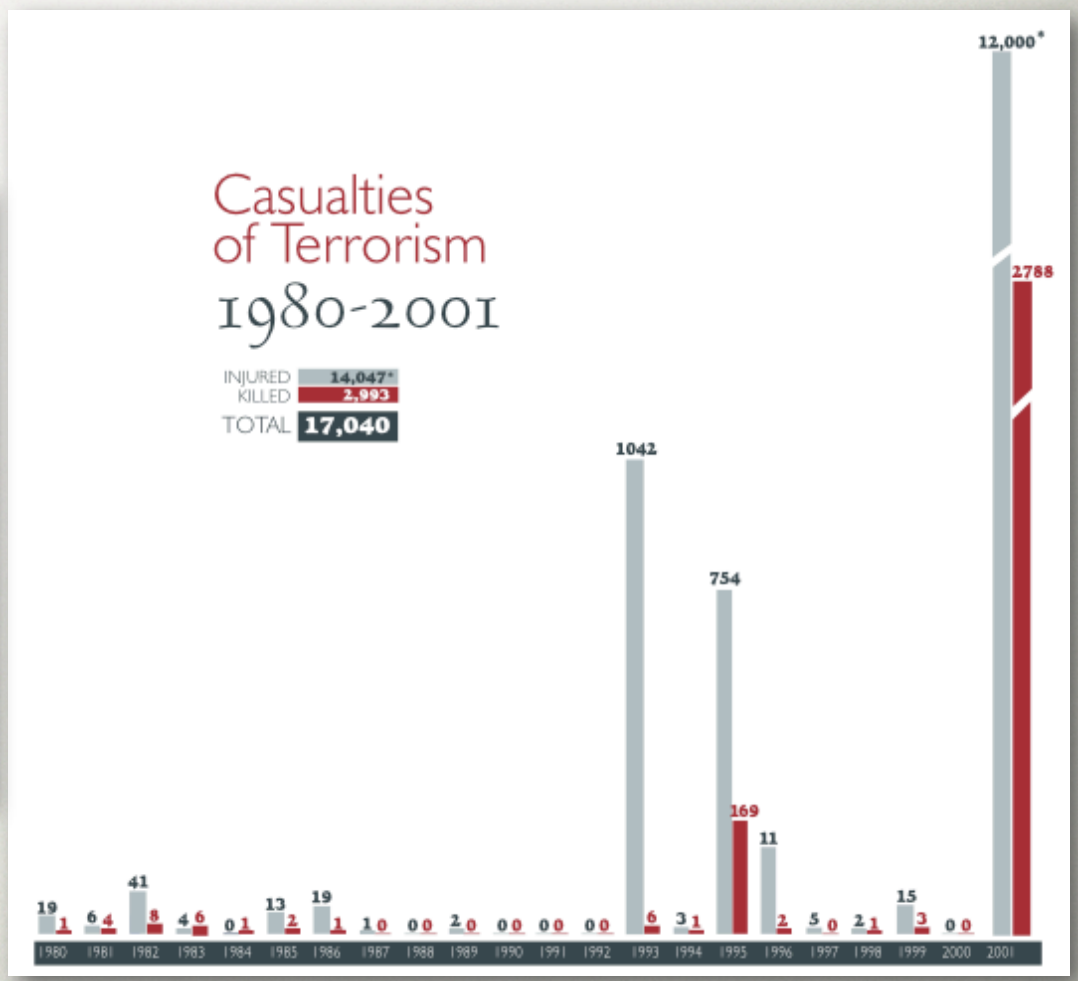
- Hard to make a **big** event (prevents aggregations)
- **Big + rare** events seem inherently interesting

