Challenges and Innovations in Statistics Education
Edited by Péter Kovács. University of Szeged, January 2018
ISBN 978-963-306-575-4
Proceedings of Challenges and Innovations in Statistics Education Multiplier Conference of ProCivicStat

ACTIVITIES FOR PROMOTING CIVIC STATISTICAL KNOWLEDGE OF PRESERVICE TEACHERS

Daniel Frischemeier, Susanne Podworny and Rolf Biehler
Paderborn University
dafr@math.upb.de

Critical statistical thinking and a sustainable knowledge in civic statistics is inevitably to become a concerned citizen. This kind of thinking and knowledge can already be enhanced in secondary school. For the implementation of civic statistics in mathematics classrooms in secondary school, teachers themselves have to be well educated in the field of civic statistics. For this purpose, we have designed and realized a university course about civic statistics in the winter term 2016/2017 at the University of Paderborn, where preservice teachers worked on projects and activities in regard to civic statistical contexts. For instance, they have analyzed official open data of the German Statistical Office on the German gender pay gap with digital tools, explored the distribution of net assets in German households and investigated the unemployment situation in different countries in the European Union via Google Public data files. In this paper, we will present some activities and derive first implications for re-designing these activities.

INTRODUCTION

Since decisions in society, politics and economy are often based on civic statistics, responsible citizens in the sense of "Mündigkeit" (responsibility, emancipation) (see Schiller 2017) need statistical and also, more specifically, civic statistical skills. When we talk about civic statistics we mean statistics about key phenomena in society such as employment, health, education, social welfare or inequality (Ridgway 2016; Engel 2017). The process of critical thinking in regard to civic statistics contexts is preferably supposed to start already at school level. In Germany mathematics teachers are not necessarily educated in this special area of statistics education. To teach civic statistics, teachers need not only statistical content knowledge but also pedagogical content knowledge, technological knowledge and a positive stance towards civic statistics. The project ProCivicStat, funded by the ERASMUS+ program of the European Commission, aims at supporting teachers with specific courses, materials, tools, and datasets for teaching civic statistics (www.procivicstat.org).

At the University of Paderborn, we have the following situation. There is a compulsory course on elementary statistics and probability and a compulsory course on didactics of statistics in students' bachelor studies. In their master studies, preservice teachers can choose a seminar which is supposed to deepen and expand their knowledge they have gained in the compulsory courses.

GENERAL INFORMATION

We have designed the seminar "Statistical literacy in mathematics classroom" in the frame of a Design-Based-Research setting (Cobb, Confrey, diSessa, Lehrer, & Schauble 2003) and have realized the first cycle in winter term 2016/2017, the next -second- cycle will take place in winter term 2017/2018. In this paper, we refer to the cycle of winter term 2016/2017 only. Our seminar had 21 participants in the winter term 2016/2017. All participants have been preservice teachers for mathematics in lower secondary school and have been at the end of their studies, having successfully attended the course on "Elementary statistics" and "Didactics of statistics". The seminar consisted of 15 sessions, each session lasted 90 minutes. The main idea of the seminar was to build on the statistical content knowledge our participants have gained in the previous courses so that the participants on the one hand can apply their statistical content knowledge in civic statistics contexts and on the other hand develop a specific pedagogical content knowledge, so that the preservice teachers get prepared to implement civic statistics ideas into mathematics classroom at secondary school. For distinguishing the several knowledge domains, see for example (Wassong and Biehler (2010)).

So, there are learning goals on two dimensions learning goals in regard to statistical content knowledge, and learning goals in regard to pedagogical content knowledge.

Learning goals of the course

Our learning goals with regard to statistical content knowledge are

- to deepen students´ knowledge about reading and interpreting summary statistics and graphical displays (also in the sense of reading beyond data of Friel, Curcio, & Bright 2001).
- to introduce students into statistical concepts and constructs (like correlation and causality or Simpson´s paradox) relevant in civic statistics.
- to introduce into the definition and operationalization of concepts such as unemployment.
- to explore multivariate datasets on the base of given and self-generated statistical questions.

