Videos as formative and summative assessment of students' learning in civic statistics

Joachim Engel
Ludwigsburg University, Germany
engel@ph-ludwigsburg.de

Formative and summative assessments provide essential feedback to both teachers and learners, highlighting what is valued in the learning process and measuring the extent of student progress. In this study, we analyze 37 short videos created by teams of two or three students in a Civic Statistics course. Each video follows a structured script that includes: introducing a hot sociopolitical topic, describing the structure and origin of a rich dataset, presenting data visualizations and analyses leading to well-founded conclusions, and offering a critical review. The video analysis provides insights into students' abilities to analyze data using suitable graphical presentations and data moves, as well as their capacity to connect statistical findings with sociopolitical contexts.

STUDENTS CREATED VIDEOS AS PROJECT-BASED LEARNING

The project method has a long-standing tradition in education, dating back to William Kilpatrick's work in 1918. In the field of statistics, however, it was not until around the year 2000 that renewed efforts and a shift in perspectives on statistical learning and content revitalized the concept of project-based learning – PBL (Moore, 1997; Holmes, 1997; Nolan & Speed, 1999). A collaborative, project-based approach to learning statistical applications — especially at the undergraduate level — offers numerous benefits. PBL is inherently complex and learner-centered, relying heavily on formative assessment and ongoing feedback to support students throughout the learning process and foster the co-construction of shared understanding. This feedback is often immediate and depends on both the teacher's ability to monitor student progress and the student's readiness to seek support. PBL emphasizes teamwork and collaboration while cultivating real-world problem-solving skills that are essential for students' future careers (see, e.g., MacGillivray & Pereira-Mendoza, 2011; Porciúncula Moreira da Silva & Samá Pinto, 2014; Santos & César, 2006; Sisto & Petocz, 2012; Phillips & Weldon 2007).

Statistical projects offer unique opportunities for monitoring and assessing student learning. Assessment plays a central role in supporting learning, and any conceptualization of assessment necessarily reflects an underlying view of learning itself (Petocz & Reid, 2007). What educators choose to assess signals to students what is valued in the learning process and thereby shapes their engagement (Chance, 2002). Both formative and summative assessments provide critical information to teachers and learners about which competencies are prioritized and how learning is progressing. For assessment to be pedagogically effective, however, it must be aligned with clearly articulated learning goals. This principle is particularly salient in the context of Civic Statistics, where the educational objective extends beyond technical proficiency to include the capacity to engage with societal data and participate meaningfully in civic discourse. Consequently, assessment strategies in Civic Statistics need to be purposefully designed in order to capture students' ability to interpret data in relation to social and policy-relevant issues. In the context of Civic Statistics, effective assessments should measure students' ability to connect data analysis with its societal and policy implications (Gal et al., 2022a). Incorporating societal and policy relevance into assessments ensures a more comprehensive evaluation of students' statistical reasoning and civic engagement. Assessments should evaluate students' ability to critically investigate and interpret data-driven messages about key social issues. The use of formative and summative assessment to review and improve student learning aims to ensure that students develop the ability to draw implications for society and policy from their data analysis (Gal et al., 2022b).

One innovative approach involves the use of student-produced statistical videos (Schiller & Engel, 2022). In small teams, learners introduce a socially relevant topic, formulate a question, identify and analyze suitable datasets, and draw evidence-based conclusions and summarize their presentation in a short video. Online platforms such as *Gapminder* and *Our World in Data* provide low-threshold access to rich, multivariate datasets and contextual resources that support this process. Throughout the project, the instructor acts as a learning facilitator, offering formative feedback and guidance to support the development of both statistical reasoning and communication skills. Through the creation of "data-driven stories," students learn to frame statistical findings in a manner that is meaningful to a specific

In: Birk, L., Loth, G., Jotzo, L., Binder, K., & Frischemeier, D. (Eds.) (2026). *Statistics and Data Science Education in STEAM. Proceedings of the Satellite Conference of the International Association for Statistical Education (IASE), September/October 2025, Münster, Germany*.

audience, develop didactic strategies for presenting data-based arguments, and begin to critically evaluate the quality and limitations of the data they use. The video production process encourages active inquiry and deeper engagement, serving both formative functions—through instructor feedback during development—and summative functions, as the completed video can be evaluated based on transparent assessment criteria. The purpose of this study is to investigate, (1) how students connect contextual information with statistical reasoning in short self-produced videos, (2) the extent to which they are aware of assessing data quality, and (3) what types of data moves, visualizations, and analyses are evident in their presentations. In addition, the study explores the viability of student-produced team videos as a method of assessment.

