

# GitHub ETL Project

## Activity Report

I have chosen to create an activity report of the public GitHub repository [tensorflow/tensorflow](https://github.com/tensorflow/tensorflow), an end-to-end open-source platform for machine learning. In particular, I chose activities defined as commits, contributors, issues, and pull requests.

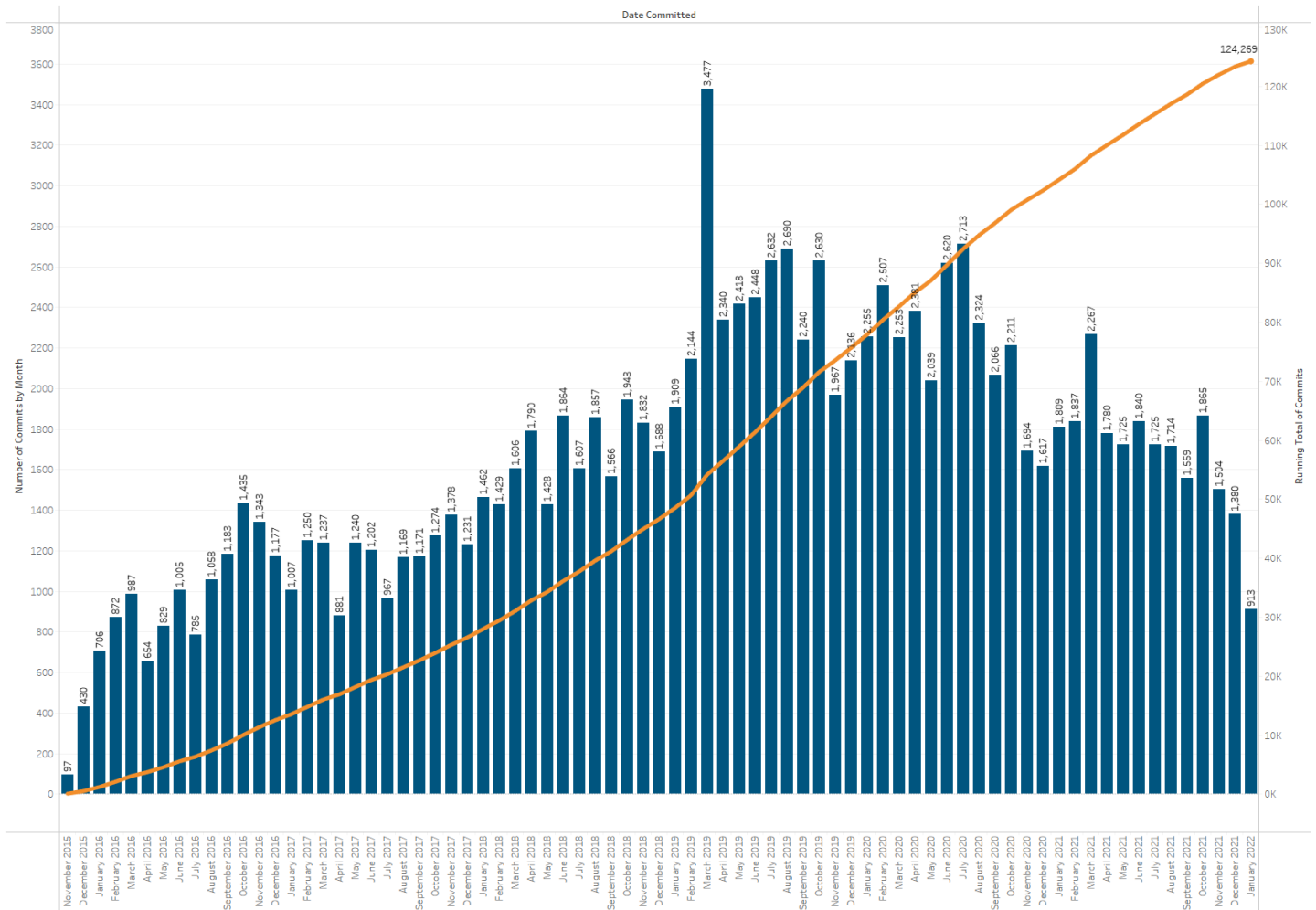


Fig.1 The number of commits made to tensorflow/tensorflow; the number of monthly commits as a bar chart and the running total of monthly commits as a line.

Since the initial commit made on November 7, 2015, there have been 408 distinct contributors who (as of January 25, 2022) have created 124,269 commits, 53,624 issues, and 19,770 pull requests for tensorflow/tensorflow.

One particular contributor, aptly-named with the login tensorflow-gardener, is responsible for 34,722 of the 105,226 total contributions (33% of total contributions). They made their first commit to the repository in June 2016 and, ever since, they have made significantly more contributions than any other contributor. There are other contributors who have been consistently active for similar lengths of time. For example, "benoitsteiner" made their first commit to the repository in December 2015, then made commits every month before suddenly stopping in June 2018. Shortly after, "mihaimaruseac" made their first commit to the repository in September 2018, and has made commits every month since.

