USC Dornsife / Los Angeles Times 2016 Election "Daybreak" Poll

Data Documentation and Codebook

Version: 9/20/2017 8:12:00 PM

USC Dornsife Dana and David Dornsife College of Letters, Arts and Sciences

Los Angeles Times

2016 Presidential Election "Daybreak" Poll

Table of Contents

USC Dornsife
Dana and David Dornsife College of Letters, Arts and Sciences
Los Angeles Times

2016 Presidential Election "Daybreak" Poll

Introduction

The 2016 USC Dornsife / LA Times Presidential Election "Daybreak" poll represents a pioneering approach to tracking changes in Americans' opinions throughout the 2016 campaign for the White House. The poll used non-traditional *probabilistic polling* methods which are an alternative to traditional polling's categorical vote questions and likely voter screens.

The 2016 Daybreak Poll is one of many projects associated with the ongoing Understanding America Study (UAS) at the University of Southern California's (USC) Dornsife Center for Economic and Social Research (CESR), and was conducted in partnership with USC's Jesse M. Unruh Institute of Politics and the Los Angeles Times.

The Daybreak Poll election panel was based in the UAS, a probability-based internet poll established in 2014. At the time of the 2016 election, the UAS consisted of about 6,000 individuals whose household addresses were randomly selected from zip codes across the United States. Detailed information about UAS methods, as well as access to UAS publically available datasets, is available online.

The Daybreak Poll's methods were developed based on findings from two prior implementations of probabilistic polling: a pilot study in 2008 (Delavande & Manski, 2010), and a full tracking poll in 2012, when members of our team, then at RAND, conducted the successful 2012 RAND Continuous Presidential Election Poll (Gutsche, Kapteyn, Meijer, & Weerman, 2014; Kapteyn, Meijer, & Weerman, 2012).

The aims of the poll were to track change in voting preference in an election panel over time; Investigate associations of candidate preference and intent to vote with voter characteristics and attributes; to continue to investigate the ability of the probabilistic polling method to accurately estimate popular vote outcome in the election; and to contribute to the field of election research by continuing evaluation of alternative methods of tracking public opinion during an election. The Daybreak Poll's tracking poll chart, data files, and methodology documentation are posted on the website election.usc.edu.

This document summarizes the Poll's various data files, methodological reports and other programs and documentation. It includes full question text, overviews and links to more detailed information regarding the poll's weighting and estimation schemes, and a guide to the various pre and post-election data collections. These election data and methodology documents, including this document, are available at https://uasdata.usc.edu/page/UAS+Daybreak+Election+Poll.

Daybreak Poll-related news stories, op-eds, blog posts, and reports are listed on our poll coverage page.

CESR's Daybreak Poll team included the Center's Executive Director Arie Kapteyn, Managing Director Tania Gutsche, Survey Director Jill E. Darling, Senior Economist Erik Meijer, and IT Director Bas Weerman.

Understanding America Study
Dornsife Center for Economic and Social Research
University of Southern California

PO Box 77902
Los Angeles, CA 90007-9983
Contact us at uas-l@mymaillists.usc.edu

To look at other survey data sets, visit https://uasdata.usc.edu

Follow us on twitter

Probabalistic Polling Questions

Each day, we invited one-seventh of the members of the UAS election panel to answer three predictive questions. As shown in Figure 1 (click here for a larger view) those questions asked:

What is the percent chance that…

1. **you will vote in the presidential election?**
2. **you will vote for Clinton, and for Trump, and for someone else (answers sum to 100%)**
3. **Clinton, Trump or someone else will win (answers sum to 100%)**

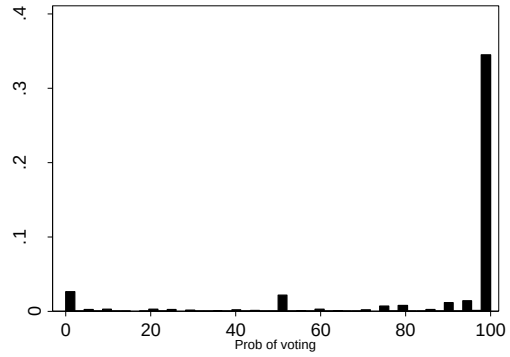
Figure 1: Screenshot of the Daybreak Election Poll vote questions (order of candidates was randomized)

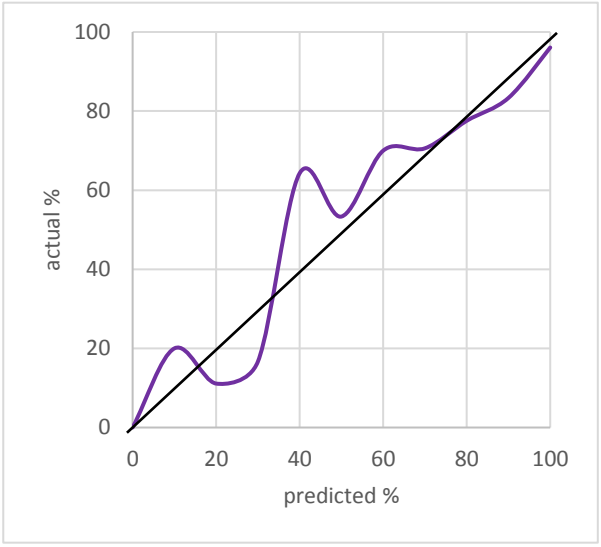
To find out more about what lay behind the vote, each week, following the three vote questions, we asked respondents one or two questions about their preferences and values. The data and documentation associated with these topical questions are described on page 14. The data were released on March 1, 2017.

Performance of the probabilistic polling questions

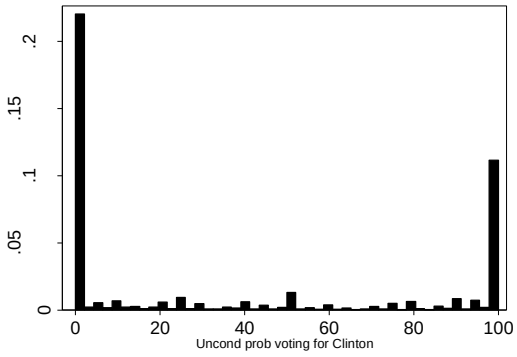
This section provides histograms and last predicted likelihood of voting v. actual vote using linear regression,, controlling for date of last participation wave, for each of the three vote questions.

