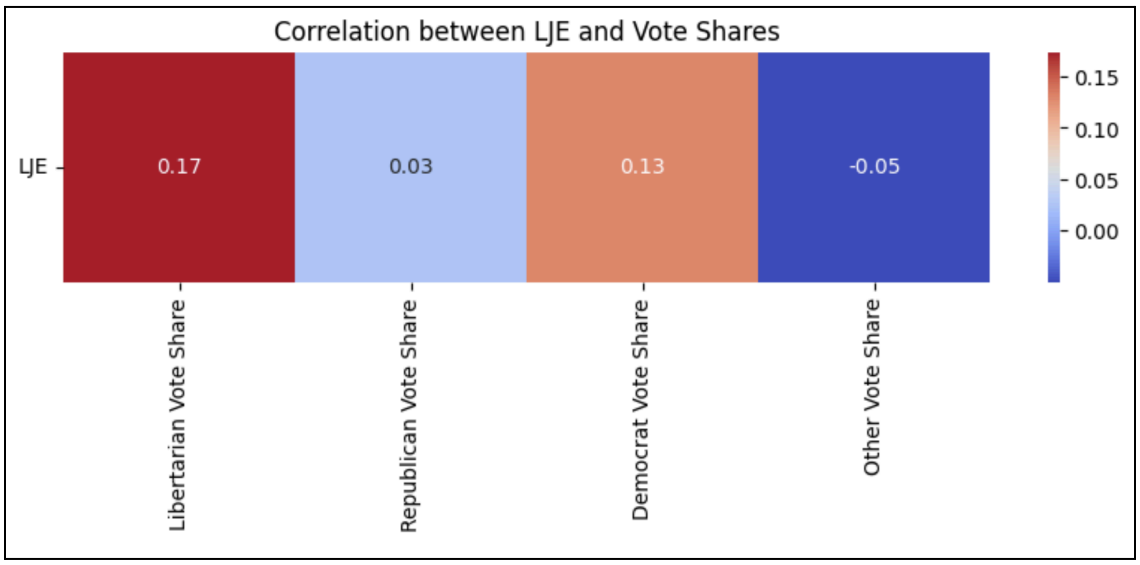
Abraham Moore Odell
August 2025



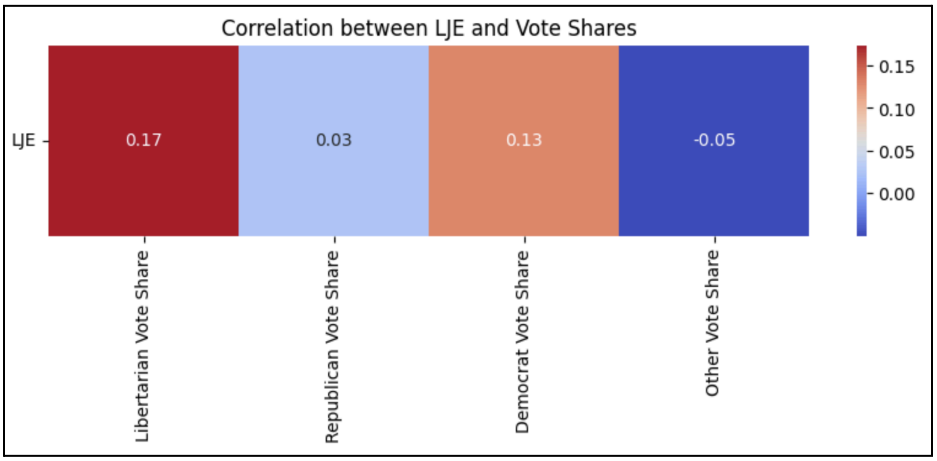
An independent analysis based on Rebuild Local News' *Local Journalism Index* exploring the relationship between journalism employment and political outcomes.