It is relatively rare that a commit receives a comment: There have only been a total of 699 comments made across all commits. In contrast, a total of 293,619 comments have been made across all issues; on average, every issue receives approximately 5.5 comments.

Of the 53,624 issues in tensorflow/tensorflow, 2,574 are open and 51,050 have been closed. The longest an issue has been open before being closed is 2,128 days (issue 360 was created on November 26, 2015 and closed on September 24, 2021). The oldest issue that is still open was created 2,262 days ago (issue 252 was created on November 17, 2015).

Of the 19,770 pull requests in tensorflow/tensorflow, 214 are open and 19,556 have been closed. The longest a pull request has been open before being closed is 764 days (pull request 31,750 was created on August 19, 2019 and closed on September 21, 2021). The oldest pull request that is still open was created 1,010 days ago (pull request 28,049 was created on April 22, 2019).

As a popular open-source platform, tensorflow/tensorflow has been active since its conception, showing no signs of slowing down. At any given time, there are many willing to make contributions in the form of commits, issues, and pull requests.

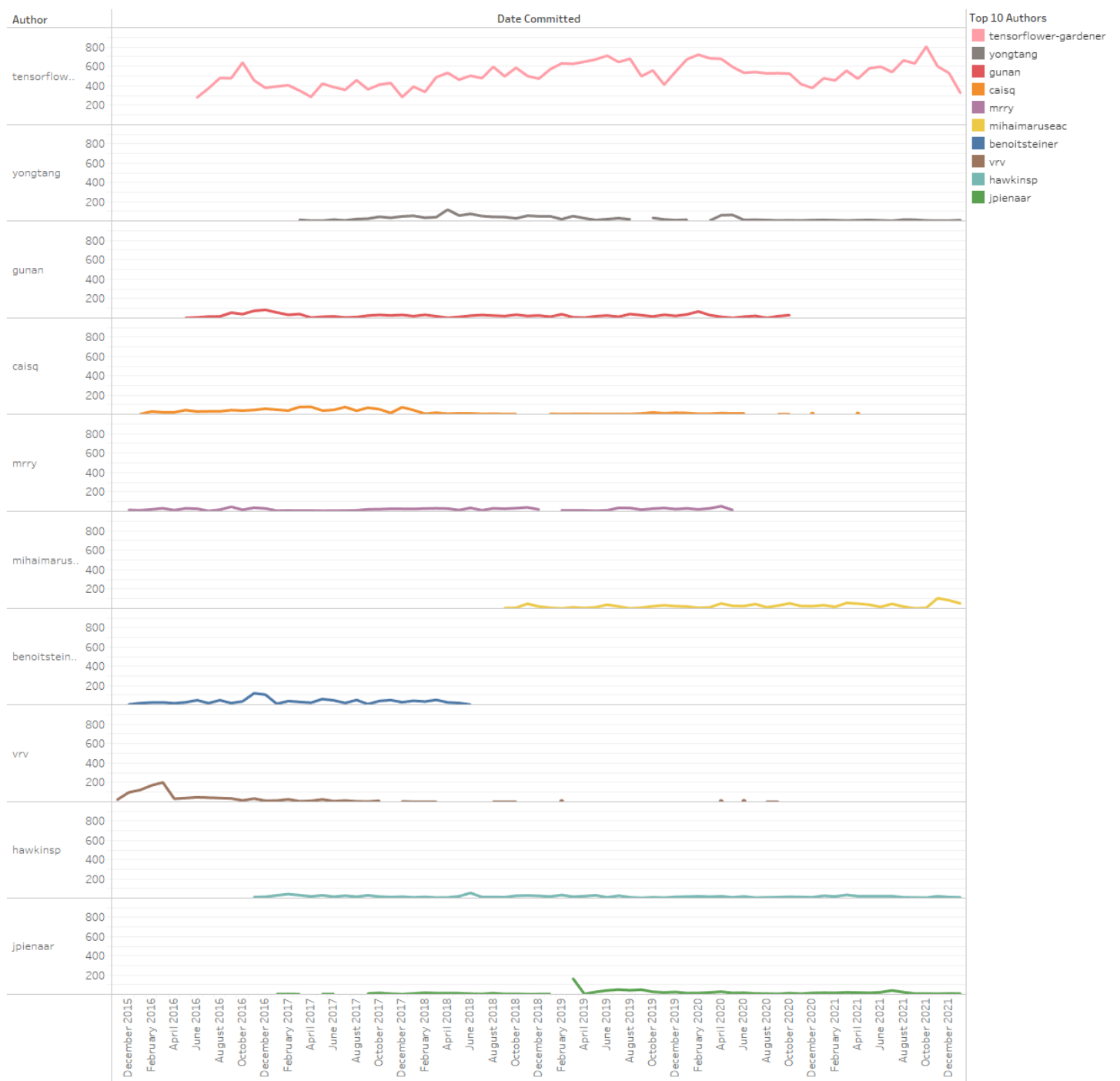


Fig.2 The number of commits made over time by the 10 most prolific contributors to tensorfloww/tensorflow.