Learning a Codebase

Derek Somerville

November 11, 2025

Contents

1 Repository

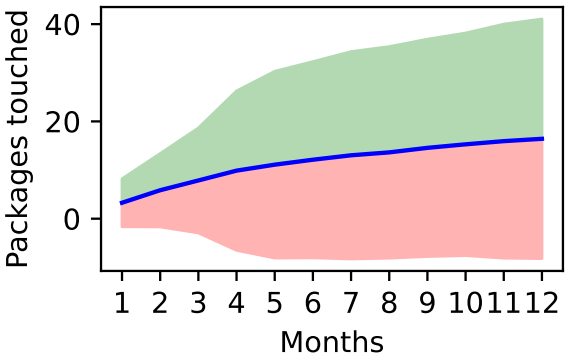
1.1 Glossary - Summary

- The founder developer starts in the first six months of a project.
- The late joiner developers start after six months.
- Sustained developers make 50 or more commits and commit for 250 days or more.
- Transient developers have fewer commits or commit for a shorter period.

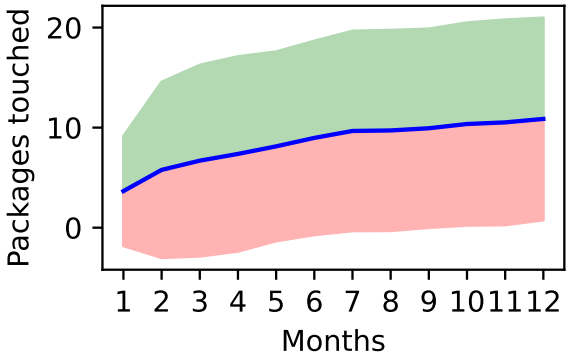
2 Repository: 14

2.1 Time series developer - 14 For packages touched for each period

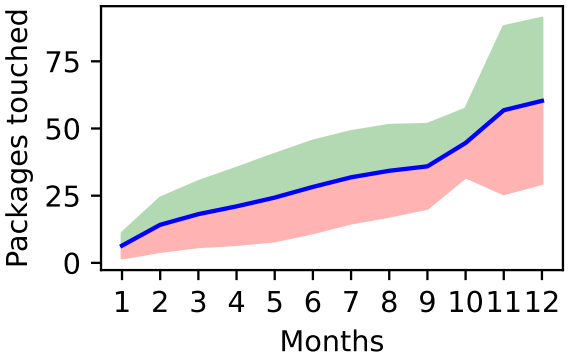
A time series of packages touched on average each month.



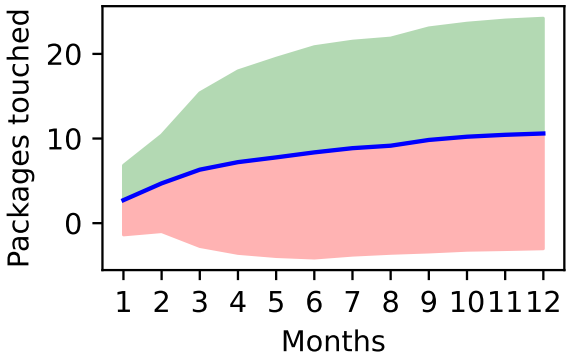
(a) All developers (244) showing packages touched



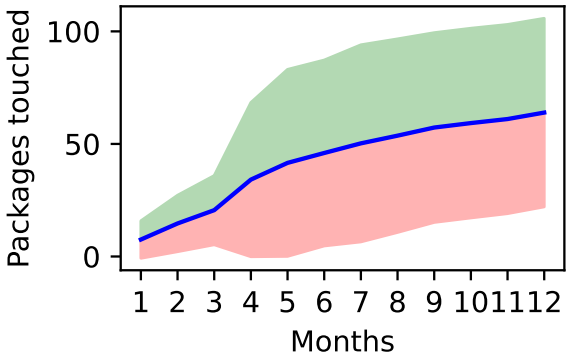
(b) Transient founder developers (13) showing packages touched



(c) Sustained founder developers (6) showing packages touched

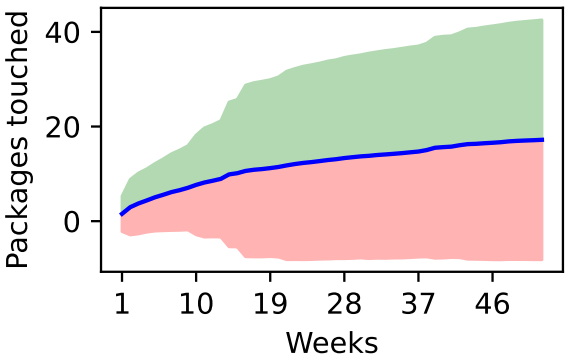


(d) Transient later joiner developers (204) showing packages touched

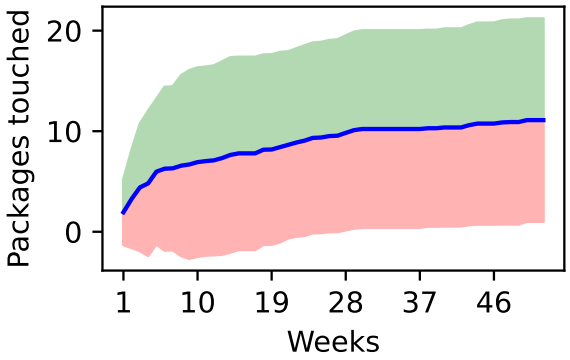


(e) Sustained later joiner developers (21) showing packages touched

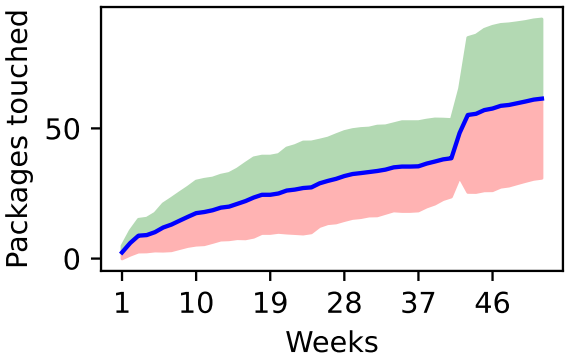
Figure 1: A time series of the average (mean) total packages touched on average each month, with positive (orange) and negative (red) filled standard deviation.



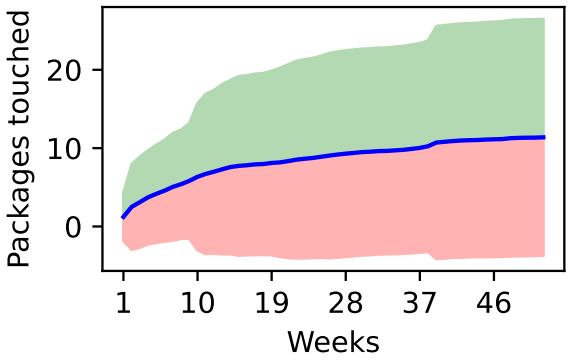
(a) All developers (244) showing packages touched



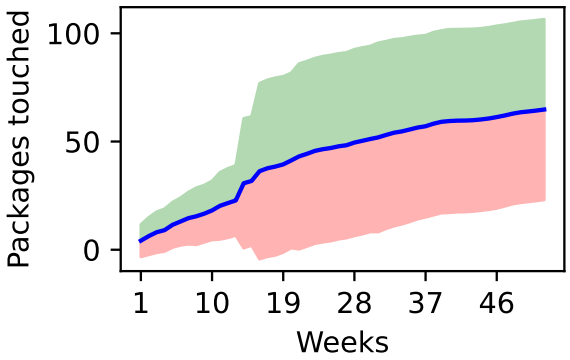
(b) Transient founder developers (13) showing packages touched



(c) Sustained founder developers (6) showing packages touched

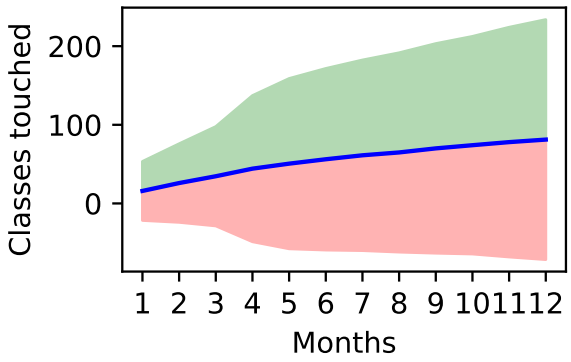


(d) Transient later joiner developers (204) showing packages touched

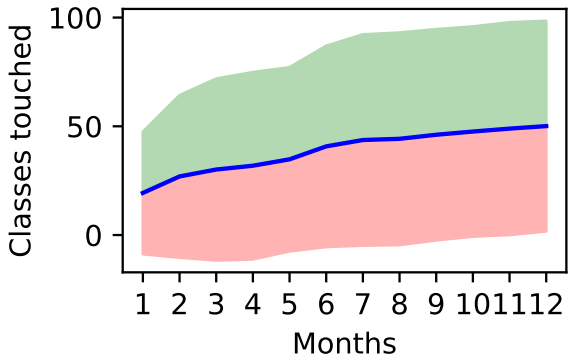


(e) Sustained later joiner developers (21) showing packages touched

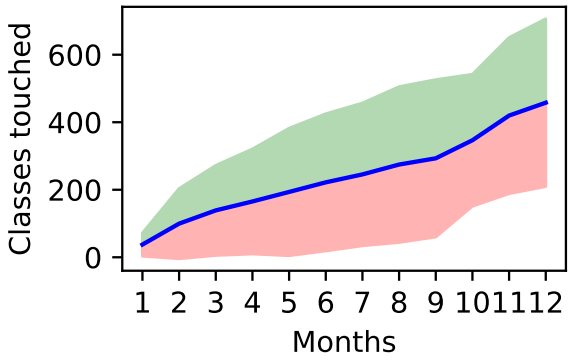
Figure 2: A time series of the average (mean) total packages touched on average each week, with positive (orange) and negative (red) filled standard deviation.



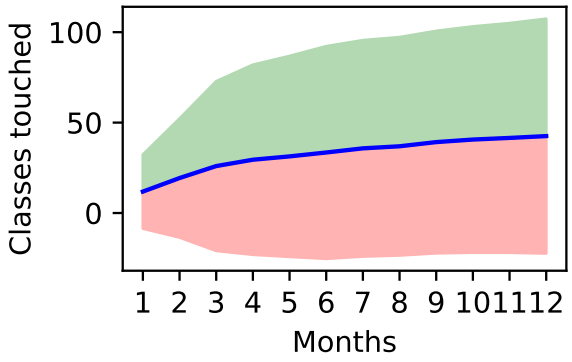
(a) All developers (244) showing classes touched



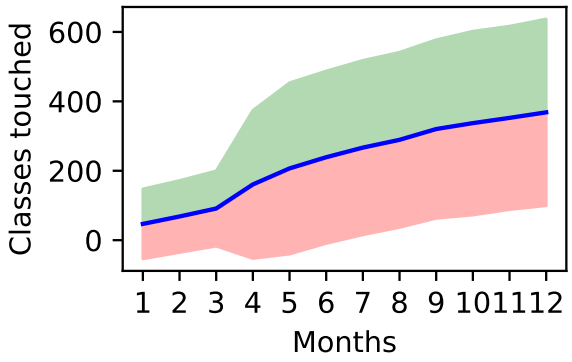
(b) Transient founder developers (13) showing classes touched



(c) Sustained founder developers (6) showing classes touched

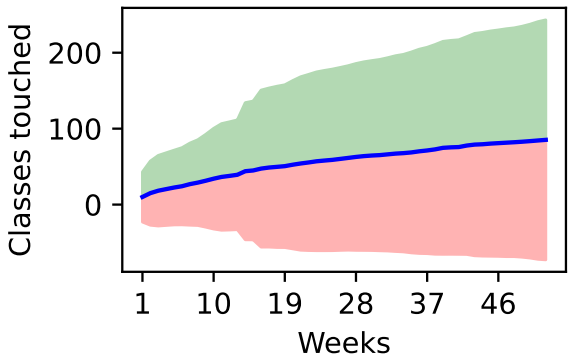


(d) Transient later joiner developers (204) showing classes touched

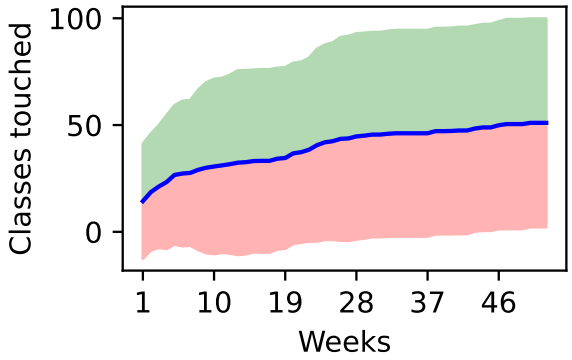


(e) Sustained later joiner developers (21) showing classes touched

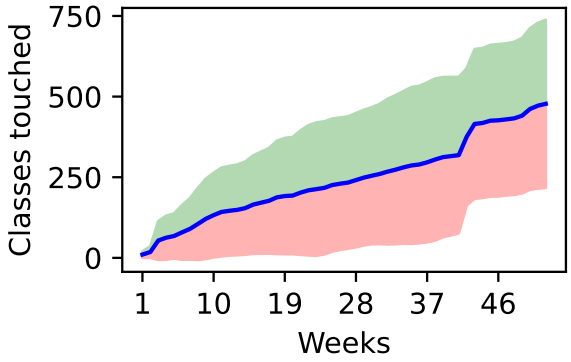
Figure 3: A time series of the average (mean) total classes touched on average each month, with positive (orange) and negative (red) filled standard deviation.



