DEMOGRAPHY AND TOURISM

LESSON 1 *[Prof. Vignoli]*

In this subject we're going to study demography, but what is demography? Is the statistical study of human population. And what do we call population? It's a group of individuals defined by the same criteria who are together to set a common goal. We study demography related to tourism because they have a strong relationship. In fact demography is the most important factor that influence tourism. Population and the knowledge around it is fundamental for tourism and for investments in that sector. Demographic trends will change the demand for tourism and the available workforces, which will fundamentally impact on how the industries are structured, how they operate, and how they develop in a sustainable manner.

When we go study a population, the things we look are:

- Its SIZE: is the number of total individuals (with no distinction from themselves, young old…)
- Its COMPOSITION: here we look at the characteristics of those individuals (so if they are men, women, young, old…)

*Ex: population of France is everyone living in France. If we look at the **size**, the 1999 census, for instance, recorded 58,518,748 people. The most important characteristics of the population are the **composition** (or structure) by sex and age*

Another basic concepts that we need to know when we're talking about a population are:

- The BIRT COHORT (or generation): that concern all the people born in the same calendar year. But we can talk about cohort also in the case were there are people that experienced the same event within a given time interval [*ex: cohort of people that get married in the same year, or also a cohort f people that graduated the same year…*]

One of the main reasons that demography is used for is to study the changes of a population and of the individuals in general. It's obvious that a population is in a continuous state of change and the demographic analysis studies the mechanisms of that change. But how and why a population change ? what are the driving factors of those changes? → BIRTHS, DEATHS, IMMIGRATION, EMIGRATION. So that means that factors like fertility, mortality and migration influence the population and are the basic factor at the bottom of the concept of 'dynamic population'. And those events that we were talking before can be: REPEATABLE or NON REPEATABLE → for ex death is not repeatable because is unique; while migration and birth are repeatable events in the sense that individuals can experience these events several times during their lifetime. But, at the same time, those repeatable events, if we consider them in a certain way, can become non repeatable too [*from all the birth that a woman can give in this world, if we consider only the first birth, it appear obvious that a woman can have that only once, so in that way we made a repeatable event in a non-repeatable event*].

When we analyse events like the one told before, we generally look two things:

- The INTENSITY of the event: that means how much and how many that events happen
- The TIMING of the event: we look at when that event happen

We are able to apply those measures at all the events except from the death, because we cannot measure the intensity of death [*ex: let's look at the fertility. The intensity of fertility means how many children every woman has in their productive life; while the timing means when, so the average age when every woman get pregnant and has a baby*].

SO, we can resume what we said just in one equation → the **GENERAL POPULATION EQUATION**, that is used to discover how and how much a population change from day x to day y:

$$P_{t+n} = P_t + B - D + I - E$$

B = birth

D = death

I = immigration

E = emigration

- The **Balancing Equation of Population Change** can also be written as:

$$P_{t+n} = P_t + NI + NM$$

- **Population change** = natural increase + net migration:

$$P_{t+n} - P_t = + NI + NM$$
$$\Delta P = + NI + NM$$

NI = Natural Increase *(birth - death)*

NM = Net Migration *(immigration - emigration)*

An example – Italy 2002

	Births 2002	Deaths 2002	Natural Increase	Immigrations 2002	Emigrations 2002	Net Migration	Balance
ITALY	538,198	557,393	-19,195	1,650,961	1,304,438	346,523	327,328

This is a specific example of the data in Italy in 2002, just to understand better the equation.

The equation that we saw up before, count the migration in its general terms. If we are looking to a population – for example – in a specific region inside a country (let’s for example take Tuscany and Italy), we must also consider another data, the internal migration flows (from that region to another place in the county and reverse).

We also have to look at demography not only in a quantitative way, but also in a QUALITATIVE one, because all the events that concern a population effect on demography and on tourism also. So, thing like marriages, divorces have to be consider in the study of demography because they are such important as the quantitative changes and data.

What are we going to study about demography is what is called APPLIED DEMOGRAPHY a specific area of demography where we lead analysis and studies to understand the change of a population, the behaviour, the future prospective and to gain knowledge with the goal to take good decision related to the planning and the strategies in a specific place. Moreover, since we’re studying tourism, all this thing are applied in a tourism sector, in tourism locations. Applied demography is the subfield of demography, focusing on practical applications of demographic materials and methods. We can say that demography is very important to understand, to predict the future in some ways, so ate the end is fundamental to save time and money. So if we want to make an investment in a specific tourist area, the right thing to do is to lead a demographic study in that area to understand the chance of growth that we have, and if we would be able to satisfy the clients. And also the possible earning that we can have from that investment in that area.

LESSON 2 *[Prof. Vignoli]*

So, we said that we study population. But since the population is different, it could be difficult to understand how to study it. So, that's the reason why, in a so heterogeneous population (different sex, color, religion, age, nationality…), we put some limits and some elements to help us study the population in an easier way. We put order by creating some groups in order to make the population more homogeneous. The main groups that are often used to study population are: biology (age), genetics (sex), anthropometrics (wight, height…), and similar. We'll put our focus only on the two most important: SEX and AGE. Historically there are more baby boys than baby girls around the globe. In fact the ratio is 106 baby boys for 100 baby girls. But, in the adult age, the mortality rate for male is higher than for girls; that means that we're going trough a feminization of population in older age. Women live more than men. Why? Because of lifestyle that men have. Men drink more, smoke more, do dangerous job, and in general the lifestyle of a man is more dangerous than a woman lifestyle (also because women are less stupid than men).

We use an indicator to measure the number of men compared to the woman. This is the GLOBAL **SEX RATIO** that tell us the total number of males per 100 females:

Mtot / Ftot * 100 = Global Sex Ratio

Then, if we want to measure this ratio related to a specific age of a human, the formula is the same, but we do not consider the total number of males and females, but of course the number of individuals in that specific age. Equally, the same data could be calculated for the number of births. The formula is the same, but we consider only the births of both sex.