Whether it was a clipping from a school band concert or clippings of my tennis match results on the fridge, growing up in Vermont local news was an important part of my childhood and my community, and I'm very proud to see Vermont has the highest LJE (local journalist equivalent, a metric of the number of reporters per 100k residents) of any state. This connection to local news drove me to read the Rebuild Local News "Local Journalist Index" report and methodology, and search for my own trends within the data. The 2025 LJE report can be found here.

I took the future work section of the report's methodology as a challenge: is there a link between LJE and political outcomes? What might it reveal? At its core, the question I tried to answer was simple: *Does the number of local journalists in a community influence how people vote?* After a thorough analysis and a couple of days writing code, my conclusion was that no direct connection can be found between LJE and political outcomes.

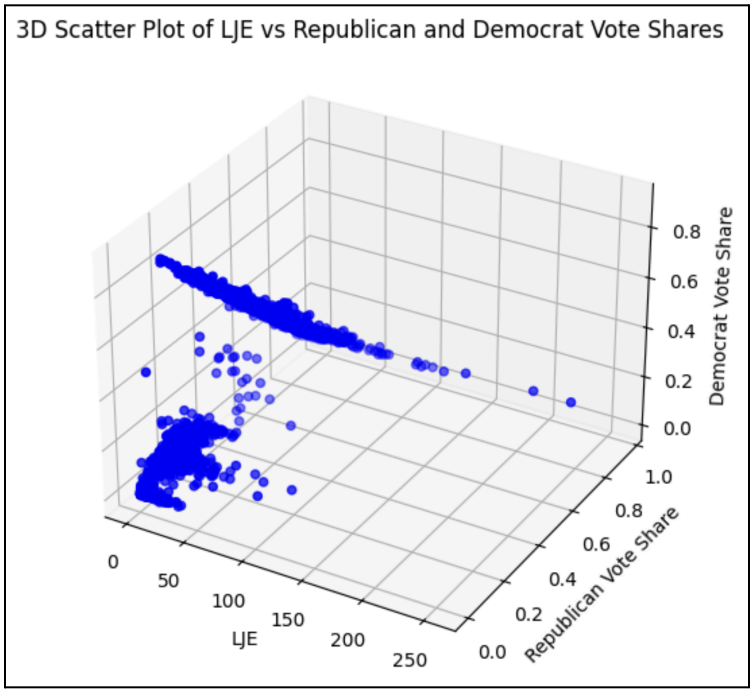
Before I give a play-by-play of my process, here is my publicly available code, including the data sources that I utilized: https://github.com/abemo/news. Feel free to download, edit, and play around with it. Additionally, I've included lots of charts and graphs in this report. I think each one tells an interesting story about journalism in our country.

I began by bringing in the 2025 LJE dataset, and paired it with population data from 1900–2020, and county-level general election results from 2000–2024. I then compared these metrics, at both state and county levels, to understand whether LJE influenced vote share and overall civic engagement. My first experiment was to examine the correlation between county-level LJE and vote share (Libertarian, Republican, Democratic, or Other votes). To accomplish this, I wrote code to calculate the percentage votes for each party. I then used the programmatic data analysis tool "pandas" to calculate the correlation between LJE and vote share for each county and generate a graphic.



A correlation score of +1 or -1 shows a strong correlation, scores close to zero (either positive or negative) mean there is likely no correlation. In this case, a +1 in any of the fields would mean an increase in LJE correlates to an increase in that party's vote share, and a -1 would represent a decrease. However, my graphic shows that there is no solid correlation between LJE and vote share, as each party's LJE correlation score is well below an absolute value of 0.25 which is the threshold for even a weak correlation.

