California Ballot Results Viewer: 2008-2018
A Web GIS Application for Viewing Ballot Proposition Results in California

by

Adam Araza

A Thesis Presented to the
Faculty of the USC Graduate School
University of Southern California
In Partial Fulfillment of the
Requirements for the Degree
Master of Science
(Geographic Information Science and Technology)

December 2019

To my parents, Renato and Rosario Araza

Table of Contents

List of Figures vii

List of Tables ix

Acknowledgements x

List of Abbreviations xi

Abstract xii

Chapter 1 Introduction 13

1.1. A Background on Ballot Propositions 13

1.1.1. What is a Ballot Proposition? 14

1.1.2. A Brief History of Ballot Propositions in California 15

1.1.3. Recent Propositions Passed in California 15

1.2. Motivation 16

1.2.1. Voter Bubbles 16

1.2.2. Proposition Consolidation 17

1.3. Application Objectives and Overview 18

1.3.1. Target Users 18

1.3.2. Pilot Study Area and Scope 19

1.3.3. Software and Platform Requirements 19

1.3.4. User Requirements 20

1.3.5. Development Process 21

1.4. Thesis Structure 21

Chapter 2 Related Work 22

2.1. Ballot Proposition Studies 22

2.1.1. The Value of Studying Proposition Results 22

2.1.2. The Value of Mapping Proposition Results 24

2.2. Dashboards and GIS 25
2.2.1. GIS on the Desktop 25
2.2.2. GIS on the Web 26
2.2.3. Dashboards 27
2.3. Current Applications 29
2.3.1. Web Maps Displaying Presidential Election Results 29
2.3.2. Web Maps Displaying Propositions Results 32
Chapter 3 Methodology 36
3.1. Requirements and Objectives 36
3.1.1. Functionality 36
3.1.2. Accessibility 37
3.1.3. Software and Platform 37
3.2. Data 37
3.2.1. Spatial County Data 38
3.2.2. Proposition Data 39
3.2.3. Transforming and Merging the Data 39
3.3. Software 40
3.3.1. Microsoft Excel 41
3.3.2. Github Pages 41
3.3.3. Tableau Public 42
3.4. Application Development 47
3.4.1. Data Preparation 48
3.4.2. Data Visualizations and Filters 52
3.4.3. Dashboard Configurations 61
3.4.4. Sharing the Dashboard 65

Chapter 4 Results 68
4.1. Dashboards for Proposition-Based Workflow 68
4.1.1. Filtering Proposition Data 69
4.1.2. Filtering Proposition Data Results 70
4.1.3. Mobile Capabilities 72
4.2. Dashboard for County-Based Workflow 73
4.2.1. Filtering County Data 74
4.2.2. Filtering County Data Results 75
4.2.3. Mobile Capabilities 76
Chapter 5 Discussion 78
5.1. Challenges 78
5.1.1. Layout Challenges 78
5.1.2. Mobile Limitation 80
5.2. Comparison to Available Applications 80
5.2.1. Secretary of State Map – The Basic Information 81
5.2.2. New York Times Map – Symbolizing the Counties by Voter Preference 81
5.2.3. Election Results in 3D – Visualizing the Number of Voters within a County 82
5.2.4. CNN Map of Results– Side-by-Side Visualizations 82
5.2.5. Coloradoan Map– Providing Text Context 82
5.3. Future Development 83
5.3.1. Formalized User Group and Testing 83
5.3.2. Expanding the Data 84
5.3.3. Additional Developments to Support Existing Studies 85
References 87

List of Figures

Figure 1 An Extremely Detailed Map of the 2016 Election by the New York Times 31

Figure 2 Election Results in 3D by Esri... 32

Figure 3 Secretary of State Proposition Results Map - Accessed January 14, 2019 33

Figure 4 Map of Proposition Results from Coloradoan.com 34

Figure 5 CNN Map Displaying Votes on Map and Using a Bar Chart 35

Figure 6 ER Diagram 40

Figure 7 A Screenshot of Data Source Joins in Tableau 43

Figure 8 Data Visualization Choices in Tableau 45

Figure 9 Shared onto Tableau Public Website.. 47

Figure 10 Results Per Proposition (left) and Results by County (right) 49

Figure 11 Raw Data (left) and Formatted Data (right) 50

Figure 12 A Screenshot of a Short Analysis of Proposition 1 from the LAO 51

Figure 13 A Screenshot of Ballotpedia.com 52

Figure 14 Screen Capture of Sheet with Proposition List... 53

Figure 15 Label Screen to Create Proposition Description 54

Figure 16 Screenshot of Map with Lassen County Selected 55

Figure 17 Treemap with Lassen County Selected 57

Figure 18 Horizontal Bar Chart with Results for Each County 58

Figure 19 Vertical Bar Chart with Results per Region 59

Figure 20 Pie Charts Comparing Rural Votes Cast to Metropolitan Votes Cast.......................... 60

Figure 21 Full Proposition Results List 61

Figure 22 General Results Graphics and County-Specific Results Graphics 64

Figure 23 County View, no Filters Applied or Proposition Selected 65

Figure 24 HTML Used to Share Site 67

Figure 25 Dashboard for Proposition-Based Workflows.. 69

Figure 26 Filters Applied to Proposition List Creates Shorter Lists for Users............................ 70

Figure 27 Proposition View with Regional and Designation Overview 71

Figure 28 Proposition View with County Results Filtered to Bay Area Counties 72

Figure 29 Mobile View for Both Views 73

Figure 30 Applying Filters and the Resulting Map 75

Figure 31 A Proposition that had Opposite Statewide Results from the Selected Region 76

Figure 32 Screenshots of Mobile View 77

List of Tables

Table 1 A Select List of Impactful Propositions Passed between 2008 and 2018........................ 16

Table 2 Summary of Data.. 38

Acknowledgements

I would like to thank everyone in the Spatial Sciences Institute faculty and staff for putting together this program. Everything from onboarding, to coursework, to networking events has been of great value. Thanks to my family, my colleagues at the City of Cupertino, and my friends for the encouragement and the support. Lastly, I am thankful for Melly and Waffles for their extraordinary patience.

List of Abbreviations

ALA	Alameda County
CC	Contra Costa County
CSV	Comma-separated value
CSS	Cascading Style Sheets
ER	Entity-Relationship
GIS	Geographic information system
HTML	Hypertext Markup Language
LAO	Legislative Analyst’s Office
SCL	Santa Clara County
SF	San Francisco County
URL	Uniform Resource Locator
USC	University of Southern California

Abstract

Ballot propositions have an important role shaping California's laws. Through these propositions, California voters can directly influence changes to the state's legislation. The importance of understanding ballot propositions has made them the subject of numerous studies. Examples of these studies include research about the influence of voter location, the influence of funding, and the impact of elected official support. Understanding election results can provide value to different groups and organizations; and while there are many different maps and visualizations that help users understand presidential election results, comprehensive data visualizations displaying ballot proposition results are difficult to locate. This thesis details the creation of a web Geographic Information Systems (GIS) application created to help users understand ballot propositions by visualizing results from recent statewide elections (2008-2018). The application incorporates data from multiple sources and presents the data on a single platform in the form of an online dashboard. Users can search for propositions, view a detailed description of propositions, and visualize how different propositions performed across the state.

Chapter 1 Introduction

In California, ballot propositions, also called ballot measures, are proposals to alter legislation with statewide implications. Over the years, California residents have voted on many types of propositions, with impacts including the changing of Daylight Savings Time, legalizing marijuana, adding kindergarten to public schools, adjustments to the state budget, and modifications to the death penalty. This project aims to create an online application that utilizes web Geographic Information Systems (GIS) to help users visualize ballot proposition results in California. These visualizations are intended to help users find insights about proposition results that may not have been discovered from only reading results from a table. The application allows users to search for different propositions and visualize statewide support or opposition by county. Some capabilities and abilities of the application include understanding where a proposition was supported or opposed by visualizing voting results within counties, identifying propositions from multiple elections with similar subjects, and understanding how a county has voted for different propositions over time. This chapter provides more background about ballot propositions, and an overview of the project goals, users, and development process.

1.1. A Background on Ballot Propositions

This application will effectively visualize ballot proposition results in California. To better understand the context of this application, this section provides background information about ballot propositions in California. The section covers three topics. First, a definition of ballot propositions and an overview of the current process to pass a ballot proposition in California. Next, a brief history of ballot propositions in California. The section concludes with an overview of ballot propositions with major impacts in California.

1.1.1. What is a Ballot Proposition?

A ballot measure, or proposition, is a proposal that allows voters to approve new laws or constitutional amendments or repeal existing laws or constitutional amendments (Initiative and Referendum Institute 2019). There are several types of propositions, including initiatives, referenda and legislative measures. As described by University of Southern California's Initiative and Referendum Institute, initiatives are new laws or constitutional amendments proposed by the people, referenda propose repealing an existing law, and a legislative measure is a measure placed on the ballot by the legislature.

The California Secretary of State indicates there are two ways a measure can be placed on the ballot (Secretary of State 2019). First, the legislature can place constitutional amendments, bond measures, and proposed changes in law on the ballot. Second, by following a specific process, any California voter can put an initiative or referendum on the ballot. The California Department of Justice indicates a six-step summary of the process (Department of Justice 2019). The six steps are listed below:

1. Write the text of the proposed law (initiative draft)
2. Submit initiative draft to the Attorney General for official title and summary.
3. Collect enough signatures from registered voters
4. County election officials verify signatures
5. Secretary of State determines whether initiative is qualified for ballot or failed
6. Initiative is approved or denied during election

Between 1912 and 2017, the California Secretary of State reports that 1,996 initiatives were titled and summarized by the Attorney General. Of those, only 376 qualified for the ballot, and of the 376, voters approved 132 (Secretary of State 2019).

1.1.2. A Brief History of Ballot Propositions in California

California adopted state-level initiative and referendum in 1911, after a shift of power from corrupt representatives of Southern Pacific Railroad (Initiatives and Referendum Institute 2019). The Initiatives and Referendum Institute details how prior to 1911, representatives of the Southern Pacific Railroad held positions in the state legislature where they used their positions to accept bribes as a means of doing business. In an attempt to legitimize and create a fair process, doctor and real-estate entrepreneur, Dr. John Randolph Haynes formed political groups with the goal of winning initiative, referendum, and recall in local jurisdictions. After successfully establishing initiative, referendum, and recall in local jurisdiction, Haynes formed another group with the goal of reforming the legislature away from the railroad. A major success came when his group successfully passed a bill that required party candidates to be selected via direct election rather than the previous method of state conventions. The passing of this bill led to the election of governor Hiram Johnson and the ratification of initiative, recall, and referendum at the state and local level in October 1911. This time is known as the progressive movement in California as Johnson and his followers worked to a more equal dispersion of power (Haveman, Rao, and Paruchuri 2007).

1.1.3. Recent Propositions Passed in California

Some ballot measures have had major impacts in shaping California's political landscape. As noted above, propositions were responsible for establishing daylight savings time as well as establishing kindergarten in public schools. The table below lists a selection of impactful propositions passed as recently as 2008.

Table 1 A Select List of Impactful Propositions Passed between 2008 and 2018

Year	Proposition	Subject
2008	Proposition 8	Eliminates Right of Same-Sex Couples to Marry
2012	Proposition 28	Limits on Legislators' Terms in Office
2012	Proposition 36	Changes to Three Strikes Law
2014	Proposition 47	Criminal Sentences, Misdemeanor Penalties
2016	Proposition 56	Cigarette Tax
2016	Proposition 64	Marijuana Legalization
2016	Proposition 66	Death Penalty Procedure Time Limits
2018	Proposition 68	Natural Resources Bond
2018	Proposition 7	Change Daylight Saving Time Period
2018	Proposition 12	Farm Animals Confinement Standards

1.2. Motivation

There were two main motivations behind the development of this project: to visualize voter bubbles and to consolidate proposition results to one interface. These two motivations influenced the decision making behind the design and the function of the project.

1.2.1. Voter Bubbles

Understanding where there may be pockets of voters voting differently from the majority may provide value to different parties. After the presidential election of 2016, there was a general sense of shock for many who supported Senator Hillary Clinton’s presidential bid. This shock could be attributed to physical and digital filter bubbles: a phenomenon that places users in context where they are only interacting with people who agree with their preexisting beliefs (DiFranzo and Gloria-Garcia 2016). Using maps, we may be able to visualize how people in closer geographies vote more similarly than people in further geographies (Tobler 1970). In 2018, The New York Times published an article with a sub-header: “Do you live in a political bubble?” that used a map to visualize concentrations of Donald Trump supporters and Hillary Clinton supporters (Bloch et al. 2018). The map utilized voting results by precinct to show precincts of Trump supporters were typically flanked by more precincts of Trump supporters and the same pattern existed for precincts of Clinton supporters. This map effectively showed “bubbles” of populations that likely only interacted with people of the same belief, and thus Clinton supporters in these populations were likely to be surprised when Donald Trump won the presidency. Understanding where these bubbles exist can help the public understand why some populations were more surprised than others. In addition, understanding these bubbles exist can help a campaign staff member from becoming overly confident, or too worried about a candidate’s performance.

The New York Times map provided users with a way to visualize potential political bubbles during the 2016 Presidential Election. However, upon searching the internet, one is unlikely to find the same type of visualization for ballot proposition results. Through mapping ballot proposition results, it is possible that political bubbles can be discovered that may not appear in presidential elections. Utilizing web GIS, this project aims to fill the void of ballot proposition results maps.

1.2.2. Proposition Consolidation

During a single election, voters cast votes for numerous ballot propositions. Finding information about these propositions can prove to be a difficult task. For researchers studying

California's propositions, California's Secretary of State website at www.sos.ca.gov hosts the official results of previous elections. However, in order to find the results of a specific proposition, the user must first select the correct election years to find the appropriate results. Another resource is the University of California Hastings Law Library at www.uchastings.edu, which hosts a database that allows users to search for California's ballot propositions throughout the years. Users can use keywords or phrases to search different propositions, and the database provides publications with details containing the word or phrase. While this can be useful for providing detailed information, users looking for quick information may find this process time-consuming. This project aims to bridge the results aspect of Secretary of State website with the search functionality of the Hastings database while providing a visual component that allows users to quickly see if the proposition passed or failed.

1.3. Application Objectives and Overview

This project aims to provide a single platform where users can visualize results from different propositions. The main visualization is a map that shows a representation of where propositions were supported and opposed. An additional component is the ability to search for different propositions over the years by using keywords, year on ballot, and result. The rest of this section indicates the scope of the project, the intended users, and provide an overview of the development process.

1.3.1. Target Users

This project is designed to aid campaign researchers and journalists writing about ballot propositions. It is intended to be a supplemental tool for gathering data, putting together graphics, or collecting insights to form research questions. While the primary audience of the tool includes campaign researchers and journalists, the design is simple enough to be used by anyone interested in exploring data about California's propositions.

One specific example of a target user is a member of campaign staff researching how a particular type of proposition has performed in the past. For example, if the campaign involves funding for state roads, by using this tool, the staff member can quickly find a list of propositions involving state roads, and the results for each proposition. The staff member can use the maps provided to determine if specific areas are generally opposed or supportive to fund state roads. Using this data, researchers can dive deeper into what motivates voters in specific areas to provide more targeted advertising.

Another example of a target user is a journalist writing a piece about different political ideologies in California. Since counties in the application will be symbolized by the percentage of support, the journalist can efficiently peruse through different types of propositions to see if there are any issues that may be divisive among regions. The application is intended to lead the researcher to more specific questions based on the results they see on the map. The application can also provide graphics for their piece.

1.3.2. Pilot Study Area and Scope

The initial study area for this project is the State of California. The proposition results are displayed by county. Propositions on ballots between the years 2008 and 2018 are used. Future development of this project may include scales more granular than counties and results for propositions earlier than 2008.

1.3.3. Software and Platform Requirements

This project is intended to be an initial phase for ongoing development. Since the development can potentially span multiple years, an initial requirement is that the platform used to create the application must be free and available without a paid subscription. By using a free platform, the developer is able to continue working on the application without worry of an ongoing subscription cost. In addition, the website that hosts the application should be available for free without a paid subscription. Next, the software should be easy to use and should allow the developer to deploy robust options for data visualization with minimal coding involved. These visualizations should include options for maps, tables, bar graphs, and pie charts. Finally, the software and platform should allow users to access the final result on mobile devices as well as desktop. Users should not have to download any additional software beyond an internet browser, and no training should be required to use the application.

1.3.4. User Requirements

Users should be able to perform two types of workflows using the application: proposition-to-county and county-to-proposition. First, in the proposition-to-county workflow, a user should be able to start by applying filters to a list of propositions, select a proposition of interest, and gather insights about how different counties voted for the specific proposition. For example, a user can search for propositions about marijuana, and find a list of propositions related to marijuana. Selecting a proposition from the list, the user is then be able to visualize where the proposition was supported and opposed throughout the state. Second, in the county-to-proposition workflow, a user can start by selecting a county from the map and see how that county has voted for different propositions over time. For example, a user interested in San Diego County can select San Diego County from the map and see a detailed visualization of how the county has voted for different propositions. Users can filter this list further so a final result can be, for example, a visualization of how San Diego County has voted for bond measures over the last 10 years.

The application should allow users functionality regardless of which workflow is selected. These functionalities include filtering propositions by proposition subject, filtering counties by whether they are rural or metropolitan, seeing responsive results dependent on which filters are activated, visualizing results by the percentage of "Yes" votes, and visualizing results by total votes cast. The components of the application should work together, so any filter should alter a related visualization.

1.3.5. Development Process

This application utilizes a webpage to display spatial information and its associated tabular data. It combines shapefiles and Excel tables on a single display. Once the application is completed, it is hosted on a separate HTML page to ensure optimal usability. A broad overview of the process is listed below:

1. Data acquisition and data configuration
2. Configuration of visualizations
3. Deploying the website

1.4. Thesis Structure

Following this section, chapter 2 provides a discussion of works related to this project. Chapter 2 highlights studies that could potentially benefit from a project of this nature, a discussion about dashboards and web GIS, and concludes with an overview of current applications with similar goals. Chapter 3 covers the methodology required for this project. It reviews the project goals, describes the process of acquiring and transforming data, and provides an overview of how the application was published to the internet. Chapter 4 reports the results from the methodologies described in chapter 3. The thesis concludes with chapter 5, which discusses the project's strengths, challenges, and ideas for further development.

Chapter 2 Related Work

This chapter provides a review of studies related to elections, dashboards, and web GIS. The first sections provide a literature review of work related to the value of studying ballot proposition results. Section 2.1.1 offers a broad overview of different ballot proposition studies and section 2.2.2 reviews ballot proposition studies with a specific spatial component. The next sections discuss web GIS and data visualization dashboards. Web GIS is an essential component of the project and concepts of data visualization dashboards influenced the look and feel of the application. The chapter concludes with Section 2.3 highlighting current applications used for visualizing ballot proposition or other election results.

2.1. Ballot Proposition Studies

Over the years, there has been a wide scope of ballot proposition studies. To determine the potential value of the application, these studies must be reviewed to understand what exactly about ballot propositions is being studied. Section 2.1.1 discusses the general proposition results studies and providing insight into which groups may find value from the information gathered. Section 2.1.2 narrows the scope of studies by reviewing studies where location plays a role in the study. Since the application features a mapping component, including studies that discuss location and proposition results can provide specific direction for what components to include in the application.

2.1.1. The Value of Studying Proposition Results

Different organizations have been able to find value from studying ballot proposition results. These organizations realize the significant influence these propositions can have. This section discusses various aspects of ballot propositions that have been studied and speculates about potential beneficiaries of the information gathered.

One important component of ballot propositions that has been studied is campaign spending. In 1986, a study about the impacts of campaign spending shed light on the relationship between spending and proposition success (Owens and Wade 1986). According to the study, the side which spent more money was more likely to see the results. Understanding how campaign spending impacts proposition success was found to be so valuable, that several additional studies expanded on the work, including a study that analyzed the impacts of spending on opposition advertising versus spending on advocacy advertising (Stratmann 2004). Clearly, organizations deciding whether they want to finance certain campaigns can see the benefit in these studies.

Another aspect of ballot propositions that has been studied is the relationship between politicians and ballot propositions. One study focused on how a candidate's position on a proposition can influence the success of his or her campaign (Ensley and Bucy 2009). According to this study, certain topics yield predictable individual voting behaviors based on the individual's political party. This type of knowledge may prove to beneficial to a candidate who may be on the fence about on a divisive proposition.

A final piece of ballot propositions that has been studied is the relationship between consumers and businesses openly supporting propositions. Numerous studies have indicated that consumers are more willing to reward businesses if the business and consumer have social beliefs that are aligned (Roberts 1995; Becker-Olsen, Cudmore, and Hill 2006). Businesses that choose to publicly support or even fund propositions must be aware of the impacts those decisions will have on their customers.

2.1.2. The Value of Mapping Proposition Results

This section provides an overview of literature that discuss location as a prominent variable in proposition results. Understanding where proposition support or opposition is located can provide valuable data to researchers, campaign managers, and policy-makers.

