

Homework and midterm review

Overview

Answering additional questions about the exam material

Quick homework 6 review

Permutation tests for multiple means using an F-statistic

Parametric test for multiple means using an F-statistic (ANOVA)

Review of probability functions in R

Questions about anything?



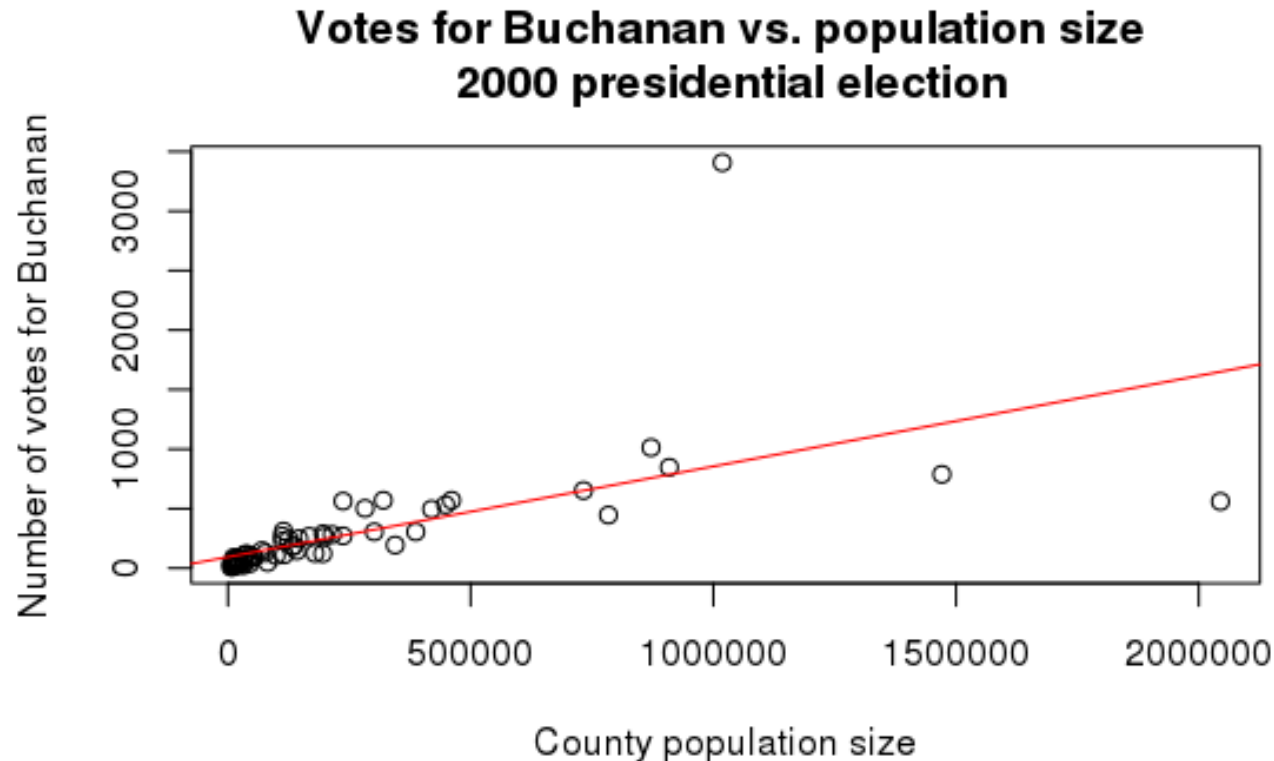
Homework 6 review

COVID-19 county trends app:

https://asterius.hampshire.edu:3939/covid_usa/

Linear regression

In **linear regression** we fit a line to the data, called the **regression line**



