

Department of Statistics
Faculty of Informatics and Electronic Economy



POZNAŃ UNIVERSITY
OF ECONOMICS

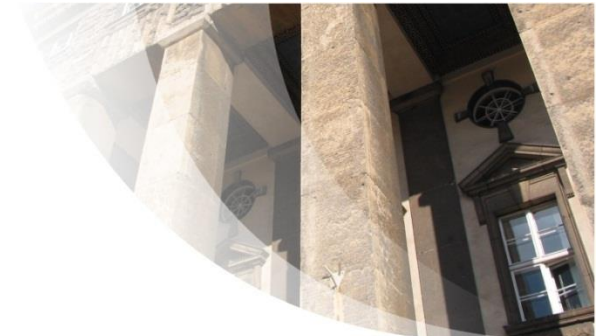
New paradigm in statistics and population census quality

Elżbieta Gołata

elzbieta.golata@ue.poznan.pl



POZNAŃ UNIVERSITY
OF ECONOMICS



Aim of the study

Discussion on the quality
of information derived from the population census
in particular
traditional versus register-based

Outline of the presentation

1. Shift in paradigm

- Interpreting the data
- Availability of the data
- Reorganization in statistics

2. Population census methodology

- Transition in methodology
- Data sources
- Criteria for success
- Evaluation methods

3. Evaluation of 2011 Population Census in Poland

- Coverage errors
- Coverage survey
- Sample survey and non-response

4. Coverage assessment

- Children 0 - 4 years old
- Young people: studying and starting their professional career
- Working age population
- Elderly

5. Conclusion

Shift in paradigm – interpreting the data

PARADIGM [gr. parádeigma 'pattern'],

as introduced by Thomas Kuhn a set of concepts and theories forming the basis of the (natural) science
"universally recognized scientific achievements that, for a time, provide model problems and solutions for a community of practitioners"

CONTROVERSIES IN THE FOUNDATION OF STATISTICS

- ❑ classical statistics or error statistics,
- ❑ Bayesian statistics,
- ❑ Likelihood-based statistics,
- ❑ Akaikean-Information Criterion-based statistics

Issues that have been debated for years without resolution

Shift in paradigm – availability of the data

For centuries, in order to provide statistical analysis,
data needed to be collected

CENSUS

Domesday Book, record of survey completed in 1086 on order of William the Conqueror - **census**
other data sources, like **books of the parish**, or other **administrative records** were also used
Graunt John, *Natural and Political Observations Made upon the Bills of Mortality* (1662)
used the mortality rolls in London to construct first life tables

Middle of XX century,
to save time and money,
thanks to the development of representative methods,

SAMPLE SURVEY

Last decades,
look what data we have,
use of government data and other administrative records,

DATA INTEGRATION

Shift in paradigm – reorganization in statistics

Modern **I**nformation and **C**ommunication **T**echnology -
challenges for statistics and achievements :

- e-surveys, e-questionnaires
- Statistical surveys in modern information infrastructure and ICT
- Statistical frames
- Variety of data sources:
 - Administrative Registers
 - Other Records Available On-line
 - Internet Data
 - Big Data
- Statistical metadata
- **GIS** in statistics
- **S**mall **A**rea **E**stimation
- Variety of methods for data dissemination

Population census – transition in methodology, data sources

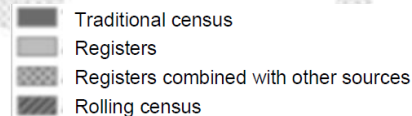
Transition from a **traditional** census in 2002 to **mixed method** in 2011

Data sources :

- Administrative registers
PESEL - Central Population Register
- Internet self-enumeration
- Sample survey long form 20% sample

UNECE countries by Census method for 2010 round

Method	2000-2002	2010-2014
Traditional	40 (80%)	31 (56%)
Register-based	4 (8%)	8 (15%)
Mixed method	6 (12%)	14 (25%)



Population census – criteria for success

According to UNECE 2013 survey on national practices in the 2010 round of population and housing censuses - no full agreement to criteria for success

„Overall user and stakeholder support”

public support

improved response/participation rates,

improved outputs

improved coverage rates,

cost saving

staff expertise,

software

Difference by methodology:

Traditional census: user and stakeholder support, improved coverage rates, government support,

Combined census: improved response/participation rates, user and stakeholder support, **cost savings**

Register-based census: **cost savings**, government support, improved outputs

Population census – criteria for success

International standard ISO 9000/200522 defines quality as:
“**the degree to which a set of inherent characteristics fulfils requirements.**”

The main objective and basic tasks of population census

Essential features of a Population and Housing Census

Individual enumeration

Simultaneity

Universality

Small-area data

Defined periodicity

The population census

produces at regular intervals the **official counting** of the population in the territory of a country and in its **smallest geographical sub-territories** together with information on a selected number of **demographic and social characteristics** of the total population.

