A DEVELOPING COUNTRY

A nation, once underdeveloped, is going through a phase of relative prosperity.
Thanks to these mutated social and economic conditions, the population is more confident about the future, and the government is able to plan long-term politics aimed at giving welfare to the people.
This is reflecting in an improvement of several demographic indicators: for instance, the life expectancy is increasing, and the child mortality is steadily decreasing.

The nations in the same geographical area are glad to see this success hoping to being able to replicate it, but at the same time they have a bit of apprehension, since there are open disputes on some border territories, and becoming a more developed country can also signify being able to put more pressure (also under the military point of view), thus potentially causing clashes with their interests.
In order to deal with this situation, these bordering nations are working on predictions regarding the future of the watched neighbour.
One of these projections concerns the population evolution, since the more inhabitants the country has, the more it can build hard and soft power.
Given that this country possesses also an high fertility rate, it is deemed appropriate to predict the habitancy on a yearly basis, by considering the beginning of each year (that is, at the 1st of January).

Task 1: the population at the start of 2025 was 10.5 million, with an increase of 5% if compared to the inception of 2024, when the population was 10 million.
Since, due to the consideration of several factors, analogous increments are predicted for the following years, two models are considered.
The first one always applies the percentage of 5% to the initial population, i.e. the one of 2024 (10 million), regardless of the years passing by.
The second one applies the same percentage, but each time to the population relative to the previous year, thus taking into account the variations that have already occurred.
Which model will result in a higher prediction? Try to reason both numerically (by testing for a certain number of years) and algebraically (by generalizing and possibly motivating the findings).

Solution: the first model can be represented by $P_n = 10 \cdot \left(1 + \frac{5\,n}{100}\right)$, where n represents the number of years elapsed starting from 2024 (i.e., $n = 0$ refers to 2024, $n = 1$ to 2025, and so on).
The second model can be represented by $Q_n = 10 \cdot \left(1 + \frac{5}{100}\right)^n$, with the same meaning for n. In both representations, the population is in millions.
It could be easily seen by expanding the n-th power as Newton binomial that, for $n > 2$, $Q_n > P_n$.
Related theory: arithmetic and geometric progressions

Task 2: in addition to the overall population, it is figured important to make predictions on its composition by demographic segmentation according to the age.
Indeed, the buildable power depends also on how many people can be conscripted, thus excluding from

the total count both the youngest and the oldest.
If we divide the population in four segments spanning 20 years each (0-20, 20-40, 40-60, 60-80 years old, not considering elderly being older than 80), some factors allow us to estimate that 98% of the people now in the first segment will reach the second one, 95% of the people now in the second segment will reach the third one, 90% of the people now in the third segment will reach the fourth one.
Furthermore, the fertility rate of women in the age range 20-40 is 4, while for women aged 40-60 is 1 (suppose that the women account for half of the population).
If these values will remain stable over time, how will evolve the population aged 20-60 (conscriptable for direct military service or indirect efforts favoring the military) during the following generations?
Assume that initially 4 million people are 0-20, 3 million are 20-40, 2 million are 40-60, 1 million are 60 -80; compare also the overall evolution with the findings of the previous Task.

Solution: if the initial vector is (in millions) $v = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$ and the Leslie matrix is

$A = \begin{bmatrix} 0 & 2 & 0.50 & 0 \\ 0.98 & 0 & 0 & 0 \\ 0 & 0.95 & 0 & 0 \\ 0 & 0 & 0.90 & 0 \end{bmatrix}$, then the vector after $20n$ years is $A^n v$.

The conscriptable population after $20n$ years is given by the sum of the two central elements, i.e. the second and the third one, of the vector $A^n v$.
For the first nonzero natural values of n, we obtain

$Av = \begin{bmatrix} 7.0000 \\ 3.9200 \\ 2.8500 \\ 1.8000 \end{bmatrix}$, $A^2 v = \begin{bmatrix} 9.2650 \\ 6.8600 \\ 3.7240 \\ 2.5650 \end{bmatrix}$, $A^3 v = \begin{bmatrix} 15.5820 \\ 9.0797 \\ 6.5170 \\ 3.3516 \end{bmatrix}$, corresponding to 6.7700, 10.5840, 15.5967 millions of conscriptable population after 20, 40, 60 years.

