



UNIVERSIDADE FEDERAL DE MINAS GERAIS

INSTITUTO DE CIÊNCIAS EXATAS

DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO

BANCO DE DADOS GEOGRÁFICO

TRABALHO FINAL

Precipitação na região Sudeste nos anos de
2010, 2013, 2016, 2019 e 2022.

Aluno: Leonardo Miguel de Sousa Jorge

Matrícula: 2023655620

Professor: Clodoveu Davis

BELO HORIZONTE

Dezembro/2023

1 INTRODUÇÃO

O presente estudo consiste no trabalho final da disciplina Banco de Dados Geográfico (DCC831), ministrada pelo Prof. Dr. Clodoveu Augusto Davis Junior, ofertada pelo Departamento de Ciência da Computação da Universidade Federal de Minas Gerais (DCC UFMG). O objetivo do trabalho é realizar o pré-processamento e análise de dados meteorológicos fornecidos pelo Instituto Nacional de Meteorologia (INMET)¹.

Devido ao elevado volume de dados, medidos e armazenados diariamente por estações meteorológicas distribuídas em todo o território nacional, a diretiva é escolher um grupo de estações correlacionadas geograficamente, bem como um dos diversos tipos de dados coletados, para a realização das análises.

2 PROPOSTA DE ANÁLISE

Como proposta de análise, pretende-se avaliar os dados de precipitação total nos estados da região Sudeste do Brasil, no período entre 2010 e 2023, espaçados de 3 em 3 anos. Em suma serão utilizados os dados pluviométricos dos anos 2010, 2013, 2016, 2019 e 2022.

O objetivo da análise é obter um panorama geral do regime de chuvas nos estados de Minas Gerais, São Paulo, Rio de Janeiro e Espírito Santo durante o período, através de comparações e correlações entre os dados dos quatro estados, bem como avaliação do cenário em diferentes anos para um mesmo estado.

3 TRABALHOS RELACIONADOS

Em um estudo realizado em 2007, Minuzzi et al. [1] analisou dados diários de 203 estações pluviométricas na região Sudeste do Brasil, visando caracterizar e analisar tendências climáticas no período chuvoso. Os resultados revelaram que, desde a década de 1970, o período chuvoso na região Sudeste iniciava-se mais cedo, entre 13 e 22 de setembro, com maior duração especialmente na região litorânea de São Paulo. Em contraste, nas regiões Norte e Vale do Jequitinhonha, em Minas Gerais, o período chuvoso inicia, em média, entre 23 de outubro e 2 de novembro, e com menor duração. Além disso, maiores quantidades de chuva foram observadas apenas no sudeste de Minas Gerais e no estado de São Paulo.

Uma análise da variabilidade da precipitação no sudeste durante o período chuvoso foi conduzida por Malvestio (2013) [2] utilizando dados de 177 postos pluviométricos com séries de 35 anos (1976-2010). Os resultados destacaram o verão como o período mais propenso a deslizamentos de terra e inundações na região, especialmente nas áreas metropolitanas de Belo Horizonte, São Paulo e Rio de Janeiro. A pesquisa visava relacionar essa variabilidade de precipitação com a incidência de desastres naturais.

Já em estudo realizado por Rao et al. [3] em 2016, foram analisadas as características das chuvas no Brasil no período de 1979 a 2011. No Sudeste do país, os pesquisadores observaram uma redução significativa das chuvas tanto na estação chuvosa quanto na seca. Essa diminuição foi confirmada pela análise das vazões fluviais nos estados de São Paulo e Minas Gerais. Os autores sugeriram que a queda nas chuvas na região sudeste do Brasil está associada ao declínio

¹[urlhttps://portal.inmet.gov.br/](https://portal.inmet.gov.br/)

dos sistemas produtores de chuva do sul, atribuído ao aumento da pressão média do nível do mar na região sul.

