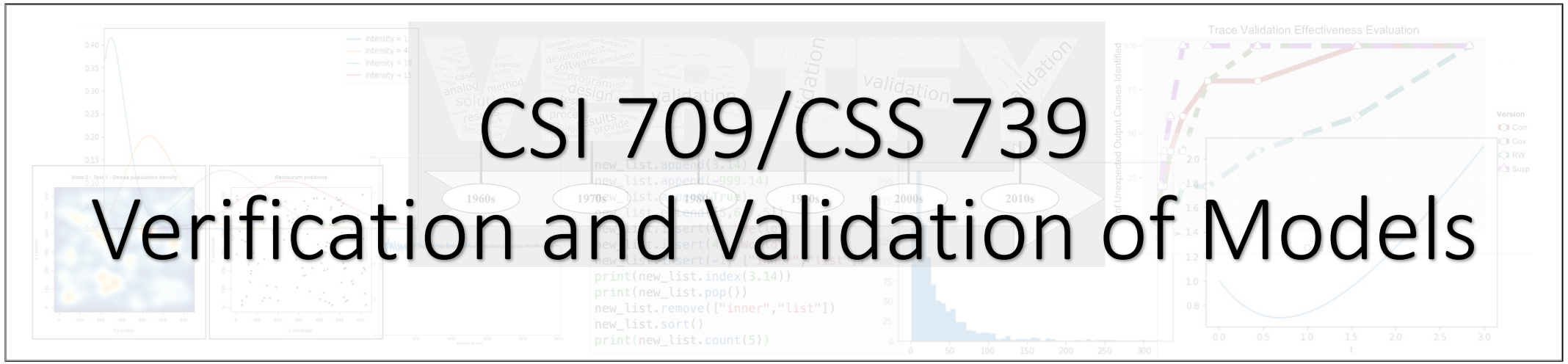


CSI 709/CSS 739
Verification and Validation of Models

Models in Science

Dr. Hamdi Kavak
Computational and Data Sciences Department

http://www.hamdikavak.com
hkavak@gmu.edu

GEORGE MASON UNIVERSITY

Center for Social Complexity

Goals of this lecture

- To illuminate the characteristics of (computational) models.
- To articulate why we need computational models.
- Develop an understanding why all models are wrong and why we need verification and validation efforts to credible models.

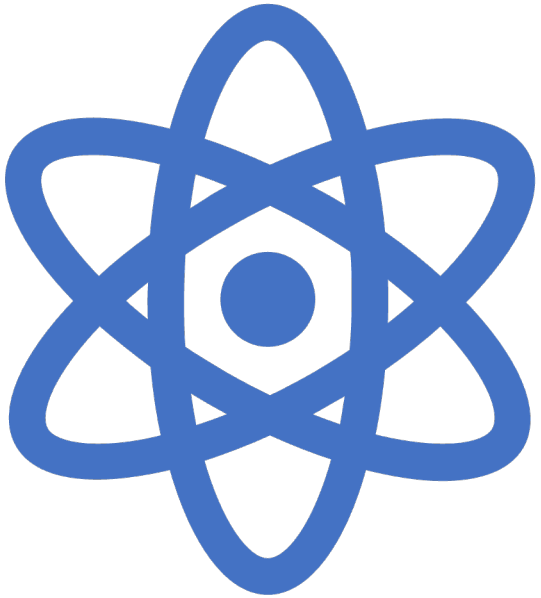
Models

- “something built or drawn esp. to show how something much larger would look”
- “a representation of something in words or numbers that can be used to tell what is likely to happen if particular facts are considered as true”
- “a description or analogy used to help visualize something (such as an atom) that cannot be directly observed”
- “a system of postulates, data, and inferences presented as a mathematical description of an entity or state of affairs”

Sources: https://dictionary.cambridge.org/us/dictionary/english/model and https://www.merriam-webster.com/dictionary/model

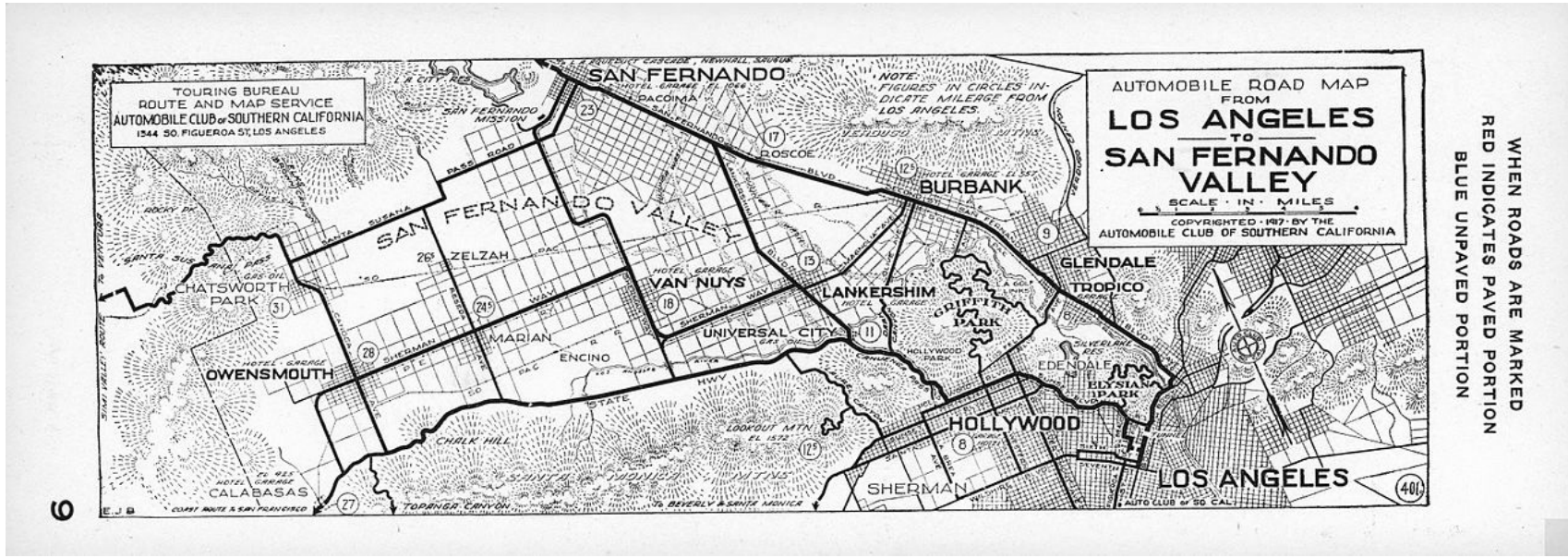
Models

- **Represents** something real or imaginary.
- Built for a **purpose**
- Has **detail** or **abstraction level**



Models in daily life

- Printed maps



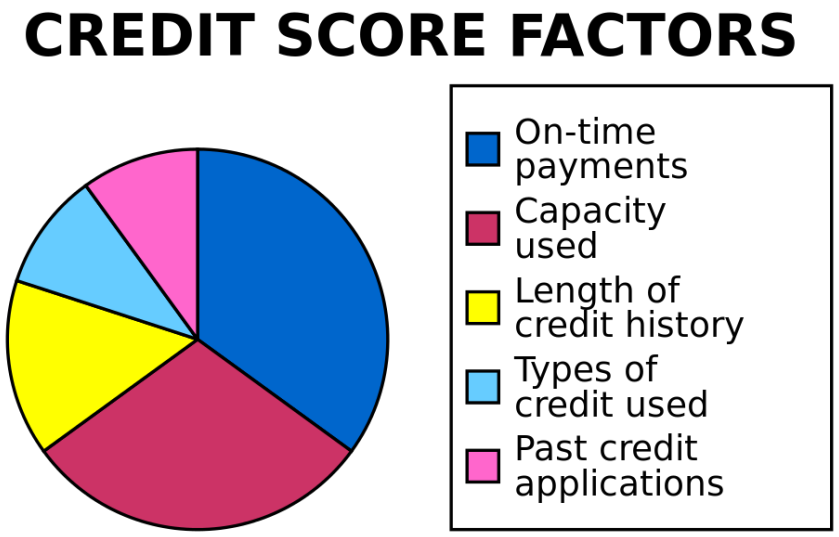
- Weather forecasters

DAY	DESCRIPTION	HIGH / LOW	PRECIP	WIND	HUMIDITY
TODAY JAN 23	Mostly Cloudy	44°/32°	0%	SE 3 mph	51%
FRI JAN 24	Cloudy	51°/46°	20%	ENE 8 mph	77%
SAT JAN 25	AM Rain	52°/30°	80%	WNW 10 mph	72%
SUN JAN 26	Partly Cloudy	47°/29°	10%	W 10 mph	60%
MON JAN 27	Partly Cloudy	46°/30°	10%	NW 9 mph	63%

