PROJECT 3 - MSR

NINJA-BUILD

Natalie Simtion, Dustin Nguyen, Sawyer Jones & Youa Yang

AGENDA

Stage 1 – Issues

Stage 2 – Commits

Stage 3 – Code-size

Stage 4 – Issue Density

Introduction

- Our repository for this project is Ninja-Build
  - https://github.com/ninja-build/ninja

- Ninja is a small build system with focus on speed

- 10+ years on GitHub

ISSUES

Table – slice of issues data

	#	Title	URL	State	Labels	# of Labels	Assignees	# of Assignees	# of Comments	Closed at	Created at	Locked
1	2117	Segfault when openi…	https://github.com/nin…	closed	[]	0	[]	0	1	4/2/2022 20:40	4/2/2022 20:13	FALSE
2	2116	linux: `pool = console…	https://github.com/nin…	open	[]	0	[]	0	1		4/2/2022 17:03	FALSE
3	2113	ninja -t clean can del…	https://github.com/nin…	open	[Label(name="bug")]	1	[]	0	0		3/31/2022 15:25	FALSE
4	2112	Buildroot | ninja build …	https://github.com/nin…	open	[]	0	[]	0	0		3/31/2022 9:53	FALSE
5	2111	RSP file not found du…	https://github.com/nin…	open	[]	0	[]	0	0		3/30/2022 20:28	FALSE
6	2105	Ninja and PDB file size	https://github.com/nin…	closed	[]	0	[]	0	1	3/17/2022 7:20	3/16/2022 17:44	FALSE
7	2100	vdf_proofs_ecaf65ad…	https://github.com/nin…	closed	[]	0	[]	0	0	3/10/2022 6:43	3/10/2022 5:11	FALSE
8	2099	[![FOSSA Status](http…	https://github.com/nin…	closed	[]	0	[]	0	0	3/10/2022 6:43	3/10/2022 5:10	FALSE
9	2098	Error 6593	https://github.com/nin…	closed	[]	0	[]	0	1	3/10/2022 6:42	3/9/2022 16:58	FALSE
10	2097	File 'GNUmakefile' do…	https://github.com/nin…	closed	[]	0	[]	0	2	3/8/2022 10:09	3/8/2022 9:31	FALSE
11	2096	host-ninja 1.10.2 buil…	https://github.com/nin…	closed	[]	0	[]	0	3	3/7/2022 16:06	3/7/2022 15:40	FALSE
12	2095	Is there an upcoming…	https://github.com/nin…	closed	[]	0	[]	0	1	2/28/2022 12…	2/28/2022 12:32	FALSE
13	2091	ninja_test fails with q…	https://github.com/nin…	open	[]	0	[]	0	1		2/19/2022 20:54	FALSE
14	2090	Ninja doesn't seem t…	https://github.com/nin…	closed	[]	0	[]	0	0	2/19/2022 5:35	2/19/2022 5:29	FALSE
15	2084	dry-run builds are ~1…	https://github.com/nin…	closed	[]	0	[]	0	4	2/17/2022 22…	2/16/2022 1:19	FALSE
16	2083	Add fail fast mode tha…	https://github.com/nin…	open	[Label(name="feature"), Label(name="frontend")]	2	[]	0	4		2/15/2022 15:55	FALSE

