BIOS27816: Population biology of infectious diseases of global health concern

Introduction to the Population Biology of Infectious Disease

UChicago Center in Paris

Paris, France

January 2024

Goals for this lecture

- To explain what we're doing here
- To define "science"
- To define "data"
- To define "models"
- To introduce many different types of models
  - Statistical
  - Mathematical
- To explain the application of these models to the fields of:
  - Classical epidemiology
  - Population biology of infectious diseases

What is global health?

- the health of the populations in the worldwide context (WHO 2006)
- an area of study, research, and practice that places a priority on improving health and achieving equity in health for all people worldwide
- focus on problems that transcend national borders or have a global political and economic impact are often emphasized (reduction of disparities, protection against global threats)
  - e.g. COVID (vaccination), polio eradication
- Different from international health, the branch of public health focused on foreign aid efforts led by high income countries to improve health improve health in low- and middle-income countries (LMICs)

What are the major threats to global health?

(e.g. globally, the leading causes of death)

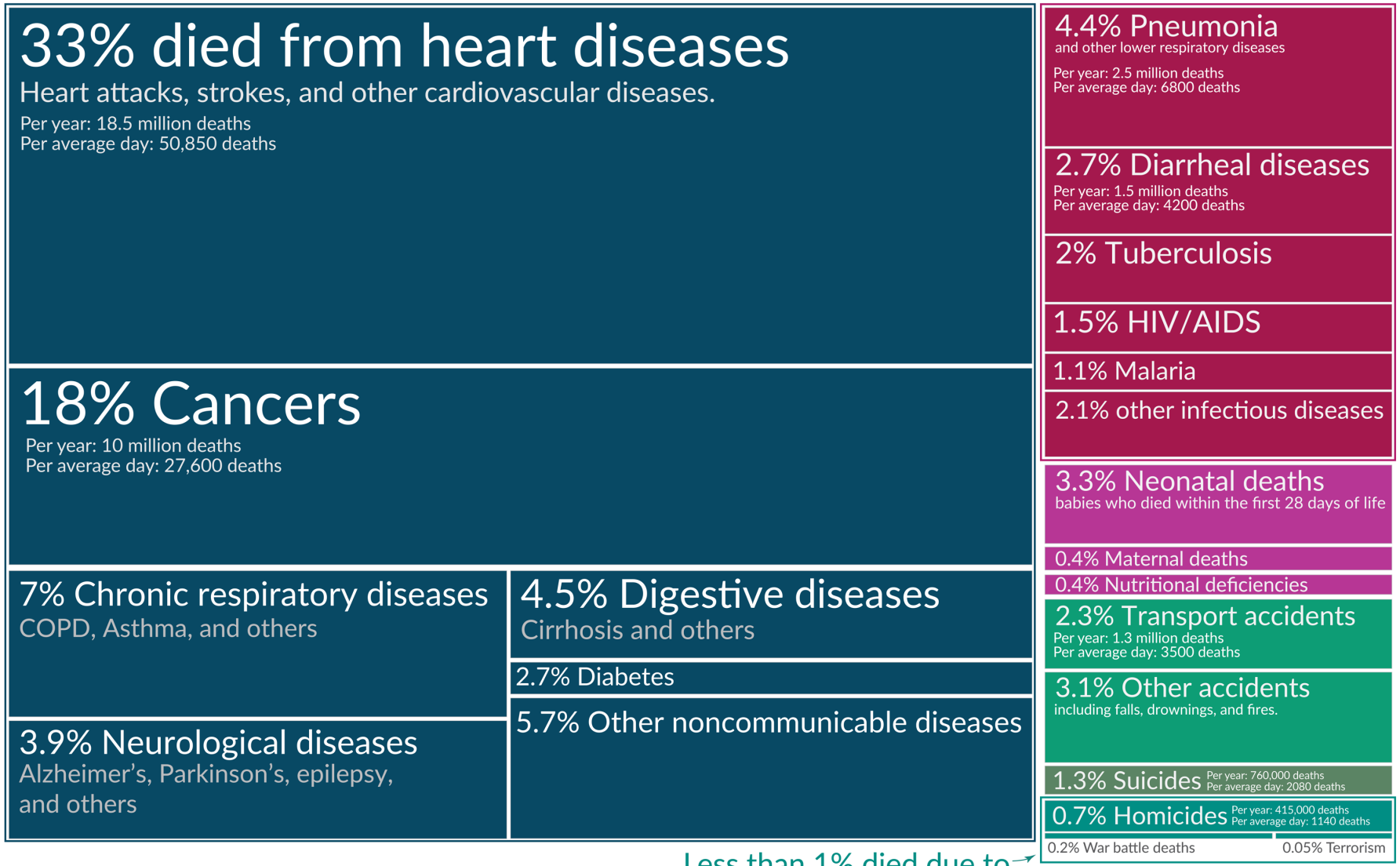
What do people die from? Causes of death globally in 2019

The size of the entire visualization represents the total number of deaths in 2019: 55 million.
Each rectangle within it is proportional to the share of deaths due to a particular cause.

Our World in Data

74% died from noncommunicable diseases

14% died from infectious diseases



Less than 1% died due to interpersonal violence

Data source: IHME Global Burden of Disease and Global Terrorism Database

OurWorldinData.org – Research and data to make progress against the world's largest problems.

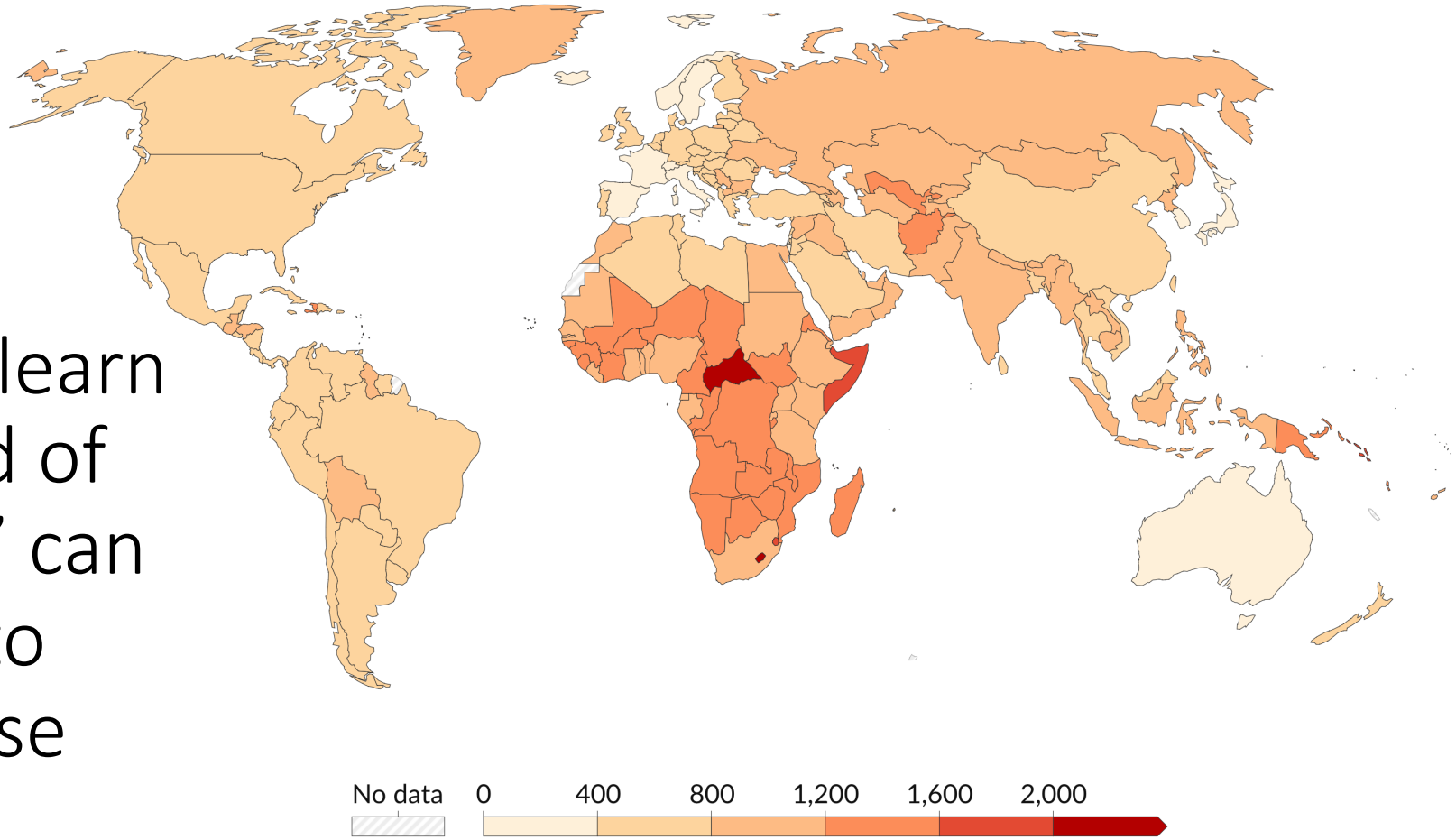
Global death rates are far from equitable.

We're here to learn how the field of 'global health' can intervene to reverse these trends.

Annual death rate from all causes, 2019

The estimated annual death rate from all causes per 100,000 people.

Our World in Data



Data source: IHME, Global Burden of Disease (2019) OurWorldInData.org/causes-of-death | CC BY

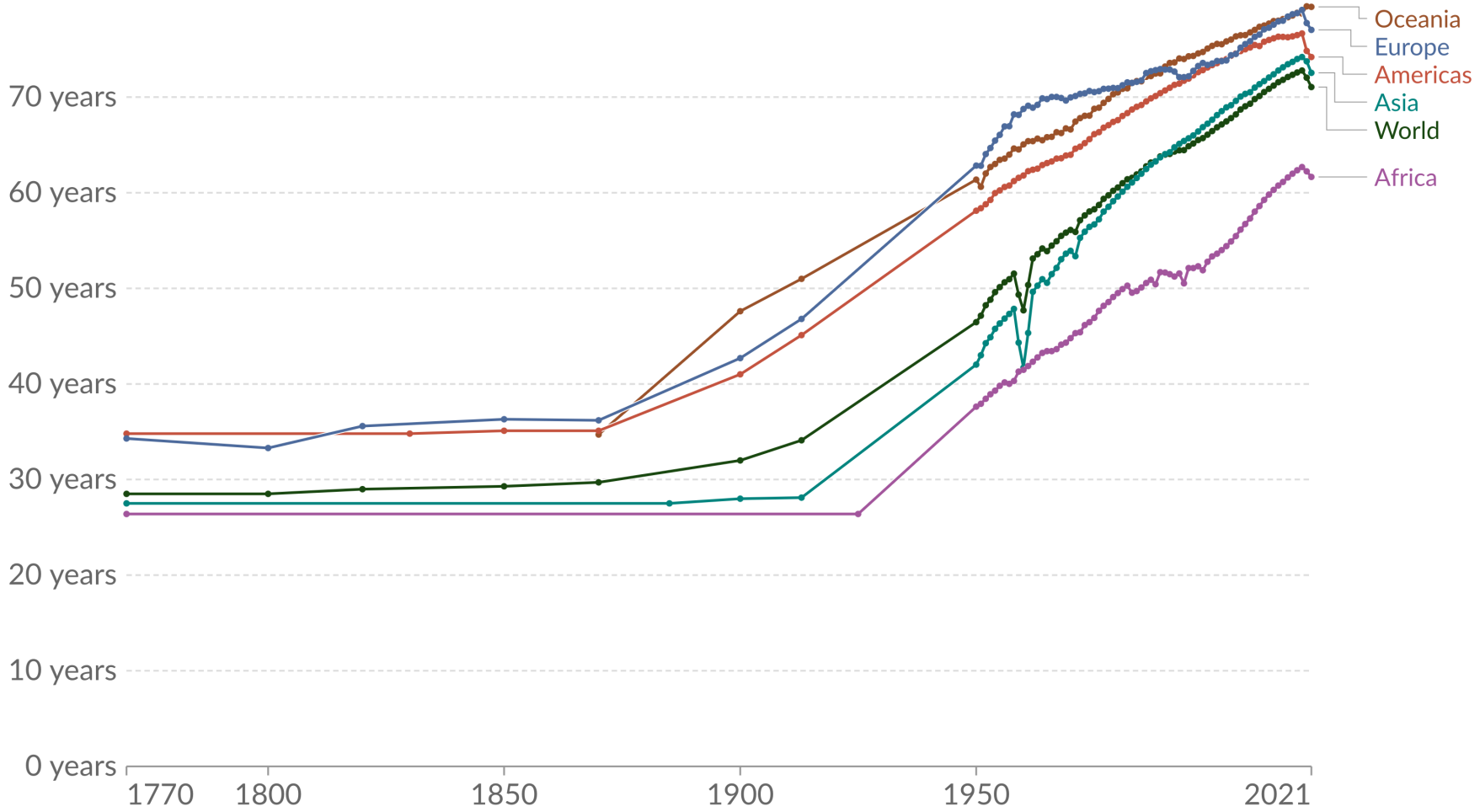
Note: To allow for comparisons between countries and over time, this metric is age-standardized[1].

1. Age standardization: Age standardization is an adjustment that makes it possible to compare populations with different age structures by standardizing them to a common reference population. Read more: How does age standardization make health metrics comparable?

Life expectancy

The period life expectancy[1] at birth, in a given year.

Our World in Data

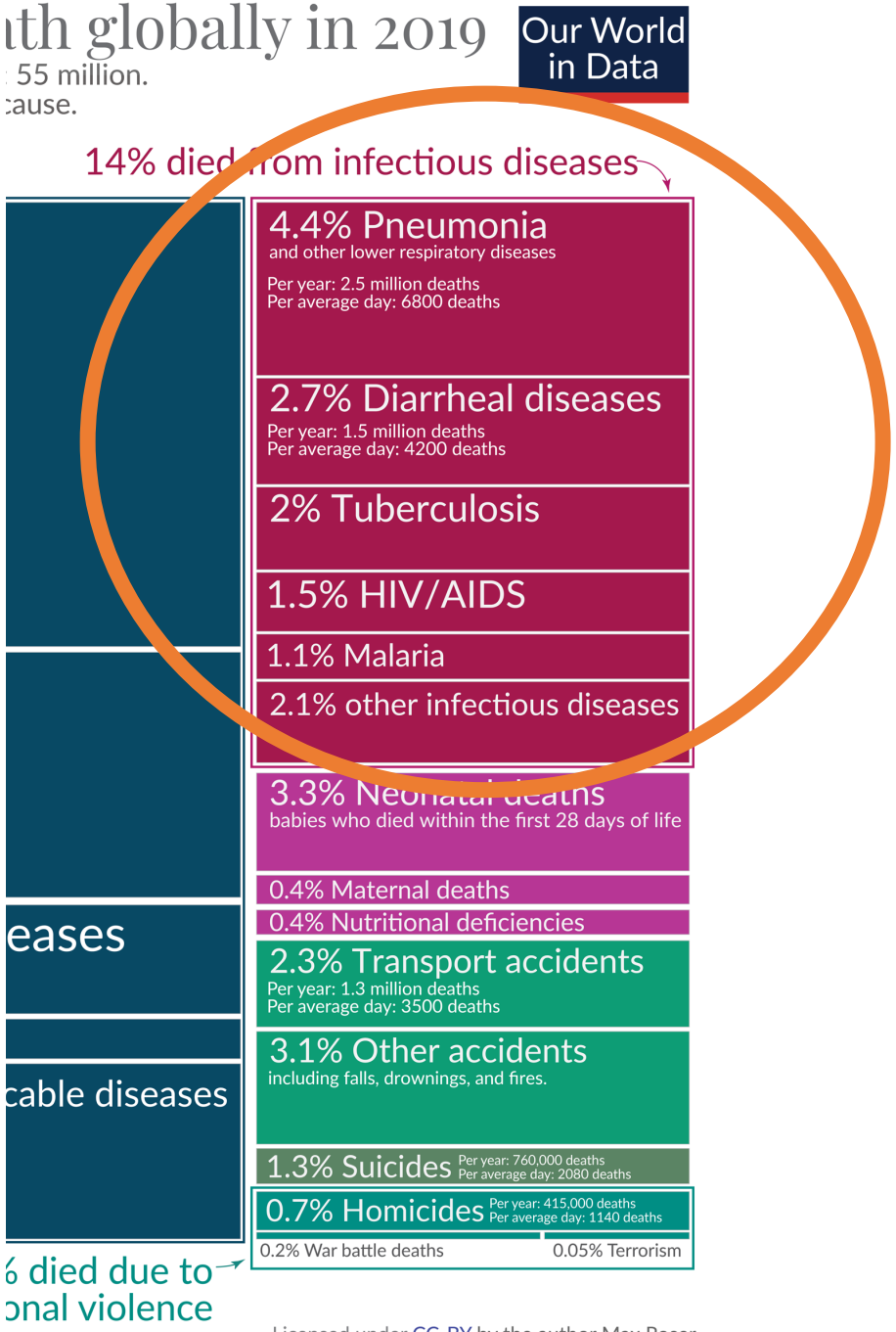


Data source: UN WPP (2022); HMD (2023); Zijdeman et al. (2015); Riley (2005)
OurWorldInData.org/life-expectancy | CC BY

1. Period life expectancy: Period life expectancy is a metric that summarizes death rates across all age groups in one particular year. For a given year, it represents the average lifespan for a hypothetical group of people, if they experienced the same age-specific death rates throughout their whole lives as the age-specific death rates seen in that particular year. Learn more in our article: "Life expectancy" – What does this actually mean?