Related theory: Leslie matrices (students can perform the task *before* being instructed on the tool, being it solvable also without matrix algebra, and then again *after*)

Task 3: the following system of interactive components allows for generalizing what developed in the previous Task:

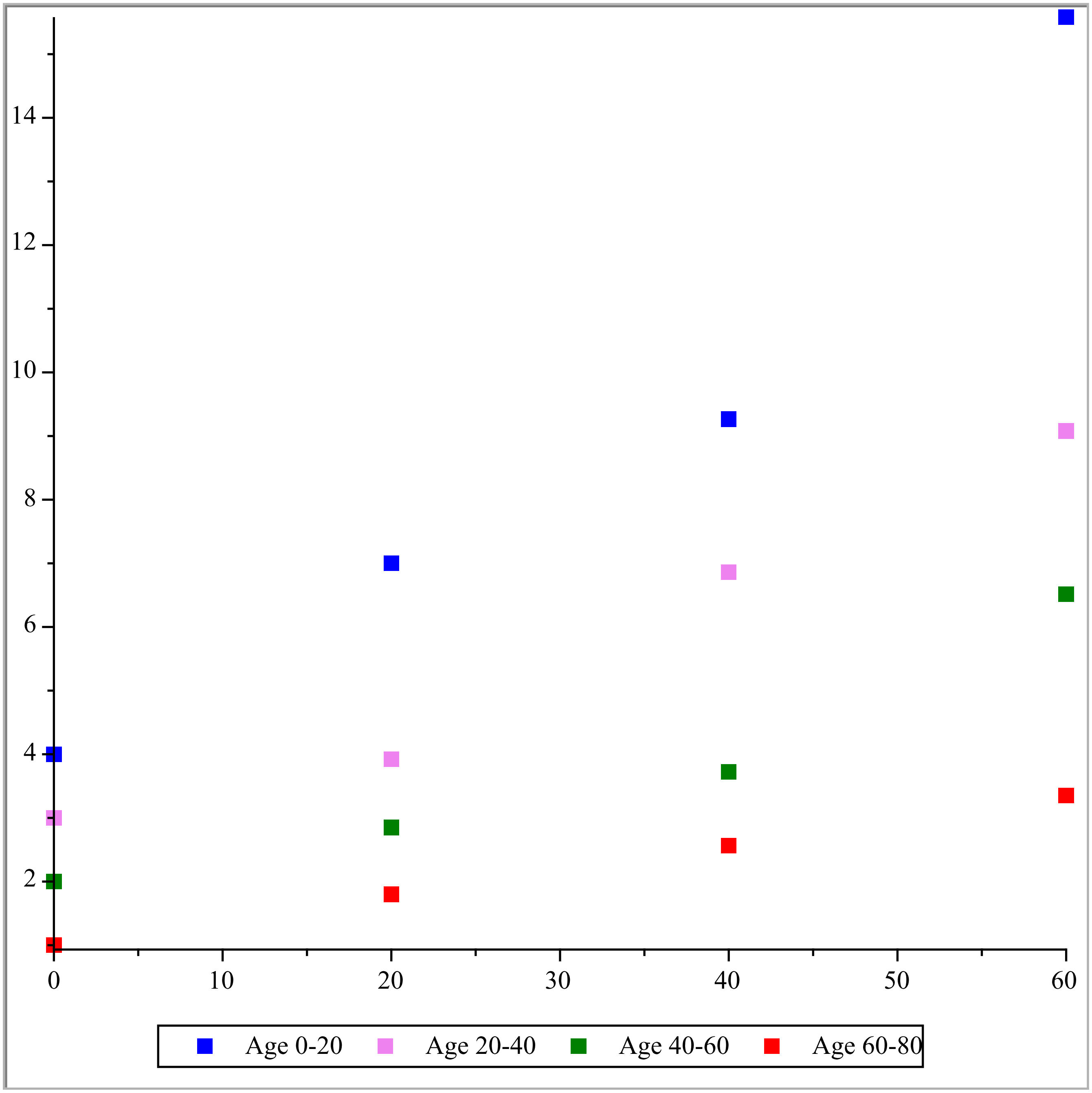
Initial population: [4] aged 0-20, [3] aged 20-40,
[2] aged 40-60, [1] aged 60-80.

Survival rates (%): 98 from 0-20 to 20-40, 95 from 20-40 to 40-60, 90 from 40-60 to 60-80.

Fertility rate: 4 in the age range 20-40, 1 in the age range 40-60.

Compute!

	Initial population	Pop after 20 years	Pop after 40 years	Pop after 60 years
Age 0-20	4	7.	9.2650	15.582
Age 20-40	3	3.9200	6.8600	9.0797
Age 40-60	2	2.8500	3.7240	6.5170
Age 60-80	1	1.8000	2.5650	3.3516



Try to vary the quantities according to proper contexts, and comment the results.

Task 4: the procedure of Task 2 allows us to predict the population evolution on a scale of several generations, but by steps of 20 years.
In order to make yearly predictions, it can be appropriate to interpolate those data, that is to infer how many people will belong to each segment on a certain year, starting from the information available.
How can this interpolation be performed? Try to reason both geometrically (possibly by starting from a dataplot) and algebrically (by finding functions that can be constrained to pass through proper points).

