

The top of the slide features a decorative header. It includes a line graph with a blue line and circular markers, set against a background of green and blue wavy shapes. The main background of the slide is a solid teal color. On the right side, there is a photograph of a wooden spoon holding a swirl of yellow butter.

Probeurre

Software Engineering 2017
Blanquet - Dousse - Mazzoleni



Introduction

What's all of this about ?

1

THE CONCEPTS

Git

Probably the most widespread tool used by software engineers to manage versioning and coding in teams

Comments

Oldest method of explaining code, still used today for good (and bad ?) reasons





The Problem

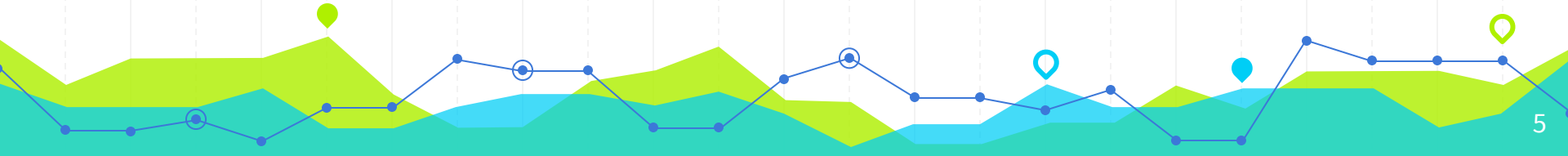
What are we trying to solve ?

2



TODOs

// TODO Find a relevant subtitle



TODO TAGS

Simple & effective

(Good) developers comment their code to explain things.
“TODO” marks the need of an action.

Language independent

Used everywhere
Now even IDEs detect comment tags such as “TODO”.

TODO TAGS

Accumulates quickly

It's easy to just leave a TODO and never look at it again

Negative connotation

Represents a lack or need of something





Our solution

What is our idea ?

3

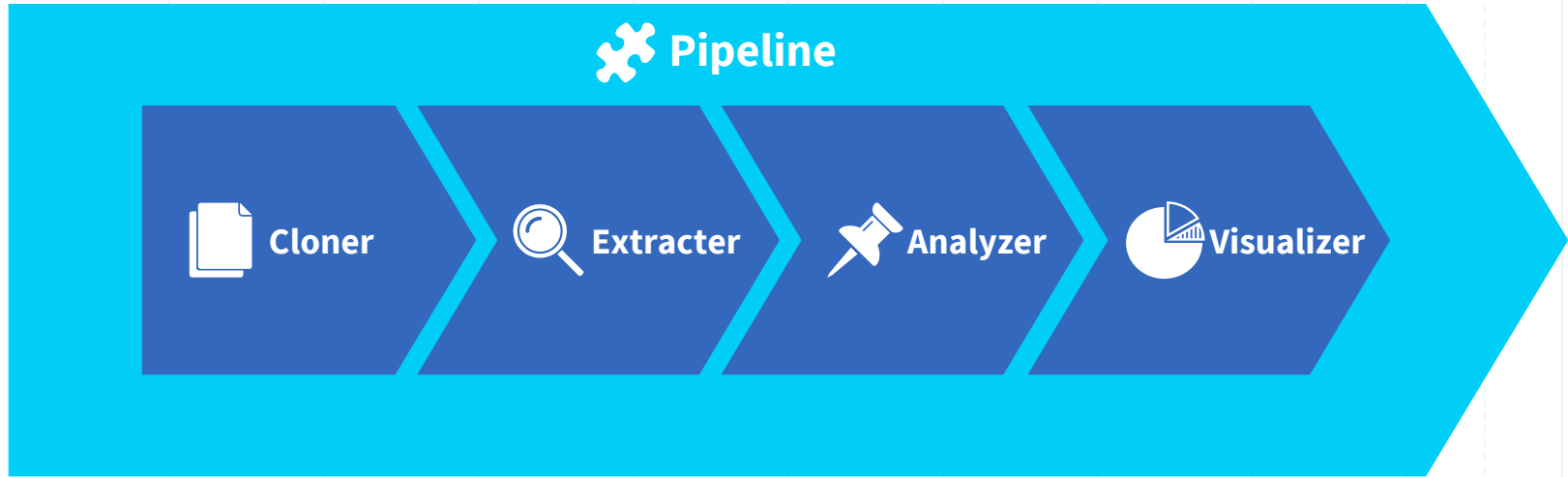
OUR IDEA

- **GitHub** is a gigantic pool of Git repositories
- Repositories probably have a lot of TODOs

Goal : Locate and count TODOs on Git repositories !



