

WORKLOAD DISTRIBUTION

Select ten popular collaboratively developed software projects on GitHub to identify how much work each developer has done. Mainly, look for the following:

- a. Code reviews by a developer and its distribution in the team (e.g., 80% of reviews are done by 10% of the team).
- b. Code contribution count by a developer and its distribution throughout the project history.
- c. Code contribution count by a developer and its distribution in the last year.

10 GITHUB REPOSITORY

1. scrapy/scrapy
2. fastapi/fastapi
3. psf/requests
4. pallets/flask
5. celery/celery
6. python-pillow/Pillow
7. python-poetry/poetry
8. psf/black
9. pydantic/pydantic
10. encode/django-rest-framework

PROCESS AND INFORMATION

- Firstly, fetching data process was conducted to get 10 GitHub repository data formed in CSV files.
- For getting information about code reviews, the fetching process focused on reviews data. While, for getting information about code contribution count, both reviews and commits data were used. So that there are two kinds of CSV files which are about reviews and commits.
- To answer **the first question (a)**, there are three kinds of visualization based on reviews data such as:
 - Distribution chart all reviewers and their reviews count throughout the project history.
 - Distribution chart of top 20 reviewers who have highest reviews count differentiated by four kinds of states (approved, changes requested, commented, dismissed). (Reason: The distribution chart of all reviewers in the project show that some reviewers dominate the review job. So that to show the significance trend, it is decided to get the top 20 reviewers and how many their reviews count.)
 - Distribution chart of top 20 reviewers in percentage. (Reason: To show the trend in percentage thus it can be easy to make some insights.)

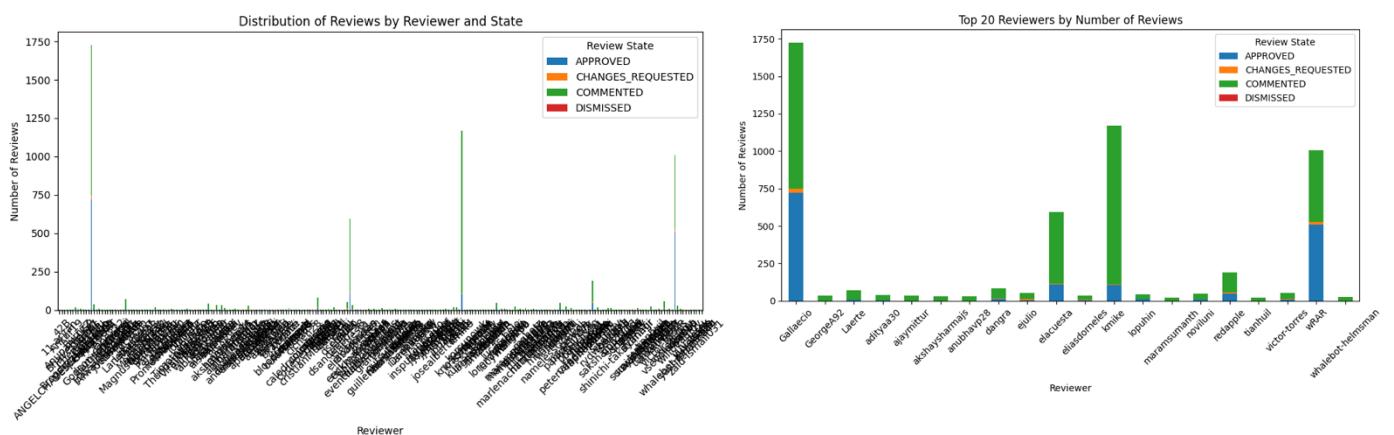
- To answer **the second question (b)**, there are some versions of visualization since it is based on the definition of "code contribution count". Here, there are two versions of definitions about that terms. Those definitions regard to review **or** commit contribution and commits **and** reviews contribution done by each developer.
 - In terms of reviews or commits contribution. (This part also include visualization such as distribution chart of all reviews/commits and distribution chart by top 20 reviewer/committer. However, since the distribution of review chart is discussed in the first question result, thus in this section there will be not discussion anymore.)
 - In terms of reviews and commits. There are two kinds of visualization formed which are:
 - Distribution chart of number of commits versus number of reviews based on top 20 committer in percentage. (Reason: To show how much workload as reviewer of the top 20 committer)
 - Distribution chart of number of reviews versus number of commits based on top 20 reviewers in percentage. (Reason: To show how much workload as committer of the top reviewers)
- To answer the **third question (c)**, it is assumed that "the last year" refers to year 2024. The visualization variations are the same with the distribution charts applied throughout the project history. So that, there will be 3 kinds of visualizations that can be analyzed.

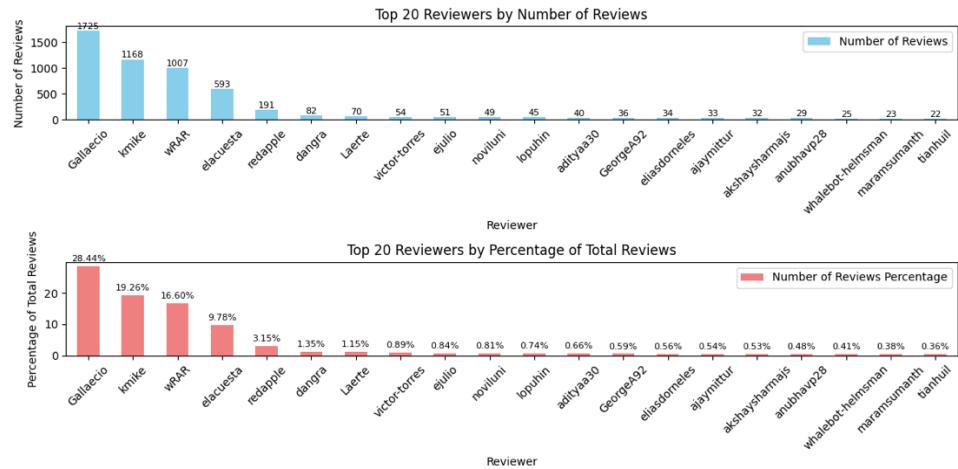
CHARTS AND INSIGHTS

1. *scrapy/scrapy*

a. Code reviews by a developer and its distribution in the team throughout the project history

- Number of reviewers: 242
- Number of reviews: 6065
- Number of committers: 747
- Number of commits: 14114



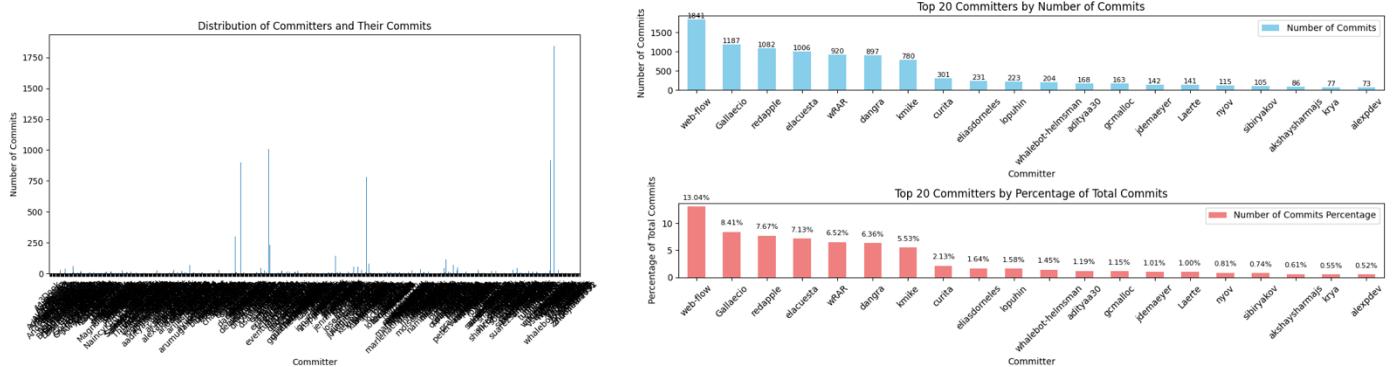


Insights (a):

- There were a few reviewers who were dominant for reviewing codes in this project.
- Regarding four kinds of state, overall, commented state was the most dominant which was higher than approved state. The portion of changes_requested state was small and the dismissed state was almost not-existent.
- Around 75% of total reviews count throughout the project history was reviewed by 4 out of 242 reviewers (about 1.7%).

b. Code contribution count by a developer and its distribution throughout the project history.

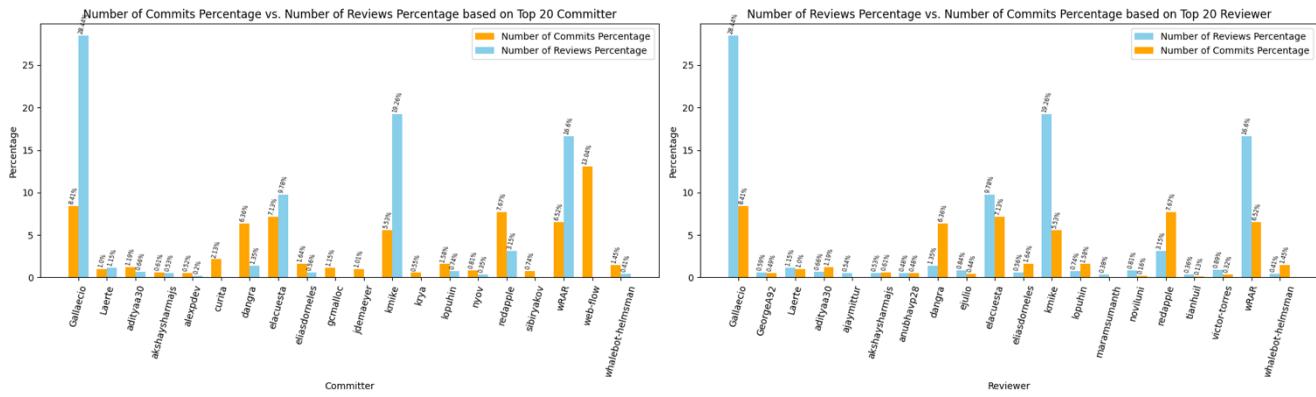
b1. In terms of commits count



Insights (b.b1)

- Overall, out of 747 committers throughout the project history, only a few committers were dominant. But it was still higher number compared to reviews distribution since the number of committers was also higher.
- Around 59% of total commits count was committed by 15 committers out of 747 (about 2% only).

b2. In terms of number of commits versus number of reviews based on top 20 committer and number of reviews versus number of commits based on top 20 reviewer.

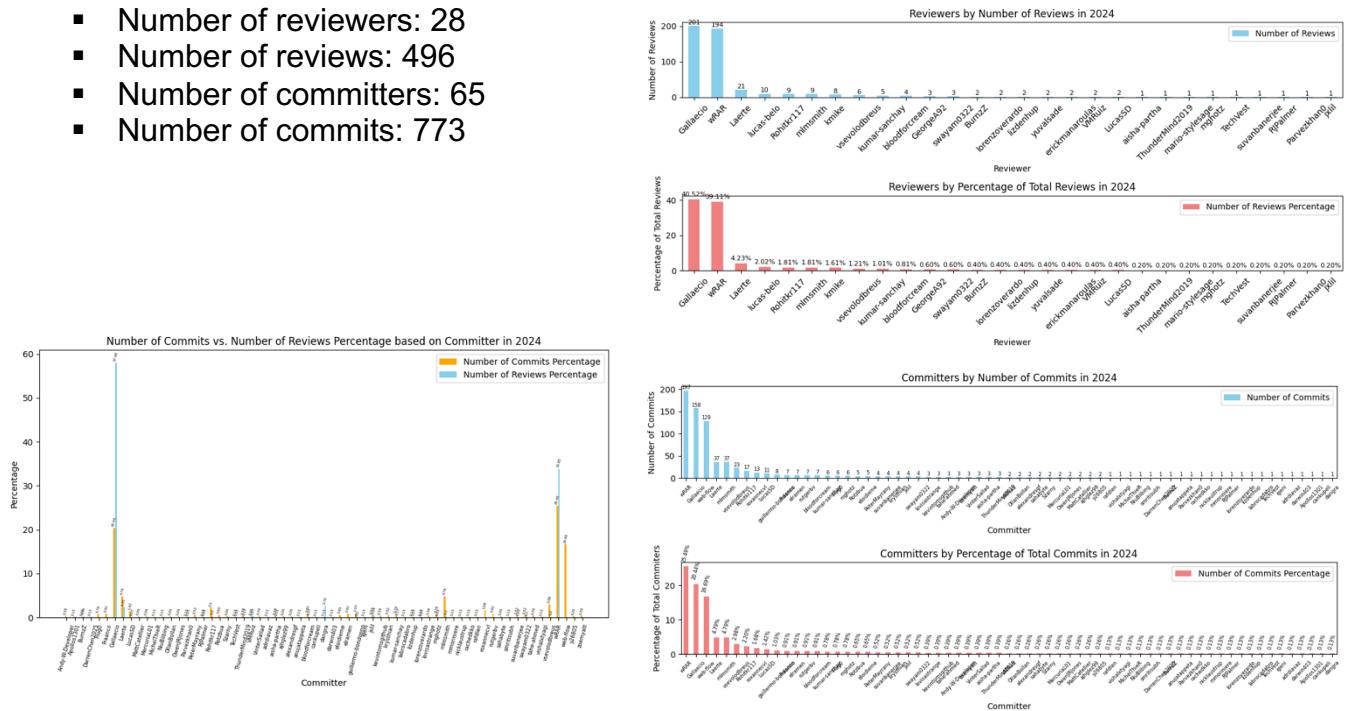


Insights (b.b2):

- Based on top 20 committers, not all committers had job as a reviewer too. However, there was a few committers who had job as reviewers which their reviews count was higher than their commits count.
- Based on top 20 reviewers, not all reviewers had job as a committer too.
- Looking at top 20 committers and reviewers' username, the chart shows that only a few usernames (6 usernames) who were dominant in this project both as a reviewer and a committer.
- There was one username (web-flow) who was a committer only with high workload of commits without contributing review work. While there was no reviewer username with high workload who did not have high commit workload, if that exists, its portion was under 1% only. So that, in most cases, a reviewer can be a committer too, but a committer is not always be able to be a reviewer at the same time.

c. Code contribution count by a developer and its distribution in the last year (2024).

