

Análise do número de atendimentos da UPA Caxangá:

O que o serviço público pode revelar sobre a situação socioeconômica e cultural de uma região?

Analysis of the number of attendances at UPA Caxangá: What public service attendances reveal about the socioeconomic and cultural conditions of a region?

A.R.P. Carolino e M.C.A.P de Melo

Universidade Federal de Pernambuco

Recife, Pernambuco, Brasil

arpc2; mcapm / cin.ufpe.br

Resumo - A relevância do serviço público de saúde brasileiro é inquestionável para o funcionamento do país. Entender o número de atendimentos diários em uma Unidade de Pronto Atendimento (UPA) pode fornecer insights valiosos sobre a dinâmica cultural e socioeconômica de uma localidade. Este estudo tem como objetivo analisar os números de atendimentos na UPA da Caxangá e compreender como eventos externos podem influenciar a demanda por serviços de saúde. Foi realizado o teste de Shapiro-Wilk para verificar a normalidade dos dados, constatando-se que a amostra não segue uma distribuição normal. Adicionalmente, a análise do gráfico Boxplot revelou que os dados não apresentam uma caixa simétrica. Logo, foi investigado e confirmado que a principal causa da não normalidade dos dados é a presença de numerosos outliers, que são resultantes de fatores externos à comunidade, como endemias, festividades ou escassez de recursos médicos.

Palavras-chave: saúde pública, análise de dados, UPA, normalidade, outliers, eventos externos, socioeconomia, cultura.

Abstract - The significance of the Brazilian public health service is undeniable for the functioning of the country. Understanding the daily number of attendances at an Emergency Care Unit (UPA) can provide valuable insights into the cultural and socioeconomic dynamics of a locality. This study aims to analyze the attendance numbers at the UPA of Caxangá and understand how external events can impact the demand for emergency services. The Shapiro-Wilk test was conducted to verify the normality of the data, revealing that the sample does not follow a

normal distribution. Additionally, the analysis of the Box Plot indicated that the data do not present a symmetrical box. It was investigated and confirmed that the main cause of the non-normality of the data is the presence of numerous outliers, resulting from external factors in the community, such as epidemics, festivities, or resource shortages.

Keywords: public health, data analysis, UPA, normality, outliers, external events, socioeconomic, culture.

I. INTRODUÇÃO

A criação de Unidades de Pronto Atendimento foi um case de sucesso do Sistema Único de Saúde no combate à mortalidade e na melhoria da qualidade de vida da população. A presença de tal instituição em um bairro pode revelar inúmeros aspectos daquela região, sendo este o objetivo de estudo deste artigo. A UPA da Caxangá fica localizada no coração do bairro da Caxangá, sendo responsável por centenas de atendimentos, funcionando 24 horas todos os dias, sendo uma UPA tipo III, com capacidade de atender até 450 pacientes por dia.

De tal modo, seria prudente afirmar a hipótese nula ($H_0 \leq 450$) seria que a UPA atende até 450 pacientes por dia, enquanto a hipótese alternativa ($H_A > 450$) seria de que a UPA pode atender mais que 450 pacientes por dia.

Com o intuito de compreender melhor essa questão, foram realizados testes de normalidade, juntamente com uma análise descritiva dos dados e a visualização de

gráficos. A coleta de dados foi conduzida ao longo de 100 dias, registrando a quantidade de pacientes atendidos em um intervalo de 4 horas diárias. Para fins de estudo e análise, esses números foram ajustados proporcionalmente para representar um período de 24 horas, permitindo a entrega de resultados equivalentes à quantidade de dias no período observado.

Os dados coletados passaram por diversos testes estatísticos para verificar a normalidade e outras características descritivas. A análise envolveu a aplicação do teste de Shapiro-Wilk para normalidade, a geração de gráficos Quantil-Quantil (QQ Plot) para inspeção visual, e a remoção de outliers para avaliar o impacto nos resultados. Além disso, estatísticas descritivas como média, mediana, moda, desvio padrão, variância e coeficiente de variação foram calculadas para proporcionar uma visão abrangente do comportamento dos dados.

A análise detalhada dos dados da UPA da Caxangá permitiu identificar padrões e discrepâncias na distribuição dos atendimentos diários, contribuindo para uma melhor compreensão da demanda por serviços de saúde na região. Através da aplicação rigorosa de métodos estatísticos, pode-se entender melhor o funcionamento da dinâmica da região, observando que fatores externos afetam a demanda de atendimento, gerando informações valiosas para a preparação da instituição aos pontos de pico, bem como o descaso governamental no fornecimento de insumos de saúde foram também evidenciados, permitindo uma visão mais abrangente e precisa do funcionamento da UPA.