$$\hat{y} = b_0 + b_1 \cdot x$$

$$R: \text{lm}(y \sim x)$$

$$b_0 = 47$$

$$b_1 = .0049$$

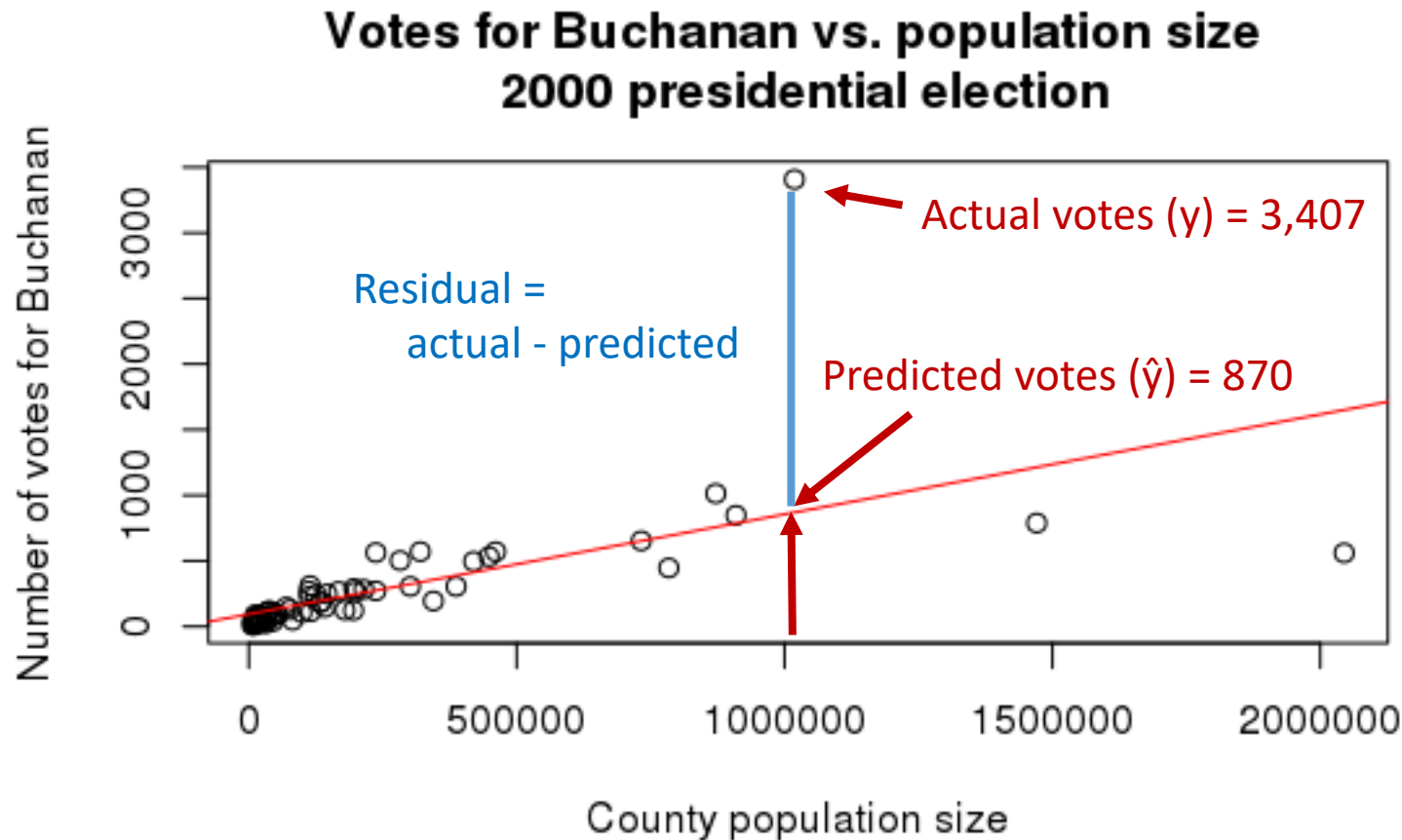
$$\hat{y} = 47 + .0049 \cdot x$$

Residuals

The **residual** at a data value is the difference between the observed (y) and predicted value of the response variable

$$\text{Residual} = \text{Observed} - \text{Predicted} = y - \hat{y}$$

Regression analysis



What is the residual for Palm Beach county?