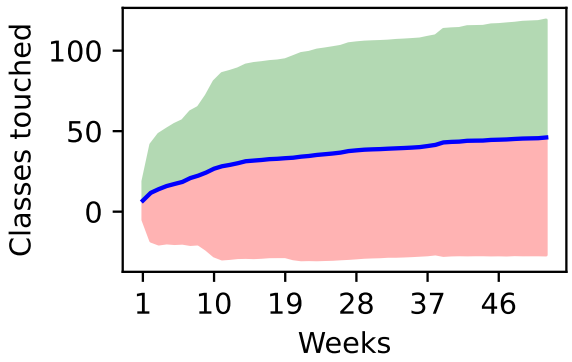
(a) All developers (244) showing classes touched



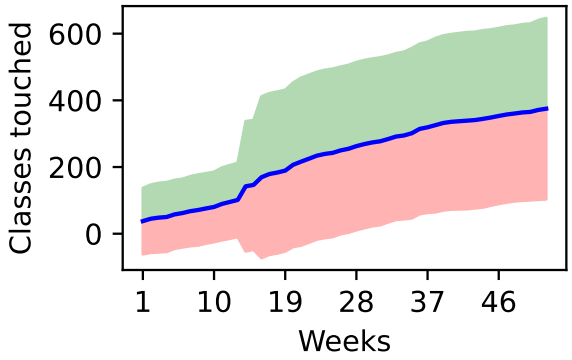
(b) Transient founder developers (13) showing classes touched



(c) Sustained founder developers (6) showing classes touched

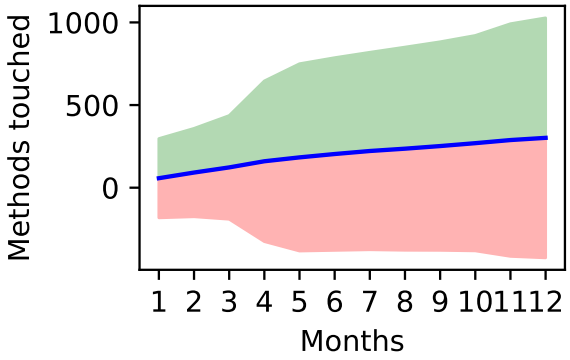


(d) Transient later joiner developers (204) showing classes touched

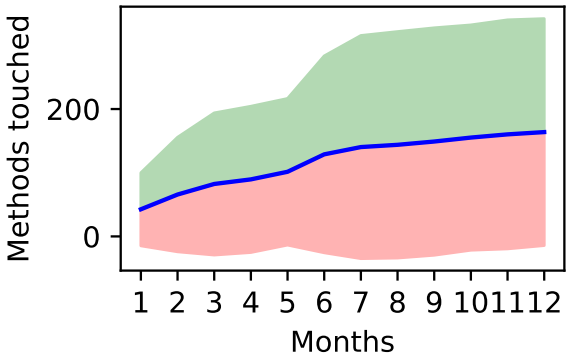


(e) Sustained later joiner developers (21) showing classes touched

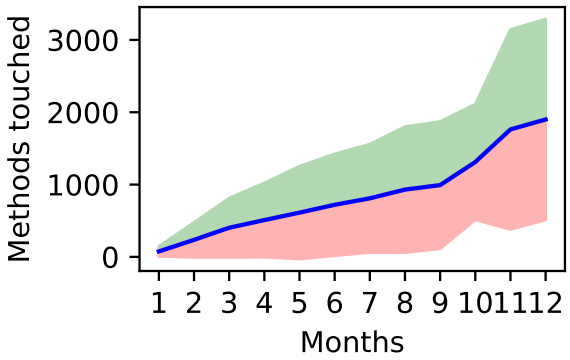
Figure 4: A time series of the average (mean) total classes touched on average each week, with positive (orange) and negative (red) filled standard deviation.



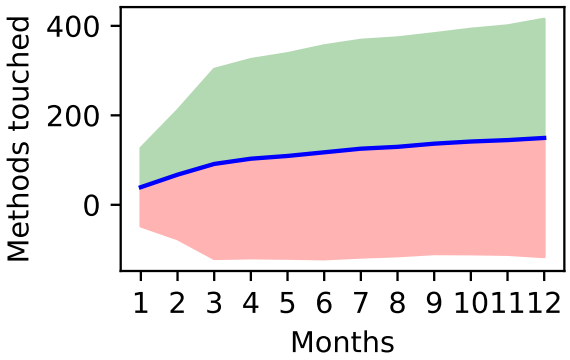
(a) All developers (244) showing methods touched



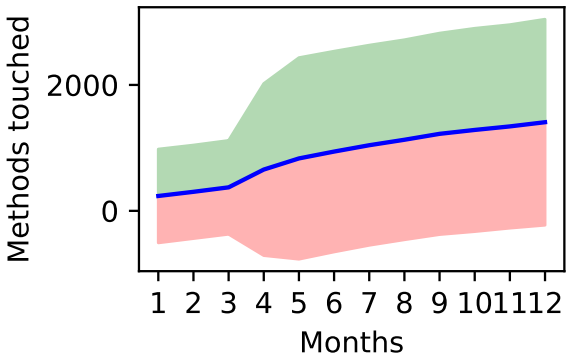
(b) Transient founder developers (13) showing methods touched



(c) Sustained founder developers (6) showing methods touched

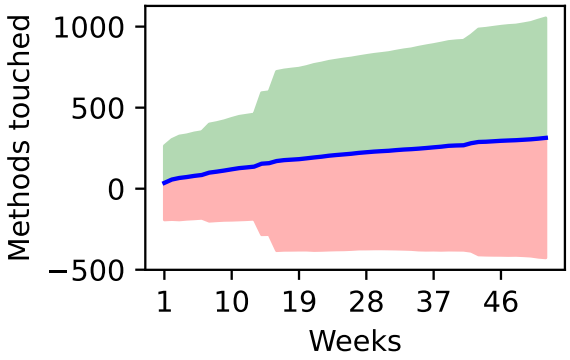


(d) Transient later joiner developers (204) showing methods touched

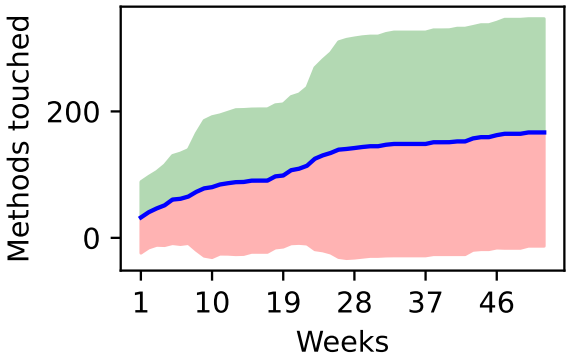


(e) Sustained later joiner developers (21) showing methods touched

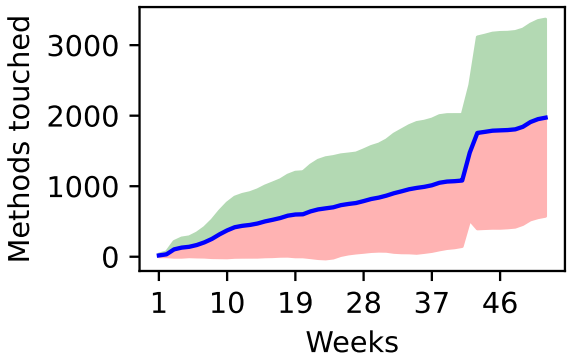
Figure 5: A time series of the average (mean) total methods touched on average each month, with positive (orange) and negative (red) filled standard deviation.



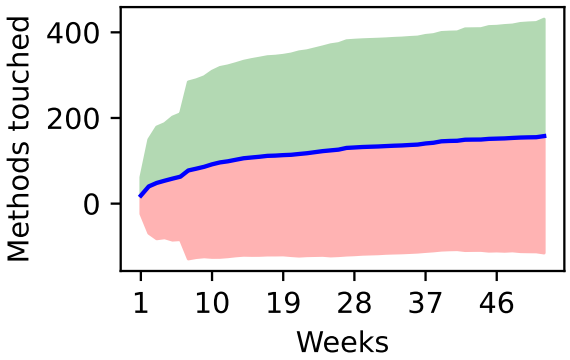
(a) All developers (244) showing methods touched



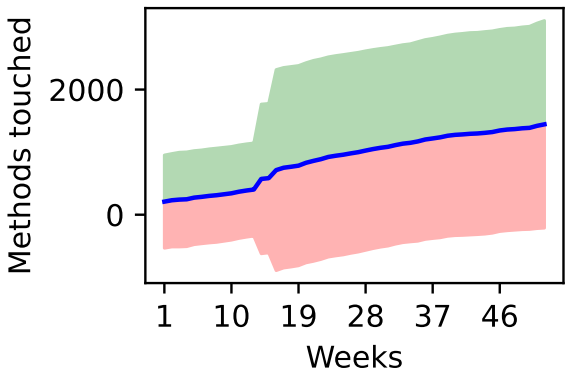
(b) Transient founder developers (13) showing methods touched



(c) Sustained founder developers (6) showing methods touched



(d) Transient later joiner developers (204) showing methods touched

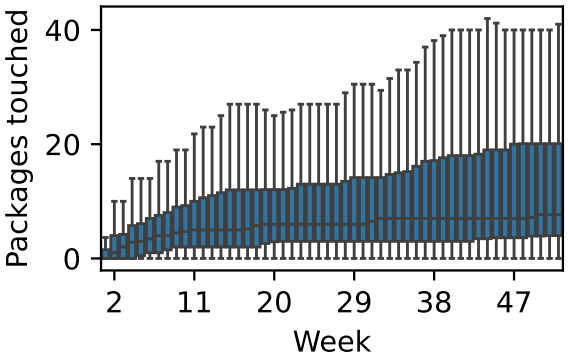


(e) Sustained later joiner developers (21) showing methods touched

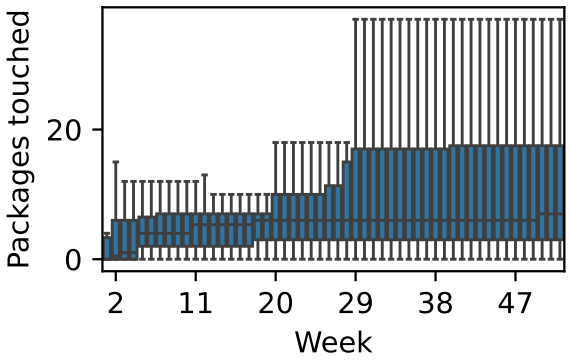
Figure 6: A time series of the average (mean) total methods touched on average each week, with positive (orange) and negative (red) filled standard deviation.

2.2 Box plot developer - 14 For packages touched for each period

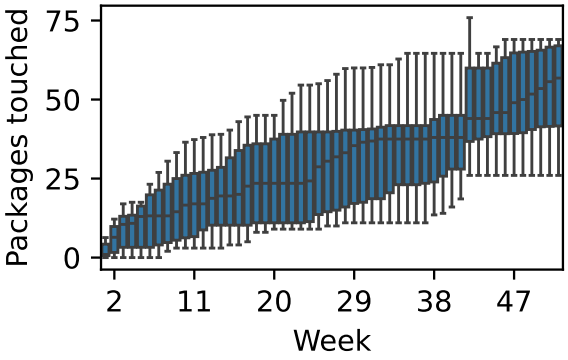
A box plot of packages touched on average each period the number of commits to components touched.



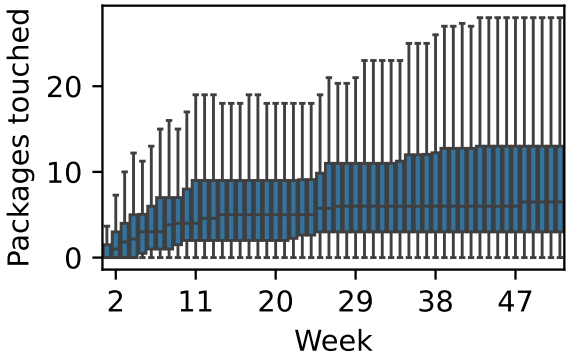
(a) All developers (244) showing packages touched



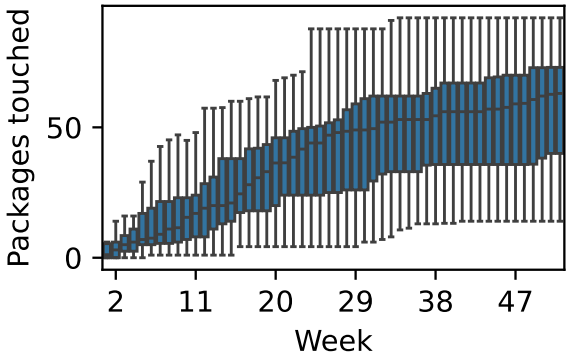
(b) Transient founder developers (13) showing packages touched



(c) Sustained founder developers (6) showing packages touched

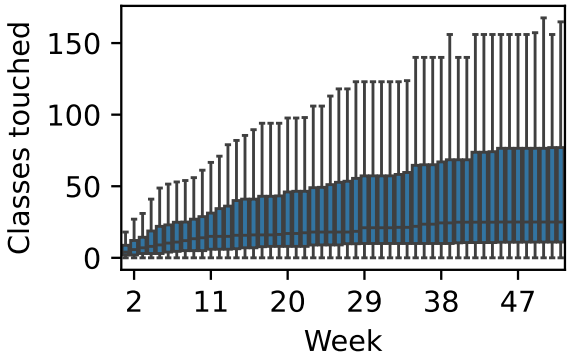


(d) Transient later joiner developers (204) showing packages touched

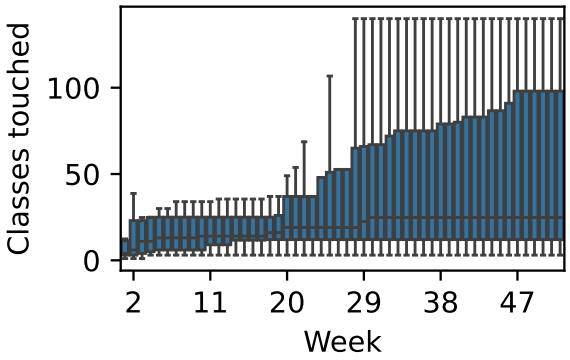


(e) Sustained later joiner developers (21) showing packages touched

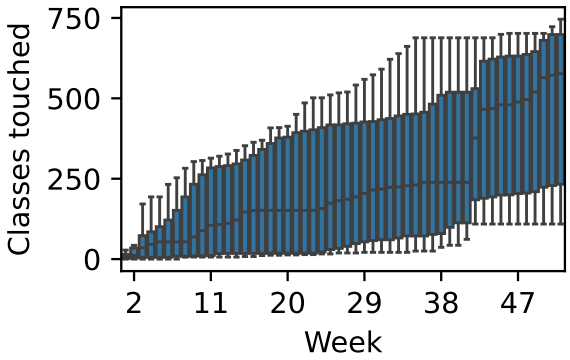
Figure 7: A box plot of total packages touched on mean each month, with quartile shading.



