Statistical Science for a Better World.

JUNE 2026

vISIon

The Magazine of
The International Statistical Institute

OFFICIAL STATS

Race, ethnicity and data:
what governments choose
to measure, and why it matters.

DATA FOR ALL

Building statistical literacy
for everyone, everywhere,
at every stage of life.

REDUCED INEQUALITIES

Data for a fairer world:
who counts, and
who is counted.

IN THIS ISSUE

Ethnicity & official data • Statistical Literacy • AI & slavery

Using secondary data to engage learners in researching social issues

The case of the CR2 Climate Explorer and water scarcity

Abstract. Drought and water scarcity are phenomena that are often confused, yet they have different causes: drought has a natural origin, while water scarcity is caused by water use. In this article, we present a proposal to address these issues through the analysis of secondary data from the CR2 Climate Explorer. By proposing different types of research questions and following statistical research processes, we illustrate how the problem of drought and water scarcity could be addressed using data from the city of Petorca (Chile). We believe that this type of activity fosters student engagement and action regarding socio-environmental issues affecting different communities, allowing them to reflect on how data serves as a source of information for decision-making.

Francisca Ubilla
Assistant Professor at Pontificia Universidad Católica de Chile. Researches statistics education and civic statistics in teacher training.

Valentina Giaconi
Associate Professor at Universidad de O'Higgins. Holds doctorates in Mathematical Modelling and Education Sciences.

Raúl Valenzuela
Assistant Professor at Universidad de O'Higgins. Specialist in extreme hydrometeorological events, atmospheric rivers and radar meteorology.

Felipe Ruz
Associate Professor in the Department of Statistics at the Pontificia Universidad Católica de Valparaíso. Researches statistical and civic literacy in maths education.

Introduction

Education has a key role in forming citizens, present and future, who can make sense of complex social phenomena, with statistics offering a powerful interdisciplinary lens for such development (Engel et al., 2021). The analysis of different phenomena using data should occur through the connection between global and local issues; that is, we should understand the phenomena that affect us globally through the critical interpretation of data, while simultaneously reflecting on and gathering local evidence regarding how these issues impact the places we inhabit daily.

One of the phenomena that has been affecting a large part of the population for several years is water scarcity and drought, phenomena that appear similar, but are fundamentally different. "Drought" is a natural phenomenon caused by large-scale climate variability and cannot be prevented through local water management. "Water scarcity" refers to the unsustainable long-term use of water resources, which water managers can influence (Van Loon & Van Lanen, 2013). This issue also aligns with Goal 6 of the United Nations' Sustainable Development Goals (SDGs), which calls to "Ensure availability and sustainable management of water and sanitation for all." Water also cuts across nearly every other SDG as well. This makes water an ideal topic for engaging learners in that is related to both global and local issues. Based on this, the Center for Climate and Resilience Research (CR2) offers the opportunity to analyze the phenomenon of drought using a database derived from various weather stations belonging to the observation networks of the General Water Directorate (DGA), the Chilean Meteorological Directorate (DMC), and Chile's National Agroclimatic Network (RAN, Agromet).

The availability of water data in Chile has allowed us to develop activities based on the statistical problem-solving process proposed in GAISE II (Bargagliotti et al., 2021), which aims to encourage students to formulate statistical research questions based on secondary data. However, based on our experience and what the liter-

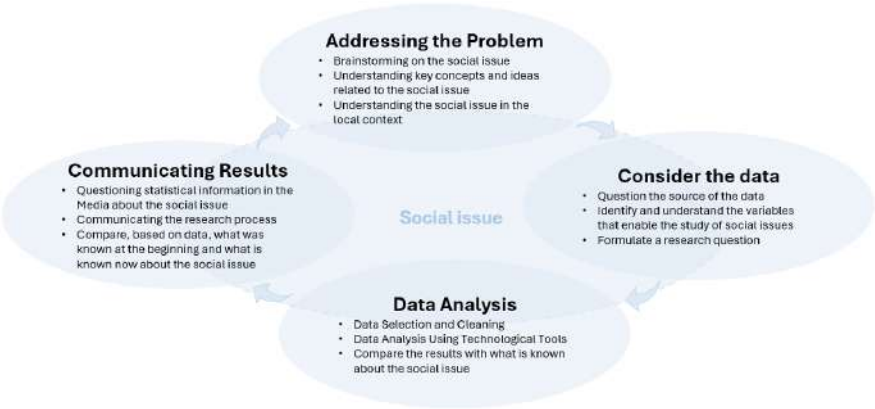

Figure 1. Statistical research process based on real data on social issues

ature reports (Ubilla and Gorgorió, 2023), formulating statistical questions is a complex process, which can be further complicated if students must generate such questions from an existing dataset. For this reason, we present ideas on how to address the issue of drought using a database by following the statistical research process, showcasing three examples of research questions that adhere to the guidelines by Arnold and Franklin (2021) on what makes a good statistical question.

