



The European Statistical Training Programme (ESTP)

**International Conference on Big Data for
Official Statistics**

Capacity Building for Innovation in Official
Statistics

What?

- The **purpose** of the ESTP is to provide European statisticians with continuous training in new methods, techniques and best practices and integrate the application of European concepts and definitions
- The programme is tailored to meet the specific needs of the European Statistical System (**ESS**)
- The programme is **managed** by Eurostat, the Statistical Office of the European Union
- The ESTP **complements** national training schemes and meets the challenges of comparable statistics at European and international level





Who?

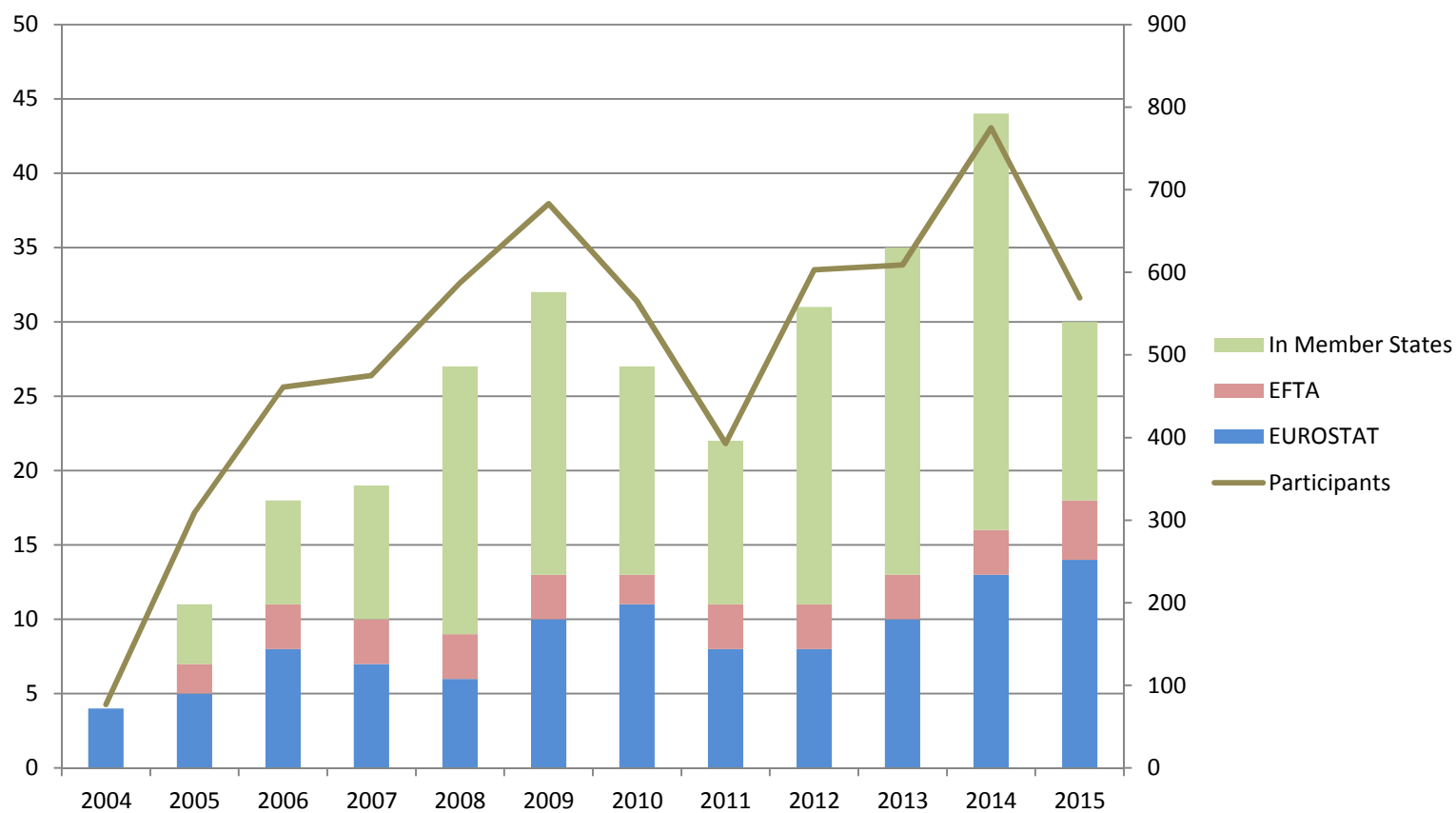
- Officials and employees of NSIs or corresponding CNAs of EU Member States, EFTA countries, candidate countries and Eurostat can apply.
- Occasionally, and on an individual basis, applicants from other administrations, international organisations and Statistical offices of non-European countries may be admitted.
- Participation to all ESTP training courses is free-of-charge. Travel expenses and daily allowances are to be paid by the participant's home organisation



Where?

