An introduction to demography

Ernesto F. L. Amaral

August 20–25, 2020
Population and Society (SOCI 312)

TEXAS A&M UNIVERSITY

www.ernestoamaral.com

Introduction

- Definition of demography
- Demographic equation
- Variables and observations
- COVID-19 pandemic
- Demographic models
- Cohorts and generations
- Lexis diagram
- Ratios, rates, probabilities

Definition of demography

- The scientific study of human population
- The term was coined by the Belgian statistician Achille Guillard in his 1855 book
  - *Éléments de Statistique Humaine ou Démographie Comparée*

Source: Poston, Bouvier, 2017.

Demography is destiny

- This phrase is attributed to the French mathematician and philosopher, Auguste Comte (1798–1857)
  - He is known as the “father of sociology”
  - Demography shapes the world, even if it does not determine it
  - Population change is an underlying component of almost everything happening in the world today, and therefore in the future as well

Source: Poston, Bouvier, 2017.

John Graunt (1620–1674)

- English statistician
  - Considered to be the founder of demography
  - Analyzed vital statistics of the London population
  - Studied the bills of mortality (weekly statistics of deaths) in early modern London
  - More specifically, studied death records that had been kept by London parishes since 1532

- Noticed certain regularities in death phenomena
  - Published in the book “Natural and Political Observations Made upon the Bills of Mortality” (1662)

Graunt's substantive contributions

- Recognized phenomenon of rural-urban migration
  - Urban death rate exceeded rural death rate
- Population was divided almost evenly by sex
  - Male birth rate was higher than female birth rate
    - Less females are born than males
  - Male death rate was higher than female death rate
    - Females live longer than males
- Presented mortality in terms of survivorship
  - He was the first to attempt to construct a life table...

Graunt's life table

Age	Number surviving	Age	Number surviving
0	100	46	10
6	64	56	6
16	40	66	3
26	25	76	1
36	16	86	0

Graunt's methodological contributions

- Paid attention to quality of data
- Exhibited a healthy skepticism
- Questioned the validity and reliability of data

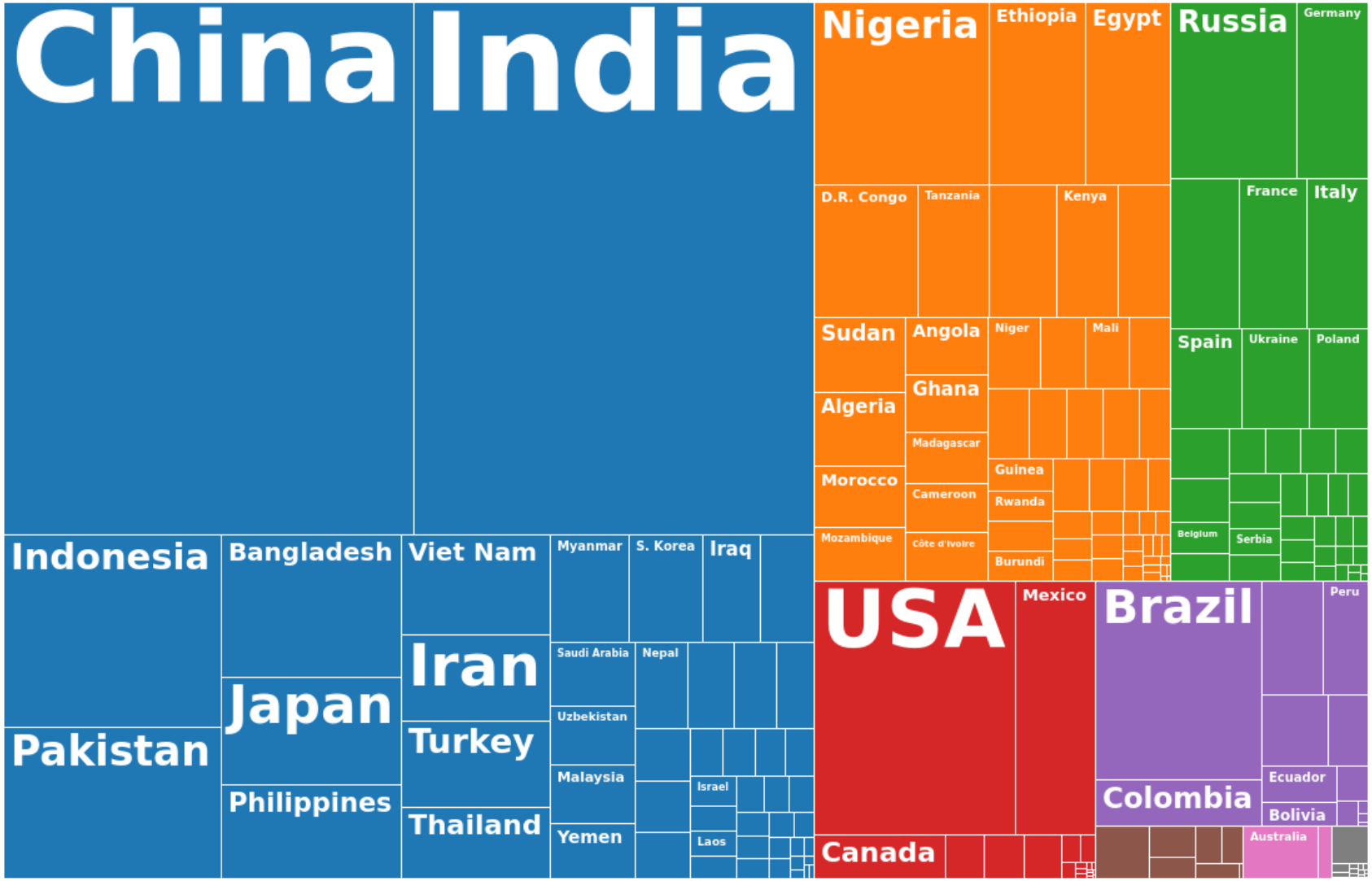
Poston's definition

- Demography is the scientific study of the size, composition, and spatial distribution of human populations
- It investigates changes in population size, composition, and distribution, resulting from fertility, mortality, and migration
- Demography helps understand what the past says about the future, given expected population changes

Source: Poston, Bouvier, 2017.

List of countries ordered by their population size

Total: 7,794,798,729 Year: 2020



PopulationPyramid.net

Concerns of demography

- Population size
- Population growth or decline
- Population processes/components
- Population distribution
- Population structure
- Population characteristics

Primary demographic questions

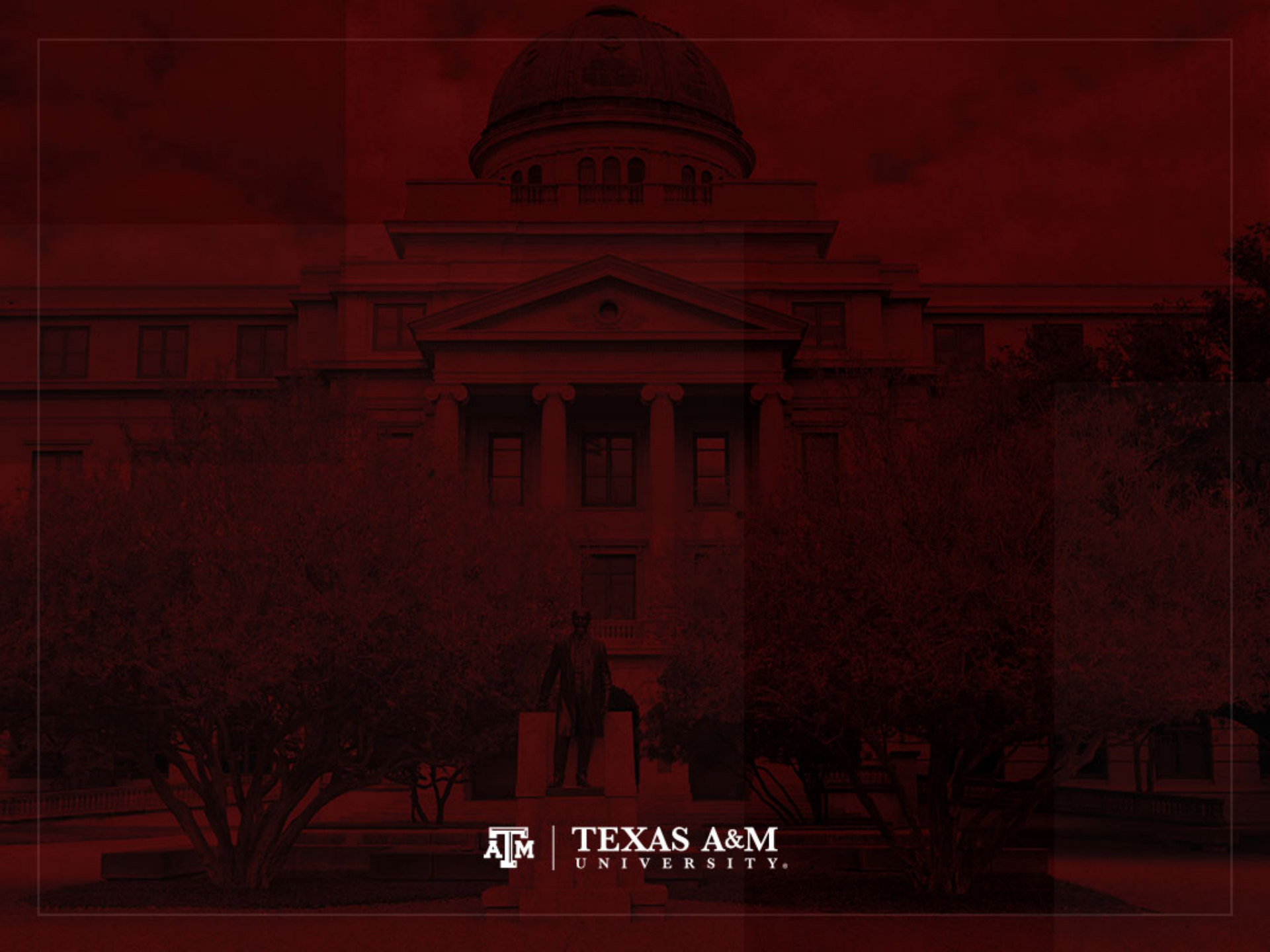
- How large (or small) is the population?
- How is the population composed, in terms of age, sex, race, marital status, and so forth?
  - What are the characteristics of the population?
- How is the population distributed spatially?
  - Populations are not randomly distributed in space
- How population changes happen over time?

Source: Poston, Bouvier, 2017.

Answers to these questions

- These demographic questions are answered in terms of the three demographic processes (components of demographic change)
  - Fertility
  - Mortality
  - Migration

Source: Poston, Bouvier, 2017.


TEXAS A&M
UNIVERSITY

Demographic equation

- Population size can change only through the processes of fertility, mortality, and migration
- Two ways of entering a population
  - Being born or moving into it
- Two ways of leaving a population
  - Dying or moving out of it
- Population can only change by way of a limited, countable number of events

Source: Poston, Bouvier, 2017.

Basic demographic equation

$$P_{t+1} = P_t + B_{t \text{ to } t+1} - D_{t \text{ to } t+1} + I_{t \text{ to } t+1} - E_{t \text{ to } t+1}$$

- P_{t+1}: population at time $t+1$
- P_t: population at time t
- $B_{t \text{ to } t+1}$: births between times t and $t+1$
- $D_{t \text{ to } t+1}$: deaths between times t and $t+1$
- $I_{t \text{ to } t+1}$: immigrants (or in-migrants) to the population between times t and $t+1$
- $E_{t \text{ to } t+1}$: emigrants (or out-migrants) from the population between times t and $t+1$

Source: Poston, Bouvier, 2017.

Components of equation

- $P_{t+1} = P_t + B_{t \text{ to } t+1} - D_{t \text{ to } t+1} + I_{t \text{ to } t+1} - E_{t \text{ to } t+1}$
- Natural increase: $B_{t \text{ to } t+1} > D_{t \text{ to } t+1}$
- Natural decrease: $B_{t \text{ to } t+1} < D_{t \text{ to } t+1}$
  - Negative natural increase

Source: Poston, Bouvier, 2017.

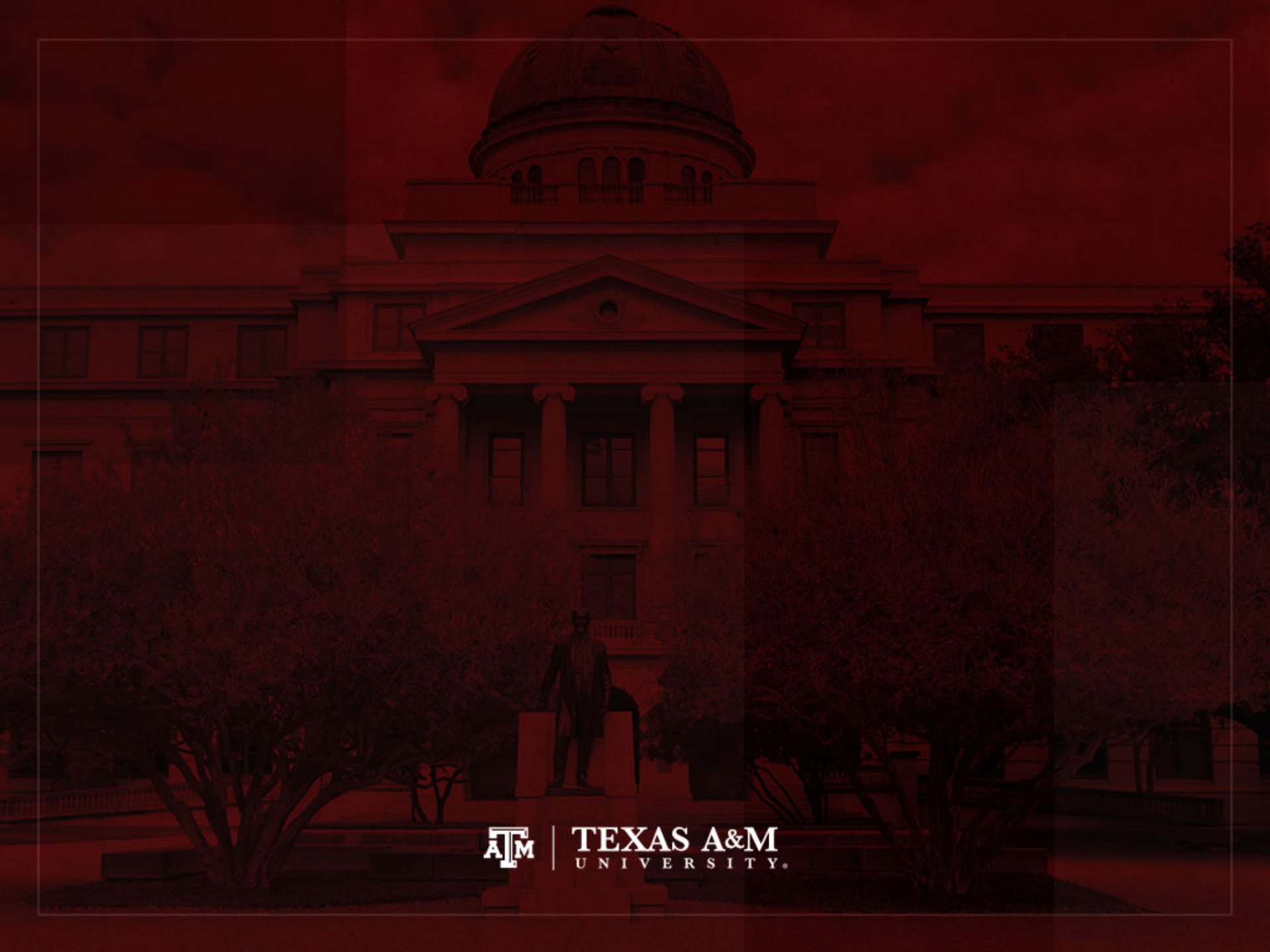
Migration components of equation

- $I_{t \text{ to } t+1} - E_{t \text{ to } t+1}$
  - Net international migration
    - Immigration minus emigration
  - Net internal migration
    - In-migration minus out-migration

- $I_{t \text{ to } t+1} < E_{t \text{ to } t+1}$
  - Negative net international migration (sending countries)
  - Negative net internal migration (net out-migration)

- $I_{t \text{ to } t+1} > E_{t \text{ to } t+1}$
  - Positive net international migration (receiving countries)
  - Positive net internal migration (net in-migration)

Source: Poston, Bouvier, 2017.