Em 2018, Marini [4] destacou eventos climáticos extremos no Brasil, incluindo secas históricas na Amazônia e no Nordeste, e inundações na região Amazônica, Sul e Sudeste. Esses eventos estão relacionados à variabilidade climática natural, podendo ser agravados pelo aumento da temperatura devido às mudanças climáticas. Foi destacado também que a seca intensa entre 2013 e 2015 na Região Sudeste foi a mais severa em 52 anos, causada por uma persistente alta pressão atmosférica que inibiu a formação de chuvas. A combinação de baixa precipitação, aumento na demanda por água e falta de planejamento público contribuíram para a crise hídrica mencionada.

Por fim, o estudo conduzido por Alves et al. (2020) [5] buscou avaliar as projeções de modelos relacionadas à mudança na variabilidade das chuvas em várias escalas temporais nas sub-regiões do Brasil. Os resultados indicaram uma concordância significativa entre os modelos, apontando para aumentos na variabilidade em todas as escalas temporais e em todas as sub-regiões analisadas. Em resumo, as estimativas sugeriram que o Brasil enfrentará maior variabilidade de precipitação no futuro, caracterizada por períodos secos mais intensos e frequentes, assim como períodos chuvosos mais abundantes em escalas diárias, semanais e mensais, descobertas que podem ter implicações relevantes para a adaptação às mudanças climáticas, especialmente nos setores agrícola e de recursos hídricos no país.

4 METODOLOGIA

4.1 DADOS UTILIZADOS

Dados Meteorológicos: O INMET disponibiliza os dados meteorológicos a partir do ano 2000, em formato csv, permitindo o download de duas formas: 1) Dados individuais por estação, automática ou convencional, em datas configuráveis para períodos limitados a 6 meses, disponível em <<https://portal.inmet.gov.br/dadoshistoricos>>. 2) Dados históricos das estações automáticas por ano, disponível em <<https://tempo.inmet.gov.br/TabelaEstacoes>>. Este trabalho possui enfoque nos dados referentes aos anos de 2010, 2013, 2016, 2019 e 2022, com comentado previamente, e por este motivo optou-se por trabalhar com os dados históricos das estações automáticas, para realizar o download dos csvs de interesse de maneira direta.

Informações das Estações Meteorológicas: As informações referentes às estações são fornecidas separadamente, também em formato em csv, disponíveis para download em <<https://portal.inmet.gov.br/paginas/catalogoaut>> para estações automáticas e em <<https://portal.inmet.gov.br/paginas/catalogoman>> para estações manuais. Há informações como localização geográfica, código de identificação, cidade, estado, data de início da operação e se a estação ainda está em estado operante.

Dados Complementares: Além dos dados meteorológicos e das informações das estações que registram esses dados, também foram utilizadas bases complementares. No caso optou-se por trabalhar com o shapefile dos estados brasileiros, com enfoque nos quatro estados da região Sudeste, disponibilizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) em <<https://www.ibge.gov.br/geociencias/downloads-geociencias.html>>.

4.2 PRÉ-PROCESSAMENTO E ANÁLISE EXPLORATÓRIA

Os dados históricos das estações automáticas são particionados por ano, porém contém todas as medidas de todas estações que operaram naquele ano no território nacional, e por isso inicialmente decidiu-se trabalhar com todo o dataset para a realização da análise exploratória, concatenando-se os dados dos cinco anos de interesse. Foi decidido que os dados das estações dentro da região de interesse seriam filtrados com base em dados espaciais, através de consultas SQL, já na etapa do banco de dados (próxima subseção do relatório).

O dataset original ocupa 1,85 GB de memória, em um total de 22.396.224 linhas x 19 colunas (data, hora e 17 tipos de dados meteorológicos medidos pelas estações). Para a organização dos dados foi necessário realizar algumas etapas de pré-processamento utilizando o pandas², uma biblioteca em python que fornece ferramentas para análise de dados. As etapas foram 1) substituição dos valores inválidos (-9999) por NaN, 2) extração do ID e do nome da estação a partir dos metadados, 3) concatenação das colunas de data e hora para uma única coluna de timestamp, e por fim, 4) unificação do dataset. Após o pré-processamento o dataset passou a ocupar 1,58 GB de memória, com 20 atributos, listados abaixo:

