

Midterm exam “Population Analysis”

- i. On Wattle, students have access to the exam and data used at 9am.
- ii. Each question (1, 2.a, 2.b, etc) is equivalent to 10% of the exam. Please answer all of them. An empty answer is a grade of zero on that question.
- iii. Calculations will be needed to solve the exam and students can use any software for that (R, excel, any...). Please do NOT include excel files or R code in your answers, just the final output that you are requested, either a Number, a Table or Figure.
- iv. Students should submit a PDF of their exam answers in turnitin before noon (12pm). Late submission (5% rate).

Name: Hannah Zhang.

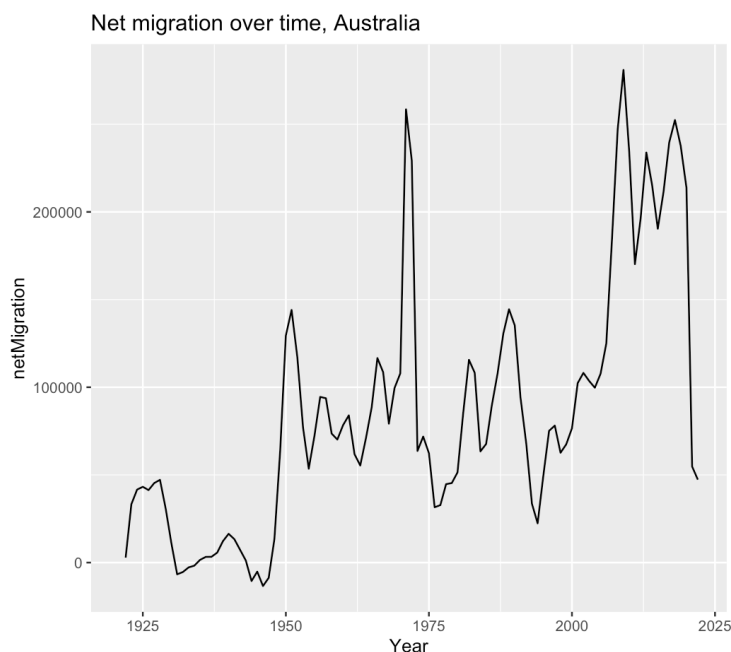
ANU u-ID: u7255134.

1. How would you describe *Demography*? Do you consider it is needed? Why? use your own words (no copy/paste from anywhere... it is really bad to have the same answer as somebody else... max 200 words)

Demography is the study of human populations. Conceptually, I currently depict it as being centred around the population balance equation (where change in population over time can be accounted for by counting births and deaths (natural causes) and immigration/emigration (net migration, either international or internal)). I do believe that Demography is important in informing policy decisions, particularly for social security. Demography enables policymakers to project current social needs and what might be expected from the population in the future. Demography allows us to quantify the need for fertility treatment and childcare, of student load and the power of the available workforce, or of the social security and pension allowances. For example, although low fertility conditions don't immediately affect a population, a sustained period of low fertility eventually increases the burden on young people to support the economy when the nation is overweighted with older, less physically able workers. Additionally, fertility rates below replacement level may cause the state to become reliant on migration for population increase. By being able to decompose and analyse the individual components that contribute to population change, Demography is special from other statistics-heavy fields by accounting for human experiences and guiding informed policies.

2) The balance equation & migration: In Wattle you are given the data “AUS.txt” with the series of Population, Births, and Deaths for Australia (female and male combined), for every year from 1921 to 2021, with population also for 2022.

2.a) Calculate the net migration counts from 1921 to 2021, and present a plot with their time trend. Mention how you calculated the net migration counts, and discuss the maximum and minimum migration counts observed in Australia, and the years when they occurred. Why do you consider the max and min values were observed in those years (max 200 words).



