

# Using Multivariate Statistics to Describe the Demographic Behaviour of Europeans

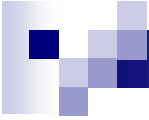
**Mare Vähi**, University of Tartu

**Ene-Margit Tiit**, Statistics Estonia,  
University of Tartu



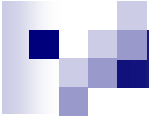
# Using statistics in demography

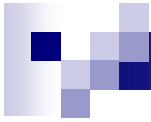
- The demographic behaviour of the world population has always been a challenging topic for statisticians.
- The first scientific forecasts in the history of social sciences were made to predict the world population size.
- Although the methodology of population projections has enormously developed, the usage of multivariate statistics in demography is not very common.

- 
- The aim of our paper is to use some traditional methods of multivariate statistics to describe the behaviour and dynamic of the world, but especially European population.
  - As data we use the official data published by Eurostat and UN, including also population projections by UN.



# Behaviour and dynamics of world population

- 
- Recently UN published the report on demographic data of the whole world, 228 countries. The data were given for years 2000 and 2005, some also for past (1950 and 1970/75) and most of them were forecasted values for up to the year 2050.
  - The list of indicators given in the report includes the crude age distribution, life expectancy, infant's and children's mortality, TFR and some indicators of population increase. For some countries also the some data about prevalence of HIV are given.



- It is evident that the demographic behavior of a population depends on multiple factors characterizing natural and social environment, cultural (including religious) traditions, economical, historical and political situation at given time and given place.
- We tried, using only demographic data, to discover the demographic behavior patterns and types that characterize different regions/ continents and to see their changes in time.



# Factor Analysis of demographic characteristics

- We used data from the period 2000—2010 (the last ones are projections).
- As research unit we used a country.
- In general, 16 variables were used, where absolute population sizes were replaced by the distribution of four main age groups .
- As a result we were able to extract 3 independent factors (further: d-factors) with summary description rate 93%, see the following Table.

	Fert_InfMort	Incr_Age Struc	Life_exp
TFR_00_05	0,827	0,504	-0,193
TFR_05_10	0,847	0,465	-0,189
InfMort_00_05	0,765	0,417	-0,42
InfMort_05_10	0,779	0,39	-0,415
Child5Mort_00_05	0,803	0,364	-0,422
ChilMort_05_10	0,809	0,341	-0,419
Share_15_59	-0,904	-0,115	0,126
Share_60_79	-0,269	-0,931	0,224
Share_80+	-0,145	-0,914	0,313
Medage_05	-0,485	-0,813	0,251
Share_00_14	0,637	0,706	-0,232
Increase_00_05	0,613	0,685	0,103
Increase_05_10	0,665	0,696	0,081
HIV_2005	0,114	0,061	-0,843
Life_exp_00_05	-0,633	-0,409	0,635
Life_exp_05_10	-0,62	-0,389	0,665





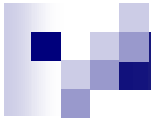
# Interpretation of factors:

- Fertility—children and infants' mortality factor (*Fert\_InfMort*), having high scores in countries, where fertility and also infant mortality and children mortality are high;
- Population increase—age structure factor (*Incr\_AgeStruc*). The factor has positive scores in countries with rapid increase of population and negative scores in ageing, older, stagnated population;
- Life expectation factor (*Life\_exp*), depending mainly on life expectation, but also on prevalence of HIV in the country and having positive scores in countries where life expectation is large.

