Demographic Research Methods and the PyCCM library: Lecture One

Charles Rahal and Jiani Yan

Leverhulme Centre for Demographic Science

A lecture delivered at the Banco de la Republica, October 2025

Who is Charlie? And What Does he Do?

A population data scientist with many interests:

- Research Methods;
- Research Software Engineering;
- Macro-orientated demographic analysis;
- Responsible Research Practice;
- and many more...

- All with a methodological focus and ‘Open’ approach.

Who is Jiani? And What Does she Do?

A population data scientist with many interests:

- Social Epidemiology
- Research Software Engineering;
- Macro-orientated demographic analysis;
- Digital Gender Gaps;
- and many more...

- All with a methodological focus and 'Open' approach.

Who Are You? And What Do You Do?

- What are you interested in?
- What's your primary research focus?
- Whats your background?
- Studied demography before?
- How much Python experience?

- Perhaps we can do a quick loop around the room!

Course Outline: Delivery

The **delivery** of this course will be structured as follows:

- Lectures: 10:00-13:00 for the first three days.
- Computer Labs: 14:00-16:00 for the first three days.
- `PyCCM` hackathon: 10:00-13:00 on day four.
- Research Talks: 14:00-16:00 on day four.
- If for any reason you're unable to attend the class (i.e. Covid), we can turn the session hybrid and you can join remotely as necessary.

There is some light optional reading at the end of days one to three. Optional discussion questions accompany it.

Key Topics

- What are we going to learn about in this course?

1. Population change
2. Periods, cohorts, rates
3. Introduction to mortality
4. Advanced life-table topics.
5. Fertility measurement
6. Heterogeneity in fertility
7. Migration
8. Policy and Projections

- In general, these topics are approached from a more ‘macro’ overview.
- We will learn about demographic measures and models, mostly as they are applied to population-level data.

Key Texts

The **key texts** of this course (perhaps in order of relevance?) are:

1. Preston, S.H, P. Heuveline and M. Guillot (2001) ‘Demography: Measuring and Modeling Population Processes’, Blackwell Publishers.
2. Livi Bacci, M., (2012), ‘A Concise History of World Population’, Wiley-Blackwell.
3. Keyfitz, N. and Caswell, H., (2005), *‘Applied Mathematical Demography’*, Springer.

This Lecture

In **this** lecture we are going to...:

- First talk about some definitions, and history of the field.
- Then we'll consider long-term population growth.
- We'll discuss some basic measures of demographic accounting.
- We'll look at the concept of growth *rates*.
- We'll talk about Thomas Malthus and low growth.
- Understand Demographic Transition Theory...
- Talk about exposures, rates, and Lexis Diagrams.

This Afternoon

In **this** afternoon's class we are going to...:

- Talk about the most common ways of installing Python.
- Introduce the basics of Python relevant to `PyCCM`:
  - Floats, integers, strings, lists, functions.
- Talk about the key libraries that `PyCCM` uses as it makes projections:
  - Numpy, Pandas, Matplotlib.

We'll end the class by loading in a key `PyCCM` dataset and plotting some of the data.

Definitions

What do these three things have in common?

1. The expanding proportion of elderly in developed societies and the associated challenges on social insurance welfare systems.
2. The Franco-Russian alliance.
3. Multiculturalism.

They each involve one of the three key demographic drivers in some way.

- Mortality (increasing life expectancies and longevity at later years).
- Fertility (population birth rates were historically a military concern).
- Migration (between countries and cultures)

Definitions

What do these three things have in common?

1. The expanding proportion of elderly in developed societies and the associated challenges on social insurance welfare systems.
2. The Franco-Russian alliance.
3. Multiculturalism.

They each involve one of the three key demographic drivers in some way.

- Mortality (increasing life expectancies and longevity at later years).
- Fertility (population birth rates were historically a military concern).
- Migration (between countries and cultures)

The ‘description’ or ‘measurement’ of ’the people’

- What is the etymology of the term ‘demography’?
- From the Greek: ‘description of the people’. (dēmos, graphō)

“Demography is a science concerned with the analysis of the size, distribution, structure, characteristics and processes of a population.”

(Weeks 1994: 25)

An Interdisciplinary Approach

- All of these things vary over time and space.
- Our aim is to explain them.

"Demography is interdisciplinary in fact, not merely in aspiration"
(Wachter 2014, 2)

- For example; the LCDS employs not just demographers, but also computer scientists, ecologists, economists, econometricians, network scientists (with chemistry backgrounds!), public health scholars, and more.

Influential Historical Demographers Over Time

We can understand Demography through its history and sub-disciplines:

1. **John Graunt**: Presented first life-table to Royal Society, Feb 27th, 1662.
    - Based on merchant's market research data pertaining to potential demand, and variations in population caused by the Plague.
2. **William Petty**: the founder of 'political arithmetic'.
    - Along with Graunt, laid the foundation for modern census techniques, as well as modern insurance.
3. **Thomas Malthus**: population will increase exponentially, food at a smaller rate. Therefore, humans need 'preventative checks' to alleviate suffering.
    - Malthus is so incredibly important that we'll return back to him!

The Entrance of Mathematical Demographers

The turn of the 18th/19th century heralded the entrance of mathematical and statistical methods:

4. **Gompertz–Makeham**: death has age dependent/independent components.
   - Age-dependent component: the (Benjamin) Gompertz function.
   - Age-independent component: the (William) Makeham term.
   - The latter is negligible in low mortality countries.
5. In 1838, **Pierre François Verhulst** introduced the logistic growth model; generalization of exponential growth with maximum value for the population.
6. **Francis Galton** and **Ronald Aylmer Fisher** began to make the discipline increasingly statistical:
   - **Galton**: Introduced correlations, and was the first to apply statistical methods to the study of human differences (somewhat controversially).
   - **Fisher**: one of the three principal founders of population genetics.

Demographic Sub-disciplines

A range of demographic sub-disciplines have emerged over the last 50-60 years:

1. **Biodemography**: the convergence of demography, evolutionary biology, and genetics (e.g. the work of Mills *et al.*).
2. **Family Demography**: A focus on co-resident groups.
3. **Historical Demography**: A focus on populations of the past (e.g. CamPop).
4. **Mathematical Demography**: Variables and relationships represented in formal terms.
5. **PaleoDemography**: Prehistoric patterns of human population growth.
6. **Social Demography:** Examines the interaction of population and society.
7. **Computational Demography:** Agent based models, social media data, etc.

. . . and many more!

Populations

- Demographers are concerned with populations; a collection of persons alive at specific point in time which meet a certain criteria.

