SOL Preprints, 2025 • doi: 10.5753/sol.2025.0002

EXPERIENCE REPORT

Promoting Data Culture: Report on the Participation of IME-UERJ in the Love Data Day Brazil 2025

Raissa Barcellos
Rio de Janeiro State University | *raissa.barcellos@ime.uerj.br*
Priscilla Abreu
Rio de Janeiro State University | *priscilla.abreu@ime.uerj.br*
Tassio Sirqueira
Rio de Janeiro State University | *tassio.sirqueira@ime.uerj.br*

Abstract. The Love Data Day, promoted by the Brazilian Computer Society (SBC), aims to disseminate the data culture and foster practical learning in Data Science. This experience report presents the participation of the Institute of Mathematics and Statistics (IME) of the Rio de Janeiro State University (UERJ), highlighting an SQL (Structured Query Language) marathon and a lecture on administrative and survey data. The marathon allowed the development of logical thinking, cooperation, and initial skills in SQL, while the lecture addressed conceptual, methodological, and ethical aspects of working with large databases. High levels of engagement and interaction were observed, demonstrating the effectiveness of strategies such as gamification and visual performance indicators, and pointing to recommendations for future editions, such as institutional support and more complex challenges.

Keywords: Data Science Education, SQL Marathon, Data Culture, Gamification

Received: 10 September 2025 • **Published:** 23 October 2025

1 Introduction

On June 3rd, 2025, the Brazilian Computer Society (SBC) promoted Love Data Day Brazil, a nationwide initiative aimed at raising awareness within the academic and professional community about the importance of data across all stages of the scientific process, as well as empowering students, researchers, and professionals to work with data effectively [Sociedade Brasileira de Computação (SBC), 2025]. In this context, the Institute of Mathematics and Statistics at the Rio de Janeiro State University (IME-UERJ) participated in the event, organizing practical activities designed to foster essential skills, collaboration, and engagement with data science.

Among the activities conducted, the SQL (Structured Query Language) marathon and a lecture were particularly notable. The marathon provided participants with opportunities to develop logical reasoning, query modeling, and relational database analysis skills through hands-on challenges, fostering insight extraction and confidence in applying these skills in professional contexts. This approach aligns with recent evidence in data science education, which suggests that gamified and collaborative environments enhance critical thinking, technical competence, active engagement, knowledge retention, and the practical application of concepts [Selvakumar *et al.*, 2024]. Furthermore, the marathon resembled datathons, collaborative competitions that encourage practical learning, peer cooperation, and the resolution of complex problems in controlled settings, distinguishing them from individual online competitions such as those hosted on Kaggle [1] [Sobel *et al.*, 2023].

The lecture provided a conceptual framework for working with large datasets, addressing topics from storage and cleaning to anonymization and de-identification strategies essential for safeguarding privacy. Participants were also introduced to statistical inference techniques applicable to administrative and survey data, enabling meaningful insights even from complex samples. Together, the marathon and lecture stimulated interest and engagement with the culture of data, promoted ethical and informed handling of information, and strengthened competencies applicable in both academic and professional contexts.

This experience report aims to describe the activities conducted at IME-UERJ during Love Data Day, highlighting the design, execution, and outcomes of the SQL marathon and the lecture. Lessons learned, challenges faced, and insights on student engagement, collaboration, and fostering a data-driven culture are discussed, with the goal of informing future initiatives and contributing to best practices for organizing hands-on, gamified learning experiences in data science.

2 Event Design and Implementation

The participation of IME-UERJ in the Love Data Day consisted of a series of in-person activities conducted over a single day, aimed at fostering hands-on learning and engagement with data science concepts. The program was structured in three stages: preparation for the SQL marathon (1:00–2:00 p.m.), the SQL programming marathon — Query Challenge by Love Data Day Brazil (2:00–5:00 p.m.), and a lecture entitled “Administrative and Survey Data: How to Structure, Process, and Analyze Large Databases” (6:00–7:00 p.m.). The preparatory stage enabled participants to familiarize themselves with the rules, objectives, and tools required for the marathon.