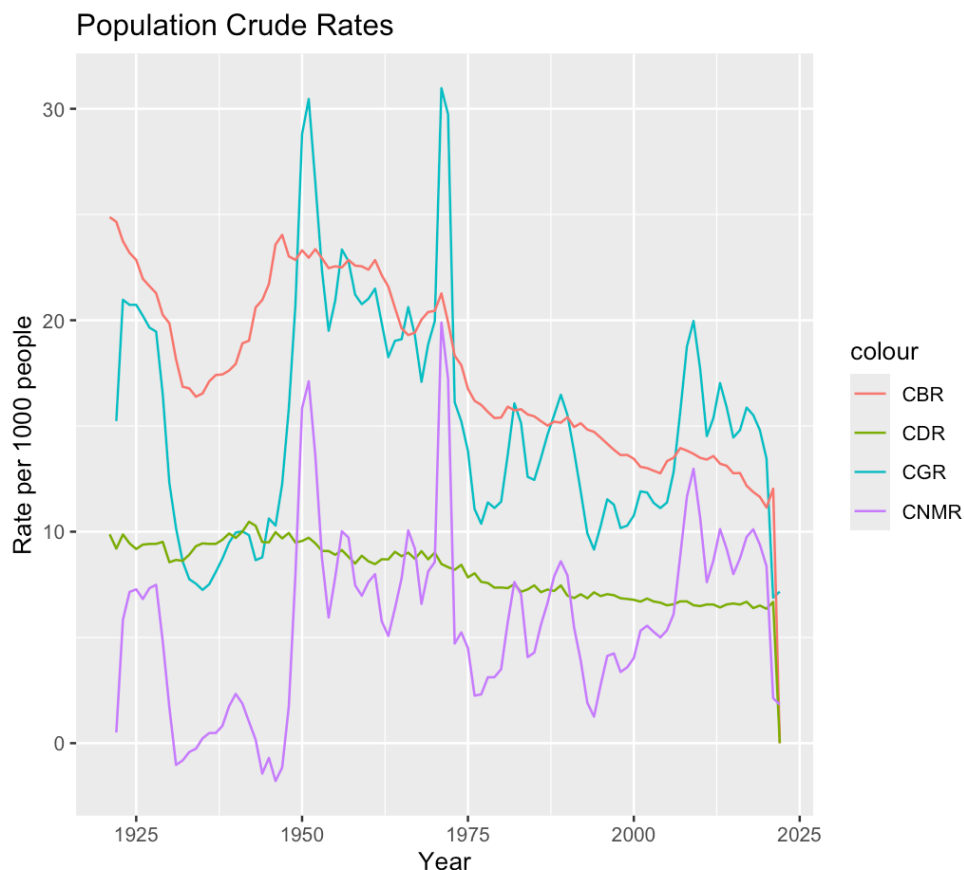
For the rows in AUS.txt, we assume that the births and deaths have not contributed to the population in that row. To calculate the net migration counts for year X, we subtract by the net natural increase based on prior fertility/mortality data and also subtract by the “initial” population from the prior year.

Net migration at year X = $\text{population}_X - \text{births}_{\{X-1\}} + \text{deaths}_{\{X-1\}} - \text{population}_{\{X-1\}}$

Following this calculation, we have a maximum of 280961.5 in 2009 and -13375.73 in 1946. So, net migration during the WW2 and the post-war period was shaky and perhaps not abundant in incentive to migrate to Australia. In fact, they may be emigrating to reunite with relatives overseas. Following globalisation and further economic development, migration came to increase rapidly after 1950s into the

2000s period as transport becomes cheaper and Australia becomes refuge for economic opportunity and public policy encourages immigration.

2.b) Calculate the Crude Growth Rate (CGR), Crude Birth Rate (CBR), Crude Death Rate (CDR) and Crude Net Migration Rate (CNMR) for the years 1921 to 2021. Present one plot including the time trends in all those rates



2.c) Mention how you calculated the rates in 2.b and discuss, from your findings in the plot in (2.b), your vision of the future for the components of the population growth rate in Australia (i.e. the future for CBR, CDR and CNMR) (max 200 words).