Our learning goals with regard to the pedagogical content knowledge are

- to consider contents in civic statistics across subjects.
- to get to know relevant material (articles, links, tools, datasets, etc.).
- to learn to "simplify" complex situations in civic statistics for classroom use.
- to develop ideas for implementing civic statistics activities in classrooms.

The general goal is to evoke a critical thinking towards statistics and analyses given for example in media reports.

General design principles

The underlying general design principles of our course are similar to the principles of the Statistical Reasoning Learning Environment from Garfield and Ben-Zvi (2008, p.48). For instance, we focus on the development of central statistical ideas, we use real and motivating data sets, we use classroom activities to support the development of our students´ reasoning, we integrate the use of technological tools, we promote classroom discourse and we also promote assessment to monitor the cognitive development of our participants.

REALIZATION OF THE COURSE

In this paragraph, we will describe the realization of some of the sessions in our course.

Sessions 1-3: Introduction

To confront our participants with civic statistics issues immediately we started the course with the task ("Interpret the statistical display in the context of inequality in German net assets") to interpret a complex statistical display showing the distribution of net assets in Germany in the years 2003 and 2012 (see Figure 1).

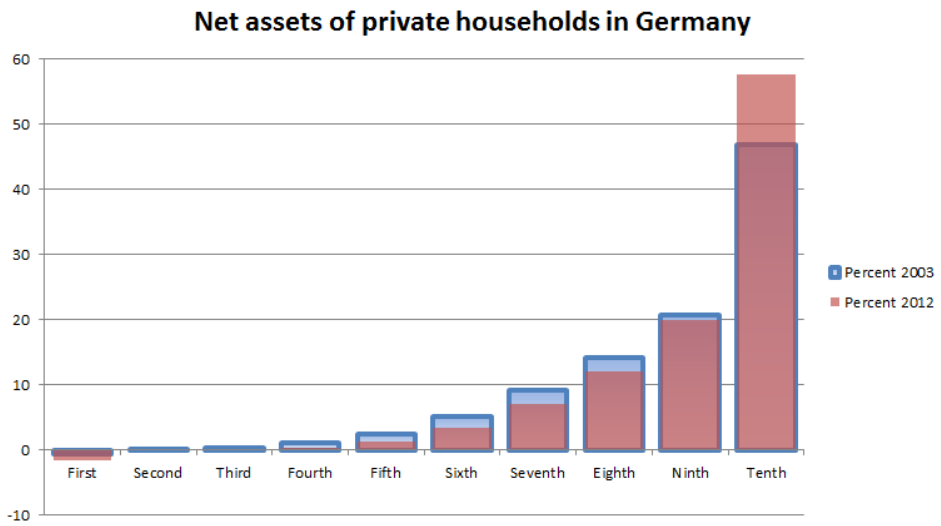
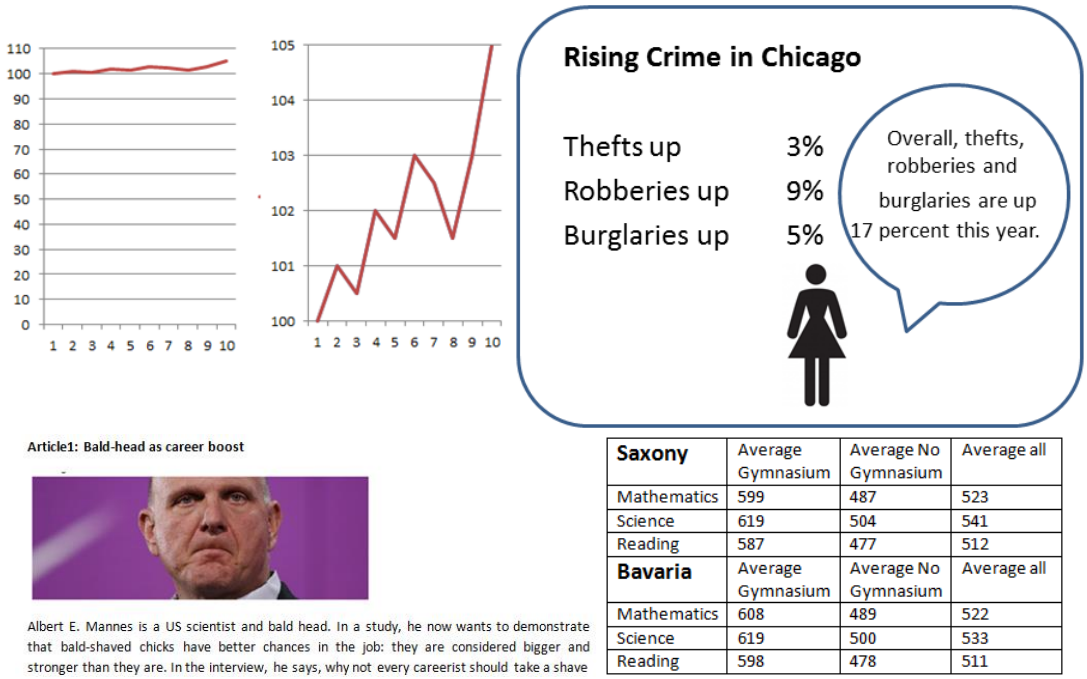