Indicators of Sex Ratio

- The **Global Sex Ratio** is the total number of male per 100 female :

  MTOT / FTOT *100

- In the absence of migration, the **Age-specific Sex Ratio** is the number of men of a given age to the number of women of same age:

  Mx / Fx *100, where *x* represents ages

- The **Sex Ratio at Birth** is the number of male births per 100 female births:

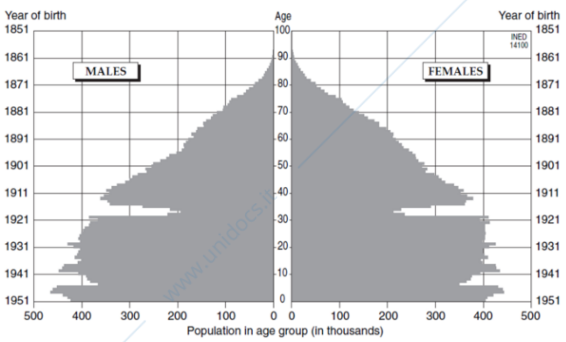
  BM / BF *100

EX: Age-Specific Sex-Ratio for Italy

ITALY 1996	M	F	M/F*100
1	271,540	256,359	105.9
1-4	1,133,319	1,072,120	105.7
5-9	1,427,562	1,358,720	105.1
10-14	1,513,464	1,447,186	104.6
15-19	1,758,695	1,686,151	104.3
20-24	2,198,980	2,120,775	103.7
25-29	2,349,650	2,289,330	102.6
30-34	2,343,271	2,319,915	101.0
35-39	2,050,500	2,047,942	100.1
40-44	1,883,155	1,900,431	99.1
45-49	1,964,182	1,996,214	98.4
50-54	1,669,927	1,726,724	96.7
55-59	1,738,958	1,850,149	94.0
60-64	1,552,371	1,727,238	89.9
65-69	1,417,710	1,683,631	84.2
70-74	1,151,438	1,541,115	74.7
75-79	627,052	955,713	65.6
80-84	498,265	869,031	57.3
85+	305,365	692,882	44.1
total	27,855,384	29,541,603	94.3

This sex-age structure that we use is fundamental in this subject, because a lot of events, demographic or socio economic, depend on this structure. Let's think about pension system, or also the productivity of a country. All these types of data depend on the sex-age of the population. So that's why these two dimensions are so important in demography and in this subject in general.

Population Pyramid, Italy, 1951



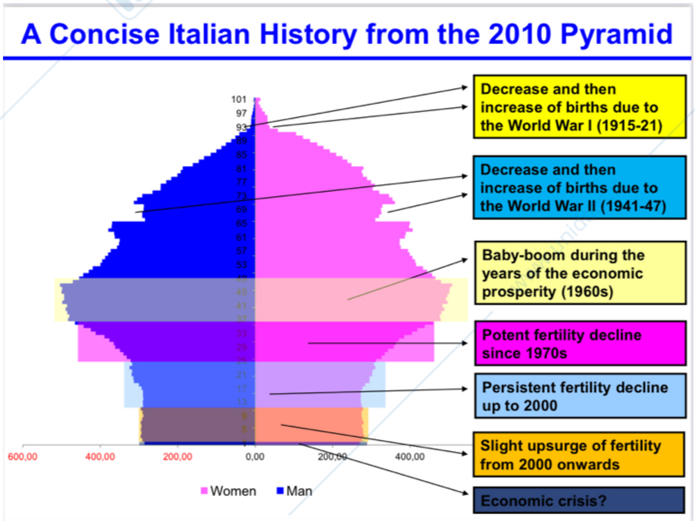
And, based on these two dimensions, we can create a graph, a pyramid graph where we put in relation age and sex of a specific population. → the result is the POPULATION PYRAMID, that is very useful to reveal a lot of phenomena in the population. The population pyramid is constructed by placing side by side the histogram representing the male age structure (on the left) and the one representing the female age structure (on the right).

As we can see, here we have the relation between these two dimensions. And it is obvious that the age of the population is illustrated in groups. This pyramid is very useful also to understand the difference of the same population in two different time period. Or also we can compare two different population. We can clearly see from this graph the process of aging of a certain population (like is happening in Italy). We can also find some shocks in the reproductive cycle in some particular periods. Like in the graph up above, where at the age of 30 y.o. more or less we have a significant contraction. But why? In that specific case, for example, the explanation is that we have less people of 30 due to the first world war. [that graph is from 1951, so the people that are 30 in that year, were born during the war, and it is obvious that during the war there were a lot less births]. Finding this shocks in the cycle of a population is important because, for the future, we'll find again those characteristics. So that contraction in 1951 will appear in all the graphs of the future. So there is a consequence. So if we want to predict the future, starting from that graph, is not too difficult, because we only have to shift up the data.

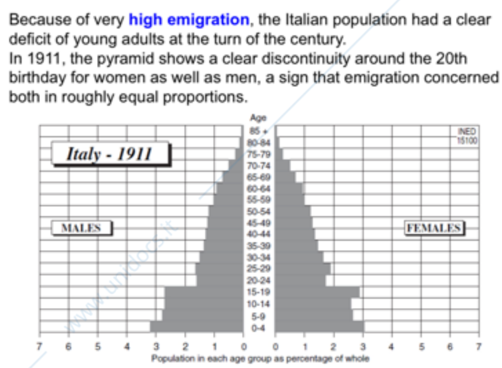
To be more correct, and to be able to compare the pyramids of different countries each other, we need to consider the data in relation to the whole population → the population of each age and sex – to be more precise – should be related to the total population of both sexes combined. So, in that case, we can compare a country with another and be able to say which has a better / worse distribution of population.

The pyramid is so called because especially in the past the graph was a perfect triangle, a perfect pyramid. Now, due to a lot of reason, is no more a perfect pyramid, but the name remains the same.

Overall, population pyramids reflect an image of the current composition of the population by age, with implications for expression of the needs of the different segments of population. But, population pyramids also illustrate the hazard of the history of a population. We can find the fluctuations trace the memory experienced by specific birth cohorts, and also the general shape of the pyramid reflects the major trends that have shaped the course of fertility, mortality, and migration.

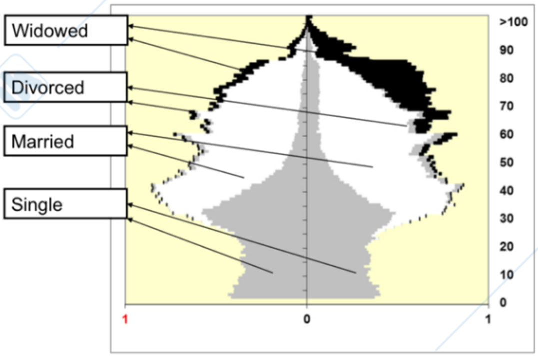


In a pyramid graphic like this, also migration matters. Yes, because we can say that migration is a shock, and an important event in a population, and thanks to this graph we can see the effects of it. Let's see the situation of Italy in 1911, when a lot of young people (about 20/30 yo) decided to emigrate for the United States. That obviously create a lack of population of that age in the data of that year →



In these types of graphics, we can also add a third dimension if we want to analyse a specific event, related to age and sex. An example could be the addition of the third dimension of '*marriages*' In that way we can find how the marriages go in a specific population (related to age and sex):

Florence, 31/12/2004



Let's now talk about the **AGING INDEX** (**AI**): is a ratio that tell us how many old people there are in the population, every 100 young. So, higher this result is, higher the number of old compared to young is in a population. [this ratio in Florence has the result of 230% → for every 100 young boys there are 230 old → this show how our city is getting older and older and how the young are less than the old].