VIDEOS AS ASSESSMENT IN A CIVIC STATISTICS COURSE

A central goal of our Civic Statistics course is the integration of methodological learning about data and statistics with subject-based investigations of "hot" real world problems, in particular the development of statistical concepts and data visualizations along the way as tools for deeper understanding of the socio-political context. Participants were teacher students in their second or third year majoring either in mathematics or civics. Main statistical contents being addressed in the course included standard elements of descriptive statistics (location and dispersion parameters, regression and correlation), operationalization of variables and indices, and multivariate data (confounders, Simpson's paradox) and their visualizations. A critical appreciation of data provenance and data collection has also been emphasized. Subject-related "hot" contents covered a variety of topics, ranging from investigating income distributions among countries, the difference between the remuneration for employed men and women (aka gender pay gap), global warming and its environmental impact, to exploring the quality of democracy. Further details about the course have been described elsewhere (Engel et al., 2025; Weber-Stein & Engel, 2025).

As a cooperative project, students were asked to investigate a socially relevant topic using data and to present their findings in a video of approximately 10 minutes. The projects followed the full investigative cycle described by Wild and Pfannkuch (1999), beginning with a socially relevant "hot topic" to introduce and motivate the inquiry. This was followed by the exploration of a rich, multivariate dataset. They then engaged in data visualization and analysis. The videos concluded with data-based answers to the original questions and a critical reflection on the entire process. Throughout the video production process, the instructors played a supportive role by offering guidance and feedback and provided formative assessment. This ensured that students received timely input and had opportunities for improvement. Work on the videos took place during class sessions toward the end of the semester and was completed outside of class.

Literature on the use of student produced videos for assessment is scarce (Schiller & Engel, 2022). This study is guided by the ProCivicStat conceptual framework (Gal et al., 2022a), adapted to the context of video production, as recommended for alternate assessment (see PCS Recommendation 5 in Gal et al., 2022b). An ideal video encompassed an introductory section that contextualized the background and provided an overview of the data by describing its source, the data collection methods, and an explanation of the variables involved. Visualizations and analysis of the data was done with tools like CODAP, Excel, Gapminder or (in a few cases) with R. Discoveries within the data, such as notable patterns, were documented and placed within the appropriate context. By drawing upon relevant calculations and graphical representations, the video derived well founded conclusions. At the conclusion, a critical assessment acknowledged the limitations of the conclusions, and proposed possible avenues for further exploration. Table 1 shows the structure of an ideal video.

Topics ranged from studies on climate change, whether people are happier in democracies, whether referees in European soccer discriminate against dark-skinned players, developments in worldwide freedom of press and many more. Materials and datasets on these and other topics can be found in the CivicStatMap from ProCivicStat, see https://iase-web.org/procivicstat. One group investigated the political and socio-economic effects of oil wealth on exporting countries. We outline this project in more detail here as a best-practice example. It was based on data from the World Bank on ten selected countries, including oil production, the Freedom House Democracy Score, the Gini score on wealth inequality, the Corruption Perception Index from Transparency International, the percentage share of income from oil exports in gross national product and the Reporters Without Borders press

freedom index. After a motivating introduction to the topic (supported by film material from TV reports), the focus was placed on ten major oil-exporting countries. With the help of time series diagrams created with R on oil production volumes, democracy values, the corruption perception index, and several other relevant variables, the group was able to confirm the thesis of the Tübingen political scientist Pawelka that many of these countries have not made the leap to democracy despite some great wealth (see https://www.tagesspiegel.de/politik/demokratie-oder-ol-1984849.html). This video is accessible on YouTube under https://youtu.be/miFDx2lg-s4.

