

Análise Exploratória de Chamados para Monitoramento de Áreas de Risco

Tassiane Barros e Priscilla Amarante

Abstract— Neste relatório, será apresentada a análise realizada nos dados de chamados para monitoramento das áreas de risco na cidade de Recife.

1 Introdução

Áreas de risco são regiões onde não é recomendada a construção de casas ou instalações, pois são muito propícias a desastres naturais, como desabamentos e inundações. Essas regiões vêm crescendo, principalmente devido à própria ação humana. No Brasil, vêm sendo realizados vários projetos no sentido de reestruturação de algumas áreas e conscientização da população [9].

As principais áreas de risco são aquelas sob encostas de morros inclinados ou à beira de rios [8]. Existem vários serviços de assistência à população em casos de emergência. A principal instituição responsável pelo monitoramento das áreas de risco é a Defesa Civil.

A Secretaria Executiva de Defesa Civil de Recife (SEDEC - Recife) desenvolve ações preventivas com o objetivo de evitar ou minimizar acidentes e situações de calamidades, a partir do monitoramento permanente das áreas de risco, tanto nos morros quanto na planície, que estejam sujeitos a alagamentos, deslizamentos e enxurradas; e em edificações que apresentem irregularidades [10].

O início do período de chuvas aumenta os riscos de inundações, desabamentos de casas e deslizamentos de terra. Existem sinais claros para identificar e evitar as grandes tragédias. Postes, cercas e árvores que começam a inclinar mostram que o terreno está se movendo. Trincas nas paredes ou no chão e degraus, junto aos barrancos, são outro sinal de alerta.

Sendo assim, a análise realizada neste trabalho busca avaliar os dados dos chamados de monitoramento das áreas de risco na cidade de Recife. A proposta visa verificar quais são os bairros/regiões com o maior número de chamados e o ano/mês de maior incidência.

A implementação desse projeto está disponível no GitHub [1].

2 Dados e Tarefas

Utilizamos duas bases de dados, uma dos chamados de monitoramento de áreas de risco de Recife [4] e outro do clima da cidade [2], que mostra uma média dos valores de cada dia do ano relativo a 20 anos. Os dados incluem registros entre 2012 e 2021, e o *dataset* resultante da união dos chamados com o clima possui 26 colunas e 61033 linhas.

Notamos que o *dataset* possui diversos atributos que podem ser utilizados. No entanto, certas informações são conhecidas por serem mais relevantes para realizar o monitoramento das áreas de risco.

2.1 Pré-Processamento dos Dados

O pré-processamento dos dados consiste em avaliar e fazer as mudanças ou remoções necessárias nos valores e colunas do

dataset, tendo em vista facilitar a análise exploratória dos dados [7]. Para este trabalho, serão considerados alguns fatores-chaves para o monitoramento das áreas de risco.

A montagem do *dataframe* foi feita a partir da remoção, inserção e alteração de tipo dos dados. Realizamos uma cópia do *dataframe* original, para que nela sejam feitas as alterações, sem que se perca os dados originais.

Algumas colunas, devido a menor relevância, serão removidas para fins de análise neste estudo. Por isso, excluímos colunas que não seriam utilizadas nesta análise ou com pelo menos 30% de dados ausentes, também removemos os dados duplicados e colunas com o mesmo objetivo.

Como os dados categóricos são valores que indicam uma característica específica do dado, transformamos alguns dos dados de *object* para *category* para auxiliar na análise. Para a coluna que representava a data, transformamos para *datetime*. Fizemos também conversão de variáveis binárias de '0' para 'Não' e '1' para 'Sim' a fim de obter uma correlação entre ela e outras variáveis.

Identificamos que as colunas de *latitude* e *longitude* possuíam muitos dados nulos, para preencher esse dados utilizamos a mediana da *latitude* e *longitude* do bairro, referente ao chamado. Notamos que em muitos dos chamados a região não era informada, mas na maioria das vezes tinha o registro do bairro, então optamos por incluir a região tendo como base o bairro do chamado.

Construímos dois *datasets* um resultado do agrupamento dos dados por região e outro por bairro, para facilitar a visualização dos dados no *Tableau Public* [6].

2.2 Questões de Pesquisa (QPs)

Utilizamos visualização para obter respostas para as seguintes questões:

QP1. Qual bairro tem mais chamados?

QP2. Quais os meses do ano que tem maior precipitação?

QP3. Qual a relação da quantidade de chamados com precipitação no mês?

QP4. Quais os atributos têm maior correlação com as precipitações?

QP5. Qual região tem mais chamados?

