May-July 2024 Northern Michigan Search Interest Forecast

Author: Dan Shaffer

================================================================================

Below are the search interest forecasts for the combined Northern Lower and combined Upper Pensinsula places for May, June, and July 2024. Note that the possible range for historical search interest is normalized to a maximum of 100 and a minimum of 0, but forecasts outside this range are permissible as these values are forecasted to be outside the historical range. In fact, both forecasts are showing historical high values for July.

Also shown in grey on these charts is the previous forecast from April. For both peninsulas, the forecasts are relatively close. However, comparing the April forecast for April and the April actual, we see that the actuals for both peninsulas are lower than the April forecast (Lower: 43 vs 36, Upper: 48 vs 40). However, some of the difference between forecasts is due to differences in the Google trends historical dataset when pulled from month to month.

Also provided are barcharts comparing the average search interest for the forecast months (May,June, July) to the same months in previous years. The forecast for both peninsulas is lower than the last two years. So while the July forecast is a historical maximum for both peninsulas, the average for May-July is lower. This is partly due to mixed weather forecasts. For May/June, accuweather shows lower average max temperatures but higher average low temperatures and weather.gov shows higher average of average temperatures. However, for July all temperature metrics are higher than average.
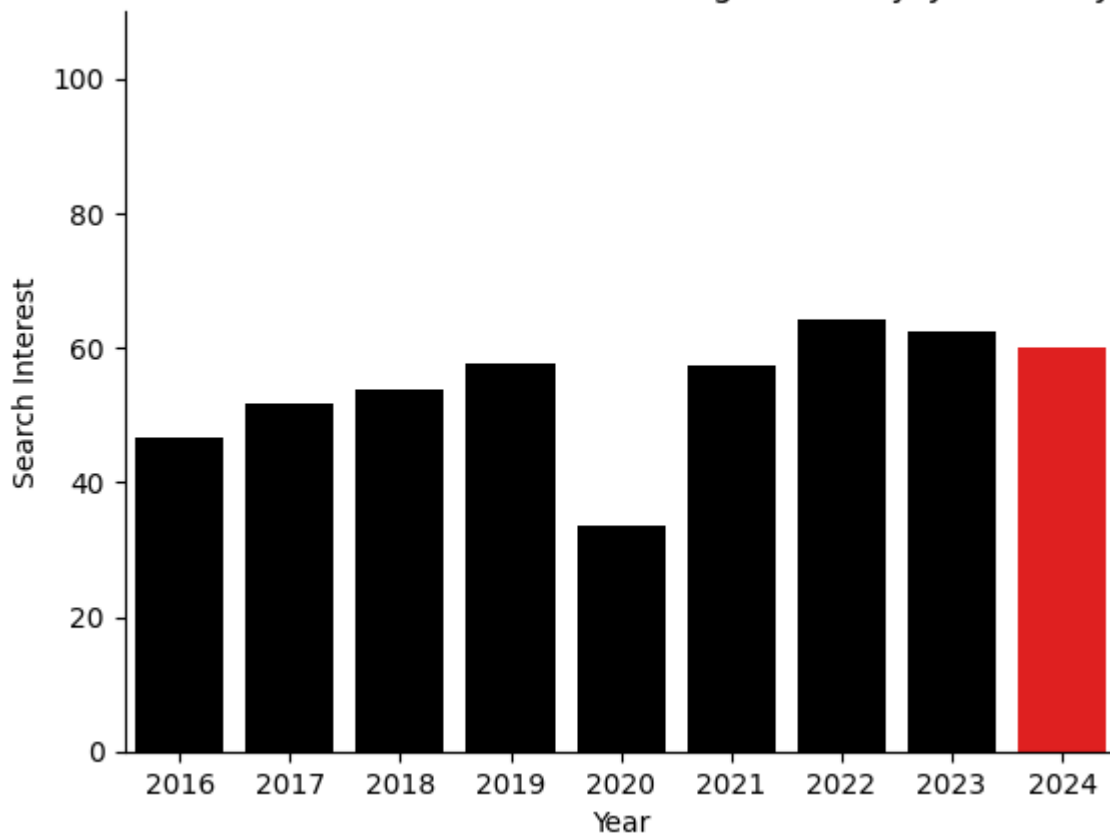
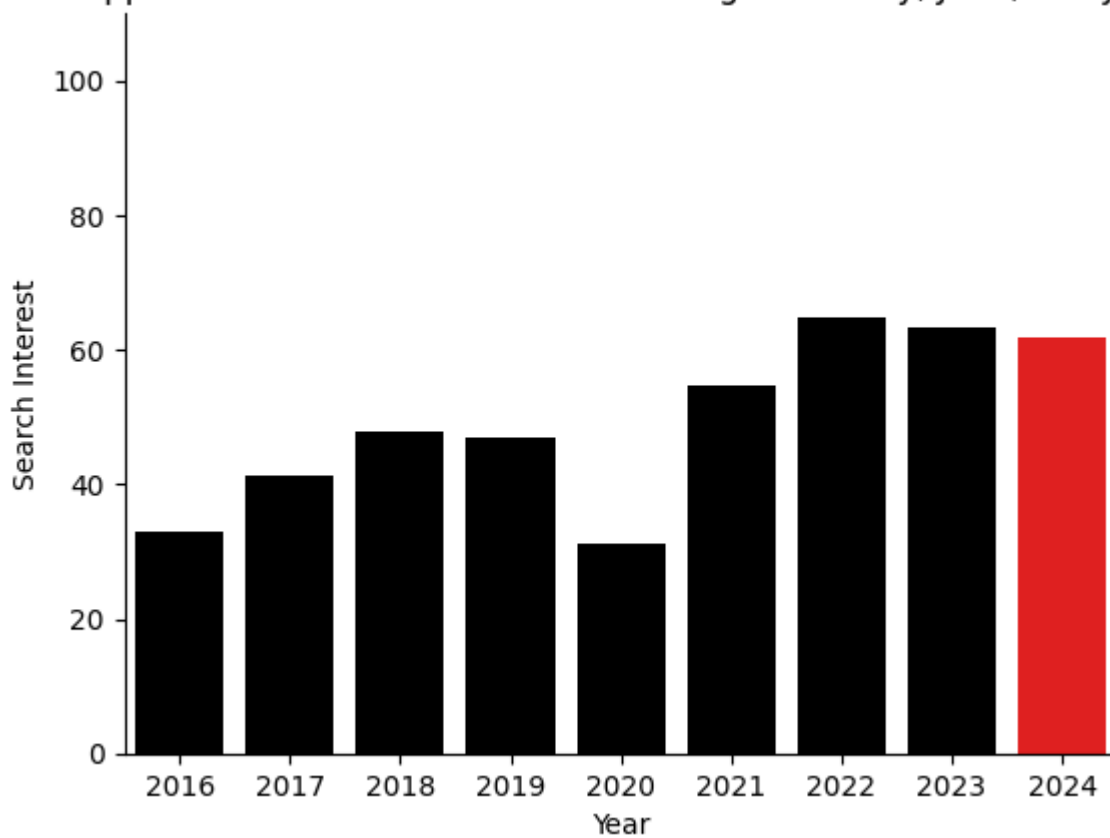Northern Lower Places Google Search Interest Forecast

