

Humanized Cryptoassets Cryptocurrency Quality Based on Source Code

Yashashree K., Océane B.
Advisors: Neha N., Rhys L.

INTRO

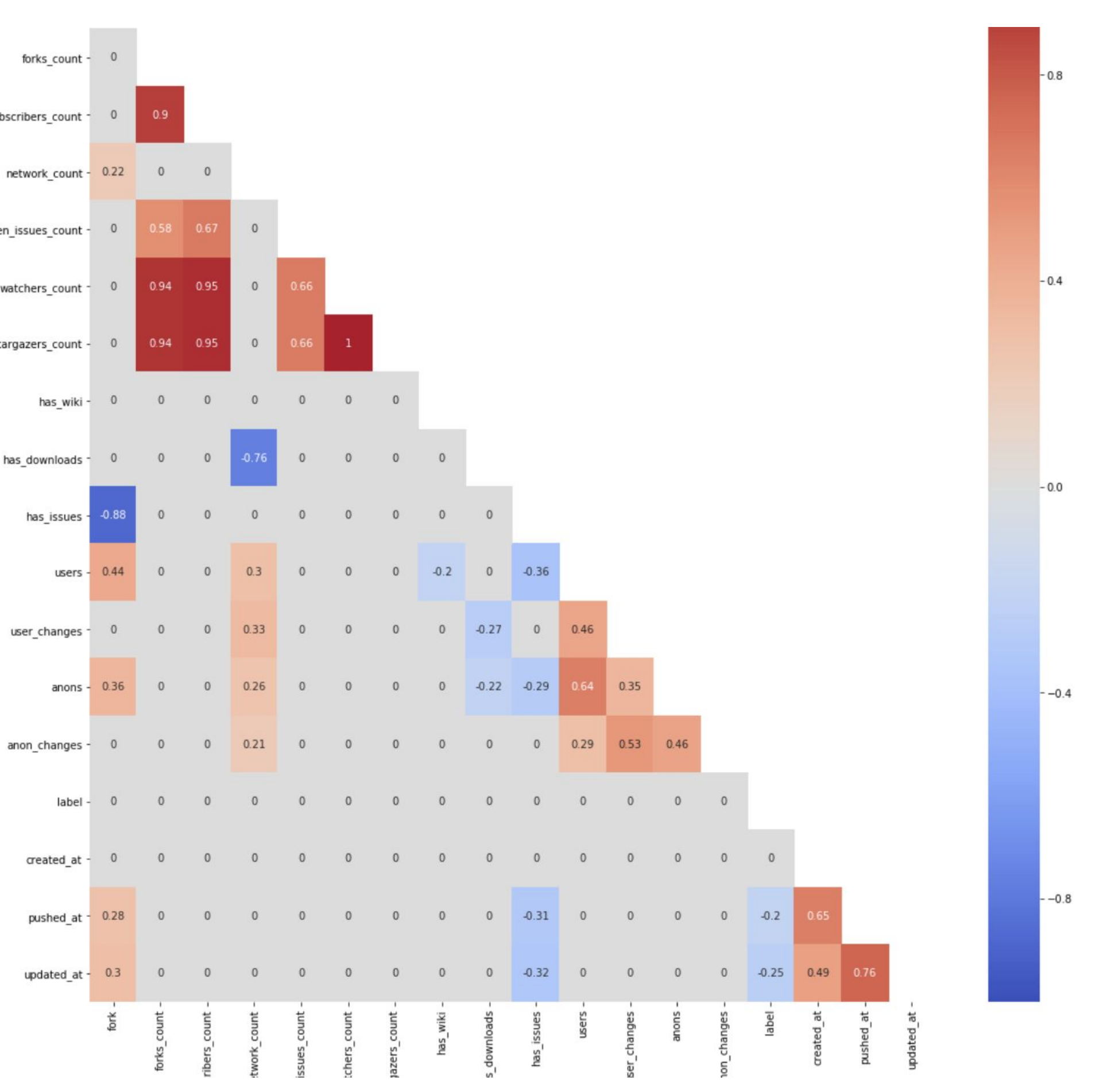
- Quantitative analysis of Github metrics collected from source code can provide rich insights into the health of the token's foundation and technical community dynamics

METHOD

- Gathered 12,141 repositories that correspond to 1011 cryptocurrencies¹
- Collected Github metadata²
- Created hypothesis metrics and build insights

