TERRABRASILIS: A SPATIAL DATA INFRASTRUCTURE FOR DISSEMINATING DEFORESTATION DATA FROM BRAZIL

Luiz Fernando Ferreira Gomes de Assis, Karine Reis Ferreira, Lúbia Vinhas, Luis Maurano, Cláudio Aparecido de Almeida, Jether Rodrigues Nascimento, André Fernandes Araújo de Carvalho, Claudinei Camargo, Adeline Marinho Maciel

Agenda

- Monitoring Large Deforestation Mapping Areas in Brazil
- Spatial Data Infrastructure
- Improving GIS Interoperability
- Transforming GIS Experts into Data Science Analysts
- Lessons Learned from the Deployment of TerraBrasilis in a Real-World Deforestation Scenario

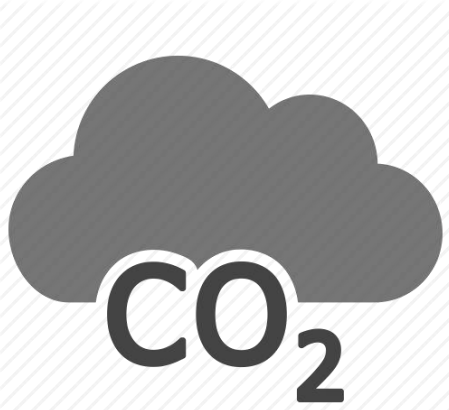
Introduction

- In order to increase Brazil's capacity to deal with environmental monitoring applications such as **deforestation detection**, **forest fire protection**, and **greenhouse gas emissions estimations**, it is essential to remove the barriers from:

- **organization**,
- **access** and
- **use** of spatial data with **temporal dynamics**.

DEFORESTATION

FOREST FIRE



GAS EMISSION

Introduction

- The demand for these capabilities can be **exemplified by scenarios** in which users need to evaluate the effectiveness of **thematic data** over time resulted from systematic environmental monitoring projects in INPE such as **PRODES** and **DETER**.

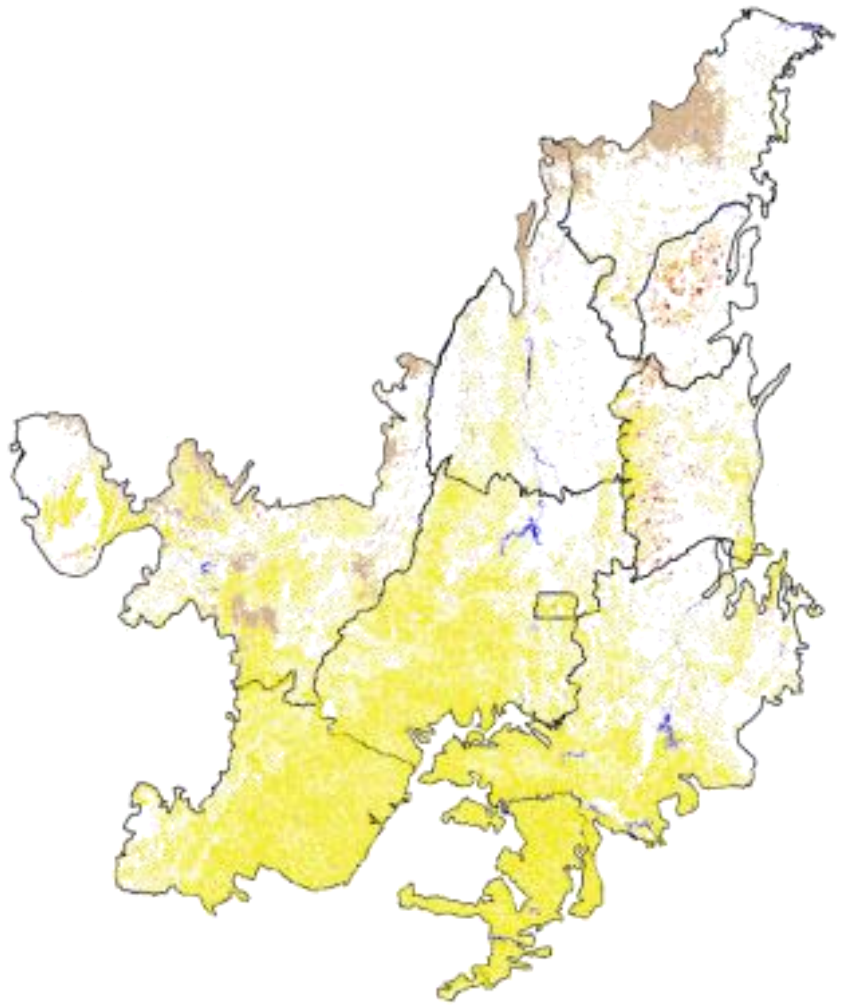
- **Distinct data characteristics such as spatial and temporal resolutions and extents**, as well as **the thematic parameters**, result in volatible requirements for analysis

Introduction

- Cerrado is the second largest biome in Brazil, covering a fourth of its territory. Over the last few years it has lost almost 24% of its original coverage due to the agriculture expansion (e.g., soybean, cotton, and corn production), supressed vegetation and pasture cattle.

- Cerrado's degree of destruction has reached such alarming rates that if it continues it will be difficult to recover its biodiversity.

- With that in mind, much of the attention that has flowed towards Amazon Forest over the last few years while other biomes stayed in the background, has cloven to Cerrado now.



Introduction

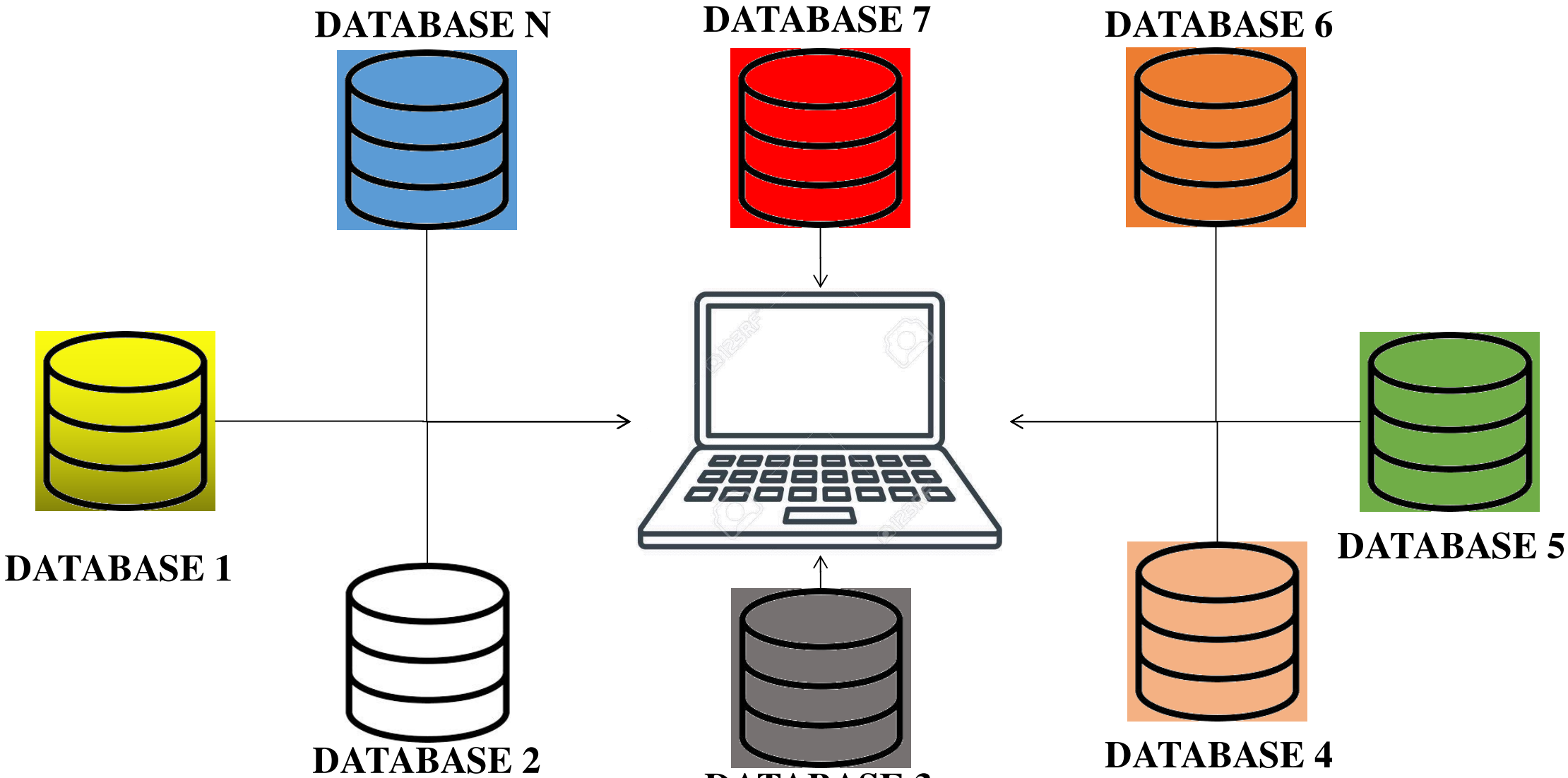
For this, a **much more generic and abstract framework** is needed, that is, not just considering **the traditional map servers to represent these kind of environments** but **visual analytics indicators and metrics** to improve decision-making.

SPATIAL DATA INFRASTRUCTURE (SDI)

Introduction

"Integrated set of technologies; policies; coordination and monitoring mechanisms and procedures; standards and agreements necessary to facilitate and order the generation, storage, access, sharing, dissemination and use of geospatial data of federal, state, district and municipal origin."

Introduction



Introduction

- The influence of regional governamental policies to increase the resilience of Cerrado biome and to preserve its biodiversity.
- The concern for handling the integrated and adaptive management of historical and near-real time deforestation-related rates, increments and alerts.
- The expensiveness to afford constantly the technology innovation transformations that often follow SDI evolution.
- The degree of SDI modularity with benefit of generic and flexible implementations to other biomes.

Introduction

- TerraBrasilis helps to **organize, access and use** spatial data produced by INPE's environmental monitoring programs, but throughout **a web portal**, it makes possible based on **customized views to aggregate other types of spatial data**.

- Rather than just **relying on geoservices**, it uses **ubiquitous clear and simple APIs** accross a cluster of virtualized machines to make spatial data analysis easier.

- TerraBrasilis **enables the management of dynamic environments** such as those found in DETER project that produces daily data.

- It allows reasonable to **trace forest degradation and fire scars areas** every day even before they are deforestated.