Moving from simple correlation to 3D visualizations, vote patterns remain scattered and no clear relationship to LJE emerged. By programmatically plotting LJE, Republican votes, and Democrat votes in 3D space, we can again see that there is no correlation between LJE and vote share. Instead, these clusters stretch across LJE and simply show how as one party's vote share increases, the other necessarily decreases. Most counties cluster around lower LJE values (under 25), and both Republican and Democrat-leaning areas exist across this range, suggesting that partisan preference is not strongly tied to the density of local journalists. A few outlier counties with very high LJE values do appear, but they do not exhibit a consistent political pattern. If LJE were a strong predictor of vote share, we would expect to see a consistent diagonal line showing that as LJE changed so did vote share. However, the clustered nature of the data points reinforces the idea that local journalism employment does not serve as a strong predictor of vote share for either major party.



Despite the lack of mathematical correlation I decided to "check my work" by training a machine learning model which may be able to pick up on a more subtle pattern of correlation. I therefore trained and ran a linear regression machine learning prediction model to examine if LJE could be used to predict the vote share of a county. Predictably (due to the low correlation), my results show that LJE alone is insufficient to predict vote share.

```
Libertarian Vote Share: RMSE = 0.0000, R² = 0.0255
Republican Vote Share: RMSE = 0.0726, R² = 0.0016
Democrat Vote Share: RMSE = 0.0341, R² = 0.0142
Other Vote Share: RMSE = 0.0161, R² = 0.0034
```

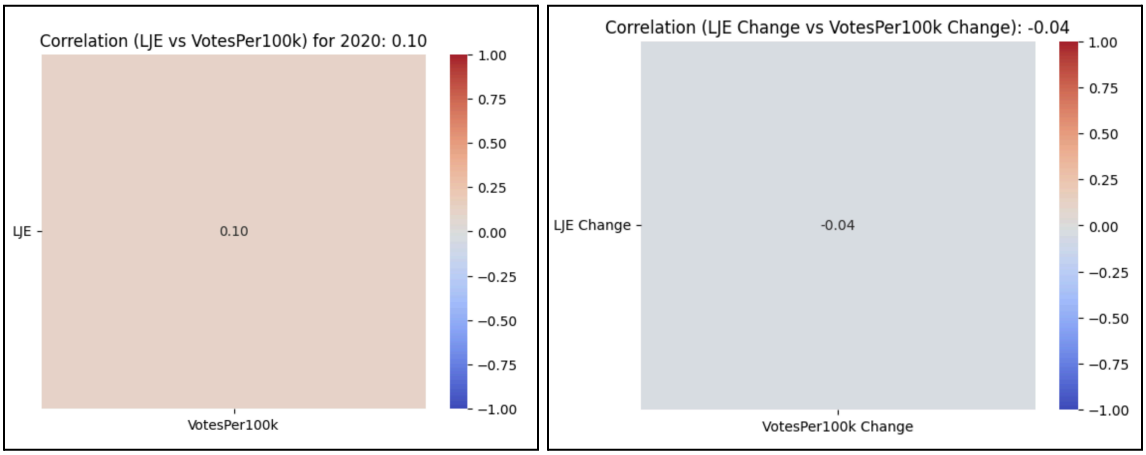
RMSE or Root Mean Square Error is how close the machine learning algorithm was to predicting the real values. The lower the RMSE score, the closer the model was to the correct values.

R^2 is a statistical measurement that tells us how much one variable is able to tell us about another variable. R^2 is typically between 0 and 1, with higher scores representing more significant relationships between the variables examined.

Although the model made predictions that were close to the actual results (low RMSE), it didn't explain much about *why* those results happened (due to the low R^2). For example, in the case of Libertarian vote share, only about 2.55% of the differences between counties could be explained

by the number of local journalists. This suggests that while the predictions weren't far off, LJE by itself isn't a strong factor in determining how people vote. I then tried flipping the analysis, and trained a new model using vote share to predict LJE, and that approach didn't work well either. But what does that tell us? The lack of correlation may be a good thing. Perhaps it simply means that our nation's local journalists do a great job of unbiased, politically neutral reporting. It also tells us a county's political leaning does not doom it to poor local journalism.

