



IASE 2020 Roundtable Conference

“New Skills in the Changing World of Statistics Education”
July 6-10, 2020 - Nanjing, China

2020 Roundtable: Rationale and Questions

1. Background

It is a critical time for statistics with new developments in digital and data technologies and the rise of big data. These changes require not only competency in statistics and statistical literacy, but also new and innovative ways of interacting with data, that involve both technical and soft skills (i.e., coding, modeling, visualization, communication, collaboration, creativity, etc.) to handle the challenges of using an enormous amount of information, and to provide insights for improving service delivery, policy making, quality of life and so on. The theme of the Roundtable 2020 “*New Skills in the Changing World of Statistics Education*” arises from the needs of the information rich society we live in today. This changing world of statistics education highlights potential curriculum issues and opportunities at school and tertiary levels.

Building on the previous IASE Roundtables, the 2020 roundtable conference will continue to focus on innovations in statistics education for a wider influence in society. Given the rapid developments in the ways of interacting with data enabled by digital technologies, this roundtable conference aims to discuss development of foundations for statistical literacy from early school years in moving towards big data and “new” skills in relation to data science and analytics, the role of different types of digital tools, and curricular issues in statistics education. The roundtable conference also intends to inspire collaborative efforts with various related disciplines and to build bridges between statisticians, statistics educators, software designers, data scientists and data providers.

2. Rationale

In the constantly changing world of statistics, due to emerging data sources available and rapidly advancing technologies over the past decade, new ways of interacting with data require (1) developing an understanding of big data/open data with strong foundations for statistical literacy and the ability to make good judgments and decisions with the efficient use of data analysis tools (known as data acumen), (2) improving skills to use technology and digital tools efficiently, (3) making curriculum at school and tertiary levels relevant to the needs of today’s information rich society, and (4) building effective collaboration among related disciplines, i.e. statistics, mathematics, data science, computer science, etc., and application areas, i.e. science, government, industry, education and so on. This conference addresses emerging opportunities and challenges in relation to these four areas that are detailed next.

2.1 Big data/open data, data acumen and statistical literacy

With recent advances in computing and communication technologies, data can be found (and collected) nearly everywhere. From various sources, including sensors and transaction systems, very large and messy data, so called “big data”, can be collected. In solving big data problems, data management and data cleaning become essentially important steps before starting the data exploration and can pose new challenges in data analysis. Furthermore the use of digital tools (see 2.2) is inevitable to cope with big data challenges. Open data refer to datasets related to health, crime, migration or economy, which are freely available to everyone, from various sources including government agencies, national statistical offices and so on. In a time of “fake news” citizens can use these datasets or web applications to inform themselves on specific topics. However, the exploration of open data also requires new skills, like downloading the data files, importing files into digital tools or interpreting the data, which are often given in aggregated form. Since usually many variables are involved in the exploration process, multivariate thinking is another relatively new skill required to explore complex datasets in an efficient and adequate way. Within this context, in order to prepare students for tackling big data problems in the future, we need to understand how to develop precursors that would allow them to use their statistical and data science skills in solving such problems. Getting useful information from big data or open data requires an understanding of key statistical concepts and ideas, and competency in statistical literacy. State-of-the-art statistics education research can offer new insights for what fundamental understandings of statistics and statistical literacy are essential in helping students develop the knowledge and skills needed in encountering and making sense of big data.

2.2 Technology and digital tools

For the exploration of big data and open data (see 2.1), the use of digital tools is essential to handle and manage complex datasets. Digital tools can help to collect, clean and explore big data and open data. However using digital tools for these purposes require further skills and, of course, different digital tools can serve different purposes. It is a matter of the complexity of the digital tool versus the potential of the digital tool. On one end of the spectrum, educational software tools like TinkerPlots can be effectively used to introduce learners in explorations of data and data modeling through simulations, and support the learning process for basic understandings of statistical concepts and ideas, but might be limited in some other features for formal statistical procedures such as regression methods. On the other end, there are professional statistical software tools (e.g. R or Python), which offer a wide range of data exploration features of large and messy data but also require efforts in learning programming commands. RMarkdown or Jupyter Notebooks provide a learning environment to structure the data analysis and data exploration workflows. In between there are tools like CODAP as online data analysis platform with the possibility to implement hierarchies in data structures or Excel as typical spreadsheet software, which requires commands for the data analysis procedure. The decision on the use of the digital tool(s) is strongly dependent to its purpose, however all these tools require new skills for learners, teachers, statisticians and statistics educators.

