

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
CURSO DE GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

JOSÉ HENRIQUE DA SILVA BRAZ

**Uma análise dos dados de queimada do
INPE no Brasil (preliminar)**

Monografia apresentada como requisito parcial
para a obtenção do grau de Bacharel em Ciência
da Computação

Orientador: Prof. Dr. Lucas M. Schnorr

Porto Alegre
2023

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos André Bulhões

Vice-Reitora: Prof^a. Patricia Pranke

Pró-Reitora de Graduação: Prof^a. Cíntia Inês Boll

Diretora do Instituto de Informática: Prof^a. Carla Maria Dal Sasso Freitas

Coordenador do Curso de Ciência de Computação: Prof. Marcelo Walter

Bibliotecário-chefe do Instituto de Informática: Alexsander Borges Ribeiro

*“If I have seen farther than others,
it is because I stood on the shoulders of giants.”*

— SIR ISAAC NEWTON

AGRADECIMENTOS

Agradeço ao L^AT_EX por não ter vírus de macro...

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS	6
LISTA DE FIGURAS	7
LISTA DE TABELAS	8
RESUMO	9
ABSTRACT	10
1 INTRODUÇÃO	11
2 VISÃO GERAL DOS DADOS.....	12
2.1 O programa DBQueimadas	12
2.2 Garimpando os dados.....	12
2.3 Como interpretar os dados.....	13
2.4 Carregando os dados para análise.....	15
2.5 Os Satélites.....	16
2.6 O que os dados gritam	18
3 TRABALHOS RELACIONADOS	21
4 APROFUNDANDO A ANÁLISE DOS DADOS.....	22
4.1 Densidade e Centrografia	22
4.2 Validade dos dados.....	22
4.3 Padronizando os dados por satélite	22
REFERÊNCIAS.....	23

LISTA DE ABREVIATURAS E SIGLAS

API	Application Programming Interface (Interface de Programação de Aplicação)
CSV	Comma Separated Values (valores separados por vírgulas).
GMT	Greenwich Mean Time
INPE	Instituto Nacional de Pesquisas Espaciais
IBGE	Instituto Brasileiro de Geografia e Estatística
URL	Uniform Resource Locator (Localizador Uniforme de Recursos)
NOAA	National Oceanic and Atmosphere Administration
MODIS	Moderate Resolution Imaging Spectroradiometer
GOES	Geostationary Operational Environmental Satellite
AVHRR	Advanced Very High Resolution Radiometer

LISTA DE FIGURAS

Figura 2.1	Relação do montante dos dados por satélite	18
Figura 2.2	Amostragem por tempo de cada satélite	19
Figura 2.3	Órbita dos satélites no dia 10 de agosto de 2022.....	19

LISTA DE TABELAS

Tabela 2.1	Significado de cada coluna dos dados de queimada do INPE	14
Tabela 2.2	Características dos satélites usados pelo INPE	16

RESUMO

Este documento é um exemplo de como formatar documentos para o Instituto de Informática da UFRGS usando as classes \LaTeX disponibilizadas pelo UTUG. Ao mesmo tempo, pode servir de consulta para comandos mais genéricos. *O texto do resumo não deve conter mais do que 500 palavras.*

Palavras-chave: Formatação eletrônica de documentos. \LaTeX . ABNT. UFRGS.

Using L^AT_EX to Prepare Documents at II/UFRGS

ABSTRACT

This document is an example on how to prepare documents at II/UFRGS using the L^AT_EX classes provided by the UTUG. At the same time, it may serve as a guide for general-purpose commands. *The text in the abstract should not contain more than 500 words.*

Keywords: Electronic document preparation. L^AT_EX. ABNT. UFRGS.

1 INTRODUÇÃO

P1. Importancia do trabalho (as maiores emissões de gases de efeito estufa do brasil são por queimadas)

P2. DBQueimadas Ferramentas opensource e códigos gerados também

P3. Introducao aos dados

2 VISÃO GERAL DOS DADOS

Neste capítulo constam algumas informações importantes sobre os dados disponibilizados pelo Instituto Nacional de Pesquisas Espaciais (INPE), que serão cruciais para compreensão dos próximos capítulos.

