**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS**

**INSTITUTO DE CIÊNCIAS EXATAS E INFORMÁTICA  
UNIDADE EDUCACIONAL PRAÇA DA LIBERDADE**

**Bacharelado em Engenharia de Software**

**Camila Silva Romagnoli**

**Leonardo Augusto Moreira Lourenço**

**Octávio Tabai Ribeiro Lage**

**Trabalho Prático Restaurante/Pizzaria**

Belo Horizonte

2022

**Camila Silva Romagnoli**

**Leonardo Augusto Moreira Lourenço**

**Octávio Tabai Ribeiro Lage**

**Trabalho Prático Restaurante/Pizzaria**

Trabalho apresentado na disciplina Laboratório de Experimentação de Software.

Professor: José Laerte Pires Xavier

Belo Horizonte

2022

# 1 INTRODUÇÃO

O relatório apresentado tem como objetivo analisar as características de qualidade de sistemas Java utilizando a ferramenta de métricas de software CK. Essa atividade foi proposta na disciplina de Laboratório de Experimentação de Software.

## 1.1 Hipóteses

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

**Hipótese 1 (Estrelas x CBO):** Repositórios populares tendem a atrair mais colaboradores externos. Dessa forma, espera-se que o acoplamento seja baixo para que os desenvolvedores possam atuar de forma eficiente no mesmo código. Portanto, espera-se que **quanto mais popular é o repositório, menor é o CBO**.

**Hipótese 2 (Estrelas x DIT):** Repositórios populares são colaborativos e precisam ser menos complexos para facilitar a contribuição externa. Portanto, espera-se que **quanto mais popular é o repositório, menor é o DIT**.

**Hipótese 3 (Estrelas x LCOM):** Repositórios populares possuem maior manutenibilidade e reusabilidade para facilitar a distribuição de esforços entre os desenvolvedores e melhorar a qualidade. Portanto, **quanto mais popular é o repositório, menor é o LCOM**.

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

**Hipótese 1 (Idade x CBO):** Repositórios mais maduros possuem classes menos acopladas por terem passado por mais evoluções. Portanto, **quanto maior a maturidade do repositório, menor é o CBO.**

**Hipótese 2 (Estrelas x DIT):** Repositórios mais maduros tendem a possuir mais classes e serem mais complexos. Portanto, **quanto maior a maturidade do repositório, maior é o DIT.**

**Hipótese 3 (Estrelas x LCOM):** Repositórios mais maduros podem ter dificuldades de refatoração conforme o seu crescimento, tornando-se menos coeso. Portanto, **quanto maior a maturidade do repositório, maior é o LCOM.**

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

**Hipótese 1 (Atividade x CBO):** Repositórios mais ativos são atualizados frequentemente e por este motivo espera-se que o código seja melhorado nesse processo. Dessa forma, os índices de qualidade devem refletir essa melhoria. Portanto, espera-se que **quanto maior a atividade, menor é o CBO.**

**Hipótese 2 (Atividade x DIT):** Repositórios mais ativos são atualizados frequentemente e por este motivo espera-se que o código seja melhorado nesse processo. Dessa forma, os índices de qualidade devem refletir essa melhoria. Portanto, espera-se que **quanto maior a atividade, menor é o DIT.**

**Hipótese 3 (Atividade x LCOM):** Repositórios mais ativos são atualizados frequentemente e por este motivo espera-se que o código seja melhorado nesse processo. Dessa forma, os índices de qualidade devem refletir essa melhoria. Portanto, espera-se que **quanto maior a atividade, menor é o LCOM.**

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

**Hipótese 1 (Tamanho x CBO):** Repositórios de menor tamanho possuem menos acoplamento. Portanto, espera-se que **quanto menor o tamanho do repositório, menor é o CBO**.

**Hipótese 2 (Tamanho x DIT):** Repositórios de menor tamanho possuem menos classes e consequentemente o número de heranças é menor. Portanto, espera-se que **quanto menor o tamanho do repositório, menor é o DIT**.

**Hipótese 3 (Tamanho x LCOM):**

Repositórios de menor tamanho são menos complexos e são mais fáceis de refatorar e realizar manutenções. Portanto, **quanto menor o tamanho do repositório, melhor é o LCOM.**

# 2 METODOLOGIA

Para realizar a análise das hipóteses, foi desenvolvido um script em Python que realiza uma consulta em GraphQL dos 1000 repositórios mais populares em Java, que foram clonados e analisados na ferramenta CK.