When studying individual voters, it can be expected that a voter's proximity to a proposition subject may influence the way the voter decides to cast a ballot. In 2007, a study showed that Democrats voted differently for propositions involving immigration depending on their proximity to the Mexican-American border (Branton et al. 2007). This finding was consistent with another study that indicated voters who were closer to Native American lands were more likely to oppose ballot propositions allowing the expansion of gaming casinos than voters who did not live near Native American land (Boehmke et al. 2012). These studies indicate that, depending on the subject, a voter's proximity to a proposition subject may have a stronger influence than the voter's political party. This knowledge can lead to stronger targeted outreach and marketing plans for campaign managers and policy-makers, as Democrats located further from a subject may respond to different outreach than Democrats located near a subject. Understanding the relationship between the individual voters and the subject of a proposition may help contribute to smarter, more targeted, campaign efforts.

Beyond individual voters, finding and determining specific areas of influence within a voting body has also been studied. A study researching ballot propositions in the state of Colorado found that urban and suburban counties had their preference reflected in the final results of ballot initiatives more often than their rural counterparts (Smith 2008). This finding illustrates how concentrated populations have the ability to pass laws that impact the entire state, regardless if the new laws will impact rural communities more than the metropolitan communities. Understanding which areas to target in a campaign can be crucial toward a proposition's success or failure. On the national level, researchers attempted to study the impact of state ballot results on similarly themed congressional votes (Huder, Ragusa, and Smith 2010). The authors indicated that the passage or opposition of a ballot measure at the state level could indirectly influence the way a member of Congress will vote for or against the law involving following the similar issue. Thus, understanding where legislation is supported locally can provide context for how legislation will be supported at a larger scale.

2.2. Dashboards and GIS

Two frameworks that drive this application are web GIS and the data visualization dashboard. GIS is a framework for gathering, managing, and analyzing data (Esri 2019). Web GIS utilizes web GIS servers, which allow web GIS to extend web applications by giving them GIS capabilities (Fu and Sun 2011, 33). A dashboard is a collection of several views, letting users compare a variety of data simultaneously (Tableau 2019). Presenting web GIS in a dashboard allows users to visualize data in multiple formats beyond a traditional web map. Utilizing tables, charts, and additional graphics, a dashboard provides users with a variety of methods to query and analyze a web GIS. Sections 2.2.1 and 2.2.2 discuss the value of GIS and explain how this value is harnessed onto the web. Section 2.2.3 provides further definition for dashboards and overviews the necessary components of dashboards.

2.2.1. GIS on the Desktop

Esri's definition of GIS as a framework provides only a broad overview of its capabilities. A versatile tool, GIS is used in a variety of fields for different applications. Local governments have utilized GIS for analysis, including monitoring land use changes in a city's master plan and determining resident access to local parks (Parsons 2014; Goldsworth 2017). It has also been used in sciences to monitor glacier movement and to determine the spatial extent of specific plant species (Davidson 2014; Klemens 2017). These examples all utilized traditional desktop GIS in the form of Esri ArcGIS Desktop to accomplish their goals. While desktop GIS is powerful and heavily in use, it is not without its limitations.

Some limitations to a traditional desktop GIS software are access and ability to share information (Alesheiskh et al. 2004). Obtaining GIS software can range from free to thousands of dollars depending on the software selected. In addition, since GIS relies on graphic representation to show geographic data, the hardware must be capable of rendering graphics and handling large datasets. Once the software and hardware have been obtained, sharing data between users will depend on how many licenses were obtained and on how many machines will have the software installed. Despite these challenges, traditional desktop GIS is still powerful and heavily in use; however, the ease of access to the internet has allowed web GIS to mitigate these limitations to an extent.

2.2.2. GIS on the Web

In 1993, the Xerox Corporation developed the first interactive spatial tool to be used over the web (Dragicevic 2004). Since then, as internet access has become more available, web GIS has allowed for easier access, sharing, analyzing, and exploring spatial data. Since web GIS is hosted online, users can interact with maps and data directly through their internet browser on their mobile device, laptop, or computer. This access greatly improves the ability to access and share data and has led to even more uses. Web GIS can be used to educate users about regulations to help make informed decisions, as Phillip Conner accomplished with his web GIS of Florida fisheries, or as Brendan Blee accomplished with his web GIS mapping drone regulations (Blee 2016; Conner 2018). Cities can also use web GIS as an asset management system; allowing employees to work on live data from the field (Gerhardt 2011). As the availability of web GIS increases, the notion of using GIS as media to communicate and share knowledge with others becomes more common as major social media platforms like Facebook and Twitter utilize location services as a means to adding value to their users (Sui and Goodchild 2011). This versatility and ease of being able to access and engage data and access from any internet-connected device increases the general public's awareness of the value of GIS (Butler 2006). As web GIS makes it easier for users to produce and share data, there is an increased importance of presenting data in a meaningful way.

2.2.3. Dashboards

As businesses and organizations generate more data, decisions must be made with how data can be efficiently presented. Smart data visualizations allow users of different levels of expertise to obtain insights efficiently and effectively (Wijk 2005). In the GIS community, a popular data visualization is the map, and often, the same data can be visualized off the map as pie charts, bar graphs, scatter plots, etc. To provide additional value in data exploration, studies have shown the value of combining multiple representations of the same data on a single screen (Tufte 2006; Roberts 2007). This combination of multiple data visualizations can be achieved by using a data dashboard.

There have been various definitions for data dashboards; however, most sources agree that a dashboard simultaneously presents different views of the same dataset on a single screen (Few 2006; Pappas and Whitman 2011). Pappas and Whitman categorize dashboards into three categories dependent on the dashboard's purpose: analytical, operational, and strategic. Strategic dashboards are intended to assist executives understand a company's performance relative to the company's goals, operational dashboards are intended to monitor a company's goods or activities at a current moment to ensure operations activities are happening as expected, and analytical dashboards are intended to allow for more in-depth exploration of data, prompting users to interact with the multiple visualizations to find insights into a dataset. The dashboard for exploring proposition results is an analytical dashboard, and this distinction is important for deciding how the data is displayed and what choices are made regarding user interaction.

Incorporating the different data visualizations can greatly impact the usefulness of a dashboard. Studies have shown that different visualizations perform better than others when searching a dataset for anomalies in a dataset like gaps, outliers, or spikes (Correll et al. 2019). In the case of the proposition dataset, visualizations like pie charts may be useful for displaying 'yes' votes against 'no' votes but may not be as effective as displaying the dispersion of those votes over a region, or even over a selected time period. Proper care must be taken when deciding which types of visualizations to use, what colors to incorporate, and how to arrange the visualizations on the dashboard. These decisions of visualization can lead to more insightful interactions with the data.

The ability for users to interact with a dashboard to find new insights is a powerful feature. As with visualizations, there are different types of interactions that can be incorporated into a dashboard. One study classified common interactions as follows: selection, exploration, reconfiguration, encoding, abstraction/elaboration, filtering, and connection (Yi, Kang, and Stasko 2007). In their study, Yi et al. provide the following quick explanations for each interaction:

- Select: mark something as interesting
- Explore: show me something else
- Reconfigure: show me a different arrangement
- Encode: show me a different representation

- Abstract/Elaborate: show me more or less detail
- Filter: show me something conditionally
- Connect: show me related items

Understanding how and when to use these interactions contributes to an effective dashboard. However, equally important is signifying how the interactions, or affordances, can be accomplished. A common error seen in design is not providing users enough instruction, clues or direction to perform the affordances provided by the object (Norman 2013). On a dashboard, these signifiers can be simple text on the dashboard indicating where or what a user is to do. For example, in a text box that allows for filtering of propositions by name, signifying text can be placed above the text box that reads "filter proposition by name." As multiple visualizations allow users to quickly interpret data on a dashboard, signifiers allow users to quickly and easily decipher the capabilities of the dashboard.

2.3. Current Applications

This section provides an overview of general web maps displaying election results and potential dashboard solutions that can potentially be used for this application. Election results are often displayed in web maps. Maps covering both the presidential election and proposition results are reviewed, and the strengths and weaknesses of each application are covered.

2.3.1. Web Maps Displaying Presidential Election Results

Web maps displaying presidential election results are a familiar sight to most voters. During most presidential elections, news broadcasters often use maps to indicate which candidate is leading in a particular area. Once the elections have been finalized, other media sources often use the results to create their own maps. The New York Times published an aptly titled map called "An Extremely Detailed Map of the 2016 Election" which displays presidential election results by precinct (Bloch, et al. 2018). This map, as seen in Figure 1, uses a red to blue diverging color ramp, with higher percentages of voters represented by deeper shades of red or blue. Inspection of the site's source code reveals the map was made using Mapbox GL, a JavaScript library that utilizes vector tiles and Mapbox Styles (Mapbox 2019). One potential area lacking in this map is the ability to quickly visualize the number of votes cast. Though users can hover over a precinct to see how many people voted for each candidate, since the color is only dependent on the percentage of votes for a candidate, two precincts can look the same but have huge differences in the number of votes cast. This potentially could lead to misinformation, if precincts with lower population densities are voting one way, while precincts with high population densities are voting another, as mentioned regarding the rural vs metropolitan counties in Colorado (Smith 2008).

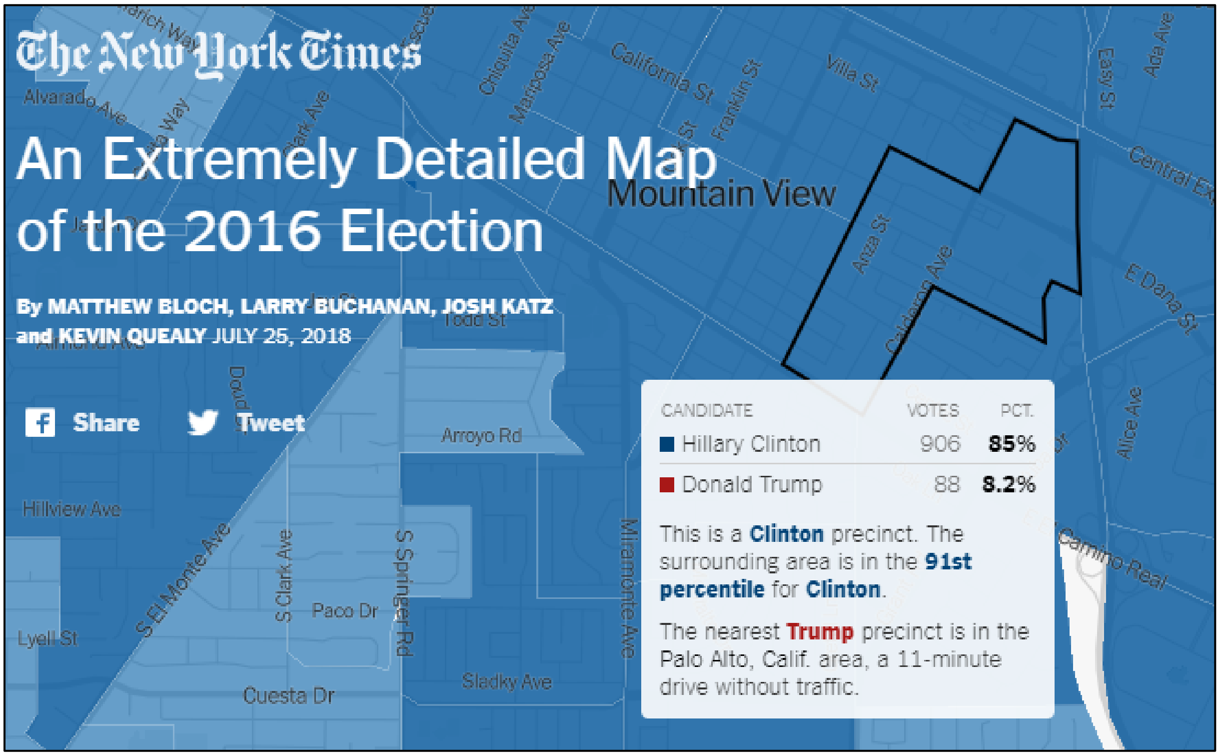


Figure 1 An Extremely Detailed Map of the 2016 Election by the New York Times

Visualizing the number of votes cast can create a more complete visualization of voting results. One way to visualize the number of votes cast is through 3D representations. Using a 3D modeling engine called City Engine, Esri produced a map that uses heights to visualize the number of voters while using a similar red to blue color ramp to indicate candidate preference (Field 2017). The result, as seen in Figure 2, is a map that allows users to visualize where there were many votes cast and which candidate was preferred. However, as this is a 3D model, loading this map takes a noticeably longer to load in a browser than the New York Times map, as well as putting more strain on computational resources. In addition, users may not be as familiar with maneuvering around a 3D map as opposed to the more common 2D map. Regardless, this map still shows the value of visualizing the number of voters.

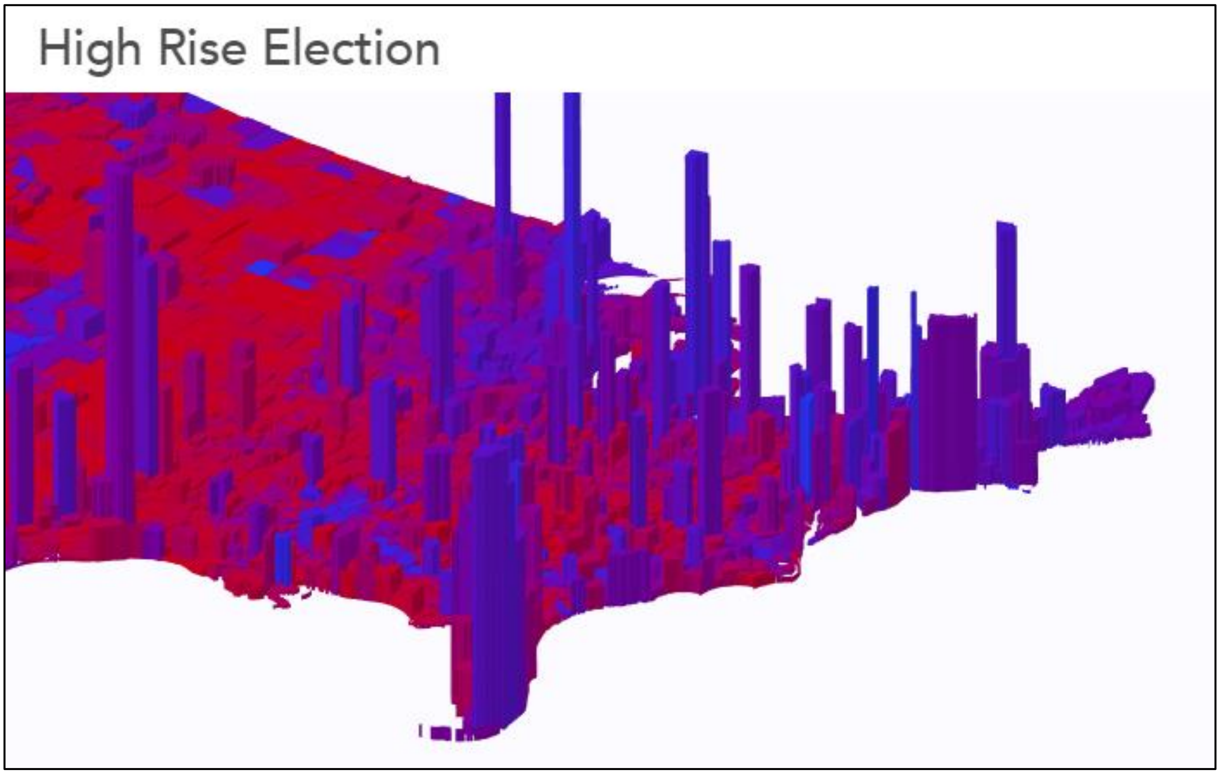


Figure 2 Election Results in 3D by Esri

2.3.2. Web Maps Displaying Propositions Results

While not as common as presidential results, various sites also publish web maps displaying proposition results. In California, the Secretary of State website actually publishes its own map displaying voting results by county. Unfortunately, this map is only available temporarily while the votes are still being counted, and the map is removed from the site once the election results are final. The map, as seen in Figure 3, shows California's counties as red or green depending on the percentage of votes for or against the proposition. There are only two colors, with no variance in hue as seen in the New York Times map. Selecting a county on the map displays a graphic of the map indicating the number of votes cast in the county. Users have to navigate to each proposition map through a homepage that lists the propositions on the ballot. This adds a level of difficulty for users looking to jump from one proposition to another. In fact, in their study, Huder et al. (2010) describe difficulty in navigating proposition results due to the lack of ability to search for propositions that fit the criteria they were exploring. Consolidating the propositions onto a single page and adding more hues to further indicate voting percentages could make this map stronger.

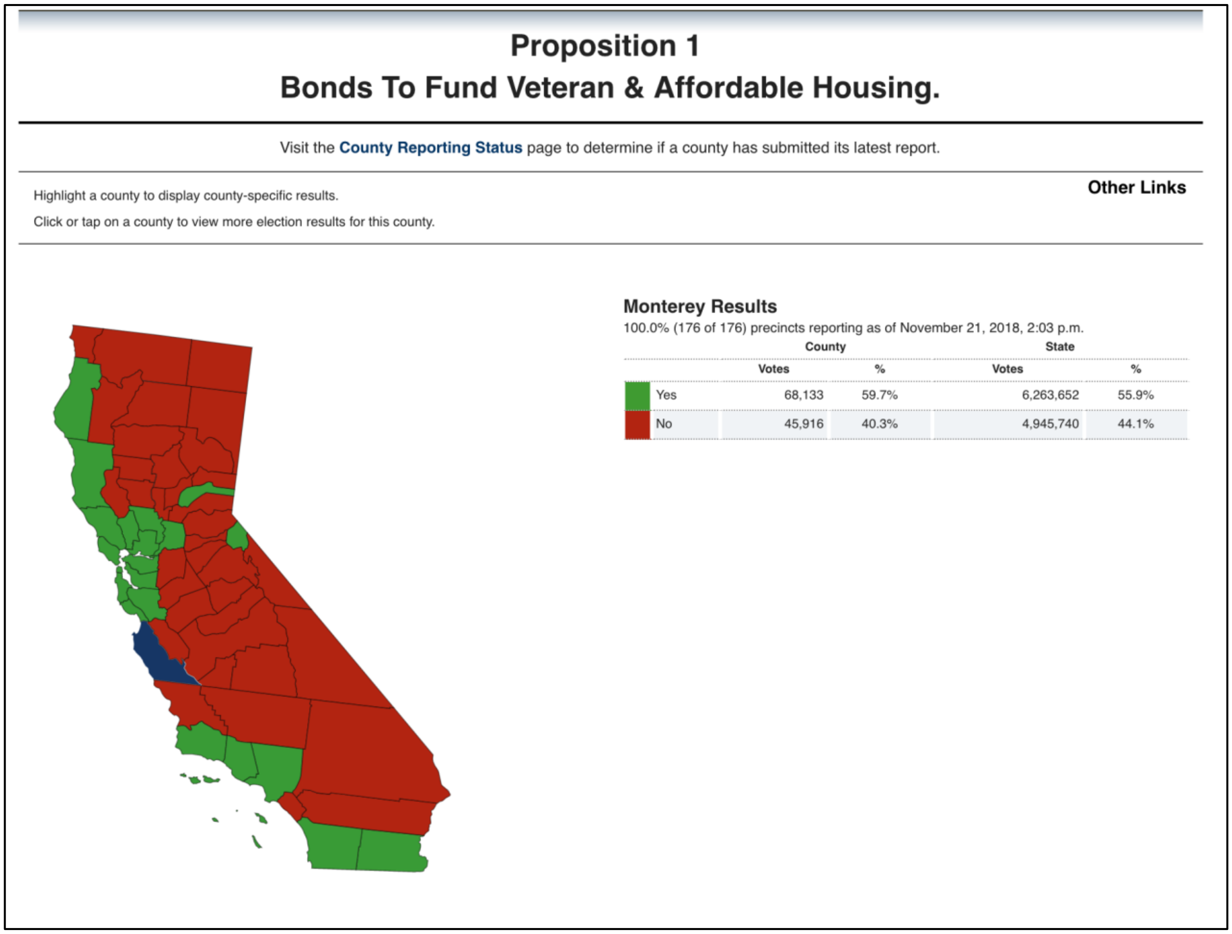


Figure 3 Secretary of State Proposition Results Map - Accessed January 14, 2019

Proposition maps may need more context for users to understand the data. Some sites utilize web pages with maps embedded into articles to provide more information or commentary about the proposition. An article from Coloradoan.com embeds Leaflet maps into their story about five propositions from their 2018 election (Powell 2018). This context is useful because it allows readers to understand a proposition, and then visualize how the state voted. These maps are also broken down by county with a simple color scheme of dark blue for supporting counties and light blue for opposing counties.

