Laboratório de Medição e Experimentação Eng. de Software - PUC Minas

UM ESTUDO DAS CARACTERÍSTICAS DE QUALIDADE DE SISTEMAS JAVA

Apresentação de Resultados



Objetivo

Analisar aspectos da **qualidade** interna de repositórios desenvolvidos em **Java**, correlacionando-os com **características** do seu processo de desenvolvimento.

Tecnologias







Questões de Pesquisa (Research Questions – RQs)

RQ01. Qual a relação entre a **popularidade** dos repositórios e suas características de qualidade?

RQ02. Qual a relação entre a **maturidade** dos repositórios e suas características de qualidade?

RQ03. Qual a relação entre a **atividade** dos repositórios e suas características de qualidade?

RQ04. Qual a relação entre o **tamanho** dos repositórios e suas características de qualidade?

Hipóteses Informais (Informal Hypotheses -IH)

IH01. Repositórios mais populares tendem a apresentar melhor legibilidade e modularidade, já que atraem mais colaboradores e passam por revisões frequentes.

IHO2. Projetos maduros, mantidos por mais tempo, possuem métricas de qualidade mais consistentes, refletindo evolução gradual e práticas consolidadas de desenvolvimento.

IH03. Repositórios com maior atividade (commits e pull requests frequentes) apresentam maior manutenibilidade, uma vez que o código é constantemente atualizado e ajustado.

IH04. Repositórios maiores tendem a apresentar desafios na manutenção e modularidade, já que o aumento de tamanho pode impactar negativamente a simplicidade e legibilidade do código.

Metodologia



Coleta de Dados

- 1000 repositórios
 Java mais
 populares.
- Paginação e tratamento de limites da API.

4

Pré processamento

- Normalização dos dados coletados.
- Organização dos dados em arquivos CSV.



Extração de métricas

- Execução do CK
 Tool para extrair
 métricas de
 qualidade.
- Geração de arquivos CSV para cada repositório.



Sumarização/ Limpeza

- Consolidação dos dados em tabelas únicas.
- Limpeza de linhas vazias e filtragem de métricas relevantes.



Cálculo/ Agregação

- Cálculo de estatísticas descritivas.
- Agregação das métricas por repositório e por classe.