Source: https://weather.com/weather/5day/l/e8321c2fb1f8234f40bf92ce494921d94e657d54cc2c01f1882755e04b761dee

Models in daily life

- Credit score calculators



Source: https://commons.wikimedia.org/wiki/File:Credit-score-chart.svg

- GPS navigation devices

Source: https://pixabay.com/vectors/gps-navigation-garmin-device-304842/

Models in daily life

- Model cars

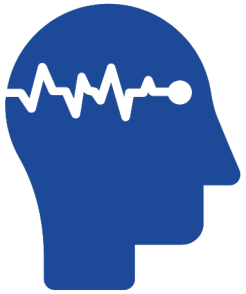
Source: https://pixabay.com/photos/model-car-ford-ford-capri-model-2093815/

- Scarecrows

Source: https://commons.wikimedia.org/wiki/File:Little_Bo_Peep_Scarecrow.JPG

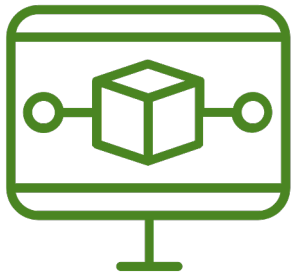
Models can be of different types

Mental model:
lives in one's brain, implicit, can't effectively run



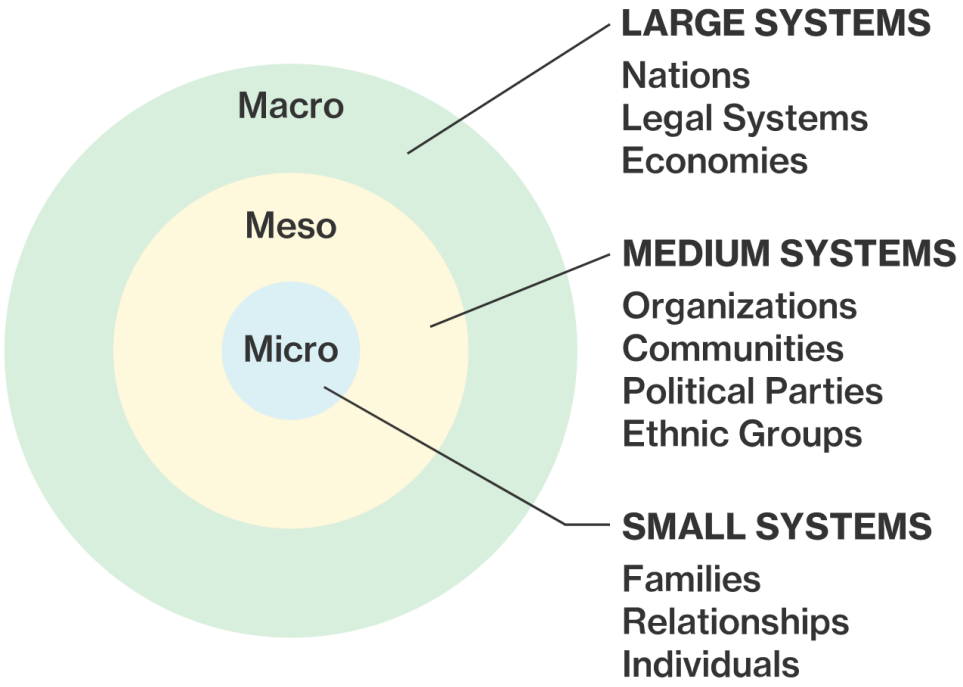
Conceptual model:
a sketch or written description of the mental model

Computational model:
codified version of a conceptual model that can run on a computer

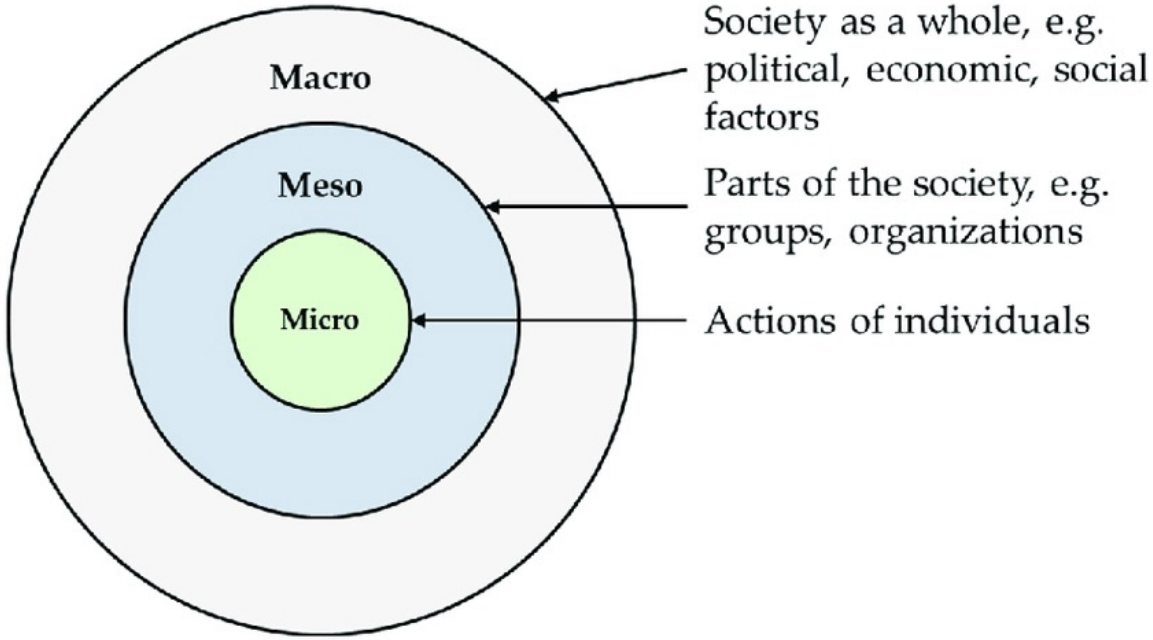


and many more

Models can represent different scales



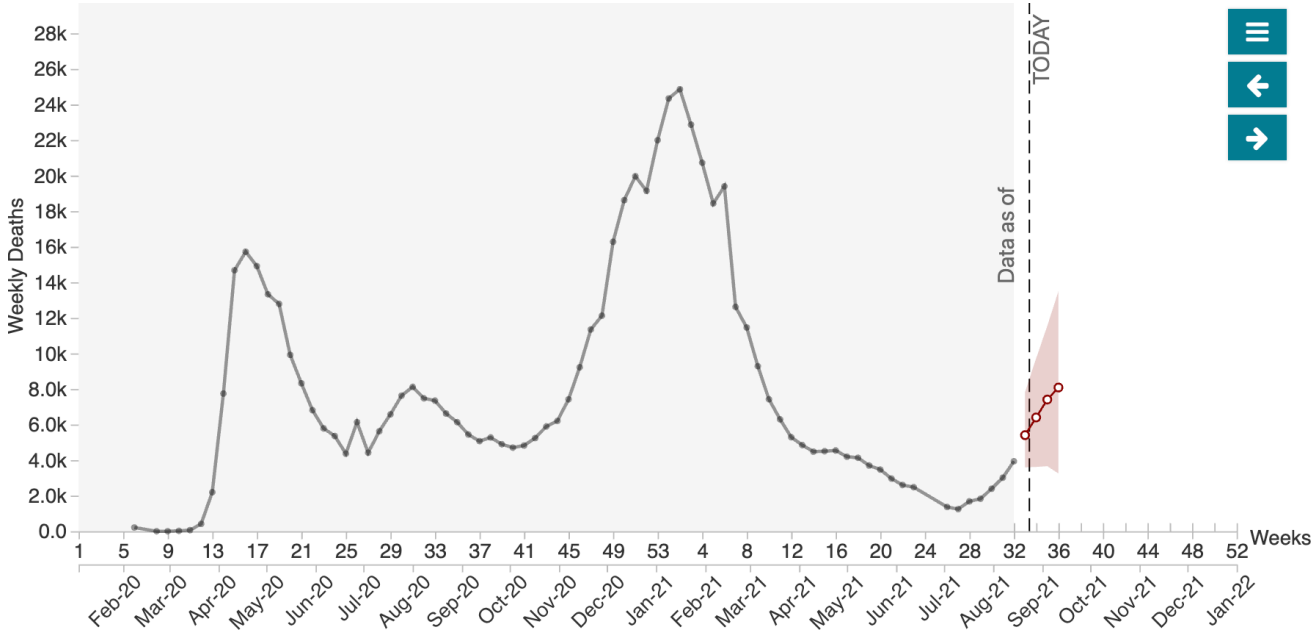
Source: https://www.coursehero.com/sg/introduction-to-sociology/macro-level-meso-level-and-micro-level-analysis/



Source: Javaid, A., Javed, A., & Kohda, Y. (2019). Exploring the role of boundary spanning towards service ecosystem expansion: A case of careem in pakistan. *Sustainability*, *11*(15), 3996.