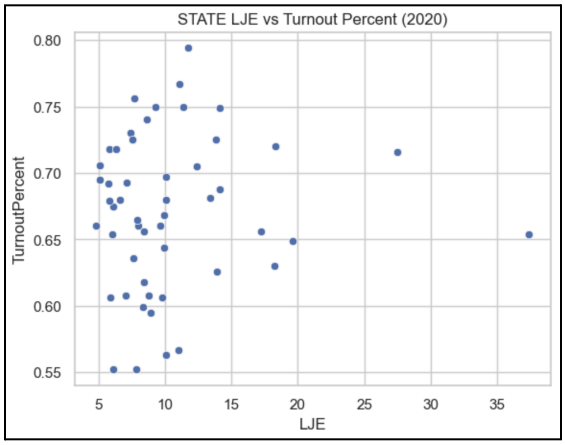
While that helps to answer the question posed in the report, I still wanted to dig deeper. I asked the question, does the LJE of a county affect total voter turnout? To answer this, I created a new voter-turn out score for each county in the US. To create this metric, I calculated the total votes per 100k residents in each county, and compared this number to the LJE score of each county. In the figure below, titled "Correlation (LJE vs VotesPer100k) for 2020" I looked across all US counties using the 2025 LJE index and 2020 population data. Again, I found no correlation between the counties with more journalists and the counties that received the most votes.

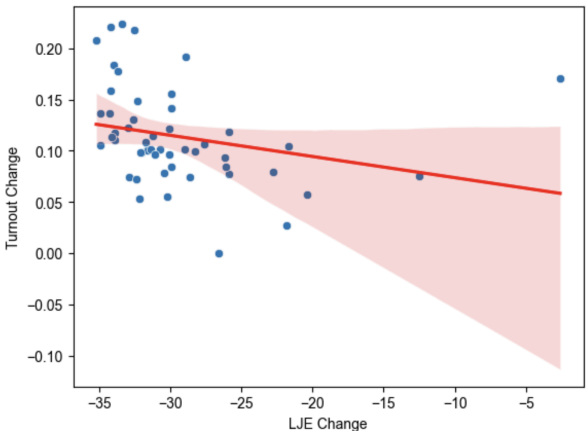


In an attempt to find a more clear understanding of the relationship between LJE and voter-turnout I wrote a computer program to look at the change in voters per 100k residents between the 2000 to 2020 general elections in relation to the change in LJE from 2000 to 2025. I used 40 as a general estimate for LJE in the year 2000, and, because we don't have LJE statistics for 2020 I used the 2025 numbers as a close estimate.

While I found that the metrics of LJE and the total-voter-turnout-per-100k-residents are uncorrelated (increasing one doesn't necessarily increase the other), this prompts some new and interesting research questions. For example: I examined presidential general election voter turnout, but with so much national news focus on elections, does local reporting have an impact on local elections even though it seems to not affect national elections?

