

TIS 6 - Análise comparativa de linguagens tipadas e não tipadas mais populares

Belle Nerissa Aguiar Elizeu

Ian Asenjo Dominguez Cunha

Jully Ketely Alves da Silva

Laura Lourdes Coutinho Rodrigues

Warley Leandro dos Anjos

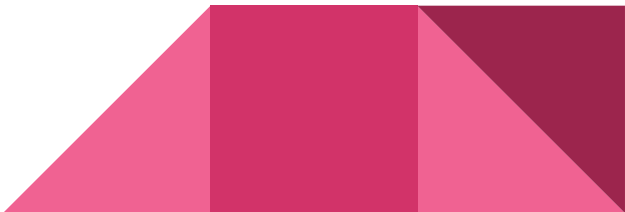
Sumário

1. Introdução
 2. GQM
 3. Questões Principais
 4. Metodologia
 5. Resultados
 6. Ameaças
 7. Discussão
 8. Conclusão
 9. Referências
-

Introdução

A escolha da linguagem de programação pode ter um grande impacto no desenvolvimento de software. Com o surgimento de diversas linguagens de programação ao longo do tempo, pode ser desafiador escolher a melhor opção para uma determinada aplicação.

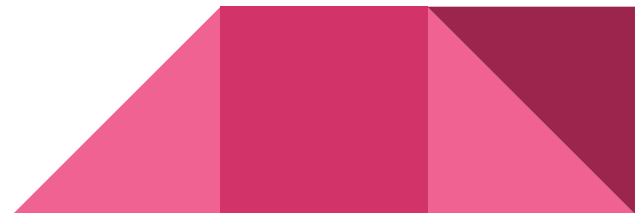
Nesse contexto, a pesquisa TIS 6 propõe uma análise comparativa entre as linguagens tipadas e as não tipadas mais populares, com o objetivo de avaliar as diferenças, vantagens e desvantagens de cada uma delas. Essa análise pode auxiliar desenvolvedores e arquitetos de software na escolha da linguagem mais adequada para seus projetos, considerando fatores como **qualidade de desenvolvimento, tamanho da participação da comunidade e facilidade de correção de bugs.**






GQM (Goal, Question and Metric)

Análise comparativa de repositórios de linguagens tipadas e não tipadas existentes no GitHub **com o objetivo de** avaliar o nível de qualidade de desenvolvimento, participação da comunidade e performance na correção de bugs **com relação** aos resultados obtidos pela aplicação de métricas de qualidade, popularidade e performance de correção de bugs **do ponto de vista** arquitetos de software **no contexto** dos repositórios mais populares open-source do Github de cada tipo.



Questões Principais



[QP.1] Repositórios escritos nas linguagens
tipadas tem melhores qualidade de
desenvolvimento de código comparado a não
tipados?

[M.1] Taxa de Code Smells: Mediana das taxas de code smell obtidas na análise estática de código-fonte providas pelo SonarQube de cada grupo de repositórios

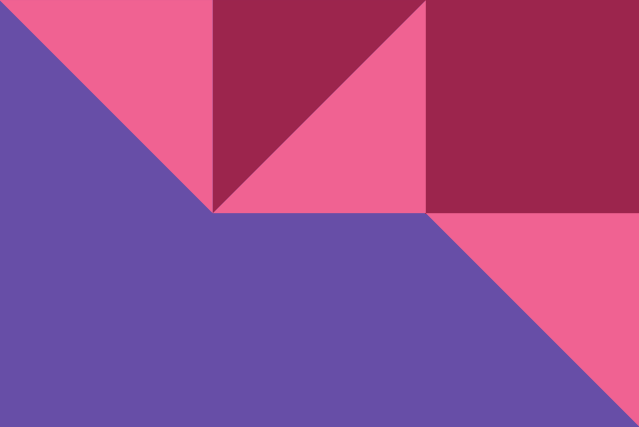
[H.0] Os repositórios de linguagens tipadas apresentaram taxa de code-smell menores que não-tipadas, ou seja uma qualidade superior em relação aos repositórios de linguagens não tipadas.

[H.1] Os repositórios de linguagens tipadas apresentaram taxa de code-smell maiores que não-tipadas, ou seja uma qualidade inferior em relação aos repositórios de linguagens não tipadas.

[M.2] Complexidade Ciclométrica: Mediana dos valores de Complexidade Ciclométrica obtidos na análise estática de código-fonte providas pelo SonarQube de cada grupo de repositórios

[H.0] Os repositórios de linguagens tipadas apresentaram valor de complexidade-ciclométrica menores que não-tipadas, ou seja complexidade superior em relação aos repositórios de linguagens não tipadas.

[H.1] Os repositórios de linguagens tipadas apresentaram valor de complexidade-ciclométrica maiores que não-tipadas, ou seja, complexidade inferior em relação aos repositórios de linguagens não tipadas.



[QP.2] Como os repositórios de linguagens tipadas e não tipadas diferem em termos de tamanho da participação da comunidade?

[M.3] Quantidade de contribuidores: mediana da quantidade de contribuidores de cada um dos grupos de repositórios

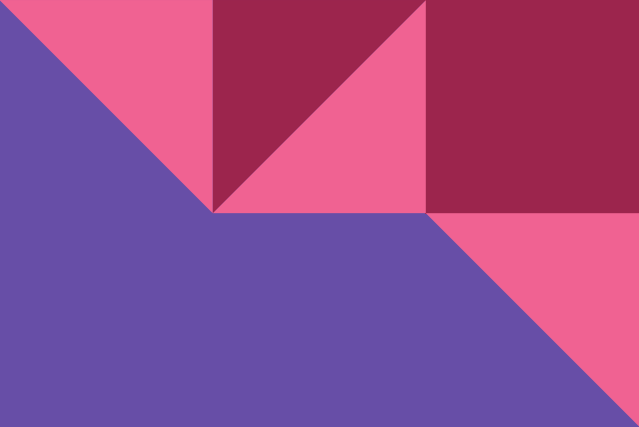
[H.0] A quantidade de contribuidores em repositórios de linguagens tipadas será superior do que as não tipadas.