Introduction

- **Engineering requirements**, **designing**, **implementing**, and **evaluating** an open-source SDI to organize and disseminate deforestation data obtained from consolidate thematic mapping projects such as DETER and PRODES;

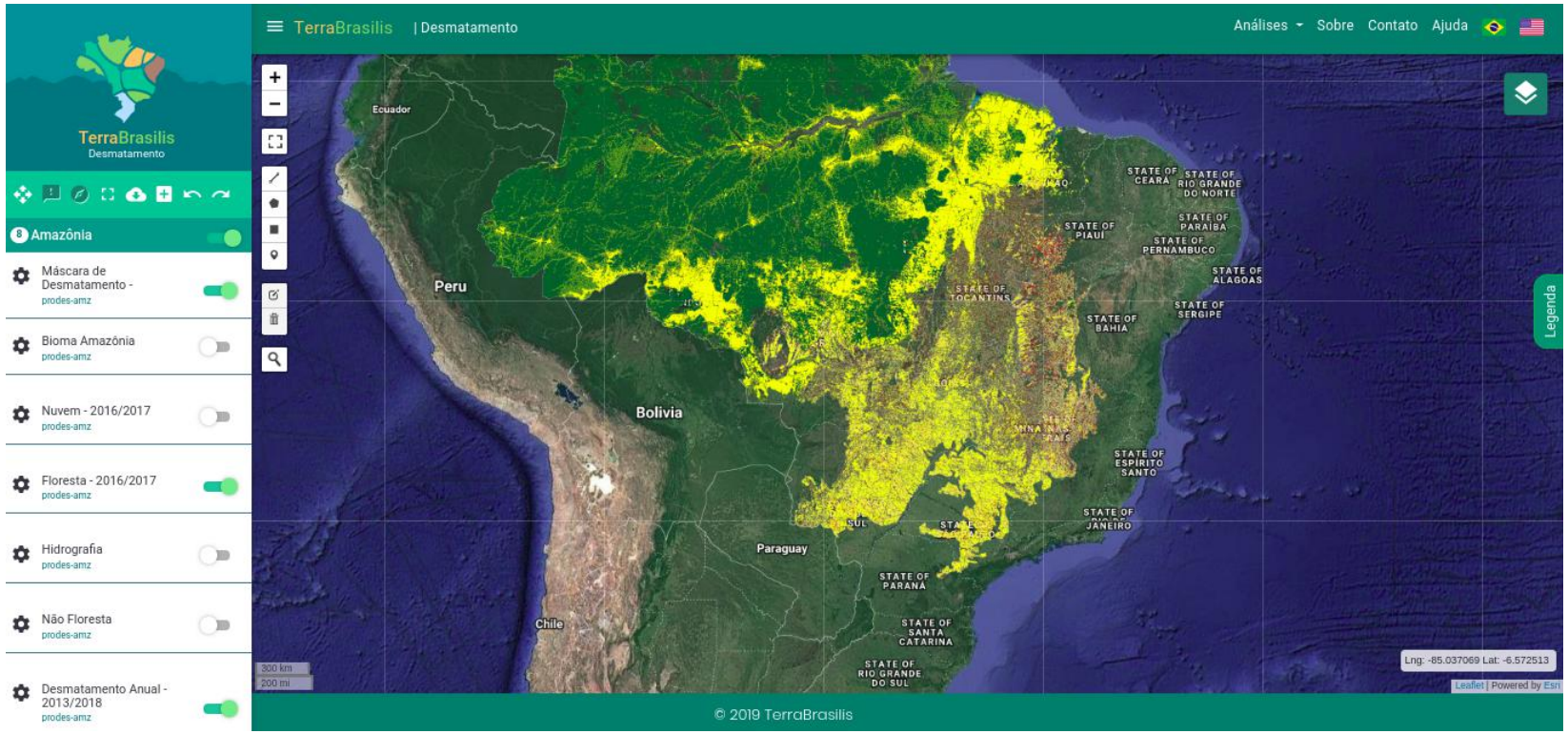
- Learning lessons from the application of the proposed approach **in a real-world deforestation scenario** that has called attention for its fast natural anthropological conversion, complex formation and high correlation to soybean cultivation in Cerrado biome, Brazil.

Terrabrasilis

Combining Web Services for Maps

- Stateful vs Stateless applications
  - Reduce ram usage on server.

- Monolithic vs microservices
  - The microservice architecture contains small services and each one runs in its own process and are independently deployable, as well as communicates with lightweight resource API.



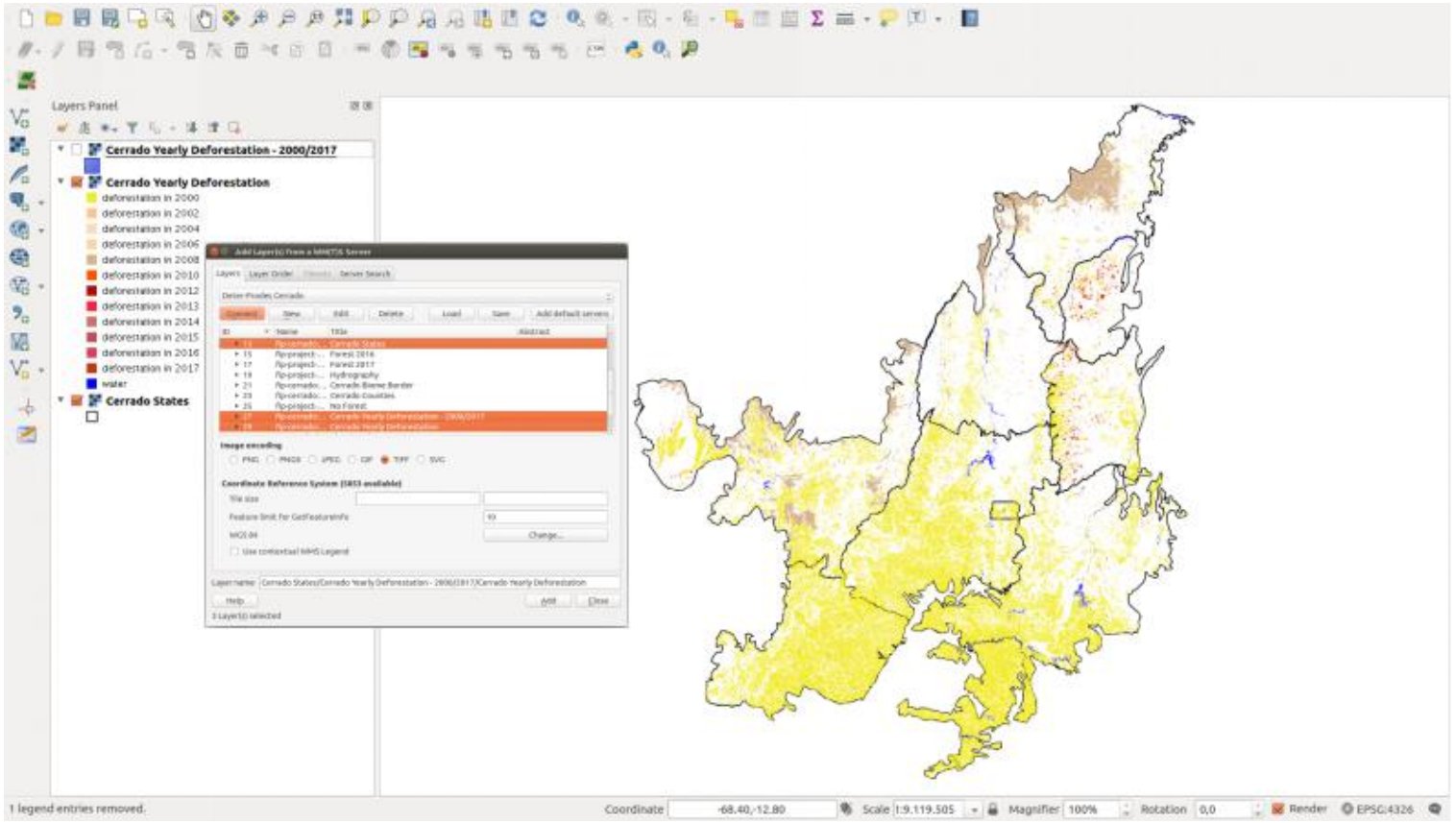
TerraBrasilis - GIS Interoperability

- The importance of OGC services
  - An international non-profit organization for the creation of spatial data dissemination standards.

- Web Map Service
  - Retrieve maps via the internet (http)

  - Combine maps from several sources regardless of the implementation

- Web Feature Service
  - Retrieve geographical features via the internet (http)



TerraBrasilis - Analytics API Environments

Optimizing writes and reads for Dashboards

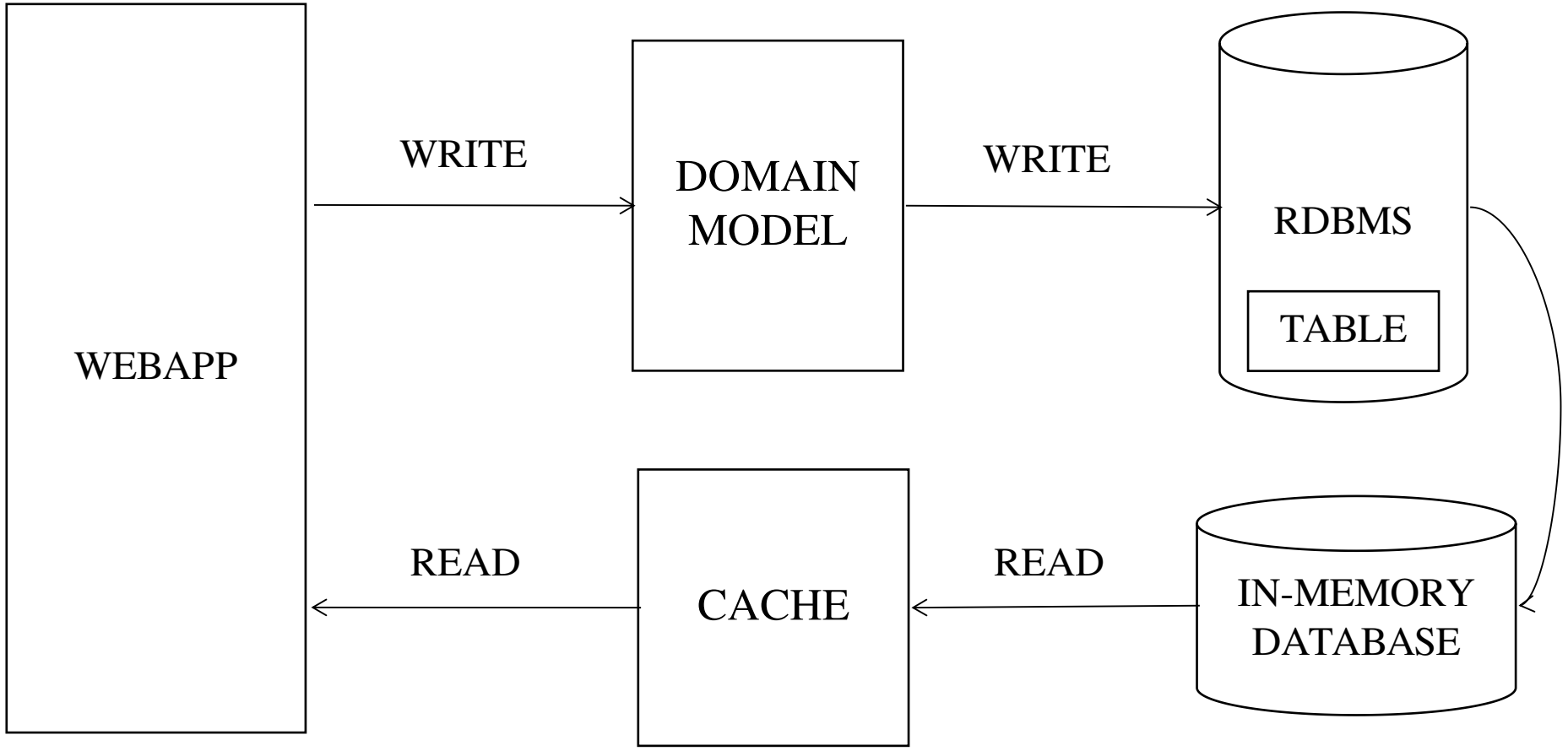
- Domain Model depicts the conceptual representation of the domain.

