FLORIDA'S PARTISAN LEANING*

Using logistic regression model to predict the tendencies of different political parties in Florida and their influences in 2023.

Heng Ma

2024-03-19

In this study, we delve into the dynamics of partisan alignment in Florida—a key swing state known for its fluctuating political orientation in 2023. Amidst a climate of intense political mobilization and rising voter registrations, our investigation targets the influence of minor party affiliation and the proportion of voters with no party affiliation on a county's inclination towards the Republican Party. Utilizing a logistic regression model, apt for dichotomous dependent variables, our findings inch toward statistical significance, suggesting nuanced impacts of these factors on Republican predominance. Despite not reaching conventional significance levels, our model, with an AIC of 56.824, hints at underlying patterns: minor party affiliation has a coefficient estimate of 6.853e-04 ($p = 0.0646$), and no party affiliation shows a coefficient estimate of -6.171e-05 ($p = 0.0641$). These results suggest a marginal yet discernible effect on GOP leanings, providing a framework for understanding the complex influence of voter registration on political tendencies within Florida's counties.

Table of contents

*Code and data are available at:https://github.com/MaEasonH/FLORIDA-S-PARTISAN-LEANING.git

introduction

In the intricate mosaic of American politics, the determination of partisan alignment within various demographic substrata offers invaluable insights for political strategists, policymakers, and social scientists in 2023. As voter registrations surge in the wake of heightened political engagement, understanding the factors that sway county-level party allegiance becomes pivotal. This analysis aims to dissect the affiliations of voters in the state of Florida, a renowned battleground with a propensity to oscillate between Republican and Democratic leanings. Utilizing logistic regression, a statistical technique well-suited for binary outcomes, we examine the extent to which affiliation to minor parties and the presence of non-affiliated voters predict a county's propensity towards the Republican Party. While traditional statistical thresholds of significance remain elusive, our model broaches the cusp of revealing tendencies, offering a narrative on the subtle interplays of voter registrations on party dominance. The article, on the other hand, lists reasons for choosing to use logistic regression curves instead of Poisson and negative binomial regression, based on the analysis results.

we use R [@citeR] for all data wrangling and analysis and R packages tidyverse [@tidy]

Data

Data Sources and variables interest

The data sources it from the official website Florida Department of State. It is responsible for managing statewide election data. It provides election results, voter registration data, various forms, and information on election day, among other things. This data comes from the section "Voter Registration - By County and Party" on the website, and it pertains to the number of election votes for each county and each political party in Florida. For the variables, County: Name of the county, Republican Party of Florida: Number of registered voters for the Republican Party of Florida, Florida Democratic Party: Number of registered voters for the Florida Democratic Party, Minor Party: Number of registered voters for minor parties, No Party Affiliation: Number of voters registered with no party affiliation and Totals: Total number of registered voters.

Results Analized

The results indicate that the Intercept has a coefficient of 1.709e+00 and a standard error of 0.3745, suggesting a higher probability of leaning towards the Republican Party when no predictive variables are considered. For Minor.Party, the coefficient is 6.853e-04 with a standard error of 3.708e-04, a z-value of 1.848, and a corresponding p-value of 0.0646, which is near significance at the 0.1 level. This implies that an increase in the number of minor party voters raises the probability of leaning towards the Republican Party. For No.Party.Affiliation, the coefficient is -6.171e-05 with a standard error of 3.332e-05, a z-value of -1.852, and a p-value of 0.0641, also approaching significance at the 0.1 level. This suggests that an increase in the number of non-affiliated voters reduces the probability of leaning towards the Republican Party.

In conclusion,This model suggests that the number of minor party voters may have a slight positive effect on the partisan leaning at the county level, while the number of non-affiliated voters may have a slight negative effect.None of the coefficients for the variables reached statistical significance at the traditional 0.05 level, but they are close to significance at the 0.1 level, indicating that these variables could have some predictive power regarding partisan leanings.In interpreting these results, we must also consider other potential explanatory variables and conduct appropriate model testing and diagnostics to ensure the robustness and predictive strength of the model.Given the model's AIC value and residual deviance, there may be room for improving the fit of the model, such as by adding additional predictive variables or considering interaction terms.

```
[1] "County"                      "Republican.Party.of.Florida"
[3] "Florida.Democratic.Party"    "Minor.Party"
[5] "No.Party.Affiliation"        "Totals"
```

```
[1] 0 1 1 1 1 0
```

