SDS 291: Case Study #1

Due Friday- October 13th End of Day

Data Description

The US presidential election of November 7, 2000 was one of the closest in history. As returns were counted on election night it became clear that the outcome in the state of Florida would determine the next president. At one point in the evening, television networks projected that the state was carried by the Democratic nominee, Al Gore, but a retraction of the projection followed a few hours later. Then, early in the morning of November 8, the networks projected that the Republican nominee, George W. Bush, had carried Florida and won the presidency. Gore called Bush to concede. While on route to his concession speech, though, the Florida count changed rapidly in his favor. The networks once again reversed their projection, and Gore called Bush to retract his concession. When the roughly 6 million Florida votes had been counted, Bush was shown to be leading by only 1,738 votes, and the narrow margin triggered an automatic recount. The recount, completed in the evening of November 9, showed Bush's lead to be less than 400 votes.

Meanwhile, Democratic voters in Palm Beach County complained that a confusing "butterfly" lay-out ballot caused them to accidentally vote for the Reform Party candidate Pat Buchanan instead of Gore. The ballot, as shown below, listed presidential candidates on both a left-hand and right-hand page. Voters were to register their vote by punching the circle corresponding to their choice, from the column of circles between the pages. it was suspected that since Bush's name was listed first on the left-hand page, Bush voters likely selected the first circle. Since Gore's name was listed second on the left-hand side, many voters—who already knew who they wished to vote for—did not bother examining the right-hand side and consequently selected the second circle in the column; the one actually corresponding to Buchanan. Two pieces of evidence supported this claim: Buchanan had an unusually high percentage of the vote in that county, and an unusually large number of ballots (19,000) were discarded because voters had marked two circles (possibly by inadvertently voting for Buchanan and then trying to correct the mistake by then voting for Gore).

The table below shows the first few rows of a data set containing the numbers of votes for Buchanan and Bush in all 67 counties in Florida, followed by a scatterplot of all 67 data points.

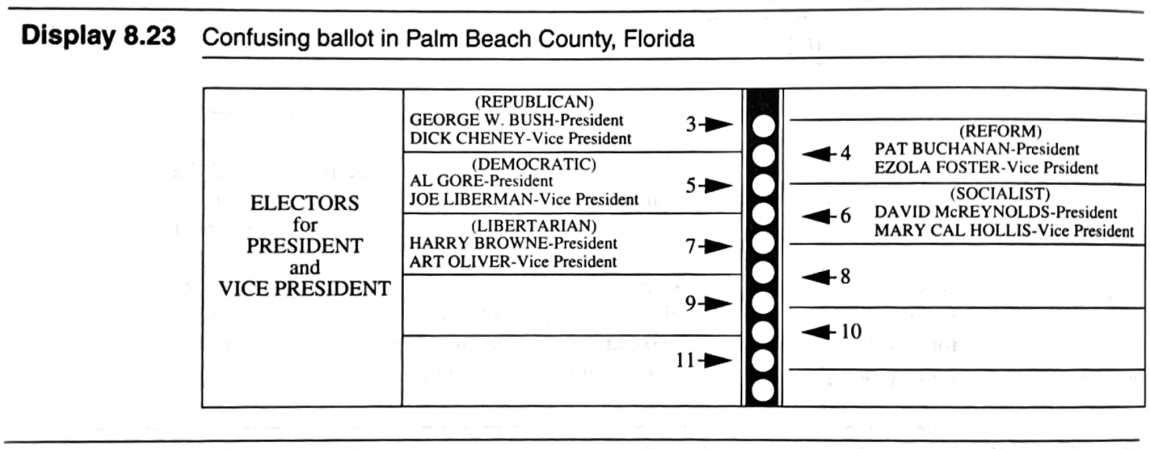
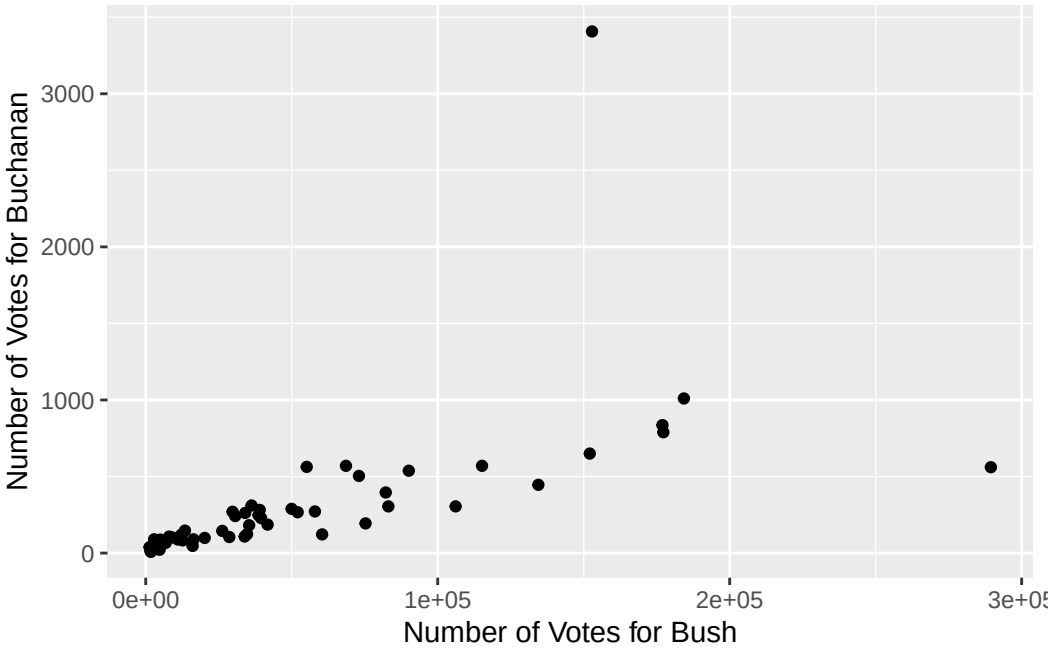


Figure 1: Butterfly Ballot

	County	Buchanan2000	Bush2000
1	Alachua	262	34062
2	Baker	73	5610
3	Bay	248	38637
4	Bradford	65	5413



Your task: Determine whether there is a significant relationship between the number of votes

for Bush and Buchanan.Your final submission must include a linear model with a formal write up interpreting the results. Additionally, you must check for all regression conditions (include code used to check the conditions in an appendix). You must consider whether there are any outliers that should be dropped and whether either of the variables requires a transformation (such as log, square root, etc). Only when you feel the conditions of regression have been met can you estimate the linear model. Don't forget if the transformation is required, then your interpretation of the regression line needs to change accordingly (see week 4 notes and textbook). This final product should be 1-2 pages for the text/output, followed by an appendix that includes a chunk with your condition checks. Your report must include:

- Written overview of the question, descriptive statistics of the variables
- Justification for inclusion/exclusion of outliers
- Decision on whether to transform the data to meet conditions
- Nice regression output (see sample code) and correct interpretation

Note: To access the dataset for this case study, you will need to run the following lines in R; they are also included in the template file for the first case study.

```
# Loading the Statistical Sleuth data library
library(Sleuth2)
# Reading in and saving the data
election <- Sleuth2::ex0825
```