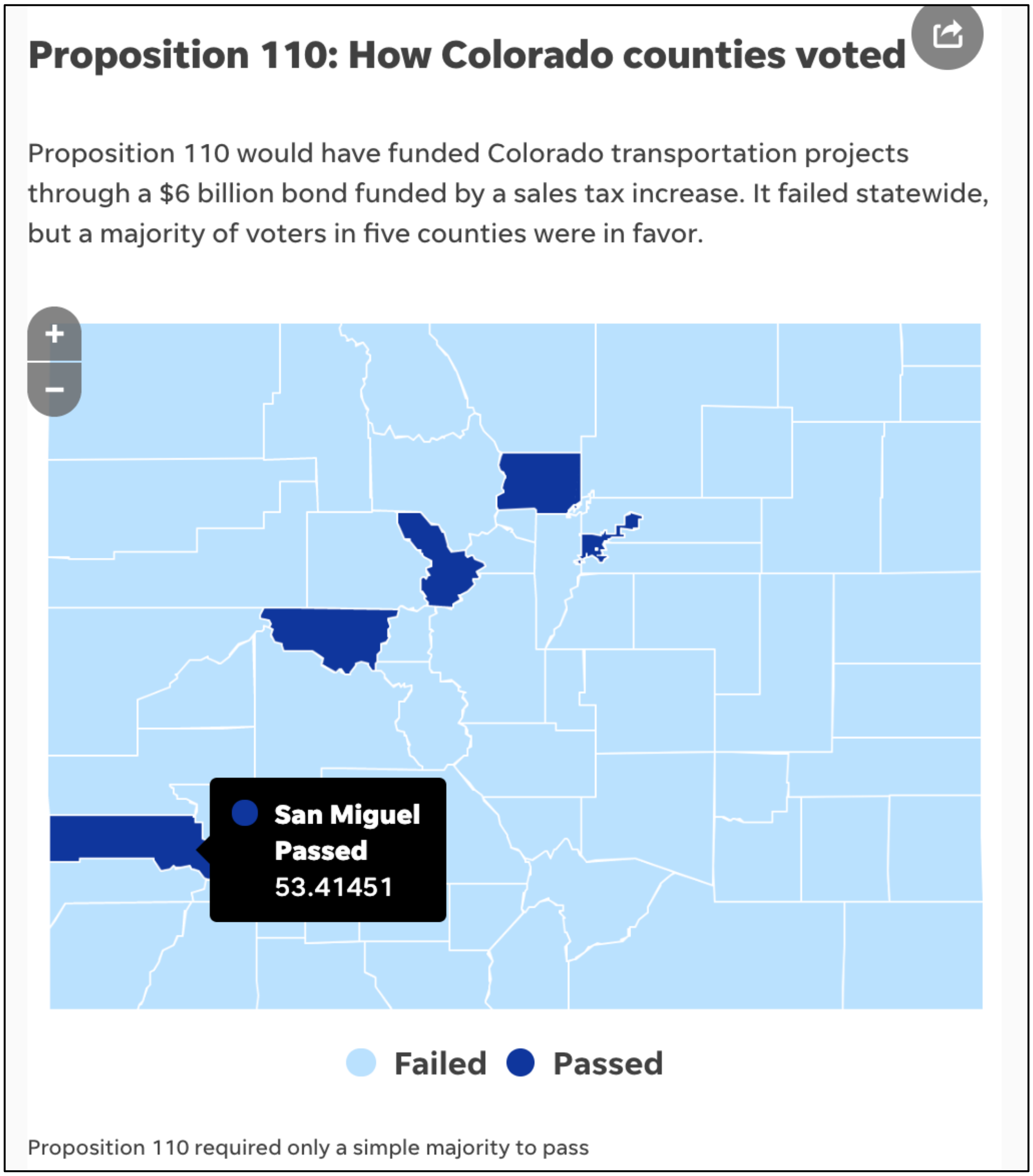


Figure 4 Map of Proposition Results from Coloradoan.com

Other ways to provide context include the use of multiple visualizations. The CNN website utilizes bar graphs, in addition to a map, to show the number of voters supporting or opposing a proposition (CNN 2018). Their map also uses different shades of colors to allow users to quickly identify counties where races were tighter. However, their site also requires users to return to a main page listing all the propositions, rather than providing an option to jump to a different proposition from the same page.

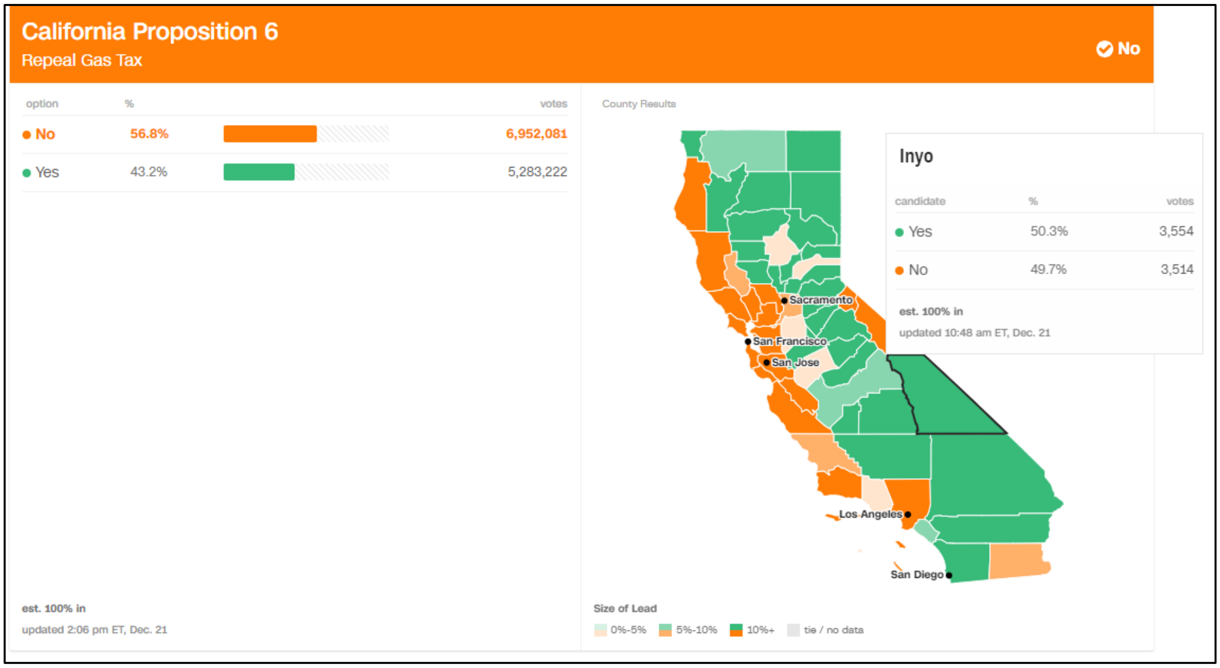


Figure 5 CNN Map Displaying Votes on Map and Using a Bar Chart

Chapter 3 Methodology

This chapter reviews the application’s methodology and requirements. The chapter opens with a review of the application's requirements and objectives. It is these requirements and objectives that shaped the decision-making involved in the application's development. The next three sections cover the steps necessary to develop the application. Section 3.2 covers the sources of data required and the necessary transformations to create usable data. Next, Section 3.3 discusses the software used to create the application. Finally, Section 3.4 describes the specific process of creating the application.

3.1. Requirements and Objectives

The objective of this application was to create a platform that allowed users to visualize California’s ballot proposition results quickly, fairly, and without bias. This section outlines the requirements necessary to achieve this objective. A goal of the application is to display ballot propositions results on a map by county and allow users to quickly change visualizations to different propositions without having to leave a single screen. By using a map, users have the ability to find potential voter bubbles, and by bringing different propositions to a single screen, users can identify if certain topics produce different results in specific counties. To accomplish these objectives, the following requirements below were set.

3.1.1. Functionality

The application must be capable of the three functionalities. First, the application must be able to visualize results by seeing which counties were in favor of the vote and how many total votes a proposition received per county. Next, the application must allow users to find a proposition by allowing users to search by name, subject, result or year on the ballot. Finally, the application must allow users to see results of a county over specified time periods.

3.1.2. Accessibility

Accessibility requirements for the application are straight-forward. First, users must be able to access the application through desktop or mobile device. Second, outside of an internet browser, no additional applications must be downloaded to the device. Third, users should be able to use the application with no training.

3.1.3. Software and Platform

This application is intended to be the first iteration of an application that sees more development over time. This consideration must be kept in mind when selecting the software used to create the application. There were five requirements for the software selection. First, the software must be capable of producing a web map. Next, the software must be available for free. Third, the application must be able to be shared and accessed for free. Fourth, the software must support robust data visualizations. Finally, the software must be one the developer has experience with, or can learn quickly.

3.2. Data

There were two sets of data required to create this application: data about the counties and data about the propositions. The required data associated with counties were geometries and county classifications. The required data associated with propositions included the proposition name, a description of the proposition, and the results of the proposition. Before developing the application, each of the two datasets had to be joined into one. Once joined, the application was able to visualize spatial data displaying where propositions were supported as well as text information describing what the proposition was proposing. The following sections describe the requirements for each dataset more thoroughly and then explain how the data was obtained and ultimately joined together. Table 2 provides a quick summary of each dataset.

Table 2 Summary of Data

Data Type	Data Source	Original Format
Proposition Results and General Information	CA Secretary of State	CSV
Extended Proposition Descriptions	CA Legal Analyst's Office	HTML
Generalized Proposition Subjects	Ballotpedia.com	HTML
County Polygons	US Census	Shapefile
County Classifications: Rural or Metro	The Office of Rural Health Policy	PDF
County Classifications: Region	CA Department of Social Services	PDF

3.2.1. Spatial County Data

The first required dataset was the spatial data for the county. This dataset had four requirements. First, each of the counties had to have a geometry in order to be properly represented on a map. Second, the counties needed to have a unique identifier field to allow this dataset to be joined with a results dataset. Third and fourth, to add additional context, each county needed to be identified as metropolitan or rural and assigned a region in California. The first two requirements were obtained by downloading a California counties shapefile from the United States Census website (Census 2019). The shapefile contained the geometries for each county, and each county was assigned a unique identifier. A document classifying each of the counties was obtained from The Office of Rural Health Policy to classify each county as rural or metropolitan (Health Resources and Service Administration 2018). Each of the counties was assigned a value of rural or metropolitan according to the document. With the geometries obtained, uniquely identified, and classified, the task of assembling the first dataset required was completed. Similarly, to classify the counties by region, a document from the California Department of Social Services was used (Tran, Jones, and Smilanick 2001). The document, called "Recommended Grouping of the Counties for Regional Studies," provides different county groupings based on purpose. For this application, the classification listed as the California Data Analysis and Publications branch was used.

3.2.2. Proposition Data

The second required dataset described the propositions. For this dataset, multiple sources were consulted to ensure the application provided relevant information about each proposition. For this dataset, the result of the proposition, the result of the proposition broken down by county, a proposition title, information about what the proposition was proposing, and a generalized proposition subject was required. For the proposition results, official data was obtained from the California Secretary of State website. The Secretary of State provided users with a title of the proposition, the result of the proposition, and a breakdown of how each of the counties voted. The data was available and downloaded in a CSV format (California Secretary of State 2019). For a more thorough description of what the proposition was proposing, data was obtained from the California Legislative Analyst's Office or LAO (Legal Analyst's Office 2019). Finally, to create a generalized search function in the application, a generalization of the proposition subject was acquired from the website Ballotpedia.com (2019).

3.2.3. Transforming and Merging the Data

With the data sources identified, the data had to be transformed into a usable format before application development. After identifying the sources of data, an entity-relationship (ER) diagram was drawn to determine an appropriate data model. Once obtained, the data was transformed to match the data model drafted in the ER diagram seen in Figure 6. For the proposition descriptions and subjects, text was extracted directly from the LAO's website and the Ballotpedia website in July 2019 and imported into a spreadsheet. County classifications for rural and metropolitan and region were extracted in a similar fashion, with text from a report exported into a spreadsheet. Proposition results were downloaded as CSVs from the Secretary of State website and input into a spreadsheet and transformed to match the data model (Secretary of State 2019). Each dataset was given a column for a primary key to allow joins between the datasets. With primary keys assigned, the data was ready to be used to develop the application.

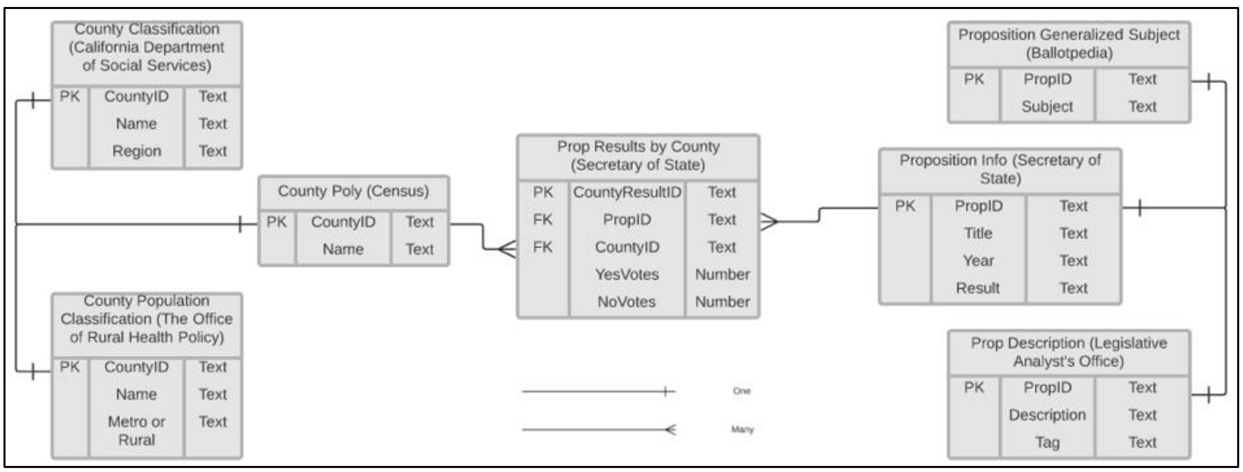


Figure 6 ER Diagram

3.3. Software

This section describes the software used to create the application. The main software used for this application was Tableau Public. In addition, Microsoft Excel was used for data clean up and processing before using Tableau Public, and GitHub Pages was used to host the application after it was fully developed. Section 3.3.1 and 3.3.2 describes the role of Microsoft Excel and

GitHub Pages, respectively, and section 3.3.3 provides a more in-depth look at Tableau Public and its capabilities.

3.3.1. Microsoft Excel

Microsoft Excel is a spreadsheet software made by Microsoft. Despite not being available to use without a subscription or license, it was used for data cleanup due to it being a commonly installed application on most machines. All data cleanup and formatting could have also been performed using a free spreadsheet application like Google Sheets or Open Office Sheets, and any updates to the data can be performed on these applications. For the purpose of this ballot application, the machines used to develop it all had Excel installed, and it was understood that only spreadsheet functions available to most spreadsheet applications were to be used while assembling the data.

3.3.2. Github Pages

Github Pages is a feature of the code repository website, Github, that allows users to publish websites for free by using code that the user uploads to Github. By using a third-party to host the application, a developer can create a neutral landing page for the dashboard free of any excess branding or text. This is needed because when a user publishes a dashboard to Tableau Public, the link to the dashboard takes the user to a Tableau branded landing pages. Other free website hosting platforms exist, including Weebly and Wix which provide a user interface for interactively designing webpages, but Github Pages was selected due to its simplicity and previous usage by the developer.

Another advantage of hosting the application outside of Tableau Public is the ability to add a developer's own tracking analytics. While not a requirement of the application, Github

Pages allowed the developer to keep track of traffic to the application through a Google Analytics code added to the HTML of the page.

3.3.3. Tableau Public

Tableau Public is the free version of Tableau, a data visualization software company. Tableau Public requires users to save and publish their work publicly on the Tableau Public website in order to save or share the work elsewhere. In June 2019, Tableau was purchased by Salesforce, but as of August 2019, the Terms of Service page indicates there have been no updates to the Terms of Service since 2015 (Tableau 2015; Lunden 2019). Work published to Tableau Public is hosted on the Tableau server and the platform provides users with links to embed dashboard applications into other websites. According to the Tableau Data Policy, published works are to remain on the site unless a formal procedure is filed to remove content from the site (Tableau 2019)

Tableau Public was selected due to its free availability, ease-of-use, familiarity to the developer, and robust options for data visualizations. With the developer's experience with different platforms an initial limiting factor, the choices of development for this application were limited to Esri products including ArcGIS Online, Web Appbuilder, and Operations Dashboard, or Tableau and Tableau Public. Due to the subscription needed to access Web Appbuilder and Operations Dashboard, the products from Esri were ultimately eliminated from consideration. The rest of this section provides an overview of a general workflow in Tableau Public starting with bringing in data, to visualizing the data, to organizing the data onto a dashboard and then sharing the completed product.

3.3.3.1. Bringing in Data

Tableau Public allows users to visualize data using different sources including Excel spreadsheets, text files, shapefiles, and JSON. These datasets are brought into a comprehensive project environment called a “workbook.” In order to use the data, each data source must have a common key linking it to the other sources on the workbook. For example, a table with county descriptions can have county name as a common key to a shapefile of the same counties. With common keys identified, the relationships between the datasets must be defined as the following:

1. Inner – only keep records that have a matching common key in both datasets
2. Left or Right – keep all records in one dataset, and only keep records in the second dataset that have a matching common key
3. Full Outer – keep all records in both datasets, regardless if there are matching common keys or not

The image below is a screenshot of Tableau’s interface to join data sources.

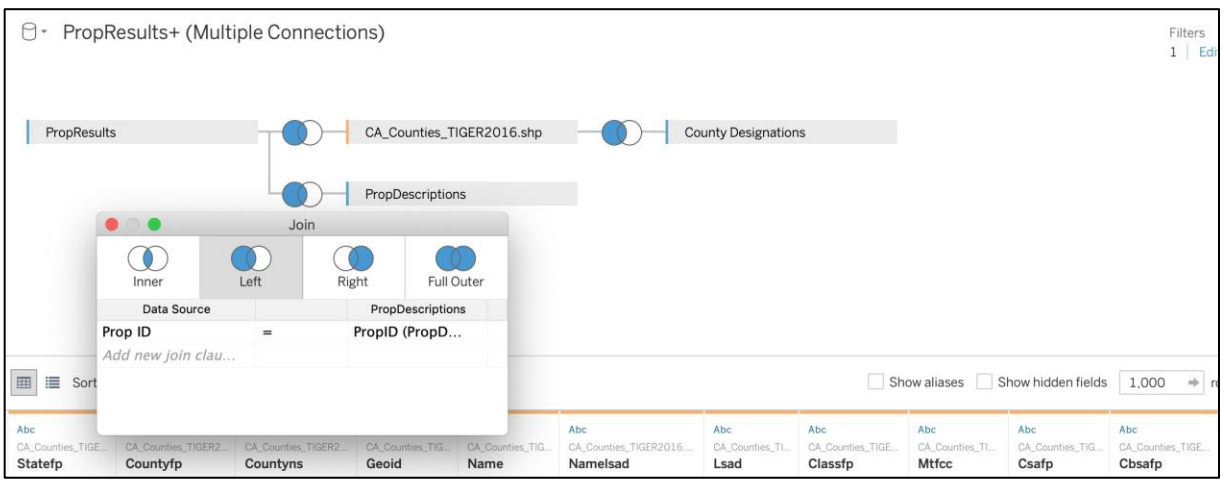


Figure 7 A Screenshot of Data Source Joins in Tableau

3.3.3.2. Visualizing Data using Sheets

After the data sources have been joined, the developer must choose the fields he or she wants to visualize. Fields from all the joined datasets can be brought into a single visualization. Fields are automatically classified as “measures” for numeric values, and “dimensions” for text values. Fields can be used as the values in a data visualization, or their values can be used as filters throughout the workbook. Fields can be duplicated and configured as measures or dimensions, and new fields can be calculated based on existing fields. For example, a “total votes” field can be created by calculating a field that adds the values in a “Yes” field to the values in a “No” field. This allows users to manipulate datasets in Tableau without having to modify the original spreadsheets or attributes of a shapefile.

Once the fields are defined, the user can drag the fields to an interface known as a “sheet.” A workbook can have multiple sheets, but each sheet can only be one visualization. These visualizations include maps, bar charts, box and whisker plots, cartograms, treemaps, and text tables. Figure 8 is a screenshot of the different types of visualizations developers can make using Tableau.

Figure 8 Data Visualization Choices in Tableau

3.3.3.3. Organizing Visualizations onto Dashboards

After creating visualizations, the user can bring the visualizations to a single interface that Tableau calls a "dashboard." A dashboard provides users a space to arrange their data visualizations. The dashboard also allows users to set options to configure interactivity between the different data visualizations. A workbook can have multiple dashboards that can be configured to interact with each other.

A single dashboard can have multiple layouts, each optimized for different screen sizes. There are three options for defining layout size: fixed, range, and automatic. Using a fixed size, the user sets a defined pixel size for the dashboard to consistently appear. A range size allows users to set a minimum and maximum pixel size for their dashboard. Automatic creates dashboard layouts that occupy the maximum space of a screen.