Series
— Actuals
— Forecasted
— April Forecast
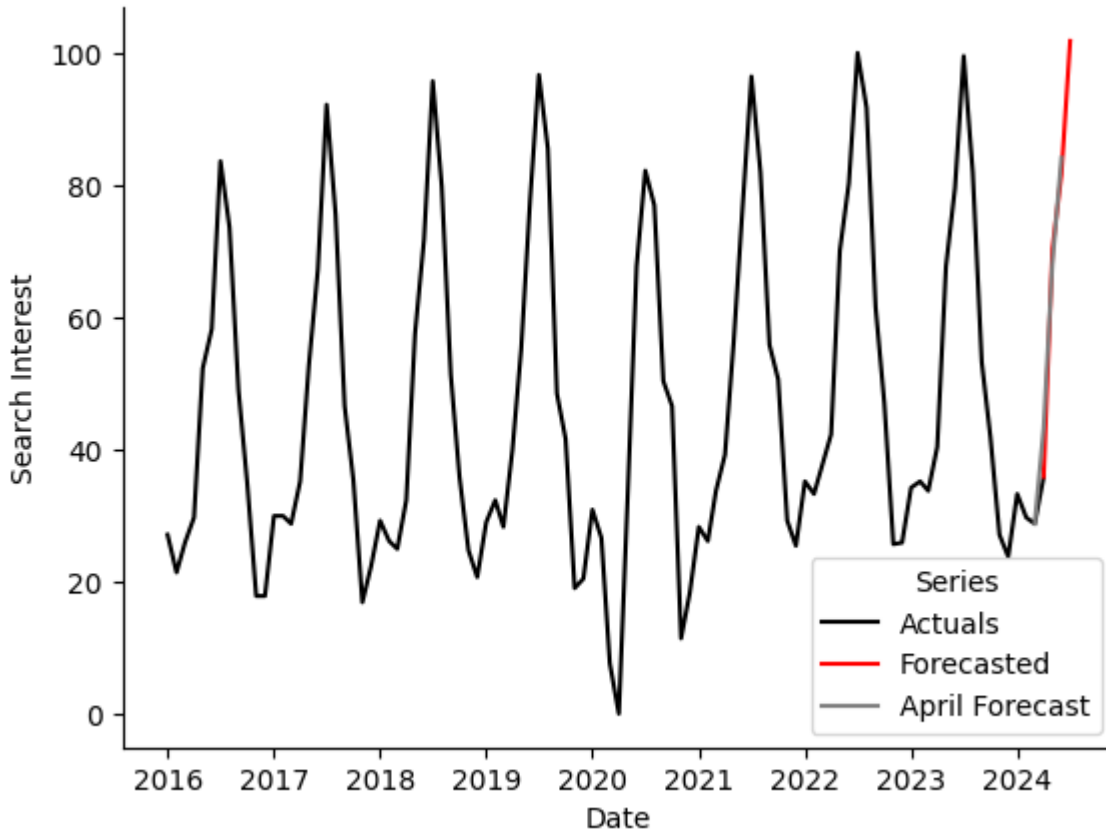