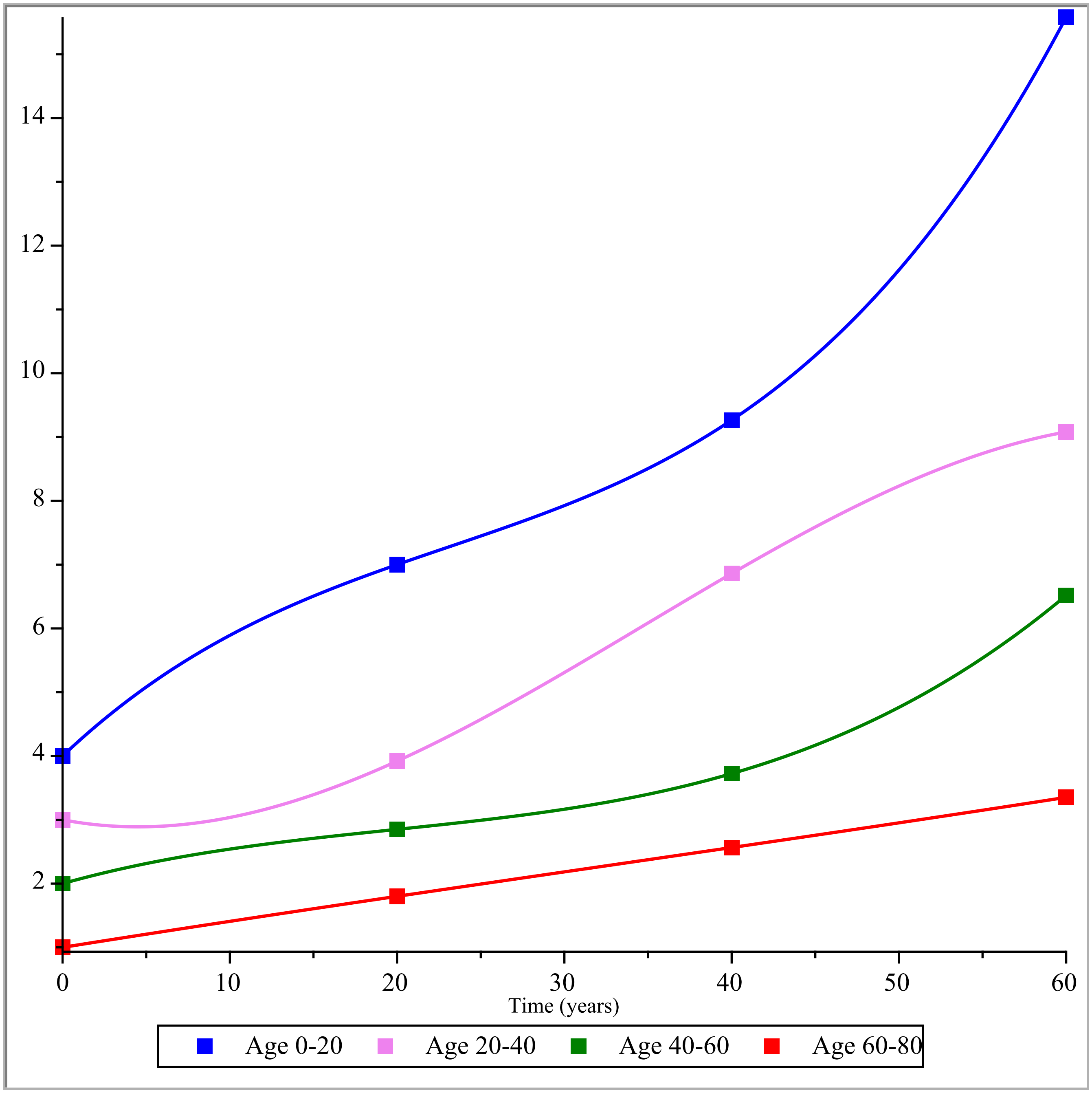
Solution: there are various means to perform such an interpolation, some of the simplest are connecting the points in the time-population plane with segments, or looking for a polynomial passing through them.
Note that using statistical methods like regression with the least squares method, apart from requiring more advanced theoretical tools (proving minimization of the residuals needs calculus/analysis), does not fit with requesting "interpolation". Indeed, in general the regression line or curve does not pass through all the dataset points, thus configuring an "approximation" which is generally not an interpolation.

Related theory: linear spline and polynomial interpolation (students can think of other ways, before knowing those tools)

Task 5: the following system of interactive components uses the data from Task 3 in order to construct noteworthy interpolations of the points:

Load da...

○ Only points ○ Linear spline ◉ Polynomial interpolation



	Age 0 -20	Age 20 -40	Age 40 -60	Age 60 -80
Population after 50 years	11.618	8.2312	4.7622	2.9521

(click again on the interpolation method in order to update)

Try to explore and comment.