Traditional Analysis

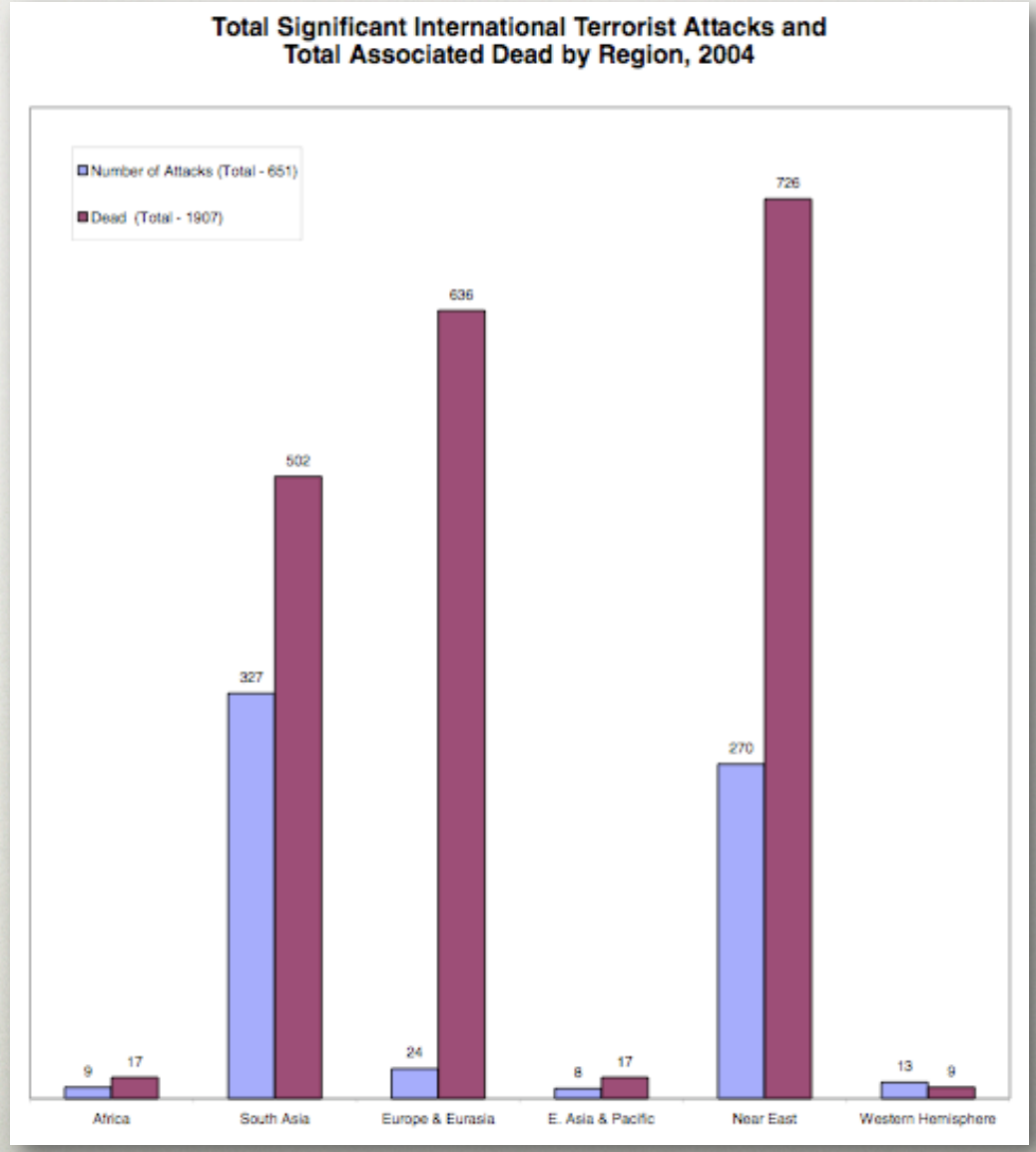
- Domestic vs. Trans-national events
- Aggregated analysis (averages, counts, etc.)
- Categorical trends
  - tactics (hijacking, assassination, bombing, etc.)
  - weapons (chem, bio, guns, “unconventional”, etc.)
  - regions (Americas, Mid East, Asia, etc.)
- **Assume: big events qualitatively different from small events**

Traditional Analysis

source: FBI



Traditional Analysis



source: NCTC Report

Traditional Analysis

Regression:

- choose dependent variable Y (e.g., incidence)
- enumerate and quantify "important" independent variables X (e.g., "freedom", income, region, etc.)
- regress on X to predict Y

$$\epsilon + \sum_{i=1}^{n} \alpha_i X_i^{\beta_i} = Y$$

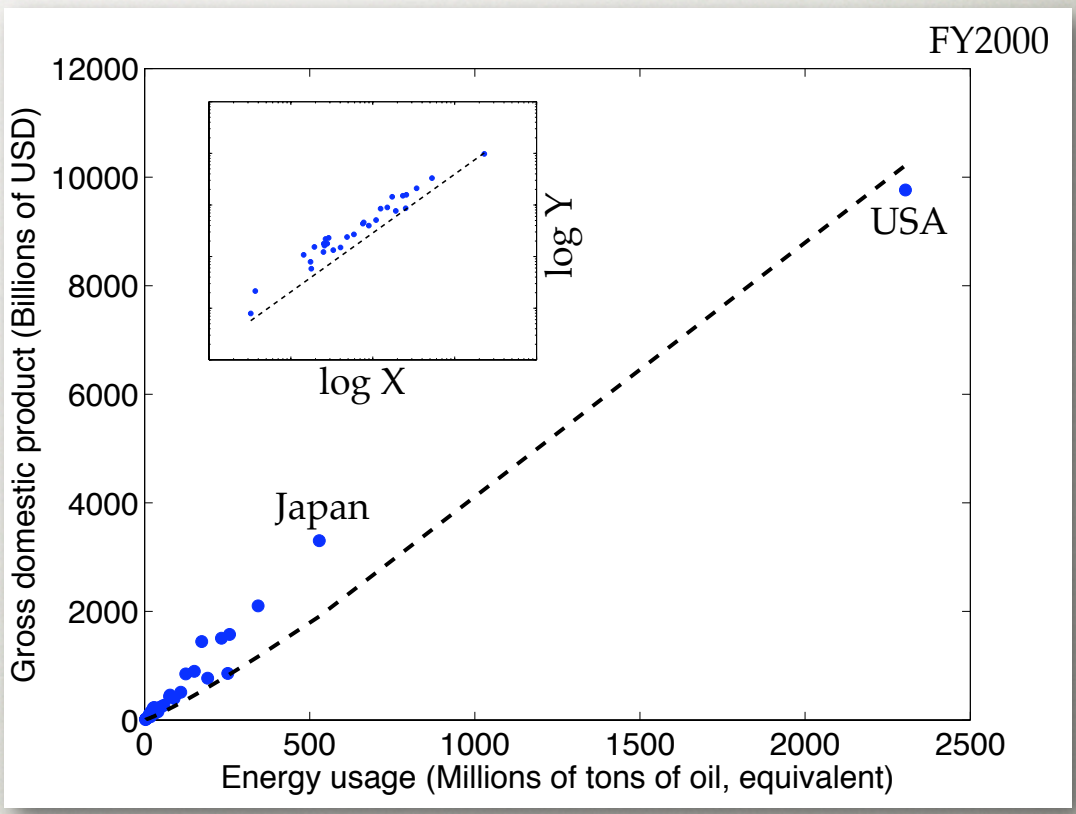
An Example

Data from economically developed countries

- predict GDP (Y) from energy usage (X)
- regression result:

$$Y = 1.5\, X^{1.14}$$

$$r^2 = 0.91$$



data source: oecd.org

Caveats

Regression:

$$\sum_{i=1}^{n} \alpha_i \, X_i^{\beta_i} = Y - \epsilon$$

- assumes linear model with normal errors

  (many variations exist)