The formula is: population in older age / population in young age x 100 (result is in %)

AGING INDEX:

$$AI = (P_{65+} / P_{0\text{-}14})*100$$

Very similar and related to this ration, is another one, the **TOTAL DEPENDANCY RATIO** (**DR**): what does it means? This ratio tell us the percentage of people that depend on the working class. So, for every 100 of working people in a population, how much is the number pf people that depend on them. So, it is obvious that is difficult to have specific number of people that are working and other there are not. So this ratio in not completely correct. Also because in the group of people that depends on the working class we put the young but also those who don't work anymore. But this is not correct because the young do not have nothing (no goods, no properties, no salary…) while the old people can have a patrimony, houses, pension and similar, so it is not a really good indicator. However, this is the formula, and we van also calculate separately two different dependency ratios (one only for the young generation, and the other only for the older one):

TOTAL DEPENDENCY RATIO

$$DR = [(P_{0\text{-}14} + P_{65+}) / P_{15\text{-}64}]*100$$

Youth Dependency Ratio

$$DR\text{-}y = (P_{0\text{-}14} / P_{15\text{-}64})*100$$

Old-Age Dependency Ratio

$$DR\text{-}o = (P_{65+} / P_{15\text{-}64})*100$$

And last but not least, the **CHILD INDEX** (**CI**) that show us how many babys are there in the population every 100 women (obviously, we calculate only the women in their reproductive age).

CHILD INDEX

$$CI = (P_{0\text{-}4} / {}^{f}P_{15\text{-}44})*100$$

ITALY	1963			1996		
	M	F	Tot	M	F	Tot
Aging Index	32.9	47.2	39.9	92.0	138.9	114.9
Youth Dependency Ratio	38.1	35.5	36.8	22.3	21.0	21.6
Old.Age Dependency Ratio	23.2	27.7	25.6	31.2	38.4	35.1
Total Dependency Ratio	61.3	63.2	62.4	53.5	59.5	56.8
Child Index			38.5			22.1

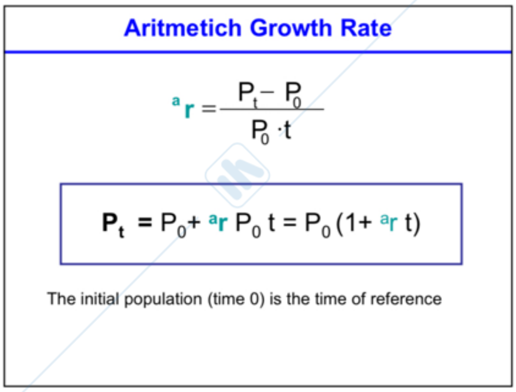
LESSON 3 *[Prof. Pirani]*

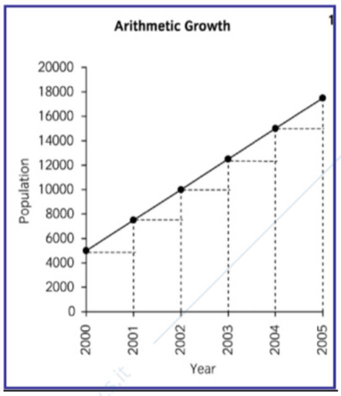
We'll talk about population increase. We'll see the POPULATION GROWT RATE that is the change in a population over time. To be more precise, the change in the number of individuals in a population, using "per unit time" for measurement. We can find the population growth of a period also from the General Population Equation of last time → **ΔP=(B-D)+(I-E)** → in that way we have the change of a population between two different moments. But this formula only gave us an absolute growth. So to be more precise we have to change the perspective and to relativize this ratio, also to be able to compare it with other states.

So that's why we introduce the **GROWTH RATE** (***r***), more specific data, that is used to find the rate at which the number of individuals in a population increases in a given time period as a fraction of the initial population. Usually, this ratio tells us the speed of growth of a population. If we want, we can also express it in percentage.

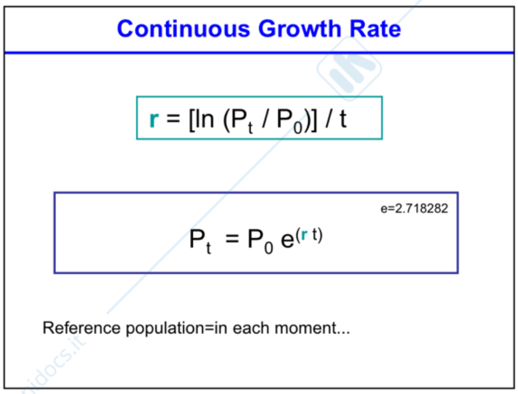
This ratio cane be expressed in two different ways:

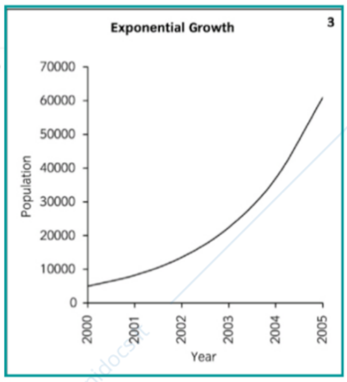
- **Arithmetic →** this is maybe the worse of the two, because is not so realistic and correct. Infact, this suppose that every year the growth of the population is the same, is constant, with no difference from the past. In fact, if we put this formula in the cartesian graph, we can see that is only a straight line. The result of the expression tell us '*how many new individuals annually we have for every 100 individuals in that population'*. The 'time' in the denominator are the years passed between the two different period that we are considering in the numerator.





- **Exponential →** the exponential one could be a better measure especially in those countries where the population is growing faster and not in a constant and continuous way. That because the exponential can give us a different growing perspective, not equal, like it was in the one before. In fact, in a cartesian graph we do not have a straight line anymore, but a classic exponential line. Also in this case the meaning of the result is the same, like the 'time' that we have at the denominator (time passed between population one an population two that we are considering in the numerator)





