



CES 2017 Seminar on the new generation of statisticians and data scientists
20 June 2017, Geneva

Skills for the new generation of statisticians



#CES2017Skills

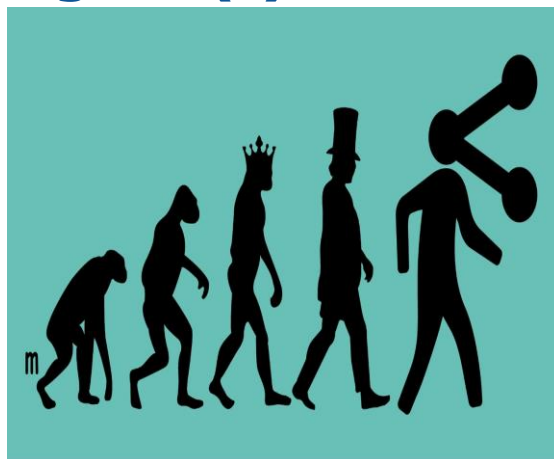
Riikka Mäkinen*, Training Manager, Statistics Finland
riikka.makinen@stat.fi  @riikkamakin

Dario Buono*, Team Leader of Methodology, Eurostat
dario.buono@ec.europa.eu  @darbuo

*The views expressed are the author's alone and do not necessarily correspond to those of the corresponding organisations of affiliation



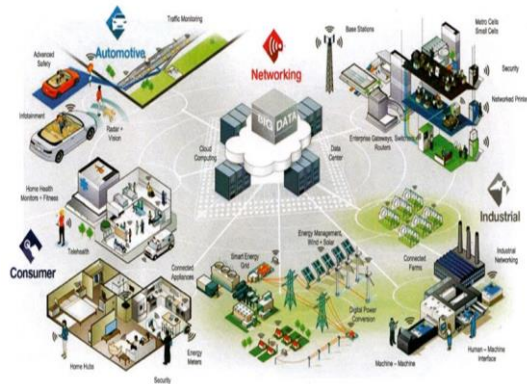
Digital (r) Evolution?



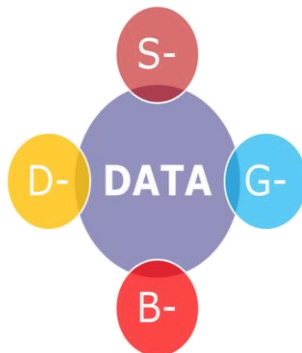


A Data-Driven Life

The Internet of Things



The New Data Lake

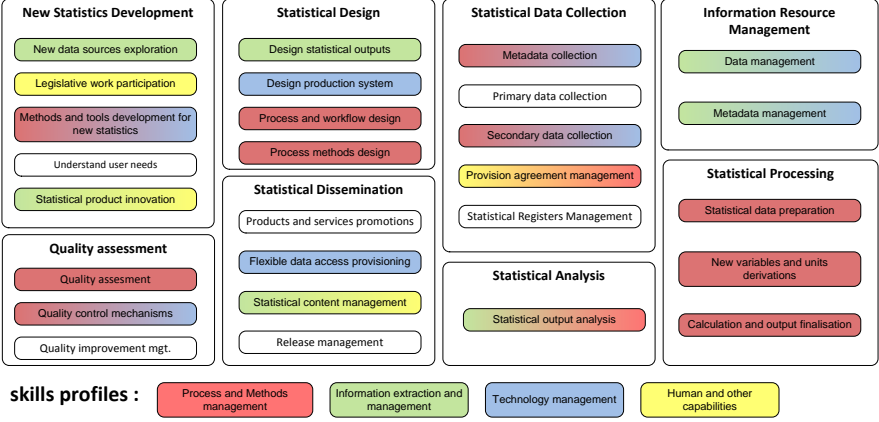


- **S-DATA**
 - surveys
- **D-DATA**
 - administrative data
- **G-DATA**
 - geospatial data
- **B-DATA**
 - Big Data