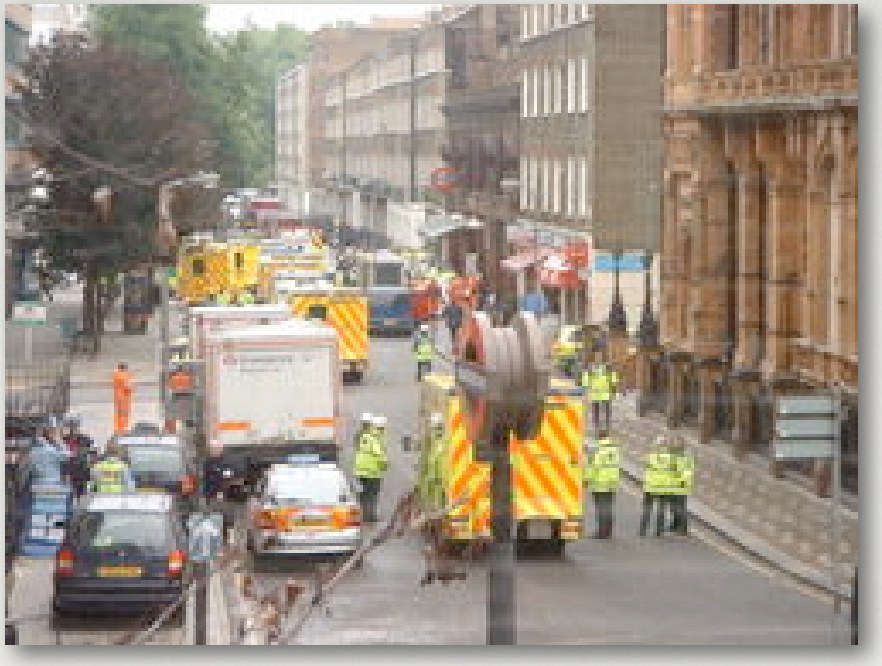
- typically sensitive to "outliers", variable selection
- r^2 value only as good as assumptions
- thinking typically not required "garbage in = garbage out"

On Severe Events

Big / severe / "outlier" events

- relatively few casualties (automobiles > terrorism)
- very infrequent
- disproportionate destabilizing effect
    - economic and political impacts

For example: major re-organization of US/UK national security apparatus after 9.11



London, 2006

NYC, 2001

Mumbai, 2006

sources: wikipedia, BBC

Oklahoma City, 1995

Bali, 2002

Lockerbie, 1988

sources: wikipedia

Terrorism Data

narrow event definition

40 years data (1968 - 2007)

32,829 events, worldwide

~5600 cities, 187 countries

14,062 with casualty > 0

MIPT TERRORISM KNOWLEDGE BASE SM

incident profile

ABU HAFS AL-MASRI BRIGADE AND SECRET ORGANIZATION OF AL-QAEDA IN EUROPE ATTACKED TRANSPORTATION TARGET (JULY 7, 2005, UNITED KINGDOM)

Incident Date: July 7, 2005

Terrorist Organization(s): Abu Hafs al-Masri Brigade , Secret Organization of al-Qaeda in Europe

Target: Transportation

City: London

Country: United Kingdom

Region: Western Europe

Tactic: Bombing

Weapon: Explosives

Fatalities: 27

Injuries: 0

alternative sources: ITERATE, GTD, ICT, etc.

source: tkb.org

Terrorism Data

narrow event definition

40 years data (1968 - 2007)

32,829 events, worldwide

~5600 cities, 187 countries

14,062 with casualty > 0

Event severity:
number of casualties,
deaths, or injuries

MIPT TERRORISM KNOWLEDGE BASE℠

incident profile

ABU HAFS AL-MASRI BRIGADE AND SECRET ORGANIZATION OF AL-QAEDA IN EUROPE ATTACKED TRANSPORTATION TARGET (JULY 7, 2005, UNITED KINGDOM)

Incident Date: July 7, 2005

Terrorist Organization(s): Abu Hafs al-Masri Brigade , Secret Organization of al-Qaeda in Europe

