Seminar BINGO!

To play, simply print out this bingo sheet and attend a departmental seminar.

Mark over each square that occurs throughout the course of the lecture.

The first one to form a straight line (or all four corners) must yell out to win!

SEMINAR BINGO

Speaker bashes previous work	Repeated use of "um…"	Speaker sucks up to host professor	Host Professor falls asleep	Speaker wastes 5 minutes explaining outline
Laptop malfunction	Work ties in to Cancer/HIV or War on Terror	"…et al."	You're the only one in your lab that bothered to show up	Blatant typo
Entire slide filled with equations	"The data *clearly* shows…"	FREE Speaker runs out of time	Use of Powerpoint template with blue background	References Advisor (past or present)
There's a Grad Student wearing same clothes as yesterday	Bitter Post-doc asks question	"That's an interesting question"	"Beyond the scope of this work"	Master's student bobs head fighting sleep
Speaker forgets to thank collaborators	Cell phone goes off	You've no idea what's going on	"Future work will…"	Results conveniently show improvement

WWW.PHDCOMICS.COM

Changing People's Minds: Understanding Social Diffusion Dynamics Using Networked Cognitive Systems

September 1, 2016

Daniel Chen, MPH
Computational. Network. Epidemiology.

VirginiaTech
Biocomplexity Institute

Social & Decision Analytics Laboratory

The Social and Decision Analytics Laboratory

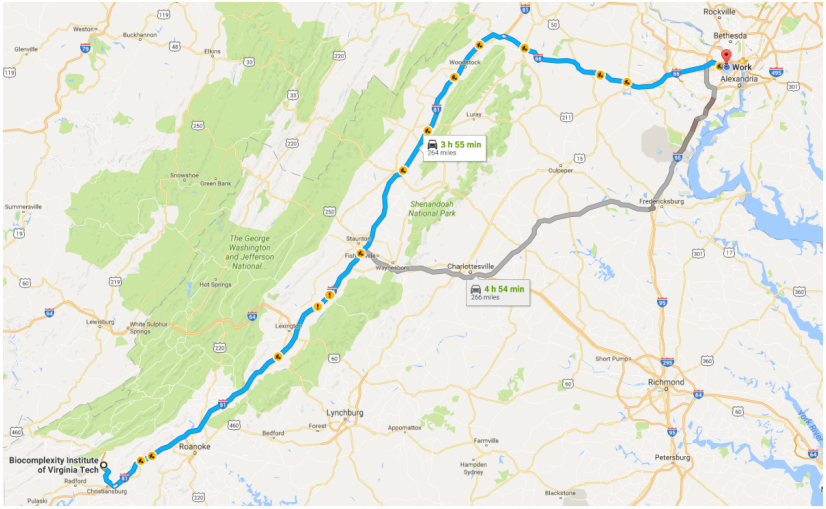
Founded in 2013

Sallie Keller
Director
Professor

Stephanie Shipp
Deputy Director
Research Professor

statisticians and social and behavioral scientists to embrace todays data revolution, developing evidence-based research and quantitative methods to inform policy decision-making.

The Social and Decision Analytics Laboratory



People

Research Faculty

1. David Higdon (Statistics)
2. Viki Lancaster (Statistics)
3. Mark Orr (Psychology)
4. Aaron Schroeder (Data Science)
5. Gizem Korkmaz (Economics)

Post Doctoral Associates

1. Kathryn Ziemer (Psychology)
2. Bianica Pires (Social Science)
3. Emily Molfino (Political Science)
4. Joshua Goldstein (Statistics)

Visiting Scholars & Collaborators: 7

Students

1. Daniel Chen (GRA: GBCB)
2. Adrienne Rogers (VT: Statistics)
3. Emily Stark (Austin Peay: Mathematics)

Administrative Staff

1. Kimberly Lyman
2. Tracie Hase

Projects

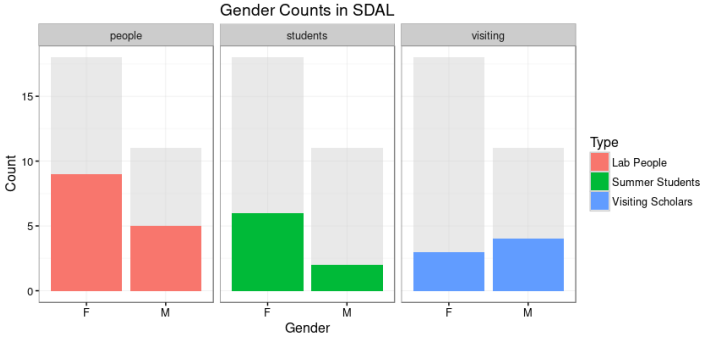
1. Assessing New Data Sources for the Federal Census
2. Simulating Urban Air Pollution Exposure
3. Collaborating to Build a Culture of Health
4. Leveraging Data to Enhance Emergency Response
5. Practicing Data Science for the Public Good
6. Modeling the Spread of Beliefs Through Social Media