Figure 1: Diagram of the task on German net assets, similar to the diagram in https://crp-infotec.de/deutschland-vermoegensverteilung/

The display in Figure 1 shows the distribution of net assets in households in Germany in 2003 (see blue bars) and 2012 (see red bars). One first competence learners need is to recognize that the bars displaying the situation in year 2003 and 2012 are overlapping. The distribution is divided in ten deciles which are ordered ascending from left to right. First interpretations might be that in most deciles the blue bars are higher than the red bars – but having a look at the tenth decile it is the other way around: the rich people have become even richer in the nine years from 2003 to 2012. Another interpretation might be that there is a big difference in net assets between 90 percent of the German population and the richest 10 percent. Our participants worked on this task in session 1 and 2 and when observing the working processes of our students, it was obvious that this task was very challenging for the students. There is a need to understand the definition of "net assets", of "household", and of "deciles". A difficulty was to compare two overlapping bar graphs showing the growing inequality between 2003 and 2012. In the next session 3, we wanted to refresh the technological Fathom knowledge. The students were familiar with using the German version of the software in their previous courses (Biehler, Hofmann, Maxara, & Prömmel 2011). Our idea was that our participants use Fathom for their explorations and that they refresh their technological Fathom knowledge when exploring a real dataset on leisure time activities of German 11th grade students (Biehler, Kombrink, & Schweynoch 2003). So, for instance in session 3, our participants had to work in pairs to investigate the question "in which way do boys and girls differ in interest with regard to different leisure time activities (e.g., in playing games on the computer)".

Sessions 4-9: Students sessions

In sessions 4-9, students as session leaders were responsible to design and moderate the sessions. We have had sessions on representation of data (session 4), percentages (session 5), percentages II (session 6), correlation & causation (session 7), Simpson's paradox (session 8) and on the concept of unemployment (session 9). In Figure 2, we get an impression on the different tasks and topics in these sessions.



Saxony	Average Gymnasium	Average No Gymnasium	Average all
Mathematics	599	487	523
Science	619	504	541
Reading	587	477	512
Bavaria	Average Gymnasium	Average No Gymnasium	Average all
Mathematics	608	489	522
Science	619	500	533
Reading	598	478	511

Figure 2. Examples of four different activities in sessions 4-9

In the upper-left corner we see an example diagram of session 4 on manipulating statistical displays similar to the diagram in Krämer (2007, p.38-39) – both graphs show the same data with different axes. In the upper-right corner we see a diagram implemented in session 5 showing the wrong use of percentages (see Bauer, Gigerenzer, and Krämer (2014, p.19)). In the lower-left

corner we see an excerpt of a German online magazine (*Spiegel online*) article (see http://www.spiegel.de/karriere/interview-mit-albert-e-mannes-glatze-hilft-der-karriere-a-876879.html - in Figure 2 we present our English translation of it) assuming a relationship between bald-heads and career boost in the session on correlation and causation. Finally in the lower-right corner we see PISA data in the subjects mathematics, science and reading from two German federal states, which turns out to be an instructive example of Simpson's paradox. The common structure of all sessions was that the student session leaders began with an introduction (~ 5 minutes), followed by a presentation that was supposed to refresh relevant statistical knowledge (~ 10 minutes) and providing examples for the theme (~15 minutes). After these inputs there was a working phase where all participants worked on activities in small groups (~ 30 minutes) followed by a plenary discussion of the results (~20 minutes). Each session 4-9 concluded with a reflection on the session (~10 minutes).