TEXAS A&M
UNIVERSITY

Variables and observations

- **Variables**
  - Characteristics that can change values from case to case
  - E.g. gender, age, income, political party affiliation...

- **Observations (cases)**
  - Refer to the entity from which data are collected
  - Also known as ”unit of analysis”
  - E.g. individuals, households, states, countries...

Source: Healey, 2015.

Variables

- **Variable:** a characteristic/phenomenon whose value varies (changes) from case to case, and is empirically quantifiable

- **Dependent variable:** a variable whose variation depends on another variable

- **Independent variable:** a variable whose variation produces ("causes") variation in another variable

Source: Poston, Bouvier, 2017.

Causation

- Theories and hypotheses are often stated in terms of the relationships between variables
  - Causes: independent variables
  - Effects or results: dependent variables

y	x	Use
Dependent variable	Independent variable	Econometrics
Explained variable	Explanatory variable	
Response variable	Control variable	Experimental science
Predicted variable	Predictor variable	
Outcome variable	Covariate	
Regressand	Regressor	

Source: Wooldridge, 2015.

Observations

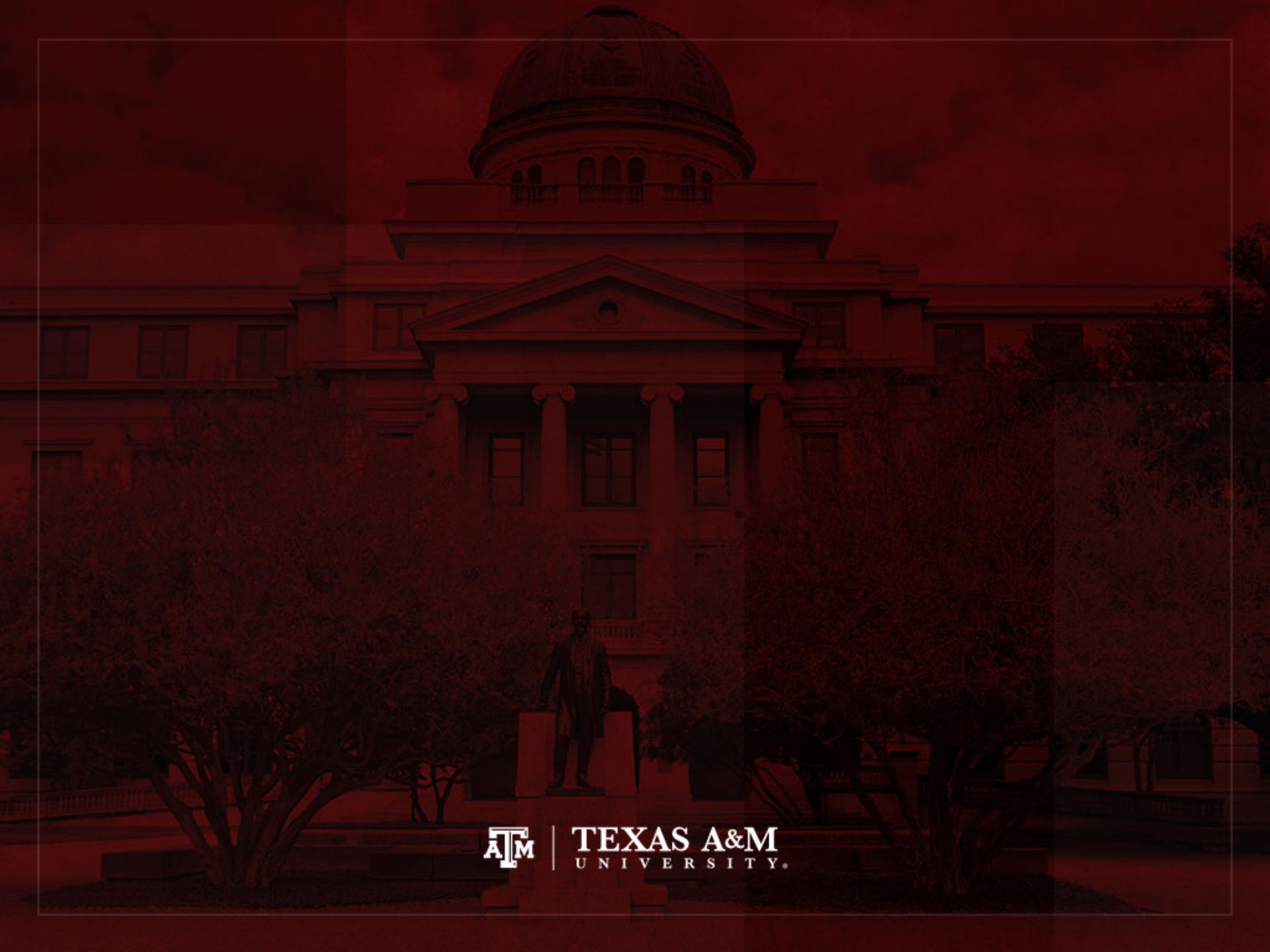
- Observations (cases) are collected information used to test hypotheses
- Decide how variables will be measured and how cases will be selected and tested
- Measure social reality: collect numerical data
- Information can be organized in databases
  - Variables as columns
  - Observations as rows

Source: Healey, 2015.

Example of a database

Observation	Salary per hour	Years of schooling	Years of experience in the labor market	Female	Marital status (married)
1	3.10	11	2	1	0
2	3.24	12	22	1	1
3	3.00	11	2	0	0
4	6.00	8	44	0	1
5	5.30	12	7	0	1
...	...	...	...	...	...
525	11.56	16	5	0	1
526	3.50	14	5	1	0

Source: Wooldridge, 2015.

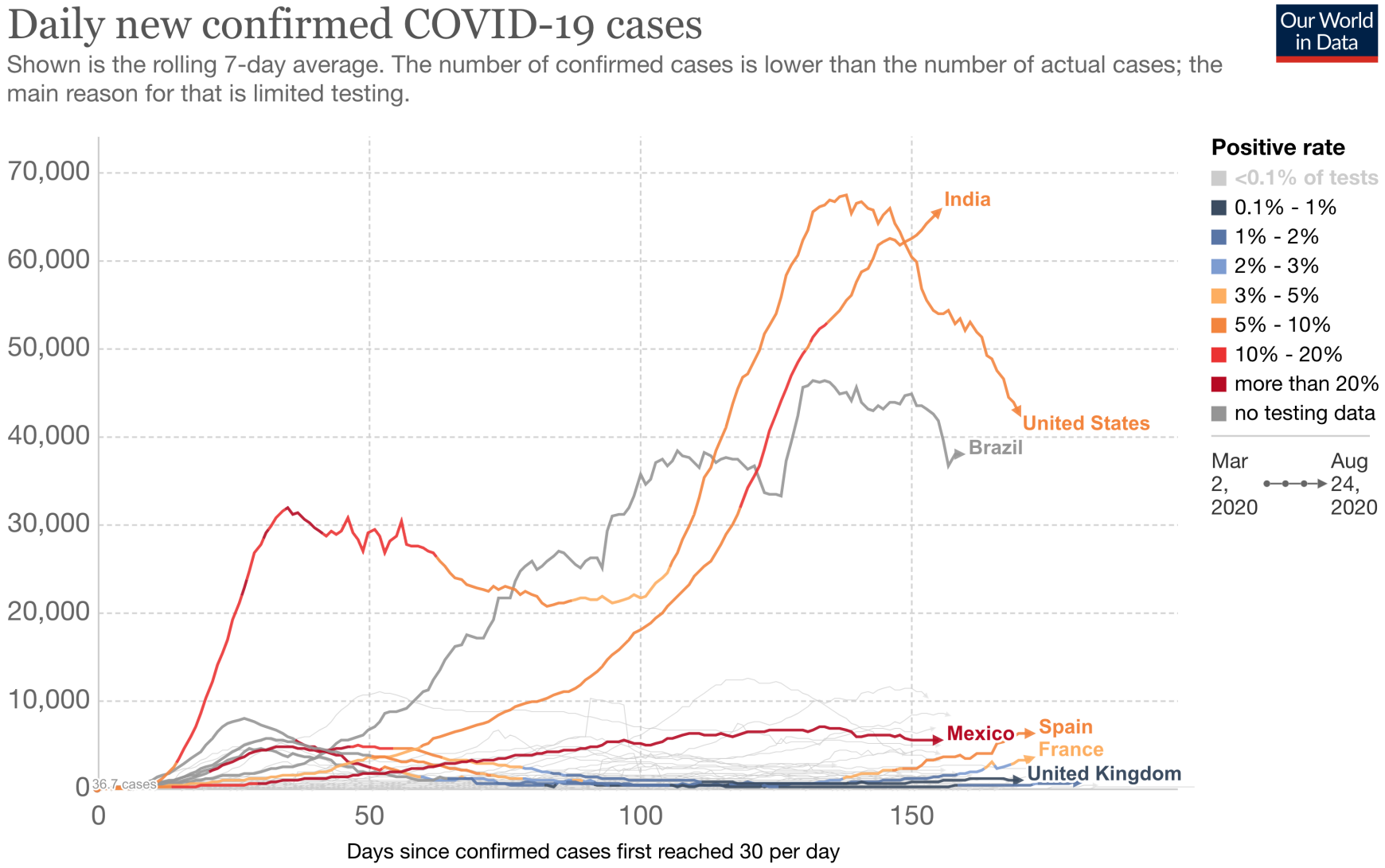

TEXAS A&M
UNIVERSITY

COVID-19 pandemic, August 24, 2020

#	Country, Other	Total Cases	New Cases	Total Deaths	New Deaths	Total Recovered	Active Cases	Serious, Critical	Tot Cases/ 1M pop	Deaths/ 1M pop	Total Tests	Tests/ 1M pop	Population
	World	23,809,061	+6,189	817,005	+431	16,358,235	6,633,821	61,715	3,054	104.8			
1	USA	5,915,630		181,114		3,217,981	2,516,535	16,483	17,856	547	76,883,479	232,071	331,293,410
2	Brazil	3,627,217		115,451		2,778,709	733,057	8,318	17,046	543	14,144,344	66,473	212,784,888
3	Mexico	563,705	+3,541	60,800	+320	389,124	113,781	3,346	4,365	471	1,263,835	9,787	129,132,739
4	India	3,164,881		58,546		2,403,101	703,234	8,944	2,290	42	35,902,137	25,978	1,382,011,722
5	UK	326,614		41,433		N/A	N/A	72	4,807	610	15,177,265	223,394	67,939,531
6	Italy	260,298		35,441		205,662	19,195	65	4,306	586	8,053,551	133,231	60,448,212
7	France	244,854		30,528		85,199	129,127	399	3,750	468	6,000,000	91,890	65,295,389
8	Spain	420,809		28,872		N/A	N/A	658	9,000	617	8,517,446	182,162	46,757,536
9	Peru	600,438		27,813		407,301	165,324	1,525	18,174	842	3,006,993	91,014	33,038,913
10	Iran	361,150		20,776		311,365	29,009	3,848	4,292	247	3,062,422	36,392	84,150,494
11	Colombia	551,696		17,612		384,171	149,913	1,493	10,825	346	2,508,972	49,231	50,962,919
12	Russia	961,493		16,448		773,095	171,950	2,300	6,588	113	34,600,000	237,077	145,943,991
13	South Africa	611,450		13,159		516,494	81,797	539	10,291	221	3,564,065	59,983	59,418,339
14	Chile	399,568		10,916		372,464	16,188	1,014	20,875	570	2,231,463	116,583	19,140,575
15	Belgium	82,092	+156	9,996	+4	18,242	53,854	89	7,079	862	2,144,563	184,921	11,597,214
16	Germany	236,117		9,336		209,600	17,181	245	2,817	111	10,197,366	121,652	83,824,401
17	Canada	125,647		9,083		111,694	4,870	62	3,325	240	5,169,166	136,782	37,791,278
18	Argentina	350,867		7,366		256,789	86,712	1,960	7,753	163	1,105,878	24,435	45,257,261
19	Indonesia	155,412		6,759		111,060	37,593		567	25	2,056,166	7,506	273,950,524
20	Iraq	207,985		6,519		150,389	51,077	661	5,154	162	1,457,665	36,125	40,350,522

Source: https://www.worldometers.info/coronavirus/.

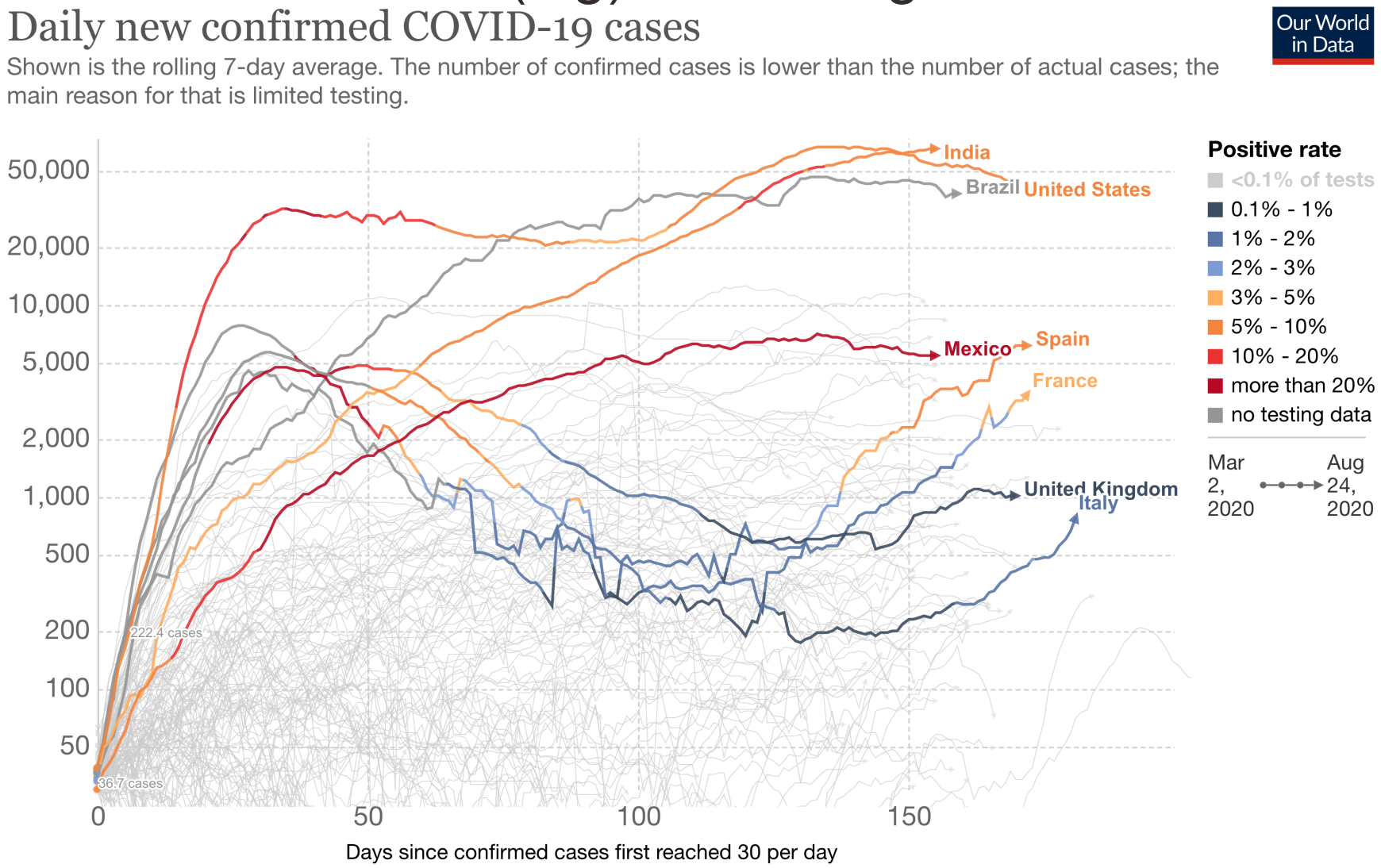
New cases (linear)



Note: Eight countries with more deaths (United States, Brazil, Mexico, India, United Kingdom, Italy, France, Spain).