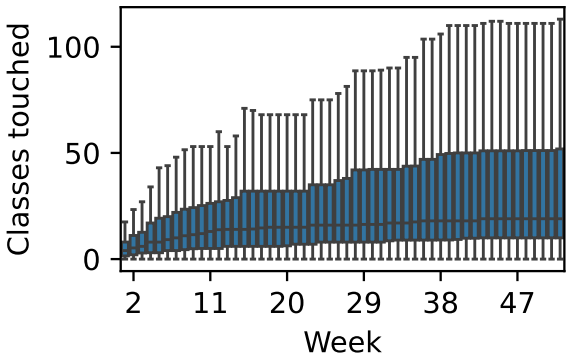
(a) All developers (244) showing classes touched



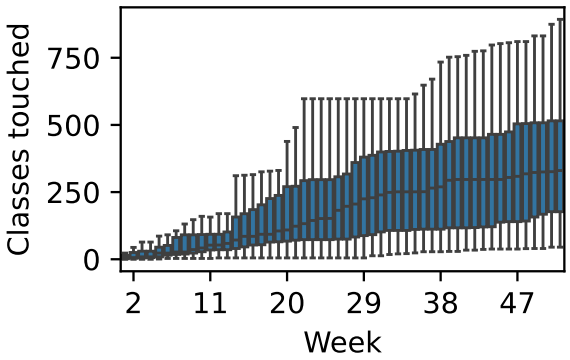
(b) Transient founder developers (13) showing classes touched



(c) Sustained founder developers (6) showing classes touched

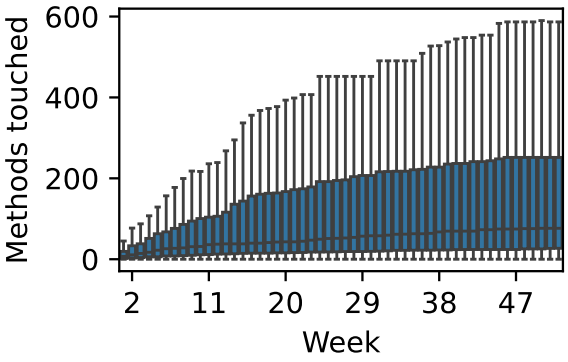


(d) Transient later joiner developers (204) showing classes touched

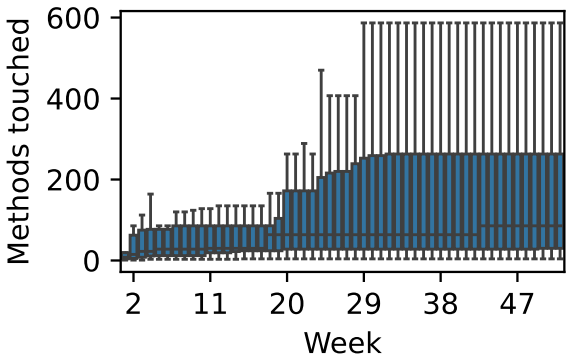


(e) Sustained later joiner developers (21) showing classes touched

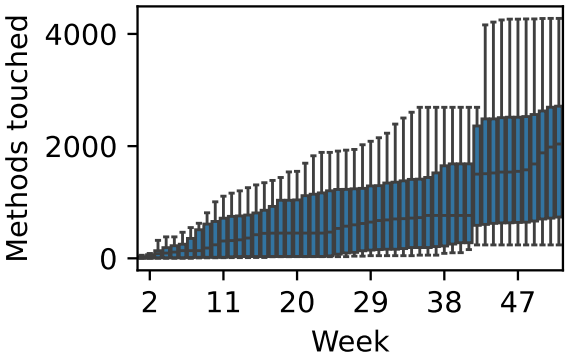
Figure 8: A box plot of total classes touched on mean each month, with quartile shading.



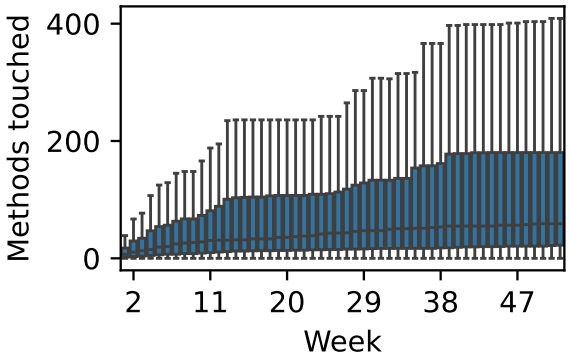
(a) All developers (244) showing methods touched



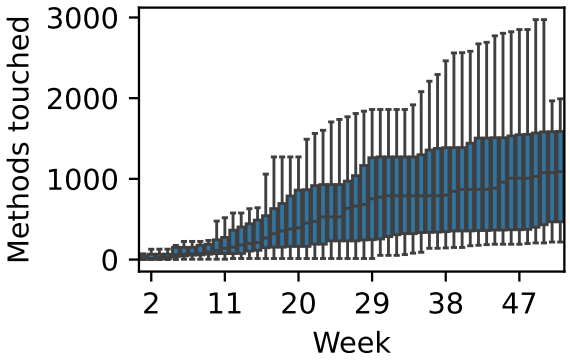
(b) Transient founder developers (13) showing methods touched



(c) Sustained founder developers (6) showing methods touched



(d) Transient later joiner developers (204) showing methods touched

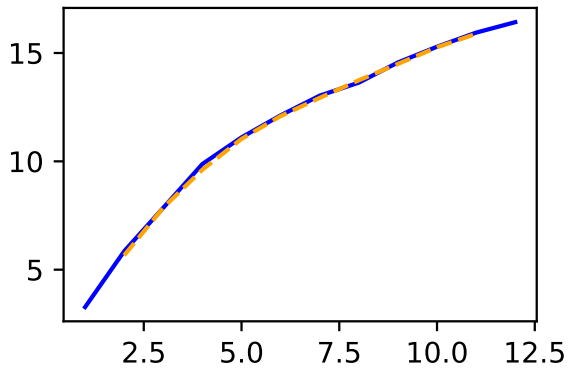


(e) Sustained later joiner developers (21) showing methods touched

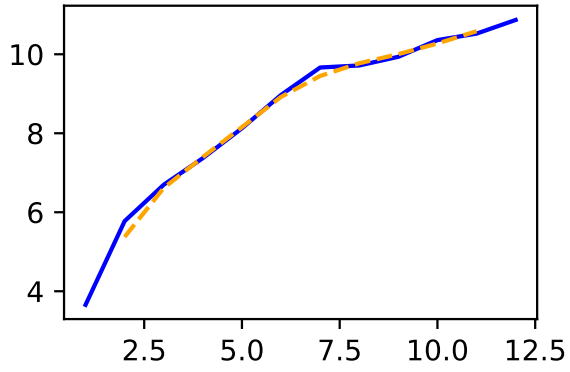
Figure 9: A box plot of total methods touched on mean each month, with quartile shading.

2.3 Smooth moving average - 14 For packages touched for each unit

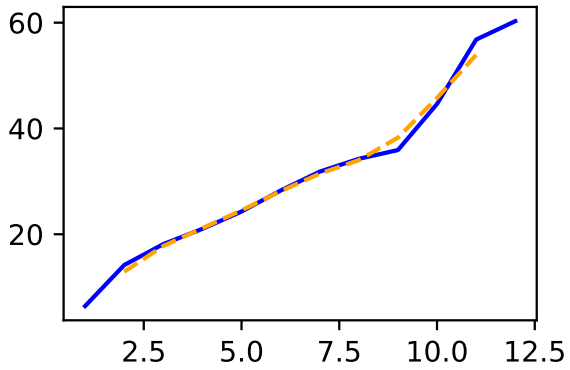
A smooth moving average packages touched on average each period the number of commits to components touched.



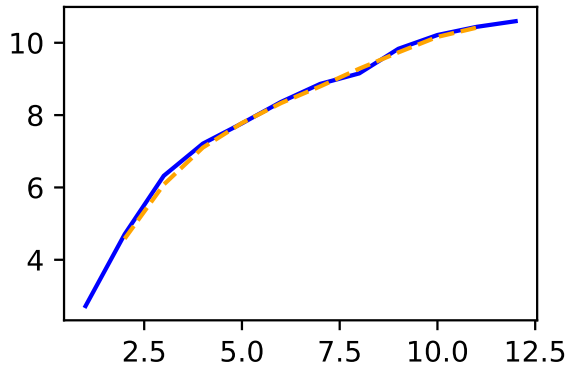
(a) All developers (244) showing packages touched



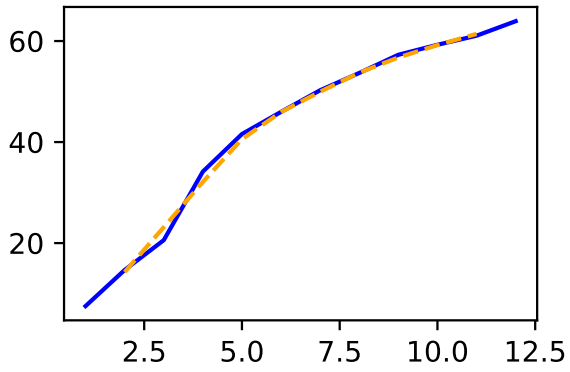
(b) Transient founder developers (13) showing packages touched



(c) Sustained founder developers (6) showing packages touched

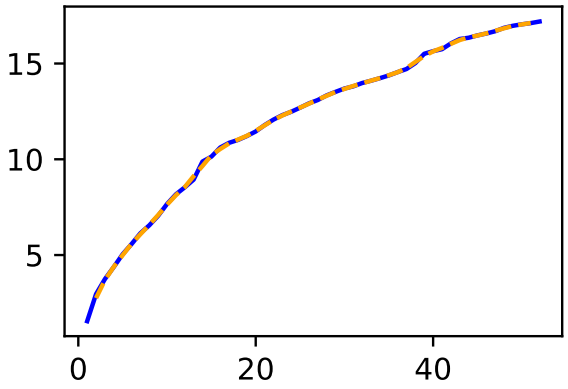


(d) Transient later joiner developers (204) showing packages touched

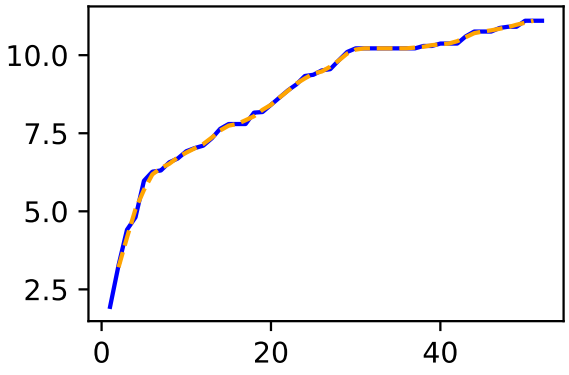


(e) Sustained later joiner developers (21) showing packages touched

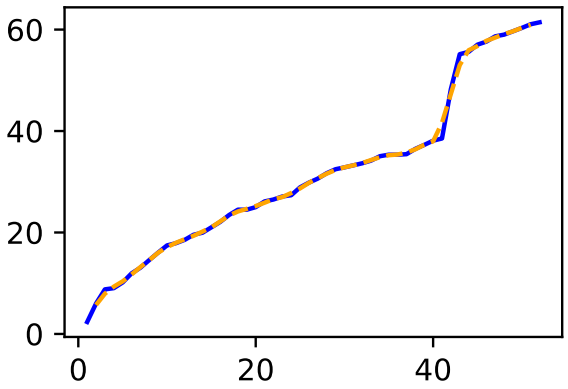
Figure 10: A smooth moving of total packages touched on average each month of a year.



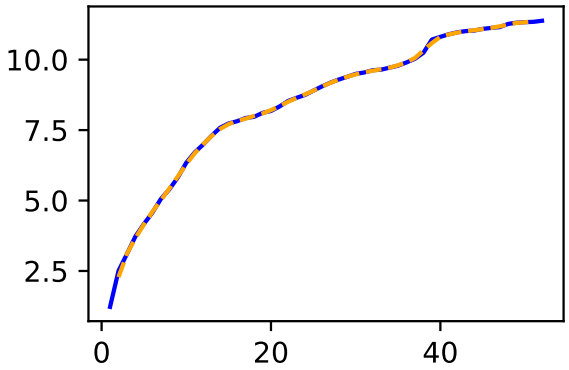
(a) All developers (244) showing packages touched



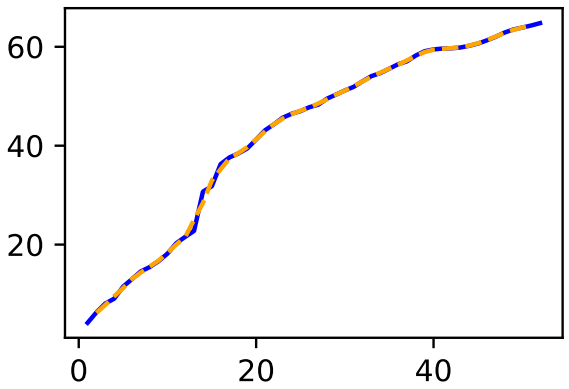
(b) Transient founder developers (13) showing packages touched



(c) Sustained founder developers (6) showing packages touched

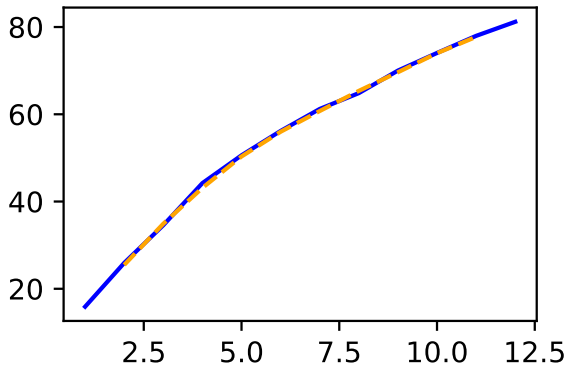


(d) Transient later joiner developers (204) showing packages touched

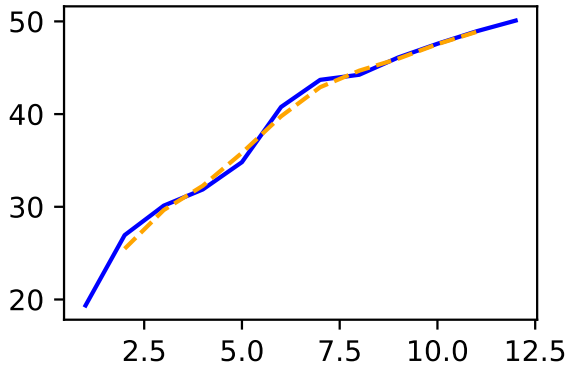


(e) Sustained later joiner developers (21) showing packages touched

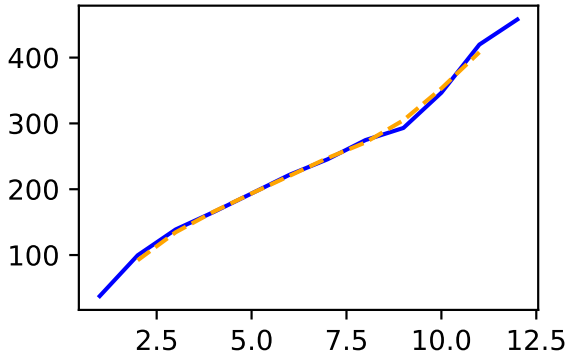
Figure 11: A smooth moving of total packages touched on average each week of a year.