- | | |
|-----------------------------------|------------------------------|
| • Id da estação | • Temperatura máxima |
| • Nome da estação | • Temperatura mínima |
| • Datetime | • Temperatura orvalho máxima |
| • Precipitação total | • Temperatura orvalho mínima |
| • Pressão atmosférica | • Umidade relativa máxima |
| • Pressão atmosférica máxima | • Umidade relativa mínima |
| • Pressão atmosférica mínima | • Umidade relativa do ar |
| • Radiação global | • Vento, direção |
| • Temperatura do ar | • Vento, rajada máxima |
| • Temperatura do ponto de orvalho | • Vento, velocidade |

Também foram necessárias algumas etapas de pré-processamento no dataset contendo informações das estações. Essas etapas consistiram em converter os dados de latitude, longitude e altitude para numéricos, e conversão da data de início de operação da estação para datetime.

Além disso, deve-se mencionar que ao longo dos anos presentes na análise, haviam 615 estações meteorológicas com registros no banco de dados. Porém, ao realizar o cruzamento do dataset de dados meteorológicos com o dataset das estações, através do ID, percebeu-se que no dataset das estações haviam apenas 566 estações registradas, contendo informações como localização geográfica. Dessa forma, na ausência desse tipo de informação para as 49 estações restantes,

²[urlhttps://pandas.pydata.org/](https://pandas.pydata.org/)

essas foram desconsideradas na análise. Após o descarte, o banco foi reduzido de 1,58 GB para 1,54 GB.

4.3 CONSTRUÇÃO DO BANCO DE DADOS GEOGRÁFICO

Pós a análise exploratória e pré-processamento dos dados, estes foram incorporados à um banco de dados do postgresSQL³. Como a tabela das estações possui dados de latitude e longitude, estes foram convertidos para geometria dentro do postgresSQL, permitindo visualização e consultas espaciais. Além dos dados meteorológicos e das estações, através do QGIS⁴, foi incorporada ao banco uma base contendo o shapefile do território nacional, particionada por estados.

Para conectar a tabela das estações com a tabela dos dados dentro do ambiente do postgresSQL, foi utilizada a coluna com o ID (chave primária na tabela das estações e chave estrangeira na tabela de dados). E para conectar a tabela das estações com a tabela contendo as geometrias dos estados, foi utilizada a coluna com sigla dos estados (chave primária na tabela dos estados e chave estrangeira na tabela das estações). Dessa forma as três tabelas foram conectadas, permitindo consultas cruzadas e extração de informações através de correlações espaciais.

Pós a construção do banco de dados, foram elaboradas queries com visualização espacial, por exemplo na Figura 1, onde A) indica a filtragem da geometria dos estados de interesse apenas, e B) indica a filtragem das estações localizadas dentro dessa região.



Figura 1: Exemplos resultantes de consultas de dados geográficos no postgresSQL. A) Filtragem da geometria dos estados da região Sudeste. B) Filtragem das estações localizadas dentro dos estados da região Sudeste.

4.4 PRODUÇÃO DAS VISUALIZAÇÕES

Para realização das análises em si, foram confeccionados dois tipos de gráfico: mapas e gráficos de barra. Os mapas foram feitos utilizando o QGIS e os gráficos de barra utilizando a biblioteca matplotlib⁵, um ferramenta de visualização gráfica do python.

³<<https://www.postgresql.org/>>

⁴<https://qgis.org/pt_BR/site/>

⁵url<https://matplotlib.org/>

Nas etapas referentes à confecção dos mapas, foi realizada a conexão do banco de dados do PostgreSQL com o QGIS através do postGIS⁶, que permite a realização de consultas SQL e visualização dos dados do banco dentro do próprio QGIS. Desta forma, as queries definidas previamente no PostgreSQL foram executadas dentro do ambiente QGIS, permitindo a conversão dos resultados em camadas do software, e posterior salvamento dessas camadas em formato shapefile.