Após a seleção dos repositórios, foi gerado um arquivo de análise com a mediana das métricas de qualidade CBO, DIT e LCOM de cada um deles. Com os valores das medianas, foram gerados gráficos Scatter Plot para analisar cada uma das hipóteses apresentadas na **seção 1.1**.

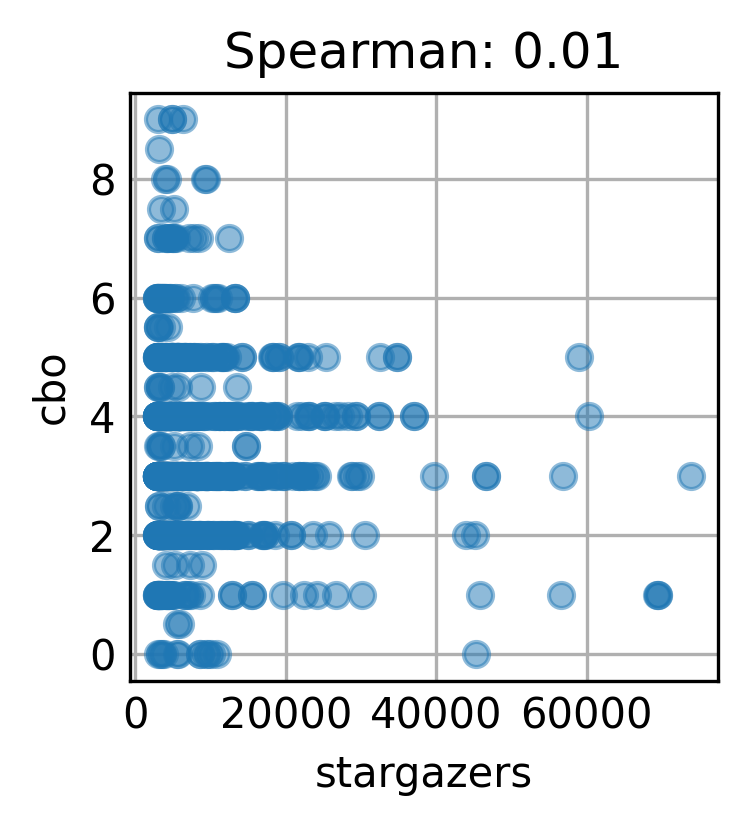
A data da coleta foi durante o período de 22/03/2022 a 26/03/2022. O código-fonte do script e o arquivo csv está contido no repositório: <https://github.com/octaviolage/ES_lab_VI_02>.

