Data Science Quarterly: Manuscript for Research Section
Volume 1, Issue 1, January 2021
ISSN 2693-8871 (print) ISSN 2693-8863 (online)

Political Horizon Charts

Thomas W. Miller

Working paper revised 11 January 2021

Abstract With the US presidential election every four years, the popular media are occupied by maps of the United States in shades of red and blue. These maps are fine for showing the status of campaigns at a single point in time as journalists anticipate the outcome in the Electoral College. But if we want to display campaign events over time there is a better tool—the political horizon chart. A Republican candidate is pitted against a Democratic candidate, and we compute differences in poll results and election forecasts. We use horizon charts to track the progress of political campaigns.

Keywords Prediction markets · Election forecasting · Data visualization · Choropleth map · Horizon chart · Data journalism · Lattice graphics

1 Introduction

Consider a new way of displaying data leading up to an election, including opinion polling, forecasting, voting, and prediction market data. Consider a method that can represent relative strengths of parties and candidates as campaigns play out over time. Data journalists would be well-advised to consider the possibilities of political horizon charts.

2 Literature Review

The horizon chart was invented more than a decade ago as a time series visualization tool (Few 2008; Heer, Kong, and Agrawala 2009). It falls within the general class of trellis or lattice graphics, which display multiple plots across a set of identically scaled panels (Cleveland 1993; Sarkar 2008).

The author thanks representatives of PredictIt (www.predictit.org) for providing market data relating to this research. He also thanks Isometric Solutions, Panel Consulting Group, and Prodege, LLC for survey implementation and panel recruitment.

3 Methods

To construct a political horizon chart, we need one or more time series, repeated observations in time. Ideally, the observations are taken at equally spaced time intervals, such as hours or days. If there are multiple time series, measures are adjusted to conform to a common scale across overlapping time intervals.

For elections, we estimate margin-of-victory or chance-of-winning percentages for the Democratic candidate minus the Republican candidate. Positive values represent an advantage for the Democrat, negative values an advantage for the Republican. At zero, the Democrat and Republican have an equal chance of winning.

Red for Republican and blue for Democrat are commonplace in popular media. We keep with that theme, using reds for Republican-advantaged, negative values and blues for Democratic-advantaged, positive values.

We define a color gradient with ten distinct colors across a 10- or 100-point range. Using a 10-point range from -5 to +5, each color encompasses a one-point interval, as we would use in comparing election margin-of-victory percentages. Using a 100-point range from -50 to +50, each color encompasses a ten-point interval, as we would use in percentage of victory comparisons between candidates. We use white at zero or the point of indifference between parties or candidates. We use five shades of red and five shades of blue on the negative and positive sides of the scale, respectively. It is most convenient to select specific hues of red and blue, varying shades by saturation or brightness. Less saturated, lighter colors are closest to the white zero-point. More saturated, darker colors are at the extreme points on the scale. Values outside the range of the color gradient are given the extreme colors—dark red and dark blue for values at the negative and positive extremes, respectively.

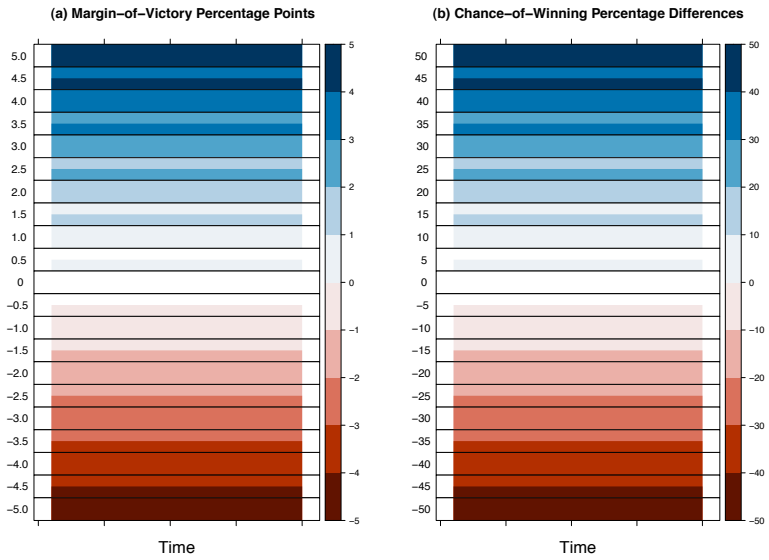
Panels in the political horizon chart are identified with various measures, geographies, voting groups, elections, or electoral markets. Each panel includes measures across time. Every value being plotted at a panel time-point is displayed as one or two colors.

