

Statistical data confidentiality and micro data in Albania

Gjergji Filipi* and Alma Kondi**

* INSTITUTE OF STATISTICS OF ALBANIA, Blvd Zhan D'Ark No 3 Tirana, Albania,
Gjergji.Filipi@instat.gov.al

** INSTITUTE OF STATISTICS OF ALBANIA, Blvd Zhan D'Ark No 3 Tirana, Albania,
akondi@instat.gov.al

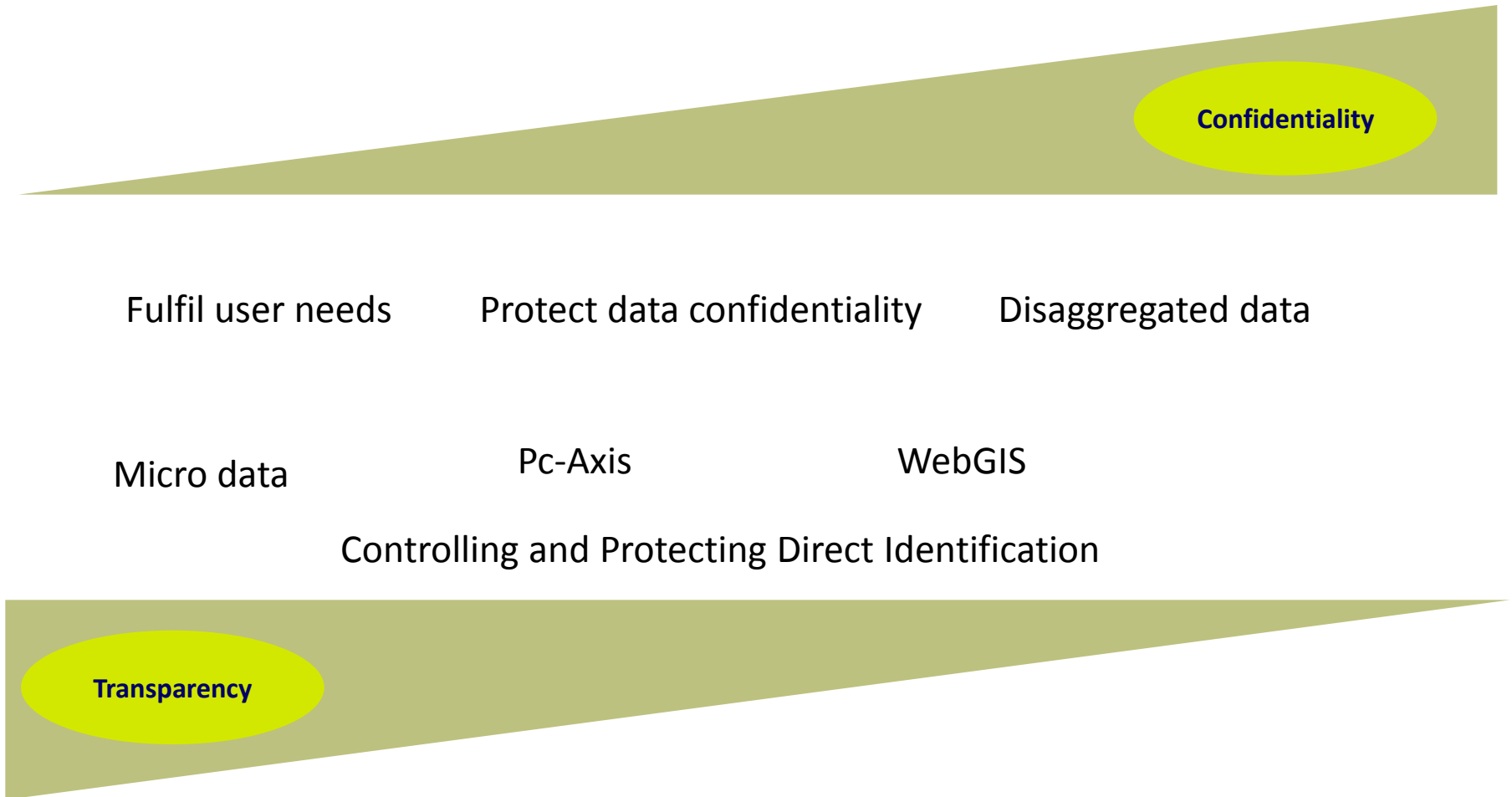
Contents

- Introduction
- Trade-off between confidentiality and transparency
- Legal Framework
- Confidentiality declaration
- Data availability for researchers
- Disclosure techniques
- Micro data access
- Disclosure techniques for Micro data
- Web-Gis
- Future Plans
- Conclusions

