NPM Ecosystem Analysis

Scott Henry

What are software ecosystems?

- Software ecosystems can be viewed as socio-technical networks consisting of technical components (software packages) and social components (communities of developers) that maintain the technical components.

- Ecosystems evolve over time through socio-technical changes that may greatly impact the ecosystem's sustainability. Social changes like developer turn-over may lead to technical degradation.

Constantinou, Eleni; Mens, Tom. (2017)

Relevant Literature

- On the Impact of Micro-Packages: An Empirical Study of the npm JavaScript Ecosystem
- Measuring Software Library Stability Through Historical Version Analysis
- A Study of Social Interactions in Open Source Component Use
- Visualizing the Evolution of Systems and their Library Dependencies
- Analyzing the evolution of social aspects of open source software ecosystems

Motivation

- How do you decide which package to use?
- What is a stable package?

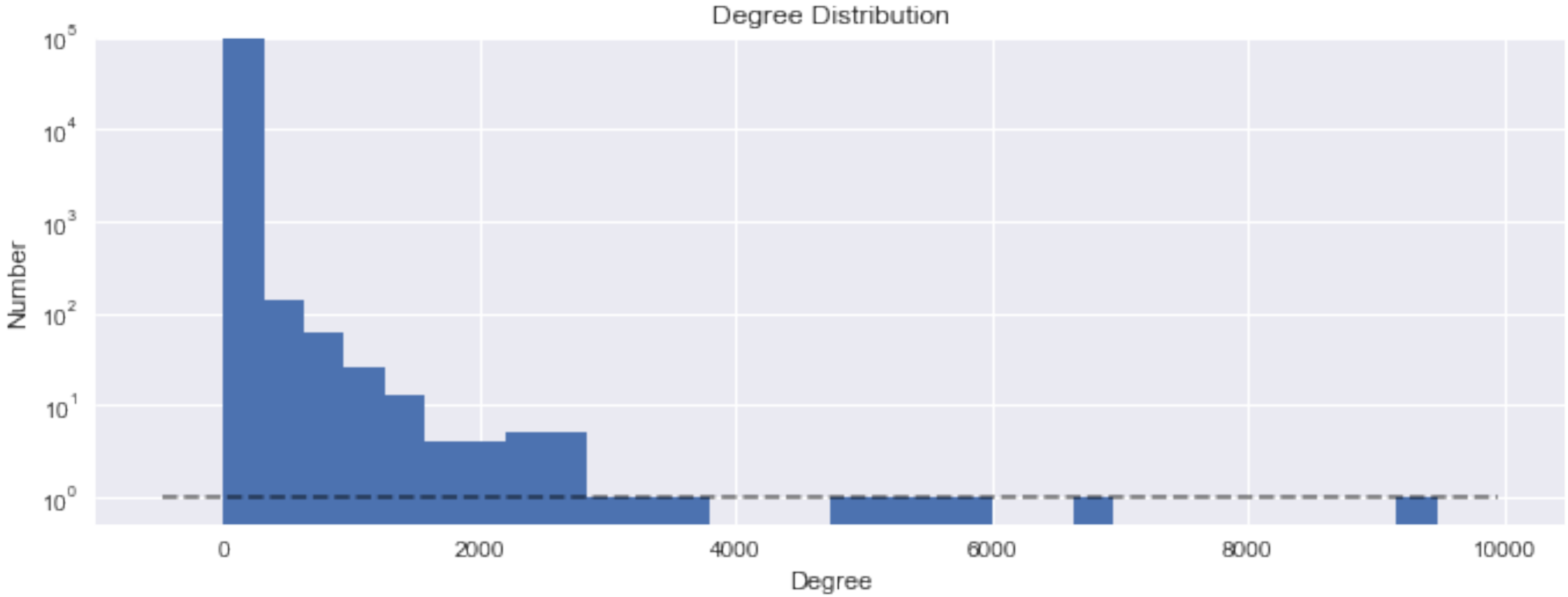
Focus: How long does it take to update a package?