2.1 O programa DBQueimadas

O DBQueimadas, Banco de Dados de Queimadas <www.inpe.br/queimadas/bdqueimadas>, é um sistema desenvolvido pelo INPE e acessível de forma aberta por meio da web. Conta com mais de 300 milhões de pontos coletados desde o ano de 1998, proveniente de vários satélites. Dentro do site é possível gerar mapas, tabelas, gráficos e exportar os dados sobre as queimadas no Brasil aplicando diferentes filtros. Todo o programa foi desenvolvido com ferramentas abertas, muitas delas criadas pelo próprio time de tecnologia da informação do INPE (SETZER; MORELLI; SOUZA, 2019). [P1. Falamos sobre o programa]

O Banco de Dados de Queimadas é um excelente caso de como os dados abertos podem ajusar a sociedade. O portal vem evoluindo muito na medida que os anos passam, sempre focado em ser simples se usar. [P2. Ressaltamos a importancia dos dados abertos para a sociedade]

2.2 Garimpando os dados

Uma parte importante do processo foi coletar os dados do DBQueimadas. Para exportar os dados pela web, é necessário preencher os campos de data inicial, data final e um endereço de e-mail, o intervalo de tempo não pode exceder 366 dias. Também é possível aplicar filtros ainda mais detalhados como: continente, país, estado, município, satélite, bioma e unidades de conservação/terras indígenas. Após clicar em "Exportar", uma mensagem é enviada para o e-mail informado com o link de download dos dados requisitados. O arquivo disponibilizado é um CSV compactado como um zip. [P1. Contextualizar como é a exportação de dados]

Apesar de ser um site com boas métrica e usabilidades, seria praticamente inviável baixar todos os dados do Brasil de forma manual. Nesse sentido, foi necessário entender

quais eventos são disparados quando solicitamos os dados pelo site a fim de automatizar o processo de download. [P2. Motivar a abordagem automatizada]

Foi identificado que na verdade o site faz uma requisição GET para a API do DBQueimadas, localizada em <<https://queimadas.dgi.inpe.br/queimadas/exportacaobdq/exportar>>, passando nos parâmetros da URL os filtros aplicados, codificados em JSON. Além dos filtros, também é necessário informar o e-mail e o formato de arquivo desejado. Um exemplo de uso dessa API, por meio de uma invocação CURL, pode ser encontrado no ?? [P3. Explicar como a api funciona]

A fim de automatizar o processo, foi desenvolvido um script em Python que solicita os dados referentes a 30 dias, totalizando 300 requisições de 1998 até 2022. Com o intuito de não sobrecarregar o servidor do INPE, foi adicionada uma espera de um minuto a cada requisição. [P4. Falar sobre os scripts de coleta dos dados]

Para o processo ser concluído, ainda seria necessário fazer o download do arquivo por meio do link enviado por e-mail. Lançou-se mão do Google Scripts, uma ferramenta que possibilita escrever programas simples, em uma linguagem parecida com JavaScript, e tem uma ótima integração com os serviços do Google (como o Gmail). A partir dessa ferramenta foi possível extrair o link de cada mensagem e finalmente salvar o dado de forma automatizada. [P5. Processo de baixar os dados para o computador]

Todos esse processo de investigação e recuperação dos dados levou cerca de uma semana. Todos os arquivos baixados ocupam pouco mais de 4 Gigabytes de armazenamento em disco e somados tem exatamente 43.782.758 linhas. Ao final, eles foram recompatados em um único zip (450 Megabytes) e estão disponíveis em <<https://bit.ly/3IgHIXH>> para download de forma independente aos servidores do INPE. [P6. Conclusão do processo]

2.3 Como interpretar os dados

Antes de mais nada, é importante entender que os focos ativos de calor, representados por uma linha na tabela dos dados de queimadas, é fruto do processamento em cima de uma imagem de satélite. O INPE ainda disponibiliza junto com as coordenadas da detecção mais alguns dados como risco de fogo, poder do fogo, precipitação e dados da região do foco. A lista completa das colunas pode ser vista na Tabela 2.1 (INPE, 2023). [P0. O que é uma observação nos dados]

