Geospatial statistical data integration for decision making and monitoring SDGs

Dr JANUSZ DYGASZEWICZ

President of European Forum for Geography and Statistics EFGS
Statistics Poland
Director of ICT Systems, Geostatistics and Census Department
Member of Executive Committee of the UN-GGIM: Europe

Geospatial World Forum 2019, Amsterdam 2-4 April 2019

European Forum for Geography and Statistics

- **Since 1998 a network of experts** – to provide spatial statistics for Europe.
- The **main goal** - to promote the integration of statistical and geospatial information and the use of GIS in NSIs in public sector.
- Mainly GIS specialists but also statisticians, researchers working in National Statistical Institutes (NSIs), employees of mapping authorities
- Today EFGS has national contact persons from more than **40** states and territories.
- Population distribution in 1 km² grid for Europe.

EFGS conferences

- The annual EFGS conferences are one of the main events for **raising awareness** among the statistical and geospatial communities.
- In recent years the conference has attracted an expanding attendance. The last **EFGS 2018** conference in Helsinki gathered more than **200 participants** from **46 countries** across all continents!
- A list of previous conferences:

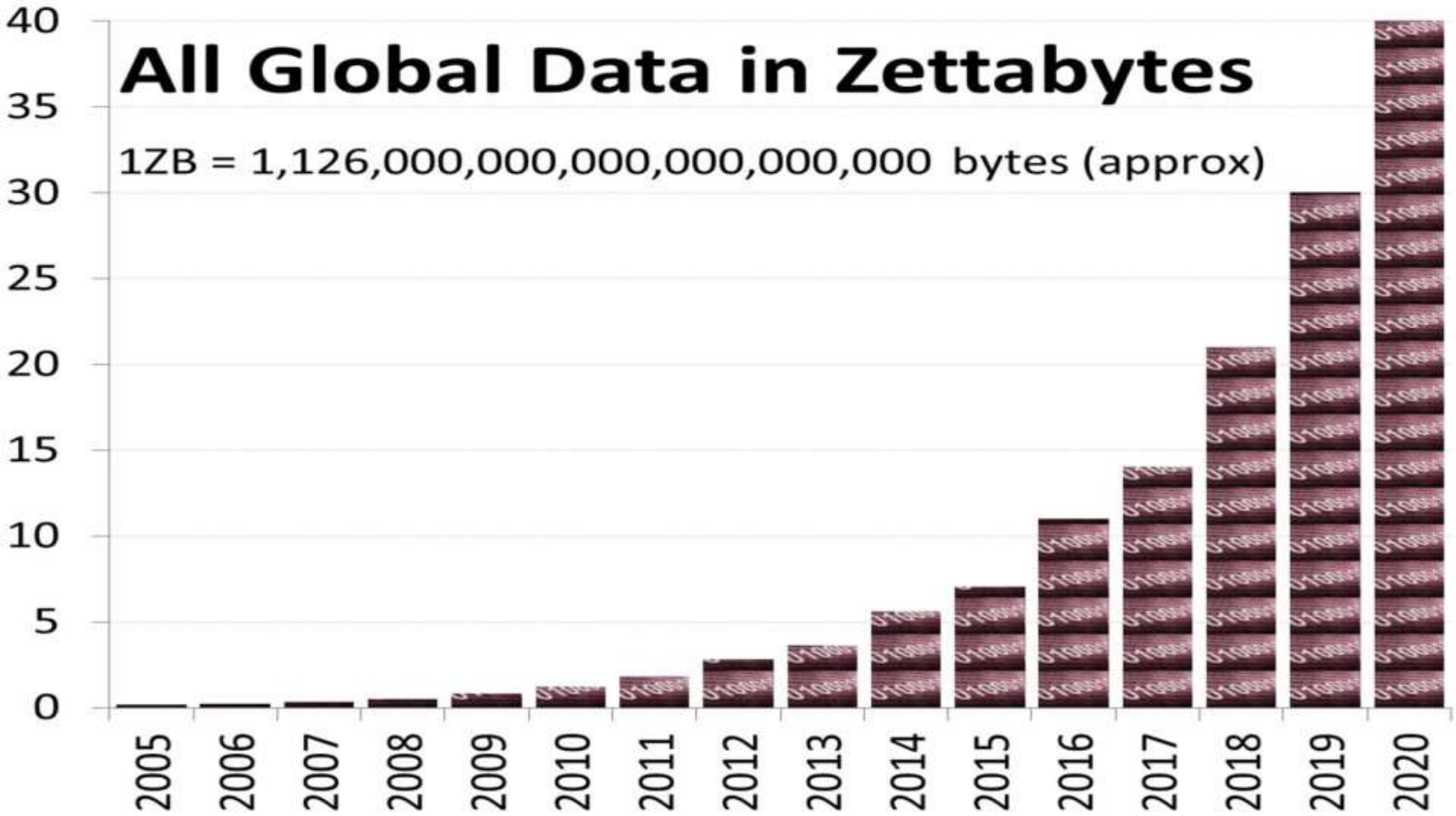
EFGS 2008 Bled, Slovenia
EFGS 2009 The Hague, The Netherlands
EFGS 2010 Tallinn, Estonia
EFGS 2011 Lisbon, Portugal
EFGS 2012 Prague, The Czech Republic
EFGS 2013 Sofia, Bulgaria
EFGS 2014 Krakow, Poland
EFGS 2015 Vienna, Austria
EFGS 2016 Paris, France
EFGS 2017 Dublin, Ireland
EFGS 2018 Helsinki, Finland