Sessions 10-13: Gender Pay Gap project

After attending to the sessions 1-9 and after gaining statistical knowledge in regard to representation of data, percentages, correlation & causation, Simpson's paradox and the concept of unemployment, our aim was to provide our participants with a more complex task. We have chosen the gender pay gap situation in Germany as we identified this as a meaningful topic for young adults.

Specifically, we wanted our students to explore the causes of the unequal pay situation between male and female employees in Germany and we wanted them to become familiar with the concept of the gender pay gap so that our participants are able to distinguish between the adjusted and unadjusted gender pay gap. The unadjusted pay gap means that it is about 23% that men earn more than women in Germany. Furthermore, we wanted our participants to explore the German income structure data set from the German statistical office, to learn to reflect reports in the media critically and to relate them to their own data explorations. In total, we had four sessions dedicated to the gender pay gap project. In session 10 our participants informed themselves by reading media and internet articles about the definition and explanations of the gender pay gap in Germany. For session 11 and 12 we provided our participants with a random stratified sample of all German employees downloaded from the German statistical office and containing about 60,000 cases with variables like gender, wage per month, region of Germany, kind of employment, age, etc. In addition, we provided five topics (profession, function, age, economy and region) according to the variables in the dataset and asked our participants to choose one of these topics – for example the topic "age" (see the precise task in Figure 3).

Project on the Gender Pay Gap - Aspect: Age

Work in teams of two!

Now, you are to carry out a project work on the gender pay gap with your knowledge gained in the seminar. In doing so, you should independently explore the data set for the 2006 Income Structure Survey and get insights into possible explanations for the gender pay gap on the basis of the available data.
You have learned that the differences in income between male and female workers, which are published in the media, have to be interpreted with caution because of the different factors that determine the difference.

Your TASK

In this article (see link below), the focus is on the factor "age", which has an influence on the differences in income. Under this perspective, examine the present data set and work out the extent to which income differences are caused by the aspect mentioned above. In addition, try to explore other aspects that affect differences in income.

Source/Link: http://www.bild.de/ratgeber/job-karriere/gehalt/wie-alter-und-geschlecht-ihr-gehalt-bestimmen-44537794.bild.html
Write a short article and create a PowerPoint presentation that you will present to your fellow students.

Figure 3. Task G*ender Pay Gap Project* (aspect: age)

In this task our participants were asked to explore the gender pay gap data in peers with Fathom and to work out in which way e.g. the aspect age might have an influence on the gender pay gap in Germany. To document their results and to be able to present them to their classmates we asked our participants to prepare a PowerPoint-presentation.

In session 13 the participants have presented their findings to their classmates via their PowerPoint presentations.

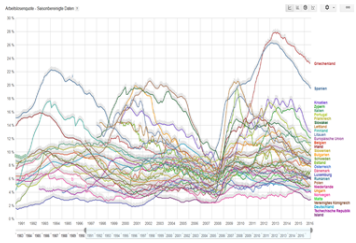
Sessions 14-15: Mini projects

Our course concluded with the sessions 14 and 15 where our participants worked in small groups on mini projects. These mini projects have the intention that the participants can apply all their competencies they have gained during the course in small projects using interactive graphs and tools that can be found on the Internet for free. One example of a mini project can be seen in Figure 4. It uses an interactive visualization from Google Public Data.