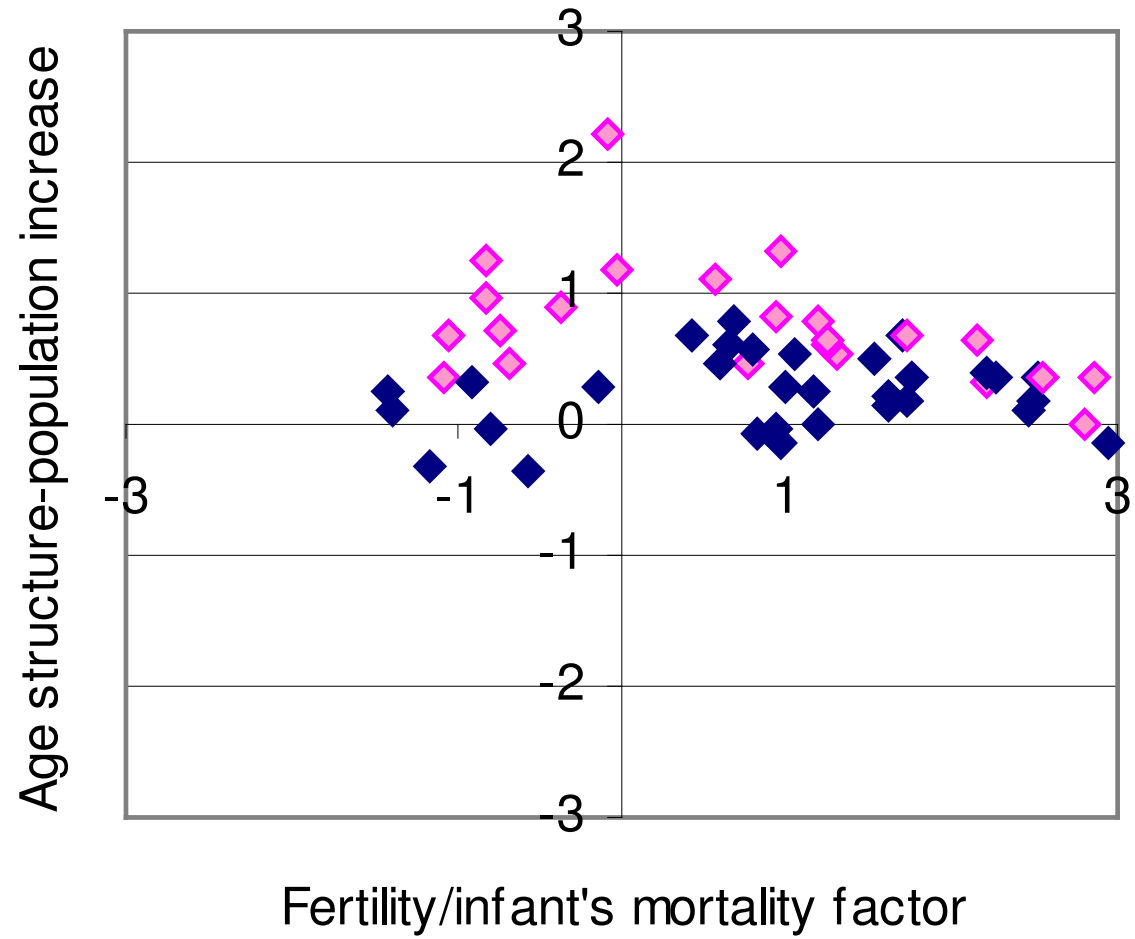


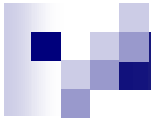
# Countries of different continents in the 3-dimensional space of d-factors

- Each country is presented as a point on the graph.
- Horizontal factor is the 1st factor. Big values (right-hand) show high fertility and children mortality.
- Vertical factor is the 2nd factor. Big values (above) show young, rapidly increasing population.
- The third factor is indicated by color: dark points show countries with considerably low, light – with rather long life expectation.

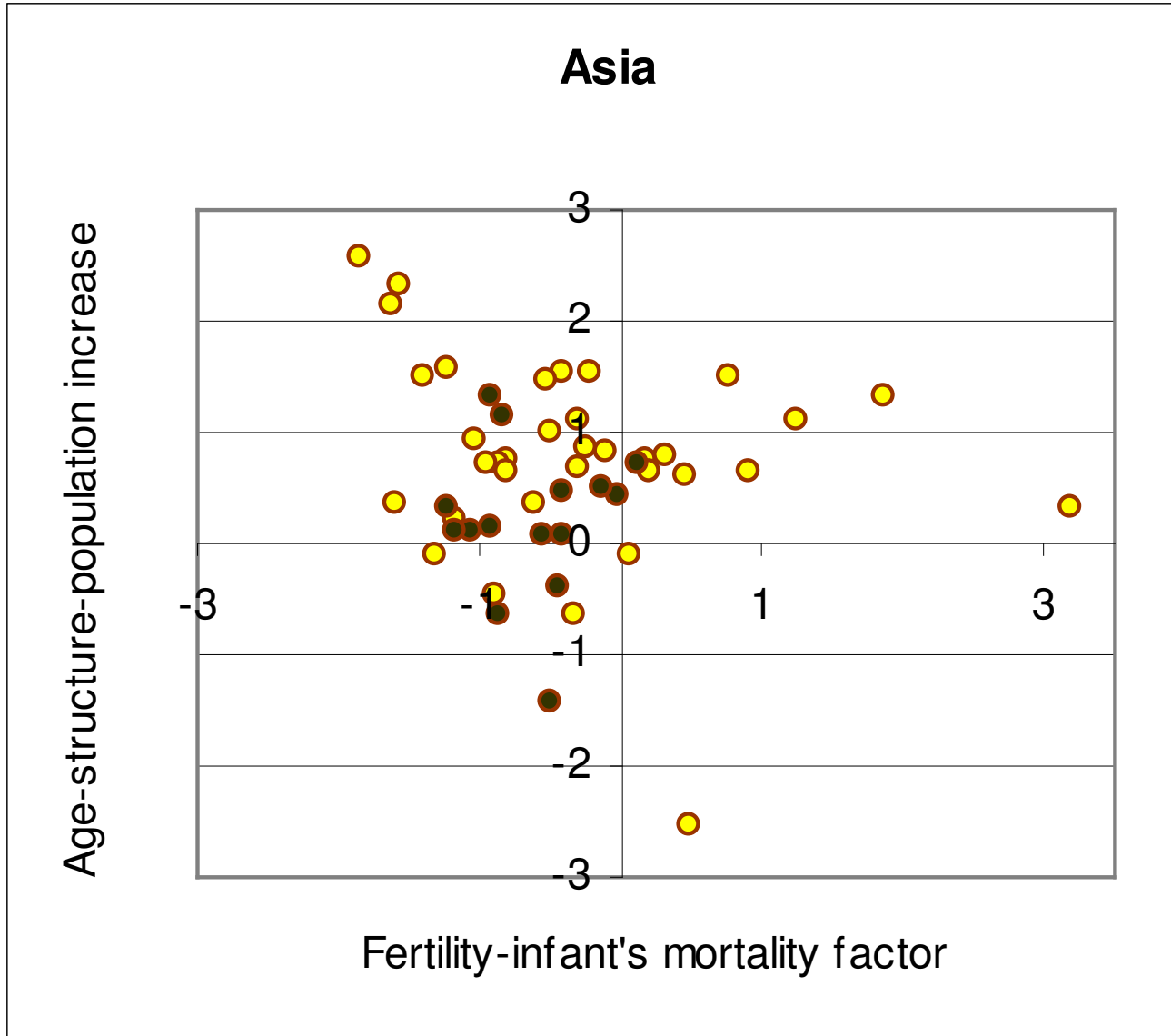
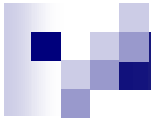