II. METODOLOGIA

A. Ferramentas tecnológicas

A análise exploratória de dados foi realizada com a utilização do Excel, bem como a linguagem Python. Ademais, as bibliotecas Pandas, Numpy, Seaborn, Matplotlib, Statistics, Scipy e Statsmodels foram utilizadas.

B. Ferramentas estatísticas

No viés estatístico, foi realizado o cálculo das medidas de tendência central e de dispersão da base de dados, além disso, para facilitar na visualização foram criados gráficos (Histograma, Boxplot e Gráfico de Linhas). Outrossim, foi realizado o teste de normalidade de dados utilizando o teste de Shapiro-Wilk.

C. Junção das ferramentas

Após o recebimento da base de dados da vazão do número de pacientes da UPA Caxangá, essa base foi importada para o Excel. Utilizando esta ferramenta, a base de dados foi tratada, uma vez que os dados recebidos eram referentes ao atendimento realizado durante 4 horas na UPA. Em seguida, para fins de estudo, a coluna de número de pacientes foi modelada para a quantidade de pacientes atendidos em 24 horas (um dia completo).

Assim, foi utilizado o Google Colab para carregar o arquivo Excel para ser processado na linguagem Python, utilizando a biblioteca Pandas. Com os data frames carregados, as informações das colunas da planilha foram transformadas em listas da linguagem de programação. Utilizando a biblioteca statistics foi possível fazer o cálculo das medidas de tendência central e de dispersão, além do uso do scipy para calcular a curtose e o numpy para os cálculos dos quartis. Os gráficos presentes neste artigo foram criados utilizando a biblioteca Matplotlib e Seaborn. Os testes de hipótese foram criados utilizando um módulo da biblioteca Statsmodels.

III. RESULTADOS DA ANÁLISE

D. Análise descritiva dos dados

Os dados coletados da amostra são quantitativos contínuos, uma vez que são valores numéricos referentes à vazão de pacientes atendidos na UPA da Caxangá. Abaixo, na tabela 1, estão as medidas de tendência central e de dispersão, utilizadas para estudar a tendência central dos dados.

Table 1: Medidas Centrais e de Dispersão

Medidas	Valores
Média	289.14
Mediana	258.0
Moda	294.0
Desvio Padrão	179.27
Variância	32136.51
Coeficiente de variação	0.62
1º Quartil	162.0
2º Quartil	258.0
3º Quartil	390.0
4º Quartil	936.0
Curtose	1.09

A média indica que, em geral, a UPA atende aproximadamente 289 pacientes por dia. A mediana, que representa o valor central quando os dados são ordenados, revela que metade dos dias tem um número de atendimentos menor ou igual a 258 pacientes, enquanto a outra metade tem um número de atendimentos maior ou igual a esse valor. A moda, ou seja, o valor mais comum de atendimentos diários, é de 294 pacientes por dia.

O desvio padrão, que mede a dispersão dos dados em relação à média, indica que os números de atendimentos

diários variam significativamente em torno da média, refletindo uma grande variabilidade nos atendimentos diários. Essa variabilidade é reforçada pela variância, que é alta, mostrando uma grande dispersão dos dados. O coeficiente de variação, que compara a dispersão relativa dos dados, sugere uma dispersão moderada em relação à média.

Os quartis mostram a dispersão dos dados em quatro partes, indicando que os valores de atendimentos variam amplamente, com 25% dos dias tendo menos de 162 atendimentos e 25% dos dias tendo mais de 390 atendimentos. A curtose, que mede a "cauda" da distribuição dos dados, sugere que a distribuição tem caudas mais pesadas e é mais pontuda que uma distribuição normal, indicando a presença de outliers.

Dessa mesma forma, a mediana sendo menor que a média sugere que a distribuição é assimétrica para a direita). O alto desvio padrão e variância indicam que há grande variabilidade no número de atendimentos diários, com muitos dias tendo atendimentos que são muito diferentes da média. Os quartis mostram uma ampla dispersão dos dados, e a curtose indica a presença de outliers, que podem ser eventos esporádicos que resultam em um número de atendimentos muito alto ou muito baixo.