3.3.3.4. Sharing a Completed Dashboard

When a user is satisfied with the dashboard or dashboards, the user can then share it to the Tableau Public website for everyone to access. Sharing to the Tableau Public website is the only way users can save their data. This means that even unfinished dashboards are published to the website. Tableau Public's website stores the user's workbook on the Tableau Servers and creates a page, as seen in Figure 9, for others to access the workbook and see accompanying metadata. Should the user wish to display the dashboard on a different site, the Tableau Public website provides a line of code that a user can add to his or her website to embed the dashboard.

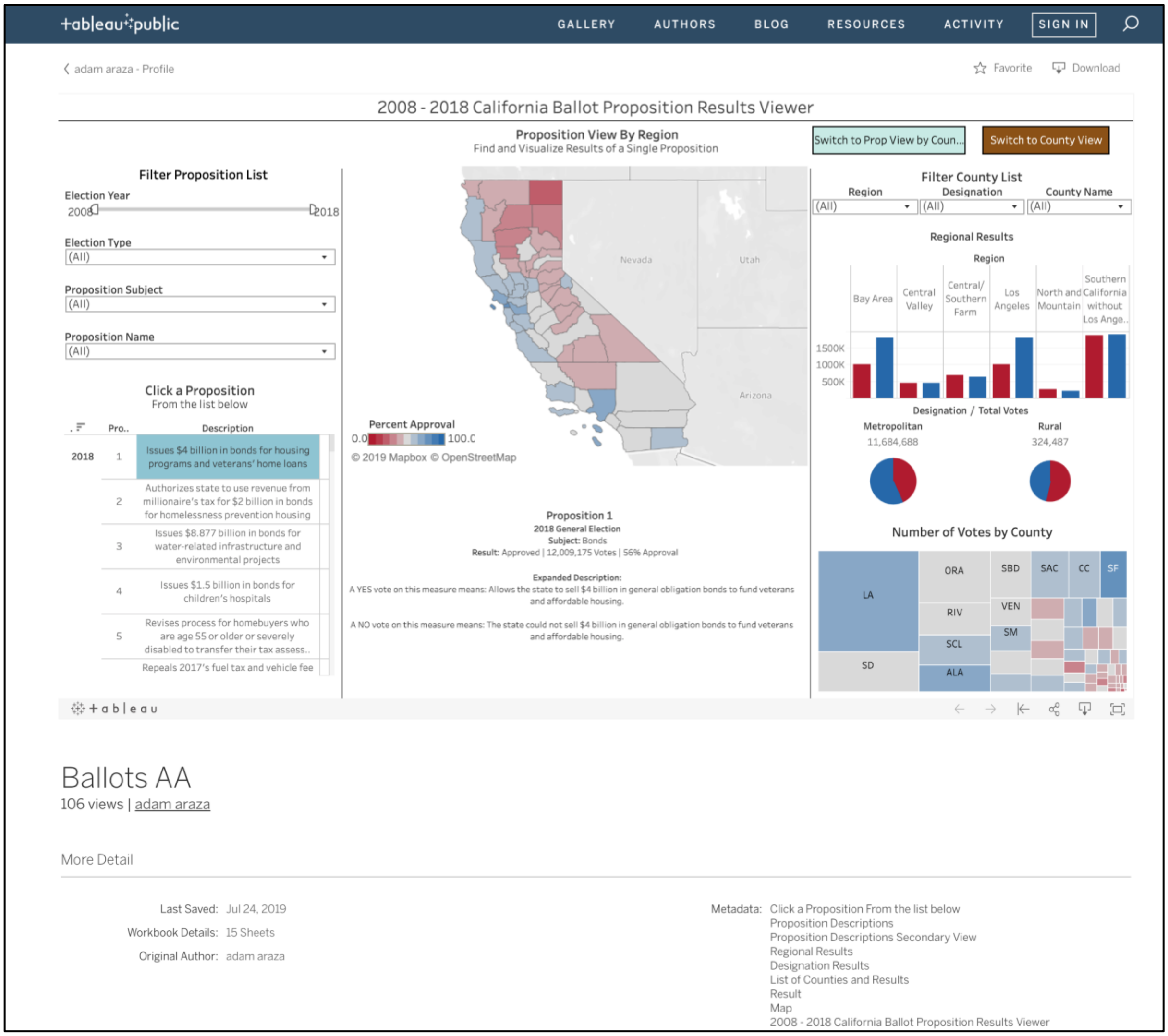


Figure 9 Shared onto Tableau Public Website

3.4. Application Development

This section discusses the specific steps taken to complete the application. The process is similar to the standard Tableau process described in the previous section. Section 3.4.1 reviews the data sources and details what transformations were necessary before bringing the data into Tableau. Section 3.4.2 reviews the different visualizations selected and the purpose of each visualization. Next, in section 3.4.3, is an overview of the dashboard layouts configured. Section

3.4.4 closes the application development section by explaining how the application was published as a website on Github Pages.

3.4.1. Data Preparation

Section 3.2.3 provided an ER diagram detailing the relationships between the different data sources used in this application. This section specifically describes the steps needed to acquire the data and transform it into data structures usable for this application. There are four main categories of data:

1. Spatial data representing the counties
2. Tabular data describing the counties
3. Tabular data describing the results of the propositions by county
4. Tabular data describing the propositions

3.4.1.1. Spatial Data

The spatial data was acquired from the State of California Open Data Portal website. The website provides users with access to the United States Census TIGER shapefiles, with national data already filtered to California. The acquired dataset reflected California county boundaries as of January 1, 2016.

3.4.1.2. Tabular Data Describing the Counties

The spatial data provided information about the county's geometric boundary and the county's name. To provide users with a starting point for filtering through the list of the counties, more data was collected to provide pre-defined filters. The data collected described the counties as either metropolitan or rural and assigned each county to a named region in the state. Each of these descriptions was taken from PDF reports, so to bring them into a usable data structure, a spreadsheet was created with a name for each county populating a row in the spreadsheet. Rural and metropolitan classifications were manually added to the spreadsheet by simultaneously comparing the spreadsheet to the report from the Office of Rural Health Policy (2018). The state regions were populated in a similar manner using a report from the California Department of Social Services. Finally, for the purpose of labeling the counties, a report from Caltrans was consulted that indicated the accepted county two or three letter abbreviations. These values were, again, added in manually by comparing the two lists.

3.4.1.3. Tabular Data Describing the Results of the Propositions by County

Due to the need for each county to have its voting results indicated, it was determined that results would need to be indicated on two separate tables. One table with a general description of whether the proposition passed or failed, and another table with specific figures for how each county voted. Using this data structure, the results by county were easily joined to the county datasets, and the results by proposition were easily be joined to the proposition description dataset. Figure 10 shows the differences between the two tables.

Title	Subject	Description	Result
Proposition 1A	State budget	Local government revenues	Approved
Proposition 59	Admin of gov't	Public records, open meetings	Approved
Proposition 60	Elections	Election rights of political parties	Approved
Proposition 60A	Admin of gov't	Surplus property	Approved
Proposition 61	Bonds	$750 million in bonds for children's hospital projects	Approved
Proposition 62	Elections	Elections, primaries	Defeated
Proposition 63	Tax increase	Mental health services expansion	Approved
Proposition 64	Business regulation	Limits on private enforcement of unfair business competition laws	Approved
Proposition 65	State budget	Local government funds, mandates	Defeated
Proposition 66	Law enforcement	"Three strikes" law, sex crimes	Defeated
Proposition 67	Taxes	Fund emergency medical services with tax increase	Defeated

	CountyName	Yes	No
Prop1	Alameda	373,847	166,156
Prop1	Alpine	339	259
Prop1	Amador	6,478	10,582
Prop1	Butte	40,578	45,578
Prop1	Calaveras	7,859	13,003
Prop1	Colusa	2,109	3,446
Prop1	Contra Costa	237,539	165,674
Prop1	Del Norte	3,385	4,697

Figure 10 Results Per Proposition (left) and Results by County (right)

To create the table that indicated results by county, a significant amount of spreadsheet data pivoting was required. This is because the California Secretary of State published the results with each proposition populating a row and each county populating a column. For the purpose of being able to assign multiple proposition results to individual counties, a proposition had to be represented on multiple rows, to create unique pairs of proposition and county name. Once the data was formatted for each of the propositions, the data was moved to a single spreadsheet and was ready to be used by Tableau. The following image shows a screenshot of a table of data as it was downloaded from the California Secretary of State website next to the result of the data to be used in the application.

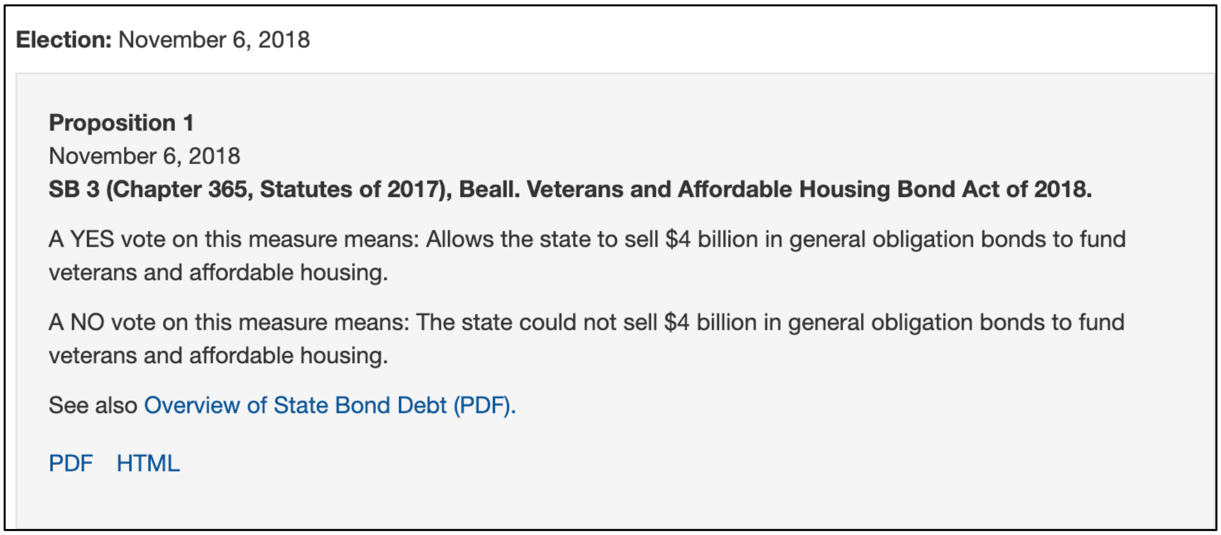
	Proposition 50	
	Suspension of Legislators.	
	Yes	No
Alameda	272,073	69,938
Percent	79.6%	20.4%
Alpine	361	90
Percent	80.0%	20.0%
Amador	9,594	2,328
Percent	80.5%	19.5%
Butte	43,973	16,417
Percent	72.8%	27.2%
Calaveras	11,558	2,891
Percent	80.0%	20.0%

PPE2016Prop50	Prop50	**Alameda**	272,073	69,938
PPE2016Prop50	Prop50	**Alpine**	361	90
PPE2016Prop50	Prop50	**Amador**	9,594	2,328
PPE2016Prop50	Prop50	**Butte**	43,973	16,417
PPE2016Prop50	Prop50	**Calaveras**	11,558	2,891
PPE2016Prop50	Prop50	**Colusa**	3,207	785
PPE2016Prop50	Prop50	**Contra Costa**	192,537	54,520
PPE2016Prop50	Prop50	**Del Norte**	4,359	1,293
PPE2016Prop50	Prop50	**El Dorado**	43,414	14,370
PPE2016Prop50	Prop50	**Fresno**	117,304	38,578
PPE2016Prop50	Prop50	**Glenn**	4,523	1,162
PPE2016Prop50	Prop50	**Humboldt**	28,806	7,365
PPE2016Prop50	Prop50	**Imperial**	18,441	3,798
PPE2016Prop50	Prop50	**Inyo**	3,943	1,105

Figure 11 Raw Data (left) and Formatted Data (right)

3.4.1.4. Tabular Data Describing the Propositions

To provide context for the propositions, data was acquired from various other sources to ensure users had more information about the prop than the proposition number and a short, general description from the Secretary of State. The Secretary of State's website provided a link to more resources for users to learn about the propositions. One of those links was the for the California Legislative Analyst's Office which provided more context about what casting a "Yes" or "No" vote meant for each proposition. Similar to the county descriptions, the descriptions from the LAO's website were manually copied and pasted into a spreadsheet with a unique list of the individual propositions. Figure 12 is a screenshot of a proposition analysis from the LAO.



Election: November 6, 2018

Proposition 1
November 6, 2018
SB 3 (Chapter 365, Statutes of 2017), Beall. Veterans and Affordable Housing Bond Act of 2018.

A YES vote on this measure means: Allows the state to sell $4 billion in general obligation bonds to fund veterans and affordable housing.

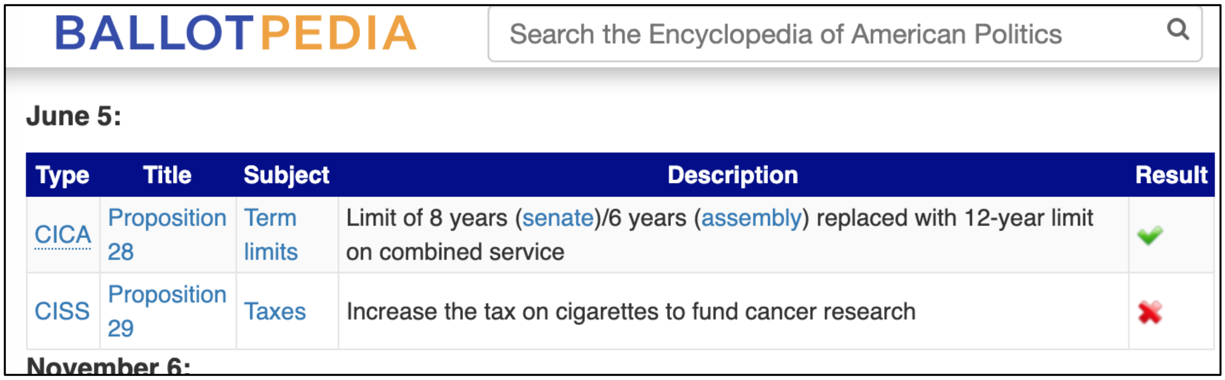
A NO vote on this measure means: The state could not sell $4 billion in general obligation bonds to fund veterans and affordable housing.

See also Overview of State Bond Debt (PDF).

PDF HTML

Figure 12 A Screenshot of a Short Analysis of Proposition 1 from the LAO

To simplify searching the propositions, a general subject field was added to the spreadsheet. This field was manually populated using data from the website Ballotpedia.com. As seen in Figure 13, Ballotpedia is a wiki site that has articles for elections at the state, local, and federal level. The website provides generalized subjects for each proposition that allow the propositions to be grouped into accessible classifications. The data from the site was used because of its understandable style of writing and broad generalizations for the propositions.



BALLOTPEDIA

Search the Encyclopedia of American Politics

June 5:

Type	Title	Subject	Description	Result
CICA	Proposition 28	Term limits	Limit of 8 years (senate)/6 years (assembly) replaced with 12-year limit on combined service	✔
CISS	Proposition 29	Taxes	Increase the tax on cigarettes to fund cancer research	✖

November 6:

Figure 13: A Screenshot of Ballotpedia.com

3.4.1.5. Loading data into Tableau Public

The formatted data was brought into Tableau Public's data loader screen. In Tableau Public, the county descriptions were joined to the county geometries in a one-to-one join using county name as a common key. Next, the proposition results by county table were left joined to the county geometries using county name as a common key. Finally, the proposition results by county were left joined to the proposition description table using a field called proposition ID.

3.4.2. Data Visualizations and Filters

Once the data was formatted, visualizations could be created in Tableau. Visualizations selected were chosen to achieve the project's goal of quickly visualizing proposition results. This section reviews each of the visualizations made for the application. Included in the discussion of the visualization is also a review of different data filters established.

3.4.2.1. Proposition List

The first visualization created was a text table indicating the proposition year, number, and a short description. This visualization, as seen in Figure 14 was intended to be a list where users can select a proposition on the list to visualize more information about the proposition.

Included as filters are fields for the subject of the proposition, the year the proposition was on the ballot, the type of election, and the number of the proposition.

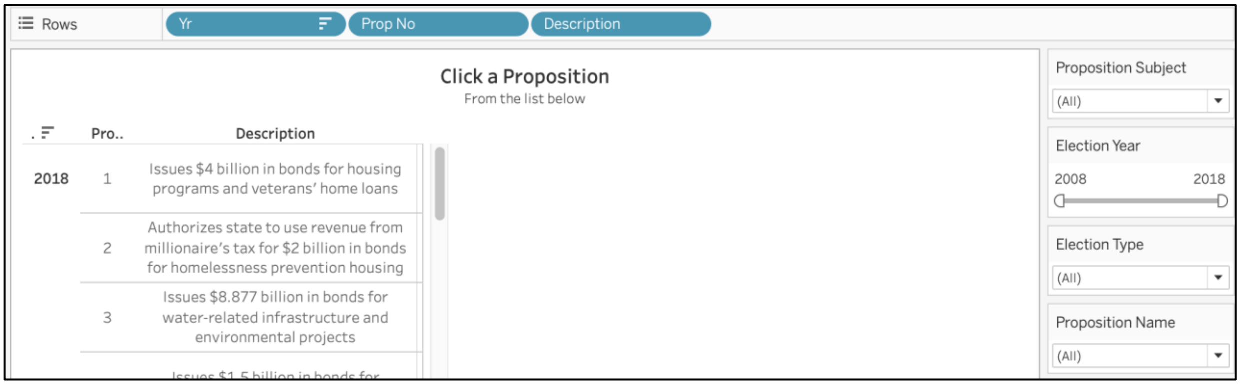


Figure 14 Screen Capture of Sheet with Proposition List

3.4.2.2. Proposition Description

The next visualization created was a card for a further proposition description. The purpose of this visualization is to provide users with more information about a proposition once one is selected from the full proposition list. This information consists of the proposition number, the year, the type of election, the generalized subject from Ballotpedia, the result, the number of votes cast, the percentage of votes that were "yes" votes and the extended description from the LAO. The visualization was created by adding the necessary fields as a label, only to display when a single proposition is selected. Figure 15 shows the configuration screen used to create this visualization.

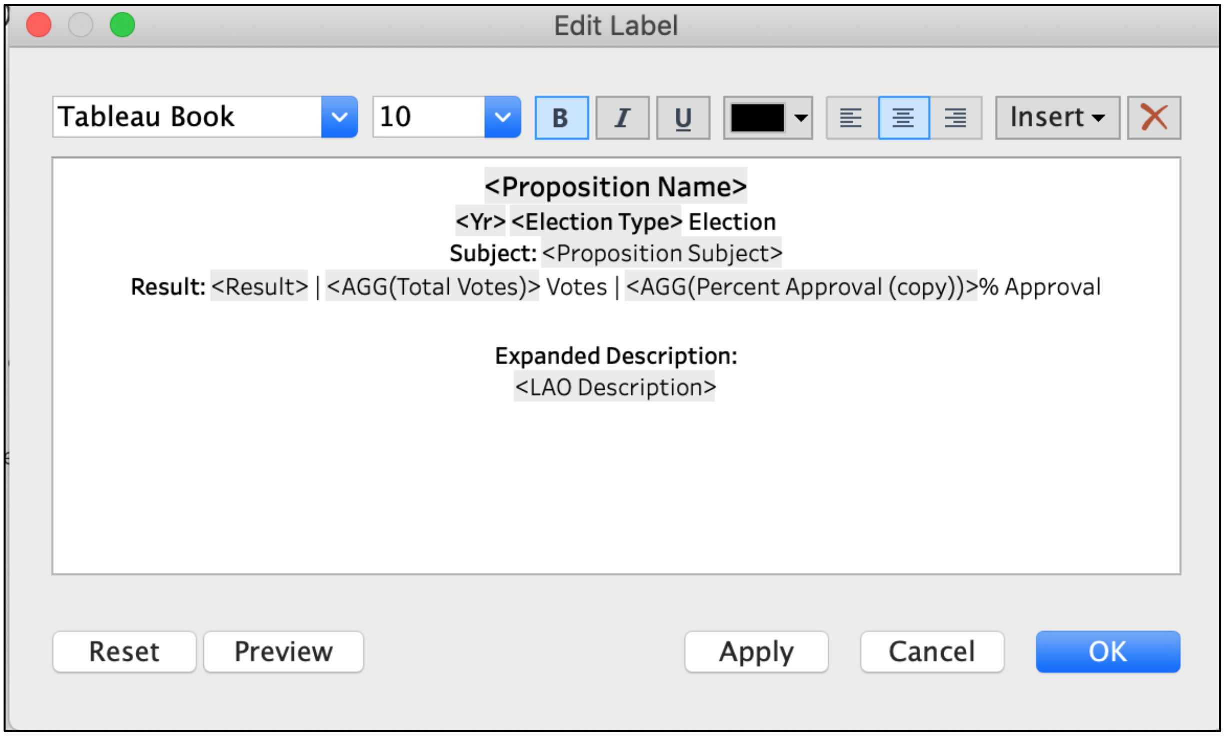


Figure 15 Label Screen to Create Proposition Description

3.4.2.3. Map

The next visualization created was the map. The purpose of the map is to provide users with a visual representation of where in the state a proposition was supported or opposed. The map is symbolized on an attribute indicating the “yes” percentage of total votes cast for that proposition. It uses a red to blue diverging scale with counties with lower “yes” percentages indicated as red and counties with high “yes” percentages indicated as blue. Counties with percentages near 50% “yes” are symbolized as grey. The basemap used is the default Tableau grey basemap. Different basemaps were considered, including customized basemaps created on Mapbox, but the simplicity of the default Tableau map provided a clean backdrop to present the data. The selected color scheme was used to best eliminate opportunities for generating maps with bias. In addition to symbolization, the map also features a tooltip that shows the name of the county, the total number of votes cast and the percentage of “yes” votes. Figure 16 shows the map visualization with a county selected for the purpose of showing the accompanying popup.