- **Eurostat premises in Luxembourg**
 - by Eurostat
- **Training sites in other EU countries**
 - by Eurostat contractors, often in cooperation with NSIs
- **Training sites in EFTA countries**
 - by EFTA through EFTA NSIs

Number of ESTP courses and participants (2004-2015)





Programme

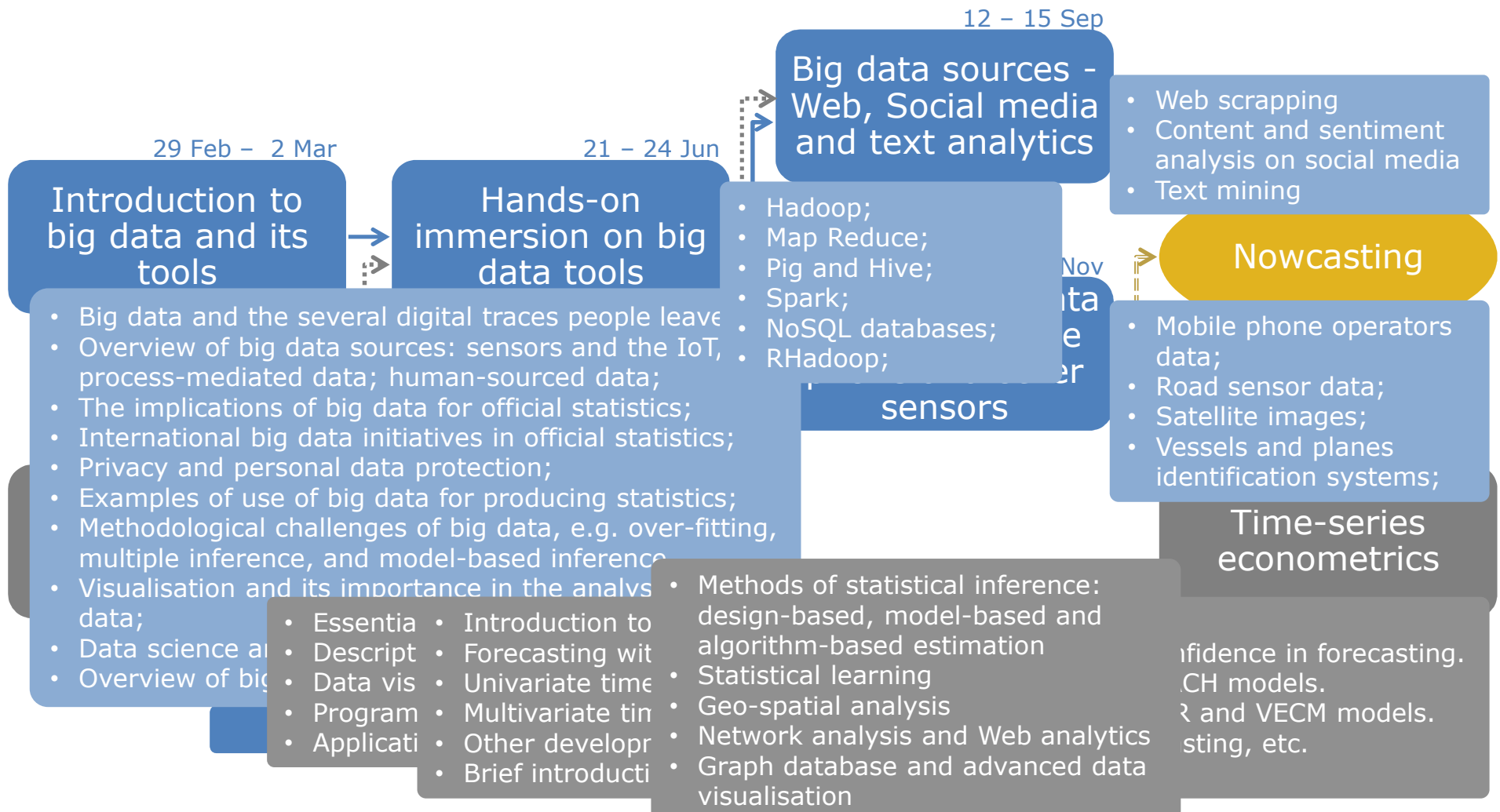
- based on training needs expressed by European NSIs and Eurostat
- in fields like data collection, survey methodology, economic and social statistics, data analysis, quality, dissemination and publication, IT applications etc.
- 2016 ESTP catalogue

