

Analysis results

This file contains analytical summary for given repository. It consists of following sections:

- **Peak contribution times** - plots and tables showing relation between number of commits and time of a day / day of a week
- **Contributors activity and productivity** - statistics allowing to assess the activity and productivity of contributors. The top n number might be changed in the config/reports_config.py file. List of tables:
 - Table showing top n contributors in terms of number of commits
 - Table showing top n contributors in terms of number of insertions
 - Table showing top n contributors in terms of insertions / deletions ratio
 - Tables showing top n contributors in terms of number of commits / insertions per day of activity
- **Messages patterns** - the most popular patterns in commit messages, both in tabular form and as 'wordcloud' plots
- **Periods of activity before merge** - relation between number of commits and time to nearest merge
- **Insertions / deletions stats** - distributions of insertions / deletions per commit in the form of histogram, plus some basic stats, such as median and avg value

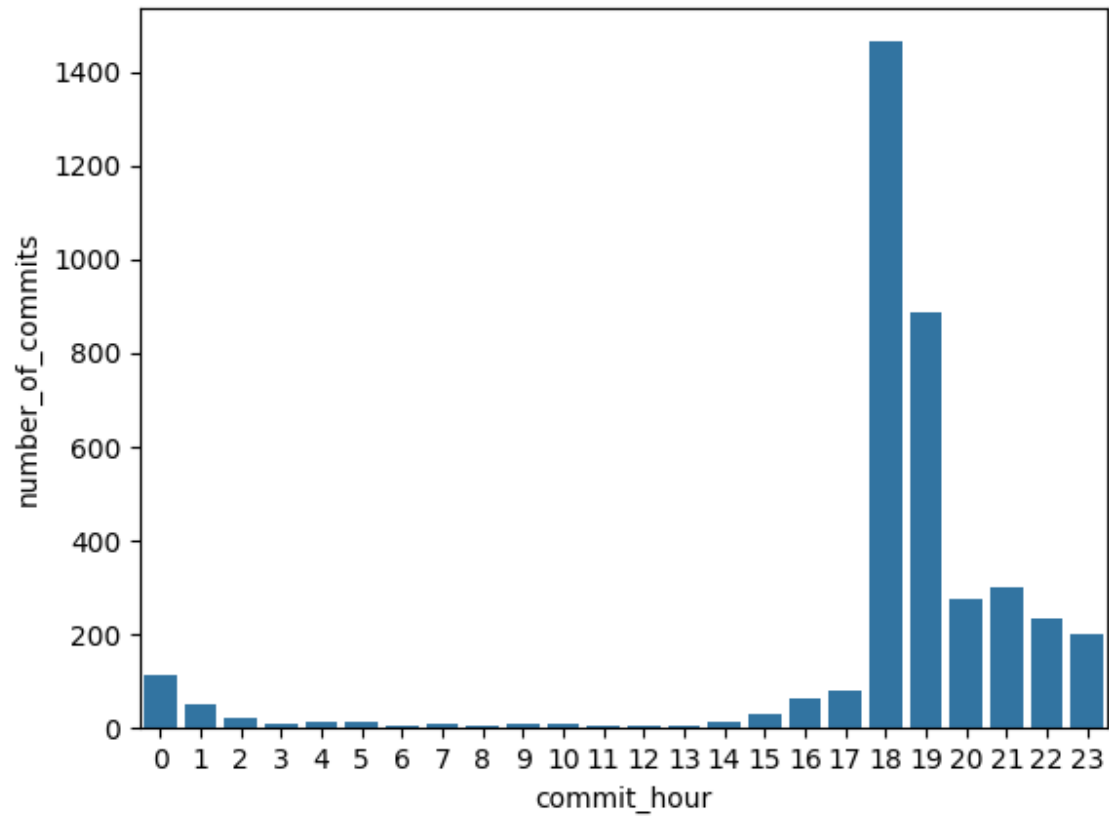
1. Peak contribution times

Table below shows the sum of commits per time of a day:

commit_hour	number_of_commits
0	114
1	50
2	21
3	11
4	14
5	14
6	6
7	9
8	6
9	9
10	8
11	4
12	3
13	6
14	13
15	31
16	62
17	80
18	1465
19	886