[H.1] A quantidade de contribuidores em repositórios de linguagens tipadas será inferior do que as não tipadas.

[M.4] Quantidade de contribuições: mediana da quantidade de contribuições calculada pelo total de commits por total de pull request de cada um dos grupos de repositórios

[H.0] A quantidade de contribuições em repositórios de linguagens tipadas será superior do que as não tipadas.

[H.1] A quantidade de contribuições em repositórios de linguagens tipadas será inferior do que as não tipadas.



[QP.3] Como os repositórios de linguagens tipadas e não tipadas diferem em termos de facilidade de correção de bugs?

[M.5] Desempenho de correção de bugs: Mediana do volume do total de issues de label “bug” de cada um dos grupos de repositórios de linguagens tipadas e não tipadas

[H.0] A mediana do volume do total de issues de label “bug” do grupo de repositórios não tipados será inferior a de tipados.

[H.1] A mediana do volume do total de issues de label “bug” do grupo de repositórios não tipados será superior a de tipados.

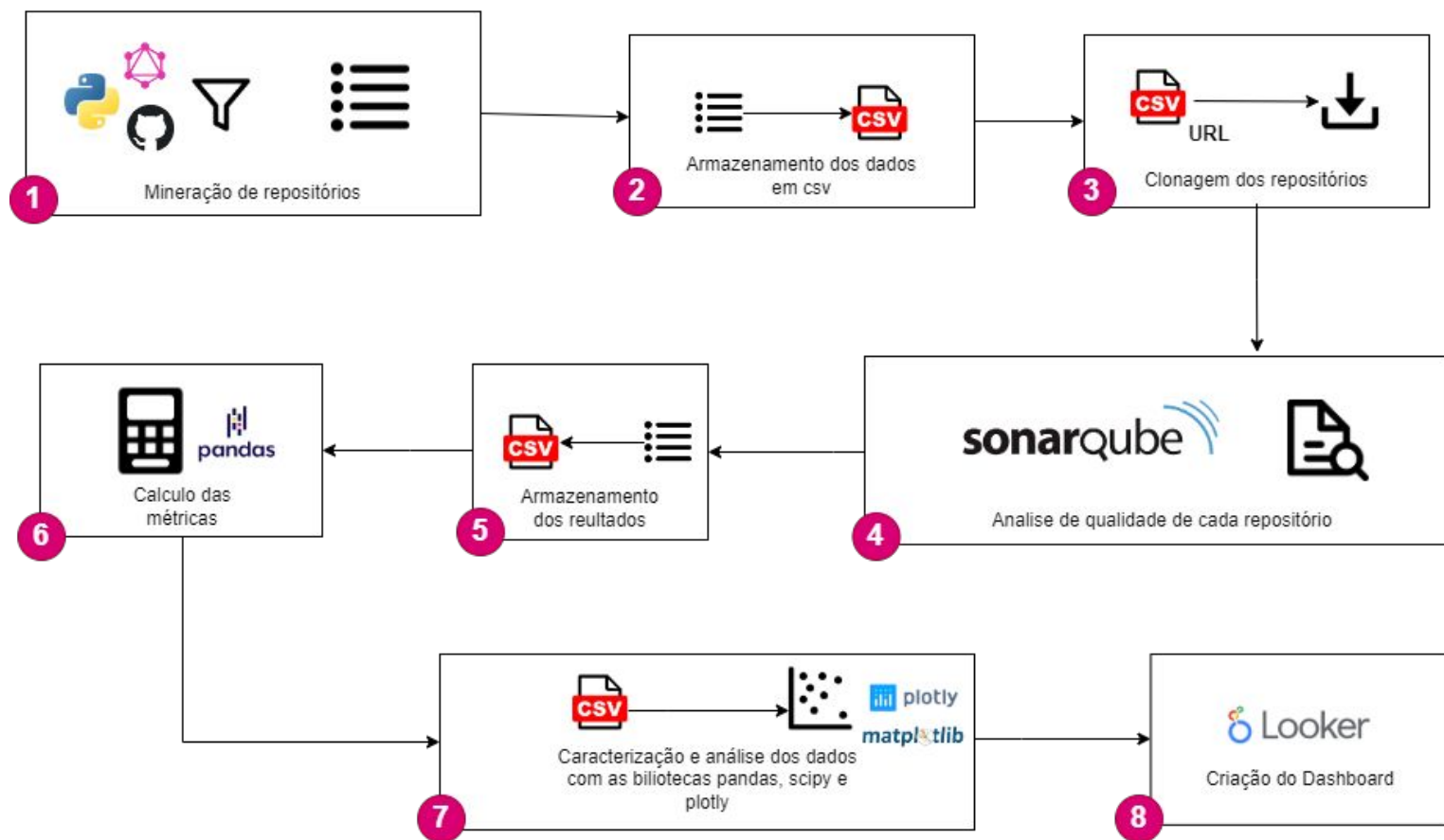
[M.6] Desempenho da correção de bugs: Mediana do tempo médio do fechamento de uma issue com label “bug” um dos grupos repositórios de linguagens tipadas e não tipadas

[H.0] A mediana do tempo médio do fechamento de issues de bug no grupo de de linguagens tipadas será inferior do que as não tipadas.

