36-402 DA Exam One

James Hawkins (JMHAWKIN)

3/24/2023

# Introduction

1. Clearly state the research questions and objectives of your study.
2. State what data you are using and what information it contains.
3. Briefly mention your final findings. State them in terms that the general public can understand, rather than using technical jargon.

A fair democratic election is at the very core of what the United States was built on. In the 2000 Presidential Election, Palm Beach County had out of the ordinary election results. The proportion of votes for Buchanan that were cast on election day was far higher than what we would expect. It’s hypothesized that this abnormal amount of election day votes is because the butterfly ballot caused voters to mistakenly vote for Buchanan instead of Gore. We are interested if this hypothesis is supported by relevant statistical analysis. For this analysis, we will use a data set containing vote counts from the 2000 presidential election for the 67 counties in Florida. What makes the issue of miscast votes in Palm Beach County so serious in this election is that the battle for Florida between Bush and Gore was close and a deciding factor in Bush’s victory. So, we are also interested in exploring how many votes Buchanan would have received in the absence of the butterfly ballot. We will use a data set containing anonymized voter-level information from Palm Beach County to obtain as estimate for this answer.

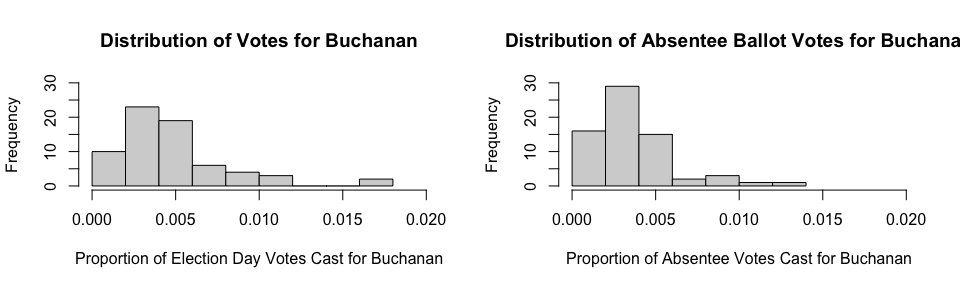
# Exploratory Data Analysis

1. Create four new variables for the county-level data:
2. totalVotes, which takes the total number of election day votes for either Bush, Buchanan, or Gore
3. buchananVotesProp which divides buchananVotes by totalVotes, this is the proportion of election day votes for Buchanan in the county;
4. absBuchananVotesProp which divides absBuchanan by absVotes, this is the proportion of absentee votes for Buchanan in the county;
5. absBuchananDiff which takes the difference between the proportion of election day votes for Buchanan and the proportion of absentee votes for Buchanan
6. Explore the key variables in the county-level data that you need to answer the research questions (e.g. the proportion of election day and absentee votes for Buchanan, the election day votes for other candidates, etc.). Describe them with any necessary univariate EDA, such as plots or summary statistics. Be selective about what exploratory data analysis you include in your report.
7. Identify and specify your response variable in the county-level data, and its distribution.
8. Do multivariate EDA on the county-level data to explore the relationship between predictors and the response.
9. Based on your univariate and multivariate EDA, are there transformations you should use before modeling your data? After any transformations, do the relationships appear to be linear? If you decide to use transformations, use them for the rest of your analysis.
10. If there are plots that would help answer the research questions, by showing key relationships, show those plots, or, if they are already part of your answer to Questions 2-5, reference and comment on those plots in your report.
11. Using the individual ballot-level data, make a table showing the total number of votes for and not for Buchanan, for absentee versus non-absentee ballots and ballots with a vote for Nelson, Deckard, or neither.
12. Describe any trends or interesting features you see in both the county-level and individual ballot-level data that suggest what you will find in the analysis.