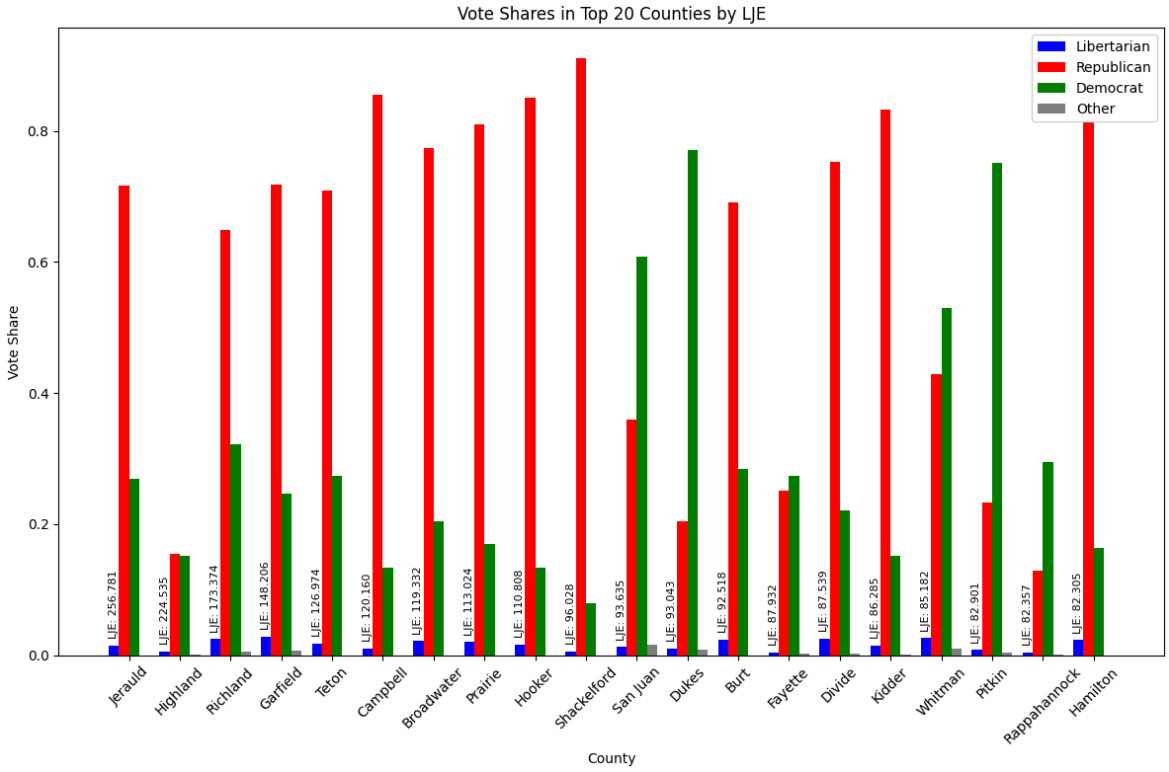
After analyzing the county level data, I took a look at the state level data in the 2025 Journalism Index report. Once again, I found that LJE is not correlated to voter turnout at a state wide level.

