R Notebook

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.4.2
```

```
## -- Attaching packages ---------------------------------------------------------------------------
```

```
## v ggplot2 3.1.0     v purrr   0.2.5
## v tibble  2.0.1     v dplyr   0.7.8
## v tidyr   0.8.0     v stringr 1.3.1
## v readr   1.1.1     v forcats 0.3.0
```

```
## Warning: package 'ggplot2' was built under R version 3.4.4
```

```
## Warning: package 'tibble' was built under R version 3.4.4
```

```
## Warning: package 'tidyr' was built under R version 3.4.3
```

```
## Warning: package 'purrr' was built under R version 3.4.4
```

```
## Warning: package 'dplyr' was built under R version 3.4.4
```

```
## Warning: package 'forcats' was built under R version 3.4.3
```

```
## -- Conflicts ------------------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(knitr)
```

```
## Warning: package 'knitr' was built under R version 3.4.3
```

Document clustering, or text clustering, is a very popular application of clustering algorithms. A web search engine, like Google, often returns thousands of results for a simple query. For example, if you type the search term "jaguar" into Google, around 200 million results are returned. This makes it very difficult to browse or find relevant information, especially if the search term has multiple meanings. If we search for "jaguar", we might be looking for information about the animal, the car, or the Jacksonville Jaguars football team.

Clustering methods can be used to **automatically group search results into categories**, making it easier to find relevant results. This method is used in the search engines PolyMeta and Helioid, as well as on FirstGov.gov, the official Web portal for the U.S. government. The two most common algorithms used for document clustering are Hierarchical and k-means.

In this problem, we'll be clustering articles published on Daily Kos, an American political blog that publishes news and opinion articles written from a progressive point of view. Daily Kos was founded by Markos Moulitsas in 2002, and as of September 2014, the site had an average weekday traffic of hundreds of thousands of visits.

The file **dailykos.csv contains data on 3,430 news articles or blogs that have been posted on Daily Kos**. These articles were posted in 2004, leading up to the United States Presidential Election. The leading candidates were incumbent President George W. Bush (republican) and John Kerry (democratic). Foreign policy was a dominant topic of the election, specifically, the 2003 invasion of Iraq.

Each of the variables in the dataset is a word that has appeared in at least 50 different articles (1,545 words in total). The set of words has been trimmed according to some of the techniques covered in the previous week on text analytics (punctuation has been removed, and stop words have been removed). For each document, **the variable values are the number of times that word appeared in the document**.

1.1) Hierarchical Clustering

Let's start by building a hierarchical clustering model. First, read the data set into R. Then, compute the distances (using method="euclidean"), and use hclust to build the model (using method="ward.D"). You should cluster on all of the variables.

Running the dist function will probably take you a while. Why? Select all that apply.

```
daily_kos = read.csv('dailykos.csv')

head(daily_kos)
```

```
##   abandon abc ability abortion absolute abstain abu abuse accept access
## 1       0   0       0        0        0       0   0     0      0      0
## 2       0   0       0        0        0       0   0     0      0      0
## 3       0   0       0        0        0       1   0     0      0      0
## 4       0   0       0        0        0       0   0     0      0      0
## 5       0   0       0        0        0       0   0     0      0      0
## 6       0   0       0        0        0       0   0     0      0      0
##   accomplish account accurate accusations achieve acknowledge act action
## 1          0       0        0           0       0           0   0      2
## 2          0       0        0           0       0           0   0      0
## 3          0       2        0           0       0           0   0      0
## 4          0       0        0           2       0           0   0      0
## 5          0       0        0           0       0           0   0      0
## 6          0       0        0           0       0           0   0      0
##   active activist actual add added addition address admin administration
## 1      0        0      0   0     1        0       0     0              1
## 2      0        0      0   0     0        0       0     0              0
## 3      0        0      0   0     0        0       0     1              0
## 4      0        0      0   0     0        0       0     0              0
## 5      0        0      0   0     1        0       0     0              0
## 6      0        0      0   0     0        0       0     0              0
##   admit advance advantage advertise advised affair affect affiliate
## 1     0       0         0         0       0      0      0         0
## 2     0       0         0         0       0      0      0         0
## 3     0       0         0         1       0      0      0         0
## 4     0       0         1         0       0      0      0         0
## 5     1       0         0         0       0      0      0         0
## 6     0       0         0         0       0      0      0         0
##   afghanistan afraid afternoon age agencies agenda agree ahead aid aim air
## 1           0      0         0   0        0      0     0     0   0   0   0
## 2           0      0         0   0        0      0     0     0   0   0   0
## 3           0      0         0   0        0      0     0     0   0   0   0
## 4           0      0         0   0        0      0     0     0   1   0   0
## 5           0      0         0   0        0      0     0     0   1   0   0
## 6           0      0         0   0        0      0     0     0   0   0   0
##   alaska allegation allegory allied allowed alternative altsite amazing
## 1      0          0        0      0       0           0       0       0
## 2      0          0        0      0       0           0       0       0
## 3      0          0        0      0       0           0       1       0
## 4      0          0        0      0       0           0       0       0
## 5      0          0        0      0       0           0       0       0
## 6      0          0        0      0       0           0       0       0
##   amendment america american amount amp analysis analyst anecdotal anger
## 1         0       0        0      0   0        0       0         0     0
```

```
## 2         0       0        0      0   0        0       0         0     0
## 3         0       0        0      0   0        0       0         1     0
## 4         0       0        0      0   0        0       0         0     0
## 5         0       0        1      0   0        1       0         0     0
## 6         0       0        0      0   0        0       0         0     0
##   angry announce annual answer apologies apparent appeal appearance
## 1     0        0      0      0         0        0      0          0
## 2     0        0      0      0         0        0      0          0
## 3     0        0      0      0         0        0      0          0
## 4     0        0      0      1         0        0      0          0
## 5     0        0      0      0         0        0      0          0
## 6     0        0      0      0         0        0      0          0
##   applied appointed approach approval apr april arab area arent arg argue
## 1       0         0        0        1   0     0    0    0     0   0     0
## 2       0         0        0        0   0     0    0    0     0   0     0
## 3       0         0        0        0   0     0    0    0     0   0     0
## 4       0         0        0        0   0     0    0    0     0   0     0
## 5       0         0        0        1   0     0    0    0     0   0     0
## 6       0         0        0        0   0     0    0    0     0   0     0
##   argument arizona arm armstrong army arrest arrive article asap asked ass
## 1        0       0   0         0    0      0      0       0    0     0   0
## 2        0       0   0         0    0      0      0       0    0     0   0
## 3        0       0   0         0    0      0      0       0    1     0   0
## 4        0       0   0         0    0      0      0       0    0     0   0
## 5        0       0   0         0    0      0      0       0    0     0   0
## 6        0       0   0         0    0      0      0       0    0     0   0
##   assess assist associate assume atlanta atrios attack attempt attend
## 1      0      0         0      0       0      0      1       0      0
## 2      0      0         0      0       0      0      0       0      0
## 3      0      0         0      0       1      0      0       1      0
## 4      0      0         0      0       0      0      2       0      0
## 5      0      0         0      0       0      0      4       0      0
## 6      0      0         0      0       0      0      0       0      0
##   attention attorney audience aug august author average avoid backed
## 1         0        0        0   1      0      0       0     0      0
## 2         0        0        0   1      0      0       0     0      0
## 3         0        0        0   1      0      0       0     0      0
## 4         0        0        0   1      0      0       0     0      0
## 5         0        0        0   1      0      0       0     0      0
## 6         0        0        0   1      0      0       0     0      0
##   background bad baghdad balance bald ballot ban bank bar barbero base
## 1          0   0       0       0    0      0   0    0   0       0    0
## 2          0   0       0       0    0      0   0    0   0       0    0
## 3          0   0       0       0    1      0   0    0   0       0    0
## 4          0   0       0       0    0      0   0    0   0       0    0
## 5          0   0       0       0    0      0   0    0   0       0    0
## 6          0   1       0       0    0      0   0    0   0       0    0
##   basic battle battleground bear beat began begin behalf behavior believed
## 1     0      1            0    0    0     0     0      0        0        1
## 2     0      0            0    0    0     0     0      0        0        0
## 3     0      0            0    0    0     0     0      0        0        0
## 4     0      0            0    0    0     0     0      0        0        0
## 5     0      0            0    0    0     0     0      0        0        0
## 6     0      0            0    0    0     0     0      0        0        0
```

```
##   ben benefit bet bid bigger biggest bill billion bin bit black blades
## 1   0       0   0   0      0       0    0       0   0   0     0      0
## 2   0       0   0   0      0       0    0       0   0   0     0      0
## 3   0       0   0   0      0       0    0       0   0   0     0      0
## 4   0       0   0   0      0       0    0       0   0   0     0      0
## 5   0       0   0   0      0       0    0       0   0   0     0      0
## 6   0       0   0   0      0       0    0       0   0   0     0      0
##   blame blast block blog blogger blogosphere bloomfield blow blue board
## 1     0     0     0    0       0           0          0    0    0     0
## 2     0     0     0    0       0           0          0    0    0     0
## 3     0     0     0    0       0           0          0    0    0     0
## 4     0     0     0    0       0           0          0    0    0     0
## 5     0     0     0    0       0           0          0    0    0     0
## 6     0     0     0    0       0           0          0    0    0     0
##   boat bob bodies bomb book boost boston bottom bounce boxblogroll
## 1    0   0      0    0    0     0      0      0      0           0
## 2    1   0      0    0    1     0      0      0      0           0
## 3    0   1      0    0    0     0      0      0      0           1
## 4    0   0      0    0    0     0      0      0      0           0
## 5    0   0      0    0    0     0      0      0      0           0
## 6    0   0      0    0    0     0      0      0      0           0
##   boxfeed_listing boxrdf_feeds boy brad bradnickel braun break. bring
## 1               0            0   0    0          0     0      0     0
## 2               0            0   0    0          0     0      0     0
## 3               1            1   0    0          1     0      0     0
## 4               0            0   0    0          0     0      0     0
## 5               0            0   0    0          0     0      0     0
## 6               0            0   0    0          0     0      0     0
##   broadcast broke brought btw budget build built bunch burn bush
## 1         0     0       0   0      0     0     0     0    0    2
## 2         0     0       0   0      0     0     0     0    0    1
## 3         0     0       0   0      0     0     0     0    0    2
## 4         0     0       0   0      0     0     0     0    0    0
## 5         0     0       0   0      0     0     0     0    0   12
## 6         0     0       0   0      0     0     0     0    0    4
##   bushcheney bushsux business buy calculate california calistan call camp
## 1          0       0        0   0         0          0        0    0    0
## 2          0       0        0   0         0          0        0    1    0
## 3          0       1        0   0         1          0        1    0    0
## 4          0       0        0   0         0          0        0    0    0
## 5          0       0        0   1         0          0        0    0    0
## 6          0       0        0   0         0          0        0    0    0
##   campaign candidacy candidate capabilities capture car card care career
## 1        0         0         0            0       0   0    0    0      1
## 2        1         0         0            0       0   0    0    0      0
## 3        0         0         0            0       0   0    0    0      0
## 4        5         0         0            0       0   0    0    0      0
## 5        2         0         1            0       0   0    0    0      0
## 6        0         0         0            0       0   0    0    0      0
##   carolina carried carson cases cash cast casualties catch caucus caught
## 1        0       0      0     0    0    0          0     0      0      0
## 2        0       0      0     0    0    0          0     0      0      0
## 3        0       0      0     0    0    1          0     0      0      0
## 4        0       0      0     0    0    0          0     0      0      1
```

```
## 5        0       0      0     0    0    0          0     0      0      0
## 6        0       0      0     0    0    0          0     0      0      0
##   caused cbs cell center central cfr chair chairman challenge chance
## 1      0   0    0      0       0   0     0        1         0      0
## 2      0   0    0      0       0   0     0        0         0      0
## 3      0   0    0      0       0   0     0        0         4      0
## 4      0   0    0      0       0   0     0        0         0      0
## 5      0   0    0      0       0   0     0        0         0      0
## 6      0   0    0      0       0   0     0        0         0      0
##   chandler change charge check chedrcheez chemical cheney chicago chief
## 1        0      0      0     0          0        0      0       0     0
## 2        0      0      0     0          0        0      0       0     0
## 3        0      0      0     0          0        0      0       0     0
## 4        0      1      0     0          0        0      0       0     0
## 5        0      0      0     0          0        0      0       0     0
## 6        0      0      0     0          0        0      0       0     0
##   children choice choose chose chris christopher cia cited cities citizen
## 1        0      0      0     0     0           0   0     0      0       0
## 2        0      0      0     0     0           0   0     0      0       0
## 3        0      0      0     0     0           1   0     0      1       1
## 4        0      0      0     0     0           0   0     0      0       0
## 5        0      2      0     0     0           0   0     0      0       0
## 6        0      0      0     0     0           0   0     0      0       0
##   civil civilian claim clark class clean clear click clinton clock close
## 1     0        0     0     0     0     0     0     1       0     0     0
## 2     0        0     0     0     0     0     0     1       0     0     0
## 3     0        0     0     0     0     0     0     0       0     0     0
## 4     0        0     0     0     0     1     0     0       0     0     0
## 5     0        0     0     0     0     0     0     0       0     0     0
## 6     0        0     0     0     0     0     0     0       0     0     0
##   closer club cnn coalition coburn cole collect college colorado column
## 1      0    0   0         1      0    0       0       0        0      0
## 2      0    0   0         0      0    0       0       0        0      0
## 3      0    0   0         0      0    0       0       0        0      0
## 4      0    0   0         0      0    0       0       0        0      0
## 5      0    0   0         0      0    0       0       0        0      0
## 6      0    0   0         0      0    0       0       0        1      0
##   combat combination coming command commercial commission commit committee
## 1      0           0      0       0          0          0      0         2
## 2      0           0      0       1          0          0      0         0
## 3      0           0      0       0          0          0      0         0
## 4      0           0      0       0          0          0      0         0
## 5      0           0      0       0          0          0      0         0
## 6      0           0      0       0          0          0      0         0
##   common communicate communities companies compare comparison compete
## 1      0           0           0         0       2          0       0
## 2      0           0           0         0       0          0       0
## 3      0           0           0         0       0          0       0
## 4      0           0           0         0       0          0       0
## 5      0           0           0         0       1          0       0
## 6      0           0           0         0       0          0       0
##   competition complain complaint complete computer con concern conclude
## 1           0        0         0        0        0   0       0        1
## 2           0        0         0        0        0   0       0        0
```