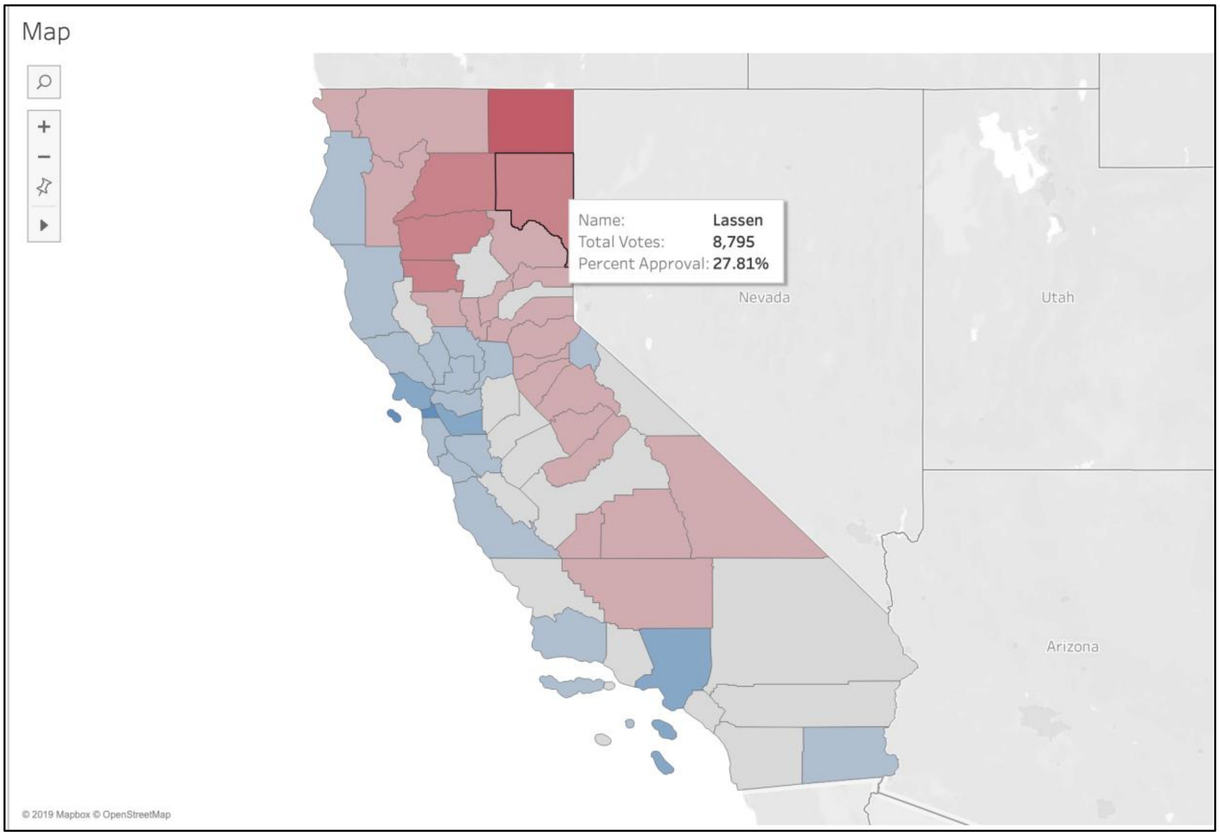


Figure 16 Screenshot of Map with Lassen County Selected

3.4.2.4. Treemap

The map in the previous section can quickly provide users with a spatial visualization of where voters support a proposition, but this visualization does not indicate how many voters are in a county. The figure above highlights Lassen County and upon first glance it may appear that a large number of supporters are in the county due to the geographic size, despite there being less than 10,000 votes cast for the particular proposition. To provide users with a visualization that shows the relationship between total number of votes cast per county and voter preference, a treemap is used.

A treemap, as seen in Figure 17, is a visualization that represents each measure of a dataset as a rectangle. Using a field as a metric, the rectangles are then sized to reflect the proportions of the total metric. For example, a county with twice as many votes as another appears as a rectangle that is twice as a large. Rectangles with the largest values are at the top-left and rectangles with the smallest values are placed at the bottom-right. For this application, the rectangles were sized using the total votes attribute, and they share the same color symbology as the map. This visualization allows users to quickly see which counties had the most voters and how those voters cast their votes for a specific proposition. Cartograms were also considered as a representation, but due to the north-south orientation of the California, a treemap was selected to better utilize the rectangular space of a viewing screen.

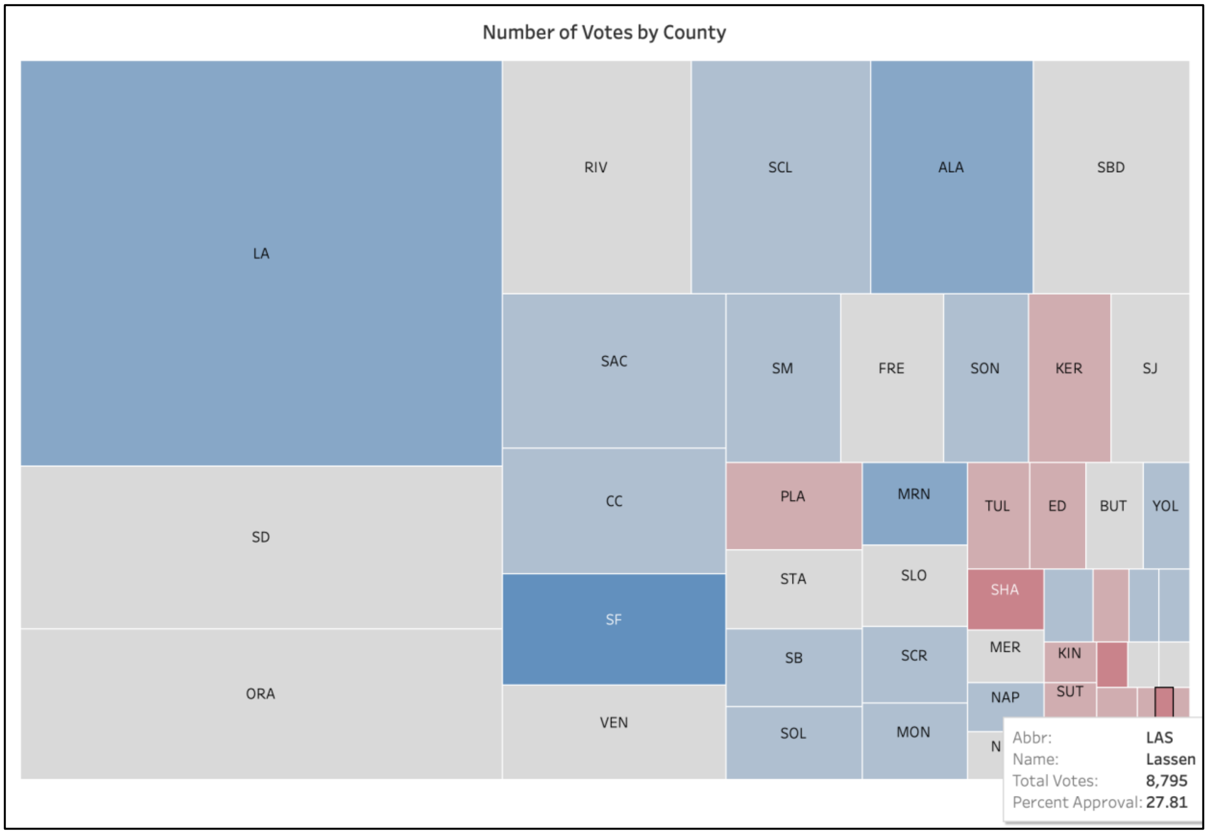


Figure 17 Treemap with Lassen County Selected

3.4.2.5. List of Counties and Results

Specific county results were displayed on a table with stacked bars representing total "yes" votes and total "no" votes respectively. These bars share the same colors as the rest of the elements in the workbook with "yes" being represented by blue bars and "no" votes represented by red bars. The purpose of this visualization is to provide a granular list of counties that users can scroll through to see the relationship of "yes" votes to "no" votes as well as a comparison of how many votes were cast in each county. This visualization, as seen in Figure 18, also serves as a filter to sift the results datasets to only show particular counties. A filter was added to allow users to only see counties designated as rural or metropolitan, and a filter to show counties based on their region.

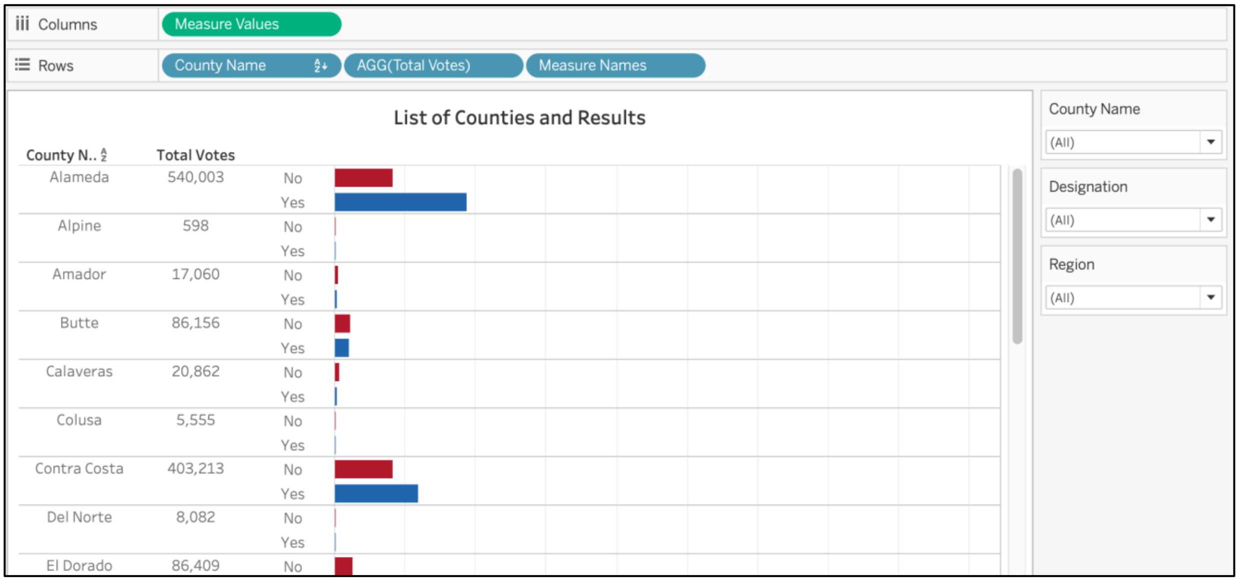


Figure 18 Horizontal Bar Chart with Results for Each County

3.4.2.6. Bar Chart of Results by Region

To provide a quick overview of how a proposition performed across the different regions, a bar chart was created to show "yes" votes and "no" votes categorized by the region. This visualization allows users to see a proposition's performance without having to scroll through individual counties. It uses the same color symbology as the bar chart for the list of counties and results visualization. Figure 19 shows the bar chart with "yes" and "no" votes categorized by region.

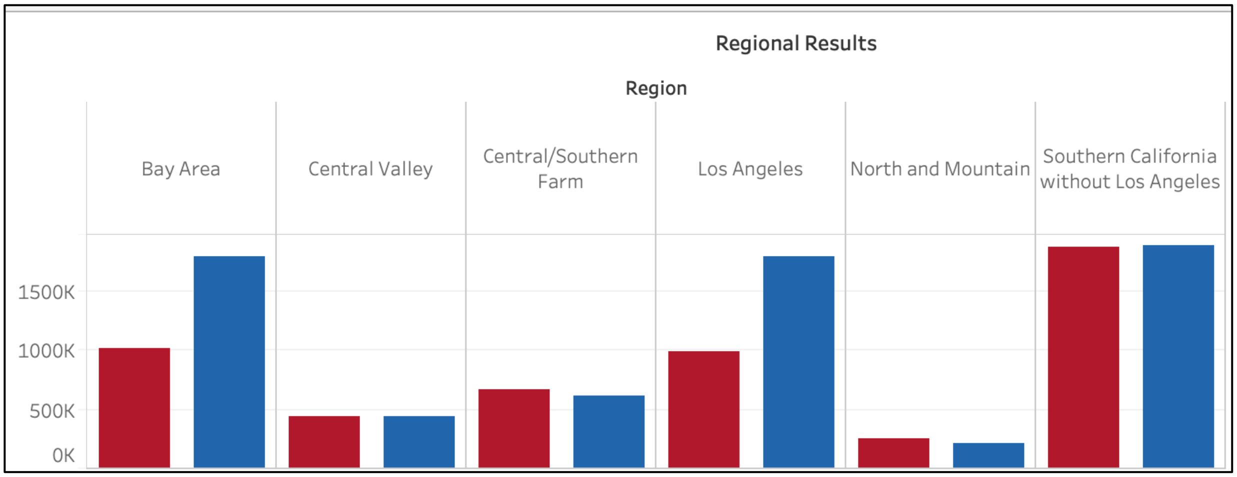


Figure 19 Vertical Bar Chart with Results per Region

3.4.2.7. Pie Charts by Rural or Metropolitan

Another visualization for seeing quick results in a generalized classification are the pie charts indicating how a proposition performed in rural and metropolitan areas. The color symbology is consistent with the rest of the workbook. Sizing for each pie was initially determined using the number of total votes, but due to the metropolitan counties consistently having significantly more votes cast, a field was calculated using the log of base 2 applied to the total votes. This calculation allowed users to still see the rural pie, but a label above the two pies still provides context with how many votes were actually cast. Figure 20 shows an example of pie charts with varying sizes comparing two variables.

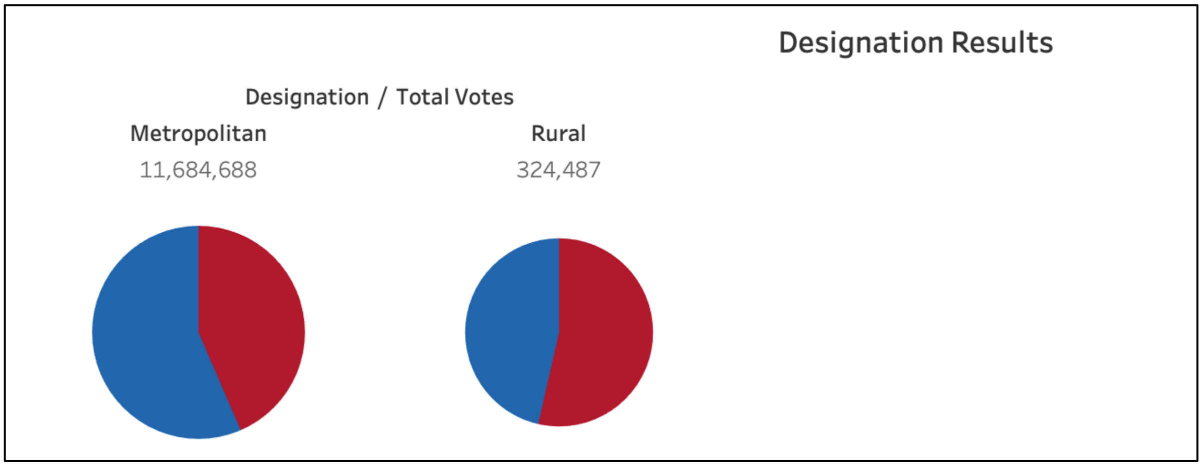


Figure 20 Pie Charts Comparing Rural Votes Cast to Metropolitan Votes Cast

3.4.2.8. List of Proposition Results by Selected County

The final visualization for this dashboard is the list of proposition results. This list includes stacked bars indicating how many "yes" votes a proposition received in relation to how many "no" votes received. The purpose of this list is to allow users to view a county or region's history of voting. With the full list of propositions, users can quickly see if there were any voting anomalies, including a proposition that was supported in a large area that ultimately failed, or a proposition that garnered significantly more votes than others in the region. The list includes the year and number of the proposition, a short description, the result of the proposition, the total votes cast, and the stacked bars mentioned previously. In addition, this visualization, as seen in Figure 21, includes filters that allow users to filter the list by proposition name, proposition subject, election type, and election year.

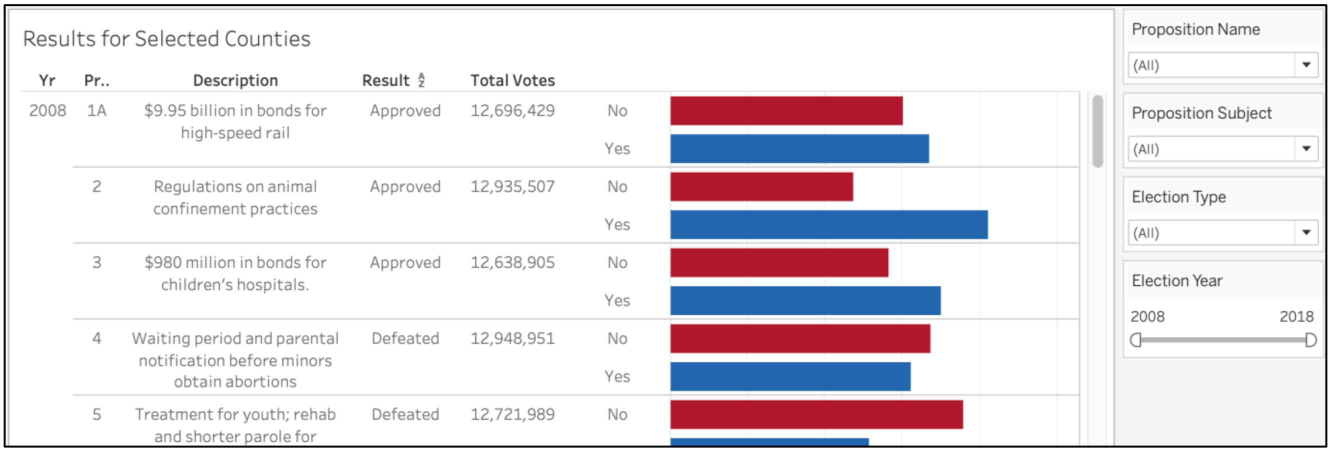


Figure 21 Full Proposition Results List

3.4.3. Dashboard Configurations

The purpose of the application was to allow users to quickly visualize ballot results by county for propositions. The application was to support two workflows. First, a proposition-based workflow that allowed users to find a proposition and see how different regions supported or opposed that single proposition. Second, a county-based workflow that allowed users to select a county or region and view how that area has supported or opposed a list of propositions over time. To support these two workflows, it was decided that two dashboards needed to be created and users needed to be able to toggle back and forth between the two from the application. This section details the composition of the two dashboards used for this application.

3.4.3.1. Layout Insights

With the visualizations completed, the next task was to place the visualizations on a layout for a single display. The first step was deciding what size dashboard to use for desktop and mobile. Tableau allows users to create a single dashboard, and then add multiple layouts to display information according to screen size. For desktops, an automatic size was selected for publishing, but a fixed size layout defined as “generic desktop” on Tableau Public or 1366px x

768px was used in development. By developing at 1366px x 768px, the developer is able to define an optimal developer size to view the dashboard. Publishing the dashboard at an automatic setting allows Tableau to appropriately scale the dashboard for screens that are larger than mobile devices but are not large enough to display the optimal viewing size. For mobile devices, a layout called "generic phone" was used with dimensions of 375px x 776px.

For desktop layouts, a three-column layout was used. This layout was selected due to the prevalence on multiple news websites and shopping websites which typically had filtering content on the left, main content in the middle, and supplemental content on the right. The main content for the dashboard is the map, so it was determined that the map and the proposition information were to be placed in the center. Menus to filter the propositions or the counties were added to the left of the dashboard depending on the workflow the dashboard was intending to support.