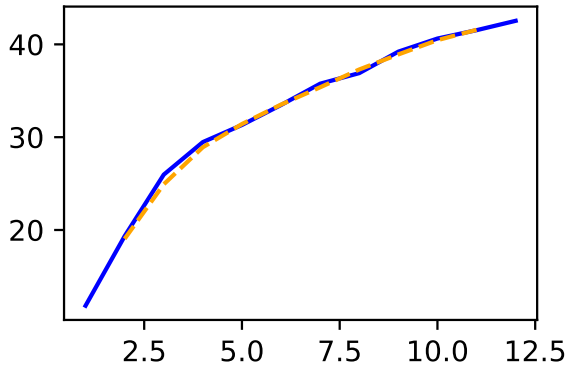
(a) All developers (244) showing classes touched



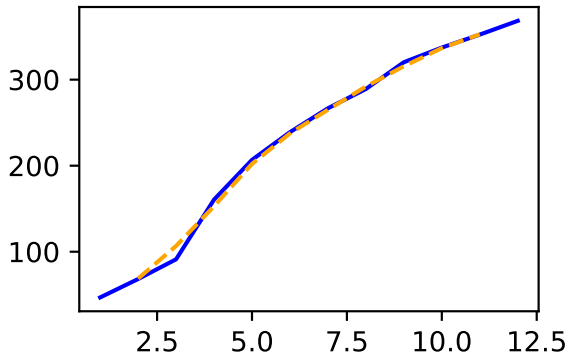
(b) Transient founder developers (13) showing classes touched



(c) Sustained founder developers (6) showing classes touched

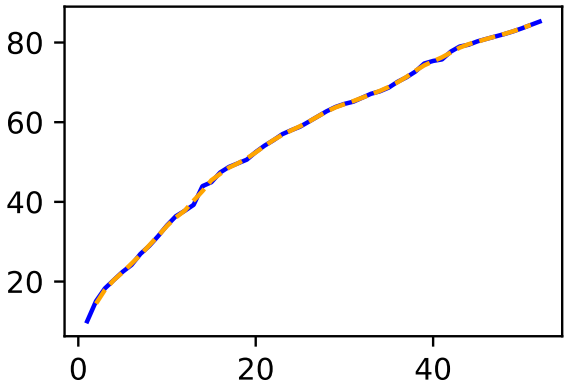


(d) Transient later joiner developers (204) showing classes touched

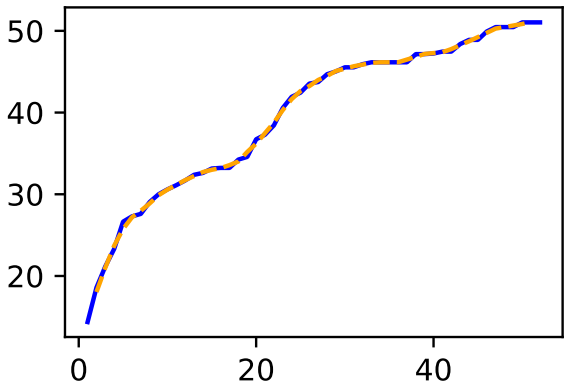


(e) Sustained later joiner developers (21) showing classes touched

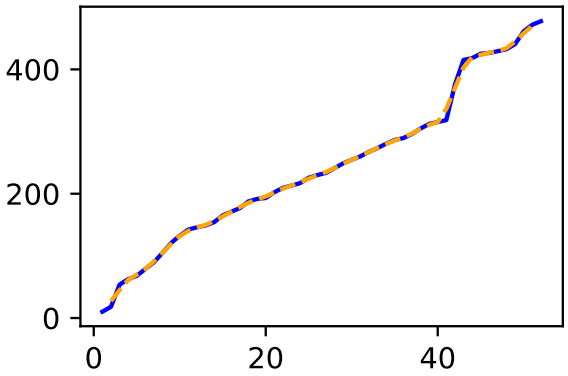
Figure 12: A smooth moving of total classes touched on average each month of a year.



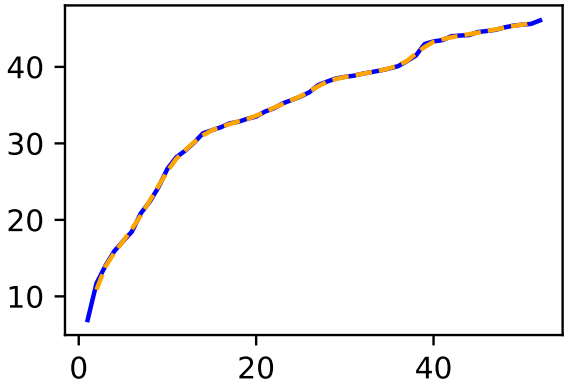
(a) All developers (244) showing classes touched



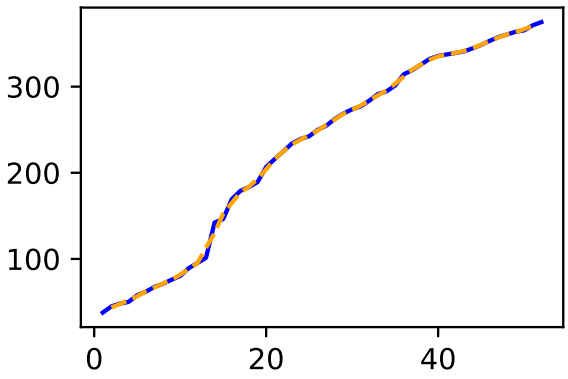
(b) Transient founder developers (13) showing classes touched



(c) Sustained founder developers (6) showing classes touched

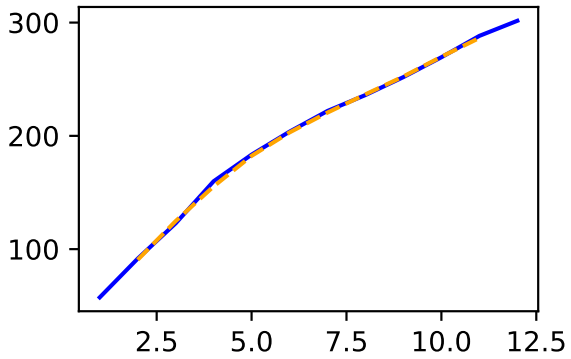


(d) Transient later joiner developers (204) showing classes touched

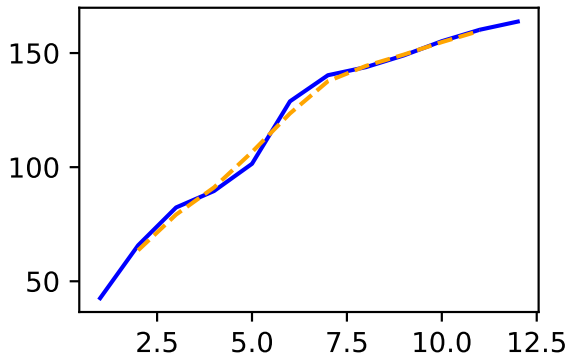


(e) Sustained later joiner developers (21) showing classes touched

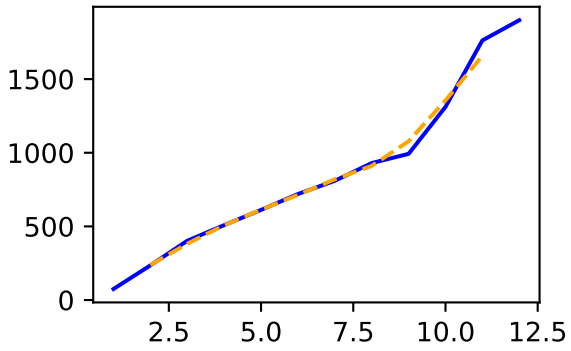
Figure 13: A smooth moving of total classes touched on average each week of a year.



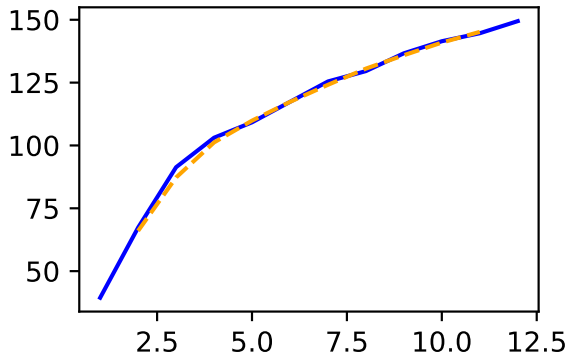
(a) All developers (244) showing methods touched



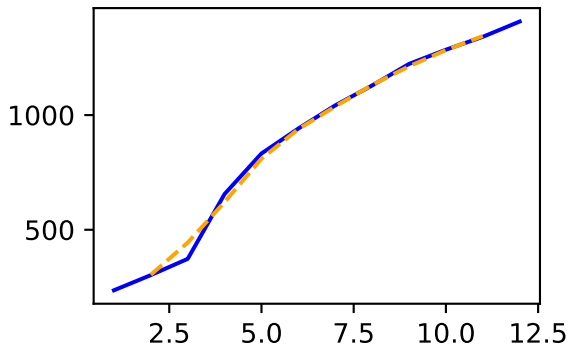
(b) Transient founder developers (13) showing methods touched



(c) Sustained founder developers (6) showing methods touched

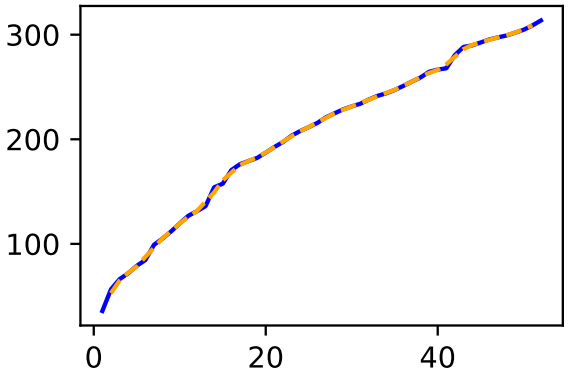


(d) Transient later joiner developers (204) showing methods touched

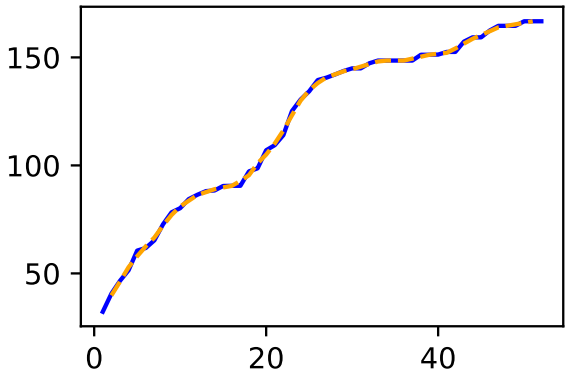


(e) Sustained later joiner developers (21) showing methods touched

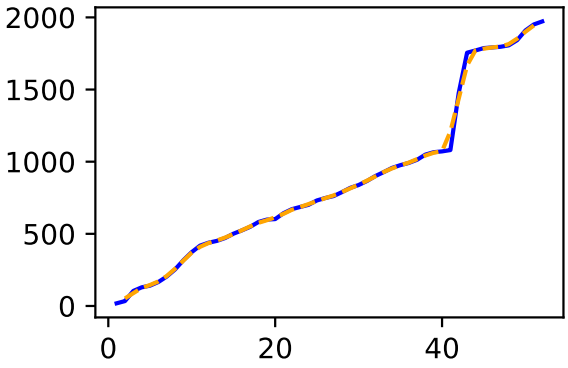
Figure 14: A smooth moving of total methods touched on average each month of a year.



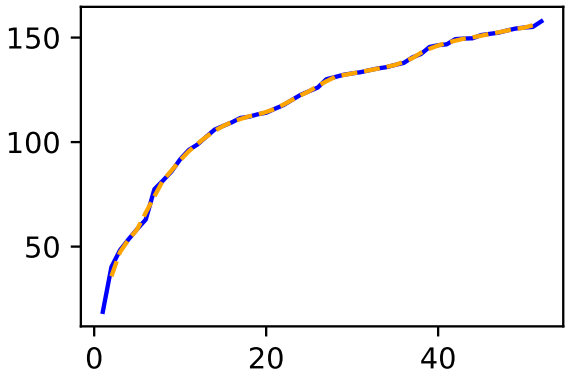
(a) All developers (244) showing methods touched



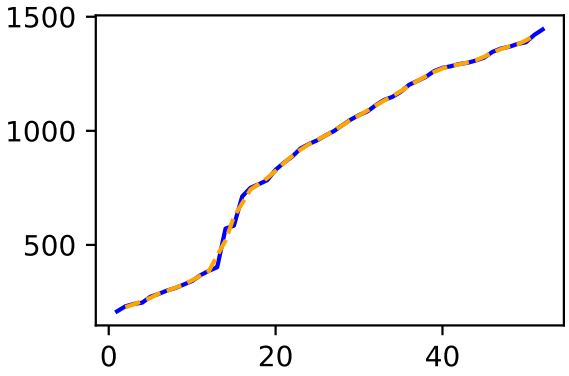
(b) Transient founder developers (13) showing methods touched



(c) Sustained founder developers (6) showing methods touched



(d) Transient later joiner developers (204) showing methods touched



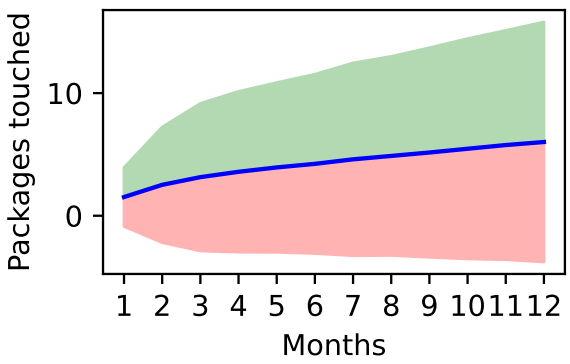
(e) Sustained later joiner developers (21) showing methods touched

Figure 15: A smooth moving of total methods touched on average each week of a year.

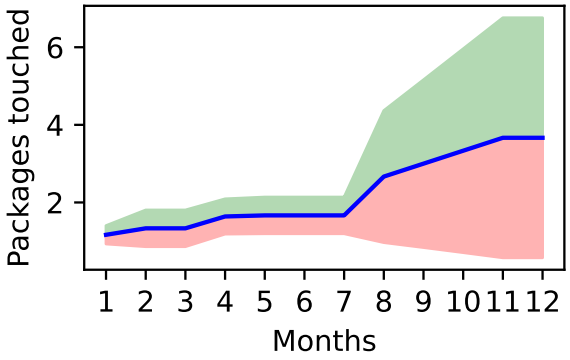
3 Repository: 21

3.1 Time series developer - 21 For packages touched for each period

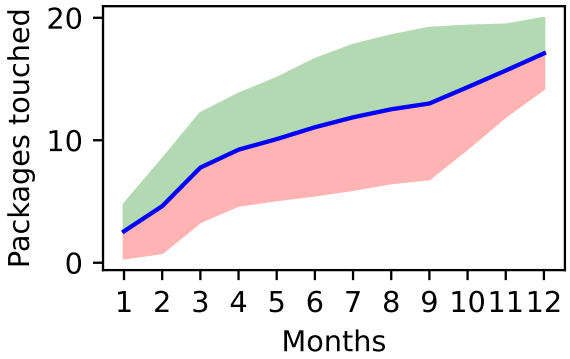
A time series of packages touched on average each month.



