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

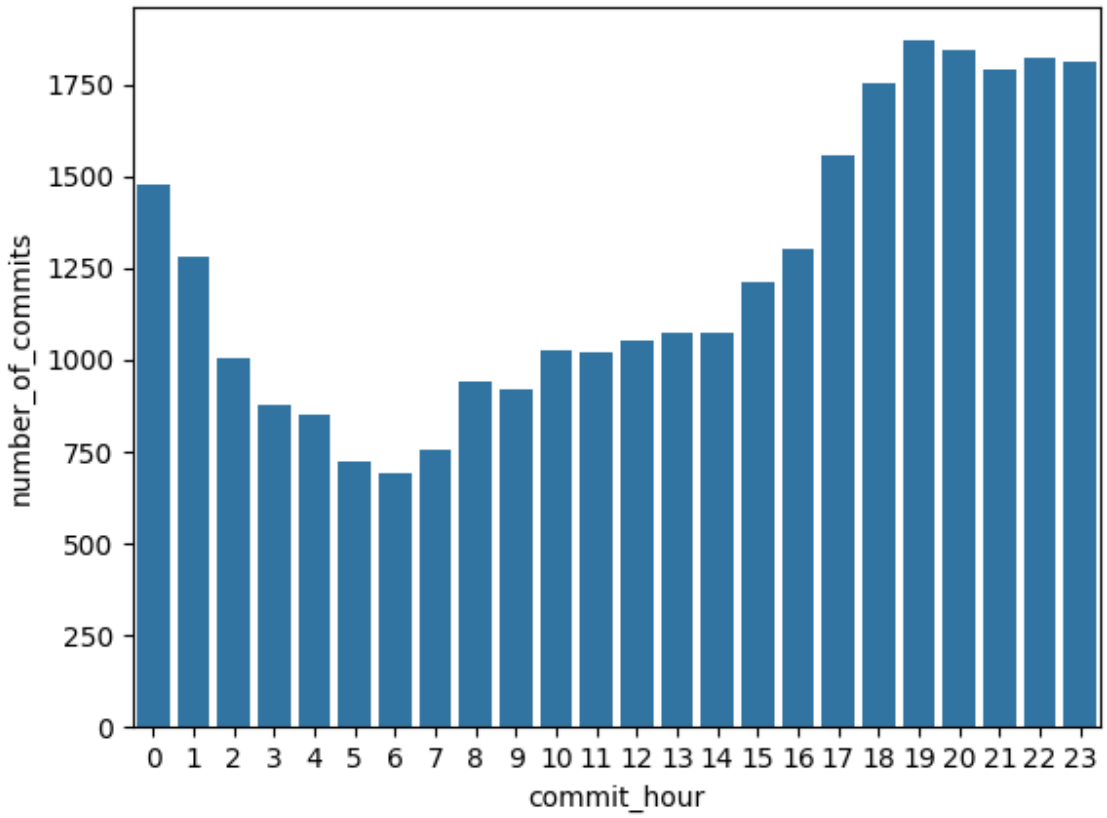
1. Peak contribution times

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

commit_hour	number_of_commits
0	1478
1	1280
2	1005
3	878
4	851
5	723
6	691
7	757
8	942
9	919
10	1029
11	1020
12	1055
13	1076
14	1072
15	1211
16	1302
17	1556
18	1753
19	1870