Next EFGS 2019 conference

- **10-11 October 2019, Manchester**
- **Motto: *Statistics + Location --> Insights + Impact***
- Planned theme e.g.*: **Leaving no one behind*** - delivering the **SDG** through geospatial case studies
- The Conference will be hosted by **United Kingdom** - Office for National Statistics (ONS) in cooperation with the Ordnance Survey (OS)
- A great **venue** in the heart of the city - **Radisson Blu Edwardian Manchester**
- More information soon: www.efgs.info

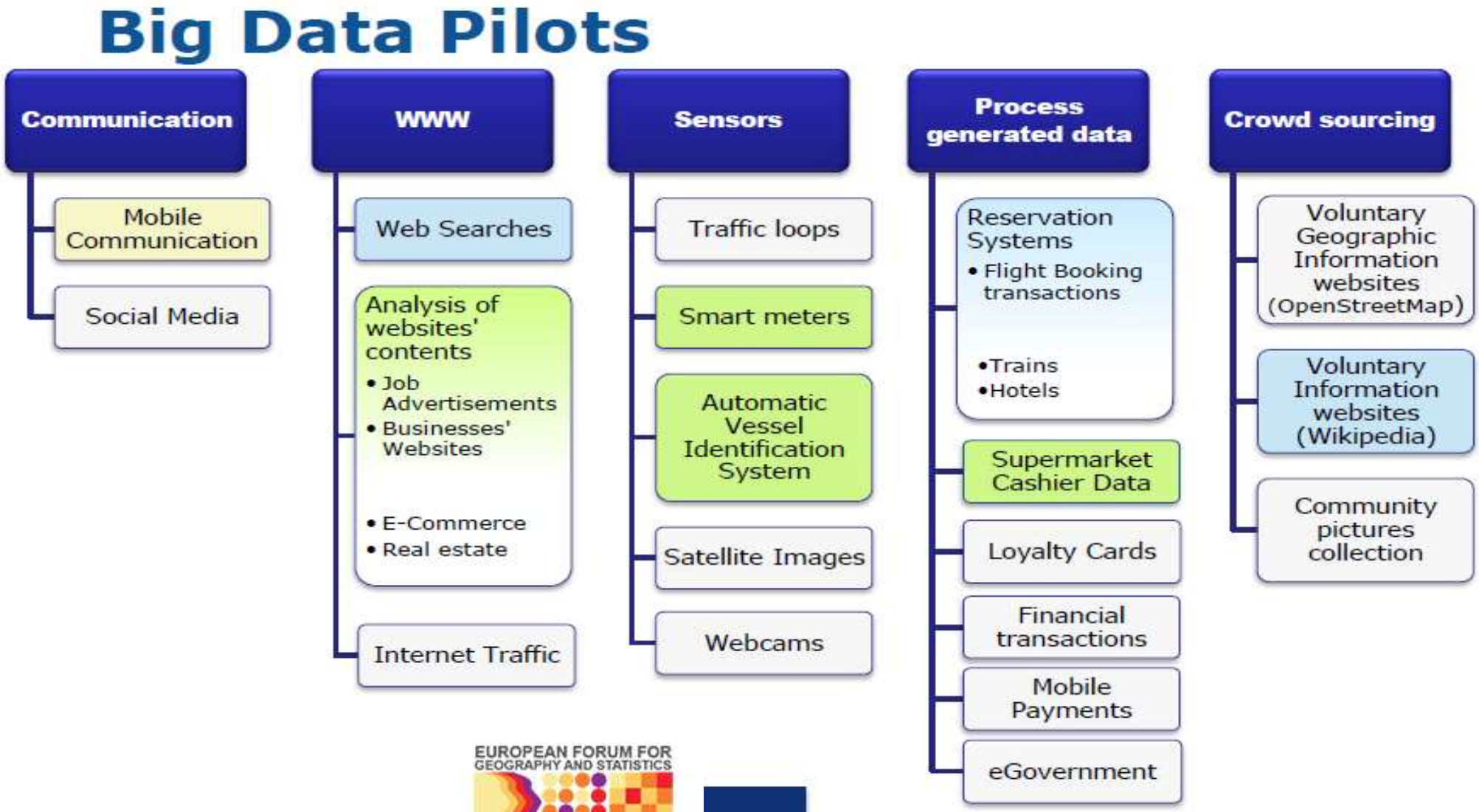
The potential usage of Big Data in Geostatistics

