

Question - 1

You are given a dataset consisting of US 1/11/2016 Presidential Election Poll Results from various pollsters.

cycle	branch	type	matchup	forecastdate	state	startdate	enddate	pollster	grade	samplesize	population	poll_wt	rawpoll_clinton	rawpoll_trump	rawpoll_johnson	adjpoll_clinton	adjpoll_trump	adjpoll_johnson	poll_id	question_id	createddate
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	U.S.	10/25/2016	10/31/2016	Google Consumer Surveys	B	24316	lv	6.139129	37.69	35.07	6.16	42.6414	40.86509	5.675099	47940	74899	11/1/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	U.S.	10/27/2016	10/30/2016	ABC News/Washington Post	A+	1128	lv	4.197292	45	46	3	43.29659	44.72984	3.401513	47881	74936	11/1/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	Virginia	10/27/2016	10/30/2016	ABC News/Washington Post	A+	1024	lv	3.676856	48	42	6	46.29779	40.72604	6.401513	47880	74934	11/1/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	Florida	10/20/2016	10/24/2016	SurveyUSA	A	1251	lv	3.403096	48	45	2	46.36931	45.30595	1.77773	47465	74252	10/25/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	U.S.	10/20/2016	10/25/2016	Pew Research Center	B+	2120	rv	3.394646	46	40	6	45.32744	42.20888	3.61832	47616	74519	10/27/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	U.S.	10/22/2016	10/25/2016	Fox News/Anderson Robbins Research/Shaw & Company Resear	A	1221	lv	3.346535	44	41	7	44.6508	42.26663	6.114222	47542	74365	10/26/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	U.S.	10/26/2016	10/31/2016	BCG/PPP	A-	1018	lv	3.275062	44.6	43.7	4.2	46.21834	43.56017	3.15359	47675	74925	11/1/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	Pennsylvania	10/25/2016	10/30/2016	Graus Marketing	B-	3017	rv	3.21775	47	44	3	46.8949	43.50333	3.466432	47967	74899	10/31/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	U.S.	10/23/2016	10/27/2016	Issos	A-	1158	lv	3.144926	41.7	36.4	6.3	41.22576	37.24948	6.420006	47643	74599	10/28/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	U.S.	10/29/2016	10/31/2016	The Times-Picayune/Lucid		2600	lv	3.044326	42	40	5	42.21983	41.6954	4.220173	47995	75055	11/1/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	U.S.	10/25/2016	10/31/2016	USC Dornsife/LA Times		3145	lv	2.956711	43.28	46.9		44.53217	43.84845		47874	74923	11/1/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	Florida	10/25/2016	10/27/2016	Siena College	A	815	lv	2.755075	42	46	4	41.81832	47.92262	2.676897	47771	74761	10/30/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	California	10/14/2016	10/23/2016	Public Policy Institute of California	A	1024	lv	2.739877	54	28	5	55.68839	29.50605	3.17051	47545	74374	10/27/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	U.S.	10/28/2016	10/30/2016	Morning Consult		1772	lv	2.729417	42	39	7	43.31551	40.34972	5.823322	47796	74895	10/31/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	North Carolina	10/25/2016	10/26/2016	Marist College	A	780	lv	2.725044	47	41	8	45.20793	42.01937	6.499082	47775	74766	10/30/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	Florida	10/25/2016	10/26/2016	Marist College	A	779	lv	2.72452	45	44	5	43.19458	45.07725	3.499082	47776	74770	10/30/16
2016	President	polls-plus	Clinton vs. Trump vs. Johnson	11/1/16	U.S.	10/20/2016	10/24/2016	GfK Group	B+	1212	lv	2.714895	51	37	6	50.18283	39.33826	5.044833	47538	74345	10/26/16

You are required to do the following:

1) Read the 2016election.csv

2) There are known swing states in the U.S. whose vote are not decisive on a candidate (such as Florida), analyze a swing state:

- Investigate polls month-by-month (beware the year!)

- Each time a candidate was not the popular option that month but becomes the next month, increase the swing of that state and show the swinging months' data.

3) Pollsters have been graded in the dataset (from max A+ to min D, including A, A- etc.), knowing that Donald J. Trump had won the election (**U.S.-wise**), re-calculate the credibility of the pollsters (increase grade if they have guessed correctly, decrease otherwise) (20 pts) and sort them from the most credible to least (15 pts) (use adjpoll column).

P.S. 1: adjpoll is just the raw poll data augmented with the historical data of that state.

P.S. 2: It is common tradition in US elections to ask voters about who they voted for in the election day when they have just exited the voting area.