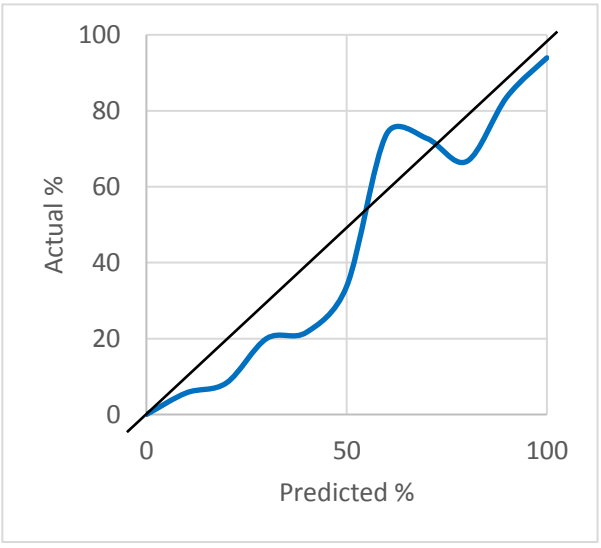
What is the percent chance that you will vote in the presidential election?



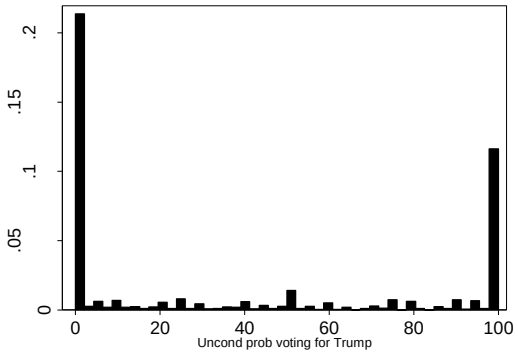


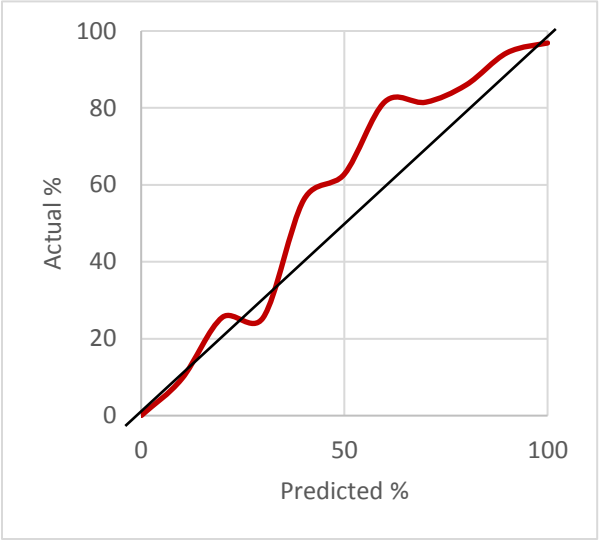
What is the percent chance that [if you were to vote] you would vote for Hillary Clinton?





What is the percent chance that [if you were to vote] you would vote for Donald Trump?





Sampling and Recruitment

This section describes methods we used to conduct the Daybreak Poll in the Understanding America Study panel. We link to more detailed or technical documents available online.

Some Important information about UAS Samples:

1. Post-election analysis indicated that the UAS and Daybreak samples at the time that the Daybreak poll was conducted included a disproportionately high number of rural respondents. Al UAS survey final weights adjust for this bias. We recommend adjusting or accounting for this bias in any unweighted analysis. Please contact us at uas-l@usc.edu with any questions.

2. The UAS panel include small oversamples of Los Angeles County residents and Native Americans. These respondents are assigned a weight of zero but you may prefer to drop them from unweighted analysis. Respondents in the representative national sample are identified as **sampletype** = 1.

Understanding America Study National Internet Panel

The Daybreak Poll is based in the UAS, a probability-based internet panel that was established in 2014. UAS members are recruited from among all households in the United States and constitute a representative sample of U.S. residents. UAS members take about one to three surveys a month and are paid an incentive of about $.66 per question for each survey they complete. Recruited households that did not have internet access were provided with a tablet and internet service to ensure coverage of includes all ages and socioeconomic statuses.

The UAS panel grew over the course of the 2016 election season from about 4000 to about 6000 respondents. Membership in the election panel grew proportionally from about 3000 to 5000 (Figure 2).

Detailed information about the overall Understanding America Study's sampling and recruitment procedures, and detailed data on UAS recruitment rates, is documented online in the UAS Panel Sample and Recruitment page.

Daybreak Poll

Recruitment

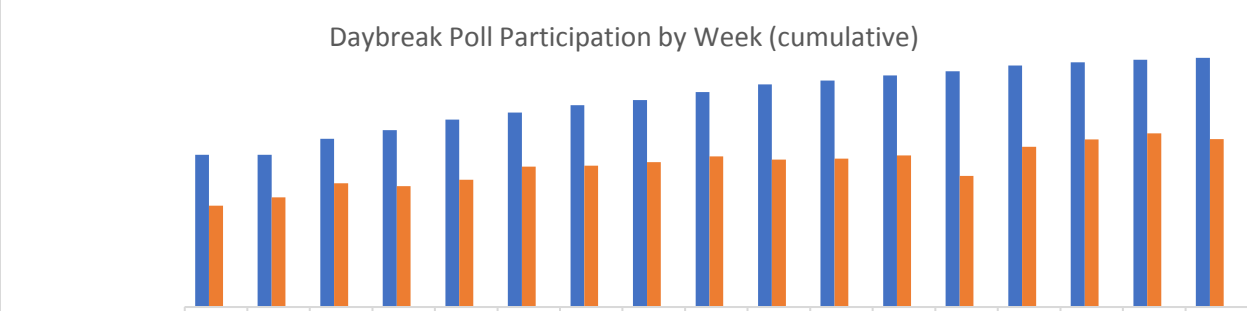
Daybreak poll's election panel members are UAS panel members who are U.S. citizens (eligible voters). The election panel was recruited via a UAS survey (UAS 47) of 2012 vote history, which included the invitation to participate in the weekly election panel. Detailed information about the Daybreak poll's sampling procedures is provided in the document: Sample selection and estimation in the Daybreak Poll.

Participation rates and patterns

We invited a total of 5706 eligible UAS members to participate between May 16 and November 1, 2016. Of those, 5007 completed the survey (88%), and 4,857 consented to participate in the election panel (85%). Figure 2 provides a weekly summary of cumulative number of participants invited, participated, and response rate.