Digital data flooding



Source: UNECE

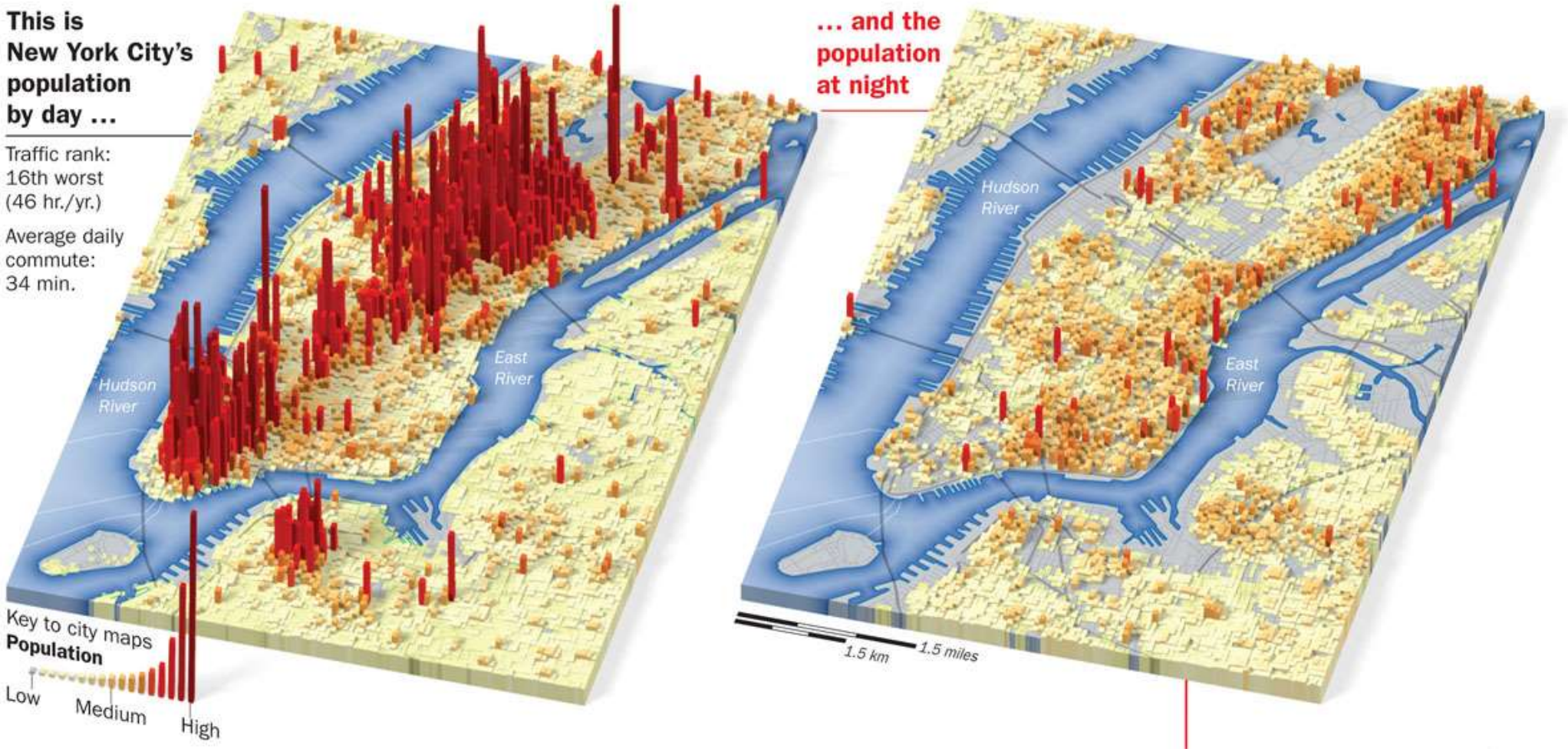
Big Data sources for Geostatistics - EUROSTAT