Upper Pensinsula Places Google Search Interest Forecast

Series
— Actuals
— Forecasted
— April Forecast
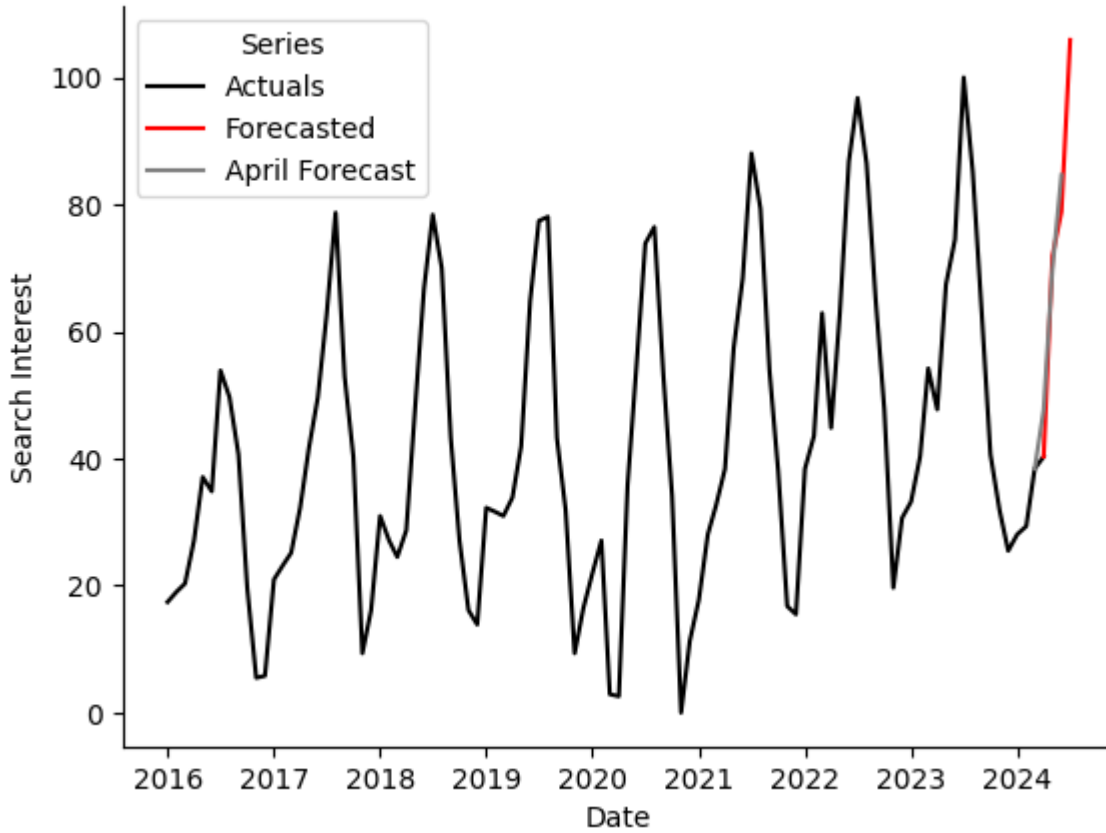
Northern Lower Search Interest Averaged for May, June, and July