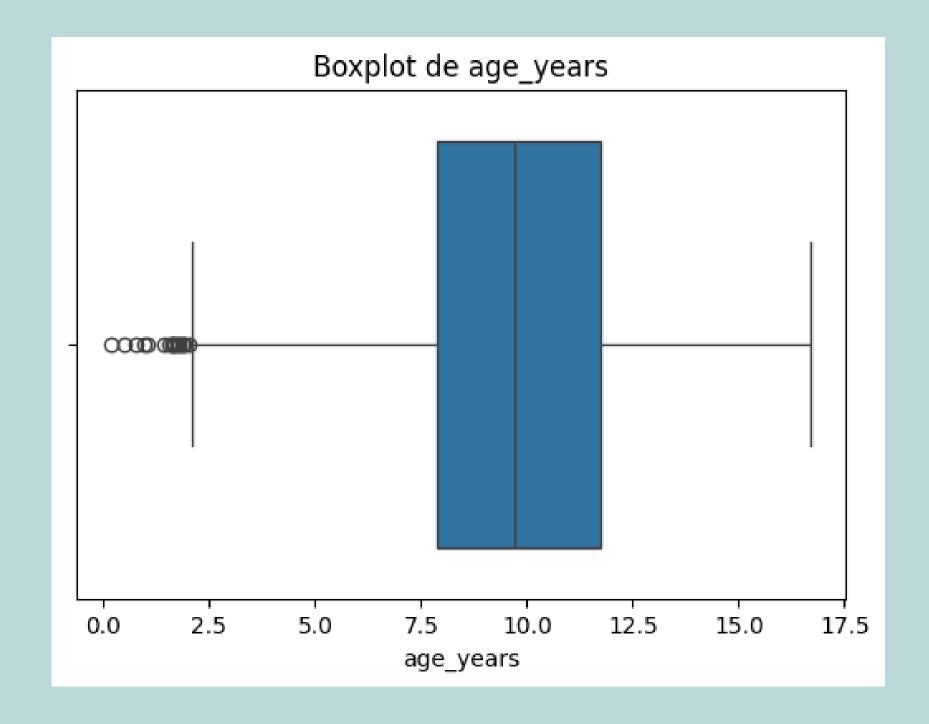


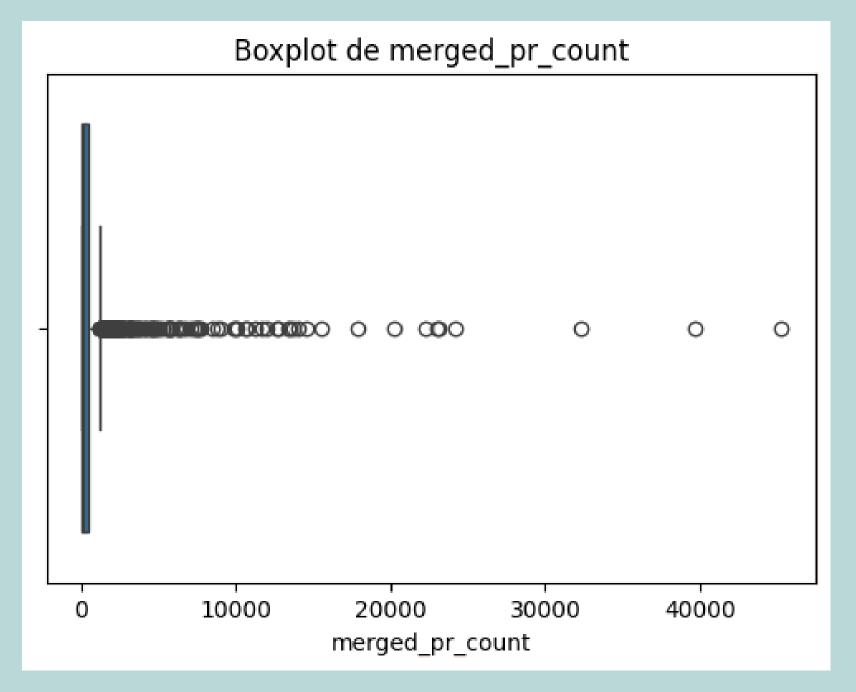
Visualização/ Análise

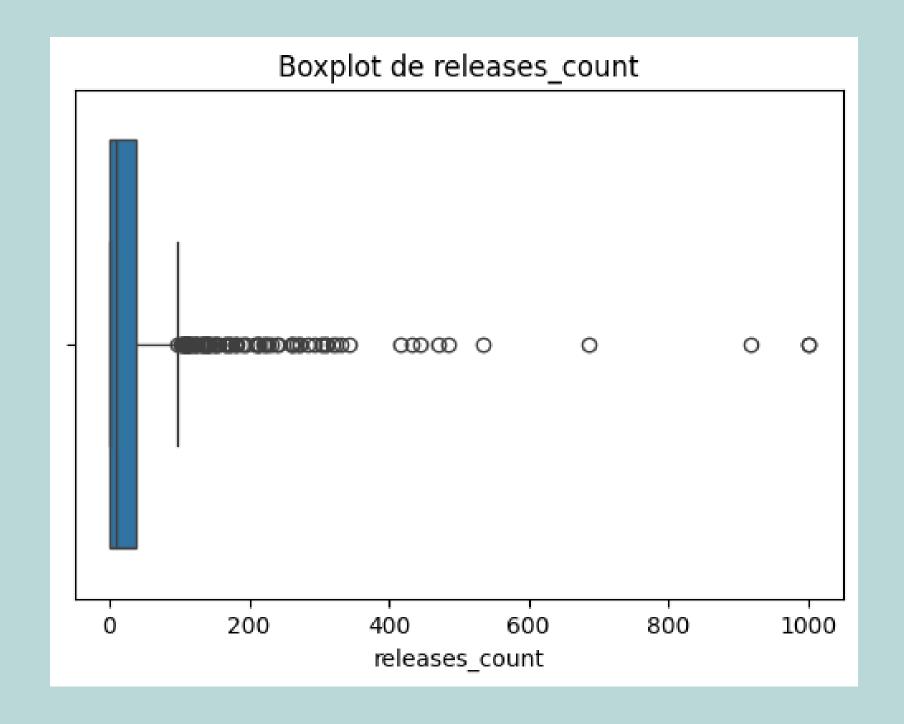
- Geração dos gráficos.
- Interpretação dos resultados para responder às questões de pesquisa.

