

**Economic and Social Council**Distr.: General
18 May 2017

English only

Economic Commission for Europe

Conference of European Statisticians

Sixty-fifth plenary session

Geneva, 19-21 June 2017

Item 4 of the provisional agenda

The next generation of statisticians and data scientists**Statistical education in times of big data****Note by the German Federal Statistical Office***Summary*

In times of big data - with the permanent growth of accessible digital data - the requirements regarding competencies of data producers and data analysts are changing. The data landscape has already started to change drastically and the production process, as well as the products of official statistics, will be next. Therefore, the future competency profiles of the staff will be different and the statistical education and training has to develop further. The statisticians of the future will be a team whose members have specialized skills. This includes the knowledge of how to produce official statistics based on new digital data sources as well as a competency profile for big data team leaders. Overall, to produce high-quality official statistics, national statistical offices will need creative concepts in order to develop and recruit the next generation of statisticians.

The document is presented to the Conference of European Statisticians' seminar on "The next generation of statisticians and data scientists" for discussion.

GE.17-08064(E)



* 1 7 0 8 0 6 4 *

Please recycle The recycling symbol, consisting of three chasing arrows forming a triangle.



I. What will be the skills for future statisticians?

1. Do all staff members of a statistical office have to be data scientists in the future? Will a data scientist be the superstar of the big data age (Davenport et al., 2012), as he or she will be able to solve all challenges coming from a big amount of data to produce insights? This includes mathematical and statistical skills, working with different programmes, organizing and mixing data and metadata, visualizing the results in a funny way, leading teams as well as writing articles in journals.

2. The real world is different to this kind of view. The data scientist adequate for official statistics will not be a single person; it will be a whole team with specialized member skills.

3. In 2016, the UNECE High-Level Group for the Modernization of Official Statistics collected a set of team skills necessary to produce official statistics based on new digital data sources as well as a competency profile for big data team leaders (UNECE, 2016).

The identified Big Data team level competencies are:

- Interpersonal and communication skills;
- Delivery of results;
- Innovation and contextual awareness;
- Specialist knowledge and expertise;
- Statistical / Information Technology (IT) skills;
- Data analytical / visualization skills.

The identified big data team leader level competencies are:

- Leadership and strategic direction;
- Judgement and decision-making;
- Management and delivery of results;
- Building relationships and communication;
- Specialist knowledge and expertise;
- Statistical / IT skills;
- Data analytical/ visualization skills.

II. Educational strategies for national statistical offices

4. The face of official statistics will change notably over the next decade. The data landscape has already started to change drastically and the production process, as well as the products of official statistics, will be next.

5. In order to steer this process, well-considered concepts will be necessary. One part of the strategy aspect is statistical education, on two sides. Concepts will need to be in place to educate data producers as well as data users. The following chapters describe how the Federal Statistical Office of Germany tries to tackle these problems.

III. E-learning and blended learning in combination with other forms of training

6. With a staff of more than 2,000 people, the Federal Statistical Office of Germany has a well-developed internal programme of advanced training. In the growing digital world, this programme has changed during the last few years. We have developed a comprehensive e-learning programme on official statistics comprising 18 modules on organisational and statistical-methodological items (see Table 1). Our e-learning platform is based on the open-source software ILIAS.

Table 1:

German e-learning programme on official statistics — 18 modules on organisational and statistical-methodological items	
Module 1:	History of Official Statistics in Germany
Module 2:	Quality of Statistical Processes and Products
Module 3:	Legal Basis of Official Statistics
Module 4:	Cooperation within the German Statistical System
Module 5:	Regional Statistics
Module 6:	European Statistical System and International Cooperation
Module 7:	Statistical Confidentiality
Module 8:	Data Protection
Module 9:	Information Security
Module 10:	Statistical Sources
Module 11:	Statistical Units and Variables
Module 12:	Statistical Indicators
Module 13:	Charts, Tables, Maps (Presentation of Statistics)
Module 14:	Average / Median
Module 15:	Data Collection
Module 16:	Measures of Dispersion
Module 17:	Coefficient of Concentration
Module 18:	Measures of Correlation (e.g. Pearson)

7. At the moment, our official statistics e-learning system is used by more than 1,300 persons (about 600 from the Federal Statistical Office, another 550 from the statistical offices of the Länder (regional offices)). Other national data producers can have access to these e-learning modules as well. In addition, we have also started to open the e-learning platform for universities. Four modules are already available on our webpage under ‘StatistikCampus’ (see <https://www.destatis.de/DE/PresseService/StatistikCampus/StatistikCampus.html>).