Target: Transportation

City: London

Country: United Kingdom

Region: Western Europe

Tactic: Bombing

Weapon: Explosives

Fatalities: 27

Injuries: 0

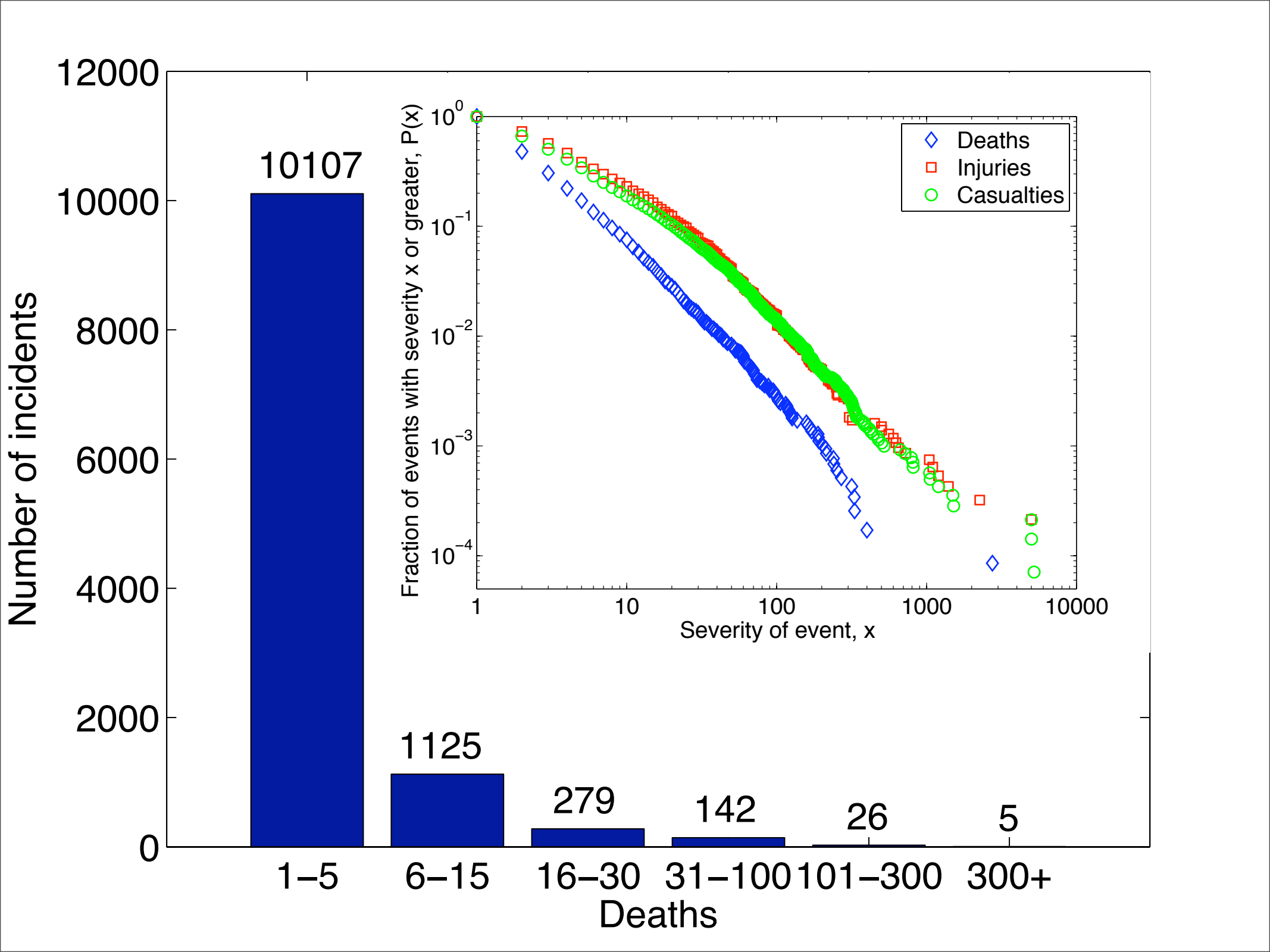
alternative sources: ITERATE, GTD, ICT, etc.

source: tkb.org

Alternative Approach

- Statistical modeling (fitting *probability* models to data)
  - data exploration -- looking for interesting patterns
  - no assumptions about causal relationships
  - model "agnostic"
    - explanatory modeling comes later
    - few / no predictions at this stage

[main article at arXiv:physics/0606007]


12000
10000
8000
6000
4000
2000
0
Number of incidents
10107
1125
279
142
26
5
1–5
6–15
16–30
31–100
101–300
300+
Deaths
Fraction of events with severity x or greater, P(x)
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
1
10
100
1000
10000
Severity of event, x
Deaths
Injuries
Casualties

Basics

	$\bar{x}$	σ	$x_{\max}$
deaths	4.22	28.21	2749
injuries	12.25	85.28	5000
casualties	11.15	83.28	5213

Basics

	$\bar{x}$	σ	$x_{\max}$
deaths	4.22	28.21	2749
injuries	12.25	85.28	5000
casualties	11.15	83.28	5213

- std much bigger than mean $\sigma > \bar{x}$

Basics

	$\bar{x}$	σ	$x_{\max}$
deaths	4.22	28.21	2749
injuries	12.25	85.28	5000
casualties	11.15	83.28	5213

- std much bigger than mean $\sigma > \bar{x}$
- max much bigger than expected $x_{\max} > \bar{x} + 3\sigma$

Basics

	$\bar{x}$	σ	$x_{\max}$
deaths	4.22	28.21	2749
injuries	12.25	85.28	5000
casualties	11.15	83.28	5213

- std much bigger than mean $\sigma > \bar{x}$
- max much bigger than expected $x_{\max} > \bar{x} + 3\sigma$
- only 880 (8%) events (deaths) with $x \geq 10$
- heavy-tailed distribution, **power law**?

Fitting and Testing the Hypothesis

Fit power-law distribution to data $P(x) = \left(\frac{x}{x_{\min}} \right)^{\alpha - 1}$

- estimate parameters $x_{\min}$ and α from the data
- test plausibility of estimated model (p-value)

[see arXiv:0706.1062 for methodology]

	$\hat{x}_{\min}$	$\hat{\alpha}$	n_{tail}	p
deaths	10	2.39	880	**0.6**
injuries	35	2.48	618	0.0
casualties	46	2.43	588	**0.9**

What does it mean?

power-law is reasonable model for deaths, casualties

- big events not fundamentally different from small events
- power-law may accurately predict risk of future severe events (additional validation needed for this)
- can use power-law model for additional analysis
  - normally blunt tool
  - with a little care, can be delicate probe

Variations

How does frequency-severity distribution vary with

- time
- weapon type
- economic development

Variation by Time

Study events in each 24 month interval

MIPT TERRORISM KNOWLEDGE BASE℠

incident profile

ABU HAFS AL-MASRI BRIGADE AND SECRET ORGANIZATION OF AL-QAEDA IN EUROPE ATTACKED TRANSPORTATION TARGET (JULY 7, 2005, UNITED KINGDOM)