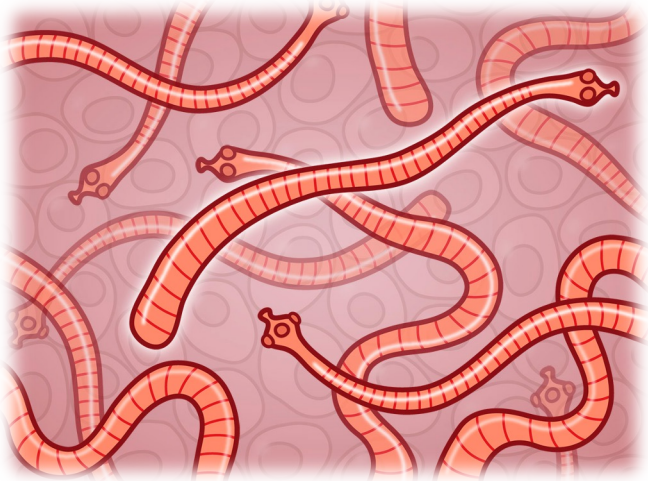
BIOS27816:

Population biology of infectious diseases of global health concern



Parasites and Pathogens

- Parasite: an organism that lives in or on another organism and benefits at the expense of others.
  - Ex: helminths (parasitic worms: tapeworms, roundworms, hookworms), ectoparasites (ticks, fleas)



Parasites and Pathogens

- Parasite: an organism that lives in or on another organism and benefits at the expense of others.
  - Ex: helminths (parasitic worms: tapeworms, roundworms, hookworms), ectoparasites (ticks, fleas)
- Pathogen: a microorganism that can cause disease
  - Ex: bacterium, virus, protozoan

Parasites and Pathogens

- Parasite: an organism that lives in or on another organism and benefits at the expense of others.
  - Ex: helminths (parasitic worms: tapeworms, roundworms, hookworms), ectoparasites (ticks, fleas)
- Pathogen: a microorganism that can cause disease
  - Ex: bacterium, virus, protozoan

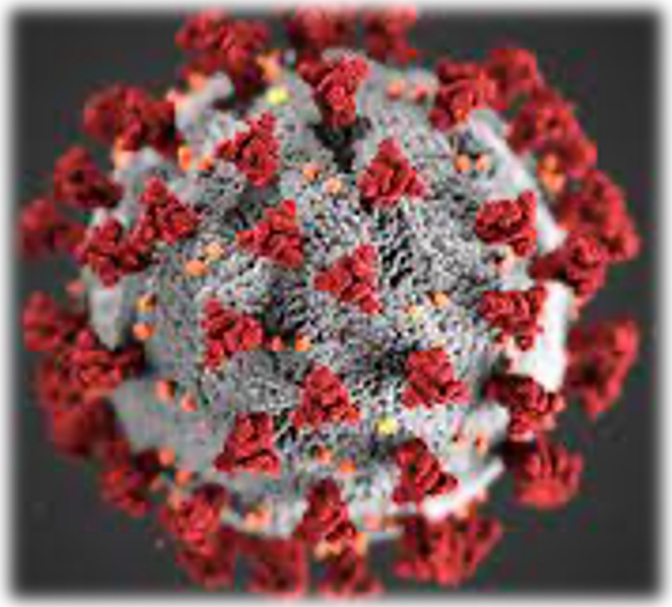
Pathogen: *Yersinia pestis*
Disease: Plague

Parasites and Pathogens

- Parasite: an organism that lives in or on another organism and benefits at the expense of others.
  - Ex: helminths (parasitic worms: tapeworms, roundworms, hookworms), ectoparasites (ticks, fleas)
- Pathogen: a microorganism that can cause disease
  - Ex: bacterium, virus, protozoan

Pathogen: *Yersinia pestis*
Disease: Plague

Pathogen: SARS-CoV-2
Disease: COVID-19



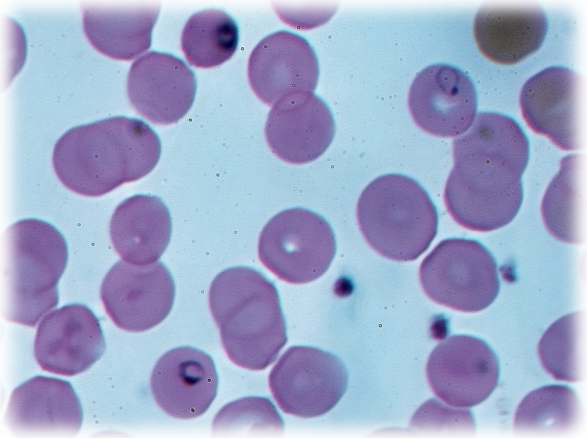
Parasites and Pathogens

- Parasite: an organism that lives in or on another organism and benefits at the expense of others.
  - Ex: helminths (parasitic worms: tapeworms, roundworms, hookworms), ectoparasites (ticks, fleas)
- Pathogen: a microorganism that can cause disease
  - Ex: bacterium, virus, protozoan

Pathogen: *Yersinia pestis*
Disease: Plague

Pathogen: SARS-CoV-2
Disease: COVID-19

Pathogen: *Plasmodium falciparum, P. vivax, P. malariae, P. ovale, P. knowlesi*
Disease: malaria

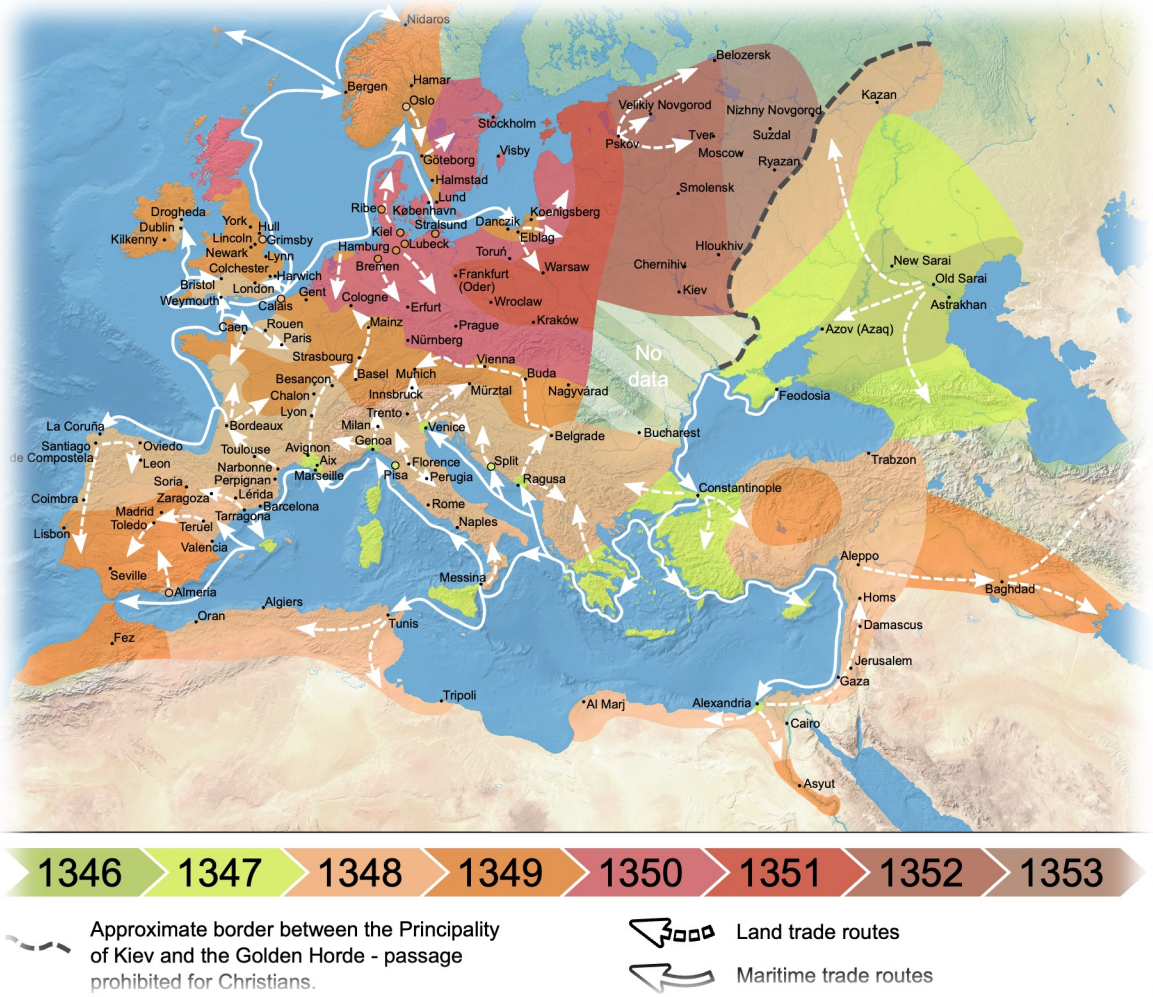


Parasites and pathogens have shaped human history.

- Plague of Justinian (541-549 AD)
  - First historically recorded pandemic of *Yersinia pestis*
  - Launched the ‘first plague pandemic’ resulting in the deaths of 15-100 million people, 25-60% of Europe’s population at the time

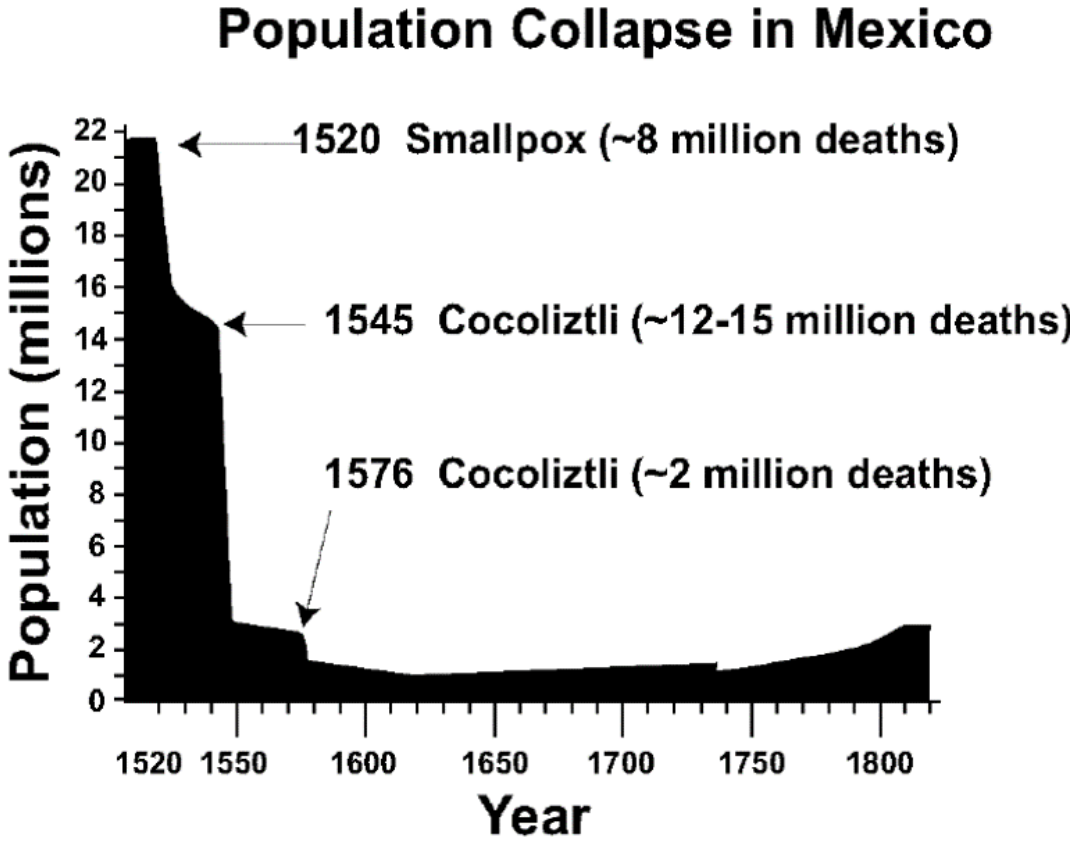
Parasites and pathogens have shaped human history.

- Plague of Justinian (541-549 AD)
- Black Death (1346-1353 AD)
  - Most fatal pandemic in human history, resulting in deaths of 75-200 million people
  - Killed 30-60% of Europe's population at the time; 17-54% of global population



Parasites and pathogens have shaped human history.

- Plague of Justinian (541-549 AD)
- Black Death (1346-1353 AD)
- Cocoliztli (1545-1548)
  - Pathogen still unknown! Maybe viral hemorrhagic fever, maybe bacterium
  - Killed 80% of the population of Mexico



Parasites and pathogens have shaped human history.

- Plague of Justinian (541-549 AD)
- Black Death (1346-1353 AD)
- Cocoliztli (1545-1548)
- Spanish Influenza (1918-1920)
  - 17-100 million deaths worldwide.
  - 1-5% of global population
  - 2nd-most devastating pandemic in history (after Black Death)

ILLUSTRATED CURRENT NEWS

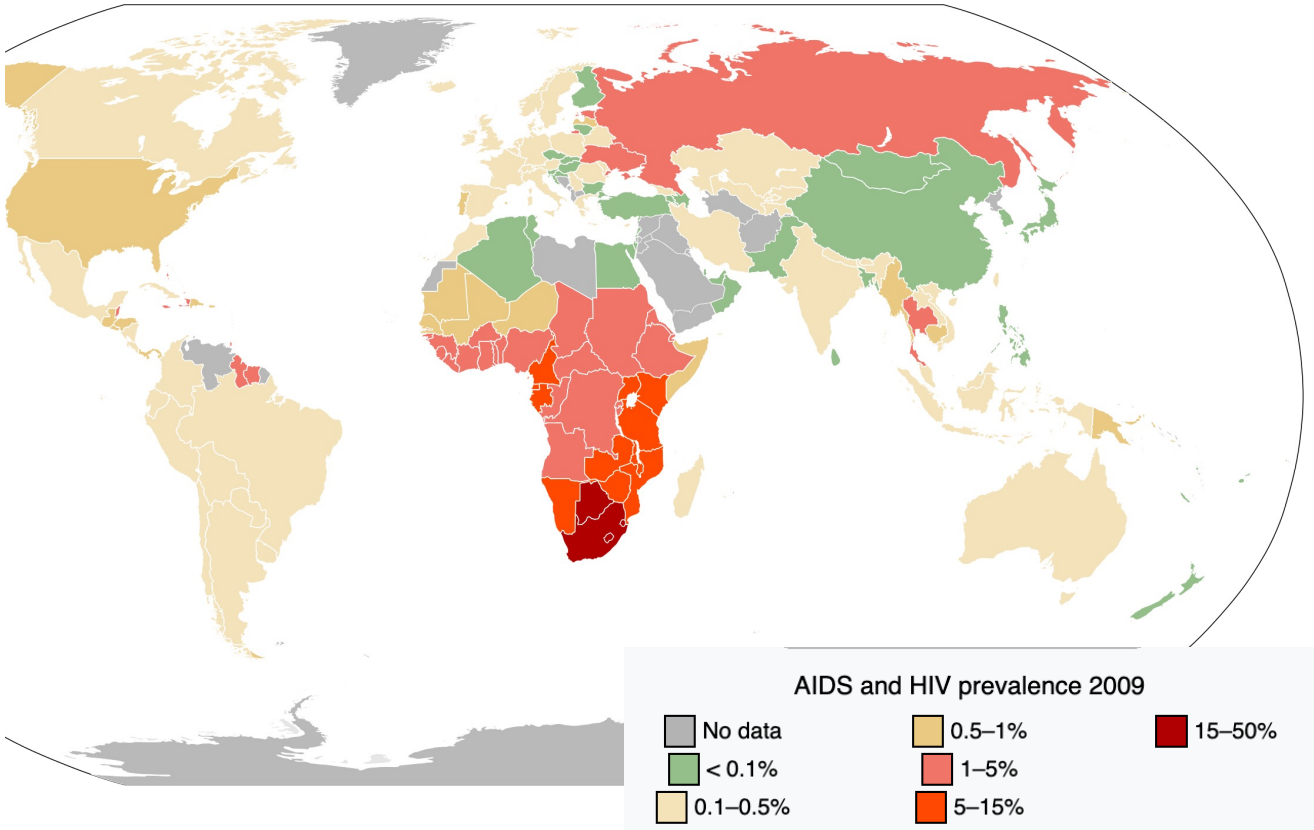
Vol. 1 No. 788
October 18, 1918

To Prevent Influenza!

Do not take any person's breath.
Keep the mouth and teeth clean.
Avoid those that cough and sneeze.
Don't visit poorly ventilated places.
Keep warm, get fresh air and sunshine.
Don't use common drinking cups, towels, etc.
Cover your mouth when you cough and sneeze.
Avoid Worry, Fear and Fatigue.
Stay at home if you have a cold.
Walk to your work or office.
In sick rooms wear a gauze mask like in illustration.

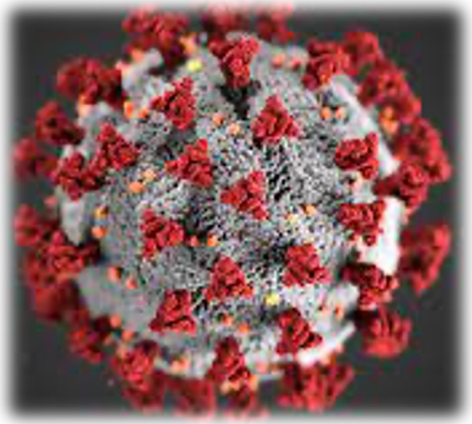
Parasites and pathogens have shaped human history.

- Plague of Justinian (541-549 AD)
- Black Death (1346-1353 AD)
- Cocoliztli (1545-1548)
- Spanish Influenza (1918-1920)
- HIV (~1960-now)
  - >40 million deaths and counting
  - Prevalence still >20% in some countries in southern Africa



Parasites and pathogens have shaped human history.