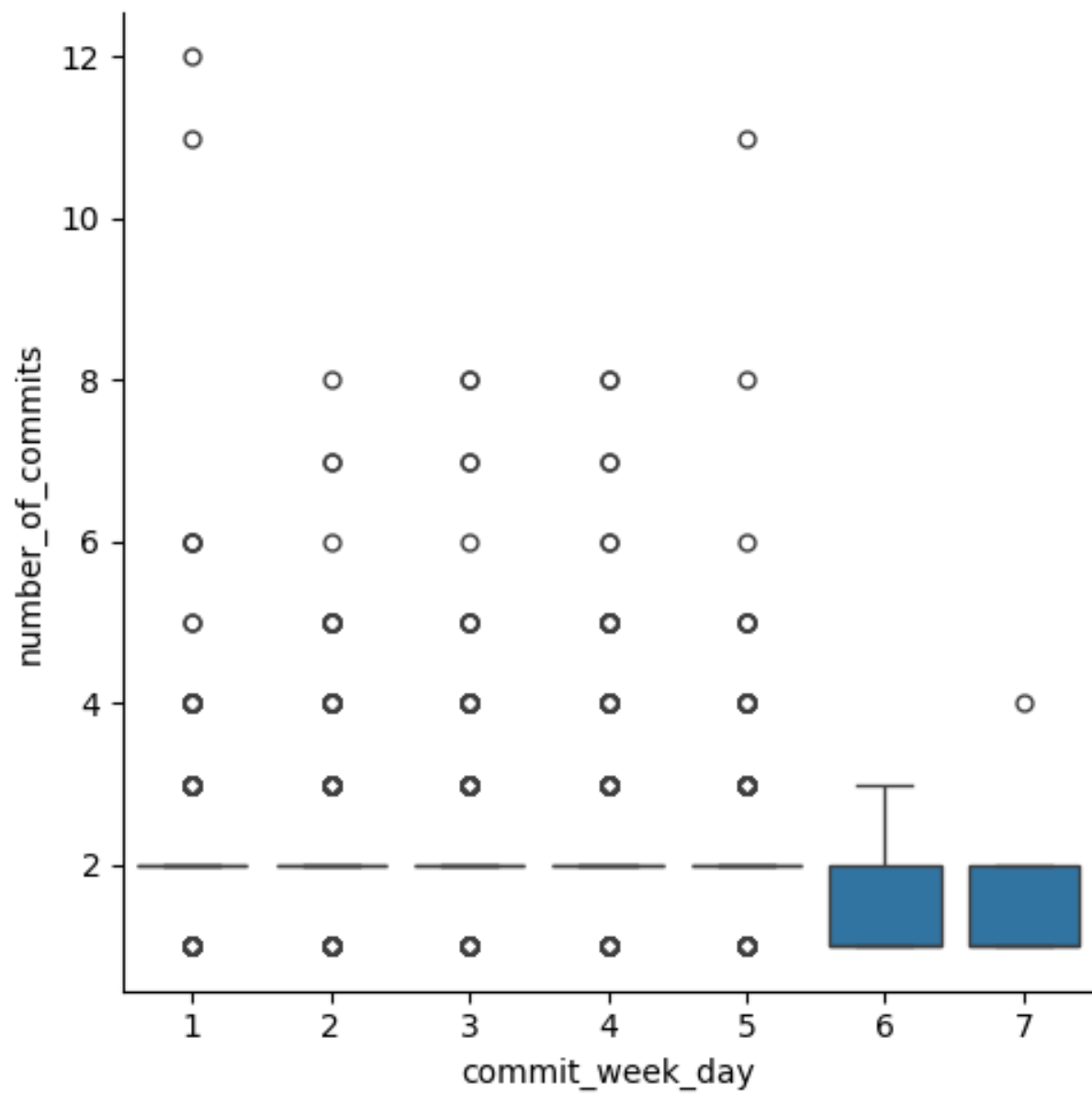
Commits time of a day table

The visualization of table above:



Commits time of day plot

Below the distribution of number of commits per weekday is shown the form of boxplot:



Commits day of week plot

2. Contributors activity

Table below shows top 10 most active contributors in terms of number of commits:

author_name	number_of_commits
aws-sdk-python-automation	2422
awstools	386
Daniel G. Taylor	223
kyleknap	186
James Saryerwinnie	123
JordonPhillips	59
Nate Prewitt	52
AWS	17
Nate Prewitt	16
Ray Luo	15