8. Our experience with the usage of our e-learning system is quite good. One of the main challenges will of course be the permanent updating of the content.

9. All in all, our aim is to offer an efficient combination of blended learning including e-learning, webinars and face-to-face training. Therefore, we have also included modern subjects like ‘Big data and official statistics’ and ‘Machine learning’ in our internal training course programme. As soft skills are needed also in the world of big data, our training programme includes seminars on intercultural competencies, knowledge transfer, change management, professional diversity management and of course on team work.

10. Interested staff members may also register for about 60 courses per year offered by the European Statistical Training Programme (ESTP) and another 60 courses offered by the

Training in Cooperation with the Statistical Offices of the Länder (called GemFo). These seminars focus on methodological issues.

11. In addition, staff members can take part in specialized courses (external seminars) if they have special training needs (e.g. in the fields of IT or New Media).

12. On the whole, the internal training programme, ESTP and EMOS in combination with specialized courses could offer the content to qualify future statisticians. But all these training opportunities at the different levels should be harmonized as far as possible.

13. What we need in the age of big data is that the processes described in the Generic Activity Model for Statistical Organizations (GAMSO) might be linked with relevant training courses. At UN and European levels the colleagues are already working on this item. Ireland is preparing this for national purposes. In Germany we are also attempting a similar approach by developing competence profiles for specific service classes; this is a big challenge indeed.

IV. Cooperation with universities and academia

14. National statistical offices (NSOs) need creative concepts in order to recruit the next generation of statisticians. Therefore, the Federal Statistical Office of Germany is working together with universities on academic programmes including the European Master of Official Statistics (EMOS).

15. Up to now, EMOS has been a network of 23 Master programmes in 15 countries providing post-graduate education in the area of official statistics at the European level. EMOS is a joint project of universities and data producers in Europe (including Eurostat). There are five German universities which currently offer an EMOS labelled Master programme.

16. Cooperation between NSOs and universities allows influencing the curriculum of university programmes. It has to be clear what the necessary skills are, and what kind of staff structure NSOs will have in the future. Most NSOs in Europe have a mix of staff members with Master and Bachelor degrees, coming from different academic fields. In particular, many of those with Bachelor degrees had only introductory courses in statistics at university. As such, Master programmes should include more aspects of new digital data sources, and introductory courses in statistics will also need to be further developed. This is the case also for the permanent internal training inside the NSOs. Introductory courses and sometimes Master courses in statistics reach students who later may not necessarily work as data producers; many may well be on the side of the data users. Introductory courses should be the focus of statistical literacy programmes run by NSOs (Forbes et al., 2011).

17. We are pleased to see that EMOS has started well. We very much welcome the fact that thanks to Eurostat EMOS webinars have also become available recently. The colleagues from the statistical offices can also participate in these webinars. Our office has successfully tested participation in the first EMOS webinars on 'Response Burden', 'Multiple Sources' and 'Big Data'.

18. We are already actively promoting EMOS. For example, in job and traineeship advertisements we explicitly encourage EMOS graduates or EMOS students to send their applications.

19. To conclude, as far as cooperation with universities is concerned, EMOS is a concrete step in the right direction and yet, for the reasons mentioned above, it is not enough. Consequently, the EMOS ideally should cover introductory statistics courses as well as PhD programmes. EMOS is not only an educational programme; it is also a network

of universities and data producers who work closely together in matters concerning official statistics. The network could and should be used to influence more than the content of master programmes.