# 

# 3 RESULTADOS OBTIDOS E ANÁLISES DOS RESULTADOS

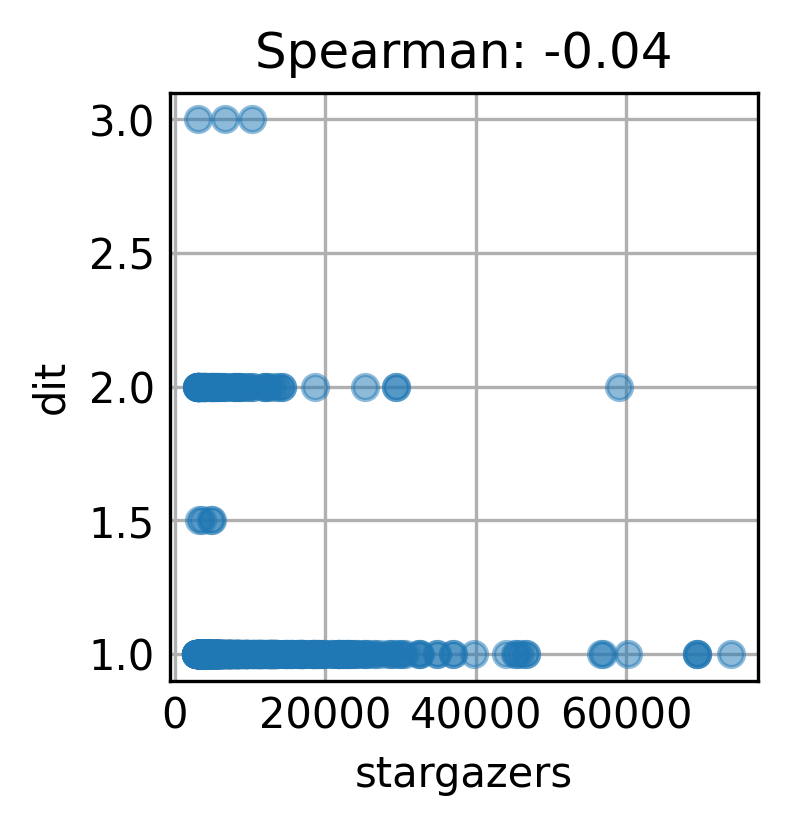
**RQ1. Popularidade**

**Estrelas x CBO**

****

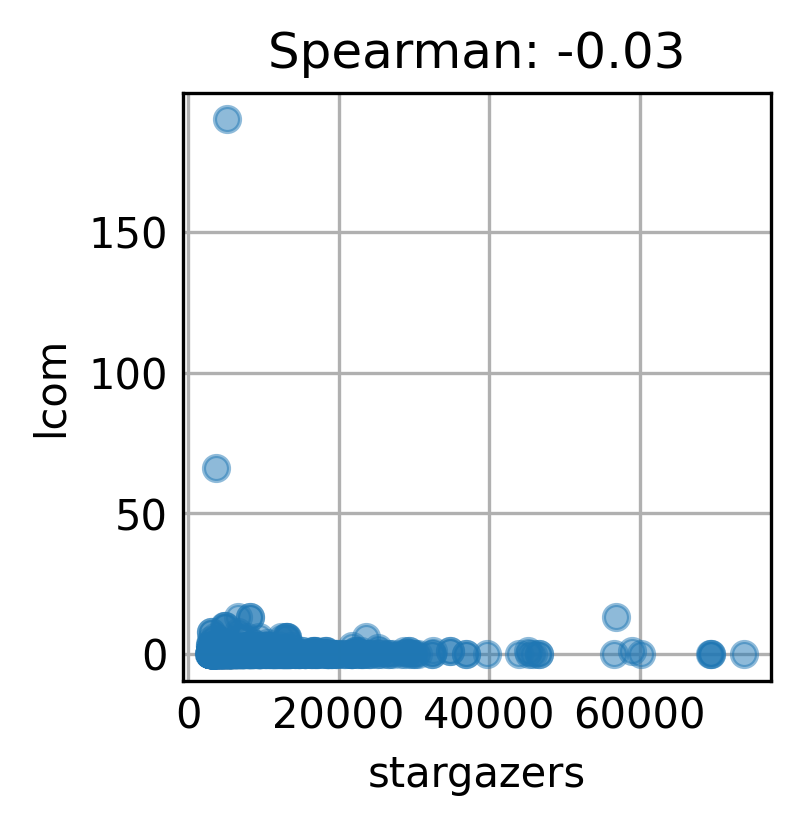
Não foi possível obter uma relação significativa entre os dois aspectos. O coeficiente de Spearman calculado mostra que existe uma correlação positiva muito fraca (0.01) entre os dados. Desta forma, não podemos assumir que existe uma correlação direta e confirmar a hipótese inicial que era: quanto mais popular é o repositório, menor é o CBO.

**Estrelas x DIT**

****

Não foi possível obter uma relação significativa entre os dois aspectos. O coeficiente de Spearman calculado mostra que existe uma correlação negativa muito fraca (-0.04) entre os dados. Desta forma, não podemos assumir que existe uma correlação direta e confirmar a hipótese inicial que era: quanto mais popular é o repositório, menor é o DIT.