We have $CGR = \text{PopChange} / PY$, $CBR = \text{Births} / PY$, $CDR = \text{Deaths} / PY$, $CNMR = \text{netMigr} / PY$. We multiple by 1000 to scale the rate to per '000 people. We estimate PY using the population average between two years (i.e. the initial year + NEXT year, divided by 2). PopChange is calculated by the difference between the initial year and last year.

Following the patterns of most developed countries, I expect that Australia's CBR will continue to face decline. After 1950, we observe that Australia's CBR has been steadily falling off throughout the years. Indeed, most developed countries have similar fertility crises, like South Korea.

Australia's CDR has largely remained steady, though possibly indicating a marginal decrease through the years. We cannot definitively infer what a shrinking CDR entails, because it is sensitive to the age structure of the population. Barring any major catastrophe, the CDR should be resilient to disturbances but may reflect changes in age structure and healthcare.

Australia's CNMR has fluctuated rapidly between 1921 and 2021 when presented in this format. Tall spikes are observed around 1950, 1975, and with a slight upward trend after 2000. Because the CGR curve closely matches the CNMR curve, it appears that migration plays a crucial role as a growth factor. I am confident that the contribution of migration to population growth will be ever-present in the future as fertility declines and migration remains flexible.

2.d) Describe the theoretical age-pattern of migration and some of the reasons why it has that shape. Also discuss if you consider if humanitarian migration shows the same age-pattern? (maximum 200 words).

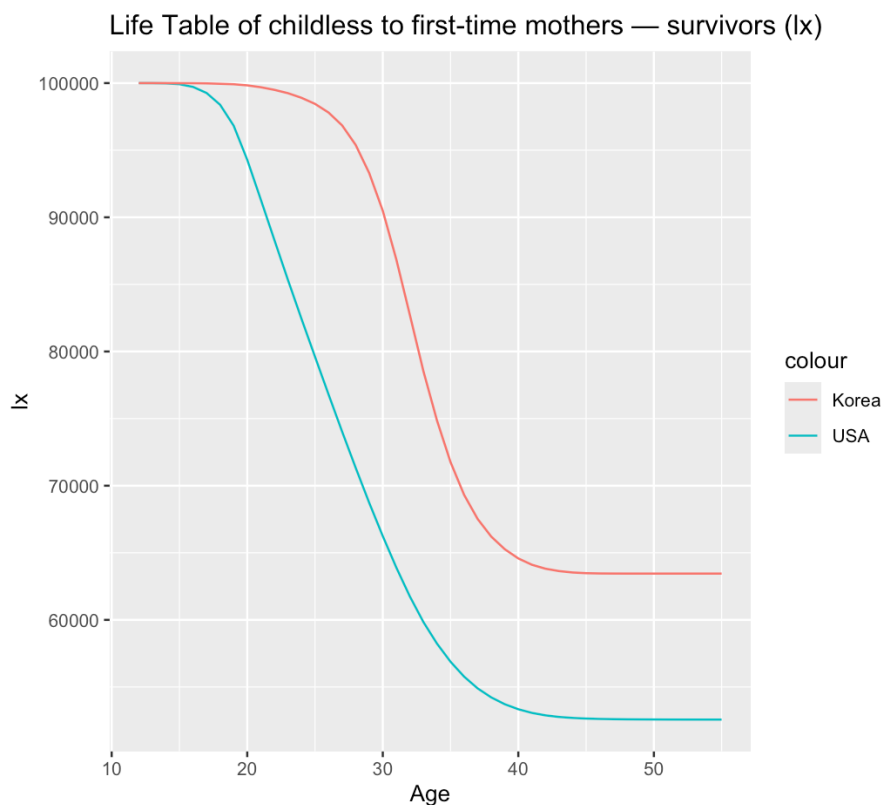
The theoretical age-pattern of migration describes the migration intensity that is typical of certain age intervals in one's life. In the infant and teenage years, the individual typically has less incentive and means to migrate. As they enter early adulthood, entry into higher education or the labour force often motivates individuals to migrate for their opportunities. Migration intensity dwindles and stabilises as the individual approaches retirement age, with a slight upward trend near the end of one's life signally a move to elder care facilities.