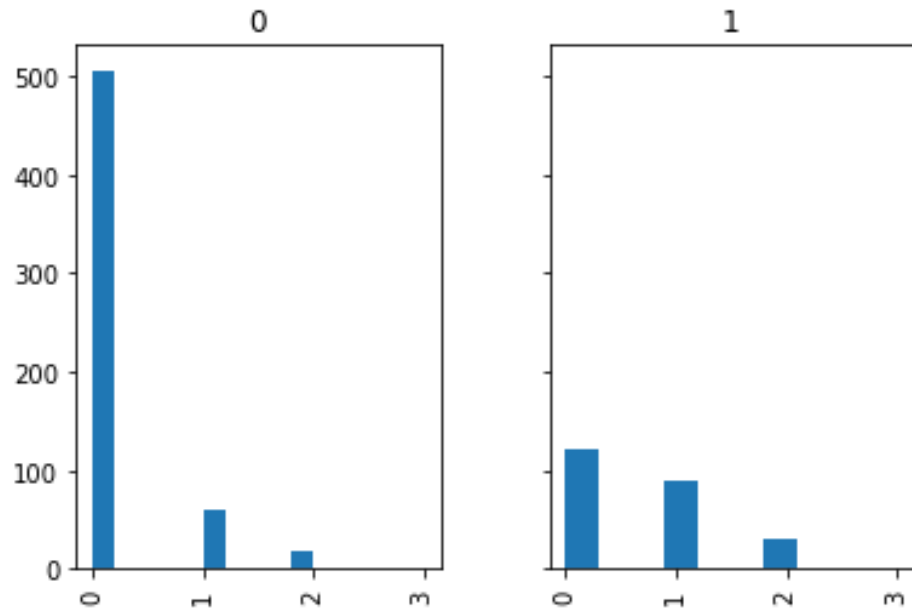Closed vs Open Issues

# of Comments vs. State
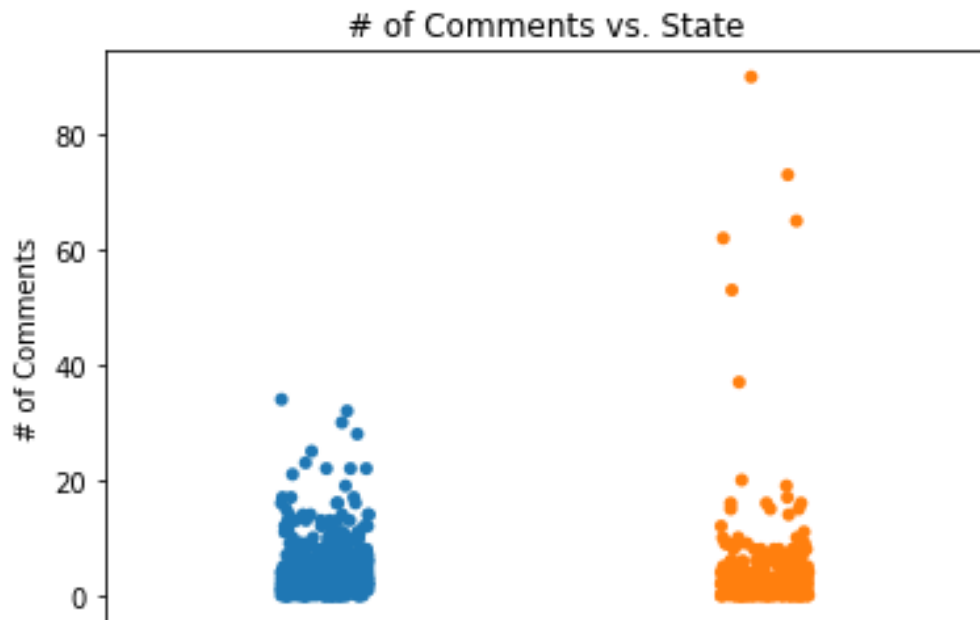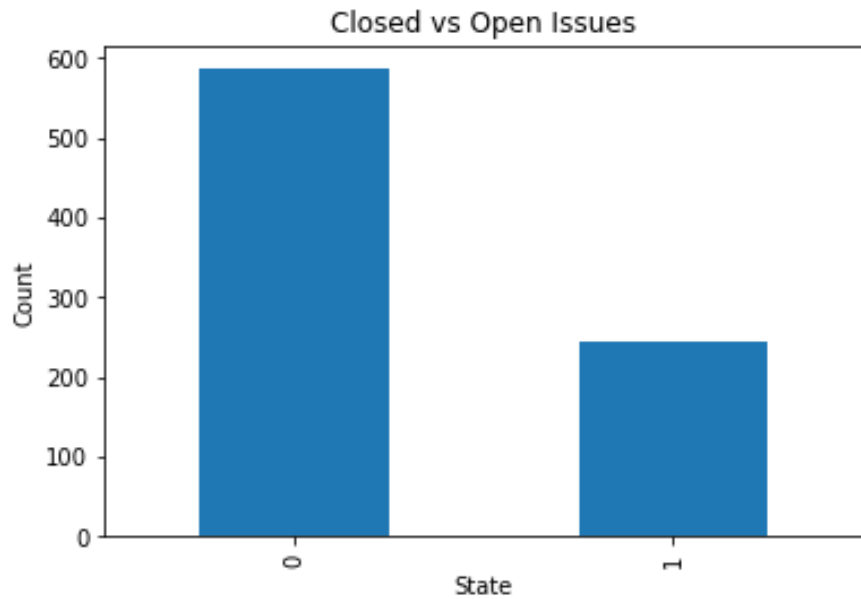
- State (0= Closed, 1= Open)
- There are more closed issues (>2X) than open issues
- Aside for those "few" outliers the # of comments doesn't seem significant whether an issue is closed or open
- Our repo did not have data for assignees or # of assignees for issues