- Normally, the RDBMS is designed as close to the domain model.

- This result in a multiple layer representation, which is harder when lots of integration is necessary.

- CQRS allowed us to leave apart reads and writes model.

Command Query Responsibility Segregation Pattern



Data Storytelling using the Grammar of Graphics

- Fit deforestation data into the most appropriate story way for your audience.

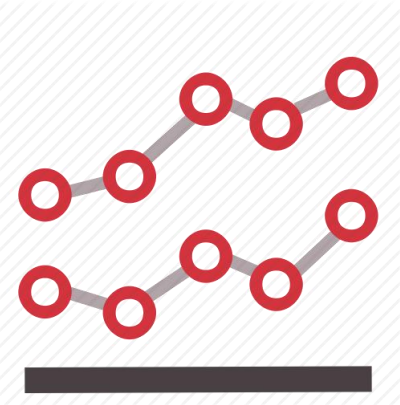
- Select visualizations metrics with clear goals that suit GIS specialists.

- Pre-process and clean the data properly.

- Get deeper into details to understand data better.

Data Storytelling using the Grammar of Graphics

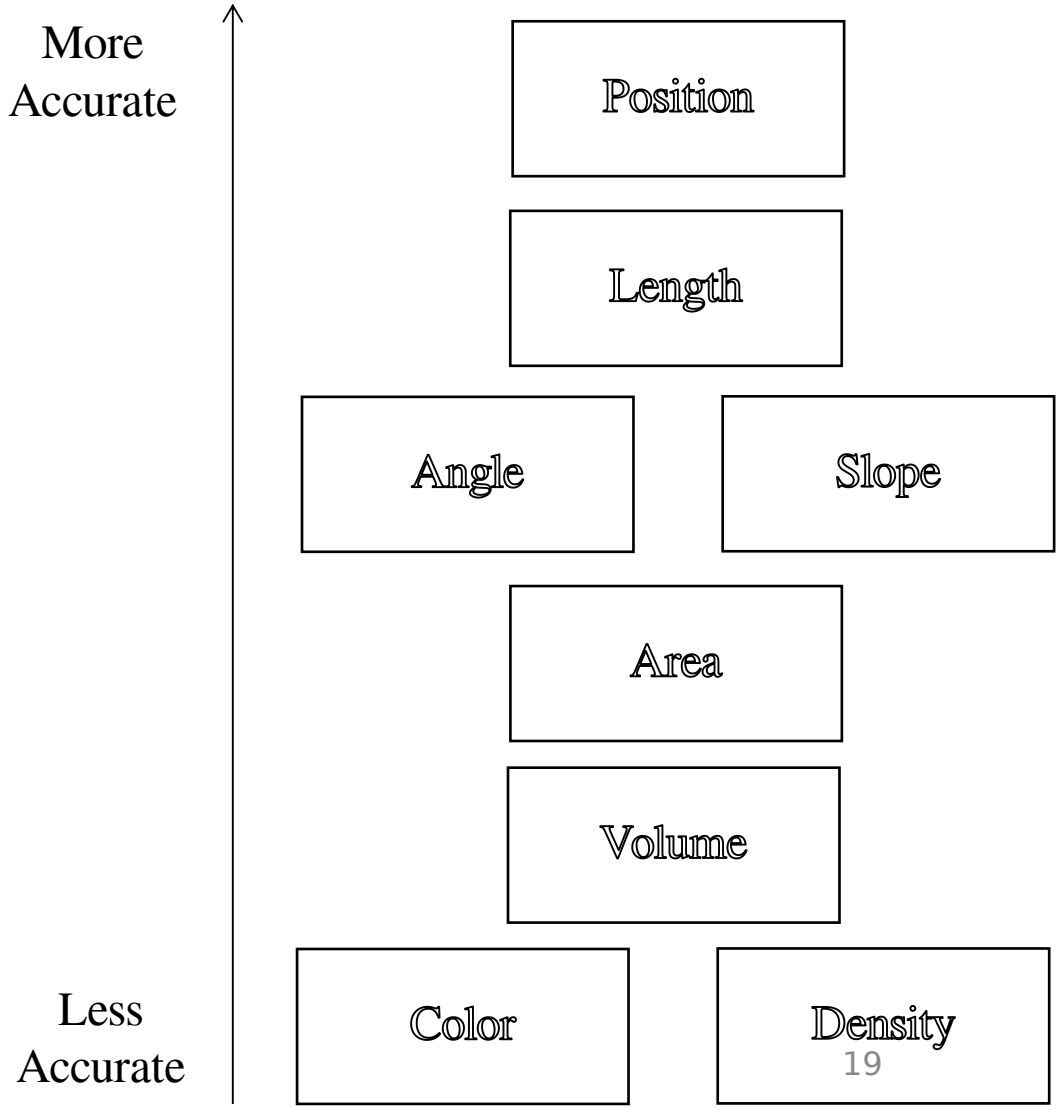
- A graph is constructed by means of quantitative and categorical information throughout position, shape, size, symbols, and color.
- "The first step is to **identify elementary graphical** perception elements that are used to visually extract quantitative information from a graph."
- This perception should come without apparent mental effort, including reading scale information.
- The ability of our preattentive visual system to detect geometric patterns and assess magnitudes.



Data Storytelling using the Grammar of Graphics

- After identifying those elementary graphical perception, they were ordered to provide a guide for **data display** that results in **more effective graphical perception**.

- **We try to avoid most graphic area** since humans' perception don't work well with attributing quantitative values in **two or three-dimensional space** (e.g., 3D pie charts).



Data Storytelling using the Grammar of Graphics

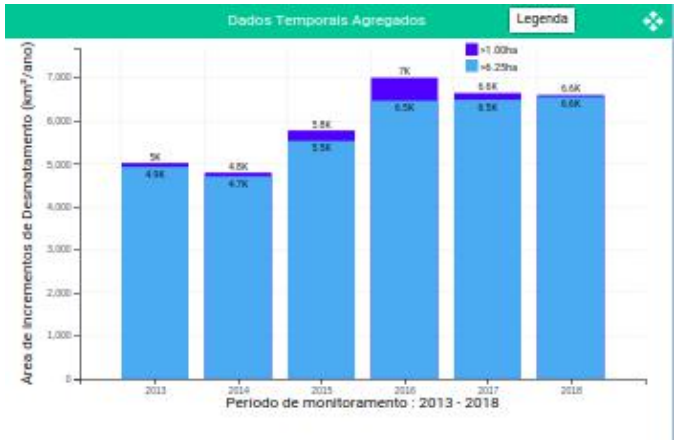
- A grammar of graphics enables the concisely description of the components of a graphic moving beyond named graphics (e.g., the “scatterplot”) into deep and formal structure that underlies statistical graphics.

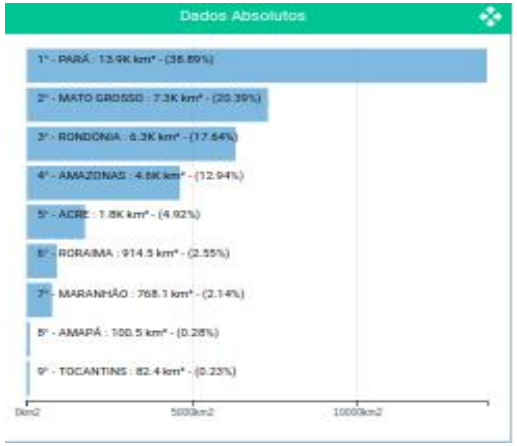
- A grammar of graphics embedds a graphical grammar into a programming language.

- A grammar of graphics helps in the convertion of such numbers measured in data units to numbers measured that the computer can display.

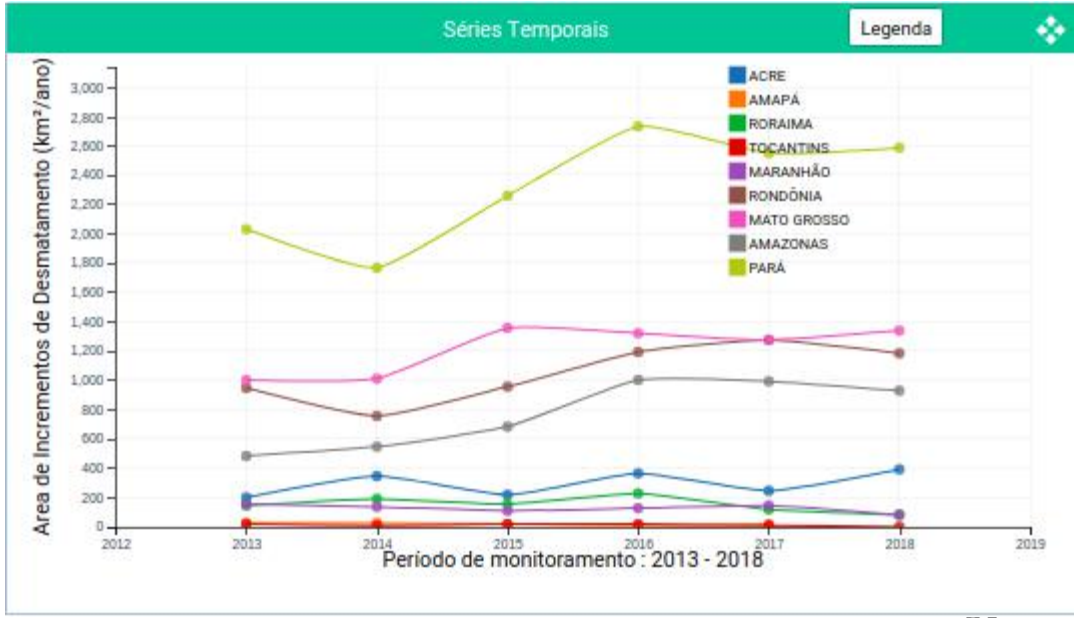
- Linear scales and a Cartesian coordinate system, which generates axes and legends so that users can read values from the graph.