For mobile layouts, lengthy text boxes were omitted. This includes the lengthier proposition description from the LAO. Filters were kept closer to the top of the screen and data visualizations were below. The rest of this section focuses on the content of each dashboard with regards to its desktop layout.

3.4.3.2. Dashboard for Proposition-Based Workflows

The proposition-based workflow was designed for users looking for specific propositions. Using a three-column layout, the middle section consisted of the map and extended proposition description, the left section contained the proposition list and the proposition filters for the list, and the right hand was composed of infographics for the proposition results. These infographics contained the following visualizations: a list of counties names aggregated into their respective regions and associated voting results, a treemap sized by the number of voters and colored by voter preference, and a pie chart to symbolize support and opposition in rural counties and support and opposition in metropolitan counties. The purpose of this view is to provide users with a quick overview of how the proposition performed without having to scroll through any of the visualizations. Should the user want to see the county names and results within each region, the user can click the region column to expand the full list of county names. On the top of this column are filters for county region, county classification and county name. These filters allow users to reduce the amount of data being shown on the screen in case the user has specific areas of interest. Users are able to select a proposition from the left column and see the results of that proposition on the map, a description of the proposition and the results of the proposition either broken down by region or by specific count, as seen in Figure 22. Filtering the counties alters the map and results visualizations, but the proposition description remains unchanged.

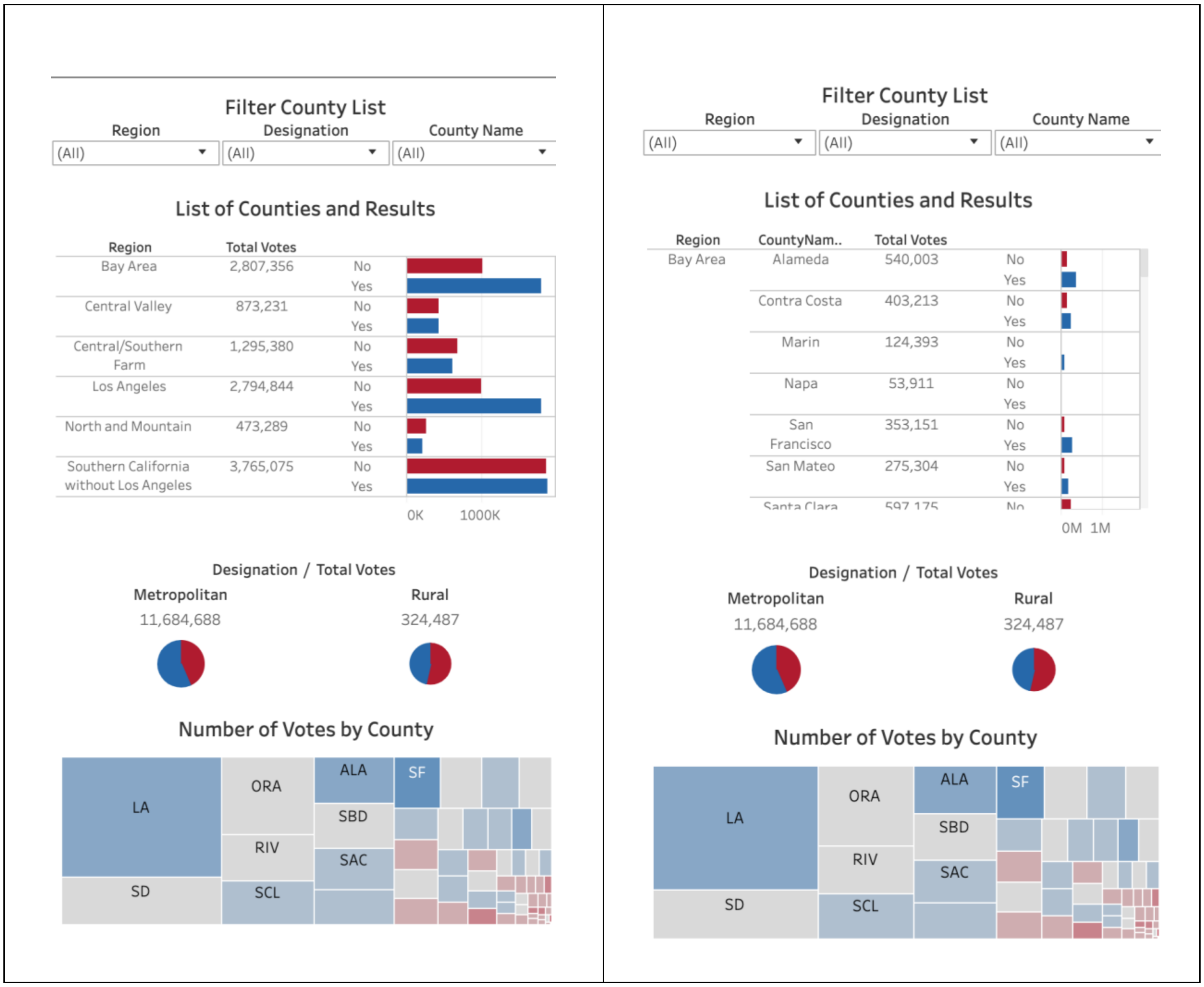


Figure 22 General Results Graphics and County-Specific Results Graphics

3.4.3.3. Dashboard for County-Based Workflows

The dashboard view for county-based workflows is designed for users looking to view how a specific county, group of counties, or region has voted for propositions over time. The same three-column principle is applied here with the left column populated with options to filter counties by name, region, or metropolitan or rural designation, the middle column populated with a map and a description if a proposition is selected, and the right section providing a full list of the propositions and their voting results. Selecting counties from the left column changes the data displayed in the map as well as the voting bars on the right column. Though the voting bars change, the result of a proposition is still displayed to avoid user confusion. Users have access to the same filters for the propositions as the previous view. To switch to the previous view, there are on-scree buttons that allow the user to switch between the various views. Figure 23 shows the county view, with no filters applied or propositions selected.

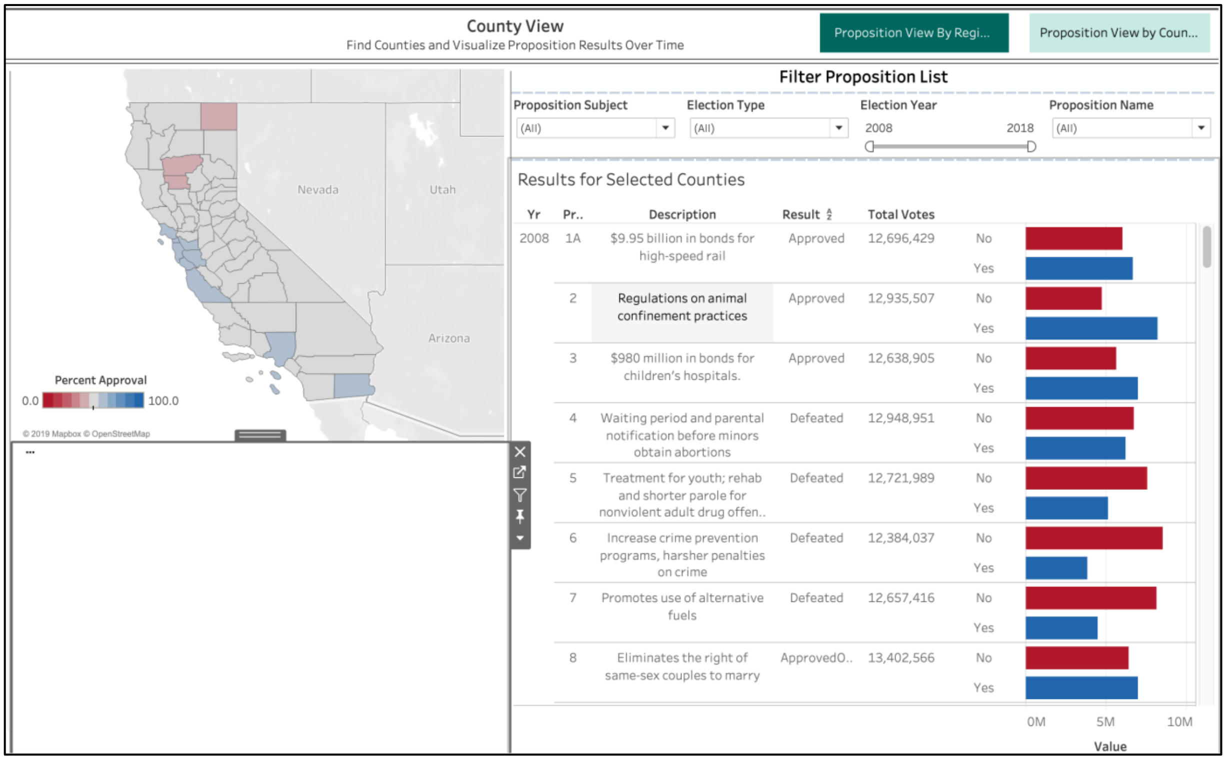


Figure 23 County View, no Filters Applied or Proposition Selected

3.4.4. Sharing the Dashboard

Once the dashboard and its views were configured and ready to be used, it was shared and published to its own website. First, as the dashboard is saved on Tableau Public's desktop interface, it was also published on the Tableau Public website. At this stage, the dashboard was now live and was able to be viewed by anyone accessing the Tableau Public website. The dashboard on the Tableau Public website has additional elements on the screen encouraging users to explore the Tableau platform, look at other dashboards, and read additional details about the project. To create a more simplified experience, the dashboard was embedded onto a personal site with a customized URL and no additional elements outside of the dashboard.

GitHub Pages was used to create the personal website. First, a repository with the project name was created under the developer's account. Next, an HTML file was uploaded to the repository. The HTML file had basic metadata in a header and simple CSS formatting to allow the dashboard to be presented in the center of the webpage. A key component of this CSS is also the hard constraints that restrict the dashboard from scaling beyond the optimal viewing size but still allow for scaling for devices smaller than the optimal viewing size. Also, in the HTML, a line of code provided by Tableau was added to embed the dashboard onto the site. To keep track of site traffic, a Google Analytics tag was added to the HTML before the site was published. Finally, the settings were configured on the project to publish the repository as a webpage. With the GitHub Page activated, the dashboard could now be accessed in a simplified webpage with a configurable URL. Figure 24 captures the HTML used for the Github Page.

```
<!DOCTYPE html>
<html>



    <head>
<!-- Global site tag (gtag.js) - Google Analytics -->
<script async src="https://www.googletagmanager.com/gtag/js?id=UA-120788146-2"></script>
<script>
  window.dataLayer = window.dataLayer || [];
  function gtag(){dataLayer.push(arguments);}
  gtag('js', new Date());

  gtag('config', 'UA-120788146-2');
</script>

        <meta charset="utf-8">
            <meta name="viewport" content="width=device-width, initial-scale=1.0">

        <meta http-equiv="X-UA-Compatible" content="IE=edge">
        <title>CA Ballot Results</title>
        <meta name="description" content="An interactive dashbaord for CA state ballot results
from 2008 - 2018.">

            <style>
body {
  background: #efefef;
}
                .content {
    max-width: 1365px;
    max-height: 768px;
  margin: 0 auto;
}

</style>
    </head>


<body>
<div class="content">
<div class='tableauPlaceholder' id='viz1564816746082' style='position: relative'><noscript><a
href='#'><img alt='2008 - 2018 California Ballot Proposition Results Viewer '
src='https:&#47;&#47;public.tableau.com&#47;static&#47;images&#47;Ba&#47;BallotsAA&#47;Californi
aBallotPropositionResults&#47;1_rss.png' style='border: none' /></a></noscript><object
class='tableauViz'  style='display:none;'><param name='host_url'
value='https%3A%2F%2Fpublic.tableau.com%2F' /> <param name='embed_code_version' value='3' />
<param name='site_root' value='' /><param name='name'
value='BallotsAA&#47;CaliforniaBallotPropositionResults' /><param name='tabs' value='no'
/><param name='toolbar' value='yes' /><param name='static_image'
value='https:&#47;&#47;public.tableau.com&#47;static&#47;images&#47;Ba&#47;BallotsAA&#47;Califor
niaBallotPropositionResults&#47;1.png' /> <param name='animate_transition' value='yes' /><param
name='display_static_image' value='yes' /><param name='display_spinner' value='yes' /><param
name='display_overlay' value='yes' /><param name='display_count' value='yes' /></object></div>
<script type='text/javascript'>                    var divElement =
document.getElementById('viz1564816746082');                    var vizElement =
divElement.getElementsByTagName('object')[0];                    if ( divElement.offsetWidth >
800 ) { vizElement.style.width='100%';vizElement.style.height='768px';} else if (
divElement.offsetWidth > 500 ) {
vizElement.style.width='100%';vizElement.style.height=(divElement.offsetWidth*0.75)+'px';} else
{ vizElement.style.width='100%';vizElement.style.height='100%';}                    var
scriptElement = document.createElement('script');                    scriptElement.src =
'https://public.tableau.com/javascripts/api/viz_v1.js';
vizElement.parentNode.insertBefore(scriptElement, vizElement);                </script></html>
```

Figure 24 HTML Used to Share Site

Chapter 4 Results

At time of publishing, the application met its accessibility requirements by being accessible via a web browser through both desktop and mobile devices. It also supported the three functionality requirements of visualizing proposition results, finding propositions, and viewing county results over time. The application is accessible through GitHub Pages at the URL: https://itsmeodom.github.io/ballots. Two dashboard views support the two workflows mentioned in the previous chapter. On-screen buttons allow the user to easily switch between the views. This chapter provides a walkthrough of the capabilities for each dashboard starting with section 4.1 about the dashboard view for a proposition-based workflow and continues with section 4.2 about the dashboard for a county-based workflow. Each section primarily describes the interactions in the desktop viewer, with supplemental screenshots of the mobile viewer provided at the end of the section.

4.1. Dashboards for Proposition-Based Workflow

The dashboard for proposition-based workflows is the view that initially shows when a user visits the page. This view is seen in Figure 25. The purpose of this dashboard is to allow users to quickly see where propositions were supported or opposed throughout the state. A proposition-based workflow allows users to learn more about counties by visualizing different types of propositions. Its layout presents the data in three columns. The left column acts as a selection column to change displays in the center and right columns. This section describes the process of filtering data and then exploring the resulting data. Screenshots of the mobile views for each of the views are provided to close the section.

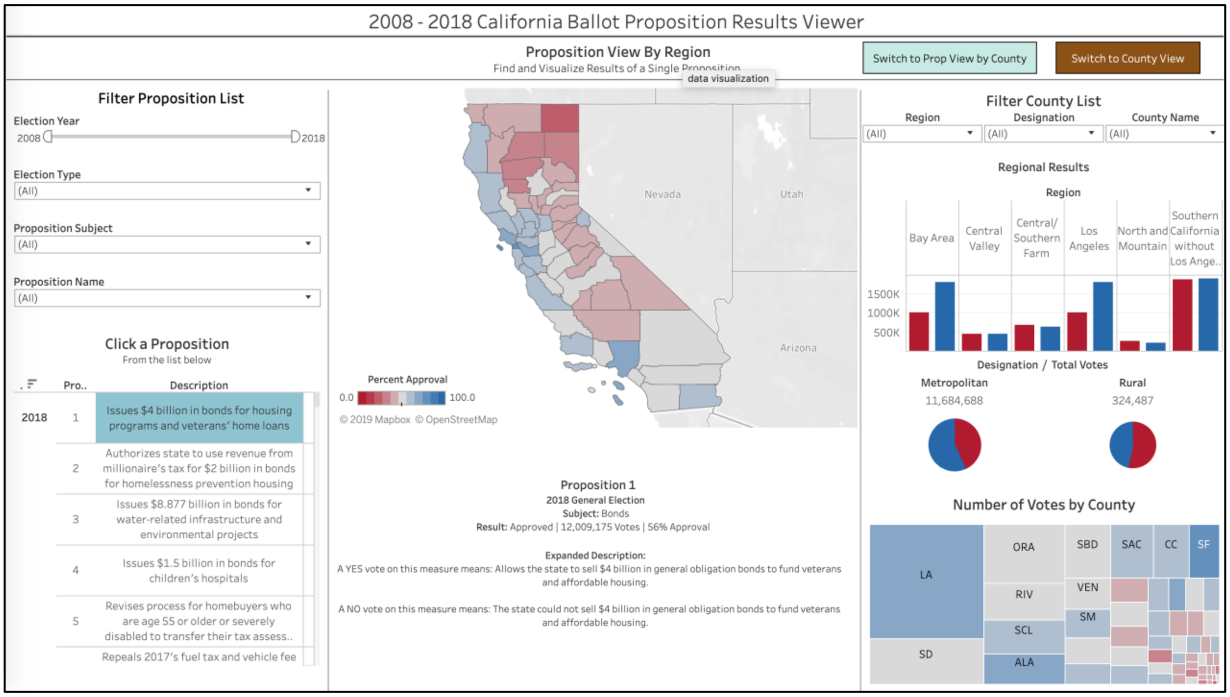


Figure 25 Dashboard for Proposition-Based Workflows

4.1.1. Filtering Proposition Data

The dashboard for the proposition-based workflow was designed for users looking to identify a proposition and visualize the results of that proposition. To help users find specific propositions, filters for proposition type, proposition number, election year, and election type are placed above a selectable list of propositions. When a user applies or removes a filter, the list of selectable propositions updates to reflect only propositions that fit in the selection criteria. In addition, each filter only shows relevant selections based on the other filters applied. For example, if a user selects propositions that were on ballots between 2014 and 2016, the list of proposition subjects is shortened to only subjects represented in that time window. These filters, as seen in Figure 26, allow users to narrow the list of propositions to propositions relevant to their study.

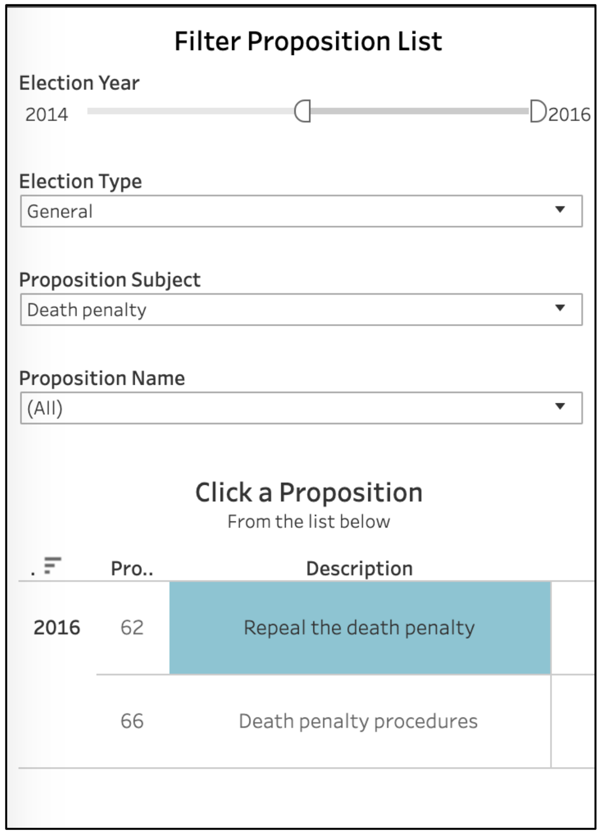


Figure 26 Filters Applied to Proposition List Creates Shorter Lists for Users

4.1.2. Filtering Proposition Data Results

Once a proposition is selected, the map in the center of the dashboard displays how the state voted for that proposition. Under the map is a more comprehensive description of the proposition from the LAO. Hovering on the map displays the county name, the total votes cast in that county, and the percentage of the votes that were cast as "yes." Selecting counties on the map changes the data in the right column to reflect the selected counties. In turn, clicking on regions, counties, or designations on the right column, either on the graphics or on the dropdown menus, also pans and zooms the map to the appropriate area. In this view, users can quickly see where a proposition was supported, identify the counties that cast the most votes, view support across a regional breakdown, and determine view results categorized by counties designated metropolitan or rural. This view, as seen in Figure 27, allows users to get quick generalizations of how the proposition was supported or opposed across the entire state.