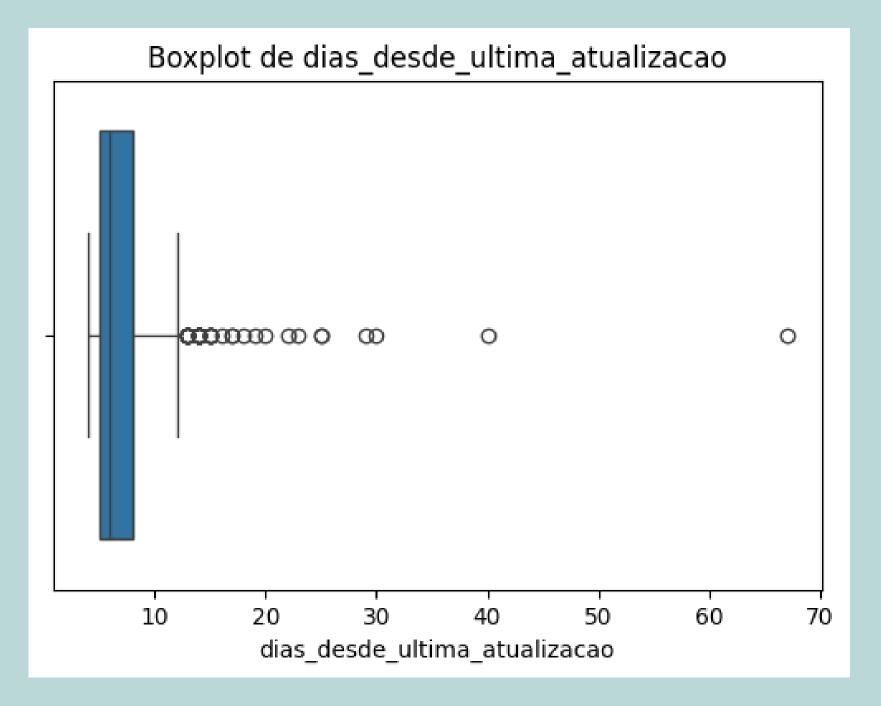
Resultados

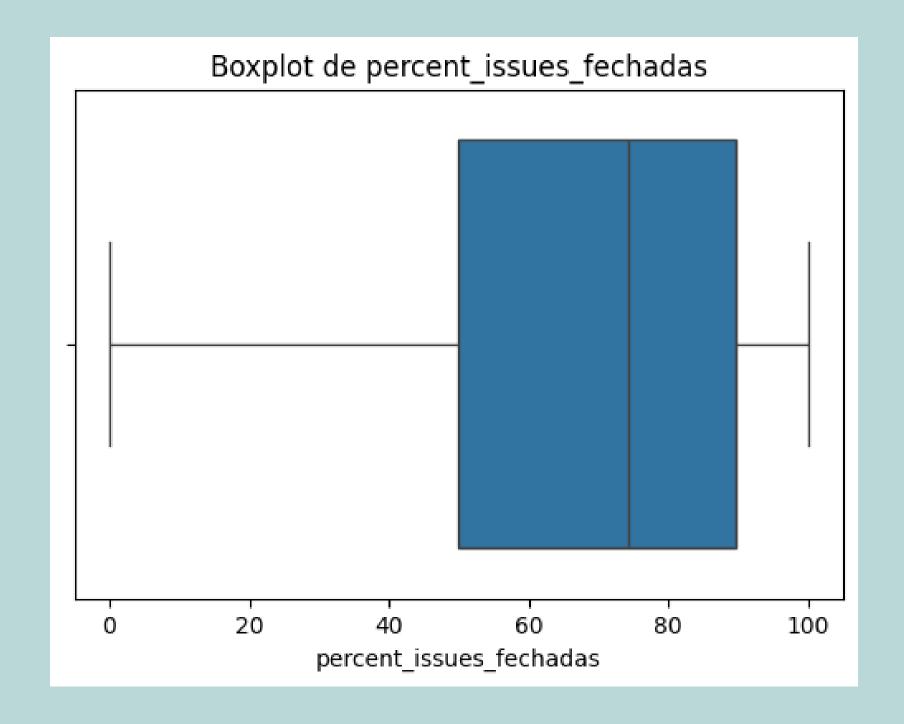
Métrica	Código	Média	Mediana	Moda	Desvio Padrão	Mínimo	Máximo
Idade do Repositório (anos)	LM01	9.61	9.71	9.68	3.04	0.18	16.69
✓ Pull Requests Aceitas	LM02	1026.93	67.00	0	3379.50	0	45219
Número de Releases	LM03	38.78	10.00	0	86.11	0	1000
Tempo desde a Última Atualização (dias)	LM04	2.08	1.00	0	3.59	0	62
Percentual de Issues Fechadas (%)	LM05	66.59	74.25	0.0	28.05	0.0	100.0
☆ Número de Estrelas (Stars)	LM06	9288.85	5716.00	3954	10594.80	3415	117052
Número de Forks	LM07	2344.96	1349.00	1051	3709.58	128	54106
Tamanho do Repositório (LOC)	LM08	50.30	43.85	5.0	31.28	2.0	406.333
⊗ CBO	AM01	5.37	5.32	0.0	1.87	0.0	21.937
■ DIT	AM02	1.46	1.39	1.0	0.35	1.0	4.388
₹ LCOM	AM03	118.24	23.60	0.0	1780.84	0.0	54799.523