‘Populations, for demographers, are an “enduring collectivity”, one that “persists through time” even though its members are continuously changing through attrition and accession.’

(Preston, Heuveline and Guillot 2001: 1)

How does Demographic Research get Done?

Demographic Research consists mainly of three things:

1. **Description:** describe characteristics and regularities with measures such as population growth rate, life expectancy, total fertility rate.
2. **Explanation:** Demographers try to explain how and why patterns and regularities come about.
3. **Projection and Policy:** What's going to happen next, and how can we affect time paths?

This links into the broader 'To Explain or Predict?' debate of Shmueli (2010).

Data for Demographic Analysis

1. **Administrative records**: Collected by governments, examples include health, court and tax records.

2. **Censuses**: "individual enumeration, universality within a defined territory, simultaneity and defined periodicity" (UN definition), dating back to Biblical times (King David) and the Romans, Domesday book, etc.

3. **Vital Registration**: births (live and fetal), deaths, marriages and divorce.

4. **Surveys**: sampling of individual units from a population with a view towards making statistical inferences.

5. **Ethnographic data**: often largely based on observer notes, logs, diaries.

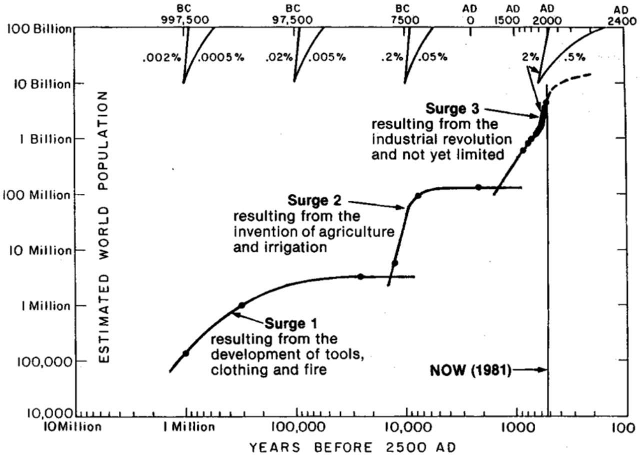
6. **'Big'** and **'Digital'**: new forms of big digital data from the web and social media, mobile phones, satellites, transactions, registrations.

Changing Data Paradigms in Demographic Research

- However, data shape, size, and provenance is changing **rapidly** in this area.
- Ever more 'micro' level data available to researchers, especially longitudinal.
- Example: the SAIL databank/safe haven of de-identified population data.
- Ever more census's (approximately two third of the world now covered).
  - For example, the ONS is increasingly open (with accreditation).
- Data linkage (e.g. UK Biobank to HES and GP records).
- Digital data made available through APIs (e.g. Facebook, Twitter, etc).
- There exist considerable inequalities: 'data rich' vs 'data poor' countries.
  - Population tools for countries thought to be accurate within 2-3% at best.

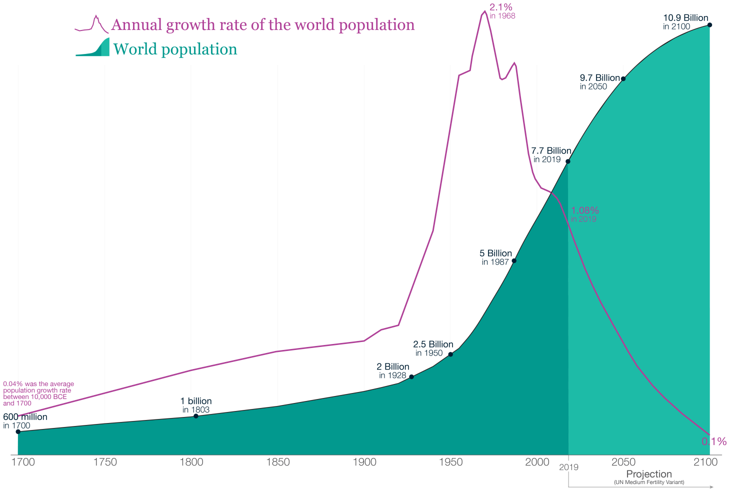
A lot of Demographic Analysis is about data visualisation.

Lets now introduce some basic key demographic concepts through this medium.

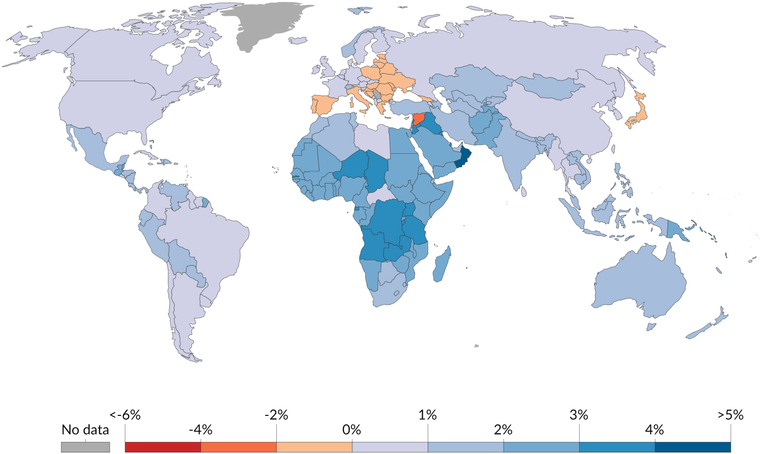


Global Population in the very Long run. Source: Tinsley (1980)

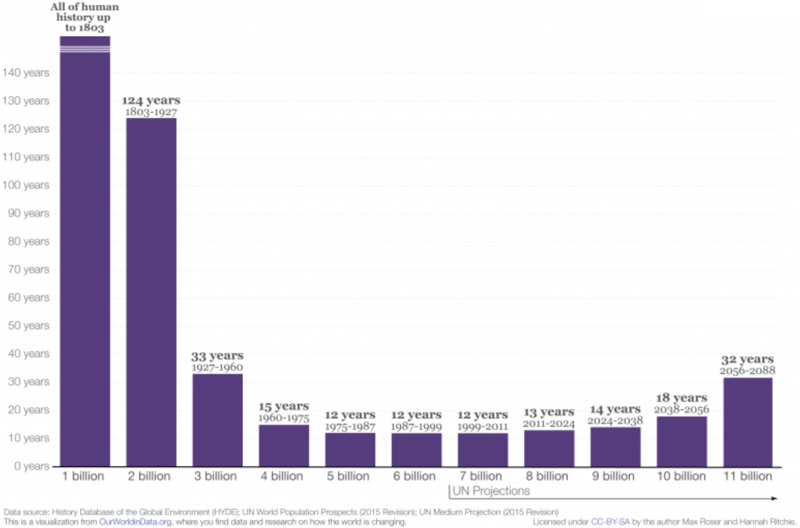
- Nobel prize in Medicine 2022 for: A previously unknown hominin, Denisova.