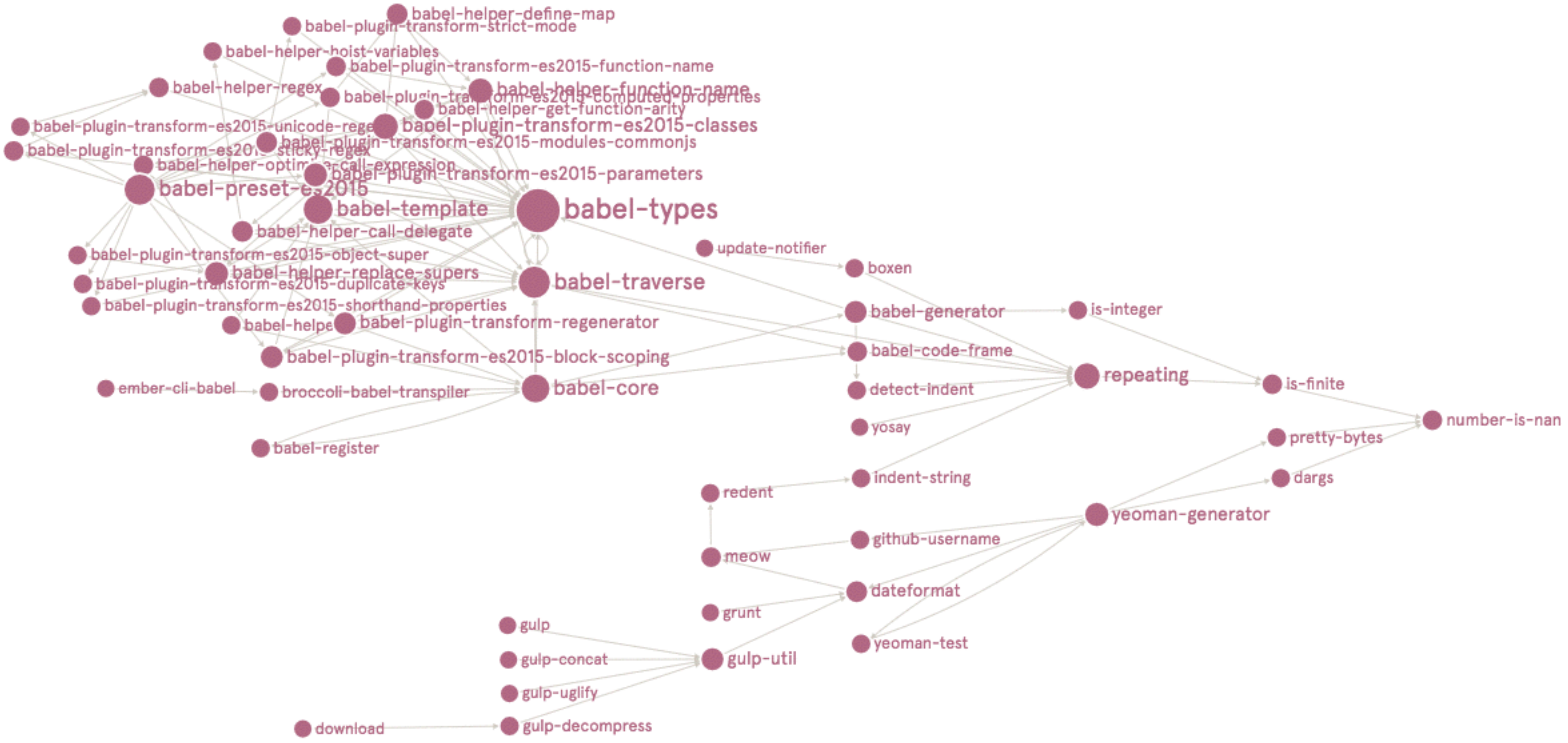
NPM Ecosystem

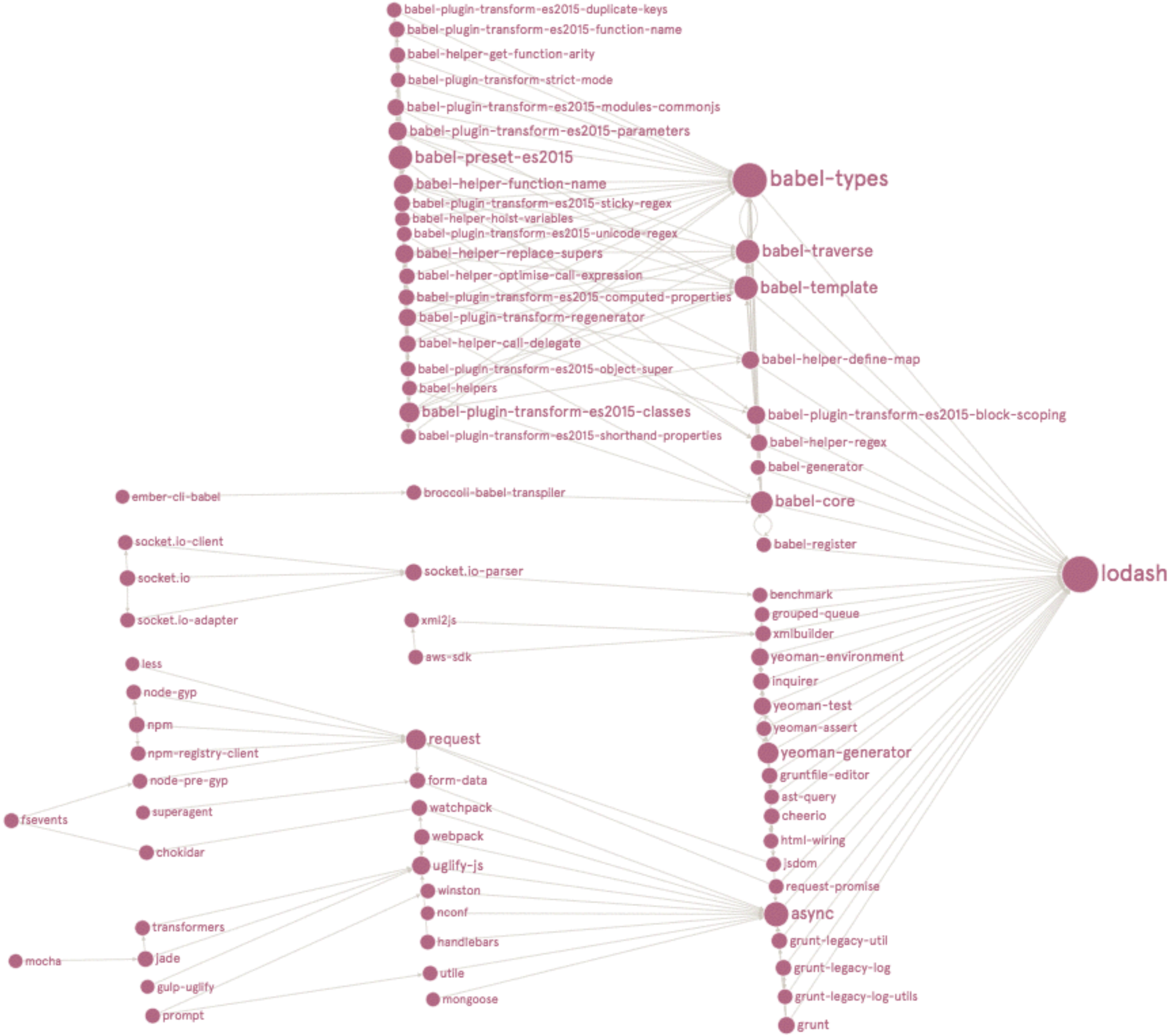
npm is a package manager for the JavaScript programming language. It is the default package manager for the JavaScript runtime environment Node.js.

- 1/2 Million Packages
- Majority of packages have under 50 dependencies.
- Average package has 1 version.

Degree Distribution




babel-helper-define-map
babel-plugin-transform-strict-mode
babel-helper-hoist-variables
babel-plugin-transform-es2015-function-name
babel-helper-regex
babel-helper-function-name
babel-plugin-transform-es2015-computed-properties
babel-helper-get-function-arity
babel-plugin-transform-es2015-unicode-regex
babel-plugin-transform-es2015-classes
babel-plugin-transform-es2015-sticky-regex
babel-plugin-transform-es2015-modules-commonjs
babel-helper-optimise-call-expression
babel-plugin-transform-es2015-parameters
babel-preset-es2015
babel-template
babel-types
babel-helper-call-delegate
update-notifier
boxen
babel-plugin-transform-es2015-object-super
babel-helper-replace-supers
babel-traverse
babel-plugin-transform-es2015-duplicate-keys
babel-plugin-transform-es2015-shorthand-properties
babel-generator
is-integer
babel-helpe
babel-plugin-transform-regenerator
babel-code-frame
babel-plugin-transform-es2015-block-scoping
repeating
is-finite
ember-cli-babel
broccoli-babel-transpiler
babel-core
detect-indent
number-is-nan
yosay
babel-register
pretty-bytes
indent-string
dargs
redent
yeoman-generator
github-username
meow
dateformat
grunt
gulp
yeoman-test
gulp-concat
gulp-util
gulp-uglify
download
gulp-decompress


babel-plugin-transform-es2015-duplicate-keys
babel-plugin-transform-es2015-function-name
babel-helper-get-function-arity
babel-plugin-transform-strict-mode
babel-plugin-transform-es2015-modules-commonjs
babel-plugin-transform-es2015-parameters
babel-preset-es2015
babel-helper-function-name
babel-plugin-transform-es2015-sticky-regex
babel-helper-hoist-variables
babel-plugin-transform-es2015-unicode-regex
babel-helper-replace-supers
babel-helper-optimise-call-expression
babel-plugin-transform-es2015-computed-properties
babel-plugin-transform-regenerator
babel-helper-call-delegate
babel-plugin-transform-es2015-object-super
babel-helpers
babel-plugin-transform-es2015-classes
babel-plugin-transform-es2015-shorthand-properties
babel-types
babel-traverse
babel-template
babel-helper-define-map
babel-plugin-transform-es2015-block-scoping
babel-helper-regex
babel-generator
babel-core
babel-register
lodash
ember-cli-babel
broccoli-babel-transpiler
socket.io-client
socket.io
socket.io-adapter
socket.io-parser
xml2js
aws-sdk
benchmark
grouped-queue
xmlbuilder
yeoman-environment
inquirer
yeoman-test
yeoman-assert
yeoman-generator
gruntfile-editor
ast-query
cheerio
html-wiring
jsdom
request-promise
async
grunt-legacy-util
grunt-legacy-log
grunt-legacy-log-utils
grunt
less
node-gyp
npm
npm-registry-client
node-pre-gyp
superagent
fsevents
chokidar
request
form-data
watchpack
webpack
uglify-js
winston
nconf
handlebars
utile
mongoose
transformers
mocha
jade
gulp-uglify
prompt