Dashboards


Dados Temporais Agregados
Legenda
Area de Incrementos de Desmatamento (km²/ano)
Período de monitoramento : 2013 - 2018

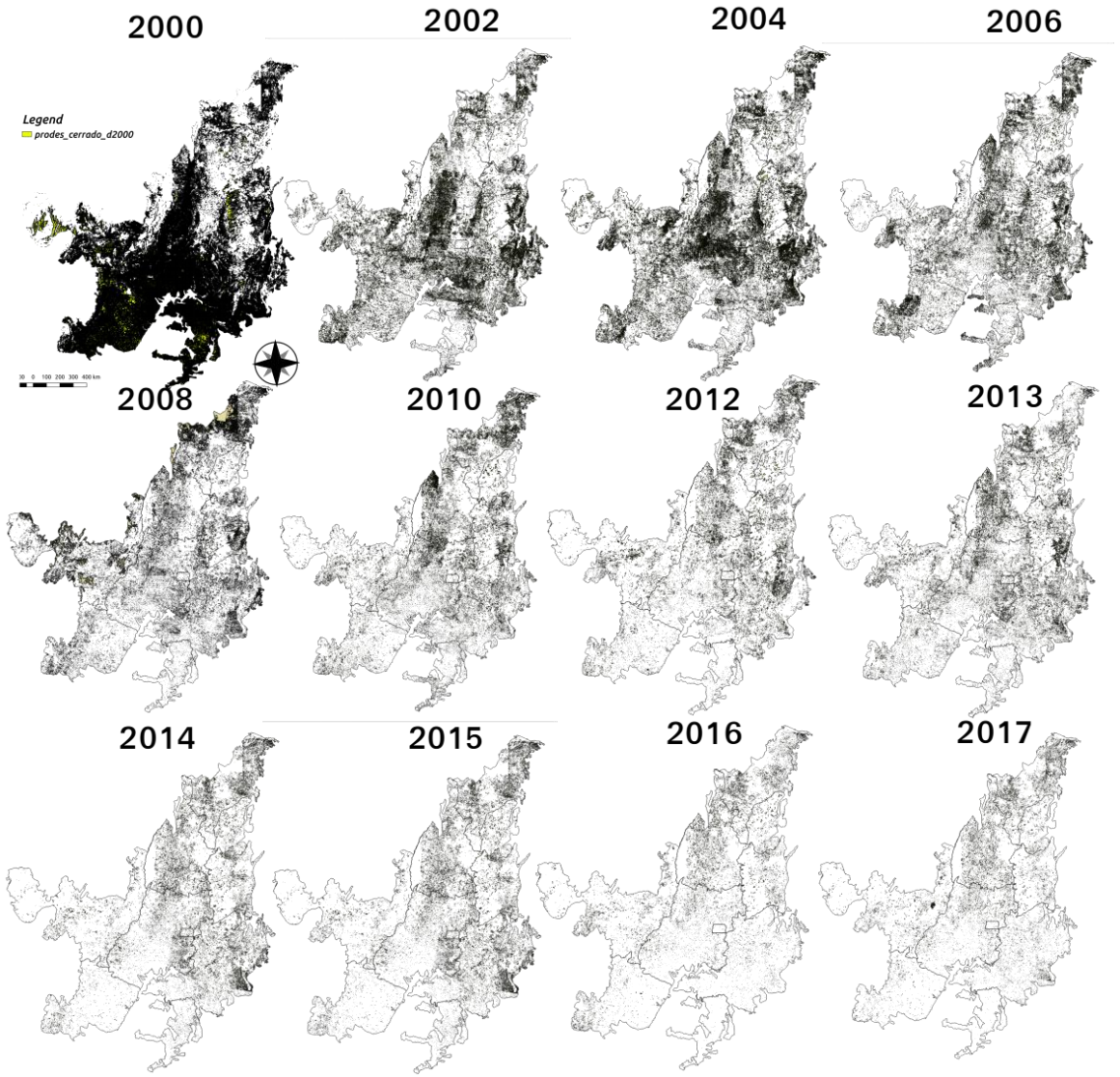

Dados Absolutos


Dados Espaciais Agregados
Caracas
Venezuela
Panamá
Medellin
Colombia
Guyana
Suriname
Ecuador
Perú
Brasil
Brasília
Bolivia
Paraguay
Min – Max (km²)
0 – 1741
1741 – 3482
3482 – 5223
5223 – 6964
6964 – 8705
8705 – 10446
10446 – 12187
12187 – 13928
13928+
Leaflet | Map data © OpenStreetMap contributors, CC-BY-SA,

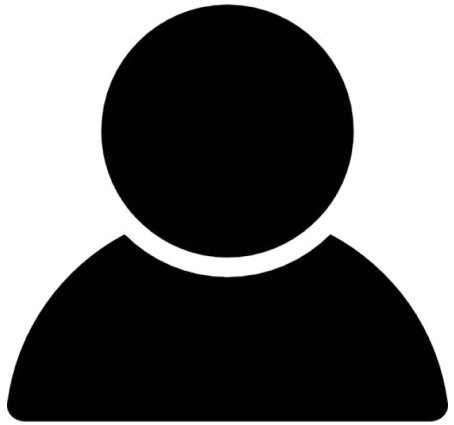

Séries Temporais
Legenda
ACRE
AMAPÁ
RORAIMA
TOCANTINS
MARANHÃO
RONDÔNIA
MATO GROSSO
AMAZONAS
PARÁ
Area de Incrementos de Desmatamento (km²/ano)
Período de monitoramento : 2013 - 2018

Results and Discussions

PRODES and DETER Cerrado Projects Data Handling using TerraBrasilis: Lessons Learned from Deforestation Data in Cerrado Biome

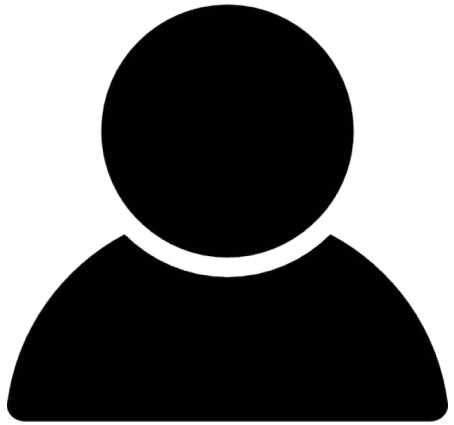


PRODES and DETER Cerrado Projects Data Handling using TerraBrasilis: Lessons Learned from Deforestation Data in Cerrado Biome



"Bring me information about deforestation!"

PRODES and DETER Cerrado Projects Data Handling using TerraBrasilis: Lessons Learned from Deforestation Data in Cerrado Biome

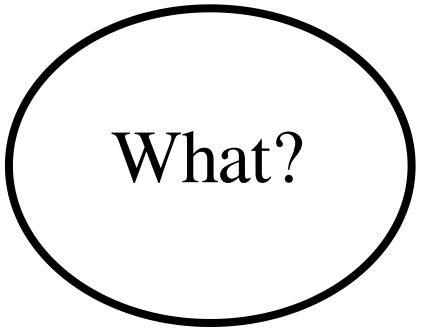


PRODES and DETER Cerrado Projects Data Handling using TerraBrasilis: Lessons Learned from Deforestation Data in Cerrado Biome

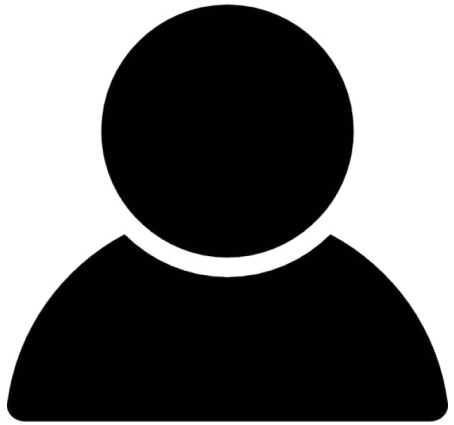
PRODES and DETER Cerrado Projects Data Handling using TerraBrasilis: Lessons Learned from Deforestation Data in Cerrado Biome

"Bring me information about deforestation in **Mato Grosso State**!"

PRODES and DETER Cerrado Projects Data Handling using TerraBrasilis: Lessons Learned from Deforestation Data in Cerrado Biome

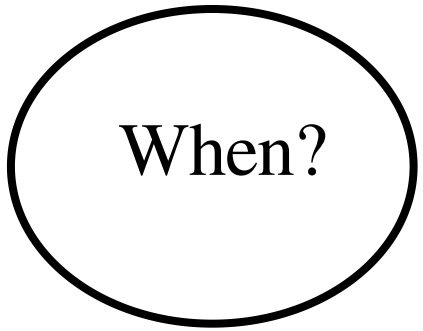


PRODES and DETER Cerrado Projects Data Handling using TerraBrasilis: Lessons Learned from Deforestation Data in Cerrado Biome

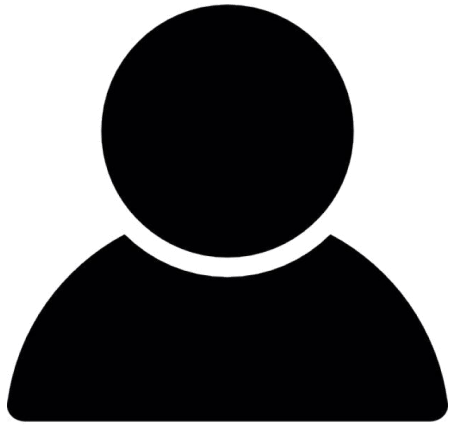


"Bring me **all the deforestation data** in Mato Grosso State!"