A Figura 2 mostra a visualização parcial do projeto no QGIS, indicando a conexão do banco de dados do PostgreSQL ao software, as camadas resultantes das queries (filtradas por ano), e as estações contidas na região Sudeste.

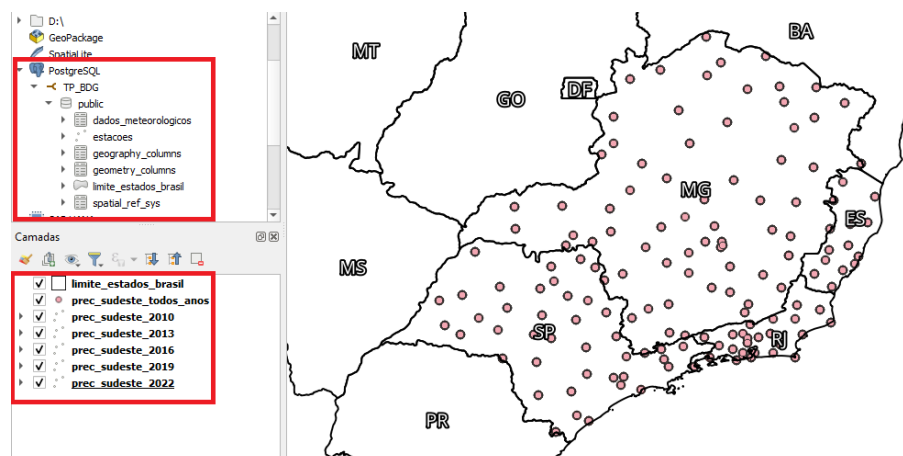


Figura 2: Visualização do projeto no ambiente GIS, indicando a conexão do banco de dados ao QGIS, as camadas resultantes das queries (convertidas para shapefile), e as estações contidas na região Sudeste, após a filtragem.

5 RESULTADOS E DISCUSSÕES

Ao todos foram produzidos quatro gráficos, sendo dois gráficos de barra e dois mapas. O gráfico de barras 1 (Figura 3) representa a precipitação acumulada por estado da região Sudeste por ano. Através dessa visualização é evidente que a precipitação anual no estado de Minas Gerais é mais elevada em comparação aos demais estados, cerca de 2 vezes maior que os estados de São Paulo e Rio de Janeiro, e 4 vezes maior que o estado do Espírito Santo.

Porém, ao realizar uma análise espacial e não apenas da quantidade de chuva, verifica-se que a extensão territorial do estado de Minas Gerais é a maior da região, principalmente se comparado aos estados do Rio de Janeiro e Espírito Santo, logo é esperado que a precipitação total seja maior nesse estado. O mapa 1 (Figura 4) consiste na precipitação acumulada por estação e não por estado, o que fornece informações complementares, visto que o número de estações é mais representativo da área de cada estado. Além disso, permite não só comparar de forma mais robusta a precipitação por estado, mas também avaliar diferentes porções do mesmo estado.

Através do mapa 1 verifica-se que, na verdade, as estações com maiores registros de pluviosidade estão localizadas mais próximo ao litoral, principalmente no estado do Rio de Janeiro. Outra observação é que a região norte do estado de Minas Gerais registra os menores valores de precipitação.

⁶[urlhttps://postgis.net/docs/manual-3.4/en/](https://postgis.net/docs/manual-3.4/en/)