QP6. Quais anos tiveram maior número de chamados?

QP7. Qual ano teve maior número de vítimas fatais? Nesse ano teve mais chamados?

3 Visualizações e Casos de Uso

Tendo como base os dados dos chamados criamos o *Dashboard* de Monitoramento de Áreas de Risco [3] com mapas, utilizando o *Tableau Public*, para melhorar a visualização dos resultados. Já para os dados que representavam series temporais optamos por utilizar o *Plotly* [5].

- Maria Tassiane Barros de Lima. E-mail: mtbl@cin.ufpe.br.
- Priscilla Amarante de Lima. E-mail: pal4@cin.ufpe.br.

3.1 Casos de Uso

Agora, apresentaremos os resultados de nossas análises:

QP1. Qual bairro tem mais chamados?

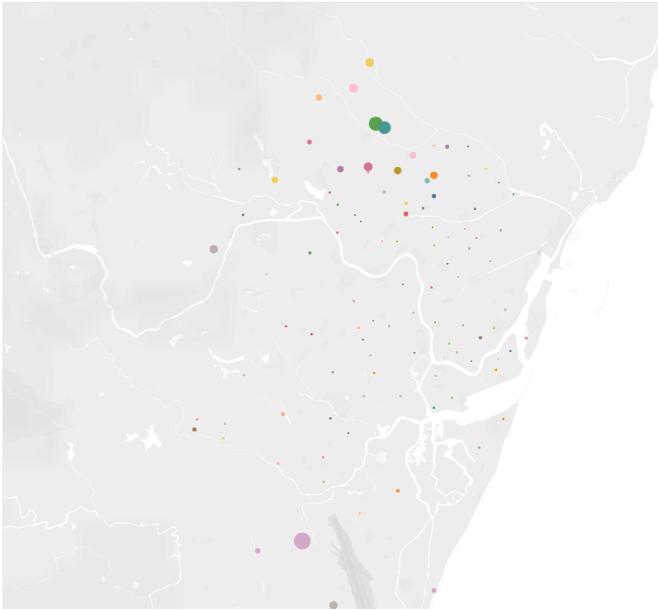


Fig. 1. Chamados por Bairro [3]

Pelo mapa da Fig. 1 pode-se afirmar que o bairro do Ibura (na cor violeta) e de Nova Descoberta (na cor verde) são os que mais possuem chamados de monitoramento. Podemos constatar que os mesmos possuem muitos morros e depressões.

QP2. Quais os meses do ano que tem maior precipitação?

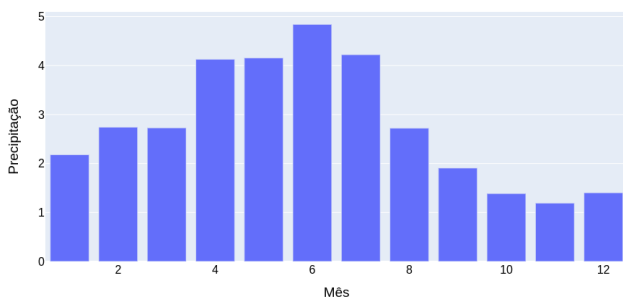


Fig. 2. Precipitação por mês

Pela Fig. 2 pode ser observado que nos meses de Abril, Maio, Junho e Julho existe maior precipitação.

RQ3. Qual a relação da quantidade de chamados com precipitação no mês?

Através da Fig. 3 pode ser notado que quando existe pico de precipitação também há picos de chamados, indicando que dias chuvosos existem mais desastres naturais.

RQ4. Quais os atributos têm maior correlação com as precipitações?

Pela matriz de correlação da Fig. 4 pode ser notado que existe uma correlação negativa da precipitação com a temperatura e as horas de sol.