Sensors in mobile phones

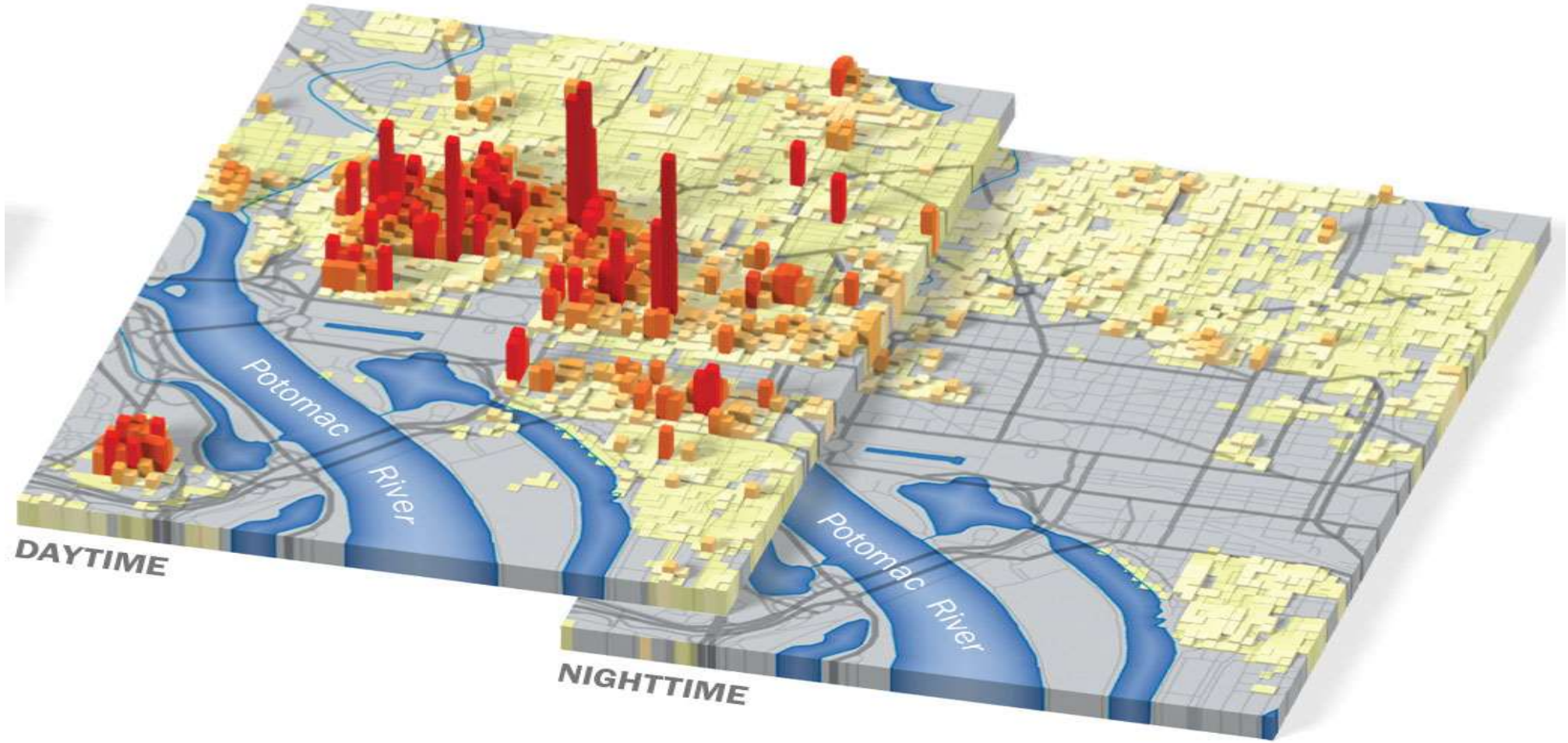
Main culprit:
Source:http://rjacquez.com