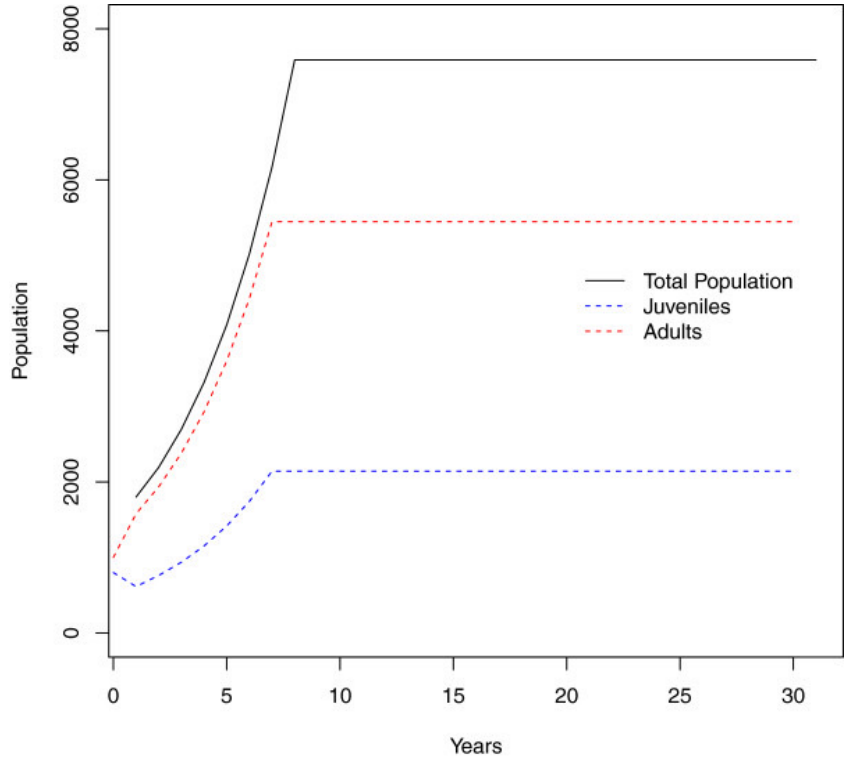
Models can vary in their certainty

Stochastic



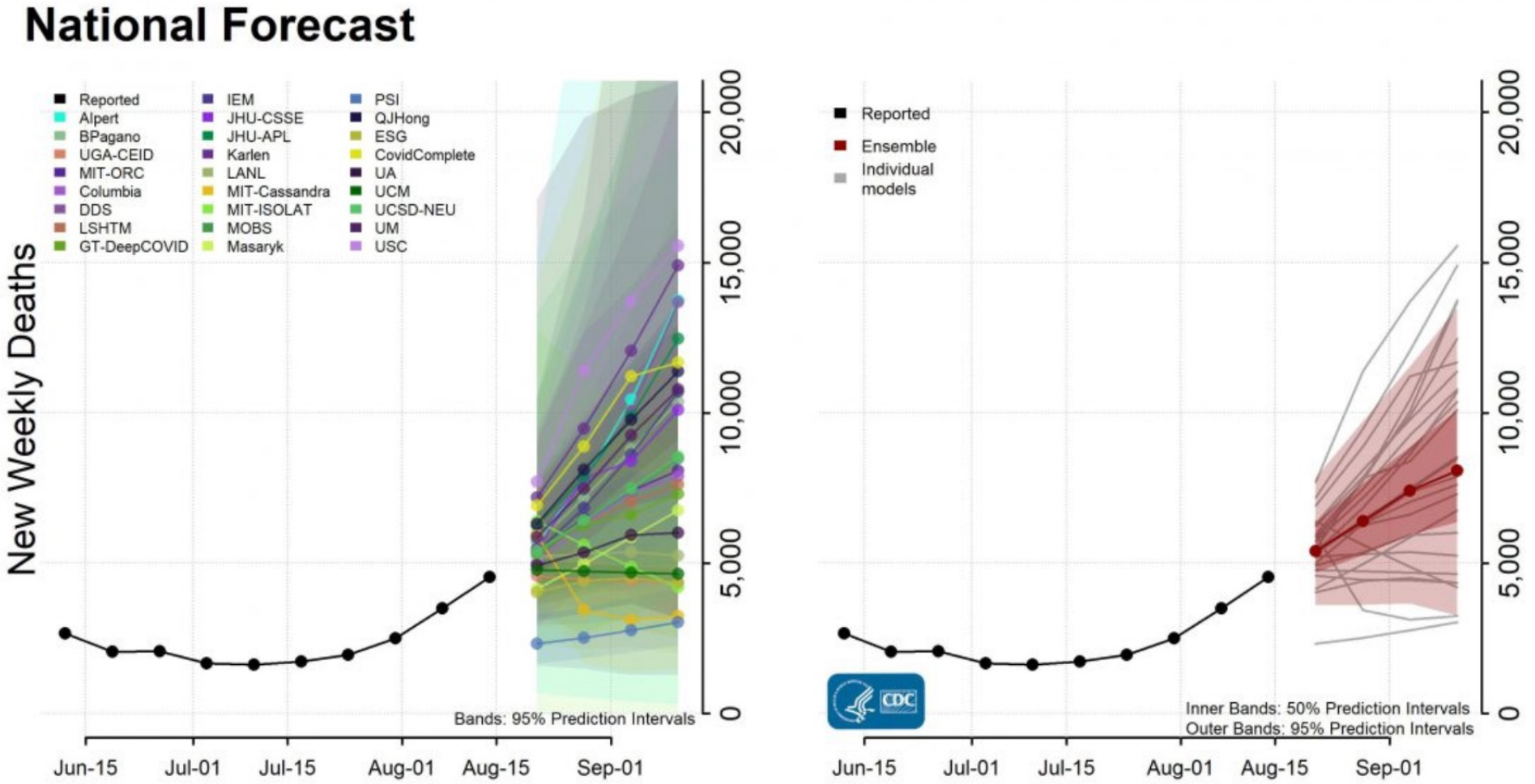
Source: https://covid.cdc.gov/covid-data-tracker/#forecasting_weeklydeaths

Deterministic



Source: Erickson, R. A., Thogmartin, W. E., & Szymanski, J. A. (2014). BatTool: an R package with GUI for assessing the effect of White-nose syndrome and other take events on Myotis spp. of bats. *Source code for biology and medicine*, *9*(1), 1-10.