## county goreVotes bushVotes buchananVotes absVotes absBuchanan  
## 15 Dixie 1826 2697 29 651 8  
## 23 Gulf 2397 3550 71 1427 16  
## 7 Calhoun 2155 2873 90 664 6  
## 61 Suwannee 4075 8006 108 1664 14  
## 67 Washington 2798 4994 88 1079 9  
## 30 Holmes 2177 5011 76 774 5  
## 66 Walton 5642 12182 120 3084 19  
## 22 Glades 1442 1841 9 507 3  
## 19 Franklin 2046 2454 33 846 5  
## 65 Wakulla 3838 4512 46 1101 6  
## 40 Madison 3014 3038 29 737 4  
## 54 Putnam 12102 13447 148 3086 16  
## 33 Jefferson 3041 2478 29 812 4  
## 42 Marion 44665 55141 563 11470 55  
## 60 Sumter 9637 12127 114 2935 14  
## 14 Desoto 3320 4256 36 1468 7  
## 47 Okeechobee 4588 5057 43 1710 8  
## 12 Columbia 7047 10964 89 2377 11  
## 34 Lafayette 789 1670 10 436 2  
## 63 Union 1407 2332 37 661 3  
## 32 Jackson 6868 9138 102 1791 8  
## 28 Highlands 14167 20206 127 4141 17  
## 9 Citrus 25525 29767 270 10191 40  
## 51 Pasco 69564 68582 570 14571 56  
## 16 Duval 107864 152098 652 31728 121  
## 1 Alachua 47365 34124 263 10694 40  
## 2 Baker 2392 5610 73 1111 4  
## 49 Osceola 28181 26212 145 5581 20  
## 4 Bradford 3075 5414 65 1126 4  
## 26 Hendry 3240 4747 22 1128 4  
## 58 St. Johns 19502 39546 229 9613 34  
## 53 Polk 75200 90295 533 20896 69  
## 21 Gilchrist 1910 3300 29 608 2  
## 37 Leon 61427 39062 282 18648 57  
## 55 Santa Rosa 12802 36274 311 7775 23  
## 3 Bay 18850 38637 248 12587 37  
## 64 Volusia 97304 82357 498 28018 82  
## 52 Pinellas 200630 184825 1013 45116 129  
## 17 Escambia 40943 73017 502 21387 60  
## 43 Martin 26620 33970 112 9980 27  
## 5 Brevard 97318 115185 570 31811 83  
## 27 Hernando 32644 30646 242 9274 24  
## 41 Manatee 49177 57952 271 13441 33  
## 10 Clay 14632 41736 186 7831 19  
## 8 Charlotte 29645 35426 182 9612 23  
## 35 Lake 36571 50010 289 11790 28  
## 36 Lee 73560 106141 305 20365 48  
## 46 Okaloosa 16948 52093 267 14245 32  
## 50 Palm Beach 269732 152951 3411 45830 101  
## 31 Indian River 19768 28635 105 6952 15  
## 18 Flagler 13897 12613 83 3773 8  
## 45 Nassau 6952 16404 90 3632 7  
## 11 Collier 29921 60450 122 13046 25  
## 29 Hillsborough 169557 180760 847 36952 70  
## 56 Sarasota 72853 83100 305 21378 40  
## 62 Taylor 2649 4056 27 1106 2  
## 59 St. Lucie 41559 34705 124 8871 16  
## 6 Broward 387703 177902 795 48525 83  
## 48 Orange 140220 134517 446 38866 64  
## 57 Seminole 59174 75677 194 15466 20  
## 44 Monroe 16483 16059 47 6732 8  
## 13 Miami-Dade 328808 289533 560 41474 45  
## 38 Levy 5398 6858 67 1516 1  
## 20 Gadsden 9735 4767 38 1827 1  
## 24 Hamilton 1722 2146 23 816 0  
## 25 Hardee 2339 3765 30 1056 0  
## 39 Liberty 1017 1317 39 211 0  
## totalVotes buchananVotesProp absBuchananVotesProp absBuchananDiff  
## 15 4552 0.0063708260 0.0122887865 -5.917960e-03  
## 23 6018 0.0117979395 0.0112123336 5.856059e-04  
## 7 5118 0.0175849941 0.0090361446 8.548850e-03  
## 61 12189 0.0088604479 0.0084134615 4.469864e-04  
## 67 7880 0.0111675127 0.0083410565 2.826456e-03  
## 30 7264 0.0104625551 0.0064599483 4.002607e-03  
## 66 17944 0.0066874721 0.0061608301 5.266420e-04  
## 22 3292 0.0027339004 0.0059171598 -3.183259e-03  
## 19 4533 0.0072799471 0.0059101655 1.369782e-03  
## 65 8396 0.0054787994 0.0054495913 2.920815e-05  
## 40 6081 0.0047689525 0.0054274084 -6.584559e-04  
## 54 25697 0.0057594272 0.0051847051 5.747221e-04  
## 33 5548 0.0052271089 0.0049261084 3.010005e-04  
## 42 100369 0.0056093017 0.0047951177 8.141840e-04  
## 60 21878 0.0052107140 0.0047700170 4.406969e-04  
## 14 7612 0.0047293747 0.0047683924 -3.901770e-05  
## 47 9688 0.0044384806 0.0046783626 -2.398820e-04  
## 12 18100 0.0049171271 0.0046276820 2.894451e-04  
## 34 2469 0.0040502228 0.0045871560 -5.369332e-04  
## 63 3776 0.0097987288 0.0045385779 5.260151e-03  
## 32 16108 0.0063322573 0.0044667783 1.865479e-03  
## 28 34500 0.0036811594 0.0041052886 -4.241292e-04  
## 9 55562 0.0048594363 0.0039250319 9.344044e-04  
## 51 138716 0.0041091150 0.0038432503 2.658647e-04  
## 16 260614 0.0025017842 0.0038136662 -1.311882e-03  
## 1 81752 0.0032170467 0.0037404152 -5.233685e-04  