PRODES and DETER Cerrado Projects Data Handling using TerraBrasilis: Lessons Learned from Deforestation Data in Cerrado Biome

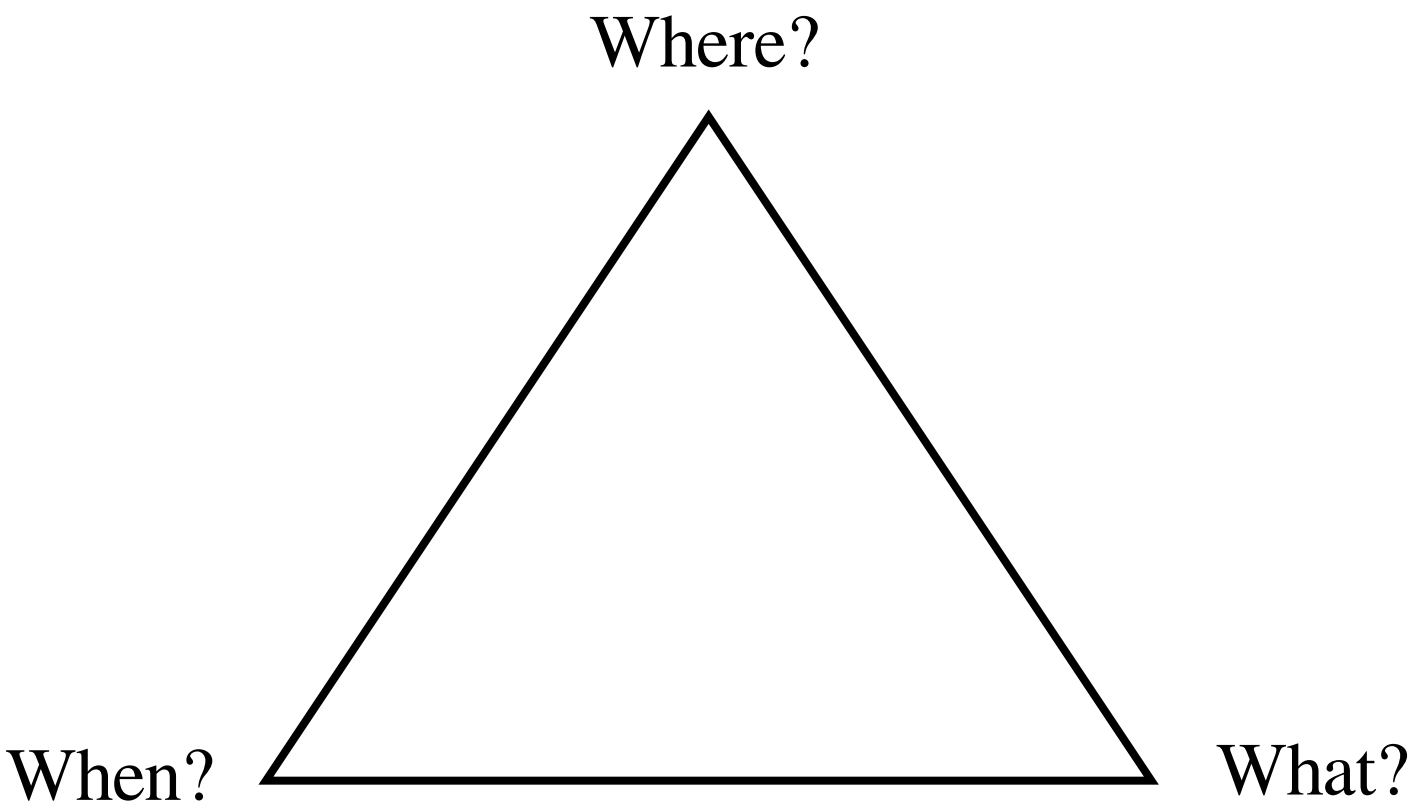


PRODES and DETER Cerrado Projects Data Handling using TerraBrasilis: Lessons Learned from Deforestation Data in Cerrado Biome



"Bring me **the deforestation data** in **Mato Grosso State in 2017**!"

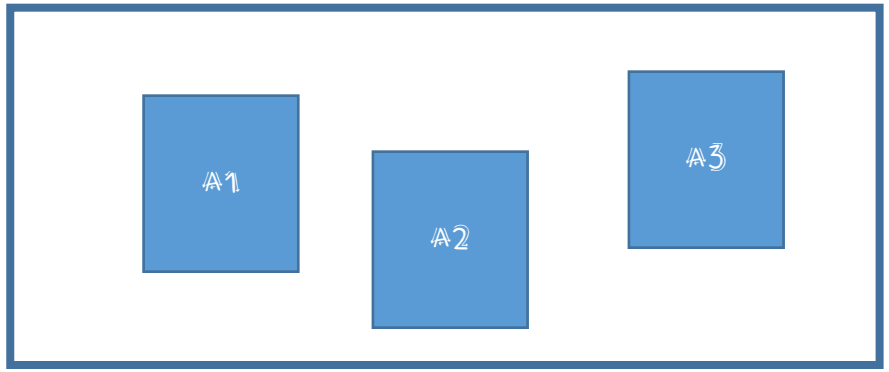
PRODES and DETER Cerrado Projects Data Handling using TerraBrasilis: Lessons Learned from Deforestation Data in Cerrado Biome



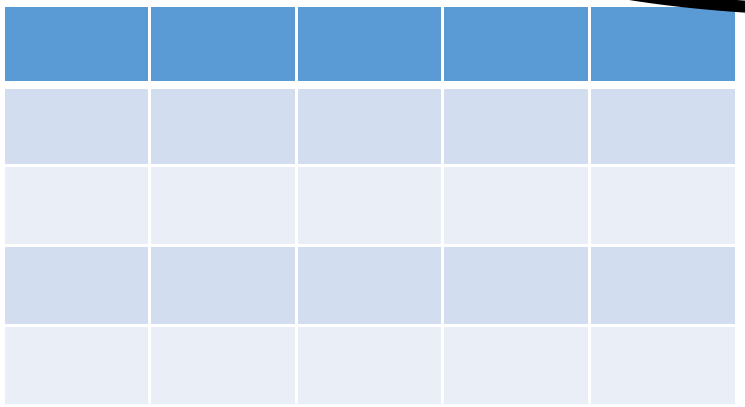
PRODES and DETER Cerrado Projects Data Handling using TerraBrasilis: Lessons Learned from Deforestation Data in Cerrado Biome

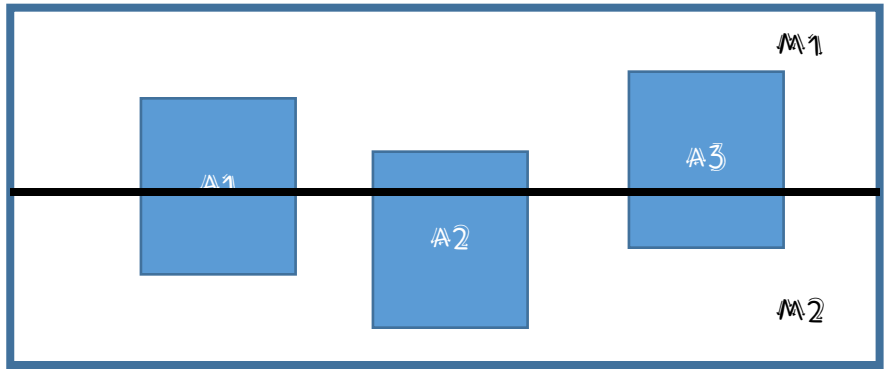
"The need to intersect everything!"

N-columns approach vs Multi-table approach

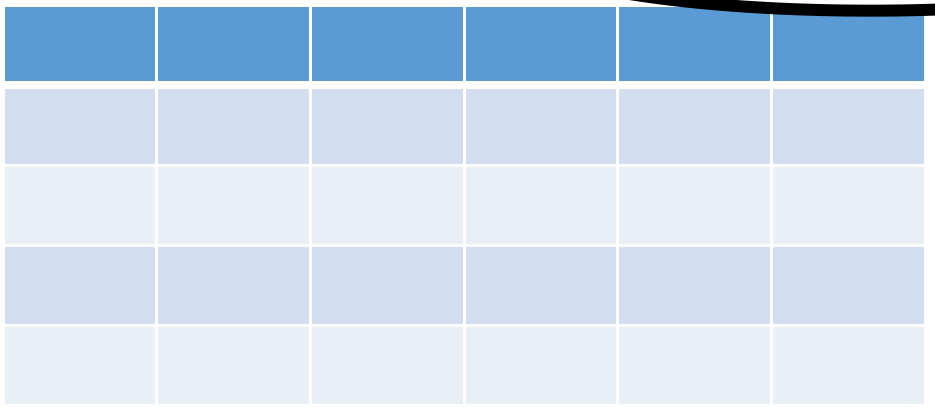


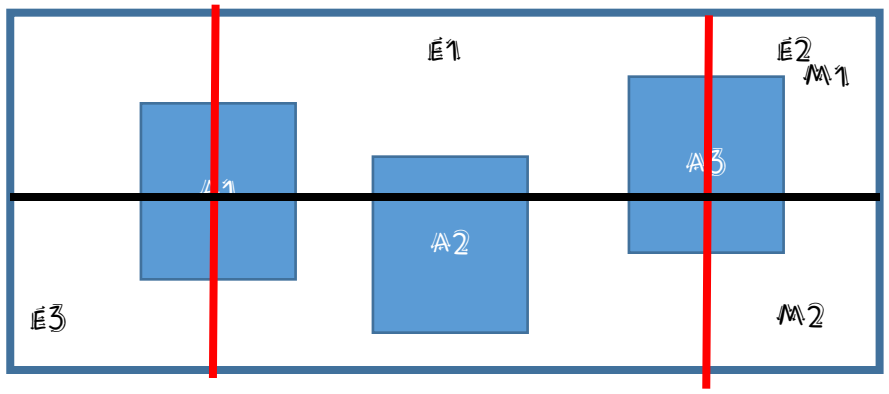
N-columns approach vs Multi-table approach



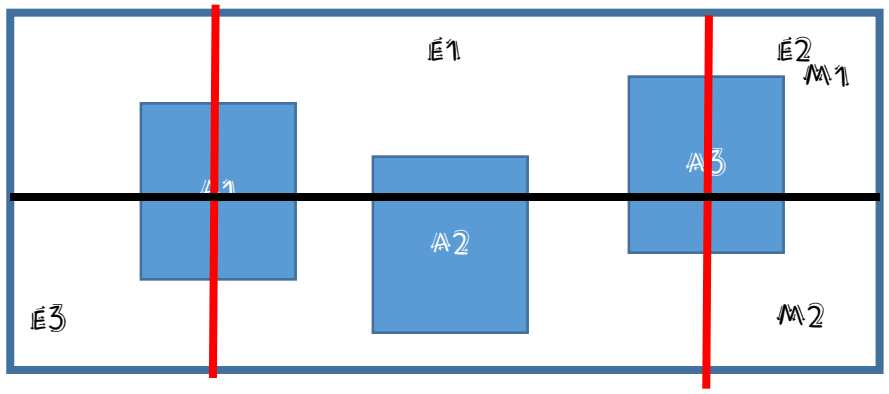


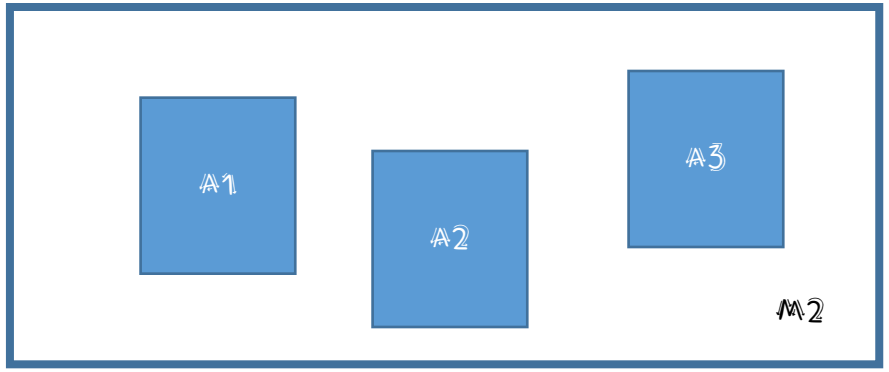
N-columns approach vs Multi-table approach