```
## 3           0        0         0        0        0   0       0        0
## 4           0        0         0        0        0   0       0        0
## 5           0        0         0        0        0   0       0        0
## 6           0        0         0        0        0   0       0        0
##   conclusion condition conduct conference confidence confirm conflict
## 1          2         0       0          0          0       0        0
## 2          0         0       0          0          0       0        0
## 3          0         0       0          0          0       0        0
## 4          0         0       0          0          0       0        0
## 5          0         0       0          0          0       0        0
## 6          0         0       0          0          0       0        0
##   congress congressional congressman connect consequence conservation
## 1        0             2           2       0           0            0
## 2        0             0           0       0           0            0
## 3        0             0           0       0           0            0
## 4        1             0           1       0           0            0
## 5        1             0           1       0           0            0
## 6        0             0           0       0           0            0
##   considerable considered consistent constitution consultant contact
## 1            0          1          0            0          0       0
## 2            0          0          0            0          0       0
## 3            0          0          0            0          0       1
## 4            0          0          0            0          0       0
## 5            0          0          0            0          0       0
## 6            0          0          0            0          0       0
##   content contest continue contrast contribute control controversial
## 1       0       0        0        0          0       0             0
## 2       0       0        0        0          0       0             0
## 3       0       0        0        0          0       0             0
## 4       0       0        0        0          0       0             1
## 5       0       0        0        0          0       0             0
## 6       0       0        0        0          0       0             0
##   convention conversation converted convince cool coordinated corner corp
## 1          0            0         0        0    0           0      0    0
## 2          0            0         0        0    0           0      0    0
## 3          0            0         0        0    0           0      0    0
## 4          0            0         0        0    0           0      0    0
## 5          1            0         0        1    0           0      0    0
## 6          0            0         0        0    0           0      0    0
##   corporate correct corrupt cost couldnt council count countdown counties
## 1         0       0       0    0       0       0     0         0        0
## 2         0       0       0    0       0       0     0         0        0
## 3         0       0       0    0       0       0     0         0        0
## 4         0       0       0    0       1       0     0         0        0
## 5         0       0       0    0       0       0     0         0        0
## 6         0       0       0    0       0       0     0         0        0
##   countries couple court courtesy cover coverage crazy create credibility
## 1         1      0     0        0     0        0     0      0           0
## 2         0      0     0        0     0        0     0      0           0
## 3         0      0     0        0     0        0     0      1           0
## 4         0      0     0        0     1        0     0      0           0
## 5         0      0     0        0     0        0     0      0           1
## 6         0      0     0        0     0        0     0      0           0
##   credit crime criminal crisis critic crooks cross crowd cultural current
```

```
## 1      0     0        0      0      0      0     0     0        0       0
## 2      0     0        0      0      0      0     1     0        0       0
## 3      0     0        0      0      0      0     0     0        1       0
## 4      0     0        0      0      0      0     0     0        0       0
## 5      0     0        0      1      0      0     0     0        0       0
## 6      0     0        0      0      0      0     0     0        0       0
##   cut cycle dakota damage damn dan danger danielua daschle data date david
## 1   0     0      0      1    1   0      0        0       0    0    0     0
## 2   0     0      0      0    1   0      0        0       0    0    0     0
## 3   0     0      0      0    0   0      0        0       0    0    0     0
## 4   0     1      0      0    0   0      0        0       0    0    0     0
## 5   0     0      0      0    1   0      0        0       0    0    0     0
## 6   0     0      0      0    0   0      0        0       0    0    0     0
##   dccc dday dead deadline deal dean death debate dec decade december
## 1    0    0    0        0    0    0     0      0   0      0        0
## 2    0    0    0        0    0    0     0      1   0      0        0
## 3    0    0    0        0    0    0     0      1   0      0        0
## 4    0    0    0        0    0    0     0      0   0      0        0
## 5    0    0    0        0    0    0     0      0   0      0        0
## 6    0    0    0        0    0    0     0      0   0      0        0
##   decide decision declaration decline deep defeat defend defense deficit
## 1      0        0           0       0    0      0      0       0       0
## 2      0        0           0       0    0      0      0       0       0
## 3      0        0           0       0    0      0      0       0       0
## 4      0        0           0       0    0      0      0       0       0
## 5      0        0           0       0    0      0      1       0       0
## 6      0        0           0       0    0      0      0       0       0
##   define delay delegate deliver dem demand democracy democrat demonstrate
## 1      0     0        0       0   0      0         0        0           0
## 2      0     0        0       0   0      0         0        0           0
## 3      0     0        0       0   1      0         0        2           0
## 4      0     1        0       0   0      0         0        0           0
## 5      0     0        0       0   0      0         0        0           0
## 6      0     0        0       0   0      0         0        0           0
##   denied department depend describe deserve design desperate destroy
## 1      0          0      0        0       0      0         0       0
## 2      0          0      0        0       0      0         0       0
## 3      0          0      0        0       0      0         0       0
## 4      0          0      0        0       0      0         0       0
## 5      0          0      0        0       0      0         0       0
## 6      0          0      0        0       0      0         0       0
##   destruction detail determination develop dhinmi dick die differ
## 1           1      0             0       0      0    0   0      0
## 2           0      0             0       0      0    0   0      0
## 3           0      0             0       0      0    0   0      0
## 4           0      0             0       0      0    0   0      0
## 5           0      0             0       0      0    0   0      0
## 6           0      0             0       0      0    0   0      0
##   difficult diplomatic direct director direwolf disaster discussed dismiss
## 1         0          0      0        0        0        0         0       0
## 2         0          0      0        0        0        0         0       0
## 3         0          1      0        0        0        0         0       0
## 4         0          0      0        0        0        0         1       0
## 5         0          0      1        0        0        0         1       0
```

```
## 6         0          0      0        0        0        0         0       0
##   dispute district divide division dkos dkosopedia dnc document dollar
## 1       0        2      0        0    0          0   0        0      0
## 2       0        0      0        0    0          0   0        0      0
## 3       0        0      0        0    0          1   0        0      0
## 4       0        1      0        0    0          0   0        0      0
## 5       0        0      0        0    0          0   0        0      0
## 6       0        0      0        0    0          0   0        0      0
##   domestic dominance donald donate donor door doubt dozen draft draw dream
## 1        0         0      0      0     0    0     0     0     0    0     0
## 2        0         0      0      0     0    0     0     0     0    0     0
## 3        0         0      0      0     0    0     0     0     0    0     0
## 4        0         0      0      0     0    0     0     0     0    0     0
## 5        0         0      0      0     0    0     0     0     0    0     0
## 6        0         0      0      0     0    0     0     0     0    0     0
##   drfranklives drive drop drug dryfly dscc duderino due duties earlier
## 1            0     0    0    0      0    0        0   0      0       0
## 2            0     0    1    0      0    0        0   0      0       0
## 3            0     0    0    0      1    0        0   0      0       0
## 4            0     0    1    0      0    0        0   0      0       0
## 5            0     1    0    0      0    0        0   0      0       0
## 6            0     0    0    0      0    0        0   0      0       0
##   early earn easier easily east easy economic economy edge edit editor
## 1     0    0      0      0    0    0        0       0    0    0      0
## 2     0    0      0      0    0    0        0       0    0    0      0
## 3     0    0      0      0    0    0        0       1    0    0      0
## 4     0    0      0      1    0    0        0       0    0    0      0
## 5     0    0      0      0    0    0        1       0    0    0      0
## 6     0    0      0      0    0    0        0       0    0    0      0
##   educated edward effect effort egon elect electoral email emerge employ
## 1        0      0      0      0    0     0         1     0      0      0
## 2        0      0      0      0    0     0         0     0      0      0
## 3        0      0      0      0    0     1         2     0      0      0
## 4        0      0      0      0    0     2         0     0      0      0
## 5        0      0      0      0    0     3         1     0      0      0
## 6        0      0      0      0    0     0         0     0      0      0
##   employee encourage end endorse endspan enemies energy engage enjoy
## 1        0         0   0       0       0       0      0      1     0
## 2        0         0   0       0       0       0      0      0     0
## 3        0         0   0       0       1       0      0      0     0
## 4        0         0   0       0       0       0      0      0     0
## 5        0         0   1       0       0       0      1      0     0
## 6        0         0   0       0       0       0      0      0     0
##   ensure enter entire entries environment equal error essential establish
## 1      0     0      0       0           0     0     0         0         0
## 2      0     0      0       0           0     0     0         0         0
## 3      0     0      0       0           0     0     0         0         0
## 4      0     0      0       0           0     0     0         0         0
## 5      0     0      0       0           0     0     0         0         0
## 6      0     0      0       0           0     0     0         0         0
##   estimate ethic evening event eventual evidence excellent excited
## 1        0     0       0     0        0        0         0       0
## 2        0     0       0     0        0        0         0       0
## 3        0     0       0     0        0        0         0       0
```

```
## 4        0     0       0     0        0        0         0       0
## 5        0     0       0     0        0        0         0       0
## 6        0     0       0     0        0        0         0       0
##   executed exist exit expect expense experience expert explain explosion
## 1        0     0    0      0       0          0      0       0         0
## 2        0     0    0      0       0          0      0       0         0
## 3        0     0    1      0       0          1      0       0         0
## 4        0     0    0      1       1          0      0       0         0
## 5        0     0    0      2       0          0      0       0         0
## 6        0     0    0      0       0          0      0       0         0
##   expose express extend extra extreme eye face fact factor fail failure
## 1      0       0      0     0       0   0    0    1      0    0       0
## 2      0       0      0     0       0   0    0    0      0    0       0
## 3      0       0      0     0       0   0    0    0      0    0       0
## 4      0       0      0     0       0   0    0    1      0    0       0
## 5      0       0      0     0       0   0    2    1      0    2       0
## 6      0       0      0     0       0   0    0    0      0    0       0
##   fair faith fall fallujah familiar families faq farmer fast father favor
## 1    0     0    0        0        0        0   0      0    0      0     0
## 2    0     0    0        0        0        0   0      0    0      0     0
## 3    0     0    0        0        0        0   1      0    0      0     0
## 4    0     0    0        0        0        0   0      0    0      0     0
## 5    0     0    0        0        0        0   0      0    0      0     0
## 6    0     0    0        0        0        0   0      1    0      0     0
##   favorite fbi fear feature feb february federal feel fellow felt fewer
## 1        0   0    0       0   0        0       0    0      0    0     0
## 2        0   0    0       0   0        0       0    0      0    0     0
## 3        0   0    0       0   0        0       0    0      0    0     0
## 4        0   0    0       0   0        0       0    0      0    0     0
## 5        0   0    0       0   0        0       0    0      0    0     0
## 6        0   0    0       0   0        0       0    0      0    0     0
##   field fight figure file fill final finance financial find finish fire
## 1     0     0      0    0    0     1       0         0    0      0    0
## 2     0     0      0    0    0     0       0         0    0      0    0
## 3     0     0      0    0    0     0       0         0    0      0    0
## 4     0     0      0    0    0     0       0         0    0      0    0
## 5     0     1      0    0    0     0       0         0    0      0    0
## 6     0     0      1    0    0     0       0         0    0      0    0
##   firm fit fix flip floor florida fly focus folks follow force foreign
## 1    0   0   0    0     0       0   0     0     0      0     0       0
## 2    0   0   0    0     0       0   0     0     0      0     1       0
## 3    0   0   0    0     0       0   0     0     0      0     0       0
## 4    0   0   0    0     0       0   0     0     0      0     0       0
## 5    0   0   0    0     0       0   0     1     0      0     0       0
## 6    0   0   0    0     0       0   0     0     0      0     0       0
##   forget form forward fought found foundation fox francisco frank free
## 1      0    0       0      0     0          0   0         0     0    0
## 2      0    0       0      0     0          0   0         0     0    0
## 3      0    0       0      0     0          0   0         0     0    0
## 4      0    0       0      0     0          0   0         0     0    0
## 5      0    0       0      0     1          0   0         0     0    0
## 6      0    0       0      0     0          0   0         0     0    0
##   freedom fresh friday friend front full fully fun function. fund
## 1       0     0      0      0     0    0     0   0         0    0
```

```
## 2       0     0      0      0     0    0     0   0         0    0
## 3       0     0      0      1     0    0     0   0         1    0
## 4       0     0      0      0     0    0     0   0         0    0
## 5       0     0      0      0     0    0     0   0         0    0
## 6       0     0      0      0     0    0     0   0         0    0
##   fundamental fundraise funny future gain gallup game garner gatana gather
## 1           0         0     0      0    0      0    0      0      0      0
## 2           0         0     0      0    0      0    0      0      0      0
## 3           0         0     0      0    0      0    0      0      0      0
## 4           0         0     0      0    0      0    0      0      0      0
## 5           0         0     0      0    0      0    0      0      0      0
## 6           0         0     0      0    0      0    0      0      0      0
##   gave gay gen general generate george georgia gephardt ghraib giving
## 1    0   0   0       2        0      0       0        0      0      0
## 2    0   0   0       1        0      0       0        0      0      0
## 3    0   0   0       2        0      0       0        0      0      0
## 4    1   0   0       0        0      0       0        0      0      0
## 5    0   0   0       1        1      0       0        0      0      0
## 6    0   0   0       1        0      0       0        0      0      0
##   global goal god gooper gop gore gotv gov govern governor grand grant
## 1      0    0   0      1   0    0    0   0      0        0     0     0
## 2      0    0   0      0   0    0    0   0      0        0     0     0
## 3      0    0   0      0   0    0    0   0      0        2     0     0
## 4      0    0   0      0   0    0    0   0      0        0     0     0
## 5      0    0   0      0   0    0    0   0      0        0     0     1
## 6      0    0   0      0   0    0    0   0      0        1     0     0
##   greater greatest green ground group grow growth guard guess guest gun
## 1       1        0     0      0     0    0      0     0     0     0   0
## 2       0        0     0      0     2    0      0     0     0     0   0
## 3       0        0     0      0     0    0      0     0     0     0   0
## 4       0        0     0      0     0    0      0     0     0     0   0
## 5       0        0     0      0     0    0      0     0     0     0   0
## 6       0        0     0      0     0    0      0     0     0     0   0
##   guy hadnt half hall hampshire hand handle hang happen happy hard harris
## 1   0     0    0    0         0    0      0    0      0     0    0      0
## 2   0     0    0    0         0    0      0    0      0     0    0      0
## 3   0     0    0    0         0    0      0    0      0     0    0      0
## 4   0     0    0    0         0    0      0    0      0     0    0      0
## 5   0     0    0    0         0    0      0    0      0     0    0      0
## 6   2     0    0    0         0    0      0    0      0     0    0      0
##   hasnt hate havent head headline health hear heard heart heat heavily
## 1     0    0      0    0        0      0    0     0     0    0       0
## 2     0    0      0    0        0      0    0     0     0    0       0
## 3     0    0      0    0        0      0    0     0     0    0       0
## 4     0    0      0    0        0      0    0     0     0    0       0
## 5     0    0      0    0        0      0    0     0     0    0       0
## 6     0    0      0    0        0      0    0     0     0    0       0
##   heavy heck held hell helped heres herseth hey hide high higher highest
## 1     0    0    0    0      0     0       0   0    0    0      0       0
## 2     0    0    0    0      0     0       0   0    0    0      0       0
## 3     0    0    0    0      0     0       0   0    0    0      0       0
## 4     0    0    0    0      0     0       0   0    0    0      0       0
## 5     0    0    0    0      0     0       0   0    0    0      0       0
## 6     0    0    0    0      0     0       0   0    0    0      0       0
```