Upper Peninsula Search Interest Averaged for May, June, and July

The following table shows the top five places that are forecasted to have the highest search interest compared to the same time period in 2023. From the original research used in this project, we know that many places in the western Lower Peninsula and some places in the Upper Peninsula are particularly sensitive to weather and/or gasoline prices. These places from the western Lower and Upper Peninsula show the largest expected changes. The exception might be free soil, which might show people literally looking for free soil for gardening in the Spring.

|   | Place | Peninsula | Difference |
|---|---|---|---|
| 0 | Atlantic Mine | Upper | 31.9 |
| 1 | Bruce Crossing | Upper | 9.5 |
| 2 | Free Soil | Lower | 9.3 |
| 3 | Tustin | Lower | 8.1 |
| 4 | Parkdale | Lower | 6.6 |

Three things impact the value of the search interest forecasts for each place.

1. The previous year's monthly value for each individual place.
2. While not directly impacting the model forecast, seasonal (12 mo) differencing accounts for the fact that search interest is higher in some portions of the year than others (so previous year actual is impacted).
3. Model difference: based on forecasted weather and gas price changes, the model will predict 12 mo changes from the previous year for each individual place. These new levels are aggregated via regression to the peninsula level series.

The two KDE plots below plot the distributions of the individual place actual values for the months of April, May, and June for the previous three years. Note that the values for 2023 serve as the last actual values for 2024. For the Lower Peninsula, the 2023 values are lower than 2022 and similar to 2021. For the Upper Peninsula, the 2023 values are higher than 2021 and 2022.