4,509 participated at least once. 6% participated once and never again. 9% dropped out before mid-October, and 64% participated every or nearly every week after joining the panel (35% every week, 29% missed 1 or 2 weeks.)

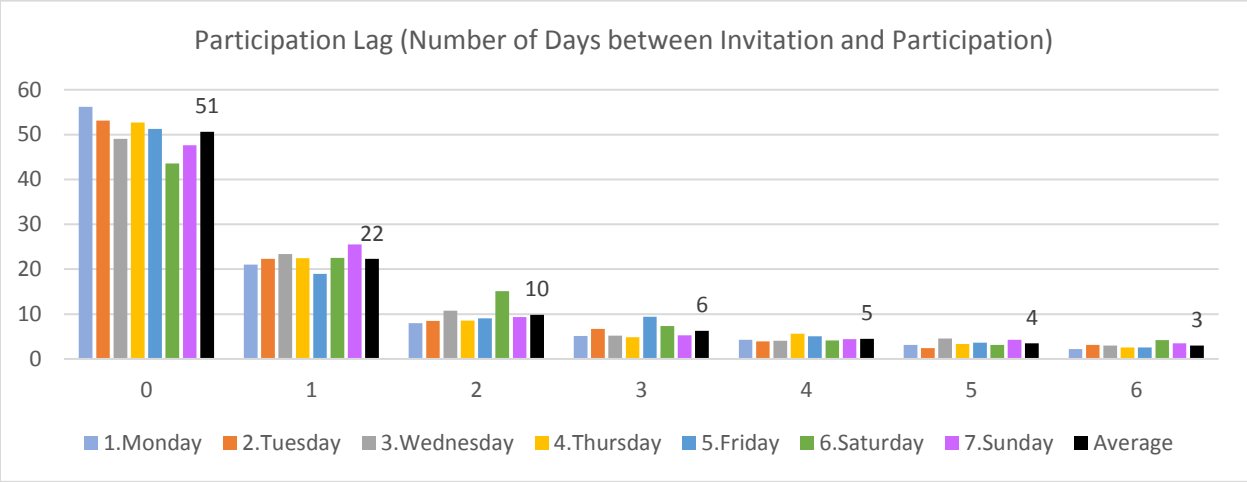
Figure 2: Number of election poll panel members invited, and participated, by week



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
invited	2957	2957	3270	3434	3641	3779	3922	4019	4180	4328	4403	4500	4583	4692	4758	4805	4842
participated	1968	2130	2404	2351	2471	2726	2748	2816	2926	2862	2882	2943	2545	3114	3254	3375	3260
%	66.6%	65.1%	70.0%	64.6%	65.4%	69.5%	68.4%	67.4%	67.6%	65.0%	64.0%	64.2%	54.2%	65.4%	67.7%	69.7%	67.3%

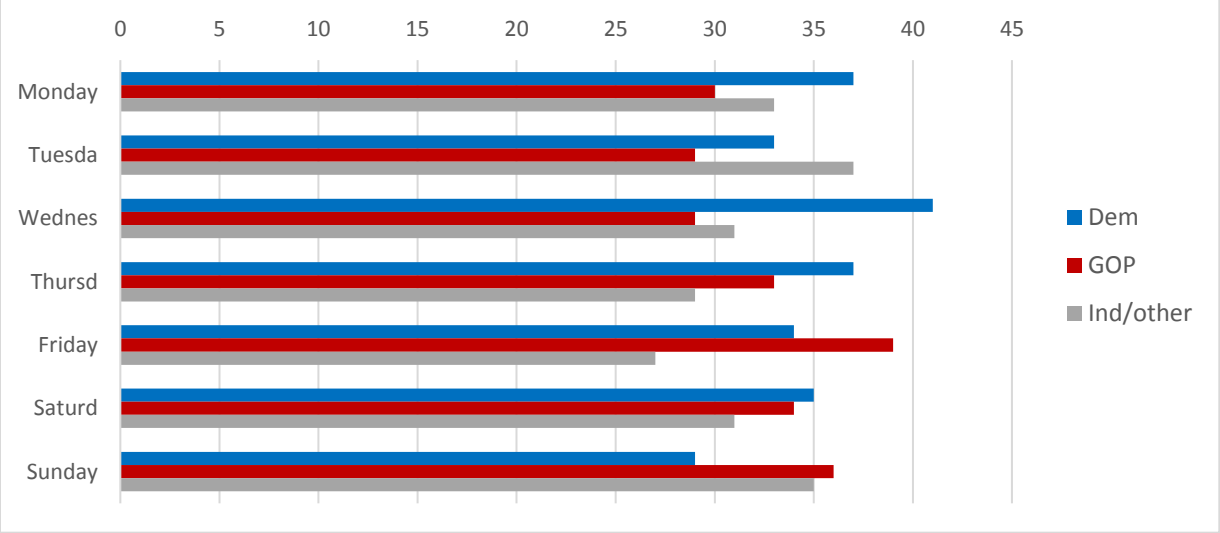
We assigned each Daybreak poll member a specific participation day of the week, so that 1/7th of the sample were invited to participate each day. Panel members could respond any day or time that was convenient before their next participation day. Figure 3 shows the "lag time" between invitation and participation, by day of week assignment. On average, about half participated on their assigned day, and more than 8 out of 10 participated within 2 days of their assigned day.

Figure 3: Participation lag by Day of Week Assignment



We made some attempt to balance the participation day assignments, and each 7-day sample was separately weighted to align with population benchmarks. However, after weighting, we continued to observe variation by party affiliation, as indicated in Figure 4.

