SCATTERGRAMS

This problem set asks you to construct a number of scattergrams using the following national Presidential election data. (Each case is a Presidential election, with some missing data.)

Election	*GDP*	*POP*	*TURN*	*D2PC*	*DEV*
1928	.	.	.	41	87
1932	.	.	.	59	472
1936	.	.	.	62	523
1940	.	.	.	55	449
1944	.	.	.	54	432
1948	6.6	39	53	52	.
1952	1.7	32	63	45	89
1956	1.2	69	60	42	74
1960	2.9	49	64	50	303
1964	5.7	74	63	62	486
1968	6.2	40	62	49	.
1972	6.1	56	56	38	17
1976	5.9	45	54	51	297
1980	-1.8	21	53	45	49
1984	7.9	52	54	41	13
1988	4.8	51	50	46	112
1992	3.3	38	55	53	370
1996	4.2	57	49	54	379
2000	4.9	58	51	50	267
2004	3.8	47	58	49	252
2008	1.2	31	59	53	365

GDP is real Gross Domestic Product (economic) growth over the Fall, Winter, and Spring quarters preceding the election (e.g., from October 1, 2007 through June 30, 2008) [U.S. Department of Commerce, Bureau of Economic Analysis]. It often serves as a measure of how "good" or "bad" times are.

POP is the incumbent President's approval rating in the first Gallup Poll taken after June 30 of the election year [Gallup Organization].

TURN (Turnout) is the total vote cast for President as a percent of the voting age population (VAP).

D2PC is the percent of the *two-party popular vote* (i.e., *excluding* votes received by *third or minor candidates*) received by the Democratic presidential candidate. Thus the Republican candidate's percent of the vote is simply 100% – *D2PC*.

DEV is the *number of electoral votes* won by the Democratic candidate. (Data for the 1948 and 1968 elections, in which third-party candidates won a substantial number of electoral votes, are excluded.)

Note that all variables other than *DEV* are continuous variables, namely percentages. *GDP* is rounded off to the nearest tenth of a percent. *POP*, *TURN*, and *D2PC* have been rounded off to the nearest whole percent. Drawing scattergrams should convince you that such rounding off hardly affects the character of the scattergram or the apparent association between the variables. *DEV* a discrete variable with a great many possible values namely — namely, all integers from 0 through 538. (Actually, 531 in 1928-1956, 537 in 1960, and 538 in 1964-2008. *Political science trivia question*: why this variation in the total number of electoral votes?) The electoral vote totals have not been adjusted to account for so-called "faithless electors."

Statements #1 through #3 below come from Problem Sets #3A and #9. Statement #4 comes from the First Midterm Test.

#1 High approval ratings boost a President's re-election performance.

#2 Close elections stimulate voting turnout.

#3 When times are bad, incumbent candidates are punished in elections.

#4 High turnout helps Democrats win elections.

1. Does the data presented on the first page provide support for statements #1 through #4 above. To find out, *construct* and *interpret* (with brief verbal comments) an appropriate *scattergram* for each of the four bivariate hypotheses.

2. Presidents are actually elected on the basis of electoral votes, not popular voters. This raises the question of the nature of the relationship between popular votes and electoral votes won by Presidential candidates over the years. Construct a *scattergram* to examine the relationship between the two variables. Comment on how this scattergram may bear on controversies concerning the Electoral College and whether it should be reformed or abolished.

Note 1. Take care to draw your scattergrams reasonably neatly. Label variables and values clearly, so that any reader can understand the scattergram. Use graph paper if possible; otherwise use a ruler to mark off the horizontal and vertical scales in equal intervals. Construct the horizontal and vertical scales so as to minimize "white space" (with no plotted points) in the scattergram.

Note 2. Each plotted point in each scattergram corresponds to one election year (1928, etc.) As a convenience to yourself in constructing the scattergram, to generic readers in interpreting it, and to me in diagnosing it, ***please label each point by election year***.

Note 3. For #3, consider the Presidential candidate of the party that controls the White House to be the incumbent candidate, even if the incumbent President himself is not running for re-election (e.g., in 2008 McCain was the "incumbent candidate"). Likewise, treat #1 as claiming that high approval ratings for the incumbent President (e.g., Clinton in 2000) will boost the election performance of his party's candidate (e.g., Gore in 2000), even if the President himself is not running for re-election. You may have to check an *Information Please Almanac* or some similar source to determine the incumbent status of candidates in some elections.

Note 4. For some of the scattergrams, you will have to "recode" one or more variables into a slightly different variable. Remember the vote data is on a strictly two-party (or major-party) basis, so that the Republican candidate's percent of the vote is always equal to 100% – D2PC.

Note 5. If you want, use SPSS, Excel, or other software to produce the scattergrams. You will find the data on the course web site, and you can download it as (i) a text file, (ii) an SPSS file, or (iii) an Excel file. But unless you are already generally familiar with (or are motivated to learn more about) SPSS, Excel, etc., you will probably find it easier to construct the scattergrams by hand.