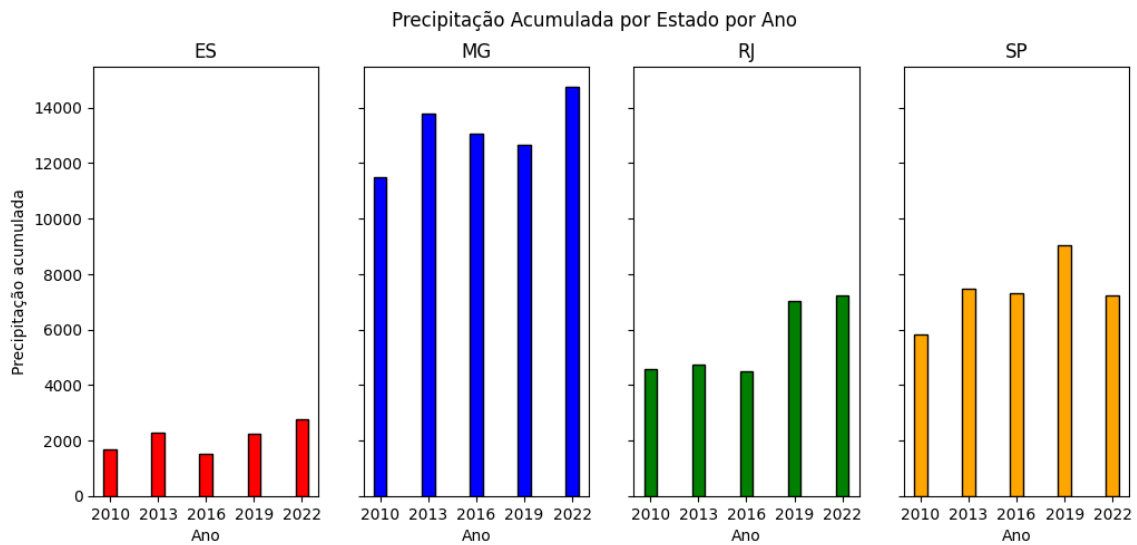


Figura 3: Gráfico de barras 1 - precipitação acumulada por estado da região Sudeste, por ano.

Precipitação Acumulada por Estação Pluviométrica - Região Sudeste - 2010 a 2022

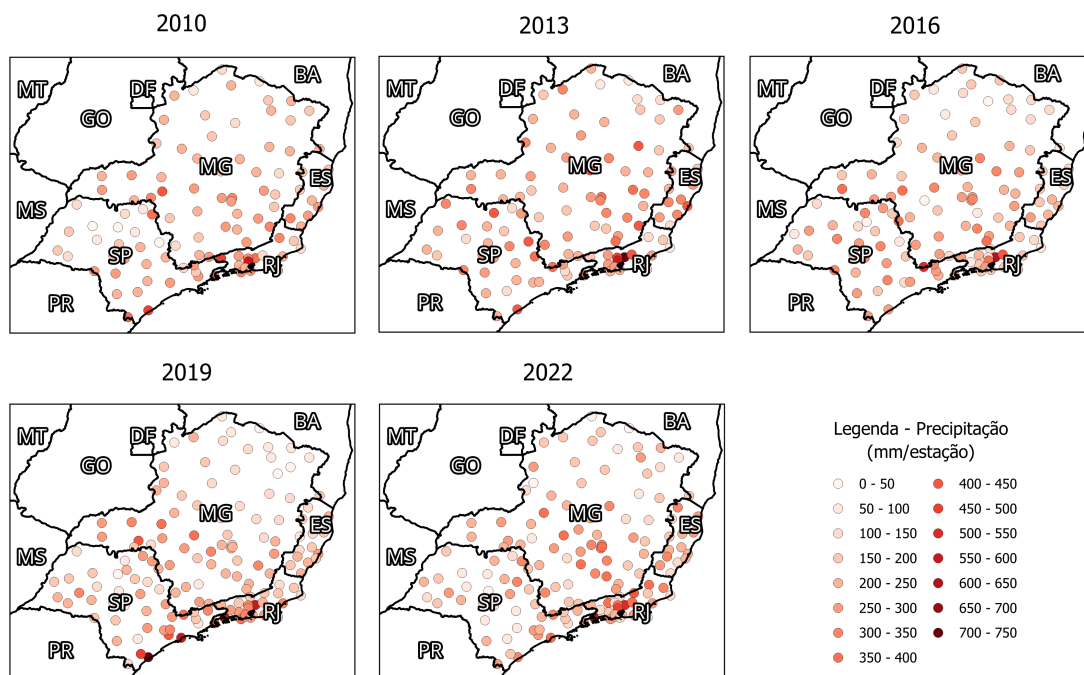


Figura 4: Mapa 1 - precipitação acumulada por estação da região Sudeste, por ano.

