



Universidade: presente!



XXXI SIC

21. 25. OUTUBRO • CAMPUS DO VALE

VISUAL ANALYSIS AND INTERACTIVE EXPLORATION OF MILLIONS OF EMPLOYMENT RECORDS

Motivation

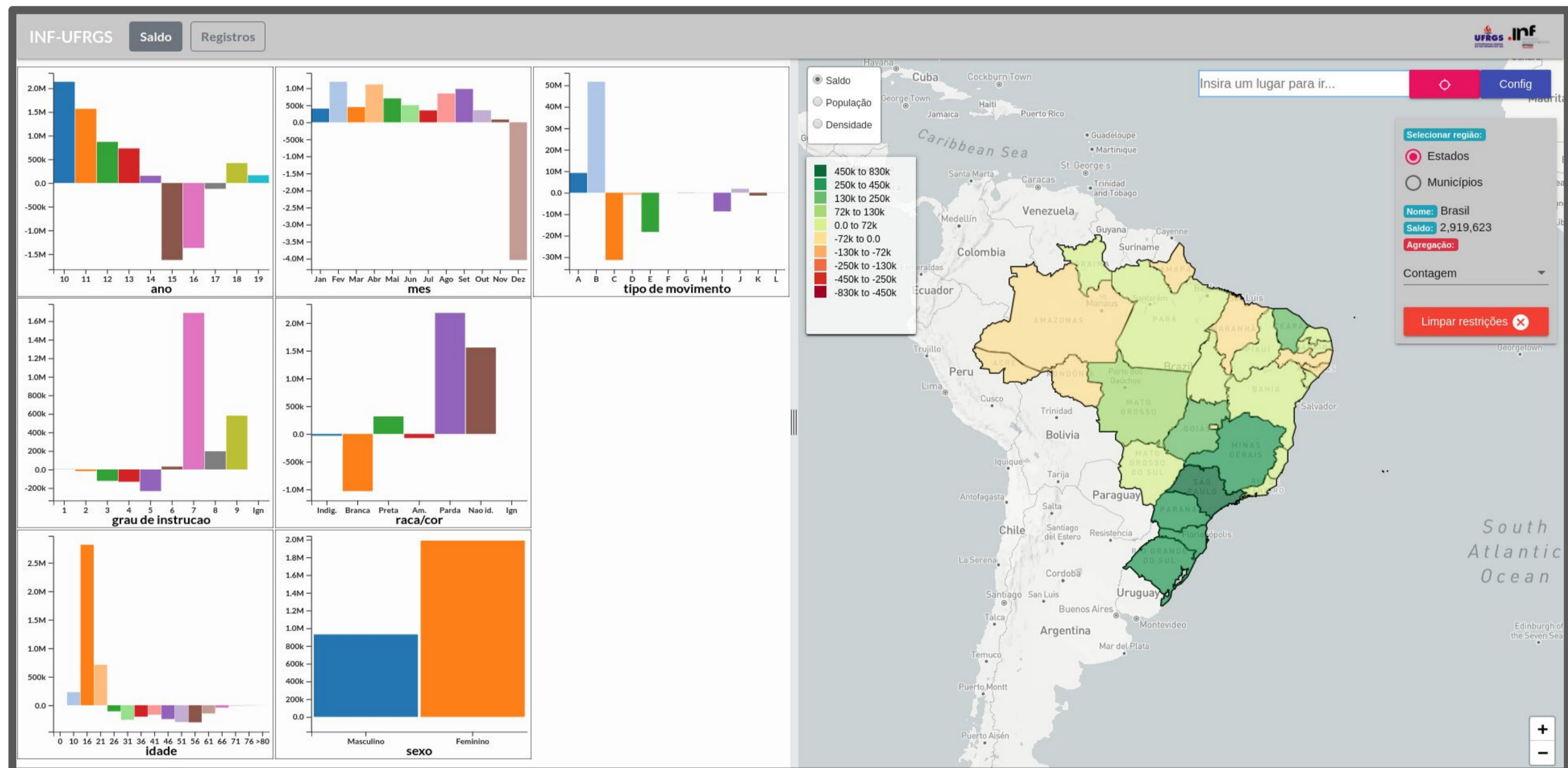
Data analysis is essential in many application areas. Business decisions, optimization strategies, among others, depend on understanding big and heterogeneous data, which is not trivial.