We can combine models



Source: https://www.cdc.gov/coronavirus/2019-ncov/science/forecasting/forecasting-us.html

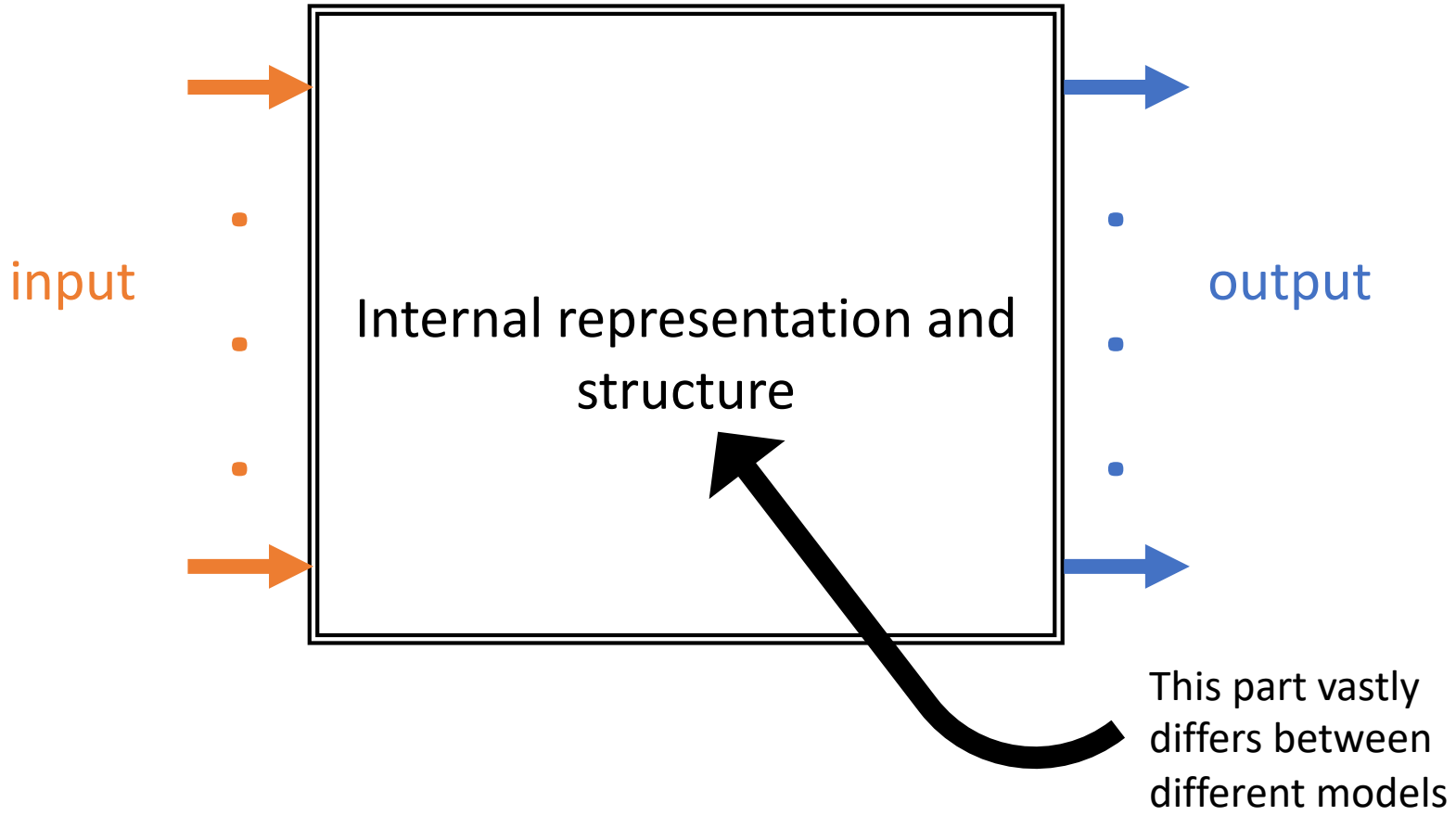
Some computational model types

- Agent-based models
- Machine learning models
- Discrete event models
- Mathematical models
  - Statistical models
  - Dynamical systems
- Network models
- Cognitive models
- …

Computational modeling involves

- A system to be modeled
  - real or imaginary
- A purpose
- Abstraction
  - we cannot include everything
- Assumptions
  - certain things are assumed true or false
- Formalism
  - Mathematical, algorithmic

A diagrammatic representation of models



Question to you: Is it a model or an algorithm?

- A software program that differentiates circular shapes from rectangular shapes.
- A software program that mimics movement of people (mobility).
- A software program that writes made-up stories.
- A software program that enables self-driving.

Why do we need models?

6 reasons among many

Why do we need models?

1. To test theories.

What is this? Any guess?

A donut?

A partly eaten donut?

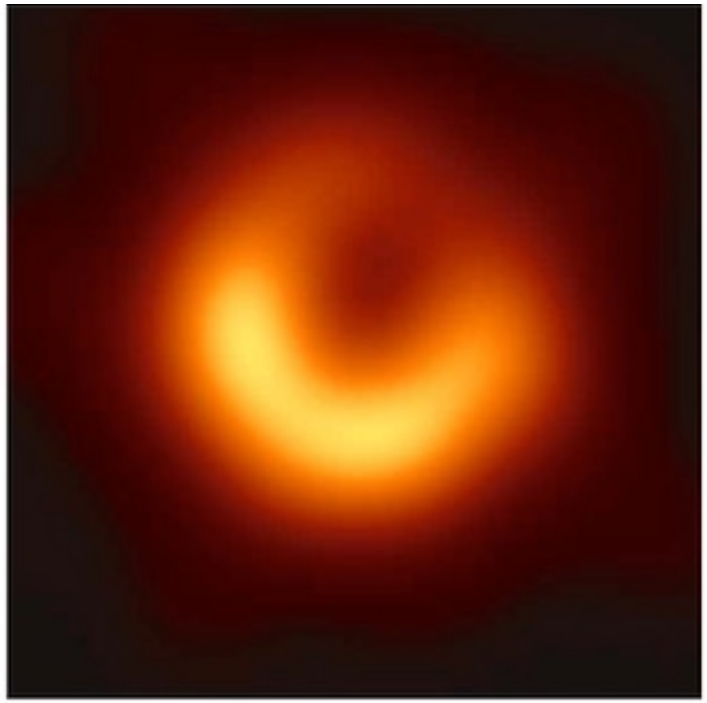
Looking at the sun during a solar eclipse?

First ever picture of a black hole from the galaxy M87 obtained using The Event Horizon Telescope.

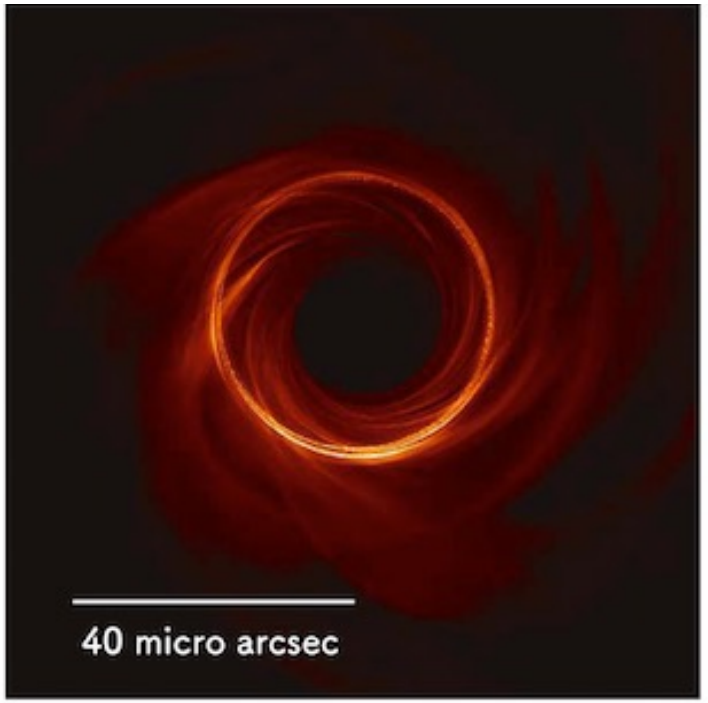
Source: https://www.jpl.nasa.gov/edu/news/2019/4/19/how-scientists-captured-the-first-image-of-a-black-hole/

How about this?

