

R & R

Jacob Oursland & Elijah Grimshaw

February 14, 2017

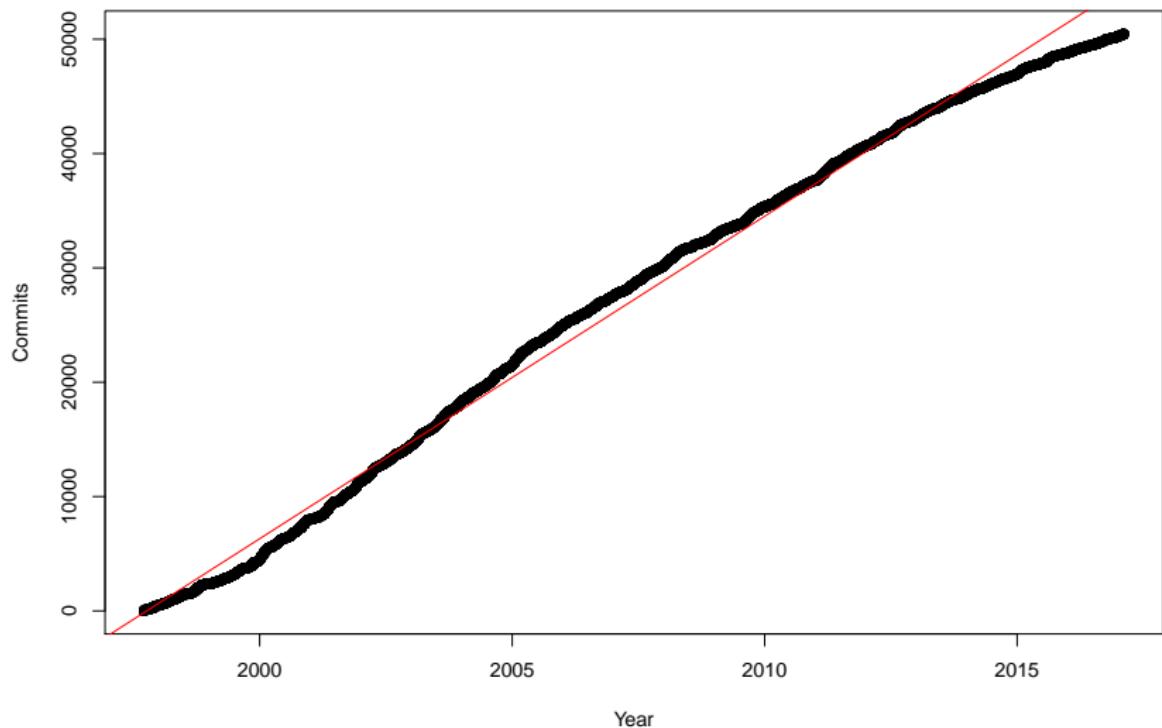
Inspiration

Video

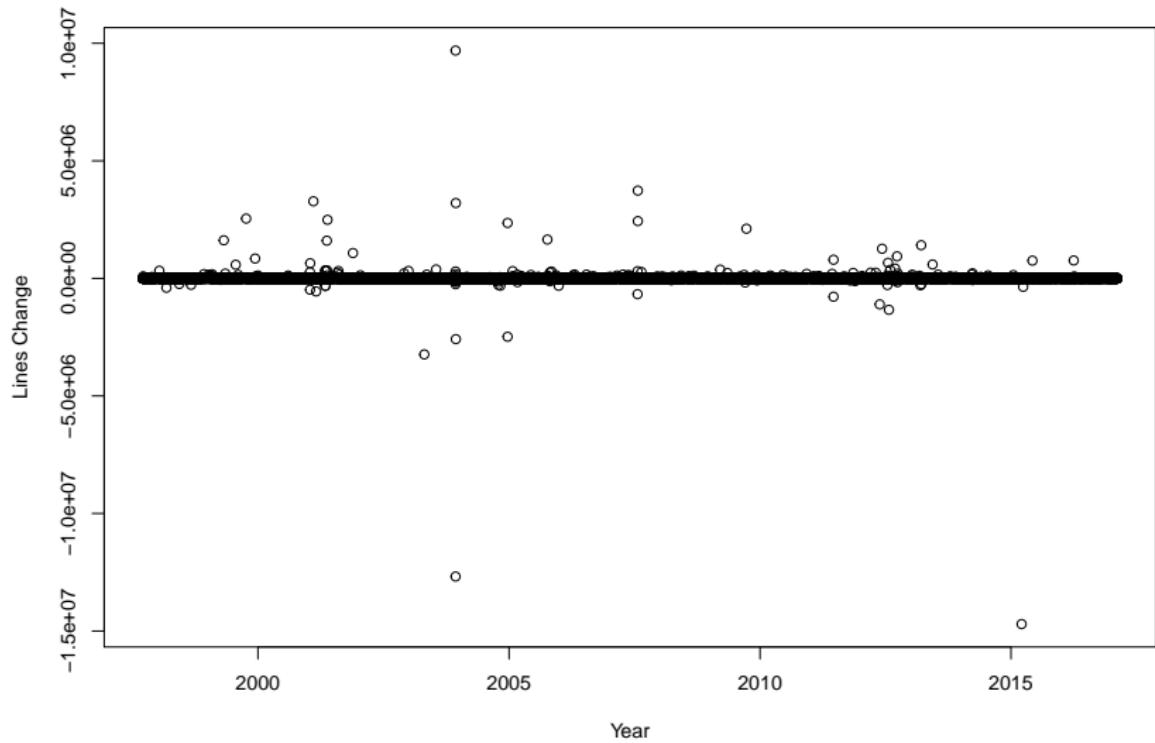
Collecting the Data

1. Clone the R source git repository:
<https://github.com/wch/r-source>
2. Use a script that parses git --log into a JSON file that contains each commit, the author, date, paths modified, and lines changed to each file.
3. Use R on R!

