ASA's 2020 Fall Data Challenge Dataset 101
About the Dataset

For *This* is Statistics' 2020 **Get Out the Vote Challenge**, all teams must utilize the IPUMS-ASA U.S. Voting Behaviors dataset (bit.ly/FDC_Dataset) in their entry.

This is a dataset about voting behavior in the U.S. over the past 14 years. It uses survey data from the Census Bureau and Department of Labor of Statistics that's collected on a sample of Americans every two years right after the November elections.

Demographic information is available for each respondent, including:
- Age
- Sex
- Race
- Ethnicity
- Marital status
- Military veteran status
- Size of community
- Employment status
- Educational background

There is also voting information for each respondent, including:
- Whether or not they were registered to vote
- If they were not registered, the reason why not
- Whether or not they voted in the most recent election
- If they didn't vote, the reason why not
- Whether they voted on Election Day or before, in person or by mail

Teams should review the complete codebook for detailed description of the variables.

For additional insights into the dataset, review the Get Out the Vote Challenge Dataset FAQ (bit.ly/FDC_Dataset_FAQ_Living).

ASA's 2020 Fall Data Challenge Dataset 101
Dataset Variable Terms Glossary

YEAR
year of the survey

STATEFIP
person's state of residence

METRO
type of metropolitan area person lives in

AGE
person's age at last birthday

SEX
person's self-reported sex

RACE
person's self-identified race

MARST
person's current marital status

VETSTAT
status as a veteran of the U.S. military forces

CITIZEN
status as native-born or naturalized citizen

HISPAN
details about Hispanic origin

LABFORCE
whether the person was in the labor force during the preceding week

EDUC99
person's highest level of educational attainment

EDCYC
number of years of college credit earned (for those with a college education but not a bachelor's degree)

EDDIPGED
completed an equivalency test such as the GED or not (for those who completed high school)

EDHGCGED
highest level of schooling attained (for those who did not complete high school)

SCHLCOLL
enrolled in high school or college or not (for those aged 16 to 24)

VOWHYNOT
reason for not voting (for those who were registered but did not vote in the most recent November election)

VOYNOTREG
reason for not registering to vote in the election (for those who were not registered)

VOTEHOW
whether the person voted in person or by mail (for those who voted)

VOTEWHEN
whether the person voted on election day or before (for those who voted)

VOREGHOW
how the person registered to vote (for registered voters)

VOTED
whether the person voted during the most recent election (for those who were registered)

VOREG
whether the person was registered to vote (for those who did not vote in the most recent election)

VOSUPPWT
the weight specific to the Voter Supplement

MARRSIMPLE, RACESIMPLE, HISPSIMPLE, EDUSIMPLE
simplified versions of MARST, RACE, HISPAN, and EDUC99