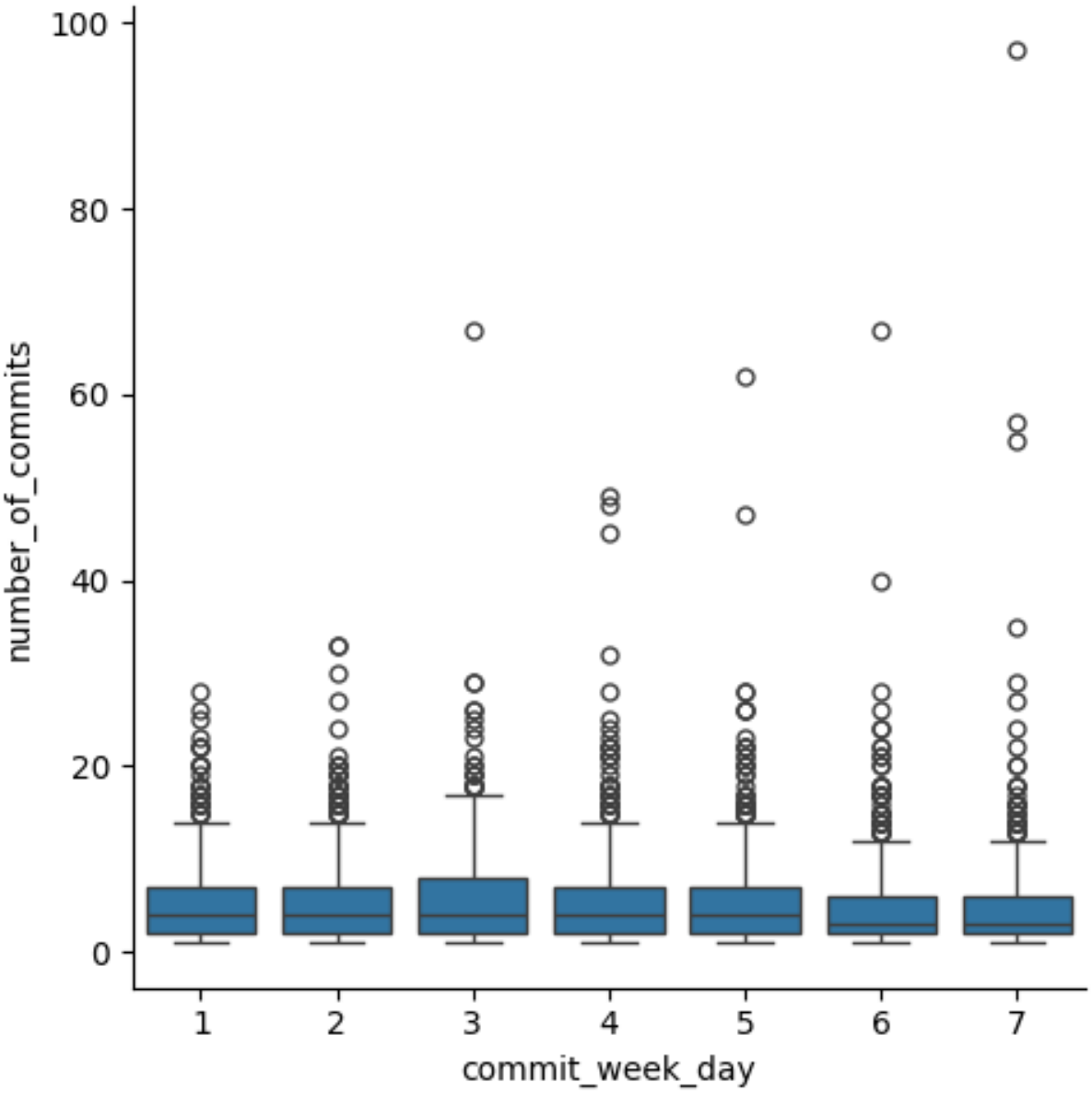
Commits time of a day table

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
Charles Harris	3055
Travis Oliphant	2033
David Cournapeau	1609
Sebastian Berg	1427
Pearu Peterson	1129
Eric Wieser	1067
mattip	1043
Pauli Virtanen	943
Julian Taylor	662
Bas van Beek	647

Top contributors number of commits

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

author_name	number_of_insertions
Travis Oliphant	402594
Eric Wieser	275643
Charles Harris	264168
Raghuveer Devulapalli	194190
ovillellas	164188
Eric Jones	144986
Sebastian Berg	136698
Pauli Virtanen	117970
Kevin Sheppard	92609
David Cournapeau	90413

Top contributors number of insertions

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

author_name	insertions_deletions_ratio
Jonathan Deng	388.7
Thomas J. Fan	77.8
prabhu	55.2
Benjamin Root	39.5
Glen Mabey	27.8
David Freese	26.3
tim cera	25.6
Stefan Otte	23.8
Matthew Rocklin	23.4
Sturla Molden	23.4

