

Relatório Final Laboratório 02 - Um Estudo das Características de Qualidade de Sistemas Java

Integrantes do grupo:

- Bernardo Cruz Rohlf
- Leonardo Augusto Pereira do Carmo
- Matheus Belo Santos Mello
- Tarcísio Ney Martins Filho

Professora: Aline Norberta de Brito

Introdução

O objetivo deste laboratório é analisar aspectos da qualidade de repositórios desenvolvidos na linguagem Java, correlacionando-os com características do seu processo de desenvolvimento, sob a perspectiva de métricas de produto calculadas através da ferramenta CK. Para isso, coletamos métricas de 1000 repositórios para realizarmos a análise e visualização desses dados.

RQ 01. Qual a relação entre a **popularidade** dos repositórios e as suas características de qualidade?

Hipótese Informal: Repositórios mais populares (medidos por estrelas, forks ou outras métricas de engajamento) tendem a apresentar melhores características de qualidade, como código mais limpo, menor número de defeitos e maior documentação.

RQ 02. Qual a relação entre a **maturidade** dos repositórios e as suas características de qualidade ?

Hipótese Informal: Repositórios mais antigos ou maduros têm maior probabilidade de possuir características de qualidade superiores, como maior estabilidade, melhor manutenibilidade e maior aderência às boas práticas de desenvolvimento.

RQ 03. Qual a relação entre a **atividade** dos repositórios e as suas características de qualidade?

Hipótese Informal: Repositórios com maior atividade (número de commits, pull requests, issues etc.) tendem a exibir melhores características de qualidade, pois a colaboração e revisão ativa podem levar a melhorias no código.

RQ 04. Qual a relação entre o **tamanho** dos repositórios e as suas características de qualidade?

Hipótese Informal: O tamanho de um repositório (medido em termos de linhas de código, arquivos ou complexidade) pode afetar suas características de qualidade. Repositórios menores podem ser mais fáceis de manter e ter menos defeitos, enquanto repositórios maiores podem ser mais propensos a problemas de escalabilidade e complexidade.

Metodologia

- **Realizamos** a coleta de informações de **1000 repositórios** hospedados em plataformas como o **GitHub** usando o **GraphQL**.
- **Realizamos** uma análise exploratória dos dados coletados:
 - **Visualização de Dados:** Criamos **gráficos** para entender a distribuição das métricas de produto. Por exemplo, plotamos histogramas para visualizar a frequência das métricas.
 - **Resumo Estatístico:** Calculamos **médias**, medianas, desvios padrão para cada métrica. Isso nos ajudou a compreender a variabilidade dos dados.
- **Exploramos** a relação entre as métricas de produto e as características de qualidade dos repositórios.
- **Inferimos** conclusões com base nos resultados obtidos, considerando as hipóteses informais e os resultados encontrados

Requisitos

RQ 01. Qual a relação entre a **popularidade** dos repositórios e as suas características de qualidade?

Gráfico:

Conclusão (Hipótese Informal x Resultados Encontrados):

RQ 02. Qual a relação entre a **maduridade** dos repositórios e as suas características de qualidade ?

Gráfico:

Conclusão (Hipótese Informal x Resultados Encontrados):

RQ 03. Qual a relação entre a **atividade** dos repositórios e as suas características de qualidade?

Gráfico:

Conclusão (Hipótese Informal x Resultados Encontrados):

RQ 04. Qual a relação entre o **tamanho** dos repositórios e as suas características de qualidade?

Gráfico:

Conclusão (Hipótese Informal x Resultados Encontrados):