Location data from mobile phones



Source: Time magazine

Day vs. night city population



Source: Time magazine

Big Data – Big Obstacles?

Law	**Data safety**	**Privacy**
Ethics	**Competence**	**Methods**
Technologies	**Quality**	**Access to the data**

Spatial databases for statistical surveys

Address points database

- address points with x,y coordinates

Administrative and Statistical division boundaries

- Statistical regions
- Census areas

- In practice they consist the geocoding frame

Point based geocoding allows a more flexible grouping of data, preparing statistics and elaborating SDG's indicators

It also makes possible carrying out spatial analyses of various phenomena concerning:

- **demography** e.g. the average distance between children's and parents' residence, commuting to work, school, distance to a hospital,
- **urbanisation and planning** e.g. useful in determining the boundaries of urban agglomerations, metropolies, and the drawing up of land development plans,
- **agriculture and environment** (analysing the structure of crops, environmental pollution),
- **the economy** e.g. analysing the effects of burdensome road and industry investments.

Results of geospatial analysis

Taking into consideration time and space, it is possible to find **answers to the following spatial questions**:

- How does a given phenomenon present itself in terms of space...?
- What is located in...?
- What caused the phenomenon of...?
- What is the relation between different location...?
- What happens if we change location in space ...?

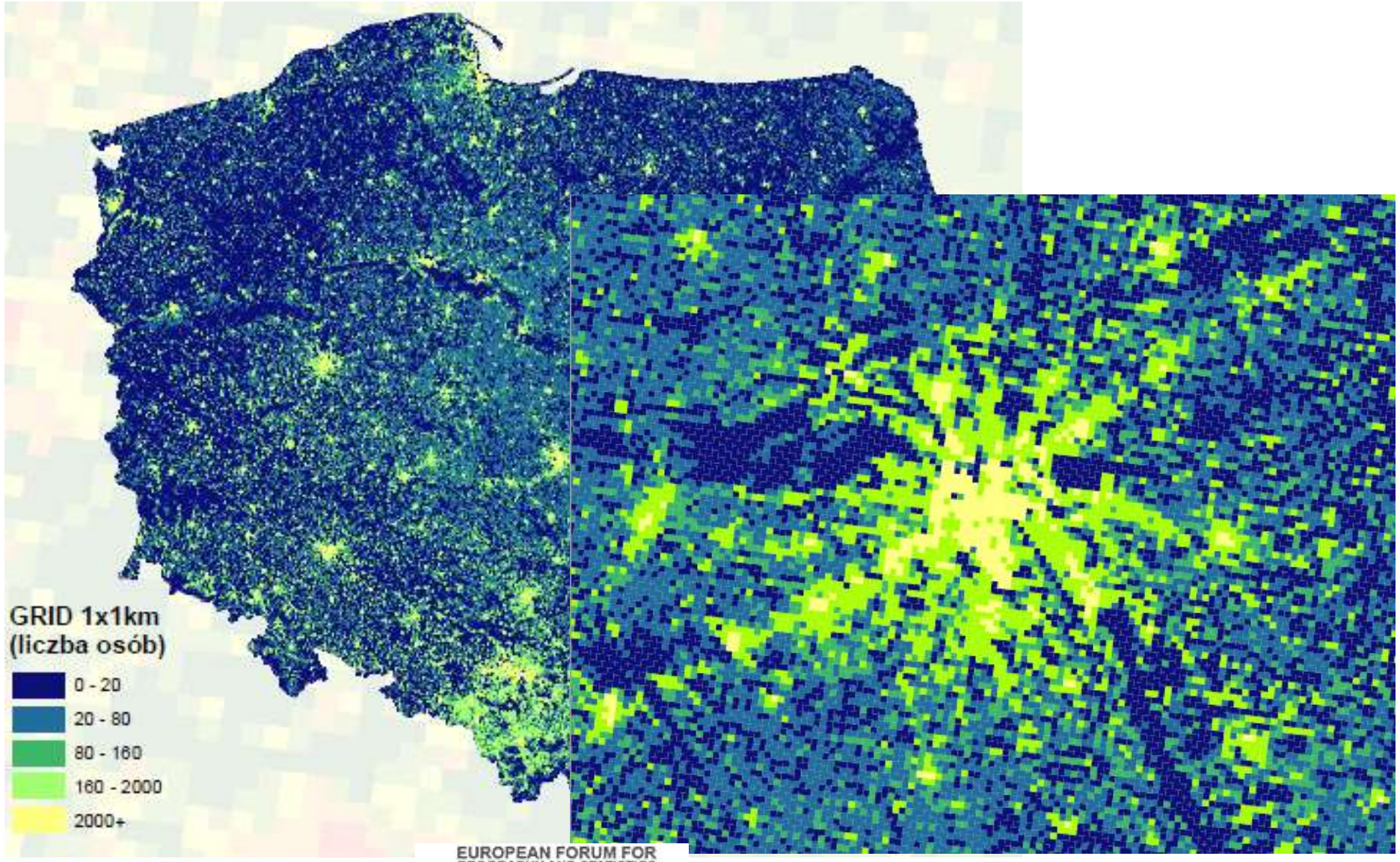
Example:

... What percentage of population living "x" kilometers from main roads suffer from "aa" disease and what is the dynamics of the phenomenon over the last "n" years...

Benefits from geocoding frames

- Point based geocoding allows publishing survey results on maps in any spatial division:
  - administrative division
  - statistical division
  - **1 km^2 grid**
  - any chosen area