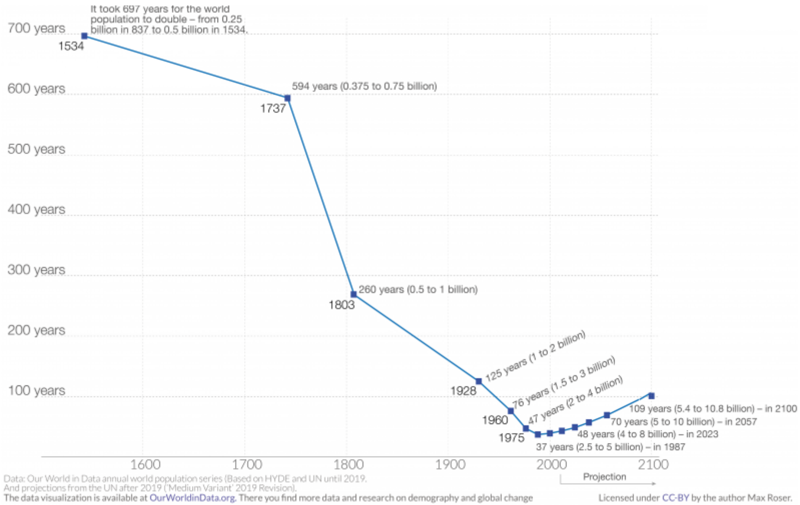
Global Population Growth. Source: Our World In Data



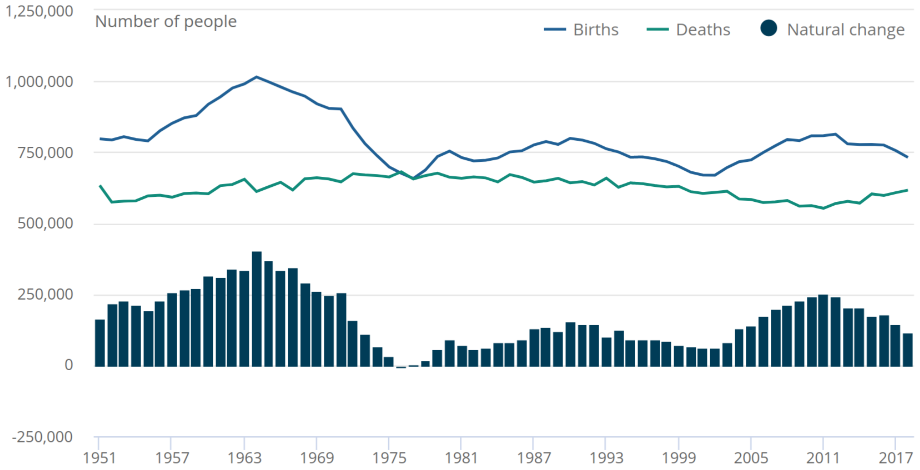
Population Growth Rate in 2015. Source: Our World In Data



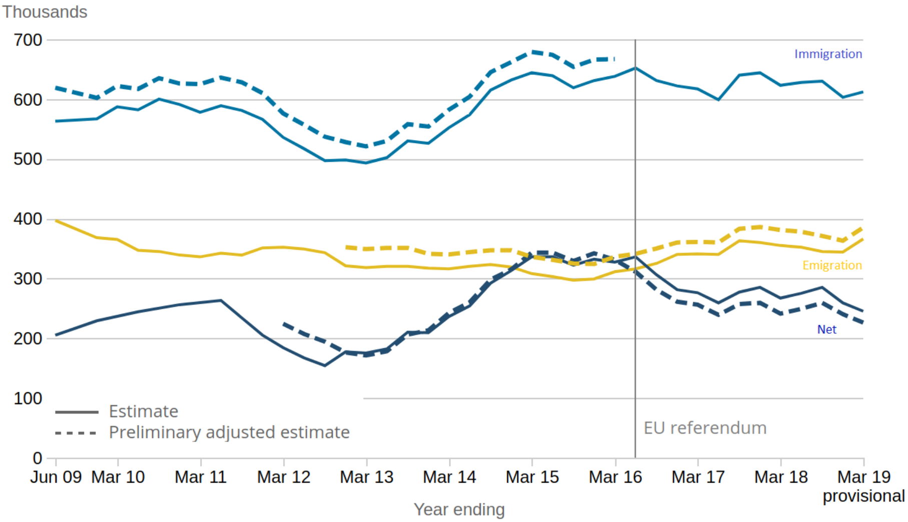
Time taken for an extra billion. Source: Our World In Data



An example of 'Doubling Time'. Source: Our World In Data



'Natural Change' in the United Kingdom. Source: ONS



Net Migration in the United Kingdom. Source: ONS

Demographic Accounting

- Lets now consider the basics of demographic accounting.
- The following equations are 'the most fundamental in social sciences'.
- The World Population in 2022:

$$\begin{aligned} K(1/1/2022) = K(1/1/2021) + B(1/1/2021 - 31/12/2021) - \\ D(1/1/2021 - 31/12/2021) \end{aligned} \tag{1}$$

- Where:
  - K is population
  - B is births
  - D is deaths
- Which parts of this equation are a *stock*, and which are a *flow*?

World Balancing Equation

- Generalising this, we get the '**World** Balancing Equation'.
- From time t to $t+n$:

$$K(t+n) = K(t) + B(t, t+n) - D(t, t+n) \tag{2}$$

- where:
  - K(t+n) is number of persons alive at time t+n.
  - K(t) is number of persons alive at time t.
  - B(t, t+n) is number of births between time t and t+n.
  - D(t, t+n) is number of births between time t and t+n.
- This is a **closed** population. How else can populations change over time?
- Note: we use t for time in general, and T to represent a particular time.

General Balancing Equation

Lets now incorporate (im)migration. From time t to $t+n$:

$$K(t+n) = \underbrace{K(t) + B(t, t+n) - D(t, t+n)}_{\text{Natural Increase}} + \underbrace{I(t, t+n) - E(t, t+n)}_{\text{Net Migration}} \tag{3}$$

- K(t+n) is number of persons alive at time t+n.
- K(t) is number of persons alive at time t.
- B(t, t+n) is number of births between time t and t+n.
- D(t, t+n) is number of births between time t and t+n.
- I(t, t+n) is number of immigrants between time t and t+n.
- E(t, t+n) is number of emigrants between time t and t+n.