Methodology

1. Developed data pipeline from NPM registry.

2. Parsing JSON into data frame.

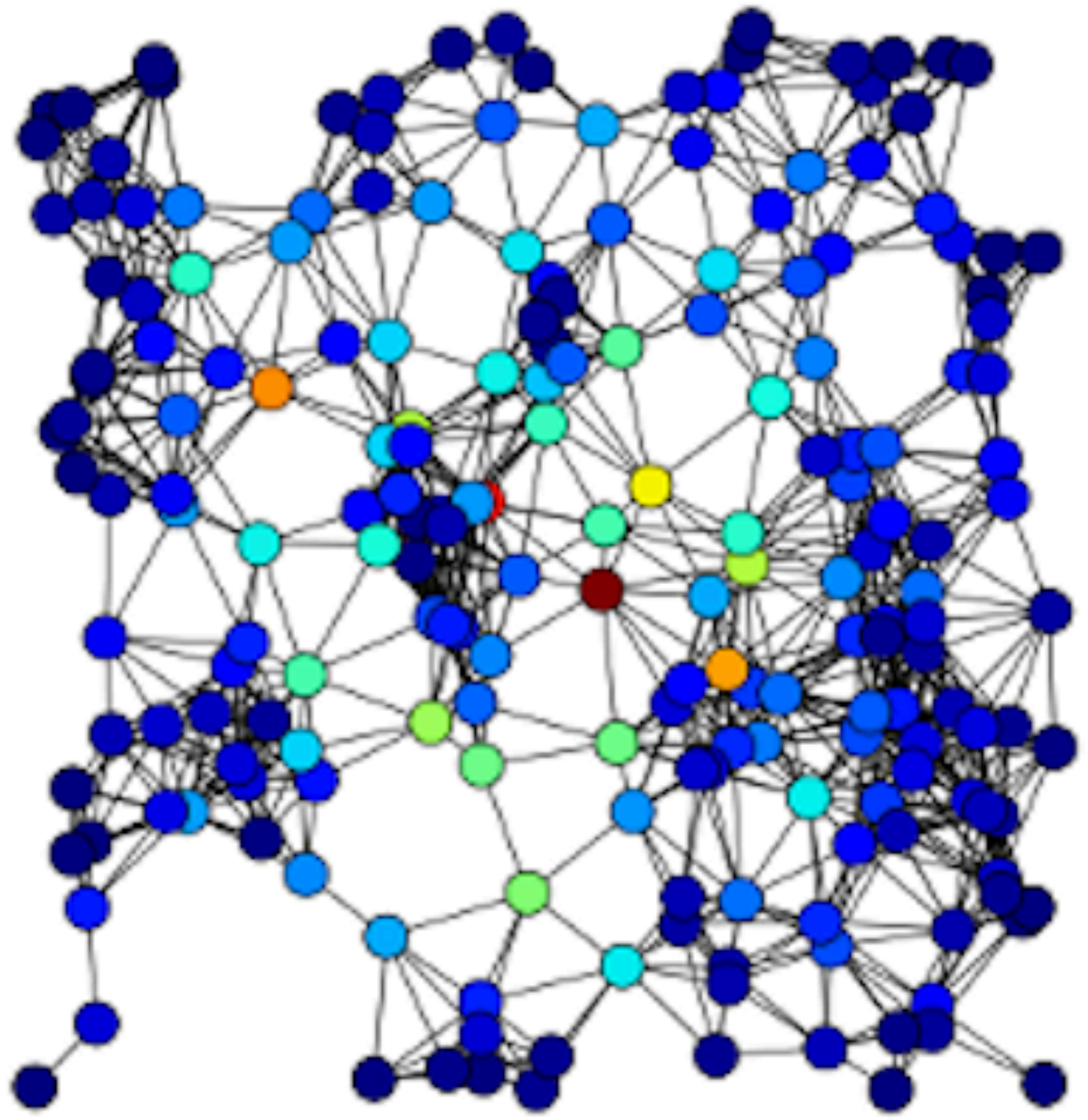
3. Assigned each version a specific version number.
   ^1.7.7
   *
   ~1.5.6

4. Created network graphs at various points in time to analyze centrality.

What is betweenness centrality?

Betweenness centrality is a measure of centrality in a graph based on shortest paths.

- For every pair of vertices in a connected graph, there exists at least one shortest path between the vertices such that either the number of edges that the path passes through (for unweighted graphs) or the sum of the weights of the edges (for weighted graphs) is minimized.
- The betweenness centrality for each vertex is the number of these shortest paths that pass through the vertex.



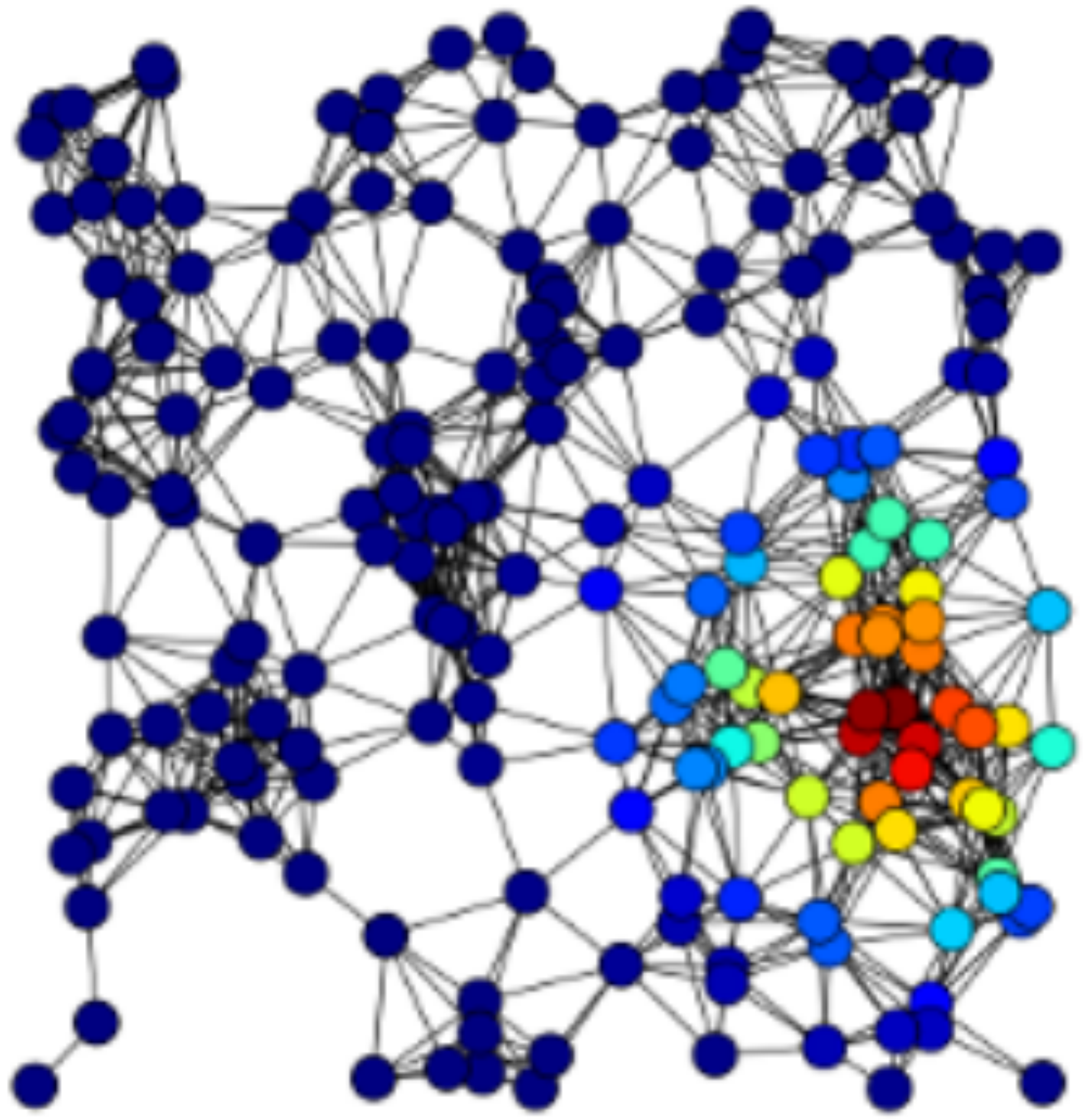
An undirected graph coloured based on the centrality of each vertex from greatest (red) to least (blue).