DEVELOPER DYNAMICS HYPOTHESES

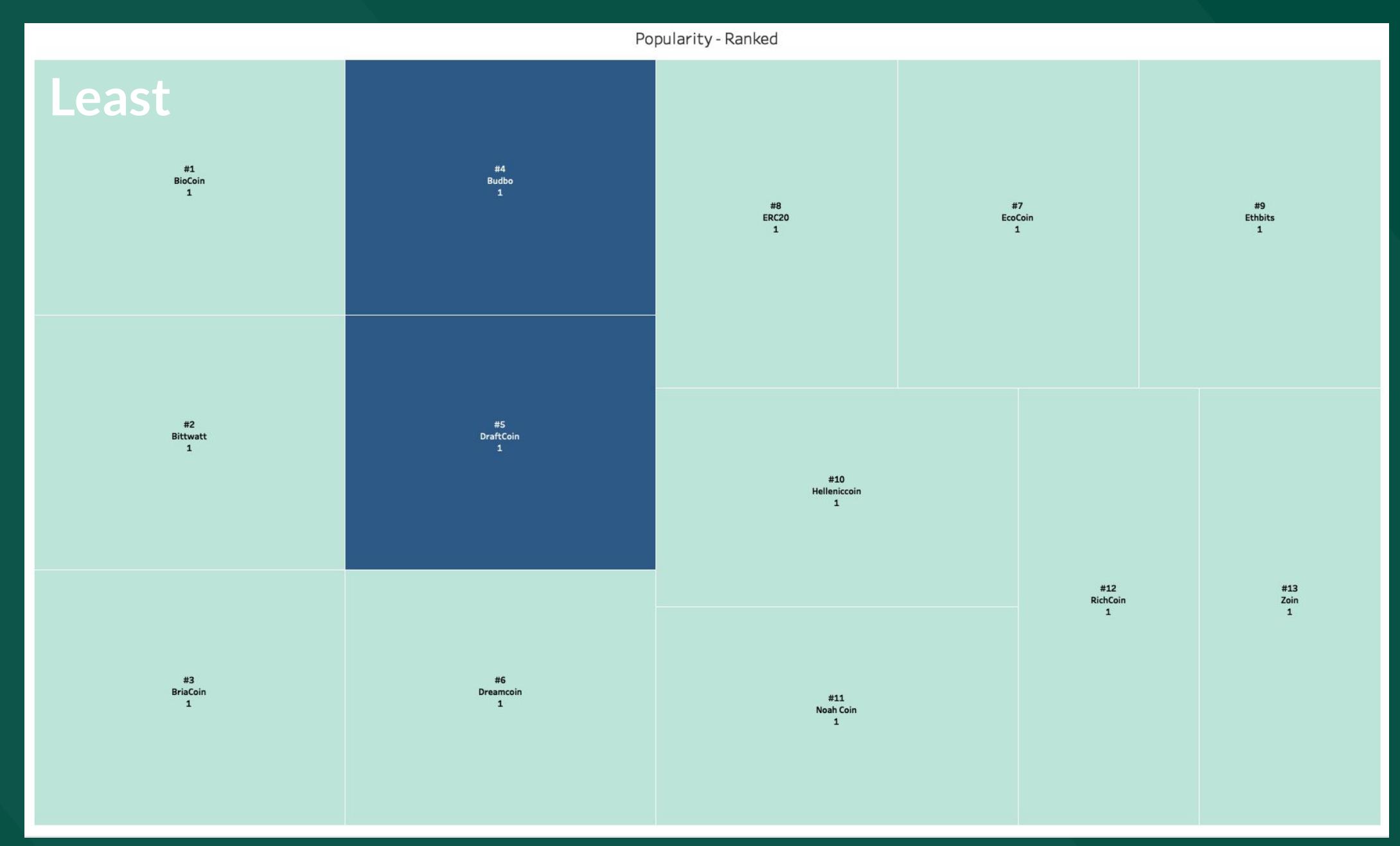
- Popularity**
 - Number of forks
 - Number of subscribers and stargazers
- Engagement**
 - Repository creation time and last pushed times
 - Monthly commit frequency
 - Number of issues
- Authenticity**
 - Disabled Issue creation
 - Anonymous user contributions