OUR PROCESS



FIVE DISTINCT REPOS + DOCKER IMAGES



Pipeline

Manages and runs the 4 other docker containers

- shell



Cloner

Takes an URL and clones the repo into a folder

- python



Extractor

Looks at files recursively to parse and extract comments

- node.js



Analyser

Looks at extracted comments and retrieves TODOs from them

- node.js



Visualizer

Takes data of Extracter and Analyzer to aggregate and display them

- python



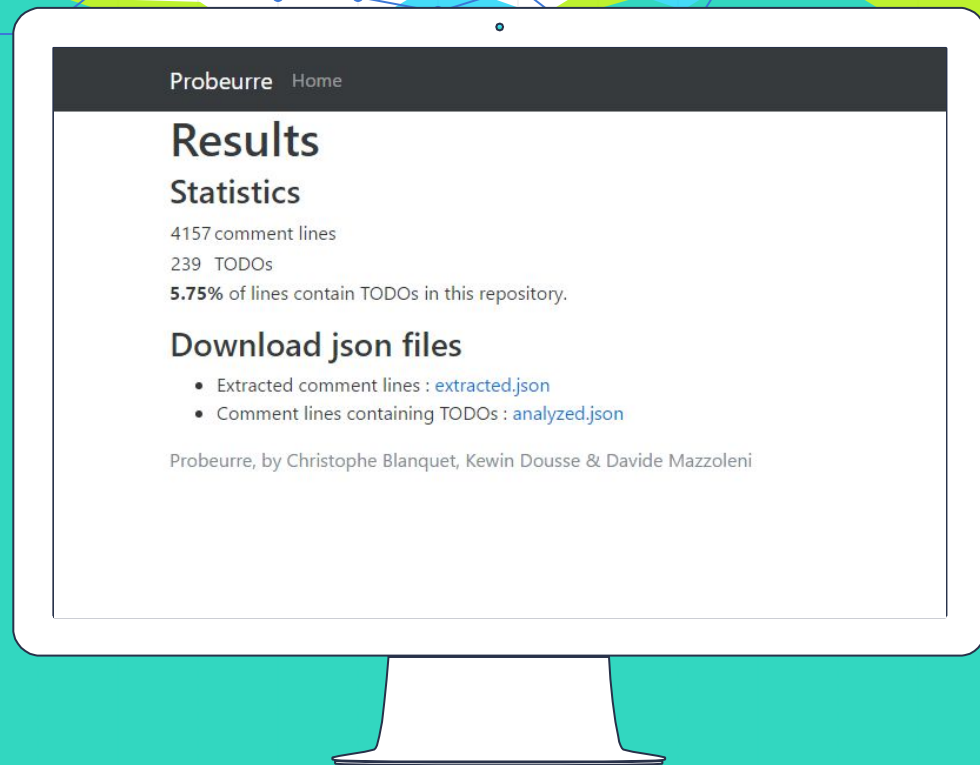
Results

4

What did we find ?