Com base nesses dados, foram criados os gráficos de eficiência da UPA (Figura 1), o Boxplot (Figura 2) e Histograma (Figura 3) da base de dados.

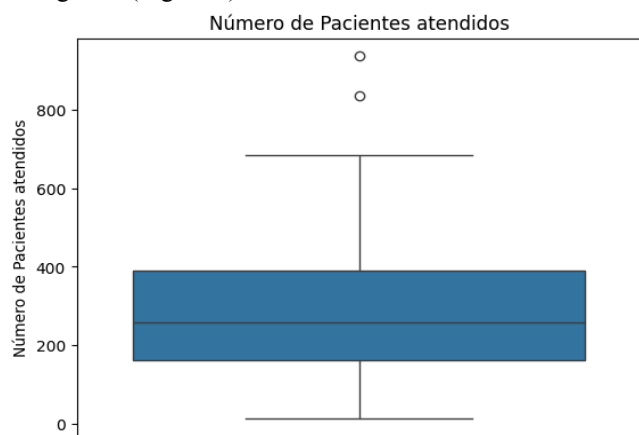


Figura 1 - Boxplot

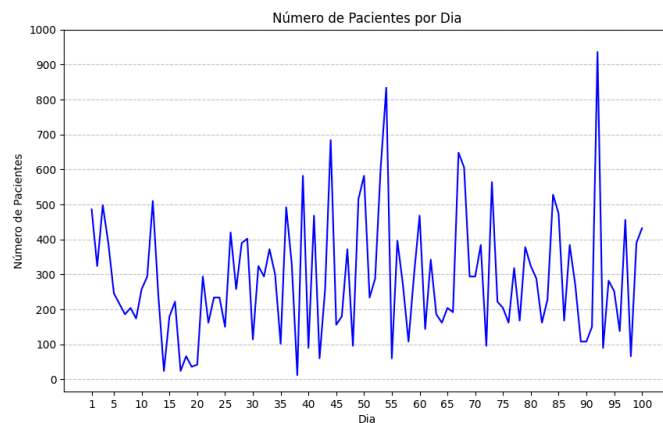


Figura 2 - Gráfico de linhas

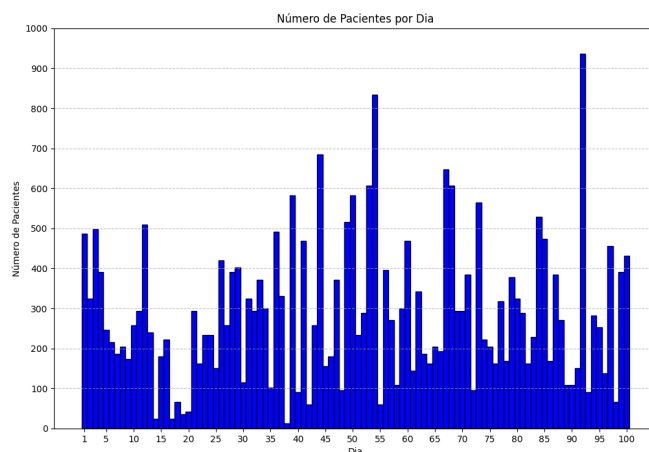


Figura 3 - Histograma

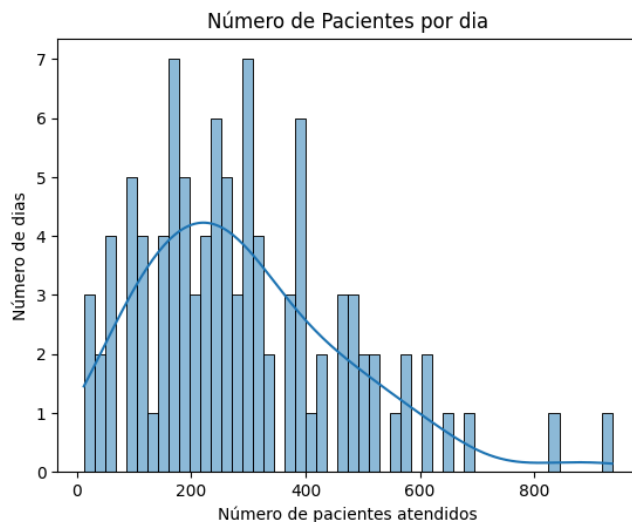


Figura 4 - Histograma com linha

Na Fig. 1 nota-se que a alta dispersão entre os quartis e a presença de outliers indicam uma grande variabilidade no número de atendimentos diários. Além disso, a assimetria da caixa do boxplot sugere uma distribuição positivamente

enviesada, indicando que há mais dias com atendimentos acima da mediana do que abaixo.