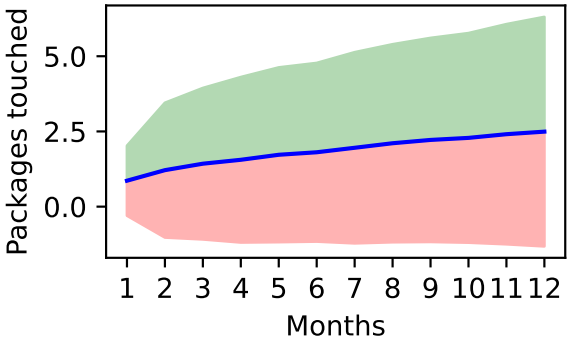
(a) All developers (300) showing packages touched



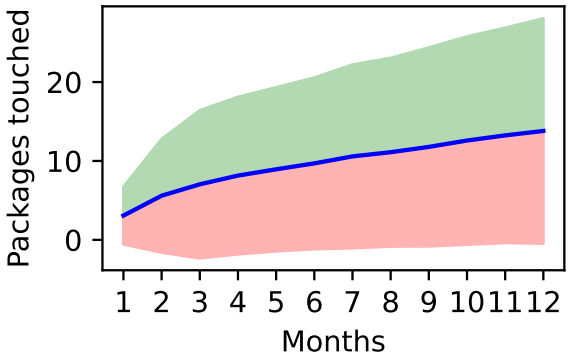
(b) Transient founder developers (3) showing packages touched



(c) Sustained founder developers (7) showing packages touched

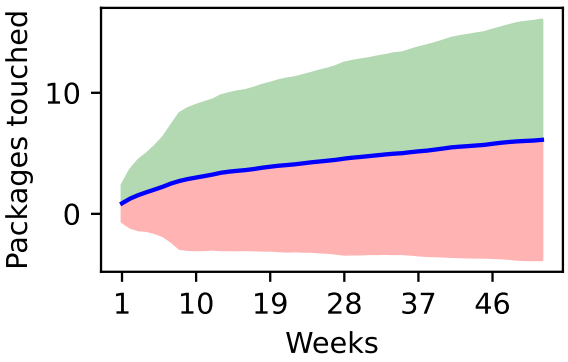


(d) Transient later joiner developers (206) showing packages touched

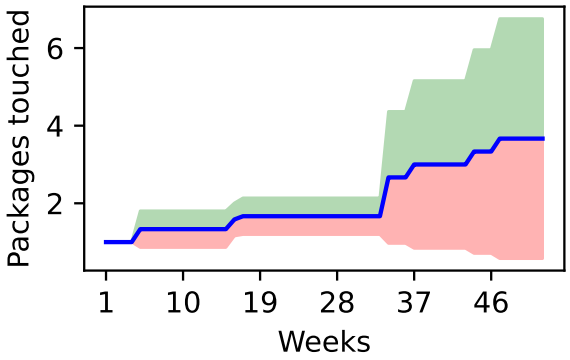


(e) Sustained later joiner developers (84) showing packages touched

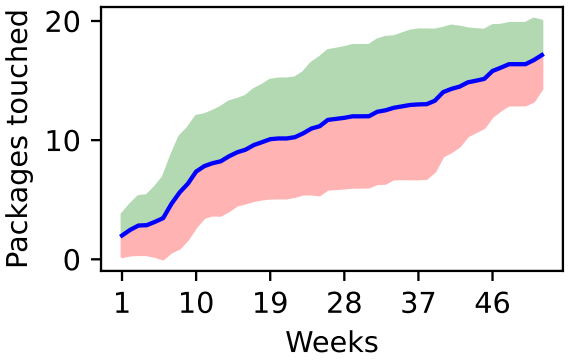
Figure 16: A time series of the average (mean) total packages touched on average each month, with positive (orange) and negative (red) filled standard deviation.



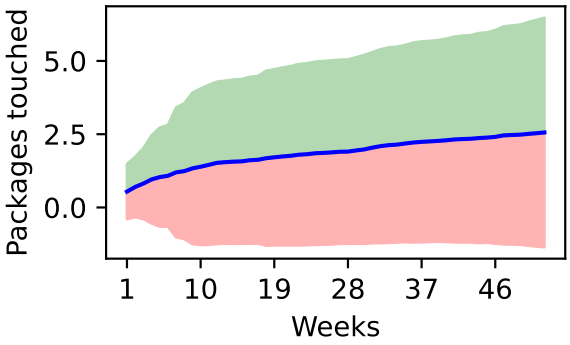
(a) All developers (300) showing packages touched



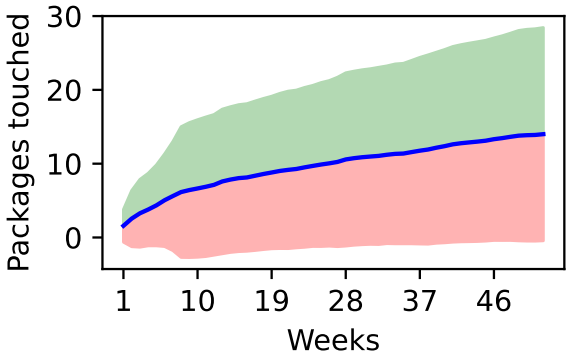
(b) Transient founder developers (3) showing packages touched



(c) Sustained founder developers (7) showing packages touched

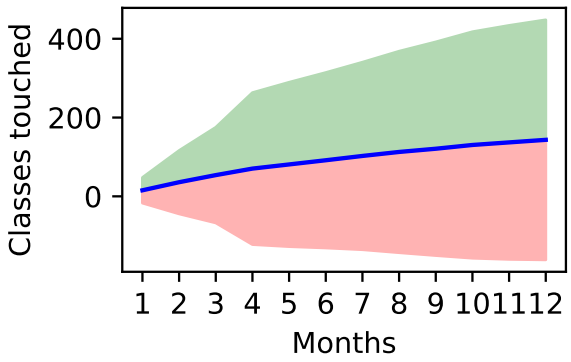


(d) Transient later joiner developers (206) showing packages touched

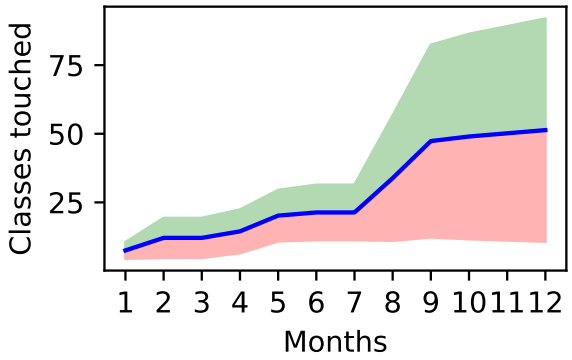


(e) Sustained later joiner developers (84) showing packages touched

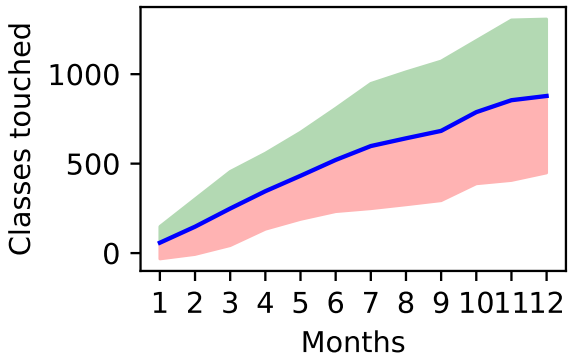
Figure 17: A time series of the average (mean) total packages touched on average each week, with positive (orange) and negative (red) filled standard deviation.



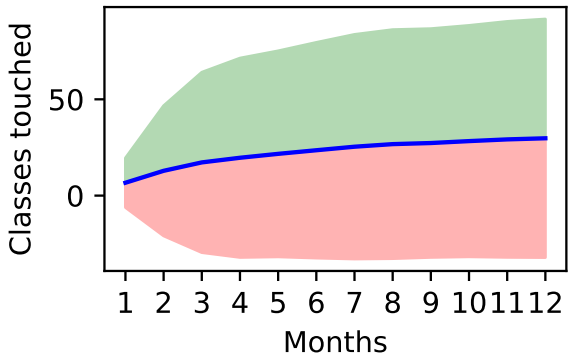
(a) All developers (300) showing classes touched



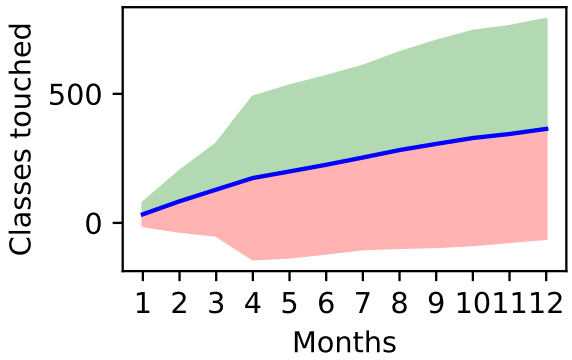
(b) Transient founder developers (3) showing classes touched



(c) Sustained founder developers (7) showing classes touched

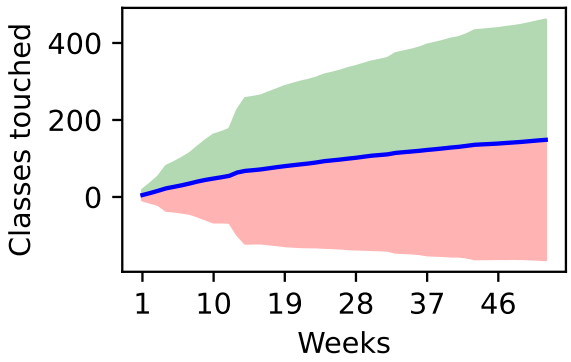


(d) Transient later joiner developers (206) showing classes touched

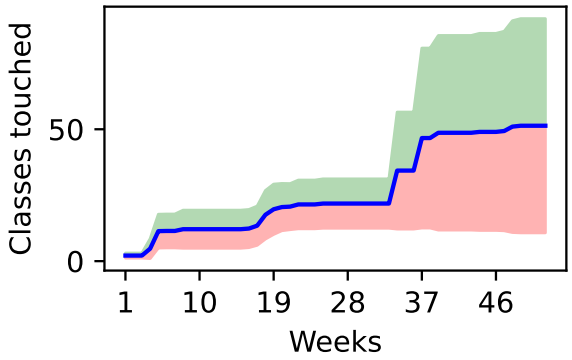


(e) Sustained later joiner developers (84) showing classes touched

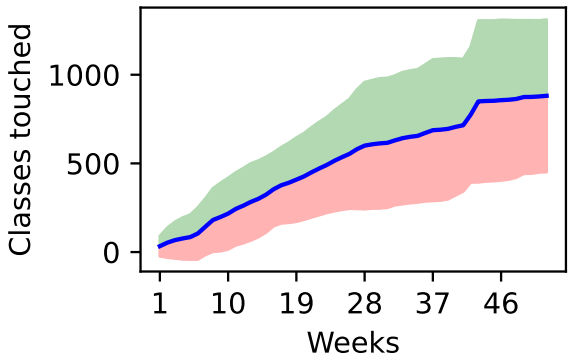
Figure 18: A time series of the average (mean) total classes touched on average each month, with positive (orange) and negative (red) filled standard deviation.



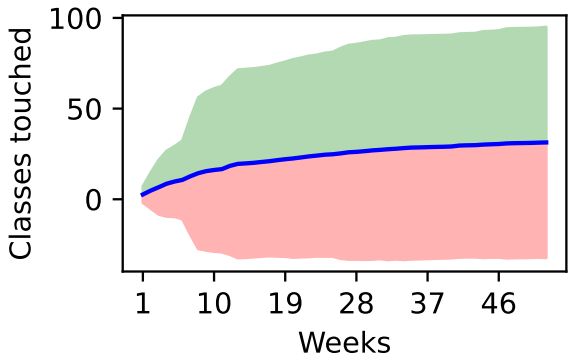
(a) All developers (300) showing classes touched



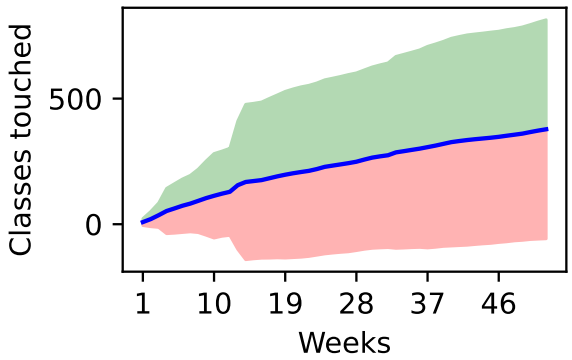
(b) Transient founder developers (3) showing classes touched



(c) Sustained founder developers (7) showing classes touched

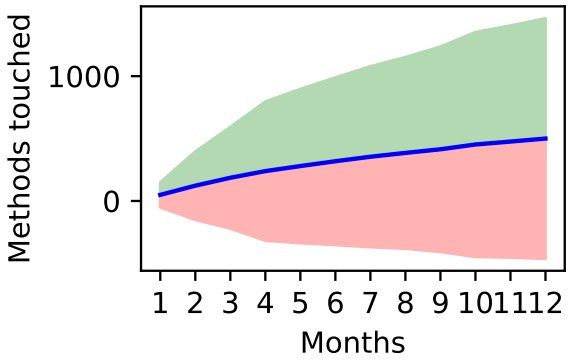


(d) Transient later joiner developers (206) showing classes touched

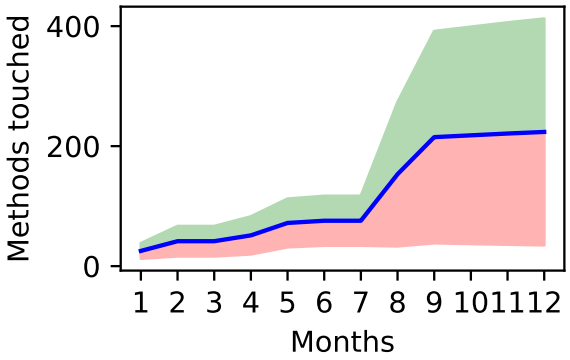


(e) Sustained later joiner developers (84) showing classes touched

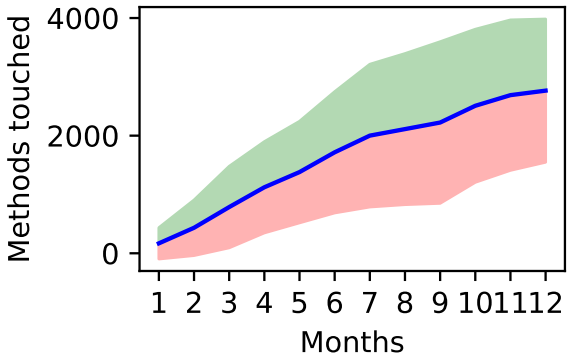
Figure 19: A time series of the average (mean) total classes touched on average each week, with positive (orange) and negative (red) filled standard deviation.



