The Challenge of Current and Future Climate Data in the Era of Digital Agriculture

PRISCILA COLTRI

RESEARCHER AND PROFESSOR AT UNIVERSITY OF CAMPINAS (UNICAMP)
DIRECTOR - CENTER FOR METEOROLOGICAL AND CLIMATIC RESEARCH APPLIED TO AGRICULTURE (CEPAGRI)
BRAZIL

Summary

1. Climate Data and Agriculture

2. Challenges and Opportunities:
   a) Data Accessibility
   b) Data Type
   c) Data Quality
   d) Data Integration
   e) Data Understanding
   f) Ethics

3. Projects Examples

Climate Data and Agriculture

Agriculture highly dependent on Weather and Climate conditions

Climate: has an impact on planning activities for long-term actions - for example, what, where and when to grow a crop

In addition, Climate change could also be an issue , understanding the impacts for production and planning activities, and also generating an huge amount of data for models and future scenarios

Climate Data

Given this dependency:

a) Monitoring and Forecasting weather and climate has always been a challenge

b) Climate data is really needed – we have a **wide range of sources for climate data**

In the era of digital agriculture, we are producing more data than ever before, and (especially) climate data.

Climate Data – Challenges

Data Accessibility

Data Type

Data Quality

Data Integration

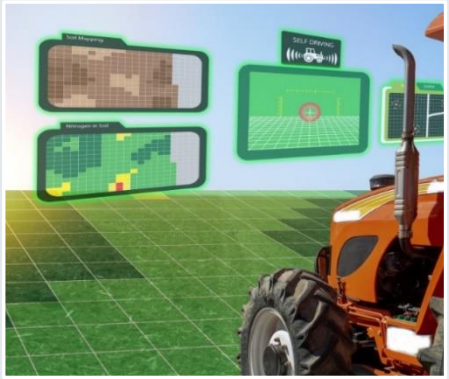
Data Understanding

Ethics

Challenges that data science and artificial intelligence could address

Data Accessibility

Different Infrastructure for data collection



Farm level

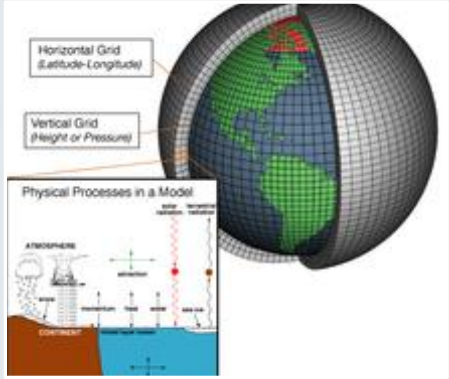
Micro Scale

Data is usually at local/farm level
The data belong to the farmer; meteorological station or sensors inside the tractors and machinery