There are two ways to enter a population, and two ways to exit. **What are they?**

Moving from Absolute to Rates of Change

Assume a closed population starting at $t = 0$, moving through time:

$$K(1) = K(0) + B(0,1) - D(0,1) \tag{4}$$

$$K(1) = K(0) \times (1 + \frac{B(0,1)}{K(0)} - \frac{D(0,1)}{K(0)}) \tag{5}$$

This term on the right hand side is our 'rate of change'.

Moving from Absolute to Rates of Change (Cont.)

Lets now see how we can iterate this forward;

$$K(2) = K(1) + B(1,2) - D(1,2) \tag{6}$$

As before, multiply our population by an update rate:

$$K(2) = K(0) \times (1 + \frac{B(0,1)}{K(0)} - \frac{D(0,1)}{K(0)}) \times (1 + \frac{B(1,2)}{K(1)} - \frac{D(1,2)}{K(1)}) \tag{7}$$

Where change is a function of of birth and death.

Growth Rates

- Population change is expressed in terms of rates of increase.
- Lets assume birth and death rates are constant, and drop subscripts:
- The rate of change is regulated by $\frac{B}{K}$ and $\frac{D}{K}$.

$$K(1) = K(0) \times \underbrace{(1 + \frac{B}{K} - \frac{D}{K})}_{\text{Let's call this A}} \tag{8}$$

$$K(1) = K(0) \times A \tag{9}$$

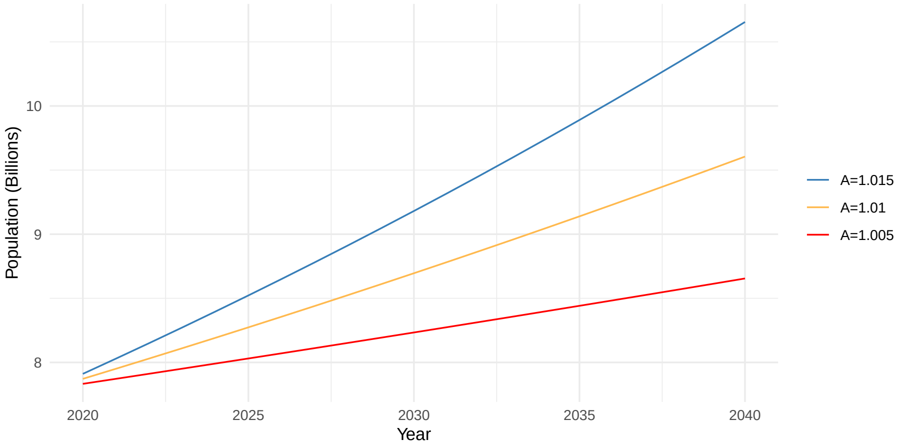
$$K(2) = K(0) \times A \times A \tag{10}$$

$$K(3) = K(0) \times A \times A \times A = K(0) \times A^3 \tag{11}$$

$$K(T) = K(0) \times A^T \tag{12}$$

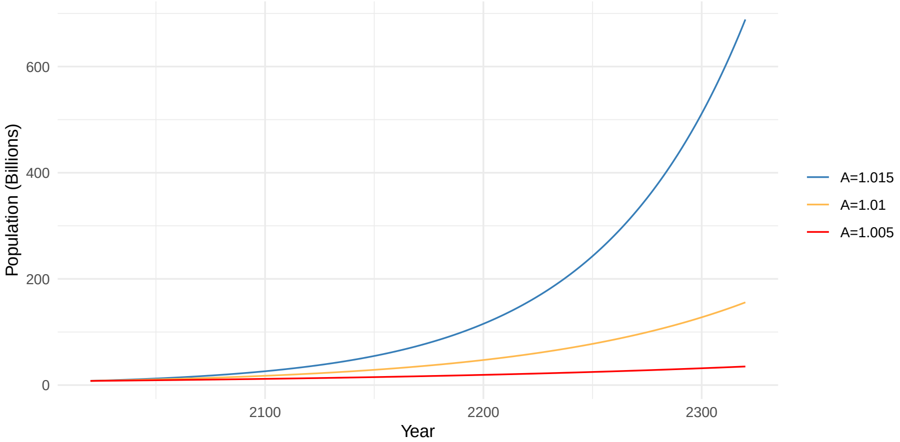
- When births exceed deaths ($B > D$), then $A > 1$.
- In this case, is population increasing or decreasing?

The Effect of Different A: 20 year horizon



Lets assume we're starting at Year 2020, when the population is 7.794bn.

The Effect of Different A: 300 year horizon



Lets assume we're starting at Year 2020, when the population is 7.794bn.

What does this simple exercise teach us?

- In the short run, there is a small difference only.
- In the long run, the exponential element creates a big difference.
  - Small differences are magnified.
- For example, by 2030:
  - When A=1.015, population projected to 9.18bn, 8.23bn when A=1.005.
- But by 2320:
  - When A=1.015, population projected to 688bn, 35bn when A=1.005.
- Note difference between *geometric* increases, and *arithmetic* increases.
  - Geometric is a fixed rate, arthmetic is a fixed increment.
  - Constancy of elements of growth translate better to geometric measures.
  - We can also have different rates within subpopulations, and over time.

Growth Rate Equations

- Lets remember our former equation for population at T:

$$K(T) = K(0) \times A^T \tag{13}$$

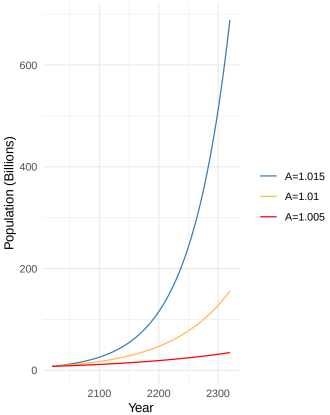
- Taking logarithms (remembering $log_a(b^c) = clog_a(b)$):

$$ln(K(T)) = ln(K(0)) + T \quad \underbrace{ln(A)}_{\text{Lets call this R}} \tag{14}$$

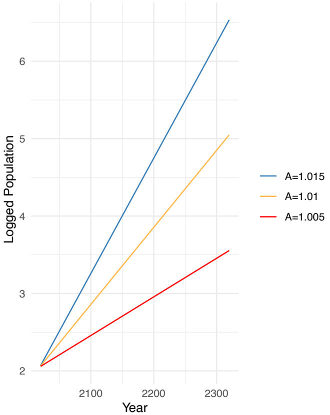
$$ln(K(T)) = ln(K(0)) + RT \tag{15}$$

- Note that this has the form $y = a + bx$!
  - Where might we – a group of economists – have seen this before?!