We take humanitarian migration to mean international migration due to crisis in their home countries (such as persecution, war, and natural disasters). For one-off and non-prolonged disasters, humanitarian entry need not be age-specific as usually any citizen may seek shelter. For systemic crisis, the desire to migrate could persist throughout their lives and may see a general upwards shift of the migration intensity curve—people already living in developed countries wouldn't have this urgency to

flee their country. In aforementioned scenarios, the same age-pattern curve shape should persist because the same migration reasons remain applicable, though significant life events perhaps might be delayed to later years as migrants “catch-up” with milestones they missed originally.

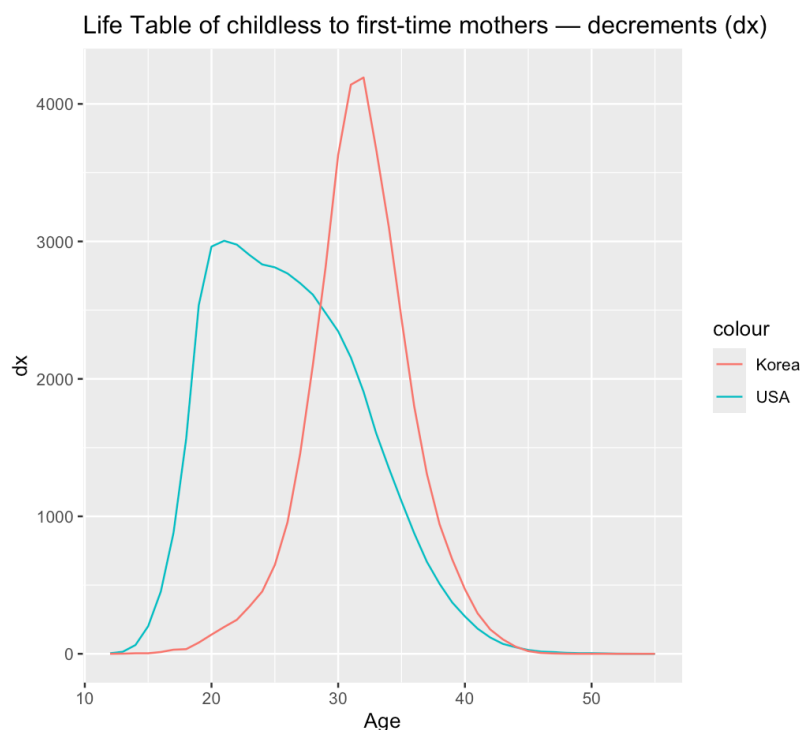
3) Life tables & fertility: you are given **age-specific fertility rates (ASFR)** and **ASFR for first births (ASFR1)** for South Korea and the USA in 2022 (from the file “ASFR1.txt”). You are asked to construct **TFR** and **MAC** for both ASFR & ASFR1, and **life tables based on the ASFR1 data**. *This ASFR1 life table assumes that we follow a cohort of childless women starting at age 12 (the radix of the population, $l_x(12)=100,000$) and that they follow the age-pattern of the fertility rates ASFR1 until age 55. That is, your life table only goes from age 12 to age 55 (Hint: you can assume that $a_x=0.5$ for each age, and your q_x in the last age group does not have to be 1). Below in the questions when I ask about life table functions, I am referring to this “ASFR1 life table” based on fertility data.*

3.a) Present a plot including both the South Korea and USA “survivors”, l_x , from your calculated ASFR1 life tables, and describe the trends. (max 200 words).