Actual - predicted

$$= y - \hat{y}$$

$$= 3,407 - 870$$

$$= 2,537 \text{ votes}$$

i.e., Gore likely should have had ~2,500 more votes

Regression caution #3: Be aware of outliers and high leverage points. They can have an huge effect on the regression line.

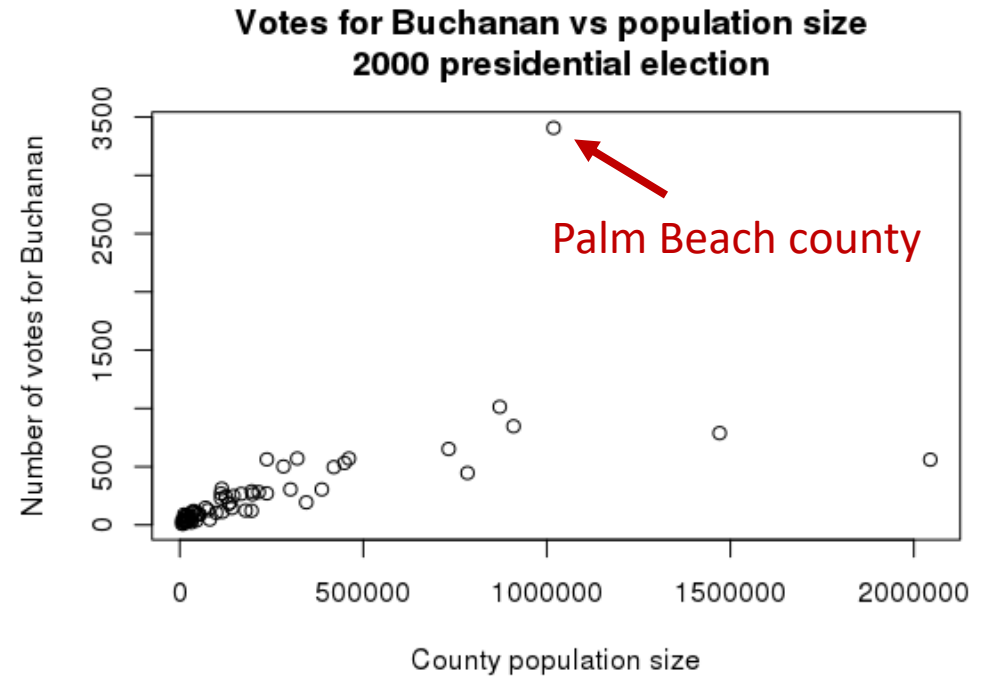
The butterfly ballot

Confusion over Palm Beach Count ballot

Although the Democrats are listed second in the column on the left, they are the third hole on the ballot.

Punching the second hole casts a vote for the Reform Party.