Introduction

- Confidentiality is one of the Fundamental Principles of Official Statistics:
 - to maintain the trust of respondents
 - to ensure the quality of data
- In Albanian Institute of Statistics (INSTAT) all individual or business data collected:
 - are strictly confidential
 - are used exclusively for statistical purposes

Trade-off between confidentiality and transparency



Legal Framework

- The main act is “On Official Statistics” :
 - The protection of data directly or indirectly violation
 - The prevention of non-statistical utilization of the data collected and unlawful disclosure
- Another important act is ” Protection of Personal Data”:
 - Define the rules for the protection and legal processing of the personal data
 - Usage and protection of micro data is not excluded

Confidentiality declaration

- In 2014, INSTAT launched a new system granting access for special authorized researches.
- The researchers have to sign a confidentiality declaration which states :
 - To not use and to not transmit confidential information to unauthorized persons;
 - To not use or disseminate any reserved information received during their work for any purpose that is in contrary to the aims of the work;
 - To maintain the confidentiality of personal information and to keep them safe;
 - To preserve the image of the institution;

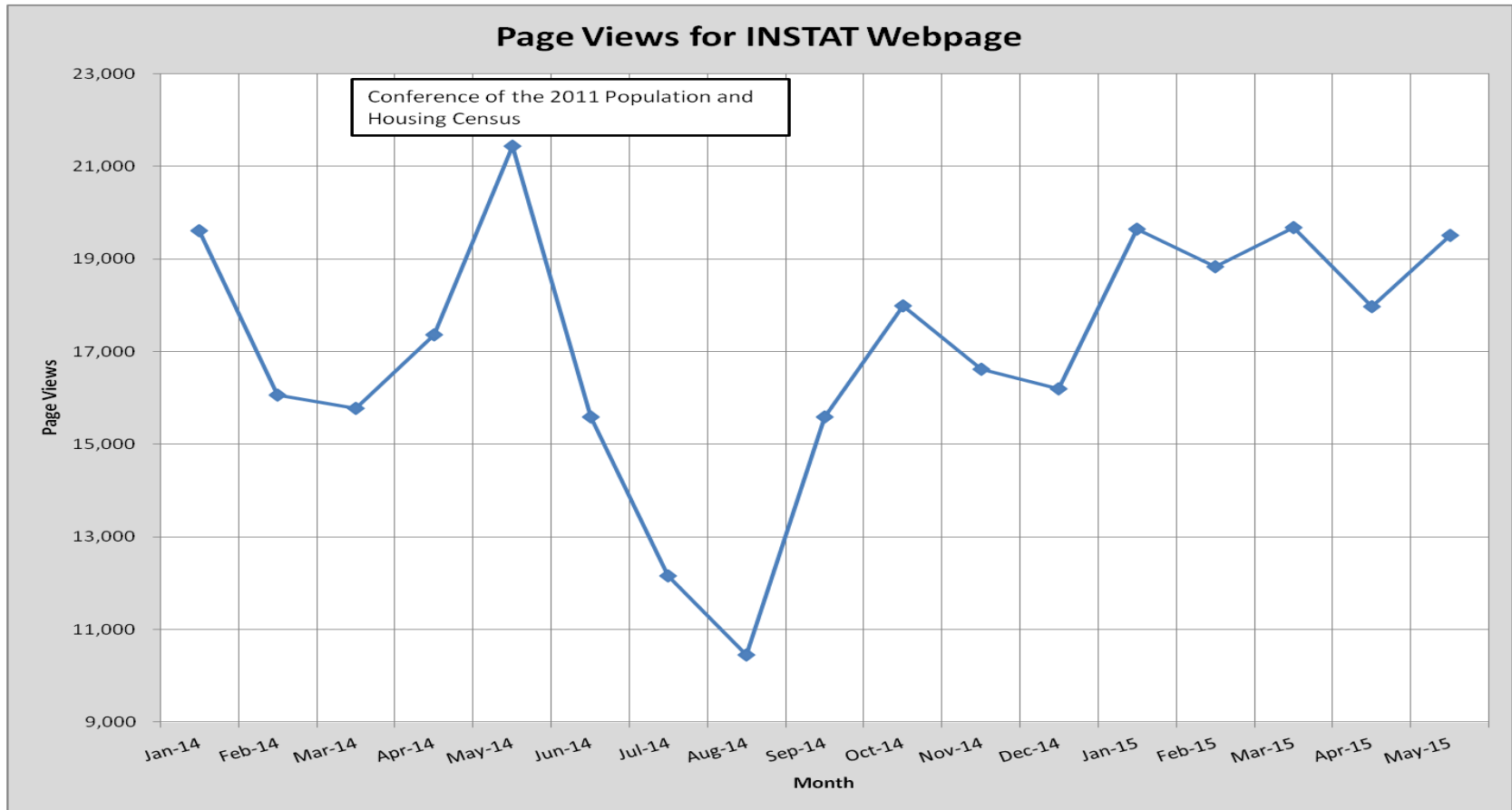
Data availability for researchers

- In September 2012 INSTAT has developed a new well-functioning website with user friendly structure.
- All statistical data produced by INSTAT are:
 - Free of charge
 - Available online at web in Excel format
 - Available online readable from different operation system
 - Can be accessed by using the PC-AXIS software in non proprietary format (CSV, XLS)
 - Published on time according to the calendar of publications
 - Access is not discriminatory

Disclosure techniques

- *Aggregation of data* - Aggregate information on entities presented in tables
- *Cancellation of individual - data for indirect identification* in statistical tables
- *Adjustment* - of the levels of details
- *Rounding* - It is used to protect small counts in tabular data against disclosure

