Exposure (23 variables)

EXPOSURE VARIABLES

* exposure\_id -> unique ID for every exposure (i.e. 9 articles that were displayed)
* exposure\_user\_id -> ID which exposure it was for that particular user
* session -> ID which session (one session: series of exposures until logout or inactivity for more than 15 minutes) it was for that particular user
* timestamp -> timestamp of the exposure
* user\_id -> unique ID of the user (can be used to merge with user info)
* majority -> what topic appeared most?
* number\_most\_freq -> how often did the most frequent topic appear?
* number\_topics\_all -> how many different topics were in the exposure?
* ties -> were there multiple topics that appeared most often?
* Match -> was the most frequently appearing topic selected?
* devices\_group -> what device (mobile or PC) was used during this exposure?
* os\_group -> what OS (Mac, Linux, Windows) was used during this exposure?
* browser\_group -> what Browser (Chrome, Safari, Firefox) was used during this exposure?
* show\_again -> whether participant opted to show the selection again

NEWS ITEM VARIABLES

* selected -> was this news item selected?
* recommended -> whether the news item was recommended
* position -> position of the news item
* news\_id -> unique ID of the news item (can be used to merge with article info)

VARIABLES SELECTED NEWS ITEM

* news\_id\_selected -> ID of the news item selected in this exposure
* rating\_selected -> rating of the news item selected in this exposure
* rating2\_selected -> other rating of the news item selected in this exposure
* topic\_selected -> topic of the news item selected in this exposure
* position\_selected -> position of the news item selected in this exposure

Users (42 variables)

* user\_id -> unique ID of the user (can be used to merge with exposure info)
* group -> experimental group
* gender
* age
* edu
* pol\_knowledge
* overconfidence -> difference between actual and assumed political knowledge
* political\_efficacy
* pol\_interest
* overall\_media\_usage
* news\_topic\_interest
* div\_trait -> how many different news topics were rated higher than 4
* div\_trait2 -> how many different news topics were rated higher than 3
* news\_pref\_strength -> strength of news preferences
* desirability\_of\_control -> desirability of control
* beh\_int -> behavioral intention (visit again)
* first\_login -> first login into the system
* last\_visit -> last visit
* number\_logins -> how often logged in in total
* number\_stories\_read -> how many stories read in total
* m\_ -> Variables about media usage (tv, newspaper etc)
* t\_ -> variables about topic interests (eight topics)
* Diversity\_ -> perceived diversity (four measures)
* System\_control/strategy\_control, control\_overall -> perceived control
* Satisfaction -> satisfaction with the system
* Avg\_rating2 -> average rating of that person

Article Features (7 variables)

* news\_id -> unique ID of the news item (can be used to merge with exposure info)
* url -> URL of the news item
* teaser\_rss -> Teaser
* title\_rss -> Title
* publication\_date -> Publication Date
* text -> Full Text
* topic -> Topic (annotated by using supervised machine learning classifier)