Top contributors number of commits

Next table shows the most active contributors in terms of number of insertions:

author_name	number_of_insertions
aws-sdk-python-automation	77605
Daniel G. Taylor	23597
kyleknap	19250
James Saryerwinnie	9808
awstools	7778
Nate Prewitt	6191
JordonPhillips	5557
Sheldon Blauman	3736
Steven Calwas	2653
jonathan343	2395

Top contributors number of insertions

Table showing top contributors in terms of insertions / deletions ratio:

author_name	insertions_deletions_ratio
Jami Schwarzwald	64.7
James Saryerwinnie	11.5
Jordan Guymon	9.1
JordonPhillips	8.7
Ryan Sonshine	6.2
Henri Yandell	5.9
Tim Finnigan	5.4
Jim Horng	4.9
kyleknap	4.4
Ray Luo	4.2

Insertions deletions ratio

Table showing top contributors in terms of number of commits per day of activity. This measure can be especially helpful for assessing the work of new developers, who joined the project quite recently:

author_name	commits_per_day
Michihito Shigemura	8.0
Henri Yandell	3.0
SamRemis	2.0
Felix Yan	2.0
bstrand	2.0
Dave Bishop	2.0
pgajdos	2.0
Rodrigo Chacon	2.0
Jussi Virtanen	2.0
mattsb42-aws	2.0

Commits number per day

Similar table showing number of insertions per day:

author_name	insertions_per_day
Sheldon Blauman	3736.0
Kyle Knapp	1403.0
Henri Yandell	201.0
Tom Keller	161.0
David L. Miller	114.0
home	107.3
Daniel G. Taylor	105.3
Jonas Neubert	96.0
Ryan Sonshine	68.0
David Huntsperger	67.0

Insertions number per day

3. Messages patterns

In this sections we shows top words from commit messages both in raw version (without any transformation) and after process of stemming (excluding plural suffix, etc.)