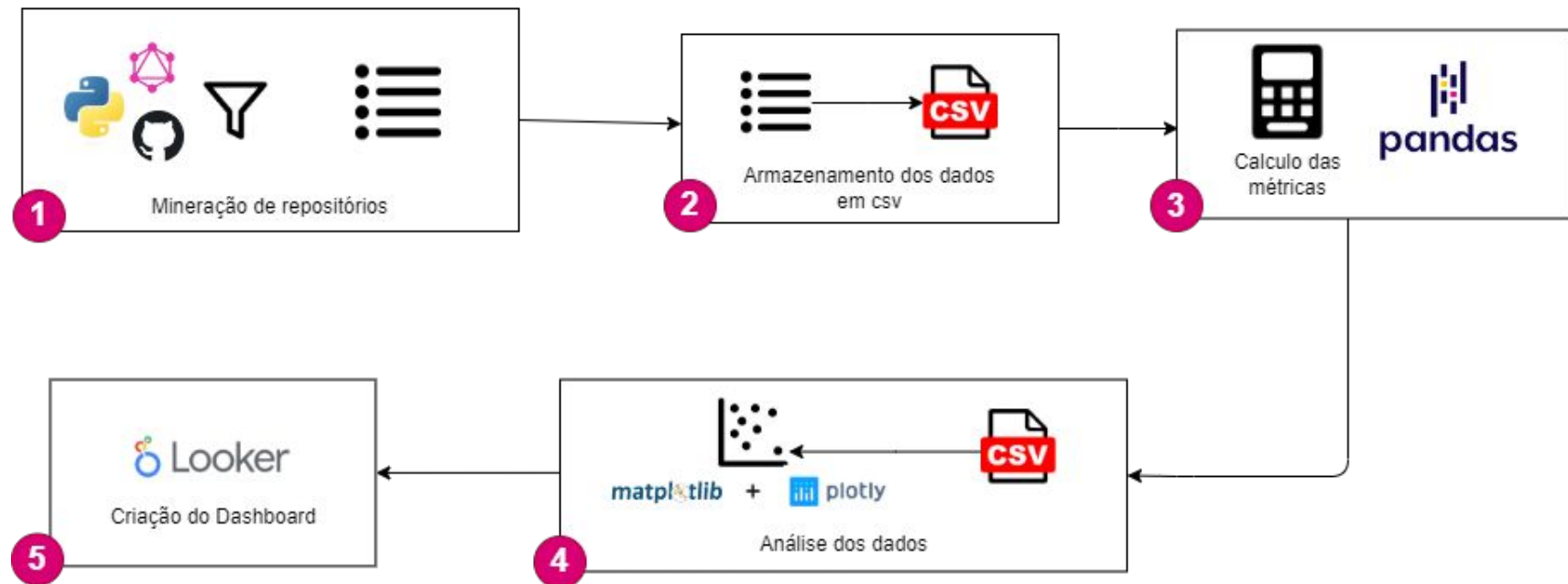
[H.1] A mediana do tempo médio do fechamento de issues de bug no grupo de repositórios de linguagens tipadas será superior do que as não tipadas.