Data Science for the Public Good

- First summer for the The Data Science for the Public Good program (DSPG)
- Research Experience for Undergraduates (REU) funded by NSF
- Bash, SSH, LaTeX, R, SQL, GIS, Literate Programming, Web Scraping, Git

My Students! :D

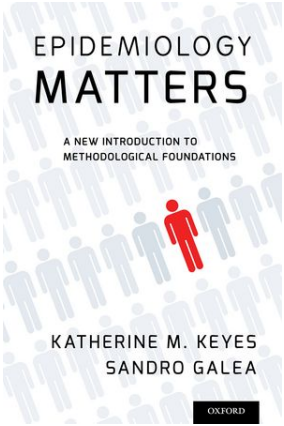
Come Visit!


Gender Counts in SDAL
people
students
visiting
Count
15
10
5
0
F
M
Gender
Type
Lab People
Summer Students
Visiting Scholars

Modeling the Spread of Beliefs Through Social Media

Epidemiology

Epidemiology is the science of population health, aiming to ***understand*** *the key causes of health and disease and doing so in a way that it may inform* ***interventions*** *so we may act*



Epidemiology

- Infectious Diseases (Ebola, Zika, Measles, etc)
- **Chronic Diseases** (diabetes, obesity, smoking, etc)

Columbia MPH Certificates ∼ GBCB

- **Advanced Epidemiology**
- **Applied Biostatistics**
- Child, Youth, and Family Health
- Climate and Health
- **Comparative Effectiveness Outcomes Research**
- Environmental Health Policy
- **Epidemiology of Chronic Disease**
- Global Health
- Health and Human Rights
- Health of an Aging Society
- Health Policy Analysis
- Health Policy and Practice
- Health Promotion Research and Practice
- History, Ethics, and Law
- **Infectious Disease Epidemiology**
- **Injury Prevention and Control**
- Molecular Epidemiology
- Public Health and Humanitarian Assistance
- **Public Health Informatics**
- **Public Health Research Methods**
- Sexuality, Sexual and Reproductive Health
- **Social Determinants of Health**
- Toxicology

(Computational) (Infectious) Disease Modeling

Two main types of models:

1. Compartmental/Mathematical/System Dynamics/ Ordinary Differential Equation (ODE) Models
2. Agent-Based Models (ABM)

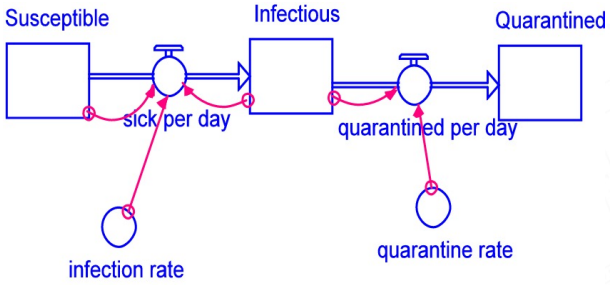
SIR Models

- **S**usceptible
- **I**nfectious
- **R**ecovered

$$\frac{dS}{dt} = -bS(t)I(t)$$

$$\frac{dI}{dt} = bS(t)I(t) - kI(t)$$

$$\frac{dR}{dt} = kI(t)$$



Example SIR model from Stella
(Maryland Virtual High School)

Compartmental Models

- SIS
- SEIR
- MSIR
- SI(CR)
- Vaccine

Pros

- Deterministic
- Overall System Dynamics
- Homogeneous (random mixing)
- Simple, easy to set up