A Fig.2 evidencia que existe uma discrepância entre a quantidade de número de pacientes atendidos durante os 100 dias analisados. Em conjunto, a Fig.3 também auxilia no entendimento dessa situação. Os dados não estão distribuídos de forma simétrica na média, indicando uma não normalidade na distribuição deles.

Ademais, é interessante notar nas Figs. 2 e 3 que ocorrem muitos outliers, representando uma grande ou pequena quantidade de pacientes atendidos, provavelmente relacionados a fatores externos, como condições climáticas, surtos de doenças, eventos locais ou ausência de materiais de saúde.

Dessa forma, pode-se inferir que os dados de atendimentos diários na UPA não seguem uma distribuição normal pela análise visual do boxplot e do histograma. A assimetria da caixa no boxplot, a presença de outliers, e a ausência de uma curva em forma de sino na Fig.4 são indicativos claros de que a distribuição dos dados não é normal.

E. Teste de normalidade

Para determinar se os dados seguem uma distribuição normal, além das inferências a partir da visualização dos grafos, foi empregado o teste de normalidade de Shapiro-Wilk nos dados de pacientes atendidos por dia na UPA da Caxangá. Os resultados do teste de Shapiro-Wilk para os dados analisados foram os seguintes:

- **Estatística do Teste de Shapiro-Wilk:** 0.9441
- **p-valor:** 0.0003

O p-valor muito menor que o nível de significância comum (0.05) indica que rejeitamos a hipótese nula. Isso sugere que os dados não são normalmente distribuídos.

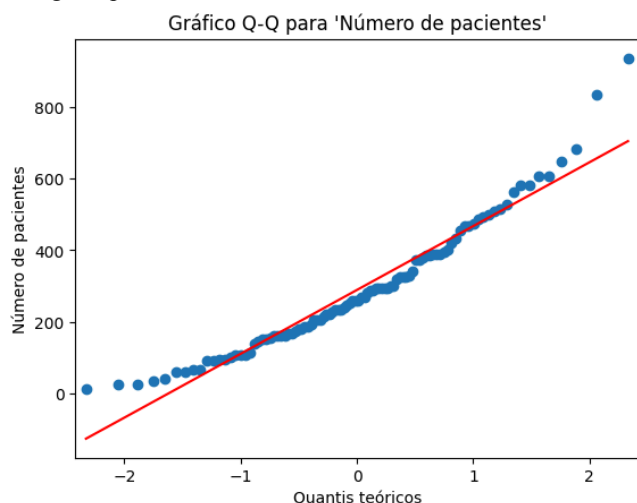


Figura 5 - Gráfico Q-Q

Para complementar o teste de Shapiro-Wilk, utilizamos um QQ Plot (Gráfico Quantil-Quantil) para uma inspeção visual da normalidade dos dados. O QQ Plot compara os

quantis dos dados com os quantis de uma distribuição normal teórica. Se os dados fossem normalmente distribuídos, os pontos do QQ Plot seguiriam aproximadamente uma linha reta.

No entanto, o QQ Plot gerado para os dados de pacientes atendidos por dia mostrou uma clara divergência dos pontos em relação à linha reta, reforçando a o teste de Shapiro-Wilk de que os dados não seguem uma distribuição normal.

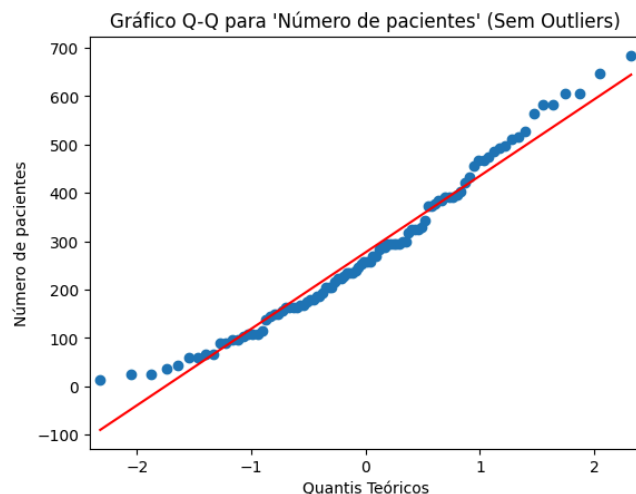


Figura 6 - Gráfico Q-Q sem outliers

Após a remoção dos outliers, os dados foram novamente submetidos ao teste de Shapiro-Wilk para verificar se houve alguma mudança significativa. Os limites para identificar os outliers foram calculados usando o intervalo interquartil (IQR). Os resultados do teste mostram variações dependendo dos níveis de significância:

- **Nível de significância de 0.05:** Os dados não seguem uma distribuição normal.
- **Nível de significância de 0.1:** Os dados não seguem uma distribuição normal.
- **Nível de significância de 0.01:** Não podemos rejeitar a hipótese nula de normalidade.