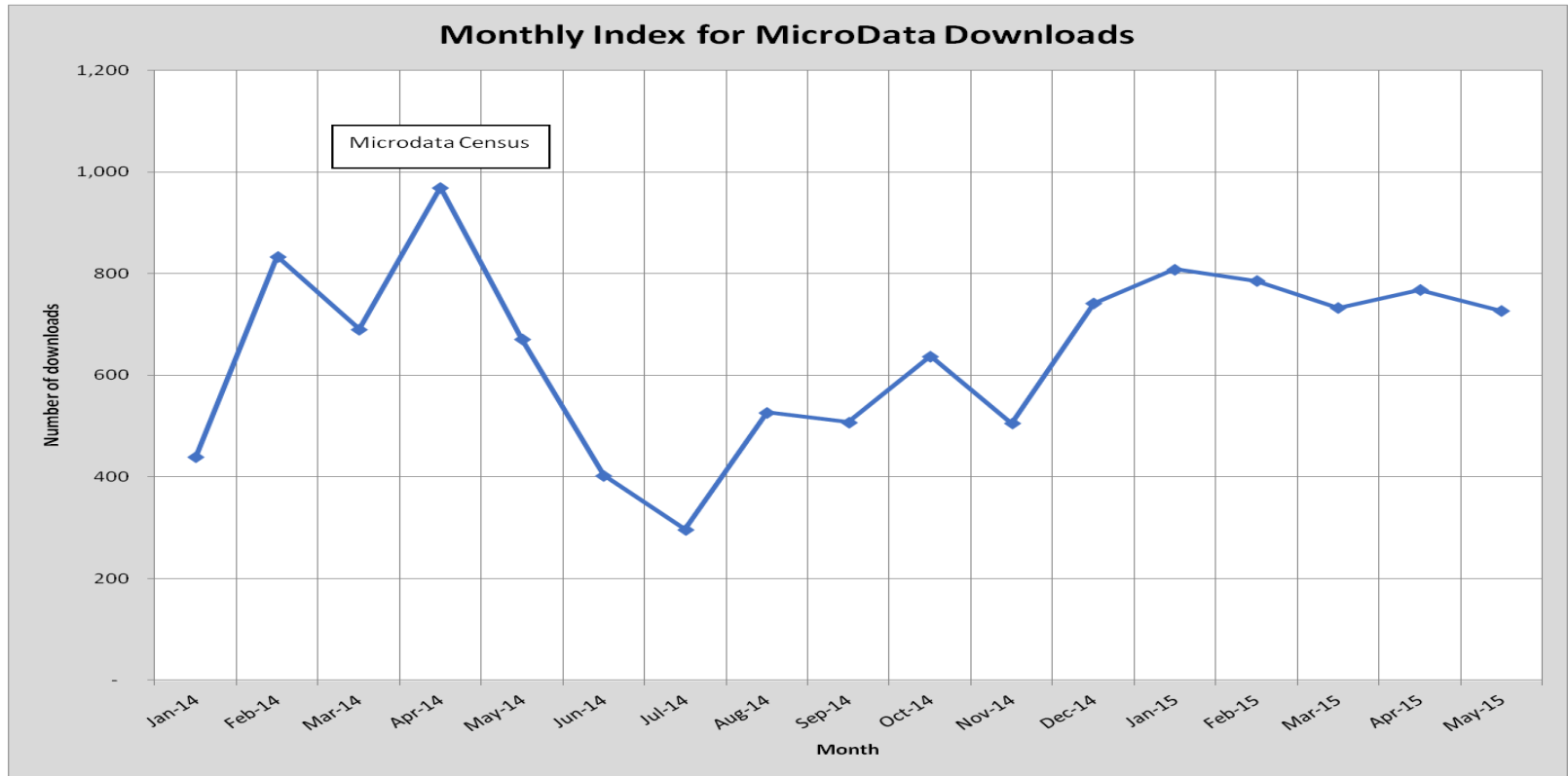
Monthly for Page Views for INSTAT Webpage



Micro data access

- Micro data has become lately (2014) an important tool of data dissemination
 - Labour Force Survey 2007 – 2013
 - Living Standard Measurement Survey 2002, 2005, 2008, 2012
 - Return migration and re-integration survey 2013
 - CENSUS : 3% of data of Population and Housing Census 2011
 - Time Use Survey 2010-2011