Demographic data in 1 km^2 grid – population distribution

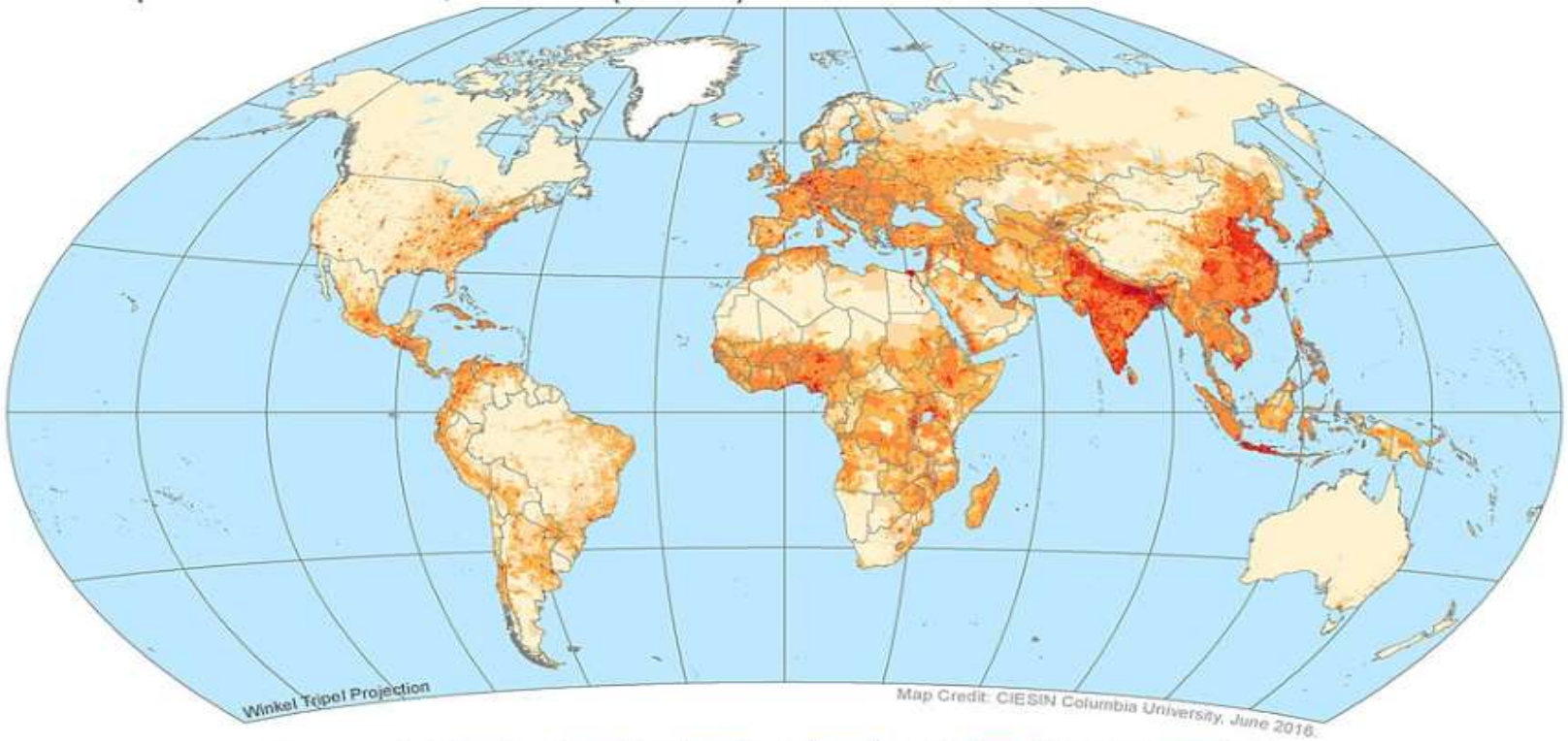


Statistics Poland

EUROPEAN FORUM FOR GEOGRAPHY AND STATISTICS

Population Density Grid, 2015: Global

Gridded Population of the World, Version 4 (GPWv4)