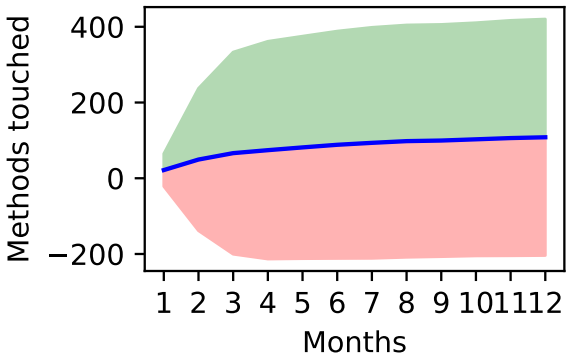
(a) All developers (300) showing methods touched



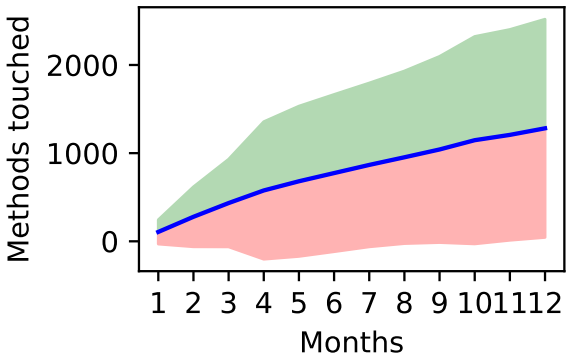
(b) Transient founder developers (3) showing methods touched



(c) Sustained founder developers (7) showing methods touched

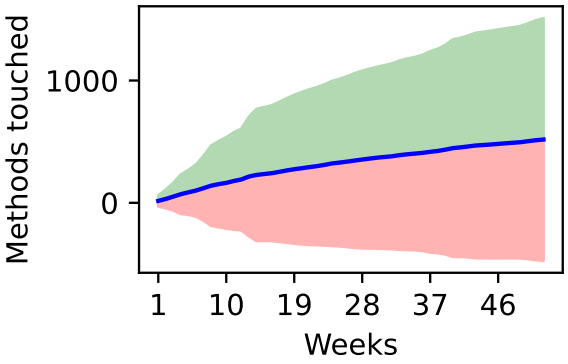


(d) Transient later joiner developers (206) showing methods touched

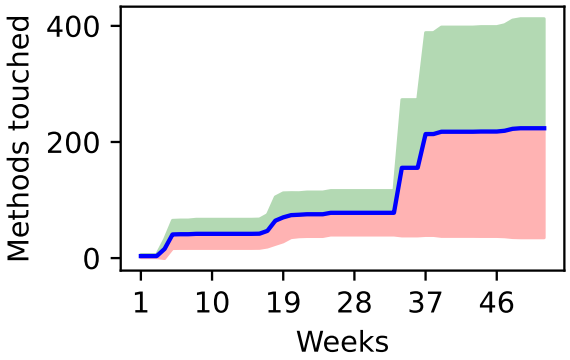


(e) Sustained later joiner developers (84) showing methods touched

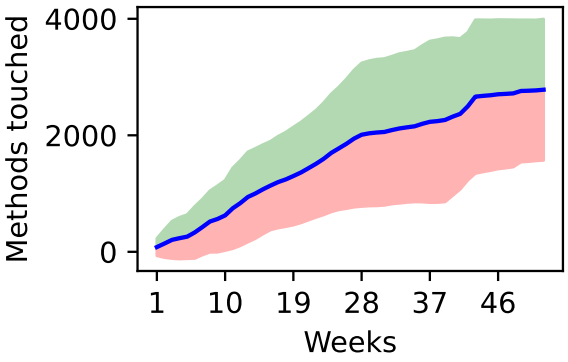
Figure 20: A time series of the average (mean) total methods touched on average each month, with positive (orange) and negative (red) filled standard deviation.



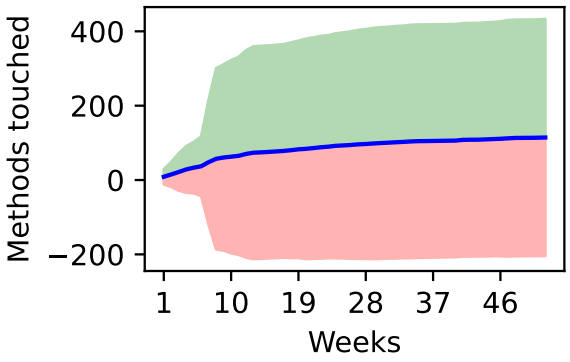
(a) All developers (300) showing methods touched



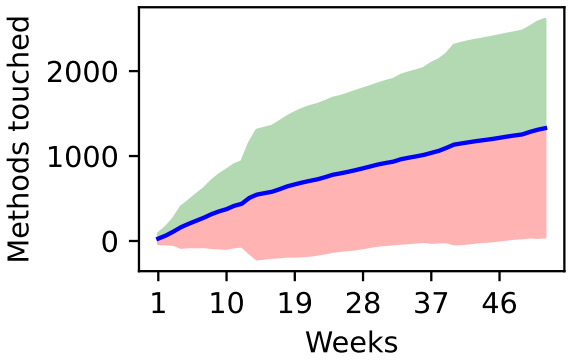
(b) Transient founder developers (3) showing methods touched



(c) Sustained founder developers (7) showing methods touched



(d) Transient later joiner developers (206) showing methods touched

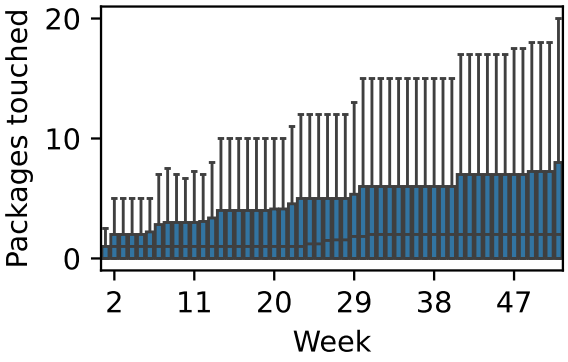


(e) Sustained later joiner developers (84) showing methods touched

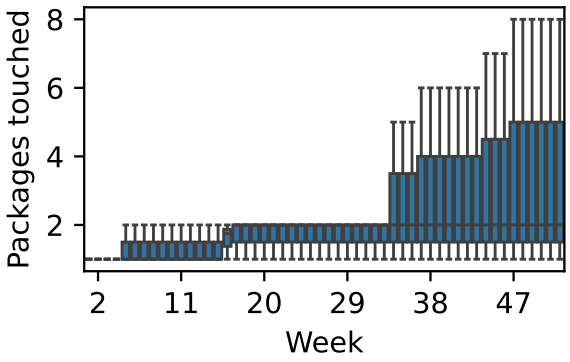
Figure 21: A time series of the average (mean) total methods touched on average each week, with positive (orange) and negative (red) filled standard deviation.

3.2 Box plot developer - 21 For packages touched for each period

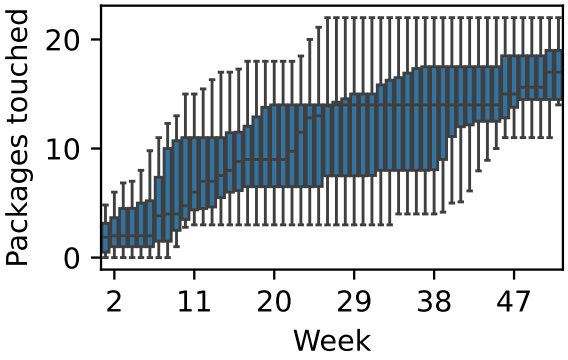
A box plot of packages touched on average each period the number of commits to components touched.



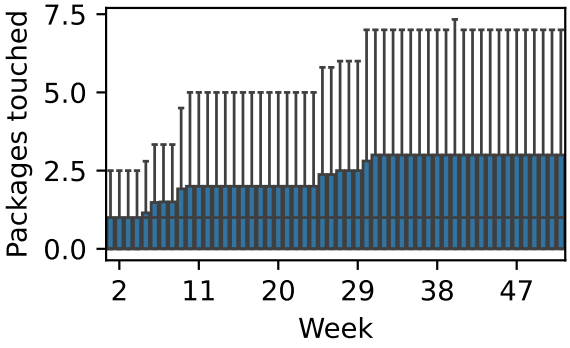
(a) All developers (300) showing packages touched



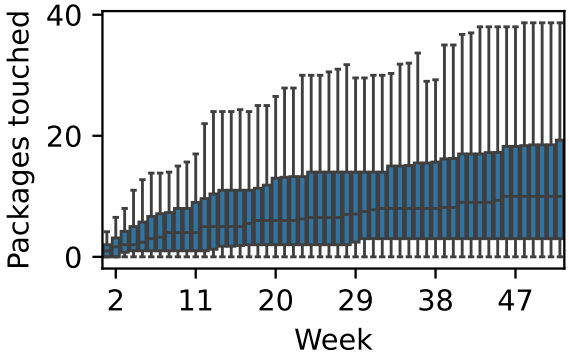
(b) Transient founder developers (3) showing packages touched



(c) Sustained founder developers (7) showing packages touched

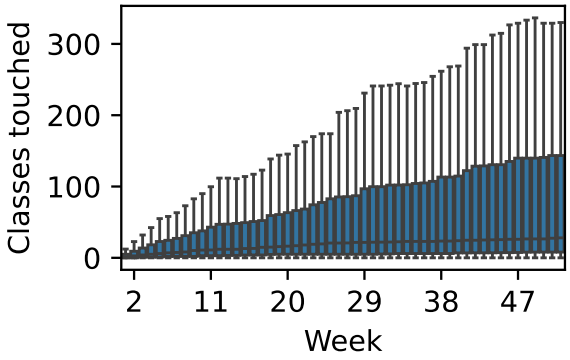


(d) Transient later joiner developers (206) showing packages touched

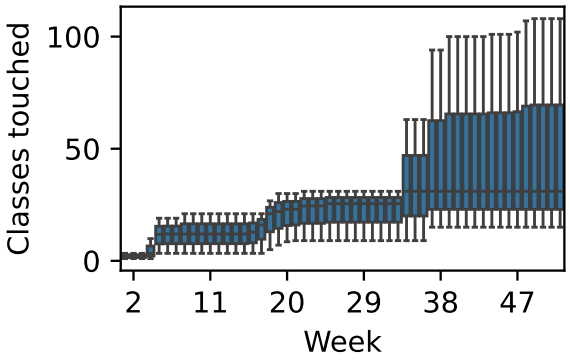


(e) Sustained later joiner developers (84) showing packages touched

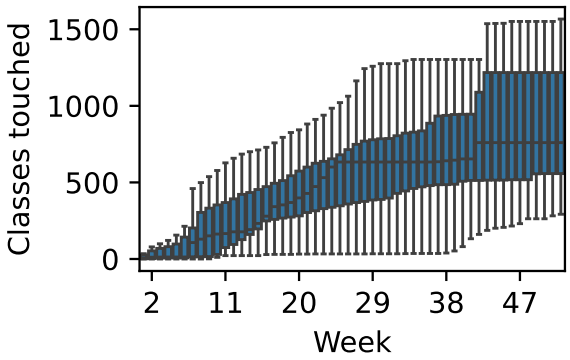
Figure 22: A box plot of total packages touched on mean each month, with quartile shading.



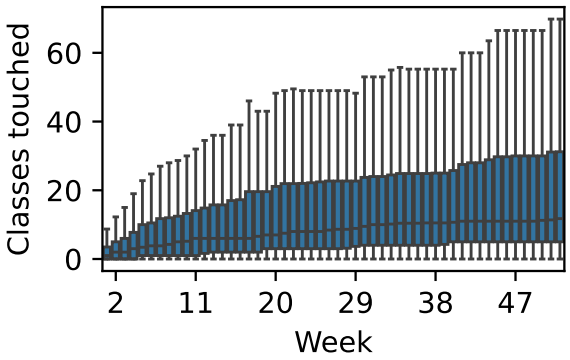
(a) All developers (300) showing classes touched



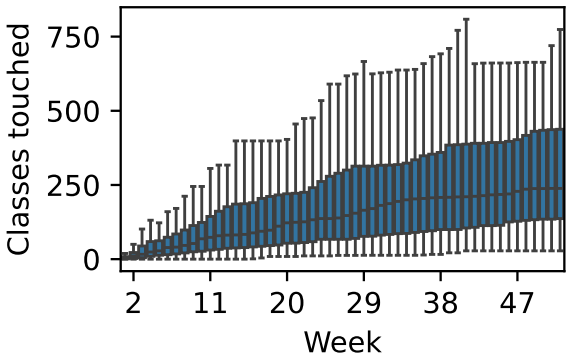
(b) Transient founder developers (3) showing classes touched



(c) Sustained founder developers (7) showing classes touched

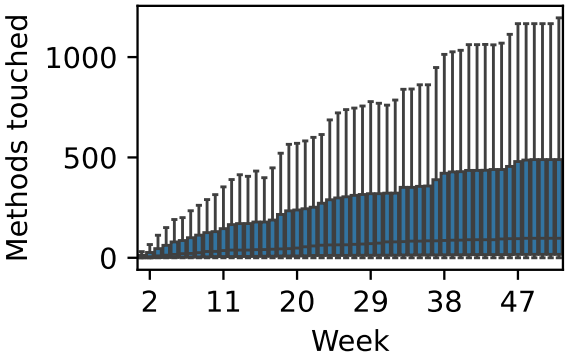


(d) Transient later joiner developers (206) showing classes touched

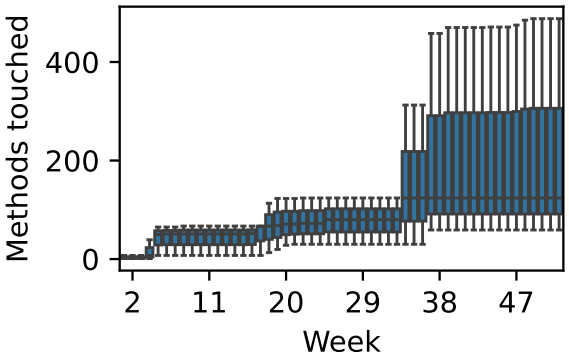


(e) Sustained later joiner developers (84) showing classes touched

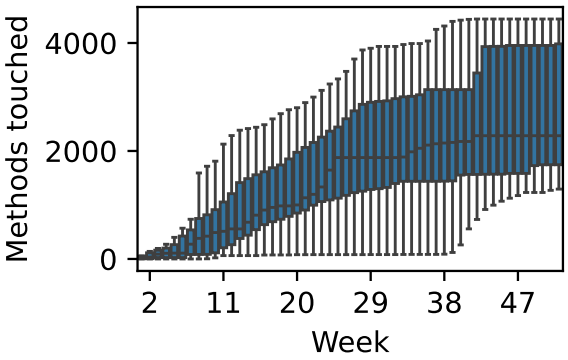
Figure 23: A box plot of total classes touched on mean each month, with quartile shading.