The five challenges of big data analysis:

- volume (the scale of data);
- velocity (the analysis of streaming data);
- variety (the different formats of data);
- veracity (the uncertainty of data);
- value (what value the analysis brings to the understanding of data).

Contribution

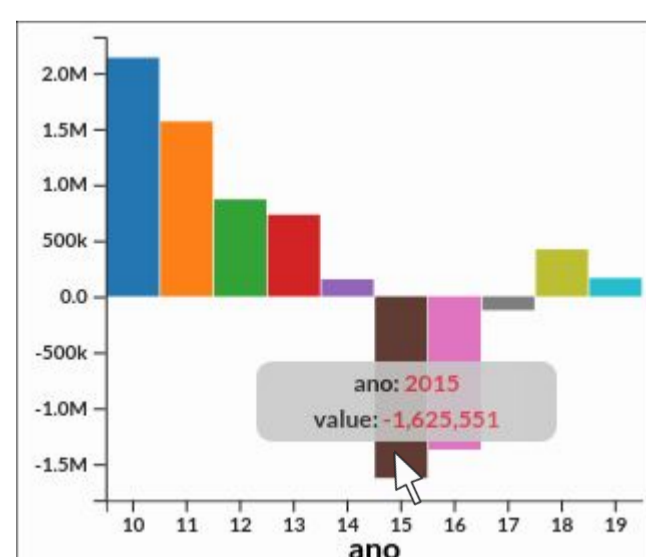
The main concern in this work is to study how visual data analysis can help to answer queries about data in a real case study. In this work, we create a prototype to understand the dataset of employment activities in Brazil, the CAGED ("Cadastro Geral de Empregados e Desempregados"). We developed our prototype in a data structure for visual big data analysis developed in our research group, called Quantile Datacube Structure (QDS). Using QDS, we demonstrate a prototype that allows the interactive visual analysis of millions of records illustrating how it can be used to validate and elaborate on different hypothesis over the data.



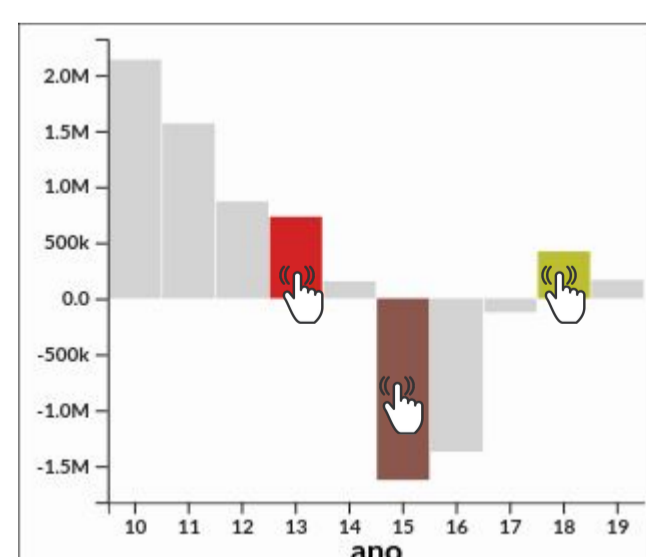
Screenshot of our CAGED prototype.

How to interact

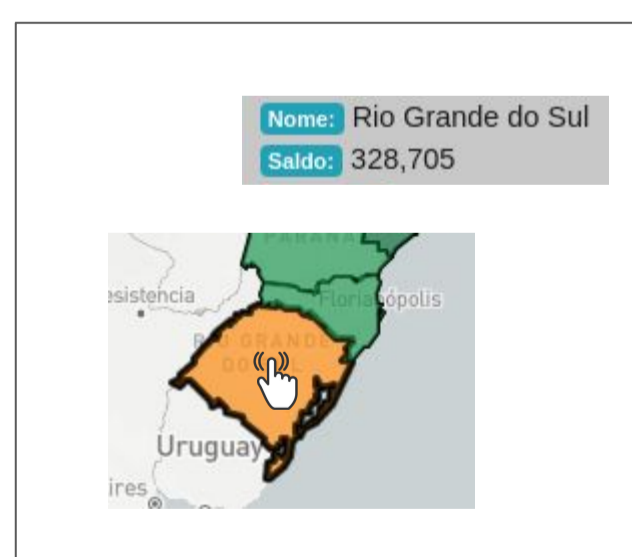
The CAGED prototype displays the balance between employee admissions and dismissals in Brazil, using interactive bar charts and region map, with data from January 2010 to March 2019.