- Number of reviewers: 28
- Number of reviews: 496
- Number of committers: 65
- Number of commits: 773



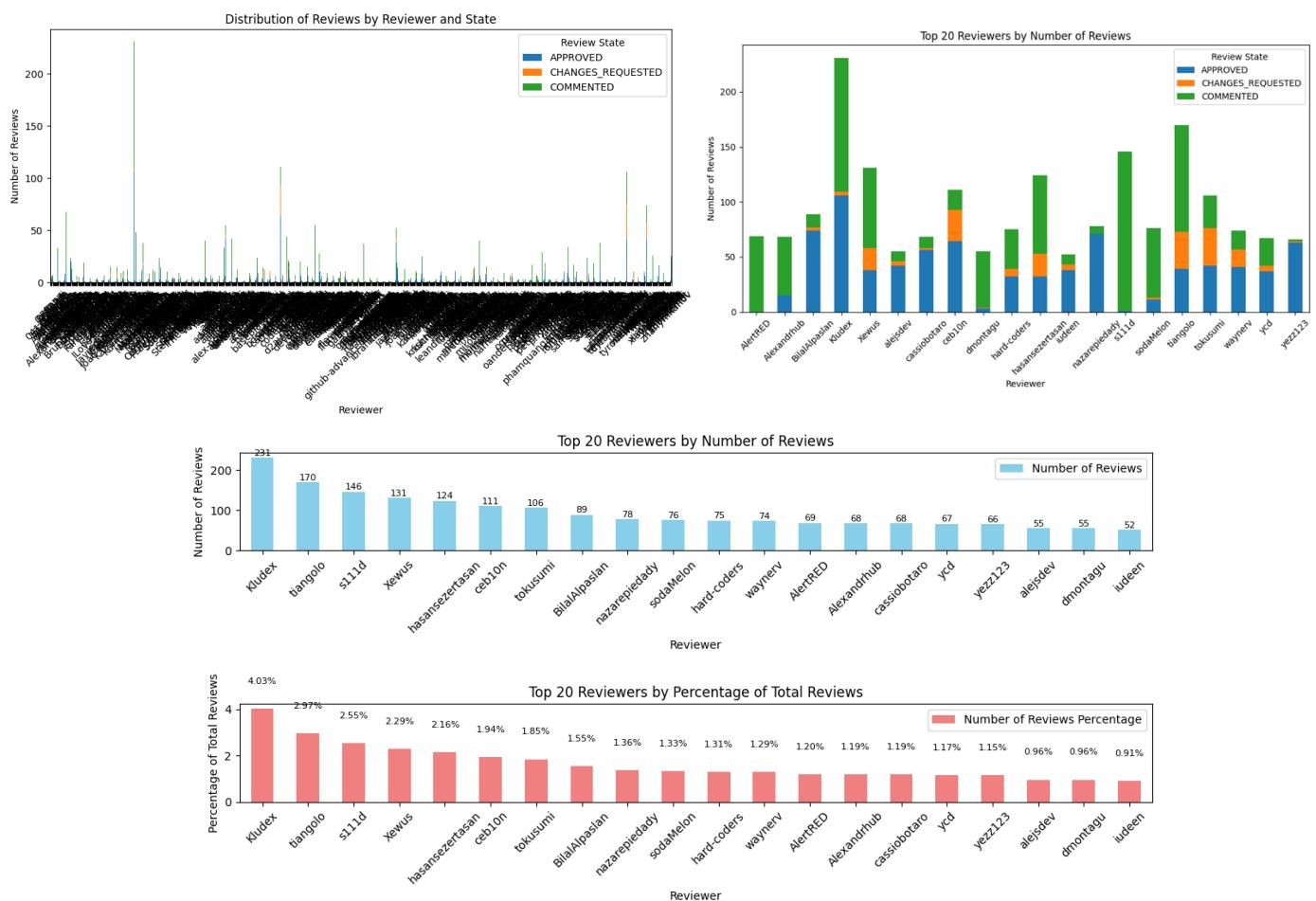
Insights (c):

- Around 79% of review workload was conducted by two reviewers only out of 28 in 2024.
- Around 62% of commit workload was conducted by three committers only out of 28 in 2024.
- Only two usernames who were active contributing in both commits and reviews work in 2024. Moreover, only one username who was active for committing code and did not review the code at all.
- Based on committer in 2024, only a few usernames who did reviewing too. Also, they who did the review job had high workload of review.

2. fastapi/fastapi

a. Code reviews by a developer and its distribution in the team throughout the project history

- Number of reviewers: 784
- Number of reviews: 5733
- Number of committers: 849
- Number of commits: 15715

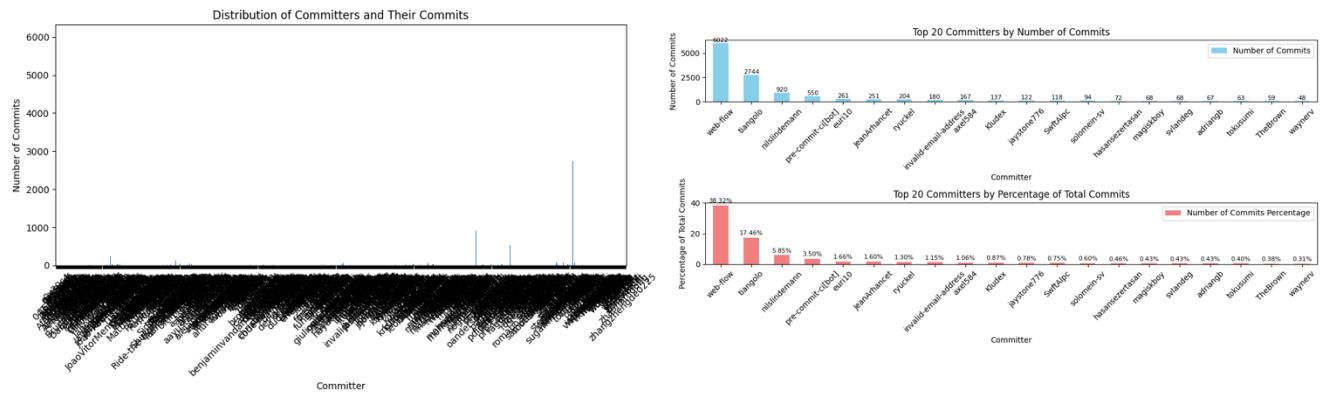


Insights (a):

- The chart illustrates a fluctuated trend between each reviewer throughout the project history, thus not just a few reviewers who were dominant to do the work.
- Based on the portion of states, there was no state of dismissed in this project showed in top 20 highest reviewer. Moreover, the portion of approved states were not always smaller than the commented state. Yet, there were a few bars having only commented state in one username of reviewer and there were a few bars which have high portion of approved state and just small portion of commented.
- The top 20 reviewer chart describes that the workload distribution of reviews was distributed well in this project, since the portion of review between reviewers differ in a small number. Around 33% of reviews were conducted by 2.5% of all reviewers.

b. Code contribution count by a developer and its distribution throughout the project history.

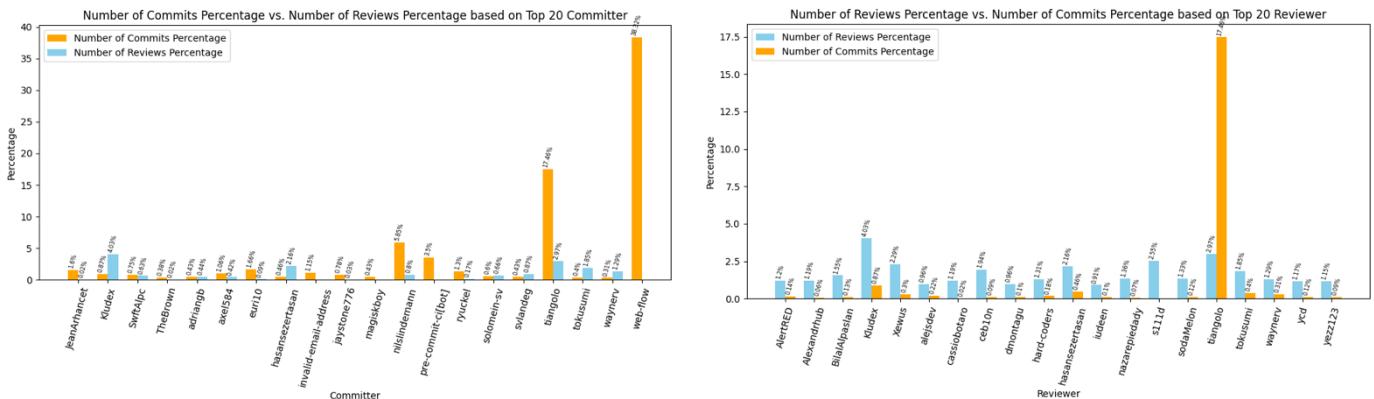
b1. In terms of commits count



Insights (b.b1)

- There were a few committers who were dominant in this project.
- About 72% of all commit workload was done by 1.06% only of all committers.

b2. In terms of number of commits versus number of reviews based on top 20 committer and number of reviews versus number of commits based on top 20 reviewer.

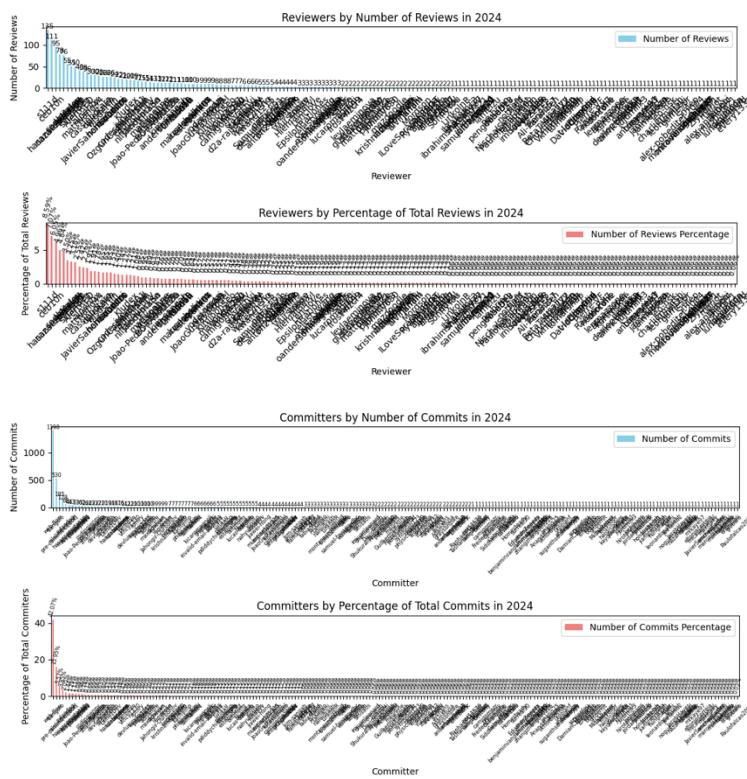


Insights (b.b2)

- Based on top 20 committers, not all committers had job as a reviewer too.
- Committers who did the review too had small portion of review workload.
- Based on top 20 reviewers, not all reviewers had job as a committer too.
- There is one reviewer username who has high portion in 17.56% of commit workload throughout the project history.
- Overall, these charts show that many of the committers and reviewers are not the same username.

c. Code contribution count by a developer and its distribution in the last year (2024).

- Number of reviewers: 176
- Number of reviews: 1571
- Number of committers: 213
- Number of commits: 3323



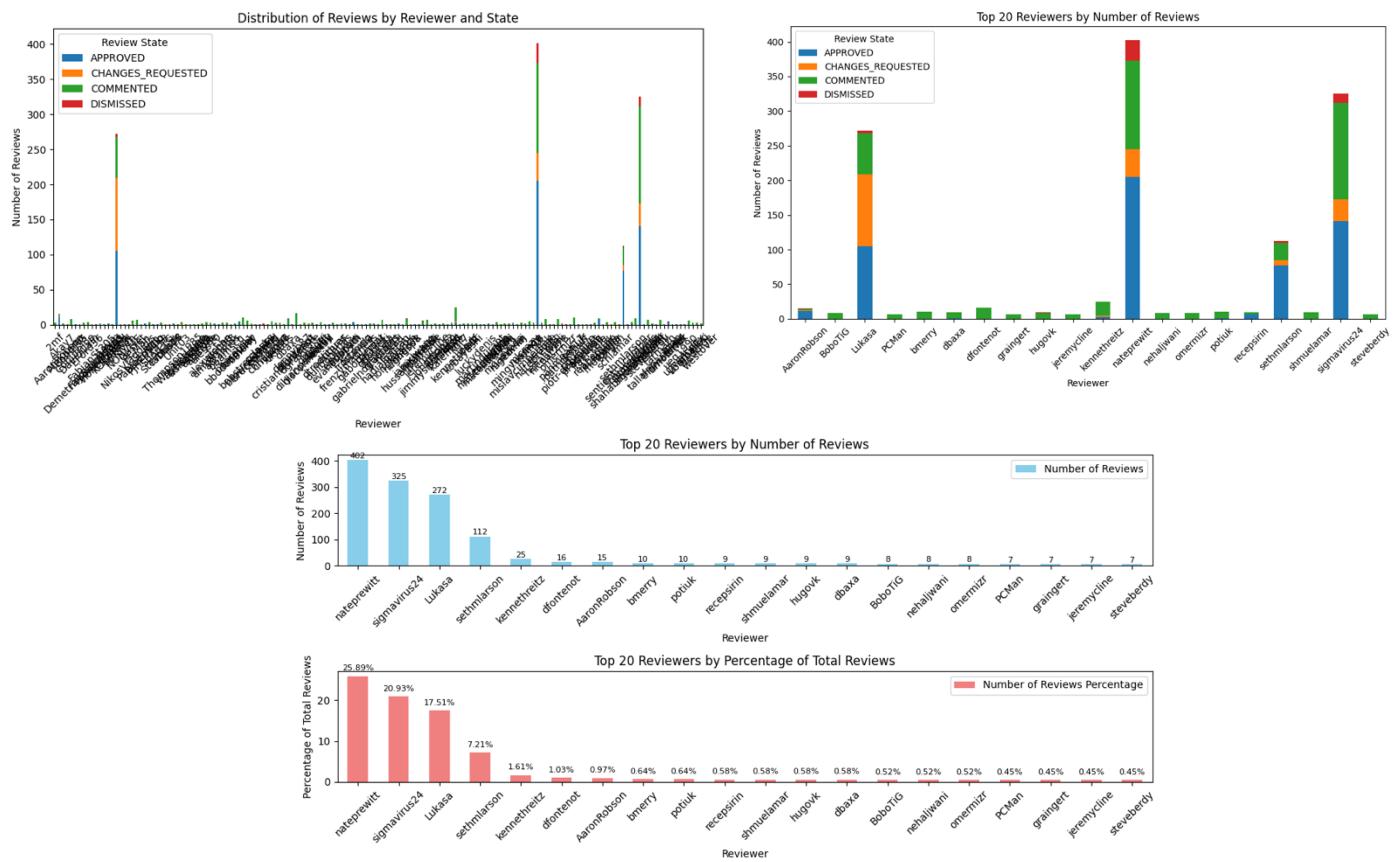
Insights (c)

- Around 54% of review workload was conducted by about 7.4% reviewers in 2024.
- Around 58% of commit workload was conducted by two committers only out of 213.
- There is a highest point of commit workload percentage which refers to a committer who did not any review work.