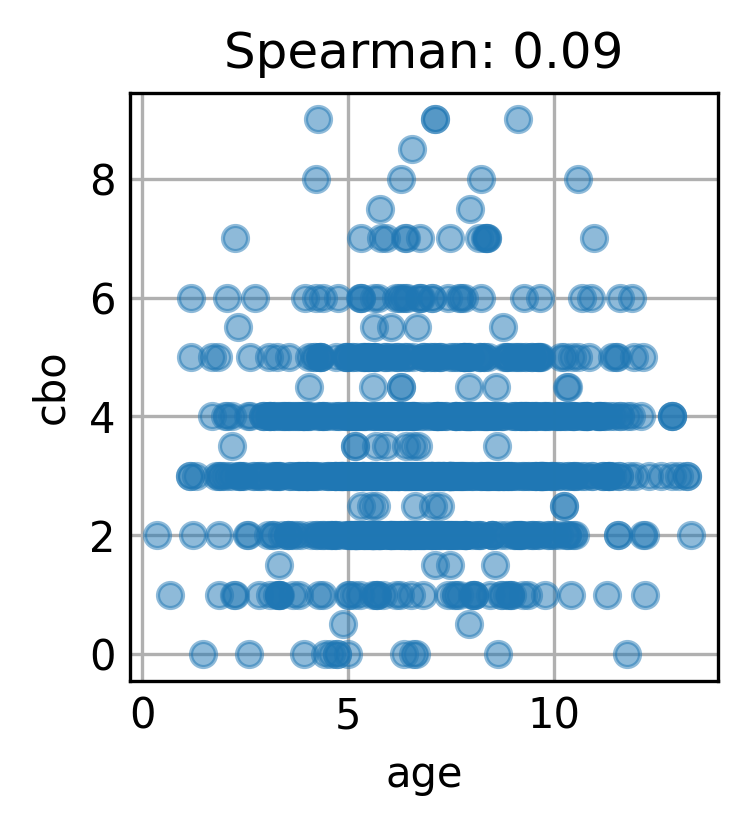
**Estrelas x LCOM**



Não foi possível obter uma relação significativa entre os dois aspectos. O coeficiente de Spearman calculado mostra que existe uma correlação negativa muito fraca (-0.03) entre os dados. Desta forma, não podemos assumir que existe uma correlação direta e confirmar a hipótese inicial que era: quanto mais popular é o repositório, menor é o LCOM.