```
##   highlight hire historic history hit hold homepage honest honor hoodies
## 1         0    0        0       0   0    1        0      0     0       0
## 2         0    0        0       0   0    0        0      0     0       0
## 3         0    0        0       0   0    0        1      0     0       0
## 4         0    0        0       0   0    0        0      0     0       0
## 5         0    1        0       0   0    0        0      0     0       0
## 6         0    0        0       0   0    0        0      0     0       0
##   hope host hostile hot hotshotxi hour house howard hstewart huge human
## 1    0    0       0   0         0    0     2      0        0    0     0
## 2    0    0       0   0         0    0     0      0        0    0     0
## 3    0    0       0   0         0    0     2      0        0    0     0
## 4    0    0       0   0         0    0     2      0        0    0     0
## 5    1    0       0   0         0    0     1      0        0    0     0
## 6    0    0       0   0         0    0     0      0        0    0     0
##   hundred hurt husband hussein idea identified ideological idetestthispres
## 1       0    0       0       0    0          0           0               0
## 2       0    0       0       0    0          0           0               0
## 3       0    0       0       0    0          0           0               0
## 4       0    0       0       0    0          0           0               0
## 5       0    0       0       0    0          0           0               0
## 6       0    0       0       0    0          0           0               0
##   ignorance ill illinois image imagination immediately impact importance
## 1         0   0        0     0           0           0      0          0
## 2         0   1        0     0           0           0      0          0
## 3         0   1        0     0           0           0      0          0
## 4         0   0        0     0           0           0      0          0
## 5         0   0        0     0           0           0      0          0
## 6         0   0        0     0           0           0      0          0
##   impossible impressed improve incident incidentally include income
## 1          0         0       0        0            1       0      0
## 2          0         0       0        0            0       0      0
## 3          0         0       0        0            0       0      0
## 4          0         0       0        0            0       0      0
## 5          0         0       0        0            0       0      0
## 6          0         0       0        0            0       0      0
##   increase incumbency independence indicating indicted individual
## 1        0          0            1          0        0          0
## 2        0          0            0          0        0          0
## 3        0          0            0          0        0          0
## 4        0          1            0          0        0          0
## 5        0          4            0          0        0          0
## 6        0          0            0          0        0          0
##   industrial inevitable influence info inform initial inside insist
## 1          0          0         0    0      0       0      0      0
## 2          0          0         0    0      0       0      2      0
## 3          0          0         0    0      0       0      0      0
## 4          0          0         0    0      0       0      0      0
## 5          0          0         0    0      0       0      0      0
## 6          0          0         0    0      0       0      0      0
##   instance institute insurgency intelligence intent interest internal
## 1        0         0          0            3      0        0        2
## 2        0         0          0            0      0        0        0
## 3        0         0          0            0      0        0        0
## 4        0         0          0            0      0        0        0
```

```
## 5        0         0          0            0      0        0        0
## 6        0         0          0            0      0        0        0
##   internet interview invasion investigate involve iowa iraq iraqi isnt
## 1        0         0        0           0       0    0    2     0    0
## 2        0         0        0           0       0    0    0     0    0
## 3        0         0        0           0       0    0    0     0    0
## 4        0         2        0           0       0    0    0     0    0
## 5        0         0        0           0       0    0    1     0    0
## 6        0         0        0           0       0    0    0     0    0
##   issue itll ive james jan january jeff jerome jersey jiacinto jim job joe
## 1     0    0   1     0   0       0    0      0      0        0   0   0   0
## 2     0    0   0     0   0       0    0      0      0        0   0   0   0
## 3     0    0   0     0   0       0    0      0      0        1   0   0   0
## 4     0    0   0     0   0       0    0      0      0        0   0   0   0
## 5     0    0   0     0   0       0    0      0      0        0   0   2   0
## 6     0    0   0     0   0       0    0      0      0        0   0   0   0
##   john join josh journal journalist juan judge jul july jump jun june
## 1    1    0    0       0          0    0     0   0    0    0   0    0
## 2    2    0    0       0          0    0     0   0    0    0   0    0
## 3    0    0    0       0          0    0     0   0    0    0   0    0
## 4    0    0    0       0          0    0     0   0    0    0   0    0
## 5    0    0    0       0          0    0     0   0    0    0   0    1
## 6    0    0    0       0          0    0     0   0    0    0   0    0
##   juppon justice karl katerina keeping ken kennedy kentucky kerry
## 1      0       0    0        0       0   0       0        0     3
## 2      0       0    0        0       0   0       0        0     7
## 3      0       0    0        0       0   0       0        0     2
## 4      0       0    0        0       1   0       0        0     0
## 5      0       0    0        0       0   0       0        0     5
## 6      0       0    0        0       0   0       0        0     4
##   kerryedwards key kick kid kill kind king kingelection kitty knew knowles
## 1            0   0    0   0    0    0    0            0     0    0       0
## 2            0   0    0   0    0    0    0            0     0    0       0
## 3            0   0    0   1    0    0    0            0     1    0       0
## 4            0   0    0   0    0    0    0            0     0    0       0
## 5            0   0    0   0    0    0    0            0     0    0       0
## 6            0   0    0   0    0    0    0            0     0    0       0
##   kossacks kucinich labor lack laden land language large larger largest
## 1        0        0     0    0     0    0        0     0      0       0
## 2        0        0     0    0     0    0        0     0      0       0
## 3        1        0     1    0     0    0        0     0      0       0
## 4        0        0     0    0     0    0        0     0      0       0
## 5        0        0     0    0     0    0        0     0      0       0
## 6        0        0     0    0     0    0        0     0      0       0
##   larry late latest launch law lawnorder lawyer lay lead leader leadership
## 1     0    0      0      1   0         0      0   0    1      0          0
## 2     0    0      0      0   0         0      0   0    0      0          0
## 3     0    0      0      0   0         0      0   0    0      0          0
## 4     0    1      0      0   0         0      0   0    0      0          0
## 5     0    0      0      0   0         0      0   0    0      0          0
## 6     0    0      0      0   0         0      0   0    0      0          0
##   leak lean learn leave led left legal legislation legislature legitimate
## 1    0    0     0     0   0    0     0           0           0          0
## 2    0    0     0     0   0    0     0           0           0          0
```

```
## 3    0    0     0     0   0    0     0           0           0          0
## 4    0    0     0     0   0    0     0           0           0          0
## 5    0    0     0     0   0    0     0           0           0          0
## 6    0    0     0     0   0    0     0           0           0          0
##   lets letter level liar liberal liberalrakkasan lie lieberman life light
## 1    0      1     0    0       0               0   0         0    0     0
## 2    0      0     0    0       0               0   0         0    0     0
## 3    0      0     0    0       0               1   0         0    0     0
## 4    0      0     1    0       0               0   0         0    0     0
## 5    0      0     0    0       0               0   0         0    0     0
## 6    0      0     0    0       0               0   0         0    0     0
##   limit line linked listed listen literally live local located login long
## 1     0    0      0      0      0         0    0     0       0     0    0
## 2     0    0      0      0      0         0    0     0       0     0    0
## 3     0    0      0      0      0         0    0     0       1     1    0
## 4     0    0      0      0      0         0    0     1       0     0    0
## 5     0    0      0      0      0         0    0     1       0     0    0
## 6     0    1      0      0      0         0    0     0       0     0    0
##   longer looked lose loss lost lot louisiana love low lower lud lzmd
## 1      0      0    0    0    0   0         0    0   0     1   0    0
## 2      0      0    0    0    0   0         0    0   0     0   0    0
## 3      0      0    0    0    0   0         0    0   0     0   1    0
## 4      0      0    0    0    0   0         0    0   0     0   0    0
## 5      0      0    0    1    0   0         0    0   0     0   0    0
## 6      0      0    0    0    0   0         0    0   0     0   0    0
##   machinations magazine main mainstream maintain major makes man manage
## 1            0        0    0          0        1     0     0   0      0
## 2            0        0    0          0        0     0     0   0      0
## 3            0        0    0          0        0     0     0   0      0
## 4            0        0    0          0        0     0     0   0      1
## 5            0        0    0          0        0     0     0   0      0
## 6            0        0    0          0        0     0     0   0      0
##   mar march margin marine mark market marriage married mass massachusetts
## 1   0     0      0      0    0      0        0       0    1              0
## 2   0     0      0      0    0      0        0       0    0              0
## 3   0     0      0      0    0      0        0       0    0              0
## 4   0     0      0      0    0      0        0       0    0              0
## 5   0     0      0      0    0      1        0       0    0              0
## 6   0     0      0      0    0      0        0       0    0              0
##   massive match material materiel matter mccain meaning meant measure
## 1       0     0        0        0      0      0       0     0       0
## 2       0     0        0        0      0      0       0     0       0
## 3       0     0        0        0      0      0       0     0       0
## 4       0     0        0        0      0      0       1     0       0
## 5       0     0        0        0      0      0       0     0       0
## 6       0     0        0        0      0      0       0     0       0
##   media medical meet melanie member memo memorial men mention menu mess
## 1     0       0    0       0      1    0        0   0       0    0    0
## 2     0       0    0       0      1    0        0   0       0    0    0
## 3     1       0    0       0      0    0        0   0       0    1    0
## 4     1       0    0       0      1    0        0   0       0    0    0
## 5     3       0    0       0      1    0        0   0       1    0    0
## 6     0       0    0       0      0    0        0   0       0    0    0
##   message met meteor meter mexico mich michael michigan midday middle mike
```

```
## 1       0   0      0     0      0    0       0        0      0      0    0
## 2       0   0      0     0      0    0       1        0      0      0    0
## 3       0   0      0     1      0    0       0        0      1      0    0
## 4       0   0      0     0      0    0       0        0      0      0    0
## 5       0   0      0     0      0    0       0        0      0      0    0
## 6       0   0      0     0      0    0       0        0      0      0    0
##   miles military miller million mind mine minister minnesota minor minute
## 1     0        5      0       0    0    0        0         0     0      0
## 2     0        0      0       0    0    0        0         0     0      0
## 3     0        0      0       0    0    0        0         0     0      0
## 4     0        0      0       0    0    0        0         0     0      0
## 5     0        0      0       0    0    0        0         0     0      0
## 6     0        0      0       0    0    0        0         0     0      0
##   miss mission missouri mistake misterajc mix model moderate mom moment
## 1    0       0        0       1         0   0     0        0   0      0
## 2    0       2        0       0         0   0     0        0   0      1
## 3    0       0        0       0         0   0     0        0   0      1
## 4    0       0        0       0         0   0     0        0   0      0
## 5    0       0        0       0         0   0     0        0   0      0
## 6    0       0        1       0         0   0     0        0   0      0
##   momentum monday money montclair month moore morning morrison mosh mother
## 1        0      0     0         0     0     0       0        0    0      0
## 2        0      0     0         0     0     0       0        0    0      0
## 3        0      0     0         1     0     0       0        0    0      0
## 4        0      0     0         0     0     0       1        0    0      0
## 5        0      0     0         0     1     0       0        0    0      0
## 6        0      0     0         0     0     0       0        0    0      0
##   motivated move movement mrliberal mydd nader nag named nancy narrow
## 1         0    0        0         0    0     0   0     0     0      0
## 2         0    0        0         0    0     0   0     0     0      0
## 3         0    0        0         0    0     1   0     0     0      0
## 4         0    0        0         0    0     0   0     0     0      0
## 5         0    0        0         0    0     0   0     0     0      0
## 6         0    0        0         0    0     0   0     0     0      0
##   nation nationwide natural nbc needed negative neighbor net network
## 1      0          0       0   0      0        0        0   0       0
## 2      0          0       0   0      0        0        0   0       0
## 3      0          0       0   0      1        0        0   0       0
## 4      0          0       0   0      0        0        0   0       0
## 5      0          0       0   0      0        4        0   0       0
## 6      0          0       0   0      0        0        0   0       0
##   nevada news newspaper newsweek newwindow nice night nominate nominee
## 1      0    1         0        0         0    0     0        0       0
## 2      0    0         0        0         0    0     0        0       0
## 3      0    0         0        0         1    0     0        0       0
## 4      0    1         0        0         0    0     0        0       0
## 5      0    0         0        1         0    0     0        0       0
## 6      1    1         0        0         0    0     0        0       0
##   north note notice nov november nprigo nro nuclear number obama object
## 1     0    0      0   0        1      0   0       0      0     0      0
## 2     0    0      0   0        0      0   0       0      0     0      0
## 3     0    0      0   1       10      0   0       0      0     0      0
## 4     0    0      0   0        0      0   0       0      0     0      0
## 5     0    0      0   0        0      0   0       0      0     0      0
```