Table 1. Structure of an ideal video on Civic Statistics.

The video aims to include the following

- Introduction and description of the context: What is it all about? Why is this important?
- Identify one or two leading questions
- Characterize the data
  - Where does it come from?
  - How are they collected
  - Why were they collected?
  - How are variables defined
- Which calculations seem interesting?
- Which graphical representations appear informative with regard to the research question?
- Carrying out the data analysis with the help of software
- Interesting discoveries and attempts at explanation, drawing well-founded conclusions.
- Critical review and limitations of the conclusions

ANALYSIS OF THE VIDEOS

What features could be observed in videos? Despite variations in the quality of the videos, several common patterns emerged across all productions. A central goal of the course was to develop students' capacity to engage with societal data as a basis to participate meaningfully in the civic discourse. The qualitative analysis of 37 videos was based on observations and an evaluation scheme comprising seven categories: introducing the topic and its relevance, data provenance, modeling, visualization, computations, discoveries, interpretation, and critical reflection. These categories were derived from the Civic Statistics framework (for details, see Gal et al., 2022a). At the center of Civic Statistics is a topic of high societal and policy relevance, with a focus on its social implications (Facet 1). Data sources used for empirical evidence require critical evaluation and reflection regarding their suitability, credibility, and provenance (Facet 2). The analysis and representation of the data involve computations, visualizations, and modeling (Facets 4–7). Drawing well-founded conclusions and presenting results require interpreting and communicating findings within the socio-political context (Facets 8 and 11), alongside maintaining a critical stance throughout the discovery process (Facet 3).

We focus on the following aspects: (1) relating the statistical explorations to its impact for society and policy, (2) taking a critical stance towards the data and their provenance and (3) developing skills in data explorations.

How did students connect the statistical findings with the sociopolitical contexts?

Students invested considerable effort in motivating their chosen topics employing creative techniques such as incorporating video clips from the internet, excerpts from TV reports and social media podcasts. Some conducted street interviews, mimicked TV newscasts, or used cartoons to introduce their subject matter. While some began their investigations with clearly defined research questions – such as "Which variables influence the Gender Pay Gap?" or "Do autocratically governed countries spend more resources on the military than democratic countries?", others approached their topics more vaguely, asking questions like "Can Germany stop worldwide CO_2 emissions?" or "Are migrants importing diseases?". A few students set out to address particularly ambitious questions, for example: "Does the Climate Change Performance Index CCPI fulfill its self-imposed goal", "Are people in democracies happier?" or "What political and economic effects does oil wealth have on exporting

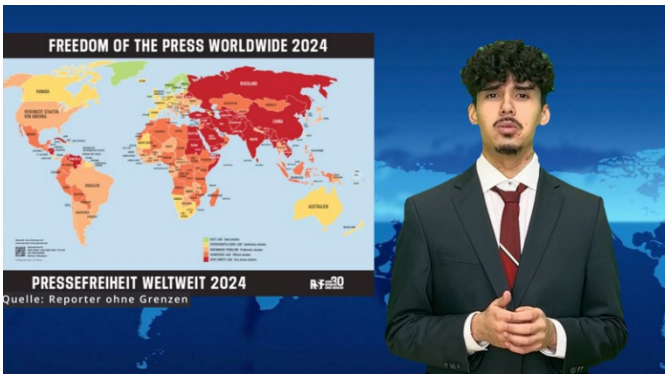

Figure 1. A student mimicking a newscaster to present the reporters without borders freedom of the press index (right). Scatterplot of safety for reporters versus RSF score, separated by continents (left).

countries?". Most data sets represented an interrelated network of many variables as this feature is typical for many data sets describing the state of society or social phenomena. Examples are world happiness data, income inequality data or data on discrimination in European soccer.

