

**Assignment 1**  
**Due: in class, Wednesday, February 18**

Here is the question: which US presidents, Republicans or Democrats, are the best stewards of the US economy?

I assembled the data set (below) from the "2008 Economic Report of the President." The report is available online if you are interested:  
[www.gpoaccess.gov/eop/index.html](http://www.gpoaccess.gov/eop/index.html)

Six variables are in the data set:

**YEAR** (1960-2007) I started the data in 1960, as a sort of general beginning to the "modern" government/budget/economic era.

**UNEMPLOY** civilian unemployment rate, in percent.

**GDPGROWTH** percent change, over previous year, in the real gross domestic product. The year 2007 figure is not yet available and was coded as missing in the data set. The SAS code for missing data is a period: "."

**INFLATION** percent change in special consumer price index, december to december.

**SURPLUS** surplus (positive) or deficit (negative) in federal budget, as percent of gross domestic product

**PARTY** political party of president responsible for that budget year, R=republican, D=democratic. One of the most important things that a president does to influence the economy is prepare/propose/haggle/& eventually sign, a federal budget. The federal budget in the first year of a president's term is set by the president in the previous year, so I used this as a time lag. For instance, 1960 was a republican year; Kennedy took office in January, 1961, but he had effectively no budgetary influence until the new federal fiscal year began in the following October. Thus, I assigned 1961 as a republican year as well.

The idea for this data set came from Michael Kinsley, writing for the Los Angeles Times, in a piece published in August, 2004 (this week he reprised his analysis in Slate.com). He took some of these variables and calculated & compared MEANS only! No consideration at all of variability, let alone any attempt at valid statistical inferences!

Your task is to conduct a more rigorous analysis. Using SAS: (1) Enter the data into a SAS data set. (2) Print the data using PROC PRINT. (3) Obtain summary statistics for each of the four economic variables (UNEMPLOY, GDPGROWTH, INFLATION, and SURPLUS) using PROC UNIVARIATE. (4) Perform a two-sample t-test for each of the four economic variables, using PARTY as the categorical (“class”) variable. Perform the tests with two-sided alternative hypotheses, using  $\alpha = 0.05$ . For each test, obtain the output table with the numerical analyses.

*Check your data typing carefully.*

Hand in ***paper copy only, stapled:***

One-page cover sheet with name and typed executive summary of the analysis results, interpreted in your own words.

Printed SAS program(s) and output of SAS analyses

The data:

YEAR	UNEMPLOY	GDPGROWTH	INFLATION	SURPLUS	PARTY
1960	5.5	2.5	1.4	0.1	R
1961	6.7	2.3	0.7	-0.6	R
1962	5.5	6.1	1.3	-1.3	D
1963	5.7	4.4	1.6	-0.8	D
1964	5.2	5.8	1.0	-0.9	D
1965	4.5	6.4	1.9	-0.2	D
1966	3.8	6.5	3.5	-0.5	D
1967	3.8	2.5	3.0	-1.1	D
1968	3.6	4.8	4.7	-2.9	D
1969	3.5	3.1	6.2	0.3	D
1970	4.9	0.2	5.6	-0.3	R
1971	5.9	3.4	3.3	-2.1	R
1972	5.6	5.3	3.4	-2.0	R
1973	4.9	5.8	8.7	-1.1	R
1974	5.6	-0.5	12.3	-0.4	R
1975	8.5	-0.2	6.9	-3.4	R
1976	7.7	5.3	4.9	-4.2	R
1977	7.1	4.6	6.7	-2.7	R
1978	6.1	5.6	9.0	-2.7	D
1979	5.8	3.2	13.3	-1.6	D
1980	7.1	-0.2	12.5	-2.7	D
1981	7.6	2.5	8.9	-2.6	D
1982	9.7	-1.9	3.8	-4.0	R
1983	9.6	4.5	3.8	-6.0	R
1984	7.5	7.2	3.9	-4.8	R
1985	7.2	4.1	3.8	-5.1	R
1986	7.0	3.5	1.1	-5.0	R
1987	6.2	3.4	4.4	-3.2	R
1988	5.5	4.1	4.4	-3.1	R
1989	5.3	3.5	4.6	-2.8	R
1990	5.6	1.9	6.1	-3.9	R
1991	6.8	-0.2	3.1	-4.5	R
1992	7.5	3.3	2.9	-4.7	R
1993	6.9	2.7	2.7	-3.9	R
1994	6.1	4.0	2.7	-2.9	D

1995	5.6	2.5	2.5	-2.2	D
1996	5.4	3.7	3.3	-1.4	D
1997	4.9	4.5	1.7	-0.3	D
1998	4.5	4.2	1.6	0.8	D
1999	4.2	4.5	2.7	1.4	D
2000	4.0	3.7	3.4	2.4	D
2001	4.7	0.8	1.6	1.3	D
2002	5.8	1.6	2.4	-1.5	R
2003	6.0	2.5	1.9	-3.5	R
2004	5.5	3.6	3.3	-3.6	R
2005	5.1	3.1	3.4	-2.6	R
2006	4.6	2.9	2.5	-1.9	R
2007	4.6	.	4.1	-1.2	R