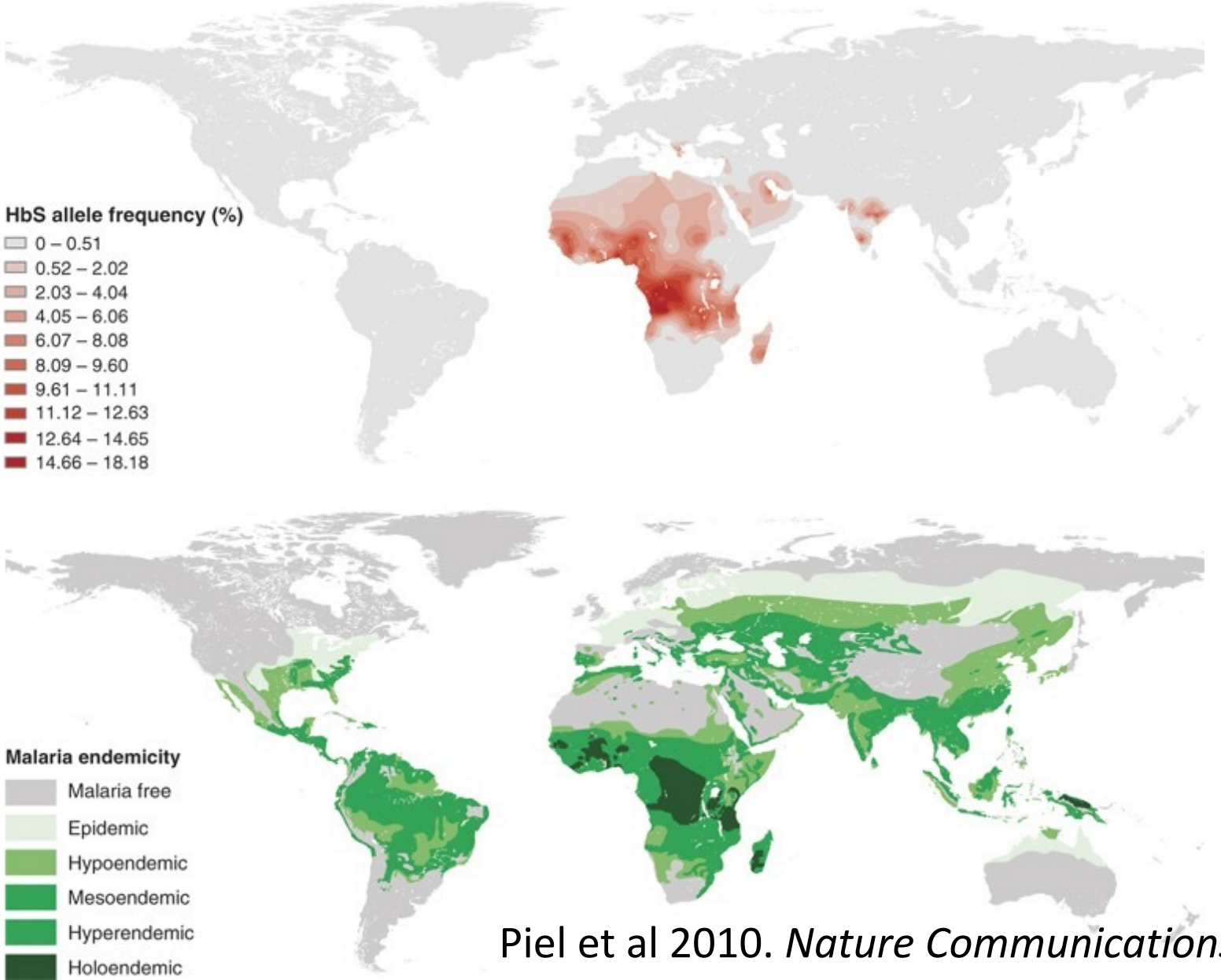
- Plague of Justinian (541-549 AD)
- Black Death (1346-1353 AD)
- Cocoliztli (1545-1548)
- Spanish Influenza (1918-1920)
- HIV (~1960-now)
- COVID-19 (2019-now)
  - ~7-30 million deaths worldwide
  - ~0.1-0.4% of population



Parasites and pathogens have also shaped human DNA.

Sickle cell anemia

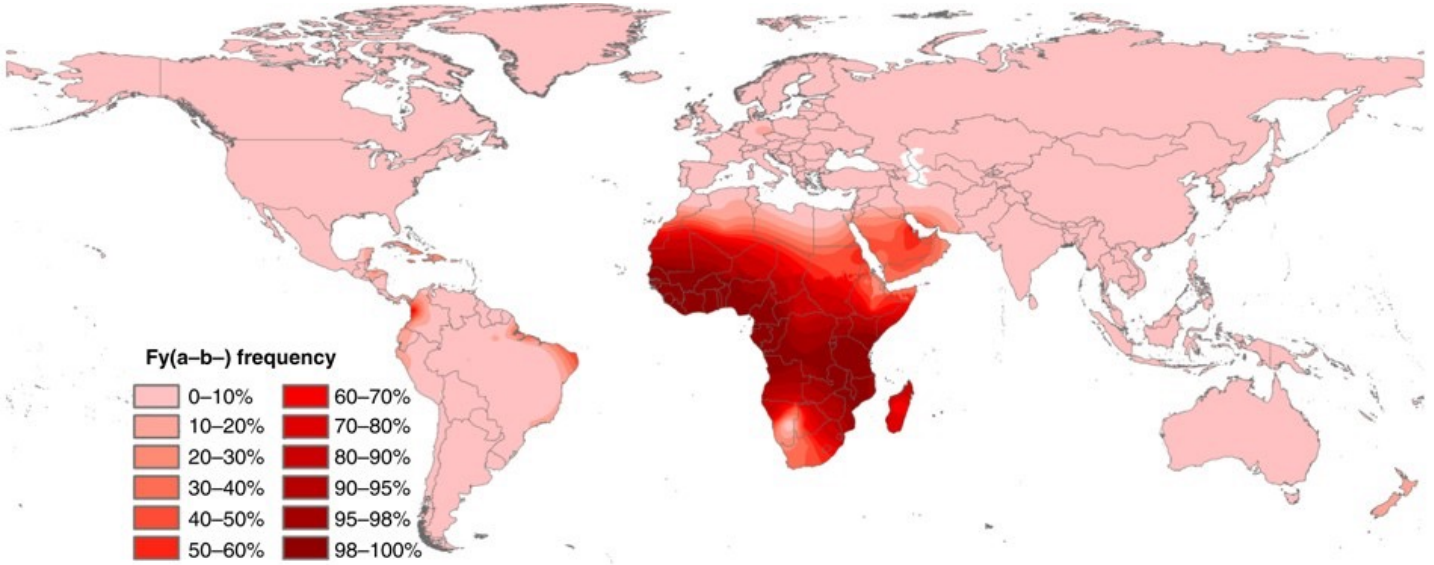
- The HbS allele confers resistance to malaria but also results in sickle cell anemia when homozygous.
- Natural selection has favored this trait in malaria-endemic regions of the planet.
- As of 2021, WHO estimates 247 million malaria cases worldwide and >600,000 deaths, 95% in Africa.
- Children <5 account for 80% of malaria deaths.



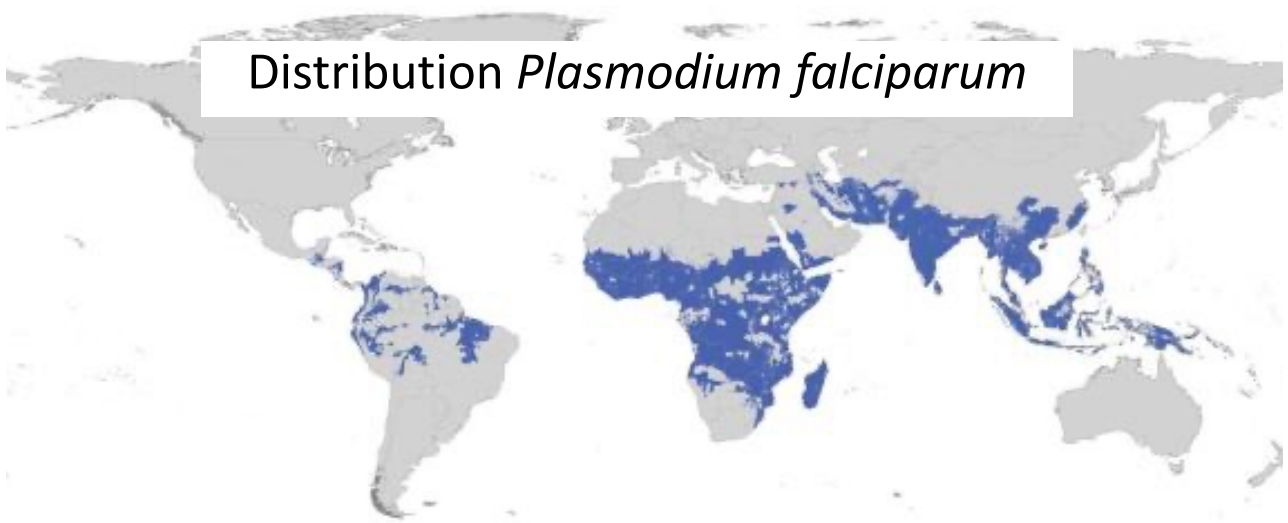
Piel et al 2010. *Nature Communications.*

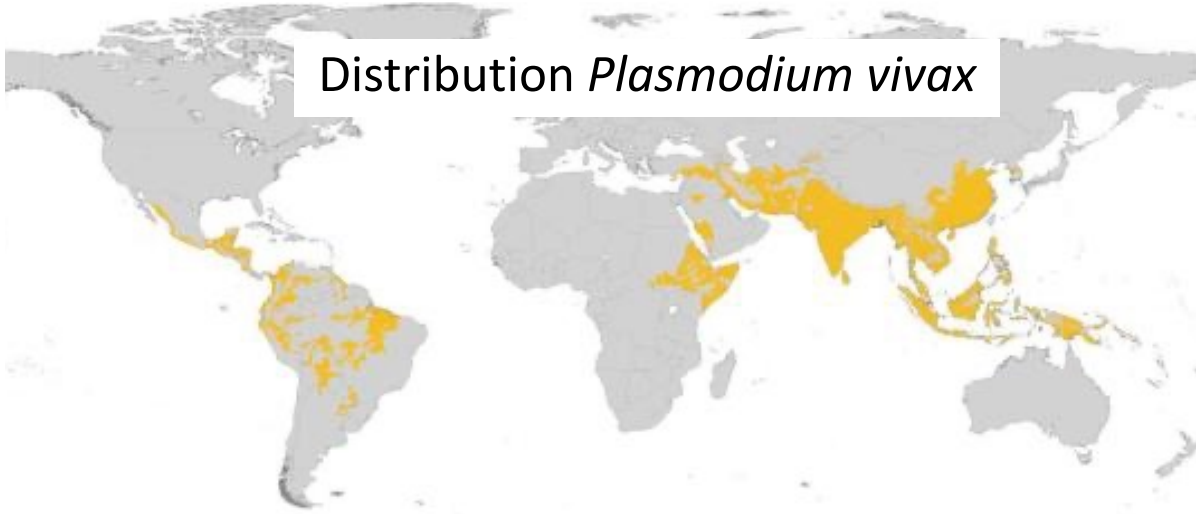
Parasites and pathogens have also shaped human DNA.

Duffy antigen



- Modeled distribution of Duffy-negative human population





Guerra et al. 2006. *Trends in Parasitology*
Howes et al 2011. *Nature Communications.*

We'll use **population biology** to try to understand their transmission.

What is science?

the systematic observation of natural events and conditions in order to discover facts about them and to formulate laws and principles based on these facts.

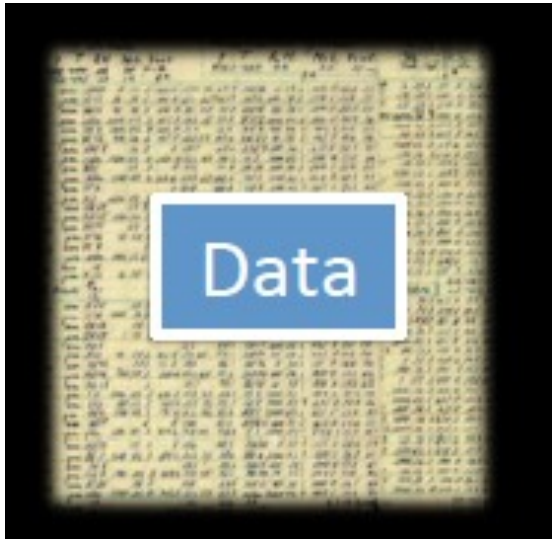
– *Academic Press Dictionary of Science & Technology*

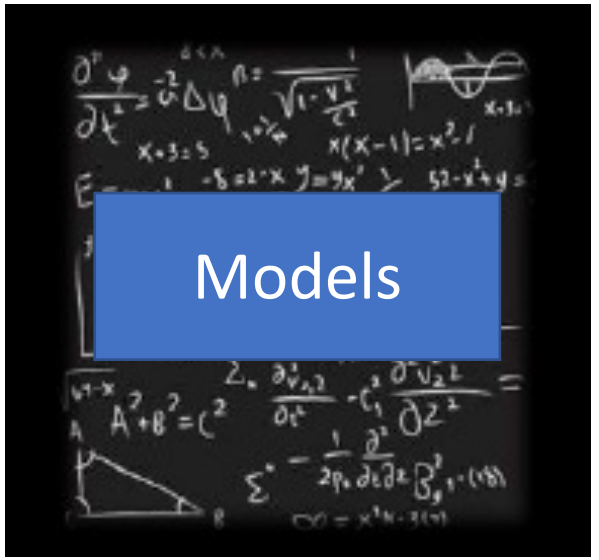
What is science?

the systematic observation of natural events and conditions in order to discover facts about them and to formulate laws and principles

Observations and Laws and Principles

Data and Models

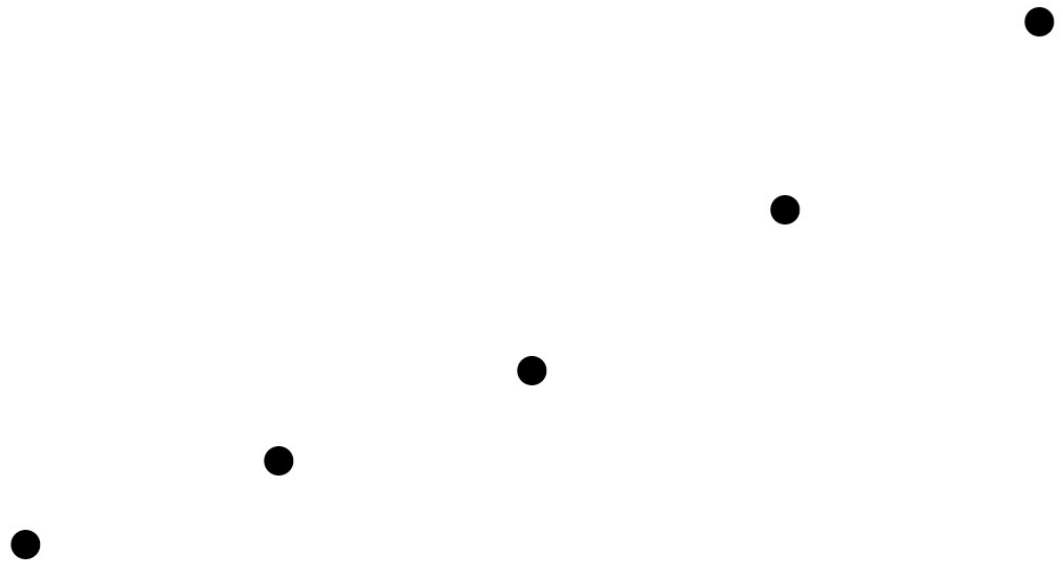




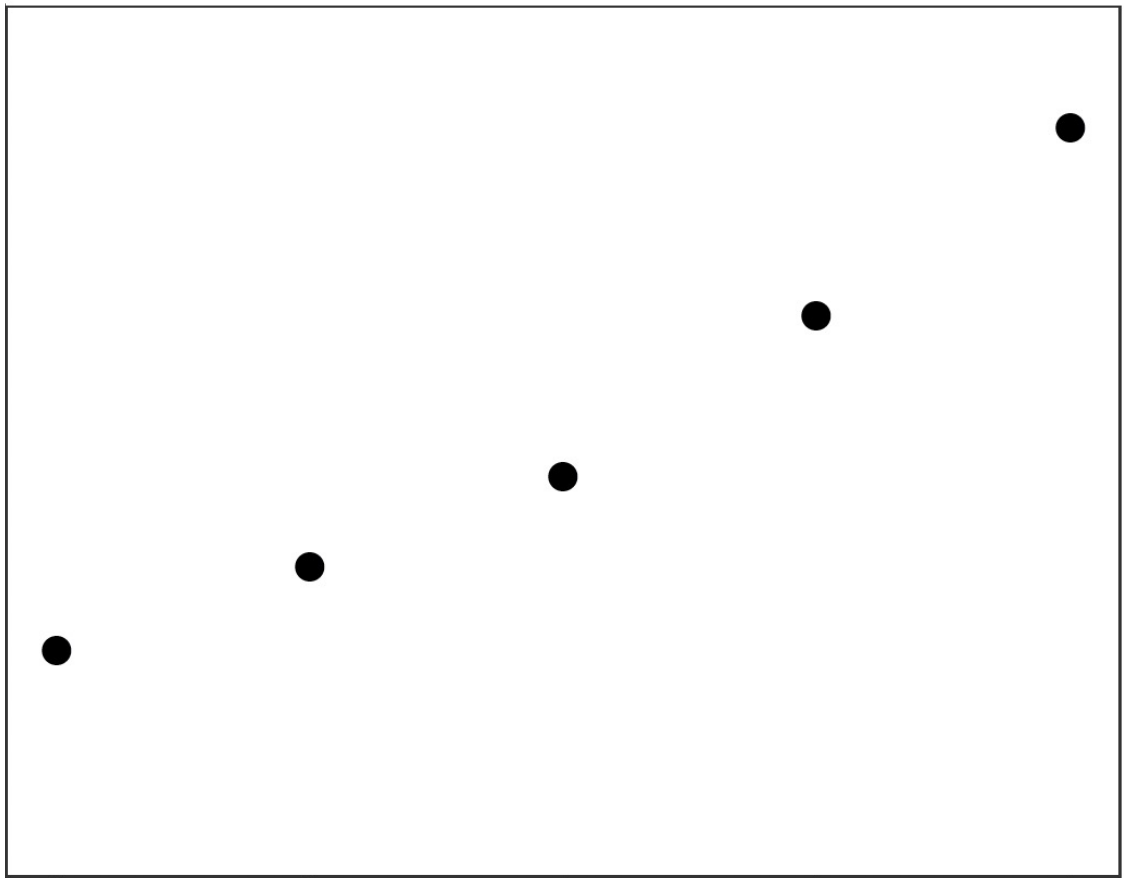
Data and Models

- What are data?
  - Evidence to support a claim

Are these data?