Persons per km² <1 1–5 5–25 25–250 250–1,000 1,000 +

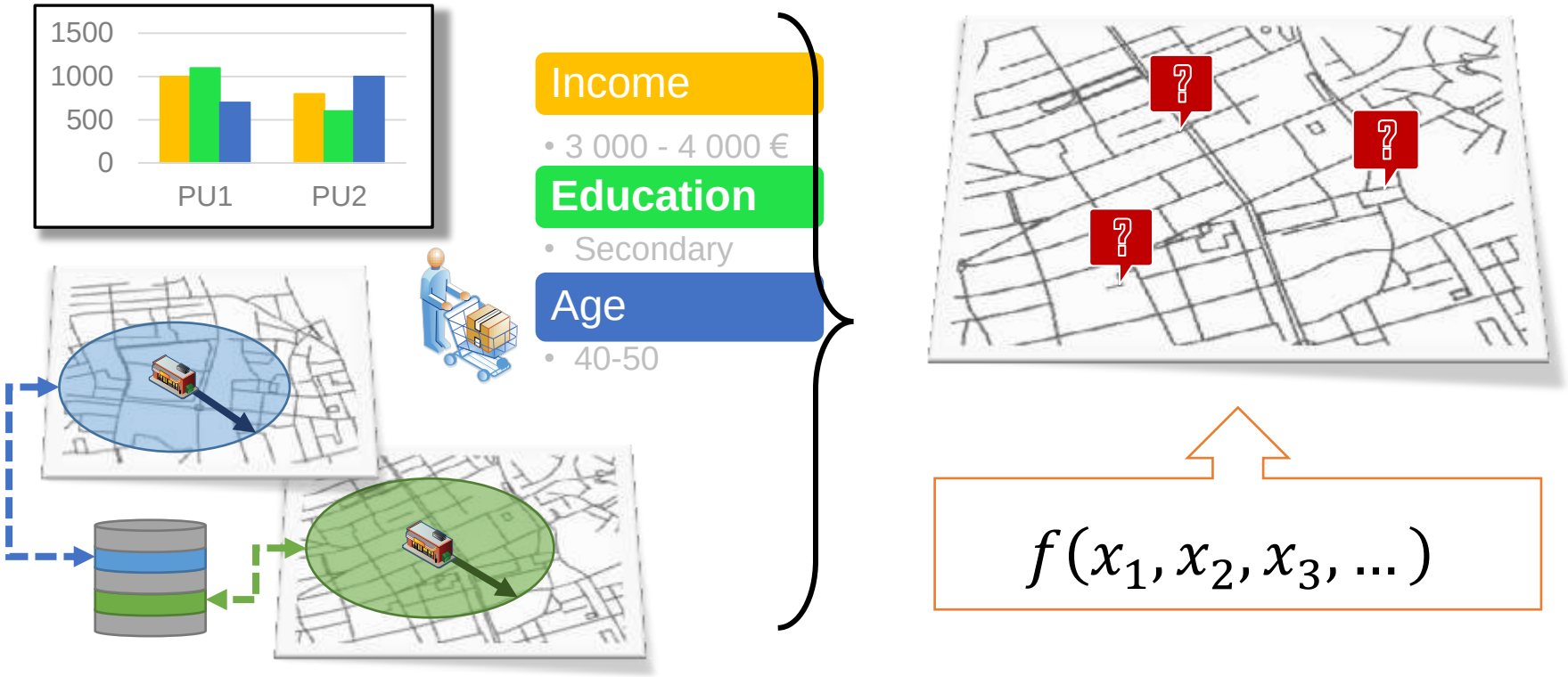
Gridded Population of the World, Version 4 (GPWv4) Population Density consists of estimates of human population density based on counts consistent with national censuses and population registers, for the years 2000, 2005, 2010, 2015, and 2020. A proportional allocation gridding algorithm, utilizing approximately 12.5 million national and sub-national administrative units, is used to assign population values to 30 arc-second (~1 km) grid cells. The population density grids are derived by dividing the population count grids by the land area grids. The pixel values represent persons per square kilometer.

Center for International Earth Science Information Network
Earth Institute | Columbia University

Data Source: Center for International Earth Science Information Network - CIESIN - Columbia University. 2016. Gridded Population of the World, Version 4 (GPWv4): Population Density. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4NP22DQ.

EUROPEAN FORUM FOR GEOGRAPHY AND STATISTICS

Development of the analytical model (regretion model)



EFGS experts are involved in the development of SDG indicators in cooperation with UN-GGIM: Europe

Geostatistics analysis on new global, regional or national SDG indicators

- Work on methodology for monitoring the Sustainable Development Goals 2030.
- Reflect spatial data aspects in SDG - min. for 15 indicators spatial data are needed.