Metodologia

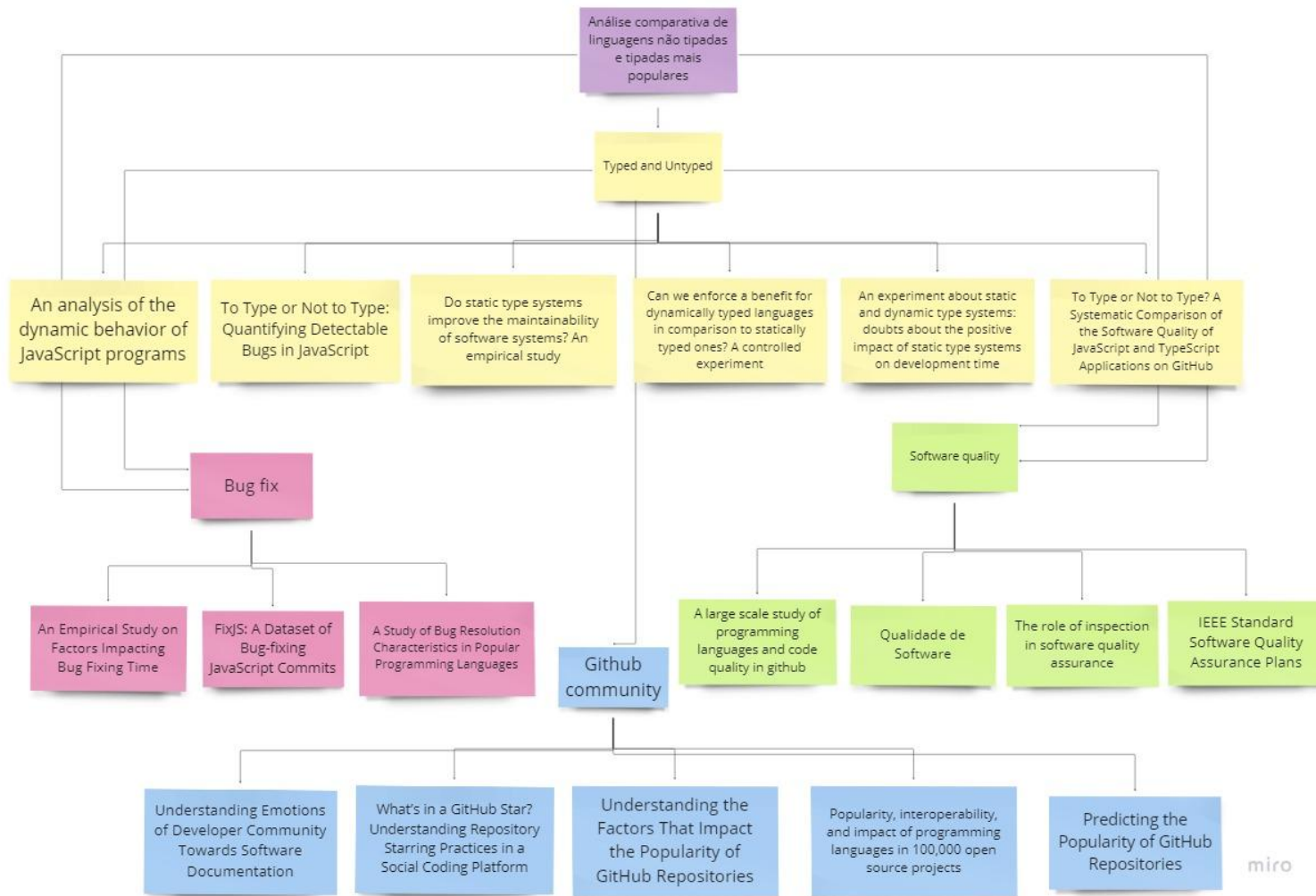
Metodologia para QP1



Metodologia para QP2 e QP3



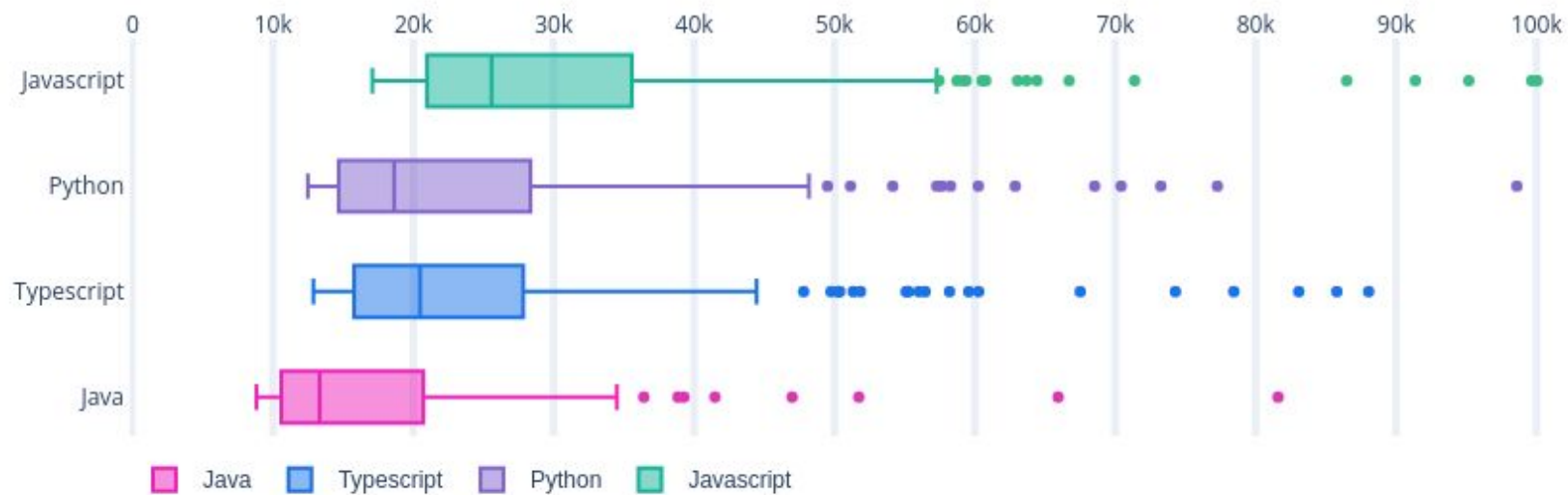
Trabalhos Relacionados





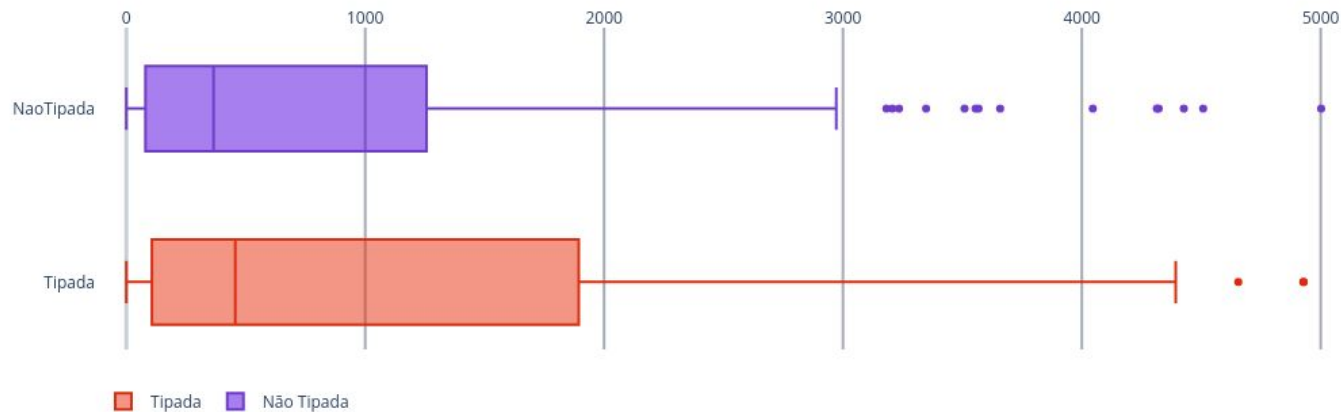
Resultados