3. psf/requests

a. Code reviews by a developer and its distribution in the team throughout the project history

- Number of reviewers: 159
- Number of reviews: 1553
- Number of committers: 900
- Number of commits: 5720

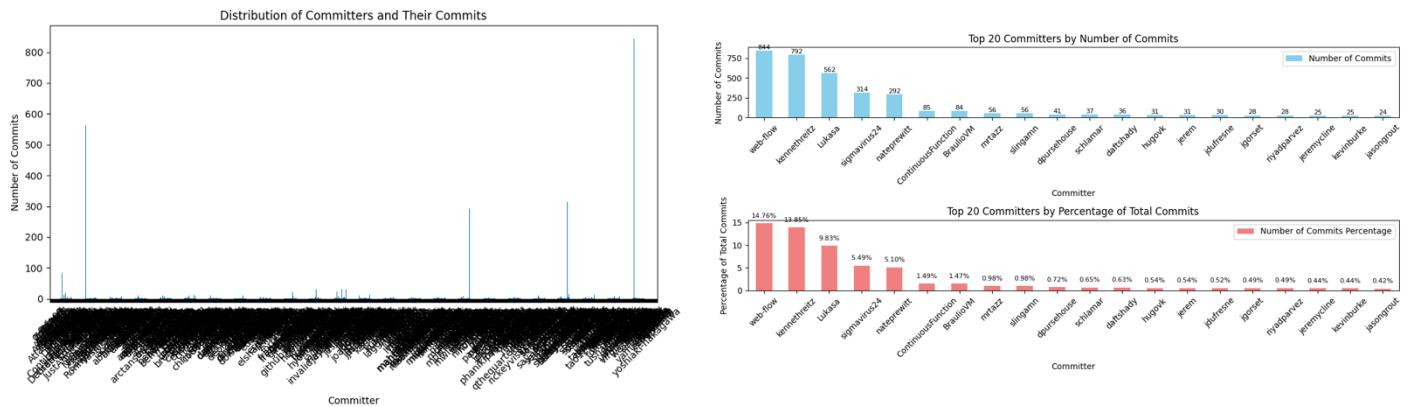


Insights (a):

- There were a few reviewers who were dominant in the review jobs throughout this project history.
- Based on the portion of states, there were portions of change_requested and dismissed state from those few reviewers who had high workload of review. While in the other reviewers, most of them were on commented state.
- Around 72% of review workload was conducted by 4 reviewers only out of 1553 (about 0.2%). So that, the review workload was not distributed evenly among reviewers.

b. Code contribution count by a developer and its distribution throughout the project history.

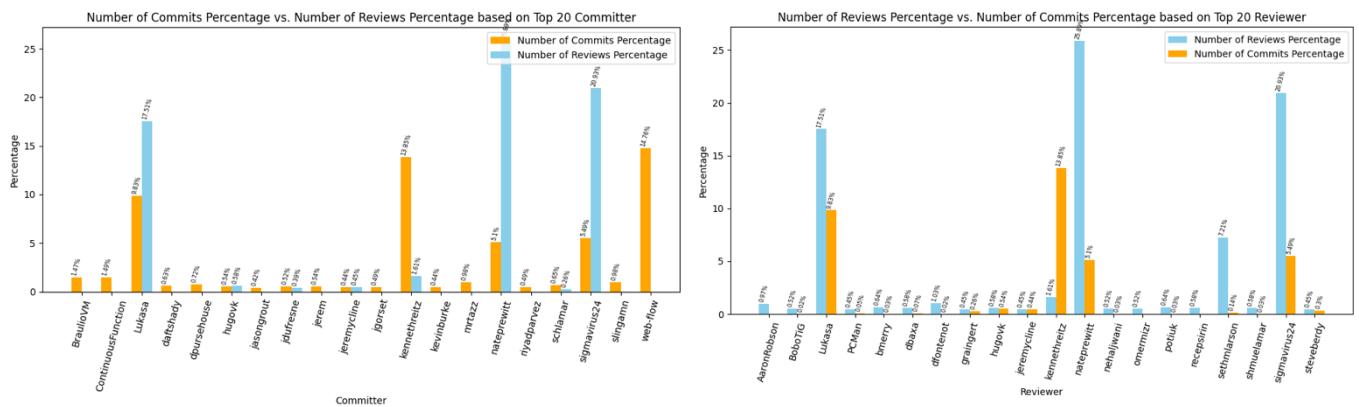
b1. In terms of commits count



Insights (b1)

- There were a few committers who were dominant in committing.
- Around 49% of commits happened in this project was conducted by 5 committers only (around 0.5%).

b2. In terms of number of commits versus number of reviews based on top 20 committer and number of reviews versus number of commits based on top 20 reviewer.



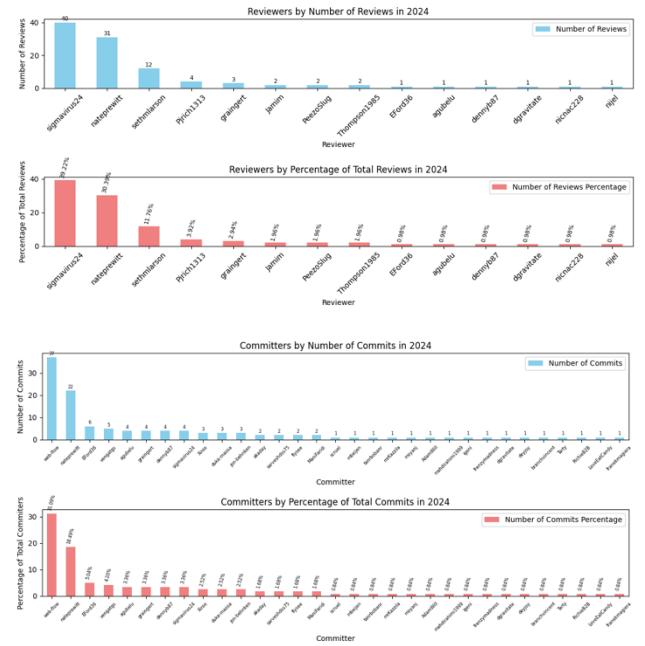
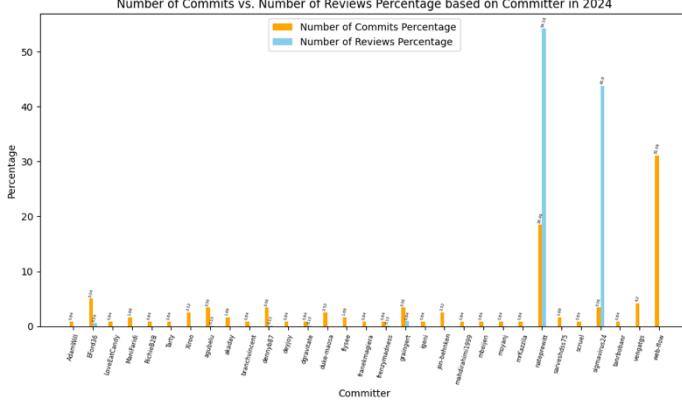
Insights (b2)

- Based on top 20 committers, not all committers had job as a reviewer too.
- There were three usernames of committers who did reviews too in a high number of review count.
- Based on top 20 reviewers, not all reviewers had job as a committer too.
- There is one username of reviewers who committed the code in a high number of commit count.
- Overall, the workload of commits and reviews were done by the same username who did high workload either for commits or reviews.

c. Code contribution count by a developer and its distribution in the last year (2024).

- Number of reviewers: 14

- Number of reviews: 102
- Number of committers: 31
- Number of commits: 119



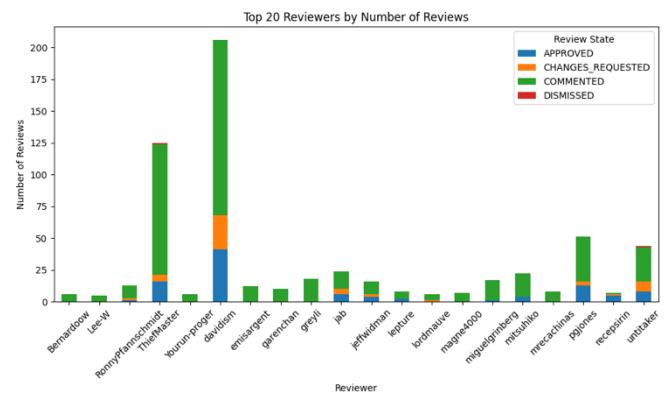
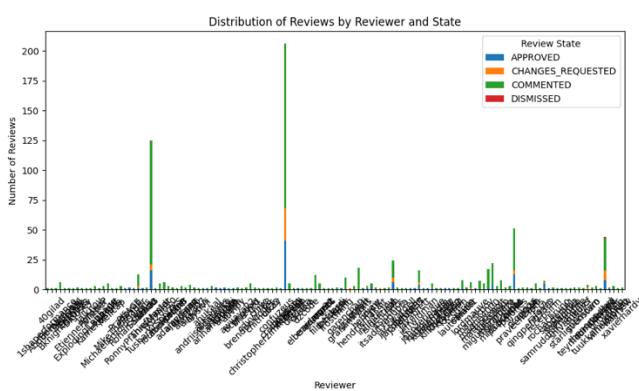
Insights (c)

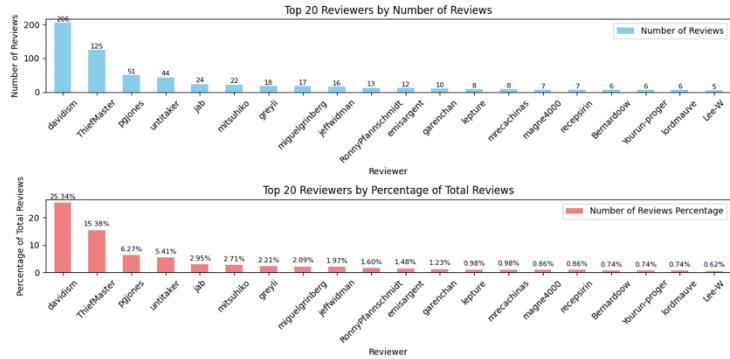
- Around 81% of review jobs was done by about 21% of reviewers in 2024.
- Around 50% of commit jobs was done by about 6% of committers in 2024.
- In most cases, the committer in 2024 did not do the review either. There were two committers who reviewed the code in a high number which showed that they are a reviewer too.

4. pallets/flask

a. Code reviews by a developer and its distribution in the team throughout the project history

- Number of reviewers: 134
- Number of reviews: 813
- Number of committers: 836
- Number of commits: 6952



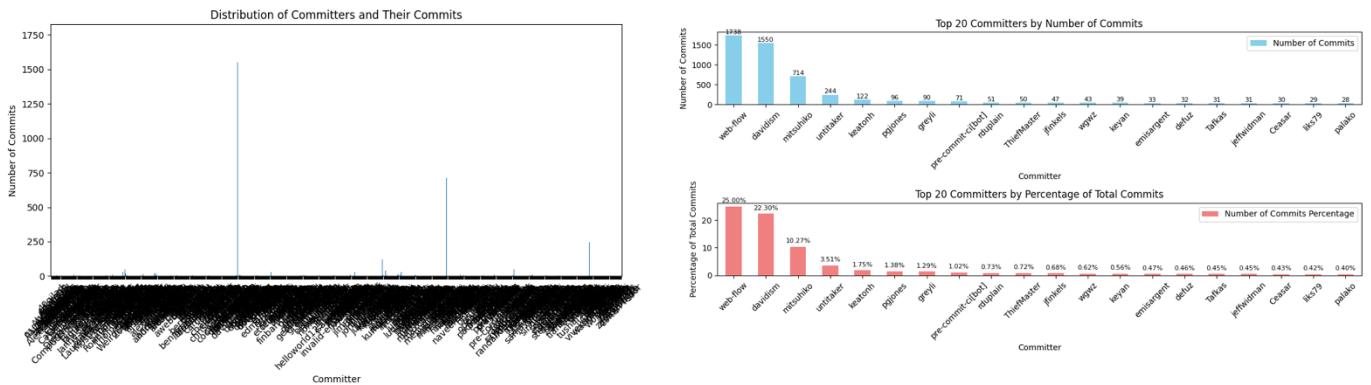


Insights (a):

- There were a few reviewers who were dominant in the review jobs throughout this project history.
- Based on the portion of the reviews' state, in most cases, the commented state was dominant in each of reviewer username. The approved states were in a smaller number than the commented one. There is a reviewer who had dismissed state too in this project, but in a small number.
- Around 60% of review works were done by 7 reviewers out of 134 in total (about 5.2%).

b. Code contribution count by a developer and its distribution throughout the project history.