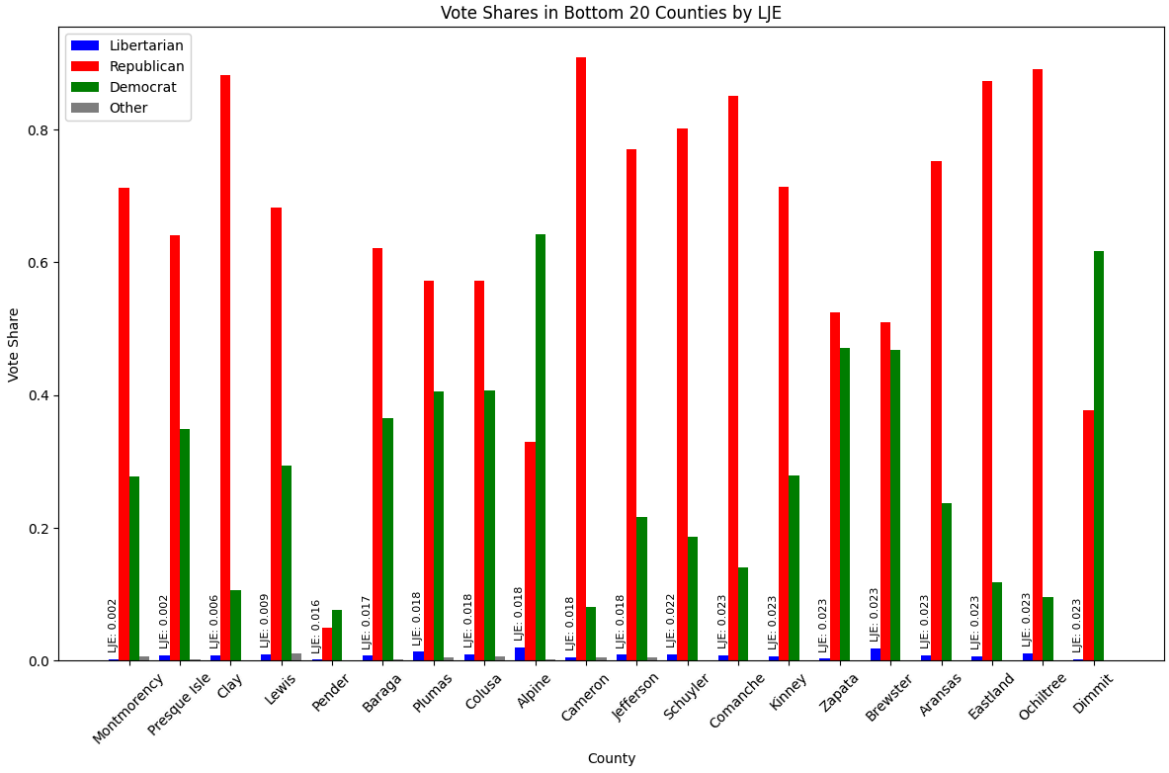




Minnesota is a clear stand out, where almost 80% of eligible voters turned out for the 2020 general election. Minnesota is also 12th on the 2025 Local Journalism ranking with a LJE of 11.7. You can easily see the lack of correlation by how spread the data points are on both graphs. If there was a strong relationship between state level LJE and voter-turnout the dots would form a more consistent shape or pattern.

To answer the question asked in the "Local Journalist Index 2025" report "what is the political breakdown of these journalist shortage areas?," I examined the top twenty and bottom twenty counties by LJE. I then used code to explore and graph the political leanings of these counties. In the charts below, each county's vote share is written alongside the corresponding bar lines.

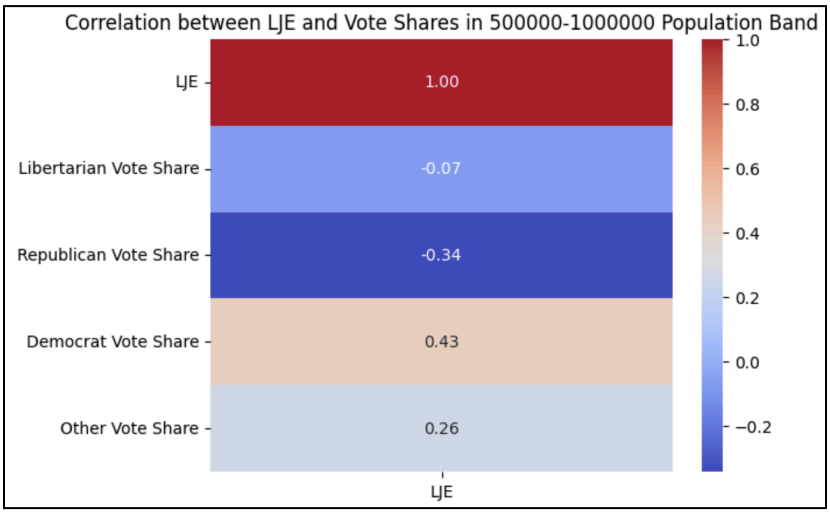




It’s important to note that these counties are the extreme outliers. While speaking to the journalists in these counties (such as Gerri Peterson in Hooker County Nebraska) would give