Número de estrelas



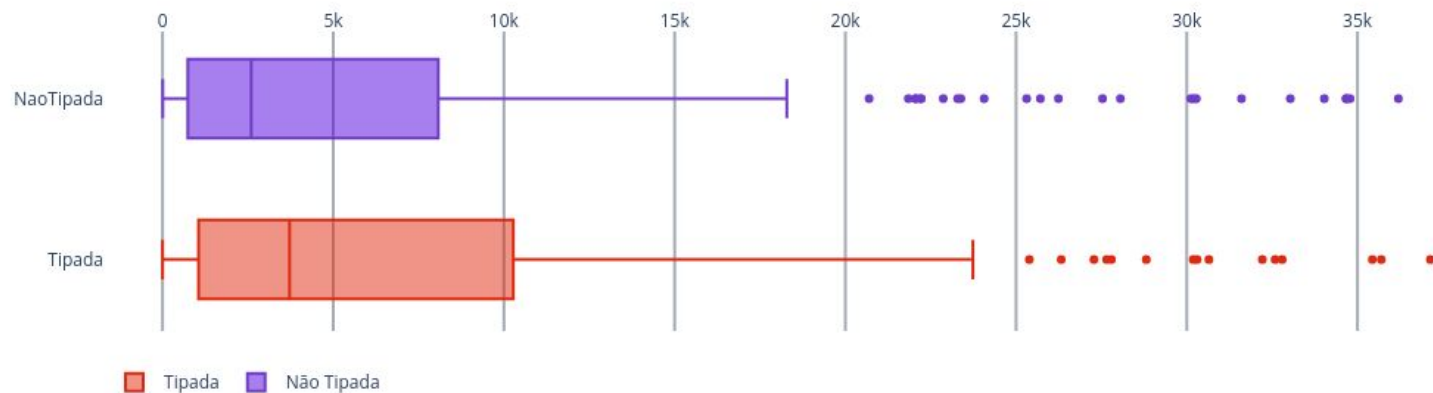
Mann-Whitney U: p-value: 0.0558

Taxa de Code Smell



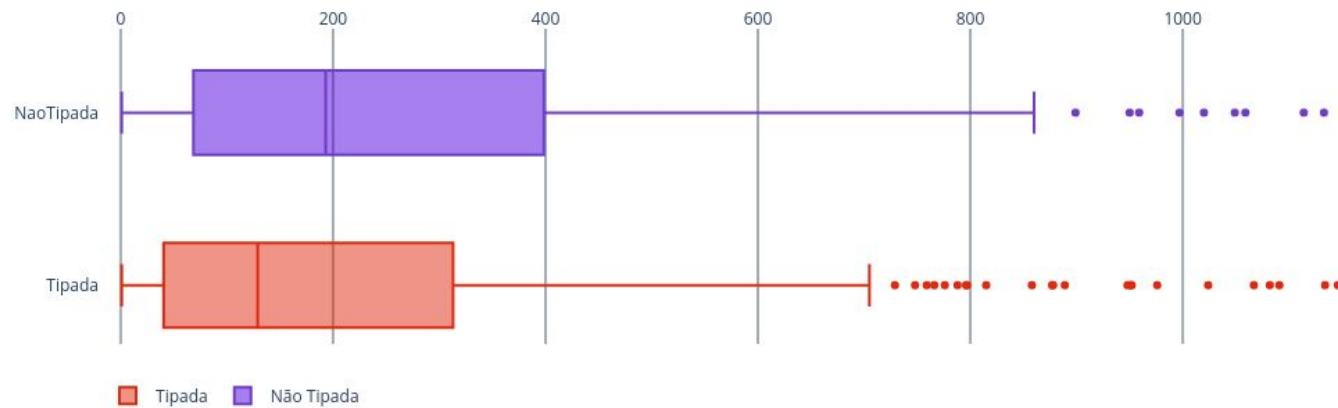
Mann-Whitney U: p-value: 0.0189

Complexidade Ciclomática



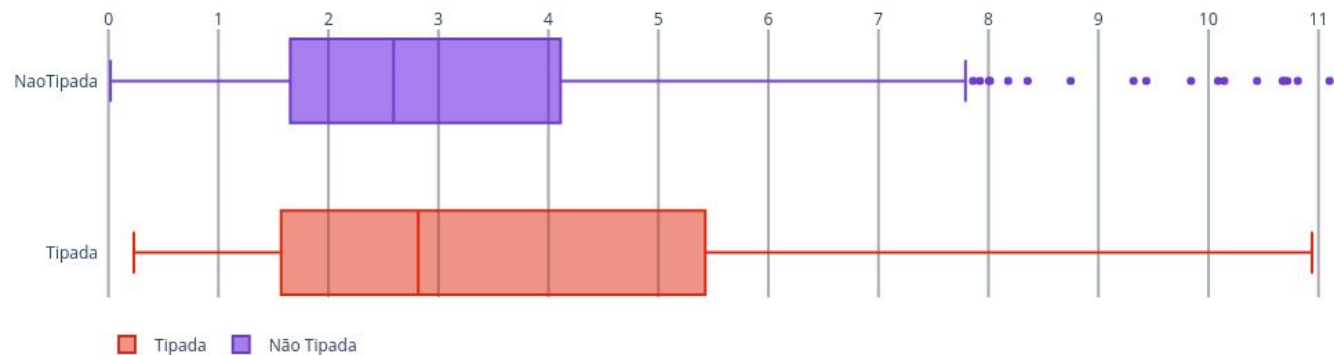