Meteorological Station

Meso Scale

Public Institutions in Brazil;
data at meteorological station level
Free access



Models

Macro Scale

Climate models
reanalyses model
Satellite images

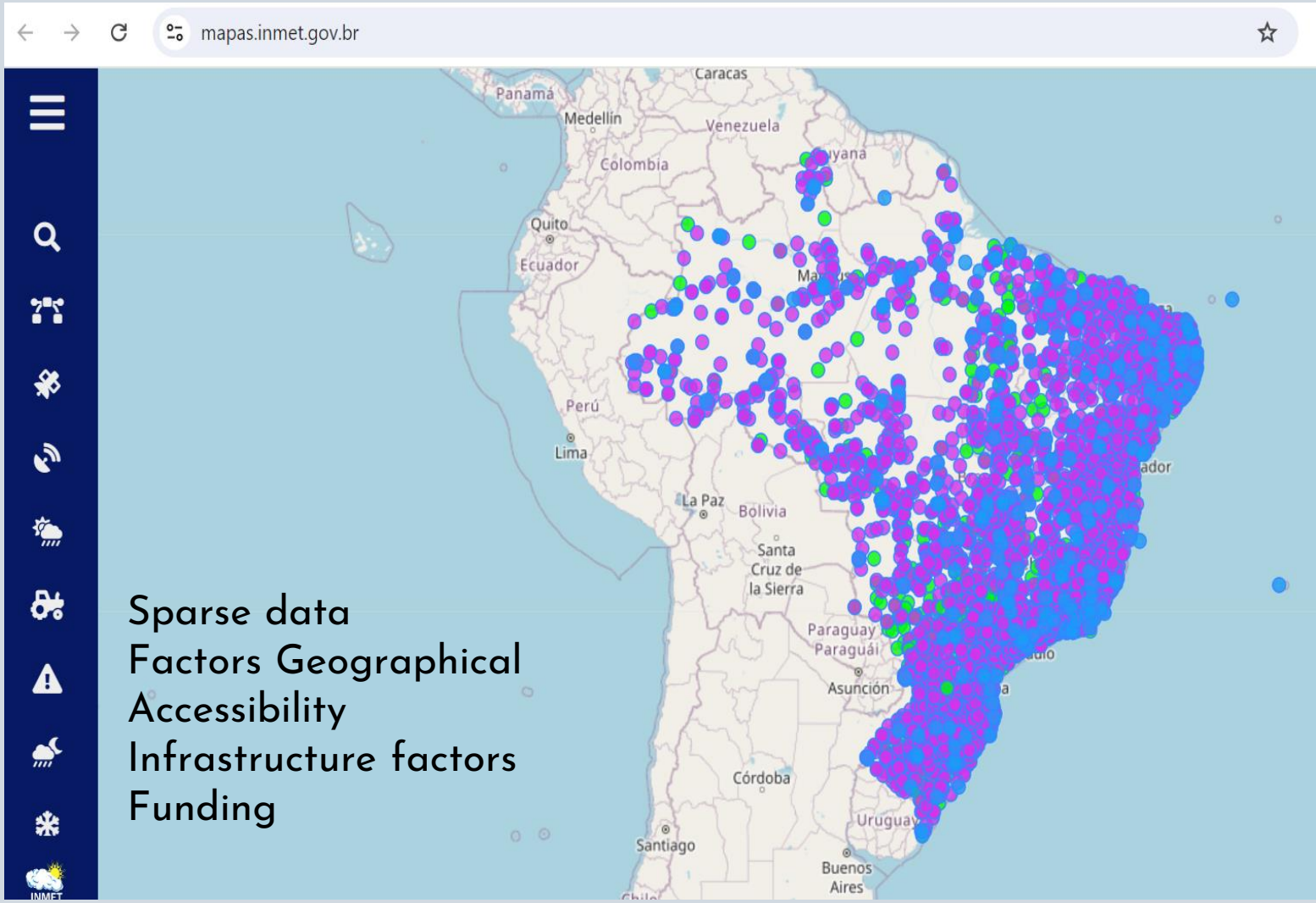
So we have different sources of data with different scales and different meanings.

Climate Data – Challenges

INMET - Instituto Nacional de Meteorologia
(National Institute of Meteorology)