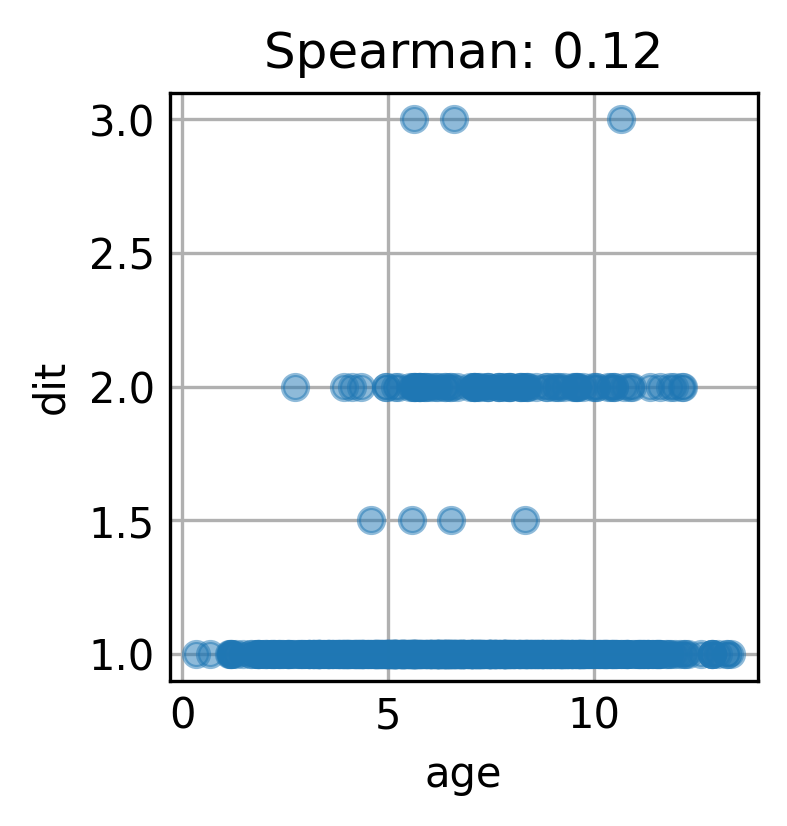
**RQ2. Maturidade**

**Maturidade x CBO**

****

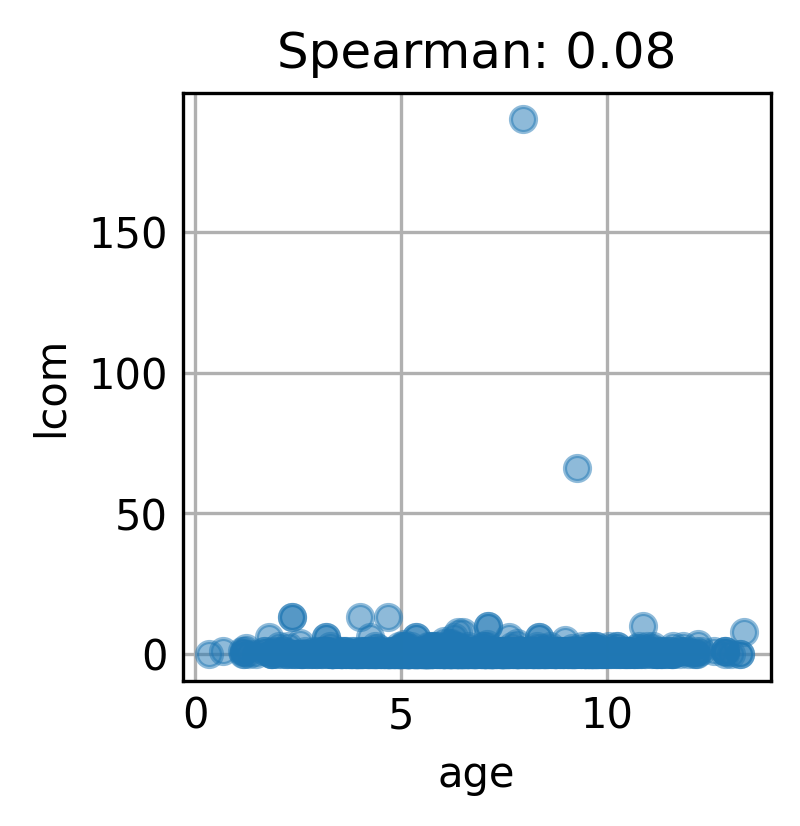
Não foi possível obter uma relação significativa entre os dois aspectos. O coeficiente de Spearman calculado mostra que existe uma correlação positiva muito fraca (0.09) entre os dados. Desta forma, não podemos assumir que existe uma correlação direta e confirmar a hipótese inicial que era: quanto maior a maturidade do repositório, menor é o CBO.

**Maturidade x DIT**

****

Não foi possível obter uma relação significativa entre os dois aspectos. O coeficiente de Spearman calculado mostra que existe uma correlação positiva muito fraca (0.12) entre os dados. Desta forma, não podemos assumir que existe uma correlação direta e confirmar a hipótese inicial que era: quanto maior a maturidade do repositório, maior é o DIT.

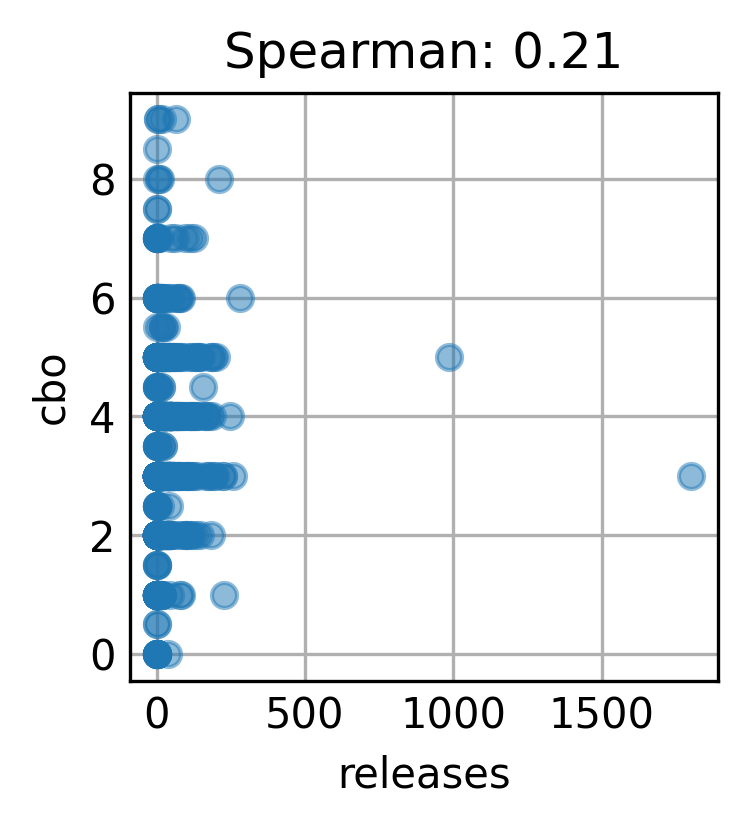
**Maturidade x LCOM**



Não foi possível obter uma relação significativa entre os dois aspectos. O coeficiente de Spearman calculado mostra que existe uma correlação positiva muito fraca (0.08) entre os dados. Desta forma, não podemos assumir que existe uma correlação direta e confirmar a hipótese inicial que era: quanto maior a maturidade do repositório, maior é o LCOM.