Além da verificação da precipitação acumulada, outra análise complementar e que torna a comparação mais fidedigna é a precipitação média. Através do gráfico de barras 2 (Figura 5), verifica-se que a diferença na precipitação média por estado é baixa, e que na verdade as médias para o estado do Rio de Janeiro são ligeiramente maiores. O mesmo é observado no mapa 2 (Figura 6), que indica médias elevadas nas estações localizadas no Sul e Leste do estado.

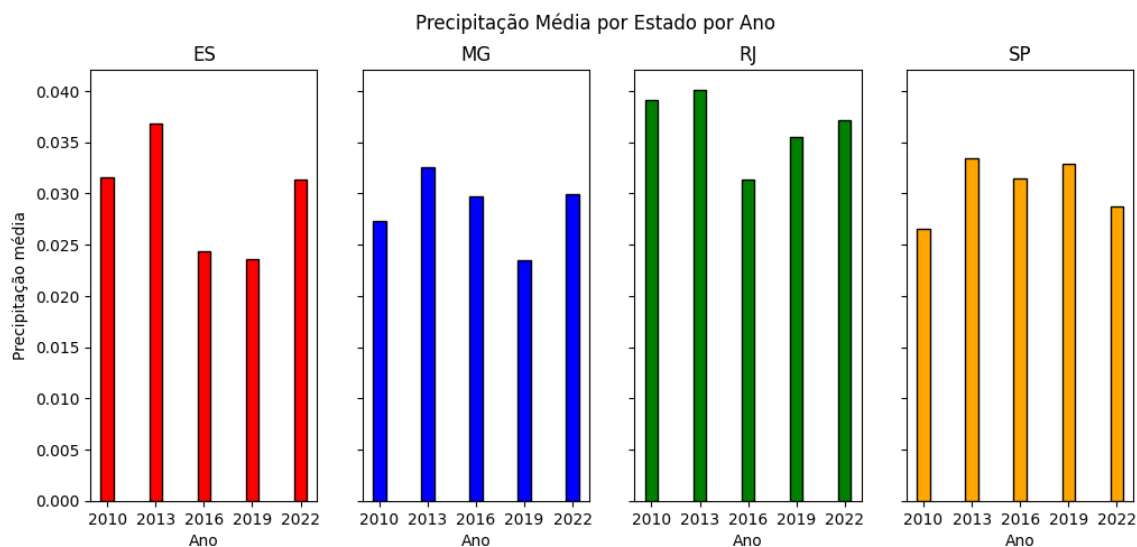


Figura 5: Gráfico de barras 2 - precipitação média por estado da região Sudeste, por ano.