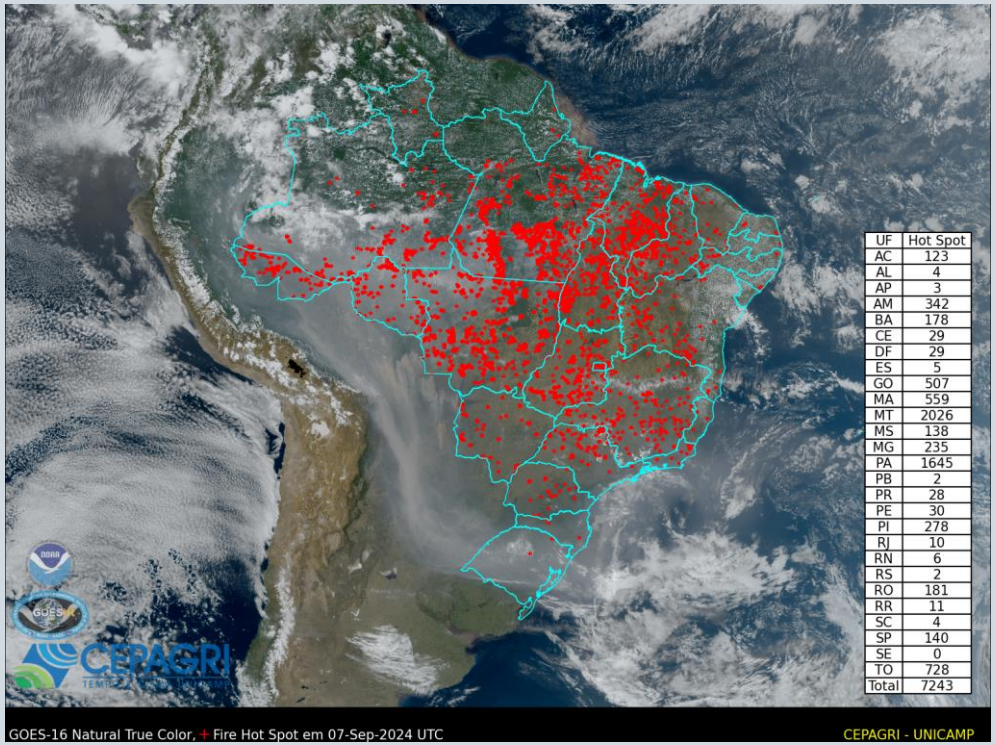
Data Accessibility

- ✓ Meteorological stations is the most popular data- freely available;
- ✓ Northern region we have few stations because of the difficulty of access
- ✓ Even in the other regions where we have a network of stations, we still have problems with funding and infrastructure, so there are gaps in the data series.