## 2 8075 0.0090402477 0.0036003600 5.439888e-03  
## 49 54538 0.0026586967 0.0035835872 -9.248905e-04  
## 4 8554 0.0075987842 0.0035523979 4.046386e-03  
## 26 8009 0.0027469097 0.0035460993 -7.991896e-04  
## 58 59277 0.0038632184 0.0035368771 3.263413e-04  
## 53 166028 0.0032103019 0.0033020674 -9.176550e-05  
## 21 5239 0.0055354075 0.0032894737 2.245934e-03  
## 37 100771 0.0027984241 0.0030566281 -2.582039e-04  
## 55 49387 0.0062972037 0.0029581994 3.339004e-03  
## 3 57735 0.0042954880 0.0029395408 1.355947e-03  
## 64 180159 0.0027642249 0.0029266900 -1.624651e-04  
## 52 386468 0.0026211743 0.0028592960 -2.381217e-04  
## 17 114462 0.0043857350 0.0028054426 1.580292e-03  
## 43 60702 0.0018450792 0.0027054108 -8.603316e-04  
## 5 213073 0.0026751395 0.0026091604 6.597915e-05  
## 27 63532 0.0038091041 0.0025878801 1.221224e-03  
## 41 107400 0.0025232775 0.0024551745 6.810300e-05  
## 10 56554 0.0032888920 0.0024262546 8.626374e-04  
## 8 65253 0.0027891438 0.0023928423 3.963015e-04  
## 35 86870 0.0033268102 0.0023748940 9.519162e-04  
## 36 180006 0.0016943880 0.0023569850 -6.625971e-04  
## 46 69308 0.0038523691 0.0022464022 1.605967e-03  
## 50 426094 0.0080052758 0.0022037966 5.801479e-03  
## 31 48508 0.0021645914 0.0021576525 6.938933e-06  
## 18 26593 0.0031211221 0.0021203287 1.000793e-03  
## 45 23446 0.0038386079 0.0019273128 1.911295e-03  
## 11 90493 0.0013481706 0.0019162962 -5.681256e-04  
## 29 351164 0.0024119784 0.0018943494 5.176290e-04  
## 56 156258 0.0019519001 0.0018710824 8.081764e-05  
## 62 6732 0.0040106952 0.0018083183 2.202377e-03  
## 59 76388 0.0016232916 0.0018036298 -1.803382e-04  
## 6 566400 0.0014036017 0.0017104585 -3.068568e-04  
## 48 275183 0.0016207397 0.0016466835 -2.594382e-05  
## 57 135045 0.0014365582 0.0012931592 1.433990e-04  
## 44 32589 0.0014422044 0.0011883541 2.538503e-04  
## 13 618901 0.0009048297 0.0010850171 -1.801874e-04  
## 38 12323 0.0054369877 0.0006596306 4.777357e-03  
## 20 14540 0.0026134801 0.0005473454 2.066135e-03  
## 24 3891 0.0059110768 0.0000000000 5.911077e-03  
## 25 6134 0.0048907727 0.0000000000 4.890773e-03  
## 39 2373 0.0164348925 0.0000000000 1.643489e-02  
## bushVotesProp goreVotesProp  
## 15 0.5924868 0.4011424  
## 23 0.5898970 0.3983051  
## 7 0.5613521 0.4210629  
## 61 0.6568217 0.3343178  
## 67 0.6337563 0.3550761  
## 30 0.6898403 0.2996971  
## 66 0.6788899 0.3144226  
## 22 0.5592345 0.4380316  
## 19 0.5413633 0.4513567  
## 65 0.5373988 0.4571224  
## 40 0.4995889 0.4956422  
## 54 0.5232907 0.4709499  
## 33 0.4466474 0.5481255  
## 42 0.5493828 0.4450079  
## 60 0.5543011 0.4404882  
## 14 0.5591172 0.4361534  
## 47 0.5219860 0.4735756  
## 12 0.6057459 0.3893370  
## 34 0.6763872 0.3195626  
## 63 0.6175847 0.3726165  
## 32 0.5672958 0.4263720  
## 28 0.5856812 0.4106377  
## 9 0.5357439 0.4593967  
## 51 0.4944058 0.5014850  
## 16 0.5836141 0.4138841  
## 1 0.4174087 0.5793742  
## 2 0.6947368 0.2962229  
## 49 0.4806190 0.5167223  
## 4 0.6329203 0.3594809  
## 26 0.5927082 0.4045449  
## 58 0.6671390 0.3289978  
## 53 0.5438540 0.4529356  
## 21 0.6298912 0.3645734  
## 37 0.3876314 0.6095702  
## 55 0.7344848 0.2592180  
## 3 0.6692128 0.3264917  
## 64 0.4571351 0.5401007  
## 52 0.4782414 0.5191374  
## 17 0.6379148 0.3576995  
## 43 0.5596191 0.4385358  
## 5 0.5405894 0.4567355  
## 27 0.4823711 0.5138198  
## 41 0.5395903 0.4578864  
## 10 0.7379849 0.2587262  
## 8 0.5429022 0.4543086  
## 35 0.5756878 0.4209854  
## 36 0.5896526 0.4086530  
## 46 0.7516160 0.2445317  
## 50 0.3589607 0.6330340  
## 31 0.5903150 0.4075204  
## 18 0.4742977 0.5225811  
## 45 0.6996503 0.2965111  
## 11 0.6680075 0.3306444  
## 29 0.5147452 0.4828428  
## 56 0.5318128 0.4662353  
## 62 0.6024955 0.3934938  
## 59 0.4543253 0.5440514  
## 6 0.3140925 0.6845039  
## 48 0.4888274 0.5095518  
## 57 0.5603836 0.4381799  
## 44 0.4927736 0.5057842  
## 13 0.4678180 0.5312772  
## 38 0.5565203 0.4380427  
## 20 0.3278542 0.6695323  
## 24 0.5515292 0.4425598  
## 25 0.6137920 0.3813172  
## 39 0.5549937 0.4285714