Os resultados do teste de Shapiro-Wilk para os dados sem os outliers foram:

- **Estatística do Teste de Shapiro-Wilk:** 0.9666
- **p-valor:** 0.0136

O p-valor (0.0136) ainda é menor que o nível de significância comum de 0.05, o que indica que rejeitamos a hipótese nula de normalidade. Portanto, mesmo após a remoção dos outliers, os dados não seguem uma distribuição normal. Além disso, o QQ Plot para os dados sem os outliers ainda mostra divergências dos pontos em relação à linha reta, reforçando que a distribuição dos dados não é normal mesmo após a remoção dos outliers.

Por fim, os resultados dos testes estatísticos e as inspeções visuais indicam que os dados de pacientes atendidos por dia na UPA da Caxangá não seguem uma distribuição normal, especialmente nos níveis de significância de 0.05 e 0.1. Mesmo após a remoção dos outliers, a distribuição dos dados permanece não normal, conforme evidenciado pelo teste de Shapiro-Wilk e pelo QQ Plot.

IV. DISCUSSÃO

A análise detalhada dos dados da UPA da Caxangá revelou padrões e discrepâncias significativas na distribuição dos atendimentos diários, proporcionando uma visão aprofundada sobre a demanda por serviços de saúde na região. A aplicação de métodos estatísticos, como o teste de Shapiro-Wilk, e a análise visual por meio de gráficos permitiram identificar a não normalidade dos dados e a presença de outliers.

Eventos externos, como jogos de futebol, shows, surtos de doenças e a ausência de material médico devido à negligência governamental, desempenham um papel crucial na variação da demanda por serviços de saúde na UPA da Caxangá. Esses eventos podem criar picos de demanda ou, inversamente, períodos de baixa procura, ambos refletidos como outliers nos dados analisados.

Por exemplo, jogos de futebol e shows são eventos que podem aumentar a demanda por atendimento médico. Esses eventos atraem grandes multidões, aumentando a probabilidade de incidentes que requerem atenção médica, como lesões, intoxicações ou desidratação. Isso resulta em outliers de alta demanda nos dias em que tais eventos ocorrem. Da mesma forma, surtos de doenças sazonais ou epidemias, como gripes ou viroses, também contribuem significativamente para a variação na demanda. Durante esses períodos, observa-se um aumento no número de atendimentos, criando outliers que refletem picos de demanda devido à alta incidência de casos que necessitam de cuidados médicos.

Além disso, a ausência de materiais médicos e insumos essenciais, frequentemente decorrente de uma gestão inadequada ou negligência governamental, pode afetar negativamente a capacidade de atendimento da UPA. Em dias com falta de materiais, a UPA pode atender menos pacientes do que o habitual, resultando em outliers de baixa demanda. Este cenário destaca a vulnerabilidade do sistema de saúde referente a questões de gestão e fornecimento de recursos, refletindo a dinâmica socioeconômica da região e a eficácia do serviço público em atender às necessidades da população.

A análise também destaca a real necessidade de atendimento por parte da população. A presença de outliers de alta demanda pode indicar momentos de necessidade urgente, onde a população depende fortemente dos serviços públicos de saúde. Esses picos podem ser associados a condições precárias de vida ou à falta de acesso a cuidados preventivos, refletindo questões mais amplas de desigualdade socioeconômica e negligência governamental.

V. CONCLUSÕES

Portanto, a análise do serviço público de saúde pode revelar a dinâmica socioeconômica e cultural de uma região. A disponibilidade e qualidade dos serviços de saúde pública, como hospitais, clínicas e centros de atendimento, refletem diretamente o nível de desenvolvimento e os investimentos governamentais na área. Ademais, no viés cultural, entender se os fatores externos atingem a necessidade da população de atendimento médico.