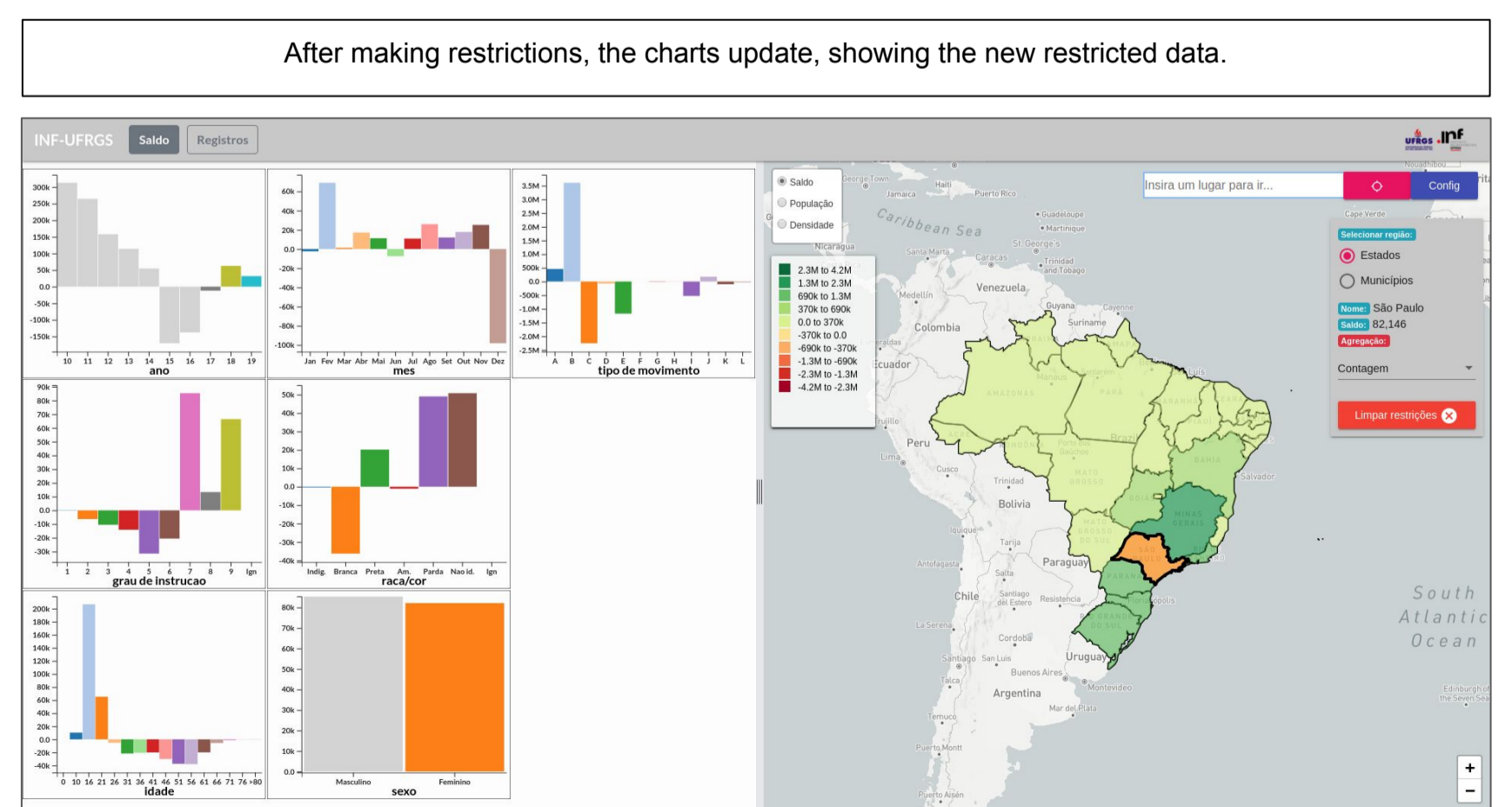
Hover on the charts to see the exact values on the bars



Click on the bars to (de)select restrictions



The previous two work in the region map as well



After making restrictions, the charts update, showing the new restricted data.