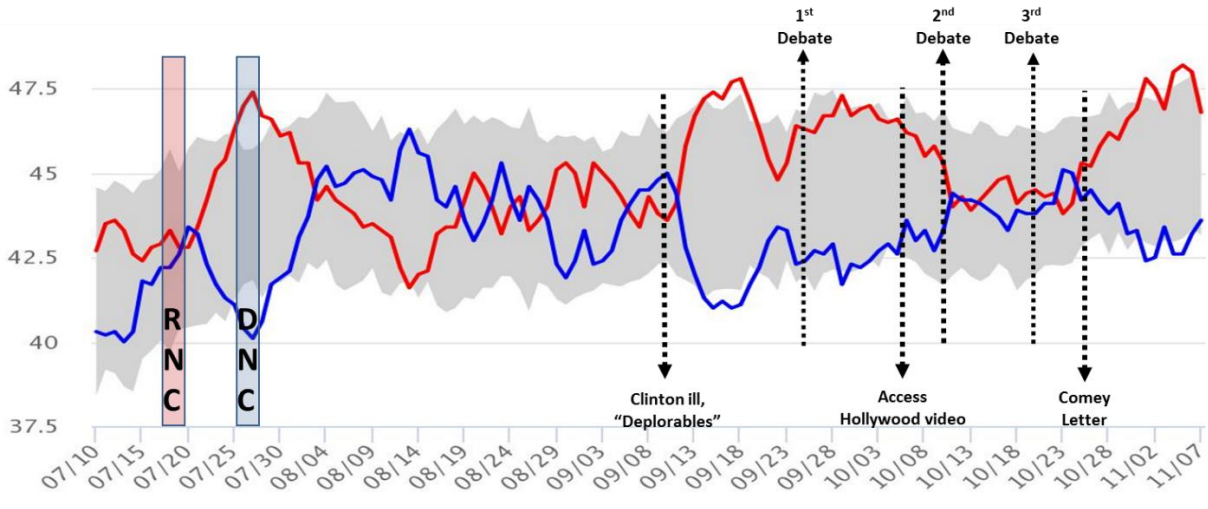
Figure 4: Party Affiliation Percentage by Assigned Day. Sample is set of final Daybreak Poll participations



The Daybreak Poll Daily Tracking Chart – Estimates of Voter Preference

At midnight each day, the daybreak poll statistician weighted the Daybreak Poll results from the previous seven days and posted the latest estimates based on that data online at election.usc.edu and on the LATimes.com Politics site shortly after midnight. Figure 5 shows the final tracking chart annotated with election events. These include the Republican National Convention (RNC); the Democratic National Convention (DNC); events around September 11th that included Clinton referring to some of Trump's supporters as "Deplorables" and concealing an illness; the three presidential debates; the release of an Access Hollywood video showing Trump making comments about women; and the release of a letter by then-FBI Director James Comey indicating that the FBI had learned of the existence of emails that appear to be pertinent" to the previously closed investigation of Clinton's emails. Preliminary analysis indicates that the 9/11 date and the date of the Comey letter are significant inflection points.

Figure 5: 2016 Daybreak Tracking Poll Annotated with election events



To obtain the values shown in the tracking chart, we weighted each respondent's likelihood of voting for a candidate with their likelihood of voting in the presidential election. We then calculated the mean of that number for all respondents during the last 7 days, taking into account respondent-level weights based on demographics and past voting behavior. This created the estimated fraction of the population that we forecasted would vote for the candidate.

Detailed information about the Daybreak poll's estimation procedures is available online in the document: Sample selection and estimation in the Daybreak Poll.

Weighting the Daybreak Poll

The Daybreak poll sample was weighted to match distributions of population of interest on a number of characteristics: sex, age, race-ethnicity, education, household size, family income, and voting in the previous presidential election (2012). We used data from the Current Population Survey (CPS), restricted to U.S. citizens age 18 or older to create the reference benchmarks for demographics and to distributions obtained from the US Election Project for 2012 vote.

Weights were constructed using a raking method, which ensures that the distributions are exactly the same in the weighted poll data as in the reference population (e.g., Lu & Gelman, 2003; Valliant, Dever, & Kreuter, 2013, pp. 358–361; Kolenikov, 2014).

For more detailed information and specific distributions used in weighting the poll, please consult the document posted online during the election: Weighting the Daybreak Poll.

Daybreak Poll Data files and Documentation

This section provides detailed information about the data files available on the 2016 Daybreak Poll Data page. All election-related data files and documentation, including this document, and data from post-election surveys are available online at https://uasdata.uas.edu under "Data/Election Poll 2016".

During the course of the election, the Daybreak Poll site's "detailed data" tab provided the day's chart data point in csv format. Detailed information about these chart files, is here: http://cesrusc.org/election/readme.txt.

The Daybreak Poll generated two sets of cumulative longitudinal microdata files: *fulldata* and *polldata*. The *fulldata* file included the cumulative microdata from all participants over the course of the election. Each *polldata* file represented a daily estimate of the vote – one point on the tracking chart. The final combined *polldata* file we have provided includes the cumulative set of 7-day aggregations over each day of the election season. Details about these files are provided in the sections below.

Daybreak Poll Cumulative Microdata: the "fulldata" file

We have provided the final longitudinal file compiled during the course of the election campaign, and an updated "corrected" version that addressed minor inaccuracies in the original file. All changes away from the final cumulative file are provided below.

The files are provided in stata format, zipped. Variable summaries for this file are provided in Appendix II: Descriptive Stats for "Fulldata" Longitudinal File, starting on on page 17. We have not provided sample weights, but strongly urge data users to construct weights that will account for sample variation away from known distributions for voting age populations. Refer to Weighting the Daybreak Poll for information on how UAS weighted the Daybreak poll.

Filenames and number of observations

Filenames	fulldata.dta (final cumulative fulldata file)
	full2016corrected.dta (corrected file)
File locations	fulldata.zip
	full2016corrected.zip
Observations	50249
Unique observations	4509

Unique identifiers for each observation

uasid	individual identifier
ts	time stamp of the observation

Panel variables

invite	Date invited to participate
pollwave/pollweek	Week invited to participate, and week the poll was completed *
polldate	Date the poll was completed
polldayno	Sequential day number survey was completed (see Appendix I, pg 16)

Notes on corrections in the full2016corrected.dta file

1. A new fulldata file was created every day during the field period. The corrected file posted online corresponds to the final Daybreak poll posted fulldata.dta file (11/08/2016), with some corrections.

2. One case responded on 07/24 before being invited:

```
     uasid                    ts   wkdaygrp      invite   pollwave    polldate
 160403807   2016-07-24 09:37:40          .   31dec1969      -2426   24jul2016
```

   (The strange values of invite and pollwave are the result of missing dates.)

   This is how this case looked like in the 07/24 version of the data. n the 07/25 and later versions of the data, ts is changed to "2016-07-24 16:37:40" and invite and pollwave are set to missing. This is how the record is in full2016corrected.dta. This has no further consequences, except that in the poll2016combined.dta file, for targetdate 07/24, ts="2016-07-24 09:37:40", so that this record does not merge with the corresponding record in full2016corrected.dta. For later target dates (07/25-07/30),ts="2016-07-24 16:37:40" so this record does merge.