How can we understand social issues through data?

Based on the GAISE II proposed statistical research process (Bargagliotti et al., 2021), in Figure 1, we outline some aspects of each phase, keeping in mind that the topic at hand concerns a social issue using secondary data.

First, we believe that when *approaching this issue*, it is necessary to identify what students already know about it and then introduce key concepts and ideas that will help them understand the phenomenon. Next, students are expected to learn how this issue affects their local community and what various institutions have done to address it. In this phase, they are integrating information about the social issue that stems from their own beliefs and experiences with theoretical, global, and local information on the issue. During the *consideration of data*, students are expected not only to question the source of the data but also to be able to identify the variables and nature of the data that help them understand the issue. The link between the previous stage and this one is established through the formulation of a research question. Because databases on social issues tend to be extensive and complex, during the *data analysis* phase, students are expected to clean the data so that, after analyzing it, they can compare the results with what they initially knew about the issue. Finally, it is hoped that, based on the results obtained, students will be able to compare them with information found in the media and integrate what they knew initially with what they observed from the data, and ultimately *communicate their process* to an audience they deem relevant.

> *"Education has a key role here, in terms of forming citizens, present and future, who can make sense of complex social phenomena, with statistics offering a powerful interdisciplinary lens."*

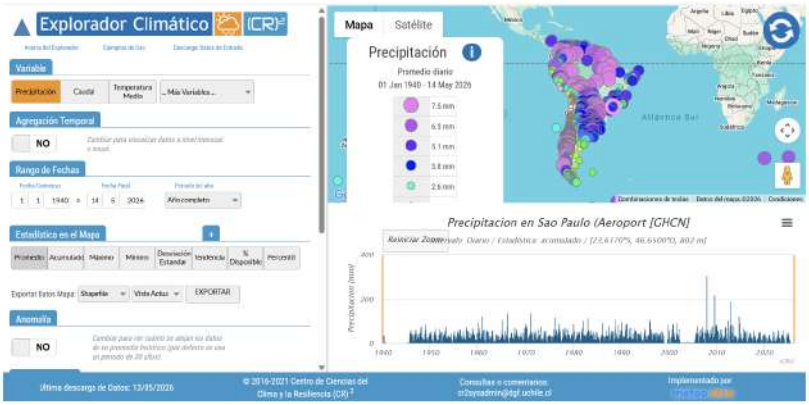

Figure 2. Home page of the CR2 Climate Explorer (https://explorador.cr2.cl/)

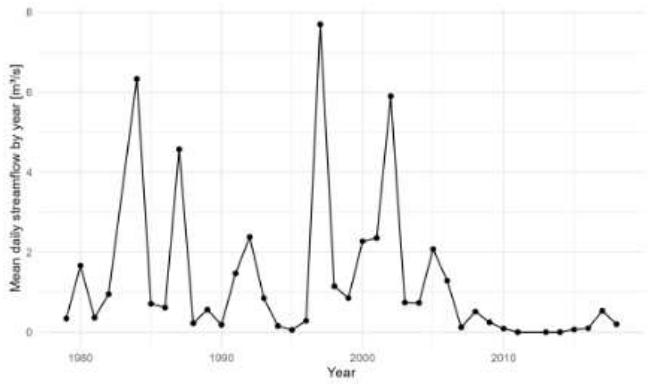

Figure 3. Annual mean daily streamflow of the Petorca River. The figure shows the yearly average of daily streamflow values, expressed in cubic meters per second [m³/s].

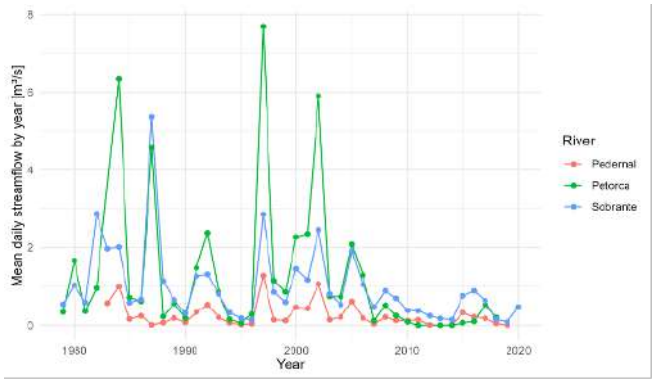

Figure 4. Annual mean daily streamflow by river. Values are expressed in cubic meters per second [m³/s].