Most presentations included thorough contextual explanations of how these variables are interrelated – such as connections between greenhouse gas emissions and global warming, the relation between quality of education, democracy, distribution of wealth, or how people's happiness is connected to social support, wealth and personal freedom. The presentation of the specific theme and the guiding research question was followed by the introduction of a rich, multivariate dataset. For instance, one group explained the role of freedom of the press in democracies, introduced the Reporters Without Borders Press Freedom Index (see Figure 1), and used visualizations to explore how press freedom correlates with various subscores of democratic quality. All videos concluded with a comprehensive summary; however, only about one-third included a critical reflection on the limitations of their analyses and conclusions. Several videos ended by placing the results obtained through data analysis into a larger framework, for examples by highlighting the importance of freedom of press for democracy in quite general terms, although the arguments put forward were not always supported by data. Some videos also ended with recommendations and rules of conduct, e.g., for ecological behavior or with policy suggestions to bridge the gender pay gap, however, without any reference to real data

To what extent did students describe the data and assessed its quality?

Students used a range of data sources for their projects including the World Bank, Eurostat, National Statistics Offices, or data collected by NGO's (e.g., freedom of press index by Reporters Without Borders or the Climate Change Performance Index by Germanwatch) or polling institutions (World Happiness Index). Five of the 37 teams neglected to give information about the data source, while another 17 of the 28 teams, who used the Gapminder platform, cited Gapminder without referencing the original source. Some teams accepted questionable data – such as homicide rates in Berlin from a tabloid newspaper or "projected years to achieve gender equality" (without any source) – without any critical evaluation. Only five teams engaged with issues of data validity or variable construction, for instance by questioning how happiness is measured in the World Happiness Report or explaining how ice age and historic CO_2 concentrations were obtained from Antarctic ice cores. While a few teams conducted basic data cleaning, others overlooked implausible values (e.g., newborns with a yearly income of $ 999,999 in the US census data, or countries reporting an average happiness score of 0.0 that would imply that virtually every person in that poor country would be totally unhappy; most likely these were just missing data). Many variables were based on some latent constructs, such as quality of democracy, level of perceived corruption, personal happiness. Even ostensibly objective measures like yearly glacier mass loss pose measurement challenges. Roughly two thirds of the teams did not address questions of data quality, whereas others demonstrated thoughtful reflection on measurement processes.

What type of statistical activities did students perform in their analysis?

The statistical procedures and data manipulation techniques – referred to as "data moves" (for a detailed definition of data moves see Erickson et al., 2019) varied according to the thematic focus and the characteristics of the multivariate datasets, which included up to approximately 2000 cases and 20 variables. All teams applied some data filtering, typically by selecting specific countries and grouping data into subcategories such as gender to facilitate comparative analysis. Only two teams needed to merge datasets, aligning them based on common variables. Several teams also defined new variables, such as calculating hourly wages by dividing annual income by hours worked.

All teams conducted some form of data summarization, applying measures of central tendency and variability, and computing correlation coefficients or fitting regression lines. A range of visual representations were employed to communicate finding, including boxplots for comparing distributions, contingency tables, and time series graphs. Table 2 gives a numeric summary of the data moves performed by the students.

Although nearly all teams used CODAP for visualization, three opted for R, and most supplemented their analysis with static displays from external sources like *Our World in Data.* Several teams integrated multiple software tool, incorporating visualizations from sources like Gapminder. Time series plots were frequently used to investigate trends. While CODAP's support for regression lines made their inclusion straightforward, only three teams critically examined the suitability of linear models, proposing smoothed curves as alternatives. In one analysis of data from the World Happiness Report (aggregated at the country level) a general linear trend was observed between happiness and the quality of democracy across most regions of the earth, suggesting higher happiness in democratic societies (https://www.worldhappiness.report/). However, in Asia, a U-shaped relationship emerged, with relatively high happiness levels also reported in autocratic regimes. The world happiness dataset asks for searching for confounders, and the observed relationship between happiness and democratic quality may well be influenced by other variables like wealth, economic equity or social support and more. Additionally, the potential impact of outliers on least-squares regression lines was largely overlooked. In particular, analyses of military spending relative to democratic quality demonstrated that the fitted regression line was disproportionately influenced by outliers.