To determine the color or colors to be plotted at a panel time-point, we find the value being plotted on the color gradient or scale. For a negative value, we look up from the value, identifying an interval of solid red, two shades of red, or light red and white.

The horizon chart gets its name from the appearance of panels. Darker colors at the bottom of panels can look like mountains, with lighter colors at the top, like a daytime sky. The rendering of negative values is straightforward, with darker reds on the bottoms of panels. The rendering of positive values is less straightforward.

Positive values are plotted as solid blue, two shades of blue, or light blue and white. When displaying positive values, we refer to the color gradient and look down (rather than up). When we see two blues within an interval, we flip those blues prior to plotting at panel time-points. In keeping with the horizon or mountain motif, we place darker blues at the bottom of panels, even though

Fig. 1 Political Horizon Chart Cheat Sheet



darker blues are associated with higher numeric values. This peculiarity in the rendering of blues can make it difficult for readers new to horizon charts to discern the exact values being displayed.

To develop a better understanding of the political horizon chart, we consult a cheat sheet showing how colors from the gradient are displayed within panels. See figure 1a for the color gradient representing the margin-of-victory scale from -5 to +5 and figure 1b representing the percentage difference scale from -50 to +50.

Referring to the cheat sheet, we note how near-zero values are translated into colors. To represent values between -1 and zero on a margin-of-victory scale, we look up on the color gradient only so far as the zero-point. We display values between -1 and zero with light red in the lower portion of the panel and white in the upper portion. In like fashion, to represent values between zero and +1 on a margin-of-victory scale, we look down the color gradient but only so far as the zero-point. Flipping colors as appropriate for positive values, we display values between zero and +1 with light blue on the lower portion of the panel and white on the upper portion. White is the filler color for panel-points with values in the -1 to +1 range on the margin-of-victory scale. No panel-point value is displayed with both light red and light blue.

In sum, we identify panel-point colors by traversing the color gradient, identifying color swatches within a one- or ten-point range for margin-of-victory or percentage-difference scales. Going from strongly Republican to strongly Democratic voters or electoral regions, we define panel-point colors of the fol-

lowing varieties: solid red, bicolor red, light red and white, solid white, light blue and white (flipped), bicolor blue (flipped), and solid blue. Solid red is strongly Republican, solid blue strongly Democratic. To identify competitive voting districts or "swing states" in a national election, we look for panel regions of white. The more white color we see, the more competitive the contest. A zero value is plotted as solid white, indicating an exact tie between the candidates or parties.

4 Results

For a first example, we examine prediction market time series from the 2020 presidential election. We have the Democratic challengers, Biden-Harris, pitted against the incumbent Republican ticket, Trump-Pence. We observe hourly prices in prediction markets for the Democratic and Republican tickets and convert these prices are into chance-of-winning percentages.