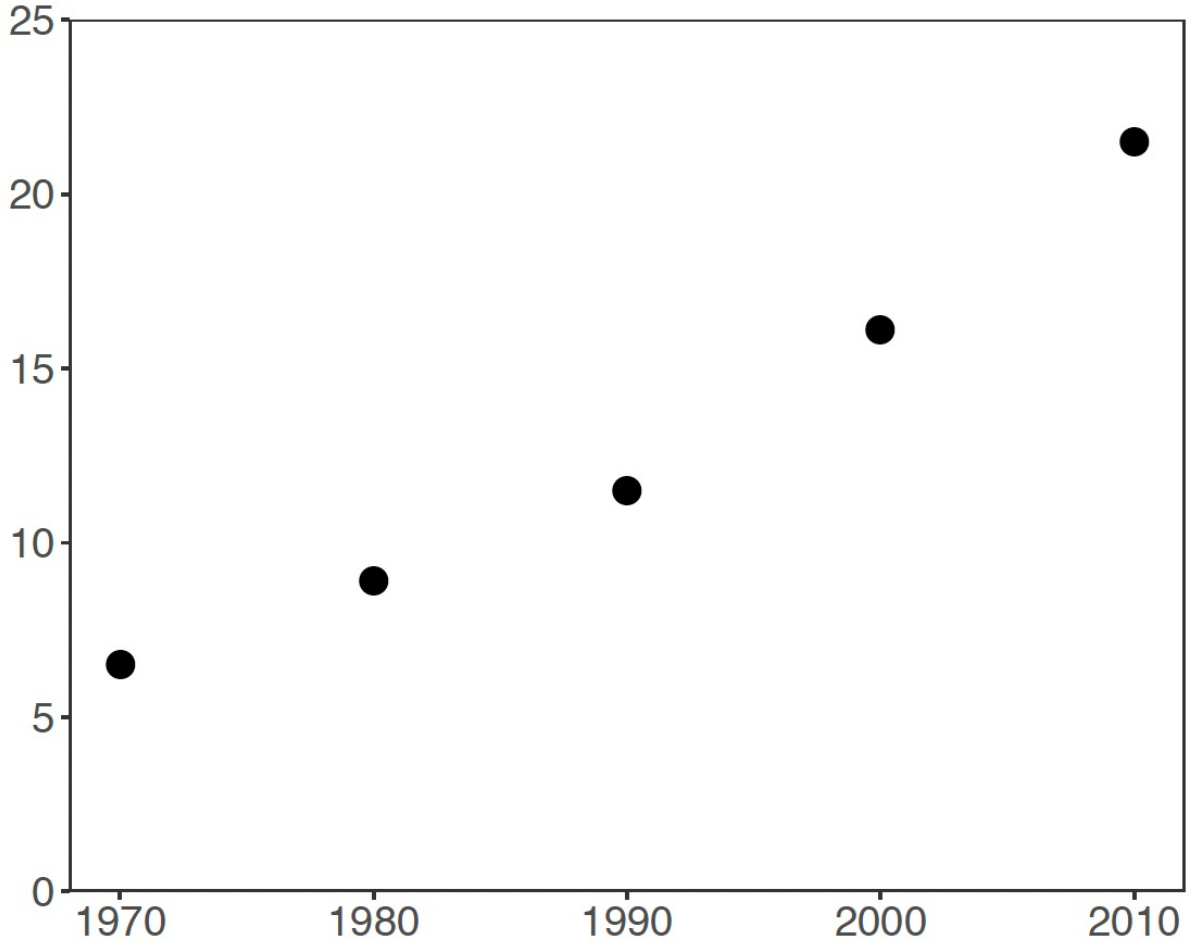


Are these data?

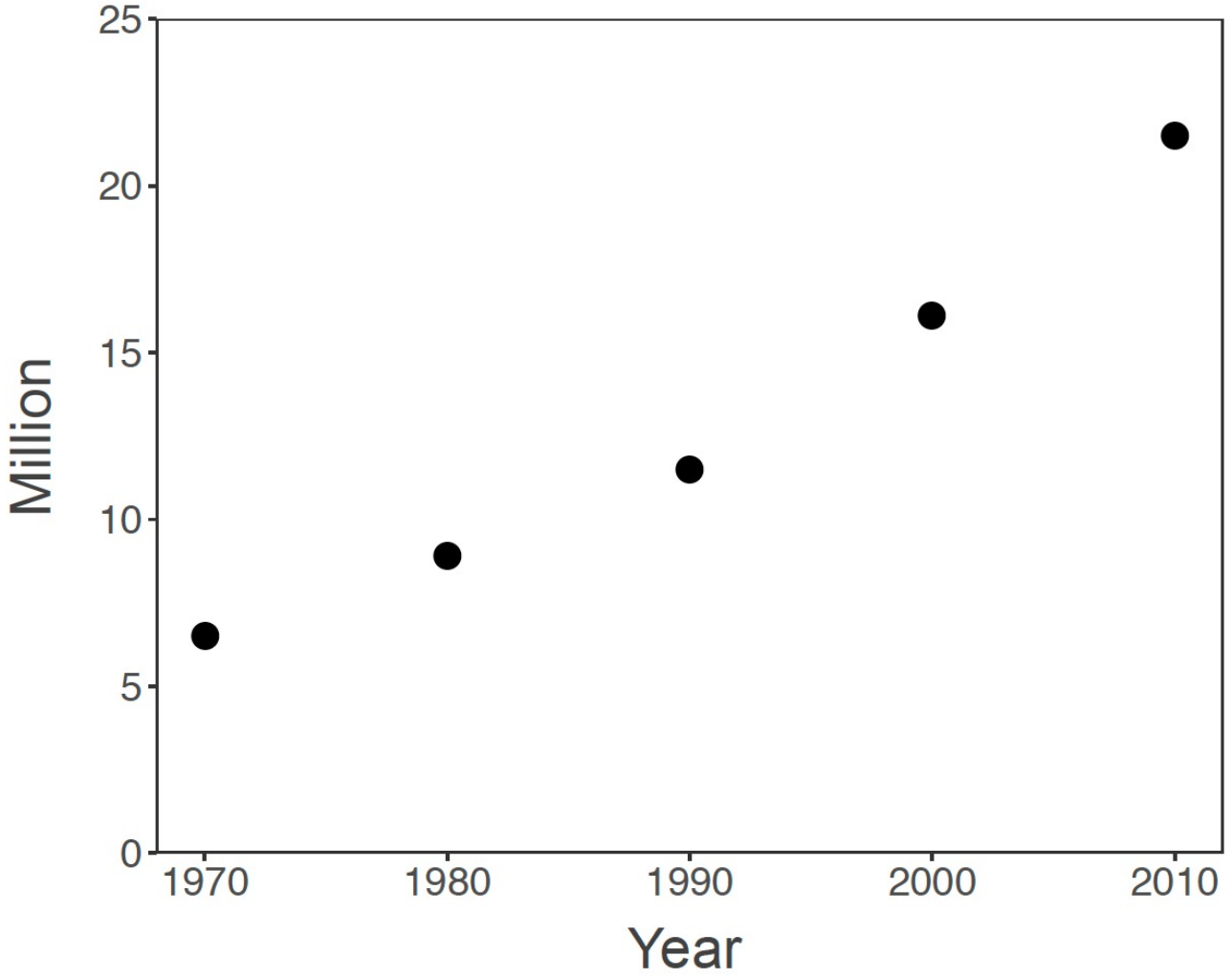


What do we need to make these data?

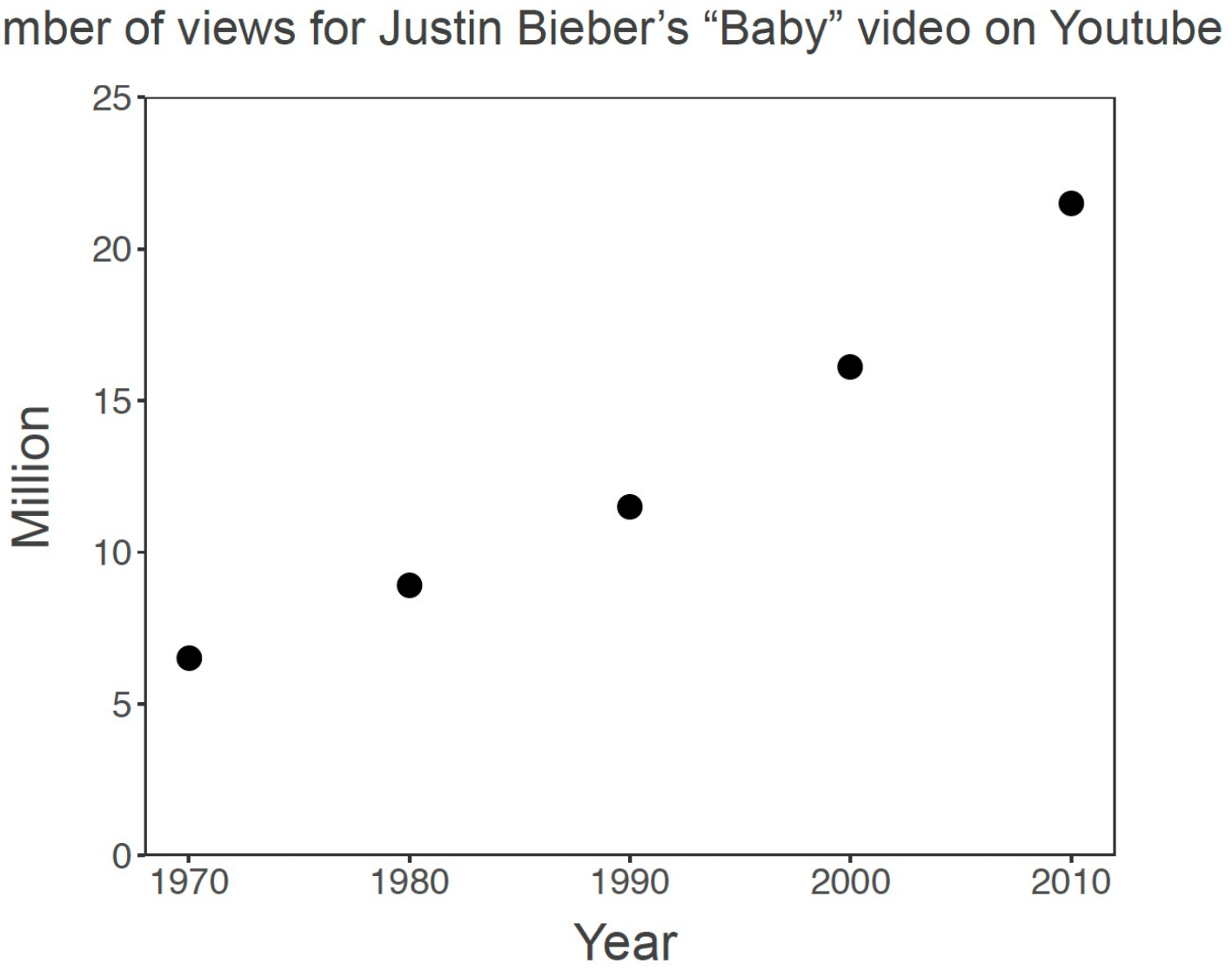
Are these data?



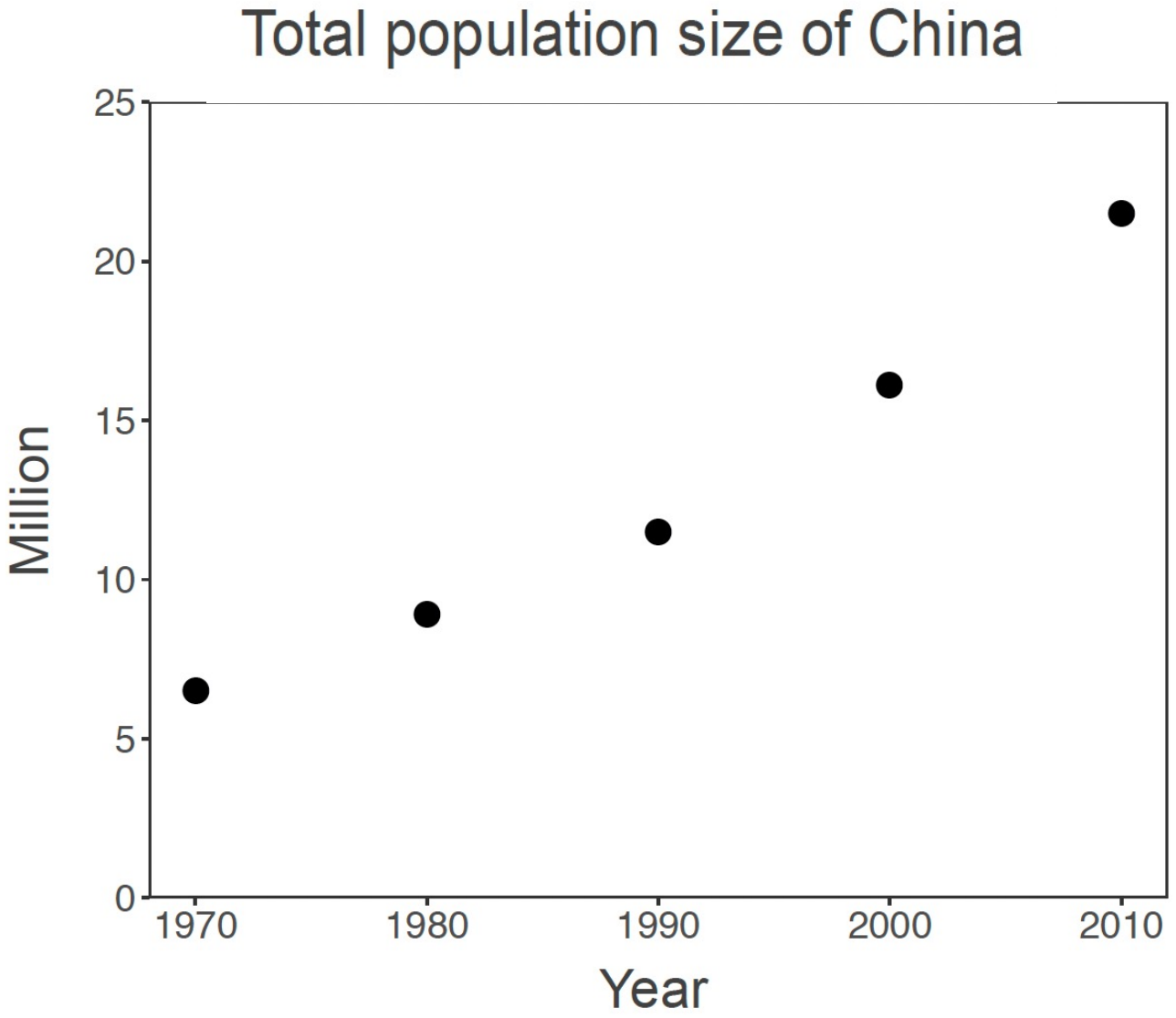
Are these data?