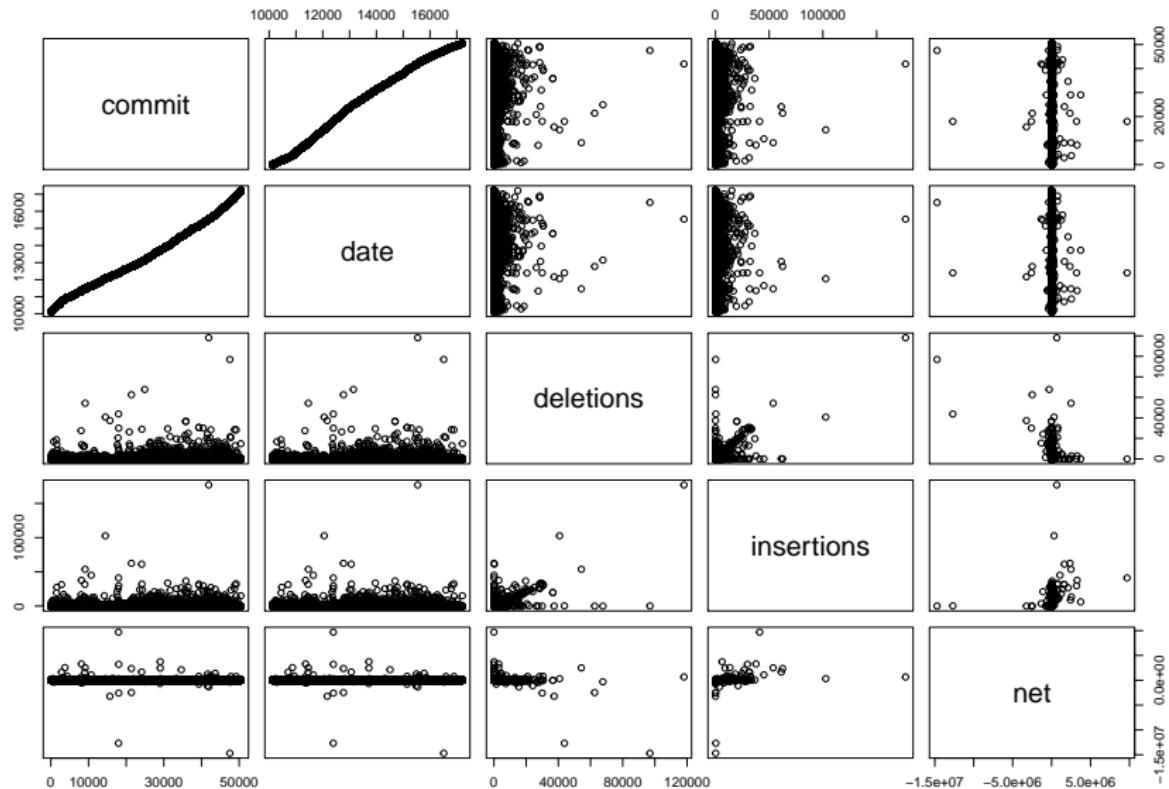
Commits Over Time



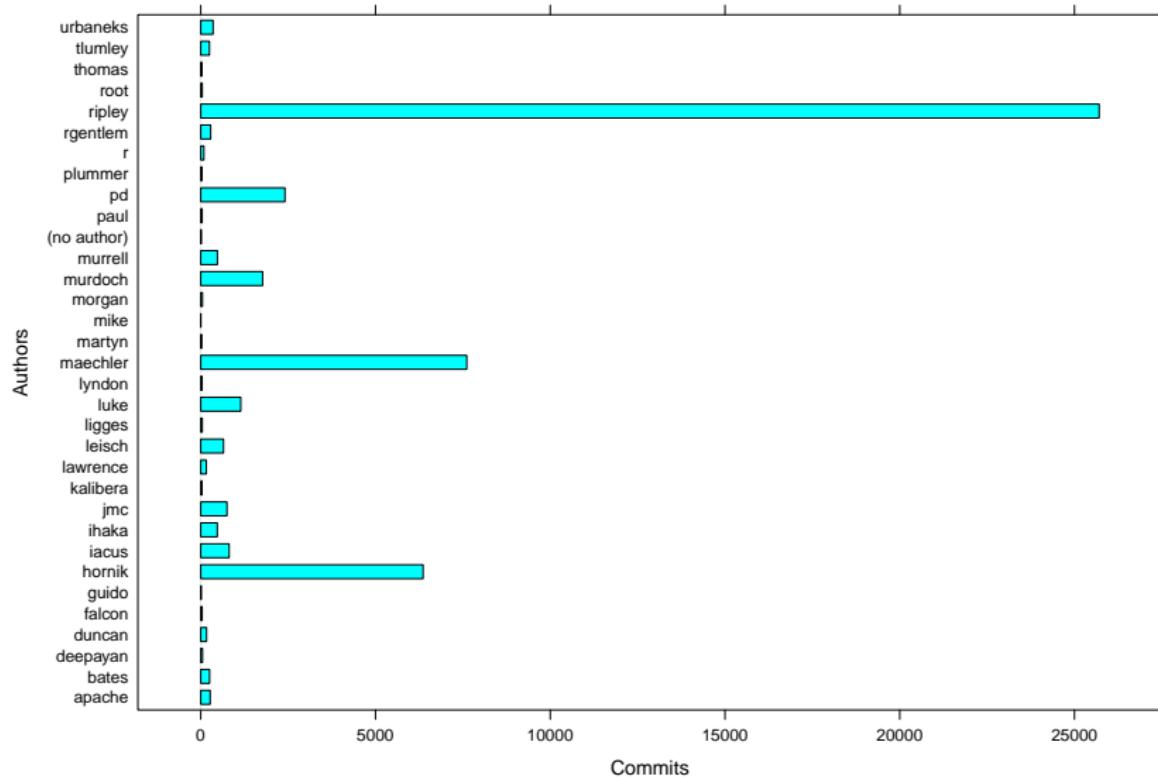
Size of Commits over Time



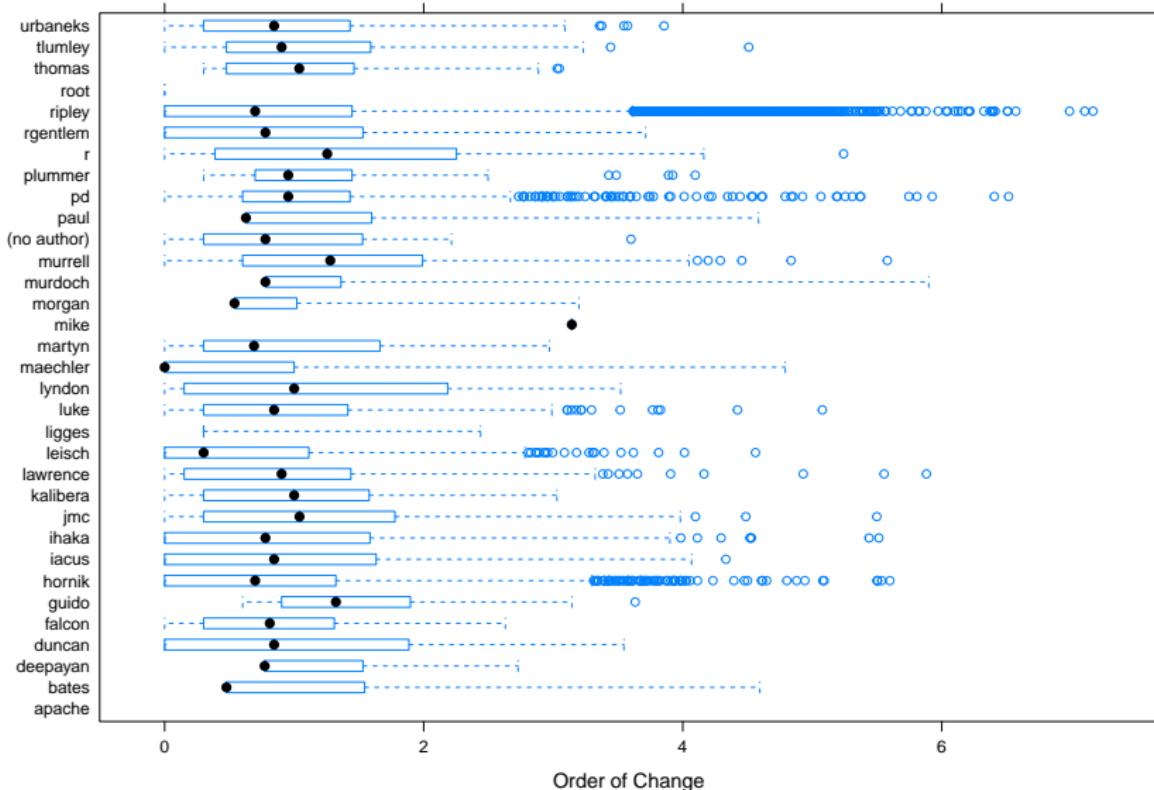
Overview



Commits per Author



Commit Sizes for Author



Conclusions

1. R is maintained by a core team, of which there are only about 4 major contributors.
2. Most large commits (features?) are by those 4 major contributors.
3. The other contributors are primarily involved in smaller commits (bugfixes?).

Further Exploration

1. Whose code experiences rapid turnover?
2. For each author, what are the statistics on average cyclomatic complexity index of their code?
3. Are 1 and 2 correlated?