The key measure of census quality is the level of response achieved
Accuracy in population estimates

Population census – evaluation methods

A. Single Source of Data

- **demographic analysis of the census**
- evaluation techniques examining the internal consistency
- interpretation studies conducted as part of the census

B. Evaluation methods that use other already existing data sources

(i) Studies using statistical matching techniques

- record checks
- comparison with existing household surveys

(ii) Non-matching studies

- **demographic analysis using previous censuses**
- **comparison with administrative data**
- **comparison with existing household surveys**

C. Evaluation methods using results of additional studies

- post enumeration survey to estimate coverage and content error
- post-censal matching surveys
- reinterview surveys
- additional research on specific populations such as minorities and ethnic groups
- focus studies on respondents' satisfaction with the data collecting process

Evaluation of 2011 census – coverage survey

Census coverage survey was conducted from 1 to 11 July 2011

A sample of 80 thousands dwellings was drawn out of 2 744 thousands dwellings drawn before to the census sample survey

Only flats with at least one person with an assigned phone or mobile number
And all the dwellings self enumerated by Internet

Census coverage survey was performed by CATI.

The form contained 14 questions

No results published yet

Evaluation of 2011 census – sample survey and non-response

Sample survey was conducted on a random sample of **20%** of dwellings on national scale

One stage sampling with deep stratification

Out of nearly 13.5 million dwellings, the sample consisted of more than **2 744** thousand

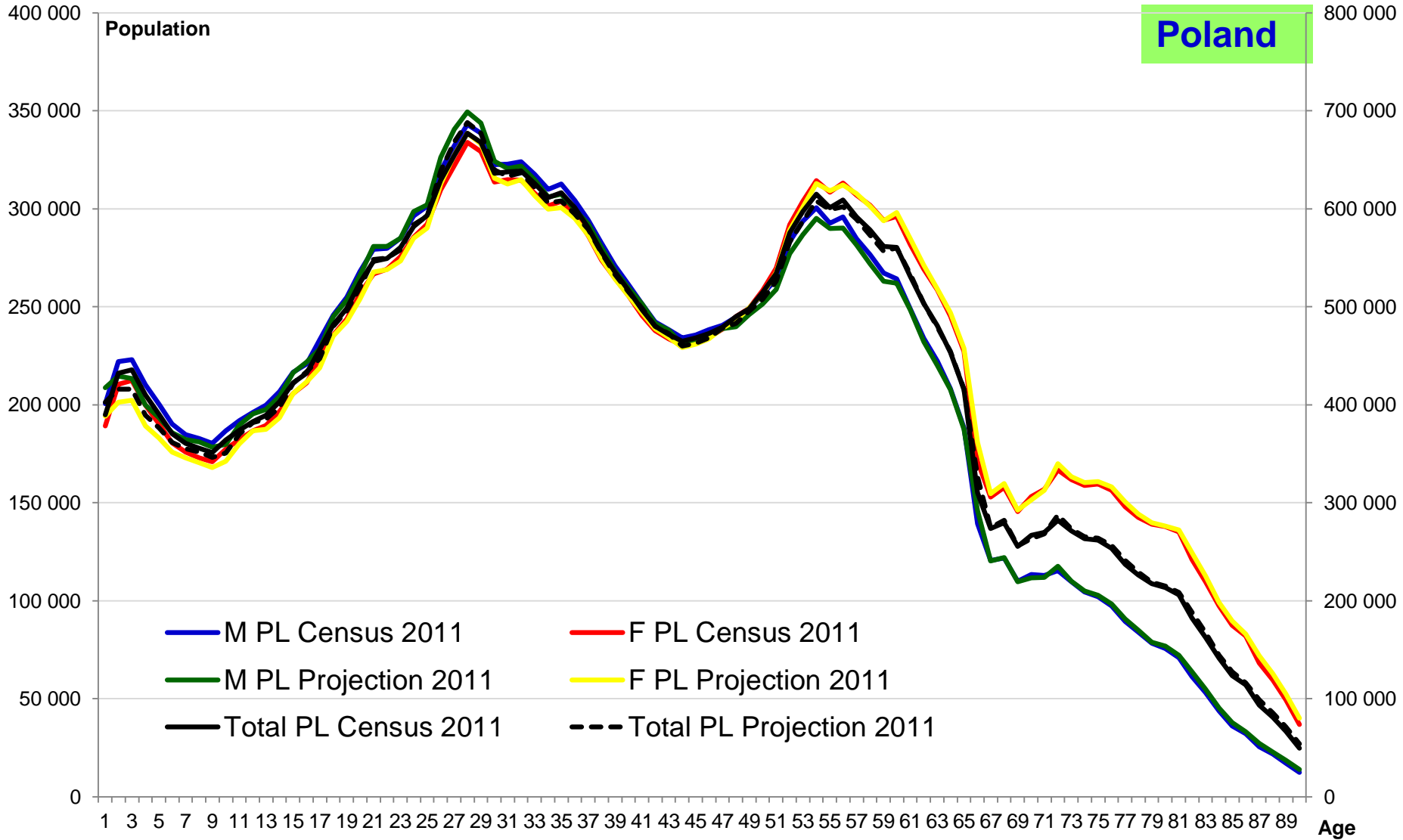
The original weights had to be adjusted due to the **13.7%** of non-response