Cons

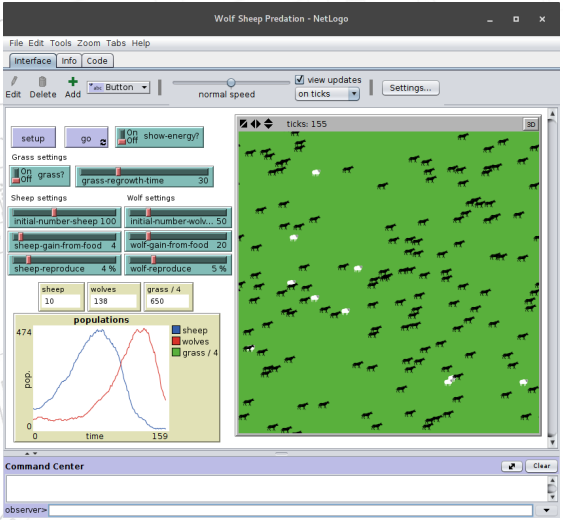
- Not stochastic
- No individual level behaviors
- No complex interactions

Agent-Based Models

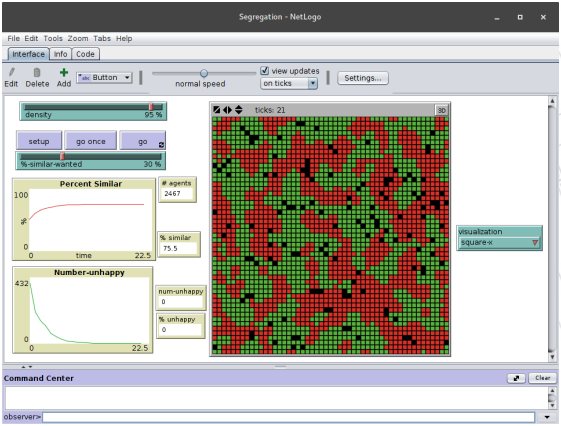
- The model is composed of individual ‘agents’
- Each agent has a set of rules
- The agents repeat these rules (ticks/cycles)

Observe complex system dynamics from the bottom-up through **emergence**

ABM Examples

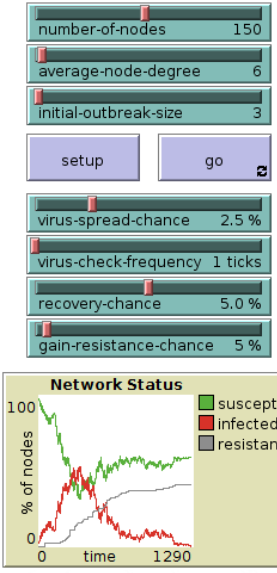


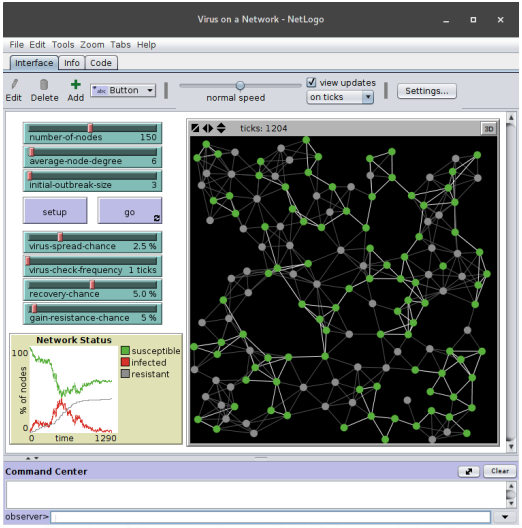
Netlogo Wolf Sheep Predation Model



Netlogo Segregation Model

Virus on a Network Model





Agent-Based Models

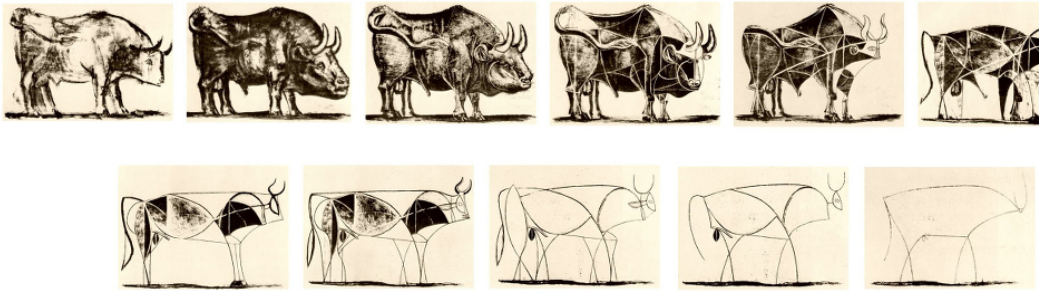
Pros

- Stochastic
- Consolidate knowledge for agent rules/behaviors
- Heterogeneous

Cons

- Needs a lot of data and time to set up
- Harder to get general system dynamics
- Needs a lot of model runs
- Resource-intensive