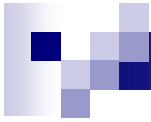
## Africa



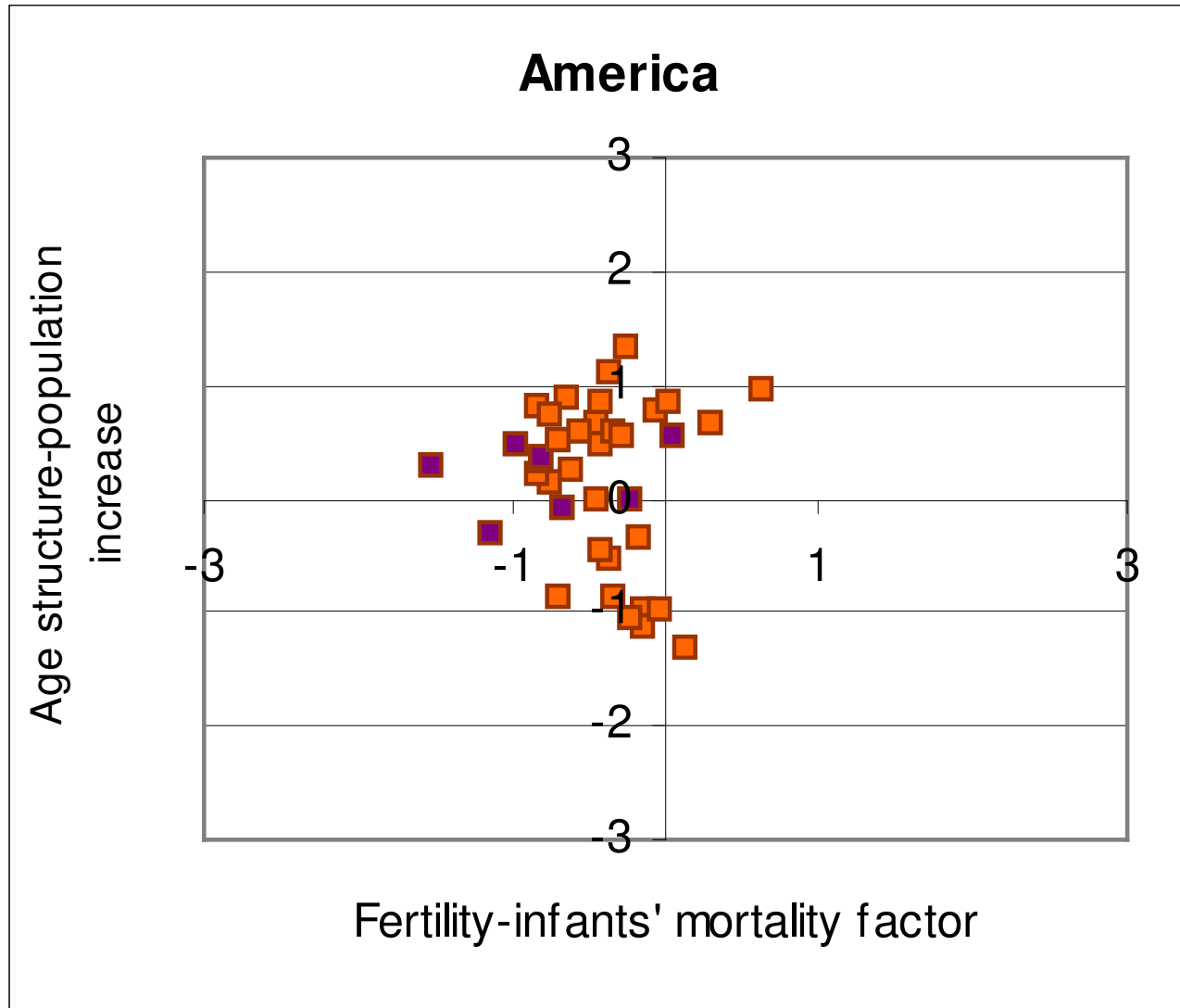
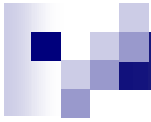


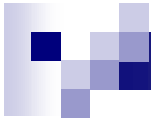
- In Africa the population is young.
- Fertility, but also children mortality are high (but here are quite big differences).
- Life expectation is in most countries low.