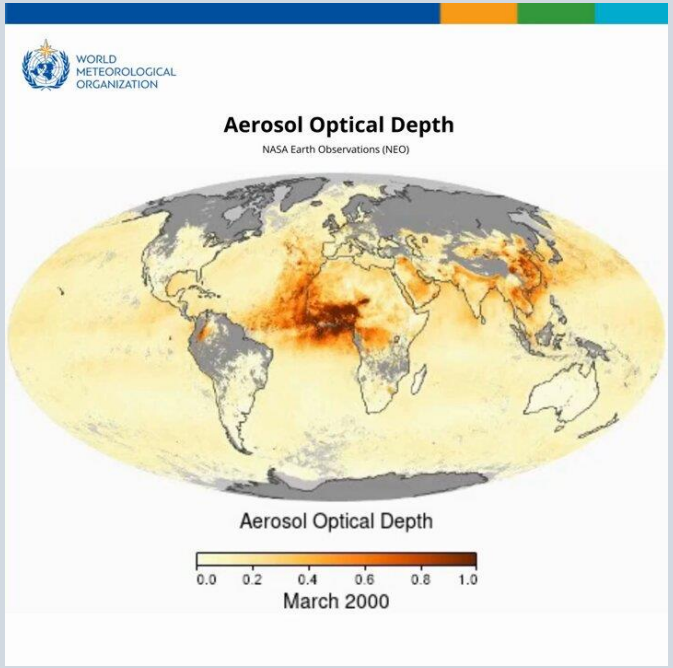


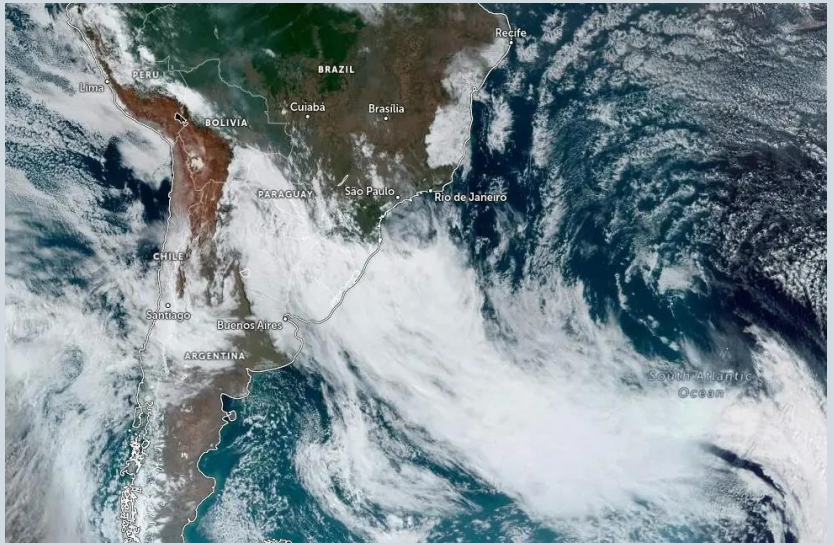
Climate Data – Challenges

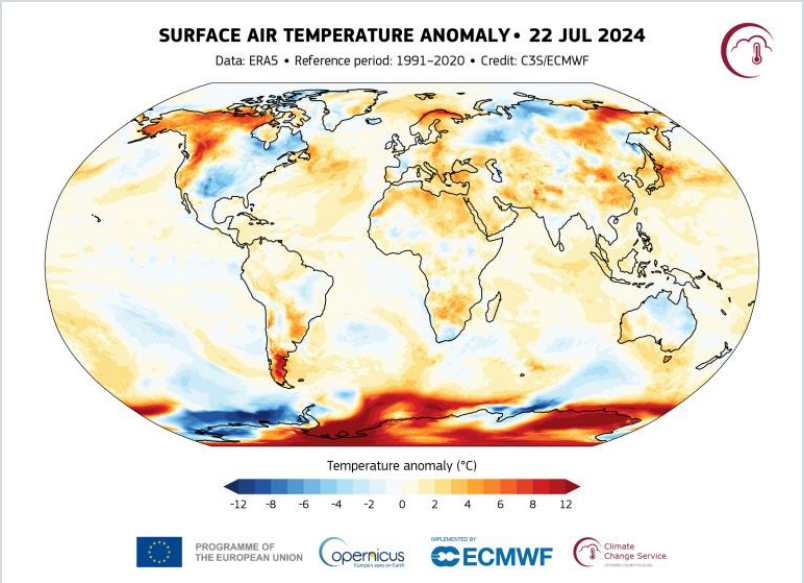
Macro Scale Examples



UF	Hot Spot
AC	123
AL	4
AP	3
AM	342
BA	178
CE	29
DF	29
ES	5
GO	507
MA	559
MT	2026
MS	138
MG	235
PA	1645
PB	2
PR	28
PE	30
PI	278
RJ	10
RN	6
RS	2
RO	181
RR	11
SC	4
SP	140
SE	0
TO	728
Total	7243