Dessa forma, é possível entender que os fatores externos que geram outliers na demanda de atendimento permite uma preparação melhor por parte da UPA. Isto é, se planejar com antecedência para eventos previstos, como shows e jogos de futebol, e implementar medidas para mitigar os efeitos de surtos de doenças ou falta de insumos de saúde, pode melhorar significativamente a eficiência e a capacidade de resposta da UPA. Antecipar picos de demanda durante eventos culturais e esportivos permite alocar recursos adicionais, como pessoal médico e insumos, para lidar com o aumento de atendimentos. Melhorar a gestão e o fornecimento de materiais médicos pode evitar quedas na capacidade de atendimento devido à falta de insumos, garantindo que a UPA esteja sempre pronta para atender a demanda diária.

Os outliers, tanto de alta como de baixa demanda, podem distorcer as medidas de tendência central e dispersão, como média, mediana e desvio padrão, dificultando a obtenção de uma visão precisa da distribuição geral dos atendimentos. A presença de outliers de alta demanda pode aumentar a média dos atendimentos, enquanto outliers de baixa demanda podem diminuí-la. A mediana, sendo uma medida mais robusta, é menos afetada, mas ainda pode sofrer variações dependendo da quantidade e extremidade dos outliers. Além disso, os outliers aumentam significativamente o desvio padrão e a variância dos dados, indicando uma maior dispersão em torno da média. Isso reflete a variabilidade introduzida pelos picos e vales extremos na demanda de atendimento.

A análise visual dos dados, por meio de gráficos como histograma, QQ Plot e boxplot, reforça a conclusão estatística da não normalidade dos dados. A presença de outliers é evidente nos gráficos, onde pontos distantes dos quartis e a assimetria das caixas no boxplot indicam a variabilidade e a dispersão dos atendimentos. O histograma com a curva de densidade KDE mostra uma distribuição enviesada e a presença de picos anômalos. O QQ Plot destaca a divergência dos pontos em relação à linha reta, confirmando a não normalidade dos dados, mesmo após a remoção dos outliers. O boxplot ilustra claramente a dispersão entre os quartis e os outliers, evidenciando a grande variabilidade nos atendimentos diários.

Concomitantemente, em vista da não normalidade na distribuição dos dados, tornou o cálculo do teste de hipótese inviável, uma vez que os métodos utilizados deveriam ser testes de hipótese não paramétricos.

Portanto, ao considerar tanto os aspectos internos quanto os fatores externos, esta análise dos dados da UPA da

Caxangá fornece uma base para aprimorar a resposta e a eficiência dos serviços de saúde na região. A compreensão dos fatores que influenciam a demanda e a variabilidade nos atendimentos diários permite uma preparação melhor e uma gestão mais eficaz. Além disso, a análise revelou como o serviço público de saúde pode ser um indicador valioso da dinâmica socioeconômica e cultural de uma região, destacando as áreas onde intervenções são mais necessárias para melhorar a qualidade de vida e reduzir a vulnerabilidade da população.

VI. ANEXOS

Material:

<https://github.com/AntonioCar0lin0/Statistics-Project->

Apresentação em slides do projeto:

<https://www.canva.com/design/DAGLhR-ePP8/KA-qVg7x-qg52yOzL8nSDQ/edit>

AGRADECIMENTOS

É de suma importância agradecer ao professor Jamilson Ramalho Dantas por ministrar a matéria de Estatística e Probabilidade para Computação, ensinando todas as ferramentas teóricas utilizadas neste artigo, bem como por

dar a oportunidade de realizarmos a análise da base de dados.

REFERÊNCIAS BIBLIOGRÁFICA

- [1] [1] D. A. Silva, H. A. Farias, H. B. A. Nascimento, I. A. F. S. Lima, K. A. L. V. Gomes, e R. A. Santana, "Análise de desempenho do CPU i5-10400 durante a execução do jogo Red Dead Redemption 2," Universidade Federal de Pernambuco - Centro de Informática, Recife, Brasil.
- [2] Ministério da Saúde. "Unidade de Pronto Atendimento (UPA 24h)." Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/u/upa-24h>.
- [3] R. Larson e B. Farber, Estatística Aplicada, 4ª Edição, São Paulo: Pearson, 2010.
- [4] Python Software Foundation. "statistics — Mathematical statistics functions." Disponível em: <https://docs.python.org/3/library/statistics.html>.
- [5] Pandas Development Team. "pandas: powerful Python data analysis toolkit." Disponível em: <https://pandas.pydata.org/>.