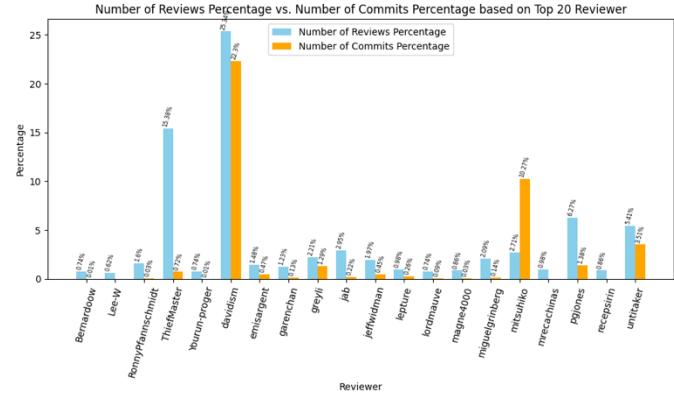
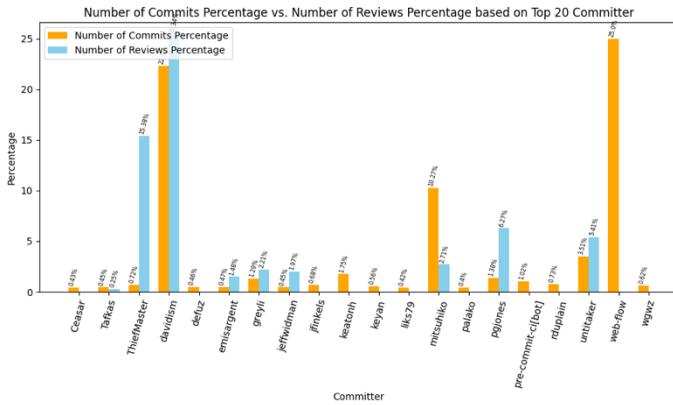
b1. In terms of commits count



Insights (b1)

- There were a few committers who were dominant in the commit jobs throughout this project history.
- Around 58% of all commits were committed by 3 committers out of 836 in total (0.35%).

b2. In terms of number of commits versus number of reviews based on top 20 committer and number of reviews versus number of commits based on top 20 reviewer.

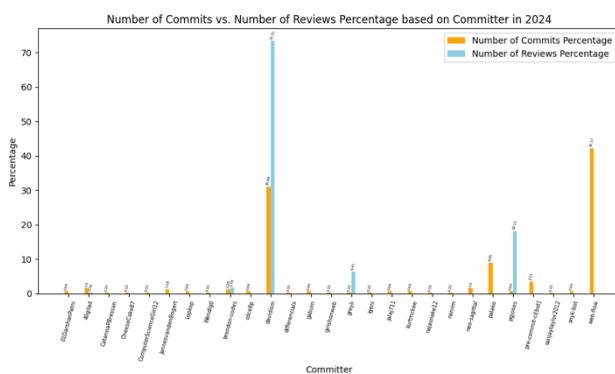
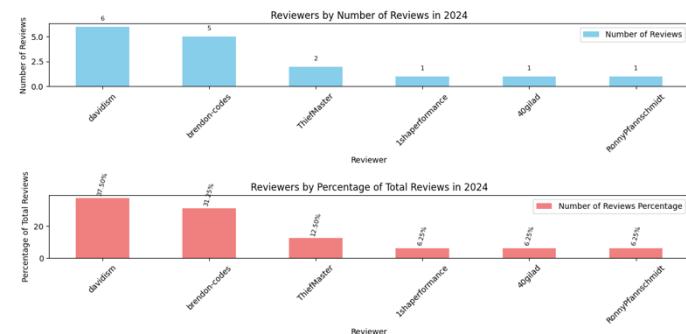


Insights (b2)

- Based on top 20 committer, not all committers had job as a reviewer too.
- There was a committer who had high number of commits did not do any review jobs. Also, there was a committer who had high number both commits count and reviews count throughout this project history.
- Based on top 20 reviewer, not all reviewers had job as committer too.
- There were 3 reviewers who committed the code in a significant number.

c. Code contribution count by a developer and its distribution in the last year (2024).

- Number of reviewers: 6
- Number of reviews: 16
- Number of committers: 27
- Number of commits: 313



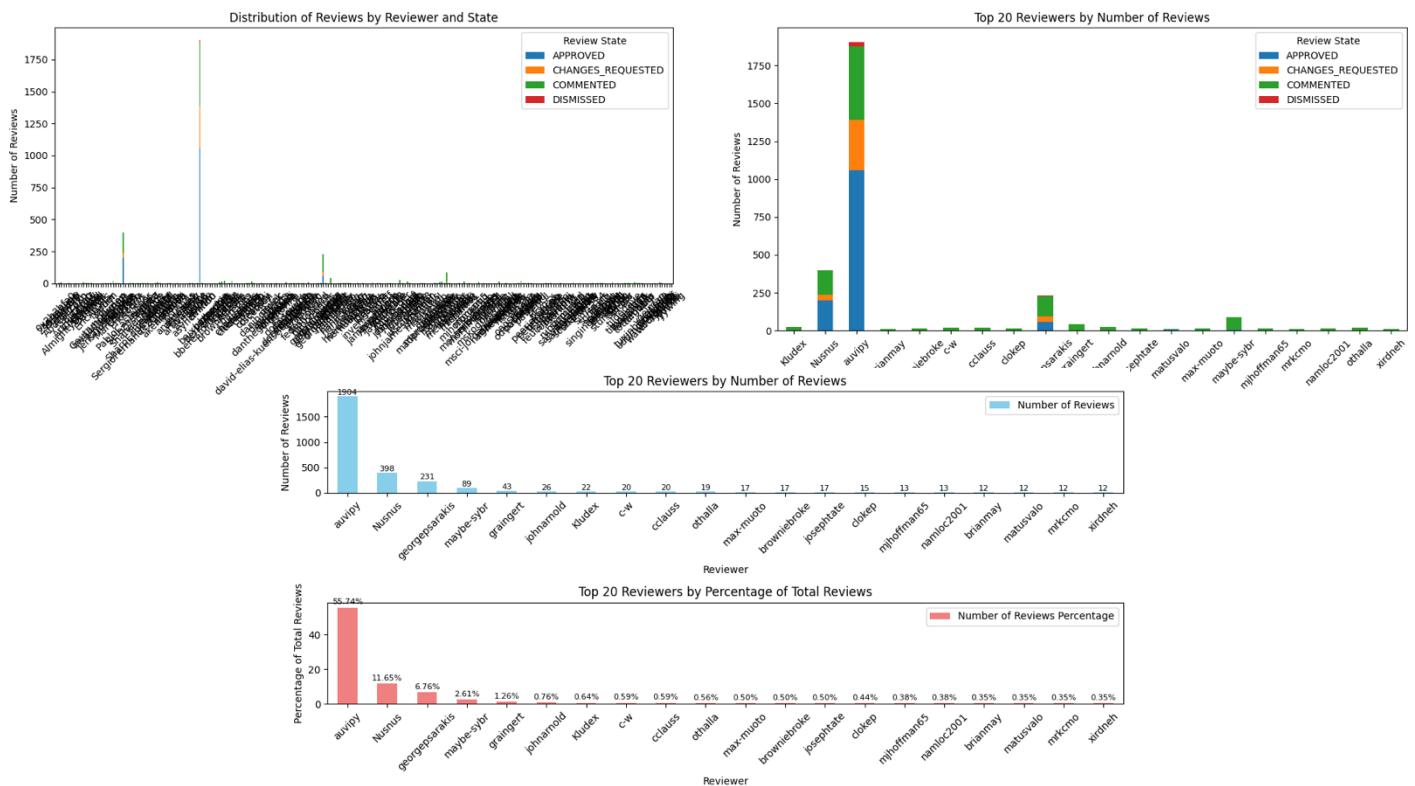
Insights (c)

- Around 81% of reviews workload was done by about 50% of reviewers in 2024.
- Around 82% of commits happened in this project was done by about 11% of committers in 2024.
- In most cases, there were a few committers who did the review jobs too. Also, there was a username who did the review jobs in a high number which was higher than his/her commits in 2024. Moreover, there was a username who committed the code in a high commit count but did not have any review works.

5. celery/celery

a. Code reviews by a developer and its distribution in the team throughout the project history

- Number of reviewers: 250
- Number of reviews: 3416
- Number of committers: 1021
- Number of commits: 12028



Insights (a):

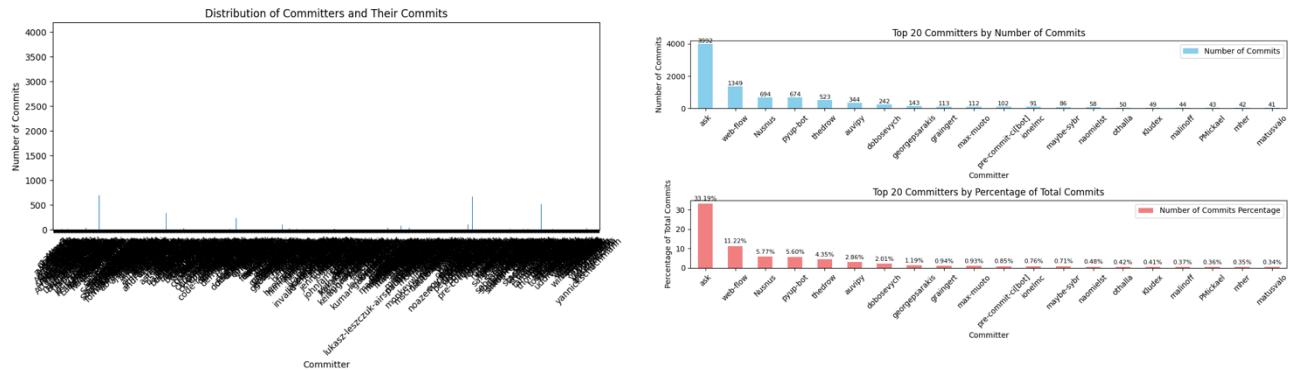
- There were a few reviewers who were dominant in the review jobs throughout this project history.
- Based on the portion of the states, in most cases, the commented states were dominant. Moreover, the highest number of review count done by one reviewer

showing the approved state portion was bigger than the others. Also, in this project history, the state of dismissed was exist.

- Almost 56% of review jobs throughout this project was conducted by one reviewer only.

b. Code contribution count by a developer and its distribution throughout the project history.

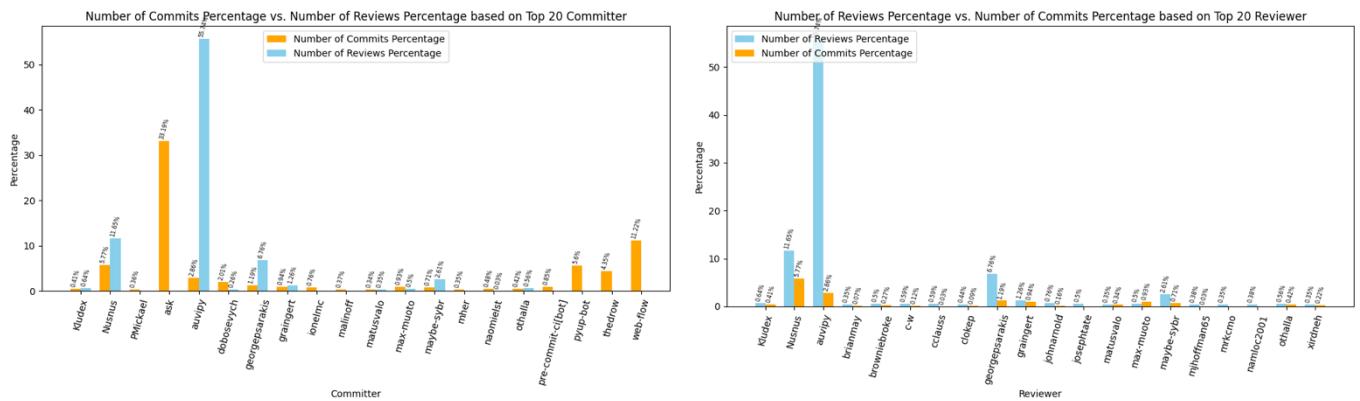
b1. In terms of commits count



Insights (b1)

- There were a few committers who were dominant in the commit jobs throughout this project history.
- Around 51% of all commits were committed by 3 committers out of 1021 in total.

b2. In terms of number of commits versus number of reviews based on top 20 committer and number of reviews versus number of commits based on top 20 reviewer.



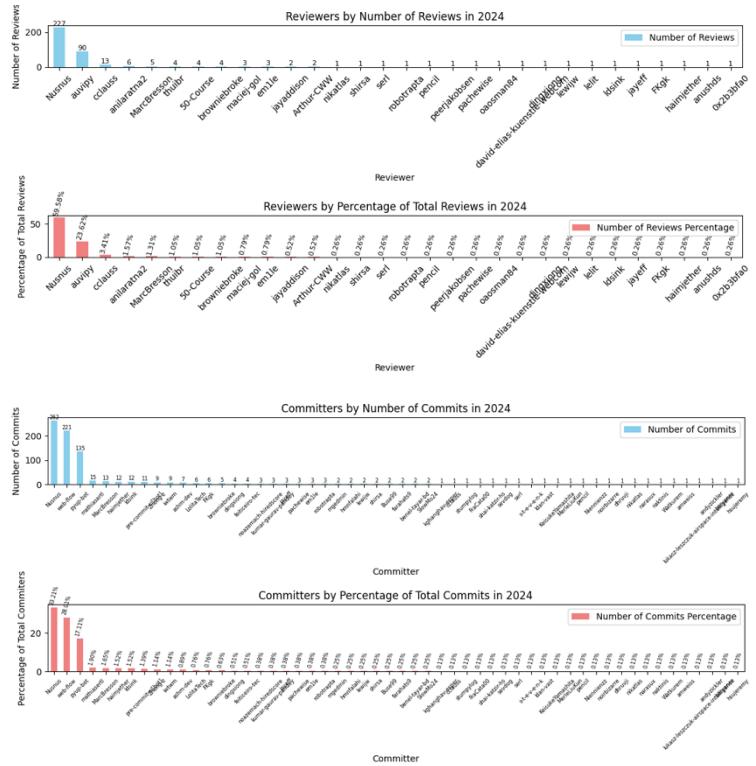
Insights (b2)

- Based on top 20 committers, not all committers had job as a reviewer too.
- The committers who did the review jobs too had a significant number of reviews count. Moreover, there were many reviewers who did not commit any code either.
- Based on top 20 reviewers, not all reviewers had job as a committer too.