Several investigations focused on comparing subpopulations. While this approach can yield valuable insights, the validity of statistical conclusions becomes increasingly uncertain as sample sizes diminish to become very small.

Table 2. Frequencies of data moves (out of 37).

	Filtering	Grouping/ Restructuring	Merging/ Joining	Summarizing	Calculating	Computing/ Visualizing
Frequencies	29	28	2	30	14	37
Examples	by country, year, glacier, work force	by continent, earlier/ later, gender, skin color, position, soccer league. ethnicity, education level	Data set arms export and military expenditures	Contingency tables, Mean, Median, Interquartile range, Correlation, Regression slope	Murder rate, Hourly wage, Red Card Rate, Oil export per capita	Least square line, boxplots, scatterplots, time series plot, barplots, smooth curves external graphics

The multivariate complexity of many datasets presented specific analytical challenges, prompting some groups to formulate hypotheses and, in some cases, assert causal relationships. Although several teams appropriately considered the role of confounding variables—for example, by examining how educational background or part-time employment might affect the gender pay gap—others inferred causality from correlations without sufficient supporting evidence or proper control for

explanatory factors. For instance, while a correlation between global temperature and atmospheric CO_2 levels is observable, it does not, on its own, constitute definitive proof that climate change is anthropogenic; such a conclusion requires corroborating evidence.

In some cases, groups based their reasoning on isolated national examples—such as asserting that improvements in educational attainment enhance democratic quality, citing South Korea as evidence—yet such generalizations remain speculative without broader empirical support.

DISCUSSION AND CONCLUSIONS

Assessing competencies in Civic Statistics requires evaluating students' ability to critically analyze and interpret data-driven messages on key social issues. The use of student-produced videos as an assessment tool demands a diverse skill set and poses unique challenges, particularly in interdisciplinary courses that bring together students from mathematics and civics backgrounds. While teams rarely included both disciplines, students demonstrated creativity in framing and presenting their topics. Course evaluations indicate that civics students gained confidence in reasoning with data and integrating statistical arguments, while mathematics students developed a deeper appreciation of the societal relevance of their analytical skills.

The videos, produced as final projects of a one-semester course, provided value insights into course design – especially the need to balance basic statistical concepts with political reasoning. Statistical themes such as multivariate reasoning, the influence of outliers on (linear) trends, and the role of sample size in subgroup analysis emerged as areas requiring more emphasis. Likewise, issues of data credibility and the treatment outliers warrant further attention.

Using videos for both formative and summative assessment, requires continuous guidance, feedback, and structured support. While the videos are produced collaboratively, grading still needs to account for individual contributions. Moreover, the final quality of the videos is heavily influenced by the instructors' feedback during the production process.

Designing a video is a creative process that reflects the entire investigative cycle: engaging with a socially significant question, reasoning with data, applying basic statistical tools for visualization and analysis, and critically evaluating findings. The overarching goal is to encourage students to explore issues of societal and political importance based on evidence, using statistical data as a foundation for analysis. A persistent challenge for instructors is to identify motivating topics and datasets rich enough to sustain meaningful inquiry. Interestingly, technical aspects of filming and editing were less problematic than anticipated, though still time-intensive. Moreover, video creation offers valuable preparation for future educators who may employ similar methods in their own teaching.

By creating "data-driven stories" in video format, students learn to adopt a didactic perspective on datasets, craft audience-tailored arguments, and—albeit to a limited extent—evaluate data quality and the soundness of conclusions. Striking the right balance between teaching statistical concepts and fostering critical reflection remains an ongoing challenge. Video projects foster teamwork, motivation, and media literacy (Schiller & Engel, 2022) but require substantial effort from both students and instructors. Their success depends on (1) the careful selection of engaging topics, (2) access to rich and meaningful datasets, (3) well-prepared background resources, and (4) proactive instructor coaching and support.