```
## 6     0    0      0   0        0      0   0       0      0     0      0
##   observation obtain obvious occupation occur oct october offense offer
## 1           0      0       0          0     0   0       0       0     0
## 2           0      0       0          0     0   0       0       0     0
## 3           0      0       0          0     0   0       0       0     0
## 4           0      0       0          0     0   0       0       0     0
## 5           0      0       0          0     0   0       0       0     0
## 6           0      0       0          0     0   0       0       0     0
##   office official ohio oil oklahoma online opened
## 1      0        0    0   0        0      0      0
## 2      0        0    0   0        0      1      0
## 3      0        0    1   0        0      1      0
## 4      0        0    0   0        0      0      0
## 5      1        0    1   0        0      0      0
## 6      0        0    0   0        0      0      0
##   openhttpwwwedwardsforprezcomdailykoshtml operate opinion opponent
## 1                                        0       0       0        0
## 2                                        0       2       0        0
## 3                                        1       0       0        0
## 4                                        0       0       0        2
## 5                                        0       0       0        2
## 6                                        0       0       0        0
##   opportunities oppose opposite optimist option order oregon organization
## 1             2      0        0        0      0     0      0            0
## 2             0      0        0        0      0     0      0            0
## 3             0      0        0        0      0     0      0            0
## 4             0      0        0        0      0     0      0            0
## 5             0      0        0        0      0     0      0            0
## 6             0      0        0        0      0     0      0            0
##   original osama ourcongressorg outrage overwhelming page paid panel paper
## 1        0     0              0       0            0    0    0     0     0
## 2        0     0              0       0            0    0    0     0     0
## 3        0     0              1       0            0    0    0     0     0
## 4        0     0              0       0            0    2    0     0     0
## 5        0     0              0       0            0    0    0     0     0
## 6        0     0              0       0            0    0    0     0     0
##   parecommend parenthesis participants parties partisan partly partner
## 1           0           0            0       0        0      0       0
## 2           0           0            0       0        0      0       0
## 3           0           0            0       0        0      0       0
## 4           0           0            0       0        0      0       0
## 5           0           0            0       0        0      0       0
## 6           0           0            0       0        0      0       0
##   pass password past pat patriot paul pay pdf peace peanut pennsylvania
## 1    0        0    0   0       0    0   0   0     0      0            1
## 2    0        0    0   0       0    0   0   0     0      0            0
## 3    0        1    0   0       0    0   0   0     0      1            0
## 4    0        0    1   0       0    0   0   0     0      0            0
## 5    0        0    0   0       0    0   0   0     0      0            0
## 6    0        0    0   0       0    0   0   4     0      0            0
##   pentagon people percent percentage perfect perform period person
## 1        0      0       4          0       0       0      0      0
## 2        0      0       0          0       0       0      0      0
## 3        0      0       0          0       0       0      0      0
```

```
## 4        0      0       0          0       0       0      0      0
## 5        0      1       4          0       0       0      0      0
## 6        0      0       0          0       0       0      0      0
##   personnel perspective philly phone photo pick pickup picture piece place
## 1         0           0      0     0     0    1      1       0     0     0
## 2         0           0      0     0     0    0      0       0     0     1
## 3         0           0      1     0     0    0      0       0     0     0
## 4         0           0      0     0     0    0      0       0     0     0
## 5         0           0      0     0     0    0      0       0     0     0
## 6         0           0      0     0     0    0      0       0     0     0
##   plan plane play plenty pointed police policies political politician poll
## 1    0     0    0      0       0      0        0         0          0    0
## 2    0     0    2      0       0      0        0         0          0    0
## 3    1     0    0      0       0      0        0         0          0    4
## 4    0     0    0      0       0      3        0         0          0    0
## 5    0     0    0      0       0      0        0         0          0    2
## 6    0     0    0      0       0      0        0         0          0    0
##   pollster poor popular population position possibilities posted potential
## 1        0    0       0          0        0             0      0         0
## 2        0    0       0          0        0             0      0         0
## 3        0    0       0          0        1             0      0         0
## 4        0    0       0          0        0             0      0         0
## 5        0    0       0          0        2             0      0         0
## 6        0    0       0          0        0             0      0         0
##   powell power practical predict prefer preparation presence present
## 1      0     0         0       0      0           0        0       0
## 2      0     0         0       0      0           0        0       0
## 3      0     1         0       0      0           0        0       0
## 4      0     1         0       0      0           0        0       0
## 5      0     0         0       0      0           0        0       0
## 6      0     0         0       0      0           0        0       0
##   presided presidential press pressure pretend pretty prevent previous
## 1        1            0     0        0       0      0       0        0
## 2        0            0     0        0       0      0       0        0
## 3        0            0     0        0       0      0       0        0
## 4        0            0     1        1       0      1       0        0
## 5        7            1     0        0       0      1       0        0
## 6        0            0     0        0       0      0       0        0
##   price pride primaries prime prior priorities prison private problem
## 1     0     0         0     0     0          0      0       0       0
## 2     0     0         0     0     0          0      0       0       0
## 3     0     0         1     0     0          0      0       0       0
## 4     0     0         0     0     0          0      0       2       0
## 5     0     0         0     0     0          0      0       0       0
## 6     0     0         0     0     0          0      0       0       0
##   process produce product profile program progress project promise promote
## 1       0       0       0       0       0        0       0       0       0
## 2       0       0       0       0       0        0       0       0       0
## 3       0       0       0       0       0        0       0       0       0
## 4       0       0       0       0       0        0       0       0       0
## 5       0       0       0       0       0        0       0       0       0
## 6       0       0       0       0       0        0       0       0       0
##   proof proper proposal prospect protect protest proud prove provide
## 1     0      0        0        0       0       0     0     1       0
```

```
## 2     0      0        0        0       0       0     0     0       0
## 3     0      0        0        0       0       1     1     0       0
## 4     0      0        0        0       0       0     0     1       0
## 5     0      0        0        0       0       0     0     0       0
## 6     0      0        0        0       0       0     0     0       0
##   public publish pull pundit punkmonk purpose push put qaeda qaqaa quarter
## 1      0       0    0      0        0       0    0   0     0     0       0
## 2      0       0    0      0        0       0    0   0     0     0       0
## 3      0       0    0      0        0       0    0   0     0     0       0
## 4      3       0    3      0        0       0    0   0     0     0       0
## 5      0       0    1      0        0       0    0   0     0     0       0
## 6      0       0    0      0        0       0    0   0     0     0       0
##   question quick quiet quote race racine rad radio raise rallies ralph ran
## 1        0     0     0     0    0      0   0     0     0       0     0   0
## 2        0     0     0     0    0      0   0     0     0       0     0   0
## 3        0     0     0     0    2      1   1     0     0       0     0   0
## 4        1     0     0     0    2      0   0     0     0       0     0   0
## 5        0     0     0     0    1      0   0     0     0       0     0   0
## 6        0     0     0     0    0      0   0     0     0       0     0   0
##   range rank rapid rare rasmussen rate reach reaction read reader ready
## 1     0    0     0    0         0    0     1        0    0      0     0
## 2     0    0     0    0         0    0     0        0    0      0     0
## 3     0    0     0    0         0    0     0        0    0      0     0
## 4     0    0     0    0         0    0     0        0    0      0     0
## 5     0    0     0    0         0    1     0        0    0      0     0
## 6     0    0     0    0         0    0     0        0    0      0     0
##   reagan real realities realization reason receive recently recognize
## 1      0    1         0           0      0       0        0         0
## 2      0    0         0           0      0       0        0         0
## 3      0    0         0           0      0       0        0         0
## 4      0    0         0           0      0       0        0         0
## 5      0    0         0           0      0       0        0         0
## 6      0    0         0           0      0       0        0         0
##   recommend record recruit red reduce reelect refer reflect reform refusal
## 1         0      0       0   0      0       0     0       0      0       0
## 2         0      0       0   0      0       0     0       0      0       0
## 3         1      0       0   0      0       0     0       0      0       0
## 4         0      0       0   0      0       0     0       0      0       0
## 5         0      0       0   0      0       1     0       0      0       0
## 6         0      0       0   0      0       0     0       0      0       0
##   regime region register registrants regular reject related release
## 1      0      0        0           0       0      0       1       0
## 2      0      0        0           0       0      0       0       0
## 3      0      0        0           0       0      0       0       0
## 4      0      0        0           0       0      0       0       0
## 5      0      0        1           0       0      0       0       0
## 6      0      0        0           0       0      0       0       0
##   relevance religious remain remark remember remind remove rep repeat.
## 1         0         0      0      0        0      0      0   1       0
## 2         0         0      0      0        0      0      0   0       0
## 3         0         0      0      0        0      0      0   0       0
## 4         0         0      1      0        0      0      1   0       0
## 5         0         0      1      0        0      0      0   0       0
## 6         0         0      0      0        0      0      0   0       0
```

```
##   replace report represent repub republican republicansforkerry request
## 1       0      0         0     0          2                    0       0
## 2       0      0         0     0          0                    0       1
## 3       0      1         0     1          2                    1       0
## 4       0      3         0     0          0                    0       0
## 5       0      0         0     0          1                    0       0
## 6       0      0         0     0          0                    0       0
##   require research reserve resign resist resource respect respond response
## 1       0        0       0      0      0        0       0       0        0
## 2       0        0       0      0      0        0       0       1        0
## 3       0        0       0      0      0        0       0       0        0
## 4       0        0       0      0      0        0       0       0        0
## 5       0        0       0      0      0        0       0       0        0
## 6       0        0       0      0      0        0       0       0        0
##   rest result retire return reuters reveal reverse review rhetoric rice
## 1    0      1      0      0       0      1       0      0        0    0
## 2    0      0      0      0       0      0       0      0        0    0
## 3    0      0      0      0       0      0       0      0        0    0
## 4    0      0      0      0       0      0       0      0        0    0
## 5    0      0      0      0       0      0       0      0        0    0
## 6    0      3      0      0       0      0       0      0        0    0
##   richard ridiculous rightly rightwing rise risk rival rnc road robert
## 1       1          0       0         0    0    0     0   0    0      0
## 2       0          0       0         0    0    0     0   0    0      0
## 3       0          0       0         0    0    0     0   0    0      0
## 4       0          0       0         0    0    0     0   0    0      0
## 5       0          0       0         0    0    0     0   0    0      0
## 6       0          0       0         0    0    0     0   0    0      0
##   role roll room root rough round rove rule rumor rumsfeld running saddam
## 1    0    0    0    0     0     0    0    0     0        0       0      1
## 2    0    0    0    0     0     0    0    0     0        0       0      0
## 3    0    0    0    0     0     0    0    0     0        0       0      0
## 4    0    0    0    0     0     0    0    0     0        0       0      0
## 5    0    0    0    0     0     0    0    0     0        0       0      0
## 6    0    0    0    0     0     0    0    0     0        0       0      0
##   safe salazar sample san sappy saturday save scandal scare scene schaller
## 1    0       0      0   0     0        0    0       0     0     0        0
## 2    0       0      0   0     0        0    0       0     0     0        0
## 3    0       0      0   0     0        0    0       0     0     0        0
## 4    0       0      0   0     0        0    0       0     0     0        0
## 5    0       0      0   0     0        0    0       0     0     0        0
## 6    0       1      0   0     0        0    0       0     0     0        0
##   schedule school science scientist scoop score scott screen seamus search
## 1        0      0       0         0     0     0     0      0      0      0
## 2        0      0       0         0     0     0     0      0      0      0
## 3        0      0       0         0     1     0     0      0      1      0
## 4        0      0       0         0     0     0     0      0      0      0
## 5        0      0       0         0     0     0     0      0      0      0
## 6        0      0       0         0     0     0     0      0      0      0
##   season seat secret secretaries section secure seek select sen senate
## 1      0    1      0           0       0      0    0      0   0      0
## 2      0    0      0           2       0      0    0      0   0      0
## 3      0    0      0           0       1      0    0      1   0      2
## 4      0    0      0           0       0      0    0      0   0      0
```