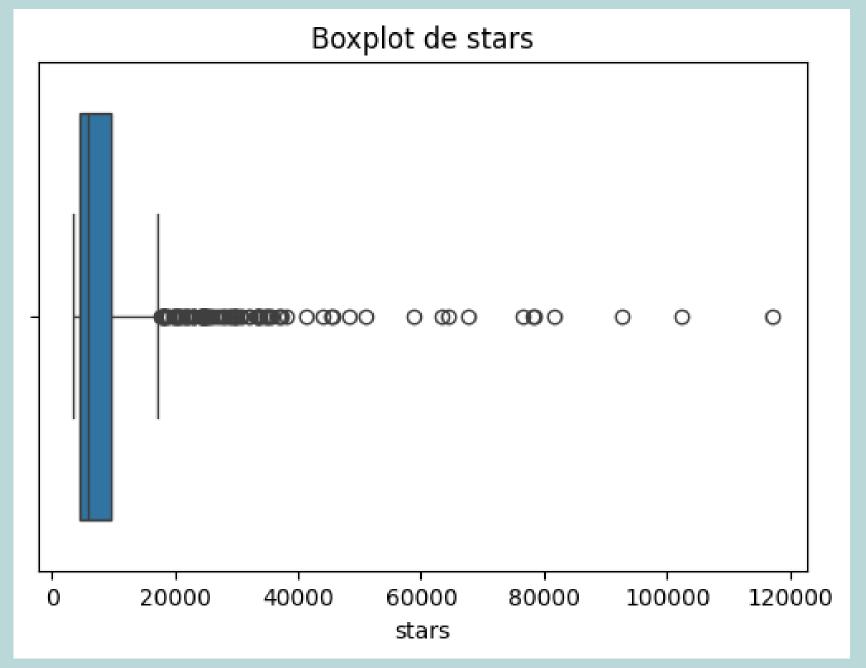


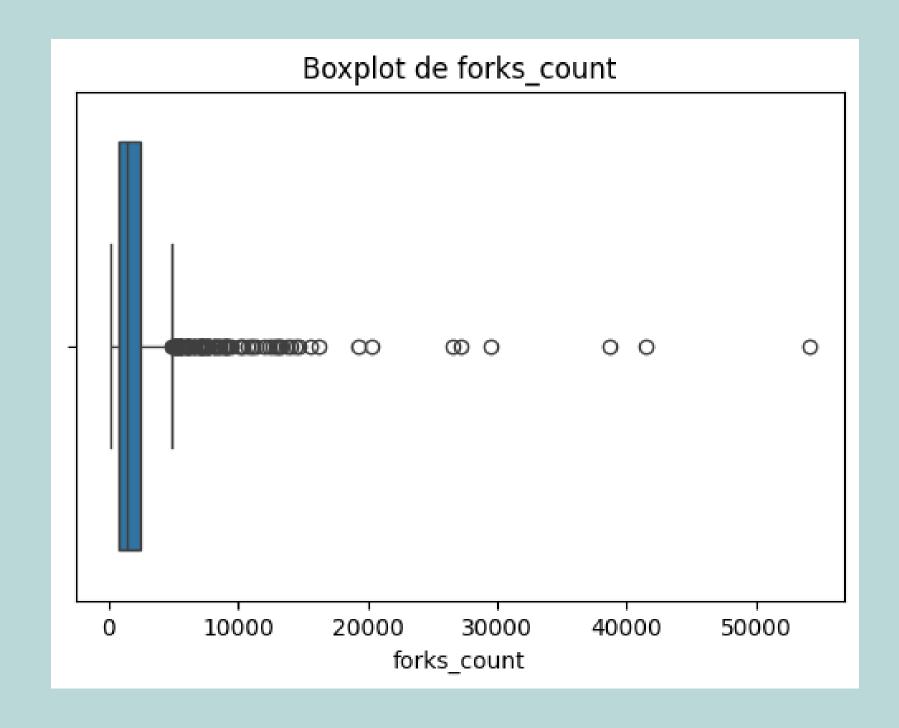


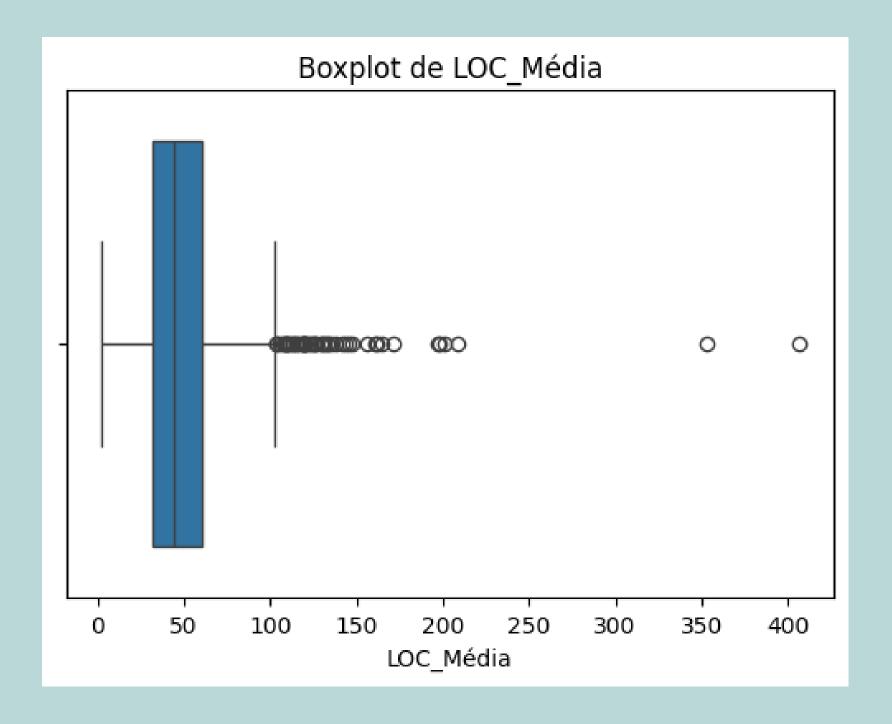


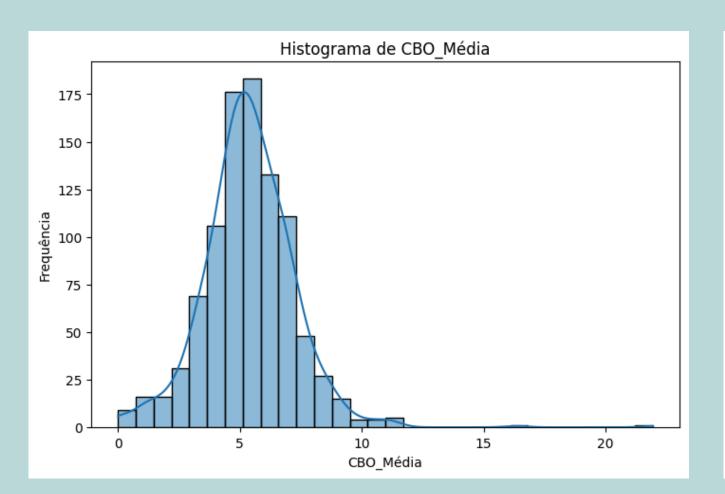


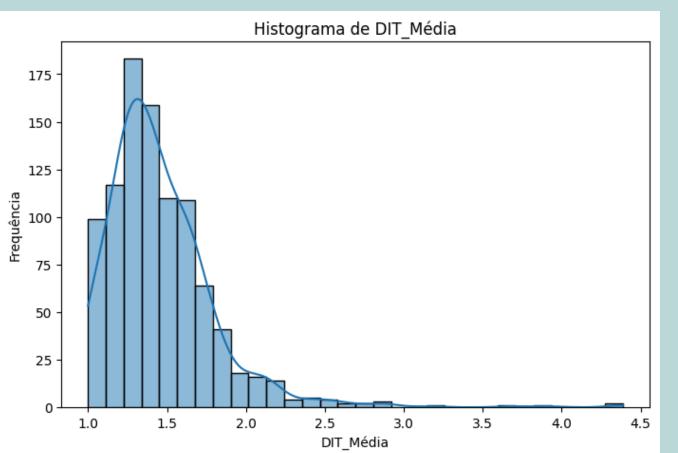