Piraveenan, Mahendra (2013).

What is eigenvector centrality?

Eigenvector centrality is a measure of the influence of a node in a network.

- It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes



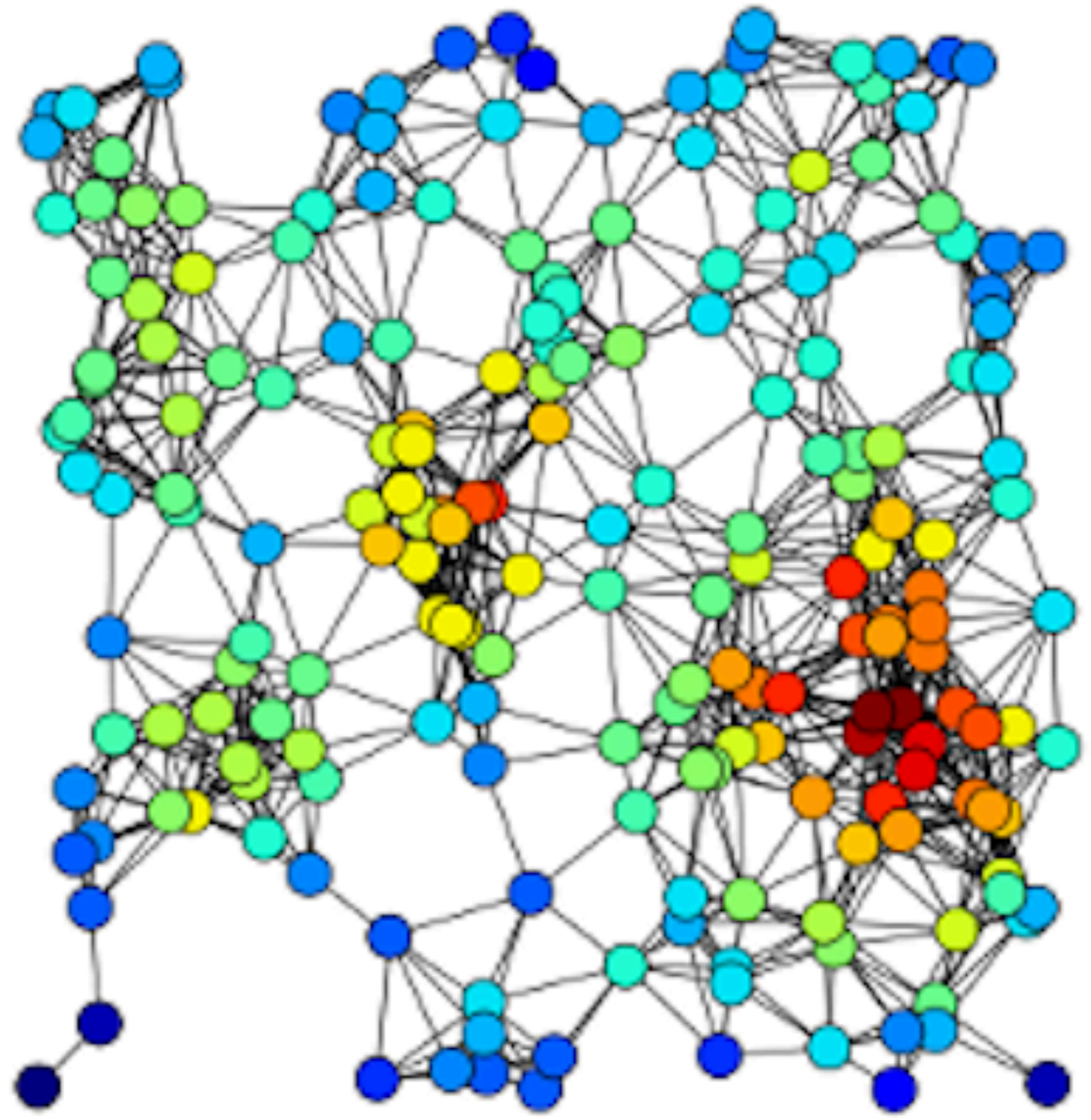
An undirected graph coloured based on the centrality of each vertex from greatest (red) to least (blue).

Fletcher, Jack McKay and Wennekers, Thomas (2017).

What is degree centrality?

Degree centrality, is defined as the number of links incident upon a node (i.e., the number of ties that a node has).

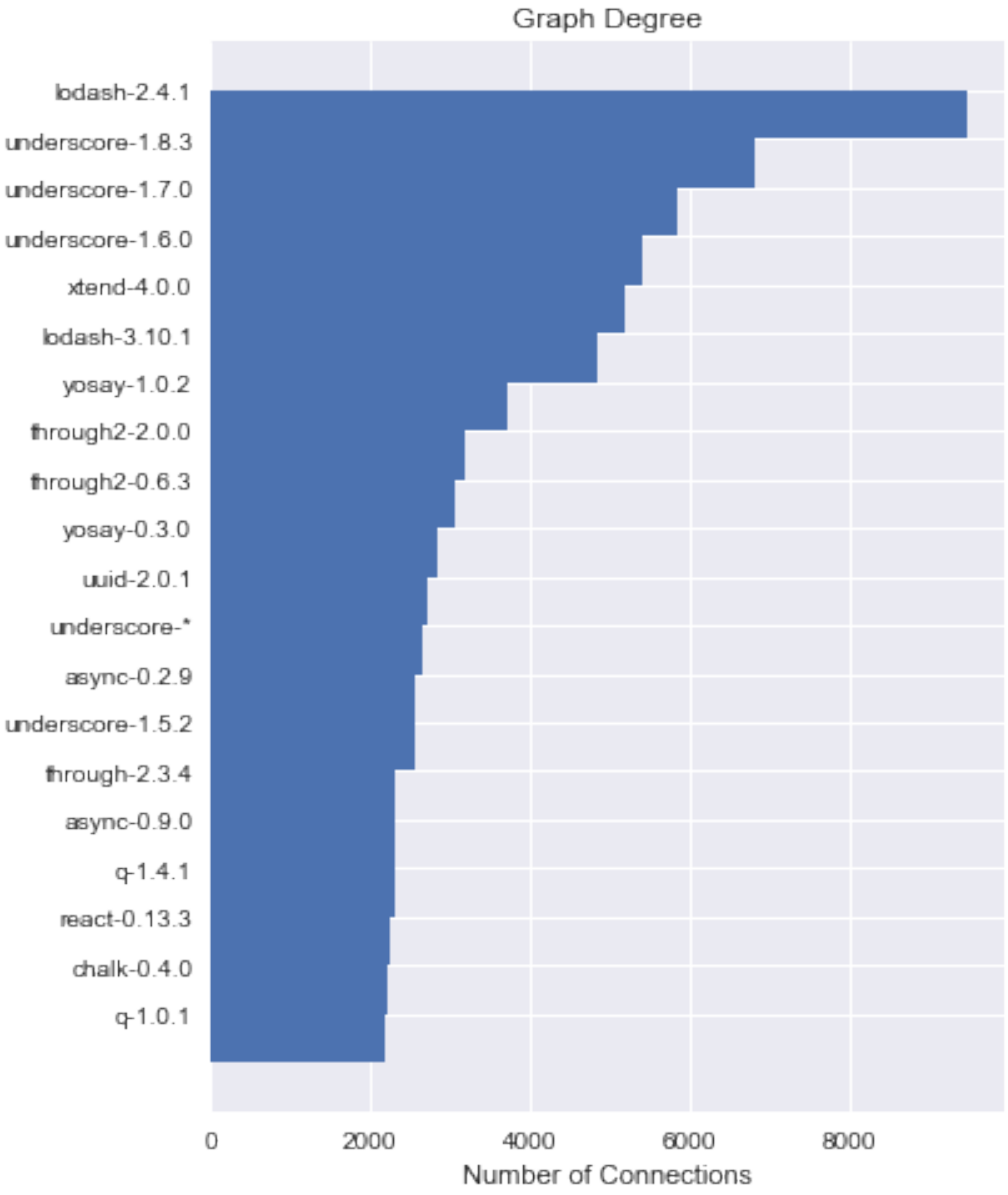
- It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes



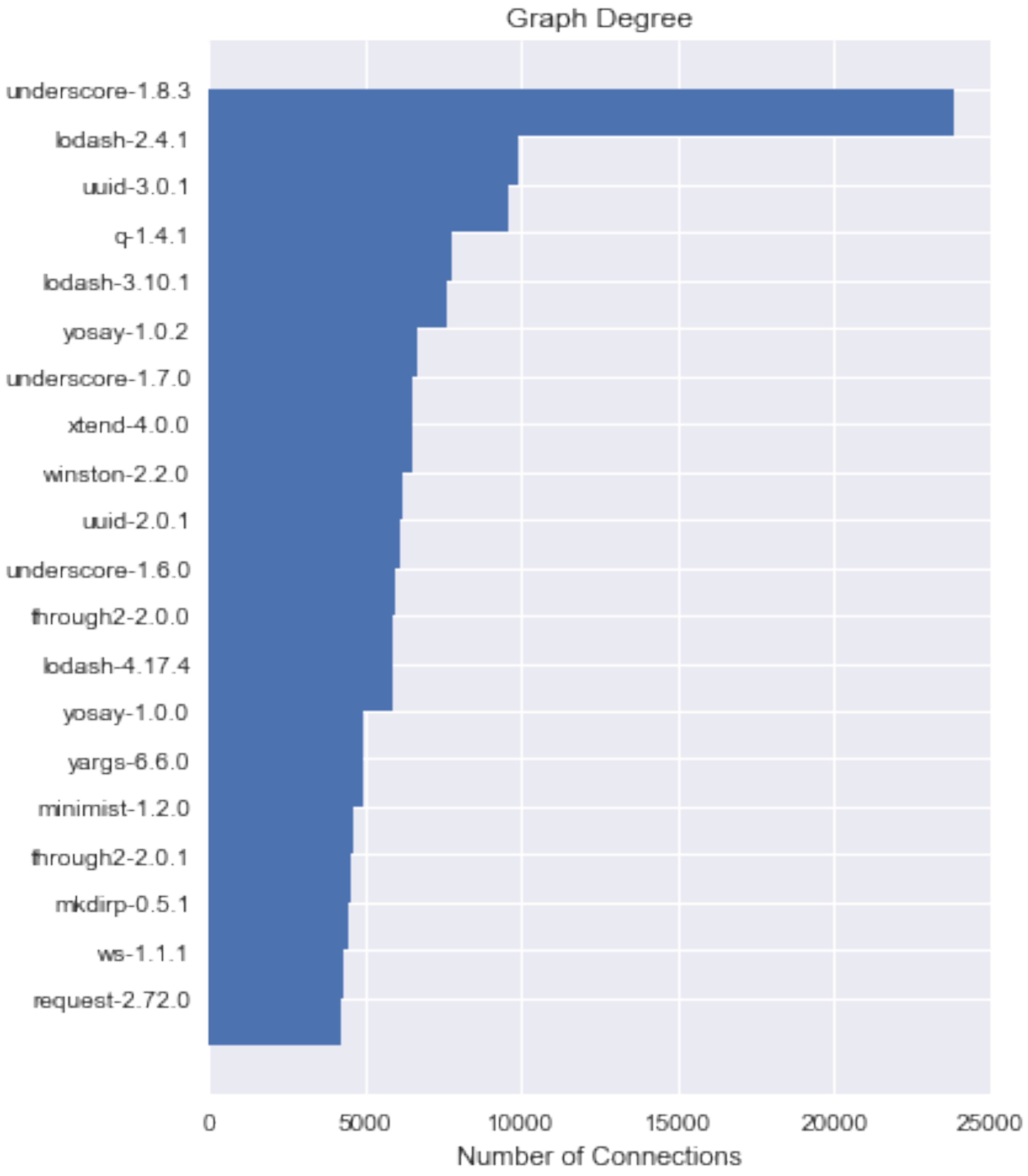
An undirected graph coloured based on the centrality of each vertex from greatest (red) to least (blue).

Freeman, Linton C. (1979)