Source: https://ourworldindata.org/coronavirus.

New cases (log): Flattening the curve

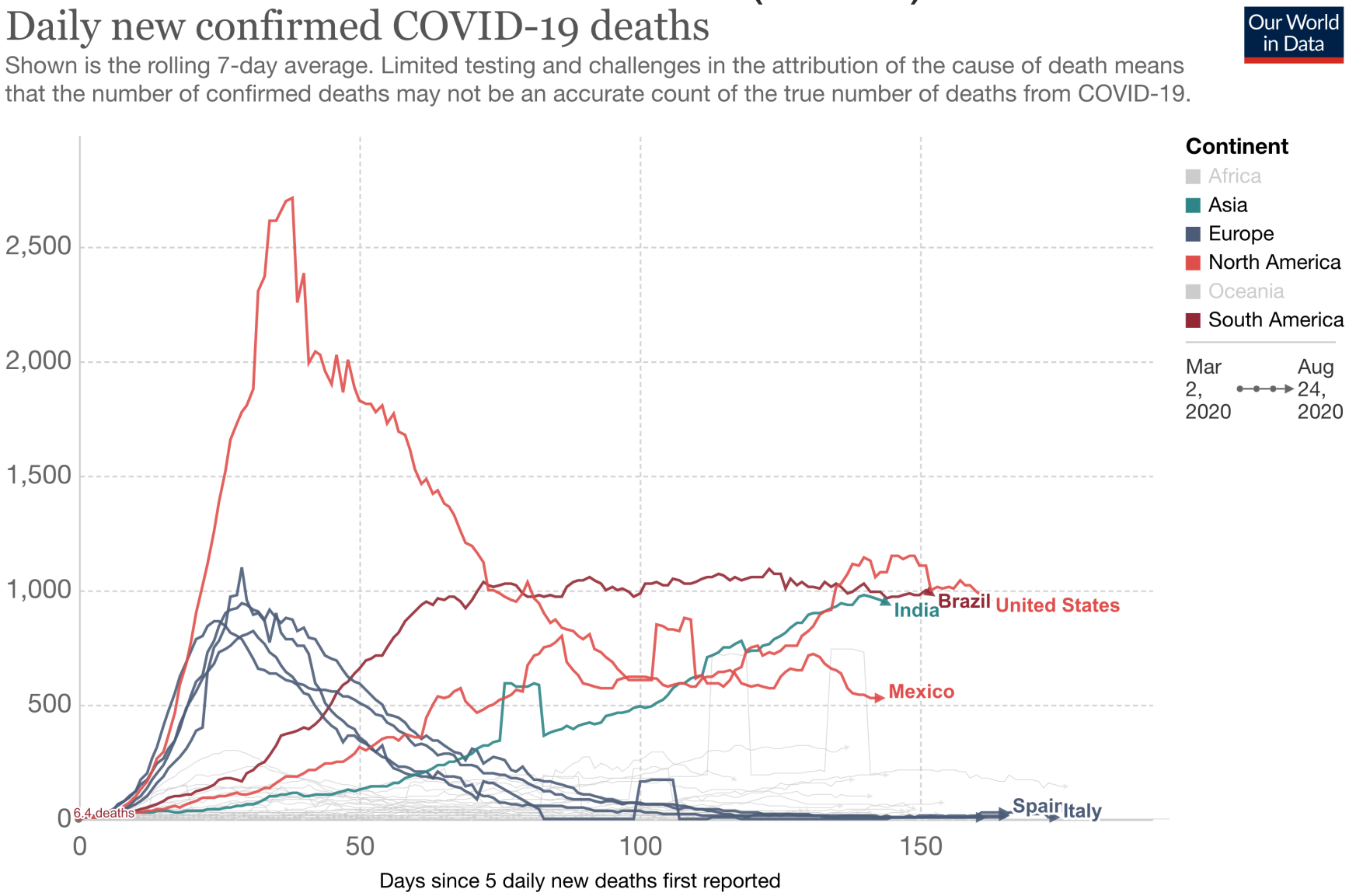


Source: European CDC – Situation Update Worldwide – Last updated 24 August, 10:04 (London time), Official data collated by Our World in Data
CC BY

Note: Eight countries with more deaths (United States, Brazil, Mexico, India, United Kingdom, Italy, France, Spain).

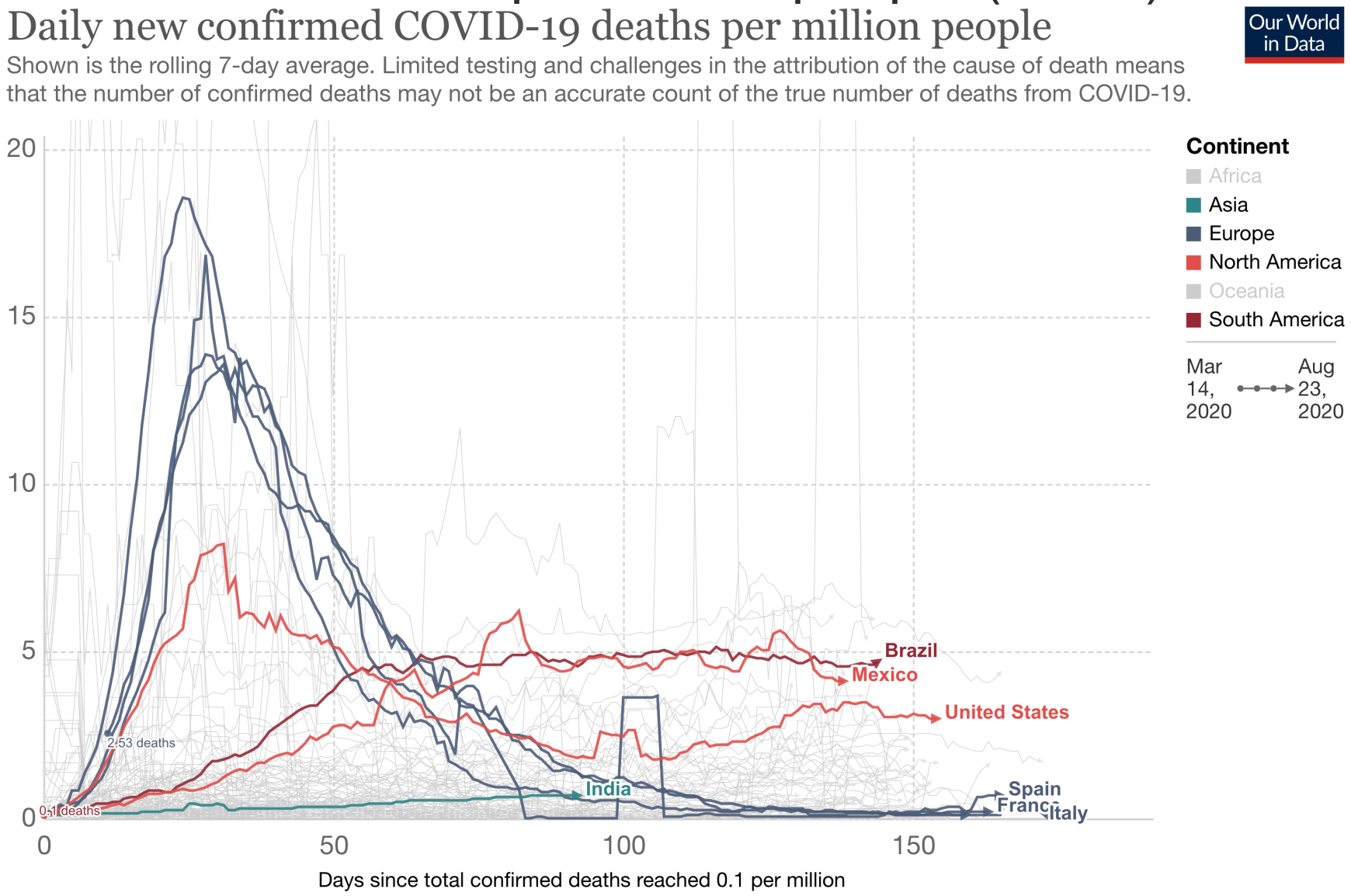
Source: https://ourworldindata.org/coronavirus.

New deaths (linear)



Source: https://ourworldindata.org/coronavirus.

New deaths per million people (linear)

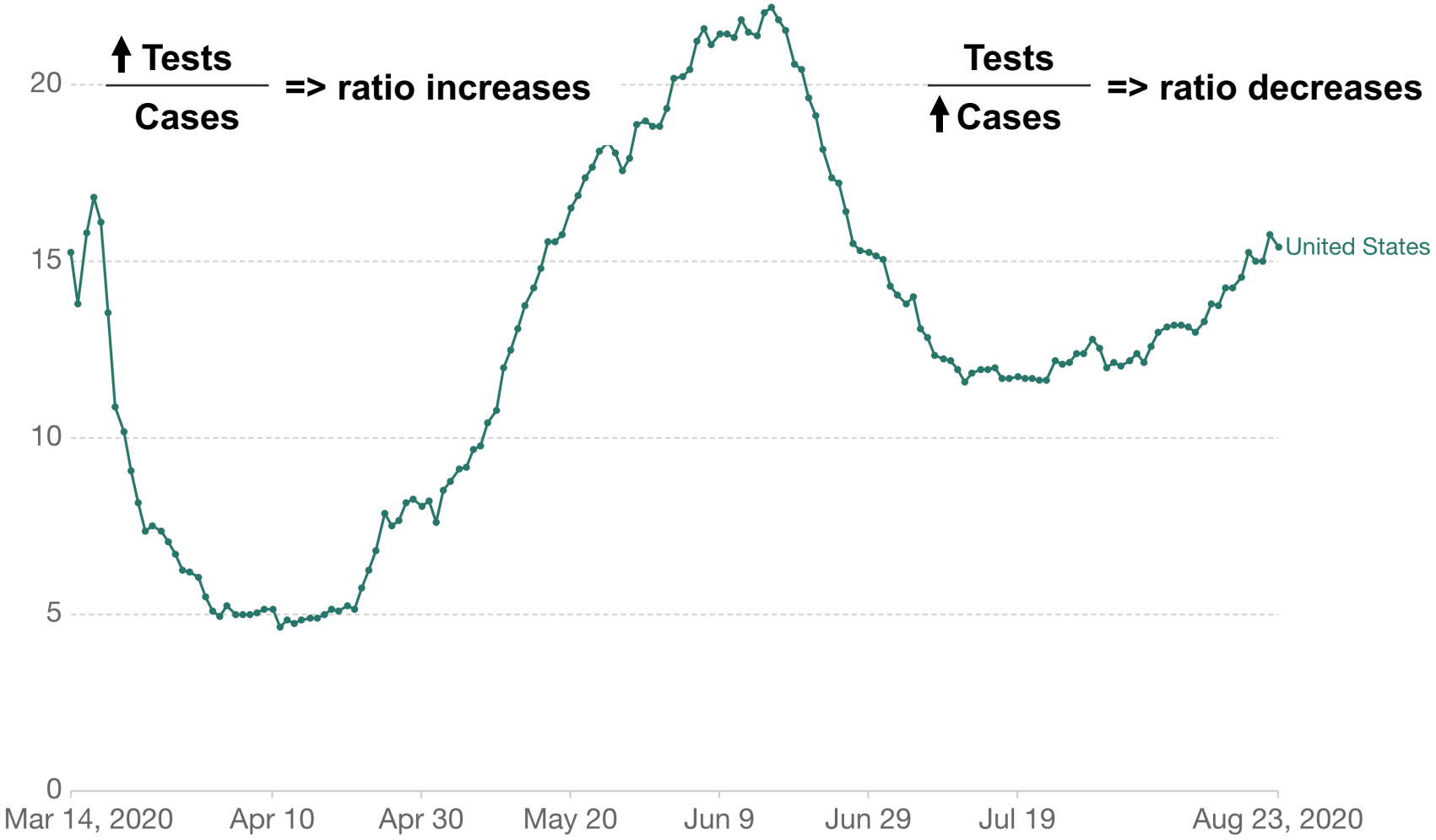


Note: Eight countries with more deaths (United States, Brazil, Mexico, India, United Kingdom, Italy, France, Spain).

Source: https://ourworldindata.org/coronavirus.

Tests conducted per confirmed case of COVID-19

Shown is the rolling 7-day average. The number of tests divided by the number of confirmed cases. Not all countries report testing data on a daily basis.