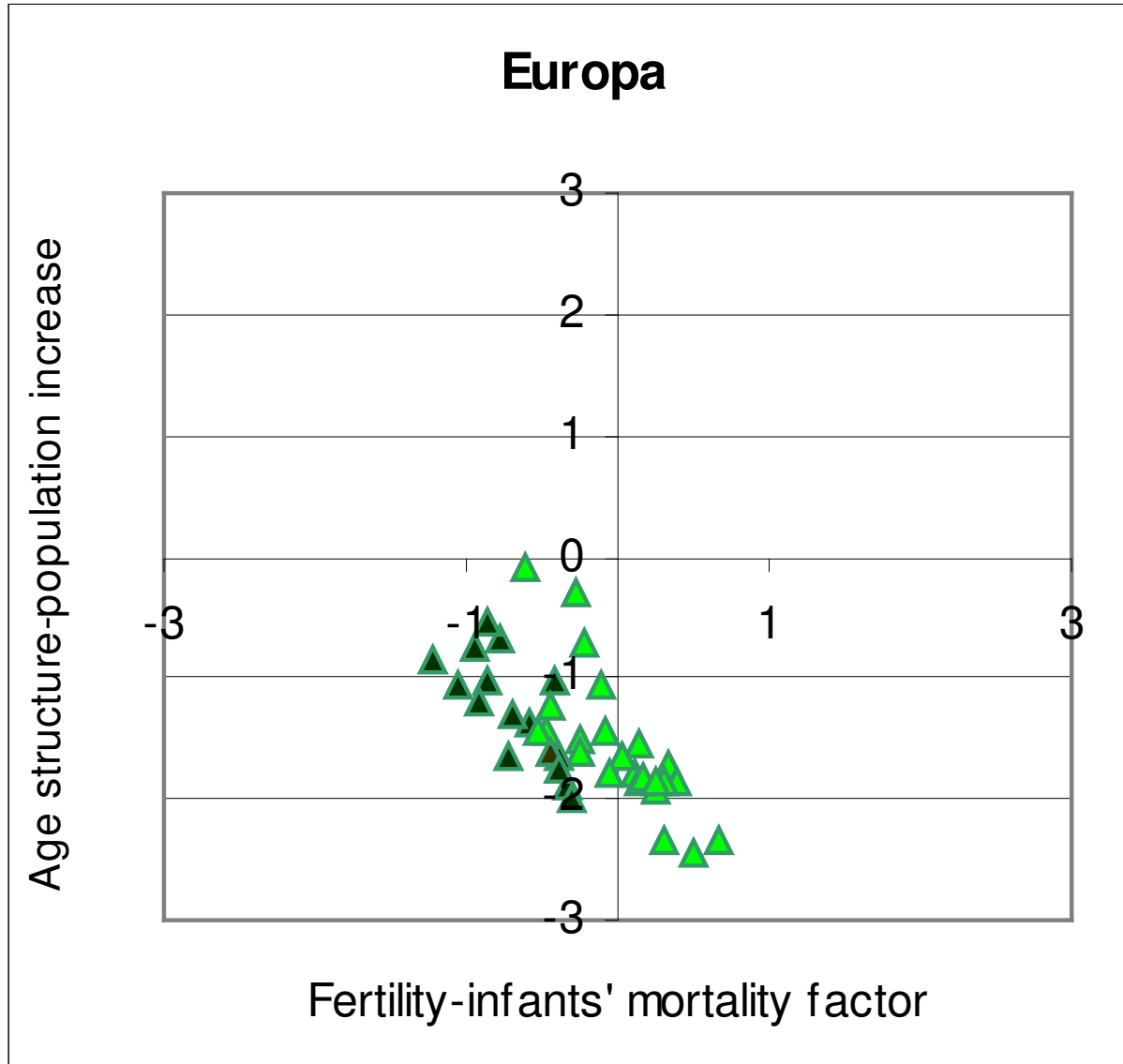
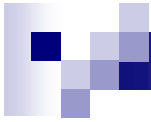
- Demographic behaviour in Asia is very diverse, especially in the case of Fertility factor.
- In most countries the population is rather young and increasing.
- The life expectancy is quite often rather high.

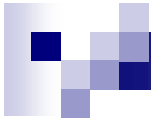




- In America (both Northern and Southern including also Caribbean) the demographic behavior is rather similar.
- Both 1st and 2nd factor are balanced and quite on medium level (for the world as whole).
- In most cases life expectancy is rather comparatively high.



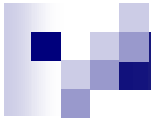




- Compared with all other continents the demographic behavior of Europeans is quite different:
  - Fertility and children mortality are low;
  - Population is aged, natural increase is almost zero;
  - There exist countries with high, but also with low life expectancy.



# Demographic behaviour of Europeans



- Although compared with countries from other continents, especially Africa and Asia, the demographic behavior of Europeans seems to be very uniform, we can see here many different tendencies and trends.



# Defining European regions

- We define six European regions by geographical and political characteristics:
- Geographical:
  - N (North),
  - C (Central),
  - S (South);
- Political:
  - D (Democratic or West),
  - S (Socialist/Transition or East).



- Our hypothesis is, that in these regions are different by the demographic behavior of population and dynamic of demographical trends.
- Mainly, we are interested in fertility and all characteristics connected with fertility indicators.