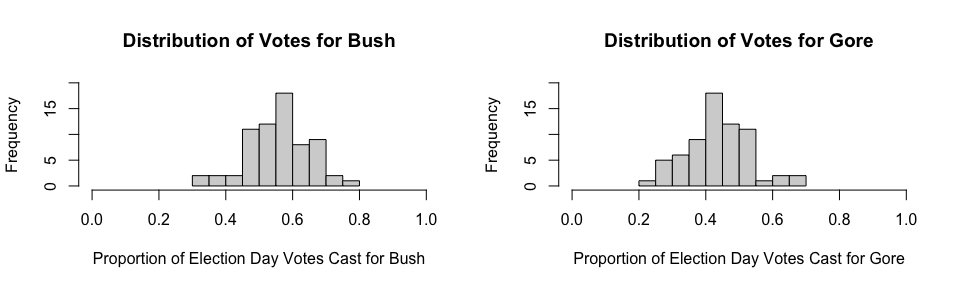
## county goreVotes bushVotes buchananVotes absVotes absBuchanan  
## 39 Liberty 1017 1317 39 211 0  
## 7 Calhoun 2155 2873 90 664 6  
## 24 Hamilton 1722 2146 23 816 0  
## 50 Palm Beach 269732 152951 3411 45830 101  
## 2 Baker 2392 5610 73 1111 4  
## 63 Union 1407 2332 37 661 3  
## 25 Hardee 2339 3765 30 1056 0  
## 38 Levy 5398 6858 67 1516 1  
## 4 Bradford 3075 5414 65 1126 4  
## 30 Holmes 2177 5011 76 774 5  
## 55 Santa Rosa 12802 36274 311 7775 23  
## 67 Washington 2798 4994 88 1079 9  
## 21 Gilchrist 1910 3300 29 608 2  
## 62 Taylor 2649 4056 27 1106 2  
## 20 Gadsden 9735 4767 38 1827 1  
## 45 Nassau 6952 16404 90 3632 7  
## 32 Jackson 6868 9138 102 1791 8  
## 46 Okaloosa 16948 52093 267 14245 32  
## 17 Escambia 40943 73017 502 21387 60  
## 19 Franklin 2046 2454 33 846 5  
## 3 Bay 18850 38637 248 12587 37  
## 27 Hernando 32644 30646 242 9274 24  
## 18 Flagler 13897 12613 83 3773 8  
## 35 Lake 36571 50010 289 11790 28  
## 9 Citrus 25525 29767 270 10191 40  
## 10 Clay 14632 41736 186 7831 19  
## 42 Marion 44665 55141 563 11470 55  
## 23 Gulf 2397 3550 71 1427 16  
## 54 Putnam 12102 13447 148 3086 16  
## 66 Walton 5642 12182 120 3084 19  
## 29 Hillsborough 169557 180760 847 36952 70  
## 61 Suwannee 4075 8006 108 1664 14  
## 60 Sumter 9637 12127 114 2935 14  
## 8 Charlotte 29645 35426 182 9612 23  
## 58 St. Johns 19502 39546 229 9613 34  
## 33 Jefferson 3041 2478 29 812 4  
## 12 Columbia 7047 10964 89 2377 11  
## 51 Pasco 69564 68582 570 14571 56  
## 44 Monroe 16483 16059 47 6732 8  
## 57 Seminole 59174 75677 194 15466 20  
## 56 Sarasota 72853 83100 305 21378 40  
## 41 Manatee 49177 57952 271 13441 33  
## 5 Brevard 97318 115185 570 31811 83  
## 65 Wakulla 3838 4512 46 1101 6  
## 31 Indian River 19768 28635 105 6952 15  
## 48 Orange 140220 134517 446 38866 64  
## 14 Desoto 3320 4256 36 1468 7  
## 53 Polk 75200 90295 533 20896 69  
## 64 Volusia 97304 82357 498 28018 82  
## 13 Miami-Dade 328808 289533 560 41474 45  
## 59 St. Lucie 41559 34705 124 8871 16  
## 52 Pinellas 200630 184825 1013 45116 129  
## 47 Okeechobee 4588 5057 43 1710 8  
## 37 Leon 61427 39062 282 18648 57  
## 6 Broward 387703 177902 795 48525 83  
## 28 Highlands 14167 20206 127 4141 17  
## 1 Alachua 47365 34124 263 10694 40  
## 34 Lafayette 789 1670 10 436 2  
## 11 Collier 29921 60450 122 13046 25  
## 40 Madison 3014 3038 29 737 4  
## 36 Lee 73560 106141 305 20365 48  
## 26 Hendry 3240 4747 22 1128 4  
## 43 Martin 26620 33970 112 9980 27  
## 49 Osceola 28181 26212 145 5581 20  
## 16 Duval 107864 152098 652 31728 121  
## 22 Glades 1442 1841 9 507 3  
## 15 Dixie 1826 2697 29 651 8  
## totalVotes buchananVotesProp absBuchananVotesProp absBuchananDiff  
## 39 2373 0.0164348925 0.0000000000 1.643489e-02  
## 7 5118 0.0175849941 0.0090361446 8.548850e-03  
## 24 3891 0.0059110768 0.0000000000 5.911077e-03  
## 50 426094 0.0080052758 0.0022037966 5.801479e-03  
## 2 8075 0.0090402477 0.0036003600 5.439888e-03  
## 63 3776 0.0097987288 0.0045385779 5.260151e-03  
## 25 6134 0.0048907727 0.0000000000 4.890773e-03  
## 38 12323 0.0054369877 0.0006596306 4.777357e-03  
## 4 8554 0.0075987842 0.0035523979 4.046386e-03  
## 30 7264 0.0104625551 0.0064599483 4.002607e-03  
## 55 49387 0.0062972037 0.0029581994 3.339004e-03  
## 67 7880 0.0111675127 0.0083410565 2.826456e-03  
## 21 5239 0.0055354075 0.0032894737 2.245934e-03  
## 62 6732 0.0040106952 0.0018083183 2.202377e-03  
## 20 14540 0.0026134801 0.0005473454 2.066135e-03  
## 45 23446 0.0038386079 0.0019273128 1.911295e-03  
## 32 16108 0.0063322573 0.0044667783 1.865479e-03  
## 46 69308 0.0038523691 0.0022464022 1.605967e-03  
## 17 114462 0.0043857350 0.0028054426 1.580292e-03  
## 19 4533 0.0072799471 0.0059101655 1.369782e-03  
## 3 57735 0.0042954880 0.0029395408 1.355947e-03  
## 27 63532 0.0038091041 0.0025878801 1.221224e-03  
## 18 26593 0.0031211221 0.0021203287 1.000793e-03  
## 35 86870 0.0033268102 0.0023748940 9.519162e-04  
## 9 55562 0.0048594363 0.0039250319 9.344044e-04  
## 10 56554 0.0032888920 0.0024262546 