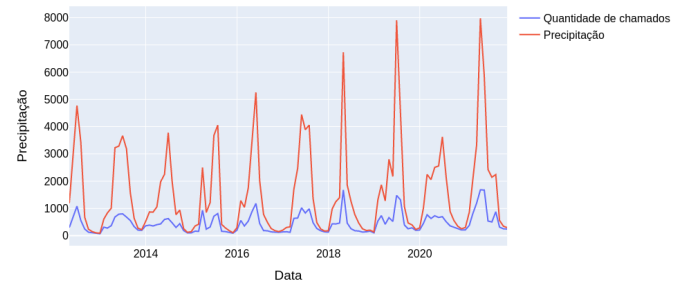


Fig. 3. Quantidade de Chamados e Precipitação

solicitacao_plantao	1	0.012	0.0058	-0.1	-0.1	0.01	0.016	0.016	0.016	0.016
latitude	0.012	1	0.41	0.023	0.023	0.037	-0.012	-0.01	-0.012	-0.01
longitude	0.0058	0.41	1	0.012	0.01	-0.017	-0.0051	-0.008	-0.0051	-0.008
solicitacao_vitimas	-0.1	0.023	0.012	1	0.99	0.092	-0.084	-0.093	-0.084	-0.093
solicitacao_vitimas_fatais	-0.1	0.023	0.01	0.99	1	0.091	-0.083	-0.092	-0.083	-0.092
precipitacao	0.01	0.037	-0.017	0.092	0.091	1	-0.52	-0.38	-0.52	-0.38
horas_de_sol	0.016	-0.012	-0.0051	-0.084	-0.083	-0.52	1	0.77	1	0.77
temperatura_max	0.016	-0.01	-0.008	-0.093	-0.092	-0.38	0.77	1	0.77	1
horas_de_sol_norm	0.016	-0.012	-0.0051	-0.084	-0.083	-0.52	1	0.77	1	0.77
temperatura_max_norm	0.016	-0.01	-0.008	-0.093	-0.092	-0.38	0.77	1	0.77	1

Fig. 4. Matriz de Correlação

RQ5. Qual região tem mais chamados?

Segundo a lei municipal nº 16.293, de 22 de janeiro de 1997, o Recife se divide em seis regiões político-administrativas (RPAs): Centro, Norte, Noroeste, Oeste, Sudoeste e Sul. As RPAs, por sua vez, dividem-se em microrregiões, que agrupam os bairros.

Conforme ilustrado no mapa da Fig. 5, a região noroeste (na cor vermelha) do estado é a que possui maior número de chamados.

A Região Político Administrativa III (RPA3) é composta pelos bairros: Afritos; Alto do Mandu; Alto José Bonifácio; Alto José do Pinho; Apipucos; Brejo da Guabiraba; Brejo de Beberibe; Casa Amarela; Casa Forte; Córrego do Jenipapo; Derby; Dois Irmãos; Espinheiro; Graças; Guabiraba; Jaqueira; Macaxeira; Monteiro; Nova Descoberta; Parnamirim; Passarinho; Pau-Ferro; Poço da Panela, Santana; Sítio dos Pintos; Tamarineira; Mangabeira; Morro da Conceição; Vasco da Gama.

RQ6. Quais anos tiveram maior número de chamados?

Podemos observar na Fig. 6 que os anos de 2019, 2020 e 2021 tiveram grande destaque. O ano de 2021 registrou a maior número de chamados.

RQ7. Qual ano teve maior número de vítimas fatais? Nesse ano teve mais chamados?

Podemos observar na Fig. 6 que os anos de 2019, 2020 e 2021 tiveram um maior número de chamados e consequentemente um maior número de vítimas fatais, como mostra a Fig. 7.

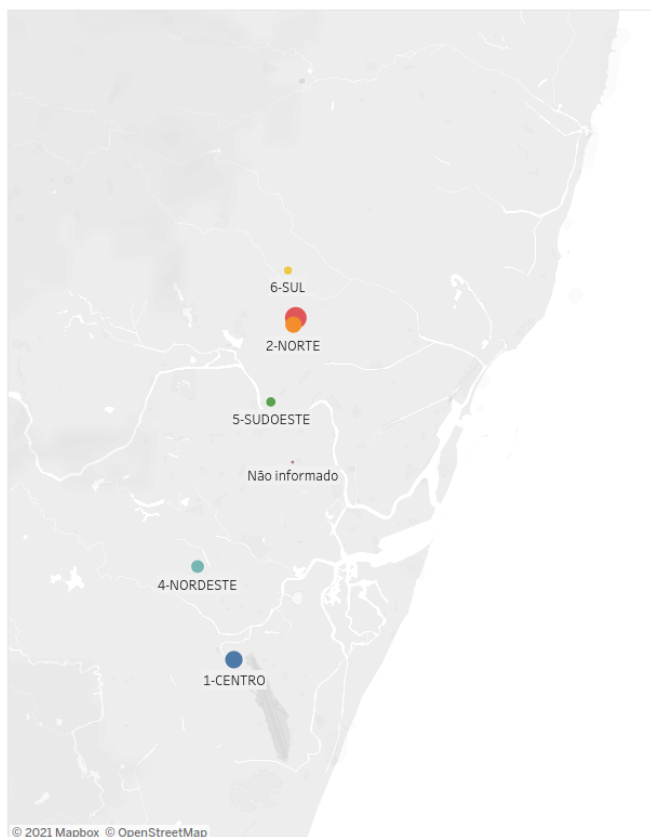


Fig. 5. Chamados por Região [3]