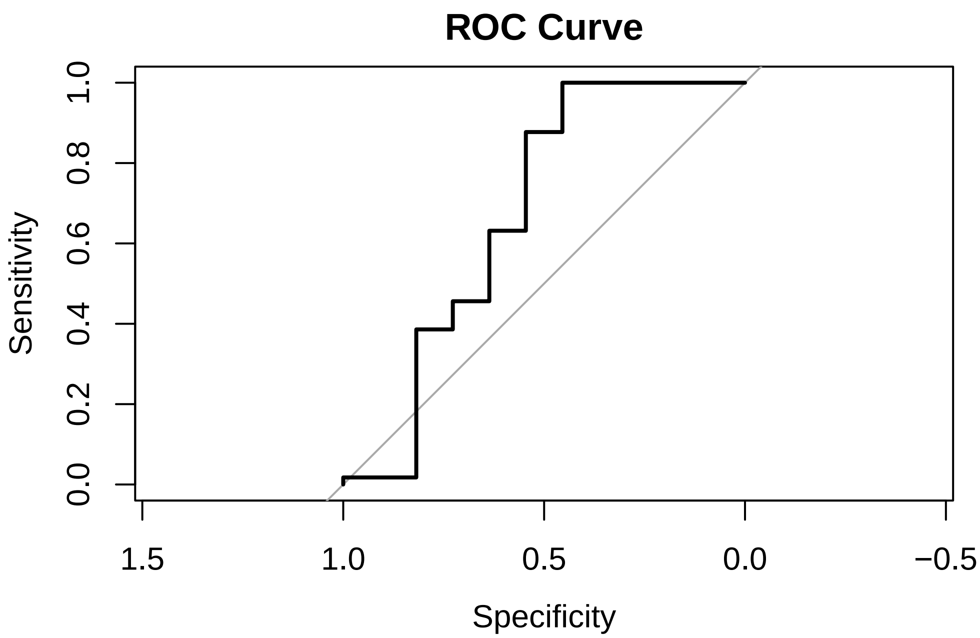
```
Call:
glm(formula = PartyTendency ~ Minor.Party + No.Party.Affiliation,
    family = binomial(), data = data)

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)          1.709e+00  3.745e-01   4.565 5.01e-06 ***
Minor.Party          6.853e-04  3.708e-04   1.848   0.0646 .
No.Party.Affiliation -6.171e-05  3.332e-05  -1.852   0.0641 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

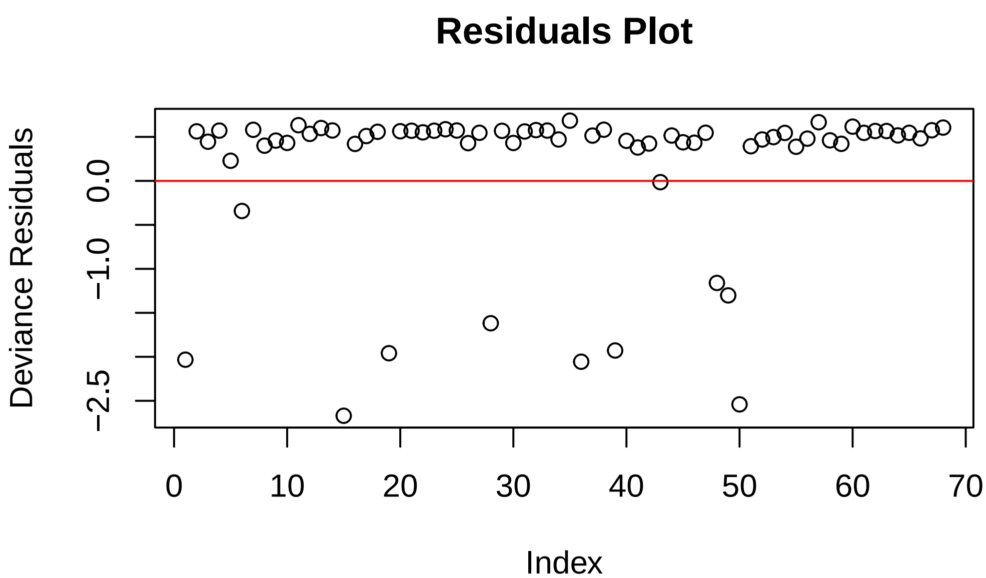
```
(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 60.192  on 67  degrees of freedom
Residual deviance: 50.824  on 65  degrees of freedom
AIC: 56.824