Mann-Whitney U: p-value: 0.0003

Quantidade de Contribuidores

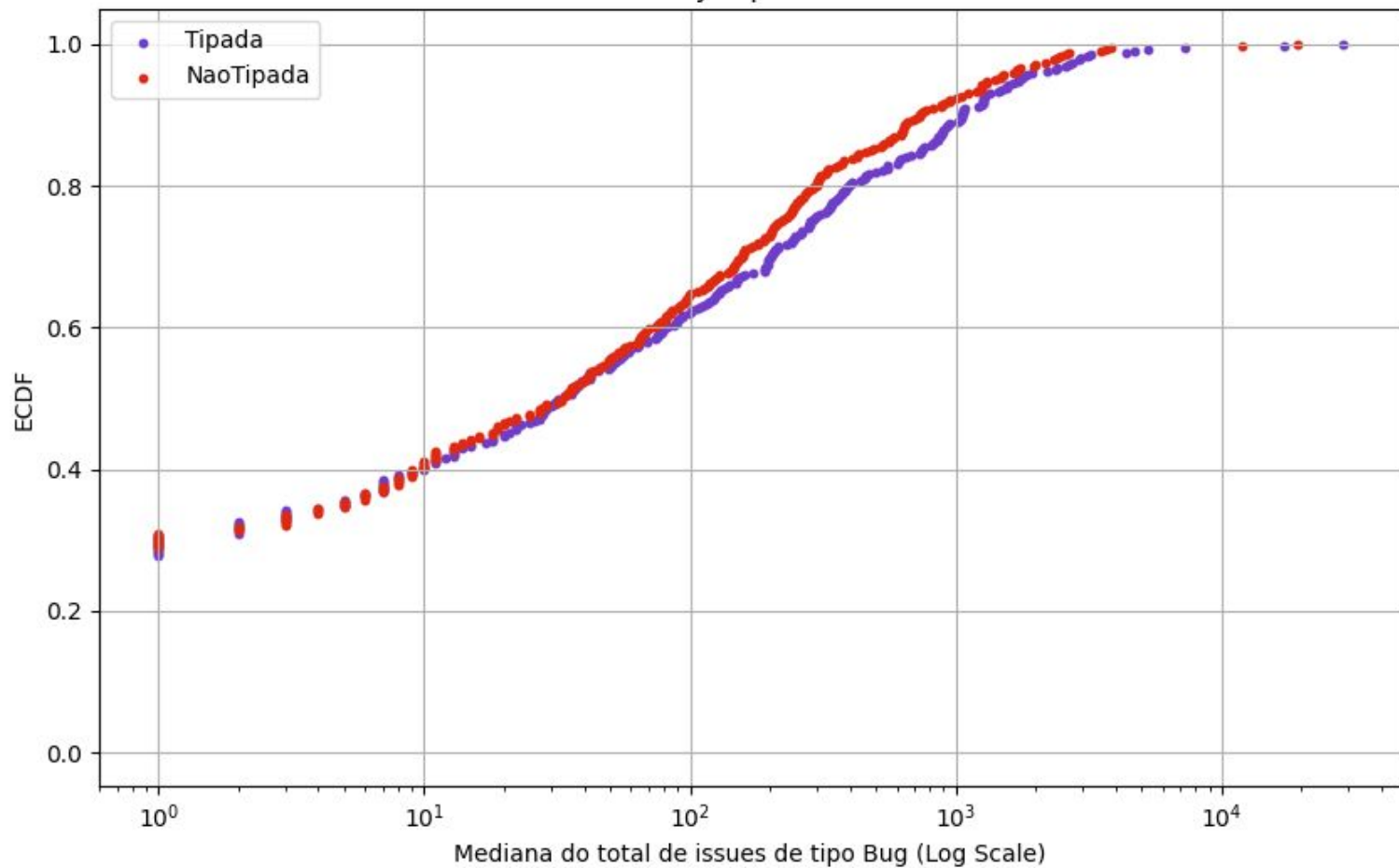


Mann-Whitney U: p-value: 0.0768

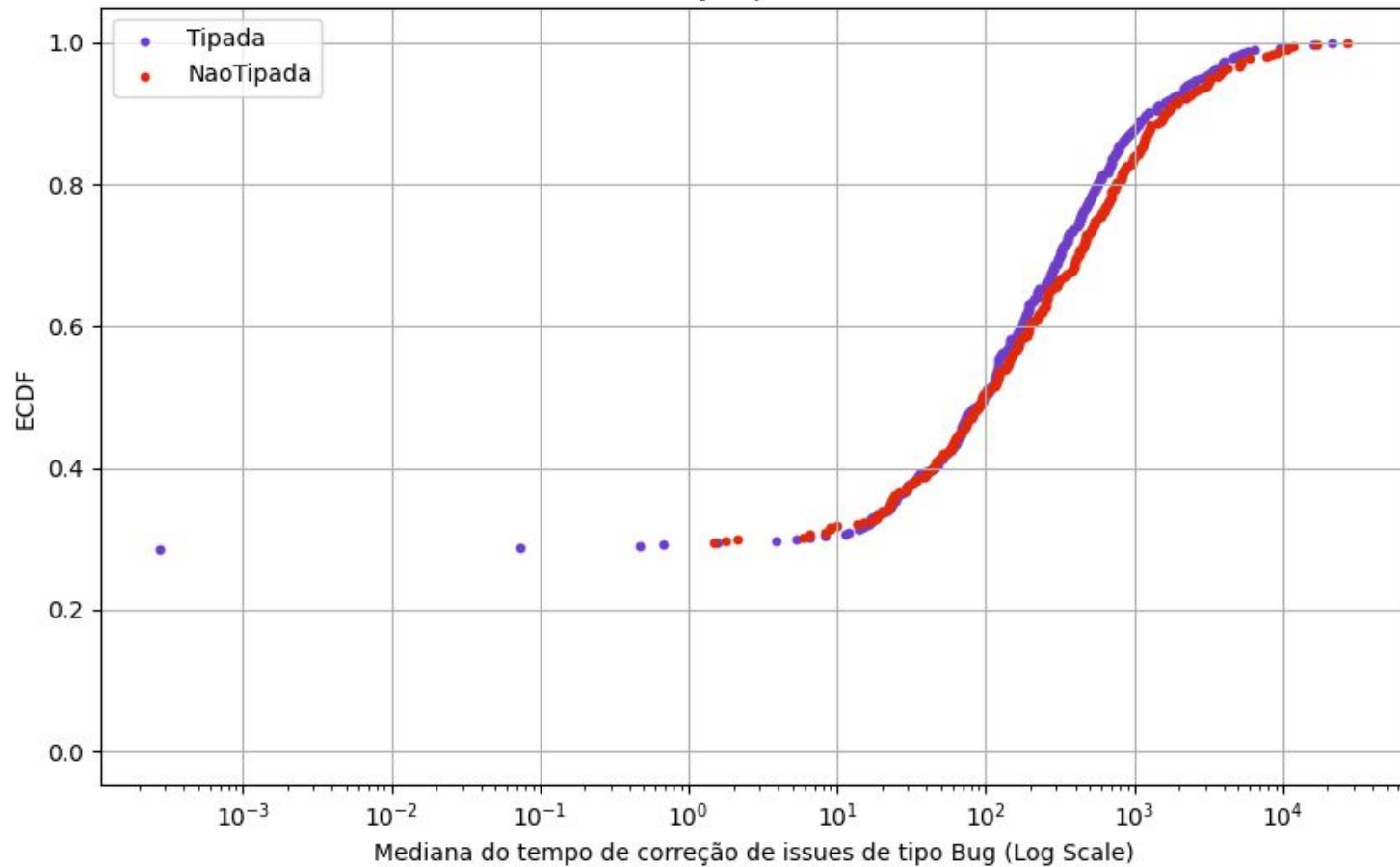
Total de *commits*/ total de *pull requests*