3. uasid == "150800005" & ts == "2016-10-15 09:23:28":
   Voted early (dp004=1), but did not say for whom (dp005=0). In the 10/15 and 10/16 versions of the data, this was coded as having voted for "someone else" (dp002_3_ = 100, other_vote=100). From 10/17 onward, this was corrected in the derived data (consistent=0, trump_vote=clint_vote=other_vote=.e, although still dp002_3_ = 100). The current file also includes this correction. This implies a discrepancy with the records in the poll2016combined.dta file with target dates 10/15 and 10/16, but not later dates in which this record was used (10/17-10/21).

4. The following variables were removed
   Per apparent new policy (myhousehold.dta):
   - dateofbirth_day [already removed in later fulldata files]
   - dateofbirth_month

   Variables without information:
   - merge47 [= always 3]
   - en002_order [= always missing; should have been en003_order, but this didn't exist either]
   - uashhid_raw

   Variables that had coding errors and were not used in the poll results:
   - batch
   - hhbatch

- primary_respondent

the variable sampletype was retained, even though it contains errors, because this affects the poll result and thus this is necessary to be able to replicate our
results. (See below.)

5. the following variables were added to the corrected file:
   - en003_order: candidate order for en003 in UAS47
   - uashhid_corr: corrected version of uashhid
   - batch_corr: corrected batch variable
   - prim_corr: corrected primary_respondent variable
   - sampt_corr: corrected sampletype variable

6. The election data was combined with the latest recruitment.dta, myhousehold.dta, and myhousehold_first.dta files to construct the corrections to the four variables above. These corrections would not affect the poll results, except sampletype. There were some observations in which sampletype was incorrect:

```
. tabulate sampletype sampt_corr, missing

                   |     Sample type (corrected)
      Sample type | 1 Nationa  2 Native   3 LA Coun |     Total
------------------+---------------------------------+----------
       1.National |    46,134          26         285 |    46,445
2.Native American |         4       2,572           0 |     2,576
      3.LA County |         1           0       1,227 |     1,228
------------------+---------------------------------+----------
            Total |    46,139       2,598       1,512 |    50,249
```

The main errors were the classification of about 80 members from the LA County 2 batch (13) as being from a "Nationally Representative" sample. Hence, these were incorrectly included in the poll, from September onward. Given that LA County is more Democratic than the rest of the country, I conjecture that this led to a slight trend toward Clinton in September, but I don't know how big this effect may have been.

Daybreak Poll Daily Estimates: the "polldata" file

We have retained the final polldata file produced during the election, and a file that provides all of the Daybreak poll vote estimates calculated for the daily tracking chart from July 4 to November 7th, 2016. These are provided in the poll2016combined file. Each record represents a daily estimate of the forecast vote.

Filenames and number of observations

Filename	polldata.dta (final 7-day file) and poll2016combined.dta (combined file)
File location	poll2016combined.zip (combined file, zipped)
Observations	304496

Identifiers

targetdate	date for which the analysis is used (last day of 7-day window)
uasid	individual identifier
ts	time stamp of the observation

Variable change from original: In this final data file, the *samp* variable which was included in the *polldata.dta* files posted online during the election season were replaced by *targetdate*.

Because most observations are used for 7 samples, i.e. each observation is included in seven daily tracking estimates, most (uasid, ts) observations occur 7 times, with different weights.

Weight variables

wgt_main	main weight
wgt_1--wgt_80	replication weights

Important information about weighting and analysis using the combined polldata file

The data have been declared and saved as survey data, with the proper weights, replication weights, and cluster info. However, the replication weights were developed for separate analyses of the different target samples, and in particular do not take dependence of observations across target samples into account.

For this reason, it is inappropriate to use this file for analyses that involve more than one target date. Hence, all analyses should be done with either ... if (targetdate == ...) or bysort targetdate: ... Any analyses (e.g., panel data analyses) that use samples that differ in some way from the target samples (e.g., use subsets or data that span more than 7 days) will need to compute dedicated weights. For analysis like these it is better to start from the fulldata file.

Also note that is difficult to impossible to recreate the polldata files from the fulldata file, because inclusion in any given day depends on the order of the observations in the raw dailypoll/fulldata file (which differed from day to day). It also varies by the Stata version used (July 10-15 were run in Stata

14, which uses a different random number generator, which affects the bootstrap replication weights). This should not matter for the estimates (which use wgt_main) but it slightly affects standard errors.

To use weights but do not want standard errors, you only need the main weight.

If you want standard errors but don't use Stata and don't know how to use replication weights, you can act as if the weights are design weights and compute robust s.e.'s using these; the calculation will typically be conservative.

Election Topics File

Each week, we asked election poll respondents to answer a very few election-related questions, directly after the vote questions. Some questions were authored by the CESR team, others were asked by non-CESR researchers. These microdata were embargoed until February 28th, 2017 and released on the Daybreak data site on March 1.

We have provided an example Stata "do" file with code for processing and cleaning which data users may freely use, reference or adapt. We have also provided full question wording and context information for all questions. These documents are available at uasdata.usc.edu under "Data/Election Poll 2016".

Each weekly set of questions was answered only once by each respondent who participated in the poll that week. Note that some sets of questions were repeated at intervals. These data may be merged with the *fulldata* file to compare respondents' responses to their candidate preference and turnout predictions from that week. Link these data to the longitudinal vote file data using variables *uasid* (respondent coded ID) and *ts* (time stamp).

Unique identifiers for each observation:

uasid	individual identifier
ts	time stamp of the observation

Panel variables

pollwave/pollweek	Week invited to participate, and week the poll was completed
polldate	Date the poll was completed
polldayno	Sequential day number survey was completed (see Appendix, pg 16)

Data file location: Weekly Election Topics File

Weighting the election topics file: these data can be weighted as they are, or combined with fulldata.dta before constructing weights specific to that week's sample.

Demographics, party affiliation, and other useful analytic measures

Daybreak's data files include basic demographics. Measures such as party registration, party affiliation, political ideology, religiosity, candidate job ratings, etc. were collected at various points in time during the election season, and in several post-election polls, but are not included in *polldata* or *fulldata*.