# Factor Analysis

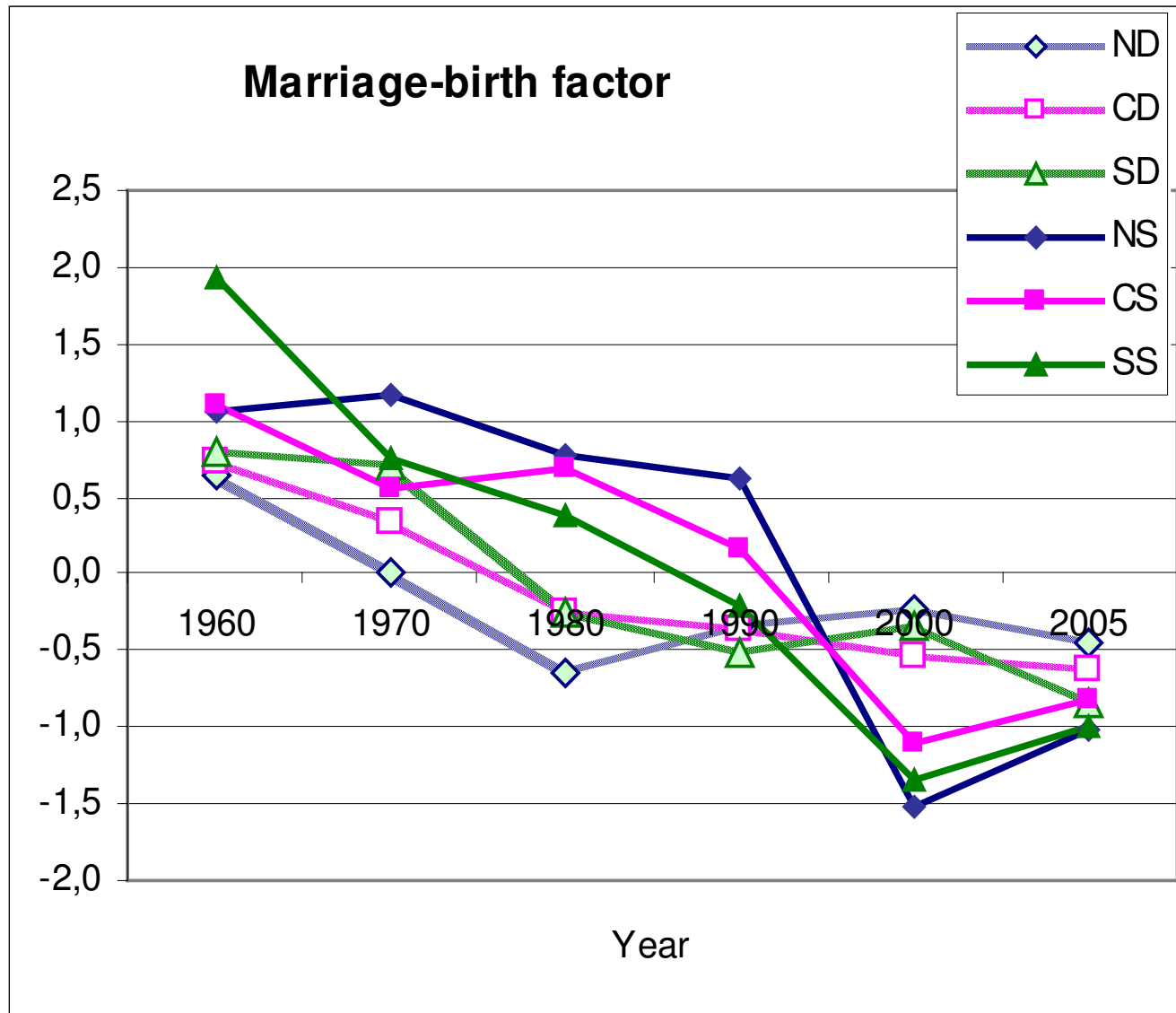
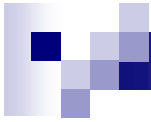
We received the following three factors:

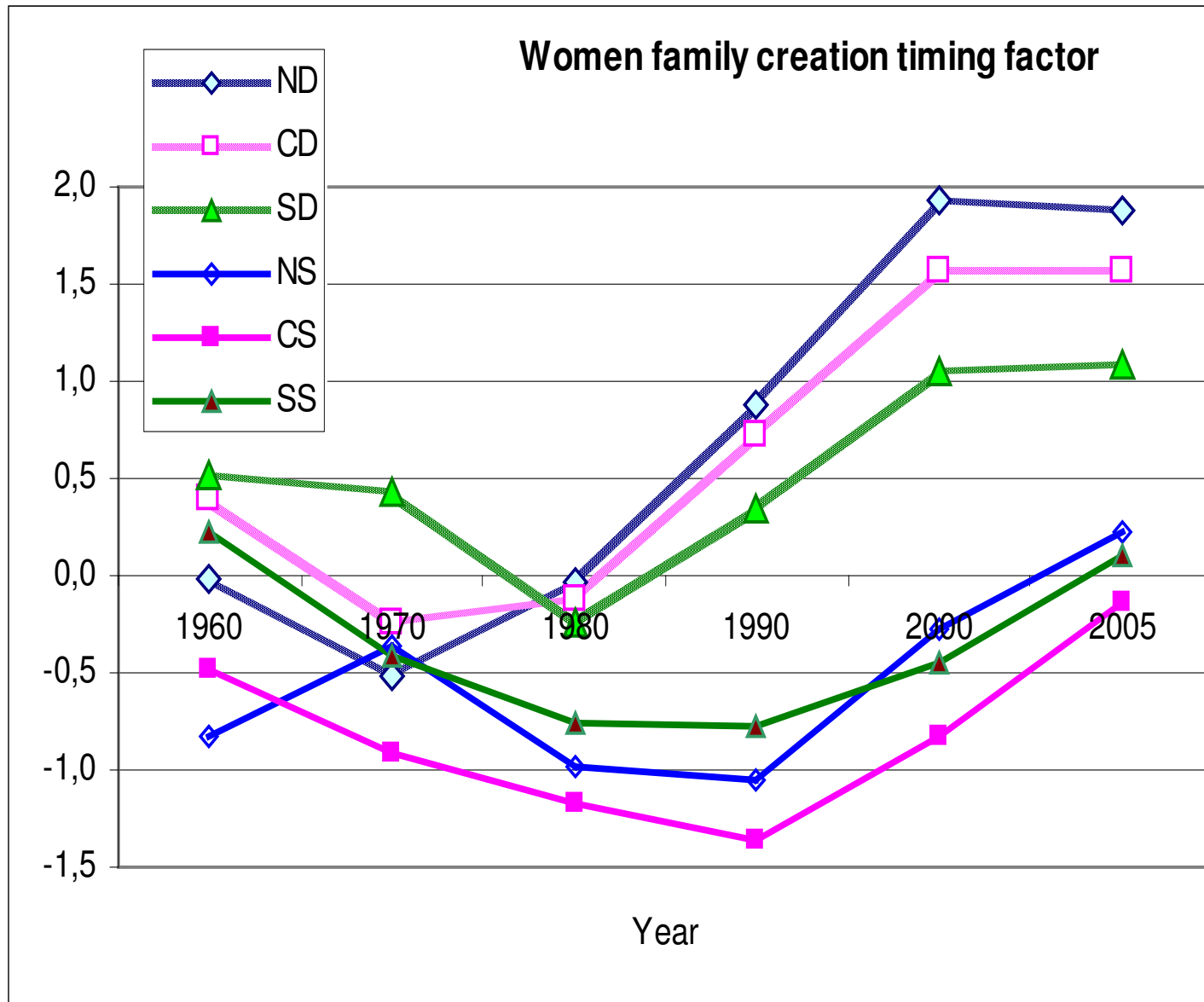
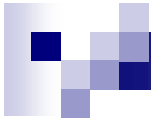
- Marriage-birth rate factor;
- Women's family creation timing factor;
- Family stability factor.

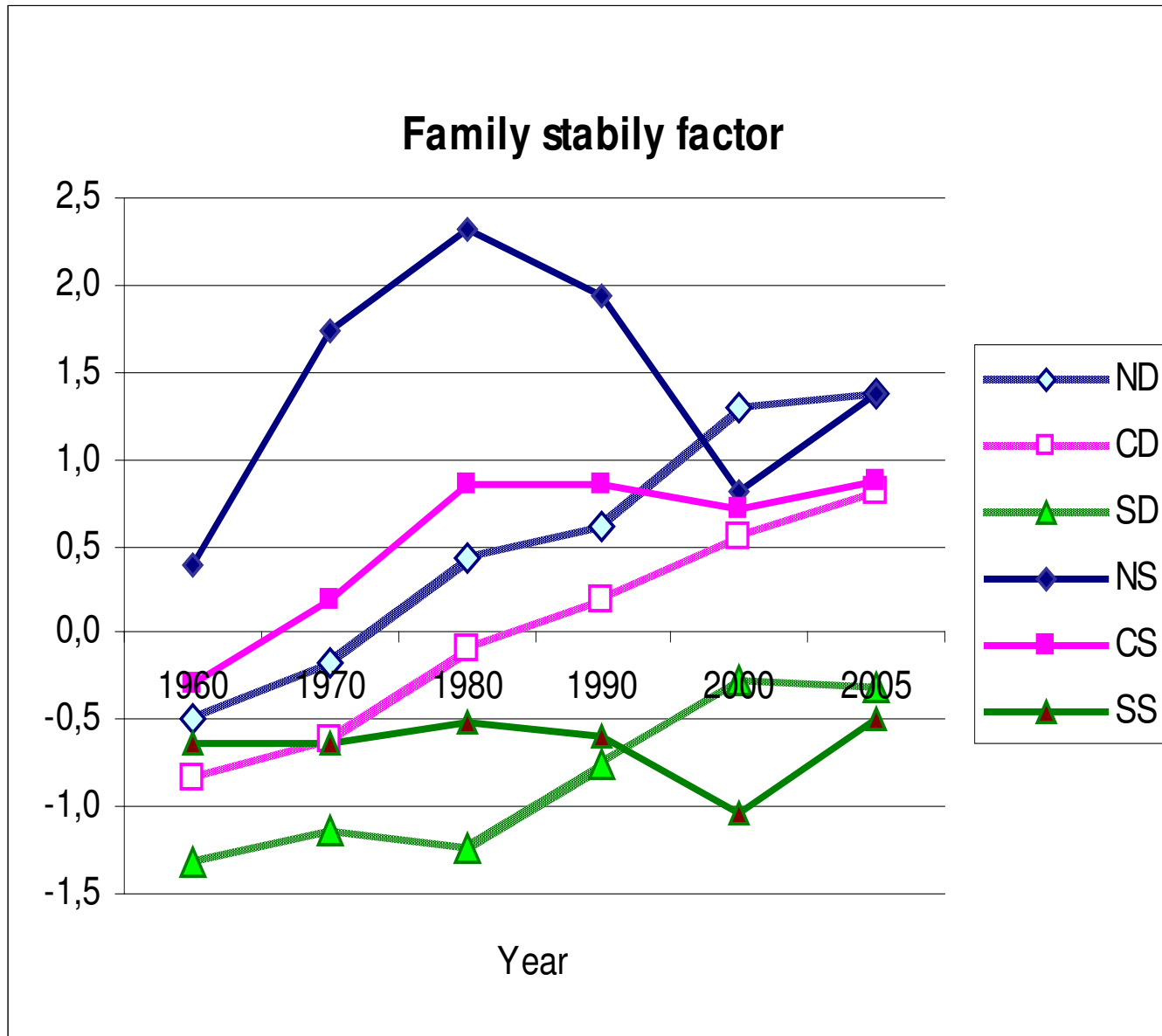
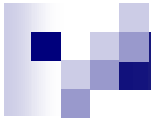