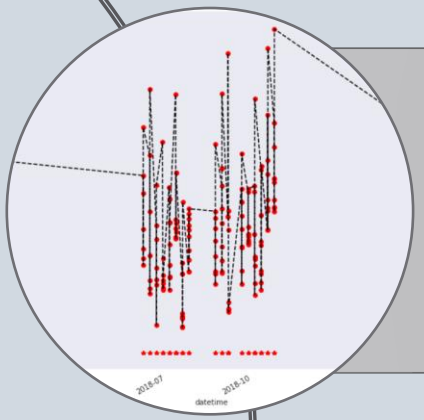




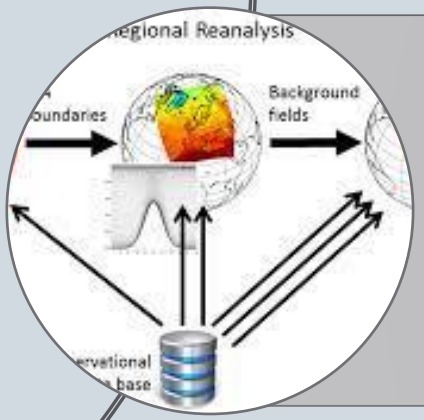


Climate Data – Challenges

Data Quality/ Consistence



Meteorological Stations present historical missing data and local measurements



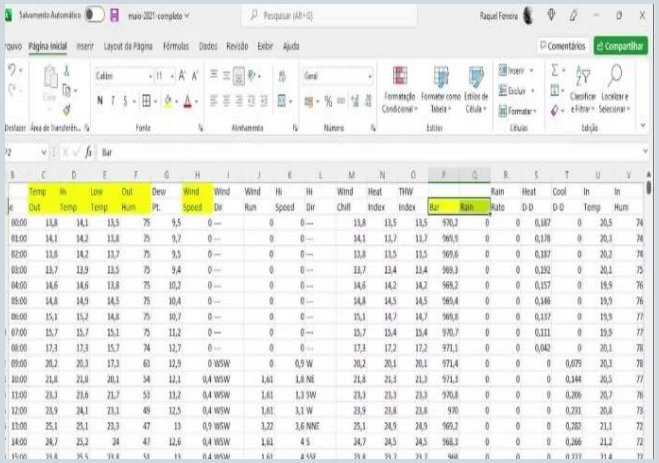
Model x measured data: require studies to understand how the model is performing local/regional/national climate