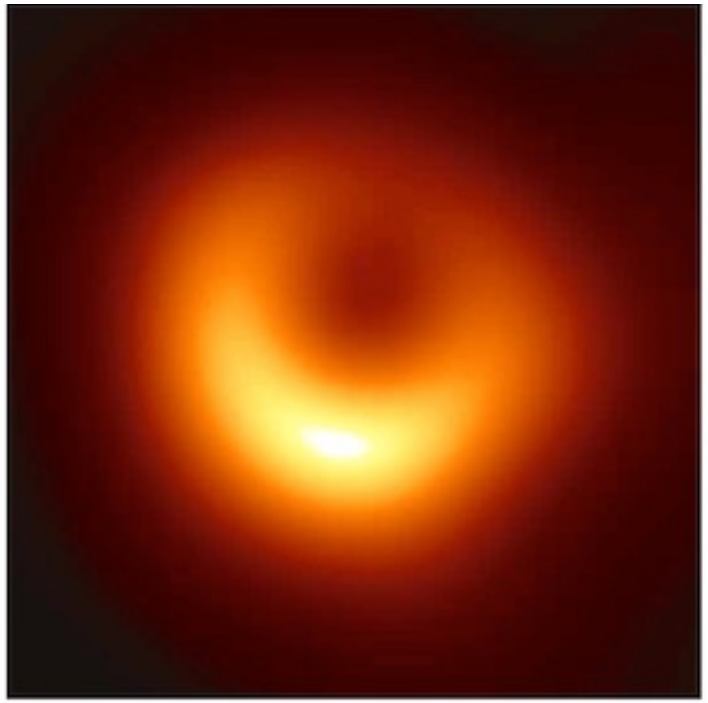
Real Black Hole (M87)



Simulation of Black Hole (M87)



Blurred Simulation

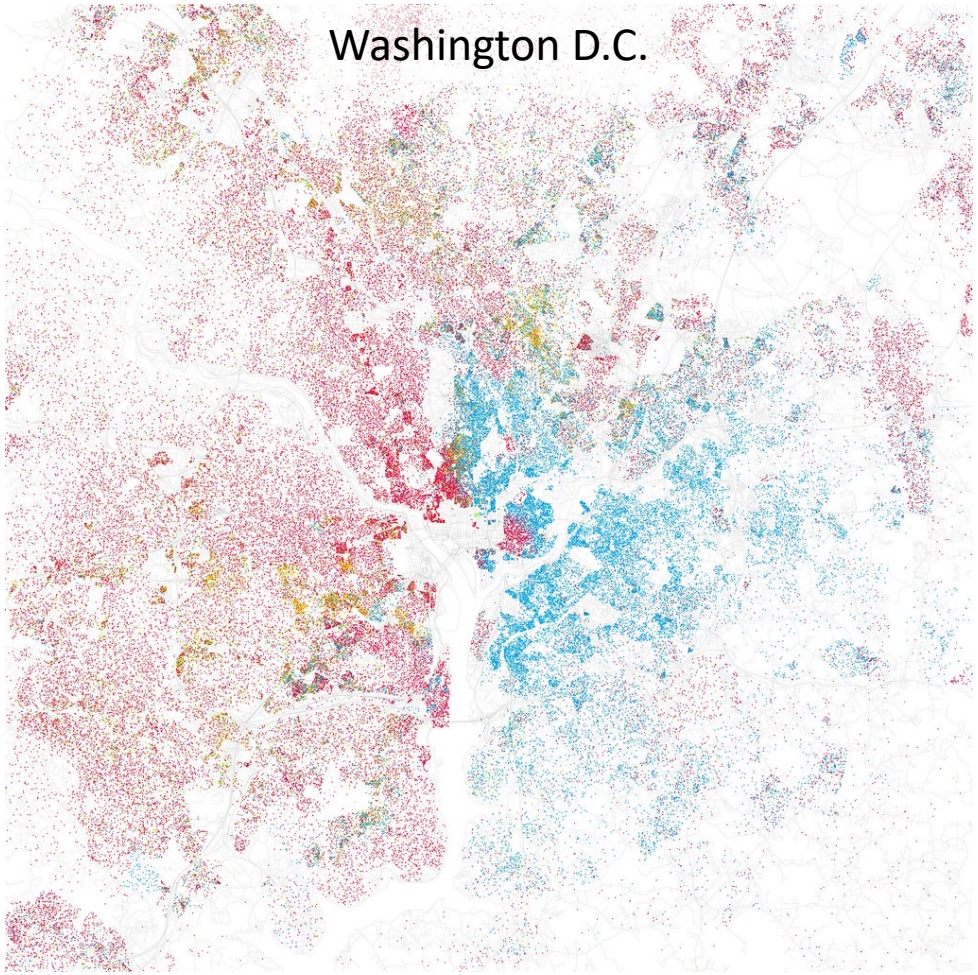


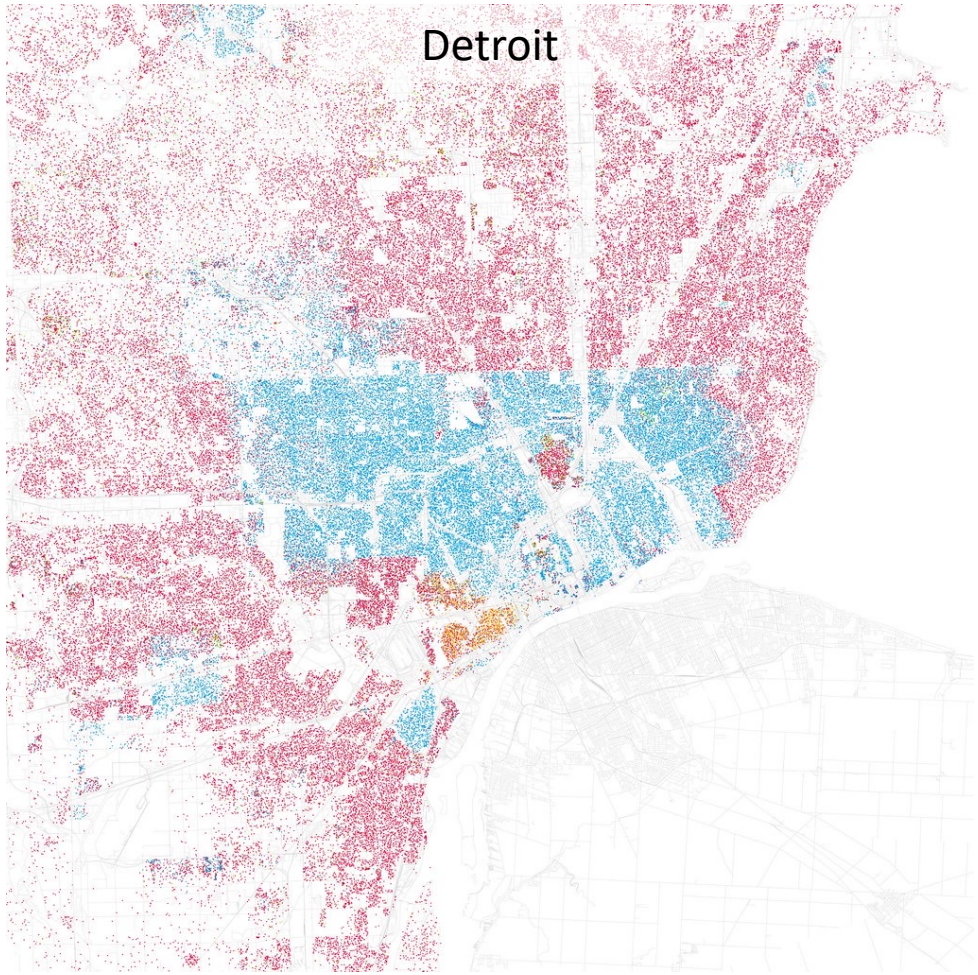
Source: https://www.cfca.nao.ac.jp/en/pr/20190410

Why do we need models?

1. To test theories.
2. To explain a phenomena.

Anyone likes maps?