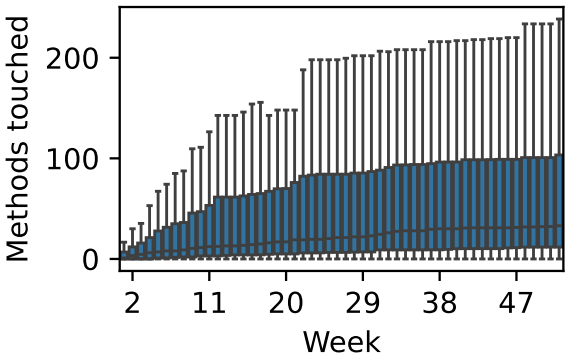
(a) All developers (300) showing methods touched



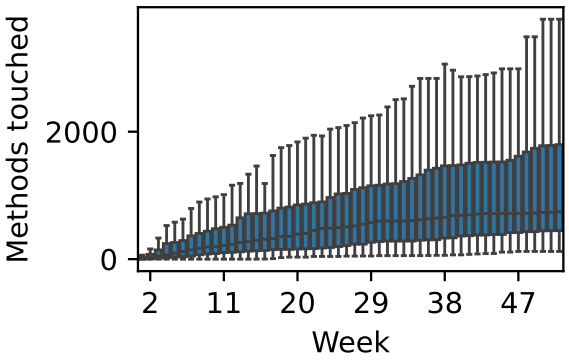
(b) Transient founder developers (3) showing methods touched



(c) Sustained founder developers (7) showing methods touched



(d) Transient later joiner developers (206) showing methods touched

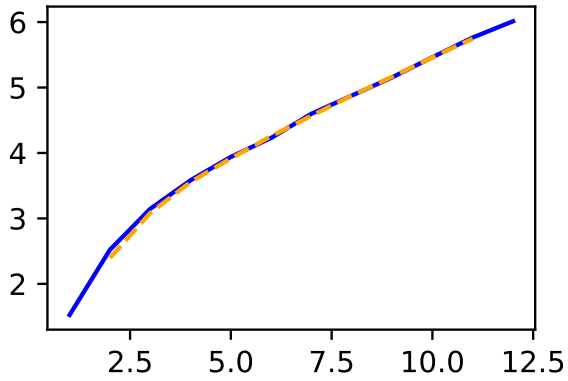


(e) Sustained later joiner developers (84) showing methods touched

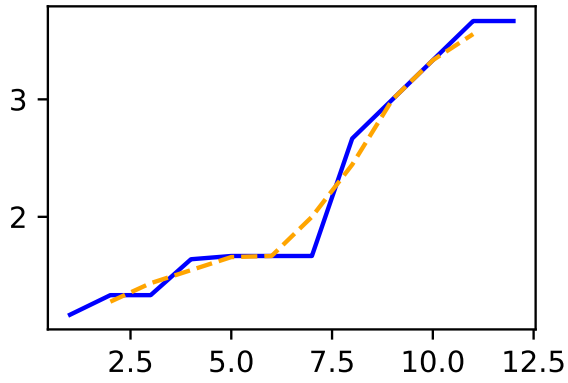
Figure 24: A box plot of total methods touched on mean each month, with quartile shading.

3.3 Smooth moving average - 21 For packages touched for each unit

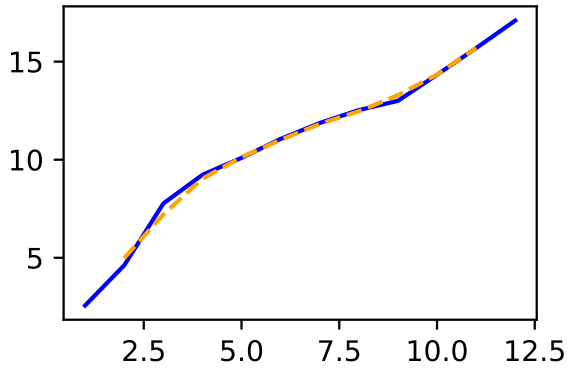
A smooth moving average packages touched on average each period the number of commits to components touched.



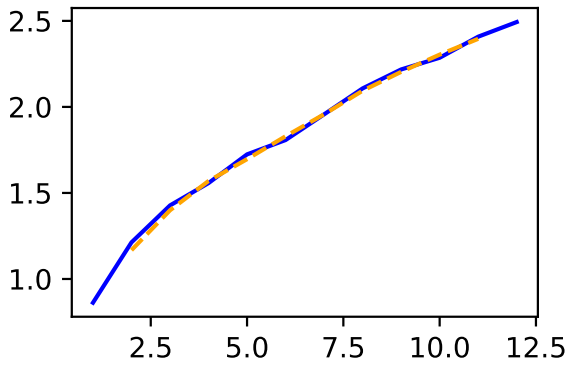
(a) All developers (300) showing packages touched



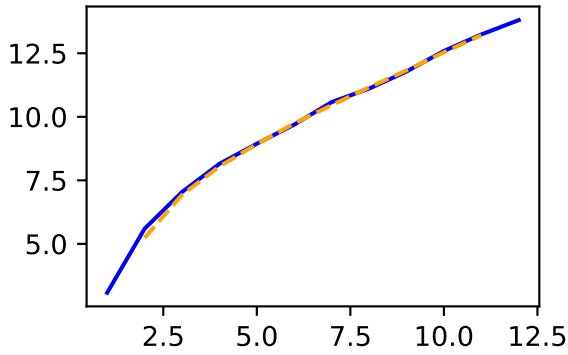
(b) Transient founder developers (3) showing packages touched



(c) Sustained founder developers (7) showing packages touched

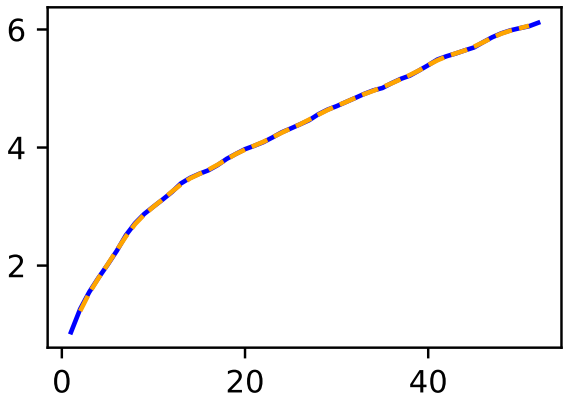


(d) Transient later joiner developers (206) showing packages touched

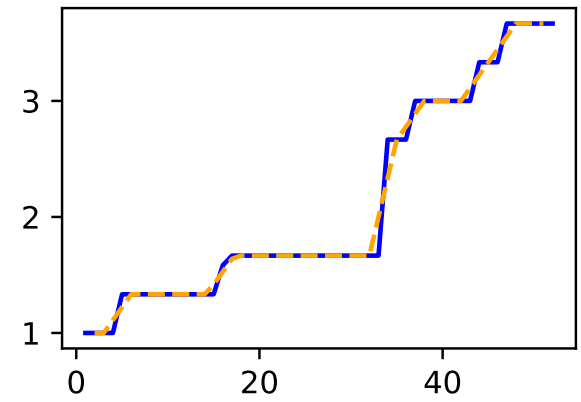


(e) Sustained later joiner developers (84) showing packages touched

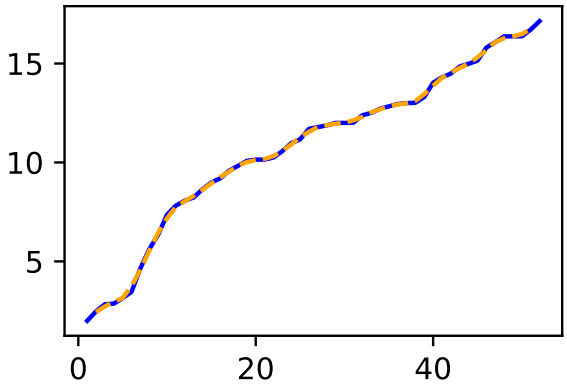
Figure 25: A smooth moving of total packages touched on average each month of a year.



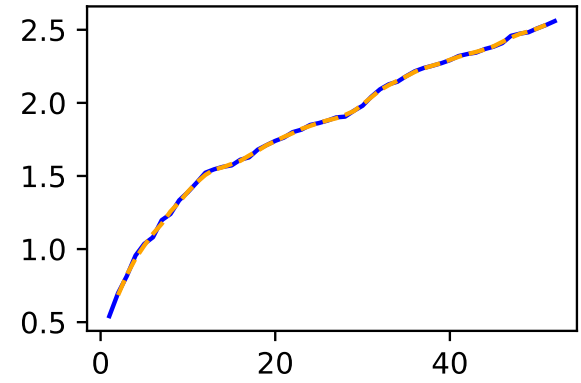
(a) All developers (300) showing packages touched



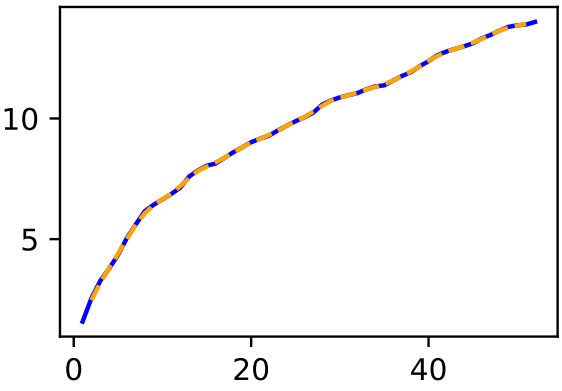
(b) Transient founder developers (3) showing packages touched



(c) Sustained founder developers (7) showing packages touched

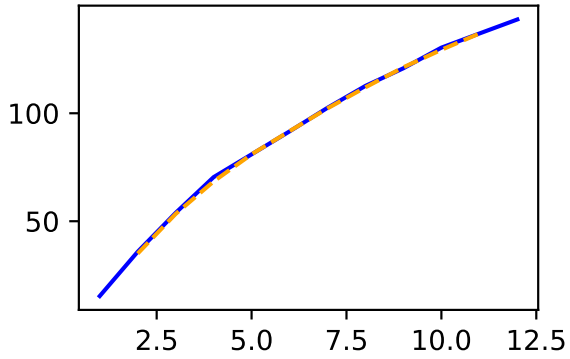


(d) Transient later joiner developers (206) showing packages touched

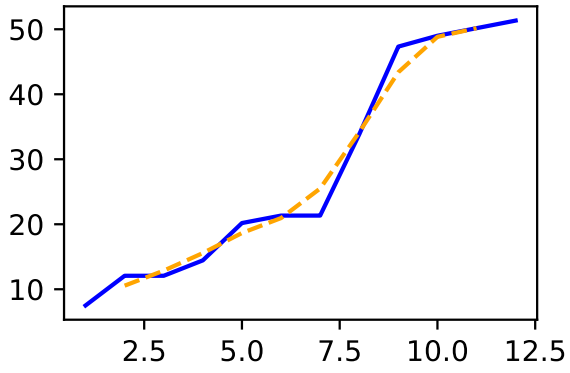


(e) Sustained later joiner developers (84) showing packages touched

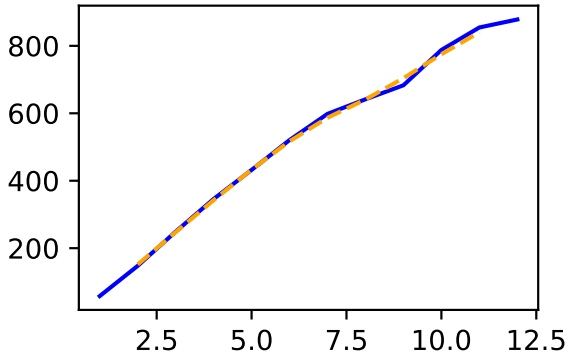
Figure 26: A smooth moving of total packages touched on average each week of a year.