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
Eric Mariasis	23.0
Chris Kerr	13.0
Dmitry Zagorny	10.0
Jérôme Kieffer	10.0
patto90	8.0
Jakob	8.0
Alex Willmer	7.5
Janus Heide	6.0
Steve Dower	6.0
peterjc	6.0

Commits number per day

Similar table showing number of insertions per day:

author_name	insertions_per_day
HaroldMills	6393.0
lapack_lite code generator	5114.0
Justin Peel	3221.0
scoder	2090.0
Anne Bonner	1694.0
ovillellas	837.7
zindy	676.0
Jonathan Deng	647.8
Jun Kudo	634.0
Mark Wiebe	573.3

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.)

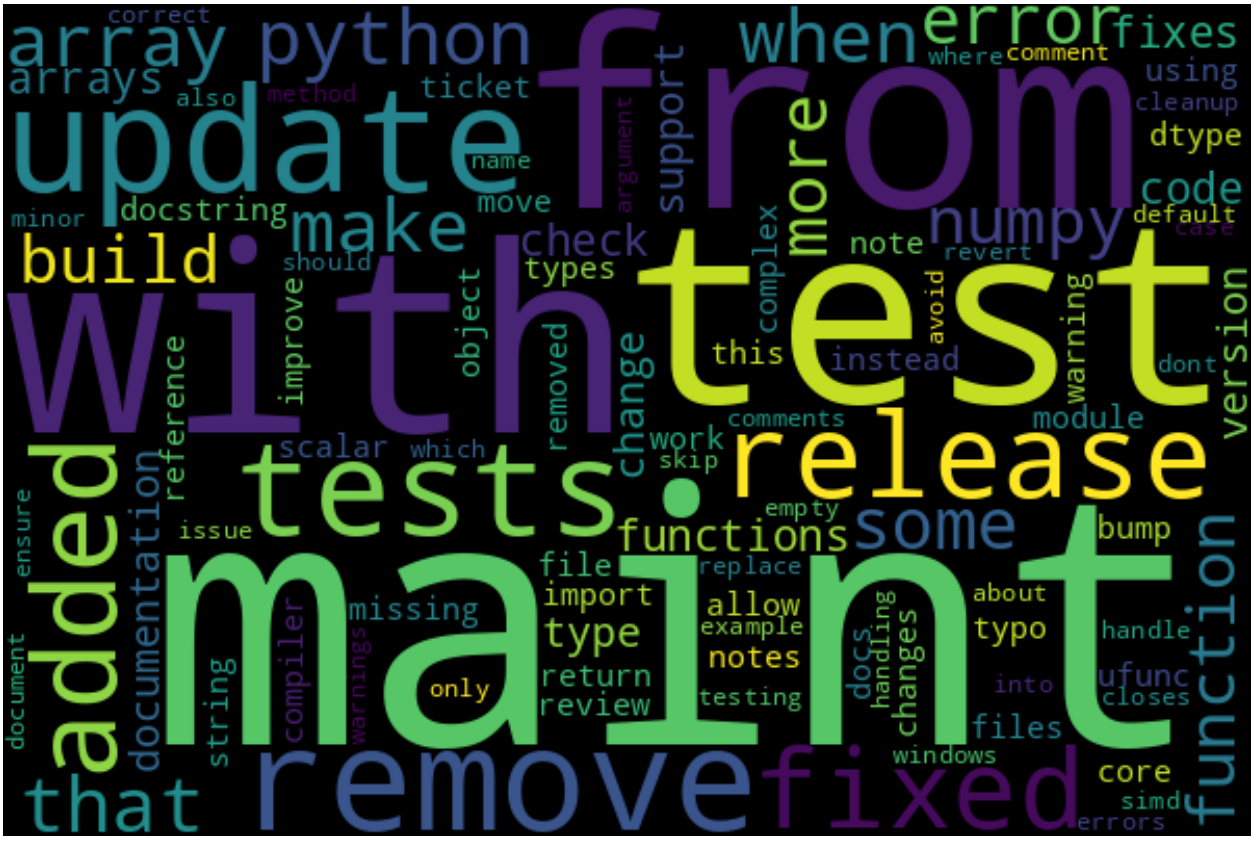
3.1 Raw words

Table showing top raw words in the commit messages:

raw_word	raw_word_freq
maint	3661
with	1765
from	1760
test	1700
update	1579
remove	1554
tests	1296
release	1022
added	997
fixed	971
python	933
that	927
when	895
array	875
make	875