Unemployment in the EU

Please go to:

https://www.google.com/publicdata/explore?ds=z8o7pt6rd5uqa6_&ctype=l&strail=false&bcs=d&nselm=h&met_y=unemployment_rate&fdim_y=seasonality:sa&scale_y=lin&ind_y=false&rdim=country_group&idim=country_group:eu:non-eu&idim=country:no:is:cy:uk:hu:cz:es:si:sk:se:ro:pt:pl:at:nl:mt:lu:lt:lv:it:ie:el:fr:fi:ee:de:dk:bg:be&ifdim=country_group&tstart=665190000000&tend=1472508000000&ind=false



and explore the interactive visualization

Questions:

- **Q1:** What trends can you see in the quote of unemployment in Germany from 1991 to 2015? Compare this with the trend of the EU and another country of the EU, chosen by yourself. Inform yourself about the definition of unemployment in the country.
- **Q2:** Think of a further question with regard to other variables (e.g. Age, Gender).
- **Q3:** What do you see? What could be causes for this?

Figure 4. Mini project on the rate of unemployment in the European Union

Mini projects had the goal to analyze a specific topic with open data and free visualizations from the internet to enhance a whole group discussion about the topic in the last session.

In the mini project of Figure 4, our participants should use Google Public Data to compare the development of unemployment rates in different countries of the European Union. In particular, our participants were asked to compare the unemployment rate of Germany to the unemployment rates of other European countries they could choose on their own. The last session covered the presentation of findings of the mini projects and whole group discussions about the context of the presented findings, for example about unemployment rates in different European countries.

CONCLUSION & FURTHER PLANS

We can state that our participants worked statistically on many civic statistics contexts. Especially in the project work of the Gender Pay Gap, our participants have been really engaged. The evaluation, which has been done in form of filling out evaluation sheets at the end of each

session (for details see Biehler, Frischemeier, & Podworny 2017), shows that our participants liked the exploration of German income structure data and the presentation of their findings via Power Point very much.

In the next winter term 2017/2018, we plan to teach a redesigned course. In this course, we will keep the general structure of the previous one, but we plan to support our students in a more concrete way, especially when designing the students´ sessions and to implement more project sessions, since our participants have worked on these activities very engaged, liked them very much and see potential in them to implement these activities in their further teaching.

REFERENCES

Bauer, T., Gigerenzer, G., & Krämer, W. (2014). *Warum dick nicht doof macht und Genmais nicht tötet: Über Risiken und Nebenwirkungen der Unstatistik*: Campus Verlag.

Biehler, R., Frischemeier, D., & Podworny, S. (2017). *Design, realization and evaluation of a university course for preservice teachers on developing statistical reasoning and literacy with a focus on civic statistics*. Paper presented at the World Statistics Congress 61, Marrakech, Marocco.

Biehler, R., Hofmann, T., Maxara, C., & Prömmel, A. (2011). *Daten und Zufall mit Fathom: Unterrichtsideen für die SI mit Software-Einführung*. Braunschweig: Schroedel.

Biehler, R., Kombrink, K., & Schweynoch, S. (2003). MUFFINS: Statistik mit komplexen Datensätzen - Freizeitgestaltung und Mediennutzung von Jugendlichen. *Stochastik in der Schule, 23*(1), 11-25.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design Experiments in Educational Research. *Educational Researcher, 32*(1), 9-13.

Engel, J. (2017). Statistical Literacy for active Citizenship: A Call for Date Science Education. *Statistics Education Research Journal, 16*(1), 44-49.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making Sense of Graphs: Critical Factors Influencing Comprehension and Instructional Implications. *Journal for Research in Mathematics Education, 32*(2), 124-158.

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning. Connecting Research and Teaching Practice*. The Netherlands: Springer.

Krämer, W. (2007). *So lügt man mit Statistik*. München: Piper.

Ridgway, J. (2016). Implications of the Data Revolution for Statistics Education. *International Statistical Review, 84*(3), 528-549.

Schiller, A. (2017). *The Importance of Statistical Literacy for Democracy – Civic-Education by Statistics*. Paper presented at the Challenges and Innovations in Statistics Education, Szeged.

Wassong, T., & Biehler, R. (2010). *A Model for Teacher Knowledge as a Basis for Online Courses for Professional Development of Statistics Teacher*. Paper presented at the 8th International Conference on Teaching Statistics, Ljubljana, Slovenia.