Degree Centrality

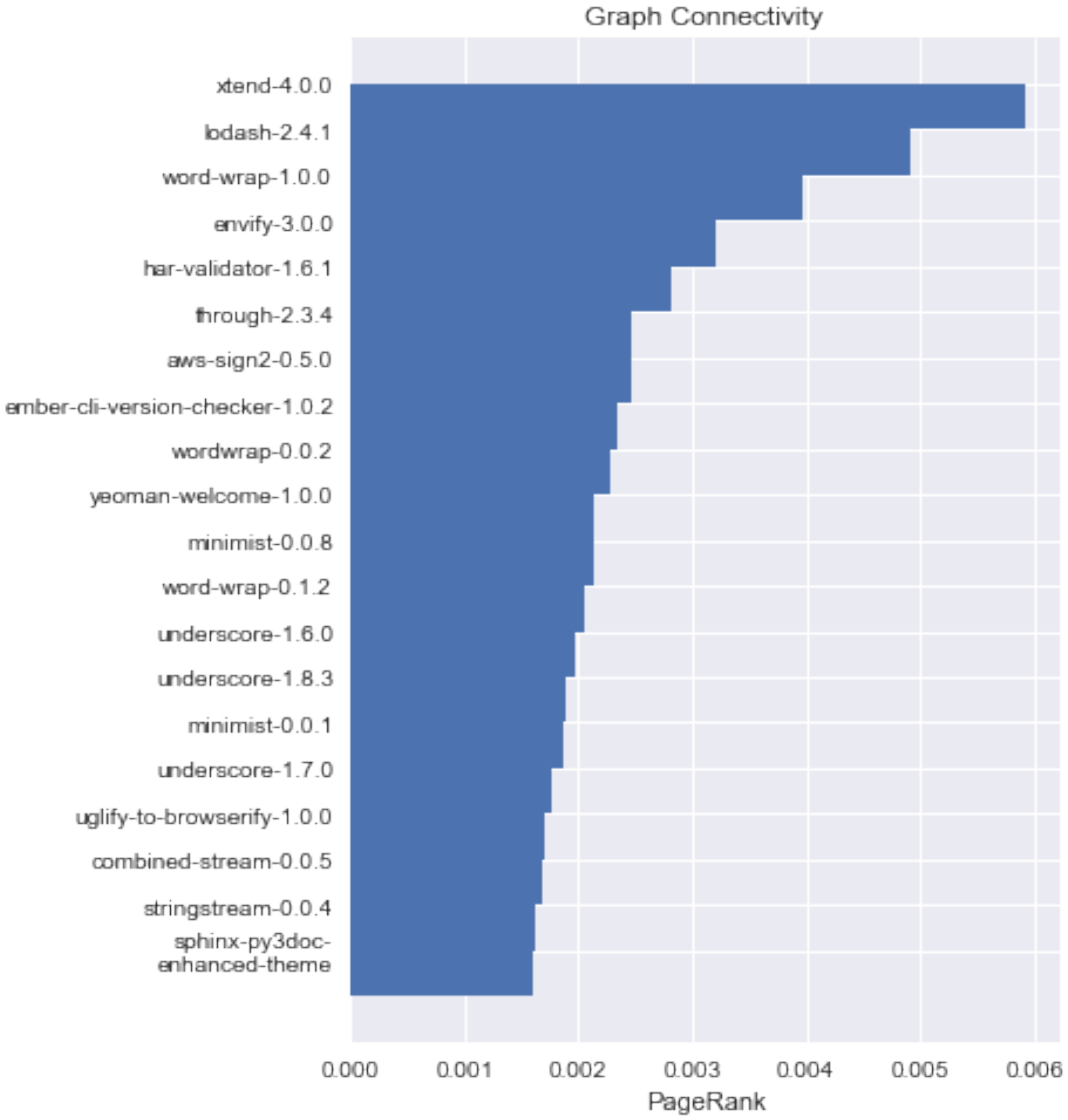


2016-01-01

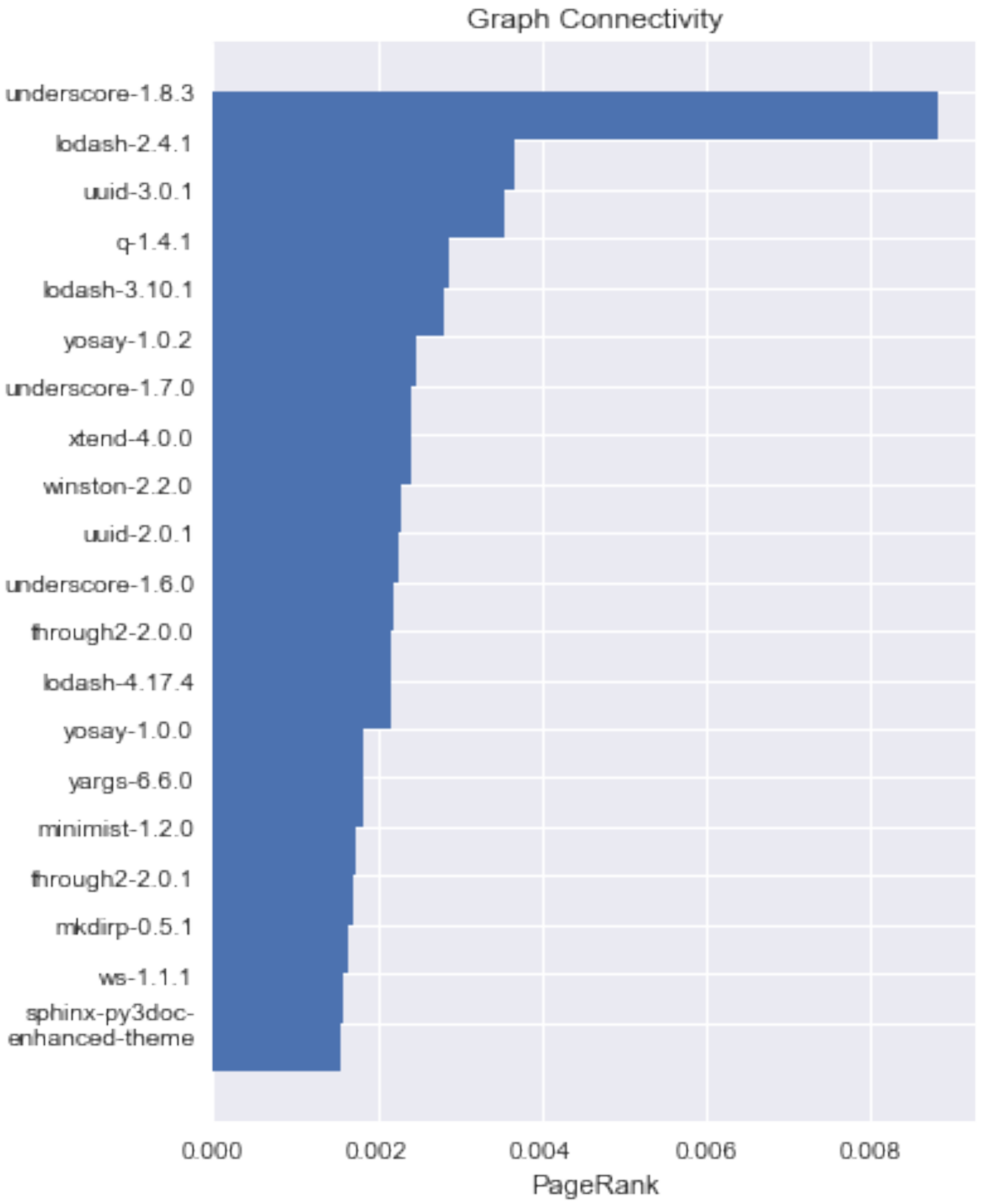


2017-06-01

Pagerank Centrality



2016-01-01

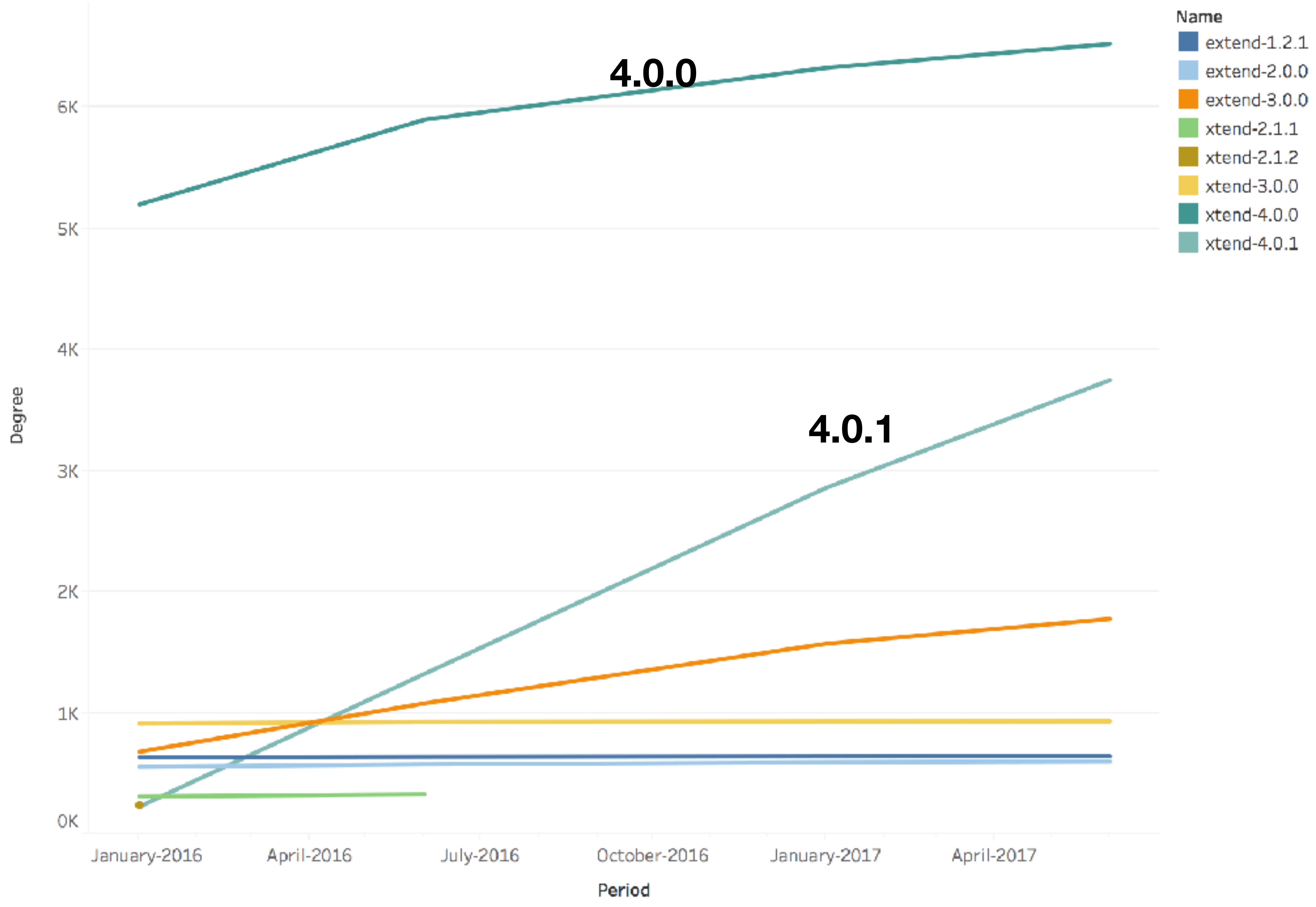


2017-06-01

Case Study: xtend

xtend is a basic utility library which allows you to extend an object by appending all of the properties from each object in a list.

