

Deregulation Technical Documentation

Purpose:

The deregulation dataset attempts to capture deregulation of the *Code of Federal Regulations* through the publication of the *Federal Register*. While RegData projects successfully track the addition of regulation, or more broadly speaking, the growth rate of regulation, RegData projects don't innately identify text that is deregulatory in nature. This dataset attempts to capture just that, deregulatory text.

Content:

The deregulation dataset is created from the same documents that support the [RegPulse](#) dataset. The RegPulse dataset tracks the all movement in the *Federal Register* and can be downloaded in bulk [here](#). However, instead of focusing on notices, rules, and proposed rules like RegPulse, the deregulation dataset looks only at final rules. The text (including the titles) of these final rules are processed, and QuantGov NLP algorithms identify specific terms that typically indicate deregulation. Just like RegData provides counts of *shall*, *must*, *required*, etc., the deregulation dataset provides counts of the following terms: *13771*, *delay*, *deregulate*, *eliminates*, *eliminate*, *unnecessary*, *obsolete*, *repeal*, *withdraw*, *withdrawing*, *remove*, and *replace*. We have identified these terms as terms that strongly proxy deregulatory activity. These word counts, as well as metadata, can be retrieved for each final rule in the *Federal Register* from January 1, 2010 through June 19, 2020.

Technical Notes:

- The deregulatory dataset is currently available for the last eight years of *Federal Register* activity. We plan to update this in a future quarterly release to include upwards of 10 more historical years.
- While we have only published the data for final rules, please contact us directly if you are interested in data pertaining to proposed rules, notices, or presidential documents which are also tracked by the *Federal Register*.

Citation:

If you use this data, please cite:

McLaughlin, Patrick A., and Stephen Strosko. Deregulation US 1.0 (dataset). QuantGov, Mercatus Center at George Mason University, Arlington, VA, 2017.
<https://quantgov.org/>

Current Version: 1.0

Release Date: 7-8-2020

Variable Descriptions:

Dataset	Variable	Description
Metadata	title	Title of the document.
	effective	Generally, the date the rule becomes effective. Data for this field is inconsistent and incomplete.
	rin	The Regulation Identifier Number (RIN) associated with the document. Only about 12% of documents have data for this field.
	docket_number	The docket number of the document. The data for this field is inconsistent and incomplete.
	agencies	The department and/or agency that issued the document. Generally, the first name is the department, and the second name is the responsible agency within that department.
	doc_type	The type of document, such as a final rule (RULE), proposed rule (PROPOSED RULE), notice of proposed rulemaking (NOTICE), or a correction to a previously published document (CORRECTION). Some documents are labelled as UNCATEGORIZED DOCUMENT if the type was not determined.
	doc_number	The document number under which the document is published within the Federal Register.
	significant	This field is marked TRUE if the document is associated with a rule deemed significant under Executive Order 12866.

	publication_date	The date the document was published in the Federal Register.
	cfr_ref	The title and part of the Code of Federal Regulations (CFR) to which the documents is referring (e.g., 14 CFR 23 = title 14, part 23 of the CFR). Only about 20% of documents have data for this field.
	html_url	The link to Federal Register website where a text version of the document may be found.
	action	The current stage of the document in the regulatory process.
	page_length	The length of the document by number of pages in the Federal Register.
	topics	Tags assigned by the Federal Register website about the associated topics of the document.