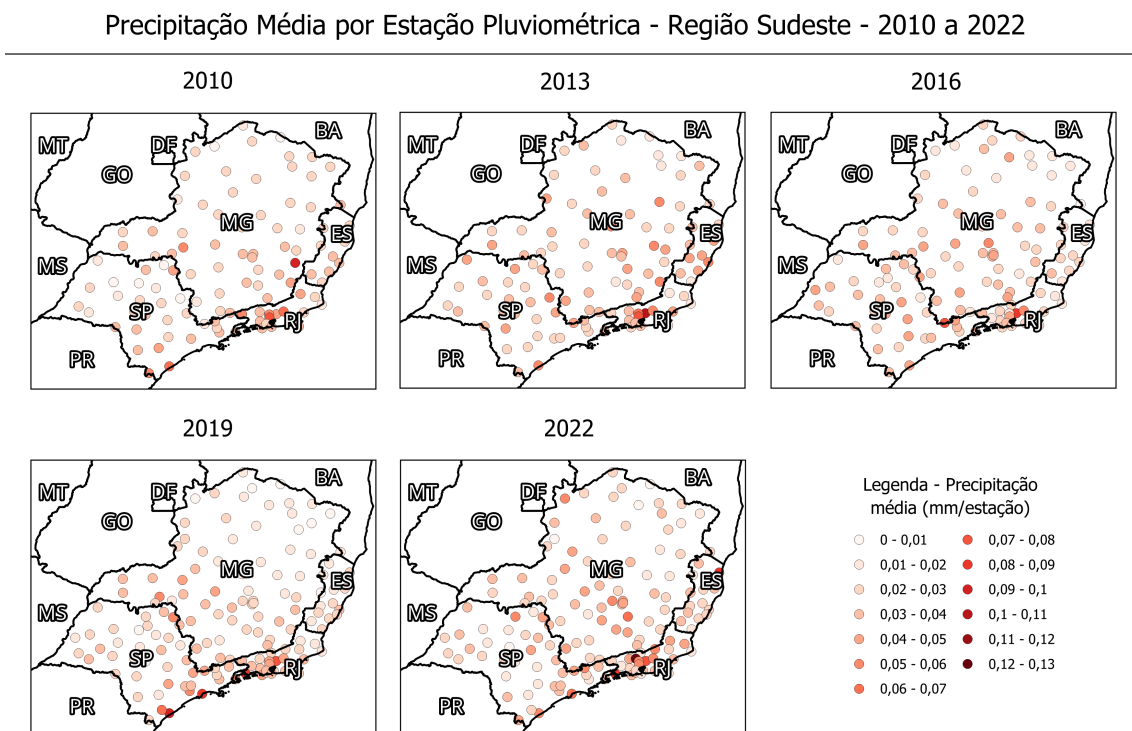


Figura 6: Mapa 2 - precipitação média por estação da região Sudeste, por ano.

Por fim, comparando-se a precipitação da série temporal de forma individual, por estado, verifica-se que não há diferenças significativas. Ou seja, dentro do período não houve aumento ou queda na precipitação ao longo dos anos.

6 CONCLUSÃO

Diante das investigações realizadas ficou evidente a importância de ferramentas como o postGIS (extensão do postgresQL) e o QGIS para lidar com dados espaciais, em complemento às ferramentas tradicionais de análises de dados, como as bibliotecas pandas e matplotlib em python e Sistemas de Gerenciamento de Bancos de Dados (SGBDs).

No caso do presente estudo, envolvendo dados meteorológicos, verificou-se que uma análise quantitativa apenas pode não ser representativa do panorama geral. Por outro lado a análise espacial pode trazer maior significância à investigação das hipóteses levantadas, levando a diferentes interpretações.

Os resultados indicaram que apesar da maior quantidade total de chuvas registradas no estado de Minas Gerais, tanto a análise espacial (por estação) quanto a verificação dos valores médios mostraram que a precipitação no Rio de Janeiro é maior, apesar de não serem diferenças bruscas. Além disso, verificou-se que numa comparação ano a ano para cada estado individualmente, nenhum deles apresentou diferenças significativas, ou seja, dentro do período não houve tendência de aumento ou queda na precipitação.

7 REFERÊNCIAS BIBLIOGRÁFICAS

[1] MINUZZI, Rosandro Boligon et al. Climatologia do comportamento do período chuvoso da região sudeste do Brasil. **Revista Brasileira de Meteorologia**, v. 22, p. 338-344, 2007.

[2] MALVESTIO, Leônidas Mantovani. Variabilidade da precipitação pluviométrica da região sudeste do Brasil no período chuvoso e suas consequências ambientais. **Universidade Estadual de Campinas**, 2013. Tese de Doutorado.

[3] RAO, V. Brahmananda et al. An update on the rainfall characteristics of Brazil: seasonal variations and trends in 1979–2011. **International Journal of Climatology**, v. 36, n. 1, p. 291-302, 2016.

[4] MARINI, Sâmia Saldanha et al. Crise hídrica e o regime de chuvas na região sudeste: a seca de 2013-2015. **Universidade Federal Fluminense**, 2018. Monografia.

[5] ALVES, Lincoln M. et al. Assessment of rainfall variability and future change in Brazil across multiple timescales. **International Journal of Climatology**, v. 41, p. E1875-E1888, 2021.