8.626374e-04  
## 42 100369 0.0056093017 0.0047951177 8.141840e-04  
## 23 6018 0.0117979395 0.0112123336 5.856059e-04  
## 54 25697 0.0057594272 0.0051847051 5.747221e-04  
## 66 17944 0.0066874721 0.0061608301 5.266420e-04  
## 29 351164 0.0024119784 0.0018943494 5.176290e-04  
## 61 12189 0.0088604479 0.0084134615 4.469864e-04  
## 60 21878 0.0052107140 0.0047700170 4.406969e-04  
## 8 65253 0.0027891438 0.0023928423 3.963015e-04  
## 58 59277 0.0038632184 0.0035368771 3.263413e-04  
## 33 5548 0.0052271089 0.0049261084 3.010005e-04  
## 12 18100 0.0049171271 0.0046276820 2.894451e-04  
## 51 138716 0.0041091150 0.0038432503 2.658647e-04  
## 44 32589 0.0014422044 0.0011883541 2.538503e-04  
## 57 135045 0.0014365582 0.0012931592 1.433990e-04  
## 56 156258 0.0019519001 0.0018710824 8.081764e-05  
## 41 107400 0.0025232775 0.0024551745 6.810300e-05  
## 5 213073 0.0026751395 0.0026091604 6.597915e-05  
## 65 8396 0.0054787994 0.0054495913 2.920815e-05  
## 31 48508 0.0021645914 0.0021576525 6.938933e-06  
## 48 275183 0.0016207397 0.0016466835 -2.594382e-05  
## 14 7612 0.0047293747 0.0047683924 -3.901770e-05  
## 53 166028 0.0032103019 0.0033020674 -9.176550e-05  
## 64 180159 0.0027642249 0.0029266900 -1.624651e-04  
## 13 618901 0.0009048297 0.0010850171 -1.801874e-04  
## 59 76388 0.0016232916 0.0018036298 -1.803382e-04  
## 52 386468 0.0026211743 0.0028592960 -2.381217e-04  
## 47 9688 0.0044384806 0.0046783626 -2.398820e-04  
## 37 100771 0.0027984241 0.0030566281 -2.582039e-04  
## 6 566400 0.0014036017 0.0017104585 -3.068568e-04  
## 28 34500 0.0036811594 0.0041052886 -4.241292e-04  
## 1 81752 0.0032170467 0.0037404152 -5.233685e-04  
## 34 2469 0.0040502228 0.0045871560 -5.369332e-04  
## 11 90493 0.0013481706 0.0019162962 -5.681256e-04  
## 40 6081 0.0047689525 0.0054274084 -6.584559e-04  
## 36 180006 0.0016943880 0.0023569850 -6.625971e-04  
## 26 8009 0.0027469097 0.0035460993 -7.991896e-04  
## 43 60702 0.0018450792 0.0027054108 -8.603316e-04  
## 49 54538 0.0026586967 0.0035835872 -9.248905e-04  
## 16 260614 0.0025017842 0.0038136662 -1.311882e-03  
## 22 3292 0.0027339004 0.0059171598 -3.183259e-03  
## 15 4552 0.0063708260 0.0122887865 -5.917960e-03  
## bushVotesProp goreVotesProp  
## 39 0.5549937 0.4285714  
## 7 0.5613521 0.4210629  
## 24 0.5515292 0.4425598  
## 50 0.3589607 0.6330340  
## 2 0.6947368 0.2962229  
## 63 0.6175847 0.3726165  
## 25 0.6137920 0.3813172  
## 38 0.5565203 0.4380427  
## 4 0.6329203 0.3594809  
## 30 0.6898403 0.2996971  
## 55 0.7344848 0.2592180  
## 67 0.6337563 0.3550761  
## 21 0.6298912 0.3645734  
## 62 0.6024955 0.3934938  
## 20 0.3278542 0.6695323  
## 45 0.6996503 0.2965111  
## 32 0.5672958 0.4263720  
## 46 0.7516160 0.2445317  
## 17 0.6379148 0.3576995  
## 19 0.5413633 0.4513567  
## 3 0.6692128 0.3264917  
## 27 0.4823711 0.5138198  
## 18 0.4742977 0.5225811  
## 35 0.5756878 0.4209854  
## 9 0.5357439 0.4593967  
## 10 0.7379849 0.2587262  
## 42 0.5493828 0.4450079  
## 23 0.5898970 0.3983051  
## 54 0.5232907 0.4709499  
## 66 0.6788899 0.3144226  
## 29 0.5147452 0.4828428  
## 61 0.6568217 0.3343178  
## 60 0.5543011 0.4404882  
## 8 0.5429022 0.4543086  
## 58 0.6671390 0.3289978  
## 33 0.4466474 0.5481255  
## 12 0.6057459 0.3893370  
## 51 0.4944058 0.5014850  
## 44 0.4927736 0.5057842  
## 57 0.5603836 0.4381799  
## 56 0.5318128 0.4662353  
## 41 0.5395903 0.4578864  
## 5 0.5405894 0.4567355  
## 65 0.5373988 0.4571224  
## 31 0.5903150 0.4075204  
## 48 0.4888274 0.5095518  
## 14 0.5591172 0.4361534  
## 53 0.5438540 0.4529356  
## 64 0.4571351 0.5401007  
## 13 0.4678180 0.5312772  
## 59 0.4543253 0.5440514  
## 52 0.4782414 0.5191374  
## 47 0.5219860 0.4735756  
## 37 0.3876314 0.6095702  
## 6 0.3140925 0.6845039  
## 28 0.5856812 0.4106377  
## 1 0.4174087 0.5793742  
## 34 0.6763872 0.3195626  
## 11 0.6680075 0.3306444  
## 40 0.4995889 0.4956422  
## 36 0.5896526 0.4086530  
## 26 0.5927082 0.4045449  
## 43 0.5596191 0.4385358  
## 49 0.4806190 0.5167223  
## 16 0.5836141 0.4138841  
## 22 0.5592345 0.4380316  
## 15 0.5924868 0.4011424