Proportional Rate of Change



(a) Levels



(b) Logarithms

- R is the slope of the graph of the natural log of population size over time.

Calculating Growth Rates from Populations

- If you know population at two points in time, we can calculate growth.
- From 6.929bn in 2010 to 7.349bn in 2015.

$$\frac{ln(7.349) - ln(6.929)}{2015 - 2010} \times 100 = 1.177 \tag{16}$$

- The population grew at an annual growth rate of 1.177.
- If you knew the growth rate, you could project the population forward:

$$\begin{aligned} K(2020) &= K(2010) \times e^{1.177} \\ &= 6.929\text{bn} \times e^{1.177} \\ &= 7.78\text{bn} \end{aligned} \tag{17}$$

- We'll learn a lot more about projections in the coming days!

Doubling Rate

- If R is constant and increasing, we can calculate a 'doubling time':

$$\text{Doubling Time } = \frac{ln(2)}{R} = \frac{0.693}{0.01177} = 58.89 \text{ years} \tag{18}$$

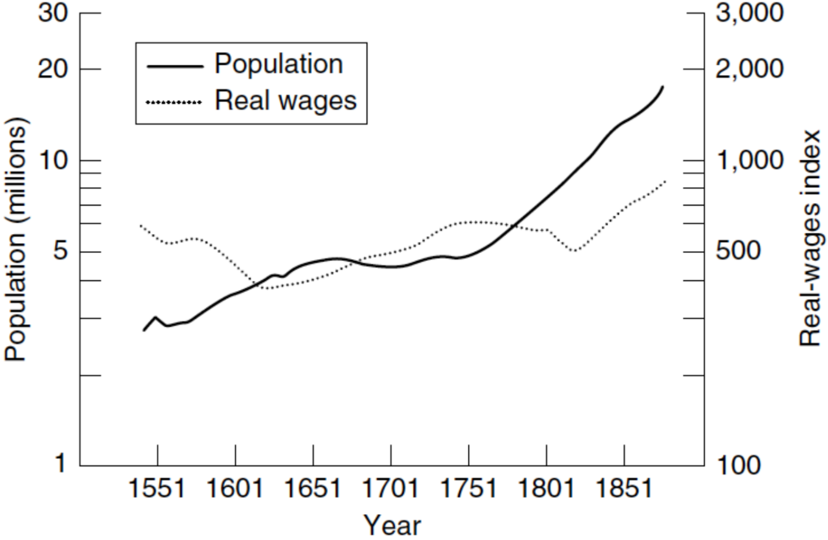
- This statistic is often more intuitive than a standard rate.
- Note: We can also think of a negative doubling time as a halflife.
- See Keyfitz and Caswell (2005, p.4) for a derivation of this.
- Note; doubling time is arbitrary.
  - Can you figure out the adjustment for trippling time?

Why Low Growth?

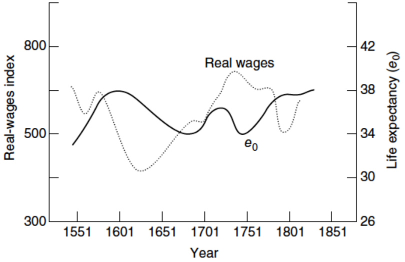
- The growth of world population was sluggish during the Malthusian epoch.
  - Average annual rate of about 0.1% over the years 0-1820.
- Key 1798 book 'An Essay on the Principle of Population'.
- An increase in food production improves well-being, but only temporarily because it led to population growth, which in turn restored the original per capita production level.

Homeostasis

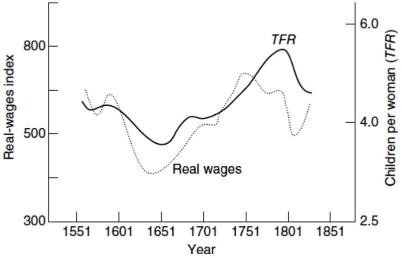
- Population multiplies geometrically and food arithmetically: when food supply increases, population will rapidly grow to eliminate the abundance.
- Population growth puts pressure on resources via a decrease in real wages.
- Larger demand of food increases prices.
- Larger labor supply decreases nominal wages.
- A slowdown in population growth occurs through two potential channels:
  1. A ‘preventive check’ (lower birth rate): marriage delayed or foregone.
  2. A ‘positive check’ (raise death rate): war, famine, epidemics.



'Was Malthus Right?' Source: Rigley and Schofield (1981)



(a) Real Wages and e_0



(b) Real Wages and TFR

'Was Malthus Right?' Source: Rigley and Schofield (1981)

- We begin to see death rates start decreasing ('Mortality Transition').
- Followed by a fall in birth rates ('Fertility Transition').

Further Comments on Malthus

- Malthus did not anticipate:
  - The substitution of tractors for horses and other draft animals
  - The comprehensive mechanization of farming
  - The use of chemical fertilizers or pesticides
  - The development of higher yielding crops
- Food production has not been growing ‘in an arithmetical ratio‘; exponentially and at least as fast as population.
- Coal, capital, and trade also played a role.

In a famous article, Notestein (1945) wrote about:

"The stage of transitional growth... in which the decline of both fertility and mortality is well established, but in which the decline of mortailty precedes that of fertility and produces rapid growth".

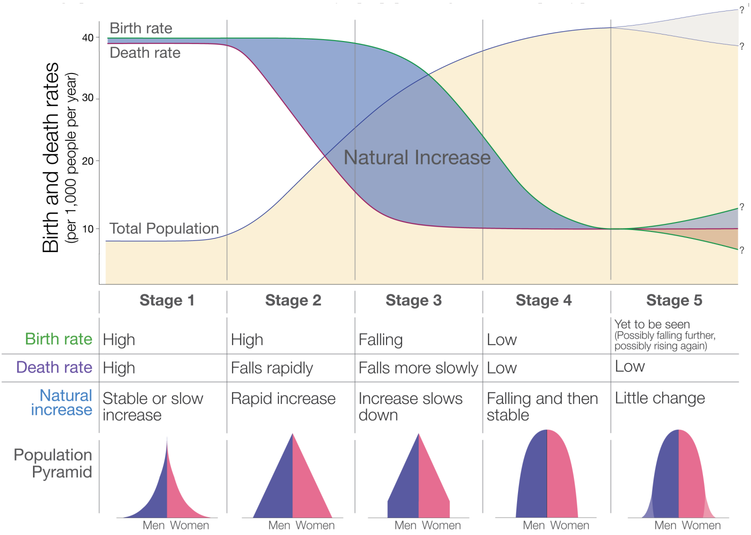
The Demographic Transition

- **Demographic Transition:** one of the best-documented generalizations in the social sciences.
- It describes the move from a pre-modern regime of high fertility and high mortality to a post-modern one in which both are low.