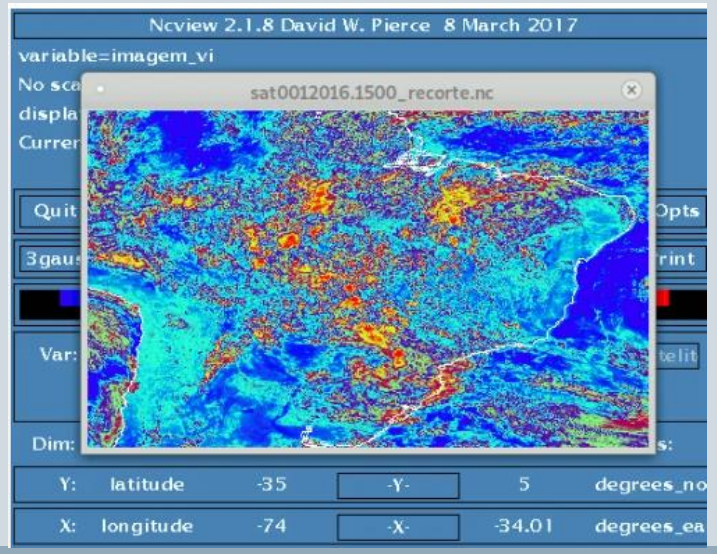
Climate Data - Challenges

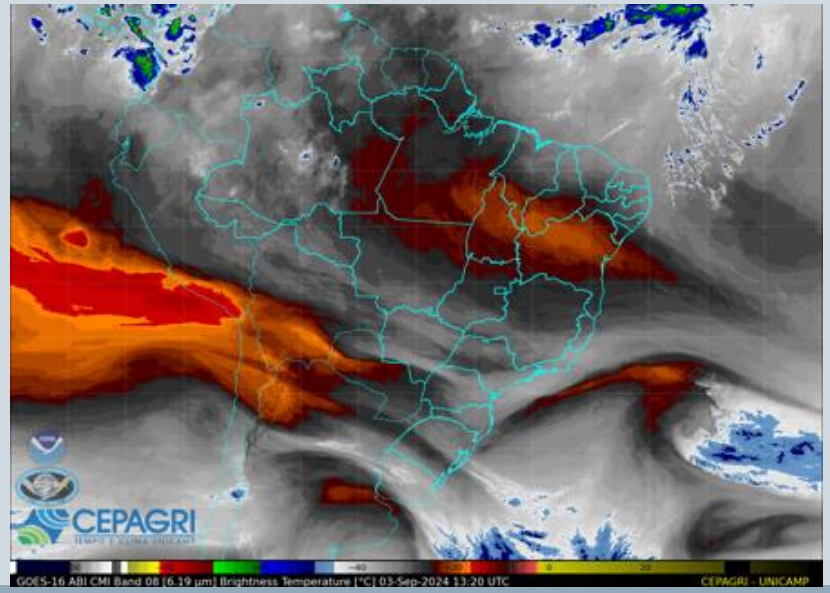
Data Type

- ✓ Each platform provides a file format
- ✓ Difficulty in working with many files: csv spreadsheets; NetCDF; TIFF
- ✓ The user is invited to choose the data source: weather station, satellite, models - the user does not always know the difference, the uncertainties

Leading us to another challange



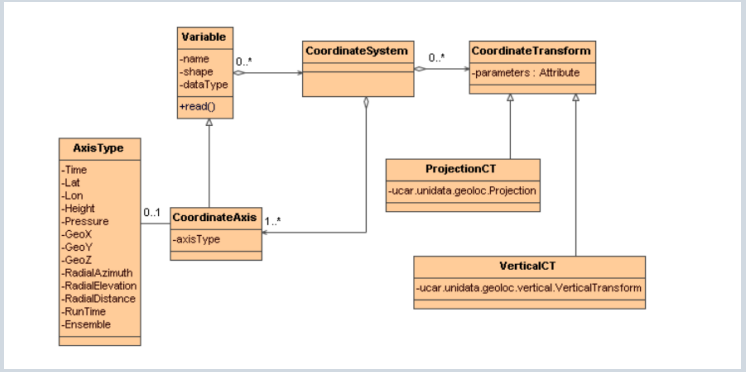




Data Integration

- ✓ A challenge with all these differences between data files is how to integrate them with other types of data: seems to be simple when working with a small dataset, but when working with a large amount of climate data, on different temporal and spatial scales, with other databases (e.g. soil, plant or other area databases), we are faced with the challenge of integrating these data.

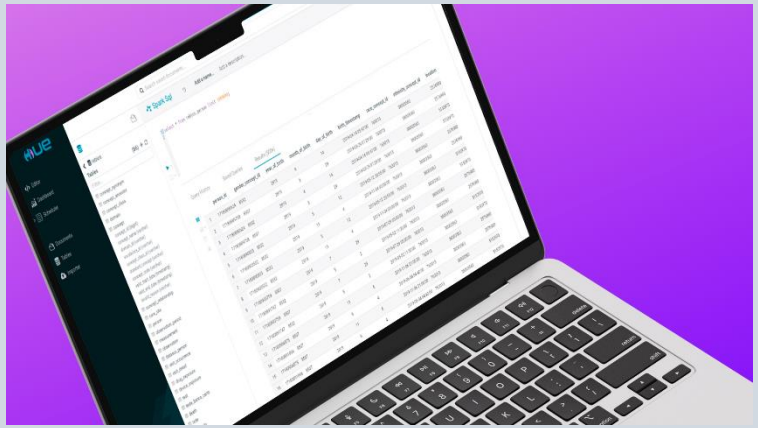
- ✓ Develop Common Climate Data Model - Assists with data description and interoperability with other databases



Data Understanding

Make data understandable

- ✓ Develop products for agricultural users, understanding what they need
- ✓ Each type of producer has a way of communication
- ✓ Develop products considering different spatial and temporal scales
- ✓ Present Data visualization and exploration tools
- ✓ Provide access to data sets for different society sectors (academic, agricultures, energy, health…)
- ✓ Encouraging citizen science



Dashboard Overview

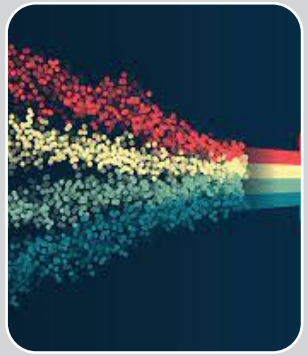
A user-friendly tool offering real-time visualization and analysis of integrated datasets, using customizable visual tools such as heat maps, time series, and predictive modeling outputs.