- 4 indicators for pilot elaboration has been chosen :
  - 11.2.1 Proportion of population that has convenient access to public transport, by sex, age and persons with disabilities
  - 11.3.1 Ratio of land consumption rate to population growth
  - 11.7.1 Average share of the built-up area of cities that is open space for public use for all, by sex, age and persons with disabilities
  - 15.1.1 Forest area as a proportion of total land area

Conlusions

- Works on SDG indicators demonstrates the great potential of geospatial-statistical integration through use of point-based geocoding for the monitoring of SDGs.
- Geospatial statistical data integration is enable only with common understaning of both communities (NSIs and NMCAs).
- For many indicators monitoring SDGs, statistical data with very precise location information is needed to conduct analyses which show territorial dimension and changes of different phenomena.
- Only official statistics has been collecting all necessary unit data with reference to points (X,Y coordinates) that is essential to produce spatial aggregates which enable reliable geospatial analysis.
- Geocoded statistical data is a key element that is required to calculate minimum 15 indicators which are crucial for informed decision making and monitoring SDGs.

Statistics Poland

Thank you for your attention

Janusz Dygaszewicz
President of European Forum for Geography and Statistics EFGS
Statistics Poland
j.dygaszewicz@stat.gov.pl

EUROPEAN FORUM FOR GEOGRAPHY AND STATISTICS