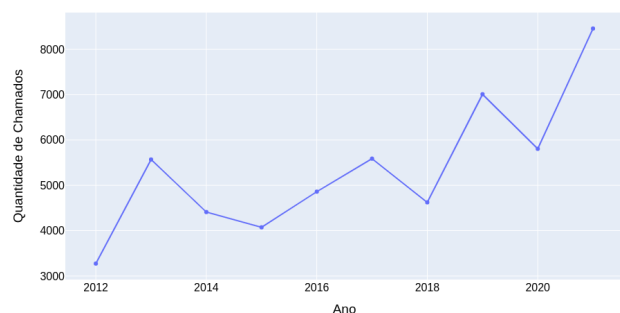


Fig. 6. Chamados por Ano

4 Conclusões e Discussões

O presente projeto teve como propósito analisar os chamados com relação aos monitoramentos das áreas de risco na cidade do Recife. Para realizar a análise foi preciso acessar os dados abertos disponibilizados no site da prefeitura de Recife na área da defesa civil. Os dados foram tratados através da exclusão de algumas colunas, para realizar uma melhor aproveitamento das suas informações.

Vale destacar a importância do monitoramento das áreas de risco na prevenção e mitigação de riscos e desastres, pois é possível realizar as vistorias (área plana x morro), a permanência do monitoramento das localidades (áreas com risco médio, alto, muito alto e alagáveis) através do mapeamento das regiões e/ou bairros da cidade.

Podemos concluir, pelas análises realizadas que durante os meses de abril a julho é a temporada de inverno, e durante

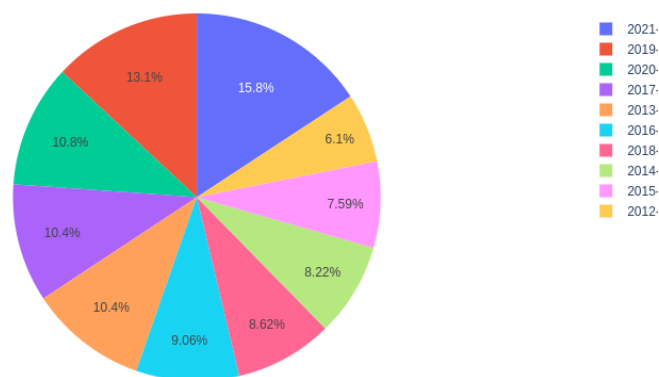


Fig. 7. Distribuição da quantidade de vítimas fatais por ano

esse período aumenta o volume de água na cidade. O trabalho da defesa civil em alertar as famílias que residem em bairros e regiões de área de morro, depressões, erosão das encostas se intensifica. Durante as fortes chuvas faz com que haja o aumento dos chamados e a equipe da defesa civil atua no monitoramento das áreas de risco a fim de evitar acidentes as famílias que residem nessas localidades.

References

- [1] Análise exploratória de chamados para monitoramento de Áreas de risco. <https://github.com/tassianebarros/monitoramento-chamados-area-risco>.
- [2] Clima de recife. <https://pt.climate-data.org/americas-do-sul/brasil/pernambuco/recife-5069/>. acessado em Novembro 2021.
- [3] Dashboard de monitoramento de Áreas de risco. <https://public.tableau.com/app/profile/tassiane2319/viz/MonitoramentoA>.
- [4] Monitoramento das Áreas de riscos. <http://dados.recife.pe.gov.br/dataset/monitoramento-das-areas-de-riscos>. acessado em Novembro 2021.
- [5] Plotly. <https://plotly.com/python/>.
- [6] Tableau public. <https://public.tableau.com/pt-br/s/>.
- [7] G. E. d. A. P. Batista et al. *Pré-processamento de dados em aprendizagem de máquina supervisionado*. PhD thesis, Universidade de São Paulo, 2003.
- [8] D. d. A. Gonzalez et al. A análise da percepção de risco e vulnerabilidade à partir da vivência dos alunos do ensino médio de nova friburgo rj. 2016.
- [9] E. H. d. NEVES and E. KOESTER. Áreas de risco geomorfológico influenciadas pelo rio piratini, na área urbana de pedro osório-rs.
- [10] A. Patrícia Nunes Bandeira. Parâmetros técnicos para gerenciamento de áreas de riscos de escorregamentos de encostas na região metropolitana do recife. 2010.