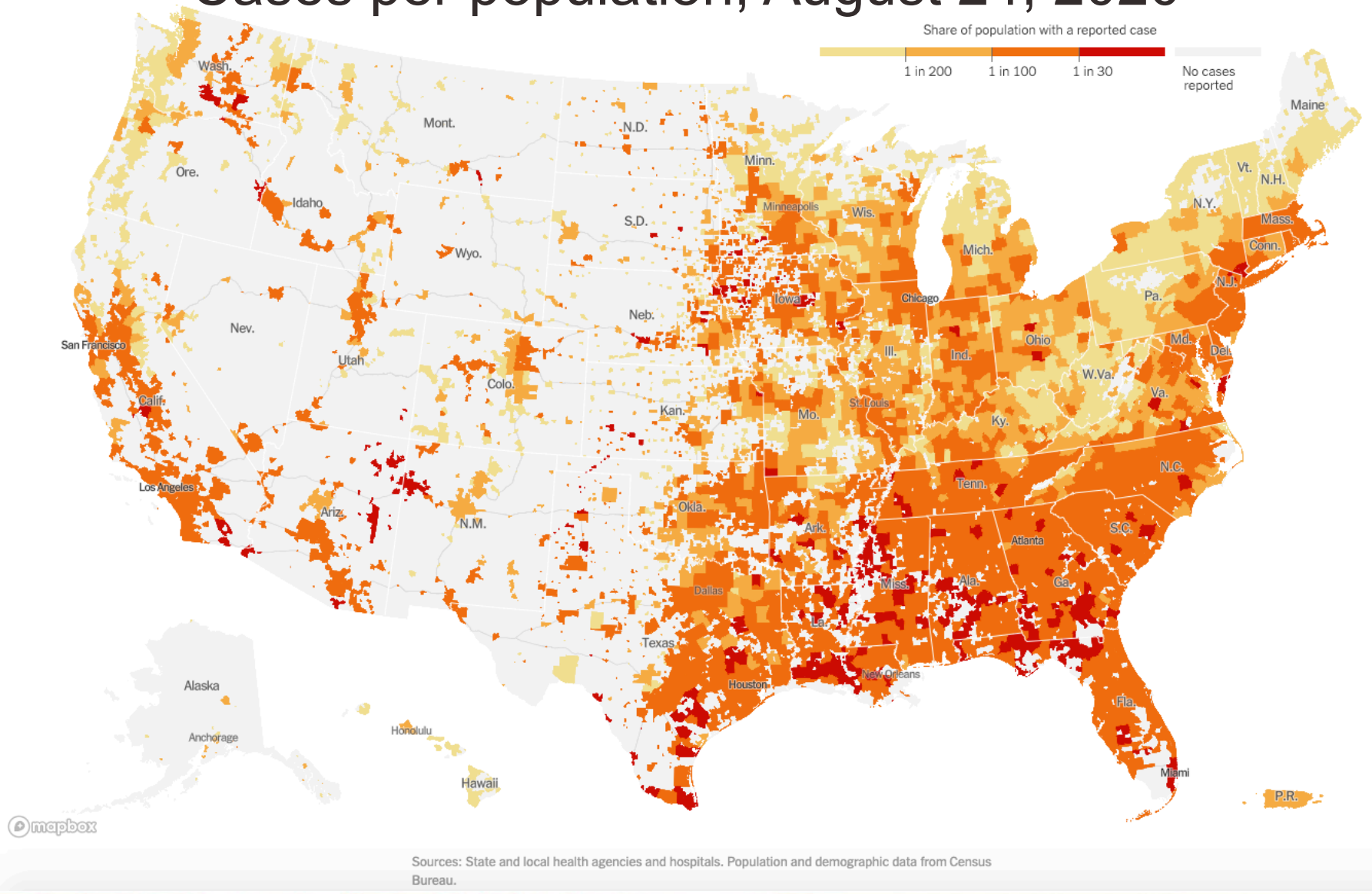


Source: Official data collated by Our World in Data

CC BY

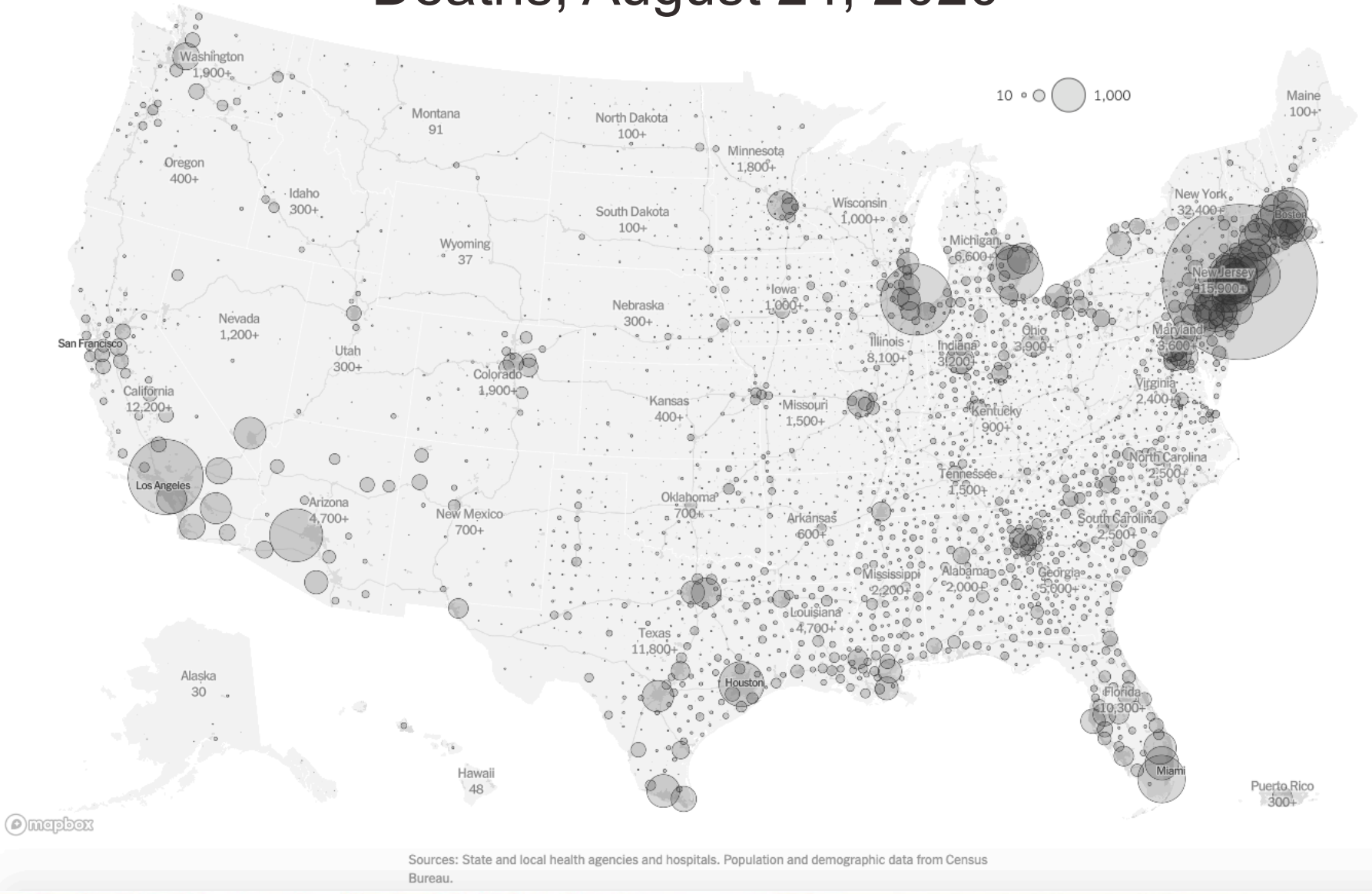
Source: https://ourworldindata.org/coronavirus.

Coronavirus in the United States, Cases per population, August 24, 2020



Source: https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html#states.

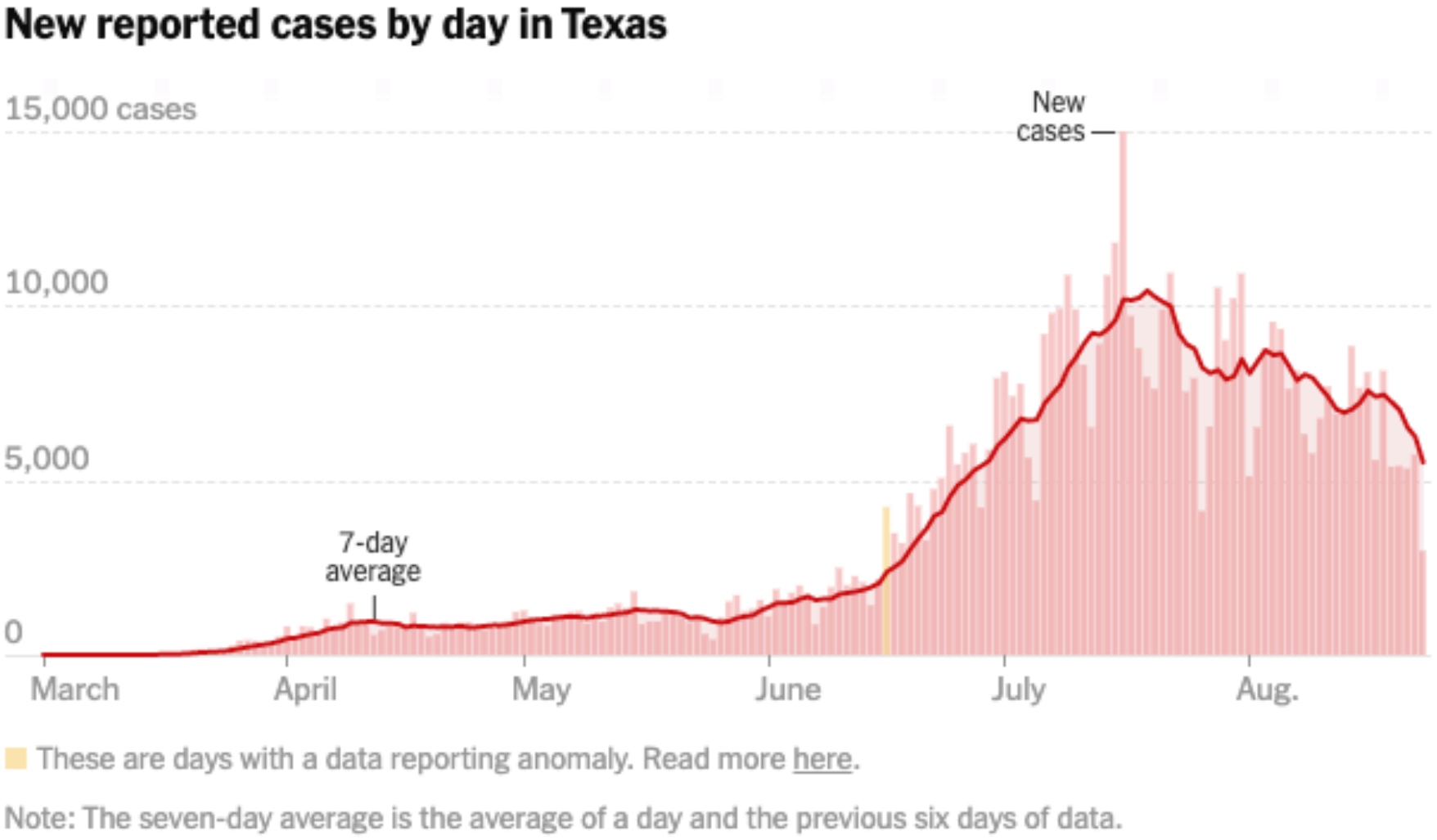
Coronavirus in the United States, Deaths, August 24, 2020



Sources: State and local health agencies and hospitals. Population and demographic data from Census Bureau.

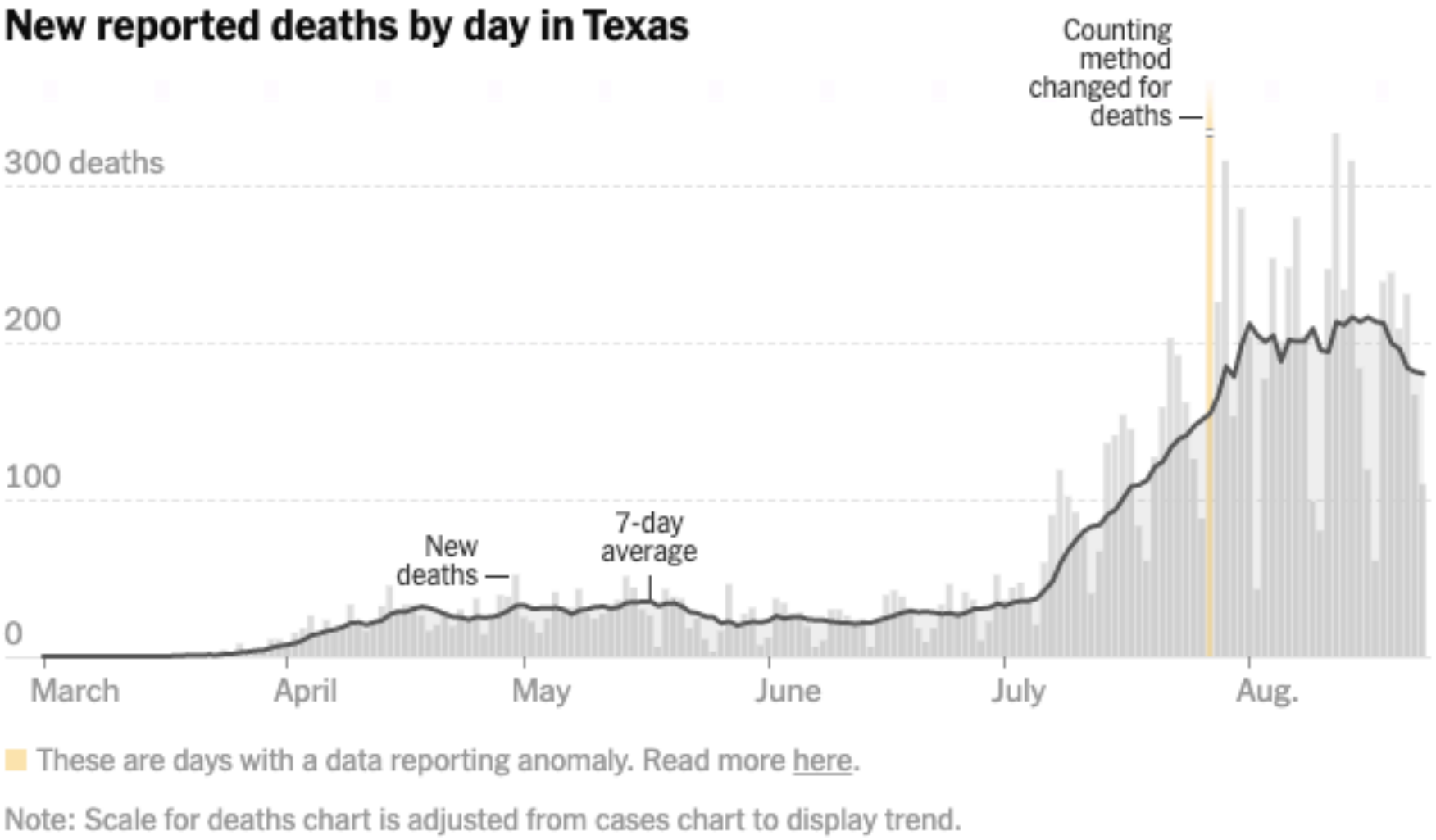
Source: https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html#states.

Coronavirus in Texas, Cases, August 24, 2020



Source: https://www.nytimes.com/interactive/2020/us/texas-coronavirus-cases.html.

Coronavirus in Texas, Deaths, August 24, 2020

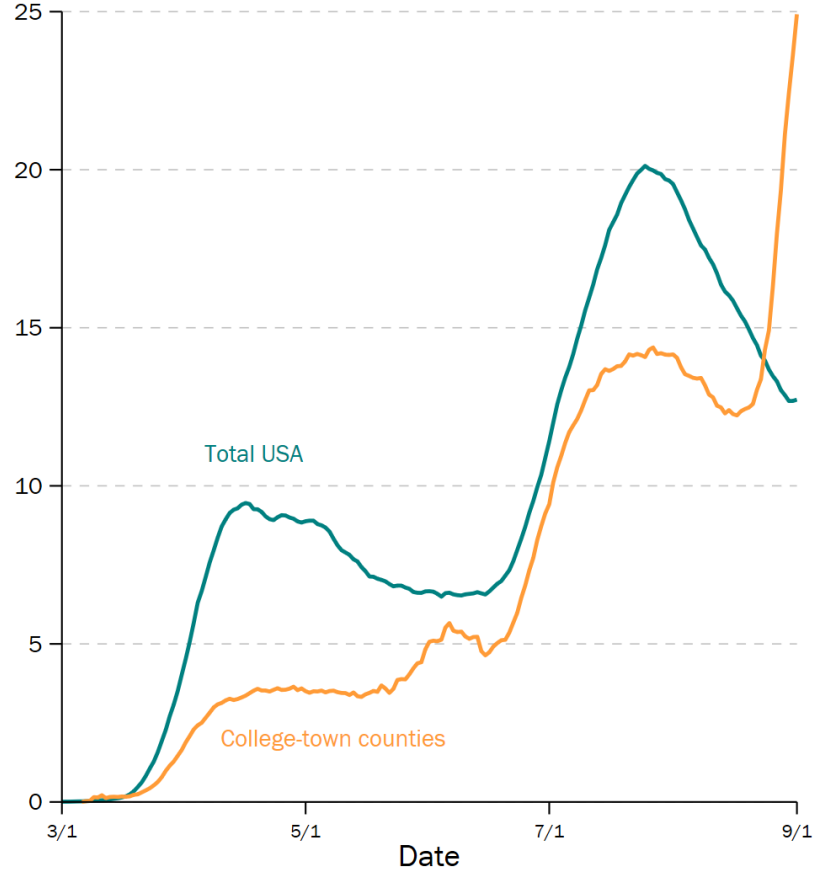


Source: https://www.nytimes.com/interactive/2020/us/texas-coronavirus-cases.html.