What data can we find in the CR2 Climate Explorer?

The Climate Explorer from the Center for Climate and Resilience Research (CR2) compiles observational records of streamflow, temperature, and precipitation from climate and hydrological stations. Streamflow is expressed in cubic meters per second (m³/s), temperature in degrees Celsius (°C), including maximum, minimum, and average temperatures, and precipitation is the daily accumulation in millimeters (mm). The platform allows users to filter data by date range and time scale, view it in map or time series format, and export it for analysis. It contains data from 669 weather stations in Chile, South America, and Antarctica that provide temperature data; 809 weather stations in Chile alone that provide flow rate data; and 1,241 weather stations in Chile, South America, and Antarctica that provide precipitation data, that have been collected over the years. Figure 2 shows the Climate Explorer interface, where each point corresponds to a data station and at the bottom, you can view a time series of precipitation data. Click the following link to access the CR2 data viewer: https://explorador.cr2.cl/

In this article, we aim to demonstrate how to use local CR2 data to answer simple research questions, thereby illustrating how to work with this data in school settings or introductory statistics courses for undergraduate programs.

With this in mind, we decided to study the Petorca River due to the documented problems regarding access to and availability of water in this region. In particular, the National Institute of Human Rights (n.d.) describes Petorca as a case of water crisis linked to the depletion of surface water and restrictions on the use of groundwater.

In the CR2 Climate Explorer, we identified one station corresponding to the Petorca River and two stations associated with the Sobrante and Pedernal Rivers, both of which flow into the Petorca River. Since the issue under analysis is directly linked to water availability in the river system, we used streamflow and precipitation as our variables.

The analysis will use the data available in the three selected time series, covering the period 1979–2018. Although the original records are daily, annual averages for streamflow and annual totals for precipitation will be used to

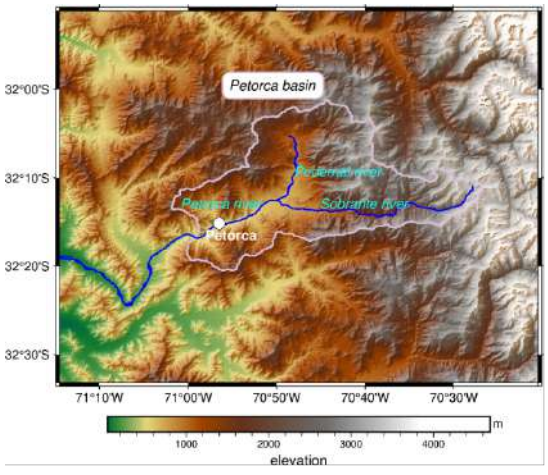

Figure 5. Topographic map showing the location of the three rivers in the mountainous landscape.

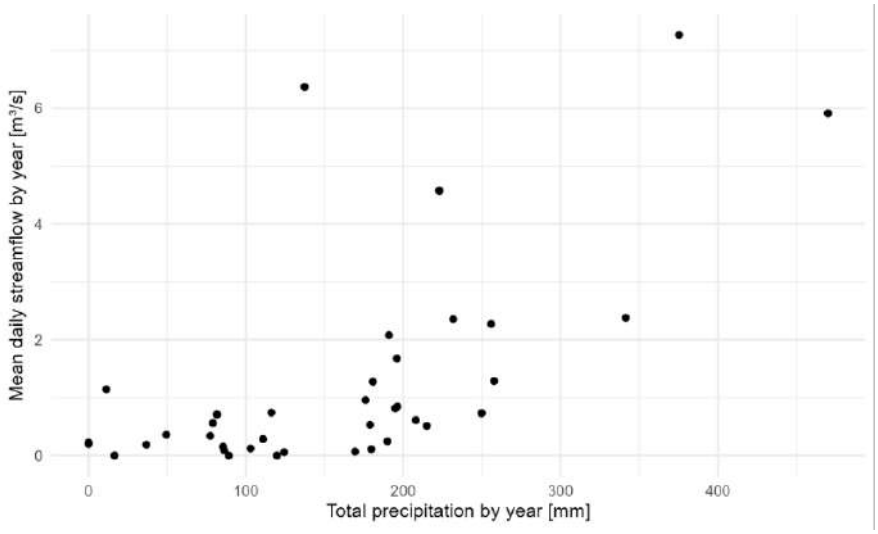

Figure 6. Relationship between annual precipitation and annual mean daily streamflow in the Petorca River.

simplify the analysis. After selecting the data, we downloaded the database and used the R software, version 4.5.3 to analyze it.