Programme 2016

- ***Statistical Methodology***
- *Economic Statistics*
- *Environmental Statistics*
- *Metadata and Classifications*
- *Quality and Dissemination*
- ***New Data Sources***
- *Business Statistics*
- *Other*



ESTP courses supporting big data (2016)





ESTP more information

ESTP National Contact Points

ESS Website

<http://ec.europa.eu/eurostat/web/ess/about-us/estp>

CIRCABC

<https://circabc.europa.eu/w/browse/6ade1ca8-6a06-44bd-bff0-498217d0ec05>



Training Strategy for Big Data

To bridge the Big Data skills gap in European official statistics

- Identification of skills required for the use of big data sources
 - Skills framework
- Inventory of existing skills in Eurostat and in the NSIs in Europe
 - Questionnaire to NSIs and Eurostat
- Analysis of the big data training needs
- To define the training objectives and content
 - Competency-Based Education approach
- Develop a training provision strategy to bridge the skill gap using different instruments



Skills and Training

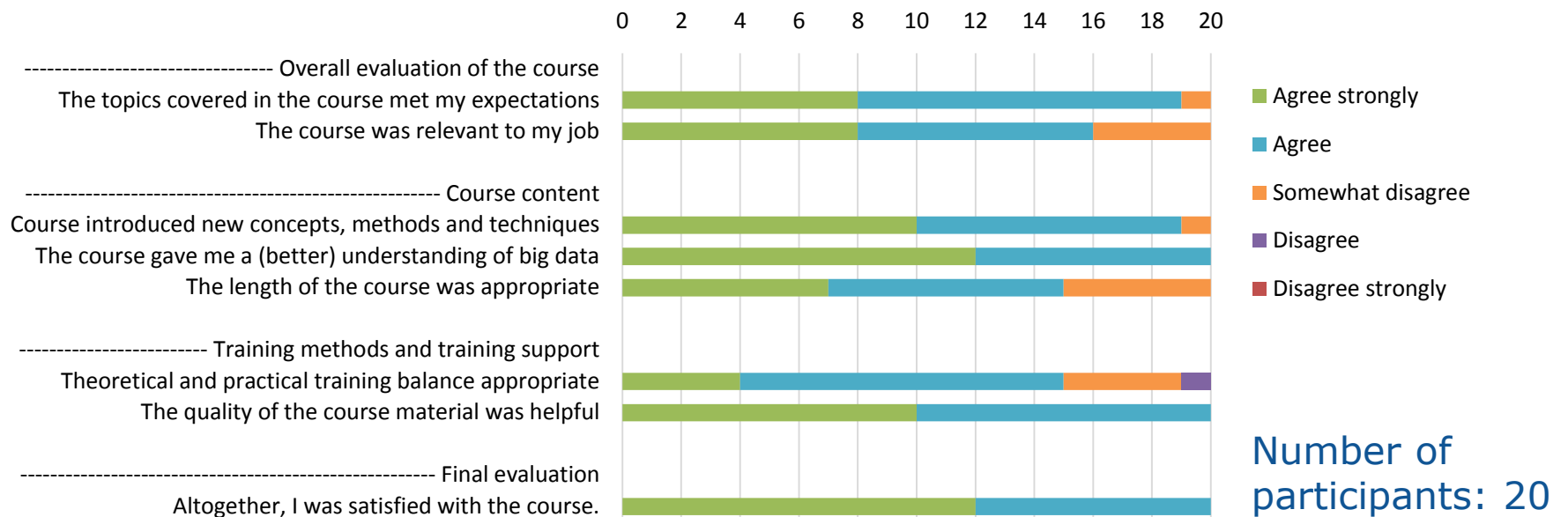
- ESTP Courses
- Hackathon
- Regular presentations
- R user group
- Sandbox environments
- Small groups for data analysis (Hands on Training)