Lower Peninsula Place Level Search Interest Averaged for May, June, and July
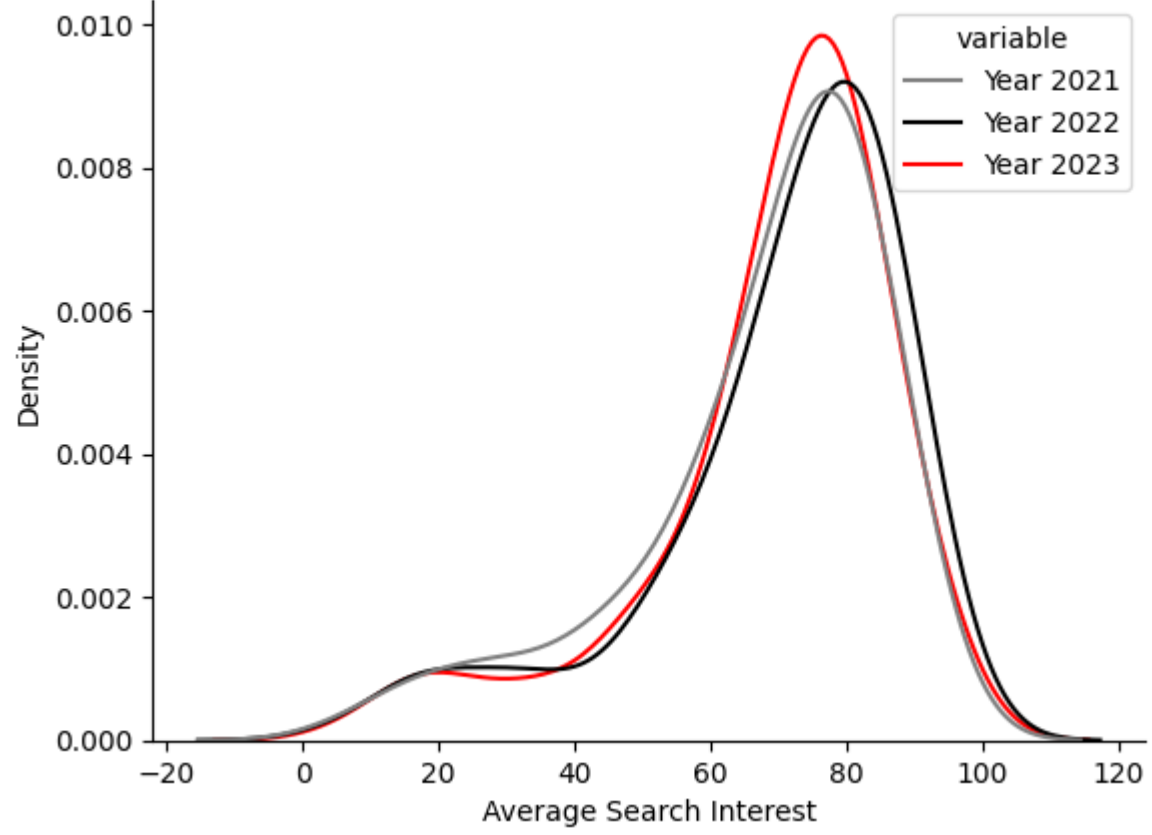
Upper Peninsula Place Level Search Interest Averaged for April, May, and June
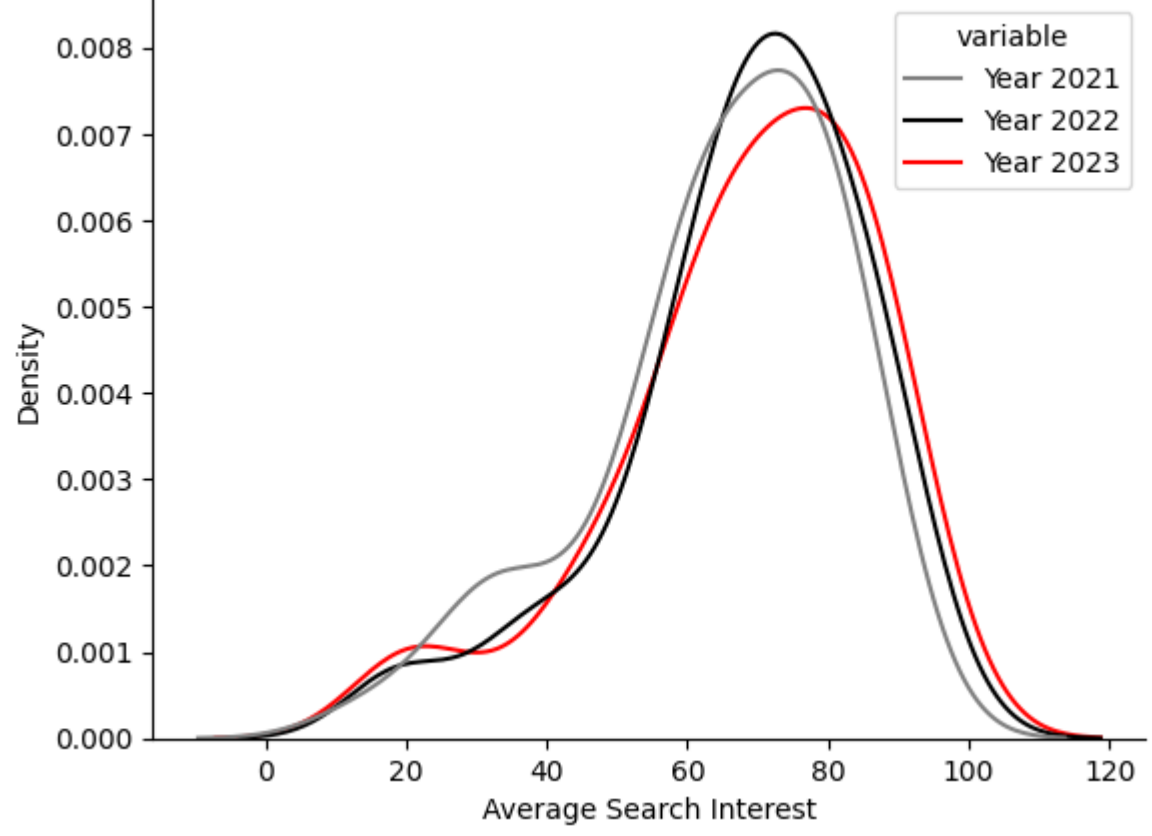
The following figure shows the aggregate average search interest for each calendar month. Both the Upper and Lower Pensinsulas have a seasonal peak in July/August with the Lower Peninsula peak solidly in July. The Upper Peninsula has higher values for January-March likely due to winter snow sports like snowmobiling. For the May-July forecast period, we should expect steadily increasing to maximum search interest.