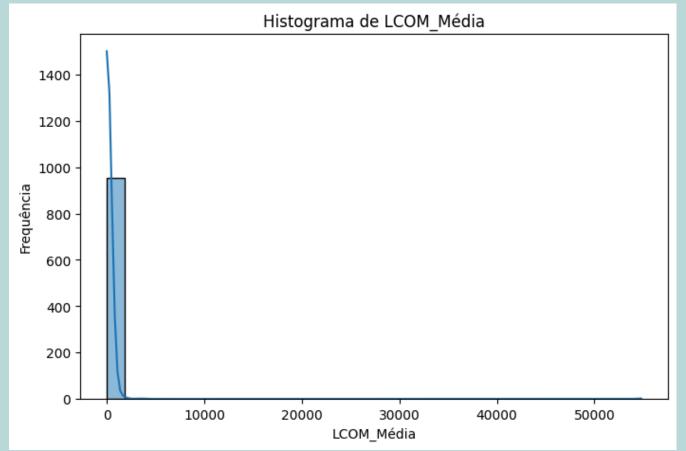




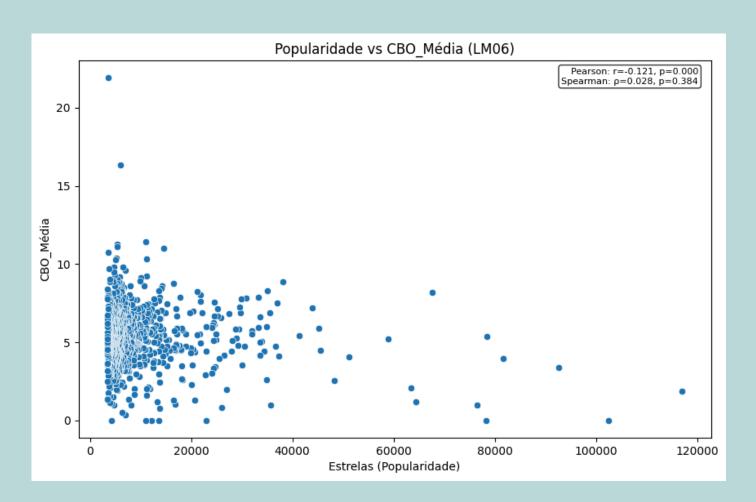


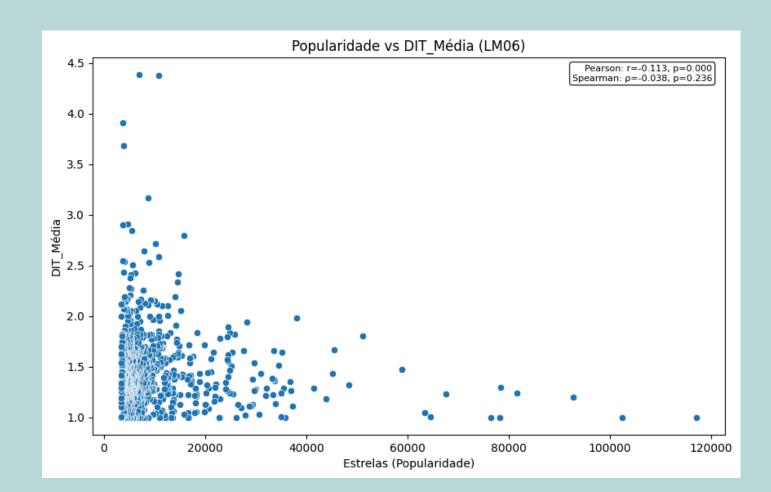


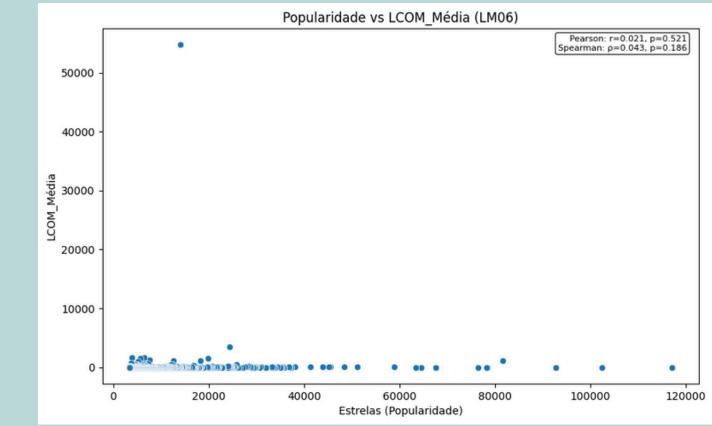




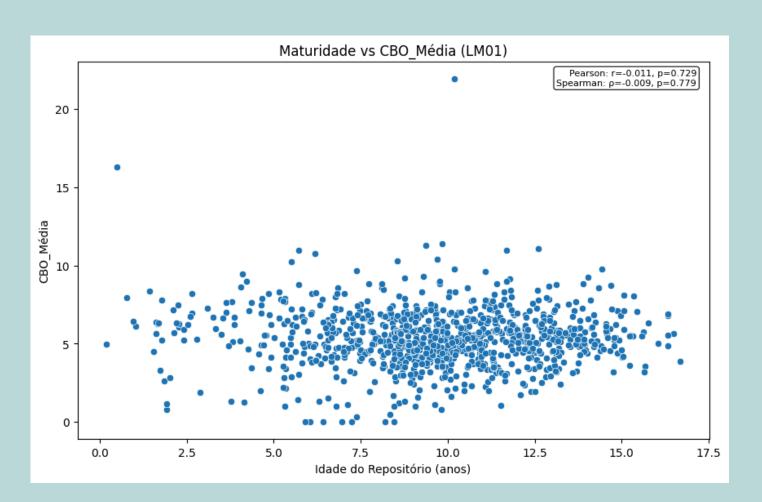