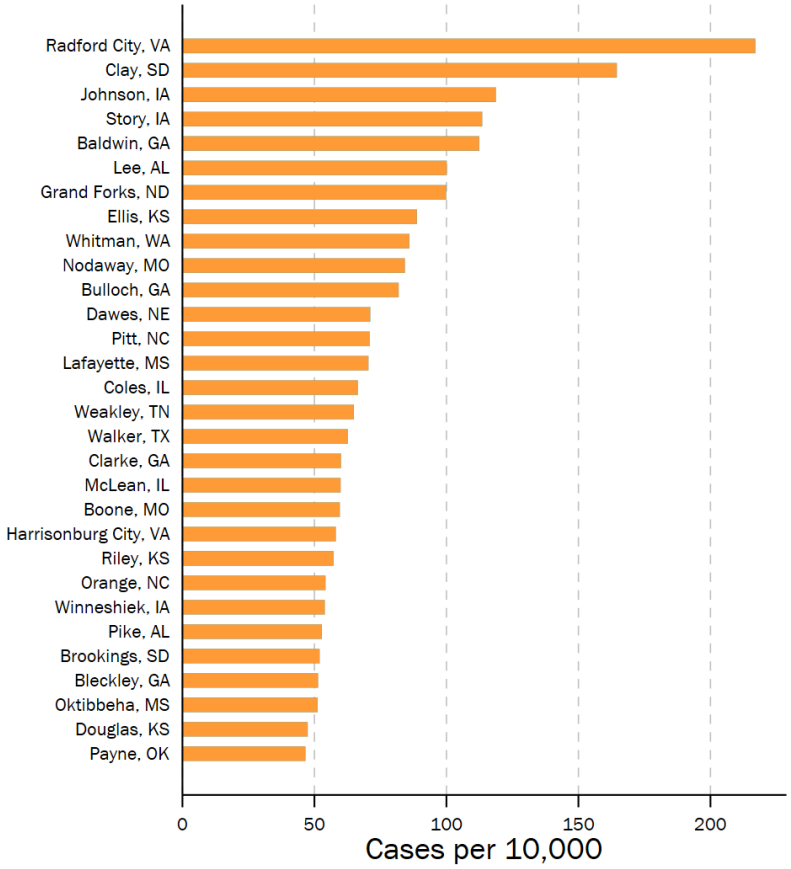
COVID-19 cases, September 02, 2020

COVID-19 cases in college-town counties

Daily cases per 100,000, two-week average

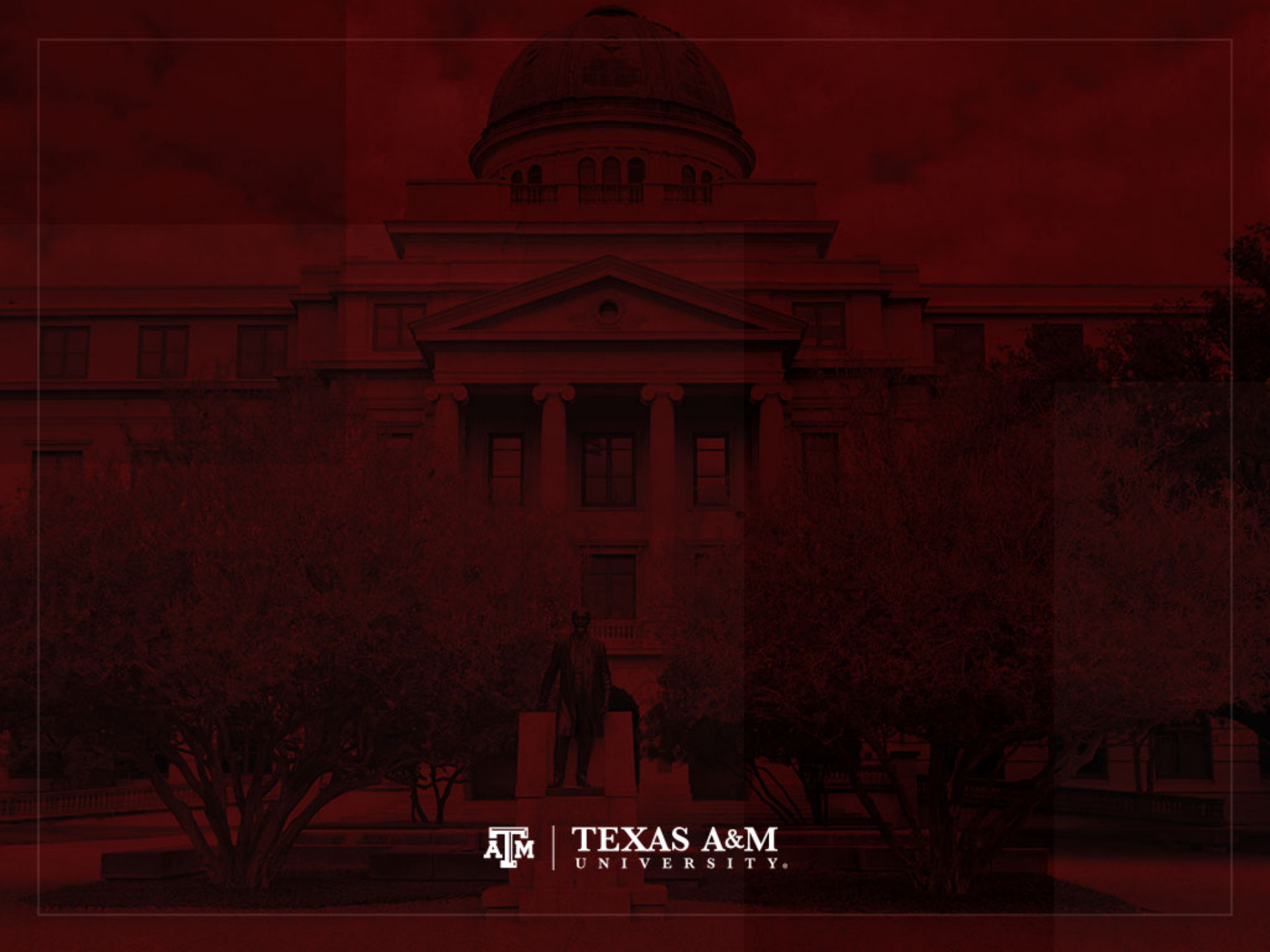


Top 30: Cases per 10,000 in the last two weeks



College-town counties (N=94) have 13% or more of population age>=3 enrolled as undergraduates.
PN Cohen analysis of New York Times and US Census Bureau data.

Source: Professor Philip N. Cohen, Department of Sociology, University of Maryland.


TEXAS A&M
UNIVERSITY

Demographic models

- Formal demography
- Population studies I
- Population studies II

Source: Poston, Bouvier, 2017.

Formal demography

Independent variable		**Dependent variable**
Demographic	→	Demographic

Examples

1. Age composition	→	Birth rate
2. Birth rate	→	Age composition
3. Sex composition of in-migrants to a city	→	Sex ratio of the total population of the city

Source: Poston, Bouvier, 2017.

Population studies I (social demography)

Independent variable		Dependent variable
Non-demographic	→	Demographic

Examples

1. Social class (sociological)	→	Death rate
2. Attitude about motherhood (social psychology)	→	Number of children
3. Annual rainfall (geographical)	→	Population density
4. Economic opportunity (economic)	→	Migration

Source: Poston, Bouvier, 2017.

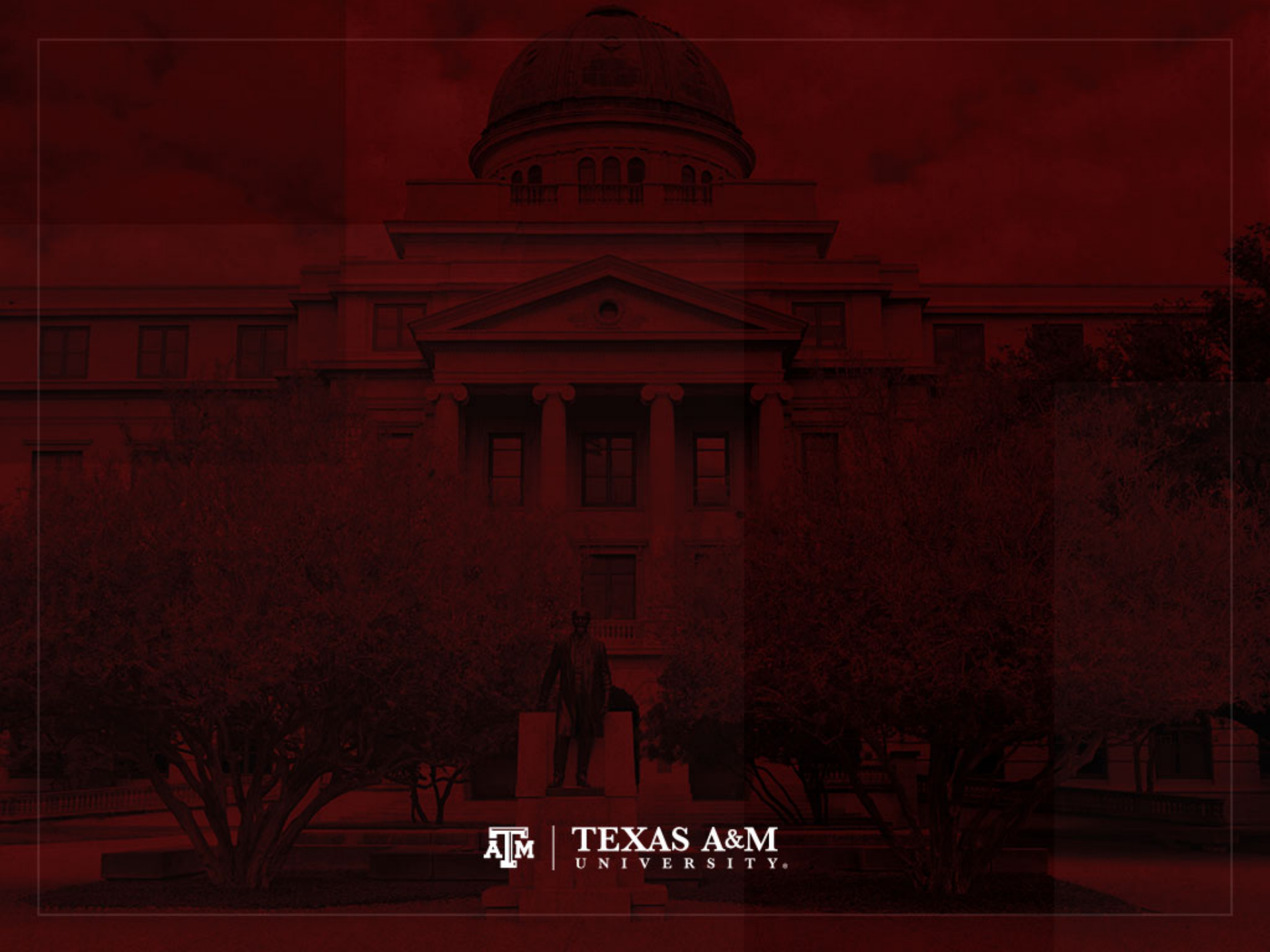
Population studies II (social demography)

Independent variable		**Dependent variable**
Demographic	→	Non-demographic

Examples

1. Age composition	→	Voting behavior (political)
2. Migration	→	Social change (sociology)
3. Birth rate	→	Need for infant & child goods/services (public health)

Source: Poston, Bouvier, 2017.


TEXAS A&M
UNIVERSITY

Cohorts and generations

- Cohort
  - Group of persons who have experienced a common event during a given time interval
  - Birth cohorts are sometimes referred to as generations

- Why study birth cohorts?
  - If you understand what distinctive opportunities and problems you have faced, you can find common ground with others in your generation and in other generations (Elwood Carlson)

Examples of cohorts

- People born during the same period who experience similar social circumstances throughout their lives
  - Lucky Few: from around 1929 through 1945
  - Baby boomers: between around 1946 and 1964
  - Baby bust cohort (Generation X): from mid-1960s to early 1980s
  - Millennials (New boomers or Generation Y): from early 1980s to early 2000s
  - Generation Z: start in early 2000s

Lucky Few cohort

- **Lucky Few cohort**, born between 1929–1945
  - They were fewer compared to the much larger number of persons in the following cohort
  - Baby Boomer cohort, born between 1946–1964

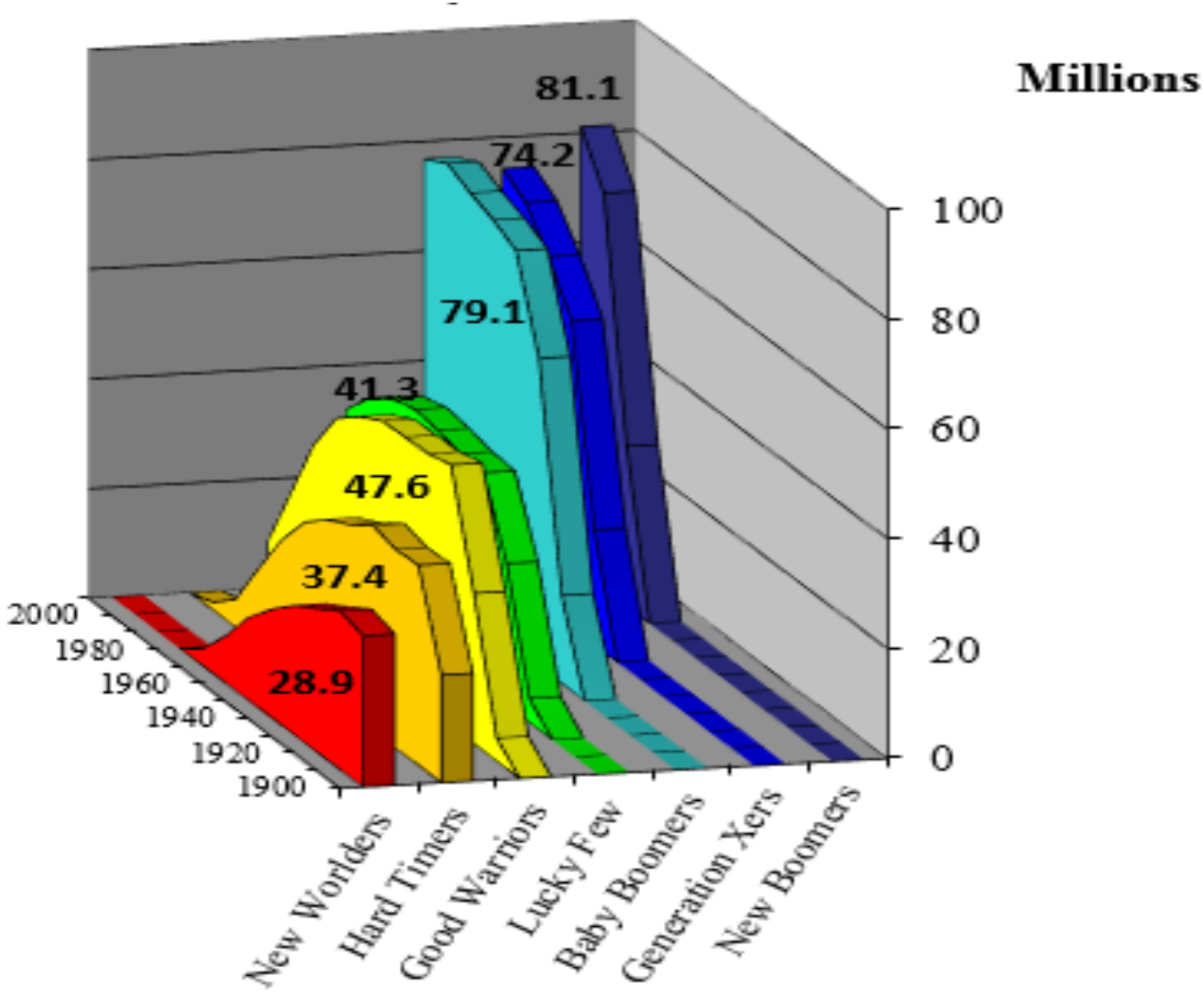
- The smaller size of the Lucky Few has enabled them to experience
  - Higher employment rates
  - Greater variety of social opportunities than members in the preceding or following cohorts