## county goreVotes bushVotes buchananVotes absVotes absBuchanan  
## 13 Miami-Dade 328808 289533 560 41474 45  
## 6 Broward 387703 177902 795 48525 83  
## 39 Liberty 1017 1317 39 211 0  
## 49 Osceola 28181 26212 145 5581 20  
## 51 Pasco 69564 68582 570 14571 56  
## 29 Hillsborough 169557 180760 847 36952 70  
## 30 Holmes 2177 5011 76 774 5  
## 50 Palm Beach 269732 152951 3411 45830 101  
## 32 Jackson 6868 9138 102 1791 8  
## 36 Lee 73560 106141 305 20365 48  
## 42 Marion 44665 55141 563 11470 55  
## 57 Seminole 59174 75677 194 15466 20  
## 21 Gilchrist 1910 3300 29 608 2  
## 59 St. Lucie 41559 34705 124 8871 16  
## 52 Pinellas 200630 184825 1013 45116 129  
## 28 Highlands 14167 20206 127 4141 17  
## 54 Putnam 12102 13447 148 3086 16  
## 40 Madison 3014 3038 29 737 4  
## 16 Duval 107864 152098 652 31728 121  
## 38 Levy 5398 6858 67 1516 1  
## 41 Manatee 49177 57952 271 13441 33  
## 20 Gadsden 9735 4767 38 1827 1  
## 53 Polk 75200 90295 533 20896 69  
## 7 Calhoun 2155 2873 90 664 6  
## 1 Alachua 47365 34124 263 10694 40  
## 65 Wakulla 3838 4512 46 1101 6  
## 12 Columbia 7047 10964 89 2377 11  
## 4 Bradford 3075 5414 65 1126 4  
## 60 Sumter 9637 12127 114 2935 14  
## 35 Lake 36571 50010 289 11790 28  
## 61 Suwannee 4075 8006 108 1664 14  
## 56 Sarasota 72853 83100 305 21378 40  
## 67 Washington 2798 4994 88 1079 9  
## 2 Baker 2392 5610 73 1111 4  
## 10 Clay 14632 41736 186 7831 19  
## 26 Hendry 3240 4747 22 1128 4  
## 48 Orange 140220 134517 446 38866 64  
## 18 Flagler 13897 12613 83 3773 8  
## 15 Dixie 1826 2697 29 651 8  
## 31 Indian River 19768 28635 105 6952 15  
## 11 Collier 29921 60450 122 13046 25  
## 27 Hernando 32644 30646 242 9274 24  
## 33 Jefferson 3041 2478 29 812 4  
## 8 Charlotte 29645 35426 182 9612 23  
## 5 Brevard 97318 115185 570 31811 83  
## 22 Glades 1442 1841 9 507 3  
## 45 Nassau 6952 16404 90 3632 7  
## 64 Volusia 97304 82357 498 28018 82  
## 55 Santa Rosa 12802 36274 311 7775 23  
## 58 St. Johns 19502 39546 229 9613 34  
## 62 Taylor 2649 4056 27 1106 2  
## 43 Martin 26620 33970 112 9980 27  
## 66 Walton 5642 12182 120 3084 19  
## 25 Hardee 2339 3765 30 1056 0  
## 63 Union 1407 2332 37 661 3  
## 47 Okeechobee 4588 5057 43 1710 8  
## 34 Lafayette 789 1670 10 436 2  
## 9 Citrus 25525 29767 270 10191 40  
## 37 Leon 61427 39062 282 18648 57  
## 19 Franklin 2046 2454 33 846 5  
## 17 Escambia 40943 73017 502 21387 60  
## 14 Desoto 3320 4256 36 1468 7  
## 46 Okaloosa 16948 52093 267 14245 32  
## 44 Monroe 16483 16059 47 6732 8  
## 24 Hamilton 1722 2146 23 816 0  
## 3 Bay 18850 38637 248 12587 37  
## 23 Gulf 2397 3550 71 1427 16  
## totalVotes buchananVotesProp absBuchananVotesProp absBuchananDiff  
## 13 618901 0.0009048297 0.0010850171 -1.801874e-04  
## 6 566400 0.0014036017 0.0017104585 -3.068568e-04  
## 39 2373 0.0164348925 0.0000000000 1.643489e-02  
## 49 54538 0.0026586967 0.0035835872 -9.248905e-04  
## 51 138716 0.0041091150 0.0038432503 2.658647e-04  
## 29 351164 0.0024119784 0.0018943494 5.176290e-04  
## 30 7264 0.0104625551 0.0064599483 4.002607e-03  
## 50 426094 0.0080052758 0.0022037966 5.801479e-03  
## 32 16108 0.0063322573 0.0044667783 1.865479e-03  
## 36 180006 0.0016943880 0.0023569850 -6.625971e-04  
## 42 100369 0.0056093017 0.0047951177 8.141840e-04  
## 57 135045 0.0014365582 0.0012931592 1.433990e-04  
## 21 5239 0.0055354075 0.0032894737 2.245934e-03  
## 59 76388 0.0016232916 0.0018036298 -1.803382e-04  
## 52 386468 0.0026211743 0.0028592960 -2.381217e-04  
## 28 34500 0.0036811594 0.0041052886 -4.241292e-04  
## 54 25697 0.0057594272 0.0051847051 5.747221e-04  
## 40 6081 0.0047689525 0.0054274084 -6.584559e-04  
## 16 260614 0.0025017842 0.0038136662 -1.311882e-03  
## 38 12323 0.0054369877 0.0006596306 4.777357e-03  
## 41 107400 0.0025232775 0.0024551745 6.810300e-05  
## 20 14540 0.0026134801 0.0005473454 2.066135e-03  
## 53 166028 0.0032103019 0.0033020674 -9.176550e-05  
## 7 5118 0.0175849941 0.0090361446 8.548850e-03  
## 1 81752 0.0032170467 0.0037404152 -5.233685e-04  
## 65 8396 0.0054787994 