- There were a few numbers of reviewer who committed the code too. Also, a reviewer who has high number of reviews count did the review work too.

c. Code contribution count by a developer and its distribution in the last year (2024).

- Number of reviewers: 30
- Number of reviews: 381
- Number of committers: 54
- Number of commits: 789



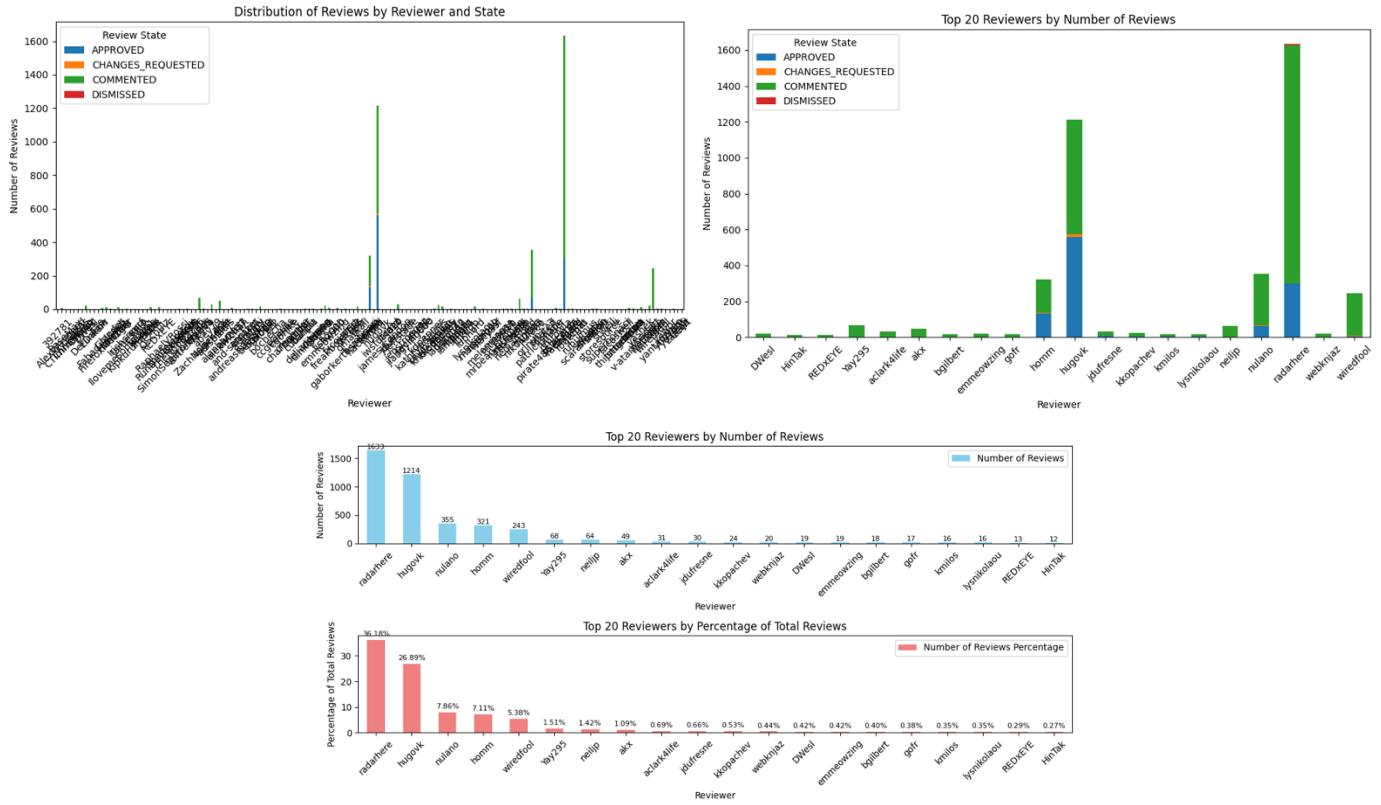
Insights (c)

- Around 83% of reviews workload was done by about 6.7% of reviewers in 2024.
- Around 78% of commits happened in this project was done by about 5.6% of committers in 2024.
- The last chart illustrates that there were 3 committers with high number of commits count including one committer only who do the review works too in 2024.

6. python-pillow/Pillow

a. Code reviews by a developer and its distribution in the team throughout the project history

- Number of reviewers: 155
- Number of reviews: 4514
- Number of committers: 407
- Number of commits: 14138

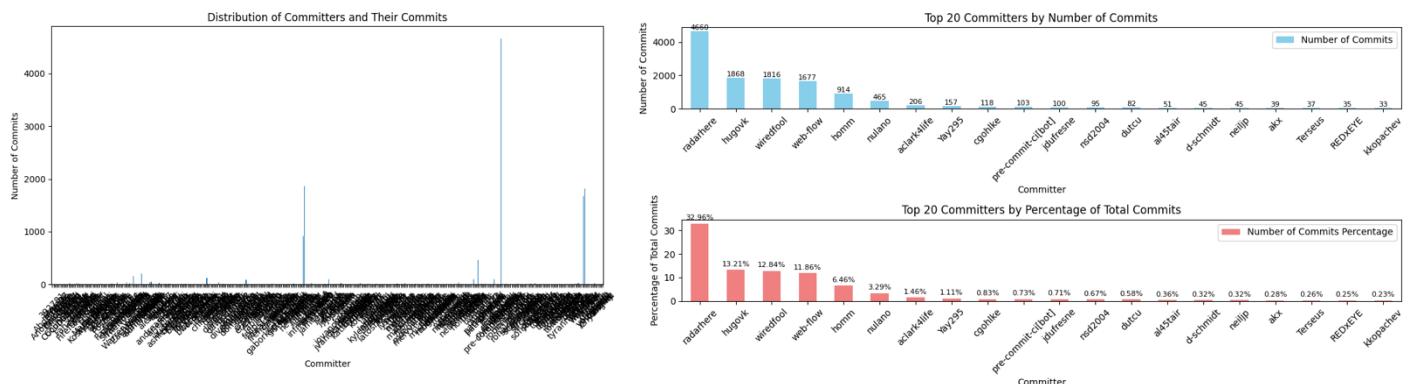


Insights (a):

- There were a few reviewers who were dominant in the review jobs throughout this project history.
- Based on the state portion, the portion of commented state in most reviewers were higher than the approved state and the others. Also, the dismissed state was existed in this project even though there is not illustration of it in the chart of top 20 reviewers.
- Around 83% of review works were done by about 3.2% of all reviewers.

b. Code contribution count by a developer and its distribution throughout the project history.

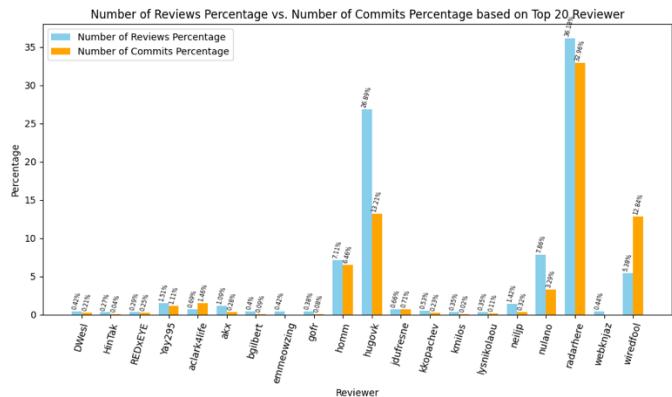
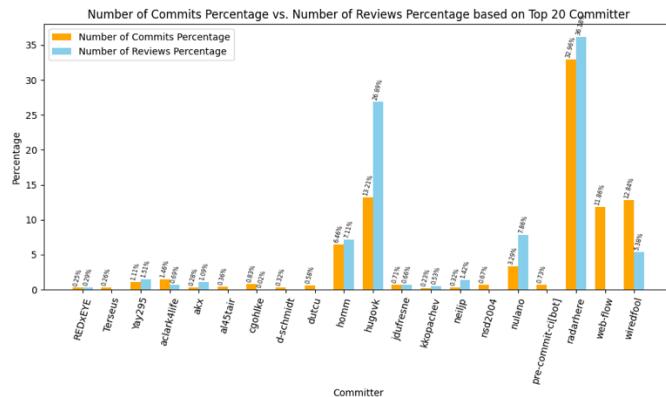
b1. In terms of commits count



Insights (b1)

- There were a few committers who were dominant in the commit jobs throughout this project history.
- Around 81% of all commits happened in this project was done by about 1.4% of all committers in this project.

b2. In terms of number of commits versus number of reviews based on top 20 committer and number of reviews versus number of commits based on top 20 reviewer.

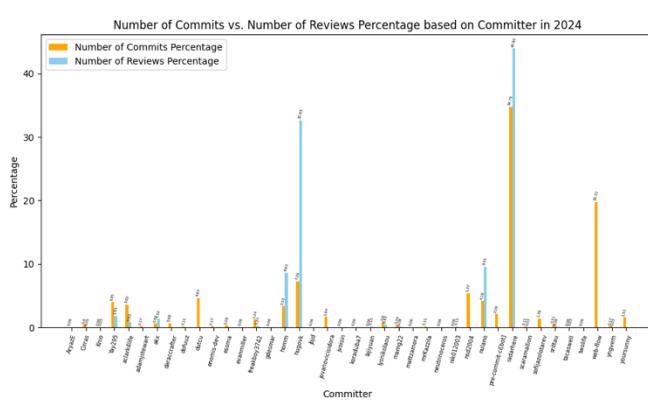


Insights (b2)

- Based on top 20 committers, not all committers had job as a reviewer too.
- There were a few committers who worked as a reviewer too and reviewed in a high number of reviews.
- Based on top 20 reviewers, not all reviewers had job as committer too.
- Overall, there were only a few usernames who were dominant in doing both commits and reviews.

c. Code contribution count by a developer and its distribution in the last year (2024).

- Number of reviewers: 35
- Number of reviews: 891
- Number of committers: 40
- Number of commits: 1770



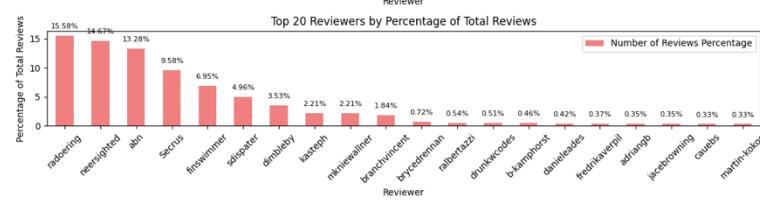
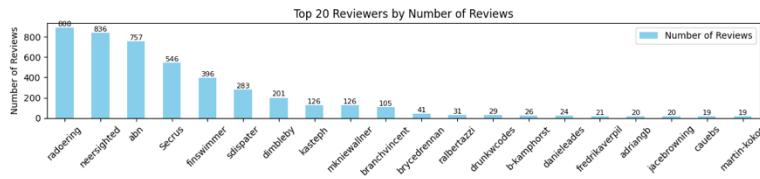
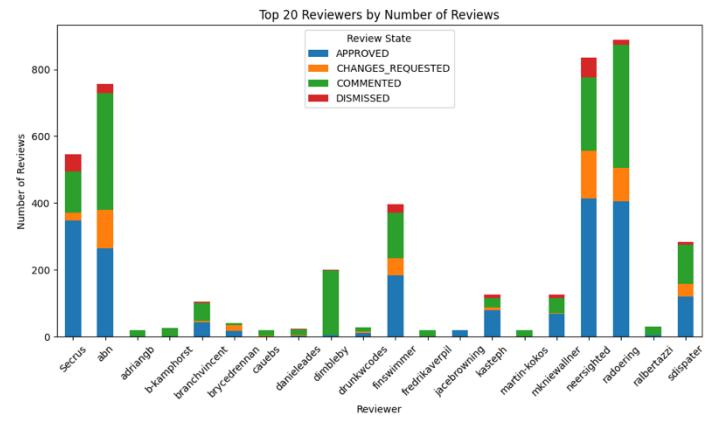
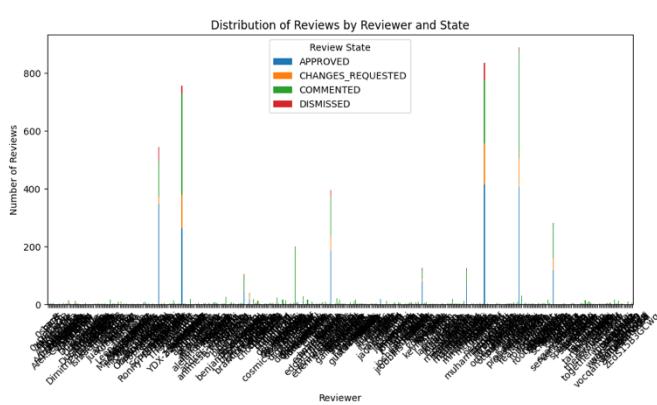
Insights (c)

- Around 81% of reviews work was done by 3 reviewers only (about 8.6% of all reviewers) in 2024.
- Around 80% of commits happened in 2024 was conducted by 17.5% of all committers in 2024. It shows that the dominance of committers is lower than the dominance of reviewers.
- There were a few committers who did the review jobs too. There were some committers who also reviewed the code had high number of review count illustrated in the last chart.

