User Groups

Misinformation sharers (MS) *(Variable name: fns)**	Users who share at least 1 URL with misinformation
Deplatformed users (DU) *(Variable name: suspended)**	Users who were removed from Twitter between January 6th and January 12th due to Twitter's enforcement of its terms of use policy
High activity users (Variable name: ha)*	Users who sent at least 3,200 tweets during a six-week collection interval between 2018 and April 2020
Moderate activity users *(Variable name: ma)**	The approximately 500,000 most active users who did not meet the criteria for high activity
Low activity users *(Variable name: la)**	All users who did not meet the criteria for the high or moderate activity groups
QAnon *(Variable name: qanon)**	Users who tweeted or retweeted at least five distinct QAnon-related hashtags. The set of hashtags was defined in the GWU-IDDP report "QAnon on Twitter: An Overview"
Antivax *(Variable name: av)**	Users who tweeted or retweeted at least five distinct anti-vax hashtags (defined internally based on a vaccine misinformation dashboard)
Supersharers 0.1% *(Variable name: ss1)**	Users in the 99.9th percentile of misinformation URL sharing/retweeting. This percentile is defined in terms of all users.
Supersharers 5% *(Variable name: ss5)**	Users in the 95th percentile of misinformation URL sharing/retweeting, among users with at least one misinformation URL shared.

See Tables Variables List

User Subgroups

Trump-only followers *(Variable name: A)**	MS non-DU who follows Trump but don't follow any other deplatformed user
Deplatformed followers, including Trump followers *(Variable name: B)**	MS non-DU who follows Trump OR any other deplatformed user; follow at least one deplatformed user, Trump or not
Four+ Deplatformed followers, including Trump followers *(Variable name: D)**	MS non-DU who follows at least four deplatformed users, including Trump or not among the three followed

Not deplatformed followers *(Variable name: F)**	MS non-DU who do not follow any deplatformed user

User-level data

- File: panel-2020-user-level-did.csv
- Note: See 'Data Privacy Statement'

Variable	Type	Description
user_id	chr	User unique identification (anonymized)
id	int	User unique identification number
date	date	Date
period	int	Date index
fake_initiation	int	Number of misinformation tweets
fake_rt	int	Number of misinformation retweets
fake	int	fake_initiation + fake_rt
fake_grinberg_initiation	int	Same as 'fake_initiation', but using Grinberg et al. (2019) list
fake_grinberg_rt	int	Same as 'fake_rt', but using Grinberg et al. (2019) list
fake_grinberg	int	fake_grinberg_initiation + fake_grinberg_rt
fake_grinberg_rb_initiation	int	Same as 'fake_initiation', but using Grinberg list
fake_grinberg_rb_rt	int	Same as 'fake_rt', but using Grinberg list
fake_grinberg_rb	int	fake_grinberg_rb_initiation + fake_grinberg_rb_rt
fake_newsguard_initiation	int	Same as 'fake_initiation', but using Newsguard list
fake_newsguard_rt	int	Same as 'fake_rt', but using Newsguard list
fake_newsguard	int	fake_newsguard_initiation + fake_newsguard_rt

not_fake_initiation	int	Number of non-misinformation tweets
not_fake_rt	int	Number of non-misinformation retweets
not_fake	int	not_fake_initiation + not_fake_rt
not_fake_conservative_initiation	int	Number of non-misinformation tweets with conservative news domains
not_fake_conservative_rt	int	Number of non-misinformation retweets with conservative news domains
not_fake_conservative	int	not_fake_conservative_initiation + not_fake_conservative_rt
not_fake_liberal_initiation	int	Number of non-misinformation tweets with liberal news domains
not_fake_liberal_rt	int	Number of non-misinformation retweets with liberal news domains
not_fake_liberal	int	not_fake_liberal_initiation + not_fake_liberal_rt
not_fake_shopping_initiation	int	Number of non-misinformation tweets with shopping-related content
not_fake_shopping_rt	int	Number of non-misinformation retweets with shopping-related content
not_fake_shopping	int	not_fake_shopping_initiation + not_fake_shopping_rt
not_fake_sports_initiation	int	Number of non-misinformation tweets with sports-related content
not_fake_sports_rt	int	Number of non-misinformation retweets with sports-related content
not_fake_sports	int	not_fake_sports_initiation + not_fake_sports_rt