To create these measures, several data sources must be combined. For example, aggregating party affiliation for most of the UAS panel requires combining three or more sources. We have provided example code for constructing such a variable on the data website.

References

Delavande, Adeline, and Charles F. Manski. 2010. Probabilistic polling and voting in the 2008 presidential election: Evidence from the American Life Panel. Public Opinion Quarterly 74:433–459. doi: 10.1093/poq/nfq019

Gutsche, T. L., Kapteyn, A., Meijer, E., & Weerman, B. (2014). The RAND Continuous 2012 Presidential Election Poll. Public Opinion Quarterly, 78, 233–254. doi: 10.1093/poq/nfu009

Kapteyn, A., Meijer, E., & Weerman, B. (2012). Methodology of the RAND Continuous 2012 Presidential Election Poll (Working Paper No. WR-961). RAND Corporation. doi: 10.2139/ssrn.2146149

Kolenikov, S. (2014). Calibrating survey data using iterative proportional fitting (raking). Stata Journal,14,22–59.

Lu, H., & Gelman, A. (2003). A method for estimating design-based sampling variances for surveys with weighting, poststratification, and raking. Journal of Official Statistics, 19, 133–151.

Valliant,R.,Dever,J.A.,&Kreuter,F. (2013). Practicaltoolsfordesigningandweightingsurvey samples. NewYork,NY:Springer.

Appendix I: Correspondence of of Daybreak polldayno and calendar dates

pollday	date	pollday	date	pollday	date	pollday	date	pollday	date
1	4-Jul	29	1-Aug	60	1-Sep	90	1-Oct	121	1-Nov
2	5-Jul	30	2-Aug	61	2-Sep	91	2-Oct	122	2-Nov
3	6-Jul	31	3-Aug	62	3-Sep	92	3-Oct	123	3-Nov
4	7-Jul	32	4-Aug	63	4-Sep	93	4-Oct	124	4-Nov
5	8-Jul	33	5-Aug	64	5-Sep	94	5-Oct	125	5-Nov
6	9-Jul	34	6-Aug	65	6-Sep	95	6-Oct	126	6-Nov
7	10-Jul	35	7-Aug	66	7-Sep	96	7-Oct	127	7-Nov
8	11-Jul	36	8-Aug	67	8-Sep	97	8-Oct		
9	12-Jul	37	9-Aug	68	9-Sep	98	9-Oct		
10	13-Jul	38	10-Aug	69	10-Sep	99	10-Oct	Polling weeks	
11	14-Jul	39	11-Aug	70	11-Sep	100	11-Oct	week	Start date
12	15-Jul	40	12-Aug	71	12-Sep	101	12-Oct	1	4-Jul
13	16-Jul	41	13-Aug	72	13-Sep	102	13-Oct	2	11-Jul
14	17-Jul	42	14-Aug	73	14-Sep	103	14-Oct	3	18-Jul
15	18-Jul	43	15-Aug	74	15-Sep	104	15-Oct	4	25-Jul
16	19-Jul	44	16-Aug	75	16-Sep	105	16-Oct	5	1-Aug
17	20-Jul	45	17-Aug	76	17-Sep	106	17-Oct	6	8-Aug
18	21-Jul	46	18-Aug	77	18-Sep	107	18-Oct	7	15-Aug
19	22-Jul	47	19-Aug	78	19-Sep	108	19-Oct	8	22-Aug
20	23-Jul	48	20-Aug	79	20-Sep	109	20-Oct	9	29-Aug
21	24-Jul	49	21-Aug	80	21-Sep	110	21-Oct	10	5-Sep
22	25-Jul	50	22-Aug	81	22-Sep	111	22-Oct	11	12-Sep
23	26-Jul	51	23-Aug	82	23-Sep	112	23-Oct	12	19-Sep
24	27-Jul	52	24-Aug	83	24-Sep	113	24-Oct	13	26-Sep
25	28-Jul	53	25-Aug	84	25-Sep	114	25-Oct	14	3-Oct
26	29-Jul	54	26-Aug	85	26-Sep	115	26-Oct	15	10-Oct
27	30-Jul	55	27-Aug	86	27-Sep	116	27-Oct	16	17-Oct
28	31-Jul	56	28-Aug	87	28-Sep	117	28-Oct	17	24-Oct
		57	29-Aug	88	29-Sep	118	29-Oct	18	31-Oct
		58	30-Aug	89	30-Sep	119	30-Oct		
		59	31-Aug			120	31-Oct		

Appendix II: Descriptive Stats for "Fulldata" Longitudinal File

```
file ../../../data/distr/20170117/full2016corrected.dta

. count
50249

. summarize _all
```

Variable	Obs	Mean	Std. Dev.	Min	Max
uasid	0				
wkdaygrp	50249	3.923859	2.011157	1	7
invite	50248	20706.37	35.83266	20639	20765
pollwave	50248	10.20679	5.116616	1	19
pollstart	50249	1.79e+12	3.09e+09	1.78e+12	1.79e+12
pollend	50249	1.79e+12	3.09e+09	1.78e+12	1.79e+12
polldate	50249	20707.51	35.77351	20639	20765
polldayno	50249	69.51386	35.77351	1	127
pollweek	50249	10.38685	5.109163	1	19
ts	0				
dp001_orig	50249	.0782105	2.341261	0	100
dp001	50249	87.14561	28.18996	0	100
dp002_order	48318	1.501118	.5000039	1	2
dp002_1_	50249	41.78063	44.01363	0	100
dp002_2_	50249	42.46972	44.01721	-50	100
dp002_3_	50249	15.73592	29.91904	0	100
dp003_order	0				
dp003_1_	50249	41.12072	24.2526	0	100
dp003_2_	50249	54.34484	24.84465	-50	100
dp003_3_	50249	4.53004	13.86194	0	100
dp004	15887	.2093536	.6428844	0	5
dp005	15887	.2200541	.6779141	0	5
dp006	15887	.148738	.7082868	0	7
dp007	11550	.1641558	.8022374	0	6
dp008	11550	.7014719	7.384078	-1	100
dp009	50249	.92772	9.162441	0	100
dp010	50249	.0149655	.2072724	0	6
dp011	0				
dp011_other	0				
dp012	50249	.0858923	.3481159	0	4
dp013	50249	.001612	.0420549	0	2
dp014	50249	-.0427272	.5086668	-1	72
dp011d1	47	.1489362	.3598746	0	1
dp011d2	47	.106383	.3116605	0	1
dp011d3	47	.0638298	.2470922	0	1
dp011d4	47	.0638298	.2470922	0	1
dp011d5	47	.106383	.3116605	0	1
dp011d6	47	.0851064	.2820567	0	1
dp011d7	47	.4042553	.4960529	0	1
dp011d8	47	.1276596	.3373181	0	1
prob_vote	50249	87.14561	28.18996	0	100
trump_vote	50237	41.79061	44.01415	0	100
clint_vote	50237	42.47987	44.01757	-50	100
other_vote	50237	15.72574	29.90566	0	100
trump_win	50249	41.12072	24.2526	0	100
clint_win	50249	54.34484	24.84465	-50	100
other_win	50249	4.53004	13.86194	0	100