WEB INTERFACE

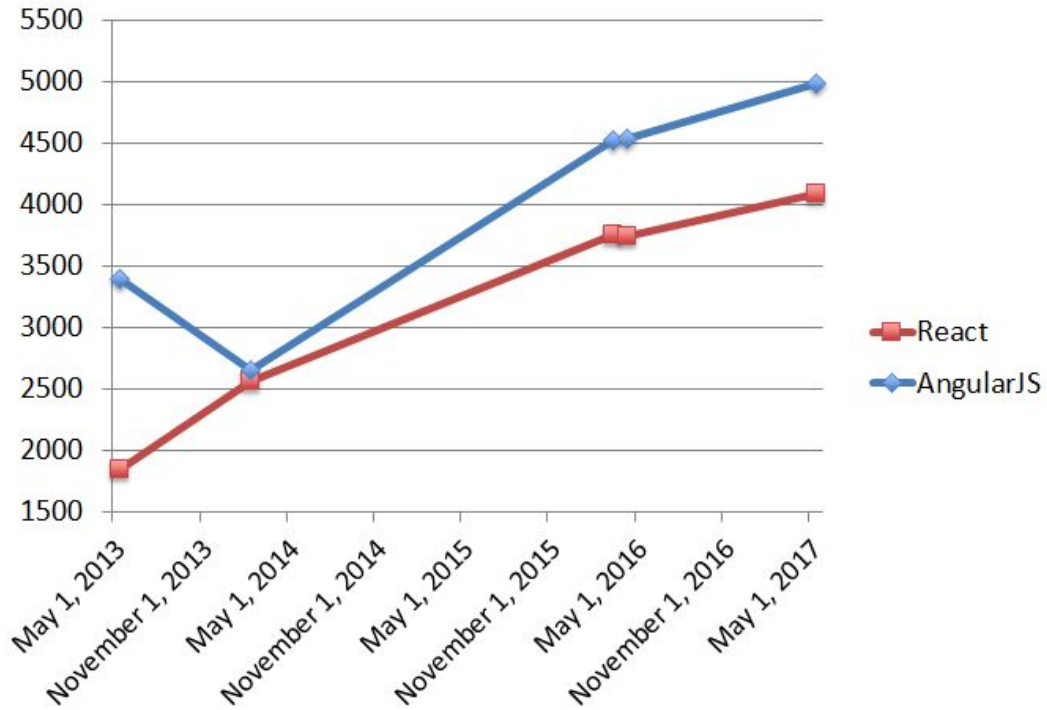
Showing the final statistics
about a repo



COMMENT LINES

	05.2013	02.2014	03.2016	05.2017
React	1'850	2'567	3'759	4'090
AngularJS	3'399	2'652	4'517	4'982