```
## 5      0    0      0           0       0      0    0      0   0      0
## 6      0    0      0           0       0      0    0      0   0      3
##   senategovernors send senior sense sep sept september series serve server
## 1               0    0      1     0   0    0         0      0     0      0
## 2               0    0      0     0   0    0         0      0     0      0
## 3               1    0      0     0   0    0         0      0     0      0
## 4               0    0      0     0   0    0         0      0     0      0
## 5               0    0      0     0   0    0         0      0     0      0
## 6               0    0      0     0   0    0         0      0     0      0
##   service session set shape share sharpton shes shift shit shock short
## 1       0       0   0     0     0        0    0     0    0     0     0
## 2       1       0   0     0     0        0    0     0    0     0     0
## 3       0       0   0     0     0        0    0     0    0     0     0
## 4       0      15   0     0     0        0    0     0    0     0     0
## 5       0       0   0     0     0        0    0     1    0     0     0
## 6       0       0   0     0     0        0    0     0    0     0     0
##   shot shouldnt showed shown side sign signature significance similar
## 1    0        0      0     0    0    0         0            0       0
## 2    0        0      0     0    0    0         0            0       0
## 3    0        0      0     0    0    0         0            0       0
## 4    0        0      1     0    0   16         0            0       0
## 5    0        0      0     0    0    2         0            0       0
## 6    0        0      0     0    0    0         0            0       0
##   simple simply single sit sites situation slate slight slip slow small
## 1      0      0      0   0     0         0     0      0    0    0     0
## 2      0      0      0   0     0         0     0      0    0    0     0
## 3      0      0      0   0     0         0     0      0    0    0     0
## 4      0      0      0   1     0         0     0      0    0    0     0
## 5      0      0      0   1     0         1     0      0    0    0     0
## 6      0      0      0   0     0         0     0      0    0    0     0
##   smear social soldier solid son sort sound source south southern space
## 1     0      0       0     0   0    0     0      0     0        0     0
## 2     0      0       0     0   0    0     0      0     0        0     0
## 3     0      0       1     0   0    0     0      0     0        0     0
## 4     0      0       0     0   0    0     1      0     0        0     0
## 5     0      0       0     0   0    0     1      0     0        0     0
## 6     0      0       0     0   0    0     0      0     0        0     0
##   speak speaker special specific specter speculate speech spend spent spin
## 1     0       0       0        0       0         0      0     0     0    0
## 2     0       0       2        0       0         0      0     0     0    0
## 3     0       0       0        0       0         0      0     0     0    0
## 4     0       0       0        0       0         0      0     0     0    0
## 5     0       0       0        0       0         0      0     0     0    0
## 6     0       0       0        0       0         0      0     0     0    0
##   split spoke spokesman spot spread staff staffer stage stand standard
## 1     0     0         0    0      0     0       0     0     0        0
## 2     0     0         1    0      0     0       0     0     0        0
## 3     1     0         0    0      0     0       0     0     0        0
## 4     0     0         0    0      0     0       0     0     0        0
## 5     0     0         0    0      0     0       0     0     0        0
## 6     0     0         0    0      0     0       0     0     0        0
##   star stark start startspan state statement statewide station statistical
## 1    0     0     0         0     0         0         0       0           0
## 2    0     0     0         0     0         2         0       0           0
```

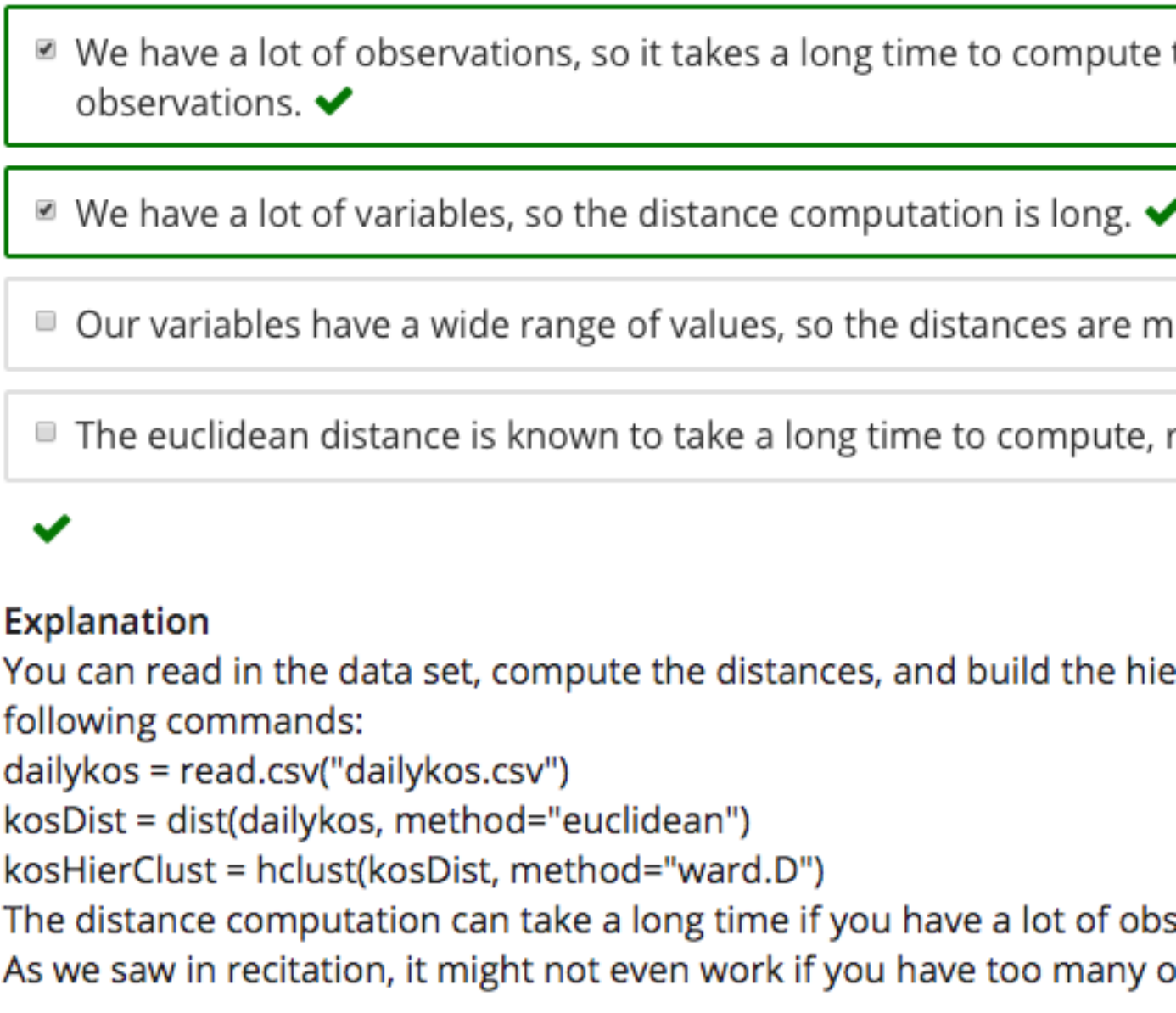
```
## 3    0     0     0         1     0         0         0       0           0
## 4    0     0     0         0     0         0         0       0           0
## 5    0     0     1         0     0         0         0       0           0
## 6    0     0     0         0     0         0         0       0           0
##   status stay steal step steve stick stock stolen stop store stories
## 1      0    0     0    0     0     0     0      0    0     0       0
## 2      0    0     0    0     0     0     0      0    0     0       0
## 3      0    1     1    0     0     0     0      1    0     1       0
## 4      0    0     2    0     0     1     0      1    0     0       0
## 5      0    1     0    1     0     1     0      1    0     0       0
## 6      0    0     0    0     0     0     0      0    0     0       0
##   straight strategic strategies strategist street strength strike strong
## 1        0         0          0          0      0        0      0      0
## 2        0         0          0          0      0        0      0      0
## 3        1         0          0          0      0        0      0      0
## 4        1         0          0          0      1        0      0      0
## 5        0         0          1          0      1        0      0      0
## 6        0         0          0          0      0        0      0      0
##   stronger struck student studies stuff stupid subject success sudden
## 1        0      0       0       0     0      0       0       0      0
## 2        0      0       0       0     0      0       0       0      0
## 3        0      0       0       0     0      0       0       0      0
## 4        0      0       0       0     0      0       0       0      0
## 5        1      0       0       0     0      0       0       0      0
## 6        0      0       0       0     0      0       0       0      0
##   suffer suggest suit summer sun sunday sunzoo support suppose suppress
## 1      0       0    0      0   0      0      0       0       0        0
## 2      0       0    0      0   0      0      0       0       0        0
## 3      0       0    0      0   0      0      2       0       0        1
## 4      0       0    0      0   0      0      0       0       0        0
## 5      0       0    0      0   0      0      0       0       0        0
## 6      0       0    0      0   0      0      0       0       0        0
##   supreme surely surge surprise survey surveyusa survival susa suspect
## 1       0      0     0        0      0         0        0    0       0
## 2       0      0     0        0      0         0        0    0       0
## 3       0      0     0        0      0         0        0    0       0
## 4       0      0     0        0      0         0        0    0       0
## 5       0      0     0        0      0         0        0    0       0
## 6       0      0     0        0      0         4        0    1       0
##   swift swing switch system table tactic takes talked tank target task tax
## 1     0     0      0      0     0      0     0      0    0      0    0   0
## 2     1     0      0      0     0      0     0      0    0      0    0   0
## 3     0     0      0      0     0      0     0      0    0      0    0   0
## 4     0     0      0      0     0      0     0      0    0      0    0   0
## 5     0     0      0      0     0      0     0      0    0      0    0   0
## 6     0     0      0      0     0      0     0      0    0      0    0   0
##   team televised telling ten term territory terror terrorist test
## 1    0         0       0   0    0         0      0         0    0
## 2    0         0       0   0    0         0      0         0    0
## 3    0         0       0   0    0         0      0         0    0
## 4    0         0       0   0    0         0      0         0    0
## 5    0         0       0   0    0         0      1         0    0
## 6    0         0       0   0    0         0      0         0    0
##   testimony texas theme theories theyll theyre theyve thinking thought
```

```
## 1         0     0     0        0      0      0      0        0       0
## 2         0     0     0        0      0      0      0        0       0
## 3         0     0     0        0      0      0      0        0       0
## 4         0     0     0        0      0      0      0        0       0
## 5         0     0     1        0      0      0      0        0       0
## 6         0     0     0        0      0      0      0        0       0
##   thousand threat threaten throw thursday ticket tie tight time tip todays
## 1        0      0        0     0        0      0   0     0    0   0      0
## 2        0      0        0     0        0      0   0     0    1   0      0
## 3        0      0        0     0        0      1   0     0    0   0      0
## 4        0      0        0     0        0      0   0     0    3   0      0
## 5        0      0        0     0        0      0   0     0    0   0      0
## 6        0      0        0     0        0      0   0     0    0   0      0
##   tom tomorrow tonight tony tool top topic total tough town track trade
## 1   0        0       0    0    0   0     0     0     0    0     0     0
## 2   0        0       0    0    0   0     0     0     0    0     0     0
## 3   0        0       0    0    0   0     0     0     0    0     0     0
## 4   0        0       0    0    0   0     0     0     0    0     0     0
## 5   0        0       0    0    0   1     0     2     0    0     0     0
## 6   0        0       0    0    0   0     0     0     0    0     0     0
##   tradesports tradition traffic trail train transfer trapper travel treat
## 1           0         0       0     0     0        0       0      0     0
## 2           0         0       0     0     0        0       0      0     0
## 3           0         0       0     0     0        0       0      0     0
## 4           0         0       0     0     0        0       0      0     0
## 5           0         0       0     0     0        0       0      0     0
## 6           0         0       0     0     0        0       0      0     0
##   treatment trend trial trip troop trouble TRUE. trust truth tuesday
## 1         0     0     0    0     0       0     1     0     0       0
## 2         0     0     0    0     0       0     0     0     1       0
## 3         0     0     0    0     0       0     0     0     0       0
## 4         0     0     0    0     0       1     0     0     0       0
## 5         0     0     0    0     0       1     0     0     0       0
## 6         0     1     0    0     0       0     0     0     0       0
##   tunesmith turn turnout type typical ultimate undecided understand union
## 1         0    0       0    0       0        0         0          0     0
## 2         0    0       0    0       0        0         0          0     0
## 3         0    0       0    0       0        0         0          0     0
## 4         0    0       0    1       0        0         0          0     0
## 5         0    0       0    1       0        0         0          0     0
## 6         0    0       0    0       0        0         0          0     0
##   unit universal unlike updated upenn urge usa username usual
## 1    0         0      0       0     0    0   0        0     0
## 2    0         0      0       0     0    0   0        0     0
## 3    0         0      0       1     1    0   0        1     0
## 4    0         0      0       0     0    0   0        0     0
## 5    0         0      0       0     0    0   0        0     0
## 6    0         0      0       0     0    0   0        0     0
##   vaantirepublican values var vast verification version veteran vice
## 1                0      0   0    0            0       0       0    1
## 2                0      0   0    0            0       0       2    0
## 3                0      0   1    0            0       0       0    0
## 4                0      0   0    0            0       0       0    0
## 5                0      0   0    0            0       0       0    0
```

```
## 6               0      0   0    0           0       0       0    0
##   victim victories video vietnam view violate violence virginia
## 1      0         0     0       0    1       0        0        0
## 2      0         0     0       2    0       0        0        0
## 3      0         1     0       0    0       0        0        0
## 4      0         0     0       0    0       0        0        0
## 5      0         0     0       0    0       0        0        0
## 6      0         0     0       0    0       0        0        0
##   virginiadem virtual visit voice volunteer vote voter vulnerable wage
## 1           0       0     0     0         0    1     2          0    0
## 2           0       0     0     0         0    0     0          0    0
## 3           0       0     0     0         0    8     2          0    0
## 4           0       0     0     0         0    0     0          0    0
## 5           0       0     0     0         0    0     1          0    1
## 6           0       0     0     0         0    0     0          0    0
##   wait walk wall wanted wapo war warn washington wasnt waste watch
## 1    0    0    0      0    0   4    0          0     0     0     0
## 2    0    0    0      0    0   2    0          0     1     0     2
## 3    1    0    0      0    0   1    0          0     0     0     2
## 4    0    0    0      0    0   0    0          0     1     0     0
## 5    0    0    0      1    0   1    0          1     0     0     1
## 6    0    0    0      0    0   0    0          1     0     0     0
##   watchers water ways wclathe weak weapon web website wednesday weekend
## 1        0     0    0       0    0      1   0       0         0       0
## 2        0     0    0       0    0      0   0       0         0       0
## 3        1     0    0       0    0      0   0       0         0       0
## 4        0     0    0       0    0      0   0       0         0       0
## 5        0     0    0       0    0      0   0       0         0       0
## 6        0     0    0       0    0      0   0       0         0       0
##   weekly werent west weve whats white whos wide wife william win wingnut
## 1      0      0    0    0     0     0    0    0    0       0   0       0
## 2      0      0    0    0     0     0    0    0    0       0   0       0
## 3      0      0    0    0     0     0    0    0    0       0   1       0
## 4      0      0    0    0     0     0    0    0    0       0   0       0
## 5      1      0    0    0     0     0    0    0    0       0   0       0
## 6      0      0    0    0     0     0    0    0    0       0   0       0
##   winner wisconsin witness woes woman women won wonderful wont word worked
## 1      0         0       0    0     0     0   0         0    0    0      0
## 2      0         0       0    0     0     0   0         0    0    0      0
## 3      0         0       0    0     0     0   0         0    0    0      0
## 4      0         0       0    0     0     0   0         0    0    0      0
## 5      0         0       0    0     0     0   0         0    0    0      0
## 6      0         0       0    0     0     0   0         0    0    0      0
##   worker worried worse worst worth wouldnt wouldve wound write writer
## 1      0       0     0     0     0       0       0     0     0      0
## 2      0       0     0     0     0       0       0     0     0      0
## 3      0       0     0     0     0       0       0     0     0      0
## 4      0       0     0     0     0       0       0     0     0      0
## 5      0       0     0     0     0       0       0     0     0      0
## 6      0       0     0     0     0       0       0     0     0      0
##   written wrong wrote yahoo yeah year yesterday york youll young youre
## 1       0     2     1     0    0    0         0    0     0     0     0
## 2       0     1     0     0    0    0         0    0     0     0     0
## 3       0     0     0     0    0    0         0    0     0     0     0
```

```
## 4       0     0     0     0     0     0         2     0     0     1     0
## 5       0     0     0     0     0     1         1     0     0     1     0
## 6       0     0     0     0     0     0         0     0     0     0     0
##   youve zogby zone
## 1     0     0    1
## 2     0     0    0
## 3     0     0    0
## 4     0     0    0
## 5     0     0    0
## 6     0     0    0
```

