Análise de READMEs

Autores

- [Hilton Pintor](github.com/hpbl)
- [Bruno Melo](github.com/bhlvm)

Motivação

"README files play an essential role in shaping a developer's first impression of a software repository and in documenting the software project that the repository hosts. Yet, we lack a systematic understanding of the content of a typical README file as well as tools that can process these files automatically."

Objetivos

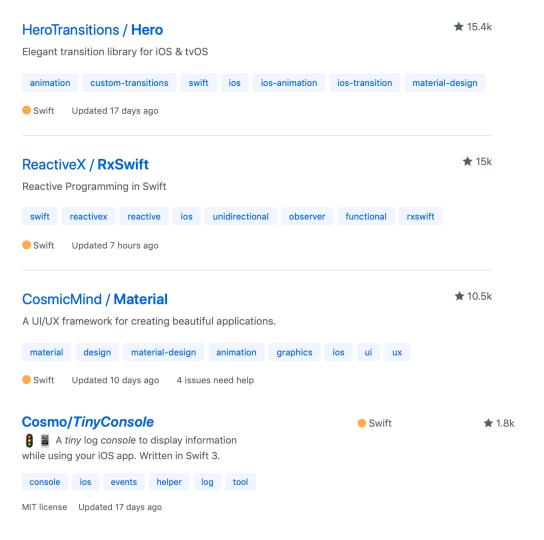
1. Mapear o estado da arte da escrita de READMEs para bibliotecas/ frameworks iOS escritos em Swift.

2. Prover um guia para desenvolvedores de boas práticas para escrita de READMEs.

Coleta de dados

- Top 100 repositórios escritos em Swift, com tag iOS (GitHub API)
- Filtro de bibliotecas e frameworks (Manual)
- 85 READMEs.md

Coleta de dados



Coleta de perguntas

- 10 desenvolvedores iOS consultados
- Quais informações/perguntas seriam de interesse?
- 44 respostas

Categorização das perguntas

1. Contributing

2. Testing

3. Status (Build/CI)

4. Package manager

5. Example project

6. Licensing/Ownership

7. Purpose

8. Versioning/Compatibility

9. Installation

10. README file structure

11. Formatting

12. Usage

13. Dependency

Selecting analysis

We selected the questions that seemed most relevant/feasible for our initial analysis, they were:

README file structure

- 27. What are the most common section titles? V
- 41. Size of the README files. V
 - Does the order of the sections follow a common order?

Package management

28. What are the most adopted package management tools (Cocoapods, Catharge, Swift Package Manager, direct installation?) 🔽

Contributing:

10. Is there a contribution guide?

Licensing/Ownership

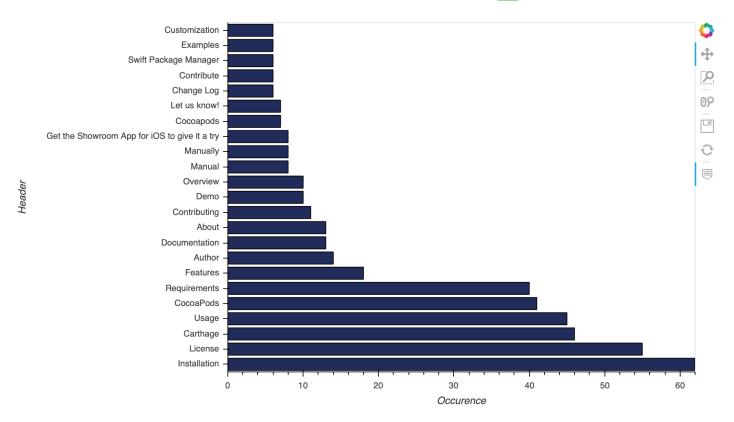
- 11. What license does it use? V
 - Do maintainers have more then one popular repository?

Status

- Is the building status present? ✓
 - Is it passing? ✓

README file structure

27. What are the most common section titles?



README file structure

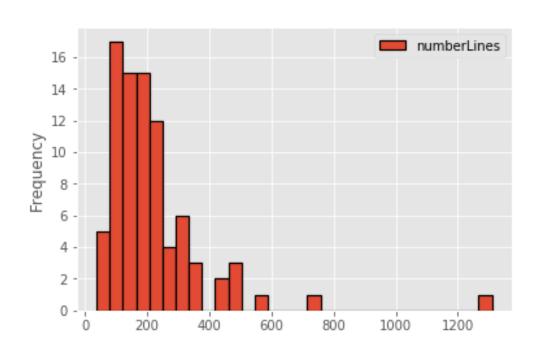
27. What are the most common section titles?



41. Size of the README files.



Number of lines



	number Lines	
ount	85.000000	

numbert ince

count	85.000000
mean	219.741176
std	171.922404
min	38.000000
25%	122.000000
50%	178.000000
75%	249.000000

max

file	num	Lines
11110	HUMIT	

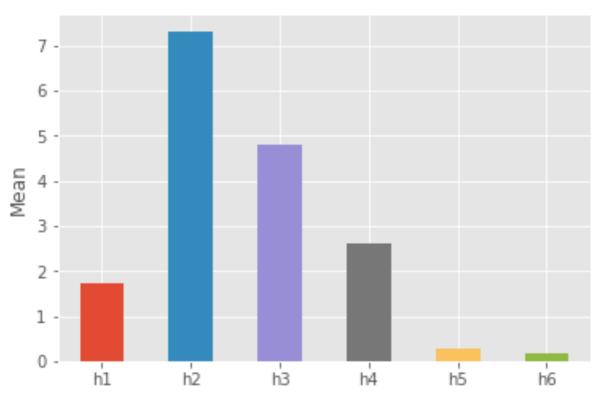
Largest	Eureka.md	1309
Smallest	ClassicKit.md	38

1309.000000

41. Size of the README files.

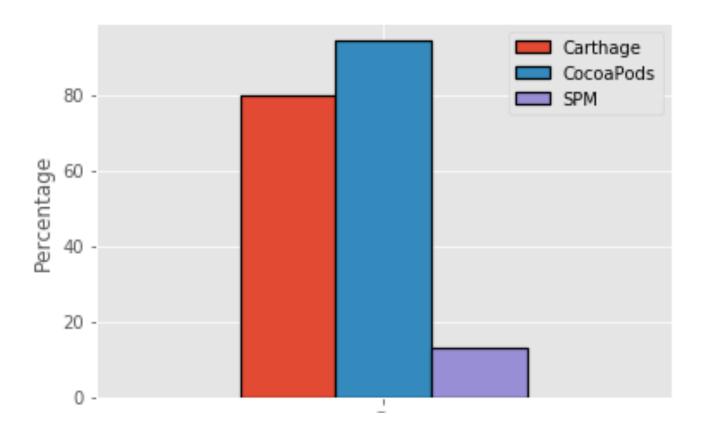


Number of sections



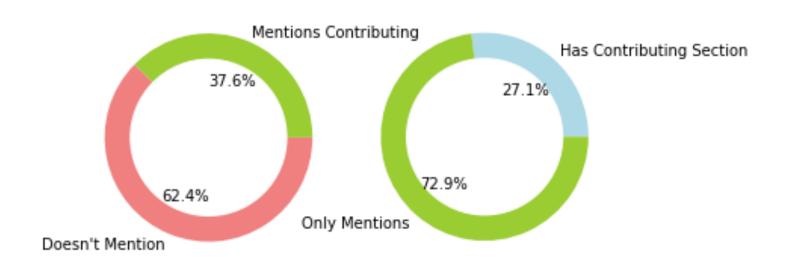
Package management V

28. What are the most adopted package management tools (Cocoapods, Catharge, Swift Package Manager, direct installation?)



Contributing:

10. Is there a contribution guide?



License/Ownership

11.What license does it use?



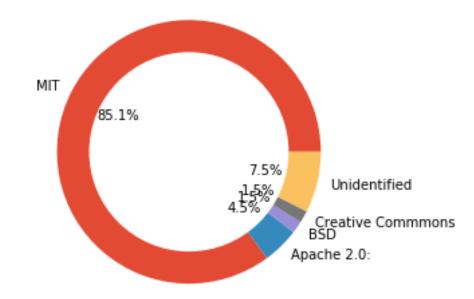
Popular open-source licenses list

MIT: 57

Apache 2.0: 3

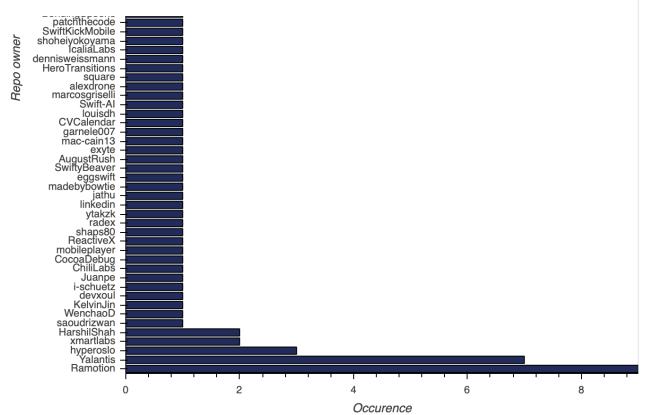
BSD: 1

Creative Commmons: 1 Total Mentions: 67



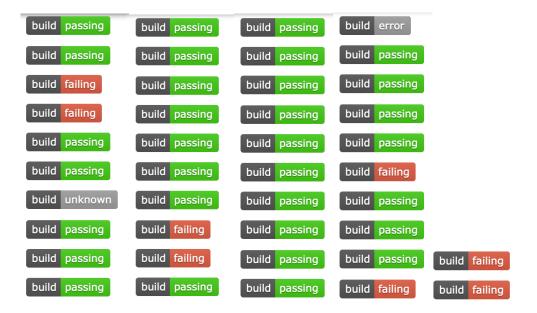
License/Ownership

- Do maintainers have more then one popular repository? V



Status

- Is the building status present? ✓
 - Is it passing?



49.41% of the repositories present a travis build status

Trabalhos futuros

- 1. Desenvolver mais análises, e de forma mais profunda.
- 2. Entrevistar mais desenvolvedores.
- 3. Utilizar ferramentas de análise de texto para categorização do conteúdo.
- 4. Desenvolver ferramenta e disponibilizar para desenvolvedores
- 5. Continuar pesquisa (TG)

Obrigado

Repo

- https://github.com/if1015-datascience-ufpe/2018-2-projeto-bubads

Autores

- [Hilton Pintor](github.com/hpbl)
- [Bruno Melo](github.com/bhlvm)