

DISCUSSANT SUMMARY: TOPIC 1– THE USE OF REAL AND MEANINGFUL DATA IN TEACHING AND LEARNING STATISTICS

Discussant: Dilcu Barnes & Chair: Sibel Kazak
The University of Alabama in Huntsville, USA & Pamukkale University, Turkey
<dh0066@uah.edu>

PRESENTERS

	Title	Presenter/Co-Author(s)
Short paper	<i>Big Data Affluence in Statistics Applications: A Comparison of Real Life and Simulated Open Data</i>	Nureni Olawale Adeboya (Nigeria)
Long paper	<i>Integrating 'Education for Sustainable Development' in statistics classes: Visual analysis of social and economic data with Gapminder</i>	Martin Andre (Austria)
Long paper	<i>Statistical Literacy, Quantitative Reasoning, and Data Science: Rethinking the Curriculum</i>	Gail Burrill (USA)
Short paper	<i>Where is Waldo in Statistic Class? Using Maps to Explore Modern Data Types</i>	Megan Mocko (USA)
Poster	<i>Using psychological testing to make statistics understandable and meaningful: An exploratory study</i>	Francesca Chiesi (Italia)
Poster	<i>2SDR methodological strategy for teaching and learning statistics in Industrial Engineering students</i>	Blanca Robles (Peru)

PRELIMINARY RESEARCH QUESTIONS

- 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?

KEY DISCUSSION THEMES

In the online asynchronous and two synchronous sessions, four papers and two posters were discussed by participants who have diverse backgrounds. After the short presentation from each author, the discussions mainly structured around the contribution and the educational goal of each paper.

At the beginning of the first Zoom session, Dr. Adeboye discussed their work evaluating relative trustworthiness of models fitted to limited real data and models fitted to simulated data. One of the main comments about this study was self-generated data has been of tremendous assistance in pedagogical teaching and learning of big data. It stimulates timely generation of suitable data and enhances students' learning of data handling procedures. However, there was also a concern about whether simulated data can be a substitute for real data. It has been mentioned that the time series data and that in real life there can be other high impact events happening in different time periods and these may not be presented in the simulated data.

The first session continued with Dr. Andre's presenting their study about a learning trajectory that they designed for middle school students to engage in statistically modelling the phenomenon of sustainability using doughnut model and Gapminder software. Students created concepts of sustainability by identifying various measurable aspects of this real-world problem. Relating data values to sustainability contributed to the process of generating statistical questions. Students also developed intuitive perceptions of statistical concepts through various visualizations. The important contribution of this study is to bring a real-life problem with various global issues within the context of sustainable development into classrooms and to engage middle school students in visualizing multivariate data by using non-conventional data representations.

After the brief overview of the first session, the second session started with Dr. Burrill's presentation explaining what content is important regarding numeracy, statistics, and data in the

school curriculum. Since textbooks often use small artificial data sets, she suggested some resources to help teachers get the real world into their classes. These resources give students opportunity about measurements in contextual situations, make decisions about appropriate data moves as they encounter missing data and data that do not seem to make sense.

After Dr. Burrill's presentation, Megan Mocko described three activities which students used geographical data and maps to interpret multivariate relationships. The activities were not only engaging but also gave students experience working with multivariate data in the data scientific cycle.

We also discussed Dr. Robles poster that highlights the importance of pedagogical approaches emphasizing 'learning to learn' in teaching statistics and students' attitude towards statistics and Dr. Chiesi's poster that used real data obtained from psychology testing to make statistics engaging and meaningful. The study draws attention to the value of engaging psychology major students in exploring kinds of data that they need to examine in their professions to promote their interest and engagement in learning statistics.

To summarize, key findings and conclusions of the discussions for this topic are as follows:

- Addressing real-world issues assist students in developing statistical conception even if they have limited statistical pre-knowledge.
- Using real data assists students in developing deeper understanding of the statistical concepts.
- Using real data help students develop awareness of the usefulness of statistics in their professional training. It is a way that students value the learning of statistics.
- The use of social and economic data in teaching is a welcoming idea. It creates a perfect understanding of relationships among the economic and social variables. It is a way of awakening the search for solutions to world problems. It's a contextualized teaching.
- The use of "big data" in the teaching and learning of statistics is a bet on the future of statistics, it will demand new techniques, so the renewal of curricula of teaching programs is imminent.
- There is a need to rethink teaching curriculum that activates statistical reasoning. Developing statistical thinking in early years is important. Major revisions at the primary, middle, and secondary levels curriculum are necessary to introduce multivariate thinking and data visualizations.