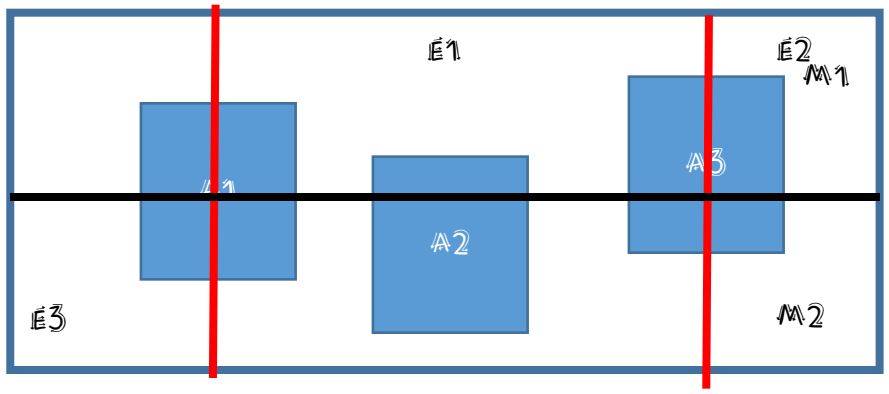


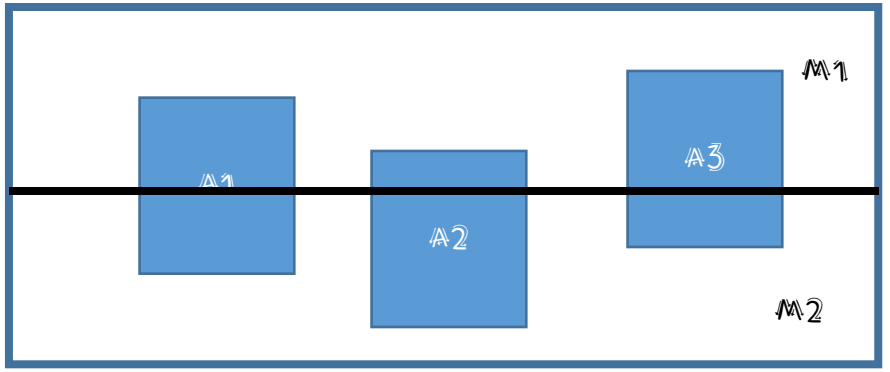
N-columns approach vs Multi-table approach



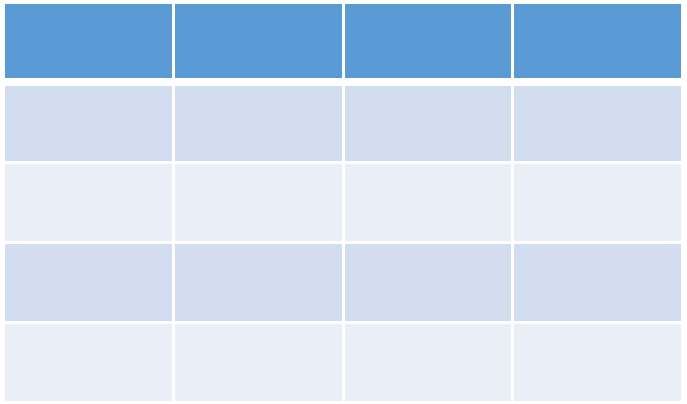


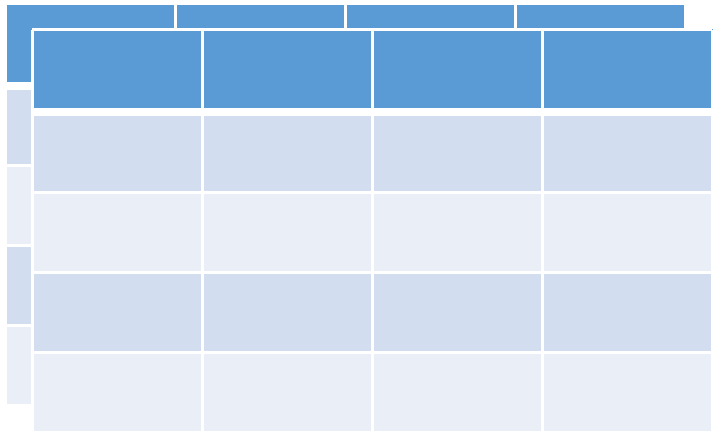
N-columns approach vs Multi-table approach

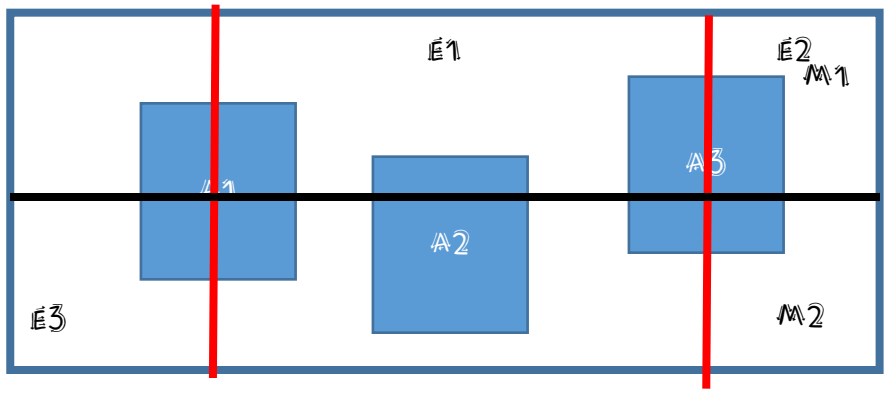


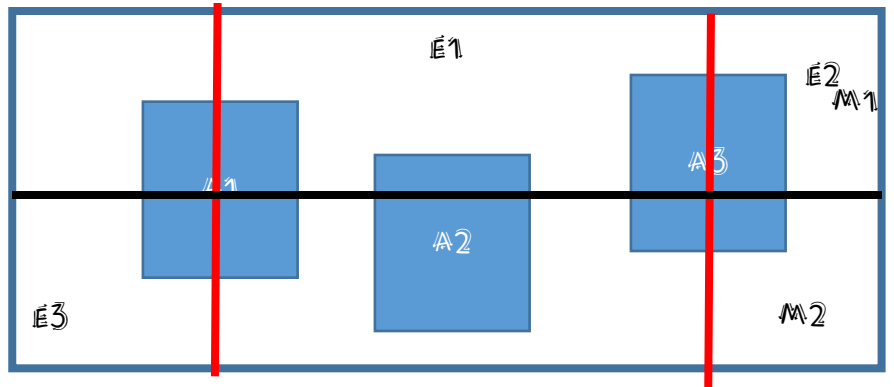


N-columns approach vs Multi-table approach

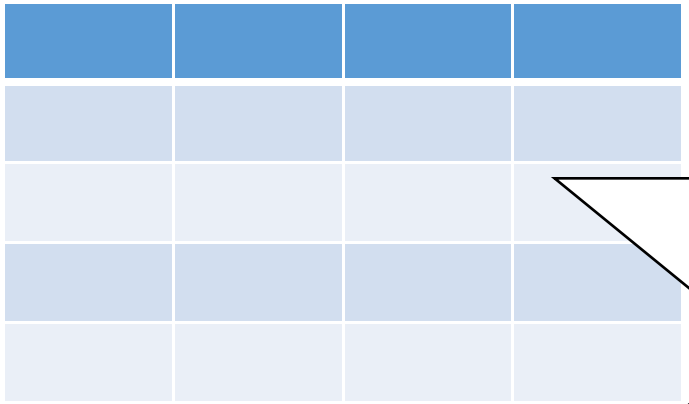




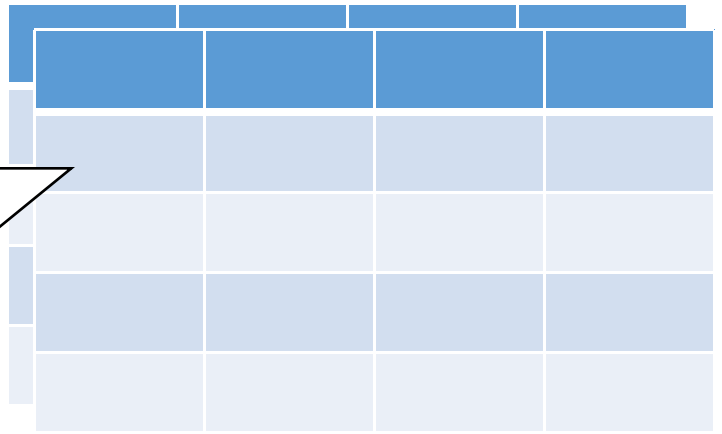


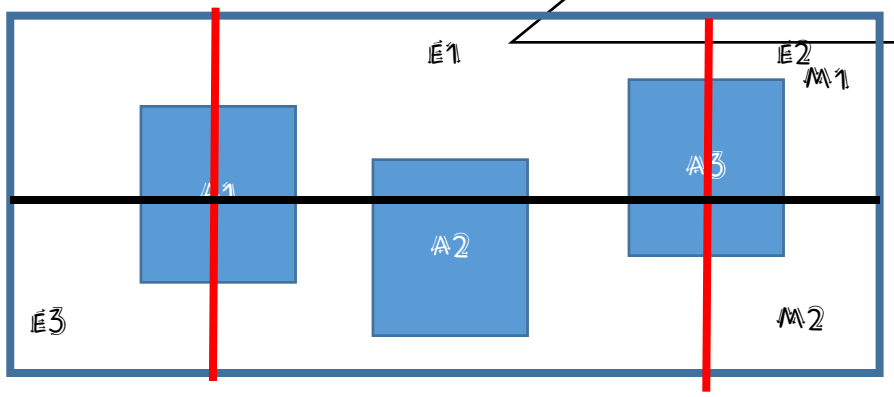


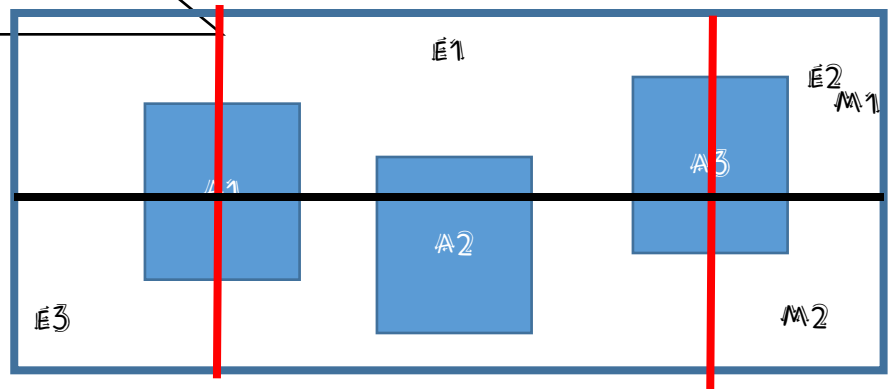
N-columns approach vs Multi-table approach