Monthly Downloads for Micro data

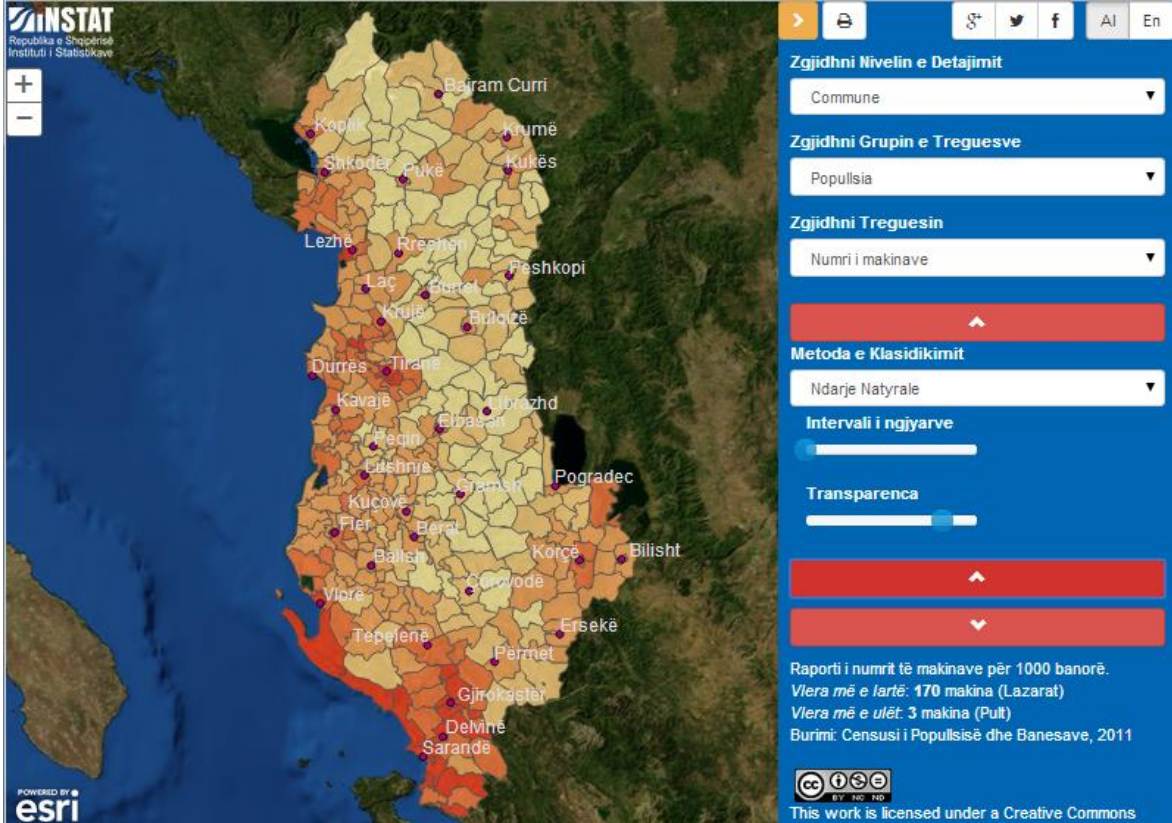


Disclosure techniques for Micro data

- *Anonymisation* – deleting direct identification data
- *Recoding*- variable for indirect identification
- *Sampling* – make a sample proportion of the original data set as a micro data file

WebGis

- A web page (WEBGIS <http://instatgis.gov.al>) is now available in the website of INSTAT from May 2014.



The screenshot displays the INSTAT WebGIS interface. On the left, a map of Albania is shown with various municipalities labeled, such as Koplik, Shkoder, Puke, Kume, Kukës, Lezhe, Rreshen, Feshkopi, Lac, Burrez, Kruje, Blloze, Durrës, Tirane, Kavajë, Elbasan, Pogradec, Fier, Berat, Korçë, Bilisht, Vlora, Tepelene, Fiermet, Gjirokastra, Delvine, and Sarande. The map is overlaid on a satellite-style background. On the right, a blue sidebar contains several sections: 'Zgjidhni Nivelin e Detajimit' with a dropdown menu set to 'Commune'; 'Zgjidhni Grupin e Treguesve' with a dropdown menu set to 'Popullsia'; 'Zgjidhni Treguesin' with a dropdown menu set to 'Numri i makinave'; 'Metoda e Klasifikimit' with a dropdown menu set to 'Ndarje Natyrale'; 'Intervali i ngjarve' with a slider; and 'Transparenca' with a slider. Below these sections, there are two red buttons with up and down arrows. At the bottom of the sidebar, there is a text block: 'Raporti i numrit të makinave për 1000 banorë. Vlera më e lartë: 170 makina (Lazarat) Vlera më e ulët: 3 makina (Pult) Burimi: Censusi i Popullsisë dhe Banesave, 2011'. The sidebar also includes a Creative Commons license logo and the text 'This work is licensed under a Creative Commons'. On the far left, the word 'Map' is written vertically. On the far right, four blue brackets group the sidebar sections: 'Print, Share and Language' (top), 'Geography and Indicators' (middle), 'Classification' (bottom), and 'Map Description' (bottom).