Incident Date: July 7, 2005

Terrorist Organization(s): Abu Hafs al-Masri Brigade , Secret Organization of al-Qaeda in Europe

Target: Transportation

City: London

Country: United Kingdom

Region: Western Europe

Tactic: Bombing

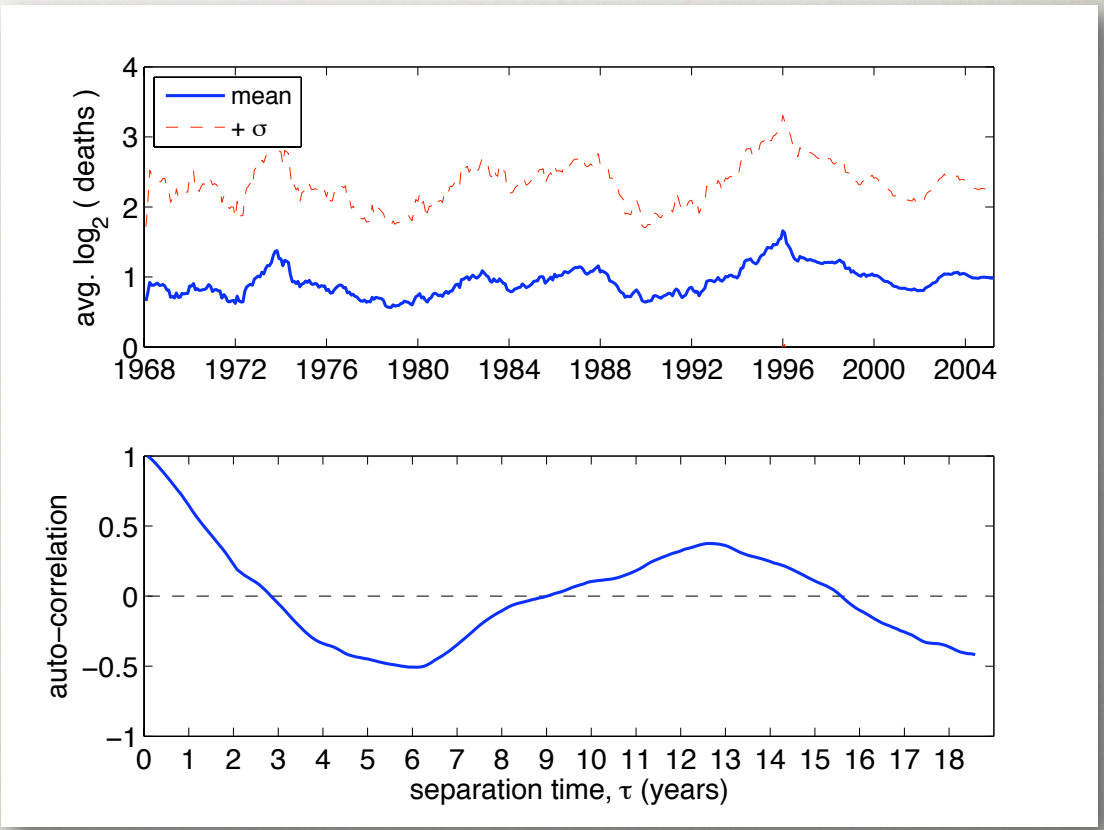
Weapon: Explosives

Fatalities: 27

Injuries: 0

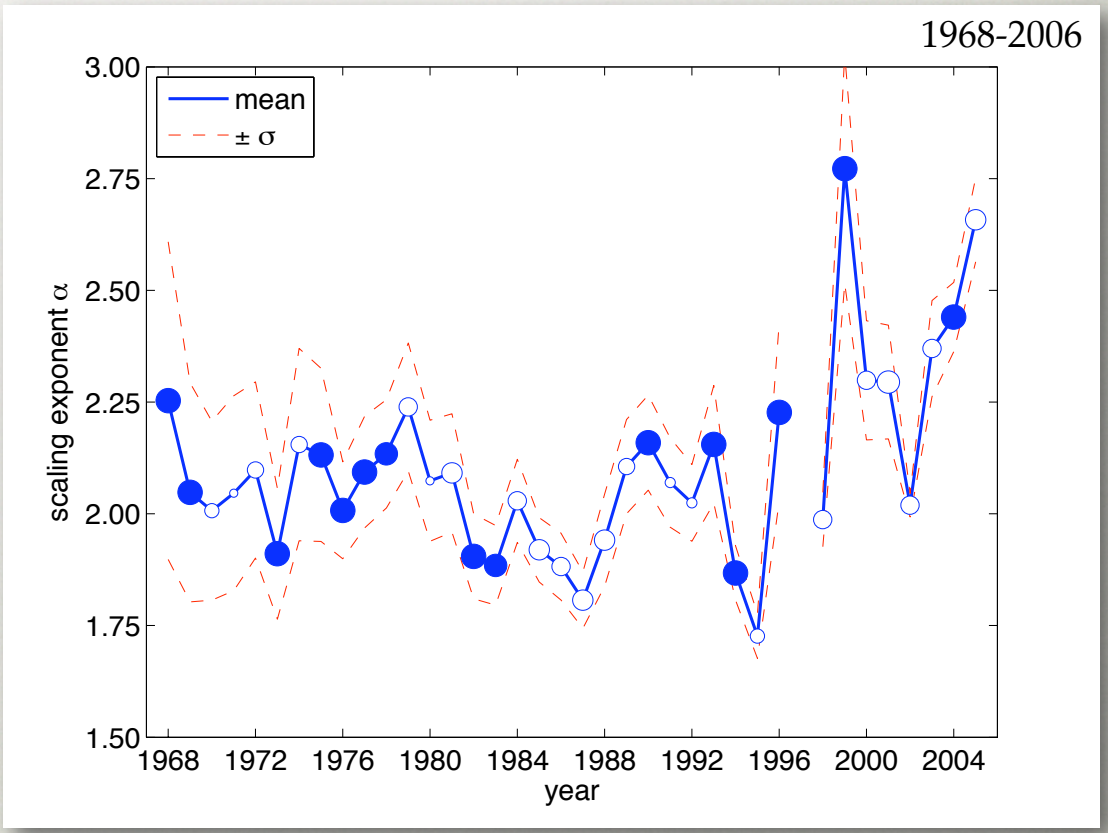
Variation by Time

- Ave. log-severity largely stable over 40 years
- Apparent periodicity in ave. log-severity at $\tau \approx 13$ years



Variation by Time

- Scaling exponent largely stable over 40 years
- Suggests severity distribution largely stable
- Main difference today: many more events



Variation by Weapon

Types:

- chem/bio
- explosives
- fire/arson
- firearms
- knives
- other/unconventional

MIPT TERRORISM KNOWLEDGE BASE℠

incident profile

ABU HAFS AL-MASRI BRIGADE AND SECRET ORGANIZATION OF AL-QAEDA IN EUROPE ATTACKED TRANSPORTATION TARGET (JULY 7, 2005, UNITED KINGDOM)

Incident Date: July 7, 2005