7. python-poetry/poetry

a. Code reviews by a developer and its distribution in the team throughout the project history

- Number of reviewers: 424
- Number of reviews: 5700
- Number of committers: 641
- Number of commits: 8384



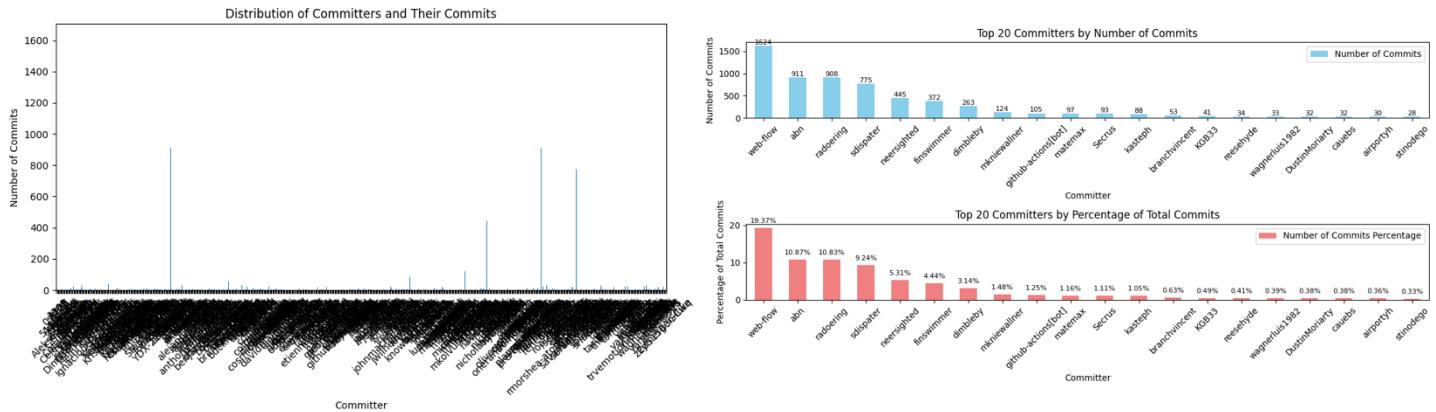
Insights (a):

- There were a few reviewers who were dominant in the review jobs throughout this project history. However, the dominance of reviewers in previous projects was still higher than this project.

- Based on the portion of states, there was significant portion of change_requested change in some reviewers' cases. Also, the state of dismissed was existed.
- Around 73% of review workload was conducted by 2.1% of all reviewers.

b. Code contribution count by a developer and its distribution throughout the project history.

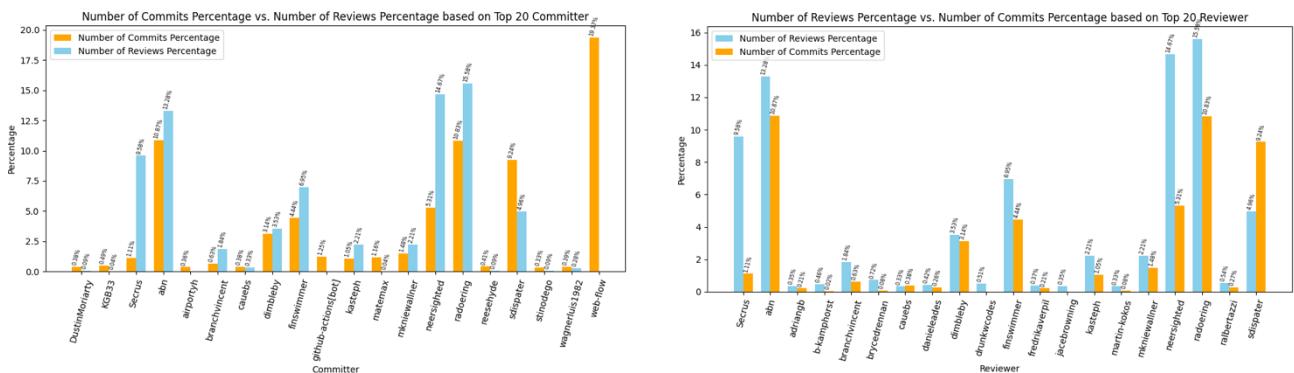
b1. In terms of commits count



Insights (b1)

- There were a few committers who were dominant in committing the code throughout this project.
- Around 63% of commits happened in this project was done by about 1.1% of all committers.

b2. In terms of number of commits versus number of reviews based on top 20 committer and number of reviews versus number of commits based on top 20 reviewer.



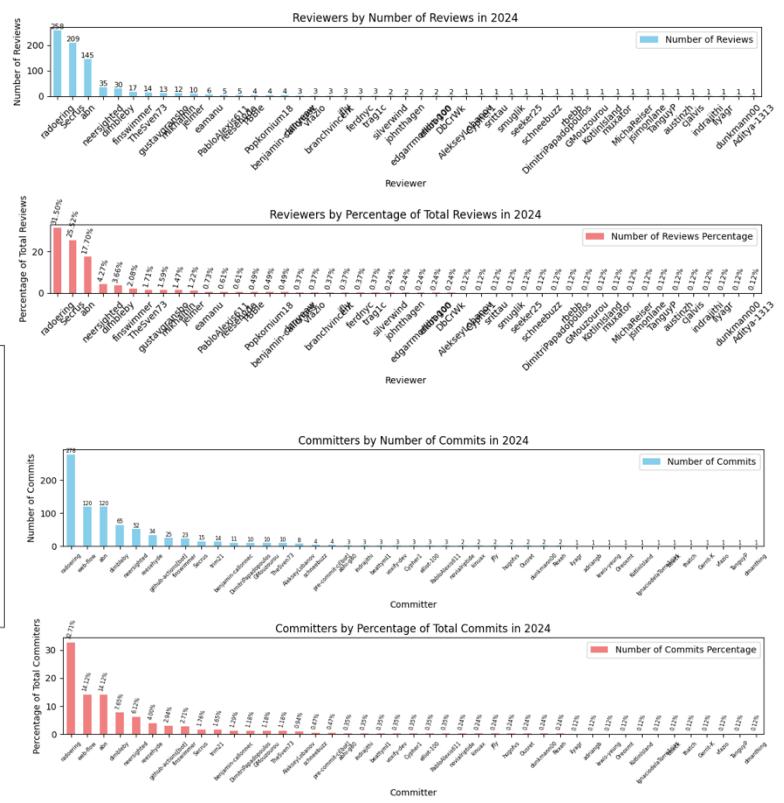
Insights (b2)

- Based on top 20 committer, there were many committers who did reviewing the code too including committers who had high number of reviews count.

- There was a username of committer who had highest number of commits but not doing any review works.
 - Based on top 20 reviewers, there were many reviewers who committed the code too. Those reviewers did not have only high number of reviews count, but also had high number of commits count.
 - Overall, there were some usernames who were dominant in doing both commits and reviews.

c. Code contribution count by a developer and its distribution in the last year (2024).

- Number of reviewers: 47
 - Number of reviews: 819
 - Number of committers: 43
 - Number of commits: 850



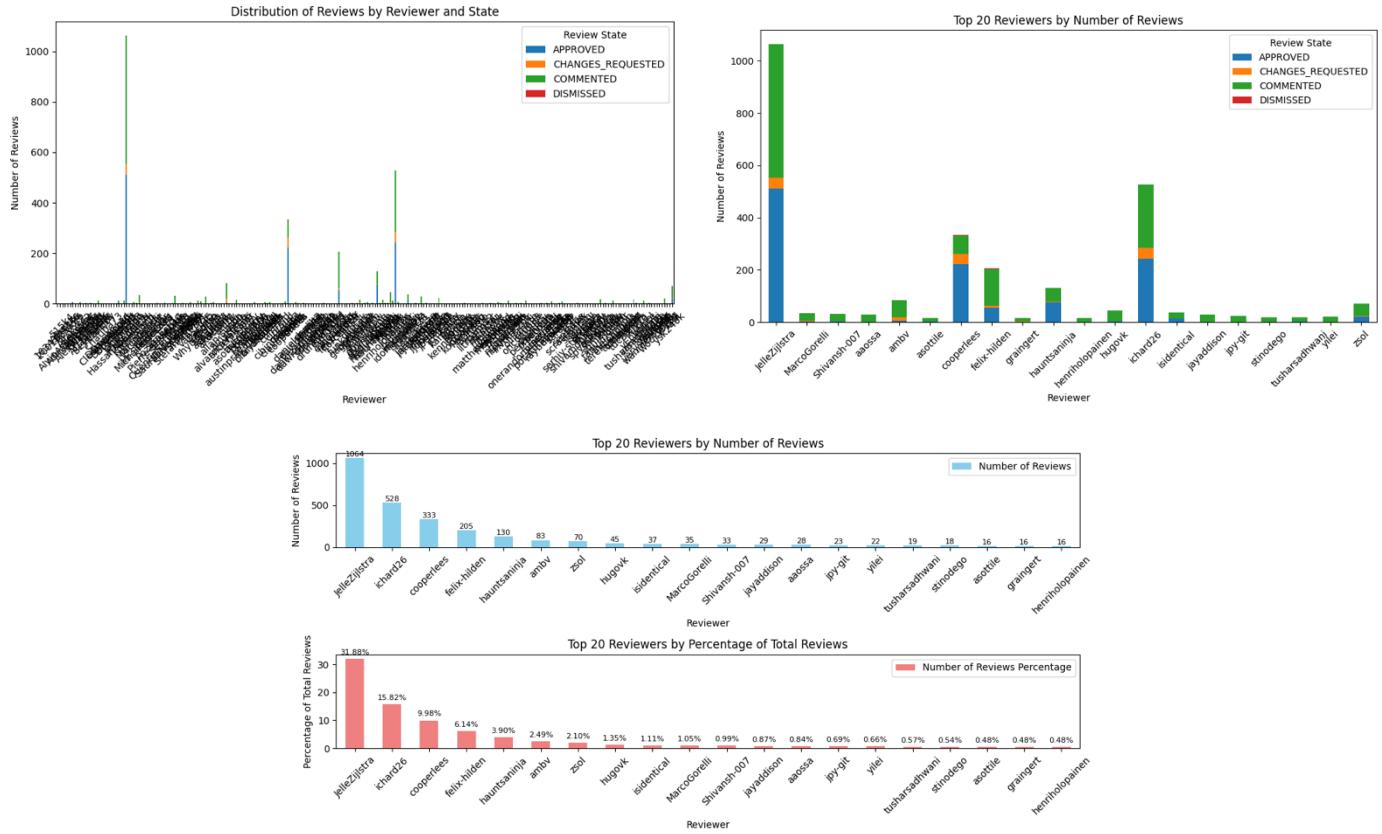
Insights (c)

- Around 75% of review jobs was done by about 6.4% of all reviewers in 2024.
 - Around 75% of commits happened in 2024 was done by about 11.6% of all committers in 2024.
 - There were a few committers who did the review jobs too. Moreover, there was a username who committed the code in a high commit count but did not have any review works.

8. psf/black

a. Code reviews by a developer and its distribution in the team throughout the project history

- Number of reviewers: 241
- Number of reviews: 337
- Number of committers: 403
- Number of commits: 5890

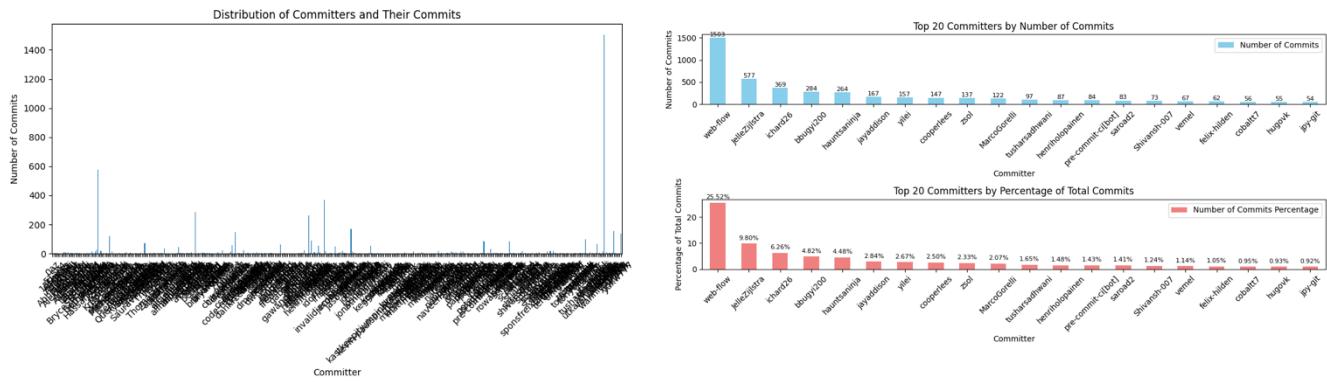


Insights (a):

- There were a few reviewers who were dominant in this project.
- Based on the portion of states, the commented states were still dominant, but in the two highest number of reviews, the portion of commented and approved were almost the same. Also, the dismissed states existed in this project.
- Around 72% of all reviews were done by about 3% of all reviewers throughout this project.

b. Code contribution count by a developer and its distribution throughout the project history.

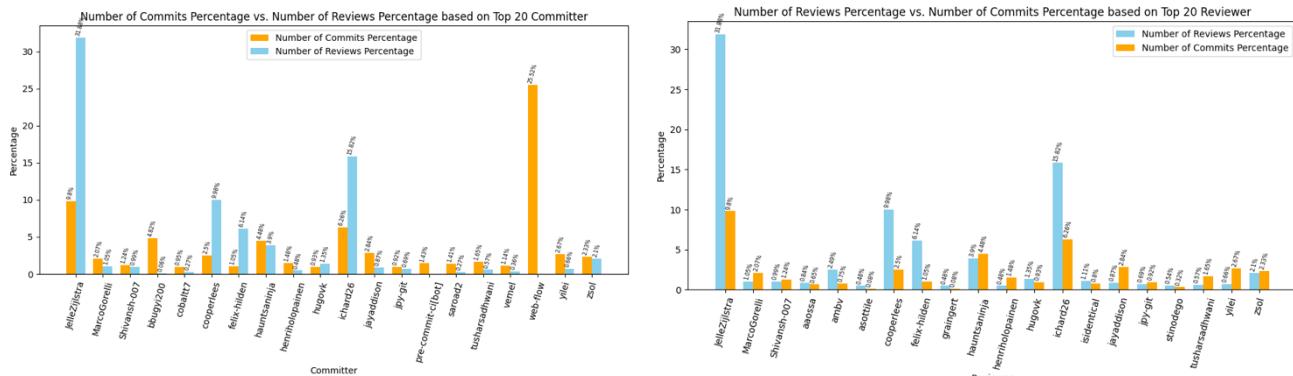
b1. In terms of commits count



Insights (b1)

- The distribution of commit works was distributed much better than previous projects.
- Around 63% of all commits were done by about 2.5% of all committers.

b2. In terms of number of commits versus number of reviews based on top 20 committer and number of reviews versus number of commits based on top 20 reviewer.