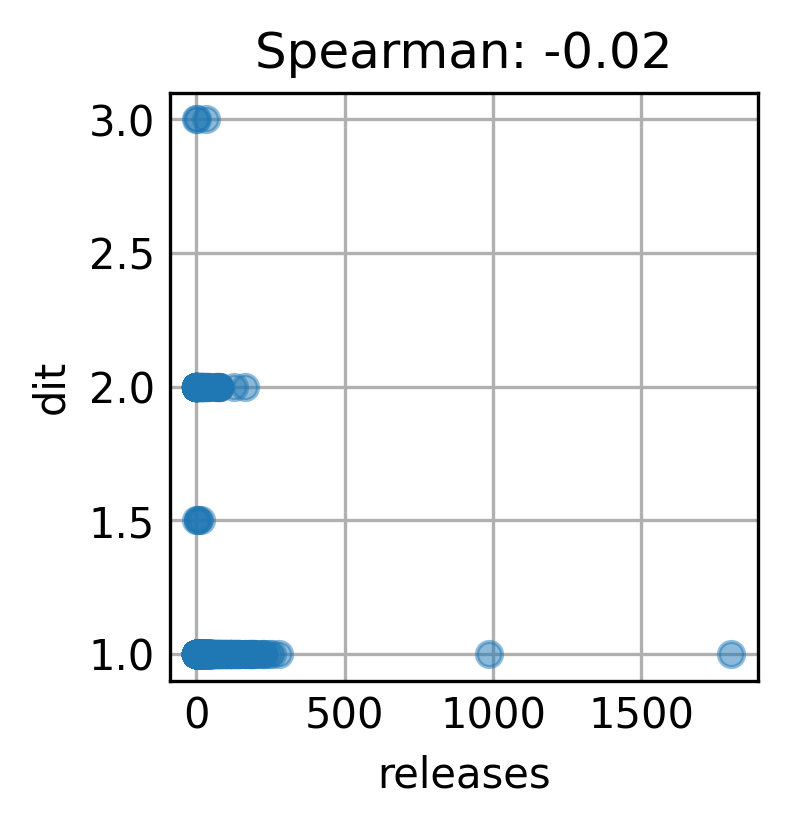
**RQ3. Atividade**

**Atividade x CBO**

****

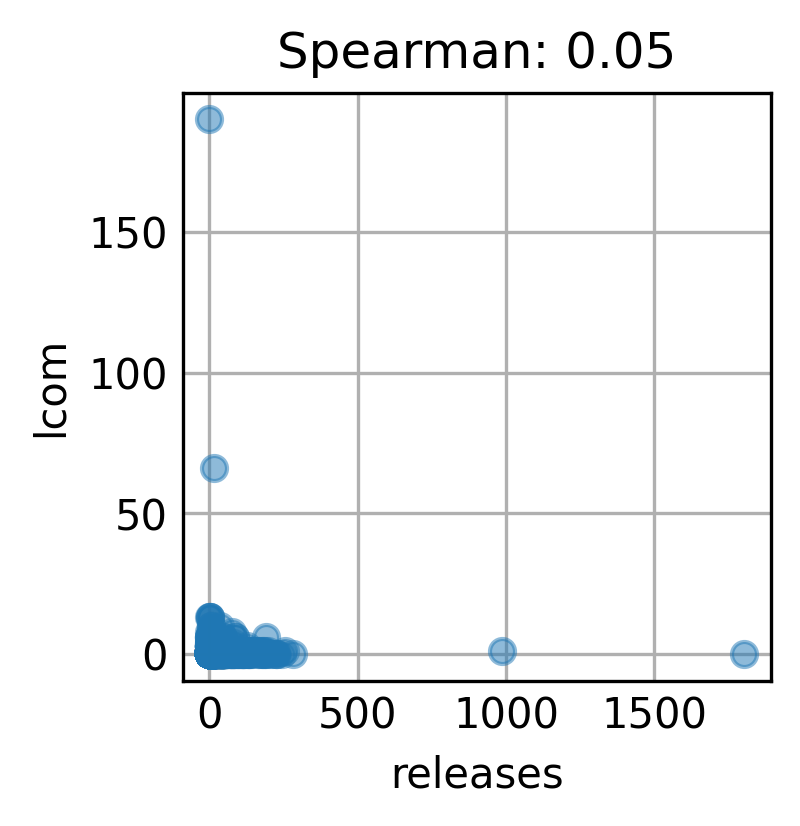
Foi possível obter uma pequena relação entre os dois aspectos. O coeficiente de Spearman calculado mostra que existe uma correlação positiva fraca (0.21) entre os dados. Desta forma, podemos negar a hipótese inicial que era: Quanto maior a atividade, menor o CBO.