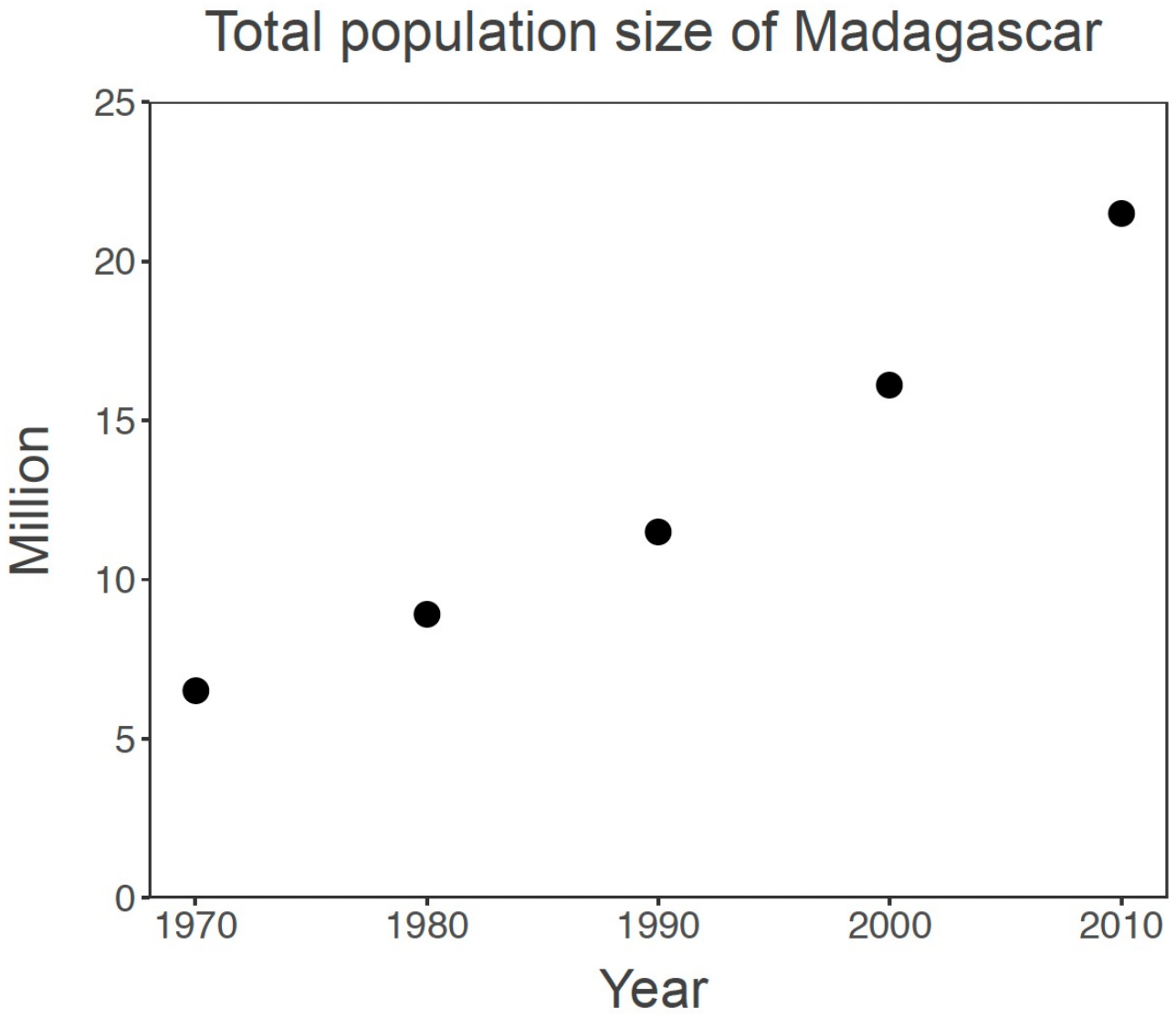
Are these data?



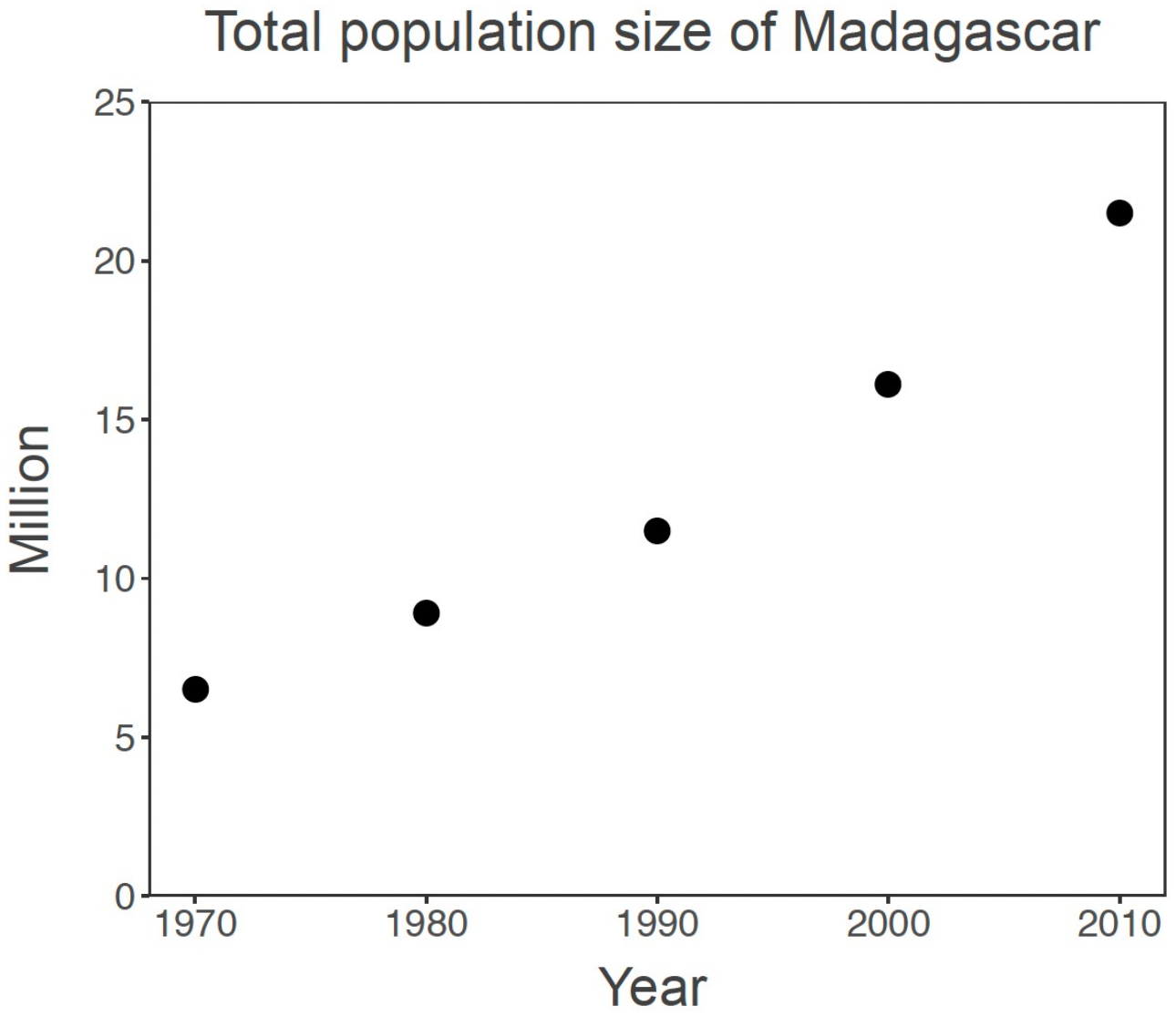
Are these data?



Are these data?



Are these data?

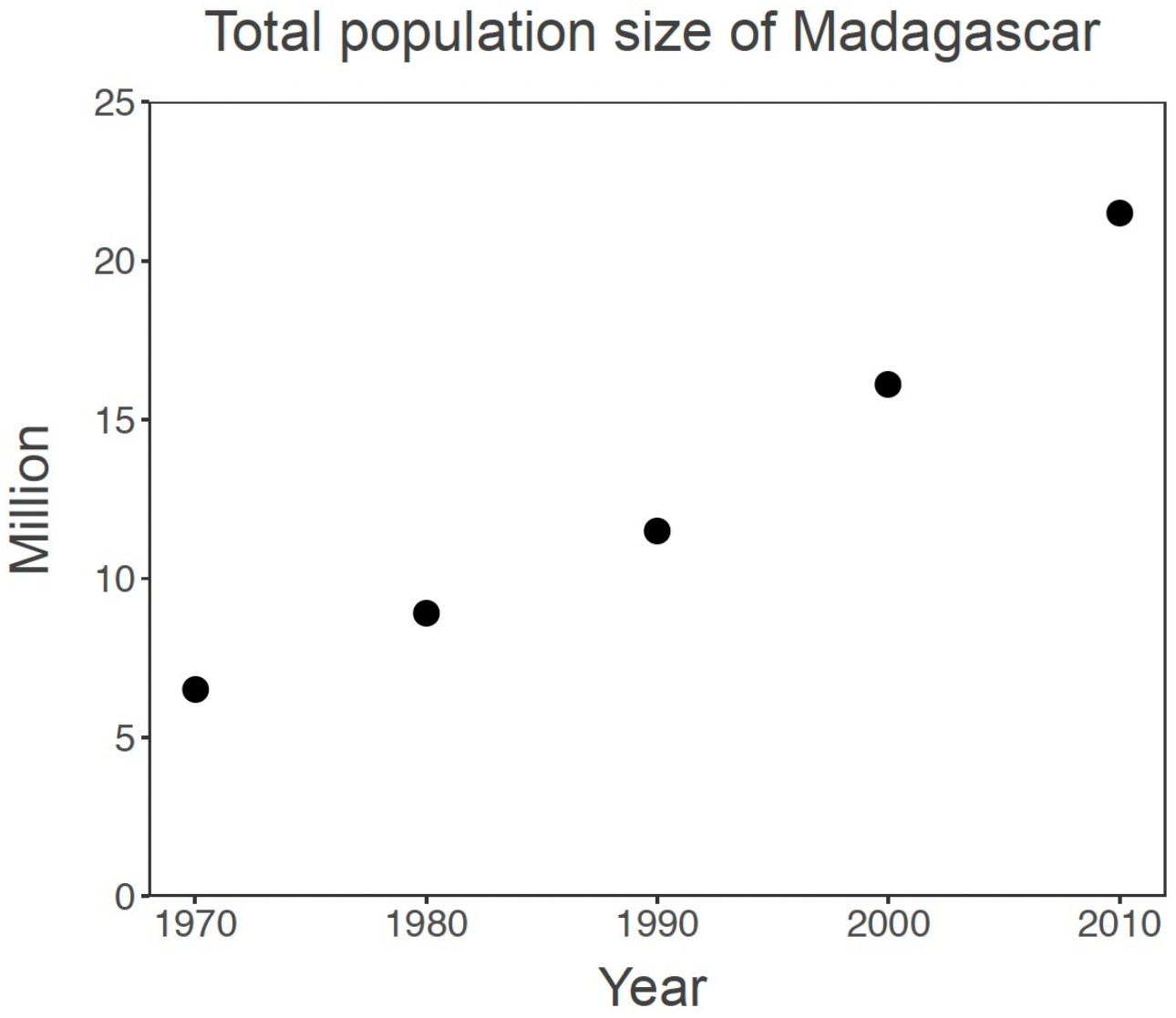


Source: World Bank (accessed 2017)

What are data?

- A relationship between at least two variables
  - x: explanatory, control, driver, independent variable(s)
  - y: response, dependent variable(s)
- x and y should be clearly defined
  - with respect to the **question!**
- **Evidence** to support a **claim**

Data provide evidence to support a claim.

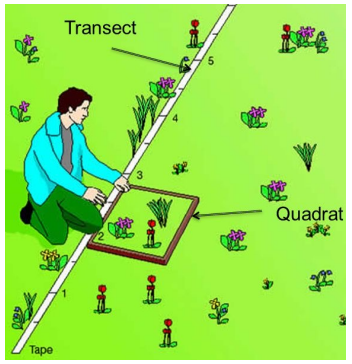


CLAIM: The population size of Madagascar has increased throughout the past 50 years

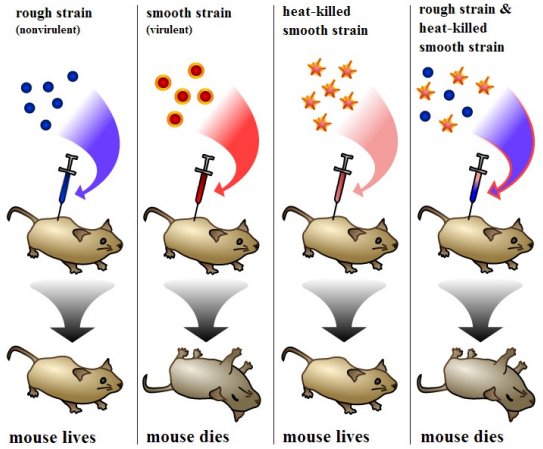
Source: World Bank (accessed 2017)

Data: Sources of x and y

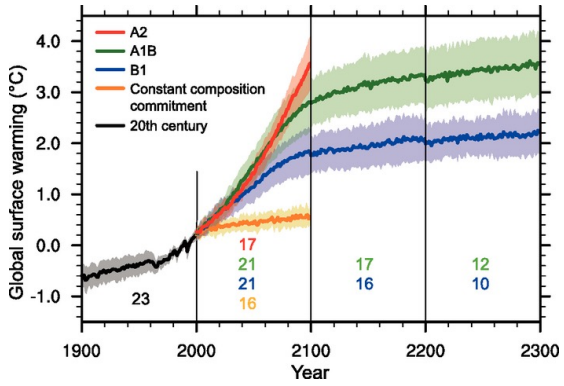
Observational



Experimental



Simulated



Empirical data

Data: Types

Numerical

- A variable is numerical when you can transform it with mathematical operation
- Examples:
- Integer, real number, multi-dimensional number

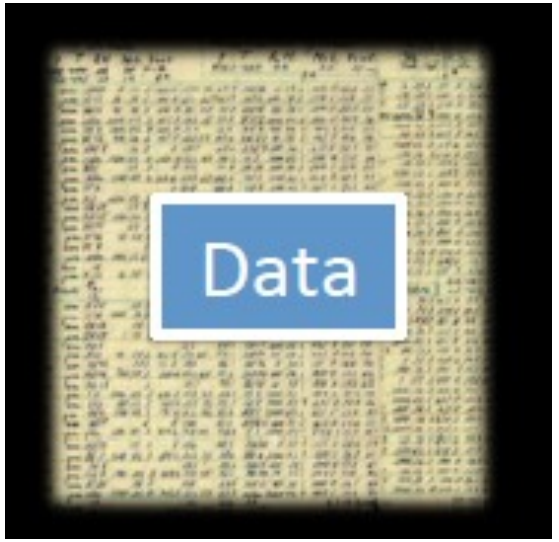
Categorical

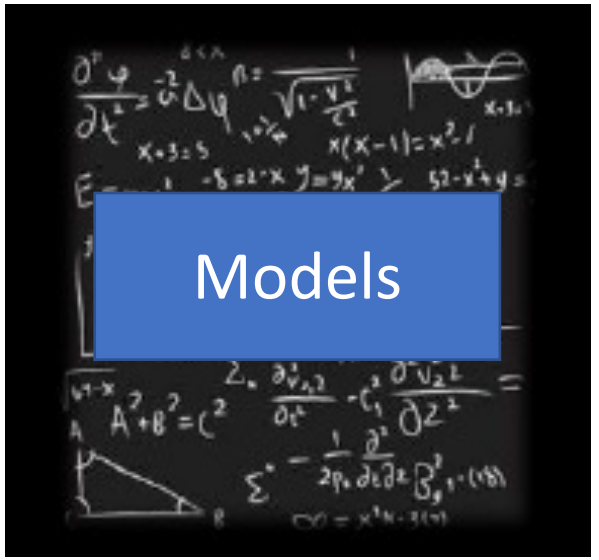
- A variable is categorical when it is not numerical but a categorical can be numerical?
- Examples:
- Colors, (blood) types, species name

Data: Things to consider

- Data acquisition
  - Impossible, example?
  - Theoretically possible but practically unfeasible, examples?
- Data quality and quantity
  - In practice there is always a trade-off
    - Example: monetary cost, human effort -> power analysis, sampling design etc.
- Reproducibility
- Measurement errors
  - Examples?

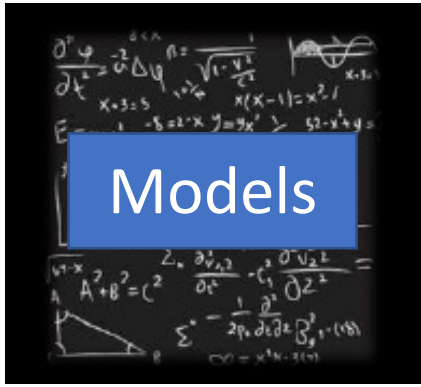
Data and Models





Data and Models

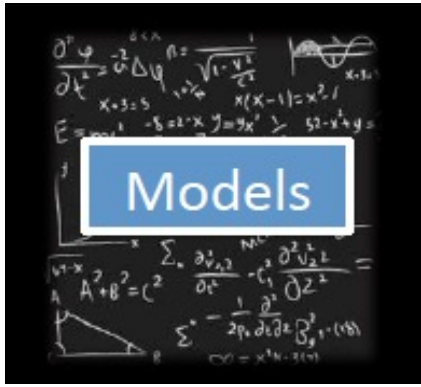
- What is a model?