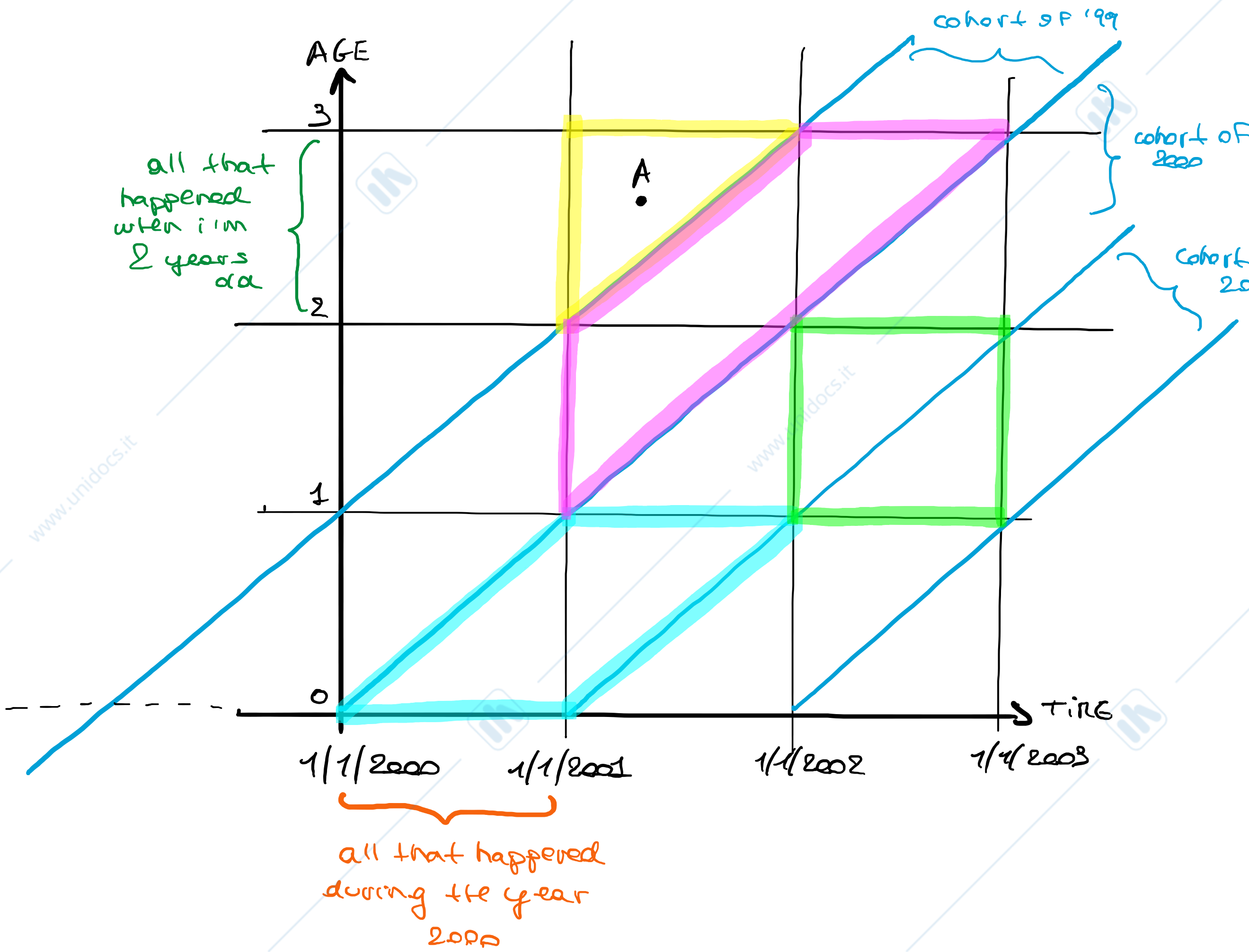
LESSON 4 *[Prof. Vignoli]*

As we also saw in the previous lessons, to study demography all the phenomena of this subject, also related to tourism, we always have to consider there three fundamental dimensions: **AGE**, **TIME**, **COHORT**. Let's have a definition of age → age is the time elapsed since birth; is something that change with the passing of time; it is not something that can go backwards; and we cane measure it in different ways. We always say our exact age (like 20 years old…) but we can express this data also in months, days etc… (for example I can say that I have 22,05 years so 22 years and some months).

About cohort, we can call it also generation, and we already saw the definition of that [the aggregate of individuals belonging to a population who experience the same event within a given time interval]; but the cohort is NOT the same thing of the age.

So when we study an event, or a specific phenomenon, we use a specific graph → the **LEXIS DIAGRAMS**

It is a diagram where all the three most important dimensions are represented. This is used to depict individual trajectories through age and time. So thanks to this graph we can draw a straight line that represents the life of an individual. And thanks to the three dimensions we can understand his life cycle and the events that he has in his life though time and age.



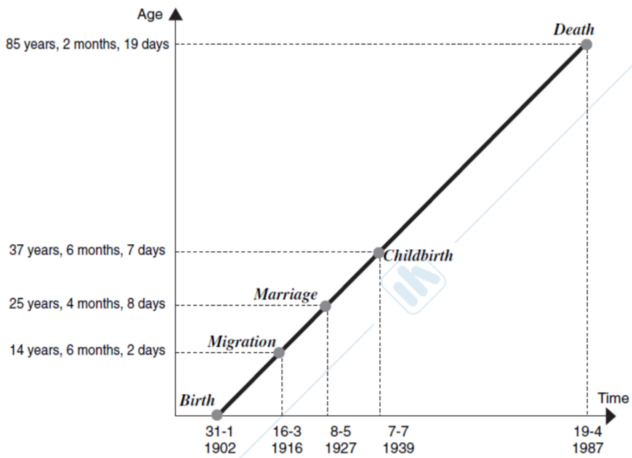
So this is what a **LEXIS DIAGRAM** is about. Three dimension that we use to discover and study some events. So, now, let's put some event into the graph and let's see how to figure it out. Thanks to this graph, if

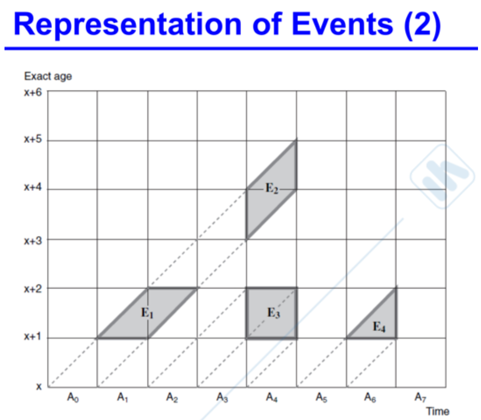
I have an event in a certain point of the life of a person [take for example the point A], I'm able to say when it happened in time and during the life of that person. → the **point A** tis an event that happened to someone who was in that period 2 years old, and the event happened in the year 2001, and the person who experienced that event was someone of the 1998 cohort/generation.

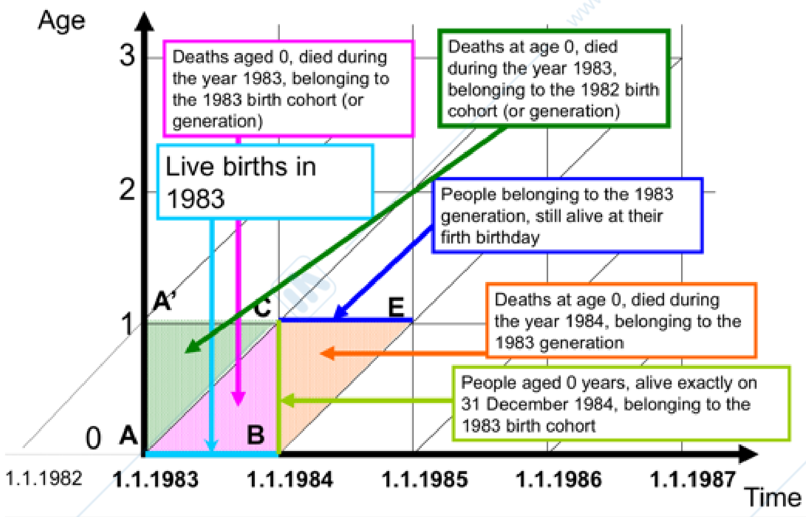
Into that diagram we can find 4 different types of events. Let's see all of them.