ABMs to Compartmental Models



Pablo Picasso, “Bull”, Plates 1-11 (Lithograph)

Initial Conception

- Mark Orr, PhD.
- Columbia University Mailman School of Public Health
  - Department of Epidemiology
    - Sandro Galea, MD, MPH, DrPH
  - Columbia University Systems Science Program (CUSSP)
    - Mailman School of Public Health (MSPH)
    - School of Engineering and Applied Sciences (SEAS)
    - Complex System Approaches in Population Health (CSAPH)
- Modeling obesity and attitude formation for teen sexual behaviors

Duncan Watts, PhD

- Currently: Principal Researcher at Microsoft Research
- Advisor: Steven Strogatz (Cornell)
- Professor of Sociology at Columbia University

1. Six Degrees: The Science of a Connected Age (2003)
2. Small Worlds: The Dynamics of Networks between Order and Randomness (1999)
3. Everything is Obvious: Once You Know The Answer (2011)

Watts 2002 Paper

A simple model of global cascades on random networks (2002)

- Binary Decisions with Externalities (general contagion model)
  - fads, riots, crime, competing technologies, spread of innovation, conventions, and cooperation
- Cascades
  - Probability of a global cascade from a single node
- Local dependencies, fractional threshold, and heterogeneity

Definitions

Blog post:

http://chendaniely.github.io/research/2016/08/31/a_simple_model_of_global_cascades_on_random_networks/

- **Cascades**: event of any size triggered by an initial seed
- **Global cascades**: a cascade that occupies a finite fraction of an infinite network. A sufficiently large cascade. More than a fixed fraction of a large, but finite network.
- **Local dependencies**: agents will incorporate information about its neighbors
- **Fractional threshold**: agents themselves a threshold that determine how it incorporates information from its neighbors
- **Heterogeneity**: every agent is different to varying degree from one another

Analogy from the paper

Diffusion of innovations:

- innovators $\sim$ initial seed
- early adopters $\sim$ vulnerable vertices (nodes)

- A cascade will occur if innovators are connected to many early adopters (connectivity).
- More early adopters, higher chance of innovation, but they need to be connected (structure).

Simulation Runs

1. n agents in a network start off with a state of 0
2. Individual agents can only have a state that is either 0 or 1
3. Each agent has k neighbors
4. An agent gets a new state of 1 if a fraction of its neighbors, ϕ, are also 1
   - Otherwise an agent gets a new state of 0.
5. During each time step, the population evolves:

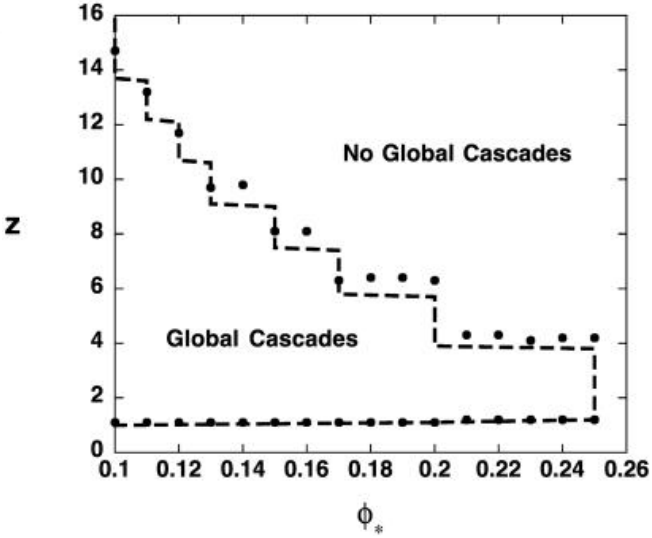
5.1 Update states in random, asynchronous order using the threshold rule