For all census results **precision tables** were provided

Analysis of non-response not available

Coverage assessment – population by age and sex

2011 census data versus projection based on 2002 census

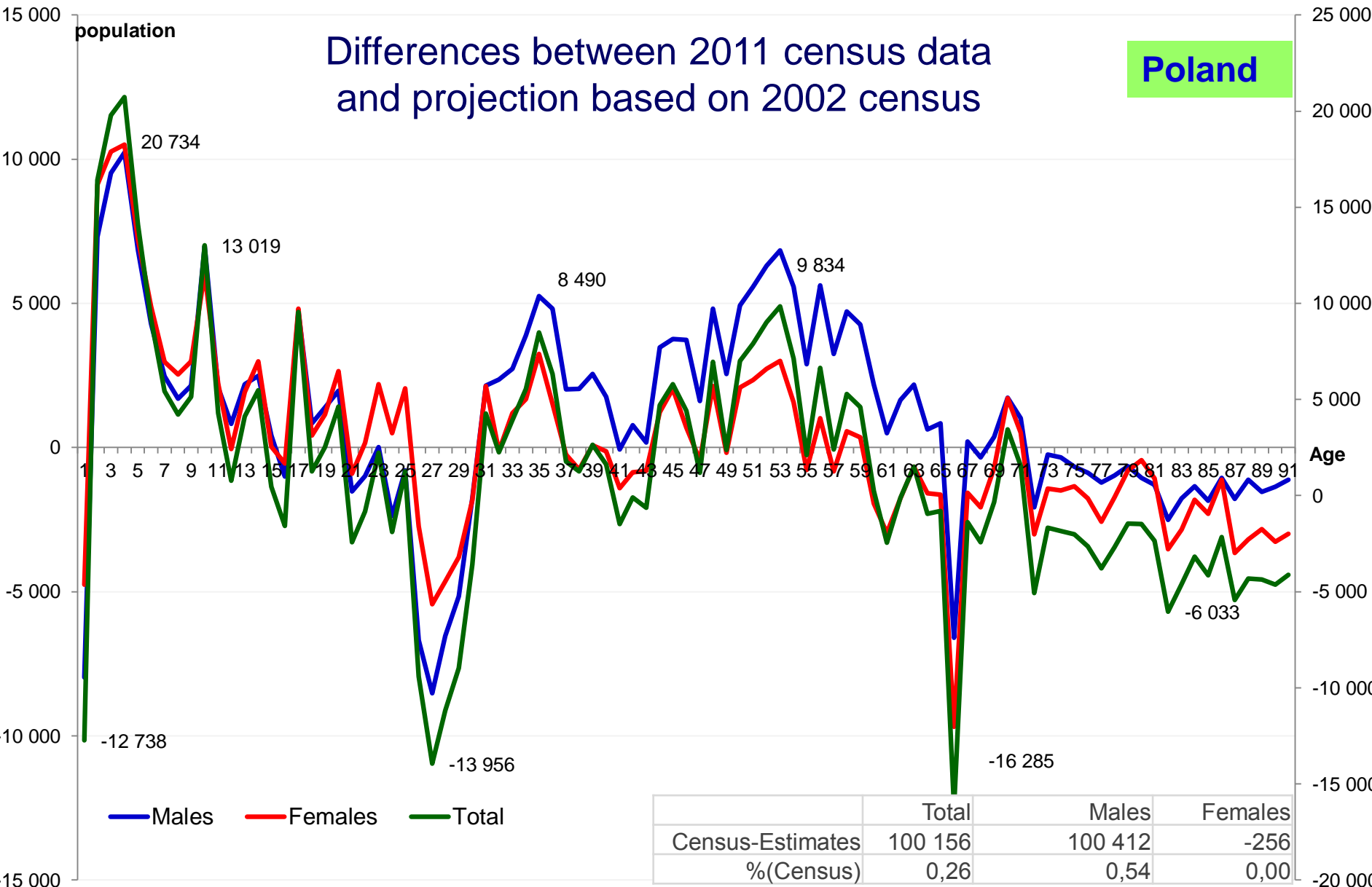


Source: Own elaboration based on 2002 and 2011 Polish Population Census data, life tables and vital statistics (records of births in the years 2002-2011), <http://demografia.stat.gov.pl/bazademografia>