The SQL marathon was organized as a group compe-

[1] https://www.kaggle.com/

tition, in which teams faced practical challenges of varying complexity without consulting external resources. The Beecrowd platform [2] was used to facilitate the exercises, being widely recognized for programming training and competitive events. Prior registration through an online form allowed efficient control of participant numbers and supported logistical planning. The lecture was delivered by an adjunct professor at UERJ, who is also a permanent member of the Graduate Program in Economics (PPGCE). It provided a comprehensive overview of the use of administrative and survey data, addressing conceptual, methodological, and ethical aspects. Topics included distinctions among data types, advantages and limitations, storage and cleaning procedures, dictionary analysis, anonymization and de-identification strategies, and the application of statistical inference techniques using R (survey and dplyr packages) and SQL (sample weights). The session concluded with a Q&A segment, promoting interaction between the speaker and the audience and encouraging deeper discussion of the presented topics.

3 Event Outcomes

The event counted with the participation of 81 students, distributed across the activities carried out. Of this total, 32 students took part in the SQL marathon, organized in groups of up to three members, with the dynamics illustrated in Figure 1 (screenshot of the Beecrowd platform) and Figure 2 (record of the in-person marathon).

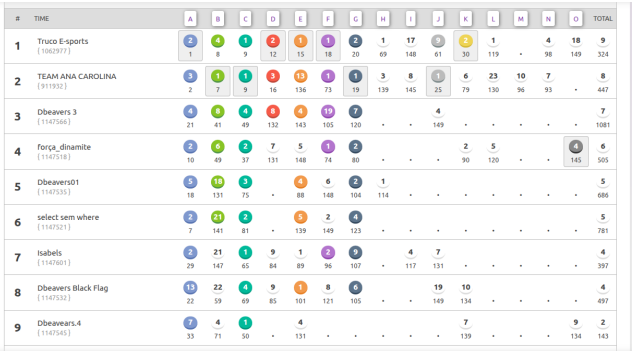

Figure 1. Distribution of groups during the SQL marathon on the Beecrowd platform.

In the lecture, 70 students participated, with partial overlap with those involved in the marathon, reflecting varied engagement and the diversity of experiences among the participants. During the marathon, a high level of cooperation and concentration was observed, with performance variations according to prior experience, as not all tasks were completed, but progress was made in logical understanding and the application of SQL syntax. Among the groups, one stood out as it was composed of students without formal knowledge in database courses, including a student not yet enrolled in the undergraduate program, who prepared autonomously.

One group, however, withdrew halfway through the activity, highlighting the complexity of the proposed challenges. Additionally, participants demonstrated progress in logical problem-solving, although difficulties related to SQL

[2] https://beecrowd.com/

Figure 2. Record of the in-person SQL marathon, with participants' faces blurred to preserve anonymity.

syntax application persisted, prompting real-time questions and clarifications that fostered active learning. In the lecture, interaction through questions and comments indicated interest in the practical application of the presented concepts, particularly among students in the early semesters of the undergraduate program. In summary, the activities were well received, contributing to the development of practical skills, peer cooperation, and engagement with data culture.

4 Community Impact

The activities conducted at IME-UERJ during the Love Data Day represented an important step in fostering a collaborative academic environment, extending beyond the scope of the scheduled activities. Students from different courses, semesters, and institutions, including the "Celso Suckow da Fonseca Federal Institute of Technology" (CEFET-RJ), engaged in structured and informal interactions, enhancing the diversity of perspectives and promoting interdisciplinary exchange. Faculty members and administrative staff provided essential support for event logistics, including the organization of the coffee break and distribution of giveaways, thereby reinforcing coordinated institutional involvement.

The activities fostered not only the development of technical skills but also peer-to-peer collaboration, cooperative problem-solving, and the establishment of a shared sense of belonging. This environment stimulated sustained engagement with data science, motivating some participants to pursue further exploration of data-related projects beyond the event. Consequently, the Love Data Day initiative demonstrated a measurable impact on both knowledge acquisition and the strengthening of a collaborative and engaged academic community.