```
include_graphics('1.1.png')
```



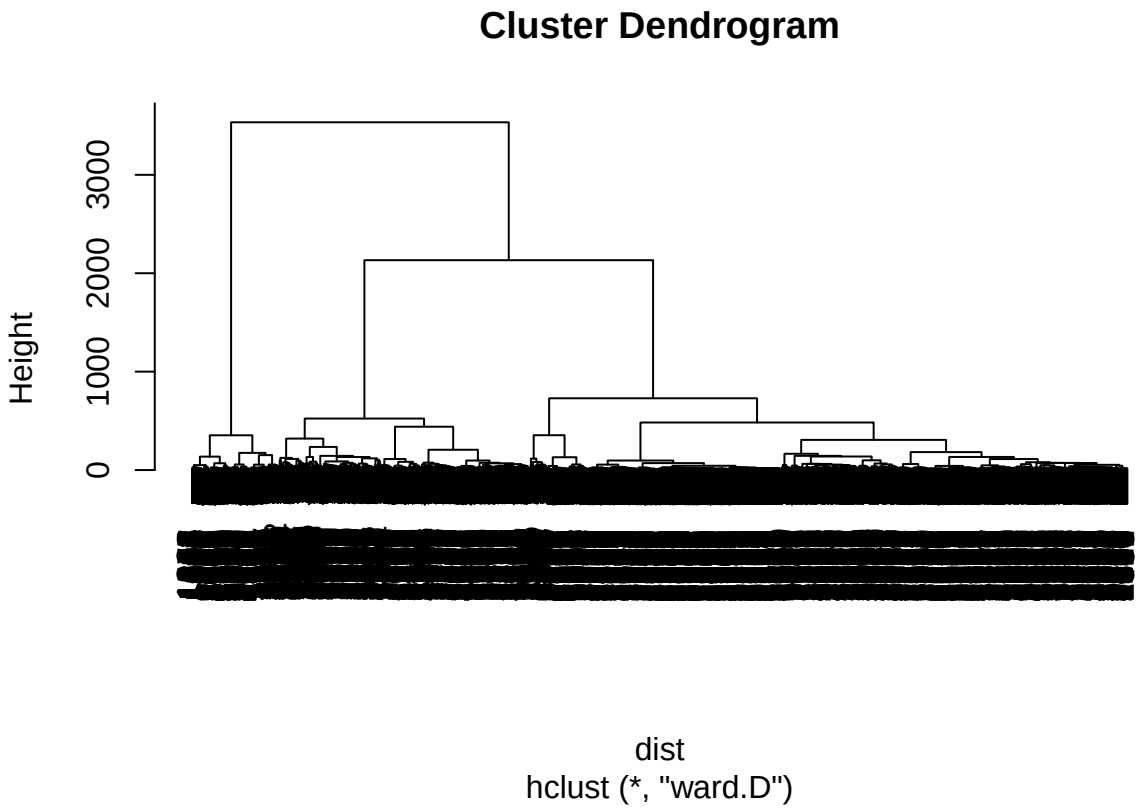

```
dist = dist(daily_kos, method = 'euclidian')

cluster = hclust(dist, method = 'ward.D')
```

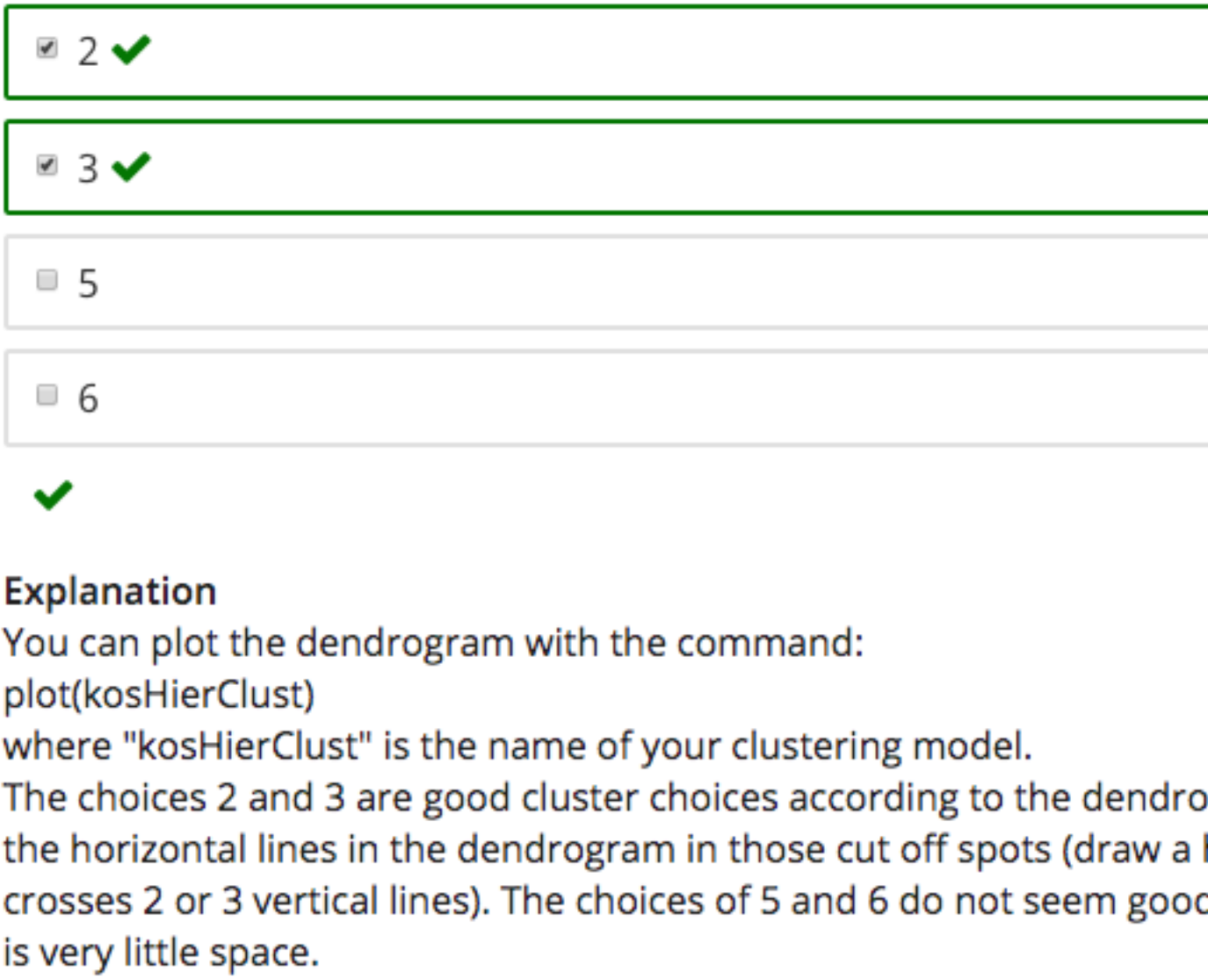
1.2) Hierarchical Clustering

Plot the dendrogram of your hierarchical clustering model. Just looking at the dendrogram, which of the following seem like good choices for the number of clusters? Select all that apply.

```
plot(cluster)
```



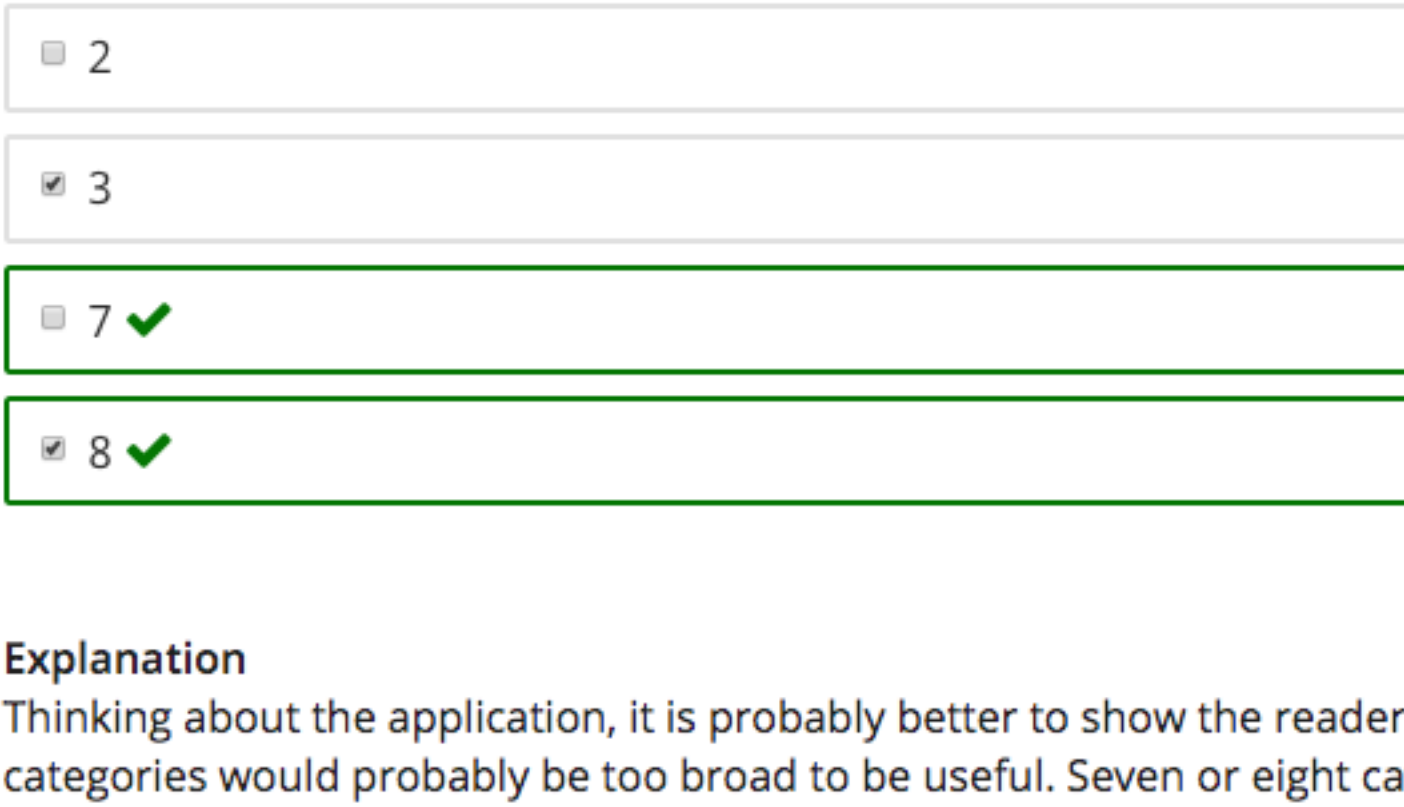

```
include_graphics('1.2.png')
```



1.3) Hierarchical Clustering

In this problem, we are trying to cluster news articles or blog posts into groups. This can be used to show readers categories to choose from when trying to decide what to read. Just thinking about this application, what are good choices for the number of clusters? Select all that apply.

```
include_graphics('1.3.png')
```



1.4) Hierarchical Clustering

Let's pick 7 clusters. This number is reasonable according to the dendrogram, and also seems reasonable for the application. Use the cutree function to split your data into 7 clusters.

Now, we don't really want to run tapply on every single variable when we have over 1,000 different variables. Let's instead use the subset function to subset our data by cluster. Create 7 new datasets, each containing the observations from one of the clusters.

How many observations are in cluster 3?

```
clusters_no = cutree(cluster, 7)
```

```
table(clusters_no)
```

```
## clusters_no
##    1    2    3    4    5    6    7
## 1266  321  374  139  407  714  209
```

```
clusters_no
```

```
##    [1] 1 1 2 3 4 1 5 6 1 6 6 2 2 6 1 3 1 1 1 1 6 6 7 2 6 5 5 1 6 1 1 3 6 6
##   [35] 2 1 6 5 3 1 5 1 1 6 2 1 6 1 1 4 4 6 6 3 1 1 1 2 3 1 1 5 6 6 1 5 6 5
##   [69] 2 1 1 1 1 6 1 4 4 3 1 2 1 1 1 1 5 3 1 5 6 2 6 1 4 1 1 6 1 6 1 2 1 6
##  [103] 1 4 4 5 1 1 6 6 6 6 2 1 1 1 1 1 1 1 1 1 1 6 6 2 1 1 2 1 1 3 1 1 1 1
##  [137] 5 1 1 6 2 5 6 1 4 1 1 2 6 6 2 1 1 1 1 5 5 6 1 4 1 1 2 1 1 5 1 6 6 4
##  [171] 3 1 5 1 2 1 6 6 1 1 1 1 3 1 4 5 2 6 6 3 6 5 1 2 3 6 3 3 6 5 3 3 2 3
##  [205] 6 5 4 6 6 1 6 2 6 1 6 6 6 6 6 2 1 3 6 6 5 5 5 1 6 3 2 3 2 1 1 1 1 1
##  [239] 4 1 1 2 6 6 3 5 1 5 6 6 3 2 3 6 6 5 6 1 6 1 4 2 1 3 4 6 4 6 3 5 6 2
```