5.2 Once an agent has a state of 1, it will stay at 1 for the remainder of the simulation

ϕ and k are 2 parameters we can change

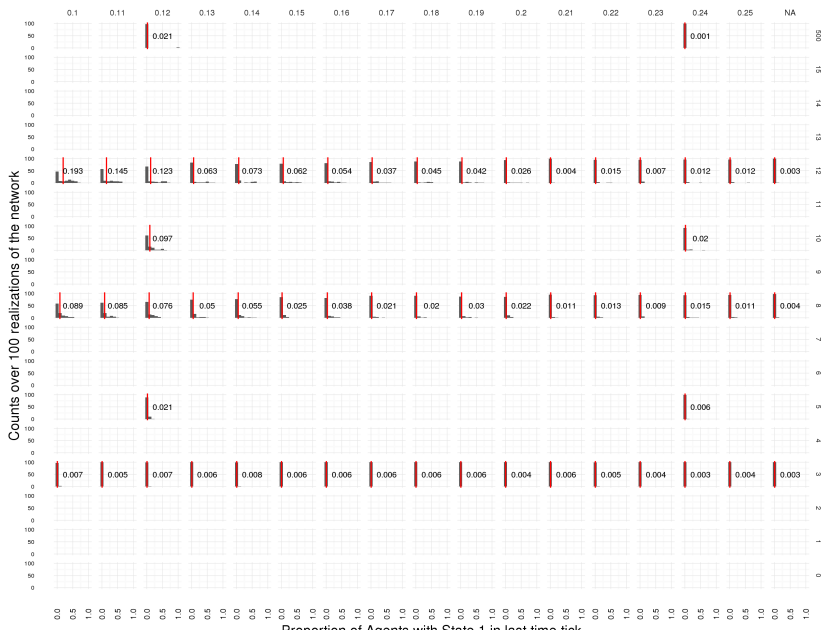
Simulation Parameterization

1. ϕ and k may be heterogeneous
   — To simplify the simulations, the paper has a homogeneous threshold, ϕ
2. The network is a uniform random graph
3. A small seed
4. Any pair of vertices is connected with probability $p = \frac{z}{n}$
   — in a uniform random graph, $p_k =$ Poisson distribution
5. $n = 10,000$
6. 100 random runs of each simulation



Watts 2002 Figure 1

Larger Image: https://github.com/chendaniely/gbcb_seminar_presentation_1/raw/master/figures/p_flipped_all.png



Expand the Watts Model

The Watts model can be used to model any **binary outcome**. From a **public health** and **epidemiology** perspective, this outcome can be a particular **behavior or action**.

- I ate a cookie. Yes/No

However, our decision making process is not that simple.

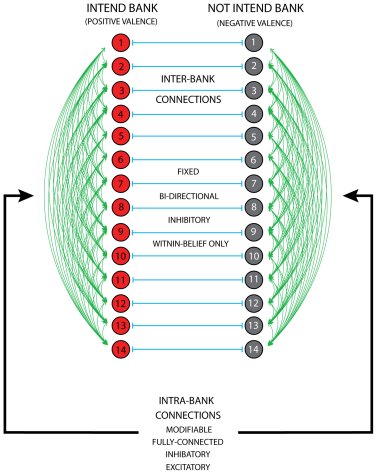
Theory of Reasoned Action (TRA)

Health behavior model

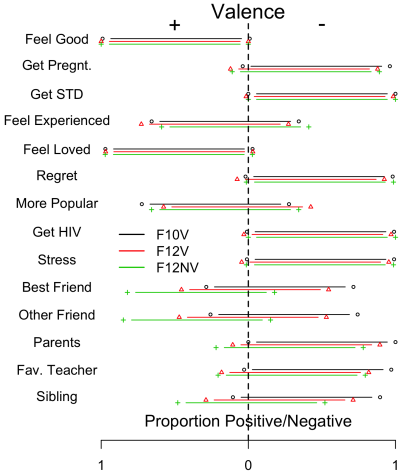
- Martin Fishbein and Icek Ajzen 1967
- Behavior is determined by an individual's intention
- Intention comes from an individual's attitudes and social context
- Attitudes and social context originates from a set of beliefs

beliefs $\rightarrow$ (attitudes & social context) $\rightarrow$ intention $\rightarrow$ behavior