What is science?

the systematic observation of natural events and conditions in order to discover facts about them and to formulate laws and principles

Laws and Principles



Theory

a declaration to explain a phenomenon

Model

an abstract representation of a phenomenon

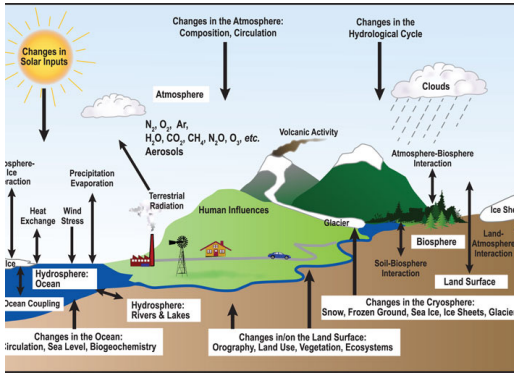
Hypothesis

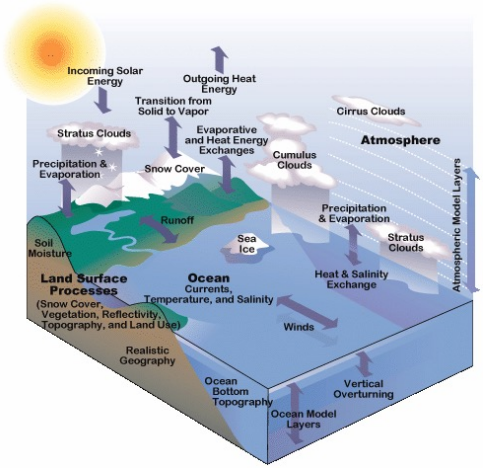
a testable declaration that is derived from a theory

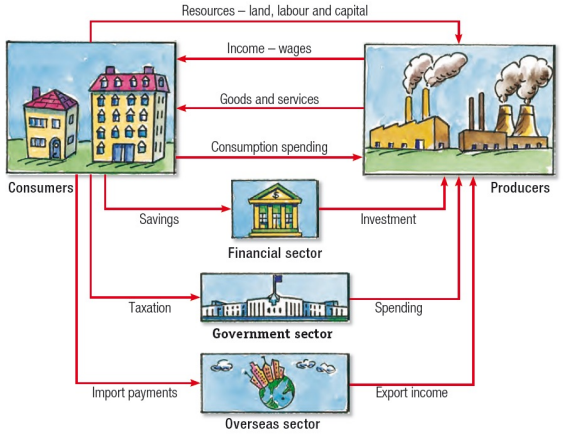
General

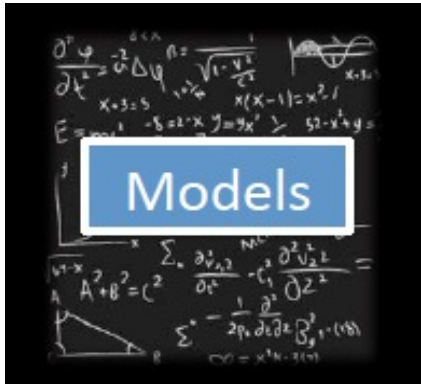
Specific

Models


Changes in Solar Inputs
Changes in the Atmosphere: Composition, Circulation
Changes in the Hydrological Cycle
Atmosphere
Clouds
Volcanic Activity
Atmosphere-Biosphere Interaction
Precipitation Evaporation
Terrestrial Radiation
Heat Exchange
Wind Stress
Human Influences
Glacier
Biosphere
Soil-Biosphere Interaction
Land Surface
Hydrosphere: Ocean
Hydrosphere: Rivers & Lakes
Changes in the Cryosphere: Snow, Frozen Ground, Sea Ice, Ice Sheets, Glacie
Changes in the Ocean: Circulation, Sea Level, Biogeochemistry
Changes in/on the Land Surface: Orography, Land Use, Vegetation, Ecosystems


Incoming Solar Energy
Outgoing Heat Energy
Transition from Solid to Vapor
Cirrus Clouds
Stratus Clouds
Evaporative and Heat Energy Exchanges
Atmosphere
Precipitation & Evaporation
Snow Cover
Cumulus Clouds
Runoff
Precipitation & Evaporation
Soil Moisture
Sea Ice
Stratus Clouds
Land Surface Processes
(Snow Cover, Vegetation, Reflectivity, Topography, and Land Use)
Ocean
Currents, Temperature, and Salinity
Heat & Salinity Exchange
Atmospheric Model Layers
Winds
Realistic Geography
Ocean Bottom Topography
Vertical Overturning
Ocean Model Layers


Resources – land, labour and capital
Income – wages
Goods and services
Consumption spending
Consumers
Producers
Savings
Investment
Financial sector
Taxation
Spending
Government sector
Import payments
Export income
Overseas sector



- When you make a model,

  you include the

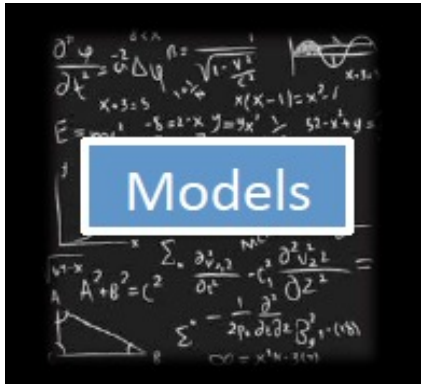
  elements that you feel are most important
  to explain a phenomenon.

- Generally, we try to make models

  that can reproduce real-world data

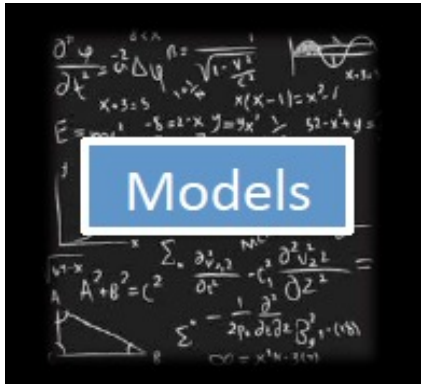
- We can distinguish between
  statistical and mechanistic models

Statistical vs. Mathematical Model



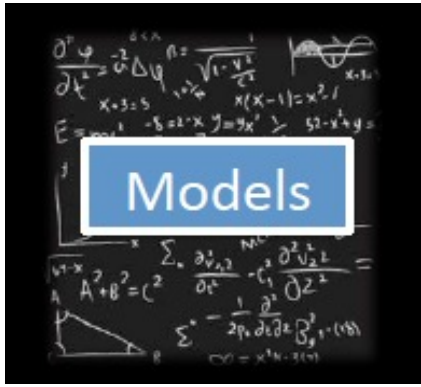
The choice depends on the research question!

Statistical Models



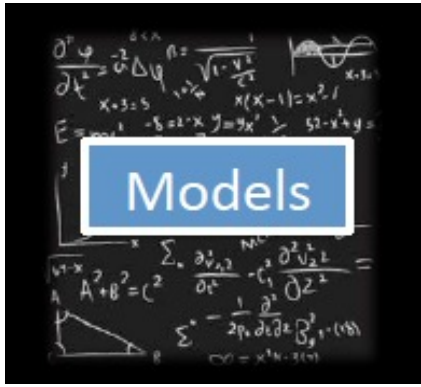
- Goal: To rigorously **assess** the strength of relationship between x and y

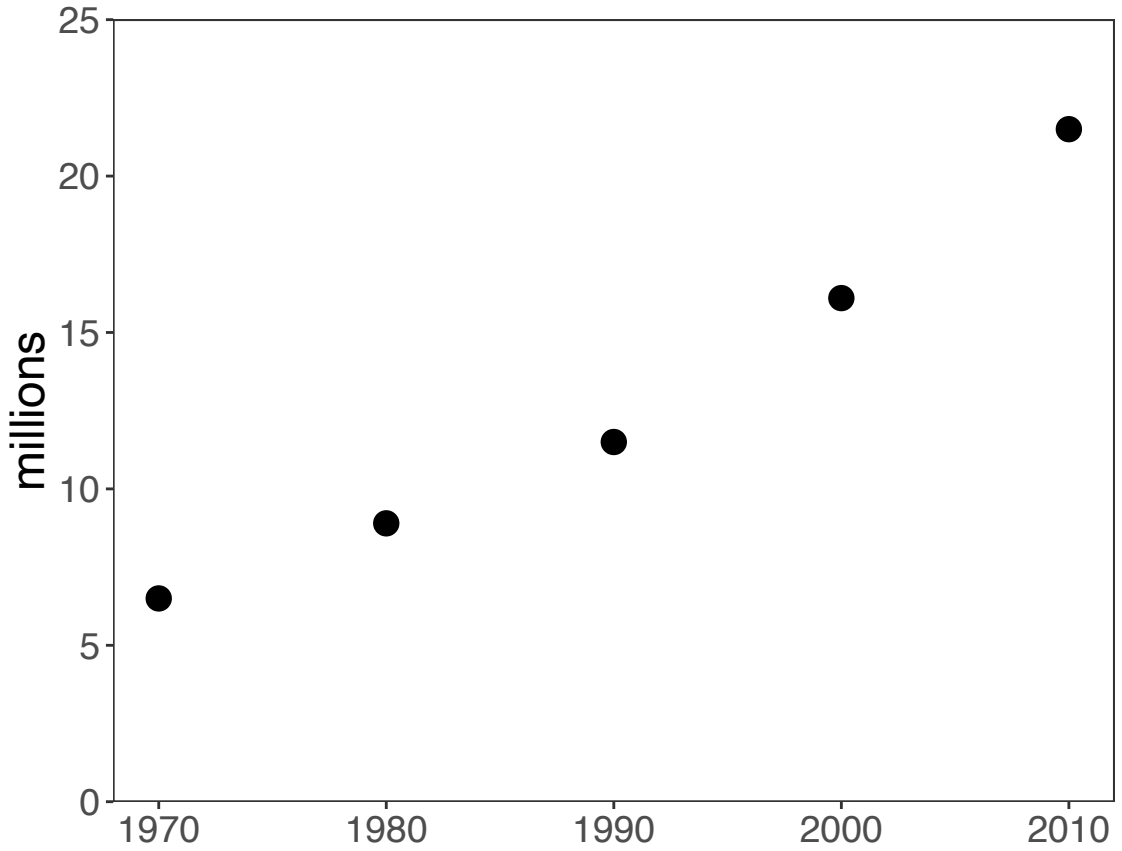
Statistical Models



- Goal: To rigorously **assess** the strength of relationship between x and y (describe patterns)
  - Find a significant relationship using a p-value as a measure of relationship strength
  - **Statistical models can demonstrate correlations.**
- Steps:
  1. Formulate a research question
  2. Formulate a hypothesis
  3. Develop a model to demonstrate your hypothesis.
  4. Collect **data (required!!!)**
  5. Evaluate hypothesis with appropriate statistical tools
     - t-test, Chi-square, ANOVA
     - Ordination (PCA)
     - Regression (LM, GLM, GLMM, GAM)

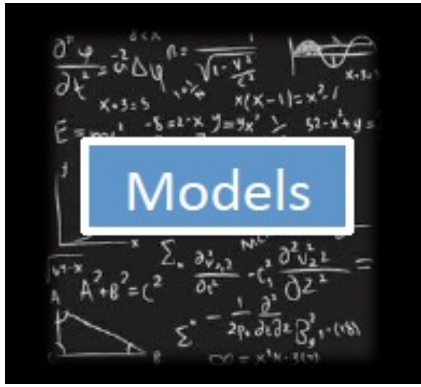
1. Example Question: What is the trajectory Malagasy population size through time?





Source: World Bank

1. Example Question: What is the trajectory of Malagasy population size through time?



2. Hypothesis: Malagasy population size increases with time

3. Statistical Model:

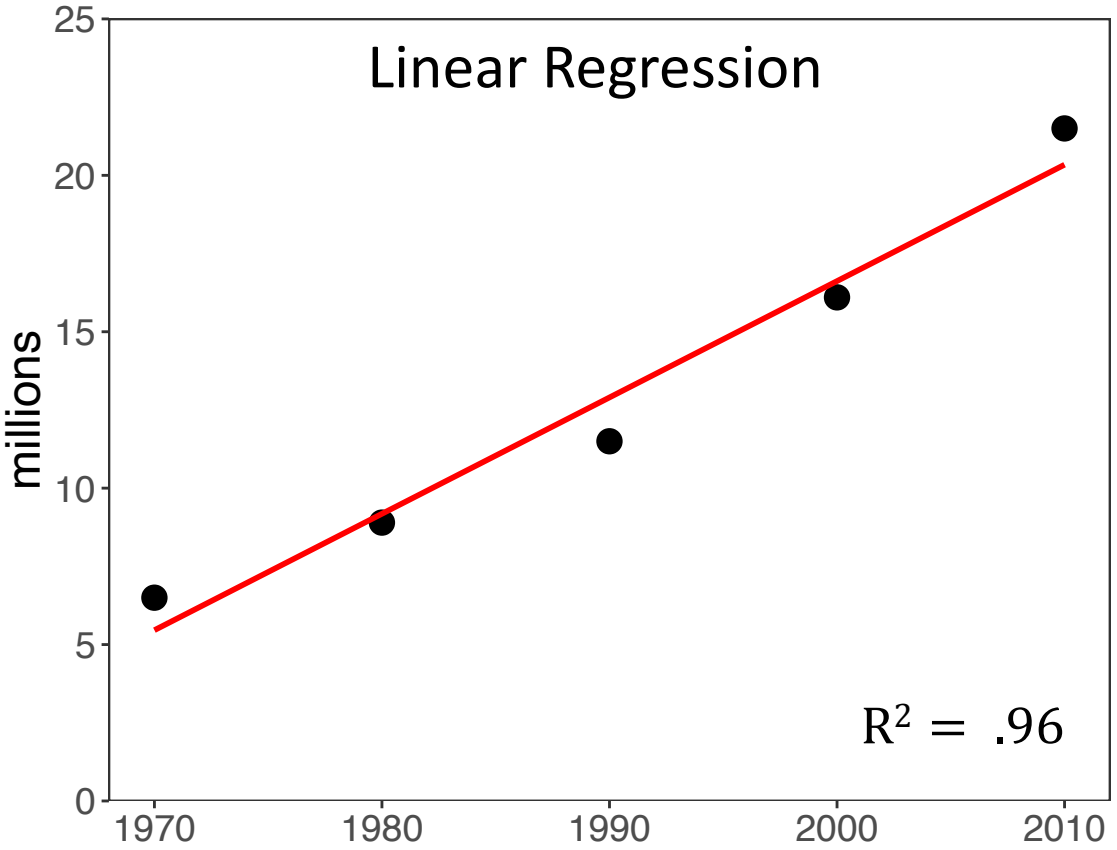
$$y = mx + b$$

5. Evaluation

$m = .372$ million

$p = .003$

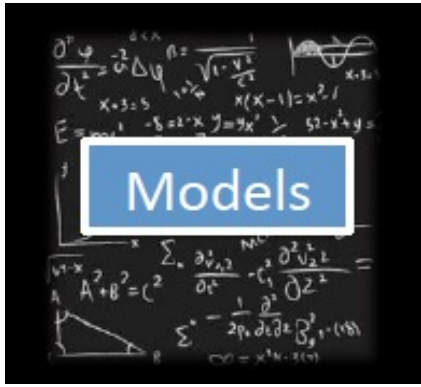
4. Data:



What can we conclude from this fitted model?

Source: World Bank

1. Example Question: What is the trajectory of Malagasy population size through time?



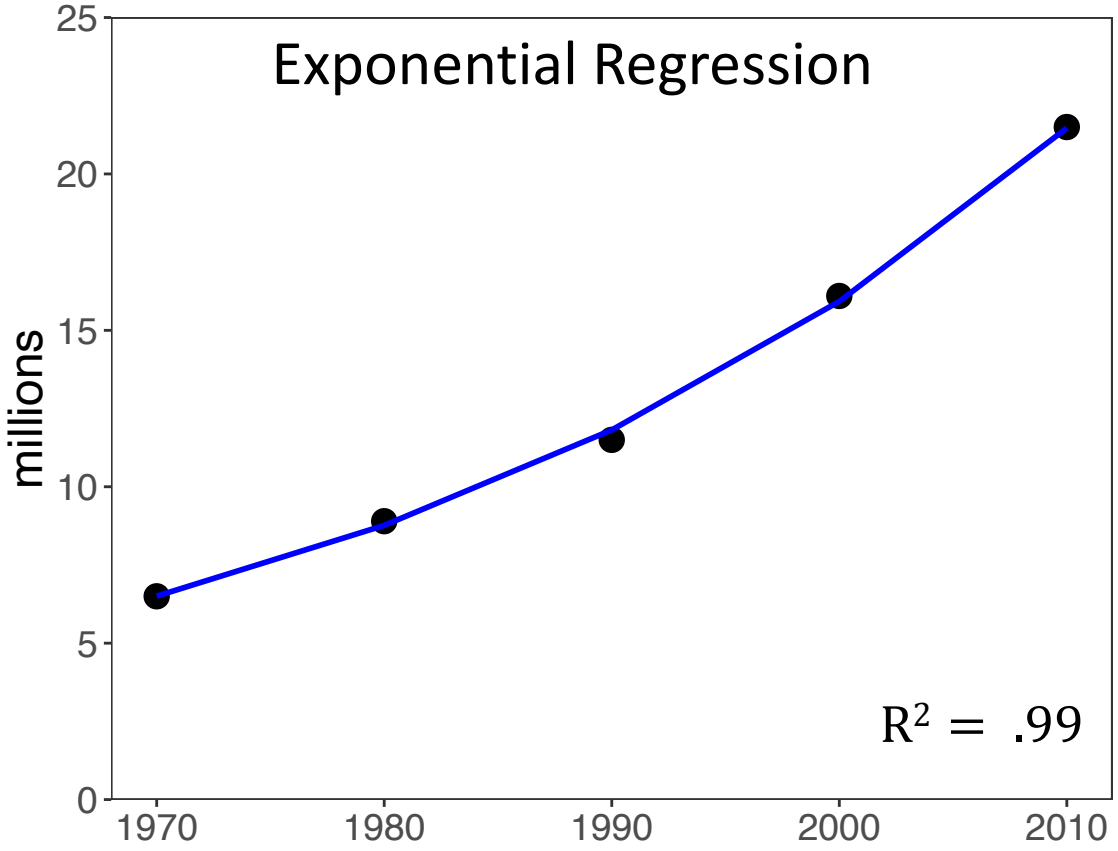
6. Adapt your model and re-evaluate:

$$y = e^{mx+b}$$

Exponential Regression

$m = 0.029$ mil.

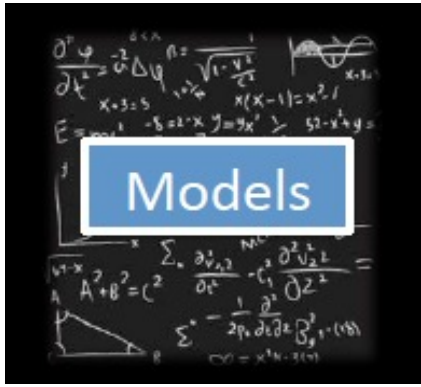
$p < .001$



What can we conclude from this fitted model?