consistent	50249	.999602	.0199466	0	1
prob_trump	50237	37.45357	43.21364	0	100
prob_clint	50237	38.41367	43.4378	-50	100
prob_other	50237	11.27546	24.45974	0	100
early_vote	15887	.2093536	.6428844	0	5
pres_vote	1919	1.821782	.9417468	1	5
partyreg	992	2.382056	1.648226	1	7
ev_method	575	3.297391	1.612314	0	6
ev_diff	575	14.09043	30.13406	-1	100
ev_conf	575	81.07304	28.98398	0	100
ch_vote	358	2.100559	1.285972	1	6
why_change1	47	.1489362	.3598746	0	1
why_change2	47	.106383	.3116605	0	1
why_change3	47	.0638298	.2470922	0	1
why_change4	47	.0638298	.2470922	0	1
why_change5	47	.106383	.3116605	0	1
why_change6	47	.0851064	.2820567	0	1
why_change7	47	.4042553	.4960529	0	1
why_change8	47	.1276596	.3373181	0	1
why_ch_other	0				
regis	3646	1.183763	.6087787	1	4
reg_state	75	1.04	.1972788	1	2
reg_state_~h	3	28.66667	37.85939	2	72
state2reg	75	31.72	18.45342	2	72
regdeadl	75	20742.88	6.843739	20737	20766
afterdeadl	75	.9733333	.1621922	0	1
language	50249	1.001692	.0410944	1	2
uashhid_corr	0				
batch_corr	50249	5.572549	3.818705	1	13
prim_corr	50249	.8589027	.3481254	0	1
sampletype	50249	1.100141	.3728168	1	3
sampt_corr	50249	1.111883	.3994353	1	3
myhhstart	50187	1.77e+12	1.90e+10	1.74e+12	1.79e+12
myhhend	50243	1.79e+12	3.84e+09	1.78e+12	1.79e+12
myhhdate	50187	20437.07	219.6606	20110	20759
citizenus	50187	1	0	1	1
bornus	50181	1.047388	.2124704	1	2
countryborn	2354	129.7256	63.94454	5	237
stateborn	47114	29.09371	15.46186	1	72
dateofbirt~r	50142	1965.817	15.31454	1910	1998
age	50142	49.65181	15.30463	18	106
gender	50186	.4413781	.4965565	0	1
hisplatino	50187	.0763943	.2656306	0	1
hisplatino~p	3791	1.909786	1.450243	1	5
race	50030	1.420907	1.172183	1	6
white	50030	.8746152	.331158	0	1
black	50030	.0937837	.2915305	0	1
nativeamer	50030	.0523486	.2227312	0	1
asian	50030	.023346	.1510014	0	1
pacific	50030	.0093544	.0962656	0	1
education	50187	11.22699	2.194204	1	16
maritalsta~s	49412	2.595867	2.038214	1	6
livewithpa~r	20450	1.728117	.4449406	1	2
anyhhmember	50188	.834562	.371579	0	1
hhmembernu~r	50188	1.716327	1.39886	0	10
hhincome	50105	10.86508	4.070408	1	16
statereside	50163	28.5471	15.91158	1	56

```
 laborstatus |     49420    2.968798    2.449778          1          8
-------------+---------------------------------------------------------
employment~e |     31459    2.150259    .8692354          1          4
   hourswork |     31477    39.61674    11.65807          0        168
workfullpart |     31486     1.18189    .3857603          1          2
     age_cat |     50142    3.024371    1.424193          0          5
  hhsize_cat |     50188    2.094206    .6451475          1          3
-------------+---------------------------------------------------------
     inc_cat |     50105    2.012733    .8131557          1          3
     racethn |     50184    1.443129    .9110631          1          4
      g_educ |     50186    8.420954    5.089606          1         15
   g_age_cat |     50141      8.6098    4.954199          0         15
  hhsize_inc |     50044    22.94645    6.686222         11         33
-------------+---------------------------------------------------------
   vote2012s |     49346    3.289203    .8788365          1          5
    nmisswgt |     50249    .0268861    .2026078          0          4
       en002 |     49574    1.227821    .4598979          1          3
 en003_order |     39299     1.48388    .4997464          1          2
       en003 |     39231    1.645408     .584014          1          3
-------------+---------------------------------------------------------
       en009 |     50249    1.176342    .5615024          1          4
    vote2012 |     48676    1.323876    .8372564          0          3
```

```
. describe _all, fullnames
```

variable name	storage type	display format	value label	variable label
uasid	str9	%9s		Respondent identifier
wkdaygrp	byte	%11.0g	wkday	Assigned weekday group
invite	int	%td		Date invited for poll
pollwave	byte	%8.0g		Week invited for poll (07/04-07/10=1)
pollstart	double	%tc		Date and time poll survey started
pollend	double	%tc		Date and time poll survey completed
polldate	int	%td		Date poll survey completed
polldayno	int	%8.0g		Sequential day number poll completed (07/04=1)
pollweek	byte	%8.0g		Week poll completed (07/04-07/10=1)
ts	str19	%19s		Poll time stamp
dp001_orig	byte	%8.0g		Chance you will vote (orig; 0 if can still vote)
dp001	byte	%8.0g		Chance you will vote (processed)
dp002_order	byte	%15.0g	order2016	Order of candidates on screen
dp002_1_	byte	%8.0g		Prob vote Trump
dp002_2_	byte	%8.0g		Prob vote Clinton
dp002_3_	byte	%8.0g		Prob vote other candidate
dp003_order	byte	%8.0g		Order of candidates on screen
dp003_1_	byte	%8.0g		Prob Trump wins
dp003_2_	byte	%8.0g		Prob Clinton wins
dp003_3_	byte	%8.0g		Prob other candidate wins
dp004	byte	%33.0g	alreadyvote	Voted already
dp005	byte	%22.0g	pres_vote	Who did you vote for
dp006	byte	%31.0g	partyreg	Party registration
dp007	byte	%63.0g	vote_method	Vote method
dp008	byte	%8.0g		Vote difficulty
dp009	byte	%8.0g		Confidence accurate counting
dp010	byte	%22.0g	pres_vote2	Who would you vote for
dp011	str15	%15s		Why change?
dp011_other	str81	%81s		Other reason: which
dp012	byte	%21.0g	regis	Registered to vote