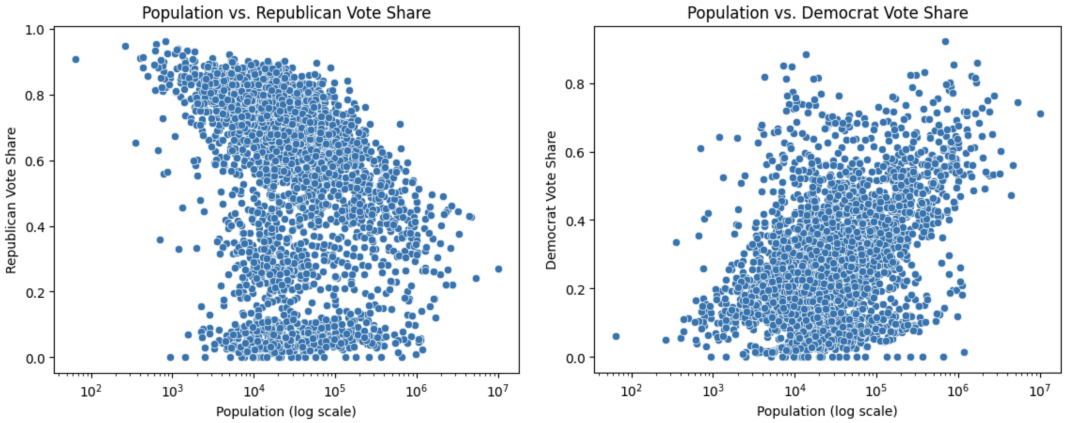
great qualitative stories, looking at only 40 of over 3000 counties could lead to misleading statistical results. For example, we can see that in both the highest and lowest LJE counties, Republican vote share is vastly over-represented. This over-representation is confounded by the small population size common amongst these counties.

Instead, by examining the relationship between LJE and vote share in the same county-size bands used in the methodology (grouping counties into: below 5k, 5k-25k, 25k-500k, 500k, 1000k, 1000k+) we see something very interesting. While LJE does not correlate with turnout within the bands, in counties between 500,000 and 1,000,000 residents, LJE is moderately correlated with vote share—particularly Republican and Democrat. This is interesting as it is observed within a specific size band and therefore cannot simply be explained by larger counties (likely containing left leaning cities) skewing votes to the Democrat party.



I understand that this result is likely not what Rebuild Local News would like to see or advertise to many of the states it seeks to serve with strong local-journalism focused legislation. However, this is a very specific case and an exception that helps to prove the rule. This correlation score of -0.34 and +0.43 is seen only in counties sized 500,000 to 1,000,000 residents. Additionally, even the highest correlation score seen of |0.43| (vertical lines used to denote absolute value, i.e. ±0.43) is still considered only a moderate correlation meaning other factors such as other public voter policies or county/city culture likely have a much larger impact on vote-share.

Despite already accounting for population size affecting vote share, I was still wary of the covariance effect of LJE and population on vote share. In other words, despite already grouping counties by population and only comparing similarly sized counties, I wanted to take a closer look at the relationship between county size and vote-share. 500,000 to 1,000,000 residents is a wide range, and as seen in the graphs below, population has a *major* effect on the voting outcome of a county.



To better understand if population changes skewed the results showing a relationship between LJE and vote-share in these mid-sized counties I trained a multiple linear regression (MLR) machine learning (ML) model to examine the effects of population and LJE on vote share in this specific population band. A MLR ML model uses multiple complex machine learning algorithms to compare the effects of multiple variables on a single outcome, for example determining if location or square-footage has a larger impact on home prices. This machine learning model found that while population seemed to have a very strong relationship with vote share, LJE also had a minor effect on vote share. In the 2020 election, and in specifically mid-sized counties, an increase of 1 in LJE correlates to a 0.16% increase in Democrat vote share. Such a small increase in counties of 500,000 to 1,000,000 residents would require about 45 more local journalists in one of these counties to increase Democratic vote share by 1 only percent, further supporting the idea that investment in local journalism is non-partisan.

These preliminary figures tell an interesting and motivational story. I had initially assumed that local journalists would be overwhelmingly liberal, and counties with the most local journalists would sway the same way. However, while billionaire owned national news sources have become increasingly partisan, and the effect of these outlets on national elections is evident, my analysis shows that while LJE may not directly influence vote share or turnout, its role in the broader civic landscape remains vital. The absence of strong political correlations could be a sign of something positive: that local journalism remains a relatively neutral, community-driven force. I see this as a foundation for deeper work. There is still much to explore about how LJE interacts with local civic health, institutional trust, and participation beyond the ballot box.