Data Governance

Ethics

Ethical Considerations: Data for all regions, without neglecting or discriminating ; generation of too much data

Data Security: protecting data from unauthorized access- need regulations



Data Consistency – Spread of data: Inconsistent data may lead to flawed models, misguided public policies

Data Stewardship: assigning responsibility for data management

Compliance and Regulation: adhering to other legal and regulatory requirements related to data management

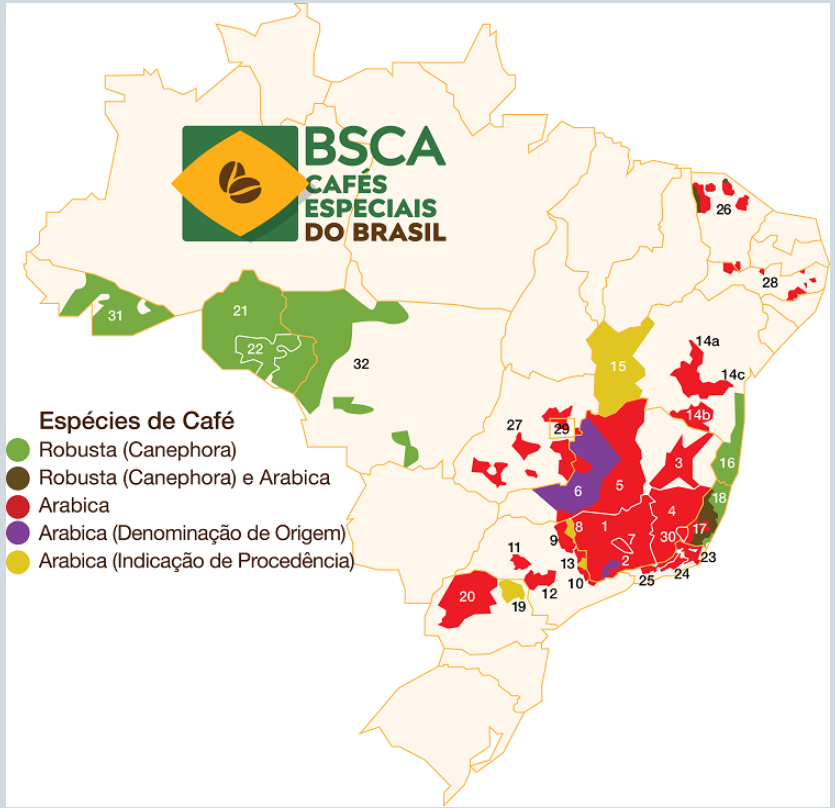
Project Examples

BRAZILIAN INSTITUTE OF DATA SCIENCE

BIOS

UNICAMP

Coffee Change



Research Questions:

- What has been happening in the climate of Brazilian coffee areas?
- What will happen in these areas?
- How will coffee react?
- What is the climate perception of coffee producers?