ECDF de Mediana do total de issues de Bug por Linguagem
Mann-Whitney U p-value = 0.4354



ECDF de Mediana do tempo de correção de issues de Bug por grupo
Mann-Whitney U p-value = 0.5119



Ameaças à validade

Construção

A quantidade de repositórios estudados pode não ser estatisticamente representativo.

Externa

Instabilidade de rede, indisponibilidade da API ou vencimento da validade do *token*.

Interna

Falha na máquina responsável pela execução dos *scripts*, falta de otimização do algoritmo de coleta e análise de dados ou indisponibilidade das ferramentas utilizadas, *SonarQube* e *Looker*.

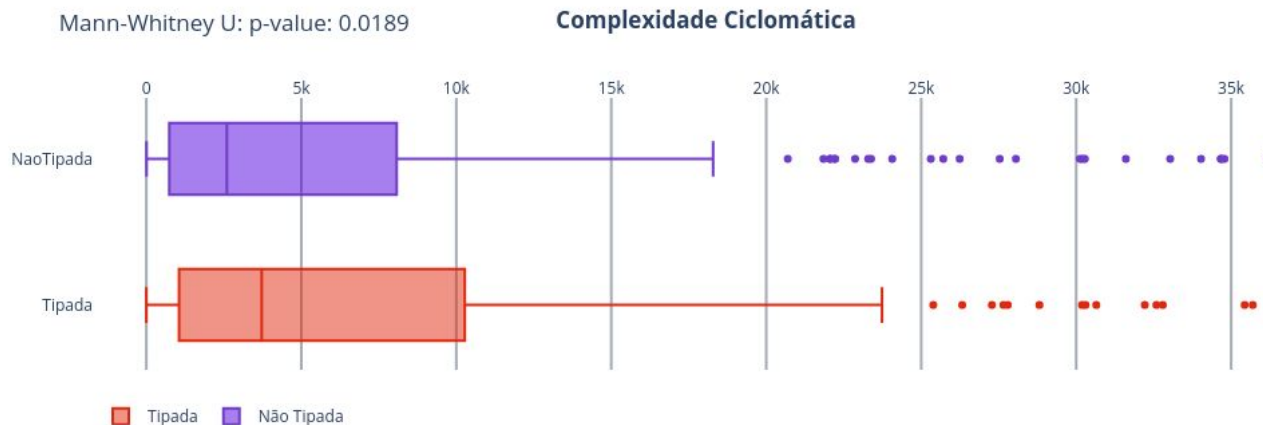
Conclusão

A maturidade das linguagens pode afetar na análise. As linguagens mais maduras (JavaScript e Java) tendem a ter mais recursos, melhor estabilidade e suporte da comunidade em relação a TypeScript e Python.



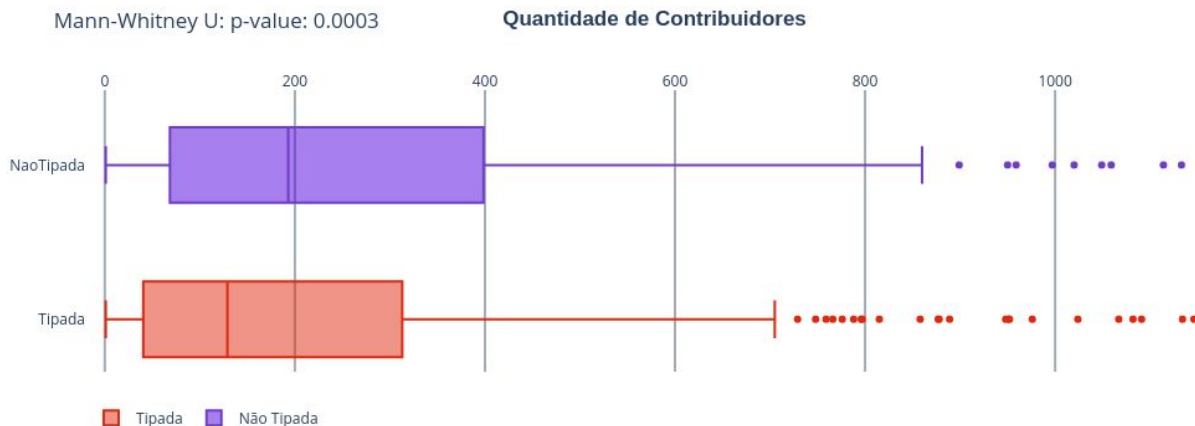
Discussão dos resultados

QP.1 Repositórios escritos nas linguagens tipadas tem melhores qualidade de desenvolvimento de código comparado a não tipados?