Popularidade vs Qualidade (RQ01)

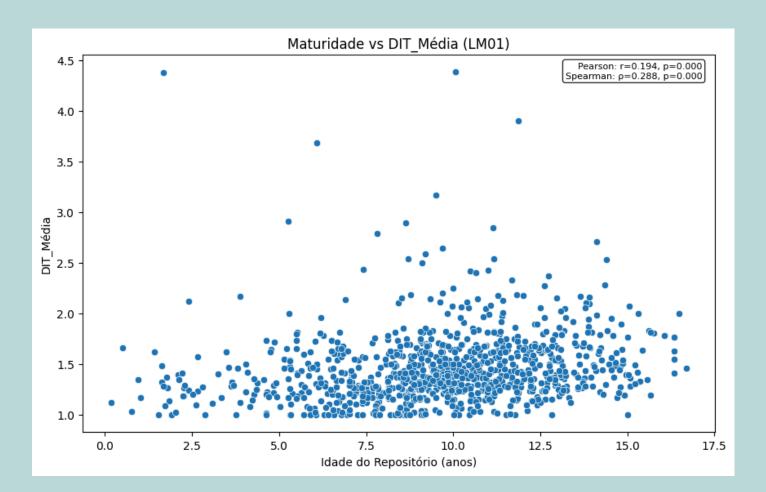


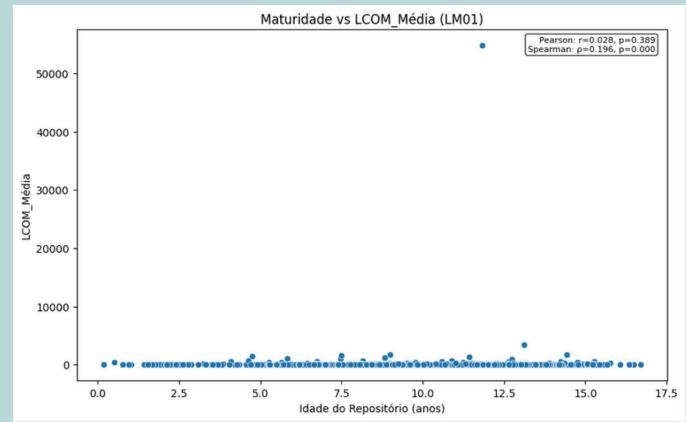




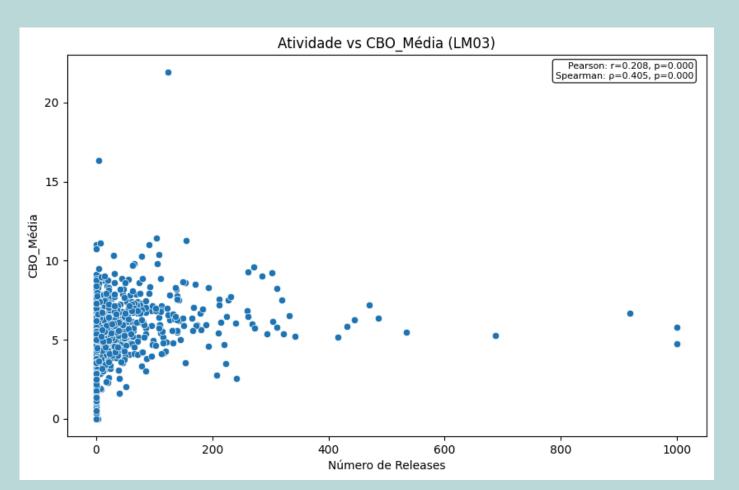
Maturidade vs Qualidade (RQ02)

