Weather Mapping Application for Brazil: Data Source and Processing

By Raphaël Faure

Introduction

This document provides an overview of the data source, processing steps, and key considerations for the "Weather Mapping Application for Brazil". This application aims to visualize weather data, specifically temperature and precipitation, across different regions of Brazil.

Data

Original Data

I retrieved the data from the website you had given me. I managed to download a zip file containing a .csv for 566 cities across Brazil with the following attributes:

The data ranges from 01/01/2023 to 31/10/2023. I wanted to retrieve even more recent data, but the download was no longer grouped. I have saved these files in the '2023' folder. I tried to do scraping with Rselenium in order to retrieve information on a daily basis, but it did not succeed due to lack of time.

Moreover, to plot the regions, I needed the coordinates of the borders. So, I retrieved a Geojson file. All sources are available at the end.

Data processing

The files are in .csv format. However, there is a sub-separator which makes processing the files very laborious. I spent some time understanding the file format and how to transform it.

First, I decided to break down the file into two subfiles. The first contains geographical data and the other contains meteorological data. Once this decomposition is done, I perform my first calculations with the following objective: I want to keep only the date column from the weather data and create two new columns, 'Temp_daily' and 'Precip_daily'. 'Temp_daily' is the daily average of the hourly averages for temperature columns, and 'Precip_daily' is the sum of

daily precipitation. Once these calculations are done, I transpose the geographical data to add it to the previous table, thus obtaining a dataframe containing all the information.

Once I have a dataframe per city, I add them recursively to create a larger dataframe containing information from all the cities. I then merge the data from the Geojson file with the previous dataframe on the 'Reg' column. The problem is that this merge significantly increases the temporal and spatial complexity of my application. Indeed, this merge, without modifying the date series of the dataframe, needs at least 49 GB to save and takes at least one hour to do so. This being really not feasible for the customer experience, I decided to limit the date range to 3 days. This simplification presents no problem for the future as my code does not depend on the size of the interval.

Finally, we obtain the desired file at the beginning.

Leaflet Mapping:

The processed data is visualized on a map using the Leaflet library. This includes: Polygons for regions with color coding based on temperature or precipitation. Je voulais que l'utilisateur ai le détail des températures par ville sur la carte en cliquant sur la région. Indeed, I find it unreadable to put the polygons of the regions with the markers for the cities. When the user launches the application, the regions are displayed with a color code that reflects the levels of heat/precipitation between regions over the time interval. Markers for cities showing specific weather data on click. In practice, due to the size of the file, the display takes a lot of time, which complicates the user experience.

User Considerations

Data Accuracy*:* The accuracy of visualizations depends on the source data. Users should be aware of potential discrepancies or data limitations.

Date Range: Users can select specific date ranges for data visualization, which can impact the aggregation and display of weather patterns.

Interactivity: The application allows interactive exploration, such as clicking on regions or cities for more detailed data.

Performance: The application's performance may vary depending on the dataset's size and the user's device capabilities.

Conclusion

This weather mapping application provides an interactive tool for visualizing temperature and precipitation data across Brazil. By integrating weather data with geospatial information, it offers a comprehensive view of weather patterns and their regional variations. I would like to add that this project was very enriching. Indeed, it was only the second project I did on R and the first one by myself. I learned a lot through all this work, but it took me a lot of time, which limited me on the application. I encountered many problems, both with the data and the

development of the shiny app. I also spent a lot of time on YouTube tutorials to train myself. Despite all this, I admit that I did not manage to do everything I wanted on the map and in the tools offered to users, and I am sorry for that.

Sources :

https://portal.inmet.gov.br/

https://cartographyvectors.com/map/1423-brazil-with-regions

https://rstudio.github.io/leaflet/shiny.html

https://gis.stackexchange.com/questions/168886/r-how-to-build-heatmap-with-the-leaflet-package

https://lrouviere.github.io/TUTO_VISU/correction/faire-des-cartes-avec-r.html

https://medium.com/swlh/creating-a-live-world-weather-map-using-shiny-f2ad05a08a13

https://github.com/Ponsoda/geoprogramming-youtube/blob/main/FirstScript/R/GettingStartedWithShiny.R