- we can have a **tringle** → in this case, if I can put an event into a triangle, it means that I'm able to know all the three dimensions of that event. So I know the year, the age and the cohort of that event
- we can have a **square** →if I have an event inside a square, that means that I know exactly the age and the time of when it happened, but I don't know the cohort who experimented that
- we can have a **normal parallelogram** → with this is similar; I know exactly the cohort and the age of that person, but I'm not able to say when this event happened in time
- we can have a **particular parallelogram** → in this case, I know the cohort that had that event but I do not know either the age or the time when this happened

SO, only if I can put an event inside a **tringle**, that means that I know everything about that event, all the three dimensions. Examples of how this diagram can be used:







LESSON 5 *[Prof. Vignoli]*

Let's now turn back just a little. The General Population Equation is a formula that works with ABSOLUTE NUMBERS but, with them, for us, is impossible to compare the results between each otherSo we cannot compare the data of two states in that case. That's why we need to relativize those numbers and those results to be able to compare them → we need to shift the focus from absolute to relative measures, so that the number of events are considered "relative" to the size of their population → we need **DEMOGRAPHIC RATES**. In demography a rate is the ratio between occurrences and exposures => number of occurrences depends on the size of a population; and the number of occurrences depends also on the time the members of the population are exposed to the risk of occurrence.

So now we introduce the **CRUDE DEMOGRAPHIC RATES** → crude rates identify how many events of a certain type (births, deaths, marriages,…) were recorded during the considered time-span (for instance a year), for each element of the population. We have a lot of crude rates:

	Formula	Definition
Crude Birth Rate ("CBR" or "n")	n= B/P *1000	Number of live births over 1000 individuals, in a specific year
Crude Death Rate ("CDR" or "m")	m= D/P *1000	Number of deaths over 1000 individuals, in a specific year

	Formula	Definition
Crude Marriage Rate ("CMR" or "m")	m= M/P *1000	Number of live marriages over 1000 individuals, in a specific year
Crude Separation Rate ("CSR" or "s")	s= S/P *1000	Number of separations (including divorces) over 1000 individuals, in a specific year

	Formula	Definition
Crude Rate of In-migration ("CRIM" or "i")	i= I/P *1000	Number of immigrants over 1000 individuals, in a specific year
Crude Rate of Out-migration ("CROM" or "e")	e= E/P *1000	Number of emigrants over 1000 individuals, in a specific year

	Formula	Definition
Crude Rate of Natural Increase ("CRNI" or "ni")	ni= (NI)/P *1000	Number of individuals that "adds" to the population by births over 1000 individuals, in a specific year
Crude Rate of Net Migration ("CRNM" or "nm")	nm= (NM)/P *1000	Number of individuals that "adds" to the population by migration over 1000 individuals, in a specific year

	Formula	Definition
Crude Growth Rate ("CGR" or "r")	r= (NI+NM)/P *1000	Number of individuals that "adds" to the population over 1000 individuals, in a specific year

As we can see, in all the formulas at the denominator we have the population. But, which population? We always have to do an average between the data of the population that we have from the 1/1/of the year that we are considering and the data of the population that we have the 1/1/of the next year → doing that we put an average at the denominator, and this is better for the calculation.

Another important thing of these rates is that they are too crude, in the sense that they don't consider the age, that as we saw is a fundamental dimension in our demographical studies. So we have to consider it because it is an important variable to describe how demographic behaviour works.

SO, in that way we have the **AGE-SPECIFIC DEMOGRAPHIC RATES** → Age-Specific Rates identify how many events of a certain type (births, deaths, marriages,…) were recorded at a certain age x, during the considered time-span (for instance a year), for each element of the population having that specific age. So all the formulas above remains the same, we only have to put in them the specific age that we want to analyse.

	Formula
Crude Death Rate	$m = (D/P)*1000$
Age-Specific Death Rate	$m_x = (D_x / P_x)*1000$ for every **x**

← This is an example with the death rate, but is like this with all the others.

And, obviously, also in that case of the specific age we have to consider, at the denominator, the average of the population, exactly as before.

We saw before the crude birth rate, and similar to that there is the **CRUDE FERTILITY RATE**, that show us how reproductive is a particular population. And to calculate that we have to consider at the denominator all the women in their reproductive age (15-49 years old). Also in that case, this do not consider a specific age, so if we want to, we have to use another specific ratio.

	Formula
Crude Birth Rate	$n = B / P *1000$
Crude Fertility Rate (or General Fertility Rate)	$GF = B / {}^{f}P_{15\text{-}49} *1000$

The **AGE SPECIFIC FERTILITY RATES**, that use at the numerator the number of births that all the women at that specific age that we are considering had, and at the denominator the average of the population of the females at the age we are looking at. The result of this formula tell us the number of children per woman of a certain age. And starting from the formula before, we can find the **TOTAL FERTILITY RATE** that is only the sum of all the specific fertility rates (obviously the sum that we do in that case is for the rates calculated on the reproductive age of women). → the total rate tell us the number of children per women. And due to the fact that the total fertility rate now consider all the specific age, this let us to compare the results of the countries all around the globe.

	Formula
Total Fertility Rate (TFR)	$TFR = \Sigma N_x / {}^{f}P_x$ $= \Sigma f_x$ → AGE SPECIFIC FERTILITY RATE for every **x**

NOTE: Generally *fx* are computed and summed up for women aged 15 to 49.

U.S. Census Bureau

IDB Summary Demographic Data for Italy

Demographic Indicators: 2005 and 2025

	2005	2025
Births per 1,000 population....................	9	7
Deaths per 1,000 population....................	10	13
Rate of natural increase (percent).............	-0.1	-0.5
Annual rate of growth (percent)................	0.1	-0.3
Life expectancy at birth (years)...............	79.7	81.9
Infant deaths per 1,000 live births............	6	4
Total fertility rate (per woman)...............	1.3	1.5

Midyear Population Estimates and Average Annual Period Growth Rates: 1950 to 2050
(Population in thousands, rate in percent)

Year	Population	Year	Population	Period	Growth Rate
1950	47,105	2005	58,103	1950-1960	0.6
1960	50,198	2006	58,134	1960-1970	0.7
1970	53,661	2007	58,148	1970-1980	0.5
1980	56,451	2008	58,145	1980-1990	0.1
1990	56,743	2009	58,126	1990-2000	0.2

← This is an example of all the ratio that we've seen in this lesson, in our country.