Coverage assessment – population by age and sex

Poland

Differences between 2011 census data and projection based on 2002 census

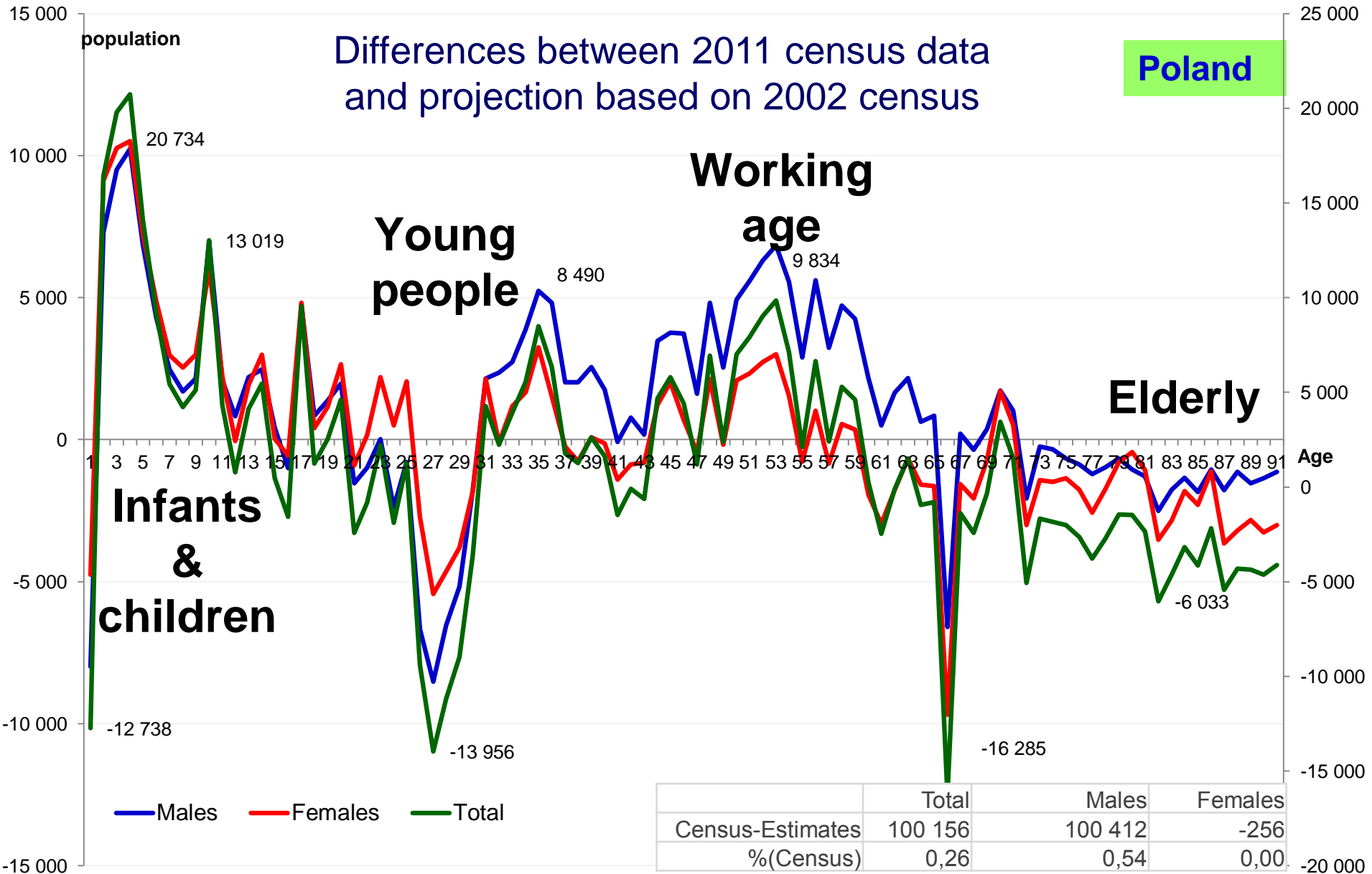


Source: Own elaboration based on 2002 and 2011 Polish Population Census data, life tables and vital statistics (records of births in the years 2002-2011), <http://demografia.stat.gov.pl/bazademografia>