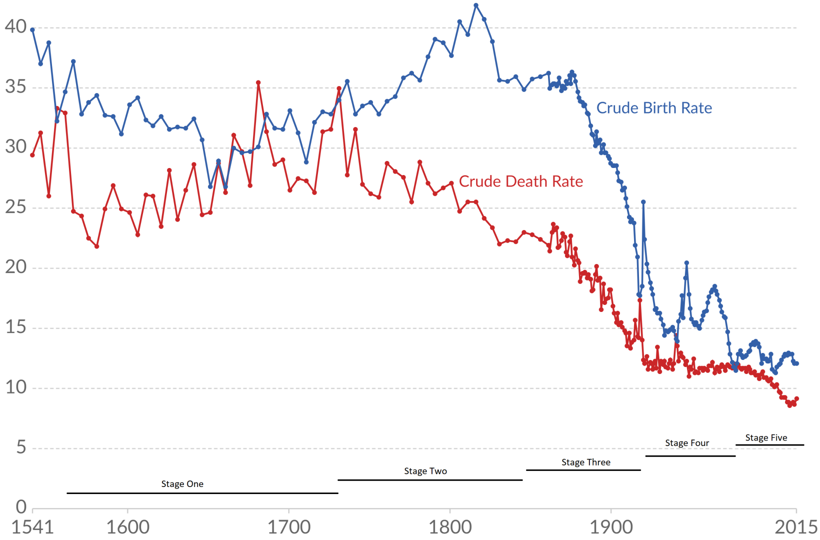
The Demographic Transition

It provides a break from the Malthusian cycle in **five** stages:

1. Death and birth rates both initially high, population determined by the food supply, evolving in accordance with the Malthusian paradigm.
2. Decline in death rate due to: a.) agricultural improvements implies less starvation/lack of water, and b.) significant improvements in public health.
3. Birth rates diminish due to: a.) urbanisation, b.) compulsory education, c.) increased literacy and employment, d.) conctraceptive technology.
4. Birth and death rates are both low, leading to a total population stability.
5. Both more-fertile and less-fertile futures have been claimed as a Stage Five.



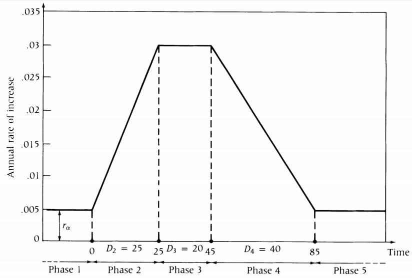
The Five Stages of Demographic Transition. Source: Our World In Data



The Transition in England and Wales. Source: Our World In Data

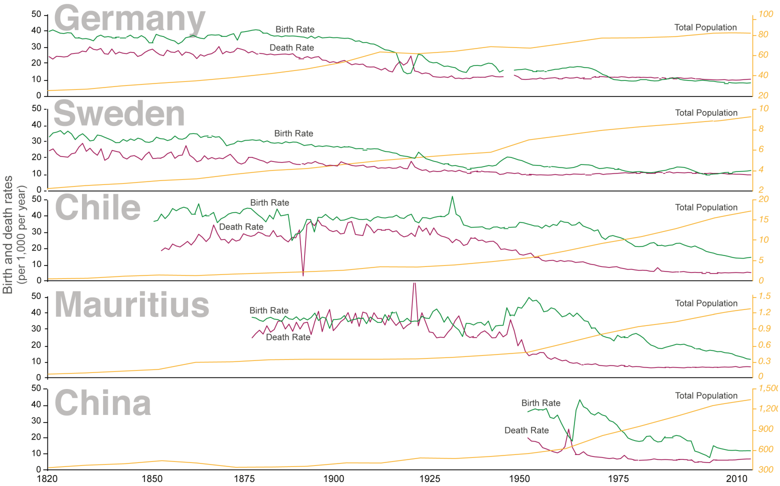
The Demographic Transition

- How much does a population increase during the demographic transition?
- The increase depends mostly on:



1. 'Length' (L) of transition: how long does it take? How long per stage?
   - Generally longer in earlier transitions.
2. 'Height' (H) of transition: how much larger is birth rate vs. death rate?
   - Bigger gap in the contemporary transitioning world.

- The above figure is taken from Chesnais (1990).
- This allows us to calculate a 'population multiplier' (M) across countries.



The transition occurs at different times. Source: Our World In Data

Time and Multipliers Differ

Country	Dates	Years	Multiplier
Sweden	1810-1960	150	3.83
Germany	1876-1965	90	2.11
Italy	1876-1965	90	2.26
USSR	1896-1965	70	2.05
France	1785-1970	185	1.62
China	1930-2000	70	2.46
Taiwan	1920-1990	70	4.35
Mexico	1920-2000	80	7.02

- The above table comes from Chenais (1990).
- It shows the large variation in timing, duration and magnitude of demographic transitions around the world.
- The diversity of cultures, socio-economic levels, and demographic factors – such as age at marriage – preclude precise prediction.
- No two countries have followed identical paths to transition.

Further Comments on the SDT

- The end point of the first demographic transition (FDT):
  - An older stationary population
  - Replacement fertility
  - Zero population growth
  - Life expectancies higher than 70.
- SDT predicts unilinear change: very low fertility/diversity of union/family.
- Shift in attitudes and norms in the direction of greater individual freedom/self-actualization
- Bounded by rational economic choice, it allows for autonomous preference drift (Maslow's shifting/hierarchy of needs).
- Related to post-materialism, with greater weight attached to individual self-realization, recognition, grassroots democracy, expression, and education.

The Importance of Age

"To represent a population as a number varying in time, and in disregard even of its age composition, is like treating the earth as a point in space: too abstract for most purposes, but still useful for some."

Keyfitz and Caswell (2005)

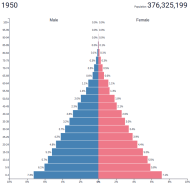
- Until now, we haven't emphasized the importance of age!
- We'll talk **a lot** more about this in Lecture 2.

Age Structure Over the Transition

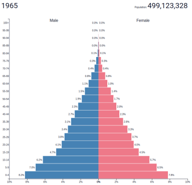
- At early stages of transition, populations show a young age structure.
- Driven by declines in infant/child mortality.
- As fertility declines, population growth slows.
- There is better survival to middle and older ages.
- We see 'base eroding' in mid-stage populations.
- Low fertility dramatically changes the shape of the pyramid.
- Lets now see some examples of this over time and across countries.