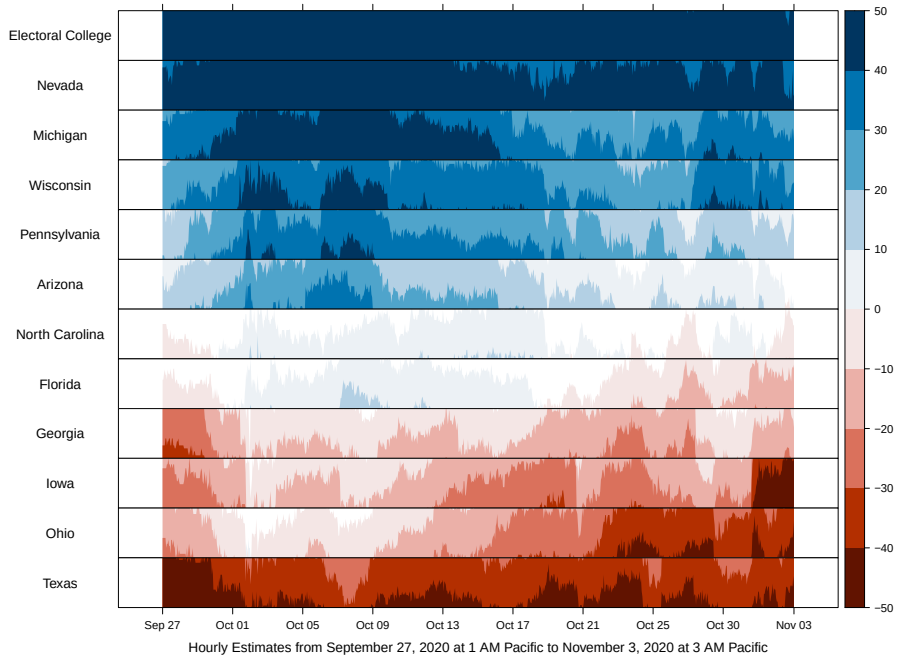
Simulating one million hypothetical elections each hour, we estimate chance-of-winning percentages for the Democratic and Republican tickets in the Electoral College (Miller 2021). We plot the Democratic percentage minus the Republican percentage as the first panel in the political horizon chart in figure 2, with additional panels for chance-of-winning percentages for eleven selected states.

The popular media uses choropleth maps of the United States to display voting projections. With 56 electoral markets, journalists can show 56 color-coded regions—56 observations. In contrast, consider the potential of the political horizon chart for electoral market data. The first example covers 27 days, 24 hours a day, across 12 measurement panels. That is 7,776 observations: $12 \times 27 \times 24 = 7776$. We could easily have represented all 56 electoral markets and the Electoral College forecast for 50 days, rendering 68,400 observations: $57 \times 50 \times 24 = 68400$. Regarding their ability to display meaningful information about elections, political horizon charts can be hundreds or thousands of times more efficient than choropleth maps.

For a second example, we turn to two US Senate runoff elections in Georgia, with a final date of in-person voting set for January 5, 2021. The regular senatorial election pitted Democratic challenger John Ossoff against Republican incumbent David Perdue. And the special senatorial runoff election pitted Democratic challenger Raphael Warnock against Republican incumbent Kelly Loeffler.

We collected hourly prediction market prices from November 11, 2020 through January 4, 2021. Margin-of-victory estimates from these data showed Republican advantages through most of the observation period. But we had the previous experience of the 2020 presidential election in which the Biden-Harris Democratic ticket narrowly defeated the Trump-Pence Republican ticket, despite prediction markets favoring the Republicans. This is reflected in the Georgia panel of figure 2.