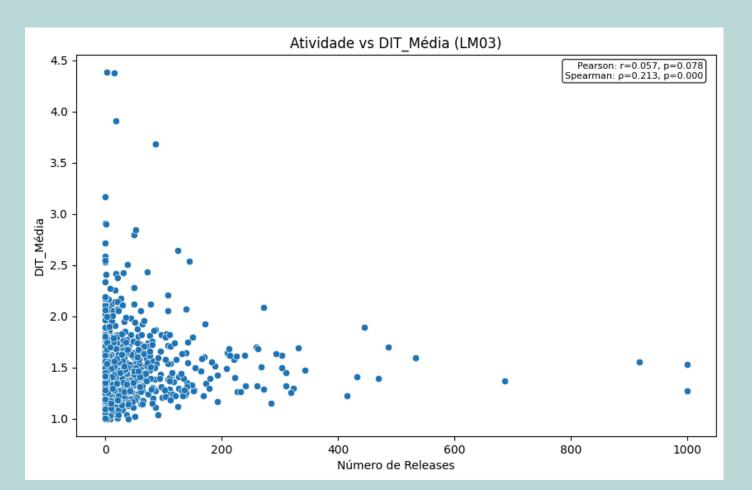


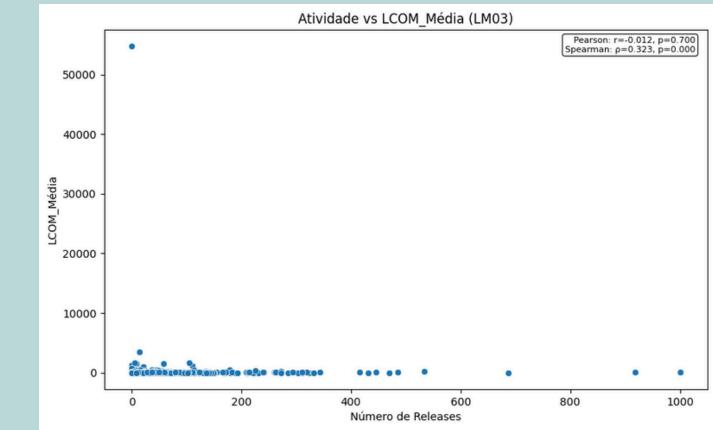




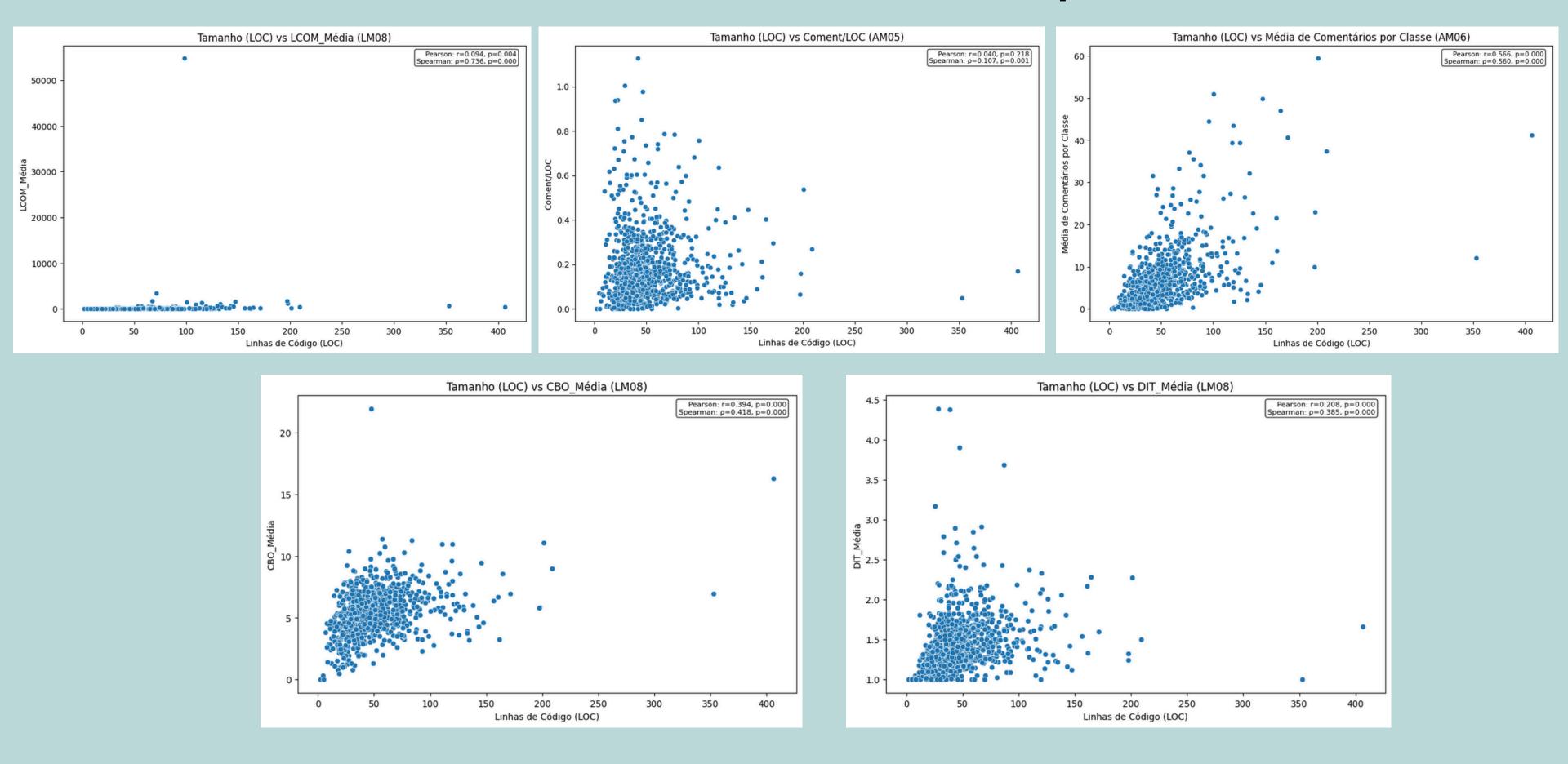
Atividade vs Qualidade (RQ03)