Average Seasonal Values by Peninsula
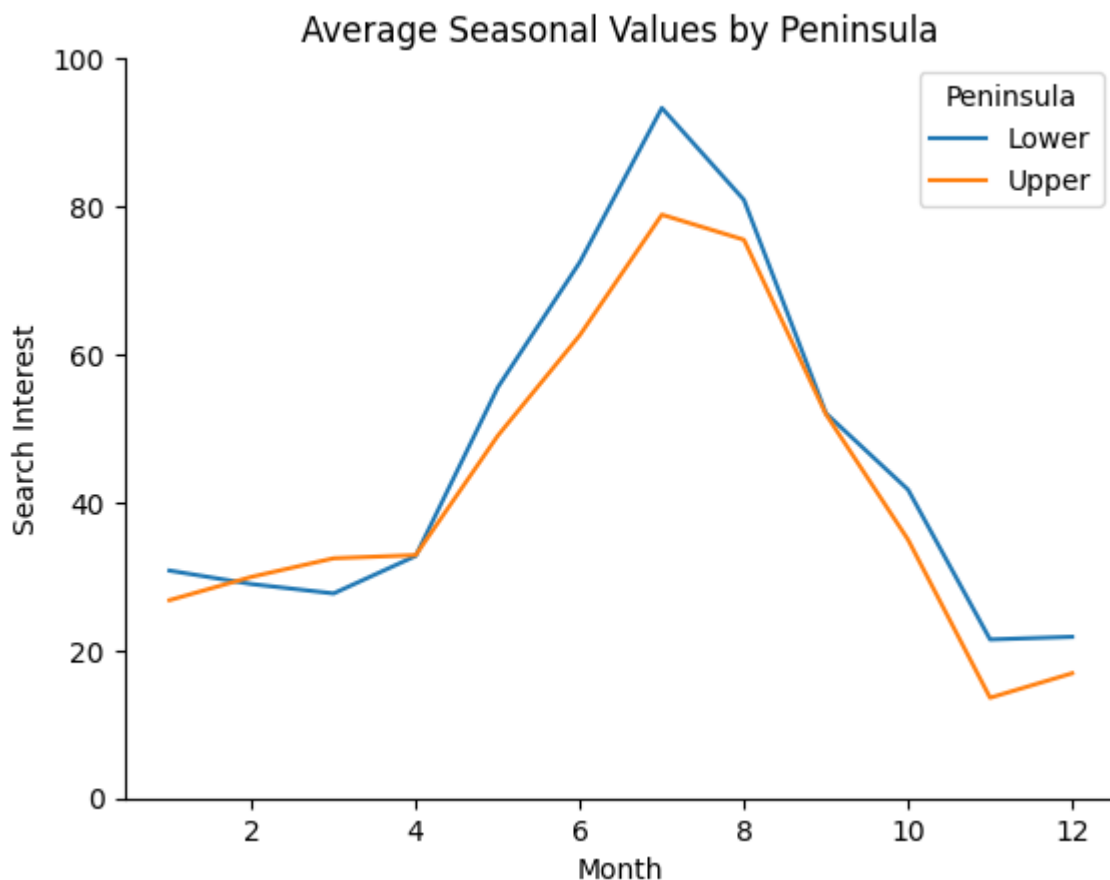
Finally, forecasts are determined by the forecasted place level change from the previous year based on weather and gasoline prices. For both peninsulas, the most likely forecasted change from the previous year is positive. However, there are significant negative values especially for the Lower Peninsula. (Note, however, that the final forecast numbers by peninsula are weighted by the size of the contribution of the place to total search interest.)