20. In addition to the formal contacts with universities, a lot of informal learning takes place in cooperation with academia, e.g. via joint annual meetings between Destatis and representatives from social and commercial research and joint scientific conferences. There are also annual scientific colloquia in cooperation with the German Statistical Society.

V. Education of data users

21. Another challenge in the world of big data with all the new digital data sources (big data as well as administrative data) is the education of data users. Therefore, in cooperation with the German Statistical Society, the Federal Statistical Office is developing new concepts to include statistics in the whole process of education, beginning early at school.

22. Statistical literacy starts with the pupils. Statistics at school has to be more than probability theory in mathematics. What we need is tailor-made products regarding official statistics for teachers and pupils, statistical material for economics, geographic and/or biology courses. One very good example is 'Bringing Data to Life in the Classroom' of the Australian Bureau of Statistics.

23. Destatis is offering tailor-made products for students and universities under the heading 'StatistikCampus' on its webpage. Here you can find for example the e-learning modules mentioned above¹.

24. On our webpage you can also find interactive charts and other visual material which can introduce data users to various themes of official statistics, such as:

- Population pyramids (age pyramids);
- Visual material on price trends, traffic accidents, housing and short-term economic indicators;
- Grid maps and atlases.

25. Statistical literacy needs tailor-made courses for journalists and politicians as well. Destatis is offering a number of such courses (e.g. after a press conference in the field of national accounts).

26. The aim of all these measures is to improve further the trust in official statistics and the quality of our data.

VI. How to get the right staff?

27. With the permanent growth of accessible digital data, commonly denoted as big data, the competence requirements of data producers as well as data analysts are changing. As mentioned above, the future competence profile will be different and the increase in job offers for data scientists as well as iStatisticians shows that also the statistical education has to develop further (J. Ridgway, 2015).

28. For official data producers, questions of human resources are essential for the future (UNECE 2013). First, NSOs need more skills in data science, in combination with other

¹ See: <https://www.destatis.de/DE/PresseService/StatistikCampus/StatistikCampus.html>.

experience, to produce high-quality official statistics. NSOs also face strong competition from companies such as Google or Amazon for the 'best brains'. Well-educated academics with an empirical background are highly sought after, and the price of such resources will increase even further.

29. At the moment, the Federal Statistical Office is changing its Human Resources recruiting system using modern ways of contacting potential talents. It will be a permanent challenge for the future to attract young people and to get the right staff.

VII. References

T.H. Davenport and D.J. Patil, Data Scientist. The Sexiest Job of the 21st Century, Harvard Business Review, October 2012, p 70 ff. <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/>, accessed 20 October 2016.

Destatis (2017): StatistikCampus (including e-learning products) (<http://destatis/StatistikCampus>)

Forbes, S., Camden. M. , Pihama, N., Bucknall, P. and Pfannkuch, M., Official Statistics and statistical literacy: They need each other, Statistical Journal of the IAOS, (2011) issue 27, p. 113 ff.

Köhler, Sabine / Bertram, Christian (2016): Fortbildung in der amtlichen Statistik, in: Qualitätsaspekte der Ausbildung in der empirischen Forschung, p. 119-126

Ridgway, J., Implications of the Data Revolution for Statistics Education. International Statistical Review (2015), 0, 0, 1–22. doi:10.1111/insr.12110.

UNECE High-Level Group for the Modernisation of Official Statistics, 2016 <http://www1.unece.org/stat/platform/display/bigdata/Competency+Profiles>, accessed 20 October 2016.

United Nations Economic Commission for Europe (UNECE), Human Resources Management and Training - Compilation of Good Practices in Statistical Offices, 2013, http://www.unece.org/fileadmin/DAM/stats/publications/HRMT_w_cover_resized.pdf, accessed 20 October 2016

Zwick, Markus (2017): Statistical Education in times of Big Data - Perspectives from an NSI point of view, Paper at the NTTTS 2017, https://www.conference-service.com/NTTTS2017/documents/agenda/data/abstracts/abstract_38.html