dp013	byte	%20.0g	which_state	Plan to register to vote
dp014	byte	%28.0g	statename	Plan to register to vote in which state
dp011d1	byte	%10.0g		Learned sth about old candidate
dp011d2	byte	%10.0g		Learned sth about new candidate
dp011d3	byte	%10.0g		So. convinced me not vote old cand.
dp011d4	byte	%10.0g		So. convinced me vote new cand.
dp011d5	byte	%10.0g		Sth old candidate did/said
dp011d6	byte	%10.0g		Sth new candidate did/said
dp011d7	byte	%10.0g		No Specific reason
dp011d8	byte	%10.0g		Other
prob_vote	byte	%8.0g		Prob of voting
trump_vote	byte	%8.0g		Cond prob voting for Trump
clint_vote	byte	%8.0g		Cond prob voting for Clinton
other_vote	byte	%8.0g		Cond prob voting for someone else
trump_win	byte	%8.0g		Prob Trump wins
clint_win	byte	%8.0g		Prob Clinton wins
other_win	byte	%8.0g		Prob someone else wins
consistent	byte	%8.0g		Consistent answers
prob_trump	double	%10.0g		Uncond prob voting for Trump
prob_clint	double	%10.0g		Uncond prob voting for Clinton
prob_other	double	%10.0g		Uncond prob voting for someone else
early_vote	byte	%33.0g	alreadyvote	Voted already
pres_vote	byte	%22.0g	pres_vote	Who did you vote for
partyreg	byte	%31.0g	partyreg	Party registration (if voted)
ev_method	byte	%63.0g	vote_method	Vote method
ev_diff	byte	%8.0g		Vote difficulty
ev_conf	byte	%8.0g		Confidence accurate counting
ch_vote	byte	%22.0g	pres_vote2	Who would you vote for
why_change1	byte	%10.0g		Learned sth about old candidate
why_change2	byte	%10.0g		Learned sth about new candidate
why_change3	byte	%10.0g		So. convinced me not vote old cand.
why_change4	byte	%10.0g		So. convinced me vote new cand.
why_change5	byte	%10.0g		Sth old candidate did/said
why_change6	byte	%10.0g		Sth new candidate did/said
why_change7	byte	%10.0g		No Specific reason
why_change8	byte	%10.0g		Other
why_ch_other	str81	%81s		Other reason: which
regis	byte	%21.0g	regis	Registered to vote (if not voted yet)
reg_state	byte	%20.0g	which_state	Plan to register to vote
reg_state_oth	byte	%28.0g	statename	Plan to register to vote in which state
state2reg	byte	%28.0g	statename	State R wants to register
regdeadl	int	%td		Registration deadline
afterdeadl	byte	%8.0g		Passed registration deadline
language	byte	%9.0g	language	Survey language
uashhid_corr	str9	%9s		Household identifier (corrected)
batch_corr	byte	%34.0g	batch	Sampling batch (corrected)
prim_corr	byte	%20.0g	primary_respondent	Primary respondent (corrected)
sampletype	byte	%17.0g	samptype	Sample type
sampt_corr	byte	%27.0g	sampletype	Sample type (corrected)
myhhstart	double	%tc		Date and time MyHH started
myhhend	double	%tc		Date and time MyHH completed

myhhdate	int	%td		Date MyHH completed
citizenus	byte	%8.0g	dummy	US citizen
bornus	byte	%8.0g	dummy	Born in the US
countryborn	int	%48.0g	myhousehold_v115	Country born
stateborn	byte	%28.0g	statename	State born - FIPS coding
dateofbirth_year	int	%8.0g		Year of birth
age	int	%8.0g		Age
gender	byte	%8.0g	gender	Gender - Male
hisplatino	byte	%9.0g		Hispanic or Latino
hisplatino_group	byte	%24.0g	hisplatino_group	Spanish/Hispanic/Latino group
race	byte	%26.0g	race	Race
white	byte	%8.0g		White
black	byte	%8.0g		Black/African Am.
nativeamer	byte	%8.0g		Am.Indian/AK Native
asian	byte	%8.0g		Asian
pacific	byte	%8.0g		HI/Pac.Islander
education	byte	%38.0g	education	Highest level of education
maritalstatus	byte	%25.0g	maritalstatus	Marital status
livewithpartner	byte	%8.0g	dummy	Living with partner
anyhhmember	byte	%9.0g		Whether any other HH member
hhmembernumber	byte	%8.0g		Number of household members besides R
hhincome	byte	%21.0g	myhousehold_v146	Household income
statereside	byte	%28.0g	statename	State residence - FIPS coding
laborstatus	byte	%24.0g	laborstatus	Labor force status
employmenttype	byte	%22.0g	employmenttype	Employment type
hourswork	int	%8.0g		Hours of work per week
workfullpart	byte	%11.0g	workfullpart	Work full-time or part-time
age_cat	byte	%8.0g	age_cat	Age group
hhsize_cat	byte	%9.0g	hhsize_cat	Household size (cat)
inc_cat	byte	%17.0g	inc_cat	Income category
racethn	byte	%30.0g	racethn	Race-ethnicity
g_educ	byte	%20.0g	g_educ	Female x education
g_age_cat	byte	%11.0g	g_age_cat	Female x age
hhsize_inc	byte	%24.0g	hhszinc	Hh size x income
vote2012s	byte	%19.0g	vote2012s	Prior voting stratum
nmisswgt	byte	%8.0g		Number of missing weight variables
en002	byte	%8.0g	YNC	Did you vote in 2012
en003_order	byte	%14.0g	en003o	Order of candidates in UAS47
en003	byte	%8.0g	RDS	Who did you vote for in 2012
en009	byte	%21.0g	regist	Registered to vote
vote2012	byte	%18.0g	vote2012	Voting for president in 2012