0.0054495913 2.920815e-05  
## 12 18100 0.0049171271 0.0046276820 2.894451e-04  
## 4 8554 0.0075987842 0.0035523979 4.046386e-03  
## 60 21878 0.0052107140 0.0047700170 4.406969e-04  
## 35 86870 0.0033268102 0.0023748940 9.519162e-04  
## 61 12189 0.0088604479 0.0084134615 4.469864e-04  
## 56 156258 0.0019519001 0.0018710824 8.081764e-05  
## 67 7880 0.0111675127 0.0083410565 2.826456e-03  
## 2 8075 0.0090402477 0.0036003600 5.439888e-03  
## 10 56554 0.0032888920 0.0024262546 8.626374e-04  
## 26 8009 0.0027469097 0.0035460993 -7.991896e-04  
## 48 275183 0.0016207397 0.0016466835 -2.594382e-05  
## 18 26593 0.0031211221 0.0021203287 1.000793e-03  
## 15 4552 0.0063708260 0.0122887865 -5.917960e-03  
## 31 48508 0.0021645914 0.0021576525 6.938933e-06  
## 11 90493 0.0013481706 0.0019162962 -5.681256e-04  
## 27 63532 0.0038091041 0.0025878801 1.221224e-03  
## 33 5548 0.0052271089 0.0049261084 3.010005e-04  
## 8 65253 0.0027891438 0.0023928423 3.963015e-04  
## 5 213073 0.0026751395 0.0026091604 6.597915e-05  
## 22 3292 0.0027339004 0.0059171598 -3.183259e-03  
## 45 23446 0.0038386079 0.0019273128 1.911295e-03  
## 64 180159 0.0027642249 0.0029266900 -1.624651e-04  
## 55 49387 0.0062972037 0.0029581994 3.339004e-03  
## 58 59277 0.0038632184 0.0035368771 3.263413e-04  
## 62 6732 0.0040106952 0.0018083183 2.202377e-03  
## 43 60702 0.0018450792 0.0027054108 -8.603316e-04  
## 66 17944 0.0066874721 0.0061608301 5.266420e-04  
## 25 6134 0.0048907727 0.0000000000 4.890773e-03  
## 63 3776 0.0097987288 0.0045385779 5.260151e-03  
## 47 9688 0.0044384806 0.0046783626 -2.398820e-04  
## 34 2469 0.0040502228 0.0045871560 -5.369332e-04  
## 9 55562 0.0048594363 0.0039250319 9.344044e-04  
## 37 100771 0.0027984241 0.0030566281 -2.582039e-04  
## 19 4533 0.0072799471 0.0059101655 1.369782e-03  
## 17 114462 0.0043857350 0.0028054426 1.580292e-03  
## 14 7612 0.0047293747 0.0047683924 -3.901770e-05  
## 46 69308 0.0038523691 0.0022464022 1.605967e-03  
## 44 32589 0.0014422044 0.0011883541 2.538503e-04  
## 24 3891 0.0059110768 0.0000000000 5.911077e-03  
## 3 57735 0.0042954880 0.0029395408 1.355947e-03  
## 23 6018 0.0117979395 0.0112123336 5.856059e-04  
## bushVotesProp goreVotesProp  
## 13 0.4678180 0.5312772  
## 6 0.3140925 0.6845039  
## 39 0.5549937 0.4285714  
## 49 0.4806190 0.5167223  
## 51 0.4944058 0.5014850  
## 29 0.5147452 0.4828428  
## 30 0.6898403 0.2996971  
## 50 0.3589607 0.6330340  
## 32 0.5672958 0.4263720  
## 36 0.5896526 0.4086530  
## 42 0.5493828 0.4450079  
## 57 0.5603836 0.4381799  
## 21 0.6298912 0.3645734  
## 59 0.4543253 0.5440514  
## 52 0.4782414 0.5191374  
## 28 0.5856812 0.4106377  
## 54 0.5232907 0.4709499  
## 40 0.4995889 0.4956422  
## 16 0.5836141 0.4138841  
## 38 0.5565203 0.4380427  
## 41 0.5395903 0.4578864  
## 20 0.3278542 0.6695323  
## 53 0.5438540 0.4529356  
## 7 0.5613521 0.4210629  
## 1 0.4174087 0.5793742  
## 65 0.5373988 0.4571224  
## 12 0.6057459 0.3893370  
## 4 0.6329203 0.3594809  
## 60 0.5543011 0.4404882  
## 35 0.5756878 0.4209854  
## 61 0.6568217 0.3343178  
## 56 0.5318128 0.4662353  
## 67 0.6337563 0.3550761  
## 2 0.6947368 0.2962229  
## 10 0.7379849 0.2587262  
## 26 0.5927082 0.4045449  
## 48 0.4888274 0.5095518  
## 18 0.4742977 0.5225811  
## 15 0.5924868 0.4011424  
## 31 0.5903150 0.4075204  
## 11 0.6680075 0.3306444  
## 27 0.4823711 0.5138198  
## 33 0.4466474 0.5481255  
## 8 0.5429022 0.4543086  
## 5 0.5405894 0.4567355  
## 22 0.5592345 0.4380316  
## 45 0.6996503 0.2965111  
## 64 0.4571351 0.5401007  
## 55 0.7344848 0.2592180  
## 58 0.6671390 0.3289978  
## 62 0.6024955 0.3934938  
## 43 0.5596191 0.4385358  
## 66 0.6788899 0.3144226  
## 25 0.6137920 0.3813172  
## 63 0.6175847 0.3726165  
## 47 0.5219860 0.4735756  
## 34 0.6763872 0.3195626  
## 9 0.5357439 0.4593967  
## 37 0.3876314 0.6095702  
## 19 0.5413633 0.4513567  
## 17 0.6379148 0.3576995  
## 14 0.5591172 0.4361534  
## 46 0.7516160 0.2445317  
## 44 0.4927736 0.5057842  
## 24 0.5515292 0.4425598  
## 3 0.6692128 0.3264917  
## 23 0.5898970 0.3983051