Top raw words table

Word cloud of raw words:



Word cloud raw words

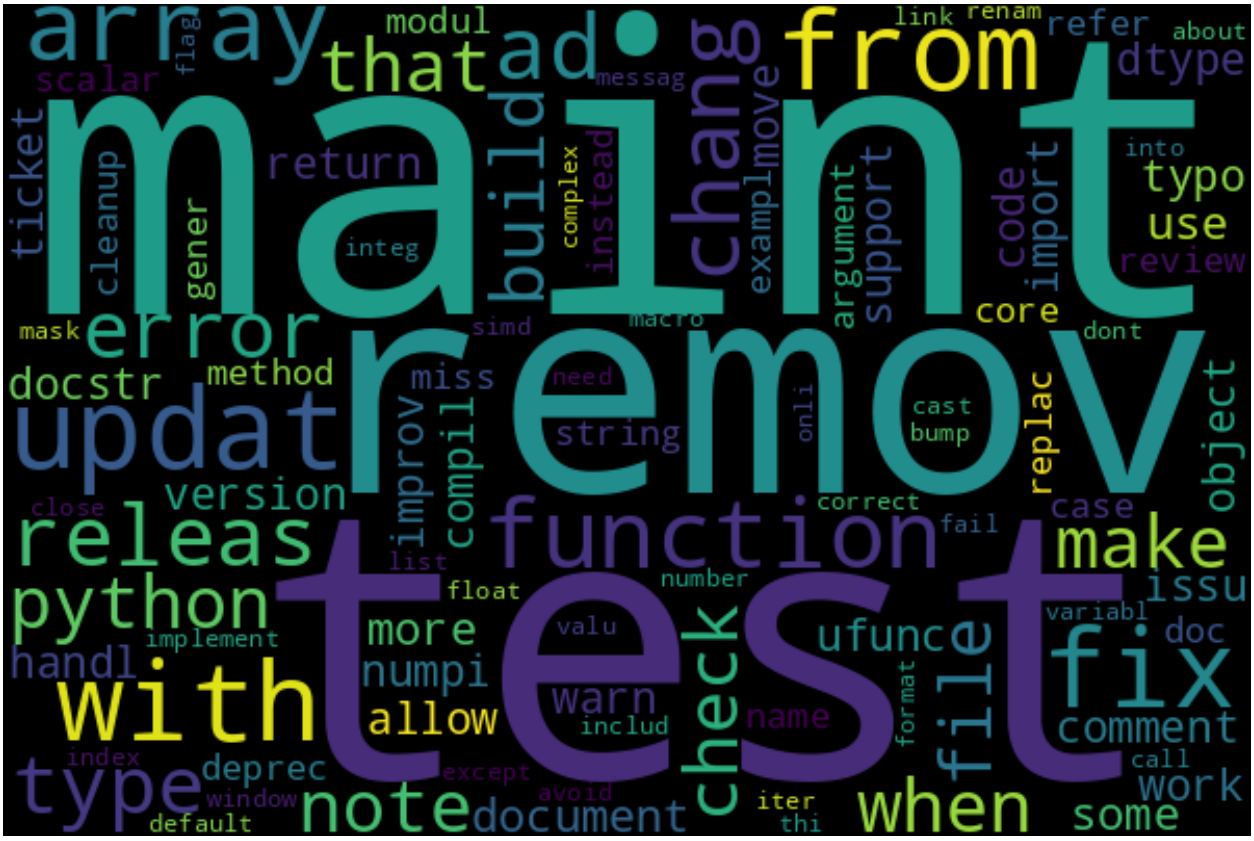
3.2 words after stemming

Table showing top stemmed words in the commit messages:

stemmed_word	stemmed_word_freq
maint	3661
test	3273
remov	1950
updat	1848
with	1765
from	1760
fix	1705
array	1476
function	1308
ad	1154
chang	1144
releas	1052
error	963
type	942
python	937