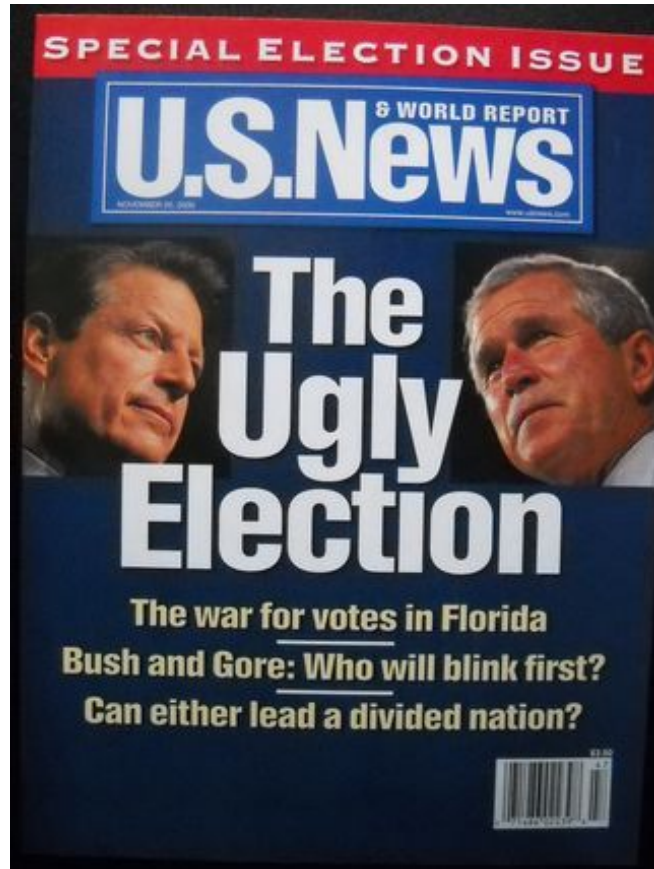
(REPUBLICAN) GEORGE W. BUSH • PRESIDENT DICK CHENEY • VICE PRESIDENT	3 →	(REFORM) PAT BUCHANAN • PRESIDENT EZOLA FOSTER • VICE PRESIDENT	← 4
(DEMOCRATIC) AL GORE • PRESIDENT JOE LIEBERMAN • VICE PRESIDENT	5 →	(SOCIALIST) DAVID McREYNOLDS • PRESIDENT MARY CAL HOLLIS • VICE PRESIDENT	← 6
(LIBERTARIAN) HARRY BROWNE • PRESIDENT ART OLIVIER • VICE PRESIDENT	7 →	(CONSTITUTION) HOWARD PHILLIPS • PRESIDENT J. CURTIS FRAZIER • VICE PRESIDENT	← 8
(GREEN) RALPH NADER • PRESIDENT WINONA LaDUKE • VICE PRESIDENT	9 →	(WORKERS WORLD) MONICA MOOREHEAD • PRESIDENT GLORIA La RIVA • VICE PRESIDENT	← 10
(SOCIALIST WORKERS) JAMES HARRIS • PRESIDENT MARGARET TROWE • VICE PRESIDENT	11 →	WRITE-IN CANDIDATE TO VOTE FOR A WRITE-IN CANDIDATE FOLLOW THE DIRECTIONS ON THE BACK OF YOUR BALLOT CARD.	
(NATURAL LAW) JOHN HAGELIN • PRESIDENT NAT GOLDHABER • VICE PRESIDENT	13 →		



Q: What do we do when we have outliers?

A: Investigate!

Side note: what was the outcome of the 2000 presidential election?



Official final Florida vote count:

Total votes Bush = 2,912,790

Total votes Gore = 2,912,253

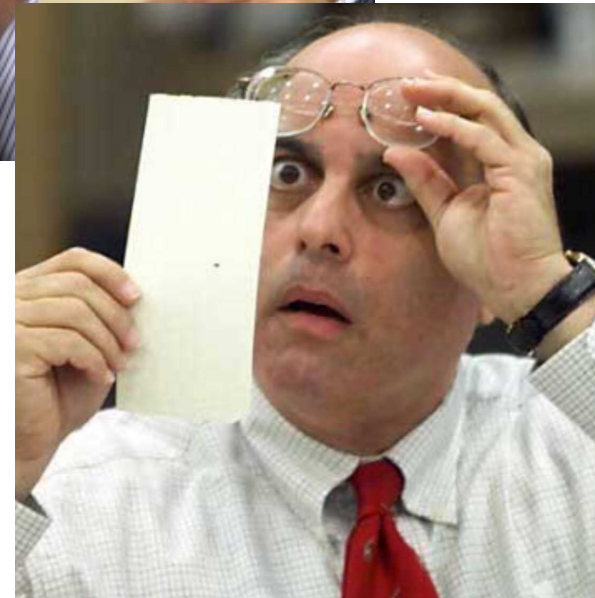
Office vote count difference:

= 2,912,790 - 2,912,253

= 537

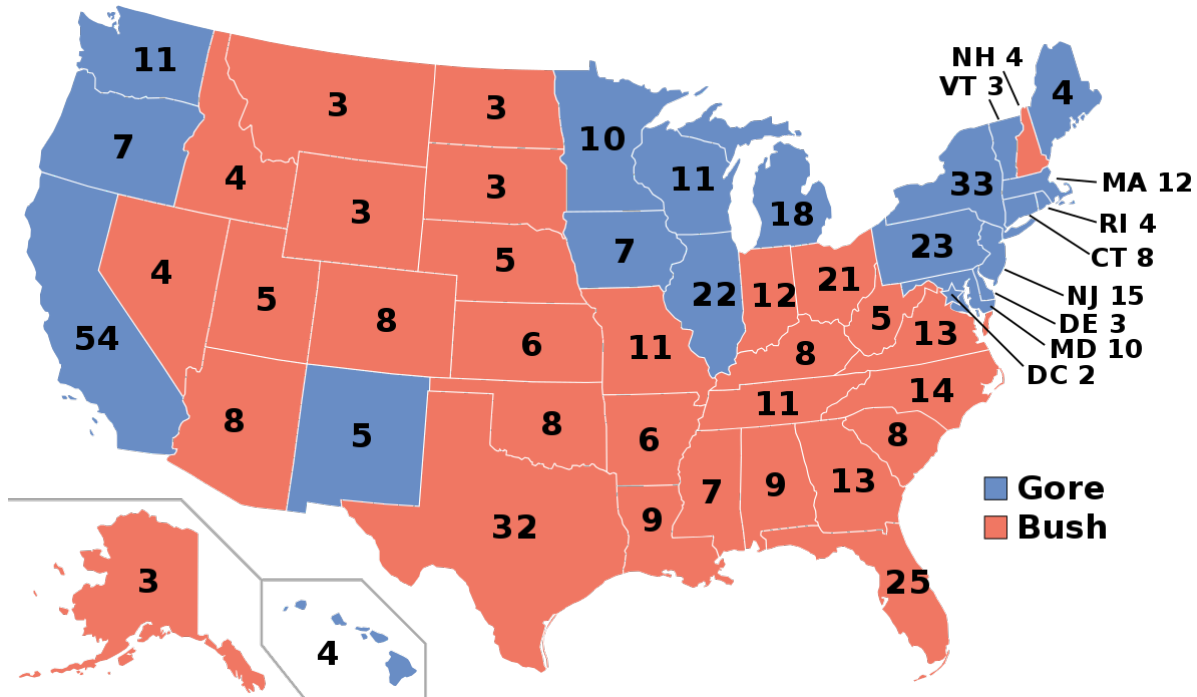
Residual votes: 2,537

Conclusions?