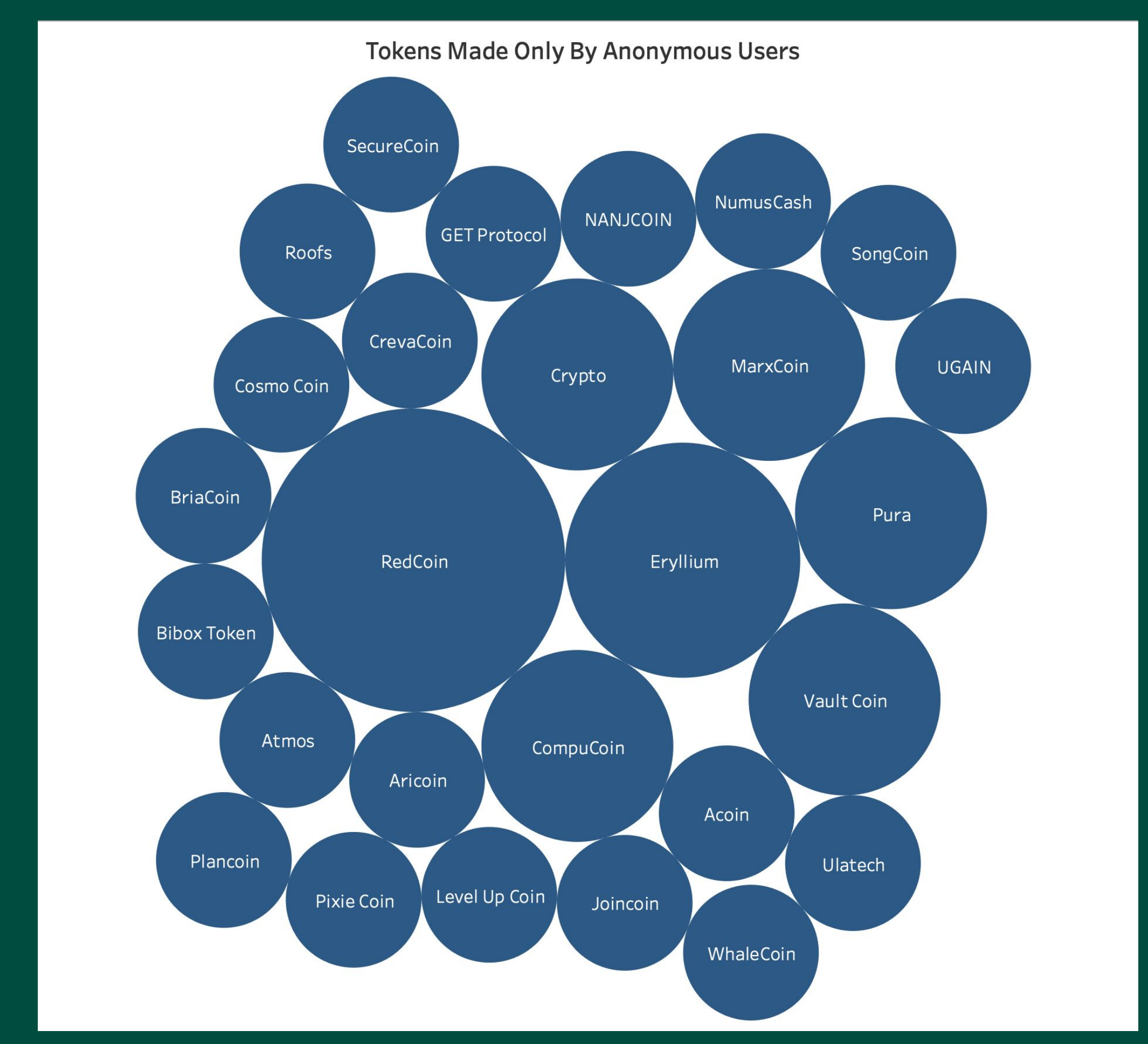
Correlation Matrix between metadata features

How do developer dynamics influence the quality of a cryptocurrency?