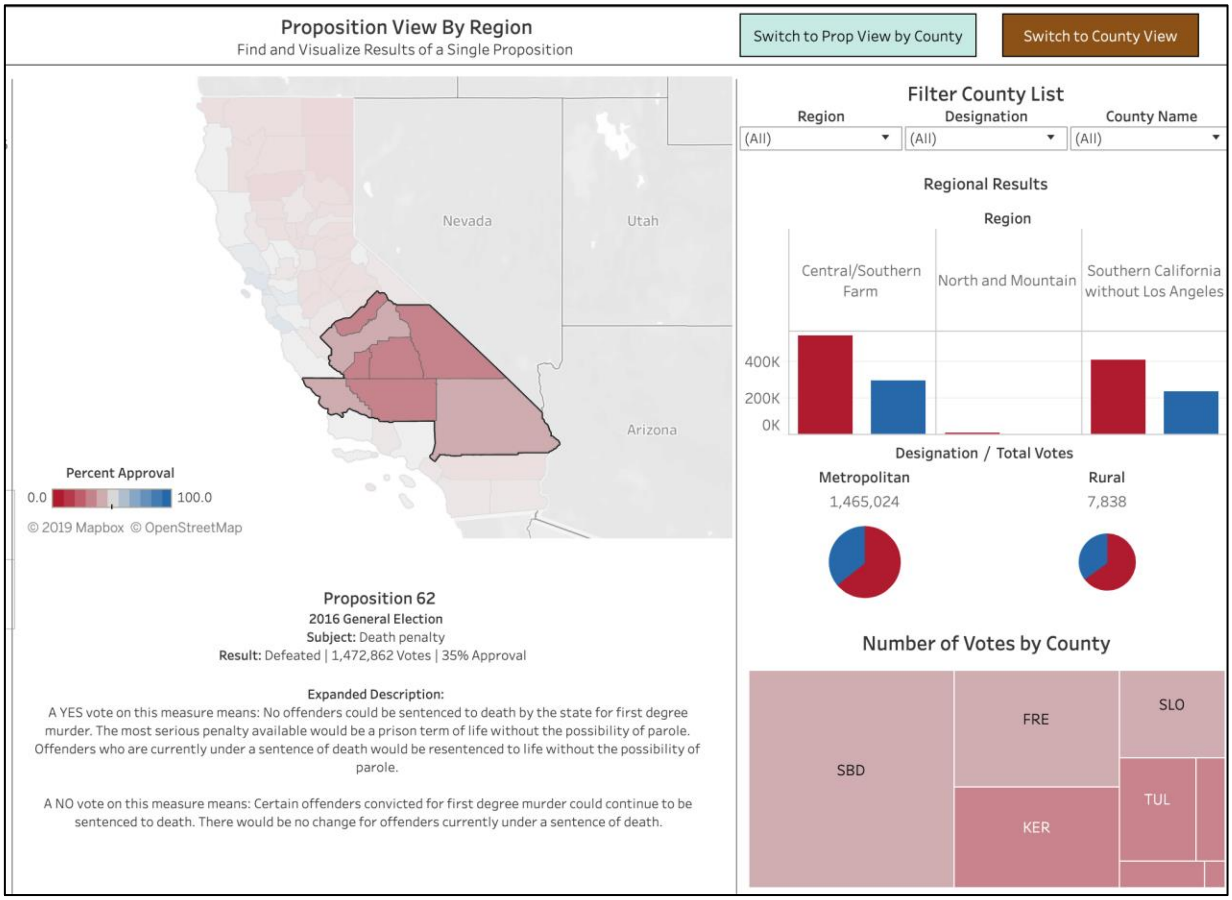


Figure 27 Proposition View with Regional and Designation Overview

If the user wanted to see a list of counties instead of the regional and designation overviews, clicking the column with the “region” header toggles the dashboard to show every county within the region. With the counties toggled on, users have a more granular view to assess the county results in specific regions. For example, in Figure 28, users can see how divisive Proposition 62 of 2016 was in the Bay Area. From the graphic, a user can see how voters in Contra Costa County (CC) and Santa Clara County (SCL) appeared to have equal support and opposition for the proposition, while San Francisco County (SF) and Alameda County (ALA) were clear supporters. In addition, geographic boundaries can easily be put into context as San Francisco County appears to be much smaller than Napa and Marin counties geographically, but using the treemap and the bar charts, users can see that San Francisco County appears to have more votes than Napa and Marin counties combined. Between these two views, users can gather insights about how different propositions were supported or opposed across the state.

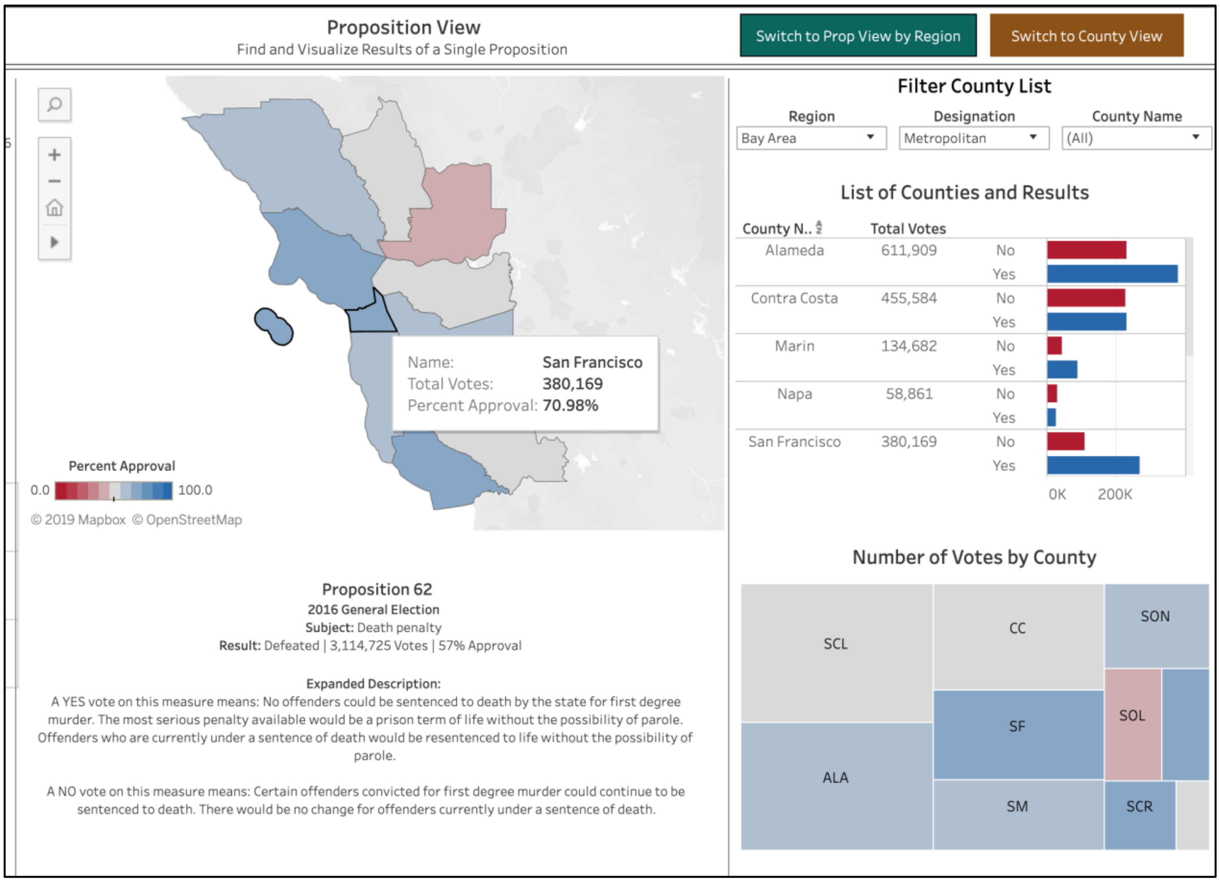


Figure 28 Proposition View with County Results Filtered to Bay Area Counties

4.1.3. Mobile Capabilities

Both of the dashboard views above support mobile access. Views were optimized to allow users to view the dashboards from phones in the portrait orientation. Users still have the ability to use filters through dropdowns or by interacting with the data visualizations. Extended proposition descriptions were omitted to optimize dashboard space. A button at the top of the screen allows users to move to the dashboard for county-based workflows and a button at the bottom of the screen allows users to toggle between the regional overview and the county list. For regional overview and county list dashboard views, the mobile version appears identical when the user loads the application. It is at the bottom of the user's screen where the differences between the regional overview and specific county list is noticed. Figure 29 shows mobile screen captures of the different views.

Figure 29 Mobile View for Both Views

4.2. Dashboard for County-Based Workflow

The purpose of the dashboard for county-based workflows is to see how a particular county, group of counties, or regions has voted for different propositions over time. Using this view, users can identify propositions that had statewide results that were opposite from the selected counties or groups of counties. A county-based workflow allows users to identify specific propositions that may contain region-based anomalies not found in other propositions.

Like the proposition-based workflow, this dashboard presents data in three columns, with the left column acting as a selection column to change displays in the center and right columns. This section describes the process of filtering counties and then exploring the resulting proposition results. Screenshots of the mobile views for each of the views are provided to close the section.

4.2.1. Filtering County Data

When this dashboard view loads, a list of propositions and their results are represented as bar charts on the right column of the dashboard. Using the filters on the left column, users can alter the bars to only show results of selected counties. In addition, as users filter counties, the map in the center column omits the corresponding counties from the map in the center column as well as the checkbox list of county names beneath the filters, as seen in Figure 30. Users can filter the counties with dropdown menus for county name, rural or metropolitan designation, and region.

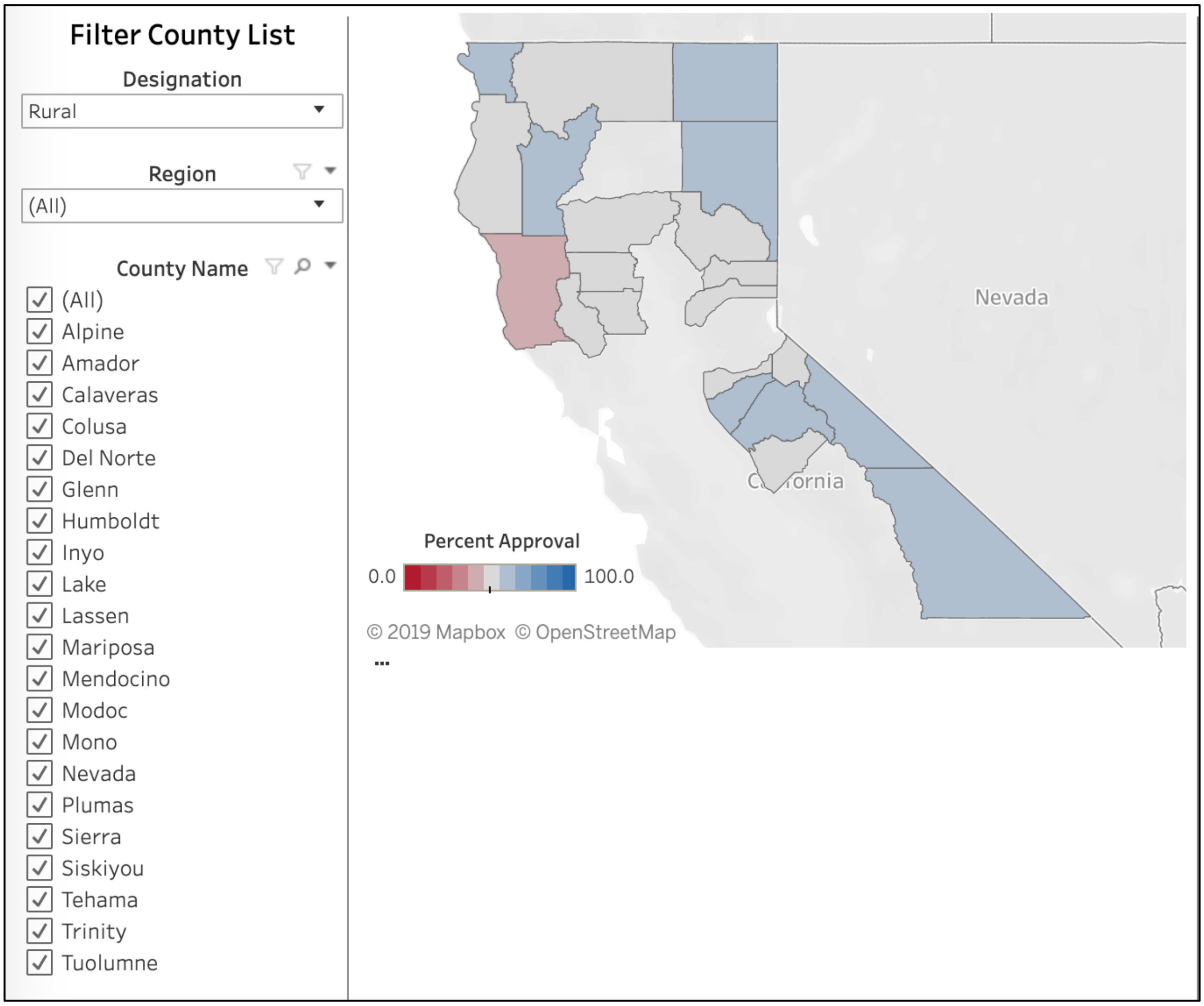


Figure 30 Applying Filters and the Resulting Map

4.2.2. Filtering County Data Results

Once the user selects an area of interest, the map displays the selected counties, and the list of propositions reflect the results for the selected counties. This view allows users to identify propositions of interest by allowing users to see statewide results next to selected county results. For example, a user is able to find propositions that have performed one way in a region but find the opposite result occur statewide. Figure 31 shows the dashboard with the Central Valley regional filter applied. As the user scrolls through the list on the right, the user may notice that Proposition 3 from 2008 was approved despite receiving more opposition than support from the

Central Valley region. Clicking on the proposition from the list displays the same extended proposition description from the previous views. With this knowledge, the user can return to the previous dashboard views to obtain further information about how the proposition performed in the rest of the state, view similar propositions to determine if there are any trends for the area or research the proposition itself to determine proposition-specific factors.

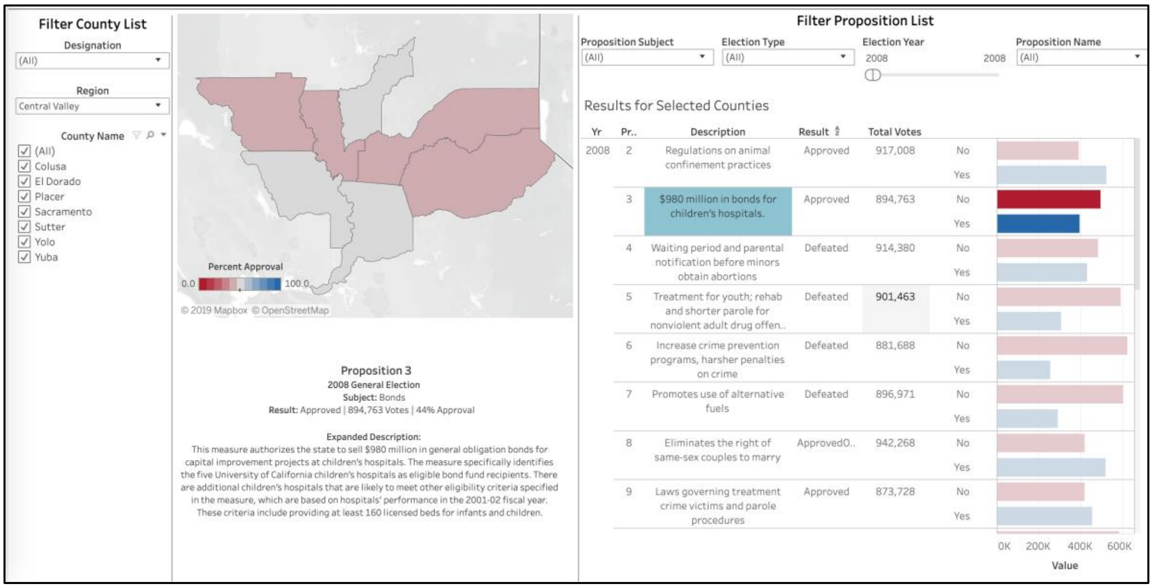
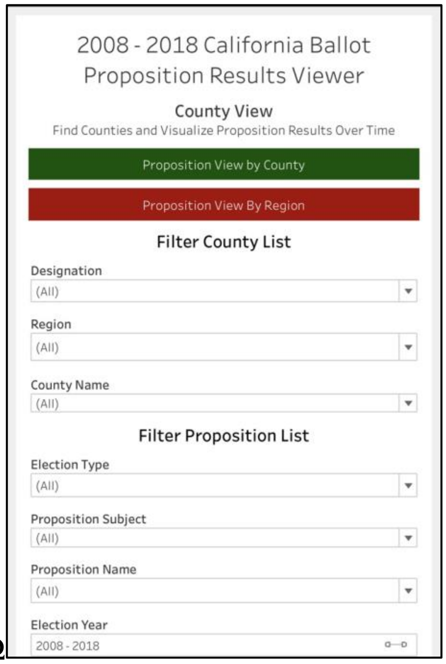


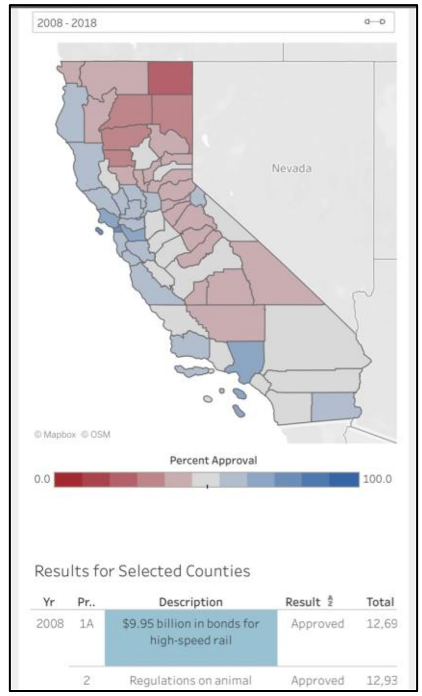
Figure 31 A Proposition that had Opposite Statewide Results from the Selected Region

4.2.3. Mobile Capabilities

This dashboard can also be accessed from a mobile device. At the top of the screen, users are given buttons to change views to the dashboards for proposition-based workflows. Additionally, at the top of the mobile screen are options to filter for counties and propositions. Past these filter options is the map and the bar charts of how each proposition performed in the selected county. To view the full data of the bar chart, users must flip their phone to landscape. Screenshots of this view are shown in Figure 32.

Ω





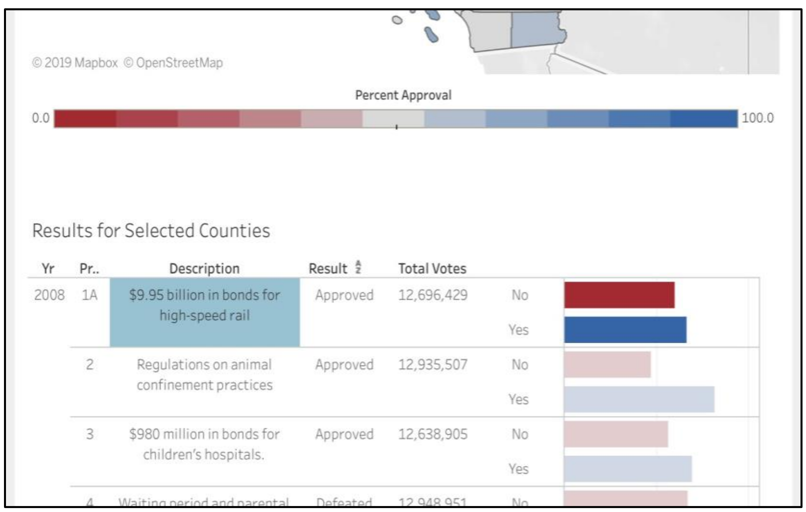


Figure 32 Screenshots of Mobile View

Chapter 5 Discussion

The application successfully achieved the goals outlined in section 3.1. In addition to the goals, the capabilities of the application are in-line with the initial motivations of creating a tool to find voter bubbles and consolidate proposition results to a single page. For discussion, this chapter contains three sections. First, section 5.1 reviews the challenges and obstacles faced during application creation. Next, section 5.2 compares the application to similar applications reviewed in 2.3. Finally, chapter 5.3 provides insight about possible opportunities for future development of the application.

5.1. Challenges

Building an application designed to meet the requirements proved to be more difficult than expected. The most prominent challenge involved the layout. An aim of the application was to present multiple views of the same data on the same screen. Determining which visualizations to include and determining where to place the visualizations on the dashboard was timely and difficult. A second challenge stemmed from the limitations in the software in mobile view.

5.1.1. Layout Challenges

Configuring a layout that users were able to understand that was also able to be viewed on various screen sizes proved to be one of the most challenging pieces of this project. This challenge can be defined in two stages. First, the challenge of creating an initial layout that met requirements, and second determining how to scale the layout to different sizes.

Before loading any data into Tableau, several wireframes were drawn as mockups to visualize potential layouts. These mockups were used as a guide as the application was developed. Unfortunately, the first iteration of the application was deemed confusing, difficult to use, and not providing enough information in informal testing through colleagues and friends of the developer. The challenge of adding more information yet simplifying the application to be less confusing led to a complete redesign of the original application, starting back from the mockups. During the second integration, the mockups were shown to the same informal user group to solicit feedback. This feedback led to the idea of using multiple views to work with the two workflows. By incorporating the feedback earlier in the process, the design changes after the dashboard was published were significantly less than the first iteration.