Para obter as imagens brutas dos satélites são necessárias antenas especiais que

Tabela 2.1 – Significado de cada coluna dos dados de queimada do INPE

Atributo	Tipo	Descrição
Id	string	Identificador único registrado no banco
Latitude	double	Graus decimais da latitude do centro do pixel de fogo ativo (valores de 90.0000 até -90.0000)
Longitude	double	Graus decimais da longitude do centro do pixel de fogo ativo (valores de 180.0000 até -180.0000)
DataHora	string	Data a hora da passagem do satélite no fuso horário de Greenwich (GMT)
Município	string	Nome do município, de acordo com os dados do IBGE 2000
Estado	string	Nome do estado
País	string	Nome do país
Bioma	string	Nome do bioma brasileiro, de acordo com dados do IBGE 2004 (para outros países o campo fica vazio)
Precipitação	double	Valor a precipitação do dia até o horário da medida (-999 para valores inválidos)
DiasSCh	integer	Dias sem chuva até a data da medida (-999 para valores inválidos)
RiscoFog	double	Valor do risco de fogo previsto naquele dia (-999 para valores inválidos)
FRP	double	Fire Radiative Power, MW (megawatts)

Fonte: O Autor com base em INPE (2023)

ficam em centros de recepção de dados. Com esse propósito, a Divisão de Geração de Imagens (DGI) do INPE possui duas Estações de Recepção e Gravação (ERG) - a primeira em Cachoeira Paulista (SP) e uma mais recente em Cuiabá (MT). Na estação de SP, é feito o processamento de mais de 200 imagens de diversos satélites todos os dias, extraindo os dados de focos ativos de queimadas que alimentam o DBQueimadas. (Divisão de Geração de Imagens, 2023) [P1. Apresentação do instituto como processador de dados]

O algoritmo para identificação de focos ativos é específico para cada tipo de sensor. No geral, eles usam diferentes canais dos sensores dos satélites, entre a luz visível e o infravermelho, e podem ter comportamentos distintos dependendo se a imagem foi gerada à noite ou de dia. Podem também usar limiares dinâmicos, de acordo com a região do planeta, que são calculados com base em uma espécie de média das temperaturas nas regiões próximas ao longo dos dias. Além disso, é comum a aplicação de máscaras para eliminar regiões submersas, costeiras, deserticas e que estavam nubladas na hora da passagem. [P2. Visão geral dos algoritmos para detecção]

Para o caso dos satélites TERRA e AQUA, que têm o sensor MODIS, o instituto mantinha seu próprio método de detecção, que produzia dados de maior confiabilidade

(INPE, 2023). A partir de 2017 o INPE migrou toda a base de dados para o "Collection 6", algoritmo aplicado pela *National Aeronautics and Space Administration* (NASA), marcando o início da chamada Base 2 de queimadas. Anteriormente, a NASA empregava o Collection 5, que gerava falsos positivos em clareiras florestais e falsos negativos para grandes queimadas obscurecidas por fumaça densa. (SCHROEDER et al., 2008). [P3. Algoritmo empregado pelo INPE]

[P4. diferença entre área queimada e focos detectados]

[P5. Falar sobre a flag risco de fogo e uma ideia de como é calculada]

2.4 Carregando os dados para análise

Para toda a análise dos dados foi utilizado algumas das ferramentas que integram o ecossistema Python para Data Science: NumPy, Pandas, Matplotlib. Além disso, algumas bibliotecas específicas para análise de dados geográficos: GeoPandas, Pysal, Xarray. Para a construção das visualizações e organização do código se optou pelo uso do Jupyter Notebook devido a reprodutibilidade das execuções. Todos os artefatos gerados podem ser obtidos em <<https://github.com/josebraz/INPE-Queimadas>> sobre a licença MIT. [P0. Visão das ferramentas utilizadas]

Foi carregado 300 arquivos das queimadas em formato CSV e depois concatenados em uma única estrutura de DataFrame. Houve a necessidade de converter o timezone das datas (que eram em UTC) para o timezone de Brasília, a fim de gerar gráficos de mais fácil entendimento para brasileiros. As colunas de texto foram convertidas para categorias, espécies de enumerações no Pandas, reduzindo o espaço ocupado de memória, uma vez que muitos valores acabavam se repetindo na mesma coluna. [P1. pré processamentos dos dados]