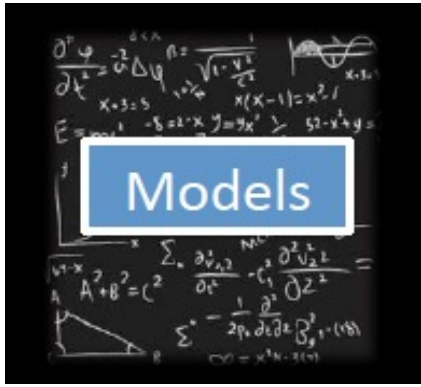
Source: World Bank

Statistical Models: Beware!



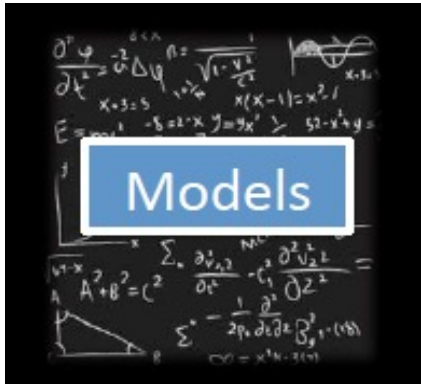
- Statistical models and tests are based on specific assumptions
  - data normally distributed
  - x and y independent
- Assessing a model means you need to make sure the assumptions are not violated.
- There are so many statistical models…

Statistical vs. Mathematical Model



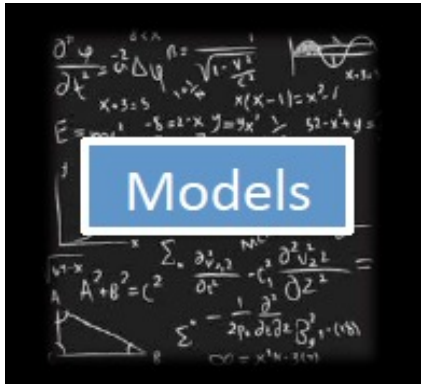
The choice depends on the research question!

Mechanistic Models



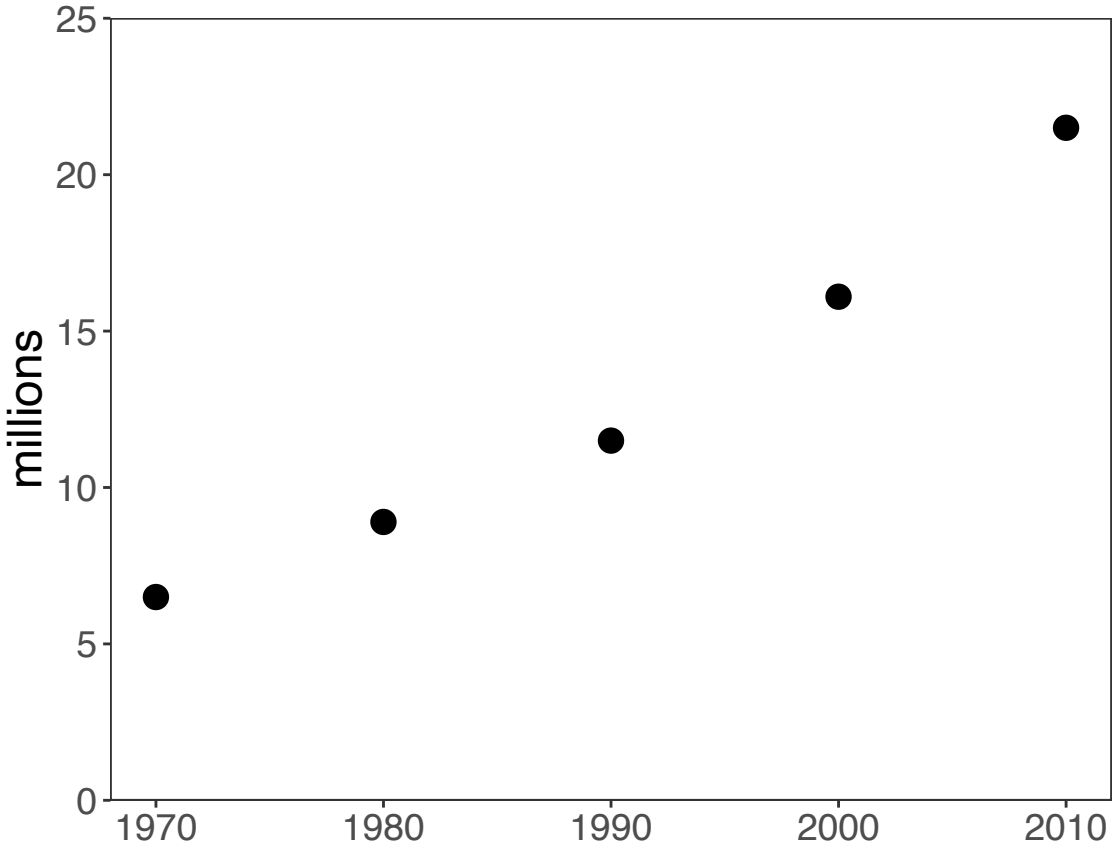
- Goal: To **demonstrate the processes** that underlie a relationship between x and y
  - Find a significant relationship using a p-value as a measure of relationship strength
  - **Mechanistic models can demonstrate causation.**
- Steps:
  1. Formulate a research question
  2. Formulate a hypothesis
  3. Develop a model to demonstrate your hypothesis.
  4. Collect **data** (for certain questions)
  5. Evaluate the extent to which your model-simulated data matches that from the real world.

1. Example Question: How does Malagasy population size change with time?



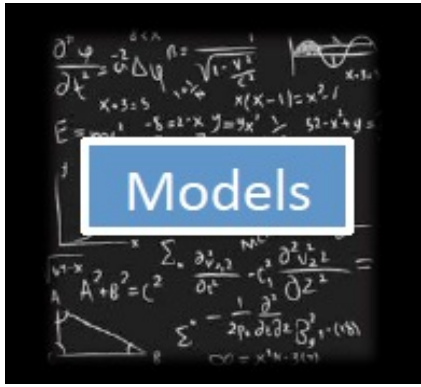
2. Hypothesis: Malagasy population size increases because people are having children.

Can you think of an alternative hypothesis?



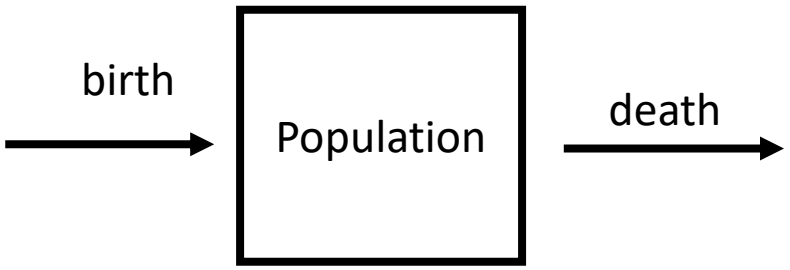
Source: World Bank

1. Example Question: How does Malagasy population size change with time?



2. Hypothesis: Malagasy population size increases because people are having children.

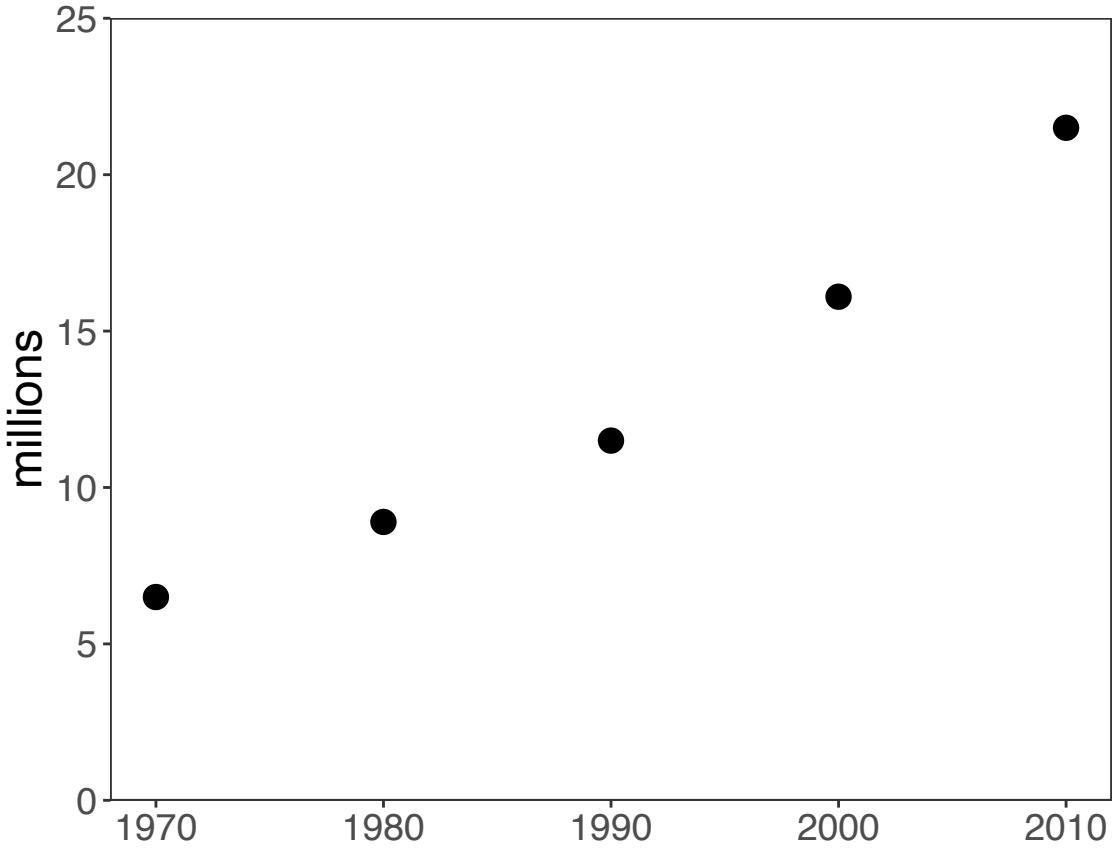
3. Mechanistic Model:



$$P_{t+1} = P_t + b * P_t - d * P_t$$

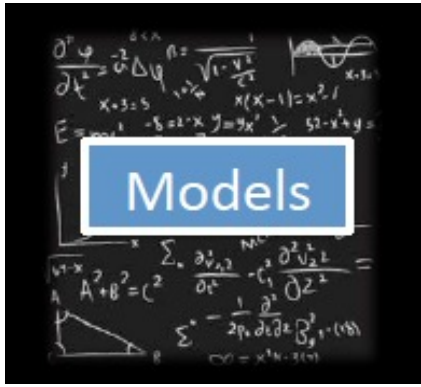
$$P_{t+1} = P_t + r * P_t$$

4. Data:



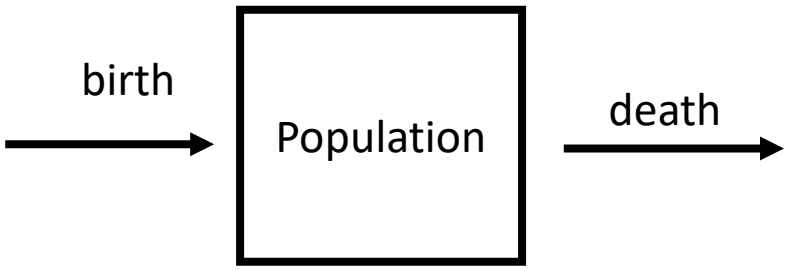
Source: World Bank

1. Example Question: How does Malagasy population size change with time?

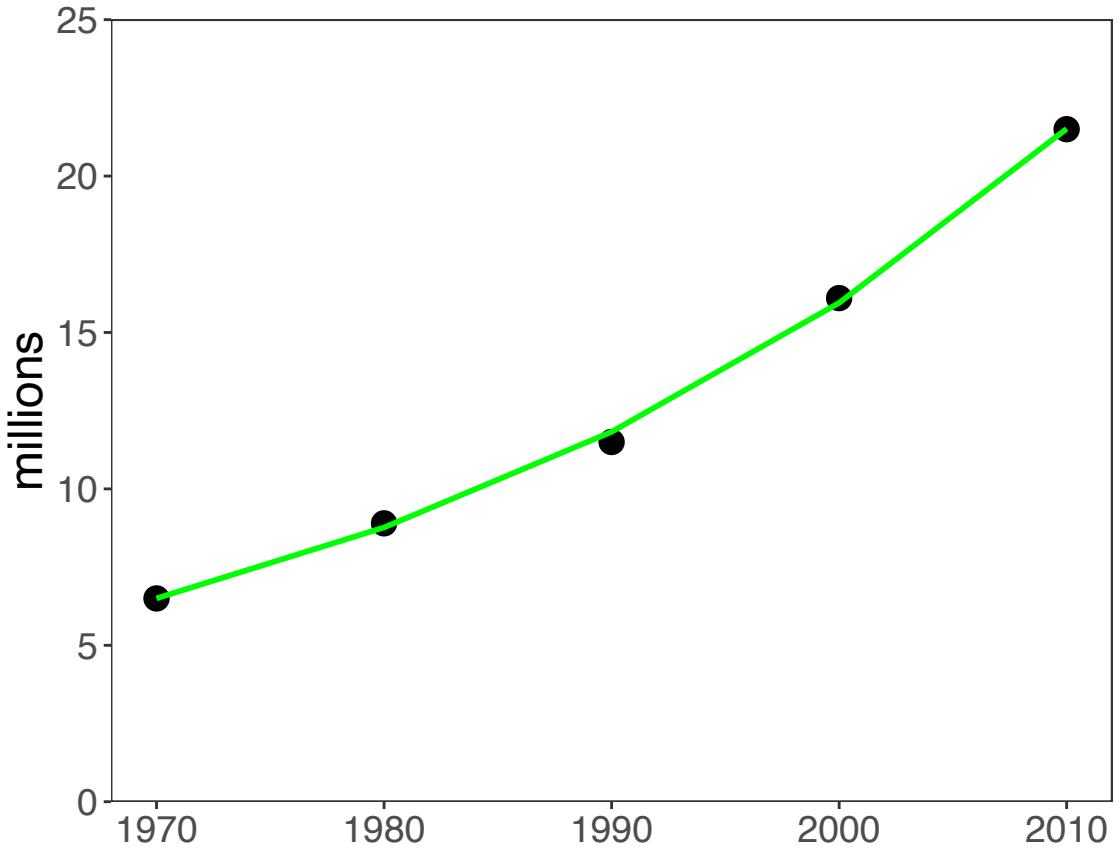


2. Hypothesis: Malagasy population size increases because people are having children.

3. Mechanistic Model:



4. Data:

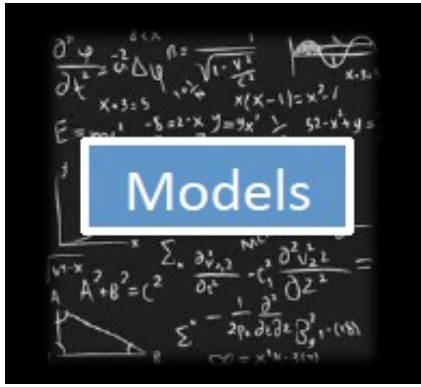


5. Evaluation:

$r = .349$/person/yr

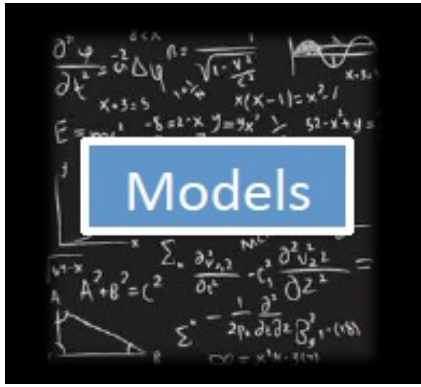
What can we conclude from this fitted model?

Mechanistic Models: Beware!



- Parameters used in the mechanistic models sometimes are not measurable!

Mechanistic Models: Beware!