What does the data tell us about the drought in Petorca?

Our goal is to illustrate the approach and methods for addressing statistical research questions. To do this, it is essential to define the focus of the inquiry. In this proposal, we present three types of research questions: descriptive, comparative, and relational. The final question is presented along with the consolidated analyses; however, students are not expected to arrive directly at that final question or those results. The idea is to arrive at these types of questions through a dialogic process with teachers and peers.

> *"We should understand the phenomena that affect us globally through the critical interpretation of data, while simultaneously reflecting on and gathering local evidence regarding how these issues impact the places we inhabit daily."*

Summary investigative question. Descriptive or summary investigative questions refer to questions that characterize how a variable behaves within a group (Arnold & Franklin, 2021). In this context, such a question would begin by considering whether we are interested in describing a specific variable in one or more specific sectors. Considering the selected issue, a relevant initial question is How has the streamflow of the Petorca River behaved over the years? For each year, we calculated the mean daily streamflow in order to obtain simpler data. We plotted this variable against the year of the data, generating the graph shown in Figure 3.

We can observe a considerable variability between 1978 and 2006, followed by a very strong downward trend, revealing an important issue. This trend may be due to broader climatic phenomena or to human management of water resources.

Comparison investigative question. Comparison investigative questions refer to comparing how a variable behaves across several groups (Arnold & Franklin, 2021). To define them, it is necessary to identify the groups of data that we are interested in comparing. In the previous example, the issue affecting Petorca was made evident; however, it is not clear whether this pattern is due to broader climatic phenomena or to water management. To explore this further, we focused on the Pedernal and Sobrante rivers, which flow into the Petorca River. These rivers originate closer to the Andes Mountains and farther from the valley and productive activities, so their behavior could be more strongly associated with climatic phenomena than with direct human intervention (Figure 5).

We therefore asked: *Does annual mean daily streamflow tend to differ among the Petorca, Pedernal, and Sobrante rivers?* To address this question, we generated a graph that shows the yearly average of daily streamflow values for each river, allowing comparison of streamflow trends over time (Figure 4).

We observe that, until 2006, the Petorca and Sobrante rivers show considerable variability, whereas the Pedernal River consistently has the lowest streamflow. After 2006, the three rivers show a tendency toward lower streamflow, which suggests the influence of a climatic drought phenomenon affecting all of them. However, until 2006, the Petorca River tends to show higher streamflow than the other rivers; after this year, its streamflow is lower

than that of the Sobrante River and similar to that of the Pedernal River. This latter phenomenon suggests an effect of human action, in addition to broader climatic factors.

Association investigative question. Association investigative questions seek to explore whether there is an association between two variables (Arnold & Franklin, 2021). Continuing with the previous account, we considered that another way to understand whether the streamflow of the Petorca River has changed due to the megadrought or to human management is to compare streamflow with precipitation. Precipitation is less directly controllable by human action. Therefore, we asked: *Is there an association between annual precipitation and annual mean daily streamflow in the Petorca River?* To address this question, we plotted yearly precipitation values against yearly average daily streamflow values for the Petorca sector (Figure 6). We also calculated the Pearson correlation between both variables, which was $r = 0.70$, $p = 0.00$. The graph and the correlation show a strong but imperfect relationship between precipitation and streamflow — one that is not only non-linear in nature, but may even suggest an exponential pattern — leaving considerable room for other factors, including human action, to influence the river's behavior.

Implications

In this article, we present an example of data analysis that school students can undertake to answer research questions related to the issue of drought using the CR2 Climate Explorer (n.d.). The initial aim of this proposal was to address a common misunderstanding among the general public regarding the phenomena of drought and water scarcity. By analyzing secondary data and posing questions that seek to address the issue from different perspectives, we have characterized the phenomenon of drought in Petorca, where the interpretation of graphs has allowed us to question the climatic nature of this phenomenon and bring human intervention into play in this water crisis. The water crisis in Petorca has been studied, and scientists specializing in climate science indicate that the crisis is due to human action and not solely to a megadrought (Álamos et al., 2023; Muñoz et al., 2020). In this vein, this activity is valuable because, through simple data analysis, it is possible to reflect on and question human action in the context of the climate crisis. Furthermore, the CR2 Climate Explorer platform offers an opportunity for the Latin American region, as it provides access to data from various weather stations. In addition, we consider it essential to encourage students to develop data cleaning skills (Wild & Pfannkuch, 1999), since databases on social issues often contain large amounts of data and variables that sometimes need to be processed in order to answer research questions. All of this is done so that students can then use software that allows them to decide which analyses are needed to answer the questions they pose, as well as to create a space for reflection and dialogue between their experiences regarding the issue and what the data reveal about it.