CORRELATION HEATMAPS

- Correlation heatmap for all numeric attributes

- Heatmaps could indicate relationships of attributes to each other

- Heatmap for our data indicates a positive relationship between State (Open/Closed) and # of Labels

- Note: not all correlation may be meaningful

STATE VS # OF LABELS

- State (0= Closed, 1= Open)
- Overall are a lot of Closed and Open issues without labels
- More Open issues have more labelling

ISSUES STATE OVER PROJECT LIFETIME
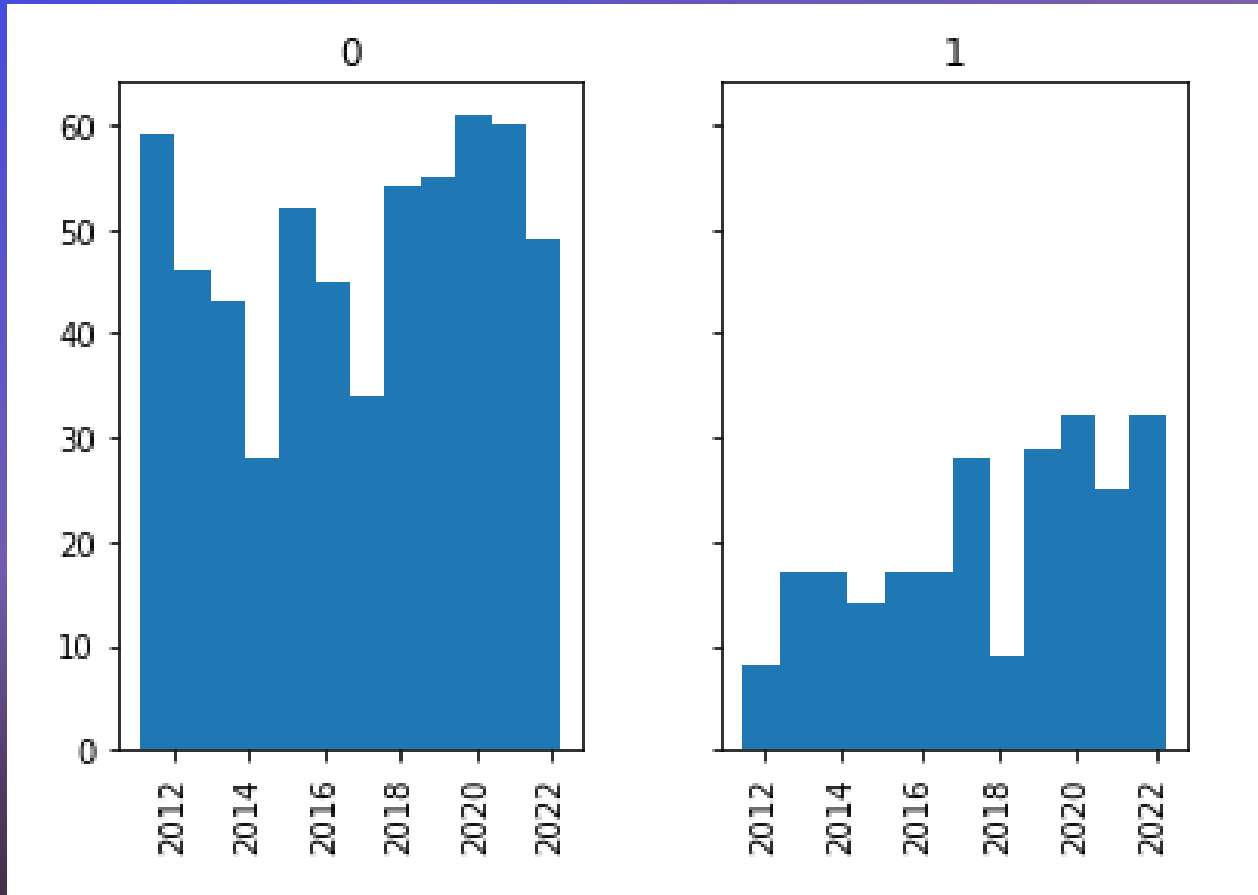
- Overall, there doesn't appear to be a trend for closed issues over the lifetime of the project

- There is a positive trend in number of open issues over the lifetime of the project

COMMITS

Table – slice of commits data

|	Date	URL
0	2022-03-22 10:39:03	https://api.github.com/repos/ninja-build/ninja...
1	2022-03-21 13:44:26	https://api.github.com/repos/ninja-build/ninja...
2	2022-03-20 10:22:00	https://api.github.com/repos/ninja-build/ninja...
3	2022-03-11 15:22:13	https://api.github.com/repos/ninja-build/ninja...
4	2022-03-11 15:22:13	https://api.github.com/repos/ninja-build/ninja...