VIDEOS

As illustration, two student produced videos are available on YouTube: The above mentioned video exploring the relationship with oil export and democracy https://youtu.be/miFDx2lg-s4 and a video on Melting of Glaciers https://youtu.be/xFBvuS_AOro.

REFERENCES

Chance, B. (2002). Components of statistical thinking and applications for instruction and assessment. *Journal of Statistics Education*, *10*(3). https://doi.org/10.1080/10691898.2002.11910677

Engel, J., Weber-Stein, F., & Kleinknecht, D. (2025). Zivilstatistische Datenkompetenz in Zeiten von Fake News: Konzept, Unterrichtsdesign und Implementierung im Lehramtsstudium. *Mathematica Didactica*, *48*(2). https://doi.org/10.18716/ojs/md/2025.2245

Erickson, T., Wilkerson, M., Finzer, W., & Reichsman, F. (2019). Data Moves. *Technology Innovations in Statistics Education*, *12*(1). https://doi.org/10.5070/T5121038001

Gal, I. & Garfield, J. (Eds.) (1997). *The Assessment Challenge in Statistics Education*. IOS Press.

Gal, I., Nicholson, J., &. Ridgway, J. (2022a). A Conceptual Framework for Civic Statistics and Its Educational Applications. In J. Ridgway (Ed.), *Statistics for empowerment and social engagement* (pp. 37–66). Springer.

Gal, I., Ridgway, J., Nicholson, J., & Engel, J. (2022b). Implementing Civic Statistics – An agenda for action. In J. Ridgway (Ed.), *Statistics for empowerment and social engagement* (pp. 67–96). Springer.

Holmes, P. (1997). Assessing project work by external examiners. In I. Gal & J. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 153–164). IOS Press.

Kilpatrick, W. H. (1918). The project method: The use of the purposeful act in the educative process. *Teachers College Record, 19*(4), 319–335.

MacGillivray, H. & Pereira-Mendoza, L. (2011). Teaching Statistical Thinking Through Investigative Projects. In C. Batanero, G. Burrill, C. Reading (Eds.), *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education. New ICMI Study Series (Vol 14)* (pp. 109–120). Springer. https://doi.org/10.1007/978-94-007-1131-0_14

Moore, D. (1997). New Pedagogy and New Content: The Case of Statistics International. *Statistical Review*, *65*(2), 123–137.

Nolan, D. & Speed, T. P. (1999). Teaching statistics theory through applications. *The American Statistician*, *53*, 370–375.

Petocz, P., & Reid, A. (2007). Learning and assessment in statistics. In B. Phillips & L. Weldon (Eds.), *Proceedings of the IASE/ISI Satellite on Assessing Student Learning in Statistics*.

Phillips, B., & Weldon, L. (Eds.) (2007). *Proceedings of the ISI/IASE Satellite on Assessing Student Learning in Statistics.* International Association for Statistical Education.

Porciúncula Moreira da Silva, M. & Samá Pinto, S. (2014). Teaching statistics through learning projects. *Statistics Education Research Journal*, *13*(2), 177–186. https://doi.org/10.52041/serj.v13i2.289

Santos, N. & César, M. (2006). Project work in statistics: Statistics learning as a tool to help students' knowledge about their educational community. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education. Proceedings of ICOTS 7.* International Association for Statistics Education.

Schiller, A. & Engel, J. (2022). Implementing Civic Statistics in Mathematics Teacher Education. In J. Ridgway (Ed.), *Statistics for empowerment and social engagement* (pp. 395–416). Springer. https://doi.org/10.1007/978-3-031-20748-8_16

Sisto, M. & Petocz, P. (2012). Engaging students with statistics using collaborative project-based approaches. In I. Solomonides, A. Reid, O. Petocz (Eds.), *Engaging with learning in higher education* (pp. 297–317). Libri Publishing.

Weber-Stein, F, & Engel, J. (2025). Implementing Statistical Literacy in Civics Teacher Education. *Journal of Political Science Education.* https://doi.org/10.1080/15512169.2025.2483786

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223–265.