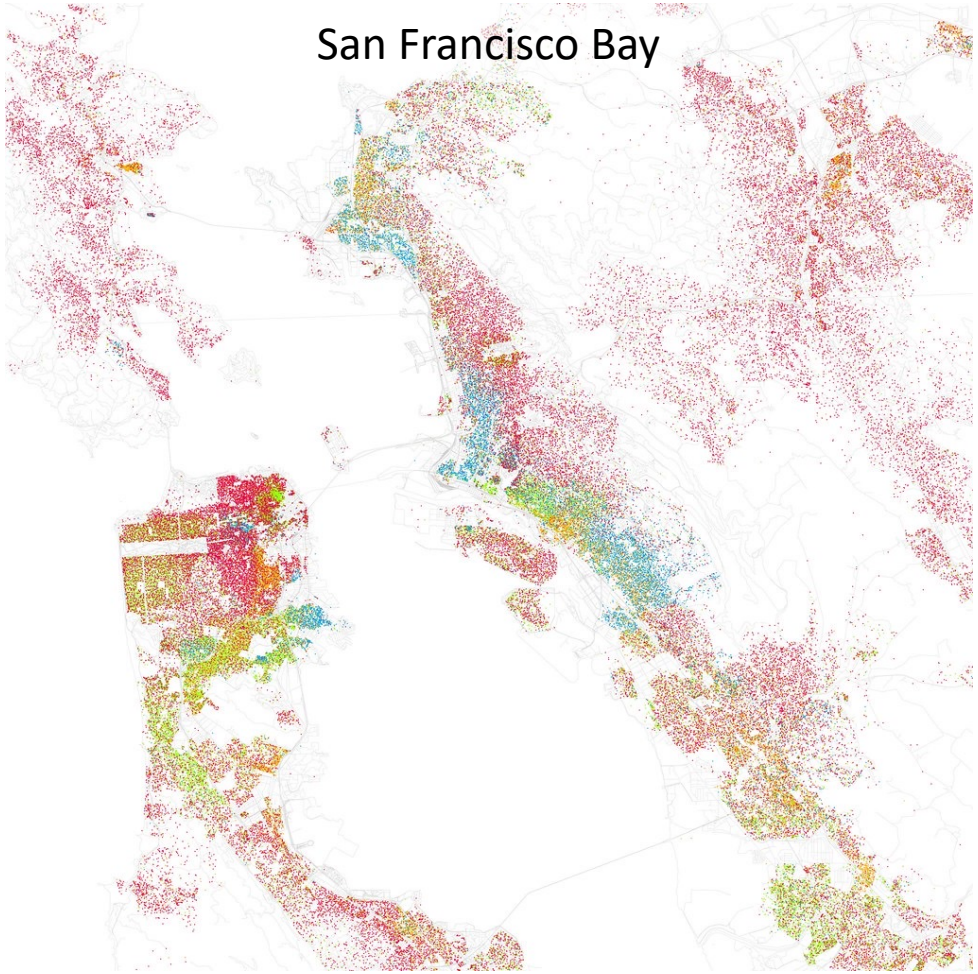




Source: https://www.fastcompany.com/1690097/infographic-day-how-segregated-your-city

More maps?

San Francisco Bay

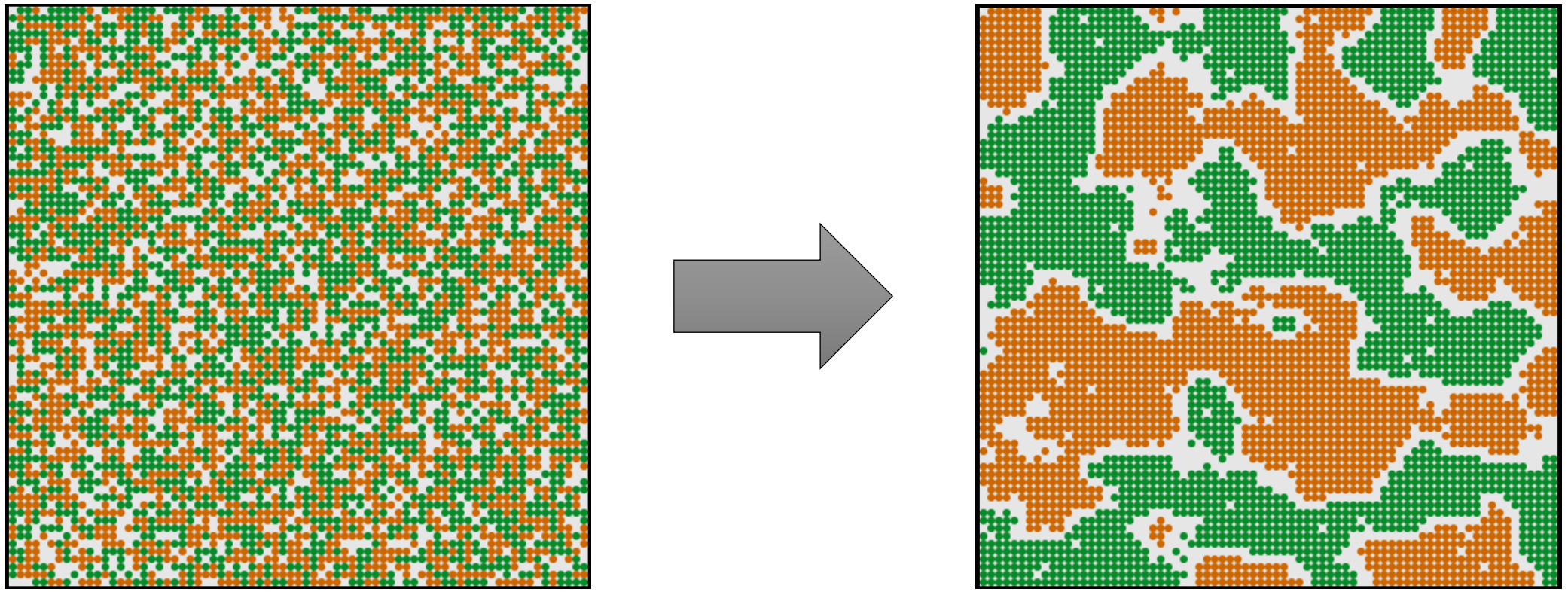


San Antonio

Source: https://www.fastcompany.com/1690097/infographic-day-how-segregated-your-city