tweet	int	Number of misinformation tweets if the user is a misinformation sharer (fns=1 or, equivalently, nfns=0) or the number or non-misinformation tweets if the user is not a misinformation sharer (fns=0 or, equivalently, nfns=1)
retweet	int	Number of misinformation retweets if the user is a misinformation sharer (fns=1 or, equivalently, nfns=0) or the number or non-misinformation retweets if the user is not a misinformation sharer (fns=0 or, equivalently, nfns=1)
trt	int	tweet + retweet
n	int	fake + not_fake
fns	bin	1=misinformation sharer; 0=not a misinformation sharer. See Table 'User Groups'
suspended	bin	1=suspended account; 0=not a suspended account. See Table 'User Groups'
qanon	bin	1=QAnon account; 0=not a QAnon account. See Table 'User Groups'
av	bin	1=Antivax account; 0=not an antivax account. See Table 'User Groups'
ss1	bin	1=Supersharer 0.1%; 0=not a supersharer 0.1%. See Table 'User Groups'
ss5	bin	1=Supersharer 5%; 0=not a supersharer 5%. See Table 'User Groups'
ha	bin	1=High activity users; 0=not a high activity user. . See Table 'User Groups'
ma	bin	1=Moderate activity users; 0=not a moderate activity user. . See Table 'User Groups'

la	bin	1=Low activity users; 0=not a low activity user. . See Table 'User Groups'
A	bin	Group A. See Table 'Subgroups'
A_ha	bin	Group A and high activity user
A_ma	bin	Group A and moderate activity user
A_la	bin	Group A and low activity user
B	bin	Group B. See Table 'Subgroups'
B_ha	bin	Group B and high activity user
B_ma	bin	Group B and moderate activity user
B_la	bin	Group B and low activity user
D	bin	Group D. See Table 'Subgroups'
D_ha	bin	Group D and high activity user
D_ma	bin	Group D and moderate activity user
D_la	bin	Group D and low activity user
F	bin	Group F. See Table 'Subgroups'
F_ha	bin	Group F and high activity user
F_ma	bin	Group F and moderate activity user
F_la	bin	Group F and low activity user
nfns	bin	1=Not a misinformation sharer user; 0 otherwise.
nfns_ha	bin	nfns and high activity user
nfns_ma	bin	nfns and moderate activity user
nfns_la	bin	nfns and low activity user
dosage_no_trump_X	bin	Same as Group B (see Table Subgroups), but follows at least X deplatformed account(s), but do not follow Trump
dosage_trump_X	bin	Same as Group B (see Table Subgroups), but follows at least X deplatformed account(s), including Trump

age	int	Age
gender	chr	Gender (male, female, unknown). Sourced from TargetSmart voter file
female	bin	Gender. 1=Female; 0=Other.
race	chr	Race. Race is self-reported from the voter file where available (typically in areas formerly under VRA preclearance), collected from commercial data sources otherwise.
white	bin	Race. 1=Causacian; 0=Other
pres_gen	dbl	TargetSmart's modeled propensity to vote in presidential general elections. Range: 0-100.
pres_pri	dbl	TargetSmart's modeled propensity to vote in presidential primary elections. Range: 0-100.
ptyid	chr	TargetSmart's modeled propensity to self-identify as a Democrat in a phone poll. Range: 0-100.

Aggregated data

- Files: decahose-daily-totals.csv, panel-2016-daily-totals.csv, panel-2020-daily-totals.csv
- Note: Variables with the same name as in the 'User-level data (panel-2020-user-level-did.csv)' contain the same information described in the codebook table for user-level data. The difference is that these files provide only aggregated data by day.

Variable	Type	Description
stat	chr	Aggregation statistics (``total'' or ``average'')
nusers	int	Number of users collected
group	chr	User group. See tables 'Groups' and 'Subgroups'