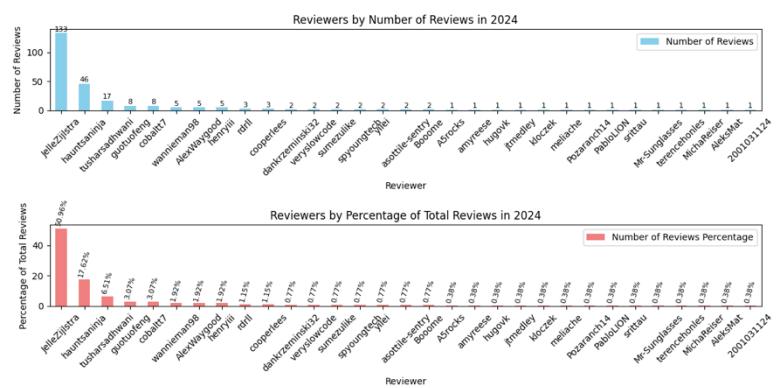


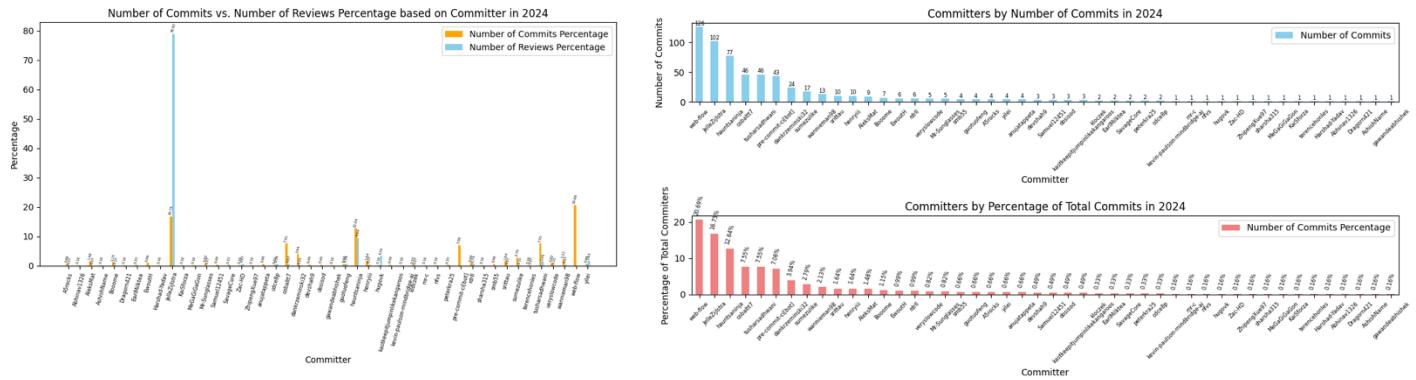
Insights (b2)

- Based on top 20 committer, not all committers had job as a reviewer too but in a small portion.
- Around 85% of all top 20 committers had review works.
- Based on top 20 reviewer, not all reviewers had job as a committer too but in a small portion.
- Around 90% of all top 20 reviewers became a committer too in this project.

c. Code contribution count by a developer and its distribution in the last year (2024).

- Number of reviewers: 31
- Number of reviews: 261
- Number of committers: 46
- Number of commits: 609





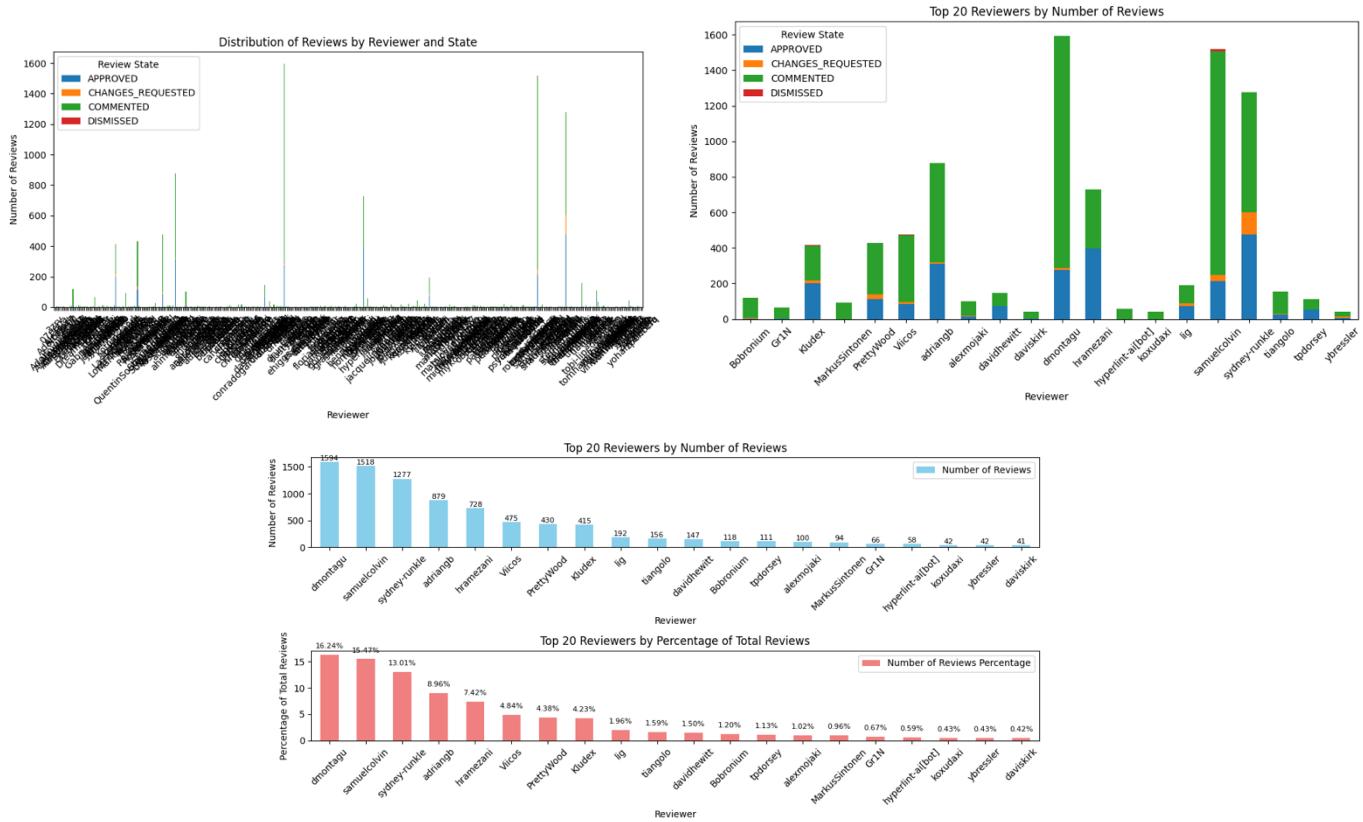
Insights (c)

- Around 75% of review works were done by about 9.7% of all reviewers in 2024.
- Around 76% of commits were done by about 15% of all committers.
- There were not many committers became a reviewer too in 2024. Also, there was a committer who had high number of reviews works too.

9. pydantic/pydantic

a. Code reviews by a developer and its distribution in the team throughout the project history

- Number of reviewers: 402
- Number of reviews: 9814
- Number of committers: 613
- Number of commits: 13903

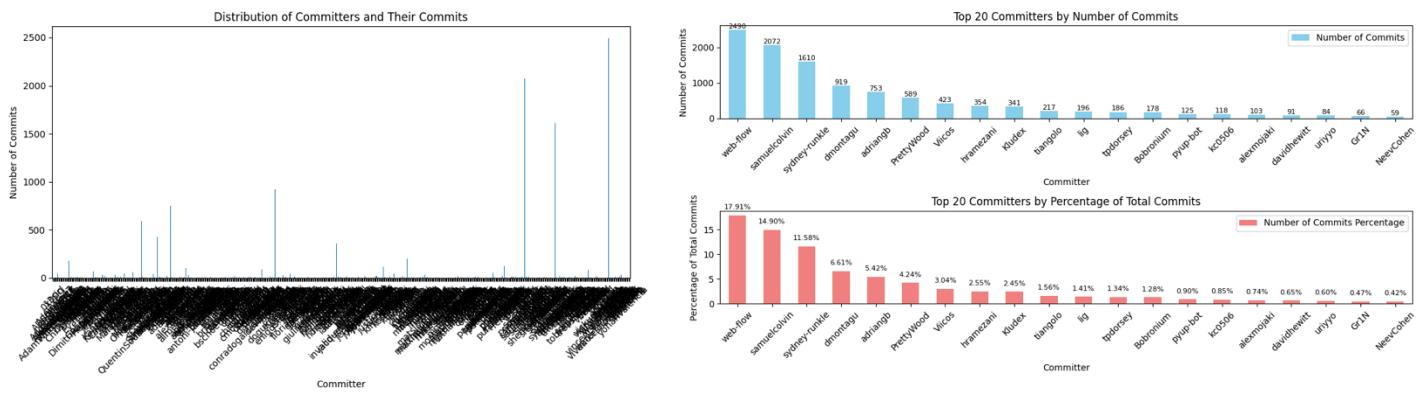


Insights (a):

- There were a few reviewers who were dominant in reviewing the code in this project.
- Based on the portion of states, the commented states were still dominant compared to other states. There was some change_requested states too in a high workload of review bar but in a small number. Also, there were dismissed states but not significant.
- Around 75% of all reviews were done by about 2% of all reviewers in this project.

b. Code contribution count by a developer and its distribution throughout the project history.

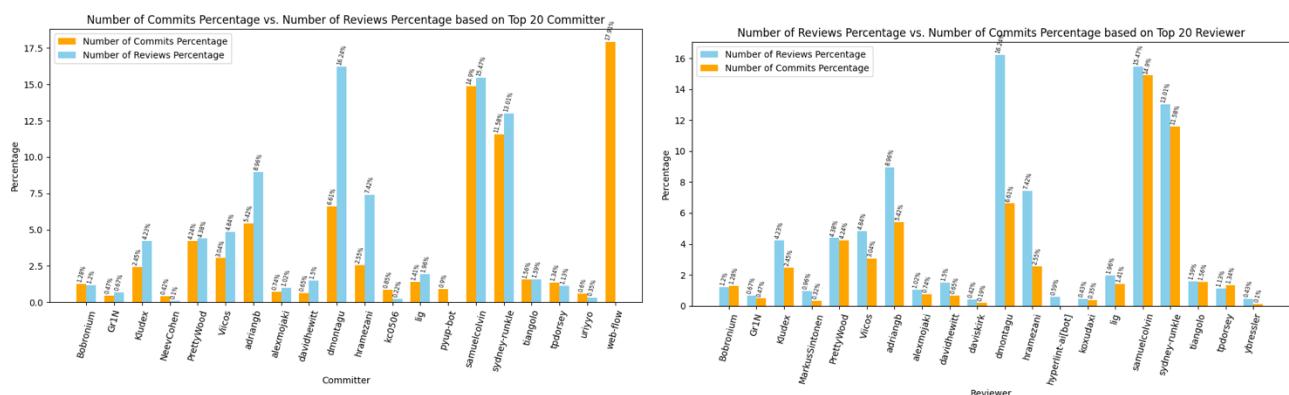
b1. In terms of commits count



Insights (b1)

- There were a few committers who were dominant in committing codes in this project.
- Around 69% of all commits were committed by about only 1.5% of all committers in this project.

b2. In terms of number of commits versus number of reviews based on top 20 committer and number of reviews versus number of commits based on top 20 reviewer.

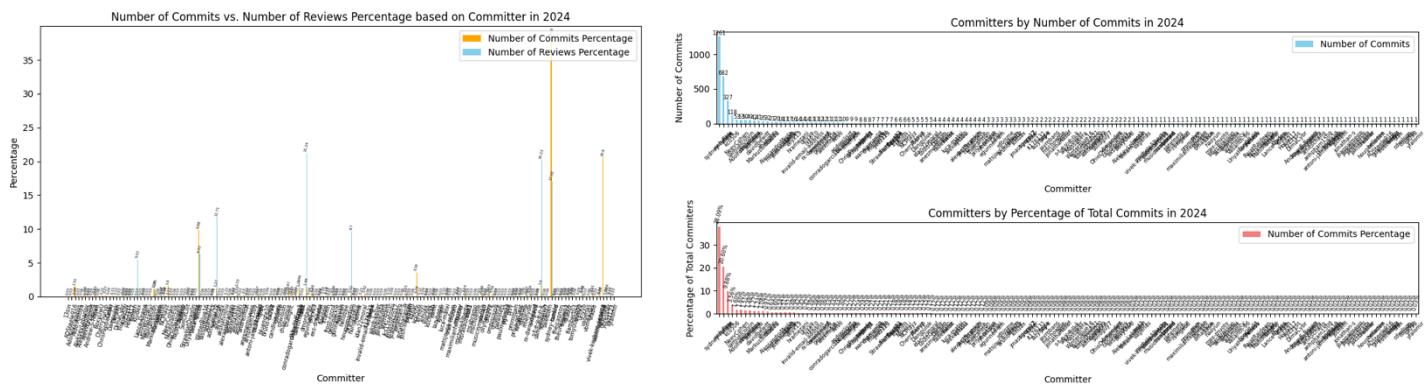
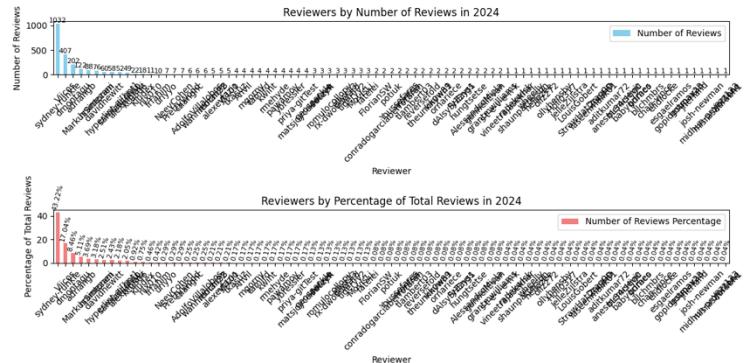


Insights (b2)

- Both charts show significant number for both commits and reviews work.
- Around 90% of all committers in top 20 committers became a reviewer too in this project.
- There was a committer who committed the code in a high frequency but did not have any review works in this project.
- Around 95% of all reviewers in top 20 reviewers became a committer too.
- In most cases, a committer who had high workload also had significant reviews work too and vice versa.

c. Code contribution count by a developer and its distribution in the last year (2024).

