State of the Union Address Data Codebook

Policy Agendas Project
2015

Overview of the Dataset

The State of the Union Address dataset provides information on each quasi-statement in Presidential State of the Union speeches. A quasi-statement simply refers to text between periods, semi-colons, and other punctuation (e.g. question marks). For example, the following text: "During the past year, we have also made a good start in providing housing for low-income groups; we have raised minimum wages; we have gone forward with the development of our natural resources; …" is divided into three quasi-statements (and thus three rows) in our dataset.

Content Coding

We have coded each entry by policy content, along with other variables of interest. We employed the standard Policy Agendas Project topic scheme in our policy content coding, and each entry was assigned one and only one content code. This allows researchers to trace activity on a particular topic across the period. The full description of our content categories may be found in the major topics codebook at http://www.comparativeagendas.info/us.

Variable Names and Descriptions:

KeyID
This column records the unique identifier for each observation. It has no substantive application.

Count
This column identifies the number of the quasi-sentence within a given year's Address. The unit of analysis in this data set is the quasi-sentence; therefore, after every semicolon and period a new entry begins.

PolicySentence
This column indicates whether the quasi-sentence has policy content. Those with policy content are coded "1"; "0" otherwise.

Date
This column indicates the date the State of the Union Address was delivered.

OralDelivery
This column specifies whether the speech was delivered orally or simply written down. Speeches delivered orally are coded "1" and "0" otherwise.

Outgoing
This column specifies whether the president delivering the speech is outgoing, giving his last speech before transition. Outgoing presidents are coded "1" and "0" otherwise.

Congress
This column denotes the Congress term in which the State of the Union Address was delivered.

President
This column lists the name of the president who delivered the State of the Union Address.

PresParty
This column records the party of the president who delivered the State of the Union Address; "100" denotes Democrat and "200" denotes Republican.

Divided
This indicates whether or not the government was divided at the time the State of the Union Address was delivered. Divided party control of legislative and executive branches is coded "1"; "0" designates unified control.

ControlHouse
This column specifies which party controlled the House of Representatives at the time the address was delivered. A Democratic controlled House is coded "100" and a Republican House is coded "200".

ControlSenate
This column specifies which party controlled the Senate at the time the address was delivered. A Democratic controlled Senate is coded "100" and a Republican Senate is coded "200".

Year
This column records the year in which the State of the Union Address was delivered.

Month
This column records the month in which the State of the Union Address was delivered.

Day
This column records the day in which the State of the Union Address was delivered.

Source
This column lists the source the Policy Agendas Project used for the State of the Union Address for a given year.

Description
This column contains the text of the given quasi-sentence.

USMajorTopic
This column records the Policy Agendas Project's major topic code that corresponds to the information found in the Description column. The codes and examples for this column are found in the Major Topics Codebook on the Policy Agendas Project website.

Note: A quasi-statement that does not contain any policy content is not given a MajorTopic code (the variable is left blank).

USSubTopicCode

This column denotes the Policy Agendas Project's subtopic code that corresponds to the information found in the Description column. The codes and examples for this column are found in the Major Topics Codebook on the Policy Agendas Project website.

Note: A quasi-statement that does not contain any policy content is given an arbitrary subTopicCode of -555.

MajorTopic

This column records the Comparative Agendas Project's major topic code that corresponds to the existing Policy Agendas Project topic code. Please visit the Comparative Agendas Project website (www.comparativeagendas.net) for more information.

Note: A quasi-statement that does not contain any policy content is not given a MajorTopic code (the variable is left blank).

SubTopicCode

This column records the Comparative Agendas Project's subtopic topic code that corresponds to the existing Policy Agendas Project topic code. Please visit the Comparative Agendas Project website (www.comparativeagendas.net) for more information.

Note: A quasi-statement that does not contain any policy content is given an arbitrary subTopicCode of -555.