Coverage assessment – population by age and sex

Poland

Differences between 2011 census data and projection based on 2002 census



	Total	Males	Females
Census-Estimates	100 156	100 412	-256
%(Census)	0,26	0,54	0,00

Source: Own elaboration based on 2002 and 2011 Polish Population Census data, life tables and vital statistics (records of births in the years 2002-2011),, <http://demografia.stat.gov.pl/bazademografia>

Coverage assessment – infants and children

Differences between 2011 census data and projection based on 2002 census
- Infants and children

Poland

Age	Total		Males		Females	
	persons	%	persons	%	persons	%
0	-12 846	-3,3	-8075	-4,0	-4 771	-2,5
1	16 414	3,8	7 297	3,3	9 117	4,3
2	19 776	4,5	9 511	4,3	10 265	4,8
3	20 734	5,1	10 229	4,9	10 505	5,3
4	14 086	3,6	6 857	3,4	7 230	3,8
5	9 192	2,5	4 304	2,3	4 888	2,7
Population aged 0 – 4 years and the difference between estimates						
2011 Census	2057998		1055902		1002096	
Birth Register	1999725		1029989		969736	
Difference	58273	2,83%	25913	2,45%	32360	3,23%

Coverage assessment – infants and children

Comparison between 2011 and 2002 census data with Birth Register
- Infants and children

Poland

Data source	2011 Census			2002 Census		
	Total	Males	Females	Total	Males	Females
Census	389903	200592	189311	351662	180116	171546
Birth Register	402641	208573	194068	357096	183440	173656
Difference	-12738	-7981	-4757	-5434	-3324	-2110
Difference (%)	3,27	3,98	2,51	1,55	1,85	1,23

Coverage assessment – infants and children

Mirror statistics: In the search for missing children

Office for National Statistics

Accessibility **A A A**

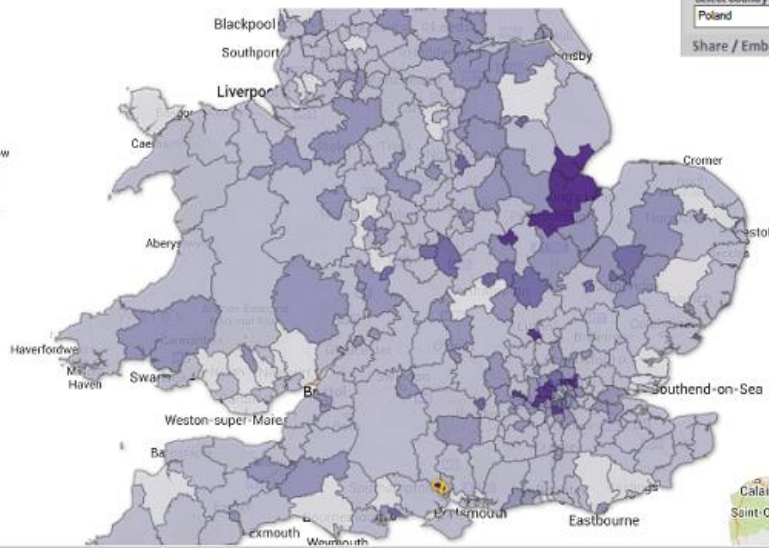
Home Browse by theme Publications Data Release calendar Guidance and methodology Media centre About ONS

You are here: Home > Interactive > Census Map 3.1 - Country of Birth Print friendly

Census map analysis - Country of Birth

Mapa Satelita

2011



Country of Birth

Select country of birth: Poland Search for a postcode or address: London

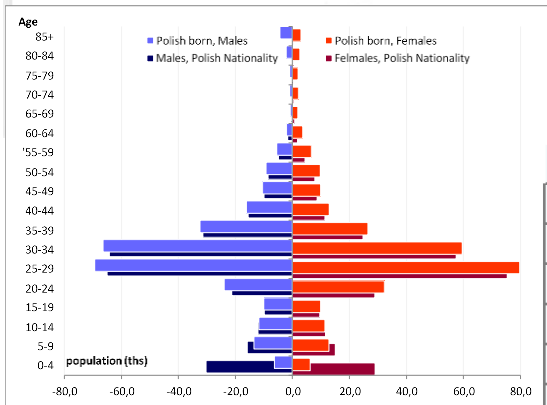
Share / Embed view Opaci Click an area to update table