Também lançou-se mão dos dados públicos territoriais do Instituto Brasileiro de Geografia e Estatística (IBGE) com o intuito de gerar gráficos delimitados em municípios, unidades federativas e biomas. Todos os arquivos baixados estão em formato Shapefile, responsável por armazenar dados vetoriais geográficos. A biblioteca GeoPandas consegue ler e gerar gráficos a partir do Shapefile. [P2. Uso dos dados de shapes do território nacional do IBGE]

2.5 Os Satélites

O INPE atualmente processa dados de vários satélites com características distintas entre si. Estão presentes nos dados desde satélites geoestacionários, como o GOES-12, que está a 29.400 km de distância da superfície (SETZER; YOSHIDA, 2004), até satélites com órbitas polares, entre 700 a 900 km de altura. [P0. Visao geral dos satelites]

Abaixo segue um resumo dos satélites usados pelos INPE desde o início da série histórica até final de 2022 (Embrapa Territorial, 2023). Os que estão em funcionamento pleno atualmente são: NOAA-20, NOAA-19, NOAA-18, GOES-16, Suomi NPP, AQUA, TERRA, MSG-03, METOP-B e METOP-C. [P1. Falar sobre os principais]

Tabela 2.2 – Características dos satélites usados pelo INPE

Nome	Sensor	Resolução esp.	Órbita	Lançamento	Passagem
METOP-C	AVHRR-3	1100m	Polar	2018	21h
NOAA-20	VIIRS	500m	Polar	2017	2h / 14h
METOP-B	AVHRR-3	1100m	Polar	2012	21h
Suomi NPP	VIIRS	500m	Polar	2011	2h / 14h
NOAA-19	AVHRR-3	1100m	Polar	2009	2h / 14h
NOAA-18	AVHRR-3	1100m	Polar	2005	Variadas
AQUA	MODIS	1000m	Polar	2002	2h / 14h
NOAA-17	AVHRR-3	1100m	Polar	2002	21h
NOAA-16	AVHRR-3	1100m	Polar	2000	Variadas
TERRA	MODIS	1000m	Polar	1999	11h / 23h
NOAA-15	AVHRR-3	1100m	Polar	1998	5h / 17h
TRMM	VIRS	2000m	Polar	1997	Variadas
NOAA-14	AVHRR	1100m	Polar	1994	21h
NOAA-12	AVHRR	1100m	Polar	1991	2h / 15h
GOES-16	ABI	2000m	Geoest.	2016	Não se aplica
MSG-03	SEVIRI	3000m	Geoest.	2012	Não se aplica
GOES-13	GOES I-M	4000m	Geoest.	2006	Não se aplica
MSG-02	SEVIRI	3000m	Geoest.	2005	Não se aplica
GOES-12	GOES I-M	4000m	Geoest.	2001	Não se aplica
GOES-10	GOES I-M	4000m	Geoest.	1997	Não se aplica
GOES-08	GOES I-M	4000m	Geoest.	1994	Não se aplica

Fonte: O Autor com base em Embrapa Territorial (2023)

Cada satélite pode ter um sensor imageador, que gera imagens, com características distintas. Neles são captados imagens não só no comprimento de onda da luz visível (de 400nm a 700nm), mas também no infravermelho (de 780nm a 1mm). Suas medições são divididas em canais, que variam em resolução espacial e espectral (intervalo de comprimento de onda). Geralmente o primeiro canal é dedicado à luz visível, entre o laranja

e o vermelho, e com a maior resolução espacial possível para o sensor. Os outros canais utilizam diferentes intervalos do infravermelho e luz visível. [P2. visão geral dos sensores e porque geram dados diferentes]