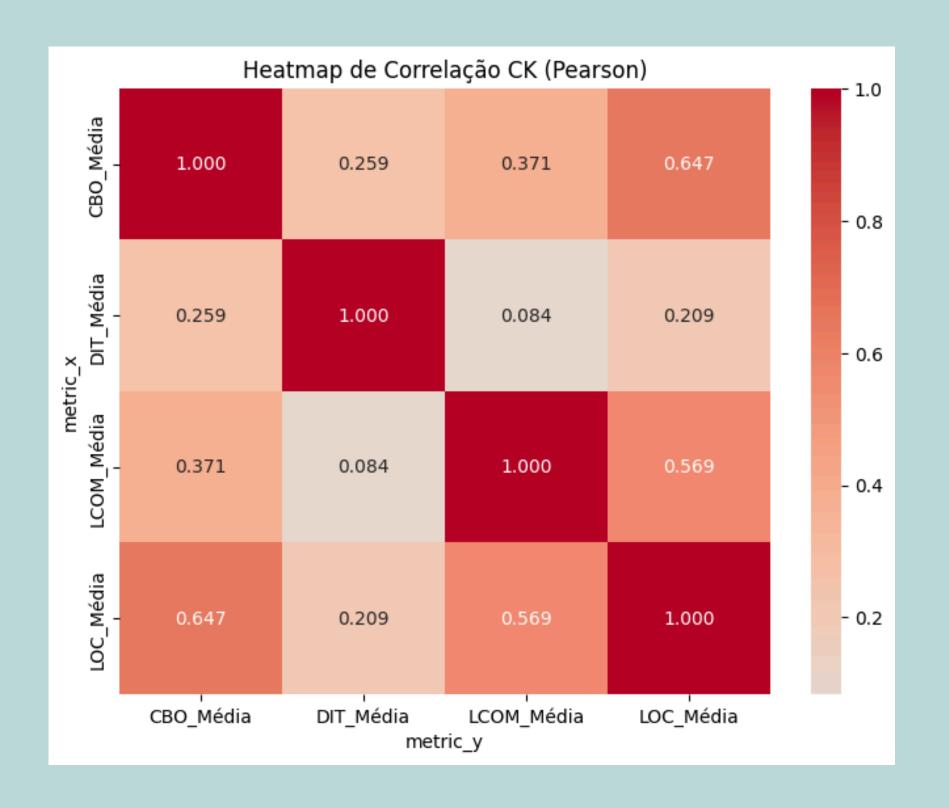


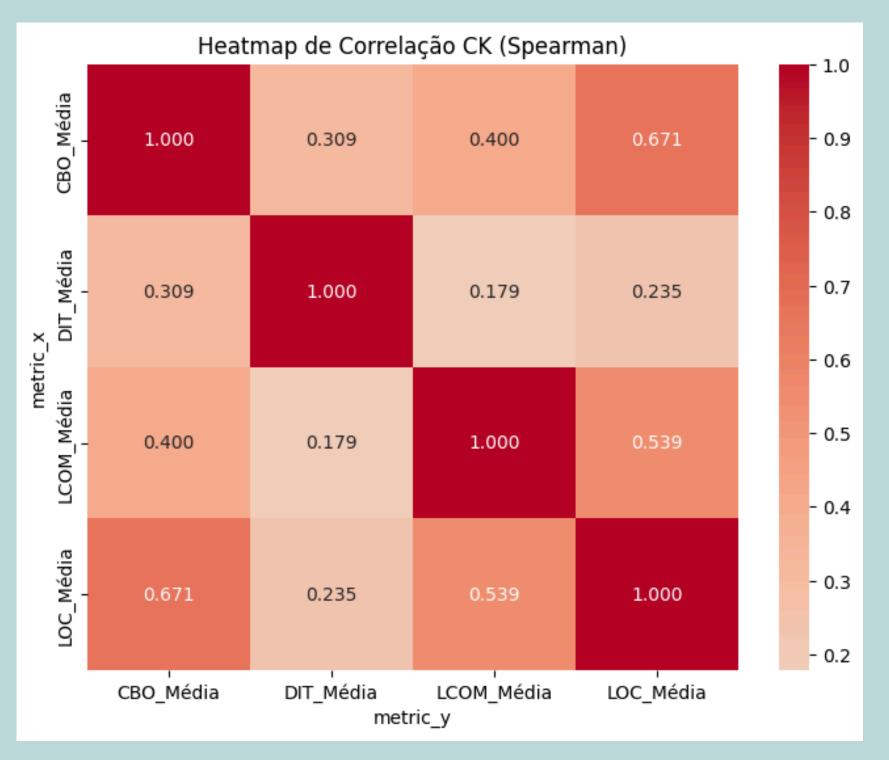


Tamanho vs Qualidade (RQ04)



Correlação entre métricas





Hipóteses vs Resultados

Hipótese	Expectativa	Resultado Observado				
IH01	Repositórios mais populares teriam melhor legibilidade e modularidade.	Parcialmente confirmada → popularidade tem baixa a moderada correlação com modularidade/coesão, mas não garante melhor qualidade interna.				
IH02	Projetos maduros manteriam métricas de qualidade mais consistentes.	X Refutada → idade não mostrou impacto significativo na qualidade.				
IH03	Repositórios com maior atividade apresentariam maior manutenibilidade.	Parcialmente confirmada → releases frequentes associadas a melhores práticas, mas a relação é baixa a moderada.				
IH04	Repositórios maiores apresentariam desafios de manutenção e modularidade.	Confirmada → maior LOC correlaciona negativamente com simplicidade e coesão.				

Obrigado!