Terrorist Organization(s): Abu Hafs al-Masri Brigade , Secret Organization of al-Qaeda in Europe

Target: Transportation

City: London

Country: United Kingdom

Region: Western Europe

Tactic: Bombing

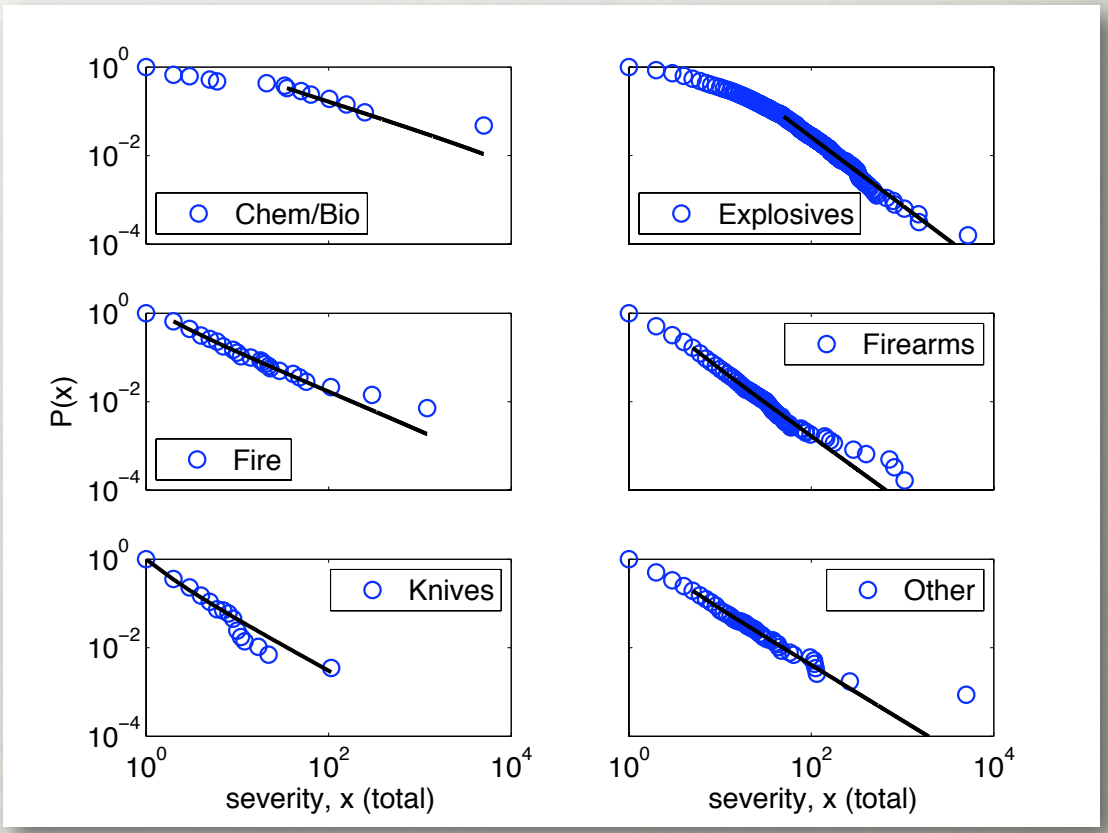
Weapon: Explosives

Fatalities: 27

Injuries: 0

Variation by Weapon

- More apparent power-law behavior
- But different $\hat{\alpha}$, $\hat{x}_{\min}$
- Not ubiquitous: no power laws by region
- Explosives most deadly, overall



Variation by Economy

Organization for Economic Co-operation and Development

- 30 countries (USA, Japan, France, UK, Turkey...)
- tracks economic statistics and data for these 30 + 70 others

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT OECD

MIPT TERRORISM KNOWLEDGE BASE℠

incident profile

ABU HAFS AL-MASRI BRIGADE AND SECRET ORGANIZATION OF AL-QAEDA IN EUROPE ATTACKED TRANSPORTATION TARGET (JULY 7, 2005, UNITED KINGDOM)

Incident Date: July 7, 2005

Terrorist Organization(s): Abu Hafs al-Masri Brigade , Secret Organization of al-Qaeda in Europe