TRA as Parallel Constraint Satisfaction



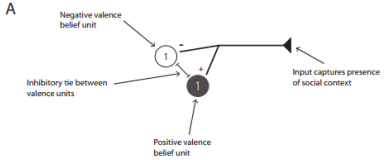
Orr 2013 Figure 1



Orr 2013 Figure 2

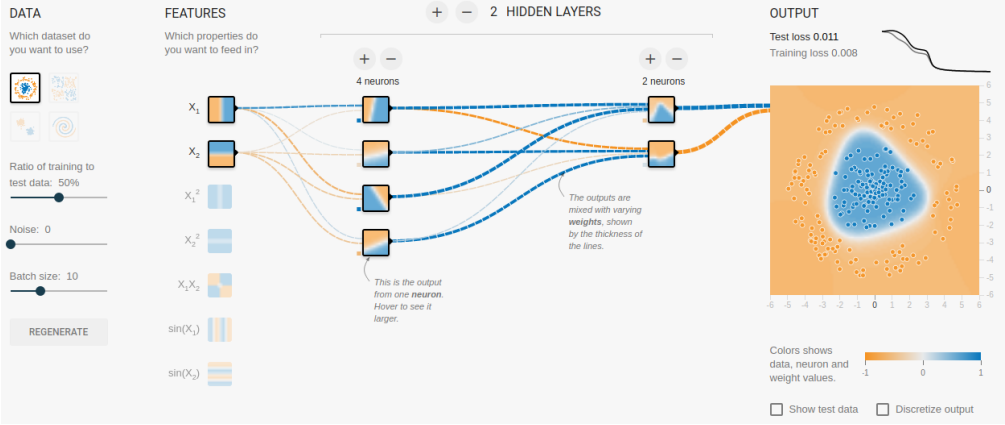
Neural Network

- **Logistic regression**
- The perceptron is a classic 'simple' neural network.
  - — Feed-forward nerual network
  - — **Recurrent**
  - — "deep learning"
    - AlphaGo vs Lee Sedol
    - Self-driving cars
- Psychological plausible decisions
  - — social processes
  - — experience/ memory
  - — influences/ dynamic



Orr 2014 Figure 1

TensorFlow Visualization



http://playground.tensorflow.org/

Results

Forthcoming Publication

- parameter sweep on NN connections
- no learning during simulation
- 7750 Simulations
- associativity
  - clustering of intention at the end of the simulation

Future Directions and Questions

NSF

- E-cigarettes
- Twitter

NYT: Shooting Scares Show a Nation Quick to Fear the Worst

Questions:

- How to summarize the NN Simulation data
  - Signal Decomposition/analysis
  - t-SNE projections to look for community movement

Forthcoming Works

1. M. Orr, K. Zeimer, and D. Chen (Forthcoming, Fall 2016). *Systems of Behavior and Human Health*. In S. Galea & A. El-Sayed (Eds.), Systems Science and Population Health. Oxford University Press: Oxford, UK.

2. M. Orr and D. Chen (Forthcoming, Fall 2016). *Computational Models of Health Behavior*. In R. Vallacher, A. Nowak, and S. Read (Eds.), Computational Models in Social Psychology. Psychology Press/Routledge: New York.

Thanks!

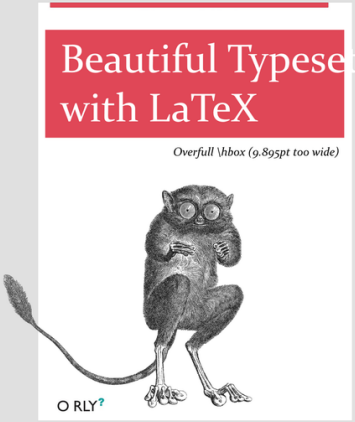
- Dennie Munson
- Social and Decition Analytics Laboratory
- Mark Orr, PhD (VT)
- Aaron Schroeder, PhD (VT)
- David Higdon, PHD (VT)
- Jacqueline Merrill PhD, MPH, RN, FAAN, FACMI (CU Nursing)

#1520359

GBCB!

Thanks!


LATEX
YOUR PAPER MAKES NO GODDAMN SENSE, BUT IT'S THE MOST BEAUTIFUL THING I HAVE EVER LAID EYES ON.


Beautiful Typesetting with LaTeX
Overfull \hbox (9.895pt too wide)
O RLY?

References I

Thanks!

Slides:
https://github.com/chendaniely/gbcb_seminar_presentation_1
Code:

- https://github.com/chendaniely/mann2
- https://github.com/chendaniely/mann2_simulations
- https://github.com/chendaniely/mann2_analysis

@chendaniely

www.github.com/chendaniely