Fig. 2 Political Horizon Chart for the 2020 Presidential Election



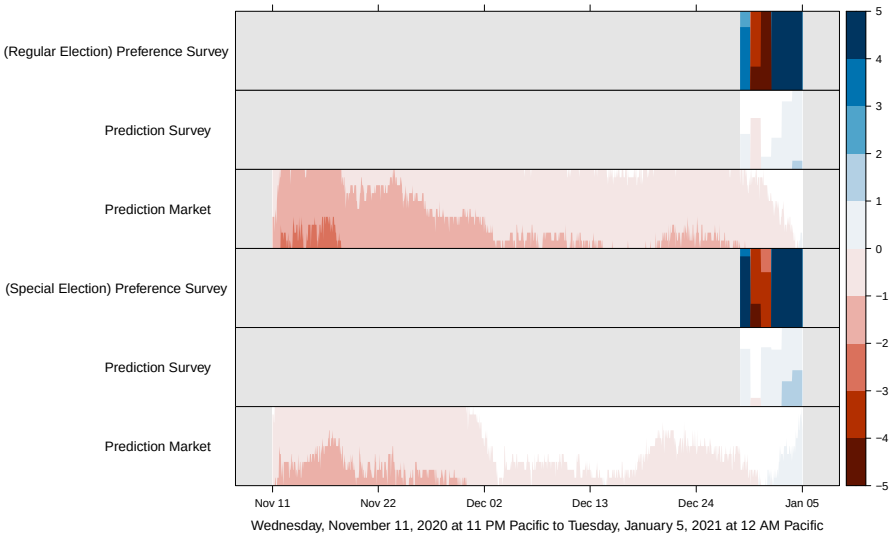
Note. Prediction market data provided by PredictIt.

In the last couple weeks leading up to the Georgia senatorial runoff elections, there had been considerable uncertainty about election outcomes. Across most of the observation period for prediction markets, margin-of-victory estimates suggested Republican wins. But opinion polls of likely voters in Georgia suggested close elections. With the help of three sponsoring firms, we implemented preference and prediction surveys for six days prior to final day of in-person voting, December 30, 2020 through January 4, 2021.

Figure 3 shows a six-panel political horizon chart for the Georgia runoff elections, three panels for the regular election and three for the special election. Panels share a common margin-of-victory scale from -5 for a strong Republican advantage to +5 for a strong Democratic advantage. The chart shows estimated margins of victory from preference surveys (political polls), prediction surveys, and prediction markets.

Many traditional political polls favor Democrats, while prediction markets often favor Republicans. Prediction surveys are able to "split the difference," providing the most accurate forecast of election outcomes. Results from the Georgia runoff elections showed that prediction surveys provided more accurate election forecasts than both traditional political polls and prediction markets.

Fig. 3 Political Horizon Chart for Georgia Runoff Elections



Note. Survey development and deployment in partnership with Isometric Solutions, Panel Consulting Group, and panel provider Prodege, LLC. Prediction market data provided by PredictIt.

5 Conclusions

The political horizon chart is an excellent tool for exploratory data analysis. Even to an untrained eye, it is easy to identify trends across time both within and between panels of a political horizon chart. This is a key advantage of this method of presenting multiple time series data.

Using political horizon charts, we can track the progress of political campaigns. We can display extensive information in a small space. Political horizon charts are hundreds or thousands of times more efficient than maps. They provide a convenient visualization of political time series, simplifying the difficult task of detecting trends and patterns across multiple time series. The benefits of political horizon charts by far outweigh the learning-curve costs associated with interpreting panel colors.

6 About the Author

Thomas W. Miller is faculty director of the data science program at Northwestern University, author of six books about data science, and owner of Research Publishers LLC.

References

Cleveland, William S. 1993. *Visualizing Data*. Summit, N.J.: Hobart Press.

Few, Stephen. 2008. "Time on the Horizon" Visual Business Intelligence Newsletter. Retrieved from the World Wide Web on December 21, 2020 from https://www.perceptualedge.com/articles/visual_business_intelligence/time_on_the_horizon.pdf.

Heer, Jeffrey, Nicholas Kong, and Maneesh Agrawala. 2009. "Sizing the Horizon: The Effects of Chart Size and Layering on the Graphical Perception of Time Series Visualizations". *Proceedings of the SIGCHI Conference*. Retrieved from the World Wide Web on December 21, 2020 from https://hci.stanford.edu/publications/2009/heer-horizon-chi09.pdf.

Miller, Thomas W. 2021. "Predicting the 2020 Presidential Election". *Data Science Quarterly* 1 (1).

Sarkar, Deepayan. 2008. *Multivariate Data Visualization with R*. New York: Springer.