5 Discussion

Students participating in the SQL marathon and the lecture experienced a dynamic and motivating environment, characterized by collaboration, active involvement, and opportunities for practical learning. The broad participation and observed enthusiasm suggest that the combination of practical activities with moments of interaction, such as the coffee break held between sessions, contributed to creating a positive and collaborative learning context. Specific strate-

gies adopted during the event proved effective in stimulating participant involvement. The use of visual indicators, such as colored balls representing the number of challenges completed by each group, promoted healthy competition and encouraged continuous participation. It was also observed that students demonstrated enthusiasm and engagement, indicating that playful elements and gamification practices can significantly enhance the learning experience.

Opportunities for improvement were, however, identified. The organization of logistical aspects, such as the coffee break and the provision of gifts, relied on voluntary contributions from faculty, suggesting that formal institutional support could ensure greater consistency and quality of these resources. Additionally, future editions of the event could incorporate more complex challenges, including exploratory analysis of real datasets and predictive activities, broadening the conceptual and practical depth of the proposed activities.

As observed in studies on statistics and data science hackathons, such events promote confidence, reduce anxiety, and stimulate the acquisition of new skills [Sharifi Far *et al*., 2025]. Similarly, participants in the SQL marathon demonstrated high engagement, effective group cooperation, and interest in applying concepts in practical situations. Strategies such as visual performance indicators and social interaction moments can be understood as aligned with best practices reported in the literature to maximize educational impact. In terms of outcomes, the event contributed to the development of initial technical skills in SQL and data analysis, providing students with hands-on experience with real-world challenges. Compared to similar initiatives, the adopted format demonstrates good practices in the data field, particularly regarding the integration of practical activities, social interaction, and playful elements—factors that enhance engagement and effective learning.

6 Conclusion

The Love Data Day held at IME-UERJ proved to be an effective initiative for promoting engagement and hands-on learning in Data Science. The SQL marathon and the lecture offered technical and conceptual challenges that stimulated practical learning, logical reasoning, and group cooperation. It was observed that strategies to encourage interaction, such as visual performance indicators and social moments between activities, contributed to making the event more engaging. At the same time, logistical challenges, such as organizing resources for the coffee break and souvenirs, highlight the need for institutional support to ensure future events are conducted consistently.

For future editions, it is recommended to include complementary activities, such as exploratory and predictive analyses of real datasets, increasing the complexity and practical applicability of the challenges. Moreover, diversifying engagement strategies can enhance student participation and experience, strengthening the data culture within the academic community. In summary, the event demonstrated itself as a valuable tool to integrate technical learning, social interaction, and playful practices, providing relevant insights for organizing similar initiatives in the field of Data Science.

Declarations

Authors' Contributions

RB, PA, and TS contributed to the organization and logistics of the LDD event at IME-UERJ, as well as writing, reviewing, and approving this experience report. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets (and/or software) generated and/or analysed during the current study will be made available upon request.

References

Selvakumar, T., Rajapakse, C., and Jayalath, N. (2024). Enhancing data science education: A systematic review of gamified learning platforms for decision-making skills development. In *2024 6th International Conference on Advancements in Computing (ICAC)*, pages 426–431. IEEE. DOI: 10.1109/ICAC64487.2024.10850893.

Sharifi Far, S., Inácio, V., Evkaya, O., and Lenzi, A. (2025). Hackathons in statistics and data science education and experiences from asa datafest. *Teaching Statistics*. DOI: 10.1111/test.12404.

Sobel, J., Almog, R., Celi, L., Yablowitz, M., Eytan, D., and Behar, J. (2023). How to organise a datathon for bridging between data science and healthcare? insights from the technion-rambam machine learning in healthcare datathon event. *BMJ Health & Care Informatics*, 30(1):e100736. DOI: 10.1136/bmjhci-2023-100736.

Sociedade Brasileira de Computação (SBC) (2025). Love data day brasil. Acessed on 01 set. 2025. Available at https://www.sbc.org.br/love-data-day-brasil.