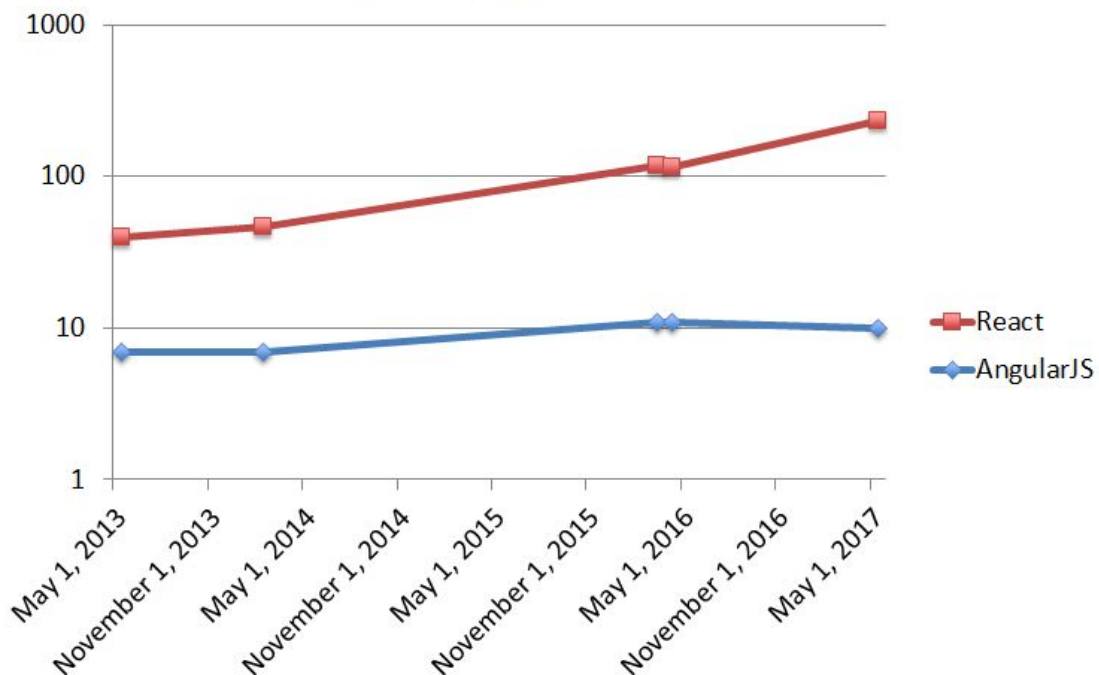
Number of COMMENT LINES over time



TODO LINES

	05.2013	02.2014	03.2016	05.2017
React	40	47	117	231
AngularJS	7	7	11	10

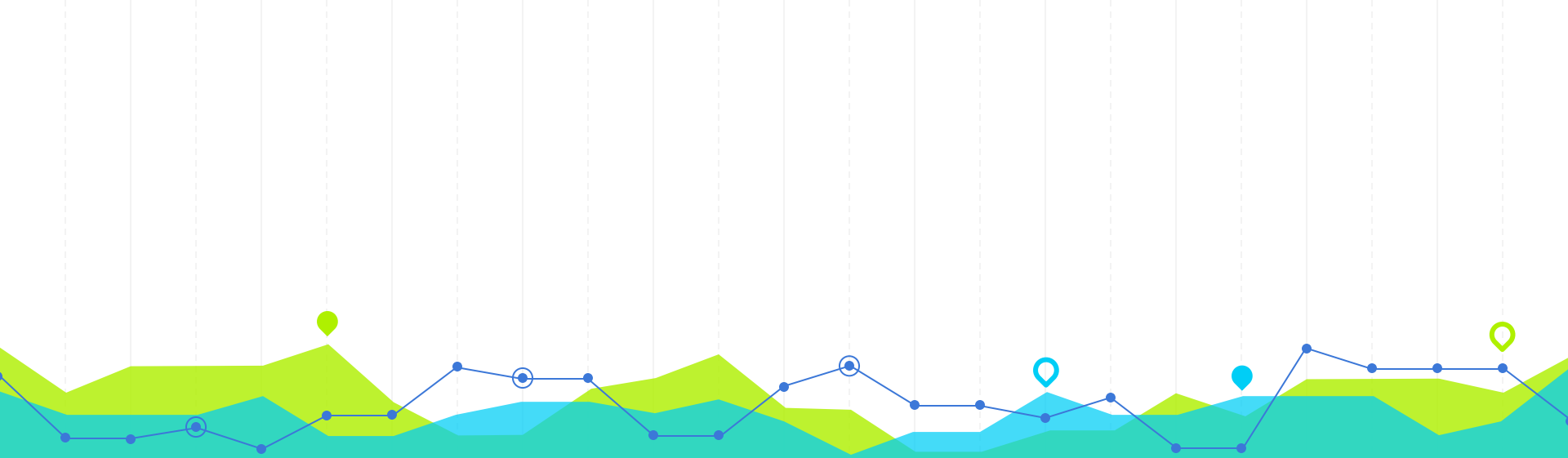
Number of TODOs over time





DEMO

Do you have a git repo in mind ?



Conclusion

Let's recap

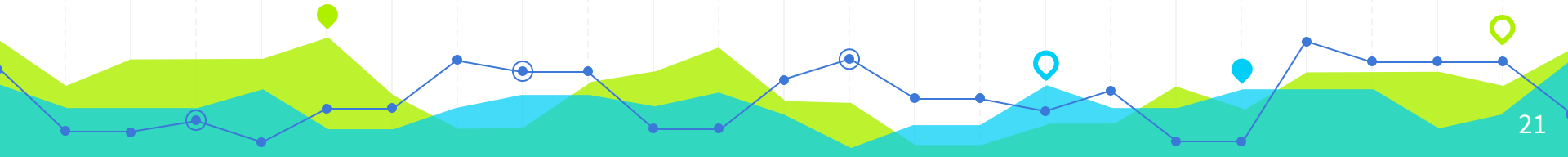
5

TO SUM IT UP

- Widely used technology
- A lot of available data
- Used **Docker** to make a scalable pipeline
- TODOs are used **everywhere**, but not in the same way
- Could use more **automation**

THANKS!

Any questions?



CREDITS

Special thanks SlidesCarnival who made and released the presentation template for free