- Number of reviewers: 87
- Number of reviews: 2388
- Number of committers: 159
- Number of commits: 3311



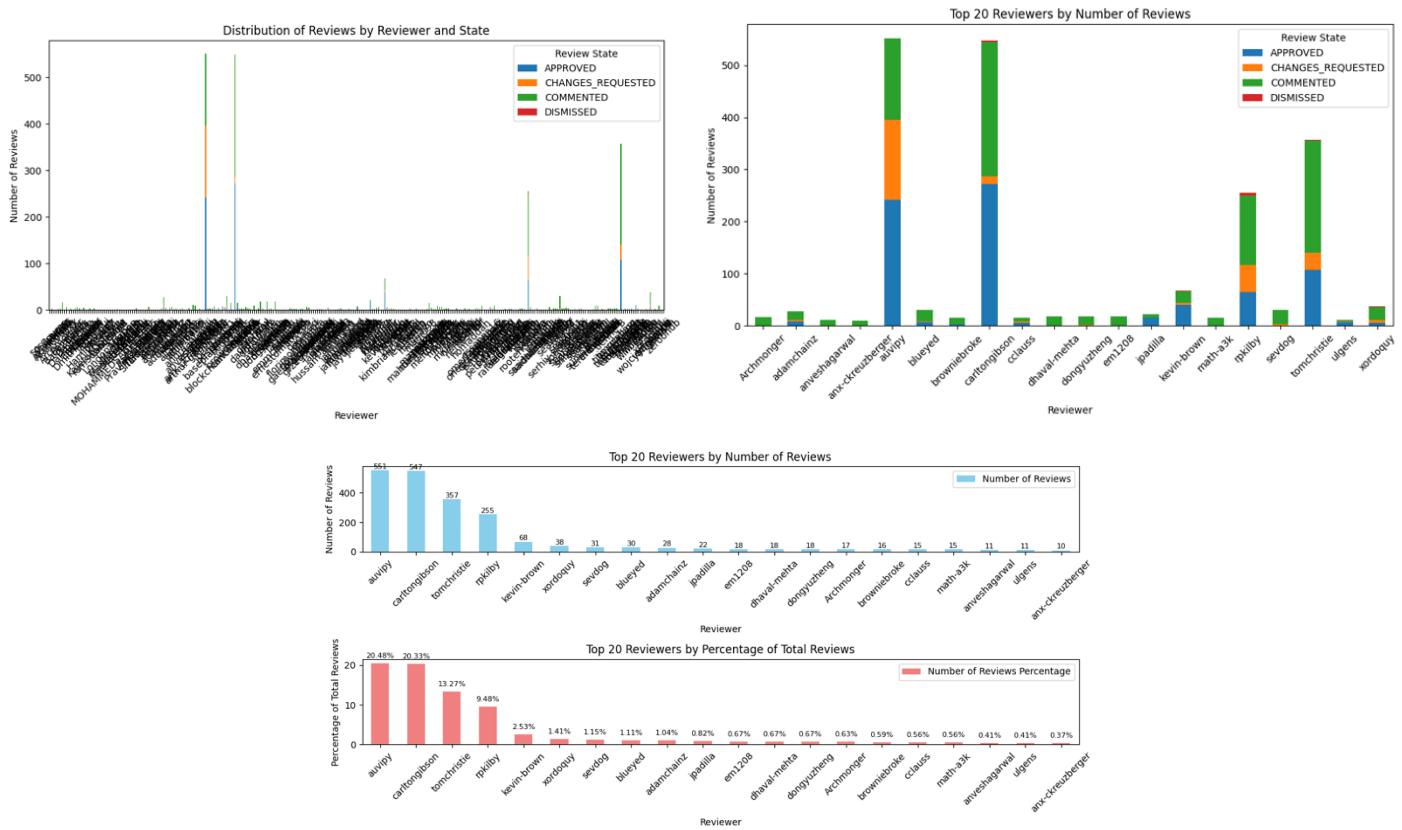
Insights (c)

- Around 77% of all reviews were done by about 5.7% of all reviewers in 2024.
- Around 62% of all commits happened in 2024 were done by 2.5% of all committers in 2024.
- There were a few usernames who were dominant in jobs of committing and reviewing in 2024. Not all committers became a reviewer too and vice versa.

10. encode/django-rest-framework

a. Code reviews by a developer and its distribution in the team throughout the project history

- Number of reviewers: 292
- Number of reviews: 2691
- Number of committers: 1468
- Number of commits: 11815

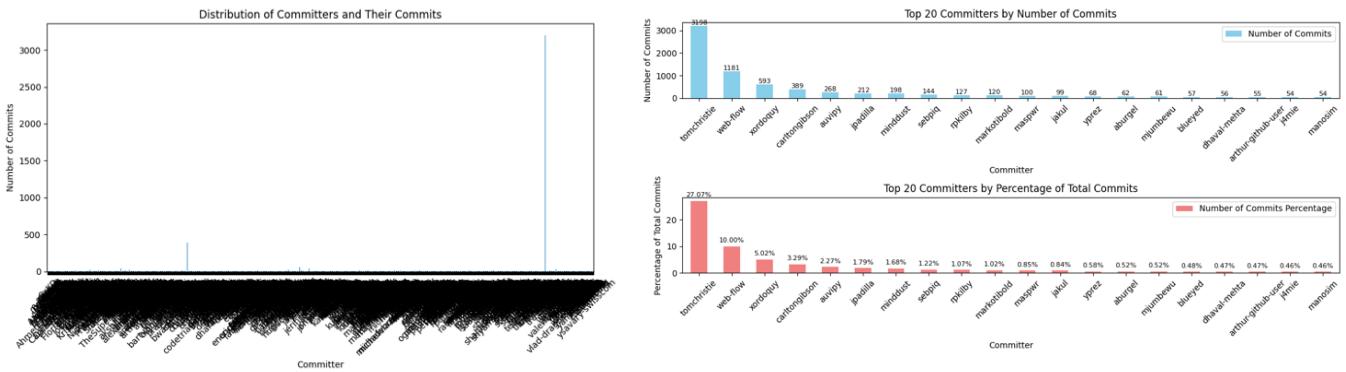


Insights (a):

- There were a few reviewers who were dominant in reviewing throughout this project.
- Based on the portion of states, state that were most happened was commented compared to the others. The dismissed state was existed in this project.
- Around 66% of all reviews were done by about 1.7% only of all reviewers in this project.

b. Code contribution count by a developer and its distribution throughout the project history.

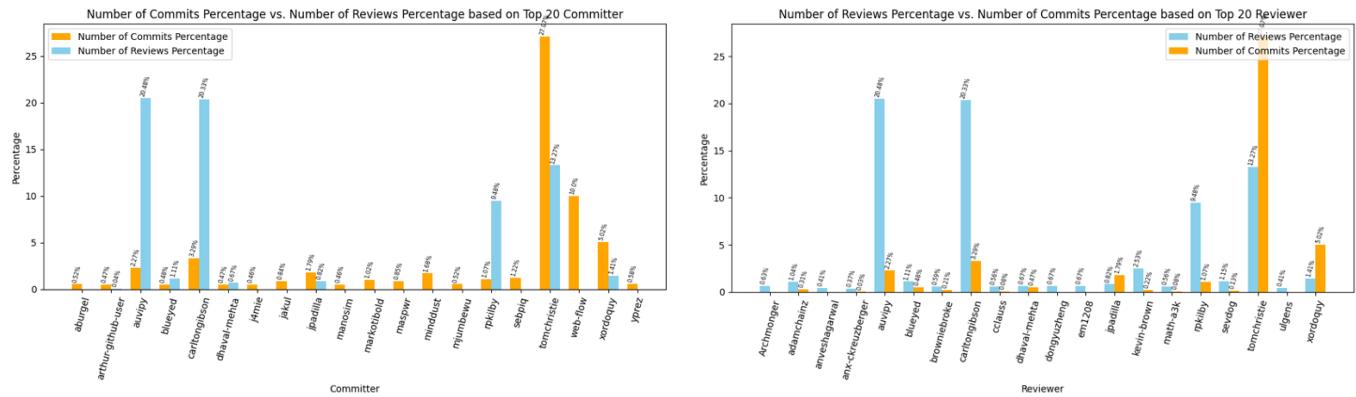
b1. In terms of commits count



Insights (b1)

- There were a few committers who were dominant in committing codes in this project.
- Around 50% of all commits happened in this project were done by only 0.5% of all committers.

b2. In terms of number of commits versus number of reviews based on top 20 committer and number of reviews versus number of commits based on top 20 reviewer.

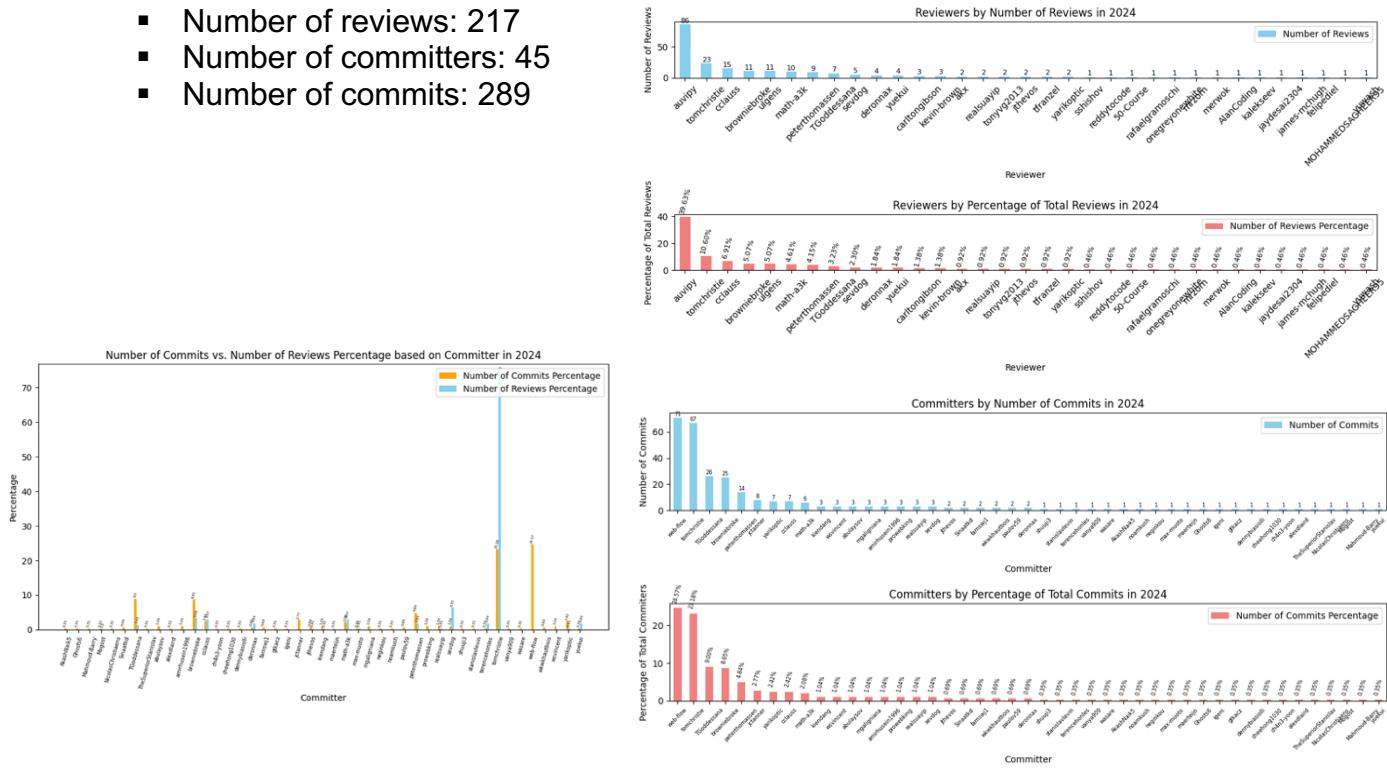


Insights (b2)

- Based on top 20 committer, there were a few committers who had review works too.
- There was a username who had high commits count becoming a reviewer with significant number of reviews works too.
- Based on top 20 reviewer, there were a few reviewers who became committer too.
- The reviewers who had high number of reviews count committed codes which were not in a high number but significant enough compared to the others.

c. Code contribution count by a developer and its distribution in the last year (2024).

- Number of reviewers: 33
- Number of reviews: 217
- Number of committers: 45
- Number of commits: 289



Insights (c)

- Around 71% of all reviews were done by about 21% of all reviewers in 2024.
- Around 70% of all commits were done by about 11% of all committers in 2024.
- There was a small number of committers who had reviews work too in 2024. Moreover, there was a username of committer who became a reviewer too with high number of reviews count.