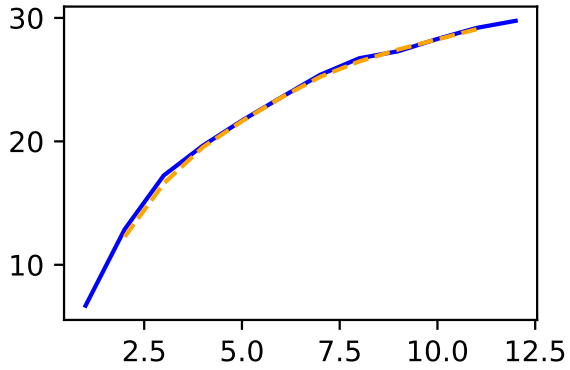
(a) All developers (300) showing classes touched



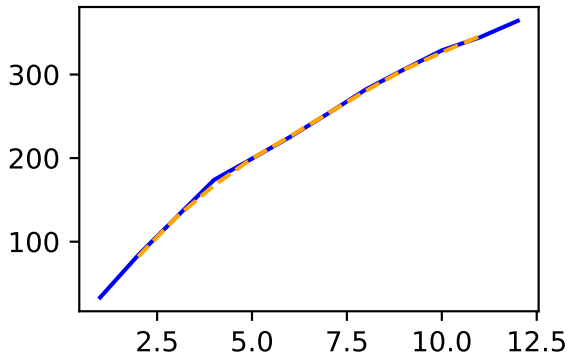
(b) Transient founder developers (3) showing classes touched



(c) Sustained founder developers (7) showing classes touched

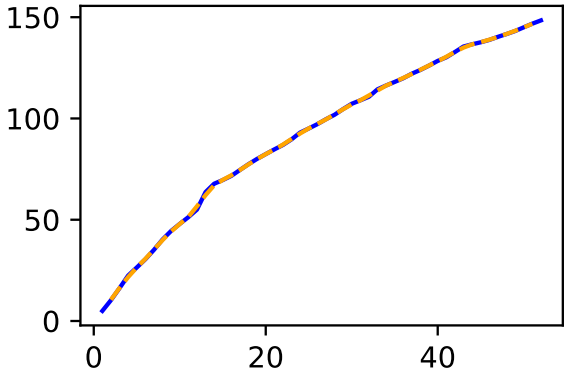


(d) Transient later joiner developers (206) showing classes touched

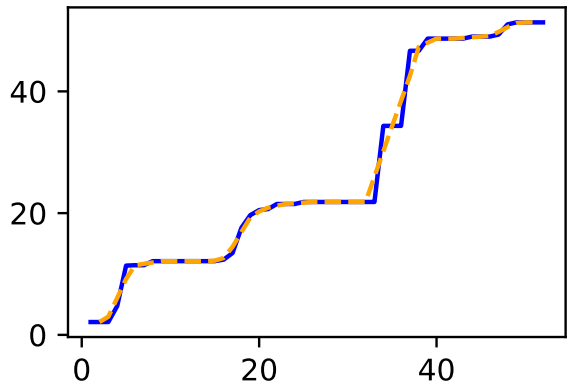


(e) Sustained later joiner developers (84) showing classes touched

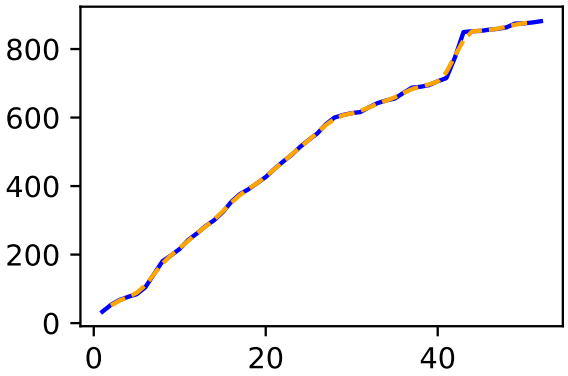
Figure 27: A smooth moving of total classes touched on average each month of a year.



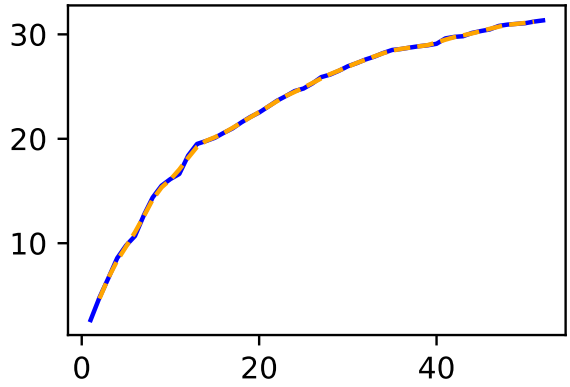
(a) All developers (300) showing classes touched



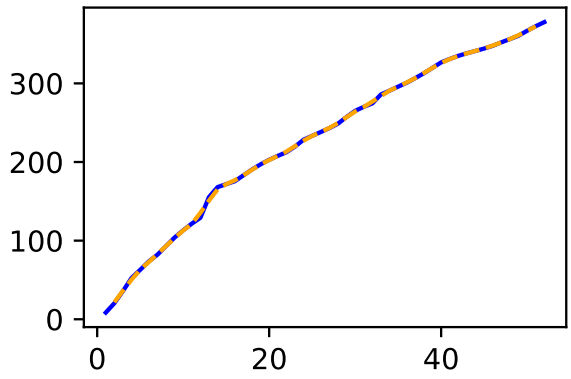
(b) Transient founder developers (3) showing classes touched



(c) Sustained founder developers (7) showing classes touched

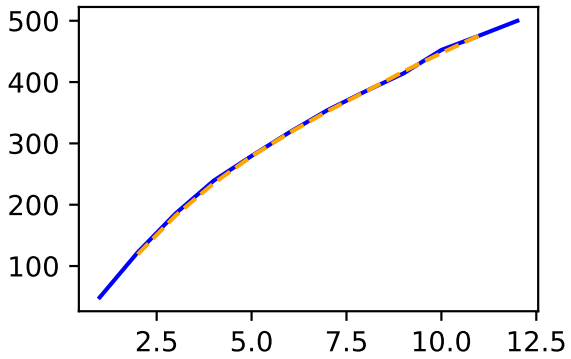


(d) Transient later joiner developers (206) showing classes touched

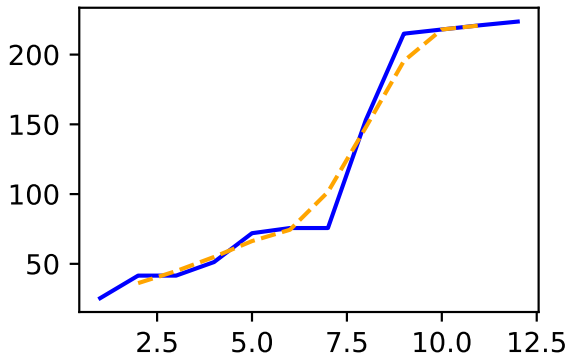


(e) Sustained later joiner developers (84) showing classes touched

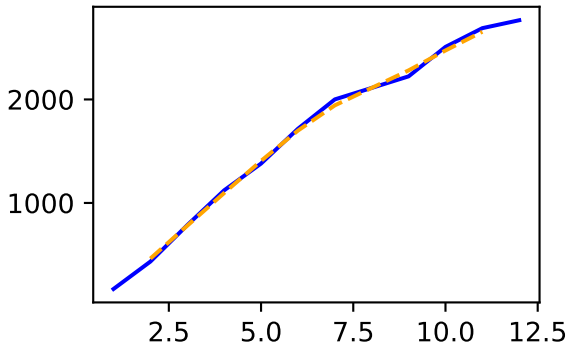
Figure 28: A smooth moving of total classes touched on average each week of a year.



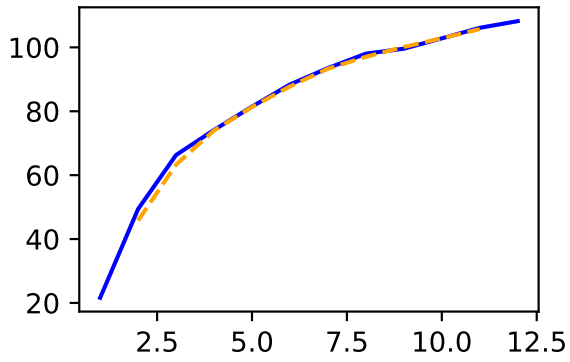
(a) All developers (300) showing methods touched



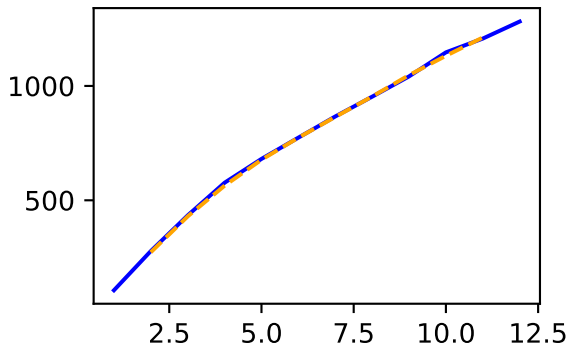
(b) Transient founder developers (3) showing methods touched



(c) Sustained founder developers (7) showing methods touched

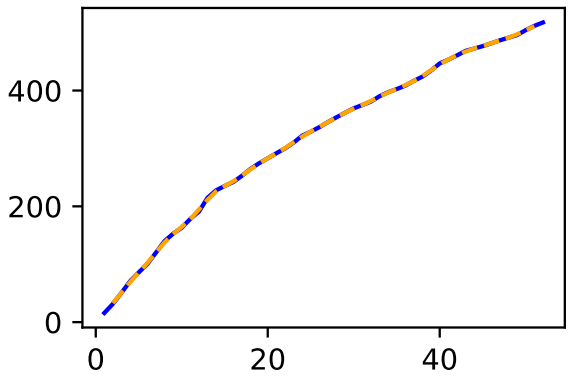


(d) Transient later joiner developers (206) showing methods touched

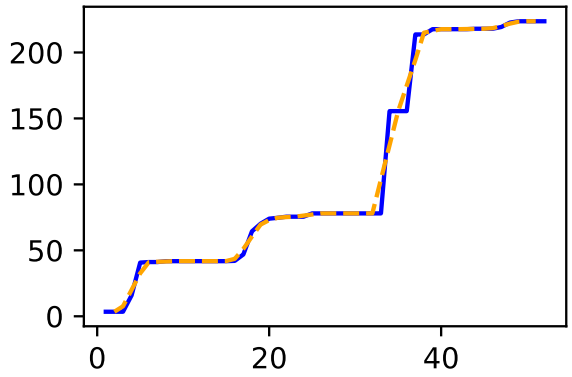


(e) Sustained later joiner developers (84) showing methods touched

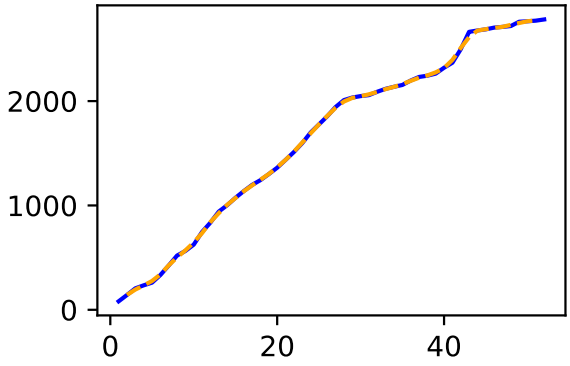
Figure 29: A smooth moving of total methods touched on average each month of a year.



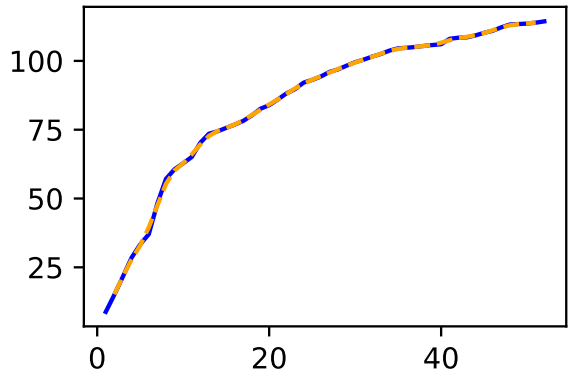
(a) All developers (300) showing methods touched



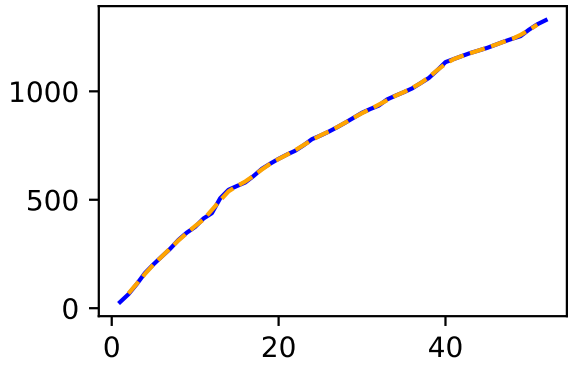
(b) Transient founder developers (3) showing methods touched



(c) Sustained founder developers (7) showing methods touched



(d) Transient later joiner developers (206) showing methods touched



(e) Sustained later joiner developers (84) showing methods touched

Figure 30: A smooth moving of total methods touched on average each week of a year.

4 **Repository: 3**

5 **Repository: 5**

6 **Repository: 7**

7 **Repository: 10**

8 **Repository: 15**

9 **Repository: 18**

10 **Repository: 20**

11 **Repository: 26**

12 **Repository: 28**

13 **Repository: 33**

14 **Repository: 36**

15 **Repository: 37**

16 **Repository: 39**

17 **Repository pull request - Pull requests**

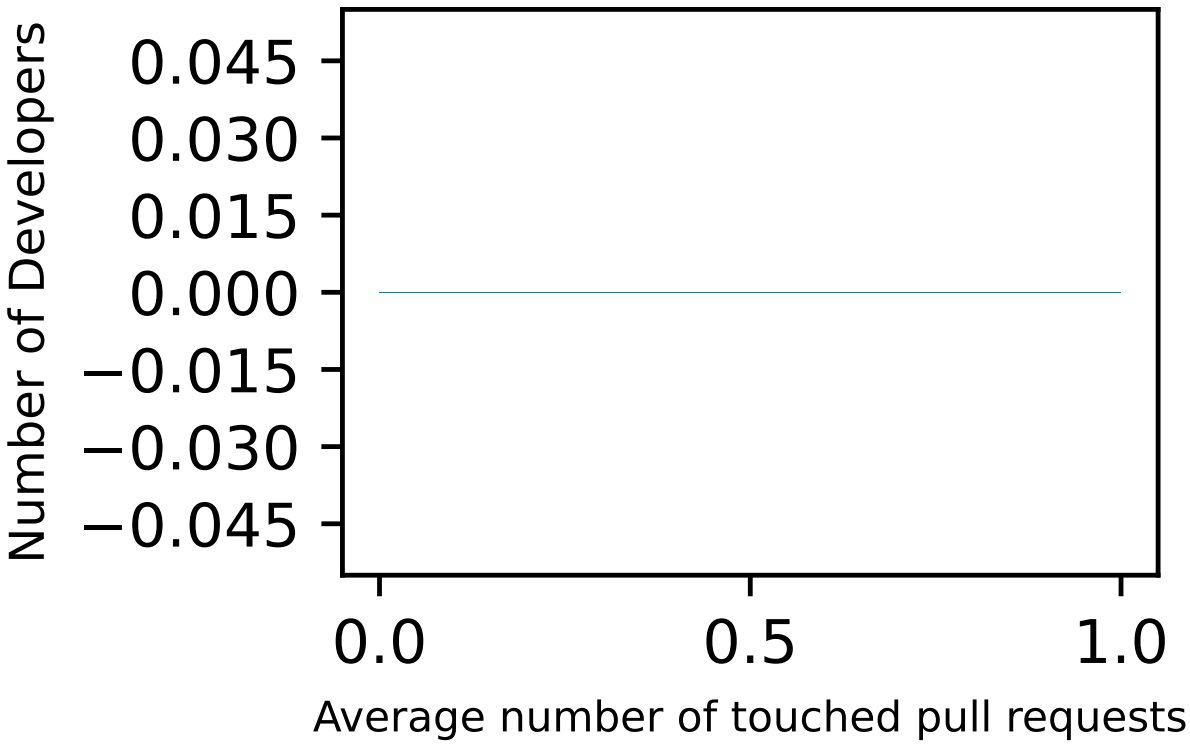


Figure 31: Average number of pull request for 0 developers.