Commits Over Most Recent Year

COMMITS

Commits Over First Year
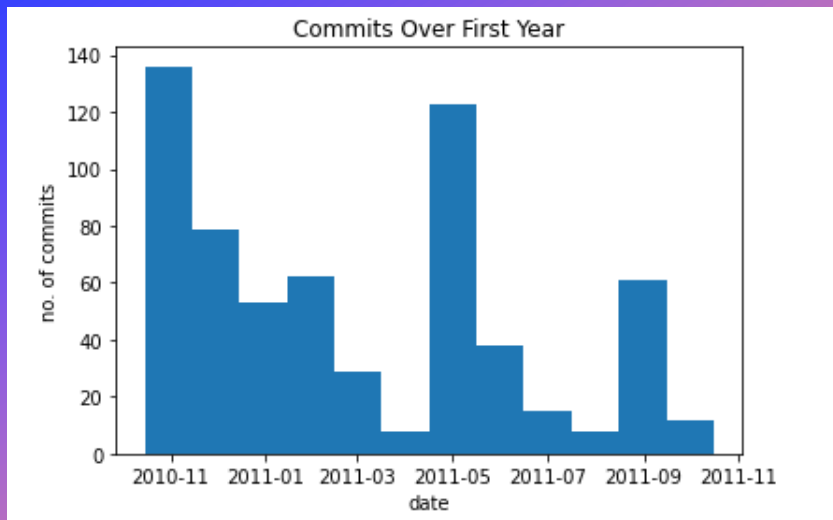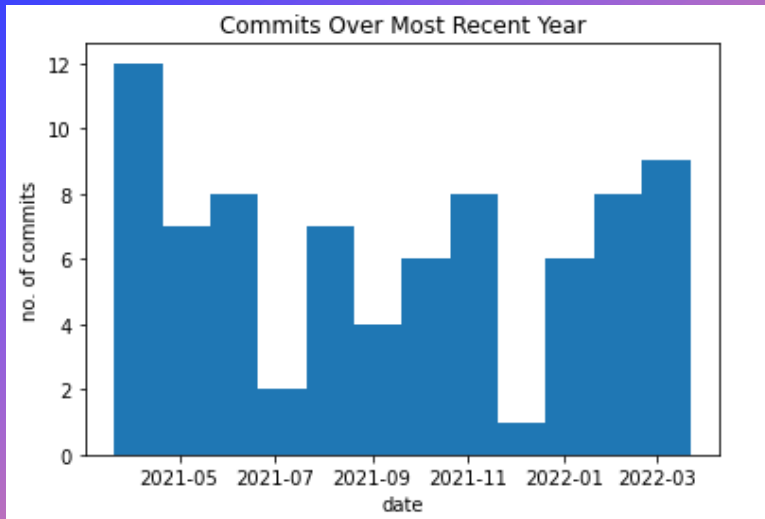
- YTD – No obvious pattern
- FIRST YEAR – Repeated downward trend with spikes
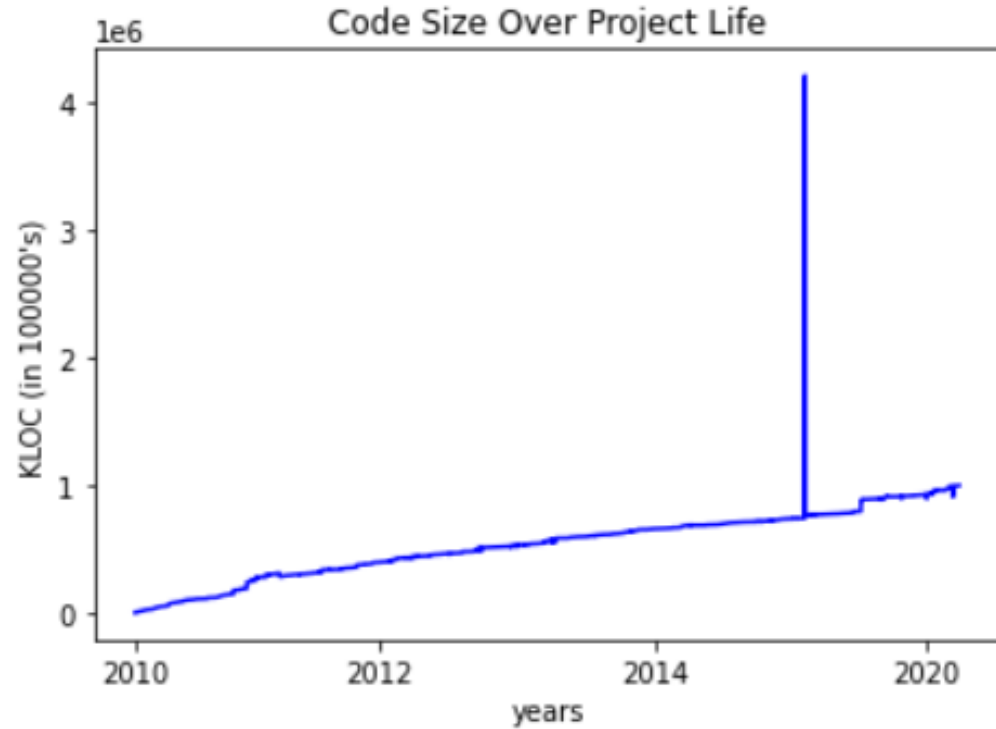
COMMITS