Click on an area to update table

Country	% of usual resident
UK	82.0
Ireland	0.5
Poland	3.5
India	1.5
Pakistan	0.5
Germany	0.5
Bangladesh	0.3
Nigeria	0.2
South Africa	0.4
United States	0.2
Jamaica	0.1
Italy	0.2
France	0.2
Portugal	0.1

Source: 2011 Census

Polish immigrants in England and Wales by sex and age



Live births to women from the top five countries of birth of non-UK mothers living in the UK, 2007 to 2011

Country	2007	2008	2009	2010	2011	Together
Pakistan	18 311	19 561	19 112	18 543	19 156	75 527
Poland	13 333	18 325	20 559	22 239	23 004	74 456
India	12 478	13 025	13 078	14 186	15 549	52 767
Bangladesh	8 850	8 710	8 546	8 440	8 458	34 546
Nigeria	6 702	7 260	7 475	7 715	7 871	29 152

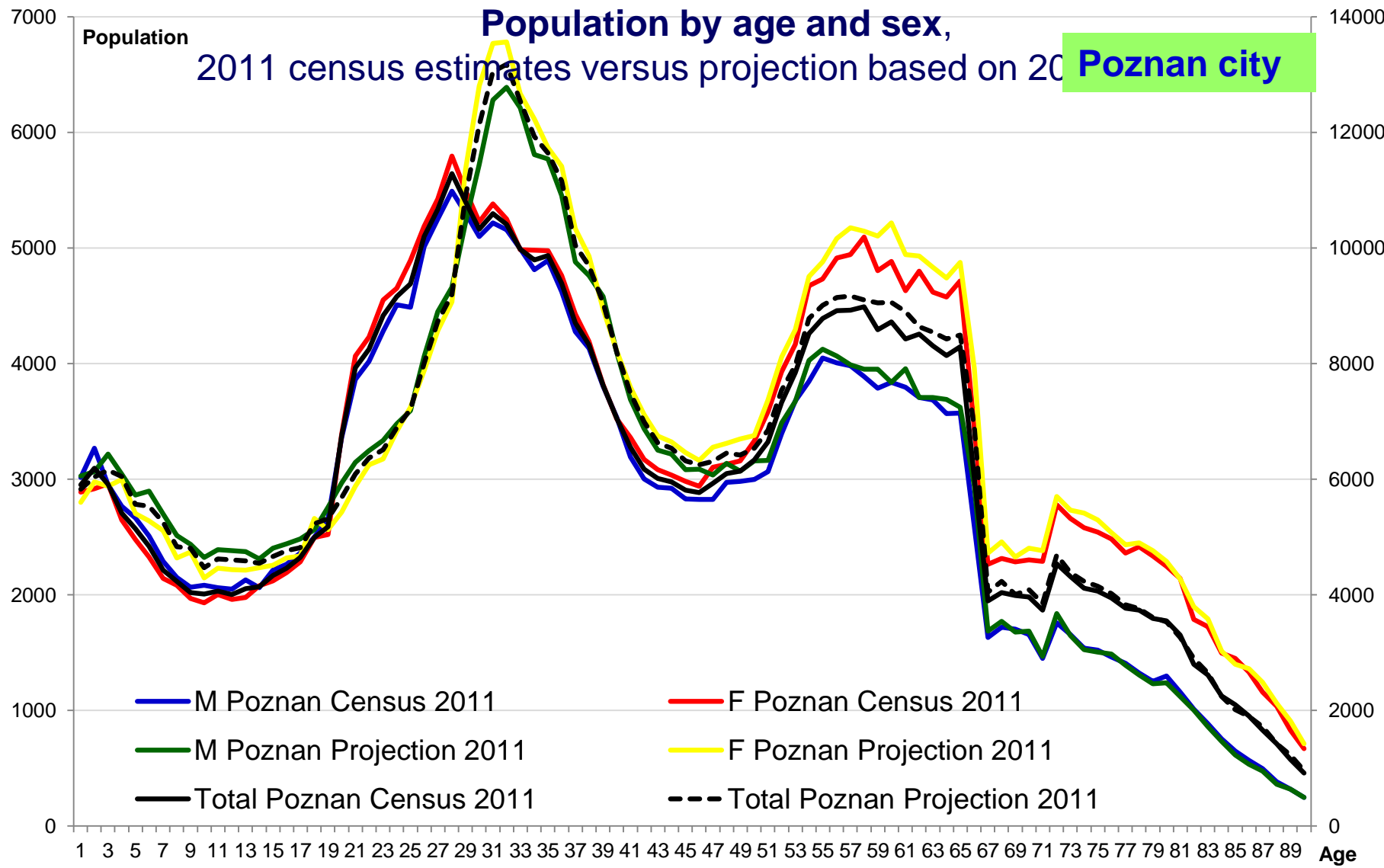
Source: Detailed country of birth and nationality analysis from the 2011 Census of England and Wales, ONS 2013