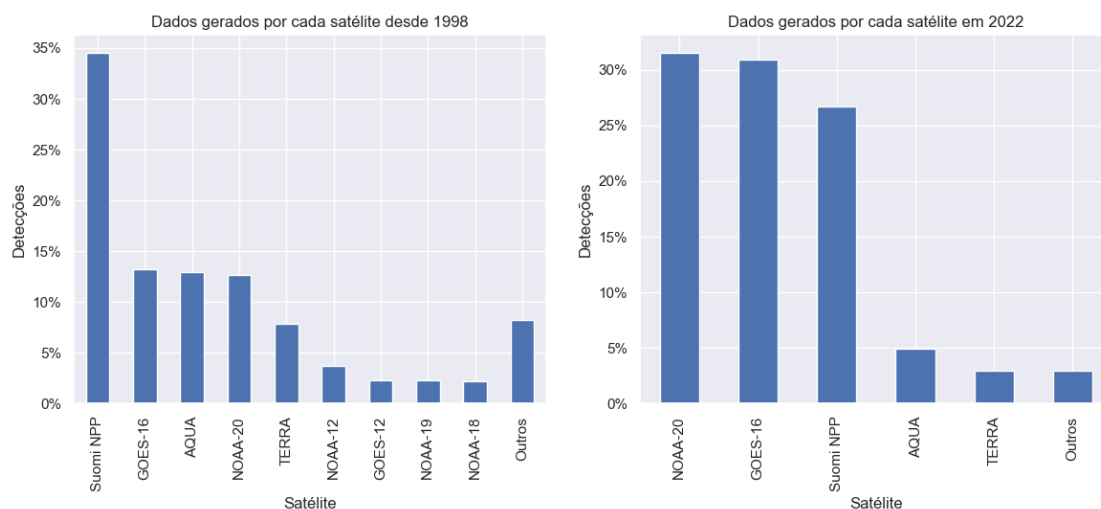
Com todas essas diferenças entre os satélites foi necessário estabelecer um satélite base, que ficou conhecido como Satélite de Referência. Ele é usado para estabelecer uma série temporal e permitir análise de tendência durante vários anos dos focos detectados para diferentes regiões. Este balizador precisa cobrir a área do país de forma satisfatória, ou seja, sua órbita não deve distorcer os dados no geral. Além disso, resoluções dos sensores muito baixas, maiores de 1 km, tornam a análise dos focos pouca precisa. [P3. Satellite de referencia]

De 01/junho/1998 a 03/julho/2002 o satélite de referência utilizado foi o NOAA-12 com passagem no final da tarde. Depois desse período passou-se a utilizar o AQUA com passagem à tarde (chamado nos dados de AQUA_M-T). O satélite AQUA já ultrapassou a data prevista de encerrar o funcionamento em muitos anos e será descontinuado em breve, quando isso acontecer o satélite Suomi NPP será o novo balizador (INPE, 2023). [P3. Satellite de referencia]

Nos dados disponibilizados pelo INPE são usadas nomenclaturas especiais para alguns satélites. Para os satélites AQUA (AQUA_M-T e AQUA_M-M) e TERRA (TERRA_M-T e TERRA_M-M), a primeira letra M representa o sensor MODIS e a última letra indica em que período do dia foi a passagem do satélite, sendo M para Manhã e T para Tarde. Outros satélites como Suomi NPP, NOAA-19, NOAA-18, NOAA-16, NOAA-15 e NOAA-12 também podem apresentar a última letra do nome sendo D para Diurno. A partir do entendimento dessa regra de nomenclatura é possível simplificar os dados relacionados aos satélites de forma a facilitar a análise, caindo de 32 para 22 valores possíveis de satélites. [P4. Explicar equivalencias entre os satelites]

Voltando para os dados, é possível perceber a partir da Figura 2.1 cinco satélites que mais identificaram focos de queimada, são eles: Suomi NPP, GOES-16, AQUA, NOAA-16 e TERRA. Todos eles estão ativos atualmente e com quantidades de coleta significativas em 2022. AQUA e TERRA são os mais antigos e possuem um sensor obsoleto. O GOES-16 é um satélite geoestacionário, conhecido por gerar dados de forma mais frequente, geralmente detecta apenas queimadas maiores devido a sua posição distante da Terra. Suomi NPP e NOAA-20 possuem um sensor que detecta 10 vezes mais focos que sensor MODIS. Por estar em atividade a mais tempo, o Suomi NPP gerou mais dados que o NOAA-20. [P5. Mostrar gráficos da quantidade de coletas por satélites]