Results of the 2000 presidential election

Electoral vote map for 2000

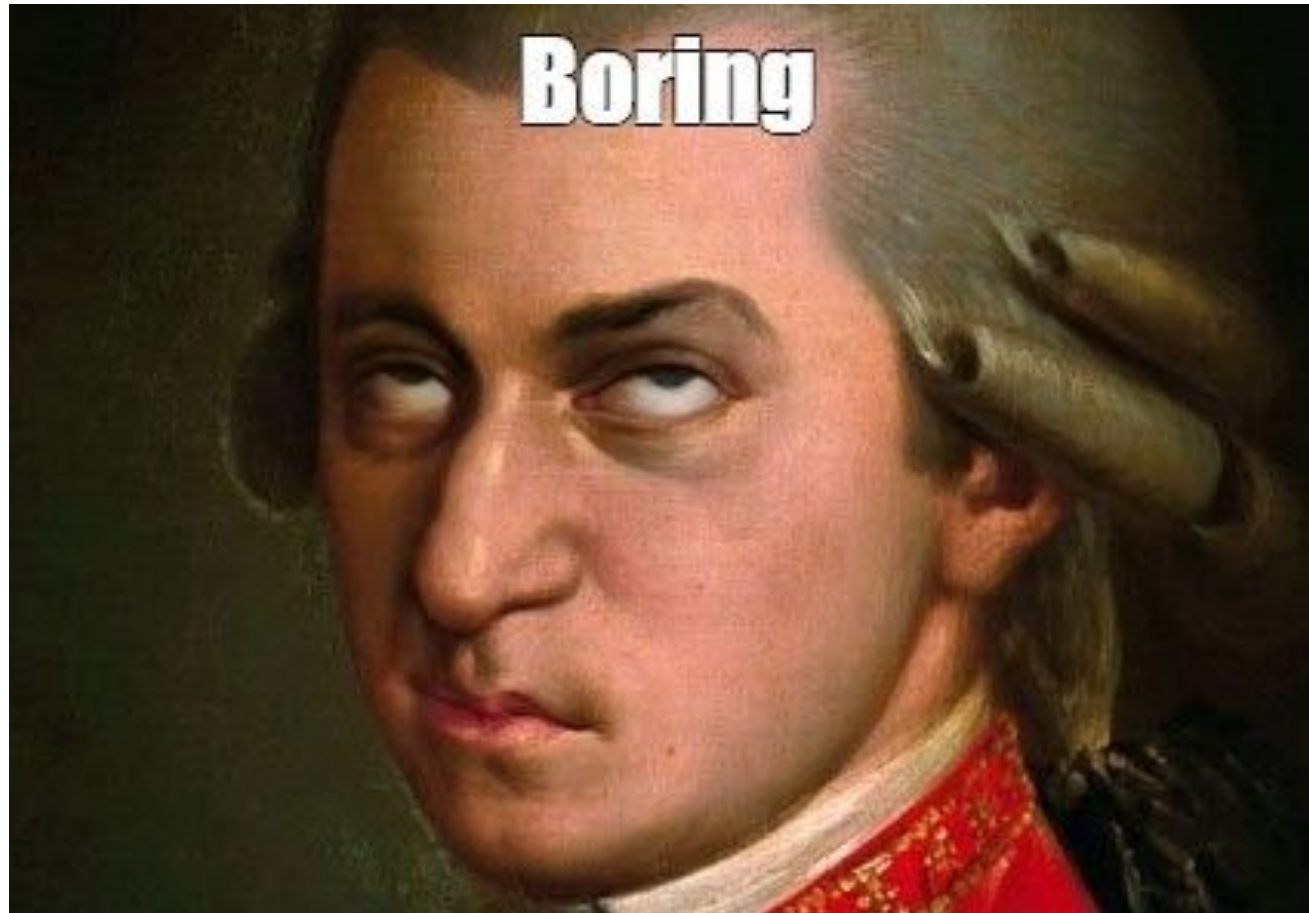


Final official electoral vote count



Conclusions?

Politics is boring!



Hope College and Sudokus!

Let's review the permutation test for comparing more than 2 means

Are the mean completion times for all majors the same?

- Applied science (as)
- Natural science (ns)
- Social science (ss)
- Arts/humanities (ah)

	5	3	2		7			8
6		1	5					2
2			9	1	3		5	
7	1	4	6	9	2			
	2						6	
			4	5	1	2	9	7
	6		3	2	5			9
1					6	3		4
8			1		9	6	7	

Reivew: Densities, probabilities and quantiles for normal distributions

We can plot the density curve using:

`dnorm(x_vec, mu, sigma)`

We can get the probability that we would get a random value less than x using:

`pnorm(x_vec, mu, sigma)`

We can get the quantile values using:

`qnorm(area, mu, sigma)`