Starting at the radix $l(12)=100000$, we see the steepest drop in childlessness during a woman's 20s (for USA). However, this drop is delayed in Korea, indicating that Korean women tend to delay their first babies. This decline tapers off, in both graphs, at around age 40s—by this time, most women have accepted their choice for childlessness. We can see that Korea remains with a larger proportion of women remaining childless by the end of the period than USA.

3.b) Present a plot including both the South Korea and USA “decrements”, dx , from your calculated ASFR1 life tables, and describe the trends. (max 200 words).



From the decrements plot, we see that a large intensity of fertility spikes in the early 30s for Korean women. However, American women have a slightly prolonged period of fertility, indicated by the curved mound spanning their entire 20s, followed by a drop off during the 30s. American women approach the 20s peak quite early beforehand, but for Korea, their fertility decrements are much more regimented and narrower before and after the peak.

3.c) For each, South Korea and USA populations, present a Table including TFR and the Mean Age at Childbearing (MAC) for each of ASFR & ASFR1, and two new measures based on the ASFR1 life table: the Proportion of Females Remaining

Childless (PFRC) and the “Life Expectancy between ages 12 and 55” (LE(12to55)).
Mention how you calculated the PFRC? (max 100 words).

I calculate the proportion of females remaining childless as the proportion of women, of the radix $l(12)$, who have not had a first baby by the end of the childbearing period (i.e. by 55). That is, $l_{55} / 100000$.

	ASFR		ASFR1			
	TFR	MAC	TFR1	MAC1	PFRC	LE(12to55)
USA	1.65686	29.90659	0.64302	27.79313	0.5256747	30.112032
Korea	0.77894	33.4096	0.45487	32.75073	0.6344973	35.316624

3.d) How big or small are the disparities between South Korean and the USA in: TFR, MAC, PFRC and LE(12to55) in (3.c) above? What is each measure telling us? (max 200 words).

Firstly, the TFR based on ASFR tells us the number of babies we expect a woman to bear if she ages and experiences each of the age-specific fertility rates we have recorded. USA's TFR is 1.7, which is much higher than Korea's TFR 0.78 (close, but not quite one baby).

Similarly, TFR1 regards only first births and has a similarly big disparity as TFR. It presumes a woman continually undergoes age-specific TFR rates.

The PFRC tells us the proportion of women we expect to remain childless after experiencing the decrements (first order fertility) of each age. Interestingly, we learn that just a slightly lower proportion of American women are expected to remain childless at 55 than Korean women (0.53 vs 0.63).

The LE tells the remaining years until we expect that a woman will give their first birth OR they will reach the end of their childbearing period — both represent a “deadline” for them to give first birth. The LE(12to55) is lower than the age distance to 55

because of the likelihood the woman might indeed give first birth earlier than age 55.

3.e) What did we gain from applying the life table to the age-specific fertility rates for first births (ASFR1), specially what are PFRC and LE(12to55) adding that we didn't have already from the TFR and the MAC? Explain your answer (max 200 words).

We can see that the TFR1 should be nonsensical, because a woman cannot again be at risk of birthing a first child after she's done so already—the calculation implies otherwise. Rather, the duration until first birth (with age-55 cutoff) is better represented by the life table method using ASFR1 as decrements. The LE(12to55) gives us the expectation of how long until we won't need to have regard for the first birth risk of a woman — either they bear their first baby or they reach the end of the period. On the other hand, the MAC1 merely considered fertility rates, but not a “deadline”, which is why LE is slightly higher (skewing towards those who would reach 55 childless).

We notice that the TFRs have a big disparity, but the PFRCs have a smaller relative disparity. This indicates that while a fraction of women in both countries similarly experience and choose childless living, the women in USA who *do* choose children possibly choose to have more than those in Korea. This could be attributed to cultural differences.