- Growth of version 4.0.0 is slowing, 4.0.1 is increasing.

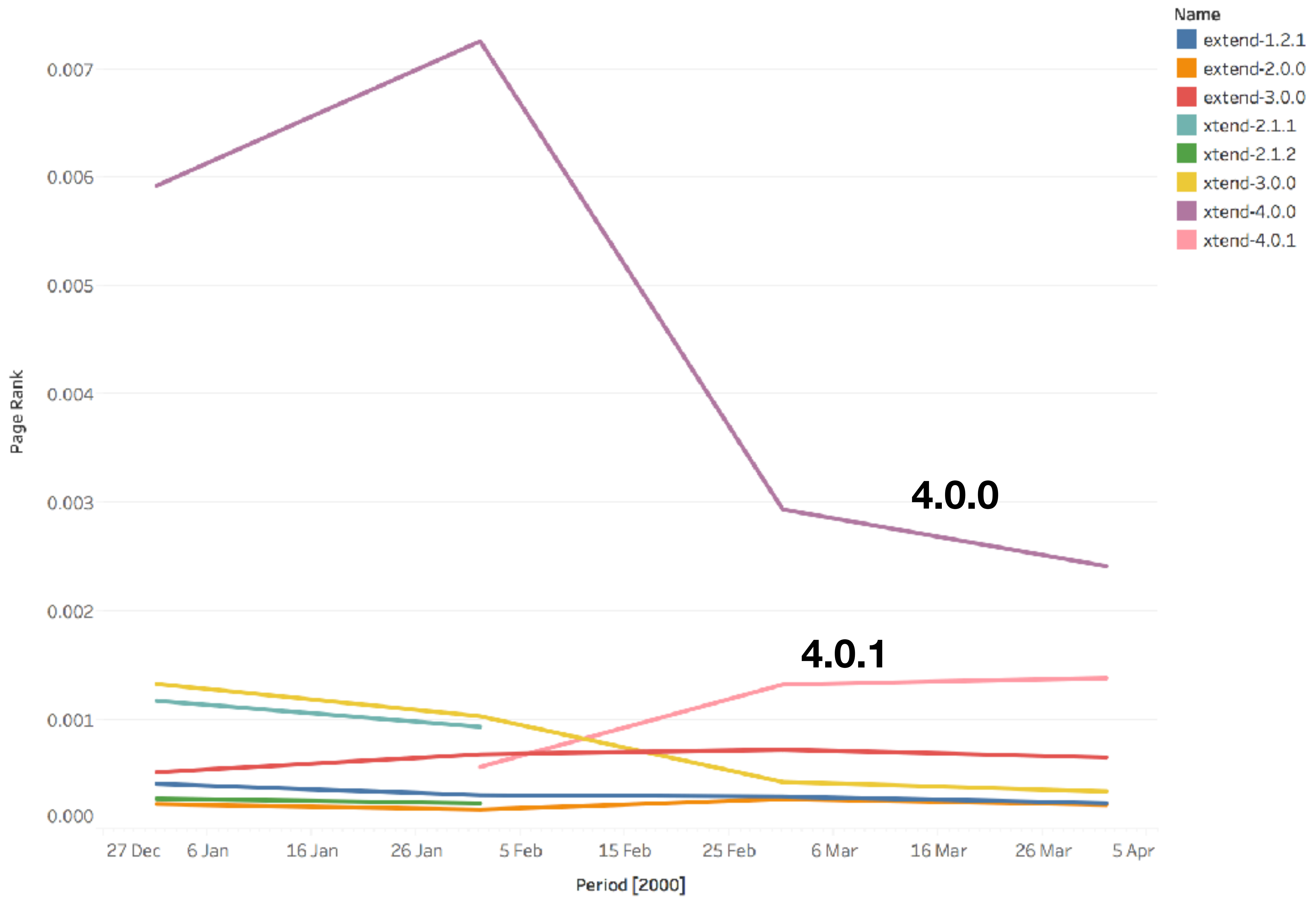


Degree Centrality

Case Study: xtend

xtend is a basic utility library which allows you to extend an object by appending all of the properties from each object in a list.

- Growth of version 4.0.0 is slowing, 4.0.1 is increasing.