Eight US birth cohorts

Birth cohort	Years of birth	Age range in 2020	Number born, total	Alive in 2019	Number born, per year
New Worlders	1871–1889	None living	~ 30 million	None	1.6 million
Hard Timers	1890–1908	None living	~ 25 million	None	1.3 million
Good Warriors	1909–1928	92–111	57.6 million	1.7 million	2.8 million
Lucky Few	1929–1945	75–91	44.1 million	20.9 million	2.5 million
Baby Boomers	1946–1964	56–74	75.8 million	69.9 million	4 million
Generation X	1965–1982	38–55	62.2 million	73.9 million	3.4 million
Millennials	1983–2001	19–37	74.5 million	84.9 million	3.9 million
Generation Z	2002–present	0–18	72.4 million	77.3 million	4 million

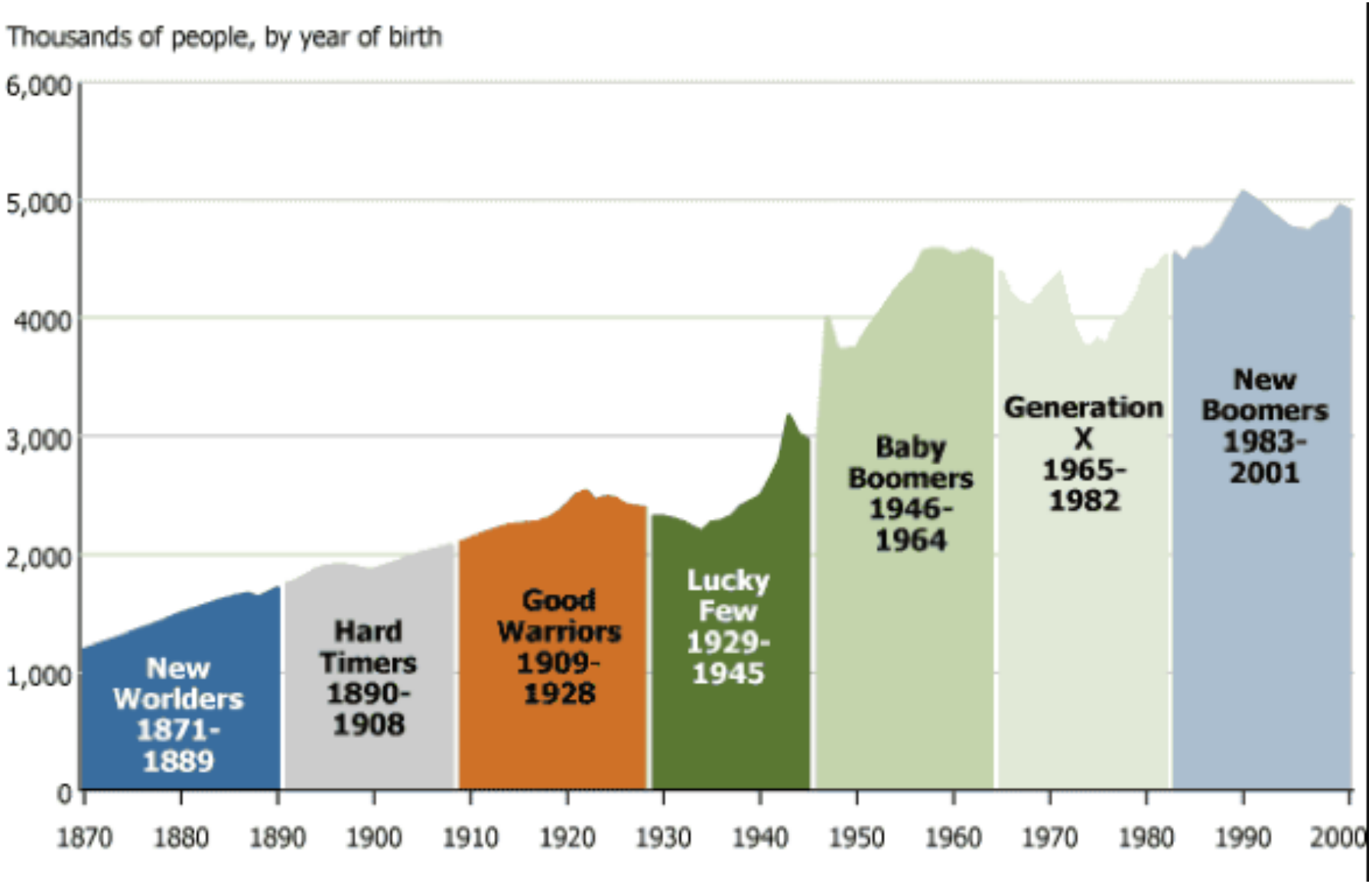
Source: Poston D. 2020. "I'm a 'Lucky Few': How About You?" Life @ The Dominion, September, p.61.

Seven US birth cohorts by size, 1900–2010

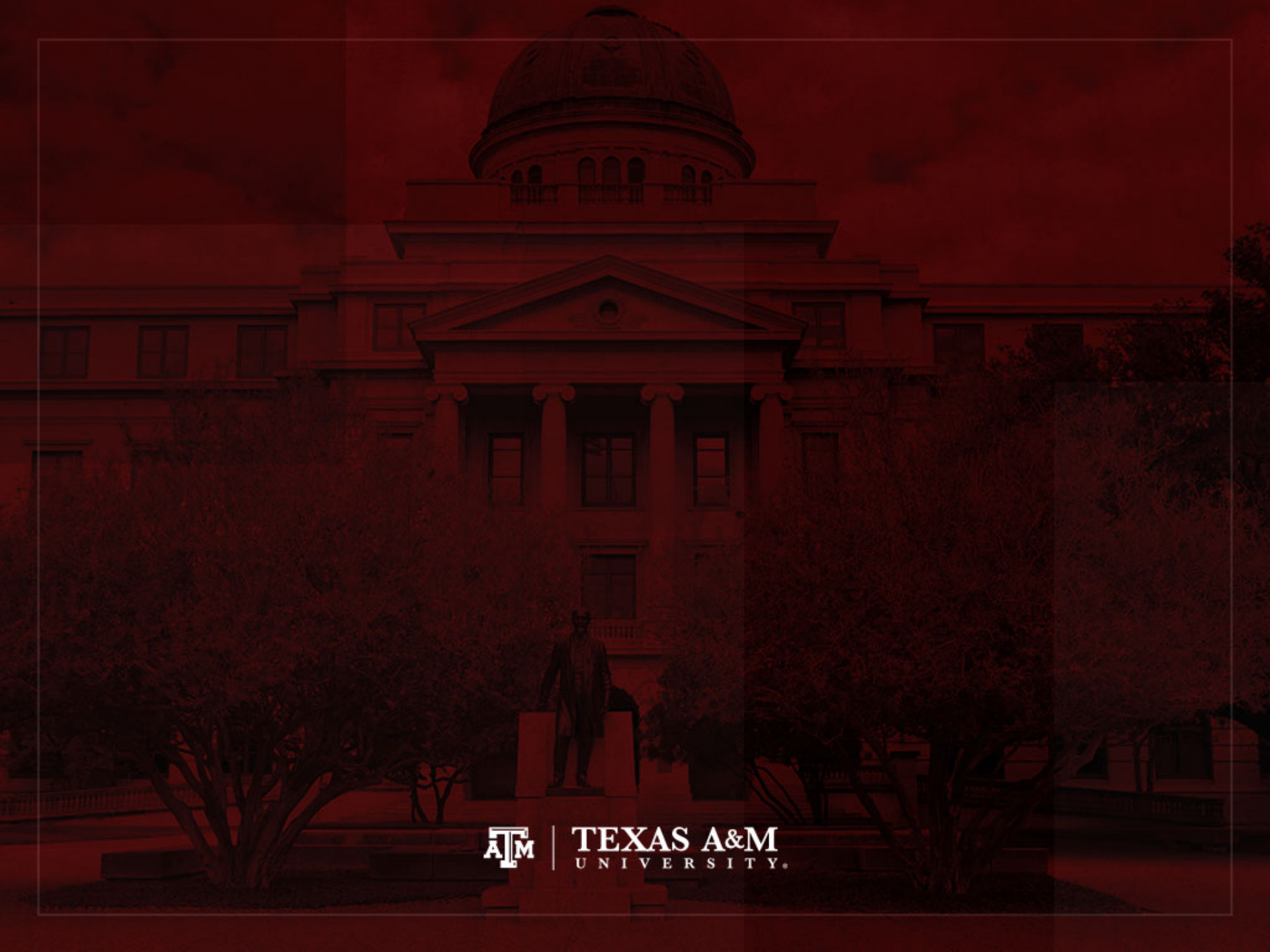


Source: Poston, Bouvier (2017).

US birth cohorts



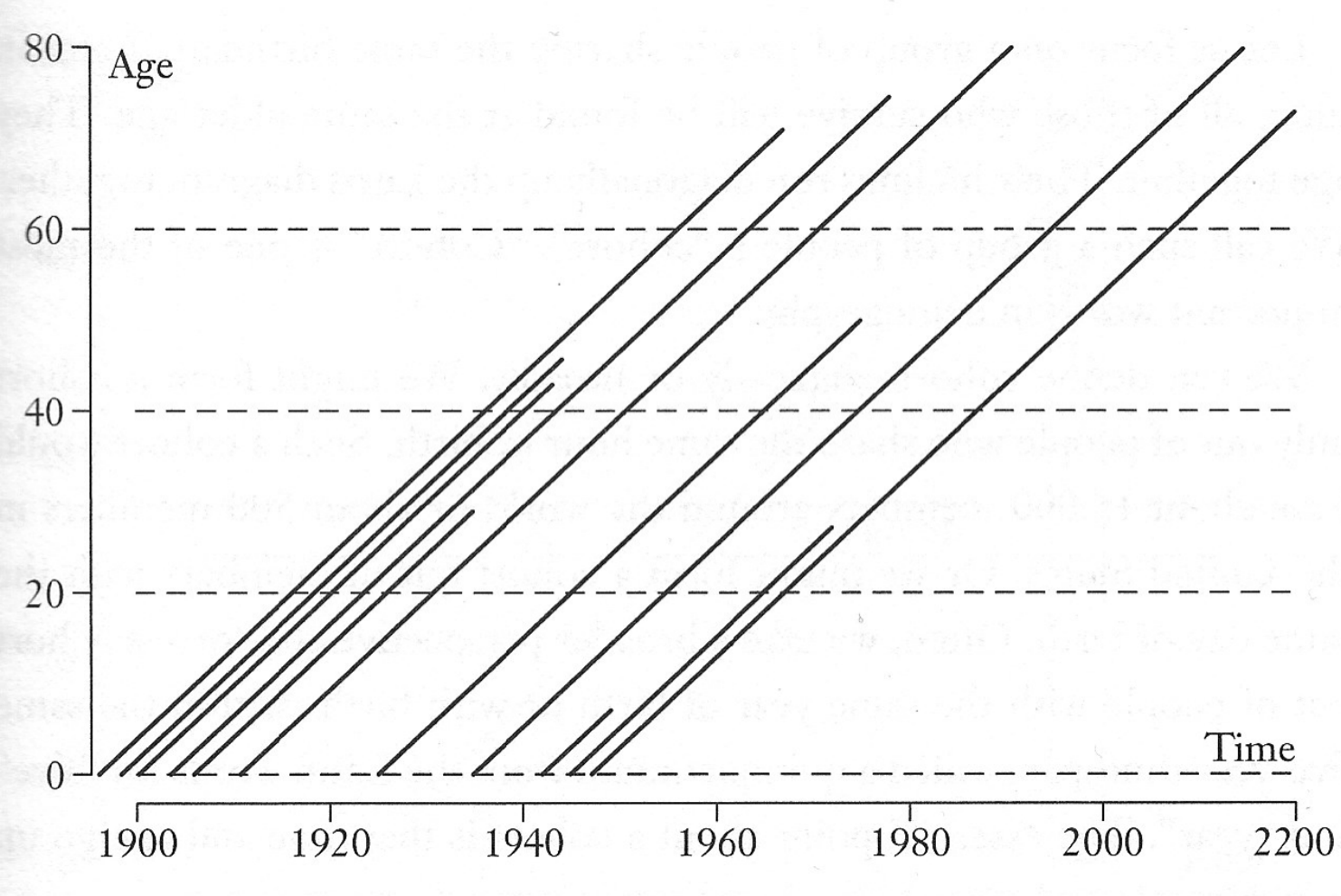
Source: Professor Dudley Poston.


TEXAS A&M
UNIVERSITY

Lexis diagram

- Lexis diagram provides relationships between chronological time *t* (horizontal) and age *x* (vertical)
- Each person has a lifeline on a Lexis diagram
  - Starting at (t_b, 0), where t_b is the person's birthdate and 0 is the person's age at birth
- Line goes up to the right with a slope equal to 1
  - People age one year in one calendar year
- Lifeline goes up until time and age of the person's death

Lexis diagram



Source: Wachter 2014, p. 31.

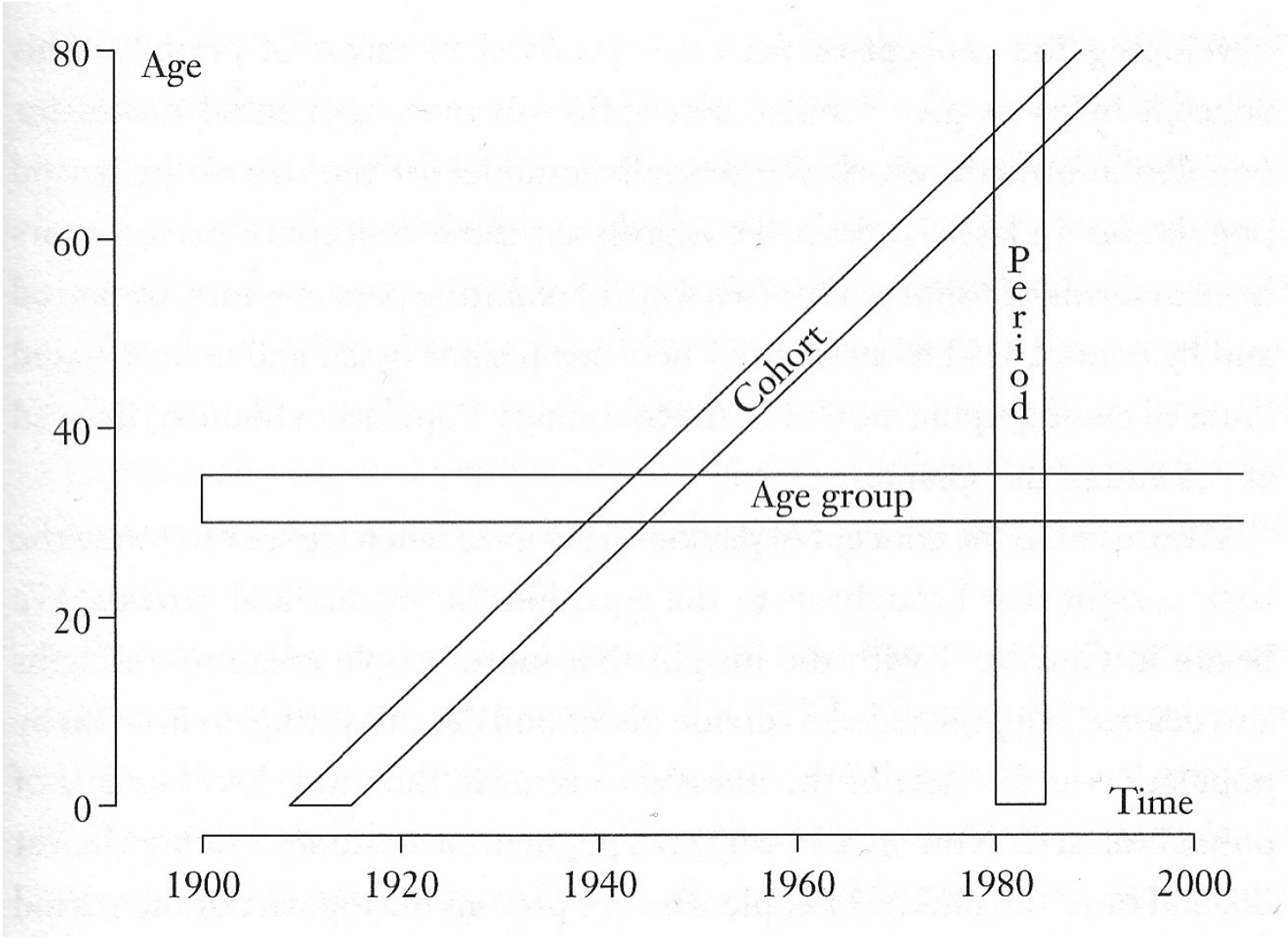
Exploring Lexis diagram

- To find population size
  - Draw vertical line upward from the time point
  - Count how many lifelines cross vertical line
- To find how many people survive to some age
  - Draw horizontal line across at the height corresponding to that age
  - Count how many lifelines cross that horizontal line
- Immigrants start at age and time of immigration

Cohort in the Lexis diagram

- Group of people sharing the same birthdate
- Group of individuals followed simultaneously through time and age
- Their lifelines run diagonally up the Lexis diagram together
- In a cohort, time and age go up together
- A cohort shares experiences

Lexis diagram: Age, period, cohort



Source: Wachter 2014, p.33.