Target: Transportation

City: London

Country: United Kingdom

Region: Western Europe

Tactic: Bombing

Weapon: Explosives

Fatalities: 27

Injuries: 0

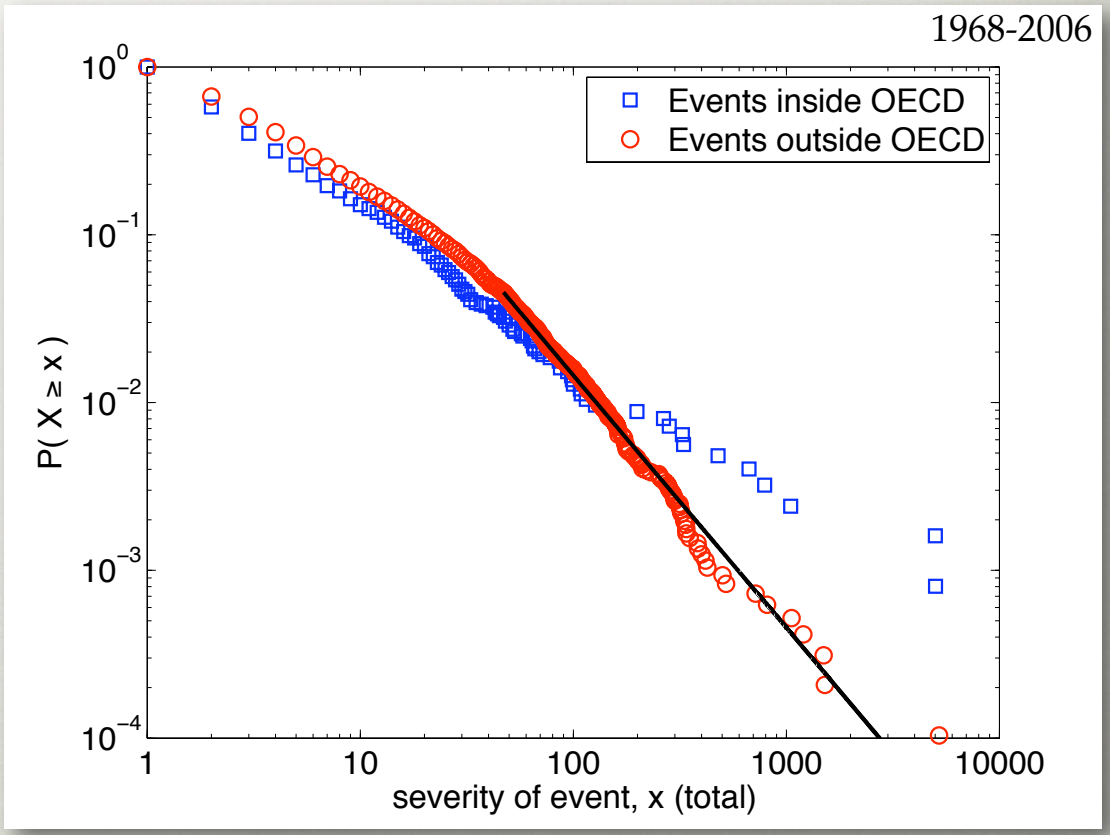
Variation by Economy

non-OECD countries

- all events more frequent
- severe events less common

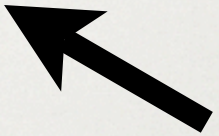
OECD countries

- all events less frequent
- severe events more common



Economy Alone?

	$x \geq x_{\min}$	of total
Turkey	335	26.9%
France	201	16.2%
Spain	109	8.8%
Germany	98	7.9%
USA	93	7.5%
Greece	76	6.1%
Italy	73	5.9%
UK	62	5.0%
total	1047	84.2%



other factors must be involved

Modeling Questions

What causes power-law distribution?

- Hyp: partly state vs. terrorist competition

What causes apparent $\tau \approx 13$ year periodicity?

What factors account for OECD / non behavior?

- Hyp: territorial disputes?

What role for technology, population density?

- E.g., lethality of explosives varies dramatically with location of detonation

On Modeling (redux)