An explanation to why cities are segregated

- http://hamdikavak.com/sims/segregation/

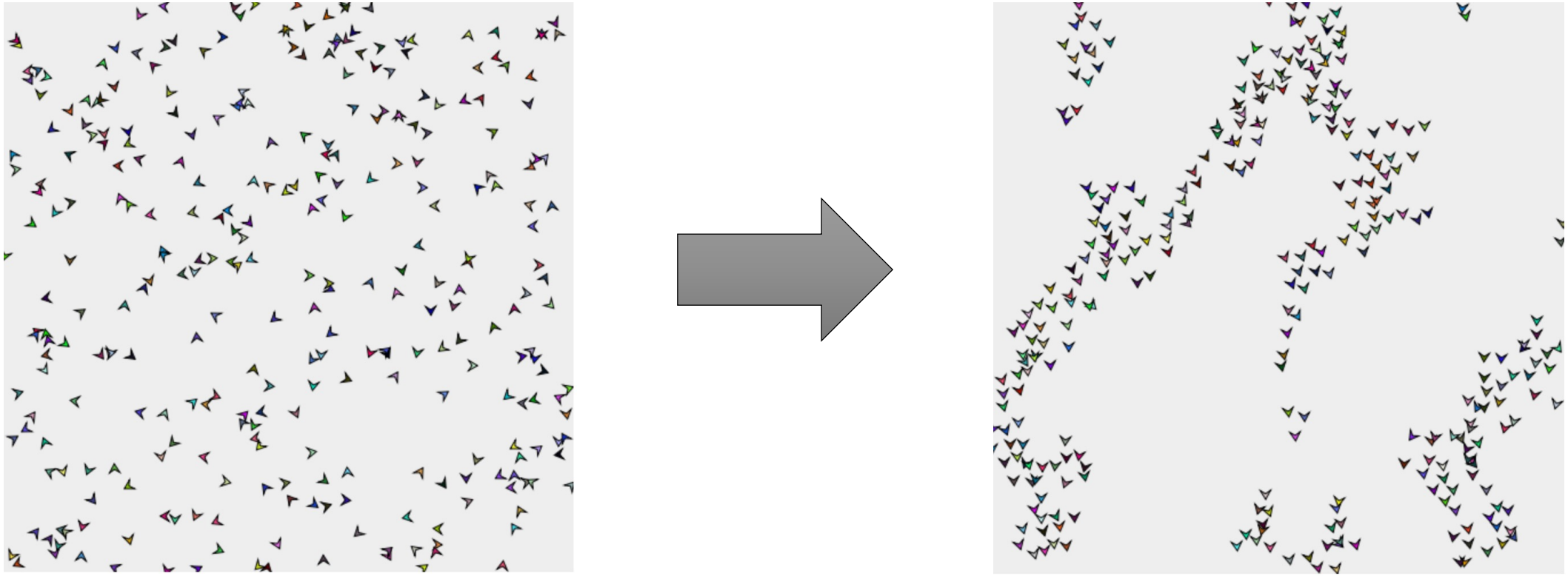


Birds

Source: https://www.youtube.com/watch?v=bb9ZTbYGRdc

An explanation to how birds flock

- http://hamdikavak.com/sims/flocking/



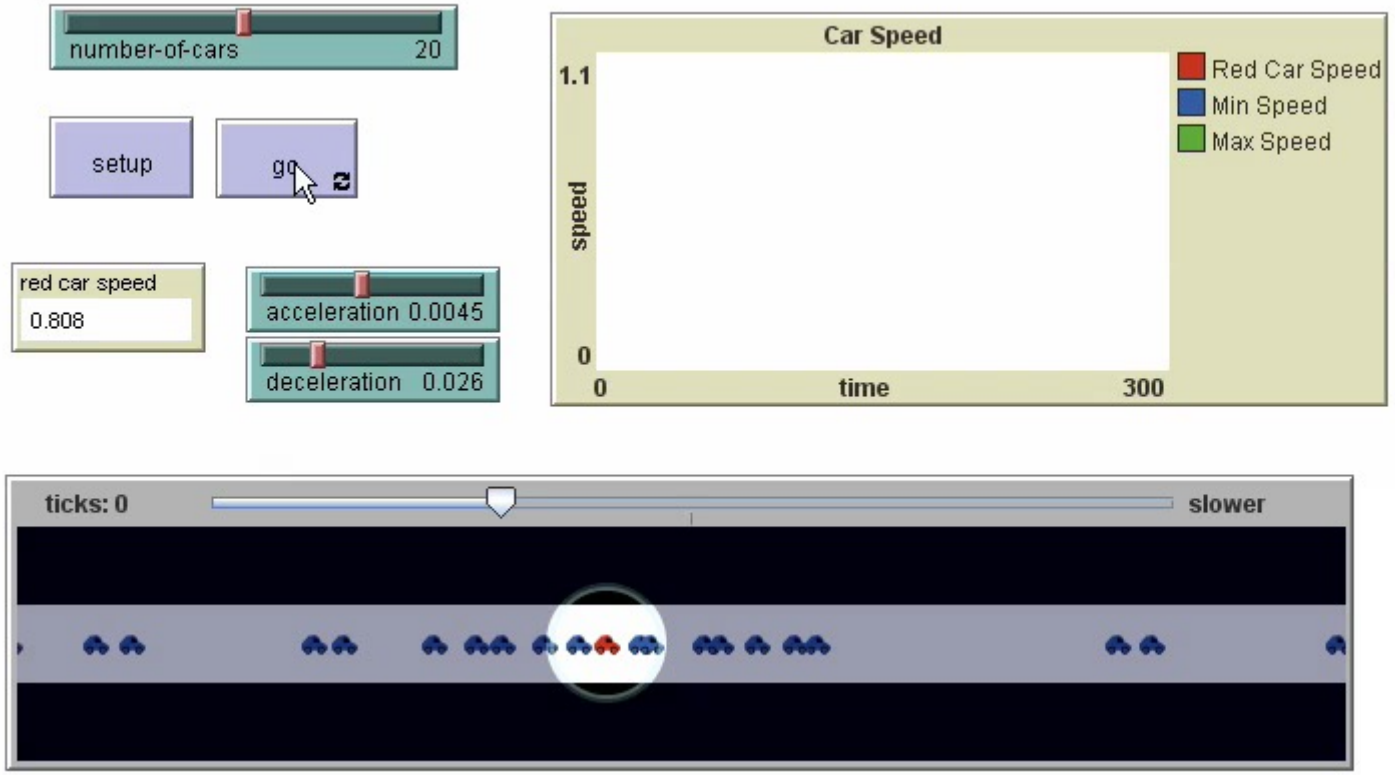
Traffic shockwave jam in real-world

- 22 cars equally spaced on a 230m single lane circle.
- Drivers asked to cruise steadily at 30km/h.
- 1st traffic moved freely.
- Disturbances/clusters soon appear.
- Causing cars to slow/stop.
- Cars at front of cluster can accelerate at 40km/h.
- But these join another jam.

Source: http://www.youtube.com/watch?v=Suugn-p5C1M
New Scientist Article: http://technology.newscientist.com/article/dn13402

Traffic shockwave jam simulated

- Example:
- Models the movement of cars on a road.
- Each car follows a simple set of rules:
  - If there's a car close ahead, it slows down.
  - If there's no car ahead, it speeds up.
- Demonstrates how traffic jams can form without any obvious incident.
- Simple rules can explain phenomena.



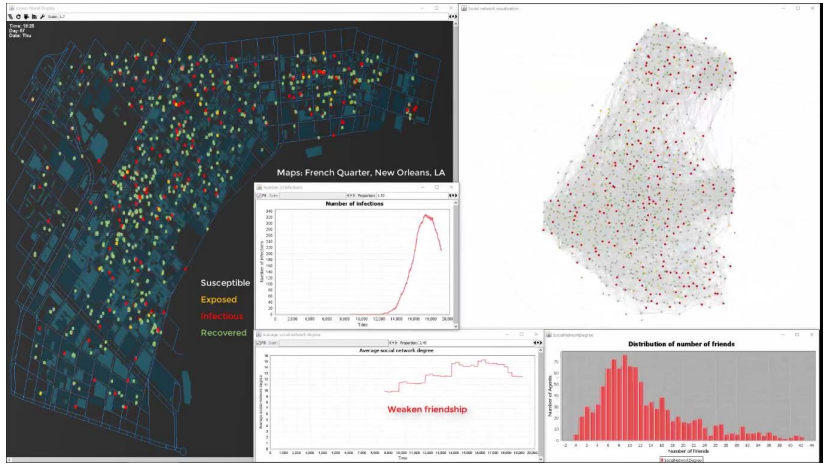
Source: NetLogo

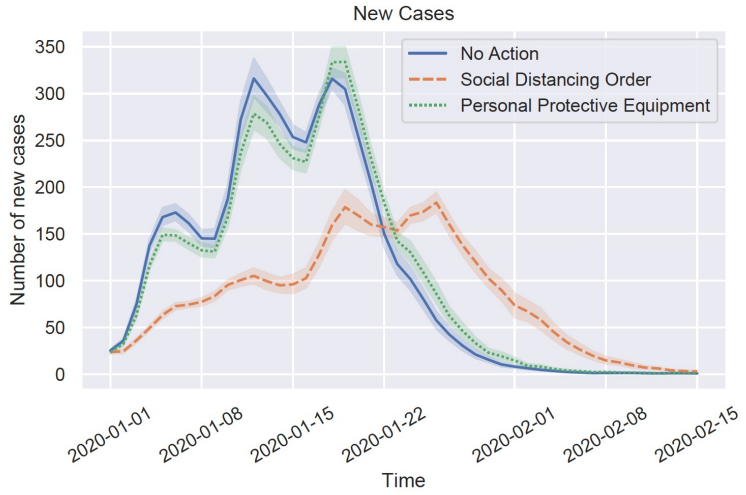
Why do we need models?

1. To test theories.
2. To explain a phenomena.
3. To predict a phenomena.

Disease spread prediction

- A simple disease model
  - http://www.shodor.org/featured/DiseaseModel
- Ebola model
  - https://www.khanacademy.org/science/health-and-medicine/current-issues-in-health-and-medicine/ebola-outbreak/pi/modelling-an-epidemic
- COVID-19 models