Game of pretend

- When we calculate a period measure, we pretend that age-specific rates we see today for different age groups continue unchanged into the future

- We are creating an imaginary cohort whose life experience is pieced together from the experiences of different people found at different ages in one period of time

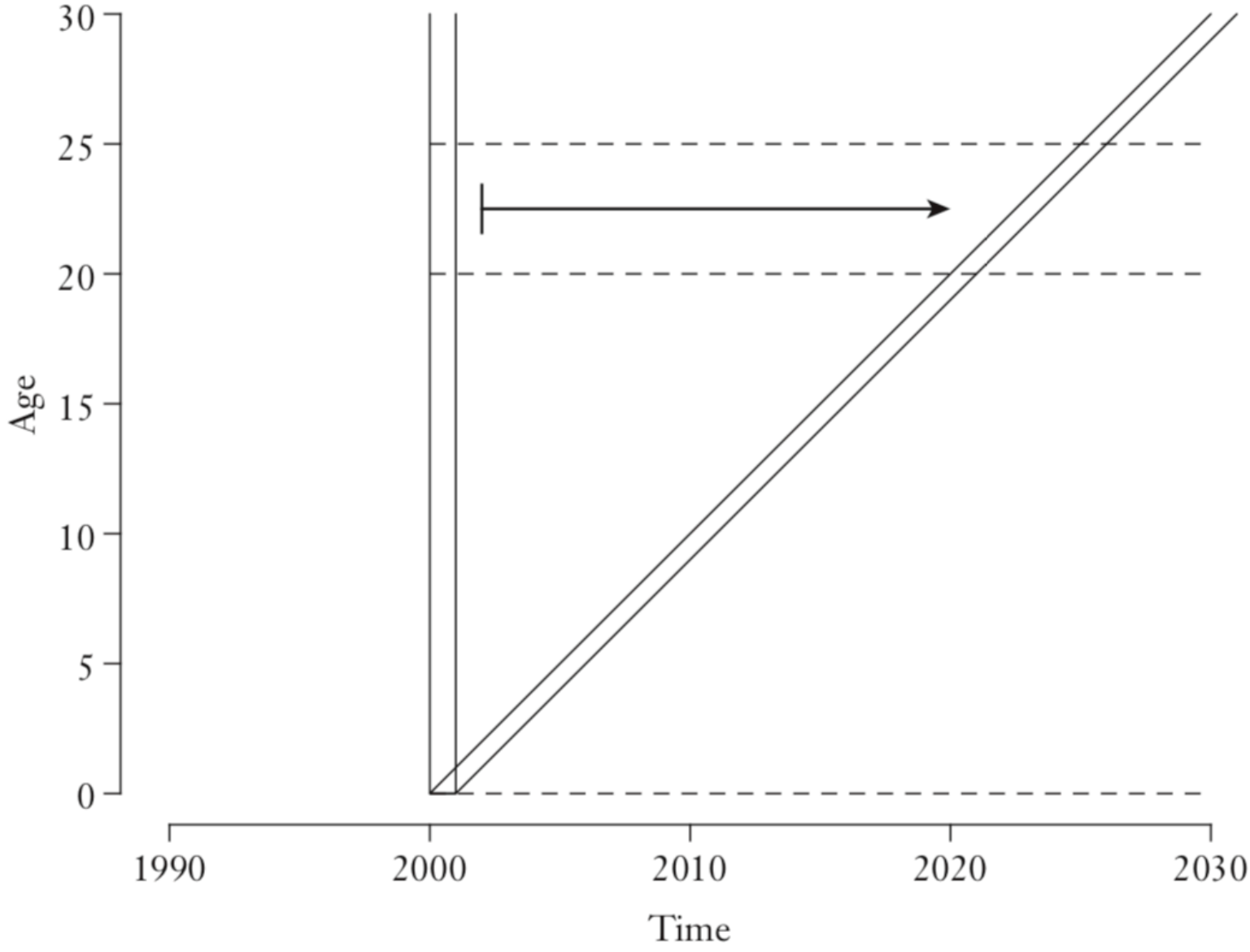
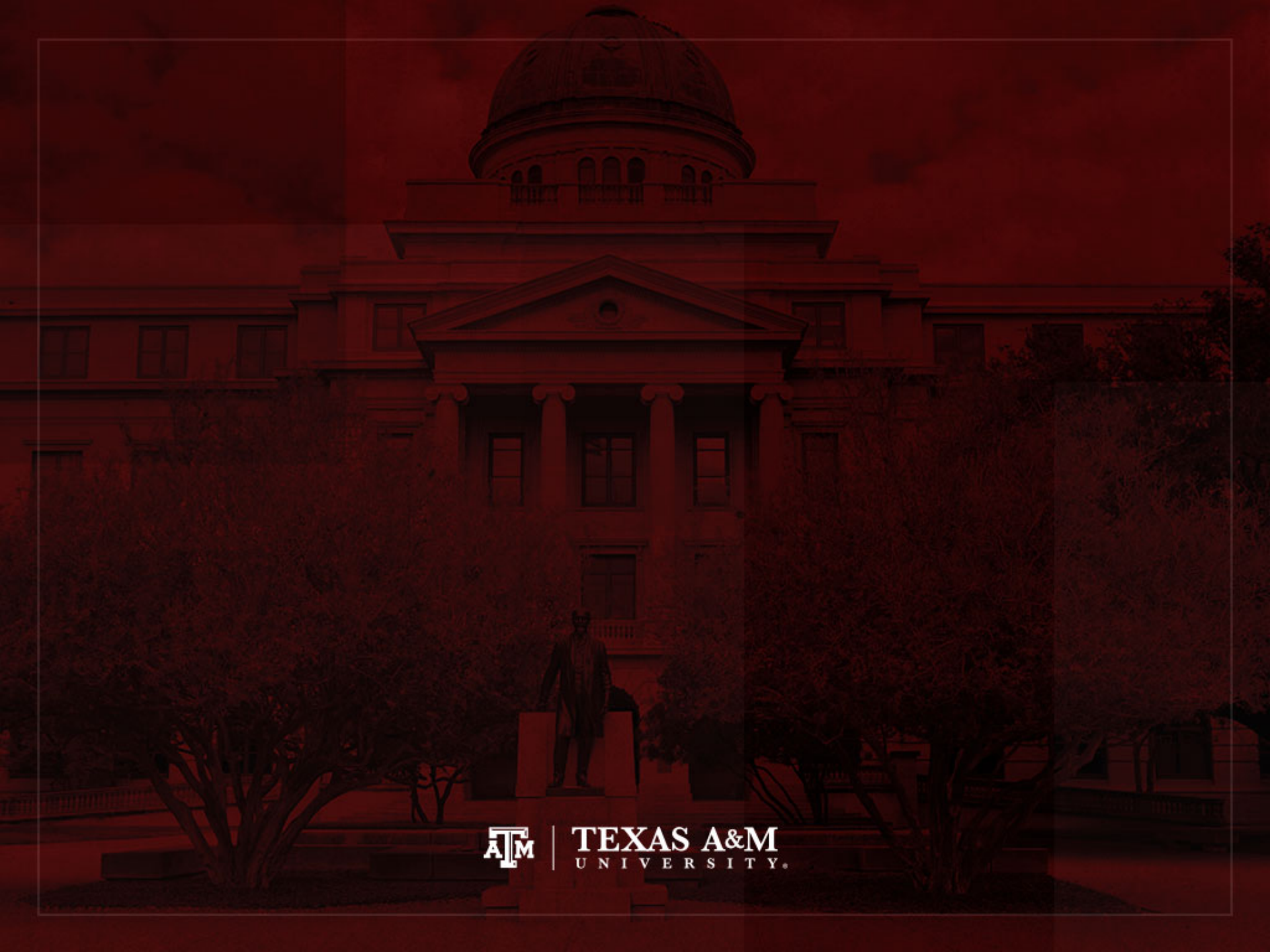


Figure 6.1 From period to cohort on a Lexis diagram

Source: Wachter 2014, p.127.

Synthetic cohort

- We call this imaginary cohort the synthetic cohort
  - *syn*: “together”
  - *thetic*: “pieced”
  - *synthetic*: “pieced together”
- Age-specific cohort rates of the synthetic cohort are the age-specific period rates of the period population
- The concept of a synthetic cohort is central to demography


TEXAS A&M
UNIVERSITY

Ratios, rates, probabilities

- Ratios
  - Compare the size of one group to the size of another group
- Rates
  - Describe the number of occurrences of an event for a given number of individuals who had the chance to experience that event per unit of time
- Probabilities
  - Divides the number of events by the total number of people at risk in the relevant time frame

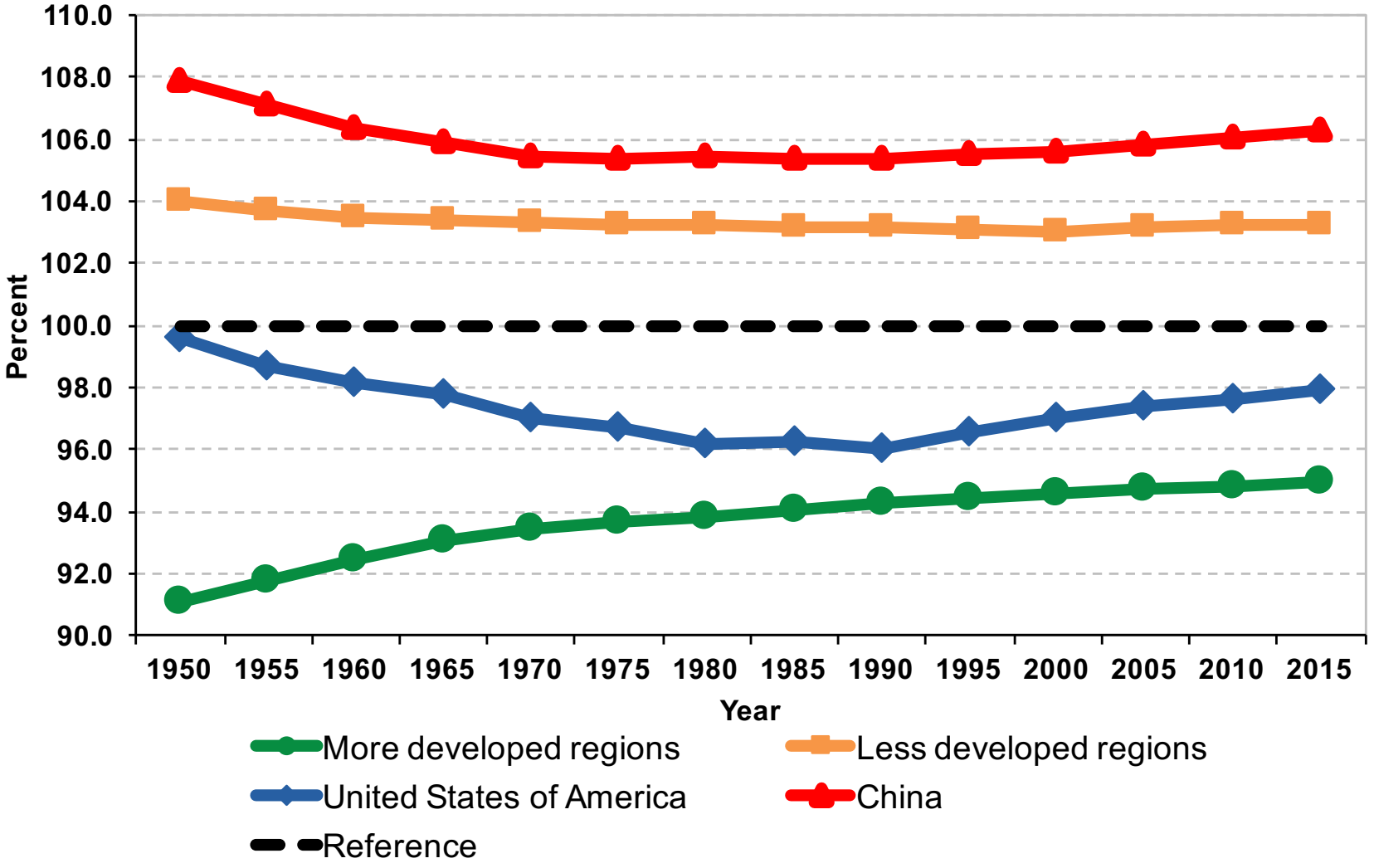
Ratios

- Describe a relationship between two numbers
  - Compare the size of one group to the size of another group
  - Compare the relative sizes of categories
  - Indicate how many times the first number contains the second
  - Denominator is not at "risk" of moving to numerator
  - Optional: multiply by 100 to get percentage

$$Sex\ ratio = \frac{Population\ of\ males}{Population\ of\ females}$$

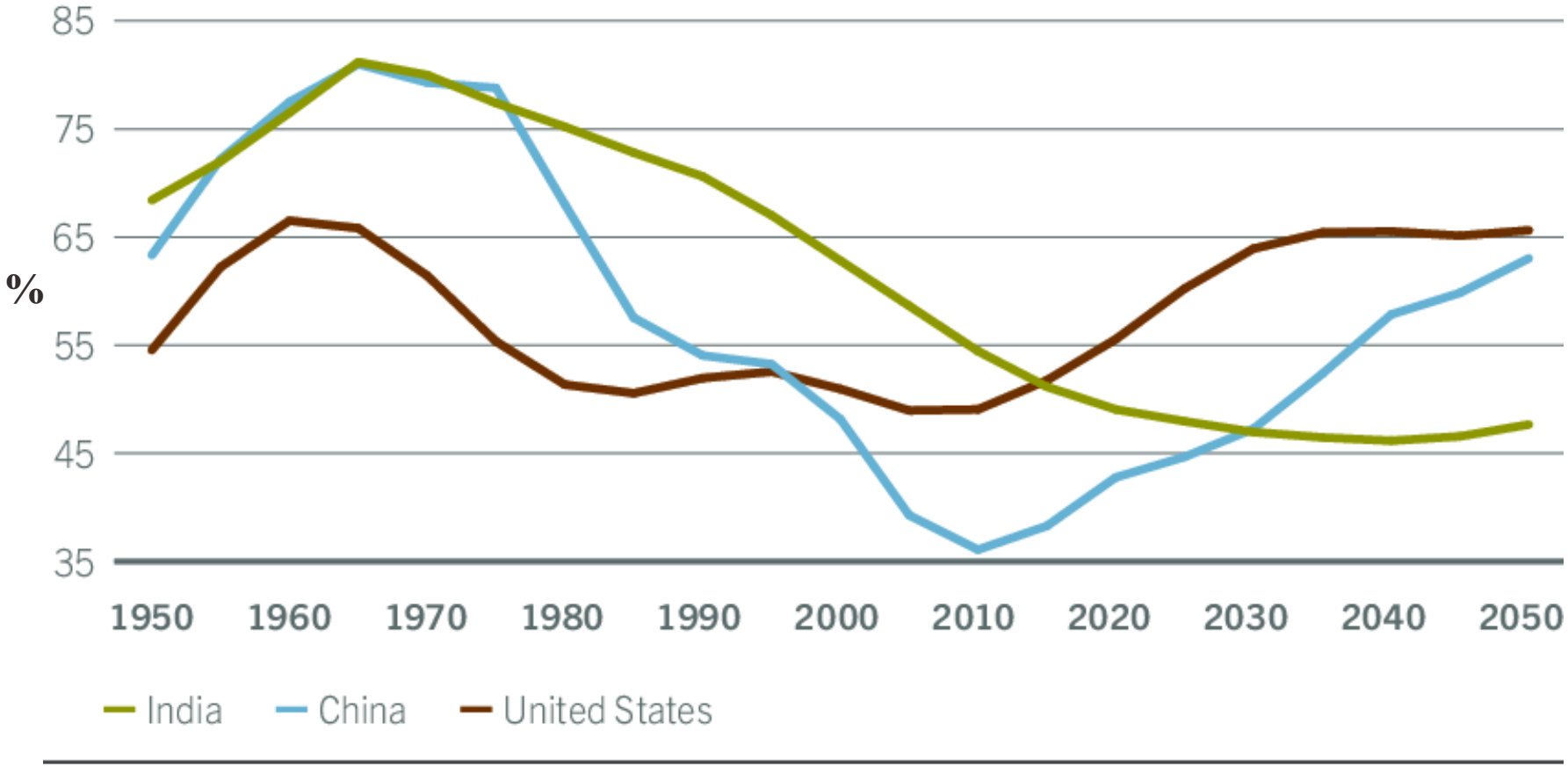
$$Total\ dependency\ ratio = \frac{Pop.\ children\ (0\ to\ 14)\ +\ Elderly\ pop.\ (65+)}{Working\ age\ population\ (15\ to\ 64)}$$

Sex ratios, 1950–2015



Source: United Nations, World Population Prospects 2017
https://esa.un.org/unpd/wpp/Download/Standard/Population/

Total dependency ratios of India, China, and the United States in %



Source: United Nations Population Division

Source: https://www.researchgate.net/figure/Total-dependency-ratios-of-China-India-and-the-US-In_fig3_275071394.