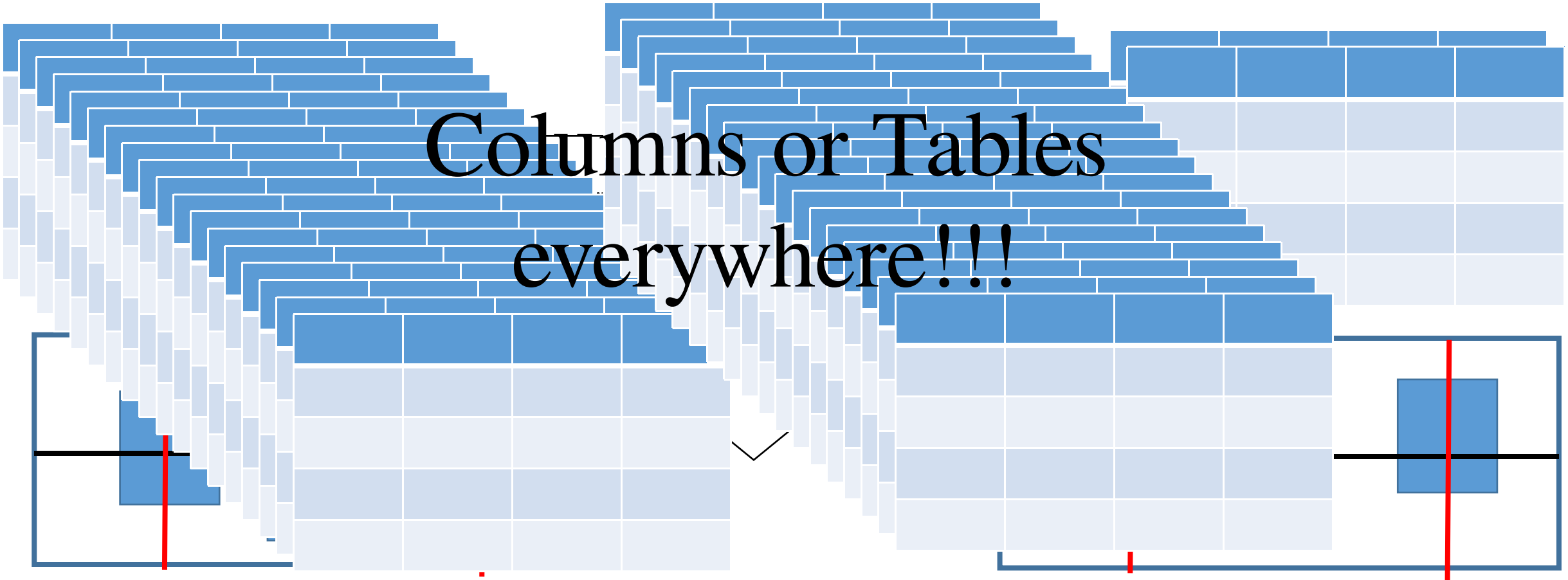
What do we do if we have many projects and local of interests??



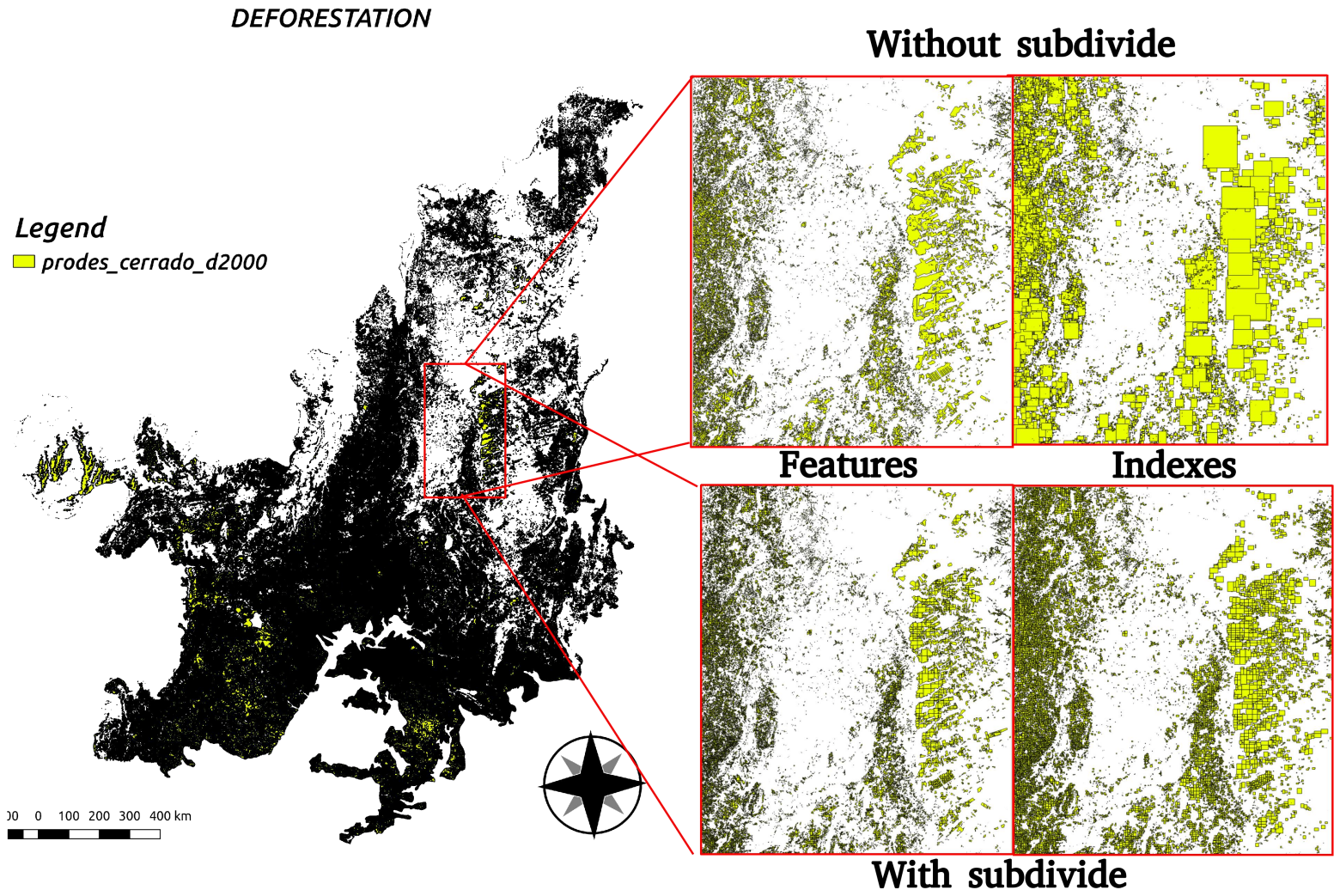




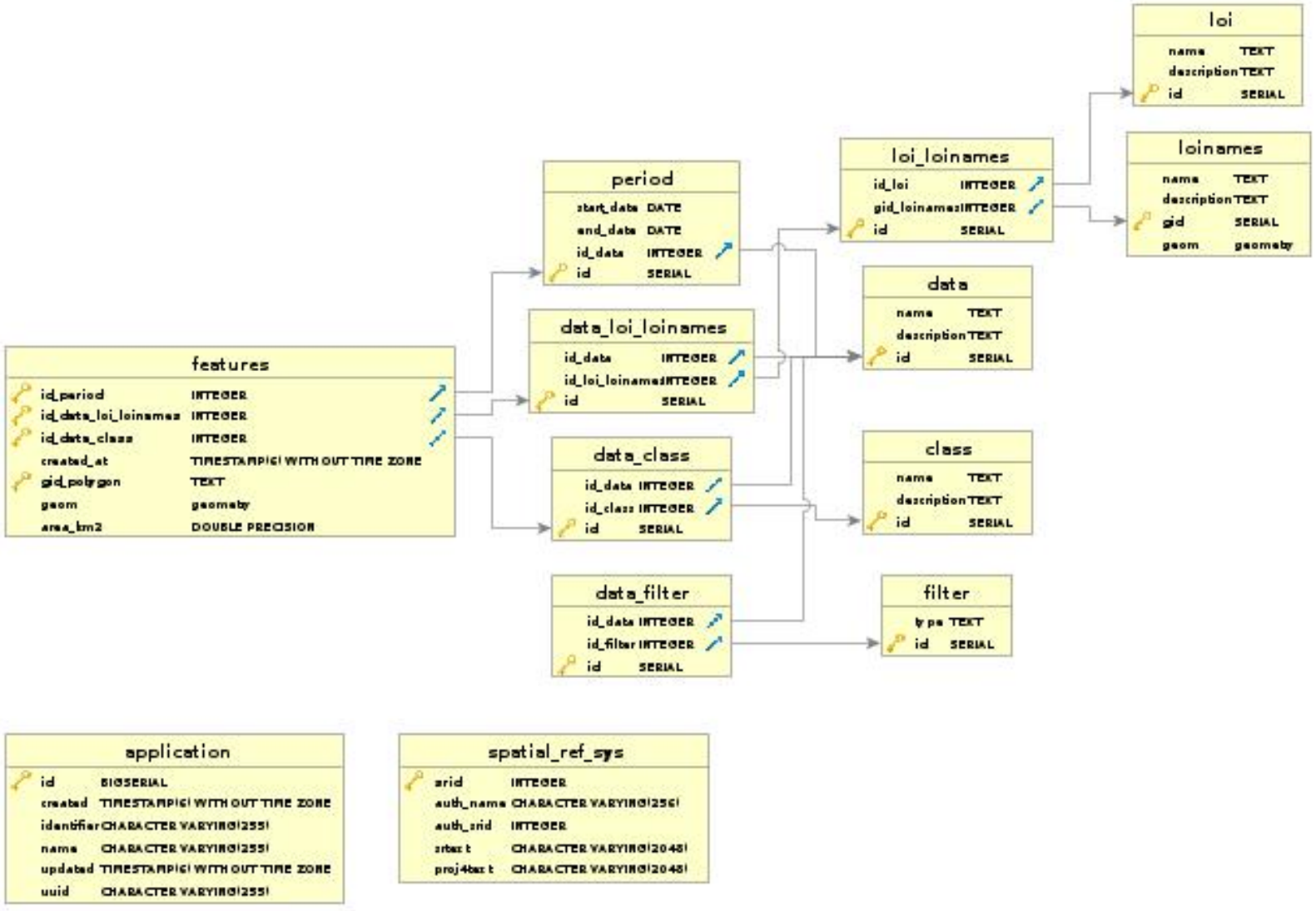
N-columns approach vs Multi-table approach