A second challenge was creating a dashboard responsive enough to display relevant data on different screen sizes for desktops, tablets, and phones. For mobile devices, this challenge was not as difficult, because Tableau allows developers to create layouts specific to phones if the display size is within a certain threshold. The larger challenge was developing a dashboard that scaled appropriately on different desktop sizes. Initially, the automatic layout defined by Tableau was thought to have been sufficient to cover different desktop sizes, when the dashboard was viewed on a screen larger than a 15-inch laptop, the dashboard scaled with a lot of empty space between the visualizations. One solution was to use a fixed layout with the viewing size defined in Tableau Public. With this solution, regardless of the screen size of the desktop, users were presented with a dashboard sized at 1366px x 768px, which Tableau defines as "generic desktop." On screens larger than 1366px x 768px, the dashboard appeared with appropriate spacing and clear visualizations. However, users with screen sizes smaller than the defined size were forced to scroll their browsers windows left and right or up and down to see all the data. In the spirit of the dashboard's definition of "displaying data on a single screen," scrolling browser windows was determined to not be an appropriate solution as well. The published solution selected utilized GitHub Pages to define max height and max-width constraints that did not allow the dashboard to be scaled larger than 1366px x 768px. This allowed the dashboard to be published using the automatic layout option but restrict the dashboard to a maximum viewing size to prevent excessive spacing between the visualizations.

5.1.2. Mobile Limitation

A requirement for the application was to be accessible by mobile and tablet devices. Due to differences in the sizing of screens, it was understood that some visualizations might need to be omitted in the mobile views, but the application should still behave the same way. One limitation in mobile view is the lack of an option to use a drill-down menu as seen on the regional list of counties. On the desktop version, while in proposition view, a user can expand the table of regions and results to reveal a list of counties within that region. On the mobile view of the application, this option is missing, and the developer was unable to add it to the visualization. A workaround was developed that required the proposition view to be duplicated, and then modified to replace the regions and counties results list with a list of only county names. This visualization can only be accessed in mobile, and it does allow the user to view the list of county names from a mobile device. However, toggling between views on mobile is noticeably slower than expanding the regions on the desktop.

5.2. Comparison to Available Applications

Section 2.3 reviewed different applications that displayed voting results. These applications inspired the different visualizations used in the dashboard. This section discusses the influence of each application.

5.2.1. Secretary of State Map – The Basic Information

The template of the basic information the application was to display was inspired from the Secretary of State website shown in Figure 3. During initial development of this application, the Secretary of State website hosted its own web GIS application that allowed users to select a proposition, see a one-sentence description of the proposition, and see which counties were generally in favor of the proposition and which counties opposed the proposition. After official election results became official, the website removed the maps from their website. Using the same data from the Secretary of State, the application was able to display the same information as well as expand the information. A major difference between the Secretary of State map and the current application is the Secretary of State map does not show multiple visualizations, does not include a detailed description of the propositions and does not allow users to easily switch to different years of ballot propositions.

5.2.2. New York Times Map – Symbolizing the Counties by Voter Preference

"An Extremely Detailed Map of the 2016 Election" from Figure 1 showed results of the 2016 Presidential Election at the precinct level. Precincts were symbolized red to indicate support for Trump and blue to indicate support for Clinton. In addition, the more support the candidate had in a precinct, the darker the precinct appeared. This concept of a diverging scale symbolization influenced the way the map was symbolized in the California Ballots application. Instead of red and blue symbolizing candidates, red symbolized opposition and blue symbolized support. Using the same logic, the more support or opposition within a county, the darker that county appeared. County data was used in the application, but the New York Times uses precincts. Expanding the application to include precinct data will be discussed further in section 5.3.2.

5.2.3. Election Results in 3D – Visualizing the Number of Voters within a County

Esri's 3D map shown in Figure 2 used height to visualize voter preference as well as the number of voters. By using a diverging color scheme, the map can indicate the voting preference of a selected area, and by using 3D, the map utilizes height to visualize the number of voters in that area. While the California Ballots application did not utilize any 3D, the 3D map inspired the need for a visualization that allowed users to easily see which county cast the most votes. The visualization selected was a treemap, and while the user has to look at multiple visualizations if they want to see number of voters and spatial context, the treemap allows users to see easily see which counties cast more votes as well as the general voting preference within each county. A key difference between the application and Esri's map is the Esri map is only focused on one election, while the application allows users to visualize multiple ballots on multiple elections.

5.2.4. CNN Map of Results– Side-by-Side Visualizations

The CNN shown in Figure 5 showed a similar map to the Secretary of State map except that the CNN map included a stacked bar graph to allow users to visually see how many votes were cast in support and opposition. The idea of adding visualizations that display the total votes categorized by support or opposition inspired the inclusion of the same bar graphs appended to the list of counties and regions in the proposition view, as well as the bar graphs appended to the propositions in county view. While the CNN visualization did not include a pie chart, the pie chart for rural and metropolitan voters was inspired by the same notion of displaying total support and total opposition in side-by-side visualizations.

5.2.5. Coloradoan Map– Providing Text Context

A series of maps published as part of a story in the Coloradoan inspired the inclusion of additional text details beyond the single sentence description provided by the Secretary of State, as seen in Figure 4. In the story, the author indicates what each proposition covered in the story would have enacted if passed, and what was the result if the proposition failed. This style of including context influenced the inclusion of similar summaries of the propositions from the California Legislative Analyst's Office. While the Coloradoan map only displays maps for select propositions from the election covered in the story, this application allows users to see descriptions and results from every proposition in the ballots covered.

5.3. Future Development

This application can be considered a successful prototype that can lead to a more robust and sophisticated application for viewing ballots. The next logical steps for this application are formalizing a user group and expanding the data. A user group would allow the developer to receive valuable feedback about the application's current state as well as insight about useful additions or modifications to the application. The pilot dataset for this application was made from county results of California's ballot propositions from 2008 to 2018. Expanding these datasets could allow users to obtain more sophisticated insights. In this section, section 5.3.1 further discusses a potential user group and section 5.3.2 discusses potential data expansions. Section 5.3.3 closes the chapter with a discussion of how the application can be further developed to fit the needs of studies from section 2.1.

5.3.1. Formalized User Group and Testing

A formalized user group would be hugely beneficial in the next stages of this application's development. The application at its current stage would provide a working proof-of-concept that would allow members of this user group to test the application in their own line of work. Ideally, the user group would have strong domain knowledge and would have members that have produced academic work as well as members that have produced work in the news media. To avoid political biases, it is important that testing is done on large, random, and diverse users. At its current stage, most decisions about application layout, data to include, and actions available on the dashboard, were made based on the developer's knowledge and informalized testing. Incorporating a user group would help ensure that as the application is further developed, it will provide greater value to specific to its target audiences.

5.3.2. Expanding the Data

The application at its current state visualizes data at the county level for propositions on ballots from 2008 to 2018. The application uses three aggregate datasets. First, a results dataset that stores voting results per proposition per county. Next, a county dataset that stores county geometries and additional county attributes. Third, a proposition descriptions dataset that stores specific information about the propositions. Three ways to expand this dataset would be to include results for more propositions, obtain more granular datasets within the spatial limits of a county, and include even more specific details about a proposition. Should application development continue, the user group described in the previous section would help decide what data could potentially add value. Before any expansion, it is important that update and maintenance procedures are properly documented to allow for development and maintenance from different developers. These documents can be stored on the application's GitHub repository.

Including more propositions could be achieved through a similar process as the initial propositions used in the application, but before 2004, the Secretary of State only provides results in PDF format rather CSV. If the user group reported a significant value in expanding the proposition list, converting results data from PDF into a spreadsheet would likely be a substantial task.

Incorporating data at smaller scales could also become a high-effort task. After 2004, the Secretary of State published results by areas within county in CSV format, but the format is not consistent throughout the years. Reformatting the data into a form that is usable by Tableau would require each results sheet to reformatted, and it is likely that similar techniques could not be used between years due to differences in formatting. Regarding voting district data, this data would likely have to be acquired by each county's registrar of voters. The outreach necessary to obtain this data will likely be a timely effort.

A consideration for expansion of the proposition dataset would be to include data for who supported the proposition. This could be from a political standpoint, where elected officials or media outlets pledge their support for a proposition, or it could be from a financial standpoint and include details about who provided financial support for a proposition. These considerations stem from the subjects of studies mentioned in section 2.1.1.

5.3.3. Additional Developments to Support Existing Studies

Section 2.1.2 described different works that included a spatial component in their ballot studies. Two of the studies describe how a voter's proximity to an issue can impact the way the voter casts their ballot (Branton et al. 2007; Boehmke et al. 2012). Another study reviews years of Colorado's proposition results data to determine whether rural counties saw propositions they supported passed by the state (Smith 2008). Finally, a study describes the impact state ballots can have on congressional votes (Huder et al. 2010). This section describes different ways this application can be used within the context of these studies.

In the studies describing how a voter's proximity to an issue impacts the voter's decision, the study by Branton et al. (2007) describes how political party influence differed between Democrat voters close to issues and Democrat voters located further from the issue. A potential development for this application could include indicators that displayed results categorized by political party. The challenge would be in data acquisition. In their study, Branton et al. (2007) collect the data using California Field Polls. Another study that examines the role of geography in voter decisions uses party data collected from state-level exit polls (Branton 2003). While these methods were deemed acceptable for their respective studies, results data collected from a source other than the Secretary of State would require careful consideration if that data is then related or joined to official data.

Another potential future development would be the inclusion of demographic data. This idea stems from the Smith (2008) study that discussed whether voters in rural counties had opinions that were weighted the same as voters in metropolitan counties. This study influenced the addition of the rural and metropolitan categorizations in the current stage of the application. Demographic data would be in the same light of exploring whether there are certain segments of the population who rarely see their opinion passed by the majority. This data collection would also be tricky, as demographic data would have to be acquired for each election to accurately represent the population of an area at the time of a vote.

Finally, another potential addition to the proposition description dataset would be to identify related propositions at local and national levels. This addition would address a challenge mentioned in the Huder et al. study (2010). In their study, the authors indicate the difficulty of finding propositions at state levels that were related to national ones. Effectively tagging the propositions and storing them in a database accessible to the application would help address their challenge, as well as provide an avenue to continue their study by examining the relationship between local ballots and state ballots.

References

Alesheiskh, Ali Ashgar, Hussein Helali, and H. A. Behroz. 2002. *Symposium on Geospatial Theory, Processing and Applications* 15.

"Ballot Initiatives and Propositions." 2019. *Ballot Analysis*. Accessed September 3. https://lao.ca.gov/BallotAnalysis.

"Ballot Initiatives." 2019. *State of California - Department of Justice - Office of the Attorney General*. March 15. https://oag.ca.gov/initiatives.

Becker-Olsen, Karen L., B. Andrew Cudmore, and Ronald Paul Hill. 2006. "The Impact of Perceived Corporate Social Responsibility on Consumer Behavior." *Journal of Business Research* 59 (1): 46–53. https://doi.org/10.1016/j.jbusres.2005.01.001.

Blee, Brendan. 2016. "Creating a Geodatabase and Web-GIS map to visualize drone legislation in the state of Maryland." Master's thesis, University of Southern California

Bloch, Matthew, Larry Buchanan, Josh Katz, and Kevin Quealy. 2018. "An Extremely Detailed Map of the 2016 Presidential Election." *The New York Times*. The New York Times. July 26. https://www.nytimes.com/interactive/2018/upshot/election-2016-voting-precinct-maps.html.

Boehmke, Frederick J., Regina P. Branton, Gavin Dillingham, and Richard C. Witmer. 2012. "Close Enough for Comfort? The Spatial Structure of Interest and Information in Ballot Measure Elections." *The Journal of Politics* 74 (3): 827–39. https://doi.org/10.1017/s0022381612000321.

Branton, Regina P. 2003. "Examining Individual-Level Voting Behavior on State Ballot Propositions." *Political Research Quarterly* 56 (3): 367–77. https://doi.org/10.2307/3219796.

Branton, Regina, Gavin Dillingham, Johanna Dunaway, and Beth Miller. 2007. "Anglo Voting on Nativist Ballot Initiatives: The Partisan Impact of Spatial Proximity to the U.S.-Mexico Border." *Social Science Quarterly* 88 (3): 882–97. https://doi.org/10.1111/j.1540-6237.2007.00488.x.

Butler, Declan. 2006. "The Web-Wide World." *Nature* 439 (16): 776–78. http://www.acsu.buffalo.edu/~lbian/nature_googleearth.pdf.

"California Proposition 6." 2018. *Cable News Network*. https://www.cnn.com/election/2018/results/california/ballot-measures/1.

"California." 2019. *Initiative & Referendum Institute*. University of Southern California. Accessed July 2. http://www.iandrinstitute.org/states/state.cfm?id=4.

Conner, Philipp. 2018. "Exploring Commercial Catch: Creating a Responsive Florida Fisheries Web GIS Using ASP .NET, the Esri JavaScript API 4.x, and Calcite Maps." Master's thesis, University of Southern California

Correll, Michael, Mingwei Li, Gordon Kindlmann, and Carlos Scheidegger. 2019. "Looks Good To Me: Visualizations As Sanity Checks." *IEEE Transactions on Visualization and Computer Graphics* 25 (1): 830–39. https://doi.org/10.1109/tvcg.2018.2864907.

"Dashboards." 2019. *Tableau*. Accessed July 2. https://help.tableau.com/current/pro/desktop/en-gb/dashboards.htm.

"Data Policy." 2019. *Tableau Public*. February 28. https://public.tableau.com/en-us/s/data-policy.

Davidson, Robert Howard. 2014. "Evaluating Glacier Movement Fluctuations Using Remote Sensing: A Case Study of the Baird, Patterson, Leconte, and Shakes Glaciers in Central Southeastern Alaska." Master's thesis, University of Southern California

Difranzo, Dominic, and Kristine Gloria-Garcia. 2017. "Filter Bubbles and Fake News." *XRDS: Crossroads, The ACM Magazine for Students* 23 (3): 32–35. https://doi.org/10.1145/3055153.

Dragićević, Suzana. 2004. "The Potential of Web-Based GIS." *Journal of Geographical Systems* 6 (2): 79–81. https://doi.org/10.1007/s10109-004-0133-4.

"Elections and Voter Information." 2019. *Elections and Voter Information | California Secretary of State*. Accessed September 3. https://www.sos.ca.gov/elections/.

Ensley, Michael J., and Erik P. Bucy. 2009. "Do Candidate Positions Matter?" *American Politics Research* 38 (1): 142–64. https://doi.org/10.1177/1532673x09337190.

Few, Stephen. 2006. "Information Dashboard Design." *Perceptual Edge*. http://blogs.ischool.berkeley.edu/i247s12/files/2012/01/Dashboard-Design-Overview-Presentation.pdf.

Field, Ken. 2017. "High Rise Election." *Arcgis.com*. http://www.arcgis.com/apps/CEWebViewer/viewer.html?3dWebScene=f034d08dbde649fe97ed2f3ec5bbf381.

Fu, Pinde, and Jiulin Sun. 2011. *Web GIS: Principles and Applications*. Redlands, CA: ESRI Press.

Gerhardt, Teri. 2011. "Utilizing Cityworks at Cupertino, California, with Many Tangible Results." *Cityworks*. August 1. https://www.cityworks.com/2011/08/utilizing-cityworks at-cupertino-california-with-many-tangible-results/.

Goldsworth, Julia. 2017. “Exploring Land Use Changes in the City of Irvine’s Master Plan.” Master’s thesis, University of Southern California’

Haveman, Heather A., Hayagreeva Rao, and Srikanth Paruchuri. 2007. “The Winds of Change: The Progressive Movement and the Bureaucratization of Thrift.” *American Sociological Review* 72 (1): 117–42. https://doi.org/10.1177/000312240707200106.

“History of California Initiatives.” 2019. *History of California Initiatives | California Secretary of State*. Accessed July 2. https://www.sos.ca.gov/elections/ballot-measures/resources-and-historical-information/history-california-initiatives/.

Huder, Joshua, Jordan Michael Ragusa, and Daniel A. Smith. 2010. “Shirking the Initiative? The Effects of Statewide Ballot Measures on Congressional Roll Call Behavior.” *American Politics Research* 39 (3): 582–610. https://doi.org/10.1177/1532673x10388137.

Klemens, Kevin. 2017. “Development and Evaluation of a USV Based Mapping System for Remote Sensing of Eelgrass Extent in Southern California.”Master’s thesis, University of Southern California

“List of Rural Counties And Designated Eligible Census Tracts in Metropolitan Counties.” 2018. Rep. *List of Rural Counties And Designated Eligible Census Tracts in Metropolitan Counties*. Office of Rural Health Policy. https://www.hrsa.gov/sites/default/files/hrsa/ruralhealth/resources/forhpeligibleareas.pdf.

Lunden, Ingrid. 2019. “Salesforce Is Buying Data Visualization Company Tableau for $15.7B in All-Stock Deal.” *TechCrunch*. TechCrunch. June 10. https://techcrunch.com/2019/06/10/salesforce-is-buying-data-visualization-company-tableau-for-15-7b-in-all-stock-deal/.

“Mapbox GL.” 2019. *Mapbox*. Accessed July 2. https://docs.mapbox.com/help/glossary/mapbox-gl/.

Norman, Donald A. 2013. *The Design of Everyday Things*. Basic Books.

Owens, John R., and Larry L. Wade. 1986. “Campaign Spending on California Ballot Propositions, 1924-1984: Trends and Voting Effects.” *The Western Political Quarterly* 39 (4): 675. https://doi.org/10.2307/448270.

Pappas, Lisa, and Lisa Whitman. 2011. “Riding the Technology Wave: Effective Dashboard Data Visualization.” *Lecture Notes in Computer Science Human Interface and the Management of Information. Interacting with Information*, 249–58. https://doi.org/10.1007/978-3-642-21793-7_29.

Parsons, Jonathan. 2014. “Mapping Uniformity of Park Access Using Cadastral Data Within Network Analyst in Wake County, NC.”Master’s thesis, University of Southern California

Powell, Rebecca. 2018. “Colorado Election: Here's Where Five Failed Statewide Ballot Issues Actually Passed.” *Coloradoan*. Fort Collins Coloradoan. November 9. https://www.coloradoan.com/story/news/politics/elections/2018/11/09/colorado-election-issues-where-failed-ballot-measures-passed/1932207002/.

Roberts, James A. 1995. “Profiling Levels of Socially Responsible Consumer Behavior: A Cluster Analytic Approach and Its Implications for Marketing.” *Journal of Marketing Theory and Practice* 3 (4): 97–117. https://doi.org/10.1080/10696679.1995.11501709.

Roberts, Jonathan C. 2007. “State of the Art: Coordinated & Multiple Views in Exploratory Visualization.” *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)*. https://doi.org/10.1109/cmv.2007.20.

Smith, Daniel A. 2008. “Representation and the Spatial Bias of Direct Democracy.” *University of Colorado Law Review* 78. https://www.researchgate.net/profile/Daniel_Smith37/publication/265659837_REPRESENTATION_AND_THE_SPATIAL_BIAS_OF_DIRECT_DEMOCRACY/links/552bf3a00cf29b22c9c242b5.pdf.

“Statewide Election Results.” 2019. *Statewide Election Results | California Secretary of State*. Accessed September 3. https://www.sos.ca.gov/elections/prior-elections/statewide-election-results/.

Stratmann, Thomas. 2004. “The Effectiveness of Money in Ballot Measure Campaigns.” *Southern California Law Review* 78 (1041). https://pdfs.semanticscholar.org/558a/efe4ebc13fc3b5db592024e9f23ee530323d.pdf.

Sui, Daniel, and Michael Goodchild. 2011. “The Convergence of GIS and Social Media: Challenges for GIScience.” *International Journal of Geographical Information Science* 25 (11): 1737–48. https://doi.org/10.1080/13658816.2011.604636.

“Tableau Software Site Usage Agreement.” 2015. *Tableau Software*. January 5. https://www.tableau.com/tos.

Tobler, W. R. 1970. “A Computer Movie Simulating Urban Growth in the Detroit Region.” *Economic Geography* 46: 234. https://doi.org/10.2307/143141.

Tran, Mary, Greg Jones, and Paul Smilanick. 2001. “Regions of California: Recommended Grouping of the Counties for Regional Studies.” Rep. *Regions of California: Recommended Grouping of the Counties for Regional Studies*. California Department of Social Services. https://www.cdss.ca.gov/research/res/pdf/multireports/RegionsofCalifornia.pdf.

Tufte, Edward R. 2006. *Beautiful Evidence*. Cheshire, CT: Graphics Press.

US Census Bureau. 2019. "TIGER/Line Shapefiles." *TIGER/Line Shapefiles*. May 6. https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html.

"What Is GIS?" 2019. *Esri*. Accessed July 2. https://www.esri.com/en-us/what-is-gis/overview.

Wijk, J.j. Van. n.d. "The Value of Visualization." *VIS 05. IEEE Visualization, 2005.* https://doi.org/10.1109/visual.2005.1532781.

Yi, Ji Soo, Youn Ah Kang, and John Stasko. 2007. "Toward a Deeper Understanding of the Role of Interaction in Information Visualization." *IEEE Transactions on Visualization and Computer Graphics* 13 (6): 1224–31. https://doi.org/10.1109/tvcg.2007.70515.