Age Structure Over the Transition

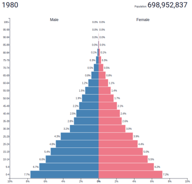
- At early stages of transition, populations show a young age structure.
- Driven by declines in infant/child mortality.
- As fertility declines, population growth slows.
- There is better survival to middle and older ages.
- We see 'base eroding' in mid-stage populations.
- Low fertility dramatically changes the shape of the pyramid.
- Lets now see some examples of this over time and across countries.



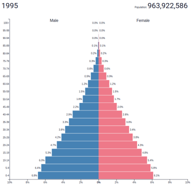
(a) India: 1950



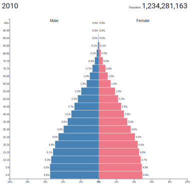
(b) India: 1965



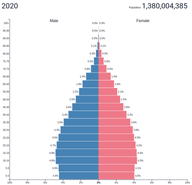
(c) India: 1980



(d) India: 1995

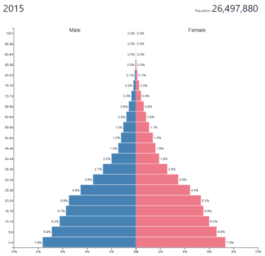


(e) India: 2010

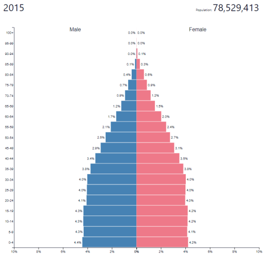


(f) India: 2020

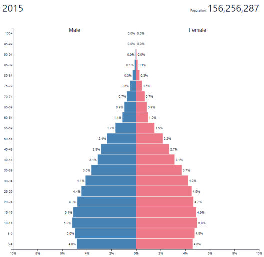
Transition Over Time. Source: populationpyramid.net



(a) Yemen: 2015



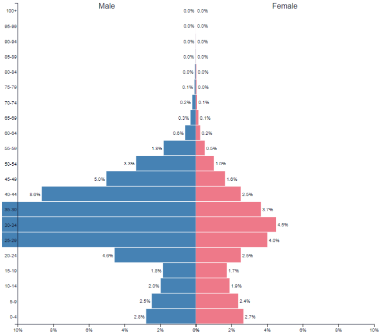
(b) Turkey: 2015



(c) Bangladesh: 2015

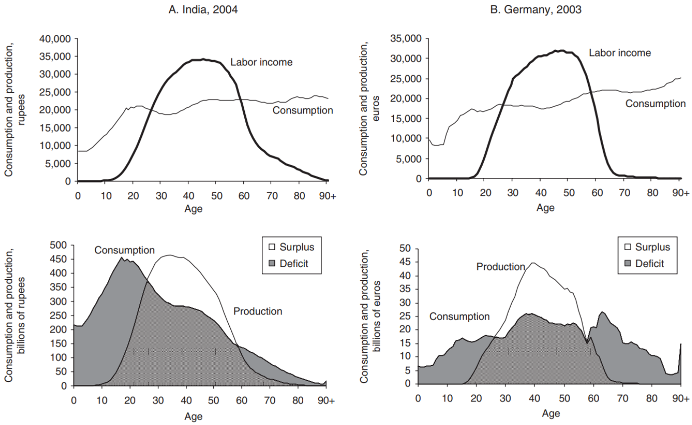
Transition Across Time and Countries. Source: populationpyramid.net

The transition story is silent about migration



UAE (2020): populationpyramid.net

- But migration also shapes age (and sex) structures of populations.
- When growth rate from natural increase is low, migration could be decisive.
- The UAE has huge immigration, primarily from India, Bangladesh, and Pakistan.
- Migrants comprise over 90% of the country's private workforce.



Why are Age Structures Important? Source: Lee and Mason, 2011

- People consume (young, old) and produce (middle age).

Demographic Dividend

- Two key ratios:

$$\text{Dependency ratio} = \frac{\text{Population aged 0-14 } \textit{or} \text{ 65}}{\text{Population aged 15-64}} \tag{19}$$

$$\text{Elderly dependency ratio} = \frac{\text{Population aged 65+}}{\text{Population aged 15-64}} \tag{20}$$

- Note: Support ratios are the inverse of dependency ratios.
- Before the transition/at early stages, age structure is too young.
- Too many people before reproductive ages creates a high dependency ratio.
- During 'transitional period', few children/few elderly: low dependency ratio.
- After the transition: high dependency.

Introducing Ages, Periods and Cohorts

- We've mostly been talking about growth in terms of births, deaths, and migrations relative to population size.
- Which part of a population, measured when, is it that we are talking about?
- The exponential model for population growth is **age-blind**.
- It treats all people as if they were alike.
    - But do all members of a population have a constant risk of death or of producing offspring?
    - Hint: Recall the Gompertz–Makeham law of mortality from Week One.
- When we calculated population sizes given specific growth rates, we talked of a particular **period**: T.

This T used previously might have been thought of as the **period**. Are there other ways we could think of time in relation to a population? What about the **age** of the group we're analyzing, or the **cohort** that they were born into?

This is going to put us in a good position to consider lifetables, and other more advanced demographic topics.

Coming up before lifetables:

We're going to learn about:

- 'Person-years' as a way to measure the population at risk for demographic events such as death or producing offspring.
- Age and demographic processes.
- Defining and differentiating between periods and cohorts in demographic measurement.
- Standardisation.
- The Lexis diagram to visualize temporal dimensions in demography.

Demographic Rates

- The **number** of events (e.g. births) occurring can be expected to be higher in a larger population, because there is more exposure.
- A **rate** relates the number of occurrences of event to the size of the population producing them.
  - For example: if there are more people of fertile age, we can expect more children to be born.
  - This would naturally result in a higher fertility rate.
- The number of events can also be expected to be higher the longer members are exposed to the 'risk' of occurrence.
  - For example: the length of the window we are analyzing.
  - If the period of study is longer, there is a higher 'risk' of the event happening.

General Rates Equation

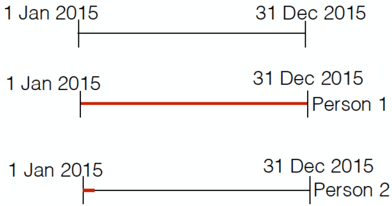
- Therefore, adjust metrics for population size/time period $\rightarrow$ rates.
- The most conventional exposure rate in demography takes the form:

$$\text{Rate} = \frac{\text{Number of Occurences}}{\text{Person-years of Exposure to the Risk}} \tag{21}$$

- It needn't be just in terms of births, deaths and immigration, though.
- The moon orbits the Earth at a rate of once every 27.322 days

Exposure

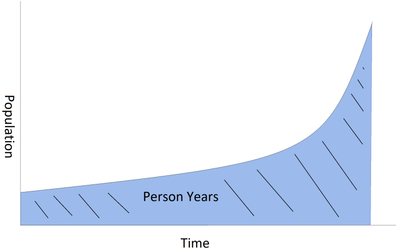
- When we defined the growth rate before, we defined it as $\frac{(B-D+I-E)}{K}$.
- We assumed K was the population size at the beginning of the interval.
- But why take the population at the start?