Figura 2.1 – Relação do montante dos dados por satélite



Fonte: O Autor

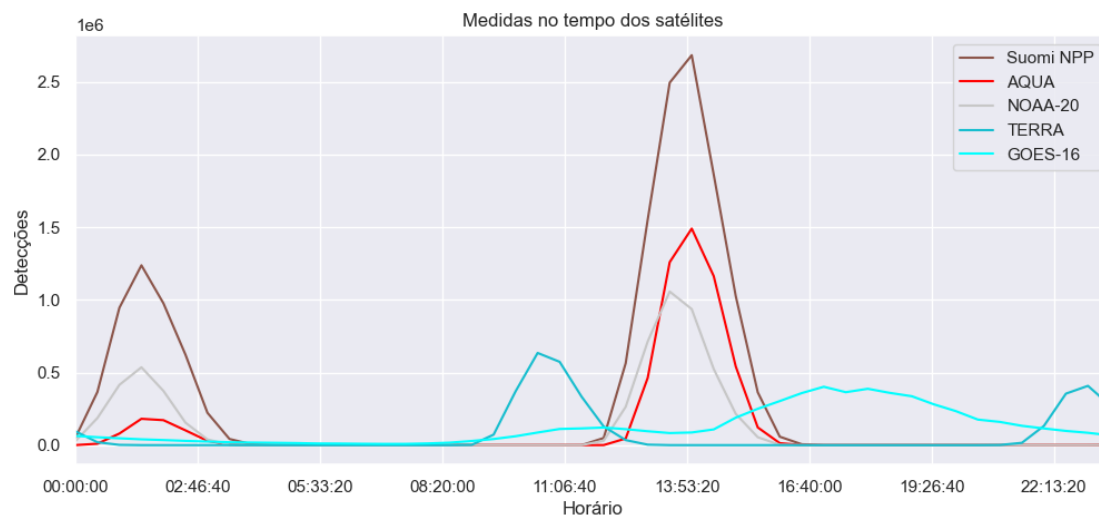
Saber em quais momentos os satélites passam também é importante para a análise. Os satélites polares passam duas vezes por dia no Brasil, variando o local exato da passagem de acordo com as características de sua órbita. Pela Figura 2.2 é possível observar esse comportamento empiricamente, em que as 5 primeiras linhas, que representam dados gerados por satélites polares, apresentam dois picos durante um período de 24 horas. Já para o caso dos geoestacionários (GOES-16), que ficam fixos em relação a uma posição na Terra, não se observou o mesmo padrão. [P6. Mostrar gráficos que indicam as horas das coletas]

Para finalizar, podemos observar na Figura 2.3 a passagem dos satélites no dia 10 de agosto de 2022, gerada a partir de dados do <<https://celestrak.org/>>. Os satélites podem levar vários dias para passar no mesmo local devido a características de sua órbita e podemos ter uma estimativa razoavelmente precisa de sua trajetória ao longo do tempo. Também é possível ver o satélite geoestacionário GOES-16, no norte do Perú, representado com um ponto no gráfico.

2.6 O que os dados gritam

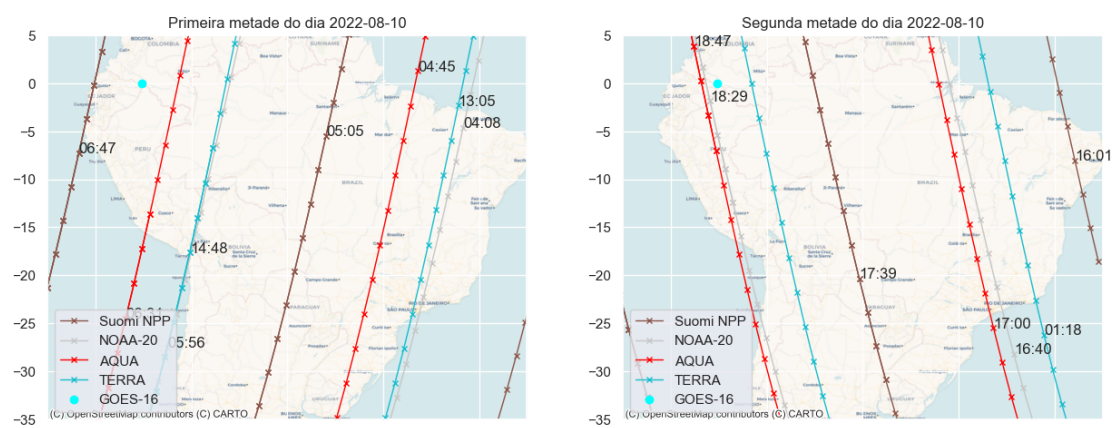
Esse capítulo vai dar uma primeira olhada nos dados, tentando extrair as informações que mais chamam a atenção. É importante tomar um certo cuidado nessa análise geral, principalmente na escolha dos satélites. Se forem usados todos os satélites, prova-