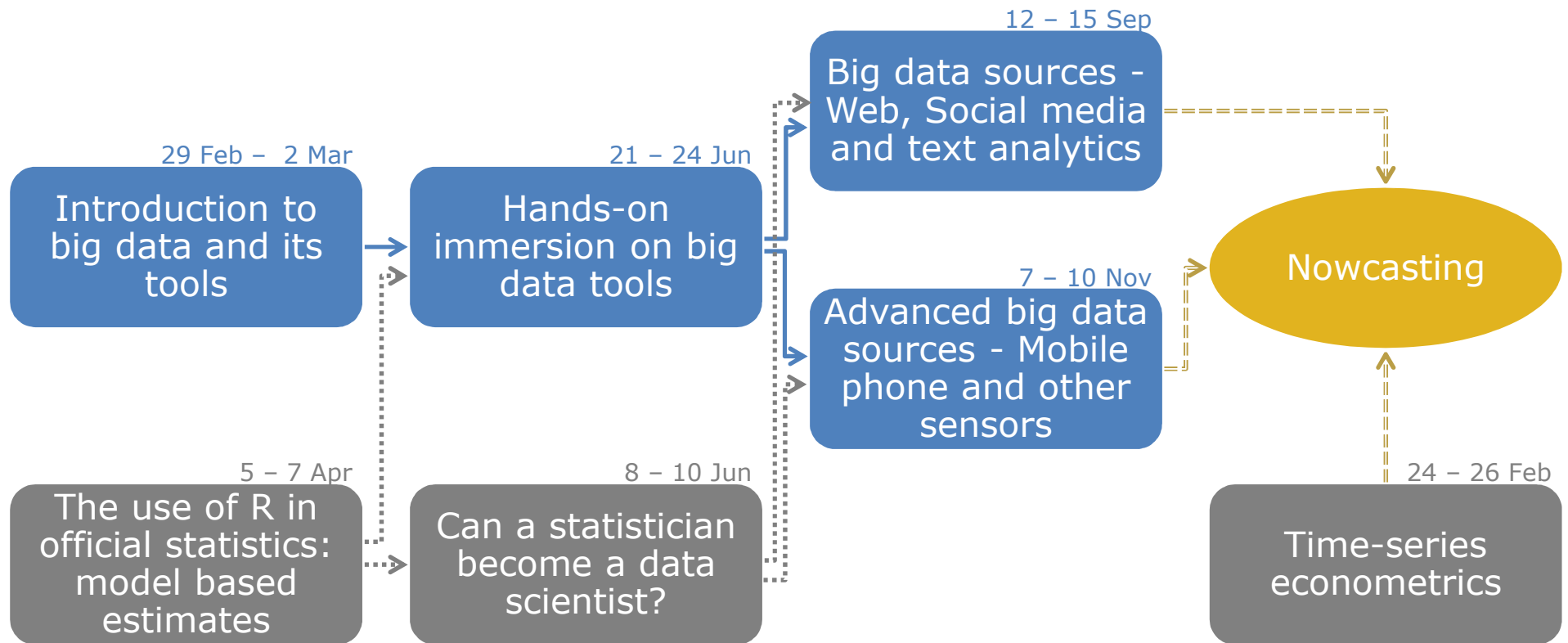
Thank you for atttention!



Introduction to big data and its tools Evaluation (2016)



ESTP courses supporting big data (2016)



 Big data courses

 Methodology courses

 Activity



ESTP courses supporting big data (2016)

Introduction to big data and its tools

- Big data and the several digital traces people leave;
- Overview of big data sources: sensors and the IoT, process-mediated data; human-sourced data;
- The implications of big data for official statistics;
- International big data initiatives in official statistics;
- Privacy and personal data protection;
- Examples of use of big data for producing statistics;
- Methodological challenges of big data, e.g. over-fitting, multiple inference, and model-based inference.
- Visualisation and its importance in the analysis of big data;
- Data science and its role in big data analytics;
- Overview of big data tools, e.g. distributed computing;

Hands-on immersion on big data tools

- Hadoop;
- Map Reduce;
- Pig and Hive;
- Spark;
- NoSQL databases;
- RHadoop;



ESTP courses supporting big data (2016)

Big data sources - Web, Social media and text analytics

- Web scrapping
- Content and sentiment analysis on social media
- Text mining

Advanced big data sources - Mobile phone and other sensors

- Mobile phone operators data;
- Road sensor data;
- Satellite images;
- Vessels and planes identification systems;





ESTP courses supporting big data (2016)

The use of R in official statistics: model based estimates

- Essentials of R
- Descriptive statistics with R
- Data visualization with R
- Programming with R
- Applications of R in an NSI

Can a statistician become a data scientist?

- Introduction to time series analysis.
- Forecasting with time series models, uncertainty and confidence in forecasting.
- Univariate time series modelling: ARIMA, ARCH and GARCH models.
- Multivariate time series modelling: cointegration and VAR and VECM models.
- Other developments : nowcasting, combination of forecasting, etc.
- Brief introduction to state space modelling;

Time-series econometrics

- Introduction to time series analysis.
- Forecasting with time series models, uncertainty and confidence in forecasting.
- Univariate time series modelling: ARIMA, ARCH and GARCH models.
- Multivariate time series modelling: cointegration and VAR and VECM models.
- Other developments : nowcasting, combination of forecasting, etc.
- Brief introduction to state space modelling;