Discovery Phase

data quality

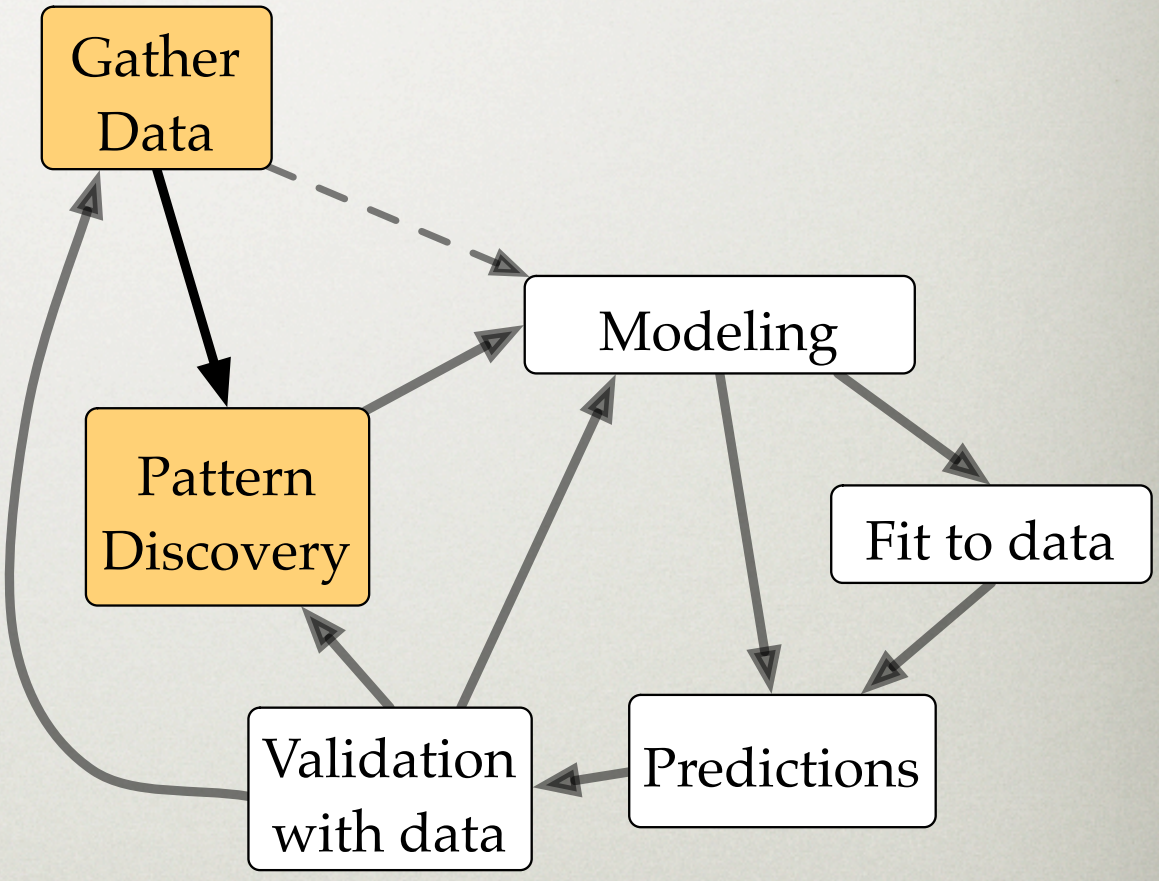
- consistency, etc.

select questions

- scope of study

data analysis

- distributions, statistics, trends



Modeling Phase

data models

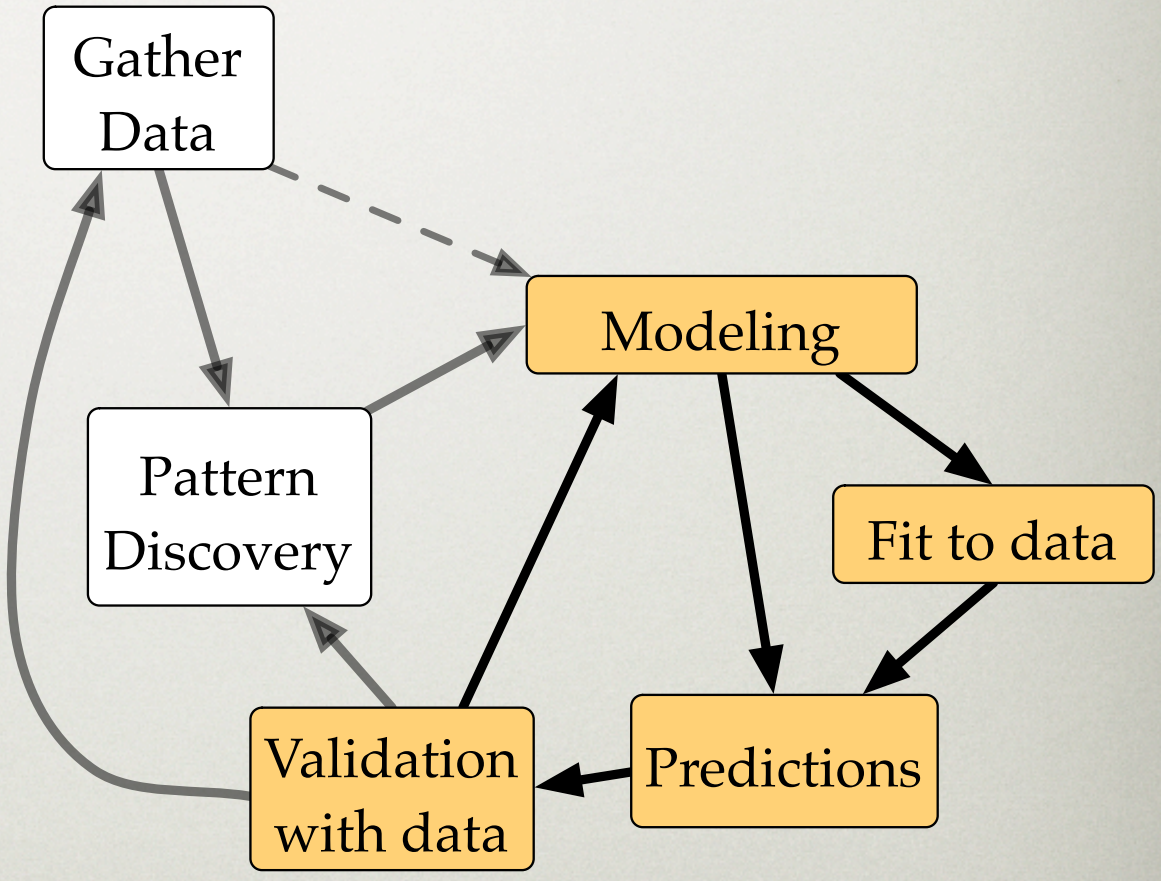
- e.g., distributions

mechanistic models

- microscopic rules

predictive models

- falsifiable
- e.g., regression and machine learning models



Closing Thoughts

Discovery phase

- know your data
- what questions are reasonable, interesting?
- practice good statistics

Modeling phase

- why, how this structure and not other structure?
- falsifiable predictions ultimate goal
- good statistics necessary for validation step

FIN