

# Does code ownership have an effect on code quality?

• • •

Daniele Montesi - Maciej Kędzierski  
Sebastian Ober - Michael Schwarz

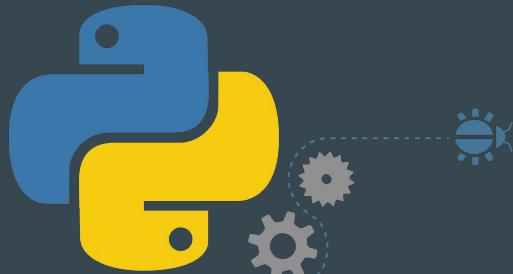
A close-up photograph of dark, textured metal that has been partially eaten away by rust, creating a mottled orange and brown pattern. Overlaid on the image are several semi-transparent white icons: a large rounded rectangle in the upper left containing the number '75,841', a heart shape in the upper right containing the number '2,039', a central gear icon, a play/pause button icon in the bottom right, and a large arrow pointing right with the number '6' next to it.

75,841

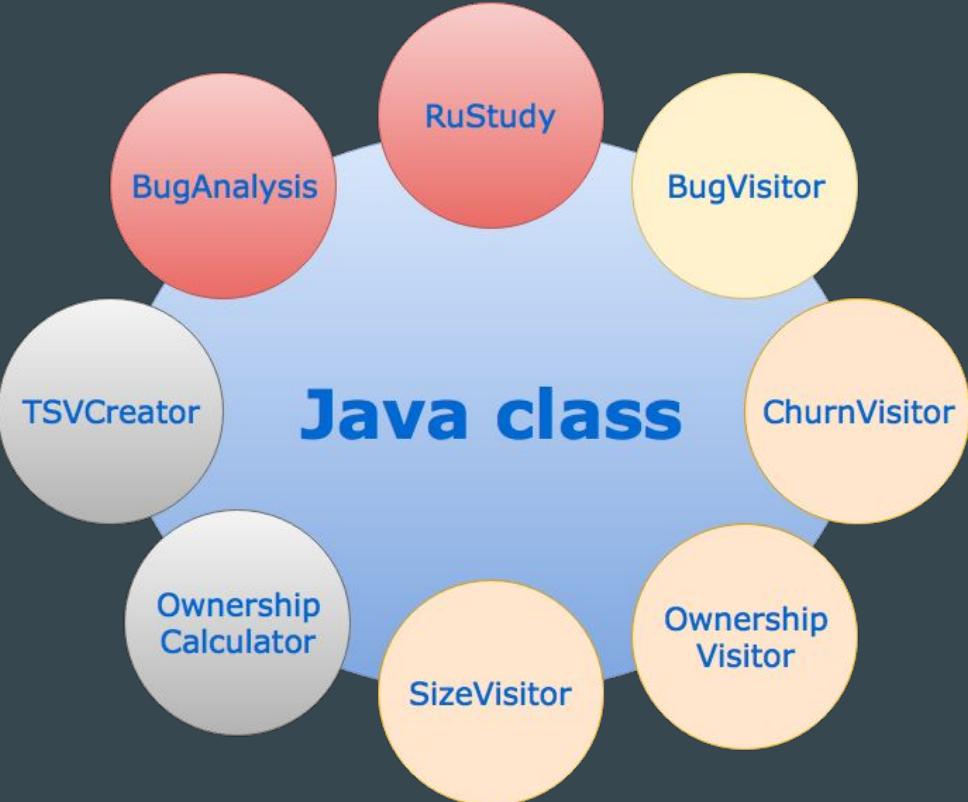
2,039

6

# From GitPython to Java RepoDriller



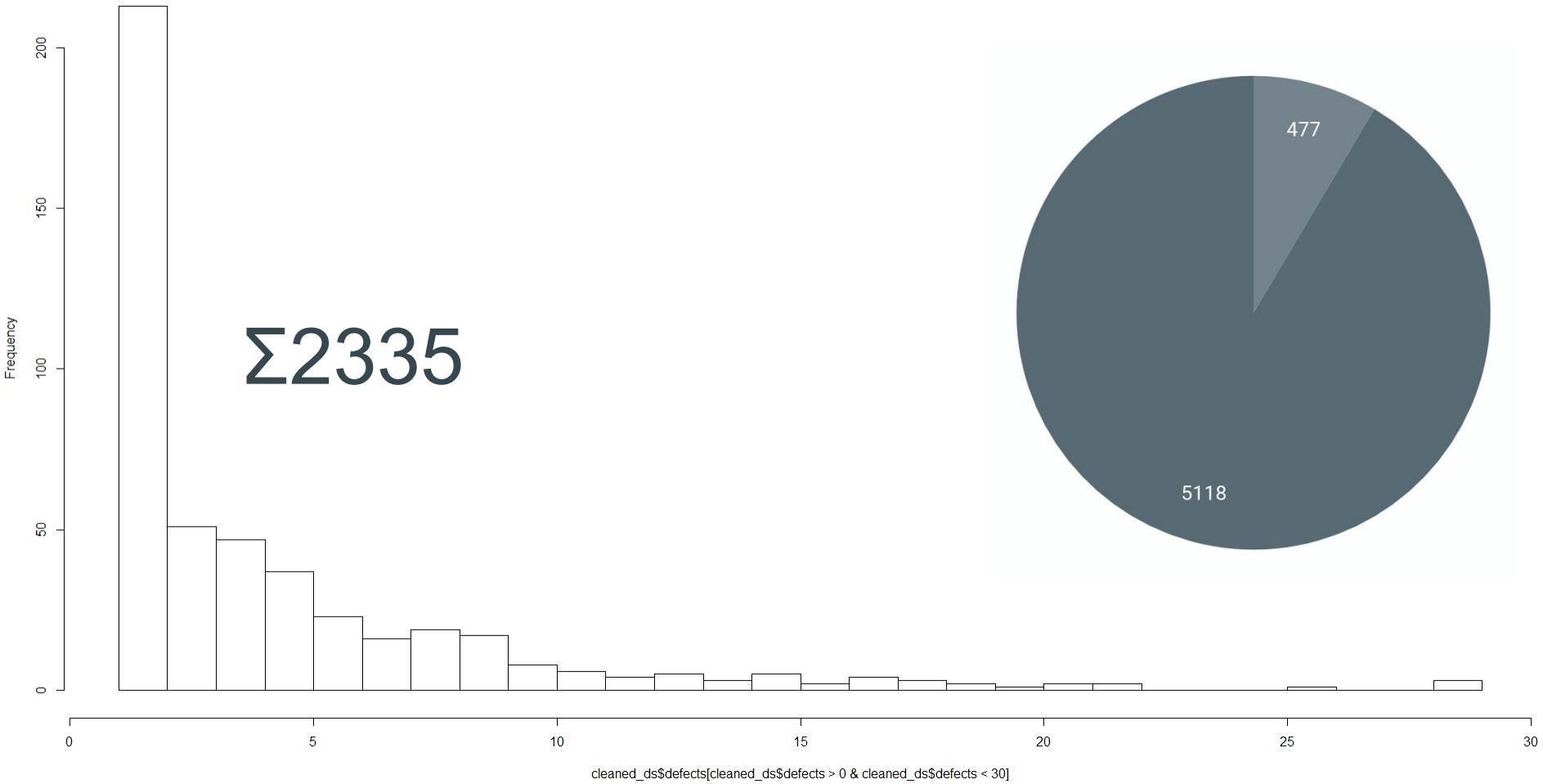
# Java class



```
public void run() {  
  
    BugVisitor bugVisitor = new BugVisitor(noBugsIntroducedBefore,  
                                         noBugsIntroducedAfter, defectsMap);  
  
    new RepositoryMining()  
        .in(GitRepository.singleProject(repoPath))  
        .through(Commits.range(commitAtRelease,  
                               commitYearAfterRelease))  
        .process(bugVisitor)  
        .mine();  
}
```

```
new RepositoryMining()  
    .in(GitRepository.singleProject(repoPath))  
    .through(Commits.range(earliestCommit, commitAtRelease))  
  
.process(ownerVisitor).process(churnVisitor).process(sizeVisitor)  
.mine();
```

Histogram of cleaned\_ds\$defects[cleaned\_ds\$defects > 0 & cleaned\_ds\$defects < 30]



## Our findings: DV in terms of CVs

```
Estimate Std. Error t value Pr(>|t|)  
(Intercept) -3.66e-02 2.26e-02 -1.62 0.1  
cleaned_ds$churn 1.04e-03 2.72e-05 38.31 <2e-16 ***  
cleaned_ds$size 1.60e-03 9.14e-05 17.52 <2e-16 ***  
---  
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Residual standard error: 1.58 on 5592 degrees of freedom

Multiple R-squared: 0.494, Adjusted R-squared: 0.494

F-statistic: 2.73e+03 on 2 and 5592 DF, p-value: <2e-16