**New data sources :
key capabilities and skills profile impacted**



MODERN DATA SCIENTIST

Data Scientist, the sexiest job of the 21st century, requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- Machine learning
- Statistical modeling
- Experiment design
- Bayesian inference
- Supervised learning: decision trees, random forests, logistic regression
- Unsupervised learning: clustering, dimensionality reduction
- Optimization: gradient descent and variants

PROGRAMMING & DATABASE

- Computer science fundamentals
- Scripting language e.g. Python
- Statistical computing packages, e.g. R
- Databases: SQL and NoSQL
- Relational algebra
- Parallel databases and parallel query processing
- MapReduce concepts
- Hadoop and Hive/Pig
- Custom reducers
- Experience with xaaS like AWS

DOMAIN KNOWLEDGE & SOFT SKILLS

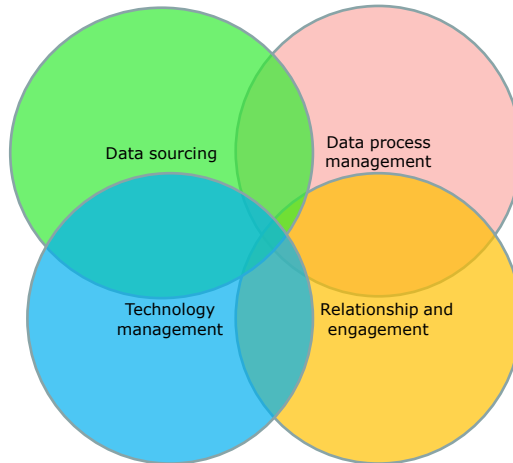
- Passionate about the business
- Curious about data
- Influence without authority
- Hacker mindset
- Problem solver
- Strategic, proactive, creative, innovative and collaborative

COMMUNICATION & VISUALIZATION

- Able to engage with senior management
- Story telling skills
- Translates data-driven insights into decisions and actions
- Visual art design
- R packages like ggplot or lattice
- Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau



Datacy Skills



Means for acquiring skilled statisticians and supporting data science-driven work

Recruiting & long-term personnel planning

- Data scientist's competence profile? Skills are seldom found in one employee
- NSIs need to have a realistic understanding of tasks and competence profiles of the future personnel

Multidisciplinary data science teams

- The competency framework for Big Data team and Big Data team leader (CSO Ireland & High-level Group for MOS)
- Challenges NSIs to find ways of building and maintaining horizontal teams





Means for acquiring skilled statisticians and supporting data science-driven work

Personnel training

- Courses and training programmes in specific competence areas
- [Statistical education in times of big data](#) (Destatis)

Collaboration with universities

- Courses and training programmes
- Trainees and thesis workers

Networks and outsourcing services

- Partnerships with research and data science networks



Means for acquiring skilled statisticians and supporting data science-driven work

Statistical offices need to adopt an innovative culture where experimental activities are commonplace; this is important since future statistical work requires creative thinking and abilities and a will to do things differently.

Supporting innovative culture

- Acknowledging current personnel's interest and potential
- Having experimental activities
- "Letting the past go", statisticians need to identify themselves with new competence requirements and new working methods





Means for acquiring skilled statisticians and supporting data science-driven work

Management and leadership has a crucial role in

- realising personnel planning
- enabling team work in organisation structure
- supporting innovative culture
- allowing time and space for experimental activities
- using participatory ways in defining the common future in statistical offices: the ways statistics are produced, what the needed competences are and how can these be acquired



How can an official statistician become a data scientist?

"Being a data scientist is not only about data crunching. It's about understanding the business challenge, creating some valuable actionable insights to the data, and communicating their findings to the business."

Jean-Paul Isson, Monster Worldwide, Inc.




"The real question is not how a statistician can become a data scientist but how statistical organisation can include "data sciences" in their toolbox and how the needed competences, skills and features can be fused into the professional identity of official statisticians."


*Skills for the new generation of statisticians by
Statistics Finland and Eurostat*





Thanks!

Riikka Mäkinen, Training Manager,
Statistics Finland
riikka.makinen@stat.fi
 @riikkamakin

Dario Buono, Team Leader of Methodology,
Eurostat
dario.buono@ec.europa.eu
 @darbuo