Model Forecasted Change by Peninsula
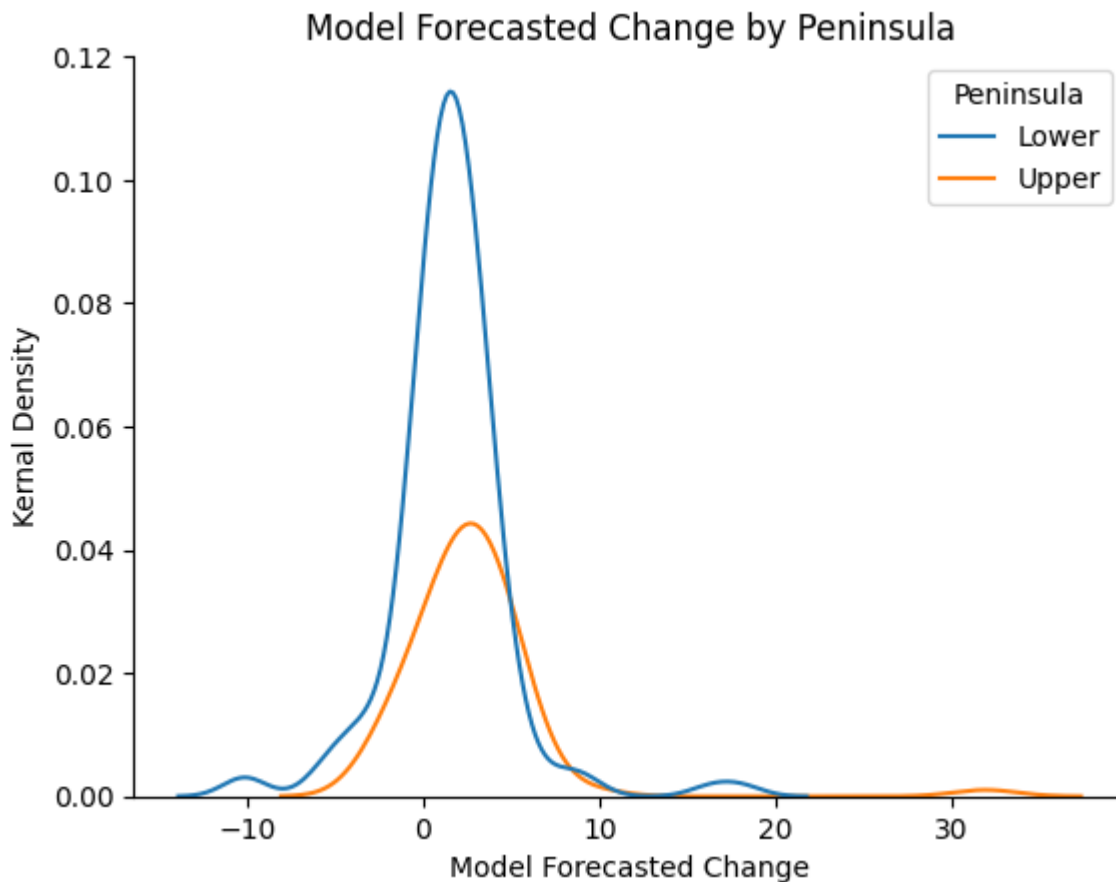
In addition to what "should" impact search interest forecasts, there is also variation between search values pulled from Google trends. As I described in the original research for this project, search interest results will vary depending on whether you pull data using the pytrends package or from the Googletrends website. Pertinent to this project, results will also vary if you pull data using pytrends at different points in time. The severity of this difference changes from month to month. However, the difference is usually most visible for the Northern Lower and during the low search interst winter months.