LESSON 6 *[Prof. Vignoli]*

We've seen all the crude rate that we have in this subject. BUT, as always, there is a problem. These crude rates are good to analyse a specific population and all the data concerning a specific place, so if we want to compare different countries themselves, we CANNOT use the crude rates. Because, in this case, they are too crude. That's why we have to do the so called **DIRECT STANDARDIZATION** → this is a process thet can help us to compare the data of two (or more) countries. So, when – for instance – I want to compare the death rate of Italy and the death rate of France, in this case I have to change the crude rates of the single countries, using the direct standardization, so I'm able to compare the two rates. But HOW? We'll see:

Starting from the Crude Death Rate formula, I have this situation (let's take Italy and France as the two countries I want to compare):

$$m_x = \frac{D_x}{P_x} \Rightarrow D_x = m_x \cdot P_x \quad \Big\} \quad m_x = \frac{\sum_{x=0}^{\omega} m_x \cdot P_x}{\sum_{x=0}^{\omega} P_x}$$

So, when I have to find the total mortality rate of a single country, the one above is the formula I have to calculate, with the sum of all the specific age that I have. In this case $\mathbf{P_x}$ is the number of individuals in the population, in that specific age (population of the country that I'm looking at, in that moment).

THEN, if I compute this calculation for both Italy and France, I will have the death rates of the two countries. BUT I can't compare themselves. And to do that I have to do now the DIRECT STANDARDIZATION, by using the SAME FORMULA, but both in the numerator and the denominator (for both countries) I have to use the SAME POPULATION. It doesn't matter which, but I have to relativize both the formulas to the same population, that can be whatever I want.

through **DIRECT STANDARDIZATION** we just weight differently the age-specific rates, not using the weights of the real population but using the weights of another population chosen as standard.

$$\overline{m}_x^{ITA} = \frac{\sum_{x=0}^{\omega} m_x^{ITA} \cdot P_x^{EU}}{\sum_{x=0}^{\omega} P_x^{EU}} \qquad \overline{m}_x^{FRA} = \frac{\sum_{x=0}^{\omega} m_x^{FRA} \cdot P_x^{EU}}{\sum_{x=0}^{\omega} P_x^{EU}}$$

Here it is the standardization process for both Italy and France → as we can see, we took the original ones, and only changed the **population**, that is the same in both the standardized formulas → we use the same population in the two formulas just to be able to compare the results themselves, and be able to say with of the country has the higher mortality rate. → the choice of the standard population *is completely arbitrary*, and therefore different choices of weights will yield different results.

Usually, we DO NOT USE direct standardization for fertility, because that does not make sense. Direct standardization process is used ONLY TO COMPARE DIFFERENT MORTALITY RATES.

LESSON 8 *[Prof. Vignoli]*

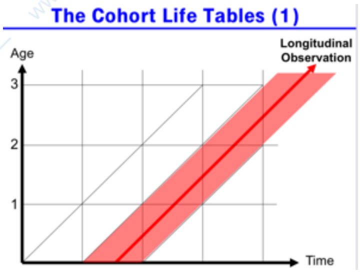
Today, we'll study life expectancy, using the life tables. So let's see them.

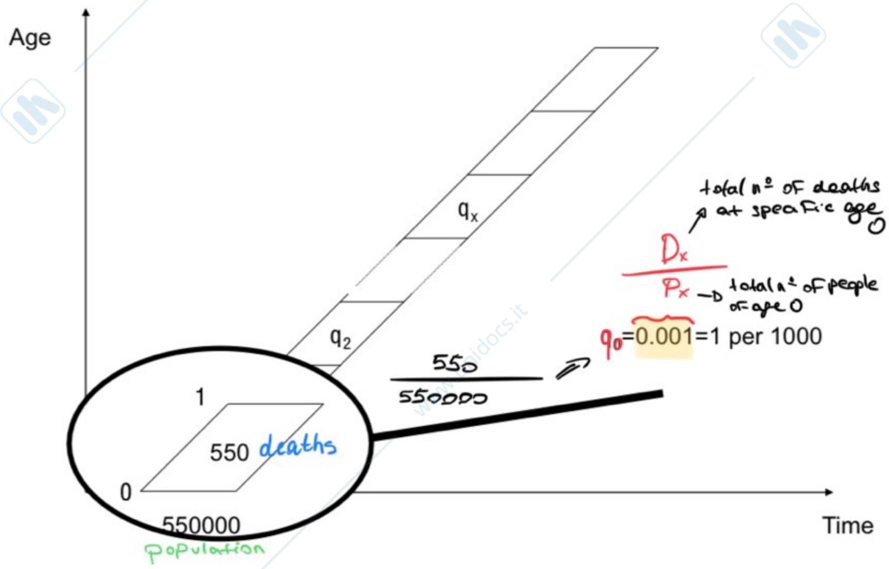
A life table is the study of the distribution of the time it takes a cohort of specific individual to experience an 'event'. Usually, we use **LIFETABLES** to see and measure the gradual extinction of a cohort, whose initial population is reduced by death from one age to the next, until its total disappearance → in the end, we will be able to describe the duration of life, synthesized in the measure called '**LIFE EXPECTANCY**' [that is the average number of years lived by a newborn, so how much is expected, that newborn, to live in that population].

The life table was originally devised by *John Graunt* (1662) using 'bulletins', containing deaths by age in London → in the London of his time he observed that starting from 100 births, about 36 died before age 6, and only one person reached age 76.

But how a life-table works and how can we fill it? → obviously, we need to start from the real data of the population (the age-distribution of population and deaths), because from that we can compute the '***probability of dying***' [q_x], and from this ratio we can find all the other data that we need to fill the table.

So, since we are looking at a specific cohort, in the lexis diagram this is translated by looking the graph with a *longitudinal perspective*. And thanks to the data and the lexis diagram, we can compute immediately the **AGE SPECIFIC PROBABILITY OF DYING** [q_x] that is the total number of death for that specific age divided by the total number of people (population) of that specific age.





Once we've found the probability of dying, as said before, we're able to compute all the others data that we need in the life-table, even the life expectancy of that cohort. For example, it is obvious that, if the probability of dying is q_x we can easily find the '***probability of surviving***' [p_x], that is the complement [$1 - q_x$] → and with the symbol l_x we represent the exact number of survivors (those of that cohort who did not die).

The life table rescales the initial size of the cohort to a round number called '***radix of the table'*** [l_x] → comparison between cohorts is then by default feasible. The radix represents, in the life table, the number of members of the cohort alive at age *x*. Usually, it is a power of 10 (1,10,1000,10000,100000…).

The tempo of deaths in the table is, in particular, the factor that enables us to summarize a cohort's mortality by determining its life expectancy at birth (average number of years lived by a newborn). *In order to obtain life expectancy we need another set of quantities*, either from data or from additional hypotheses. How many

<u>years do people dying between x and x+1 live?</u> → a_x is the **average number of persons-years lived** by individuals who die between age *x* and age *x+1* → we can thus define the **total persons-years lived** [L_x] by individuals between age x and age x+1.

a_x is the average number of persons-years lived by individuals who die between (exact) age *x* and (exact) age *x+1*

We can thus define the (total) **persons-years** lived by individuals between (exact) age *x* and (exact) age *x+1*: $L_x = l_{x+1} + a_x \cdot d_x$

The **total number persons-years** lived above age *x* and (exact) age *x+1* is: $T_x = L_x + L_{x+1} + \ldots + L_{\omega-1}$

$$T_x = \sum_{j=x}^{\infty} L_j$$

The sum of all the L_x for all the specific ages we're considering is called the **total number person-years** lived above age x and age x+1 [T_x].