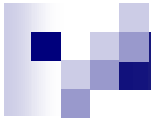
# Dynamics of fertility factors in Europe 1960—2005












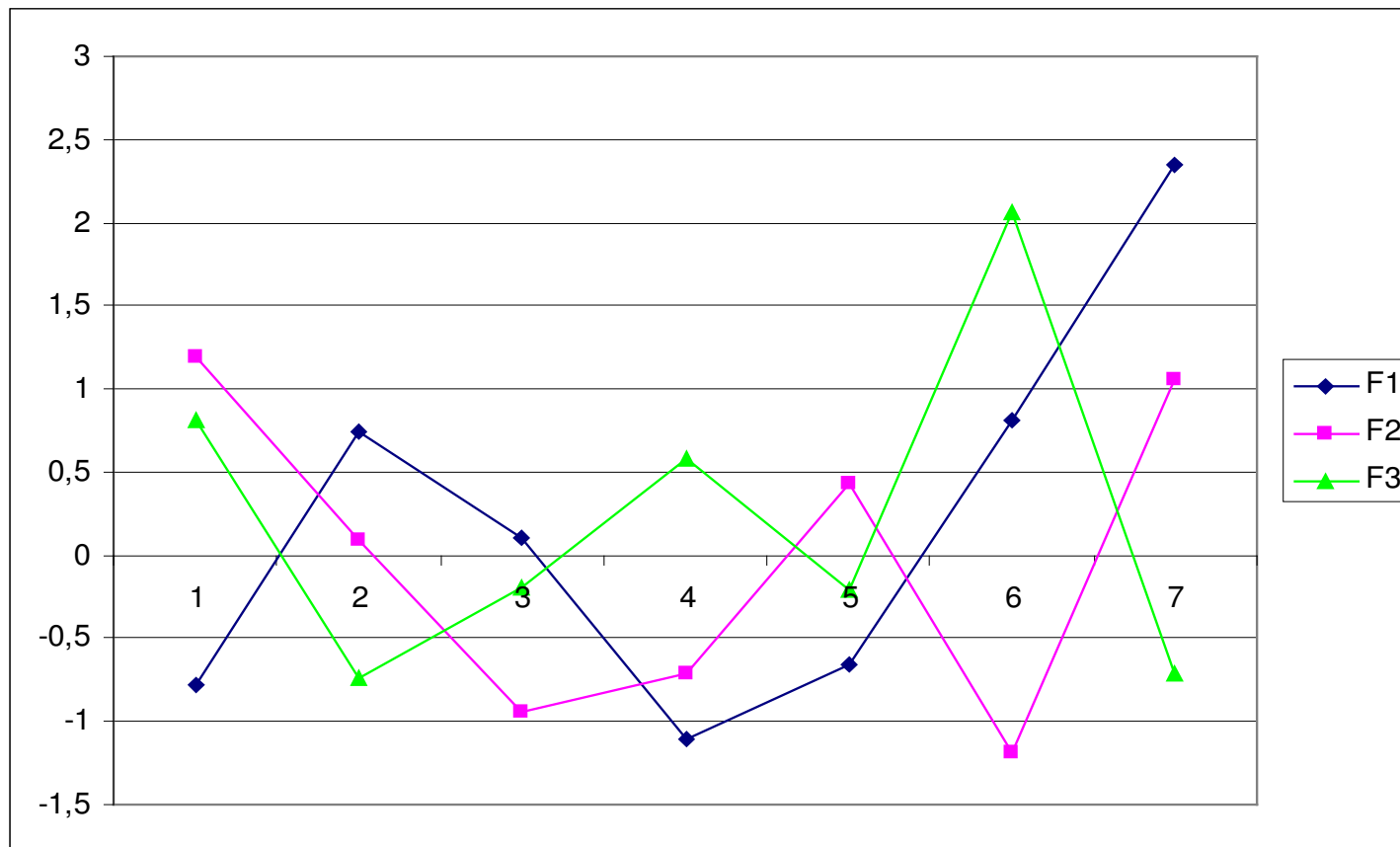
- 1st Factor has declined in all European countries very rapidly, but has stabilised during the last decade;
- 2nd factor has increased in the whole Europe, but is much higher in West compared with East.
- 3rd factor increases in the whole Europe, but its levels depend strongly on geographic variable.

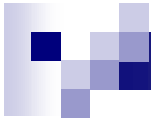


Using clustering to  
describe the trends of  
Europeans demographic  
behaviour

- 
- To check our hypotheses about grouping European countries we used the clustering methodology,
    - where 42 (EU and EFTA) countries were represented,
    - each country in 6 different time (1960, 1970, 1980, 1990, 2000, 2005).
    - Missing values of indicators were EM-imputed.
  - We used K-means method.
  - 7 groups were formed.