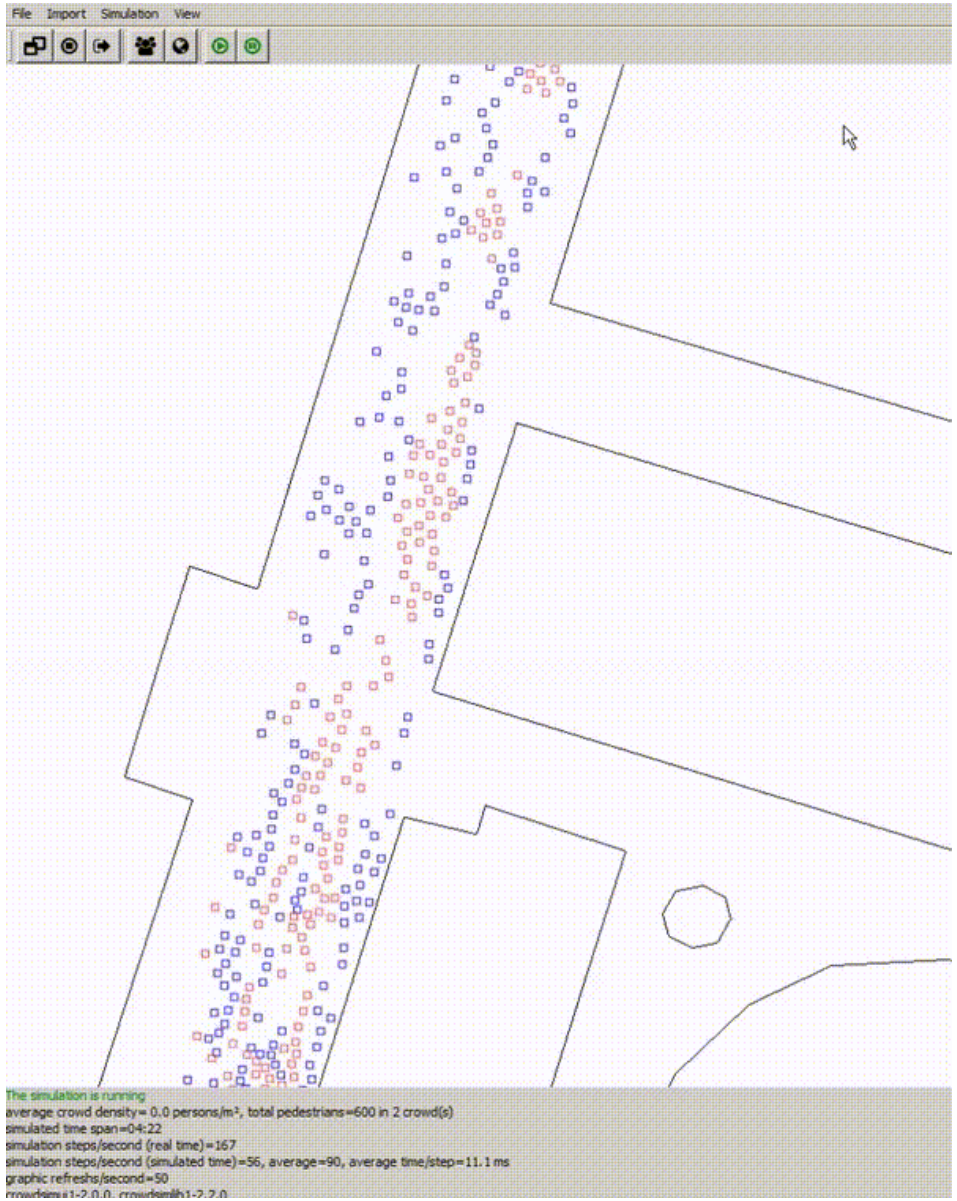
Why do we need models?

1. To test theories.
2. To explain a phenomena.
3. To predict a phenomena.
4. To test dangerous scenarios.

Evacuation Scenarios

- Office fire
  - https://youtu.be/st8HRgHOErw
- Panic evacuation
  - https://youtu.be/gDtfTV_c7Es

Pedestrian movement



Source: https://twitter.com/stefanhahmann/status/1082213811497635846

Why do we need models?

1. To test theories.
2. To explain a phenomena.
3. To predict a phenomena.
4. To test dangerous scenarios.
5. Entertainment

Entertainment

- Games and movies

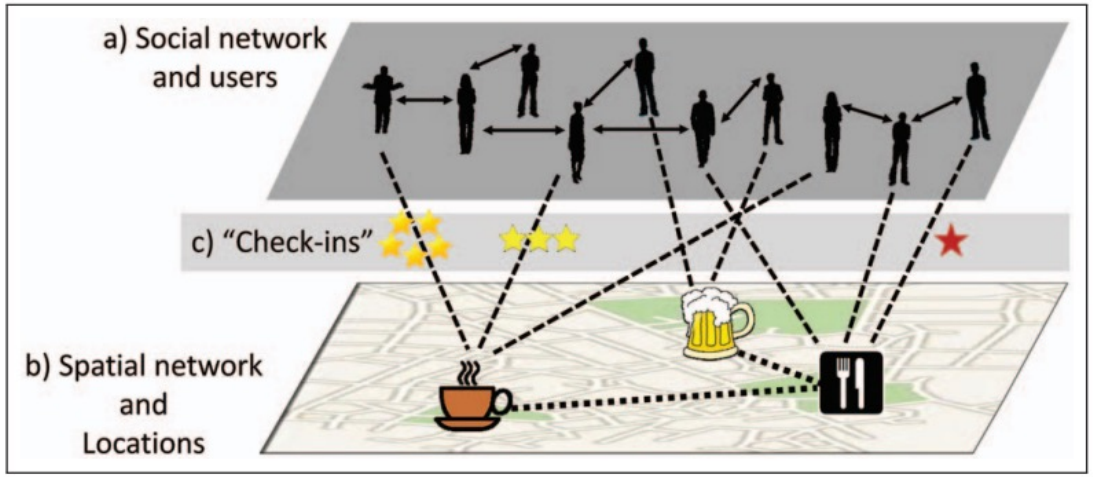
Sources: https://www.amazon.co.uk/The-Sims-4-Standard-Edition/dp/B00KHJLXN2, https://www.imdb.com/title/tt0499549/, and https://fmmvibe.com/forums/topic/42906-fantastic-4-2-3-1-perfect-tactic-for-real-madrid/

Why do we need models?

1. To test theories.
2. To explain a phenomena.
3. To predict a phenomena.
4. To test dangerous scenarios.
5. Entertainment
6. Data generation

Data generation

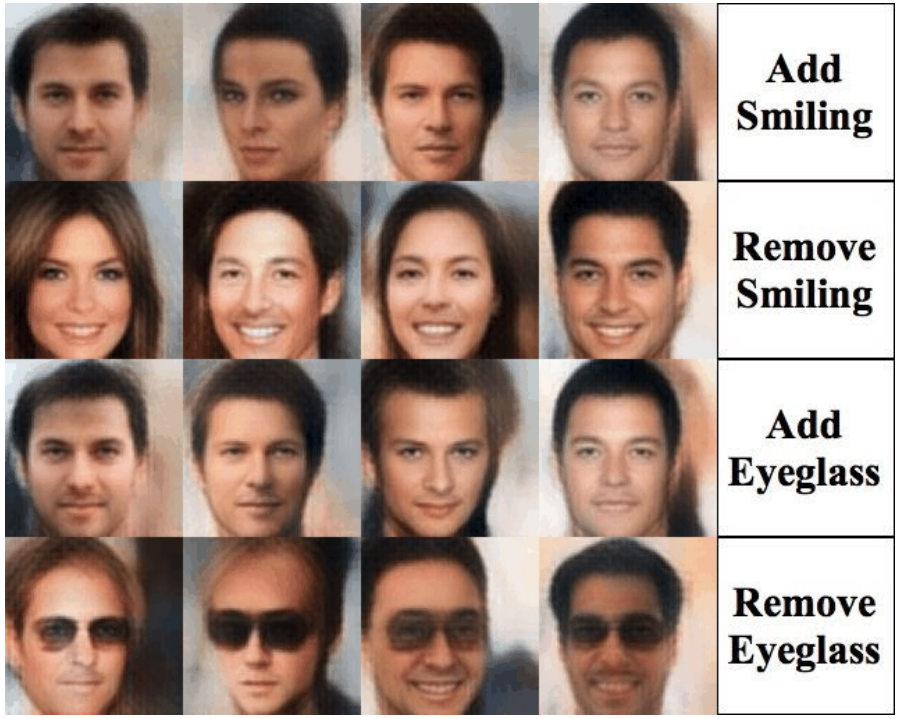
- Sometimes it's not feasible, ethical, or cheap to collect data
- We can generate data using models
- For instance, we can simulate check-in at public venues



Kim, J. S., Jin, H., Kavak, H., Rouly, O. C., Crooks, A., Pfoser, D., ... & Züfle, A. (2020, June). Location-based social network data generation based on patterns of life. In *2020 21st IEEE International Conference on Mobile Data Management (MDM)* (pp. 158-167). IEEE.

Data generation

- Generative adversarial networks are very popular to generate data



Source:https://houxianxu.github.io/assets/project/dfcvae

Why all models are wrong?

"All models are wrong, but some are useful" George Box (1978)

- Abstraction
  - E.g., Using masses vs. atoms vs. quantum particles vs. …
  - E.g., Parameters chosen
- Scope
  - e.g., Newton's Laws work at a particular scope but Einstein's general relativity covers larger spectrum.
- Uncertainty
  - Measurements from real system (e.g., weather model resolution)
  - Model outputs (e.g., stochastic models)
- Bias
  - Measurements from real system (e.g., undercounted populations)
  - Algorithmic bias (e.g., ordering of things)
- Assumptions
  - From modeling (e.g., assume people won't react to pandemic)
  - From the formalism used in the model (e.g., mathematical or algorithmic)

The bottom line

- All models are wrong in many ways.
- Modeling process involves many arbitrary decisions.
- Models are used in mission-critical applications…

- Thus, our duty is to make our models **explicit**, **transparent**, and **interrogate** them with verification and validation techniques so that they can be useful for the purpose they were built for.

- Next week we will talk about the V&V terminology and the modeling process in more detail.