So, at the end of all these formulas, we can finally compute the **LIFE EXPECTANCY** → the <u>AVERAGE NUMBER OF YEARS LIVED BY INDIVIDUALS WHO ARE ALIVE AT AGE X</u> [e_x]

Life expectancy at age x is the **average number of years lived by individuals who are alive at (exact) age x**:

$$e_x = \frac{T_x}{l_x}$$

Now, we're finally able to complete all the **LIFE-TABLE**:

We Do Have the Cohort Life Table!

$d(x,x+1) = l_x \cdot q_x = 1000 \cdot 0{,}1974$

$T_x = L_x + \ldots + L_{w-1} = 9013{,}2 + \ldots + 1{,}8$

$l_{x+1} = l_x \cdot p_x = 1000 \cdot 0{,}8026$

$p_x = 1 - q_x = 1 - 0{,}1974$

$L_x = (l_x + l_{x+1})/2 = (10000+8026)/2$

$e_x = T_x / l_x = 37714/1000$

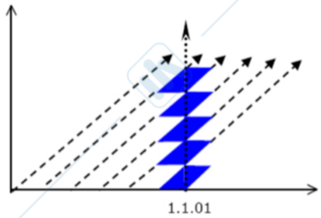
Age	D(x,x+1)	P_x	q_x	l_x	p_x	d(x,x+1)	L_x	T_x	e_x
⋮									
85	18760	95052	0.1974	10000	0.80263435	1974	9013.2	37714.0	3.8
86	15923	76292	0.2087	8026	0.79128873	1675	7188.7	28700.8	3.6
87	13167	60369	0.2181	6351	0.78189137	1385	5658.5	21512.1	3.4
88	10904	47202	0.2310	4966	0.76899284	1147	4392.3	15853.5	3.2
89	9124	36298	0.2514	3819	0.74863629	960	3338.8	11461.2	3.0
90	7106	27174	0.2615	2859	0.73850004	748	2485.1	8122.4	2.8
91	5571	20068	0.2776	2111	0.72239386	586	1818.2	5637.3	2.7
92	4336	14497	0.2991	1525	0.70090364	456	1297.1	3819.1	2.5
93	3206	10161	0.3155	1069	0.68447987	337	900.3	2522.0	2.4
94	2331	6955	0.3352	732	0.66484543	245	609.1	1621.7	2.2
95	1636	4624	0.3538	486	0.64619377	172	400.4	1012.6	2.1
96	1191	2988	0.3986	314	0.60140562	125	251.7	612.2	1.9
97	693	1797	0.3856	189	0.61435726	73	152.6	360.5	1.9
98	460	1104	0.4167	116	0.58333333	48	91.9	207.9	1.8
99	269	644	0.4177	68	0.58229814	28	53.6	115.9	1.7
100	172	375	0.4587	39	0.54133333	18	30.4	62.3	1.6
101	94	203	0.4631	21	0.53694581	10	16.4	31.9	1.5
102	51	109	0.4679	11	0.53211009	5	8.8	15.5	1.4
103	34	58	0.5862	6	0.4137931	4	4.3	6.7	1.1
104	13	24	0.5417	2.5	0.45833333	1	1.8	2.4	1.0
105	11	11	1.0000	1.2	0.00	1	0.6	0.6	0.5

Most of the times, we hear about "life expectancy in 2005" or "life expectancy in 2003-2004"

This information comes from **Period Life Table**

Main differences from cohort life tables:

- **Input data**
- **Interpretation**



The one saw before, is a COHORT life-table, but if – for example – we want to compute the life expectancy of a population in a specific year (life expectancy in Italy in 2003), we need to use a PERIOD life-table → in that case we do not need to follow the lexis diagram with a longitudinal point of view, but with a vertical point of view, as in this (←) situation.

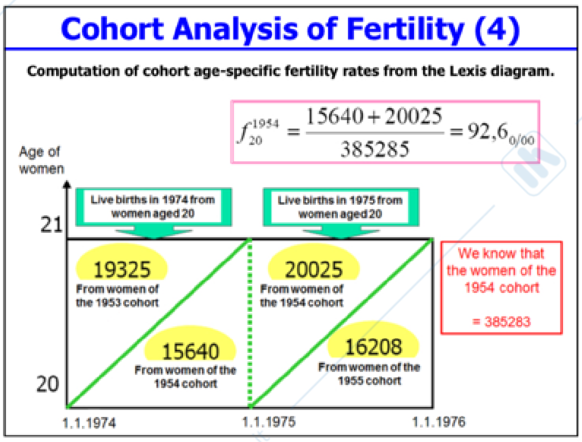
LESSON 9 *[Prof. Vignoli]*

Fertility is a very important process, also more than death rates, because is the one that, together with migration, is what make a population grow → is a key driver of population dynamics. A whole set of measures has been devised for fertility considering birth a "repeatable event". This phenomenon has a problem: it can't be studied for men, but only for women. So when we talk about fertility, we always refers to women and their ability to have children in their reproductive age (that goes from 15 years old to 49). When we talk about fertility, differently to mortality, is that we have to consider not only the timing of the phenomenon, but we have to look also at the intensity of it → that means that we're able to respond at these two question: ***How many children per women?*** & ***Which is the mean age at childbearing?***

As for the mortality, if we want to measure this phenomenon in a cohort point of view, we have to look at the lexis diagram with a longitudinal perspective. But, if we want to compute the total fertility that a specific cohort had during its history, we need to wait until all the women of that cohort is no longer fecund → only in that moment, based on the number of live-born children they have had during their lifetime, we can compute the total fertility rate of that cohort → we can define the **AGE-SPECIFIC FERTILITY RATE** at age x for those women interviewed at the end of their childbearing period (→); this is the same formula of the 'normal' fertility rate, with the only difference that is the denominator (that in this case is locked, and represent the total number of women that are no longer fecund).

$$f_x = B_x / F_{55}$$

Example of computation:



The **sum of the rates over all *x***, from puberty to menopause, will yield the **average number of children born** at the end of the childbearing period.

This **cohort total fertility** rate (**TFR**), or completed fertility rate (**CFR**) as it is sometimes called, represents the **intensity of fertility** and can be written:

$$TF = \sum_{x=15}^{54} \frac{B_x}{F_{55}}$$

So, thanks to that formula, I know the ***average number of children that women had*** during their reproductive ages. And this give me the quantity and intensity of the phenomenon.