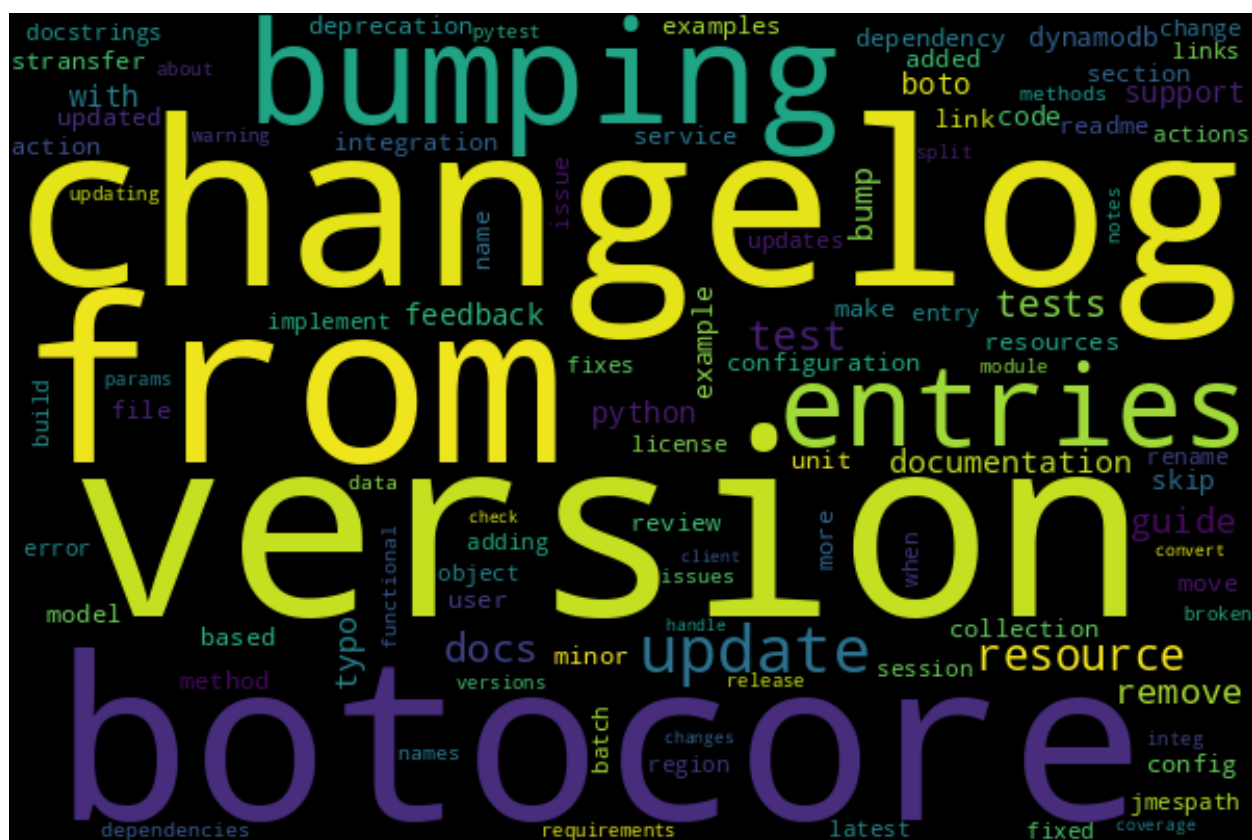
3.1 Raw words

Table showing top raw words in the commit messages:

raw_word	raw_word_freq
version	1471
changelog	1464
botocore	1437
from	1431
bumping	1429
entries	1403
update	196
resource	75
test	68
docs	58
tests	50
guide	45
remove	43
documentation	35
typo	35

Top raw words table

Word cloud of raw words:



Word cloud raw words

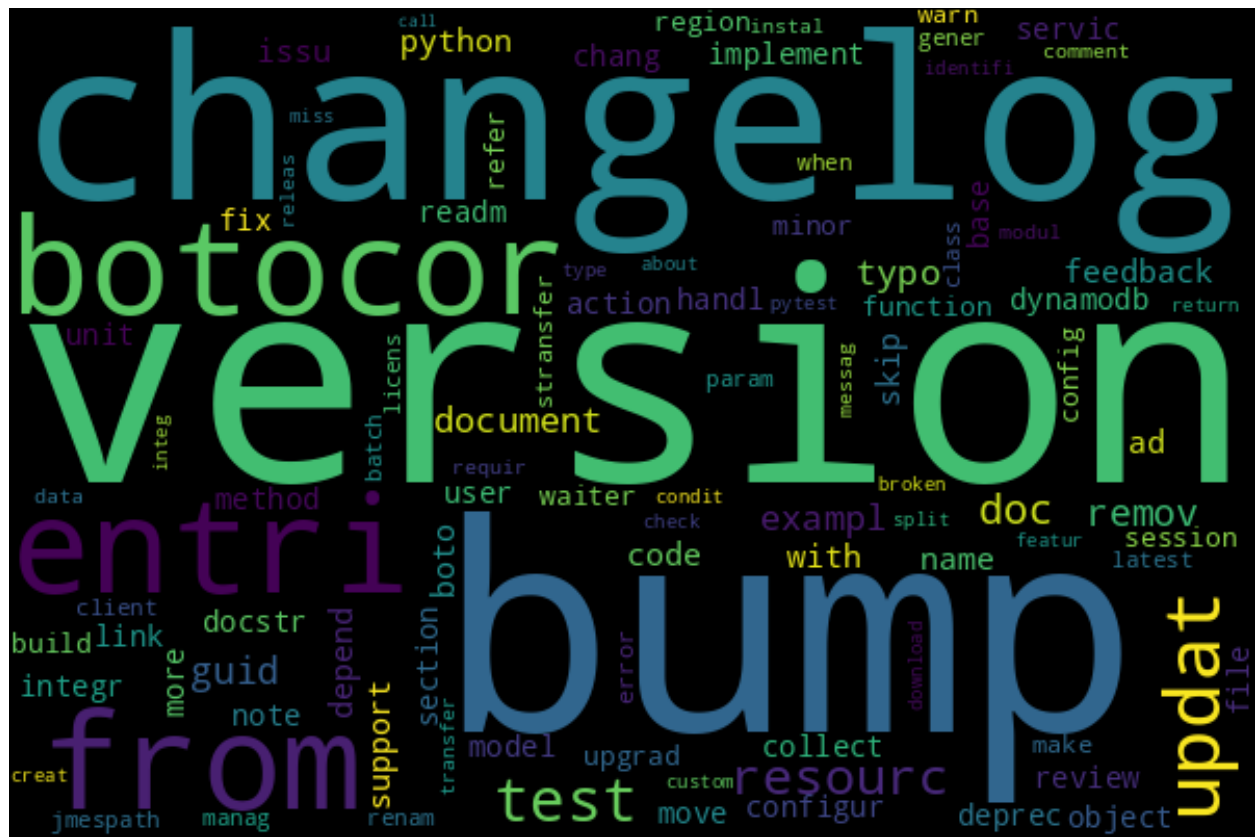
3.2 words after stemming

Table showing top stemmed words in the commit messages:

stemmed_word	stemmed_word_freq
version	1479
bump	1465
changelog	1464
botocor	1438
from	1431
entri	1413
updat	225
test	124
resourc	94
doc	58
remov	49
guid	49
document	41
exampl	40
typo	40

Top stemmed words table

Word cloud of raw words:



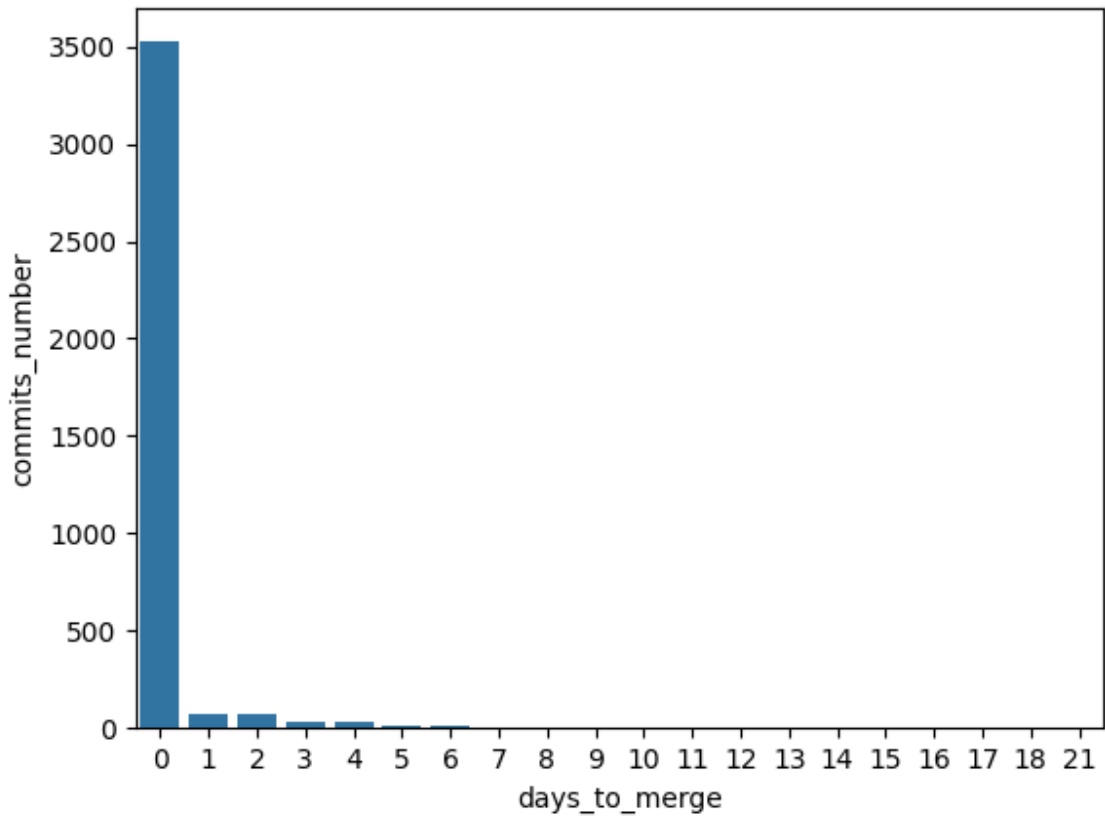
4. Periods of activity before merge

Table showing relation between number of commits and time to nearest merge:

days_to_merge	commits_number
0	3525
1	76
2	69
3	31
4	29
5	12
6	17
7	7
8	5
9	3
10	3
11	3
12	2
13	2
14	5
15	6
16	2
17	2
18	3
21	2

Time to merge table

The same relation in the form of plot:



Time to merge plot

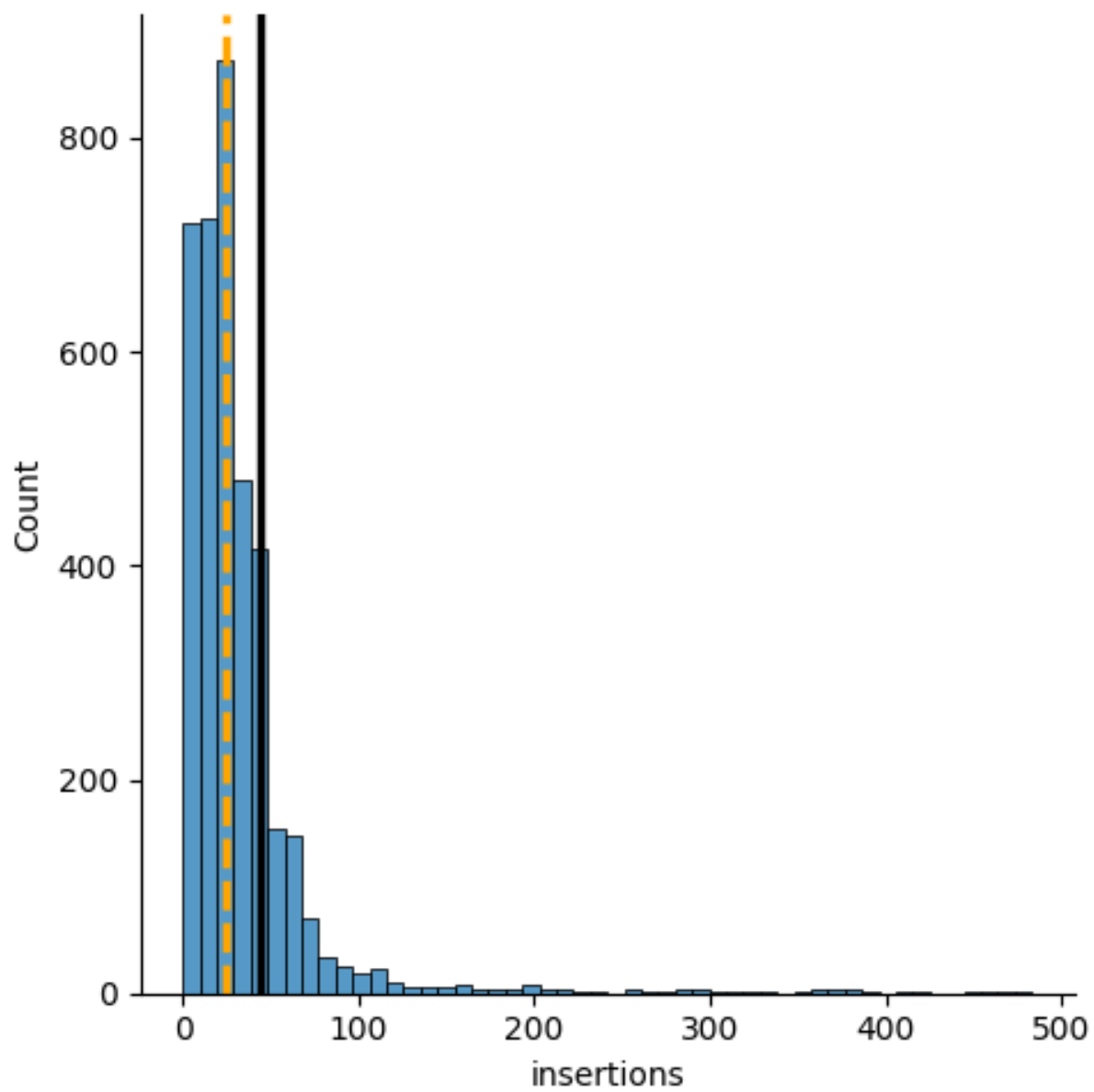
5. Insertions / deletions distributions

Table below shows the base statistical characteristics of insertions, such as mean value, standard deviation, max, min, etc:

measure	insertions
count	3819.0
mean	44.4
std	168.04
min	0.0
25%	12.0
50%	25.0
75%	40.0
max	4737.0

Insertions table

Plot below show the distributions of insertions per commit in the form of histogram. Average value is shown as black line and median is represented by orange dashed line:



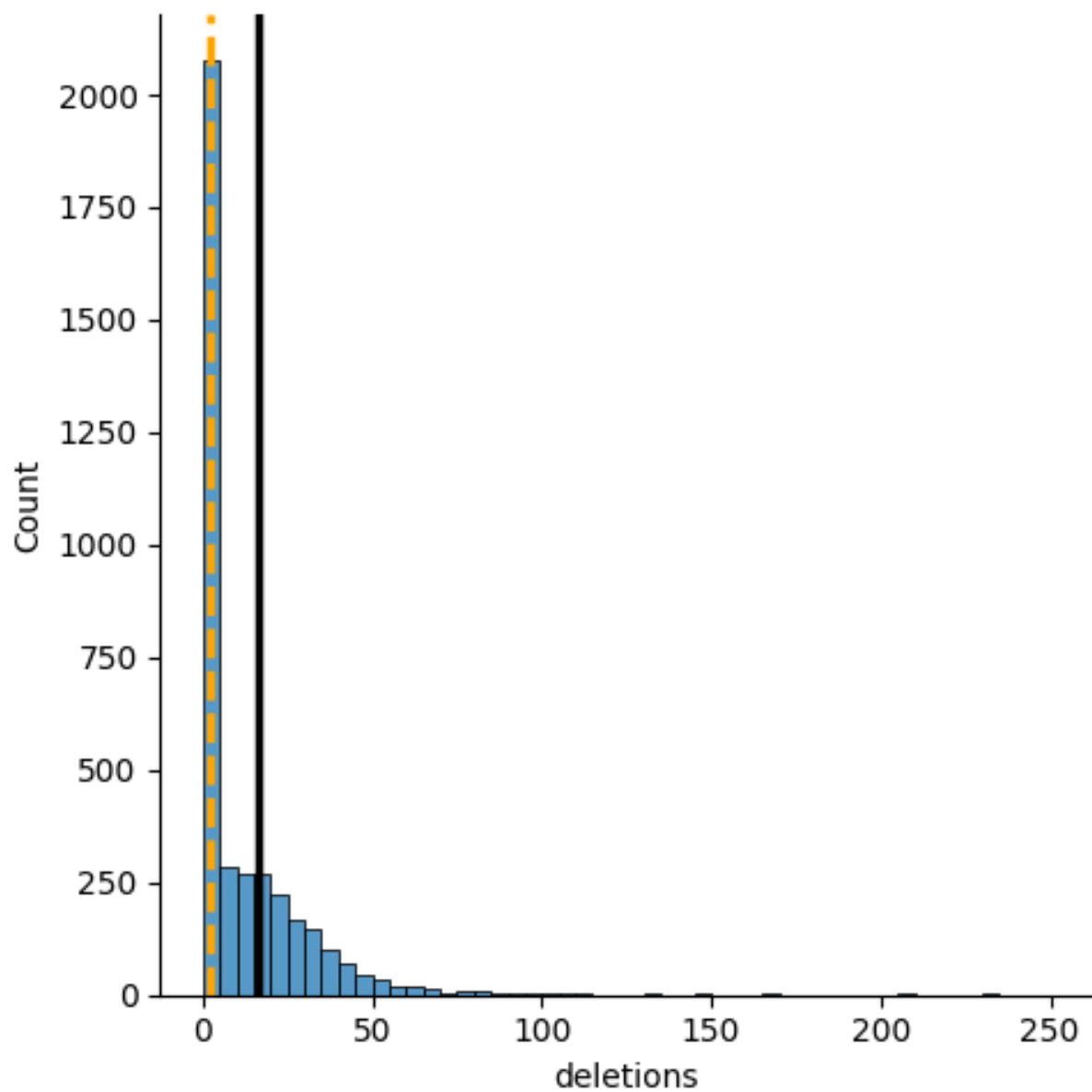
Insertions histogram

Table below shows the base statistical characteristics of insertions, such as mean value, standard deviation, max, min, etc:

measure	deletions
count	3819.0
mean	16.68
std	84.19
min	0.0
25%	0.0
50%	2.0
75%	18.0
max	3635.0

Deletions table

Histogram showing deletions distribuiton:



Deletions histogram

Summary

Hope this document was helpful. For more information please take a look at the dashboard.