Figura 2.2 – Amostragem por tempo de cada satélite



Fonte: O Autor, agrupando os dados de queimadas

Figura 2.3 – Órbita dos satélites no dia 10 de agosto de 2022



Fonte: O Autor

velmente será contado um mesmo foco várias vezes, ou ainda, para um mesmo satélite polar, contar o mesmo foco nas duas passagens dele (de dia e a noite). Para evitar esse problema, será sempre usado o satélite AQUA com passagem a tarde (AQUA_M-T) que é o satélite de referência.

P1. Fazer análise preliminar dos dados gerando alguns gráficos

P2. Gráficos geral do brasil com os focos de queimadas totais (REY; ARRIBAS-BEL; WOLF, 2020)

3 TRABALHOS RELACIONADOS

P1. falar sobre como os trabalhos relacionados podem ajudar a entender os dados

4 APROFUNDANDO A ANÁLISE DOS DADOS

aqui a gente mostra que é válido usar esses dados para análises aprofundadas

4.1 Densidade e Centrografia

P1. Verificar densidade e centrografia: tendências, dispersão, extensão

4.2 Validade dos dados

Precisamos verificar que os dados seguem algum padrão para ser possível usar eles para tomadas de decisões (garantir que não é aleatório) (REY; ARRIBAS-BEL; WOLF, 2020, Point Pattern Analysis)

4.3 Padronizando os dados por satélite

P1. Verificar relação entre dados dos diferentes satélites (se possível) e talvez restringir a análise apenas ao satélite de referencia se for identificado que são basicamente equivalentes

REFERÊNCIAS

Divisão de Geração de Imagens. **Geração de imagens**. 2023. Disponível em: <<https://www.dgi.inpe.br/>>. Acesso em: 16 de fev. 2023.

Embrapa Territorial. **Satélites de Monitoramento**. 2023. Disponível em: <<https://www.embrapa.br/satelites-de-monitoramento/>>. Acesso em: 28 de jan. 2023.

INPE. **Programa Queimadas Perguntas frequentes**. 2023. Disponível em: <<http://www.inpe.br/queimadas/portal/informacoes/perguntas-frequentes>>. Acesso em: 23 de jan. 2023.

REY, S. J.; ARRIBAS-BEL, D.; WOLF, L. J. **Geographic Data Science with Python**. [S.l.: s.n.], 2020. Disponível em: <<http://geographicdata.science/book/notebooks/>>. Acesso em: 5 de jan. 2023.

SCHROEDER, W. et al. Validation of goes and modis active fire detection products using aster and etm+ data. **Remote Sensing of Environment**, v. 112, n. 5, p. 2711–2726, 2008. ISSN 0034-4257. Earth Observations for Terrestrial Biodiversity and Ecosystems Special Issue. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425708000102>>.

SETZER, A.; MORELLI, F.; SOUZA, J. C. O banco de dados de queimadas do inpe. **Biodiversidade Brasileira-BioBrasil**, n. 1, p. 239–239, 2019.

SETZER, A.; YOSHIDA, M. C. **Deteccção de Queimadas nas Imagens do Satélite Geoestacionário GOES-12**. 2004. Disponível em: <https://queimadas.dgi.inpe.br/~rqueimadas/documentos/relat_goes12_3_4.htm>. Acesso em: 1 de fev. 2023.