Meanwhile, if I want to know the timing of that, I have to compute another calculation → The tempo of fertility is done by the **MEAN AGE** at **CHILDBEARING** (**MAC**).

$$\bar{x} = \frac{\sum_x x \cdot f_x}{\sum_x f_x}$$

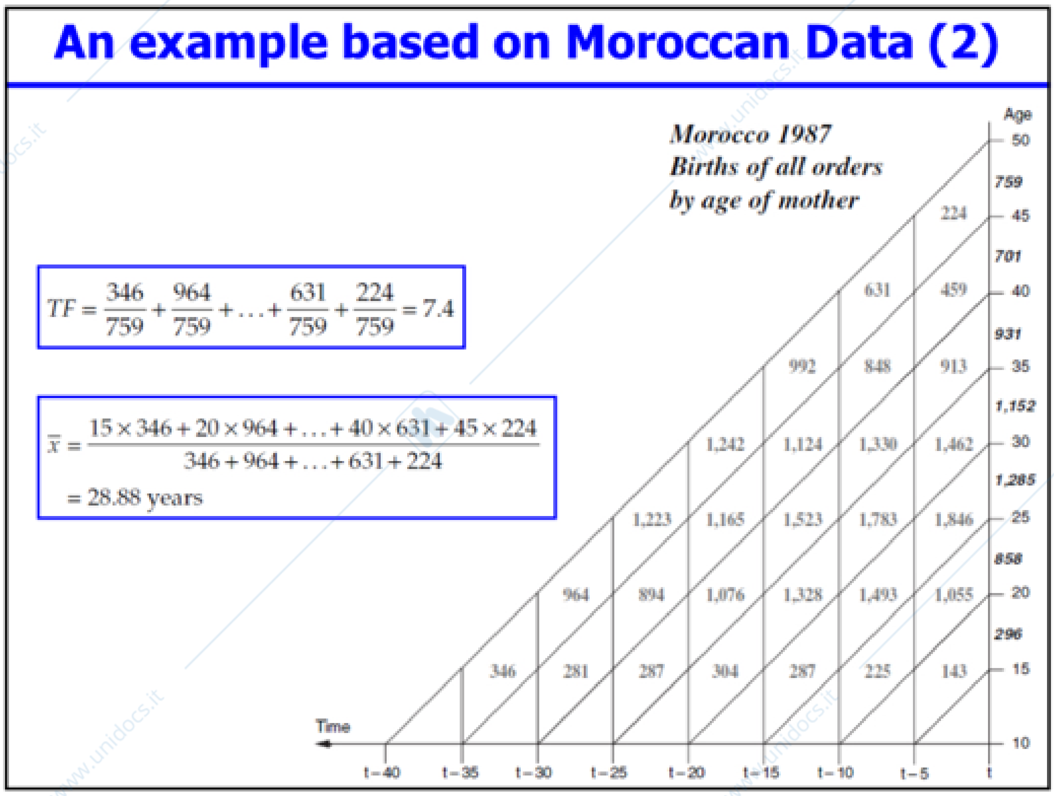
(x → AGE) remember that $f_x = \frac{\text{N° BIRTHS AT THAT AGE}}{\text{N° FEMALE OF THAT AGE}}$

If we are watching at this phenomenon at the end of the reproductive age – for the women – then the formula is like this:

$$\bar{x} = \frac{\sum x.B_x}{\sum B_x} \text{ or } \bar{x} = \frac{\sum x \frac{B_x}{F_{55}}}{\sum \frac{B_x}{F_{55}}} = \frac{\sum x f_x}{\sum f_x}$$

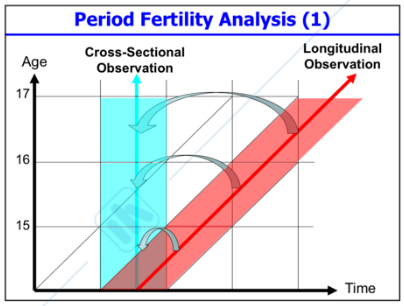
(x → age; B_x → BIRTHS)

This is a perfect and real example of the formulas we wrote before, based on Moroccan population in 1987. When we compute the calculation about fertility rates, we always have a "retrospective" view. Because, from the end of the fecundity of the women, we look at their past history, to understand how their reproductive life went:



→ N° OF WOMAN AT THE END OF THE REPRODUCTIVE AGE

→ AGE OF WOMEN



We've studied fertility from a cohort point of view, but often, in exercises and in the statistic data of countries, this is shown from a period point of view. To monitor current trends, we need to analyze the fertility, looking year per year → this will provide us a description of the behavior of a synthetic cohort of women. So, what we need to do is to change our perspective, from a longitudinal to a cross-sectional one. When we want to compute the f_x, now we do not need to take the data from the red line, but from the blue one. The formula is the same, with the births at the numerator and the female population [*average*] at the denominator → but when we need the **TOTAL FERTILITY RATE** from a **PERIOD** point of view, the sum we need to do is different. We need to consider all the blue squares and consider all the ages of the female childbearing.

Period Fertility Analysis (3)

If $f_{x,t}$ is the age-x-specific fertility rate (ASFR) in year t, the period **total fertility rate (TFR)** for the year t is, where α and w are the extreme ages of the female childbearing period:

$$TFR_t = \sum_{x=\alpha}^{\omega} f_{x,t}$$

Pay attention, because in this formula we also need to consider the size of the age groups. Usually, we consider ages of women in groups of five (15-19 / 20-24…), and when we do this - in the formula at the left – we have to introduce this size, how? → multiplying the sum by '*a*' that is the size of the group. In the case the size is the same, this is not a problem, we can multiply the final result. But if the size of the groups is always different, we have to multiply every single f_x by his specific size *a*.

Naturally, if we are working on age groups, we must multiply each rate by the dimension of the age interval *a*:

$$TFR_t = \sum_{x=\alpha}^{\omega} a f_{(x,x+a),t}$$

If *a* is the same for all age groups, we obtain:

$$TFR_t = a \sum_{x=\alpha}^{\omega} f_{(x,x+a),t}$$

The **period mean age at childbirth** (when?)

$$\bar{x} = \frac{\sum_{x=\alpha}^{\omega} x \cdot f_{x,t}}{\sum_{x=\alpha}^{\omega} f_{x,t}}$$

If we work with 5-year age groups, we have:

$$\bar{x} = \frac{\sum_{x=\alpha}^{\omega} (x+2.5) f_{(x,x+5),t}}{\sum_{x=\alpha}^{\omega} f_{(x,x+5),t}}$$

As for the cohort, also here we have a formula for the mean age → more in detail, the **PERIOD MEAN AGE AT CHILDBIRTH**, give us the measure of the tempo of our phenomenon (←).

In this case, the *x* represent the age we are considering, and in the case of an age group -for example of five – as we can see we add 2.5 just to take an average age inside the group of five (that's why we add 2.5 at the beginning age of the group we are considering).

CONCLUSIONE PARTE I del Professor VIGNOLI – INIZIO PARTE II della Professoressa PIRANI