# Factor's values of groups

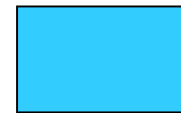
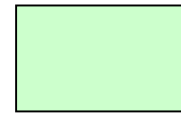
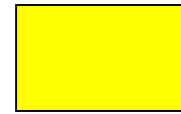


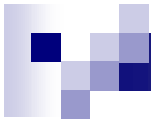


- The interpretation of groups in terms of factors.
- Also the colors of groups used in following tables are added

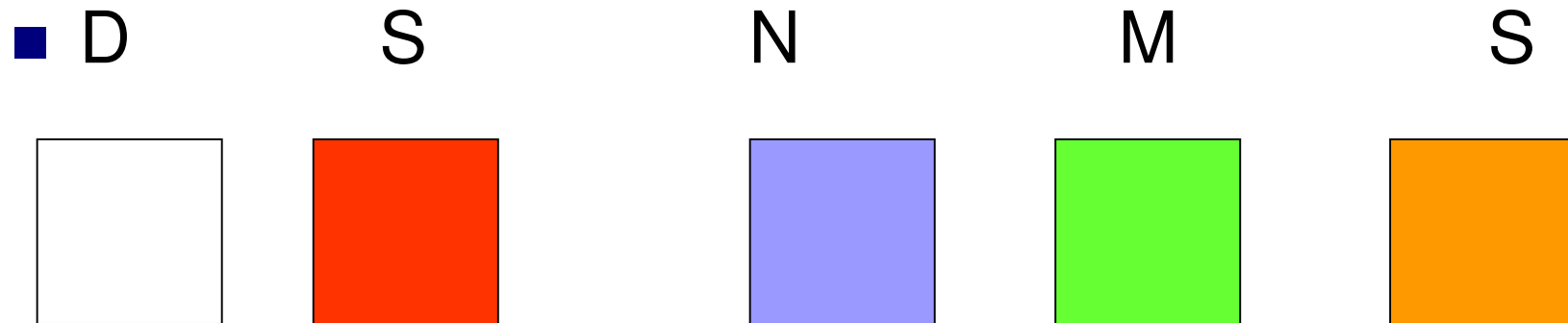


- 
- Group 7 – very traditional, extremely high fertility, stable family;
  - Group 2 – moderately traditional, rather high fertility and rather stable family
  - Group 3 – medium fertility and stability, early family creating
  - Group 5 – late family creating, medium-stable family, quite low fertility
  - Group 6 – Very unstable family, very early family formation and quite high fertility
  - Group 1 – Late family creation, family unstable, fertility low
  - Group 4 – Extremely low fertility, unstable family, quite early family formation





- The following table shows how the countries have in their demographic development moved from one group to another.
- Each line is a country, each column – year (1960, 1970, 1980, 1990, 2000, 2005);
- The last two columns indicate the groups –



	1960	1970	1980	1990	2000	2005		Pol	Geo
	3	5	1	5	1	1		D	N
	2	5	1	1	1	5		D	N
	2	3	5	1	1	5		D	N
	2	3	5	5	1	1		D	N
	7	3	3	5	1	1		D	N
	2	3	5	5	1	4		D	M
	2	2	5	5	1	1		D	M
	2	3	2	5	1	1		D	M
	2	3	5	5	1	1		D	M
	2	3	5	1	1	1		D	M
	2	3	5	5	1	1		D	M
	2	3	5	5	1	1		D	M
	2	2	5	5	5	5		D	M
	7	7	2	5	5	1		D	M
	2	2	2	5	5	5		D	S
	2	2	2	1	5	5		D	S
	2	2	2	2	5	5		D	S
	2	2	3	5	5	1		D	S
	2	2	3	3	5	2		D	S
	2	2	3	2	2	5		D	S
	2	3	3	6	4	5		S	N
	3	3	6	6	1	5		S	N
	3	6	6	6	1	1		S	N
	2	3	3	6	4	4		S	M
	2	3	6	6	4	1		S	M
	2	3	6	6	4	4		S	M
	3	3	3	6	4	5		S	M
	2	3	3	3	5	5		S	M
	2	3	3	3	5	1		S	M
	3	2	3	3	3	1		S	M
	3	3	3	4	1	5		S	M
	7	3	3	6	4	1		S	M
	2	2	3	3	4	4		S	S
	2	3	3	3	5	4		S	S
	2	3	3	4	1	5		S	S
	3	3	3	3	4	1		S	S
	3	3	3	3	5	5		S	S
	7	2	2	2	4	1		S	S
	7	2	3	3	3	4		S	S
	7	2	3	3	3	2		S	S
	7	7	2	2	4	1		S	S
	7	7	7	2	3	1		S	S



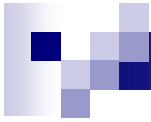
# Conclusions

- Cluster analysis supports only partly the geographical and political classification of European countries;
- It follows that in 1960 the demographic behavior of European citizens was quite uniform
- It was quite diverse in 90-ies and 2000.
- In new millennium the diversities rather diminish



# Methodological Conclusion

Using multivariate methods as Factor and Cluster Analysis it is possible to integrate demographic characteristics and see the trends not visible when using the traditional demographic methodology.



Thank you!