@coffeechange.unicamp

CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico

Project Examples

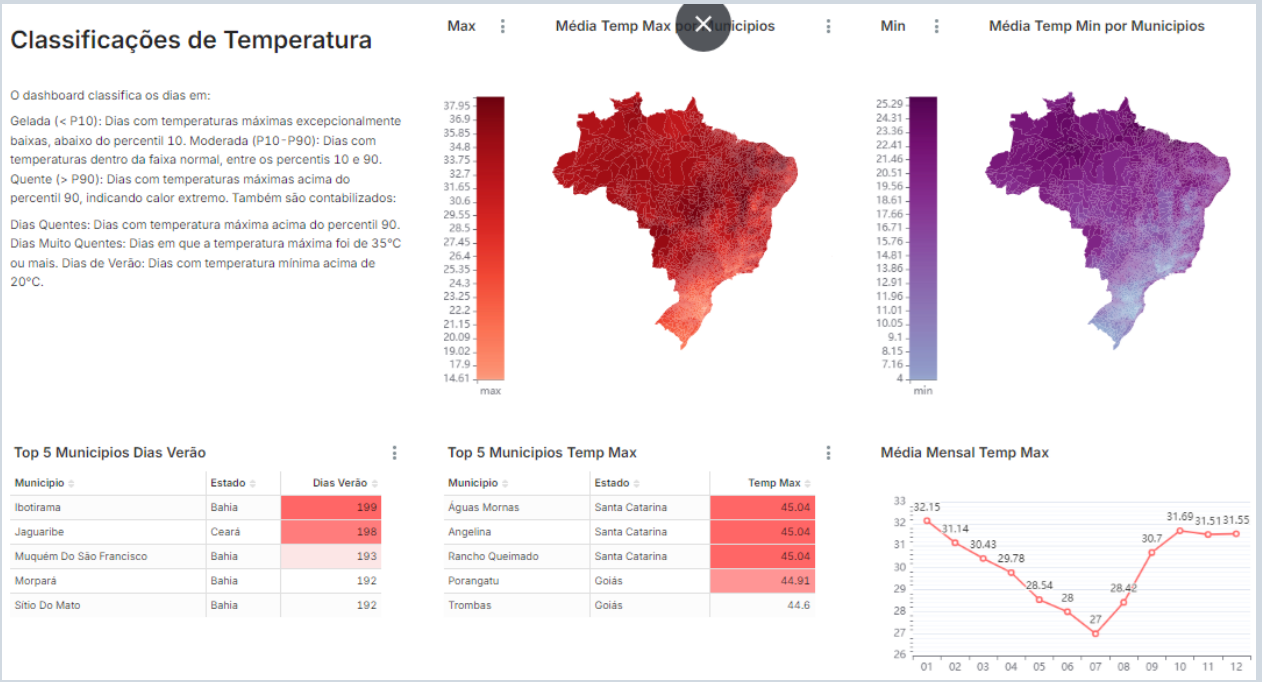
BRAZILIAN INSTITUTE OF DATA SCIENCE

BIOS

climaterna

The Climaterna project, a initiative by Brazilian Institute of Data Science (BIOS), at University of Campinas, is a cutting-edge data platform that integrates climate and maternal-perinatal health information.

This open-source based repository will serve as a permanent observatory, designed to analyze the impacts of climate change on maternal (pregnancy) and newborn health in Brazil.



Datasets

climaterna

The dashboard is formed by several public key datasets from over 5,000 municipalities in Brazil.

Environmental Data

Climate & Pollution Data:

Sourced from Brazilian Daily Weather Gridded Data (BR-DWGD), Climate Models and CETESB, it includes temperature, precipitation, air quality, and extreme weather events, which are crucial for identifying climate-related health risks.

Health Data:

SIM (Mortality Information System) Provides data on maternal and perinatal mortality, enabling analysis of how climate factors correlate with mortality rates.
SINASC (Live Birth Information System): Offers detailed birth data, essential for understanding trends in premature births and neonatal health in relation to environmental conditions.

Demographic Data:

Captured from Census, which is a comprehensive population survey conducted by the Brazilian Institute of Geography and Statistics (IBGE). It includes information such as age, income distribution, genders, races/ethnicities, education levels, and employment status.

climaterna

https://me-qr.com/l/climaterna

https://bi0s.unicamp.br/

UNICAMP

Inria

Laboratório Nacional de Computação Científica

BRAZILIAN INSTITUTE OF DATA SCIENCE

BIOS

Thank you for your attention

Merci de votre attention

pcoltri@unicamp.br

FAPESP
FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico