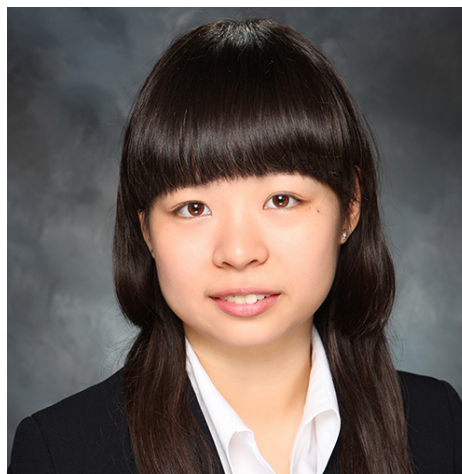
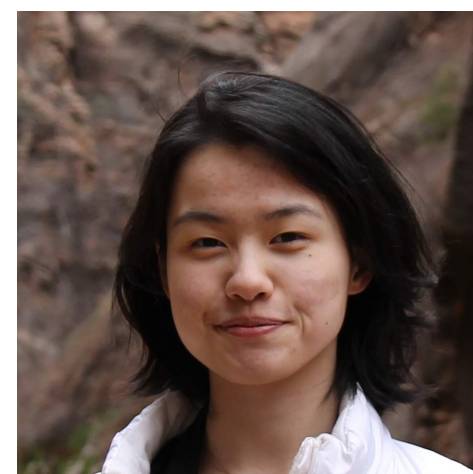


The Signals that Potential Contributors Look for When Choosing Open-source Projects



Sophie Qiu



Lily Li



Susmita Padala



Anita Sarma



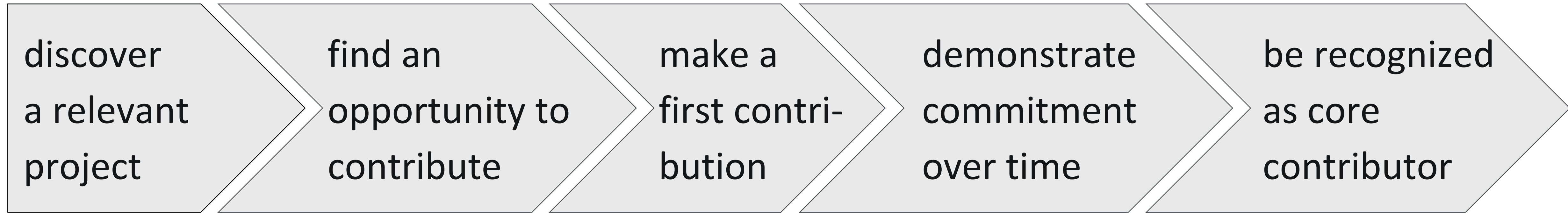
Bogdan Vasilescu

Goal: Increase participation in Open Source

- To sustain open source projects that need constant supply of effort
- To provide opportunities to learn / develop skills / showcase talent
- To increase diversity
- ...

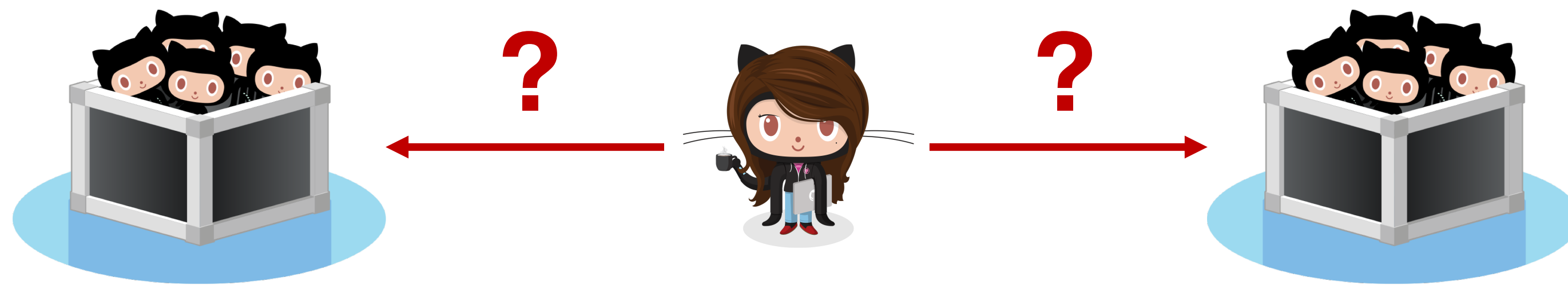
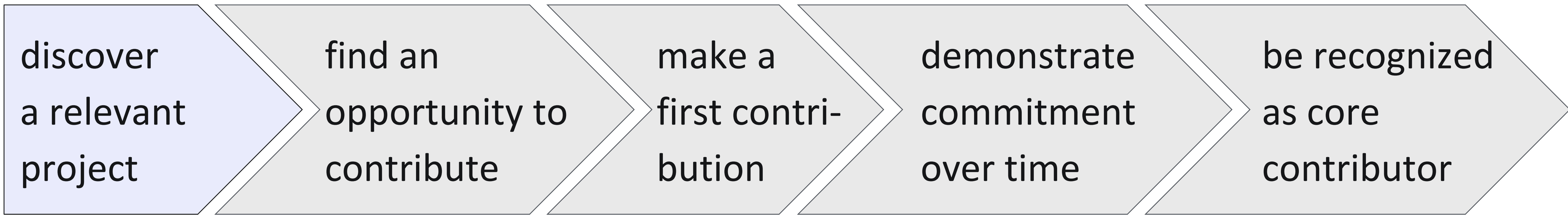


Participation in Open Source has a long history of scholarship



Yet, little is known about how people choose which project to contribute to

Especially for people who contribute for fun, to learn, etc.



**Project discovery
is especially relevant on
transparent platforms
like GitHub**

GitHub users rely on signals for many inferences

(Dabbish et al., 2012)

Pioneer / charts

Watch 1

Star 2

Fork 2

<> Code

Issues 17

Pull requests 1

Projects 6

Wiki

Security

Insights

Data Visualization Charting library for Angular deployments. <https://charts.pioneercode.com>

angular

data-visualization

d3

charts

pioneer-charts

pioneer-code

667 commits

2 branches

20 releases

2 contributors

MIT

Branch: master

New pull request

Create new file

Upload files

Find file

Clone or download

Update angular deps

Latest commit e1f3 on Jun 18

.github	Add meta	2 years ago
.vscode	Update document project	4 months ago
apps	Fix table header bug	4 months ago
build	Fox docs build	2 years ago

GitHub users rely on signals for many inferences

(Trockman et al., 2018)

Status



What's included

Within the download you'll find the following directories and files, logically grouping common assets and providing both compiled and minified variations. You'll see something like this:

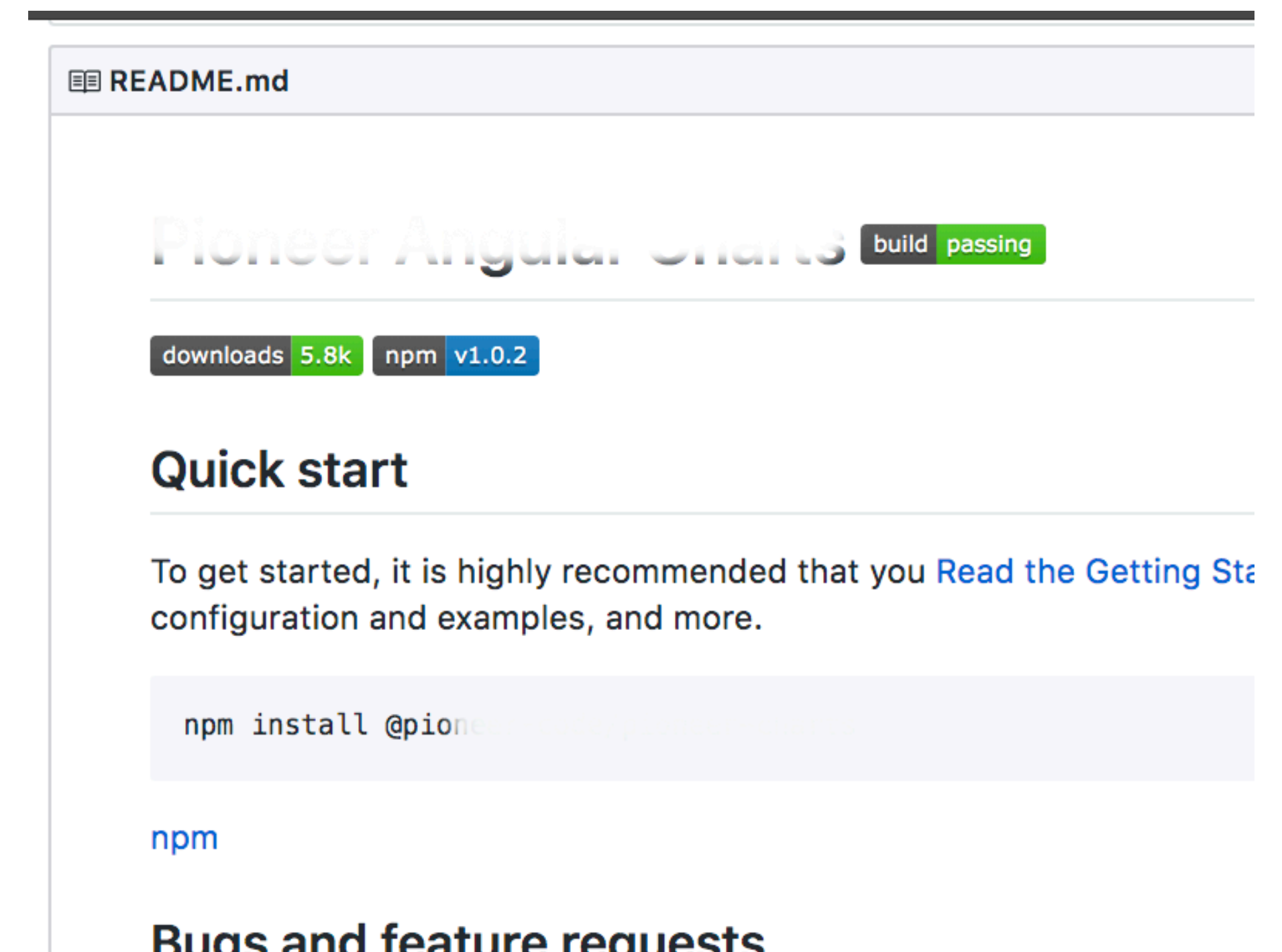
**How do people use signals
when choosing a GitHub
project to contribute to?**

Many signals used when choosing projects

Quality	README
	Contributing guideline
Community	Impolite language
	Issue template
Popularity	Stars
	Recent commits

Many signals used when choosing projects

Quality	README
	Contributing guideline
Community	Impolite language
	Issue template
Popularity	Stars
	Recent commits



Many signals used when choosing projects

Quality	README
	Contributing guideline
Community	Impolite language
	Issue template
Popularity	Stars
	Recent commits

Contributing

Please read through our [contributing guidelines](#). Included are development.

Moreover, if your pull request contains JavaScript patches or f should conform to the [Code Guide](#), maintained by [M](#).

Editor preferences are available in the [editor config](#) for easy use <https://editorconfig.org/>.

Many signals used when choosing projects

Quality	README
	Contributing guideline
Community	Impolite language
	Issue template
Popularity	Stars
	Recent commits

build	YES! FUCK YOU WEBPACK
dist	Precompile assets
docs	Experimenting with component data tables
public	Rederubfne
src	Move redirects again

Many signals used when choosing projects

Quality	README
	Contributing guideline
Community	Impolite language
	Issue template
Popularity	Stars
	Recent commits

I'm submitting a...

Please search GitHub for a similar issue or PR before submitting

<pre><code>

☐ Regression (a behavior that used to work and stopped working in a new release)

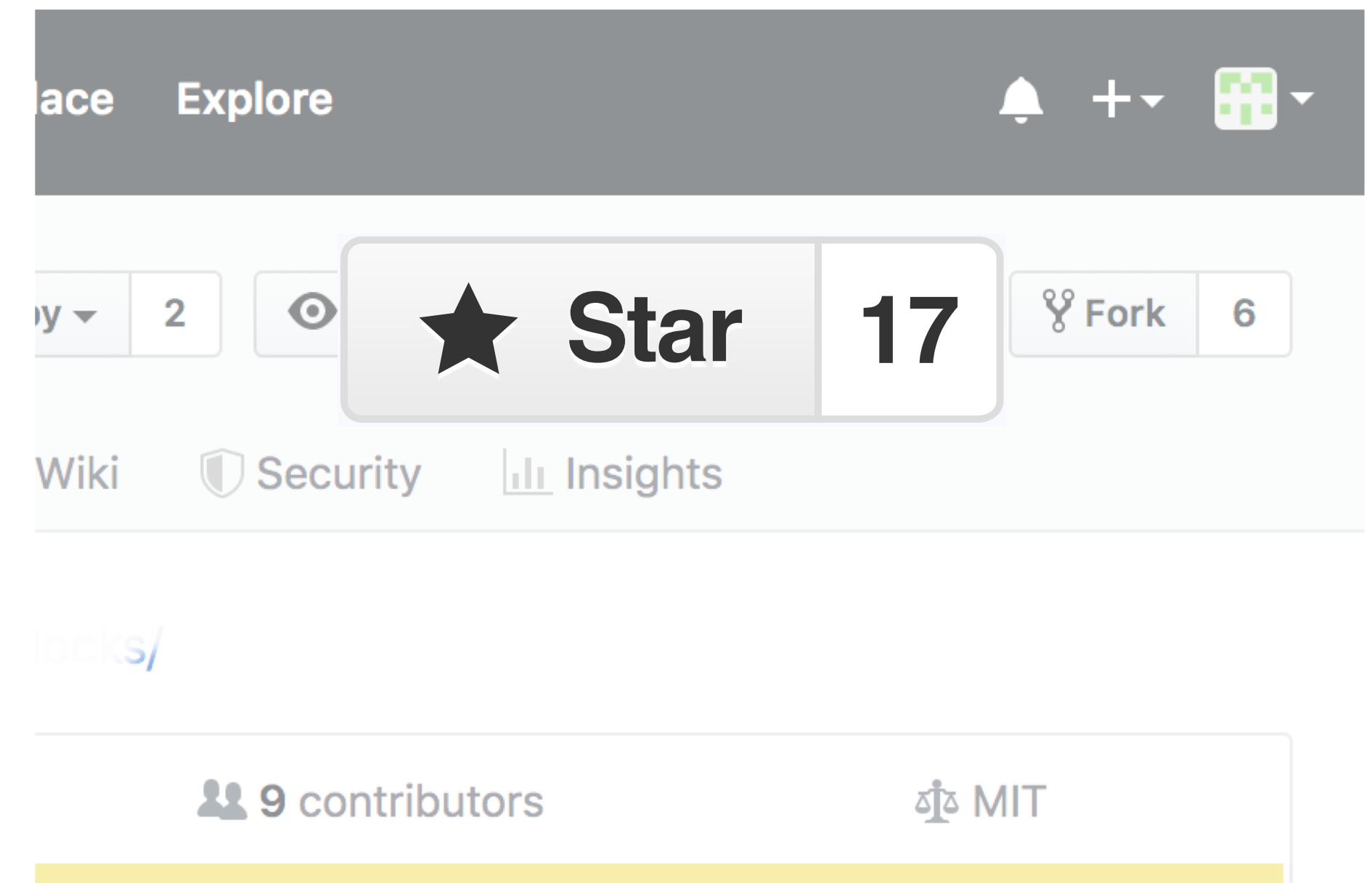
☐ Bug report

☐ Feature request

☐ Documentation issue or request

Many signals used when choosing projects

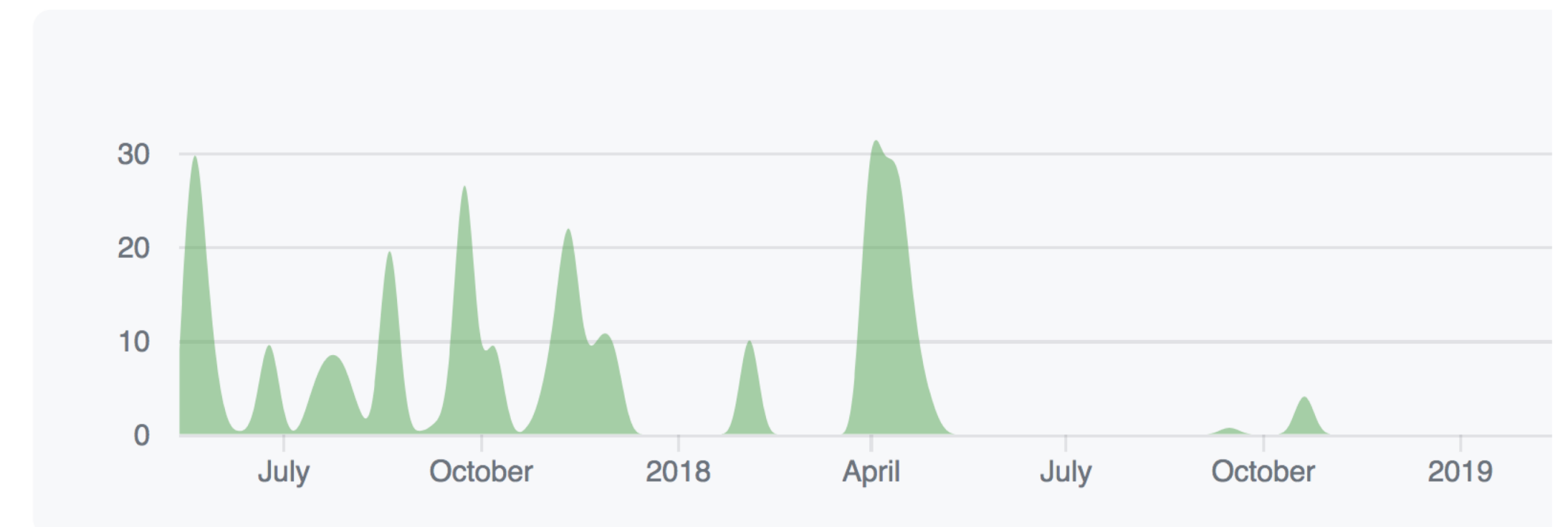
| | |
|------------|------------------------|
| Quality | README |
| | Contributing guideline |
| Community | Impolite language |
| | Issue template |
| Popularity | Stars |
| | Recent commits |



Many signals used when choosing projects

| | |
|------------|------------------------|
| Quality | README |
| | Contributing guideline |
| Community | Impolite language |
| | Issue template |
| Popularity | Stars |
| | Recent commits |

Contributions to master, excluding merge commits



Methodology

A mixed-method study design - Qualitative



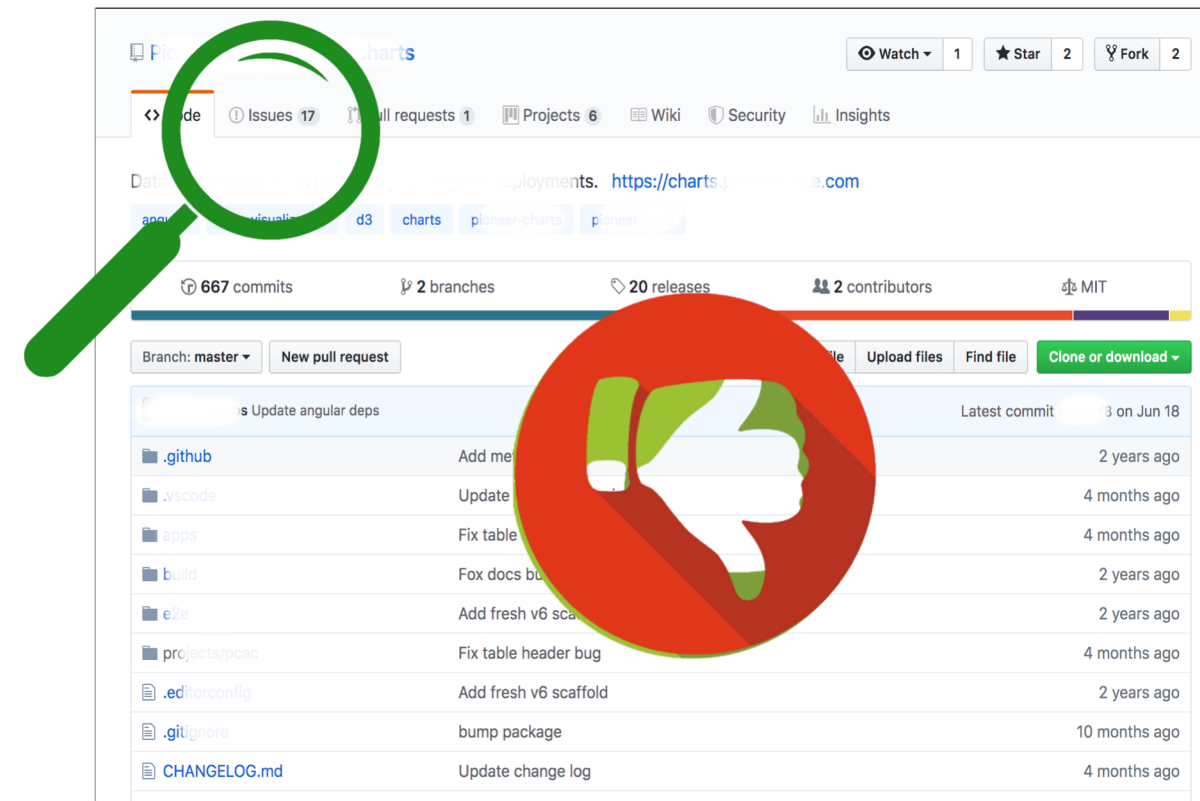
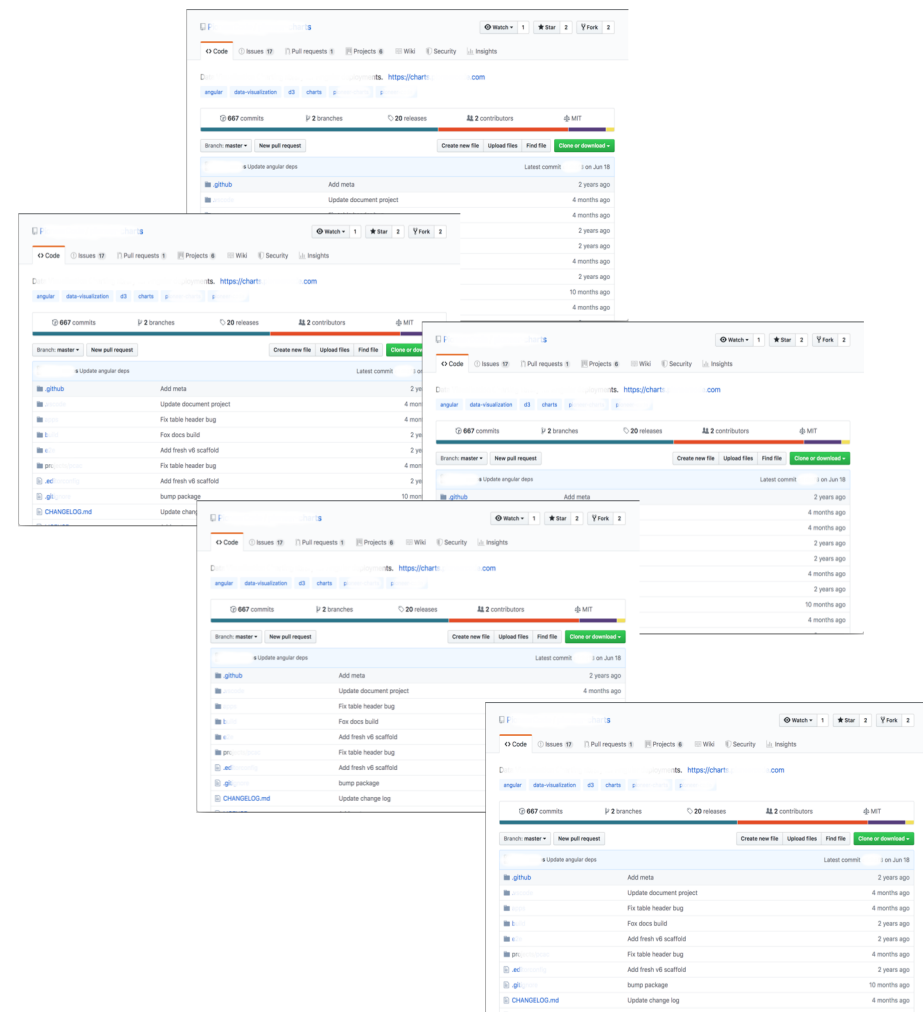
Interviews with
15 developers



Qualitative
analysis



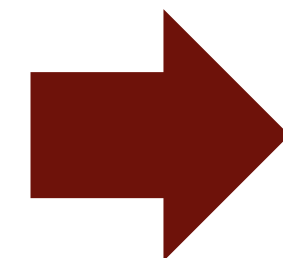
Qualitative: Review and recommend projects



Jane,
Bachelor degree in
Computer Science
First-time contributor to npm

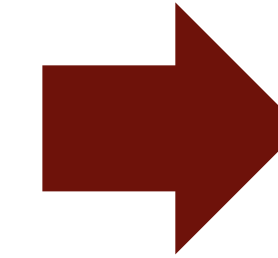
Pre-selected projects

Shown a set of NPM
projects' GitHub pages



Think-aloud interviews

Describe how they
evaluate each project



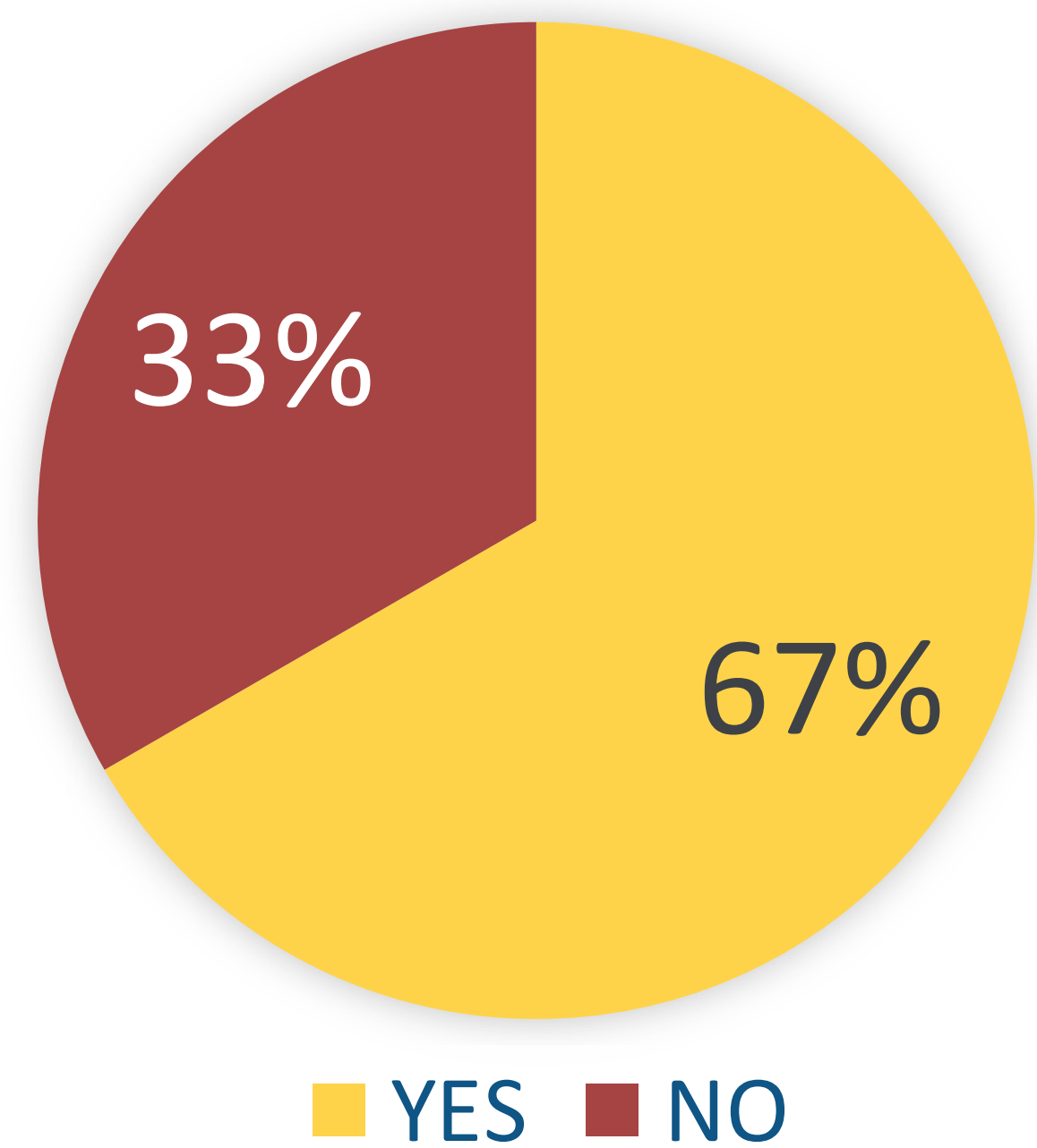
Persona

Decide for each project,
if they would
recommend it to Jane

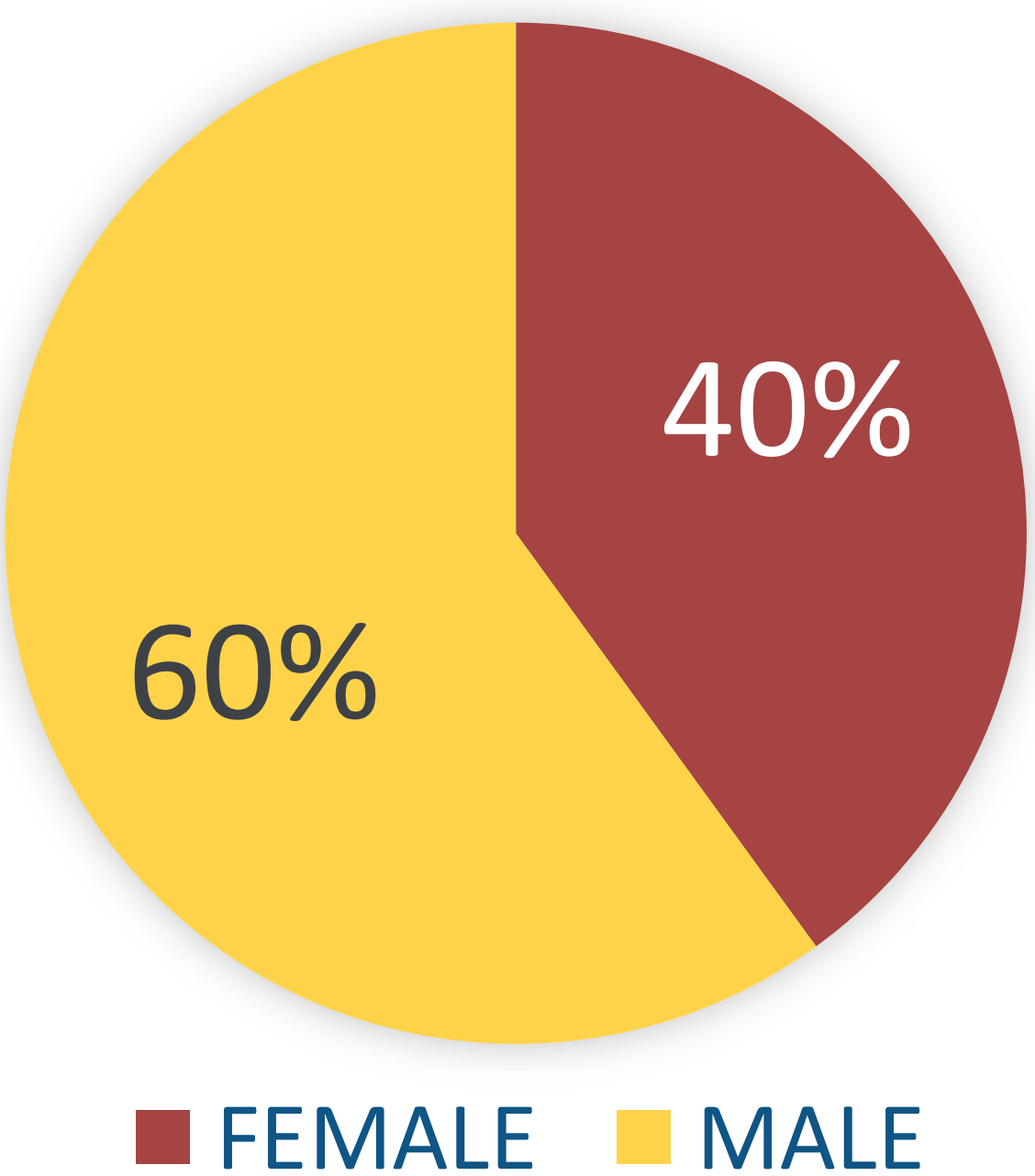


Interviewees came from diverse backgrounds

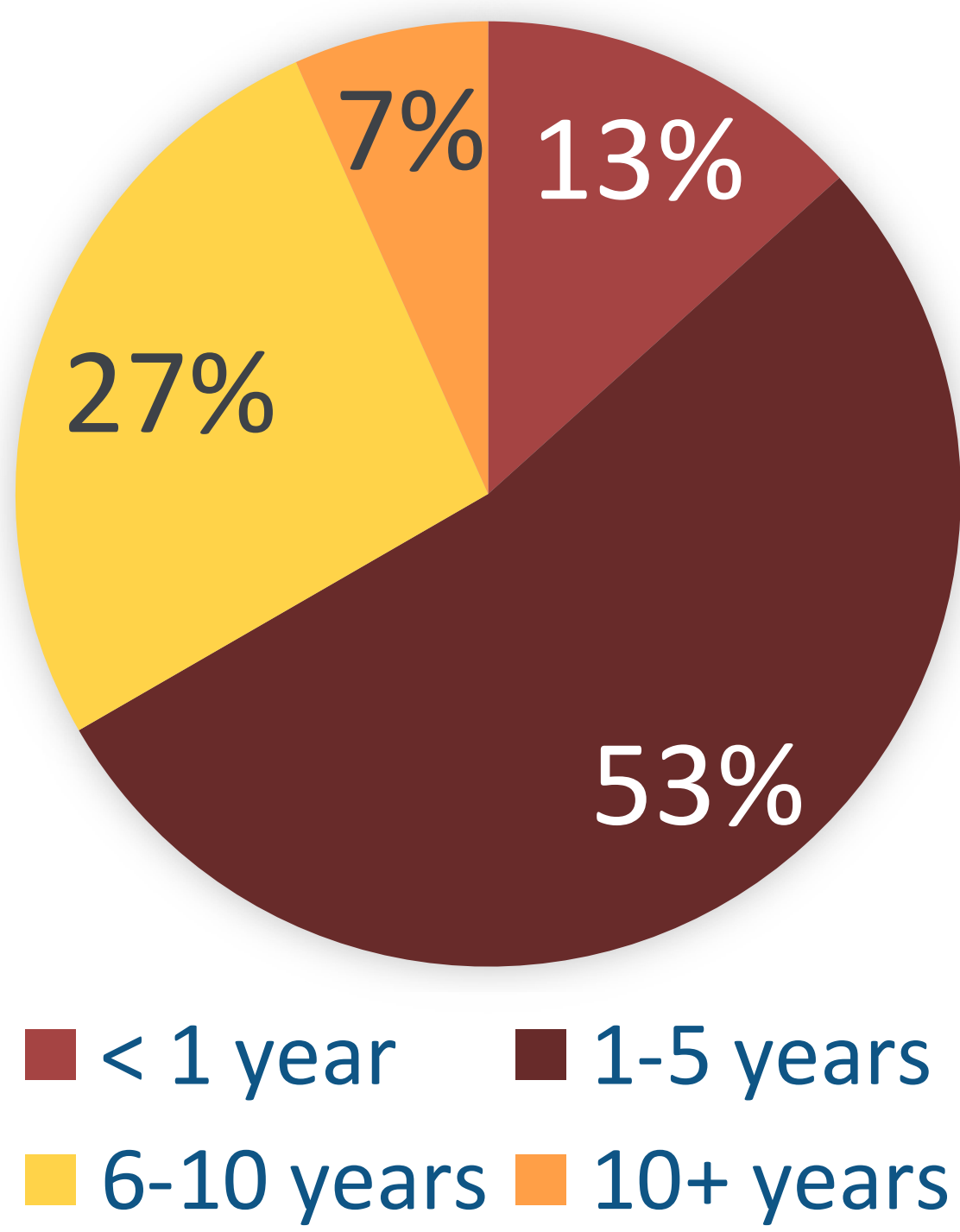
Full-time developers



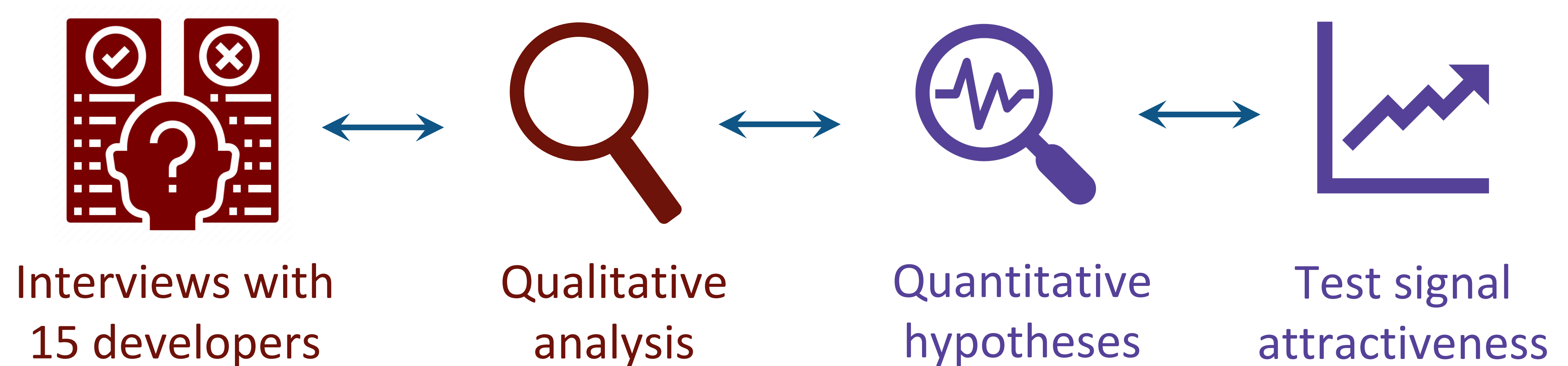
Gender



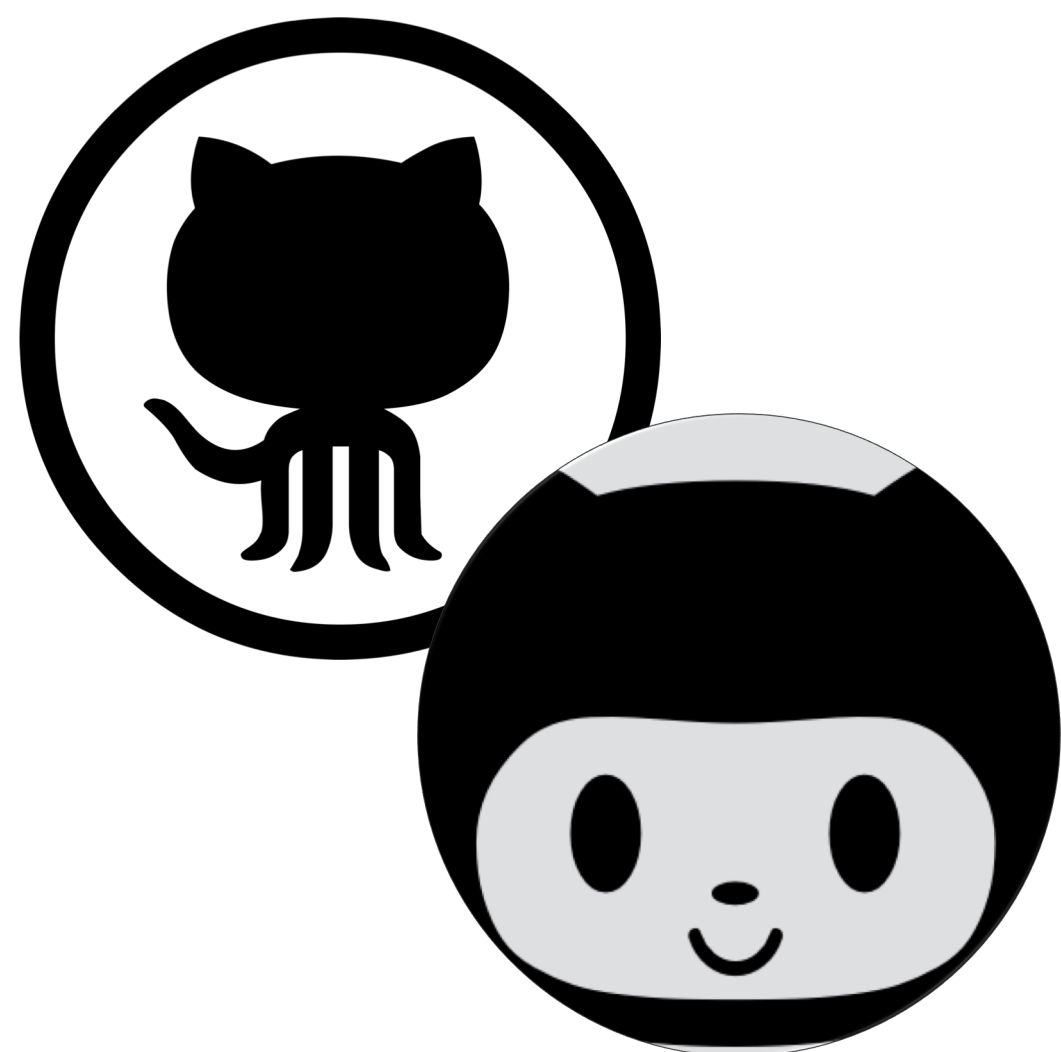
Years of OSS experience



A mixed-method study design - Quantitative

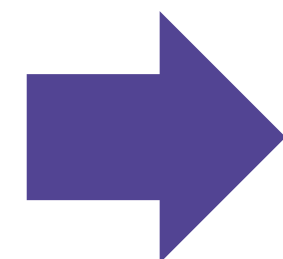


Quantitative: GitHub data mining



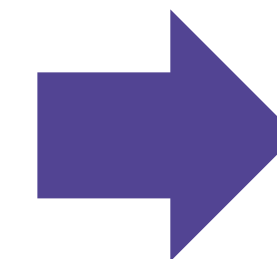
Mining trace data

~10,000
npm packages from
GitHub and GHTorrent



Operationalization

14 quantitative
measurements





Logistic regression

Associations between
signals and newcomers

Result highlights:

Many signals used when choosing projects

| | | Interviews | Logistic Regression |
|------------|------------------------|---|---|
| Quality | README |  |  |
| | Contributing guideline |  |  |
| Community | Impolite language |  |  |
| | Issue template |  |  |
| Popularity | Stars | |  |
| | Recent commits |  |  |

Three main classes of signals

Quality

Community

Popularity



Interviews:

“README is a project’s welcome mat”

Data Visualization Charting library for Angular deployments. <https://github.com/pioneerjs/pioneer-charts>

#angular #data-visualization #d3 #charts #pioneerjs #pioneer-charts

667 commits 2 branches 20 releases 2 contributors MIT

Branch: master New pull request Create new file Upload files Find file Clone or download

Update angular deps Latest commit 8 on Jun 18

| | | |
|-------------|-------------------------|--------------|
| .github | Add meta | 2 years ago |
| .vscode | Update document project | 4 months ago |
| tslint.json | Update to Angular 8 | 4 months ago |

README.md

Pioneer.js charts build passing

downloads 5.8k npm v1.0.2

Quick start

To get started, it is highly recommended that you [Read the Getting Started](#) page for information on setup, configuration and examples, and more.

```
npm install @pioneerjs/pioneer-charts
```

[npm](#)

Bugs and feature requests

Have a bug or a feature request? Please first read the [issue guidelines](#) and search for existing and closed issues. If your



“the README is most important. It should describe **without having to navigate away from that page the key information people need**”

README.md

Pioneer Charts

build passing

downloads 5.8k npm v1.0.2

Quick start

To get started, it is highly recommended that you [Read the Getting Started](#) page for information on setup, configuration and examples, and more.

```
npm install @pioneer/pioneer-charts
```

[npm](#)

Bugs and feature requests

Have a bug or a feature request? Please first read the [issue guidelines](#) and search for existing and closed issues. If your problem or idea is not addressed yet, [please open a new issue](#).

Documentation

Pioneer Chart's documentation, of which is [included in this repo](#), is built with [Angular](#) and publicly hosted on GitHub Pages at <https://charts.pioneercode.com/>. The docs may also be run locally.

Contributing

Please read through our [contributing guidelines](#). Included are directions for opening issues, coding standards, and notes on development.

Editor preferences are available in the [editor config](#) for easy use in common text editors. Read more and download plugins at <http://editorconfig.org/>.



A nice README should have: a table of contents

README.md



Bootstrap

Sleek, intuitive, and powerful front-end framework for faster and easier web development.

[Explore Bootstrap docs »](#)

[Report bug](#) · [Request feature](#) · [Themes](#) · [Blog](#)

Table of contents

- [Quick start](#)
- [Status](#)
- [What's included](#)
- [Bugs and feature requests](#)
- [Documentation](#)
- [Contributing](#)
- [Community](#)



A nice README should have: descriptions and instructions

README.md

Pioneer Charts build passing

downloads 5.8k npm v1.0.2

Quick start

To get started, it is highly recommended that you [Read the Getting Started](#) page for information on setup, configuration and examples, and more.

```
npm install @pioneer/pioneer-charts
```

[npm](#)

Bugs and feature requests

Have a bug or a feature request? Please first read the [issue guidelines](#) and search for existing and closed issues. If your problem or idea is not addressed yet, [please open a new issue](#).

Documentation

Pioneer Chart's documentation, of which is [included in this repo](#), is built with [Angular](#) and publicly hosted on GitHub Pages at <https://charts.pioneercode.com/>. The docs may also be run locally.

Contributing

Please read through our [contributing guidelines](#). Included are directions for opening issues, coding standards, and notes on development.

Editor preferences are available in the [editor config](#) for easy use in common text editors. Read more and download plugins at <http://editorconfig.org/>.



A nice README should have: contact information

Bugs and feature requests

Have a bug or a feature request? Please first read the [issue guidelines](#) and search for existing and closed issues. If your problem or idea is not addressed yet, [please open a new issue](#).

Documentation

Project's documentation, of which is [included in this repo](#), is built with [Angular](#) and publicly hosted on GitHub Pages at <https://chrisb3.github.io/strudel/>. The docs may also be run locally.

Contributing

Please read through our [contributing guidelines](#). Included are directions for opening issues, coding standards, and notes on development.

Editor preferences are available in the [editor config](#) for easy use in common text editors. Read more and download plugins at <http://editorconfig.org/>.

Creator

Chad Ramos

- [Twitter](#)
- [GitHub](#)



A nice README
should have:
code quality badges

Status



What's included

Within the download you'll find the following directories and both compiled and minified variations. You'll see something

Logistic regression: Counting the number of headers in README

README.md

Pioneer Angular Charts build passing

downloads 5.8k npm v1.0.2

Quick start

To get started, it is highly recommended that you [Read the Getting Started](#) page for information on setup, configuration and examples, and more.

```
npm install @pioneercode/charts
```

npm

Bugs and feature requests

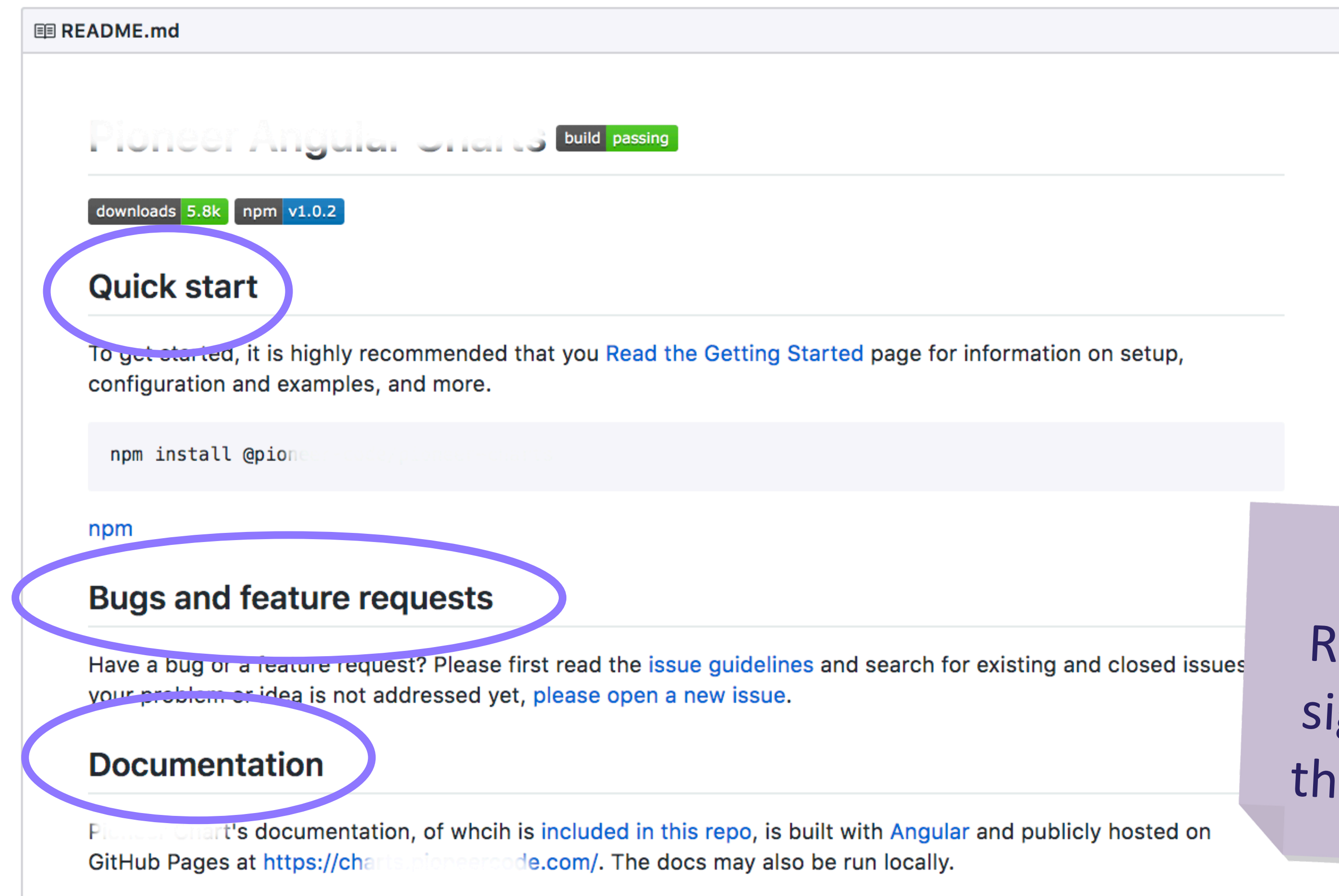
Have a bug or a feature request? Please first read the [issue guidelines](#) and search for existing and closed issues. If your problem or idea is not addressed yet, [please open a new issue](#).

Documentation

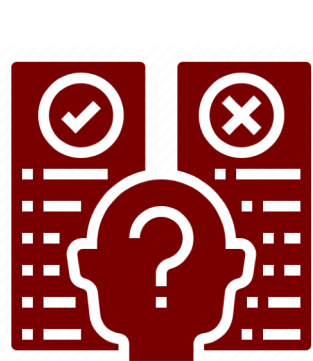
Pioneer Chart's documentation, of which is [included in this repo](#), is built with [Angular](#) and publicly hosted on GitHub Pages at <https://charts.pioneercode.com/>. The docs may also be run locally.

Logistic regression:

A well-structured README correlates to higher likelihood of attracting newcomers










Other
README
signals in
the paper



Interviews:

Contributing file “suggests they have some experience with handling new contributors”

| | |
|--|--------------------------------|
|  .test.eslintignore | test (test): add node_modules |
|  .test.eslintrc.js | cli (cli): configure |
|  .travis.yml | Removed travis node version |
|  AUTHORS | chore (AUTHORS): Fix npm use |
|  CONTRIBUTING.md | cli (cli): use yarn start to l |
|  LICENSE | dev (dev): include 2018 in |
|  README.md | cli (cli) remove greenkeeper |

Quick start

To get started, it is highly recommended that you [Read the Getting Started](#) page for information on setup, configuration and examples, and more.

`npm install @p10`
`npm`

Bugs and feature requests

Have a bug or a feature request? Please first read the [issue guidelines](#) and search for existing and closed issues. If your problem or idea is not addressed yet, [please open a new issue](#).

Documentation

Please read through our [documentation](#), of which is [included in this repo](#), is built with [Angular](#) and publicly hosted on GitHub Pages at <https://code.angularjs.org/>. The docs may also be run locally.

Contributing

Please read through our [contributing guidelines](#). Included are directions for opening issues, coding standards, and notes on development.

Editor preferences are available in the [editor config](#) for easy use in common text editors. Read more and download plugins at <http://editorconfig.org/>.



A nice contributing guideline should be: **prominent**

README.md

downloads 5.8k npm v1.0.2

Quick start

To get started, it is highly recommended that you [Read the Getting Started](#) page for information on setup, configuration and examples, and more.

```
npm install @picocolors/picocolors --save-dev
```

[npm](#)

Bugs and feature requests

Have a bug or a feature request? Please first read the [issue guidelines](#) and search for existing and closed issues. If your problem or idea is not addressed yet, [please open a new issue](#).

Documentation

Picocolors's documentation, of which is [included in this repo](#), is built with [Angular](#) and publicly hosted on GitHub Pages at <https://picocolors.github.io/picocolors/>. The docs may also be run locally.

Contributing

Please read through our [contributing guidelines](#). Included are directions for opening issues, coding standards, and notes on development.

Editor preferences are available in the [editor config](#) for easy use in common text editors. Read more and download plugins at <http://editorconfig.org/>.



A nice contributing guideline should: **explain the GitHub jargon**

Using the issue tracker

The issue tracker is the preferred channel for [bug reports](#), [features requests](#) and [submitting pull requests](#), but please respect the following restrictions:

- Please **do not** derail or troll issues. Keep the discussion on topic and respect the opinions of others.



A nice contributing guideline should:
mention ways to contribute other than code








Using the issue tracker

The [issue tracker](#) is the preferred channel for [bug reports](#), [features requests](#) and [submitting pull requests](#), but please respect the following restrictions:

- Please **do not** derail or troll issues. Keep the discussion on topic and respect the opinions of others.

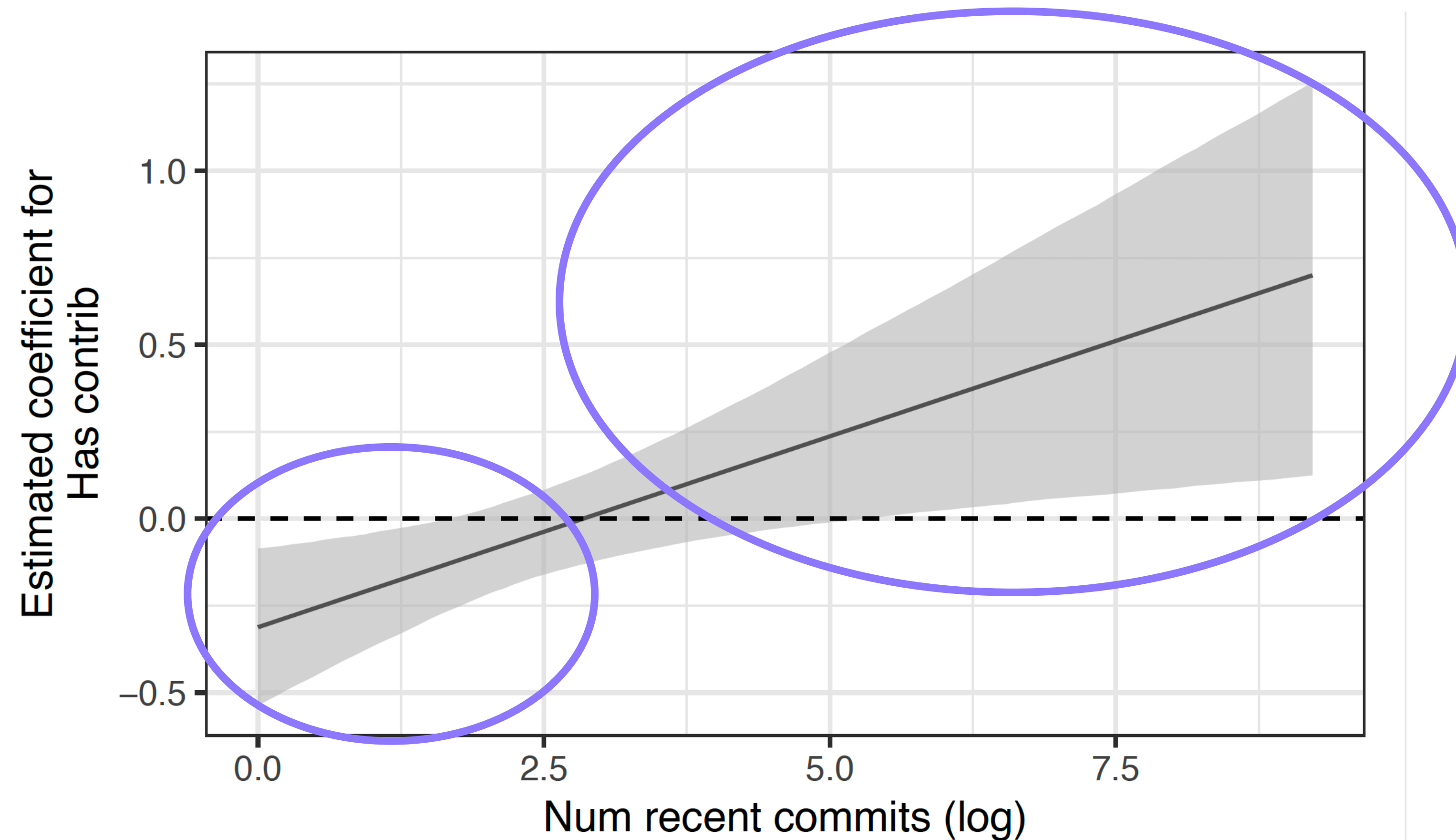
Logistic regression:

The presence of contributing guidelines correlates with **lower** likelihood of attracting newcomers

| | |
|--|----------------------------------|
|  .test.eslintignore | test(commit): add node_modules |
|  .test.eslintrc.js | chore(eslint): configure |
|  .travis.yml | Removed travis node version |
|  AUTHORS | chore(AUTHORS): Fix npm use |
|  CONTRIBUTING.md | chore(yarn): use yarn start to l |
|  LICENSE | docs(license): include 2018 in |
|  README.md | chore remove greenkeeper |



Logistic regression: Contributing Guidelines and amount of project activity interact: help vs overhead



Three main classes of signals

Quality

Community

Popularity



Interviews:

The use of language signals the community's openness



qipeng commented on Mar 7

Collaborator



Hi @g3rm4r10x , could you share a minimal code snippet with the text input that produces this error, so we can more easily diagnose it? Thanks in advance!

Contributing

1. Search the issues, if it is not already there, add one.
2. Fork the repository
3. Code Code Code
4. Submit a pull request
5. Wait for some nice guy to review and merge



Logistic regression: Measuring impoliteness in conversations

lol dude, that's not how this works.

@dinephn and I have clearly explained in [#1809](#) why we won't add this to Compose. Opening a new issue won't change that.

👎 13



Logistic regression: Impolite language correlates with **lower** likelihood of attracting new contributors



!Only for experienced contributors!

First-time GitHub
contributors:





Interviews:

Having a template signals that “there’s a good structure for contributing to [this project]”

I'm submitting a...

Please search GitHub for a similar issue or PR before submitting

<pre><code>

☐ Regression (a behavior that used to work and stopped working in a new release)

☐ Bug report

☐ Feature request

☐ Documentation issue or request

(<https://github.com/twbs/bootstrap/blob/master/>

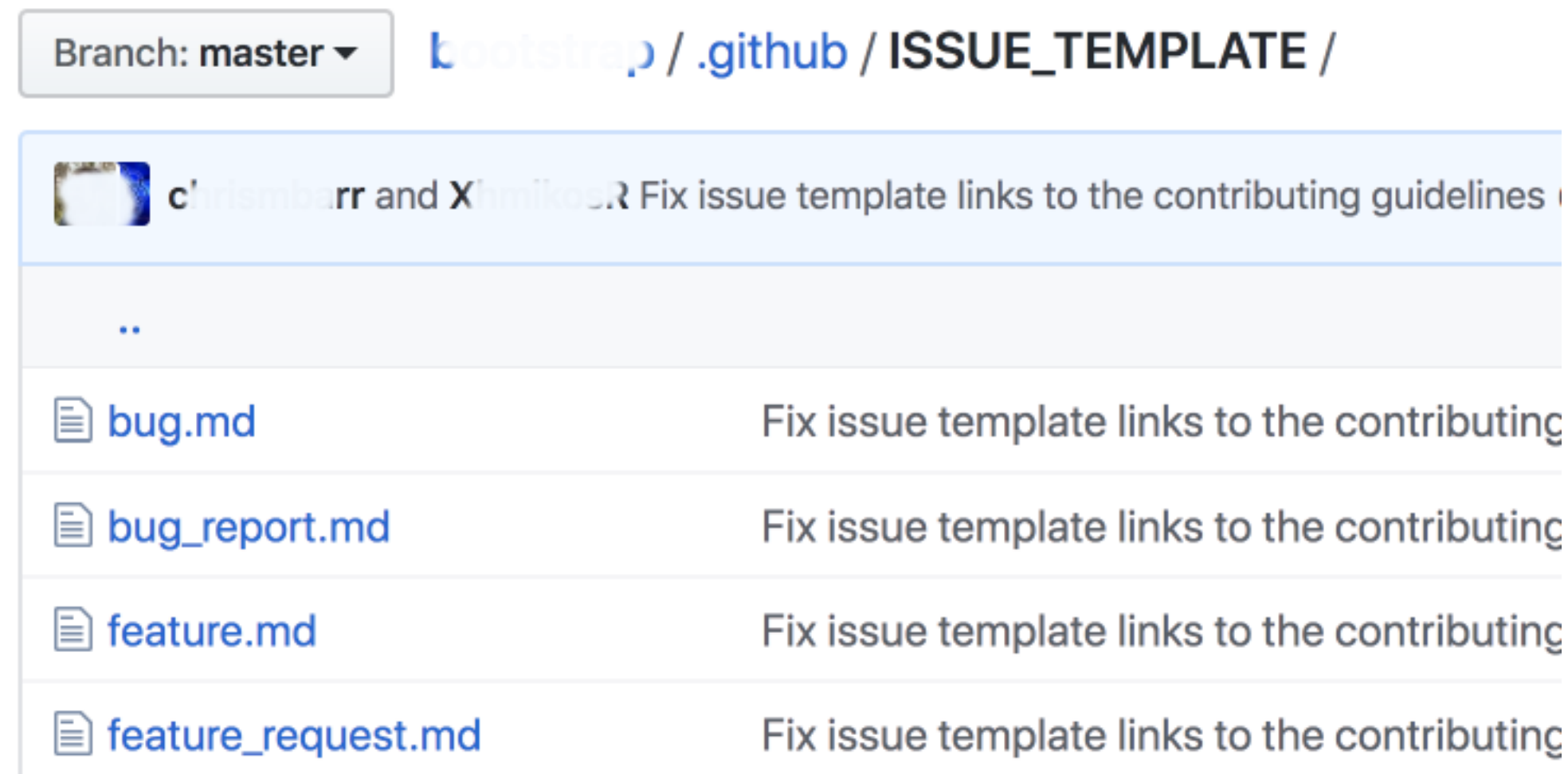
Bug reports must include:

- Operating system and version (Windows, macOS)
- Browser and version (Chrome, Firefox, Safari, IE)
- [Reduced test case](<https://css-tricks.com/reduced-test-cases/>)
- [CodePen](<https://codepen.io/>) or [JS Bin](<https://jsbin.com/>)

Logistic regression: Having a template correlates with **higher** likelihood of attracting new contributors



!Only for experienced contributors!



First-time GitHub
contributors:



Three main classes of signals

Quality

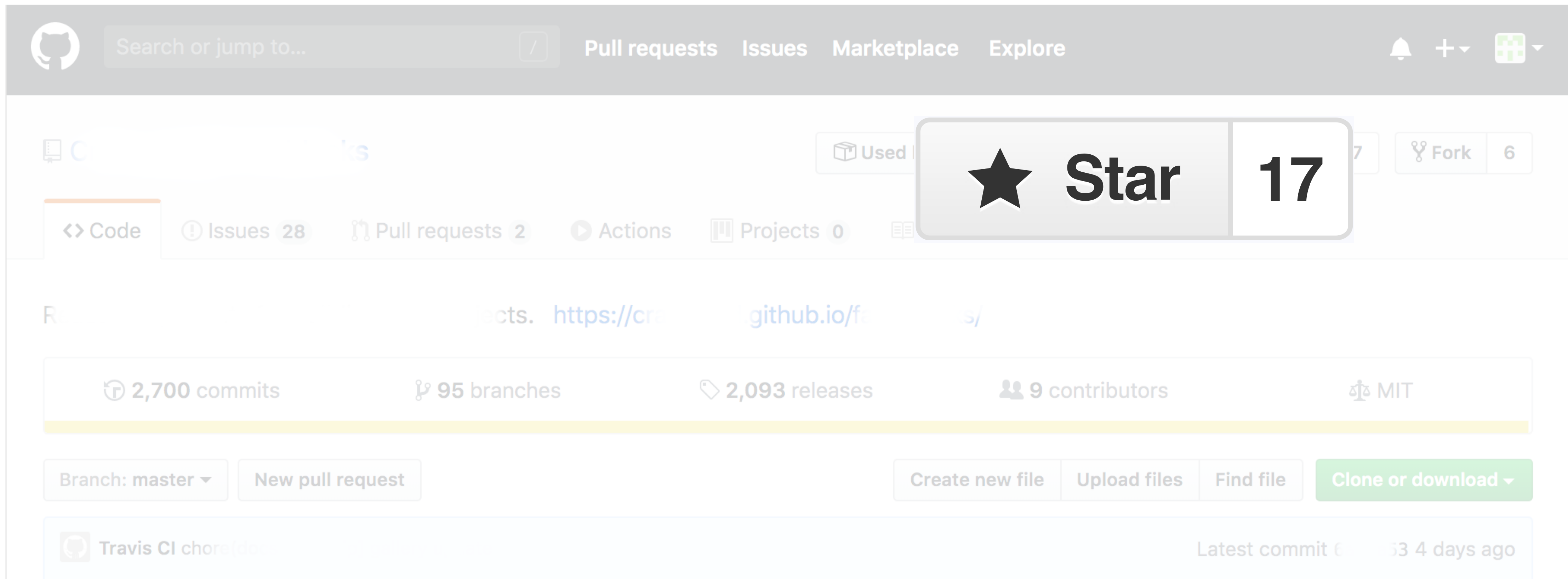
Community

Popularity



Interviews:

The number of stars signals popularity





Interviews:

Three different opinions

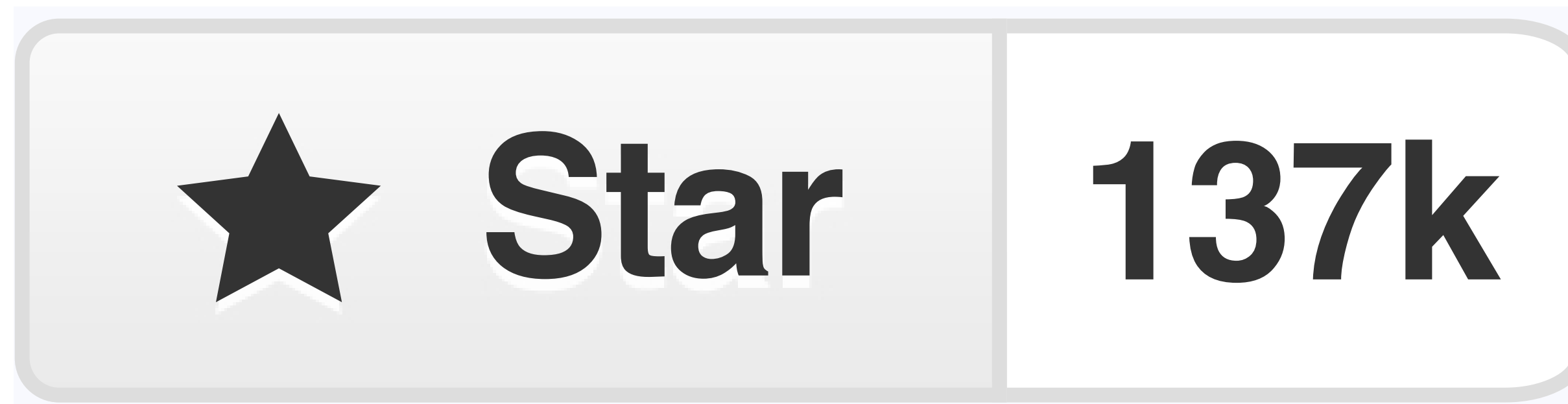
“Everyone uses [project X]. If you contribute to it your change is gonna have **a huge impact.**”



“If it has tons of stars, and there weren’t that many contributors, I would think they weren’t **friendly to new people**”.

“I wouldn’t worry about popularity, because a less popular project gives you **more self-efficacy** that forces you really to look at things, ... try everything out.”

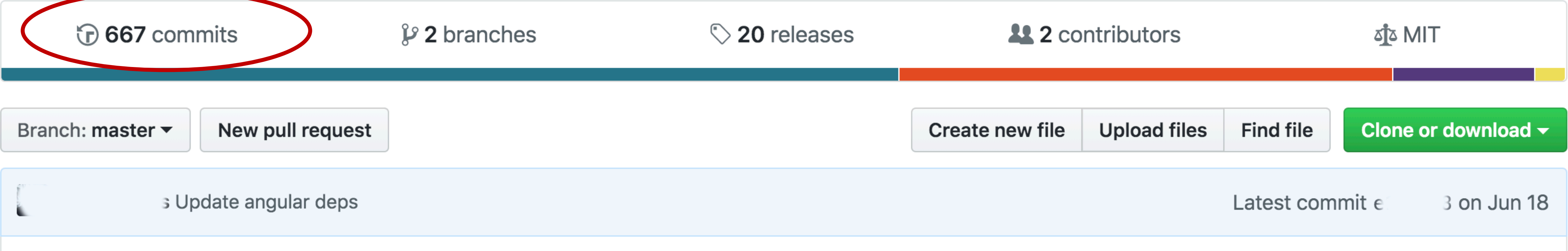
 Logistic regression:
The number of stars correlates strongly with **higher**
likelihood of attracting new contributors



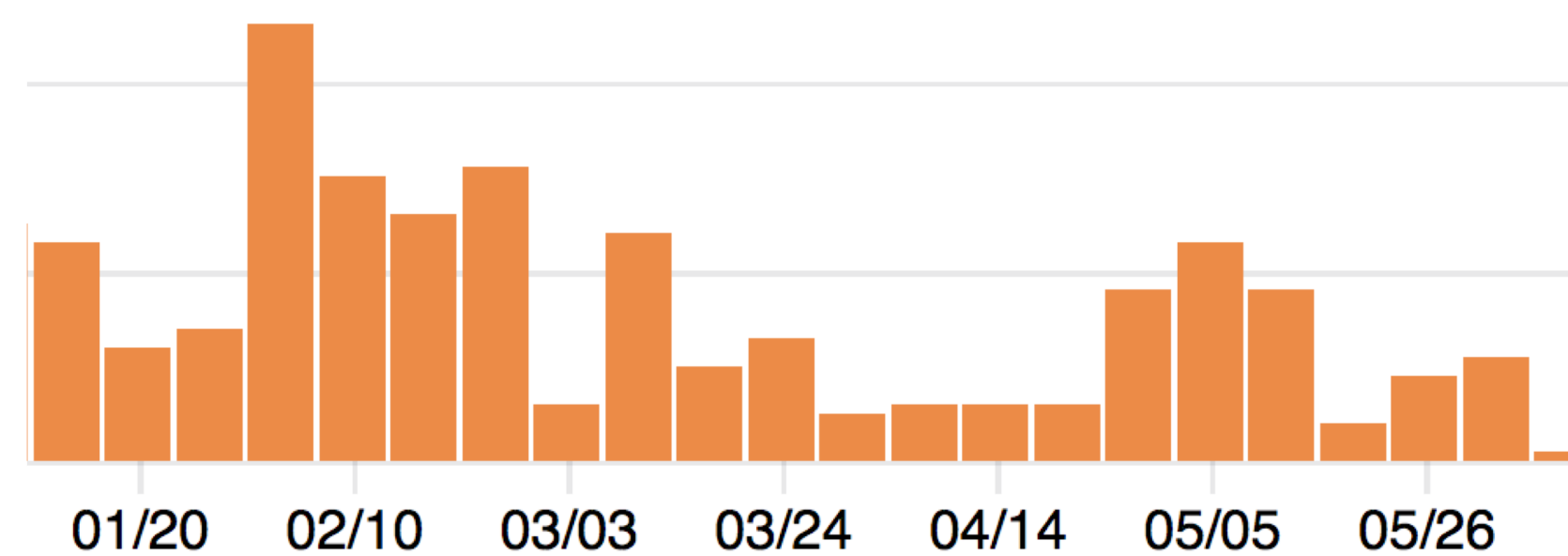


Interviews:

Recency of commits signals that the repo is active



Logistic regression: More recent commits correlates with **higher** likelihood of attracting new contributors



!Only for experienced contributors!

First-time GitHub
contributors:



Result highlights:

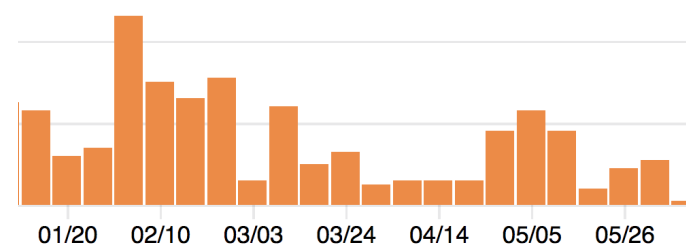
Many signals used when choosing projects

| | | Interviews | Logistic Regression |
|------------|------------------------|---|---|
| Quality | README |  |  |
| | Contributing guideline |  |  |
| Community | Impolite language |  |  |
| | Issue template |  |  |
| Popularity | Stars | |  |
| | Recent commits |  |  |

What can we do?

New signals (maybe in the form of badges)

Logistic regression:
More recent commits correlates with **higher** likelihood of attracting new contributors



!Only for experienced contributors!

First-time GitHub contributors:



maintained **no! (as of 2018)**

commit activity **457/month**

issue resolution **3 h**

updated **today**

Logistic regression:
Impolite language correlates with **lower** likelihood of attracting new contributors



!Only for experienced contributors!

First-time GitHub contributors:



commitizen **friendly**

PRs **welcome**

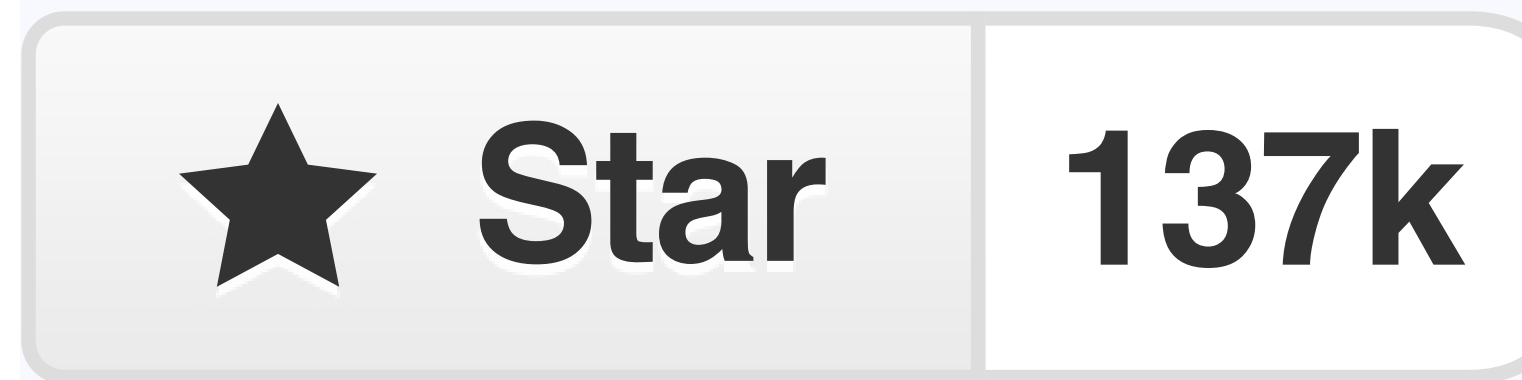
help-wanted

Personalized project recommendation

“Everyone uses [project X]. If you contribute to it your change is gonna have a **huge impact**. ”

“a less popular project gives you **more self-efficacy** that forces you really to look at things, ... try everything out.”

POPULARITY



vs



SIZE



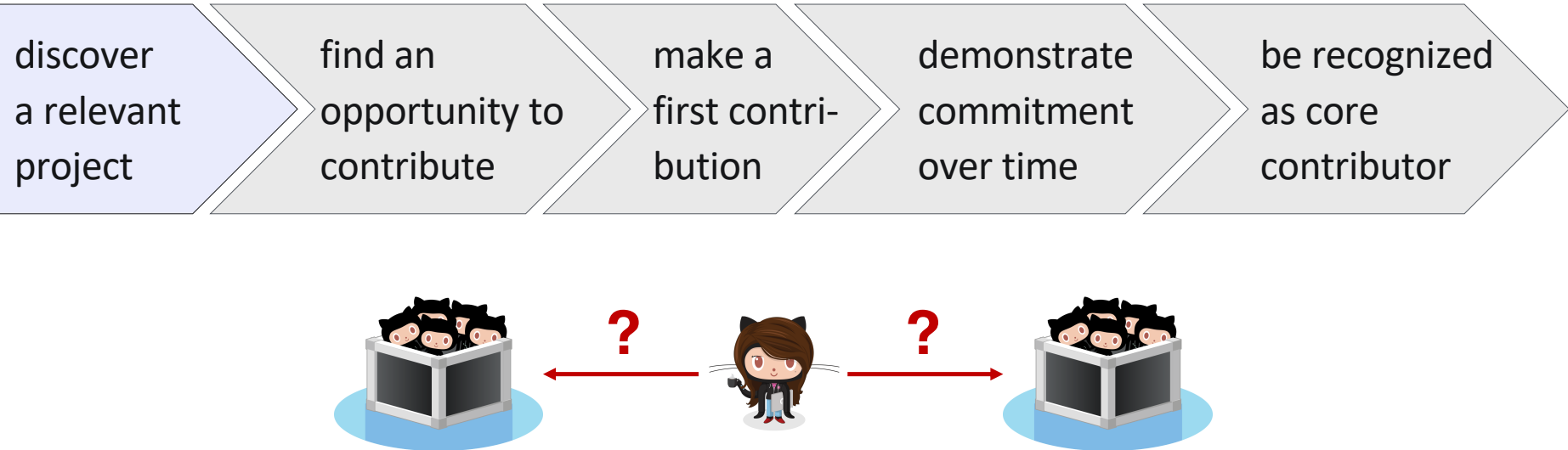
vs



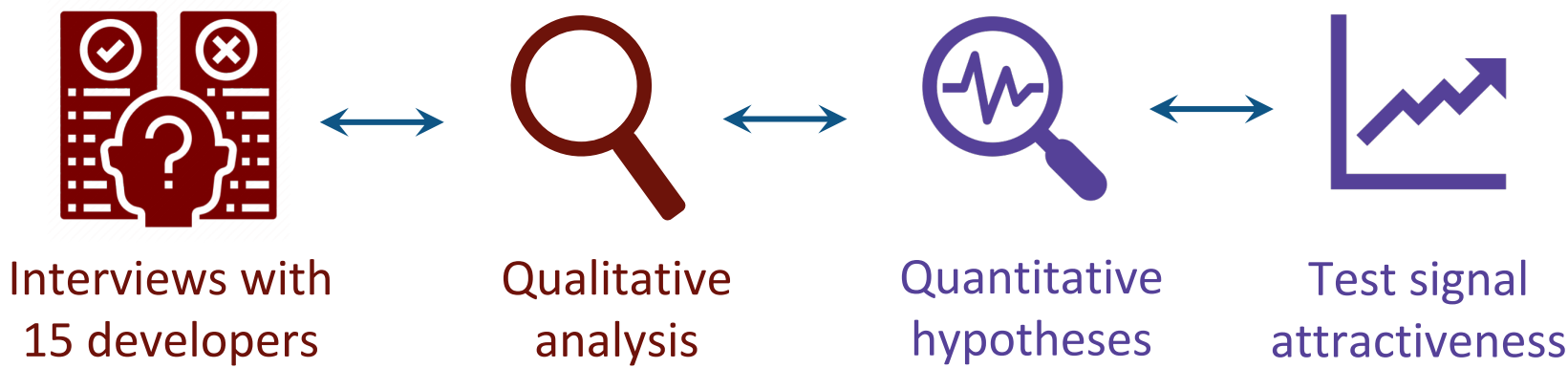
Many signals used when choosing projects

Yet, little is known about how people choose which project to contribute to

Especially for people who contribute for fun, to learn, etc.



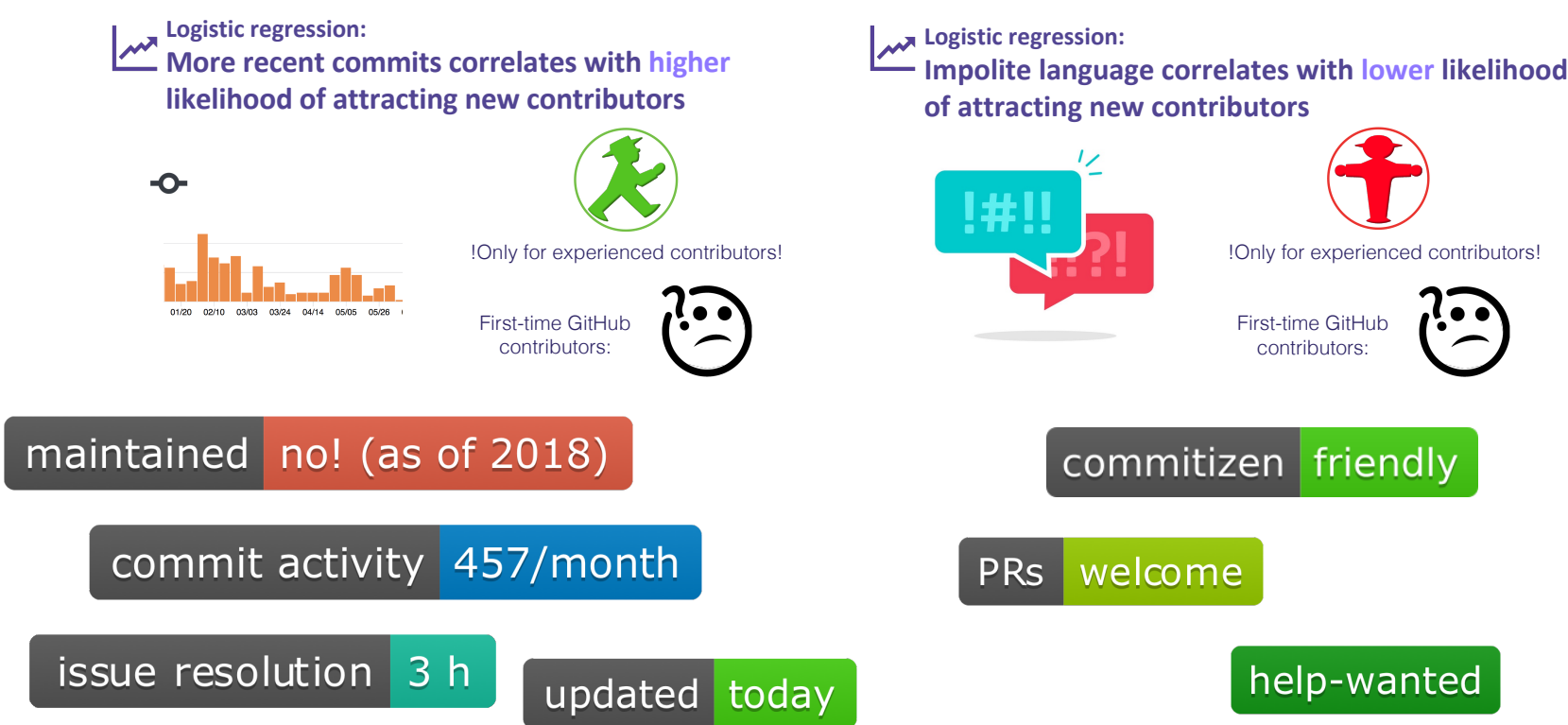
A mixed-method study design



Result highlights: Many signals used when choosing projects

| | | Interviews | Logistic Regression |
|------------|------------------------|------------|---------------------|
| Quality | README | | |
| | Contributing guideline | | |
| Community | Impolite language | | |
| | Issue template | | |
| Popularity | Stars | | |
| | Recent commits | | |

New signals (maybe in the form of badges)



Acknowledgements:



Contact: Sophie Qiu
[@sophiehsqq](https://twitter.com/sophiehsqq)
hsqq@cmu.edu

[@asarma](https://twitter.com/asarma)
Anita.Sarma@oregonstate.edu

[@b_vasilescu](https://twitter.com/b_vasilescu)
vasilescu@cmu.edu

Code and data:

DOI [10.5281/zenodo.3371186](https://doi.org/10.5281/zenodo.3371186)