```
##  [273] 3 1 1 1 1 6 6 1 2 1 3 1 1 4 1 3 6 6 5 2 4 1 3 1 6 6 6 1 1 2 6 3 6 1
##  [307] 6 6 6 1 2 6 6 6 1 5 1 6 1 1 2 6 6 1 3 6 1 1 6 4 2 5 1 2 6 3 1 1 1 1
##  [341] 1 6 2 1 1 3 1 6 6 6 5 1 2 1 1 1 1 6 1 1 6 4 3 1 2 3 1 6 1 1 1 6 1 3
##  [375] 7 1 2 3 1 4 1 1 1 1 1 1 6 2 1 3 1 1 3 6 3 1 1 2 5 1 3 1 6 1 4 5 5 1
##  [409] 2 1 3 3 1 4 1 6 1 5 6 6 6 2 1 5 1 6 1 1 1 1 1 1 1 3 1 2 6 5 5 6 6 6
##  [443] 3 4 6 2 6 6 2 1 1 6 7 6 7 1 5 2 6 6 1 7 1 6 1 5 1 1 2 1 1 1 6 1 1 6
##  [477] 1 1 1 4 2 4 6 1 1 1 5 4 1 1 6 1 2 3 3 1 5 1 4 3 5 6 1 1 2 1 4 3 1 5
##  [511] 1 1 1 6 6 2 1 6 1 1 4 1 6 1 6 1 6 5 2 4 4 1 1 5 5 6 6 1 1 2 6 3 6 1
##  [545] 5 1 6 1 4 1 6 1 2 4 1 3 3 1 1 4 5 1 1 1 1 2 1 5 2 6 1 1 1 1 3 6 3 1
##  [579] 6 2 7 1 1 6 3 1 6 6 6 1 1 1 2 6 3 1 3 3 6 1 3 1 1 4 1 2 1 5 1 1 1 4
##  [613] 1 1 1 6 1 5 2 3 1 5 5 6 1 1 1 1 2 5 5 4 6 1 1 1 1 5 1 1 6 2 1 1 5 1
##  [647] 1 6 6 1 2 6 7 6 3 1 3 1 6 1 1 2 6 6 3 5 5 1 5 3 1 1 4 2 5 1 6 6 3 1
##  [681] 1 6 3 6 2 3 2 1 6 1 1 6 1 1 6 5 3 1 2 3 6 6 1 1 1 3 1 1 1 2 6 6 6 4
##  [715] 6 1 3 1 4 3 2 3 5 3 1 6 6 6 6 6 6 6 2 1 3 6 6 5 1 5 1 1 6 2 1 6 1 1
##  [749] 1 4 1 1 5 6 3 2 1 1 4 6 1 1 4 1 1 1 2 1 3 6 3 1 1 6 3 1 1 2 1 1 1 5
##  [783] 1 3 1 6 1 1 1 1 4 2 6 6 6 1 1 1 1 1 2 6 3 2 3 1 6 6 1 3 1 3 1 2 6 1
##  [817] 1 6 6 6 5 1 1 1 2 1 1 1 1 6 1 5 6 1 2 1 3 5 6 6 6 1 6 6 5 4 2 1 1 5
##  [851] 1 6 5 1 3 1 6 2 6 6 4 3 1 4 1 1 1 2 1 6 1 6 6 6 3 3 1 1 5 2 3 6 1 1
##  [885] 6 1 1 1 6 2 6 3 3 5 6 5 3 1 1 5 1 2 1 1 6 6 6 3 5 1 1 1 2 5 1 2 1 6
##  [919] 1 6 6 1 1 1 1 3 2 1 6 3 5 1 3 1 1 1 2 1 3 3 6 5 6 6 1 2 1 6 6 1 6 5
##  [953] 6 6 1 1 2 1 1 1 1 5 1 1 1 4 6 1 5 2 1 1 3 1 1 6 5 6 5 2 5 1 6 5 5 1
##  [987] 1 5 6 1 5 1 6 5 1 4 1 5 5 1 1 2 1 1 1 1 3 5 5 3 1 2 1 1 4 6 1 1 1 6
## [1021] 1 4 1 2 1 5 2 5 6 1 1 1 6 5 6 6 3 6 3 5 6 6 5 1 1 5 3 1 1 2 5 1 1 1
## [1055] 5 3 1 1 1 1 2 1 1 6 1 1 5 5 6 1 1 5 2 6 6 1 1 3 1 5 1 4 6 6 2 3 3 1
## [1089] 1 1 1 1 1 1 6 2 6 1 1 1 1 5 1 1 4 1 3 1 2 6 5 6 6 6 5 5 1 1 1 3 2 5
## [1123] 5 5 5 5 4 5 5 1 1 1 2 5 5 1 1 1 4 6 5 1 2 6 1 2 6 4 4 3 6 1 1 1 6 6
## [1157] 2 6 5 6 1 5 5 4 1 4 1 5 1 2 1 6 1 5 5 6 3 4 5 1 3 2 1 5 1 3 1 5 1 1
## [1191] 1 1 1 5 2 6 3 1 6 1 5 4 5 6 1 5 2 5 5 1 5 3 1 1 4 1 6 1 2 1 5 3 5 1
## [1225] 5 1 1 5 2 3 1 6 1 3 5 5 1 1 6 4 2 5 1 5 1 5 6 1 1 1 5 2 1 1 1 5 3 1
## [1259] 3 1 3 3 1 2 6 3 2 5 1 6 5 3 5 1 1 6 5 1 2 6 6 5 5 5 6 1 1 4 2 1 5 2
## [1293] 1 5 5 1 1 6 5 1 1 5 5 2 1 5 4 1 5 1 5 1 6 5 6 3 6 1 5 5 1 6 5 6 1 3
## [1327] 2 5 6 1 3 4 5 6 3 3 1 6 2 1 3 3 1 6 1 1 1 1 1 1 6 2 1 6 6 1 6 5 5 5
## [1361] 1 6 6 2 1 1 1 6 3 1 1 5 6 6 1 2 5 3 3 1 6 5 1 3 3 1 1 2 1 2 1 3 1 5
## [1395] 1 3 1 6 3 1 2 1 1 5 5 3 1 3 1 1 5 2 5 3 3 1 3 6 5 1 5 2 6 6 5 5 3 5
## [1429] 1 1 1 2 1 4 6 1 6 3 1 5 3 6 1 1 2 1 1 1 5 4 1 6 1 1 6 6 6 2 1 1 5 1
## [1463] 1 1 6 6 4 6 5 4 2 1 1 4 4 1 1 4 1 1 1 5 2 1 1 6 5 5 3 5 3 1 1 1 2 3
## [1497] 5 6 5 1 5 4 3 1 1 2 5 1 2 1 1 6 6 6 4 1 3 3 4 1 2 1 1 4 7 4 5 6 3 1
## [1531] 1 3 2 3 3 5 6 5 1 1 3 3 5 1 6 2 6 5 3 1 6 1 1 1 6 1 2 1 1 1 1 1 1 7
## [1565] 3 2 1 7 1 5 1 3 1 1 3 6 7 1 1 1 3 4 6 6 2 3 3 3 3 3 1 1 2 3 1 3 6 1
## [1599] 1 5 6 3 3 2 5 6 4 7 1 6 3 1 1 2 2 1 1 6 4 1 3 6 7 4 1 6 2 3 6 3 3 3
## [1633] 6 3 6 6 2 5 3 7 5 3 3 7 1 2 7 7 6 6 6 7 7 6 7 3 2 6 6 3 7 3 6 6 5 3
## [1667] 7 6 2 1 6 4 5 5 6 3 1 5 3 2 6 3 3 6 1 1 6 6 6 6 6 6 2 7 1 1 1 7 1 6
## [1701] 4 6 2 1 3 3 6 3 3 3 1 1 1 6 3 3 2 7 1 6 6 1 1 6 3 1 1 6 2 1 2 7 1 6
## [1735] 1 5 5 1 3 2 6 6 5 1 1 5 6 3 5 6 1 2 7 6 7 6 5 1 1 6 6 5 2 1 7 1 1 3
## [1769] 6 3 7 2 5 3 1 1 3 5 7 6 2 7 3 7 7 6 6 1 6 7 6 1 7 3 6 3 7 3 7 5 6 6
## [1803] 2 4 3 6 7 7 7 5 5 3 2 3 1 1 3 7 6 7 5 1 6 7 2 1 7 7 7 7 6 3 7 1 6 7
## [1837] 2 1 7 2 6 6 6 3 7 1 7 1 6 2 6 6 6 6 6 1 1 1 3 6 7 1 2 6 7 1 6 7 6 6
## [1871] 6 5 2 6 1 1 7 7 7 7 7 1 7 1 2 5 5 6 6 6 6 1 6 7 2 4 6 7 1 6 6 7 3 6
## [1905] 7 1 2 1 1 6 5 1 6 7 6 7 7 2 1 6 7 3 5 5 5 3 7 6 2 1 7 6 7 1 7 1 1 7
## [1939] 1 1 3 7 2 6 6 7 6 3 7 6 1 1 7 7 3 6 6 2 7 1 1 7 1 1 1 1 6 7 2 1 4 7
## [1973] 7 7 6 6 1 1 6 6 2 7 3 6 1 7 7 1 7 7 1 1 3 2 1 3 6 7 7 7 3 6 1 3 7 2
## [2007] 1 7 5 6 3 7 7 7 2 3 1 7 1 6 7 7 1 7 1 2 7 7 6 6 1 7 1 6 7 7 2 3 3 1
## [2041] 1 7 7 7 6 1 2 5 5 5 5 7 6 7 7 1 7 7 2 1 3 1 1 7 1 1 1 3 5 2 1 3 2 7
## [2075] 1 5 3 1 1 3 6 5 7 1 2 7 7 6 6 3 6 6 1 1 5 1 2 6 1 7 1 1 1 6 1 1 2 7
```

```
## [2109] 3 5 5 1 5 1 1 6 5 5 2 7 5 3 7 7 7 7 3 1 1 1 2 1 3 1 5 6 1 1 1 2 6 1
## [2143] 3 7 7 5 7 1 3 2 1 1 1 1 3 5 6 7 5 2 1 1 6 3 1 1 7 1 7 4 1 3 6 2 1 1
## [2177] 6 6 1 6 4 3 1 5 1 2 1 2 1 7 1 1 7 3 5 1 7 7 2 1 7 6 7 1 3 7 4 1 3 6
## [2211] 2 7 1 7 6 6 1 1 4 1 2 1 3 7 1 6 5 1 1 7 6 1 6 2 5 1 5 7 1 1 7 7 5 3
## [2245] 1 2 7 3 7 1 6 5 5 4 6 1 2 5 7 3 7 7 6 1 3 7 2 7 3 7 1 3 5 7 7 1 2 1
## [2279] 3 6 1 1 3 5 6 1 6 3 2 7 6 5 3 7 5 3 7 6 5 5 6 2 3 6 2 1 1 1 1 1 6 6
## [2313] 1 1 2 1 1 5 6 1 3 7 1 3 6 3 2 1 3 1 7 5 6 1 1 6 3 1 2 7 5 3 1 6 1 7
## [2347] 3 6 6 2 5 1 7 3 3 3 6 1 5 3 6 2 7 3 3 6 1 7 1 6 5 1 1 1 2 1 3 5 5 6
## [2381] 7 1 1 1 7 2 1 1 1 7 7 7 1 1 3 1 6 2 1 5 3 7 1 1 3 1 1 1 6 2 4 5 1 1
## [2415] 7 1 1 3 1 1 6 2 6 7 2 5 1 1 6 6 3 3 1 6 7 2 1 1 6 6 1 1 6 1 1 3 1 5
## [2449] 1 2 7 1 6 5 7 3 3 1 1 6 1 2 1 1 6 6 1 1 5 5 5 3 7 1 2 6 1 1 7 7 7 3
## [2483] 7 1 3 2 5 5 3 6 1 6 1 5 1 5 6 2 5 1 6 5 6 5 6 6 1 5 2 1 3 5 3 6 5 1
## [2517] 6 1 1 2 6 1 5 3 5 6 6 6 6 1 1 2 7 1 1 1 1 6 3 7 7 6 3 6 2 6 2 1 7 7
## [2551] 1 5 6 6 1 7 5 2 5 6 3 1 7 5 1 5 3 2 3 1 5 1 7 5 1 6 1 5 2 1 6 1 1 1
## [2585] 1 7 1 3 1 2 6 6 1 1 1 6 5 1 3 7 7 2 7 7 6 6 3 7 1 1 3 3 2 1 5 6 3 1
## [2619] 6 5 6 1 1 1 2 1 3 7 6 6 7 7 7 5 7 2 3 5 3 3 5 1 7 1 1 2 1 4 6 3 1 5
## [2653] 1 7 5 3 7 6 2 6 1 2 6 5 1 1 3 7 6 3 6 1 2 6 5 7 1 6 7 1 6 6 6 2 1 6
## [2687] 6 1 5 6 3 5 6 2 6 1 6 5 1 6 6 6 3 1 1 2 1 6 1 3 3 6 6 1 1 5 6 1 2 1
## [2721] 6 5 3 4 1 1 6 1 2 1 4 1 5 6 1 5 3 4 1 1 2 6 1 1 1 4 4 6 5 4 1 6 2 6
## [2755] 4 1 6 1 1 1 1 6 2 6 6 1 1 4 1 1 1 6 2 1 6 2 3 6 1 6 6 1 1 1 6 1 2 6
## [2789] 6 1 6 1 1 1 1 4 4 5 6 2 1 3 6 6 6 6 4 1 6 2 6 1 1 2 6 6 1 6 5 2 1 6
## [2823] 6 6 5 6 1 1 2 1 3 6 1 1 1 1 1 6 5 1 6 2 1 1 1 1 1 1 1 5 3 3 6 1 5 1
## [2857] 4 1 6 3 1 2 5 5 5 1 1 4 4 4 4 5 2 1 6 6 6 1 6 1 1 1 2 5 2 4 1 1 1 5
## [2891] 5 5 3 5 2 1 6 6 1 1 1 1 2 6 4 1 1 5 5 1 4 1 3 2 1 1 5 3 1 1 4 1 1 6
## [2925] 6 2 1 6 1 5 6 6 1 1 6 1 2 6 6 1 1 5 1 3 1 3 6 2 1 3 1 1 6 6 1 6 6 1
## [2959] 6 1 2 1 1 3 3 5 6 5 1 1 4 2 1 3 1 1 1 1 1 1 5 2 4 1 5 1 3 1 6 3 1 1
## [2993] 5 2 3 2 1 1 1 6 3 1 1 5 6 2 3 6 6 1 1 1 1 5 1 1 2 6 6 3 6 5 5 1 1 5
## [3027] 6 1 2 3 5 5 1 3 6 6 3 6 1 1 2 4 1 1 1 1 6 5 1 3 6 2 6 3 1 5 1 6 6 6
## [3061] 1 1 1 1 2 1 6 1 5 1 1 1 1 2 1 1 5 5 1 1 1 1 3 6 2 6 3 1 1 1 1 6 4 1
## [3095] 4 2 1 1 6 5 1 1 6 5 5 6 1 2 6 1 2 1 4 4 1 1 6 1 6 3 1 2 6 1 4 6 1 1
## [3129] 3 5 1 5 2 5 5 5 3 6 6 3 5 3 1 2 1 1 6 1 1 3 5 1 4 4 2 4 6 1 5 5 1 6
## [3163] 1 6 6 1 2 6 3 6 1 5 6 1 1 1 2 1 1 3 1 1 1 6 6 1 1 1 3 3 1 6 6 1 6 4
## [3197] 1 1 6 1 3 3 6 1 1 1 6 1 5 1 2 1 1 1 1 5 4 1 6 3 6 3 3 2 1 4 2 3 6 6
## [3231] 1 6 1 6 4 1 1 5 2 1 4 1 1 3 1 1 1 5 1 2 5 1 1 1 1 1 1 5 1 5 2 1 6 6
## [3265] 5 6 6 5 2 6 1 1 1 6 1 1 1 6 2 6 6 5 5 1 1 5 1 2 1 1 6 6 6 1 1 6 2 1
## [3299] 1 1 3 5 1 1 6 2 6 6 6 5 1 6 1 1 6 2 6 1 6 1 5 1 3 6 1 1 2 5 1 2 4 1
## [3333] 1 5 6 1 5 1 4 1 1 3 1 4 1 1 1 4 6 1 6 6 2 6 1 5 5 6 1 5 4 1 6 5 2 3
## [3367] 1 1 1 1 5 4 5 5 1 1 2 6 1 5 1 6 1 1 6 6 1 3 6 6 1 1 1 5 6 3 5 5 2 1
## [3401] 6 3 6 5 1 5 1 1 6 5 3 6 1 1 1 1 5 1 1 1 1 2 4 1 1 1 6 1 1 5
```

Now, we don't really want to run tapply on every single variable when we have over 1,000 different variables. Let's instead use the subset function to subset our data by cluster. Create 7 new datasets, each containing the observations from one of the clusters.