Acknowledgements

F.U. thanks the support of ANID Fondecyt Iniciación 11250974, V.G thanks the support of ANID/PIA/Basal Funds for Centers of Excellence Center for Mathematical Modeling (CMM) FB210005., R.V. thanks the support of ANID Fondecyt Iniciación 11230184 and the Center for Climate and Resilience Research (CR2), based at the University of Chile and funded through the Actividades de Interés Nacional (AIN) program.

References

Álamos, N., Alvarez-Garreton, C., Muñoz, A., and González-Reyes, A. (2023). The influence of human activities on streamflow reductions during the megadrought in central Chile. *Hydrology and Earth System Sciences Discussions*, 2023:1–28.

Arnold, P. and Franklin, C. (2021). What makes a good statistical question? *Journal of Statistics and Data Science Education*, 29(1):122–130.

Centro de Ciencia del Clima y la Resiliencia (CR2) (2025). Inicio. https://www.cr2.cl.

Instituto Nacional de Derechos Humanos (2025). Petorca, historia de una crisis hídrica. rehuir el olvido. https://shorturl.at/OJJr2.

Muñoz, A. A., Klock-Barría, K., Alvarez-Garreton, C., Aguilera-Betti, I., González-Reyes, A., Lastra, J. A., and LeQuesne, C. (2020). Water crisis in Petorca Basin, Chile: The combined effects of a megadrought and water management. *Water*, 12(3):648.

Ubilla, F. M. and Gorgorió, N. (2023). Using Eurostat data to teach statistics to prospective primary teachers: on how the context of the task may promote their social awareness. *Journal of Mathematics Teacher Education*, 28:655–680.

Wild, C. J. and Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3):223–248.

The ISI Mission

Established in 1885, the ISI is a non-profit, non-government organisation with members in over 150 countries. Our mission is to promote the understanding, development, and good practice of statistics worldwide by fostering knowledge, sharing best practices, and creating opportunities to network. Since 1947, the ISI has held Special Consultative Status with the UN Economic and Social Council, enabling participation in the UN Statistical Commission and related activities. The ISI network spans most national statistical offices, international organisations, and professional societies around the world.

Join the ISI

We welcome members who share our mission of *Statistical Science for a Better World*. The ISI unites statisticians and data scientists across sectors and regions, including students and early-career members. As a member, you gain recognition, community, up-to-date knowledge, discounts, and a growing international network. Membership runs from 1 January to 31 December.

Become a member

Existing or former members: log in to the Membership System with your email (reset password if needed) to renew, reinstate, or add memberships. For assistance: isimembership@isi-web.org.

ISI Journals

International Statistical Review

The flagship journal of the International Statistical Institute and its Associations, publishing broad-interest work in statistics and probability: reviews and expository pieces, statistical computing and graphics, statistics education, and applications across sectors.

Stat

An ISI–Wiley journal for rapid, high-quality dissemination of new theoretical, methodological, and applied results. Compact articles with fast decisions, supported by online materials (data, code, graphics, and media) across all areas of statistics and allied quantitative fields.

Executive Committee

President (2025–2027)
Fabrizio Ruggeri

President-Elect (2025–2027)
Nalini Ravishanker

Vice President (2023–2027)
Denise Britz do Nascimento Silva

Vice President (2023–2027)
Jessica Utts

Vice President (2025–2029)
Ayşe Bilgin

Vice President (2025–2029)
Matthew Shearing

Executive Director
Conchita Kleijweg

Council 2023–2027

Kaye E. Basford
Arup Bose
Miguel de Carvalho
Ying Chen
Daniel Frischemeier
Matthew Parry
E. Ashley Steel
Linda J. Young

Council 2025–2029

Olawale Ale
Cathy Chen
Rabaa Baniyas
G. Van Haldaren
Elisabetta Carfagna
V. Gonzalez-Lopez
M. Cetinkaya-Runderl
Arun Sinha

Association Presidents

Bernoulli Society
Nancy Reid

International Association for Official Statistics
Gary Dunnet

International Association for Statistical Computing
Paulo Canas Rodrigues

International Association for Statistical Education
Bruno de Sousa

International Association of Survey Statisticians
Partha Lahiri

International Society for Business and Industrial Statistics
Matilde Bini

The International Environmetrics Society
Nathaniel Newlands