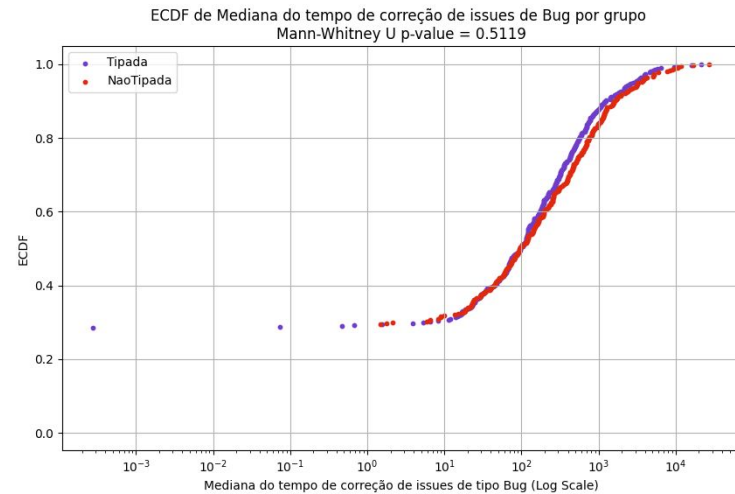
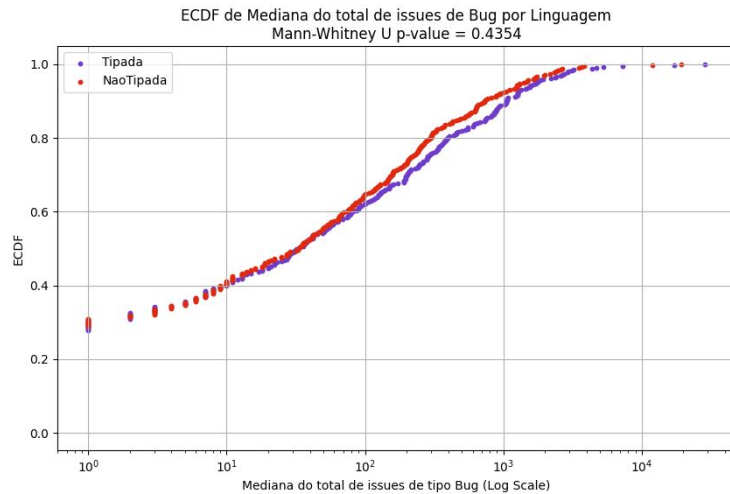
A escolha de linguagens não tipadas pode ser considerada como um fator determinante para a qualidade do desenvolvimento de código. Porém, outros fatores como, práticas de programação, experiência da equipe e boas práticas de engenharia de software, devem ser considerados para avaliar e melhorar a qualidade do código independentemente do tipo escolhido.

QP.2 Como os repositórios de linguagens tipadas e não tipadas diferem em termos de tamanho da participação da comunidade?



A quantidade de contribuidores do grupo de repositórios não tipados é maior que os de tipados. Em contrapartida, a participação é equivalente entre os dois grupos, o que significa que a decisão de escolha entre linguagens tipadas e não tipadas em projetos é irrelevante em relação à participação efetiva das comunidades.

QP.3 Como os repositórios de linguagens tipadas e não tipadas diferem em termos de facilidade de correção de bugs?



Embora os repositórios não tipados tenham um volume de correções de bugs maior, não há diferenças significativas no tempo de correção quando comparados aos repositórios tipados. Portanto, no contexto específico da facilidade de correção, não é evidente que há indícios de que o uso de linguagens tipadas ou não tipadas influencie de forma substancial na eficiência da correção.



Conclusão

- A escolha de linguagens **não tipadas** pode ser considerada como um fator determinante para a **qualidade do desenvolvimento de código, principalmente se relevar a baixa a taxa de complexidade ciclomática.**
- As **linguagens não tipadas** mais populares tendem a **atrair mais contribuidores**, mas a quantidade de contribuições não mostrou diferenças relevantes em comparação com os repositórios de linguagens **tipadas**.
- As linguagens **não tipadas** têm valores **maiores de quantidade de issues de tipo bug** comparado ao grupo de linguagens **tipadas**.
- Os valores de **tempo médio de fechamento das issues** de **tipados** demonstram ser ligeiramente **maiores** comparadas a **não tipadas**.

Portanto, com base nas medições realizadas, **não podemos afirmar que os repositórios escritos em linguagens tipadas são superiores em qualidade de código, tamanho da comunidade ou facilidade de correção de bugs** em comparação com os repositórios escritos em linguagens não tipadas. A escolha entre linguagens tipadas e não tipadas também é feita considerando outros fatores, como preferências pessoais, requisitos do projeto e ecossistema de desenvolvimento disponível.

Link Repositório:

<https://github.com/ICEI-PUC-Minas-PPLES-TI/plf-es-2023-1-ti6-3150100-pesquisalinguagemtipadaenaotipada-1>

Link Diagrama Trabalhos Relacionados:

https://miro.com/app/board/uXjVMQZIMsc=