Top stemmed words table

Word cloud of raw words:



Word cloud stemmed words

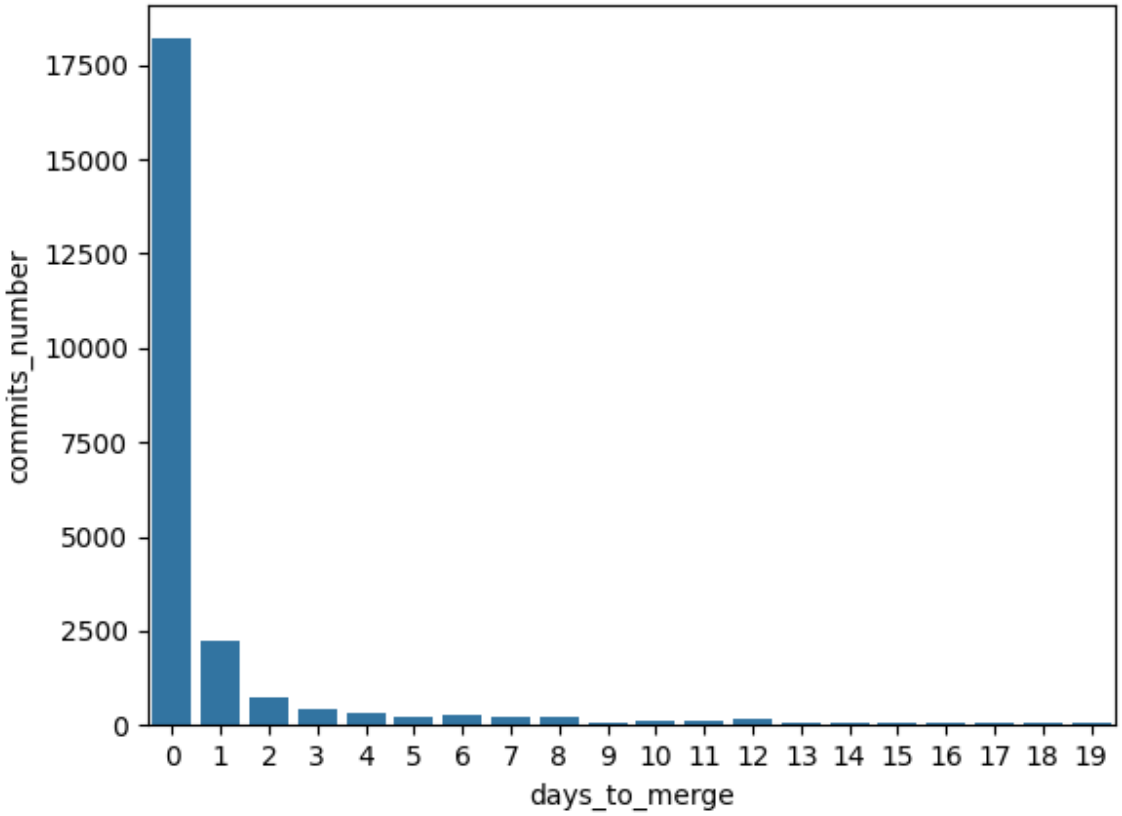
4. Periods of activity before merge

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

days_to_merge	commits_number
0	18204
1	2220
2	718
3	405
4	295
5	243
6	256
7	233
8	193
9	84
10	114
11	120
12	171
13	80
14	80
15	85
16	67
17	70
18	78
19	76

Time to merge table

The same relation in the form of plot:



Time to merge plot

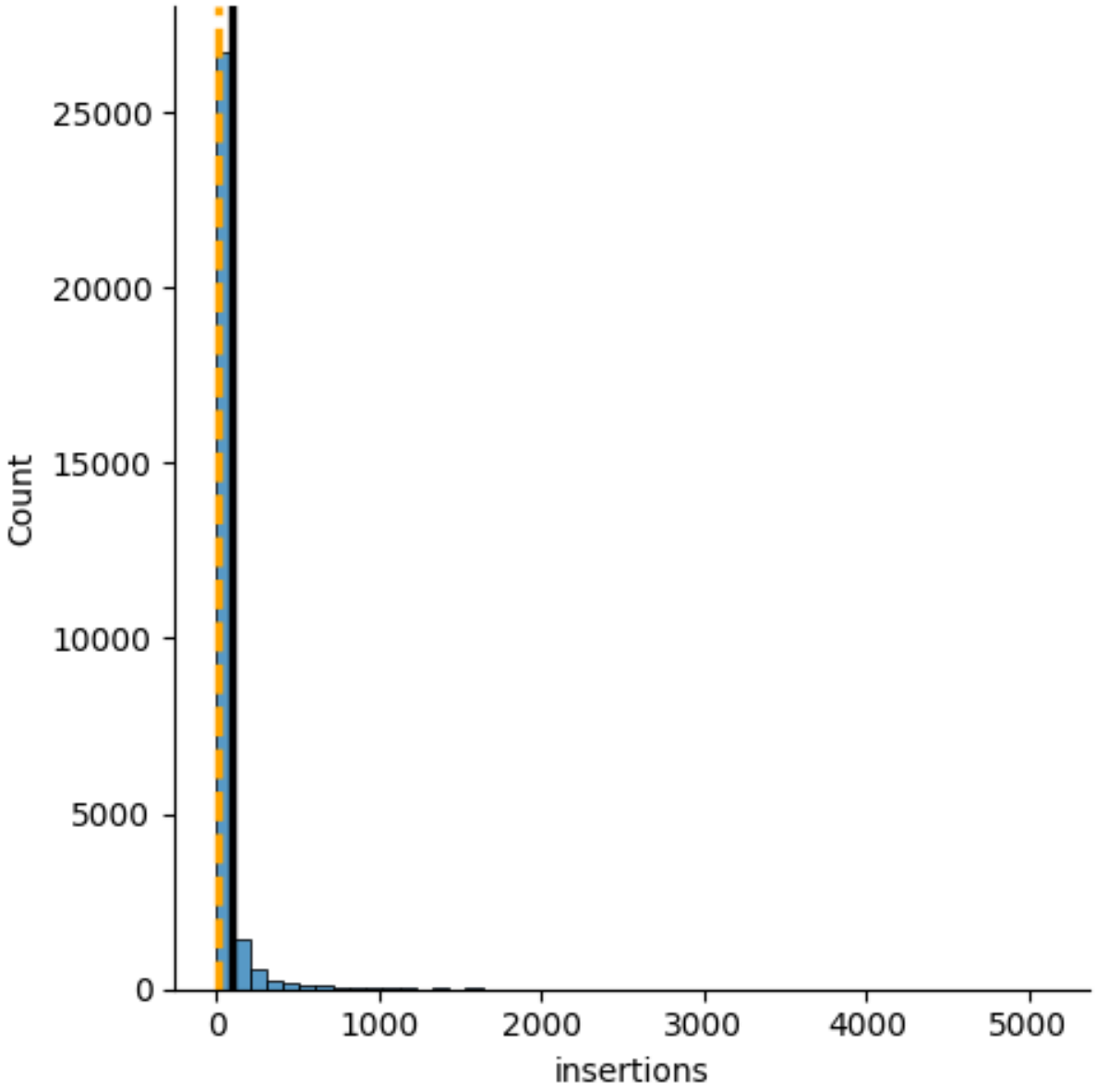
5. Insertions / deletions distributions

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

measure	insertions
count	29741.0
mean	101.85
std	1686.65
min	0.0
25%	2.0
50%	9.0
75%	34.0
max	146761.0

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:



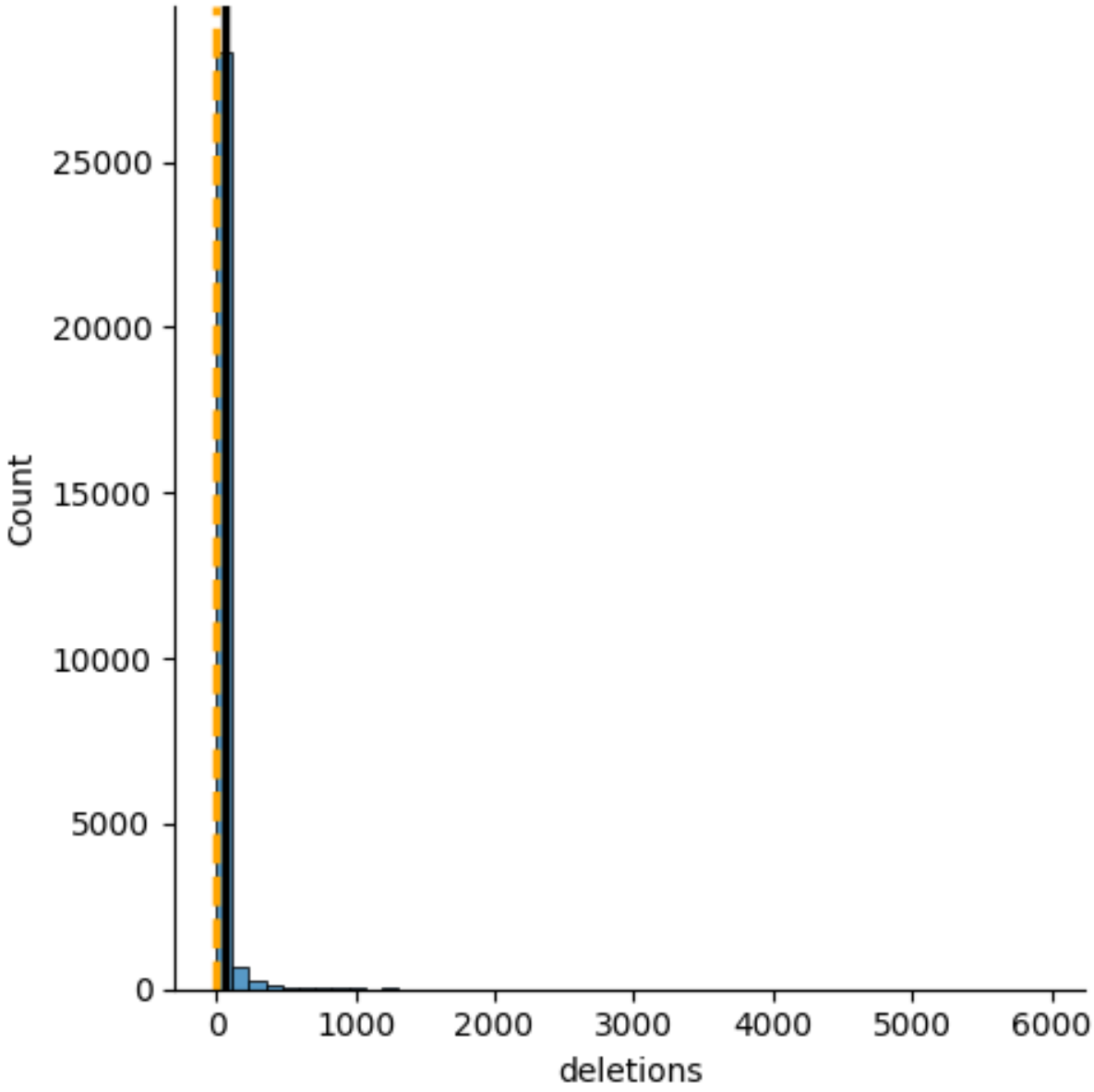
Insertions histogram

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

measure	deletions
count	29741.0
mean	70.17
std	1974.92
min	0.0
25%	1.0
50%	3.0
75%	12.0
max	232219.0

Deletions table

Histogram showing deletions distribuiton:



Deletions histogram

Summary

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