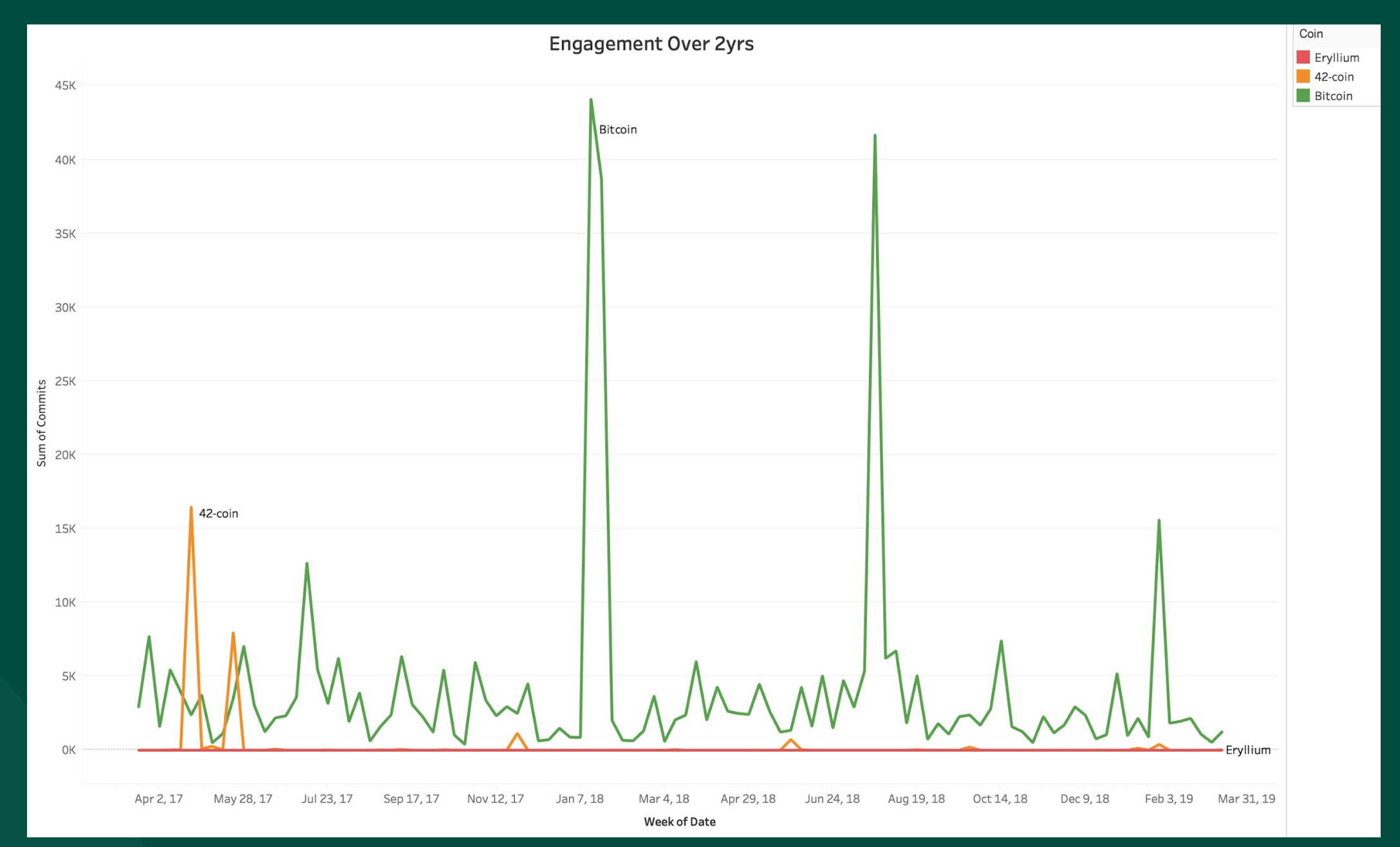
Popularity



Authenticity



Engagement



INSIGHTS

- Most Popular Cryptocurrencies**
 - Ethereum
 - Bitcoin
 - EOS
- Disabling Issue Creation**
 - Issues creation is enabled by default
 - But 2% of Repositories (4% of Tokens) have explicitly disabled this
- Most Active Repositories**
 - Ethereum
 - Bitcoin
- Anonymous User Contributions**
 - Users can register for Github accounts or contribute anonymously using an email account (unverified) as their identity
 - Coins comprised entirely of Anonymous users: Eryllium, PixieCoin, SecureCoin

DISCUSSION

- Technical innovation is the implied promise of a new coin when a whitepaper is released or ICO is launched
- Through this work, we hope to shed light on this narrative from a code-based perspective

NEXT STEPS

- Improve the baseline accuracy (70%) of the SVM (Support Vector Machine) based model using deep neural networks
- Compare and contrast with unsupervised learning methods to surface key attributes that impact cryptocurrency quality

KEY DEFINITIONS⁰

- Repository:** A repository contains source code files (including documentation) and stores each file's revision history.
- Fork:** A fork is a personal copy of another user's repository that lives on your account
- Issue:** Issues are suggested improvements, tasks or questions related to the repository.



visit PubPub to read our 'Token Funnies'

⁰ <https://help.github.com/en/articles/github-glossary>
¹ <https://github.com/manganese/alteramentum-repo-data/blob/master/repos.csv>
² url, forks_count, open_issues_count, created_at, updated_at, pushed_at, has_issues, is_fork, contributors_username, contribution count, contributor_type, commit_frequency