2.3 Curriculum

New sorts of data and data visualization tools available now offer possibilities as well as challenges for curricular development in statistics education from primary school to the tertiary level. In the world of big data, including open data, the increased size and complexity of the data along with larger number of variables and a variety of data types require specific knowledge (in mathematics, statistics and computing), data and computing skills, and dispositions. Around the world, some universities have degree programs specialized in data science/data analytics and some countries have begun to consider the integration of data science across the curriculum at school level or designing a data science curriculum for secondary schools. The common emphasis in these efforts involves solving real and meaningful problems, developing an understanding of core statistical ideas in context, multivariate thinking, statistical modeling, computing, data-driven decision making, multidisciplinary and so on. School curriculum has an important role in providing the groundwork for statistical thinking and statistical literacy needed for these in working with new sorts of data. However, the existing school curricula in many countries are not aligned with these essential ideas and approaches for dealing with the complexity of big data with the use of advancing technology. The content of these curricula needs to be reviewed and revised to reflect the changes in the knowledge, skills and dispositions required in data science. Yet, there might still be issues related to the accessibility to complex datasets and linking the use of such data with content taught within the school curriculum. Furthermore, in the era of big data, the opportunities for incorporating the use of secondary data in the curriculum both at the secondary school and tertiary levels can provide meaningful context for exploring foundational statistical ideas, such as the role of sampling theory in data collected digitally.

2.4 Collaboration

Extracting knowledge and insights from increasingly large, complex datasets for use in various application areas requires a multidisciplinary approach. While statistics is inherently multidisciplinary, data science brings statistics, mathematics and computer science together to solve big data problems in applied fields. Hence collaboration among researchers and experts from these disciplines and domain application areas becomes essential to effectively solve the problems. Collaboration and communication are also one of the 21st century learning skills in preparing students to succeed in the information age. Both the statistics and data science fields can generate new educational approaches to engage students at different ages in practices across different disciplines. Although this multidisciplinary approach provides new opportunities to build connections among mathematics, statistics, science, computing and other data-rich subjects at school level, integration of data science in existing school curriculum might bring some challenges in relation to aligning different content areas and training of teachers.

3. Audience and Potential Participants

We envision a meeting with 40-50 invited participants over 5 days, involving paper presentations accompanied by extended discussions and group work among:

- Statistics educators
- Statisticians
- Cognitive scientists
- Data scientists
- Data science educators
- Technology developers
- Data providers, i.e., national statistics offices and governments
- Computer scientists
- Computer science educators
- Mathematics educators
- Educators from other data rich subjects e.g. sciences, humanities, and social sciences

4. Topics and Preliminary Research Questions

The 2020 IASE Roundtable Conference will focus on the following areas:

4.1. The use of real and meaningful data in teaching and learning statistics

- How can real and meaningful data be used to promote student interest and engagement in statistics?
- How to access real and meaningful data to promote student interest and engagement in statistics?

4.2. The emerging role of multivariate thinking in inferential reasoning

- What are the important ideas and skills in relation to multivariate thinking in statistical inference?
- How can the use of technology promote multivariate thinking with large datasets?
- How does the use of multivariate data nurture students to become data detectives?
- What are innovative tasks or sequence of instructional activities that can be used to support students' multivariate thinking?

4.3. The influence of data science on the school curriculum and introductory statistics courses

- How do we prepare people to cope with the complexity of big data?
- What knowledge, skills and dispositions are required in data science to develop data acumen?
- What is the role of statistics, computation and domain knowledge in data science curriculum?
- What are the ways to engage students in studying data science?

- What are the challenges for integration of data science in the school curriculum/undergraduate statistics courses or designing a data science curriculum at school level/undergraduate statistics?
- What are the ways to support teachers/instructors implementing aspects of data science in schools/at the tertiary level?

4.4. Increasing power of technology and its use for doing statistics and for enhancing learning and understanding of key statistical concepts

- How can new technology tools enrich students' statistical practice, such as gathering organizing, structuring, modeling and interpreting data and making inferences?
- How can digital tools be implemented in classrooms? In which way can the orchestration take place?
- Which digital tools are adequate for different purposes? How can the balance of learning the tool and using the tool be realized?
- How does technology support pedagogical approaches that are needed to enable students to develop data acumen and deepen foundational background in statistics?

4.5. The changing nature of data visualization and implications for the curriculum

- How can different data visualizations help to develop students' conceptual understanding of key statistical ideas?
- How does student-generated data visualizations support the development of students' understanding of key statistical ideas?

4.6. Collaboration with other disciplines to enhance students' statistical understanding

- What can statistics and statistics education learn from other disciplines?
- What can other disciplines (e.g. computer science, computer science education, domain application areas, etc.) contribute to statistics education?
- What are the emerging educational approaches for engaging students at different ages in multidisciplinary of statistics and data science?
- What are the opportunities and challenges in integrating data science in existing school curriculum including mathematics and statistics, science, computing and other data-rich subjects?