Number of Fisher Scoring iterations: 6
```

Discussion

Why I choose Logistic Regression

The primary objective of this article is to predict partisan tendencies and the subsequent implications. To achieve this, we first need to define how to derive the target variable "partisan tendency" from the existing data. Given that our dataset includes the number of registered voters for different political parties, a straightforward approach would be to create a binary variable indicating whether a party holds the majority of registered voters. For instance, we might compare the number of registered voters for the Republican Party of Florida to those for the Florida Democratic Party to ascertain which party has the majority in each county. Under these circumstances, since our target variable is binary (e.g., 1 for a Republican majority, 0 for a Democratic majority), a logistic regression curve becomes the most efficacious method. This approach allows us to capture the nuances of voter registration dynamics and provides a probabilistic assessment of majority party affiliation at the county level, which is crucial for developing targeted political strategies and understanding voter behavior patterns.

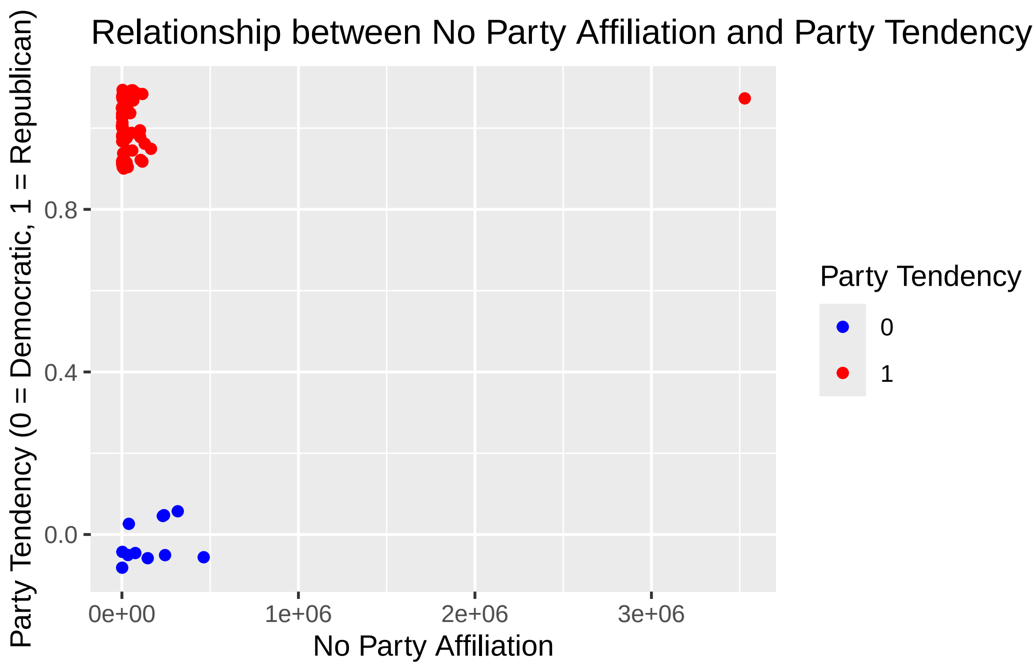




Interpretability

The interpretability of logistic regression lies in its coefficients, which can be understood through the lens of the log-odds ratio. This provides analysts with a clear framework to comprehend the magnitude and direction of the influence that different predictors exert on the target variable—in this case, partisan tendency.

We opt for logistic regression because it is particularly adept at handling binary outcome variables. It enables us to calculate changes in the probability of an event occurring (e.g., a county leaning towards a certain political party) in response to changes in predictor variables. Additionally, logistic regression does not assume a linear relationship between the independent and dependent variables, nor does it require the residuals to be normally distributed. This flexibility makes it a more suitable choice for predicting categorical outcomes based on data that may not meet the stringent assumptions of other regression models. Furthermore, the odds ratios derived from logistic regression are multiplicative, providing an intuitive interpretation of the effect size: for instance, how much more likely a county is to lean Republican with each additional registered voter from minor parties or with changes in the number of non-affiliated voters.



Flexibility

The flexibility of the logistic regression model is a substantial advantage when dealing with complex social science data. It seamlessly accommodates both continuous and categorical predictor variables, allowing for a multifaceted analysis that can reflect the intricacies of real-world data. Moreover, logistic regression can be readily extended to include interaction terms and to model non-linear relationships, which are often encountered in the social sciences where the impact of one variable may depend on the level of another.

Such versatility is particularly important when exploring the electoral dynamics within a politically diverse state like Florida. By using logistic regression, we can capture the nuanced effects of demographics, voter registration status, economic factors, and more, potentially even exploring how these effects might amplify or diminish each other. The ability to model interactions is crucial; for example, the influence of no-party-affiliation voters on partisan leaning might be different in urban counties compared to rural ones. Logistic regression allows us to investigate these subtleties systematically, making it a powerful tool for understanding and predicting the complex phenomenon of partisan alignment.

Why Poisson, and Negative Binomial Regression are not Recommended.

In the context of predicting partisan tendencies, Poisson and negative binomial regression models are not recommended due to their specific suitability for count data, where the response variables represent counts of events or occurrences. These models assume that the data follow a Poisson distribution, or in the case of overdispersion, a negative binomial distribution, which is not applicable when the outcome of interest is binary, as with party majority status. Additionally, these models typically address cases where the counts are independent of each other, an assumption that may not hold in voter data where demographic and social factors create interdependencies. In contrast, logistic regression is explicitly designed for binary outcomes, making it better equipped to handle the prediction of a dichotomous variable such as partisan leaning, where the outcome is the presence or absence of a majority, not a count of occurrences.

Conclusion

In conclusion logistic regression is the analytical method of choice for examining partisan tendencies in Florida due to its compatibility with binary dependent variables, providing clear interpretive value through log-odds ratios. It excels in elucidating the effects of predictors on political leanings and affords substantial flexibility, capable of incorporating both continuous and categorical variables, as well as interactions and nonlinear relationships. This flexibility is vital in the complex arena of social sciences, particularly in a politically variegated context like Florida, where the interplay between demographics, voter registrations, and partisan inclination is intricate. Logistic regression's comprehensive nature enables a nuanced analysis, rendering it an invaluable tool for stakeholders in the realms of political strategy and social research.