# Our findings: DV in terms of IV and CVs

```
Coefficients: (1 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) -1.69e-01  1.16e-01  -1.47   0.14    
cleaned_ds$churn    8.78e-04  3.19e-05   27.49 <2e-16 ***
cleaned_ds$size     1.23e-03  9.84e-05   12.51 <2e-16 ***
cleaned_ds$minor    6.39e-02  6.74e-03    9.47 <2e-16 ***
cleaned_ds$major    2.34e-02  1.69e-02    1.38   0.17    
cleaned_ds$total      NA         NA        NA       NA      
cleaned_ds$ownership 9.51e-02  1.14e-01    0.84   0.40    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Residual standard error: 1.57 on 5589 degrees of freedom

Multiple R-squared: 0.503, Adjusted R-squared: 0.502

F-statistic: 1.13e+03 on 5 and 5589 DF, p-value: <2e-16

## Our finding

We failed to establish any meaningful relationship between ownership metrics and the number of post-release bugs for the Rust repository when we control for size and churn.

# Possible reasons

We are correct.

We f\*\*\*\*\* up the Data Science part.

There is some preprocessing step missing.

We are not properly detecting post-release bugs.

# Learnings

```
graph LR; KISS((KISS)) --- SAUSAGE((How the  
sausage  
is made)); TRAP((The  
pancakes  
are a  
trap)) --- SAUSAGE; R((R)) --- SAUSAGE;
```

How the  
sausage  
is made

The  
pancakes  
are a  
trap

R

# Is there a correlation between code ownership and code quality?

• • •

“Any headline that ends in a question mark  
can be answered by the word *no*.”

-- Betteridge's law of headlines

# Thank you

Daniele Montesi  
Politecnico di Milano  
[daniele.montesi@mail.polimi.it](mailto:daniele.montesi@mail.polimi.it)

Maciej Kędzierski  
Politechnika Warszawska  
[m.a.kedzierski@stud.elka.pw.edu.pl](mailto:m.a.kedzierski@stud.elka.pw.edu.pl)

Sebastian Ober  
Technische Universität München  
[sebastian.ober@tum.de](mailto:sebastian.ober@tum.de)

Michael Schwarz  
Technische Universität München  
[m.schwarz@tum.de](mailto:m.schwarz@tum.de)

