

Membros: Anderson Barbosa & Henrique Alberone

### Laboratório 3 - Avaliando o Buzz de issues do GitHub no StackOverflow

#### Hipótese:

Há uma relação visível das issues dos repositórios mais populares do GitHub no StackOverflow

#### Questões analisadas:

- As linguagens com mais issues no GitHub são também as com mais issues no StackOverflow?
- As linguagens mais populares no GitHub são as com mais issues no StackOverflow?
- O número de menções de um repositório popular do GitHub no StackOverflow está diretamente relacionado ao número de issues abertas no GitHub?

#### Metodologia

Para responder às questões, utilizamos a api do GitHub e do StackOverflow para coletar os dados usando um script para automatizar o processo e retornar os seguintes dados:

- Quantidade de issues abertas no GitHub
- Linguagem primária dos repositórios do GitHub
- Quantidade de issues abertas no StackOverflow por linguagem
- Quantidade de menções aos repositórios populares no StackOverflow

#### Resultados Obtidos

Encontramos os seguintes dados com as buscas realizadas:

Tabela de issues por linguagem do StackOverflow

Linguagem	Perguntas
JavaScript	2034683
Java	3869922
Python	1866121
C	18341578
Swift	318593
PHP	1478417
Kotlin	44525

## Repositórios do GitHub & issues open vs Menções StackOverflow

Repositório	Issues Open	Menções SO
freeCodeCamp/freeCodeCamp	193	32
996icu/996.ICU	36	0
vuejs/vue	316	16292
facebook/react	492	72726
EbookFoundation/free-programming-books	28	0
tensorflow/tensorflow	3422	11866
twbs/bootstrap	292	43672
sindresorhus/awesome	14	43777
getify/You-Dont-Know-JS	73	0
jwasham/coding-interview-university	31	0

Issues Open	Menções SO
Media: 489.7	Media: 18836.5
Mediana: 242.5	Mediana: 14079

Pearson: 0.8233529804
-----------------------

## Discussão

Observando os dados da tabela das linguagens mais populares no StackOverflow e comparando aos dados retornados dos repositórios mais populares do GitHub, comprovamos que as linguagens principais dos repositórios populares estão também entre as mais buscadas no StackOverflow.

Com relação a análise utilizando o coeficiente de Pearson nos 10 primeiros repositórios, acabamos tendo como resultado uma relação forte entre ambos.

Porém, com relação às linguagens mais populares do GitHub serem as mais populares no StackOverflow, nossa busca se mostrou falha, pois apesar de C ter uma quantidade notável de issues no StackOverflow, não houve como diferenciar na query do SO se era C, C++, C#. Apenas notamos que JavaScript foi a mais recorrente dentre as 100 primeiras do GitHub e seu número de issues no StackOverflow era notável também, mas não chegamos em nenhuma conclusão.

#### **Ameaças à validade do estudo**

- Não conseguir diferenciar linguagens como C, C++, C# nas buscas do StackOverflow
- Nem todos os repositórios foram encontrados no StackOverflow
- React & ReactNative podem alterar os valores citados