``SubDivide'' and Conquer: Tunning Spatial Database Operations for Query Performance Optimization



``SubDivide'' and Conquer: Tunning Spatial Database Operations for Query Performance Optimization



TerraBrasilis - Analytics API Environments

In [1]:
```
devtools::install_github("terrabrasilis/terrabrasilisAnalyticsAPI") # github group name is terrabrasilis
library(terrabrasilisAnalyticsAPI) # R package name is terrabrasilisAnalyticsAPI
```

```
Downloading GitHub repo terrabrasilis/terrabrasilisAnalyticsAPI@master
from URL https://api.github.com/repos/terrabrasilis/terrabrasilisAnalyticsAPI/zipball/master
Installing terrabrasilisAnalyticsAPI
Installing curl
'/usr/lib/R/bin/R' --no-site-file --no-environ --no-save --no-restore --quiet  \
  CMD INSTALL '/tmp/RtmpwtejnM/devtools1a18caf118/curl'  \
  --library='/srv/rlibs' --install-tests

'/usr/lib/R/bin/R' --no-site-file --no-environ --no-save --no-restore --quiet  \
  CMD INSTALL  \
  '/tmp/RtmpwtejnM/devtools1a23df3a32/Terrabrasilis-terrabrasilisAnalyticsAPI-880e4ac'  \
  --library='/srv/rlibs' --install-tests
```

In [2]:
```
tbaAPIPath <- "http://terrabrasilis.dpi.inpe.br/dashboard/api/v1/redis-cli/"
```

In [3]:
```
appIdentifier <- tba_list_apps_identifier(tbaAPIPath)
```

In [4]:
```
appIdentifier
```

identifier	name	created
prodes_cerrado	Dashboard of the Prodes in the Cerrado	2019-03-20 23:30
prodes_amazon	Dashboard of the Prodes in the Amazon Forest	2019-03-20 23:30
prodes_legal_amazon	Dashboard of the Prodes in the Legal Amazon Forest	2019-03-20 23:37

TerraBrasilis - Analytics API Environments

In [6]:
```
prodesCerrado <- appIdentifier$identifier[1]
```

In [7]:
```
periods <- tba_list_periods(tbaAPIPath, prodesCerrado)

periods
```

startDate.year	startDate.month	startDate.day	endDate.year	endDate.month	endDate.day
1988	8	1	2000	7	31
2000	8	1	2002	7	31
2002	8	1	2004	7	31
2004	8	1	2006	7	31
2006	8	1	2008	7	31
2008	8	1	2010	7	31
2010	8	1	2012	7	31
2012	8	1	2013	7	31
2013	8	1	2014	7	31
2014	8	1	2015	7	31
2015	8	1	2016	7	31
2016	8	1	2017	7	31
2017	8	1	2018	7	31

In [8]:
```
classes <- tba_list_classes(tbaAPIPath, prodesCerrado)

classes
```

id	name	description
1	deforestation	It is the process of complete and permanent disappearance of forests

TerraBrasilis - Analytics API Environments

In [9]:
```
lois <- tba_list_lois(tbaAPIPath, prodesCerrado)

lois
```

gid	name
1	UF
2	MUN
3	ConsUnit
4	Indi
5	Pathrow

In [10]:
```
loinames <- tba_list_loinames(tbaAPIPath, prodesCerrado)

loinames[20:30,]
```

gid	loiname	loi
1481	PARQUE ESTADUAL SERRA VERDE	3
1566	RESERVA PARTICULAR DO PATRIMÔNIO NATURAL JOAQUIM THEODORO DE MORAES	3
1567	PARQUE ESTADUAL DE VASSUNUNGA	3
1404	ESTAÇÃO ECOLÓGICA ITABERÁ	3
1421	ÁREA DE PROTEÇÃO AMBIENTAL LAGO DE PEIXE/ANGICAL	3
1405	ÁREA DE PROTEÇÃO AMBIENTAL DO SALTO MAGESSI	3
1446	ESTAçãO ECOLóGICA DE SANTA BáRBARA	3
1603	RESERVA DE DESENVOLVIMENTO SUSTENTáVEL NASCENTES GERAIZEIRAS	3
1649	RESERVA EXTRATIVISTA EXTREMO NORTE DO TOCANTINS	3
1626	RESERVA BIOLÓGICA DA CONTAGEM	3
1658	RESERVA PARTICULAR DO PATRIMÔNIO NATURAL PONTE DE PEDRA	3

In [11]:
```
loiUF = dplyr::filter(lois, grepl("UF", name))$gid

loinamesByLoi <- tba_list_loinamesByLoi(tbaAPIPath, prodesCerrado, loiUF)

loinamesByLoi
```

gid	loiname
11	TOCANTINS
7	MINAS GERAIS
10	SÃO PAULO
9	PIAUÍ
12	PARANÁ
13	RONDÔNIA
4	MARANHÃO
5	MATO GROSSO
2	DISTRITO FEDERAL
1	BAHIA
3	GOIÁS
6	MATO GROSSO DO SUL
8	PARÁ

TerraBrasilis - Analytics API Environments

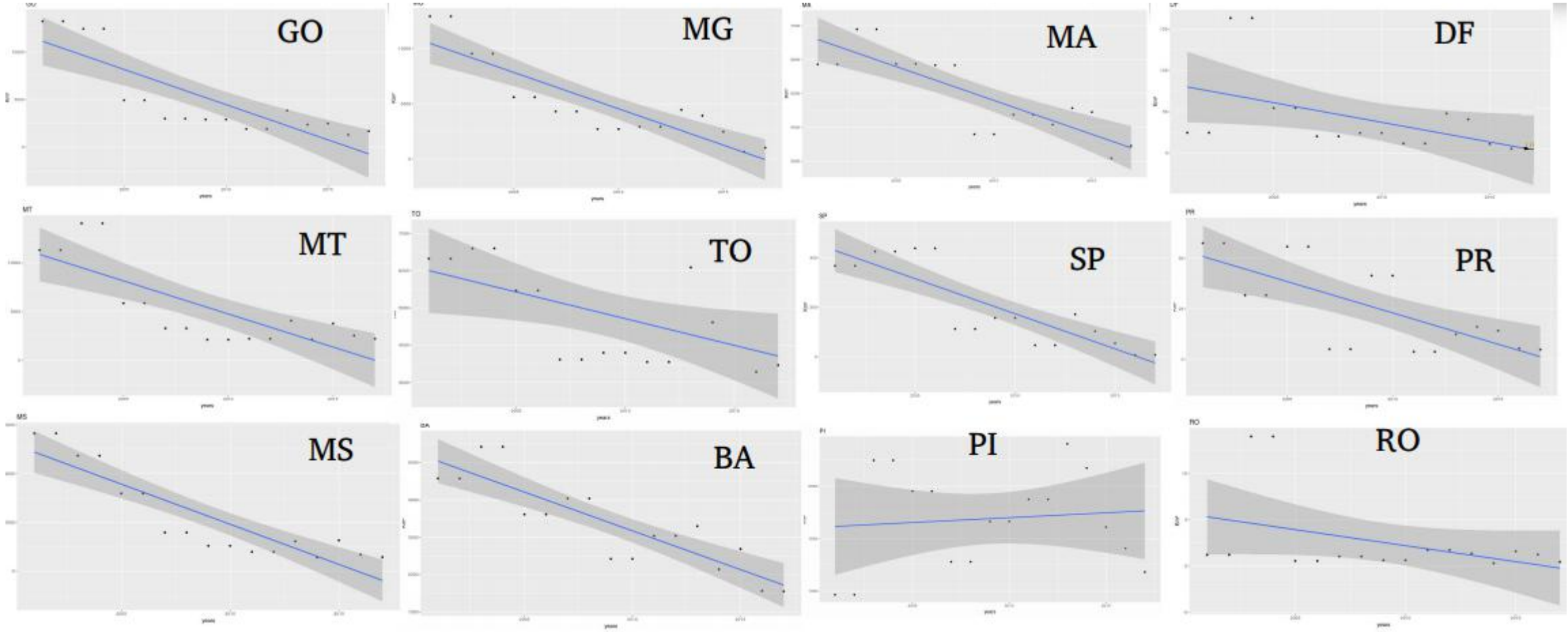
In [12]:
```
datafilters <- tba_list_filters(tbaAPIPath, prodesCerrado)

datafilters
```

id	type
1	fid_area >= 0.0625
2	fid_area >= 0.01

In [13]:
```
data <- tba_get_dataByLoiname(tbaAPIPath, prodesCerrado, classes$name, loinamesByLoi[1,]$gid)

startDate <- data$periods$startDate[1,]

endDate <- data$periods$endDate[1,]

loi <- data$periods$features[[1]]$loi

loiname <- data$periods$features[[1]]$loiname

areas <- data$periods$features[[1]]$areas

finalDF <- cbind(loi = loi,
                 loiname = loiname,
                 startDate = startDate,
                 endDate = endDate,
                 areas,
                 row.names = NULL)

finalDF
```

loi	loiname	startDate.year	startDate.month	startDate.day	endDate.year	endDate.month	endDate.day	type	area
1	11	1988	8	1	2000	7	31	1	38003.72
1	11	1988	8	1	2000	7	31	2	38286.16

TerraBrasilis - Analytics API Environments



CONCLUSIONS

Final Remarks

- Free and open source software as a paradigm.
- Moving from traditional Geoinformatics specialists into spatial data scientists.
- **Open Science and Open Data** have been increasingly deployed in the last few years.

Acknowledgements

FIP FOREST INVESTMENT PROGRAM

GRUPO BANCO MUNDIAL

FUNDEP
Fundação de Desenvolvimento da Pesquisa

Questions