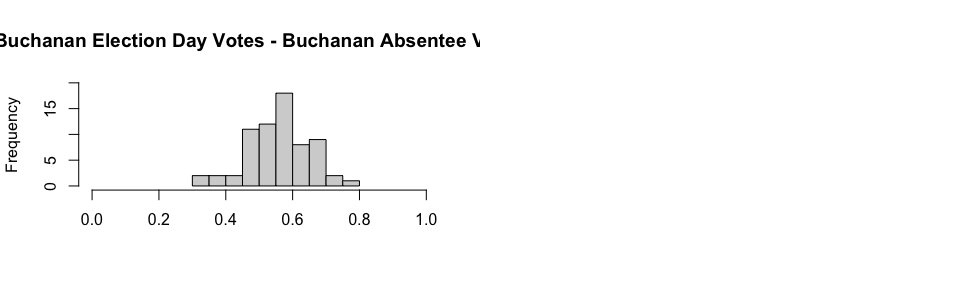
Distribution of Votes for Buchanan

We will begin our exploratory data analysis by looking at the distribution of votes cast for Buchanan on election day and via absentee ballots. In figure 1 we see that both distributions appear somewhat normal, but are right skewed. We also notice that there is an outlier on the histogram showing the distribution of election day votes for Buchanan. However, this data point is not Palm Beach County it is Calhoun county. On the histogram showing the distribution of absentee ballot votes cast for Buchanan, there aren’t any clear outliers, but we do notice a tail on the right side of the distribution. These two counties on the right side of the distribution are Dixie and Gulf county. We note that Palm Beach falls in the middle of the distribution with a proportion of absentee ballot votes going to Buchanan of around 0.0022. Palm Beach County had the fourth highest difference between its proportion of election day voters for Buchanan and proportion of absentee voters for Buchanan. However, it also had the 8th highest proportion of voters who voted on election day.



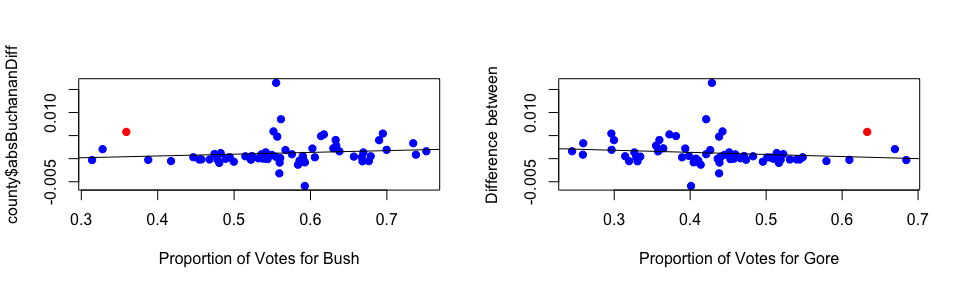
Distribution of Election Day Votes for Bush and Gore

Next, looking at the distributions of the proportions of election day votes for Bush and Gore in figure 2 we see that both appear somewhat normal with no outliers. Both distributions have similar standard deviations of .092 and .093 for Bush and Gore respectively.



Distribution of Difference in Voting Method Frequency

Our Response variable that we are interested in will be the difference between the proportion of election day votes cast for Buchanan and the proportion of absentee votes cast for Buchanan.



Bush and Gore Votes vs buchanan Votes

# Modeling & Diagnostics

1. Construct three models to predict absBuchananDiff from the predictors you have chosen in the county-level data (do not use data from Palm Beach County to fit the model)
2. a linear model (If you decided to use transformations of any variable, use those when fitting your model.)
3. a kernel regression, with bandwidths for each predictor chosen as the sample standard deviation of the predictor divided by n^(1/5), where n is the number of data points.
4. a smoothing spline with a single predictor that you decide is a good choice based on your EDA (use the default value of the smoothing parameter)
5. Present model diagnostic plots. Discuss whether the models appear to fit well, and describe any possible improvements to your models to address any violations of the model assumptions. (Do not use data from Palm Beach County.)
6. Use cross-validation to determine which of the models fit best to the data, in terms of prediction error. (You can use any type of cross-validation you think is suitable.) State the error you estimated for each model and state which model you decided is best. (Do not use data from Palm Beach County.)
7. Comment on whether the difference between the models appears significant, based on the uncertainty in your estimates of the prediction error; a formal test is not required here.
8. For your chosen model, examine the residual diagnostics to determine what type of bootstrap would be appropriate for this data.
9. With the individual ballot-level data, plot the conditional regression function for the probability of voting for Buchanan, conditioned on whether the ballot is absentee and on the senatorial vote (Nelson, Deckard, or neither). (Hint: You can show one figure with 3 regression functions; one for each type of senatorial vote)

# Results

1. Using the selected regression model for the county-level data, construct a bootstrap confidence interval for the expected difference between the proportion of election day votes for Buchanan and the proportion of absentee votes for Buchanan in Palm Beach County. More specifically, choose an appropriate bootstrap procedure to construct a 95% confidence interval for this expected difference.
2. Compare your confidence interval to the observed difference in Palm Beach County.
3. Using the individual ballot-level data, compute the effect of the election day ballot versus the absentee ballot on the proportion of votes for Buchanan, adjusting for senatorial vote and using the empirical senate vote distribution *from election day* (that is, non-absentee votes). Describe what assumptions are required for this to be a valid estimate of a causal effect.
4. Multiply your computed effect (the adjusted difference in vote proportions) by the total number of non-absentee votes to estimate the expected number more or less of Buchanan votes in Palm Beach County in the absence of the butterfly ballot. Describe what assumptions are required for this difference in vote counts to be entirely due to the butterfly ballot (and so be due to miscast votes).
5. Finally, use an appropriate bootstrap procedure to construct a 95% confidence interval of the expected number of votes more or less for Buchanan.

# Conclusions

1. Summarize your main findings about whether Buchanan recieved a surprising number of votes in Palm Beach County.
2. Discuss any limitations to the county-level analysis, such as violations of assumptions or limitations in the data.
3. Summarise how many votes more or less you conclude Buchanan received due to the butterfly ballot.
4. There are certainly limitations to the individual ballot-level analysis. Discuss these limitations and how they might affect the conclusions, such as violations of assumptions or limitations in the data.