Map

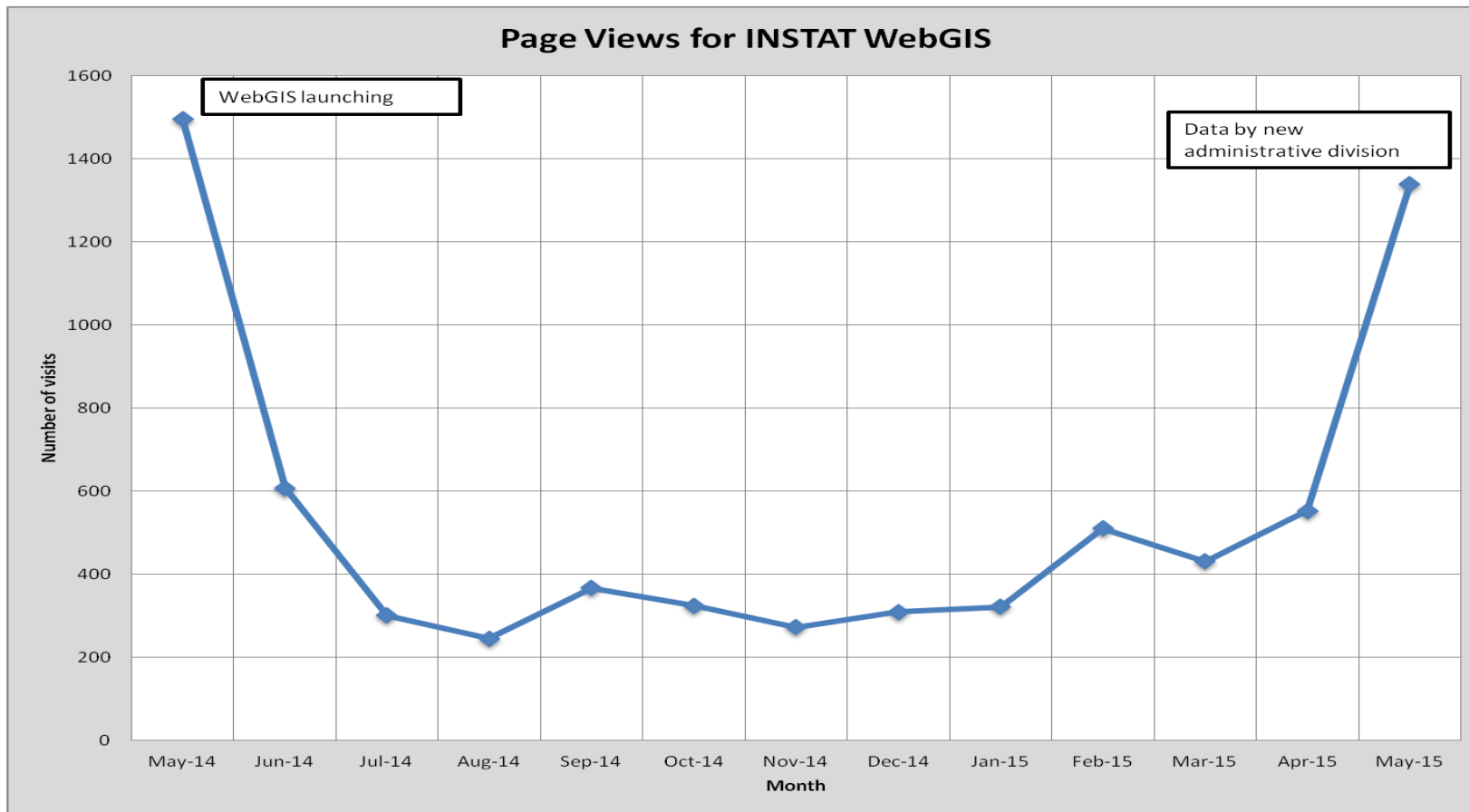
Print, Share and Language

Geography and Indicators

Classification

Map Description

Monthly Page Views for INSTAT WebGIS



Future Plans

- Establishment of a Training School
 - Data providers and users will be trained on:
 - how to access and interpret official statistics
 - use of standard methodology and classifications – understand how specific indicators are calculated and international classification are used for comparability
 - Use of microdata in the on-site laboratory
 - Increasing the microdata dissemination

Conclusions

- Users are more and more interested in disaggregated data.
- A lot of efforts are made in the direction of improving data dissemination
 - access to micro data
 - confidentiality declaration for researchers
- The challenge remain for the Institution
 - finding the right level of disaggregation for producing data that either fulfil user needs or protect data confidentiality

Thank you for your attention !