**Atividade x DIT**

****

Não foi possível obter uma relação significativa entre os dois aspectos. O coeficiente de Spearman calculado mostra que existe uma correlação negativa muito fraca (-0.02) entre os dados. Desta forma, não podemos assumir que existe uma correlação direta e confirmar a hipótese inicial que era: quanto maior a atividade, menor é o DIT.

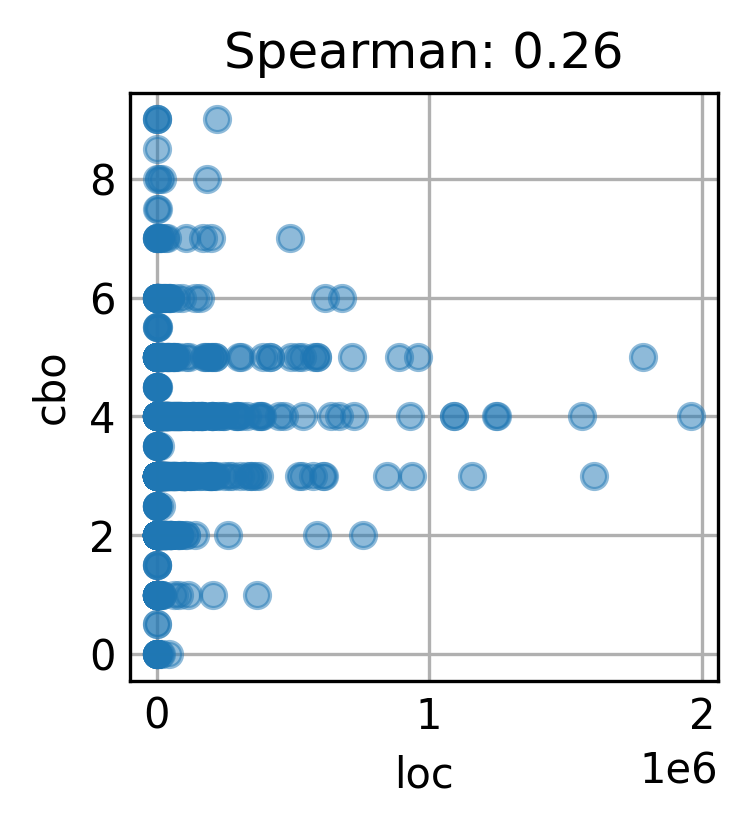
**Atividade x LCOM**



Não foi possível obter uma relação significativa entre os dois aspectos. O coeficiente de Spearman calculado mostra que existe uma correlação positiva muito fraca (0.05) entre os dados. Desta forma, não podemos assumir que existe uma correlação direta e confirmar a hipótese inicial que era: quanto maior a atividade, menor é o LCOM.