Source: Childbearing Among UK Born and Non-UK Born Women Living in the UK

Evaluation of 2011 census – data for the smallest territories

Data for the smallest geographical sub-territories

Coverage assessment – by territorial division

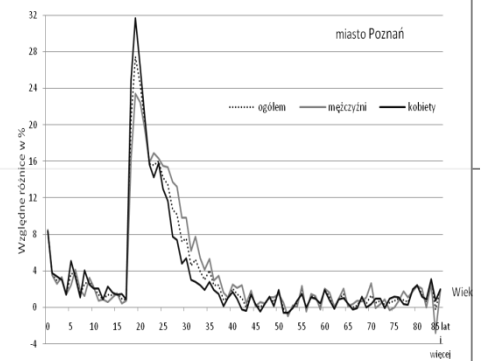
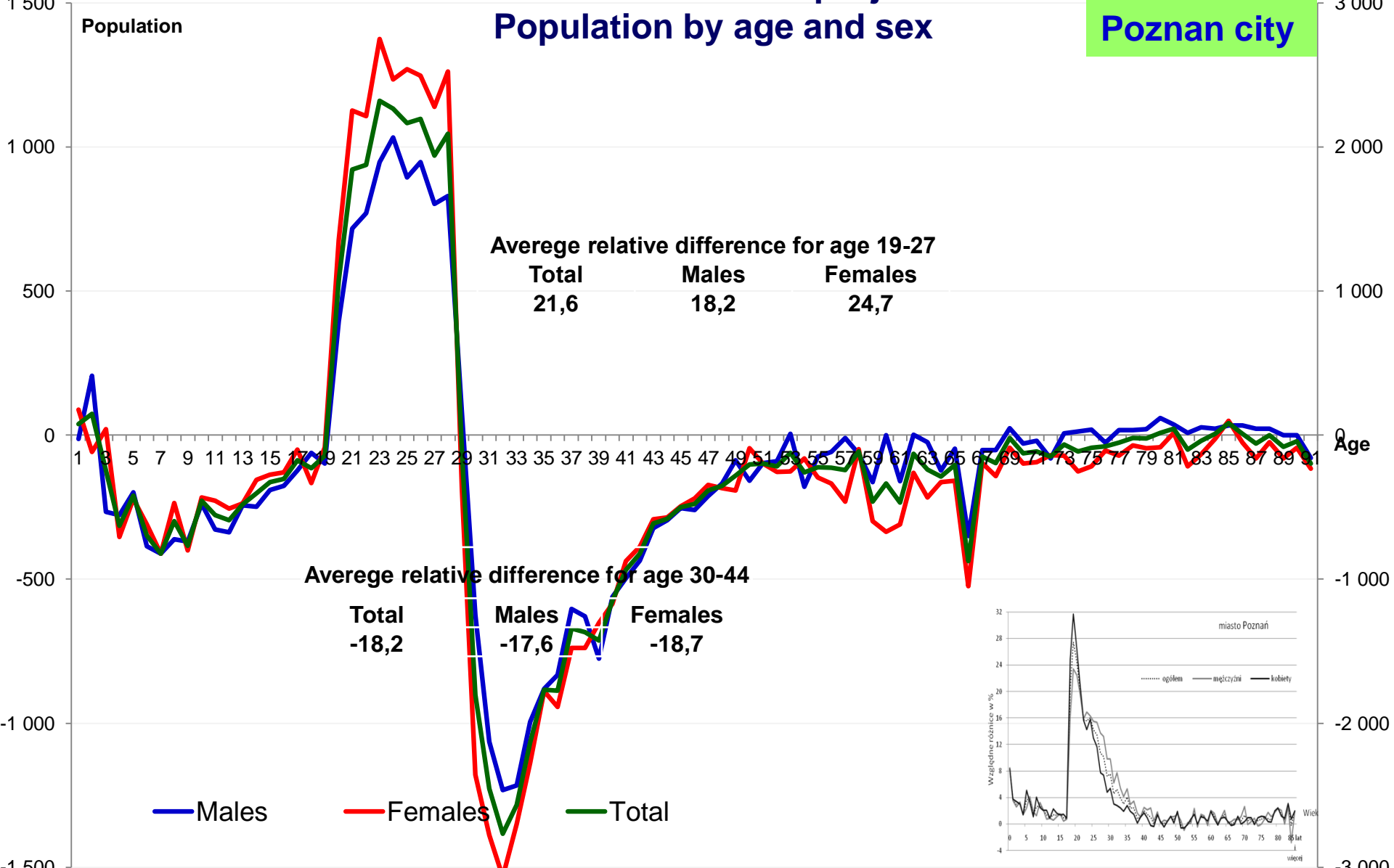