- Some people may be present at the beginning, but not at the middle of the interval, or appear sometime in the interval and stay until the end.
- Demographers approximate with 'person-years' to measure exposures.
- Population in the middle of the period multiplied by the length of the period is a common simplification. Why is this reasonable?

Person Years

- Think of a graph with population on the y-axis, and years on the x-axis.



- Each day a person is present contributes 1/365 person-years.

- If we have some discrete intervals, we can write:

$$PY[0,T] = \sum_{i=1}^{j} N_i \times \Delta i \tag{22}$$

- PY represents person years, N_i the number of people alive, and Δ_i is the change in time period.

Person Years

- If we know the whole sequence, we can just add up day by day. When subintervals are small, the sum is virtually equal to the area under the curve.

- So, as $\Delta t \to 0$, we can write:

$$\begin{aligned} PY &= \int_0^T K(t)dt \\ &= \int_0^T K(0)e^{Rt}dt \end{aligned} \tag{23}$$

- In the limit, we have a precise answer. Solving this expression gives a measure of PY. If we assume a constant growth rate, we can write:

$$PY = \frac{(K(T) - K(0))}{R} \tag{24}$$

Introducing Crude Rates

- This leads us to the concept of Crude Rates.
- Using this concept of person-years (PY), we can refine our general definition:

$$\text{Rate } [0,\ \text{T}] = \frac{\text{Number of Occurences between 0 and T}}{\text{Person-years of Exposure to the Risk between 0 and T}} \tag{25}$$

Introducing Crude Rates

- What type of rates are we most commonly talking about?

$$\text{Crude Birth Rate } [0,T] = \frac{\text{Number of births in } [0,T]}{\text{PY lived in } [0,T]} \tag{26}$$

$$\text{Crude Death Rate } [0,T] = \frac{\text{Number of deaths in } [0,T]}{\text{PY lived in } [0,T]} \tag{27}$$

$$\text{Crude Immigration Rate } [0,T] = \frac{\text{Number of in migrations in } [0,T]}{\text{PY lived in } [0,T]} \tag{28}$$

$$\text{Crude Emigration Rate } [0,T] = \frac{\text{Number of out migrations in } [0,T]}{\text{PY lived in } [0,T]} \tag{29}$$

Crude Growth Rate

- The crude growth rate (CGR) can therefore be defined similarly as before:

$$\text{CGR }[0,T] = \frac{B[0,T]}{PY\ [0,T]} - \frac{D[0,T]}{PY[0,T]} + \frac{I[0,T]}{PY[0,T]} - \frac{E[0,T]}{PY[0,T]} \tag{30}$$

$$\text{CGR}[0,T] = \text{CBR}[0,T] \text{ - } \text{CDR}[0,T] + \text{CRIM}[0,T] \text{ - } \text{CREM}[0,T] \tag{31}$$

$$\text{CGR}[0,T] = \underbrace{\text{CRNI}[0,T]}_{\text{Natural net increase}} + \underbrace{\text{CRNM}[0,T]}_{\text{Net migration}} \tag{32}$$

- It's important to realise that we need to have some *reference period*!
- Lets now turn to looking for/at some data regarding these calculations.

Year on Year Population, Both Sexes, 2010-2015

Data | Notes

Location	2010	2011	2012	2013	2014
World					
UN development groups (a)					
More developed regions (b)	1,234,768	1,239,557	1,244,115	1,248,454	1,252,615
Less developed regions (c)	5,722,056	5,801,637	5,881,713	5,962,129	6,042,676
Least developed countries (d)	836,615	856,471	876,867	897,793	919,223
Less developed regions, excluding least developed countries (e)	4,885,441	4,945,165	5,004,846	5,064,335	5,123,453
Less developed regions, excluding China	4,322,553	4,394,313	4,466,549	4,539,157	4,612,005
Land-locked Developing Countries (LLDC) (f)	420,661	430,709	441,057	451,699	462,624
Small Island Developing States (SIDS) (g)	65,253	66,035	66,779	67,491	68,180

Total population by sex (thousands).
Source: UN World Population Prospects 2019

Number of Births, Both Sexes, 2000-2020

Data | Notes

Location	2000 - 2005	2005 - 2010	2010 - 2015	2015 - 2020
World				
UN development groups (a)				
More developed regions (b)	66,014	69,917	69,133	66,942
Less developed regions (c)	600,516	616,478	628,639	634,336
Least developed countries (d)	131,688	140,532	148,952	157,956
Less developed regions, excluding least developed countries (e)	468,827	475,946	479,686	476,380
Less developed regions, excluding China	513,268	528,744	539,044	548,007
Land-locked Developing Countries (LLDC) (f)	64,585	70,177	75,557	79,401
Small Island Developing States (SIDS) (g)	6,562	6,539	6,594	6,502

Number of births, both sexes combined (thousands).
Source: UN World Population Prospects 2019

Crude Birth Rate, Both Sexes, 2000-2020

Data | Notes

Location	2000 - 2005	2005 - 2010	2010 - 2015	2015 - 2020
World				
UN development groups (a)				
More developed regions (b)	11.0	11.4	11.1	10.6
Less developed regions (c)	23.3	22.3	21.2	20.1
Least developed countries (d)	37.6	35.6	33.5	31.6
Less developed regions, excluding least developed countries (e)	21.1	20.1	19.1	17.9
Less developed regions, excluding China	27.0	25.5	23.9	22.5
Land-locked Developing Countries (LLDC) (f)	36.5	35.3	33.8	31.5
Small Island Developing States (SIDS) (g)	22.1	20.7	19.7	18.5

Crude birth rate (births per 1,000 population).
Source: UN World Population Prospects 2019

Summary of Class One

- World population growth took off sometime in early 1800s.
- The exponential growthis a simple model of population growth.
- The Transition is a stylised description of population development.
- Pathways of the transition have been diverse.
- With significant implications for population age structures.
- We also learnt about:
  - Methods for constructing rates
  - Exposure and person years
  - Importance of standardisation
  - Lexis diagrams
  - Basic foundations of age-period-cohort analysis.
- Check out Our World in Data/Population Pyramids; what can you find?