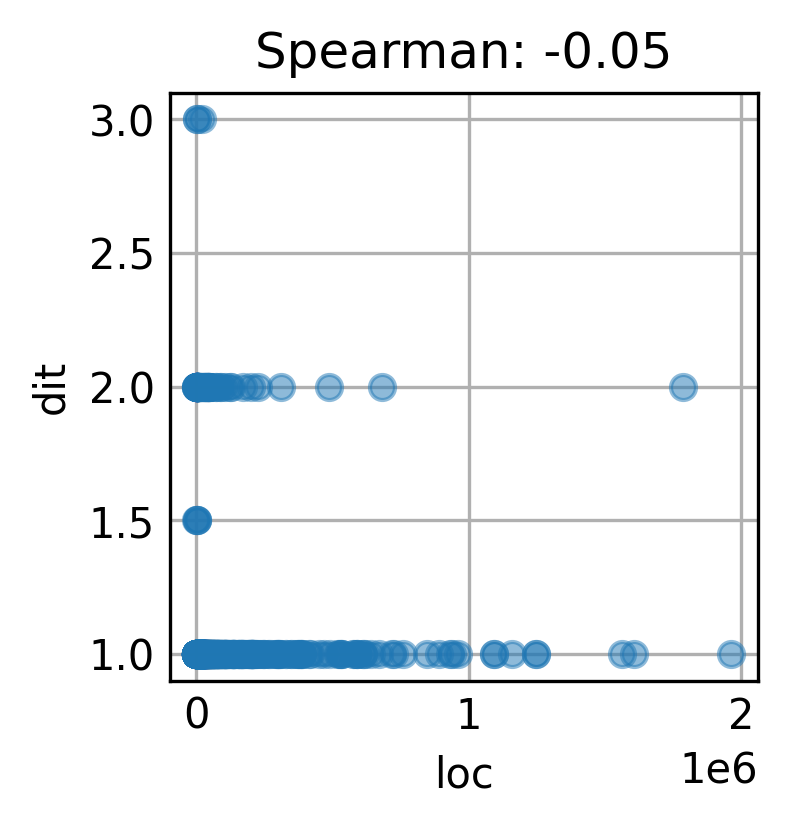
**RQ4. Tamanho**

**Tamanho x CBO**

****

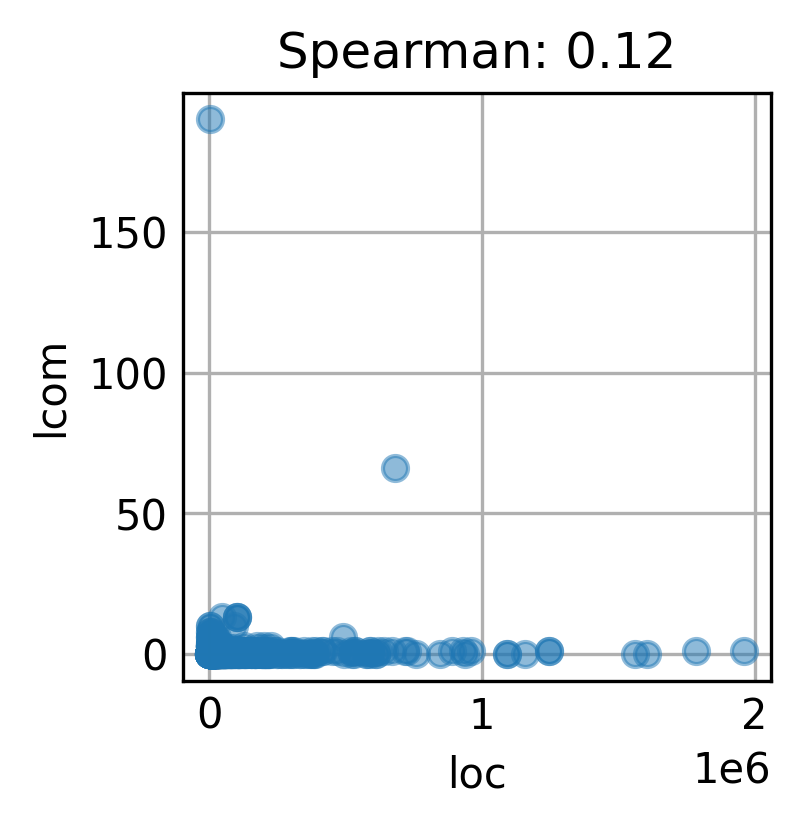
Foi possível obter uma pequena relação entre os dois aspectos. O coeficiente de Spearman calculado mostra que existe uma correlação positiva fraca (0.26) entre os dados. Desta forma, podemos afirmar a hipótese inicial que era: Quanto menor o tamanho, menor o CBO.

**Tamanho x DIT**

****

Não foi possível obter uma relação significativa entre os dois aspectos. O coeficiente de Spearman calculado mostra que existe uma correlação negativa muito fraca (-0.05) entre os dados. Desta forma, não podemos assumir que existe uma correlação direta e confirmar a hipótese inicial que era: quanto menor o tamanho do repositório, menor é o DIT.

**Tamanho x LCOM**



Não foi possível obter uma relação significativa entre os dois aspectos. O coeficiente de Spearman calculado mostra que existe uma correlação positiva muito fraca (0.12) entre os dados. Desta forma, não podemos assumir que existe uma correlação direta e confirmar a hipótese inicial que era: quanto menor o tamanho do repositório, melhor é o LCOM.

**Conclusão**

Os repositórios pesquisados raramente nos mostraram algum padrão com força o suficiente para que embasassem ou neguem as hipóteses levantadas, o que nos levou a ver a diversidade dos mesmos, com relação a implementação e qualidade.

É um pouco frustrante termos mais dúvidas do que certezas após uma busca que nos atiçou a criar possíveis padrões, entretanto, é um desafio para que melhoremos ainda mais em nossas próximas minerações.