Source: Own elaboration based on 2002 and 2011 Polish Population Census data, life tables and vital statistics (records of births in the years 2002-2011), <http://demografia.stat.gov.pl/bazademografia>

Coverage assessment – by territorial division

Differences in 2011 census estimates versus projection based on 2002 census

Poznan city



Source: Own elaboration based on 2002 and 2011 Polish Population Census data, life tables and vital statistics (records of births in the years 2002-2011), <http://demografia.stat.gov.pl/bazademografia>

Coverage assessment – by territorial division

Differences in 2011 census estimates versus projection based on 2002 census Population by age and sex, Poznań

Age	Census		Projection based on 2002 census	'Survival' ratio	Difference between			
	2002	2011			2011 census and adequate 2002 census population		2011 census population and projection based on 2002 census	
					Absolute	Relative (%)	Absolute	Relative (%)
0-4	22 858	22 682	24 923					
5-9	24 827	21 549	23 814					
10-14	31 653	20 652	23 012	0,9035	-2 206	-9,65	-2 360	-11,430
15-19	44 285	26 025	25 804	1,0483	1 198	4,83	221	0,848
20-24	63 232	43 540	33 069	1,3755	11 887	37,55	10 471	24,050
25-29	52 044	53 267	48 949	1,2028	8 982	20,28	4 318	8,106
30-34	36 352	50 637	62 321	0,8008	-12 595	-19,92	-11 684	-23,073
35-39	32 488	41 087	48 145	0,7895	-10 957	-21,05	-7 058	-17,178

Evaluation of 2011 census – coverage errors

2011 Census versus previous traditional censuses

The census carried out using the traditional method was not "better" than the one conducted with a mixed method, because of variety of sources used.

There is well documented evidence on coverage errors in Polish traditional censuses:

R. Jończy (2010), M. Kędelski (1990), J. Kordos (2007, 2008), B. Sakson (2002), P. Śleszyński (2004, 2004a, 2005)

Jan Paradysz (1989, 2010):

- **1988 Census** – shortage of up to **30%** of **women** with the shortest length of the marriage
- **2002 Census** – omission of **10%** of the youngest **infants** up to 6 months
- **2002 Census** – omission of the population with **increased mobility** (20-29)
- **2002 Census** – lack of **elderly** (over 90)

Conclusion

1. Census based on multiple data sources enforces **application of modern methodology**:
 - Statistical Data Integration
 - Estimation using multiple sources of information
 - Calibration
2. Surveys based on multiple sources include a mechanism for **mutual control**, research compliance, comparative analyses, what results in more reliable information

**Quality assessment system
is built into procedures using various data sources**

3. When using different data sources a 'natural' danger appears of obtaining **inconsistent results**. Divergent estimates force attempts to provide consistent estimates and explain the reason of differences - **calibration**
4. Use of different sources of information, including sample surveys, implies the need to consider **random errors**
5. Error estimation for analysis using integrated data (register and sample survey) requires **development of new theoretical concepts**

Conclusion

Quality evaluation of 2011 Census in Poland should include

1. Advantages of the applied methodology
estimation for small areas, calibration, GIS, statistical data integration

Development of science in response to the needs

2. An in-depth exploration of the results of demographic analyses
3. Implementation of modern information and communication technologies
website, online database, GIS, web-database
4. Cost reduction
5. The multiplicity of sources of information - Improvement of the data quality
6. Verification of quality of administrative records

The most important success 2011 censuses

**The quality of the 2011 Population Census in Poland
is the subject of discussion,
and not tacitly accepted as unquestionable**

Thanks for your attention



POZNAŃ UNIVERSITY
OF ECONOMICS