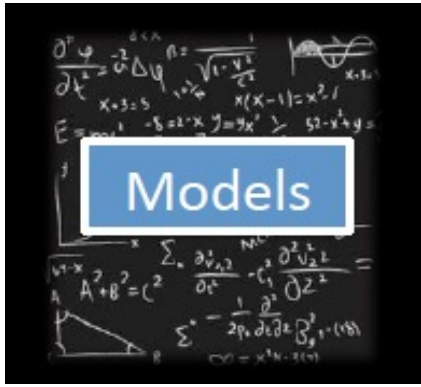
- Parameters used in the mechanistic models sometimes are not measurable!
- Simulations can be computationally intensive
- Advances in computational power often inspire development of more complex models which are not necessarily better

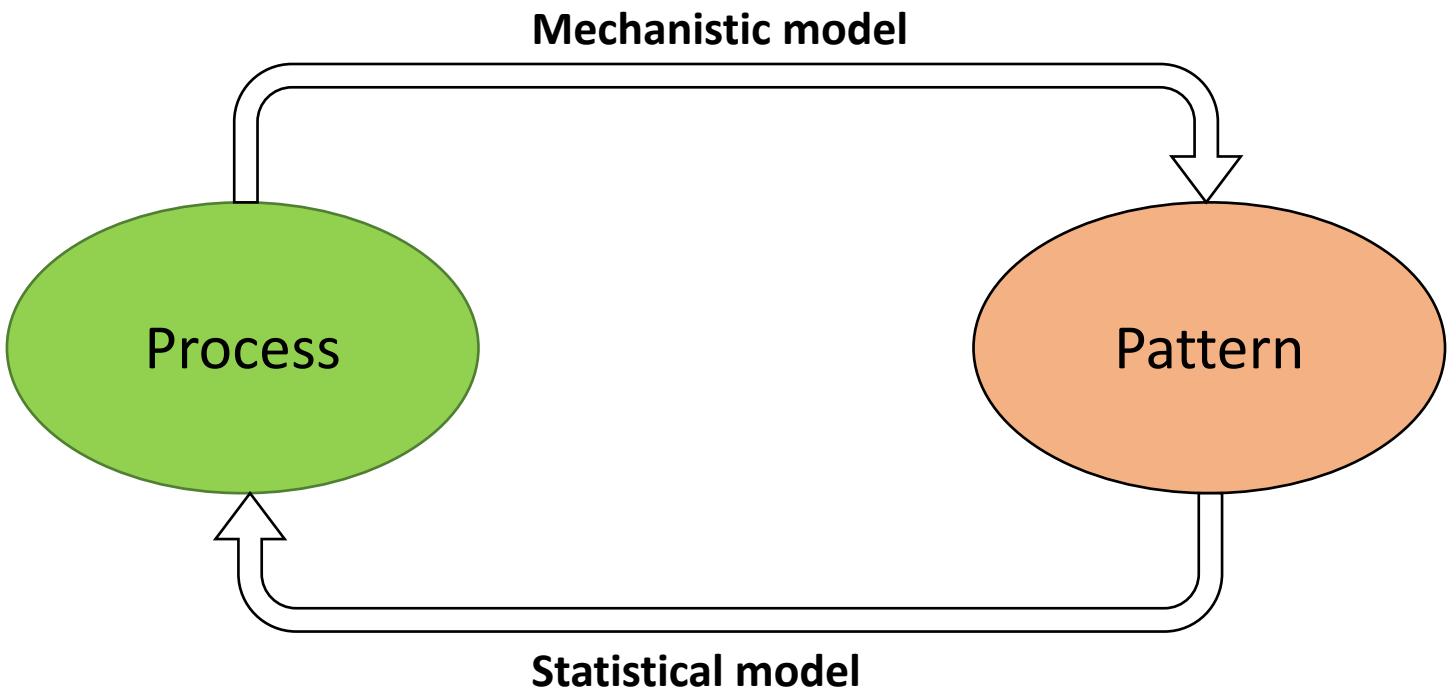
"All models are wrong but some are useful…"

-George Box

We use models to both **predict** and **explain**.

It is ideal when statistical and mechanistic models meet:





Models and Data through the ages:

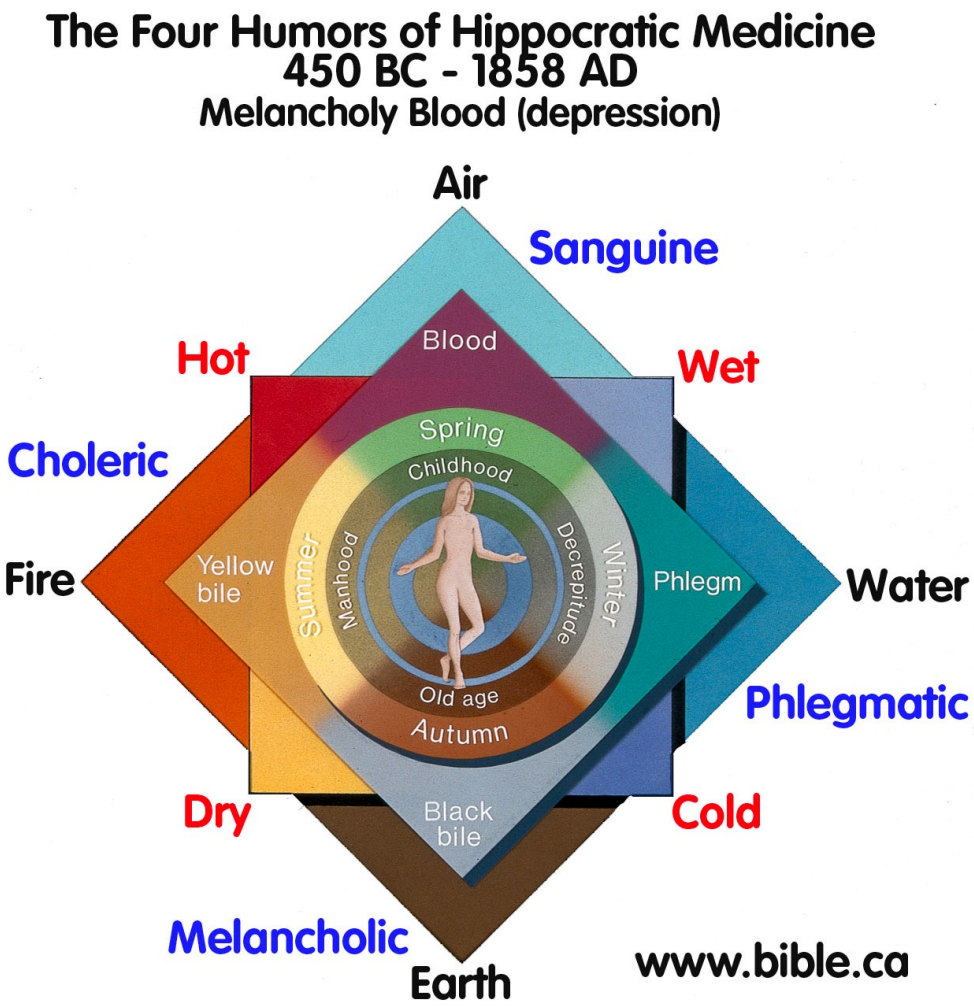
Parasites and pathogens have shaped human history.

But we have not always known they were responsible for disease!

Models and Data through the ages:

What is responsible for disease?

1. **<u>Four Humors:</u>** Hippocrates (c. 400 BC) wrote that disease results from an imbalance of the four humors



Models and Data through the ages:

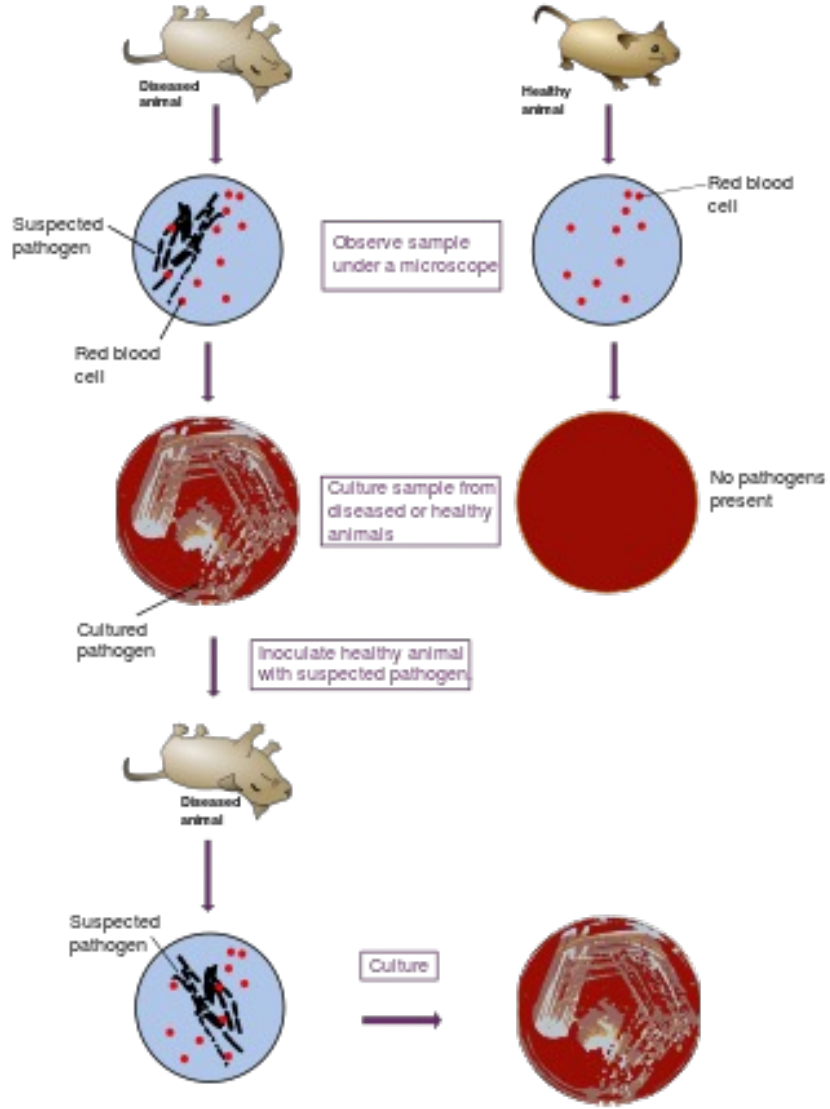
What is responsible for disease?

1. **<u>Four Humors:</u>** Hippocrates (c. 400 BC) wrote that disease results from an imbalance of the four humors
2. **<u>Miasmatic Theory:</u>** Extension of Hippocrates that lasted through the 1800s – idea that disease was caused by bad air. Popularized by Florence Nightengale

Models and Data through the ages:

What is responsible for disease?

1. **<u>Four Humors:</u>** Hippocrates (c. 400 BC) wrote that disease results from an imbalance of the four humors
2. **<u>Miasmatic Theory:</u>** Extension of Hippocrates that lasted through the 1800s – idea that disease was caused by bad air.
   - Popularized by Florence Nightengale
3. **<u>Germ Theory of Disease:</u>** Idea that disease results from germs
   - Leuwenhoek's microscope (1675)
   - Koch's postulates (1890)



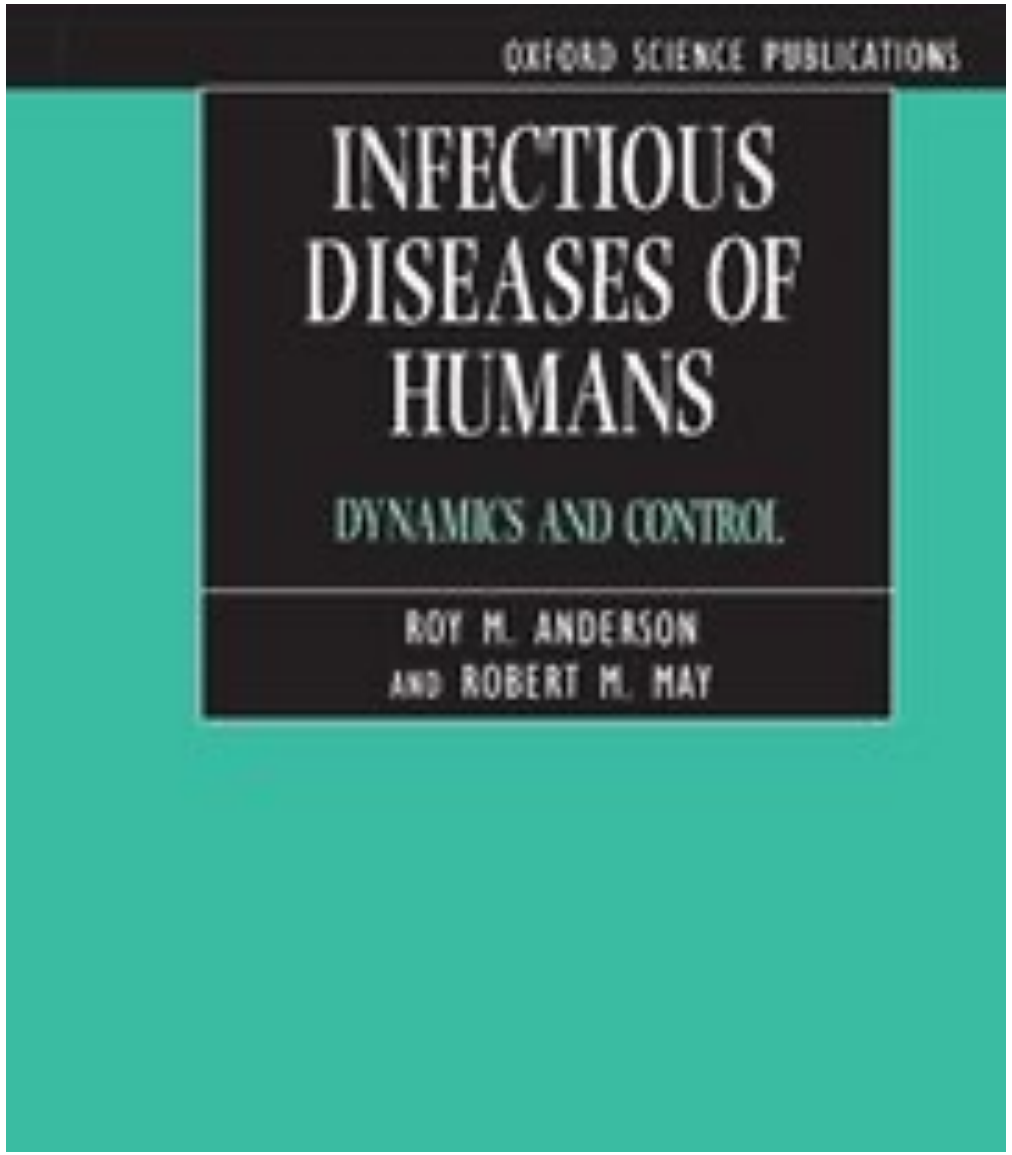
Models and Data through the ages:

What is responsible for disease?

1. **<u>Four Humors:</u>** Hippocrates (c. 400 BC) wrote that disease results from an imbalance of the four humors
2. **<u>Miasmatic Theory:</u>** Extension of Hippocrates that lasted through the 1800s – idea that disease was caused by bad air.
   - Popularized by Florence Nightengale
3. **<u>Germ Theory of Disease:</u>** Idea that disease results from germs
   - Leuwenhoek's microscope (1675)
   - Koch's postulates (1890)
4. **<u>Classic epidemiology</u>**
   - Risk factors for disease = John Snow (1854)

Models and Data through the ages:

1. **<u>Four Humors:</u>** Hippocrates (c. 400 BC) wrote that disease results from an imbalance of the four humors
2. **<u>Miasmatic Theory:</u>** Extension of Hippocrates that lasted through the 1800s – idea that disease was caused by bad air.
   - Popularized by Florence Nightengale
3. **<u>Germ Theory of Disease:</u>** Idea that disease results from germs
   - Leuwenhoek's microscope (1675)
   - Koch's postulates (1890)
4. **<u>Classic epidemiology</u>**
   - Risk factors for disease = John Snow (1854)
5. **<u>Population biology of infectious disease</u>**
   - Understanding the *process* of infectious disease *transmission.* = Kermack and McKendrick (1927); Anderson and May (1991).



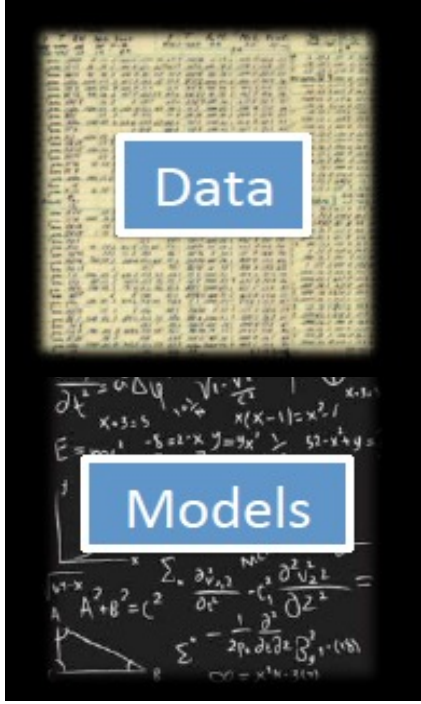
Epidemiology vs. Population Biology

- Epidemiology = “the study of **what** is on the people”
  - Coined by Spanish physician Villalba in 1802
- Emphasis on the study and analysis of the **distribution and determinants** of health and disease (“risk factors”)
  - Including chronic diseases!
- Often uses **cross-sectional** data to demonstrate associations of variables with outcome (disease)
- More statistical (y=mx+b)
  - **Pattern**

- Population biology = the study of **how** a disease spreads
  - Emphasis on the **interactions** of organisms with each other and the environment…when interactions result in disease
- Emphasis on understanding the **transmission dynamics** of infectious disease
- Typically involves fitting dynamical (population) models to **time series** data
- More mathematical (dN/dt)
  - **Process**

A Tool for Modeling

- Computer power keeps increasing
- Language/software
  - Fortran, C, C++
  - Julia, Java, Python
  - Matlab, Maple, Mathematica,
  - SAS, SPSS, Stata
- Specific programs
  - Vortex, RAMAS, NetLogo for IBM
  - NicheMapper for physiology, iLand for forest dynamics
  - MaxEnt for species distribution modeling
  - Zonation for reserve selection etc…
- The compromise: R---very powerful for
  - Visualization
  - Data formatting and sorting
  - Statistical analyses
  - Simulation (mechanistic model)



ICI3D International Clinics on Infectious Disease, Dynamics, & Data

MMED: ***Clinic on the Meaningful Modeling of Epidemiological Data***

May-June 2023, Cape Town, South Africa

DAIDD: ***Clinic on Dynamical Approaches to Infectious Disease Data***

December 2023, Virtual

South African Center for Epidemiological Modeling and Analysis (SACEMA), Director

Dr. Juliet Pulliam

University of Stellenbosch

www.ici3d.org

E^2M^2: Ecological and Epidemiological Modeling in Madagascar

www.E2M2.org

Photo: January 2023, Ranomafana National Park, Madagascar

Next Clinic: March 2024, Mantasoa, Madagascar