Commits Over Project Life

- DOWNWARD TREND OF NUMBER OF COMMITS AS THE PROJECT AGED
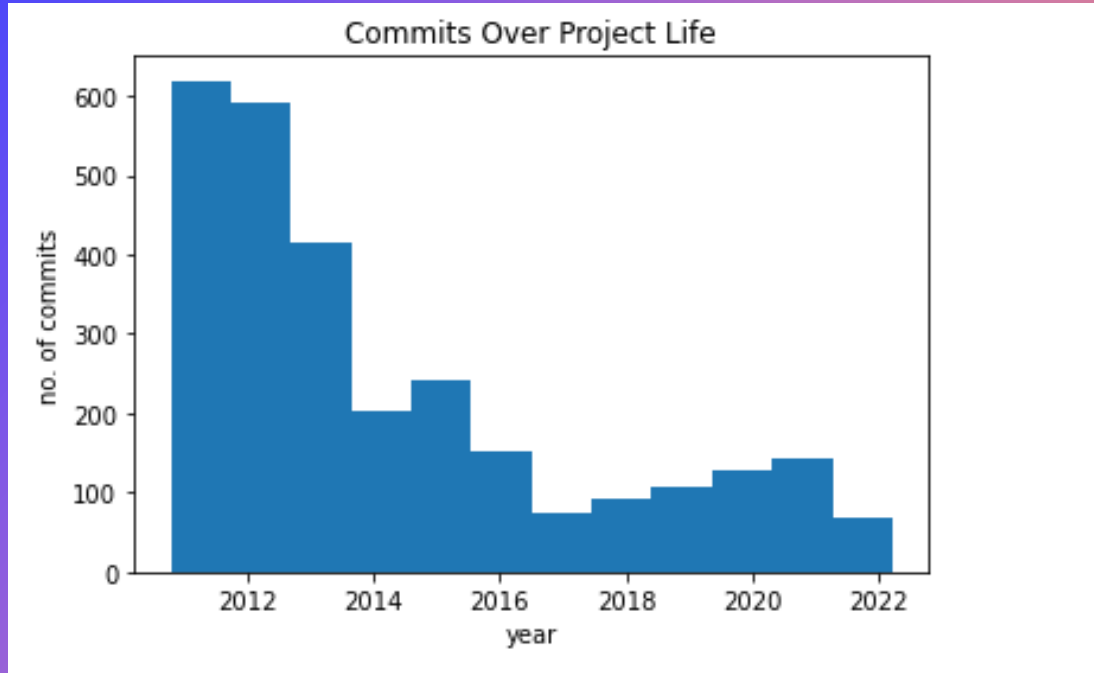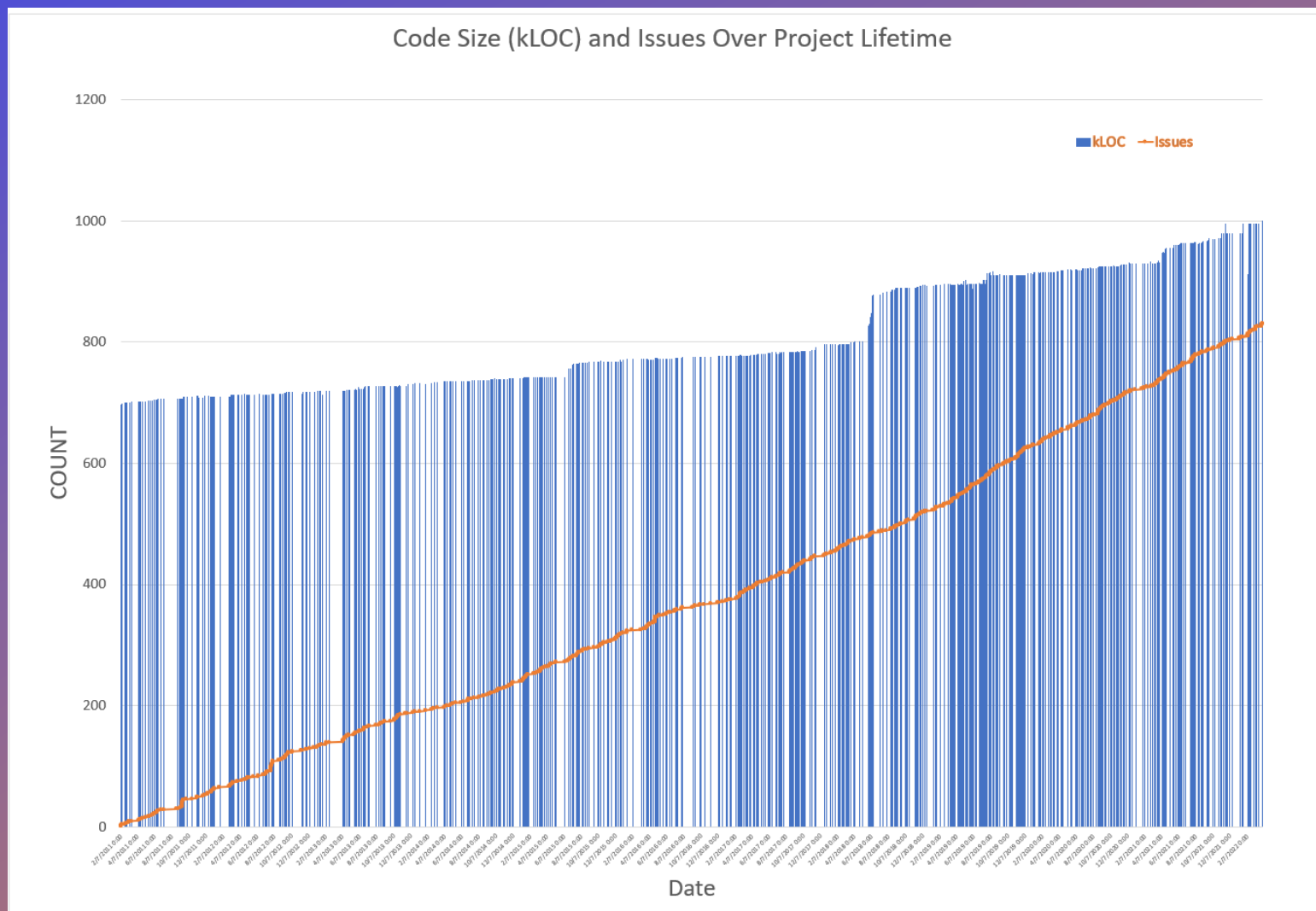
CODE SIZE

CODE SIZE

- Upward trend in code size as commits are published
- Code Size anomalies:
    - 2017-04-12 4215484 lines
    - 2017-04-12 4215477 lines

- Fairly monotonic
- Domain of date/time
- Codomain of KLOC

ISSUE DENSITY

Issue Density

Code Size (kLOC) and Issues Over Project Lifetime

Issue Density

- Upward trend in issue density over time

- Issue density seems to be monotonic

- Code size does have a steady increase with step-ups, likely due to bigger changes

- As the project aged issue density should increase, which seems counter to what we learn

- There should be a peak of issue density somewhere in the middle with a downward trend as the project age

QUESTIONS?