Rates

(Fleurence, Hollenbeak 2007)

- Rates are an instantaneous measure that range from zero to infinity
  - Rates describe the number of occurrences of an event for a given number of individuals per unit of time
  - Rates consider the time spent at risk
- Numerator
  - Number of events (e.g. births, deaths, migrations)
- Denominator includes time
  - Sum of each individual's time at risk of experiencing an event for a specific population during a certain time period (person-years)
  - We can use approximations for the denominator
    - Population in the middle of the period or
    - Average of starting and ending populations for that period

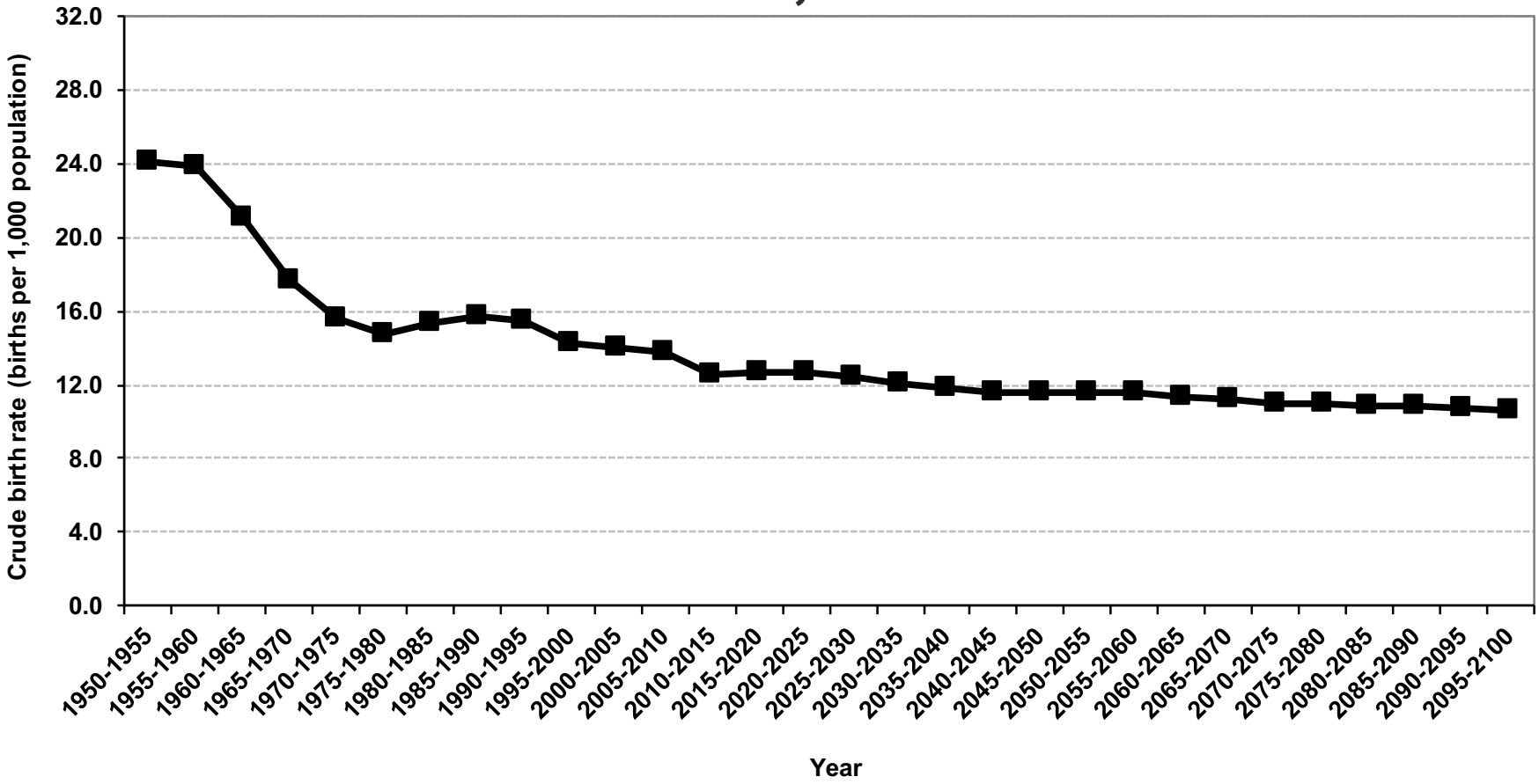
Crude birth and death rates

- Express the number of actual occurrences of an event (e.g. births, deaths, homicides) vs. number of possible occurrences per some unit of time

$$Crude\ birth\ rate = \frac{Number\ of\ births}{Total\ population} \times 1{,}000$$

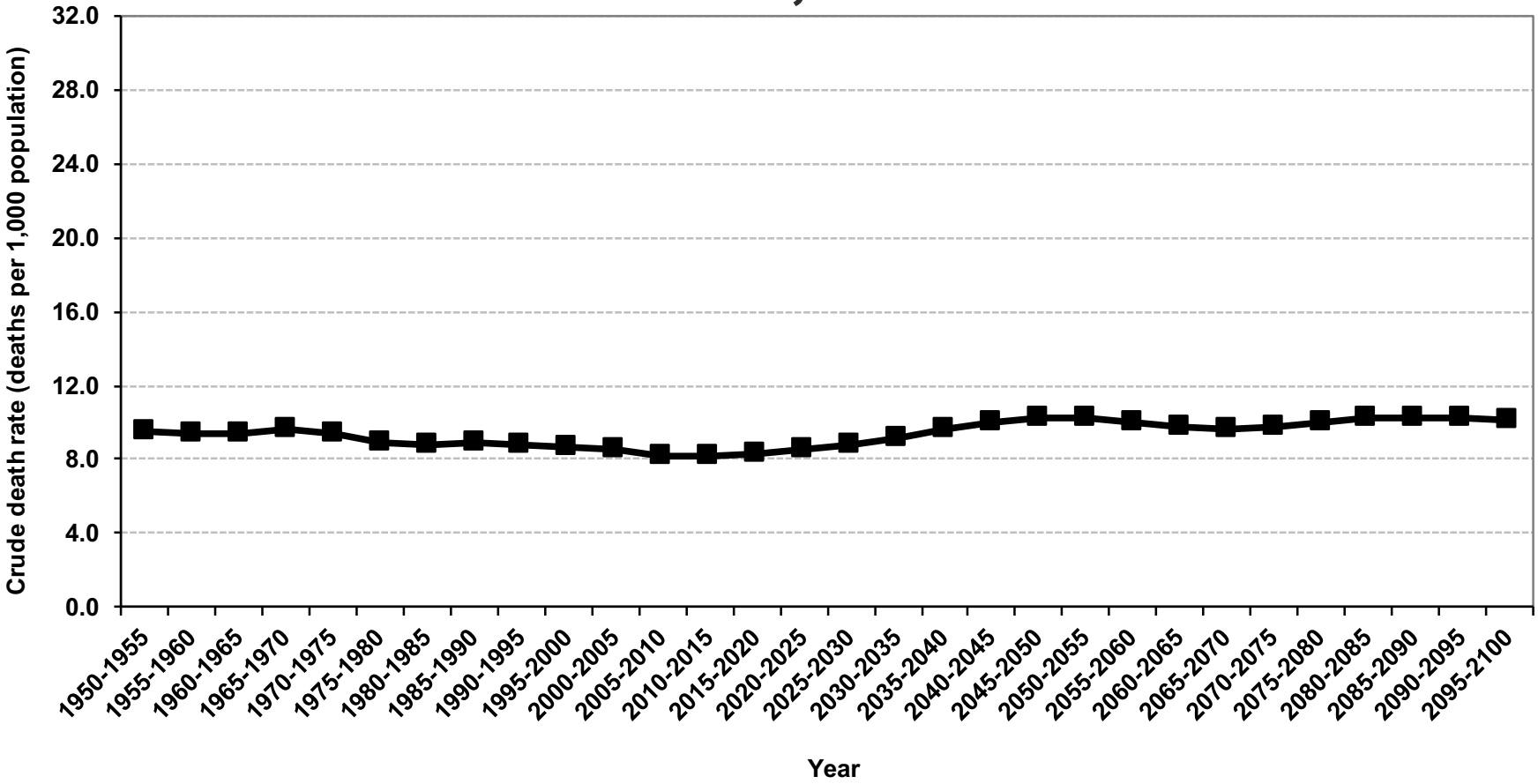
$$Crude\ death\ rate = \frac{Number\ of\ deaths}{Total\ population} \times 1{,}000$$

Crude birth rates, United States, 1950–2100



Source: United Nations, World Population Prospects 2017
https://esa.un.org/unpd/wpp/Download/Standard/Population/ (medium variant).

Crude death rates, United States, 1950–2100



Source: United Nations, World Population Prospects 2017
https://esa.un.org/unpd/wpp/Download/Standard/Population/ (medium variant).

Probabilities

(Fleurence, Hollenbeak 2007)

- Probabilities describe the likelihood that an event will occur for a single individual in a given time period and range from 0 to 1
  - Does not include time in the denominator
  - Divides the number of events by the total number of people at risk in the relevant time frame

- An approximation for the denominator is the population at the beginning of the period

References

Fleurence RL, Hollenbeak CS. 2007. “Rates and probabilities in economic modelling: Transformation, translation and appropriate application.” Pharmacoeconomics, 25(1): 3–6.

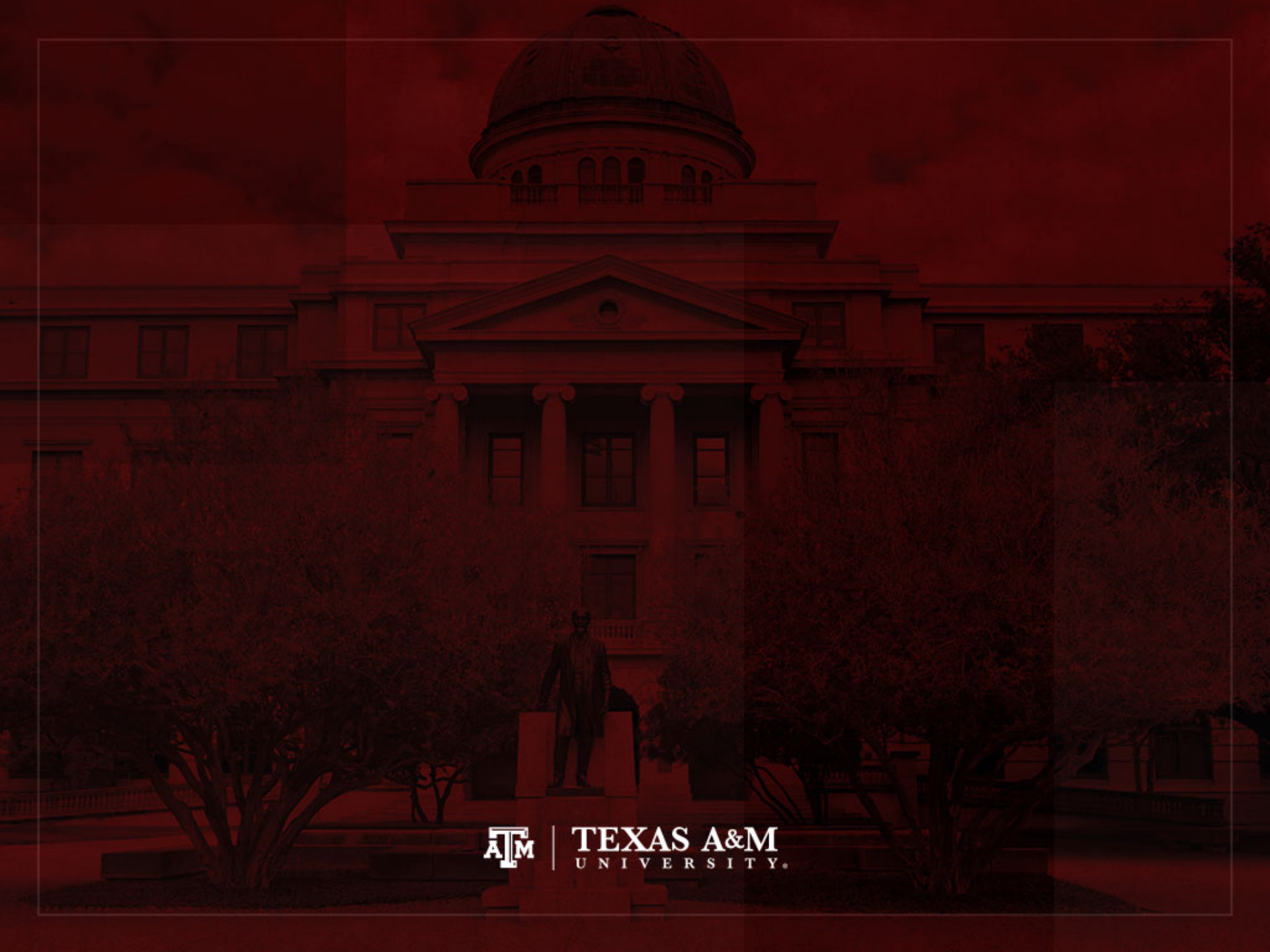
Healey JF. 2015. ”Statistics: A Tool for Social Research.” Stamford: Cengage Learning.

Hugo G. 2011. “Future demographic change and its interactions with migration and climate change.” Global Environmental Change, 21(Supplement 1): S21–S33.

Poston DL, Bouvier LF. 2017. Population and Society: An Introduction to Demography. New York: Cambridge University Press. 2nd edition. Chapter 1 (pp. 3–16).

Weeks JR. 2015. Population: An Introduction to Concepts and Issues. Boston: Cengage Learning. 12th edition. Chapters 1 (pp. 1–24), 2 (pp. 25–57).

Wooldridge JM. 2015. “Introductory Econometrics: A Modern Approach.” Boston: Cengage Learning.


TEXAS A&M
UNIVERSITY