```
cluster1 = subset(daily_kos, clusters_no == 1)
cluster2 = subset(daily_kos, clusters_no == 2)
cluster3 = subset(daily_kos, clusters_no == 3)
cluster4 = subset(daily_kos, clusters_no == 4)
cluster5 = subset(daily_kos, clusters_no == 5)
cluster6 = subset(daily_kos, clusters_no == 6)
cluster7 = subset(daily_kos, clusters_no == 7)
```

How many observations are in cluster 3?

374

Which cluster has the most observations?

Cluster 1

Which cluster has the fewest observations?

Cluster 4

1.5) Hierarchical Clustering

Instead of looking at the average value in each variable individually, we'll just look at the top 6 words in each cluster. To do this for cluster 1, type the following in your R console (where "HierCluster1" should be replaced with the name of your first cluster subset):

```
tail(sort(colMeans(cluster1)))
```

```
##      state republican       poll   democrat      kerry       bush
##  0.7575039  0.7590837  0.9036335  0.9194313  1.0624013  1.7053712
```

This computes the mean frequency values of each of the words in cluster 1, and then outputs the 6 words that occur the most frequently. The colMeans function computes the column (word) means, the sort function orders the words in increasing order of the mean values, and the tail function outputs the last 6 words listed, which are the ones with the largest column means.

What is the most frequent word in this cluster, in terms of average value? Enter the word exactly how you see it in the output:

Explanation

After running the R command given above, we can see that the most frequent word on average is "bush". This corresponds to President George W. Bush.

1.6) Hierarchical Clustering

Now repeat the command given in the previous problem for each of the other clusters, and answer the following questions.

Which words best describe cluster 2?

```
tail(sort(colMeans(cluster2)))
```

```
##      bush  democrat challenge      vote      poll  november
##  2.847352  2.850467  4.096573  4.398754  4.847352 10.339564
```

Which cluster could best be described as the cluster related to the Iraq war?

In 2004, one of the candidates for the Democratic nomination for the President of the United States was Howard Dean, John Kerry was the candidate who won the democratic nomination, and John Edwards with the running mate of John Kerry (the Vice President nominee). Given this information, which cluster best corresponds to the democratic party?

```
tail(sort(colMeans(cluster3)))
```

```
##      elect    parties      state republican   democrat       bush
##   1.647059   1.665775   2.320856   2.524064   3.823529   4.406417
```

```
tail(sort(colMeans(cluster4)))
```

```
## campaign    voter presided     poll     bush    kerry
## 1.431655 1.539568 1.625899 3.589928 7.834532 8.438849
```

```
tail(sort(colMeans(cluster5)))
```

```
##       american       presided administration            war           iraq
##       1.090909       1.120393       1.230958       1.776413       2.427518
##           bush
##       3.941032
```

```
tail(sort(colMeans(cluster6)))
```

```
##      race      bush     kerry     elect  democrat      poll
## 0.4579832 0.4887955 0.5168067 0.5350140 0.5644258 0.5812325
```

```
tail(sort(colMeans(cluster7)))
```

```
## democrat    clark   edward     poll    kerry     dean
## 2.148325 2.497608 2.607656 2.765550 3.952153 5.803828
```

Explanation

You can repeat the command on each of the clusters by typing the following:

tail(sort(colMeans(HierCluster2)))

tail(sort(colMeans(HierCluster3)))

tail(sort(colMeans(HierCluster4)))

tail(sort(colMeans(HierCluster5)))

tail(sort(colMeans(HierCluster6)))

tail(sort(colMeans(HierCluster7)))

You can see that the words that best describe Cluster 2 are november, poll, vote, and challenge. The most common words in Cluster 5 are bush, iraq, war, and administration, so it is the cluster that can best be described as corresponding to the Iraq war. And the most common words in Cluster 7 are dean, kerry, poll, and edward, so it looks like the democratic cluster.

2.2) K-Means Clustering

Now, run k-means clustering, setting the seed to 1000 right before you run the kmeans function. Again, pick the number of clusters equal to 7. You don't need to add the iters.max argument.

```
set.seed(1000)

k_cluster = kmeans(daily_kos, centers = 7)
```

Subset your data into the 7 clusters (7 new datasets) by using the "cluster" variable of your kmeans output.

```
cluster1_k = subset(daily_kos, k_cluster$cluster == 1)
cluster2_k = subset(daily_kos, k_cluster$cluster == 2)
cluster3_k = subset(daily_kos, k_cluster$cluster == 3)
cluster4_k = subset(daily_kos, k_cluster$cluster == 4)
cluster5_k = subset(daily_kos, k_cluster$cluster == 5)
cluster6_k = subset(daily_kos, k_cluster$cluster == 6)
cluster7_k = subset(daily_kos, k_cluster$cluster == 7)
```

How many observations are in Cluster 3?

```
nrow(cluster3_k)
```

```
## [1] 277
```

Which cluster has the most observations?

Which cluster has the fewest number of observations?

```
table(k_cluster$cluster)
```

```
##
##    1    2    3    4    5    6    7
##  146  144  277 2063  163  329  308
```

Explanation

You can run k-means clustering by using the following commands:

set.seed(1000)

KmeansCluster = kmeans(dailykos, centers=7)

Then, you can subset your data into the 7 clusters by using the following commands:

KmeansCluster1 = subset(dailykos, KmeansCluster$cluster == 1)

KmeansCluster2 = subset(dailykos, KmeansCluster$cluster == 2)

KmeansCluster3 = subset(dailykos, KmeansCluster$cluster == 3)

KmeansCluster4 = subset(dailykos, KmeansCluster$cluster == 4)

KmeansCluster5 = subset(dailykos, KmeansCluster$cluster == 5)

KmeansCluster6 = subset(dailykos, KmeansCluster$cluster == 6)

KmeansCluster7 = subset(dailykos, KmeansCluster$cluster == 7)

Alternatively, you could answer these questions by looking at the output of table(KmeansCluster$cluster).

More Advanced Approach:

There is a very useful function in R called the "split" function. Given a vector assigning groups like KmeansCluster$cluster, you could split dailykos into the clusters by typing:

KmeansCluster = split(dailykos, KmeansCluster$cluster)

Then cluster 1 can be accessed by typing KmeansCluster[[1]], cluster 2 can be accessed by typing KmeansCluster[[2]], etc. If you have a variable in your current R session called "split", you will need to remove it with rm(split) before using the split function.

2.2) K-Means Clustering

Now, output the six most frequent words in each cluster, like we did in the previous problem, for each of the k-means clusters.

Which k-means cluster best corresponds to the Iraq War?

Which k-means cluster best corresponds to the democratic party? (Remember that we are looking for the names of the key democratic party leaders.)

```
tail(sort(colMeans(cluster1_k)))
```

```
##          state           iraq          kerry administration       presided
##       1.609589       1.616438       1.636986       2.664384       2.767123
##           bush
##      11.431507
```

```
tail(sort(colMeans(cluster2_k)))
```

```
## primaries  democrat    edward     clark     kerry      dean
##  2.319444  2.694444  2.798611  3.090278  4.979167  8.277778
```

```
tail(sort(colMeans(cluster3_k)))
```

```
## administration          iraqi       american           bush            war
##       1.389892       1.610108       1.685921       2.610108       3.025271
##           iraq
##       4.093863
```

```
tail(sort(colMeans(cluster4_k)))
```

```
##      elect republican      kerry       poll   democrat       bush
##  0.6010664  0.6175473  0.6495395  0.7474552  0.7891420  1.1473582
```

```
tail(sort(colMeans(cluster5_k)))
```

```
##       race     senate      state    parties republican   democrat
##   2.484663   2.650307   3.521472   3.619632   4.638037   6.993865
```

```
tail(sort(colMeans(cluster6_k)))
```

```
##  democrat      bush challenge      vote      poll  november
##  2.899696  2.960486  4.121581  4.446809  4.872340 10.370821
```

```
tail(sort(colMeans(cluster7_k)))
```

```
## presided    voter campaign     poll     bush    kerry
## 1.324675 1.334416 1.383117 2.788961 5.970779 6.480519
```

Explanation

You can output the most frequent words in each of the k-means clusters by using the following commands:

tail(sort(colMeans(KmeansCluster1)))

tail(sort(colMeans(KmeansCluster2)))

tail(sort(colMeans(KmeansCluster3)))

tail(sort(colMeans(KmeansCluster4)))

tail(sort(colMeans(KmeansCluster5)))

tail(sort(colMeans(KmeansCluster6)))

tail(sort(colMeans(KmeansCluster7)))

By looking at the output, you can see that the cluster best correponding to the Iraq War is cluster 3 (top words are iraq, war, and bush) and the cluster best corresponding to the democratic party is cluster 2 (top words dean, kerry, clark, and edward).

2.3 - K-Means Clustering

For the rest of this problem, we'll ask you to compare how observations were assigned to clusters in the two different methods. Use the table function to compare the cluster assignment of hierarchical clustering to the cluster assignment of k-means clustering.

Which Hierarchical Cluster best corresponds to K-Means Cluster 2?

```
table(clusters_no)
```

```
## clusters_no
##    1    2    3    4    5    6    7
```

```
## 1266  321  374  139  407  714  209
```

```
tail(sort(colMeans(cluster1)))
```

```
##      state republican       poll   democrat      kerry       bush
##  0.7575039  0.7590837  0.9036335  0.9194313  1.0624013  1.7053712
```

```
tail(sort(colMeans(cluster2)))
```

```
##      bush  democrat challenge      vote      poll  november
##  2.847352  2.850467  4.096573  4.398754  4.847352 10.339564
```

```
tail(sort(colMeans(cluster3)))
```

```
##      elect    parties      state republican   democrat       bush
##   1.647059   1.665775   2.320856   2.524064   3.823529   4.406417
```

```
tail(sort(colMeans(cluster4)))
```

```
## campaign    voter presided     poll     bush    kerry
## 1.431655 1.539568 1.625899 3.589928 7.834532 8.438849
```

```
tail(sort(colMeans(cluster5)))
```

```
##       american       presided administration            war           iraq
##       1.090909       1.120393       1.230958       1.776413       2.427518
##           bush
##       3.941032
```

```
tail(sort(colMeans(cluster6)))
```

```
##      race      bush     kerry     elect  democrat      poll
## 0.4579832 0.4887955 0.5168067 0.5350140 0.5644258 0.5812325
```

```
tail(sort(colMeans(cluster7)))
```

```
## democrat    clark   edward     poll    kerry     dean
## 2.148325 2.497608 2.607656 2.765550 3.952153 5.803828
```

```
table(k_cluster$cluster)
```

```
##
##    1    2    3    4    5    6    7
##  146  144  277 2063  163  329  308
```

```
tail(sort(colMeans(cluster1_k)))
```

```
##          state           iraq          kerry administration       presided
##       1.609589       1.616438       1.636986       2.664384       2.767123
##           bush
##      11.431507
```

```
tail(sort(colMeans(cluster2_k)))
```

```
## primaries  democrat    edward     clark     kerry      dean
##  2.319444  2.694444  2.798611  3.090278  4.979167  8.277778
```

```
tail(sort(colMeans(cluster3_k)))
```

```
## administration          iraqi       american           bush            war
##       1.389892       1.610108       1.685921       2.610108       3.025271
##           iraq
##       4.093863
```

```
tail(sort(colMeans(cluster4_k)))
```

```
##      elect republican      kerry       poll   democrat       bush
##  0.6010664  0.6175473  0.6495395  0.7474552  0.7891420  1.1473582
```

```
tail(sort(colMeans(cluster5_k)))
```

```
##       race     senate      state    parties republican   democrat
##   2.484663   2.650307   3.521472   3.619632   4.638037   6.993865
```

```
tail(sort(colMeans(cluster6_k)))
```

```
##  democrat      bush challenge      vote      poll  november
##  2.899696  2.960486  4.121581  4.446809  4.872340 10.370821
```

```
tail(sort(colMeans(cluster7_k)))
```

```
## presided    voter campaign     poll     bush    kerry
## 1.324675 1.334416 1.383117 2.788961 5.970779 6.480519
```

Explanation

From "table(hierGroups, KmeansCluster$cluster)", we read that 116 (80.6%) of the observations in K-Means Cluster 2 also fall in Hierarchical Cluster 7.

```
table(clusters_no, k_cluster$cluster)
```

```
##
## clusters_no    1    2    3    4    5    6    7
##           1    3   11   64 1045   32    0  111
##           2    0    0    0    0    0  320    1
##           3   85   10   42   79  126    8   24
##           4   10    5    0    0    1    0  123
##           5   48    0  171  145    3    1   39
##           6    0    2    0  712    0    0    0
##           7    0  116    0   82    1    0   10
```

2.4) K-Means Clustering

Which Hierarchical Cluster best corresponds to K-Means Cluster 3?

Explanation

From "table(hierGroups, KmeansCluster$cluster)", we read that 171 (61.7%) of the observations in K-Means Cluster 3 also fall in Hierarchical Cluster 5.

2.5) K-Means Clustering

Which Hierarchical Cluster best corresponds to K-Means Cluster 7?

Explanation

From "table(hierGroups, KmeansCluster$cluster)", we read that no more than 123 (39.9%) of the observations in K-Means Cluster 7 fall in any hierarchical cluster.

2.6) K-Means Clustering

Which Hierarchical Cluster best corresponds to K-Means Cluster 6?

Explanation

From “table(hierGroups, KmeansCluster$cluster)”, we read that 320 (97.3%) of observations in K-Means Cluster 6 fall in Hierarchical Cluster 2.