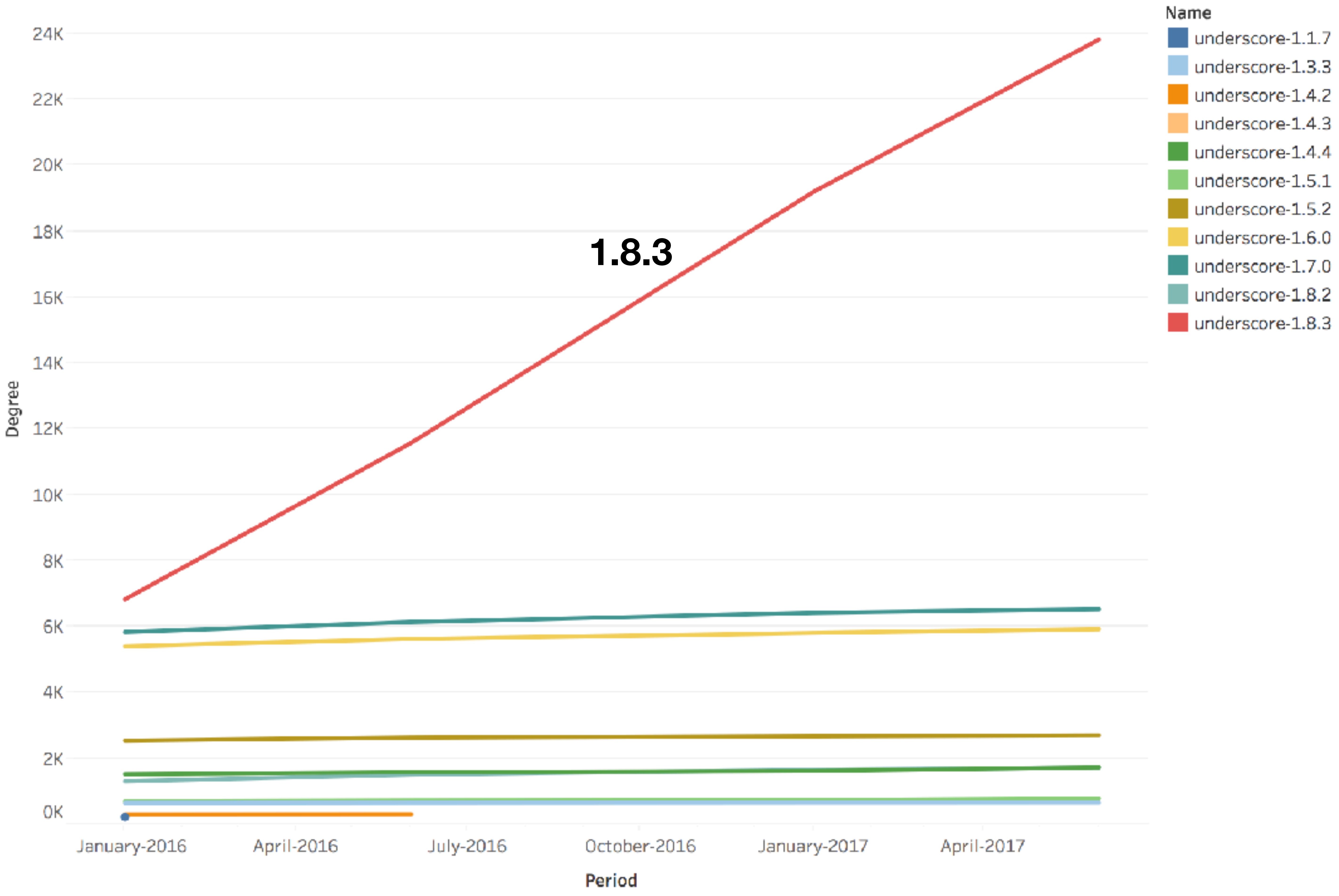


Page Rank

Case Study: underscore

Underscore.js is a utility-belt library for JavaScript that provides support for the usual functional suspects (each, map, reduce, filter...) without extending any core JavaScript objects.

- Significant growth of version 1.8.3.



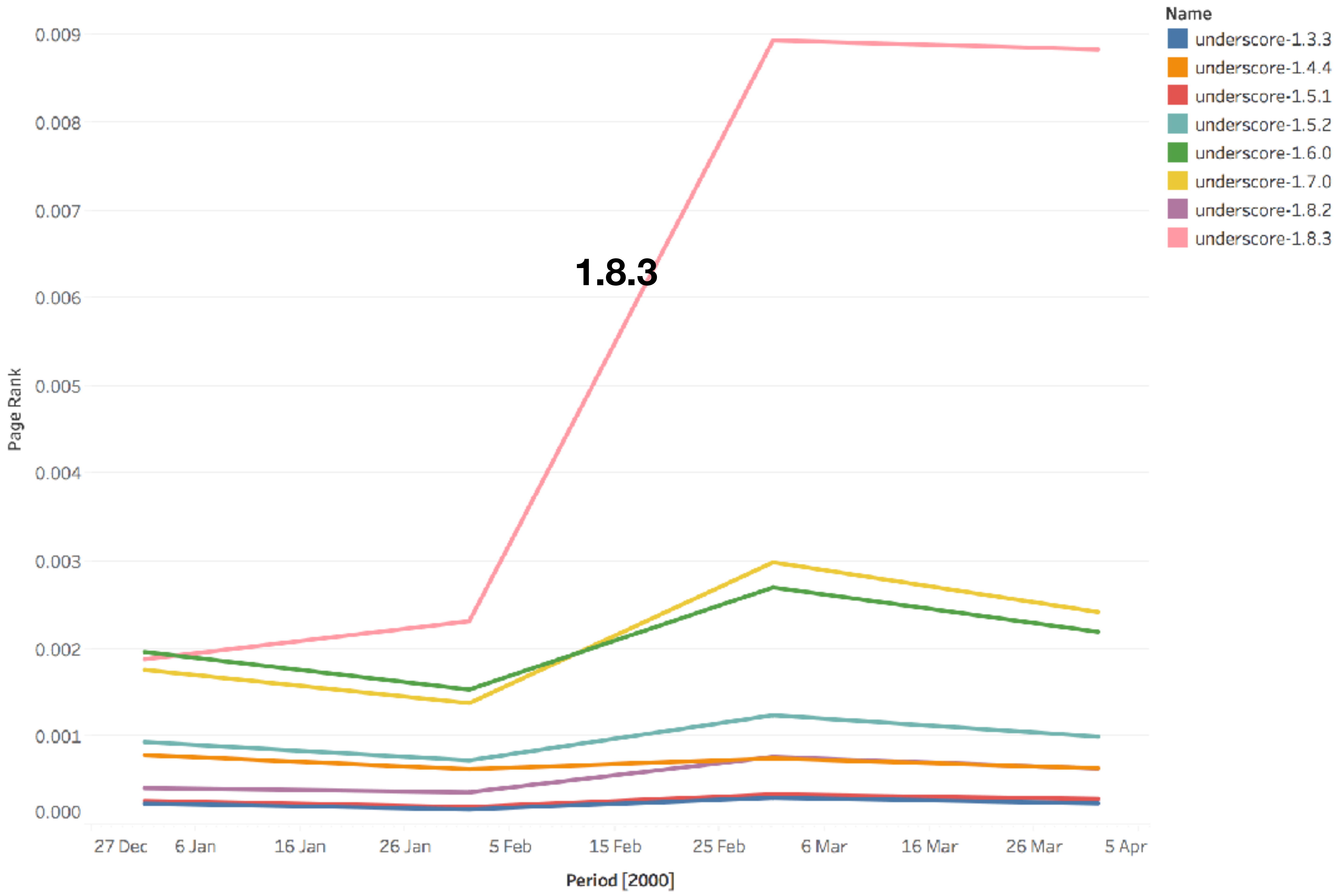
1.8.3

Degree Centrality

Case Study: underscore

Underscore.js is a utility-belt library for JavaScript that provides support for the usual functional suspects (each, map, reduce, filter...) without extending any core JavaScript objects.

- Significant growth of version 1.8.3.



Page Rank

Next Steps

- Calculate average time it takes for a package’s centrality to switch. i.e. How long does it take for the ecosystem to update to the latest version?

- Further examine case studies for packages that have a high influence over the network.

- Compute other code metrics to determine if a change in metrics correlates with a change in the nodes centrality.