While the most obvious solution to this problem is retaining historical values and keeping them constant, this is more difficult for Google trends data which is internally normalized. Thus, it's much easier to repull the entire series each time. I have a good idea of what my long term solution to this issue is, and I've started to save back each monthly data pull to help make it happen.

Compares Current and Past Month's Northern Lower Data Pull
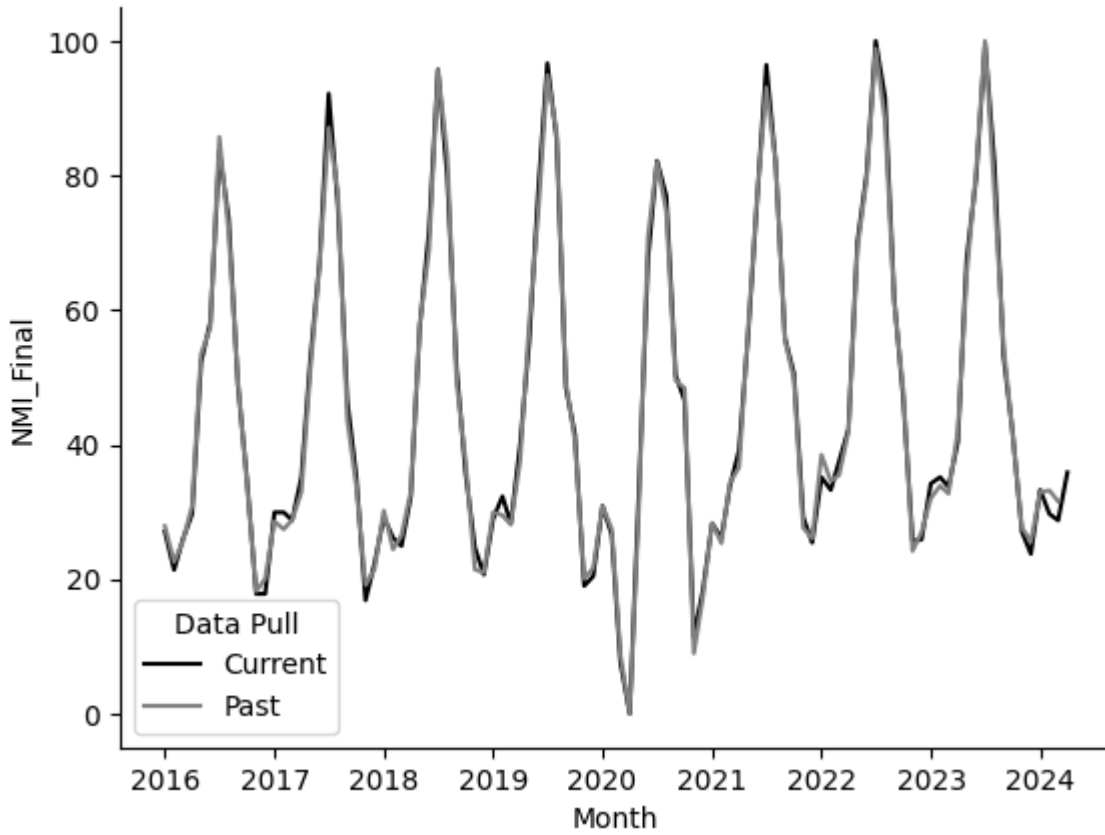


Compares Current and Past Month's Upper Peninsula Data Pull
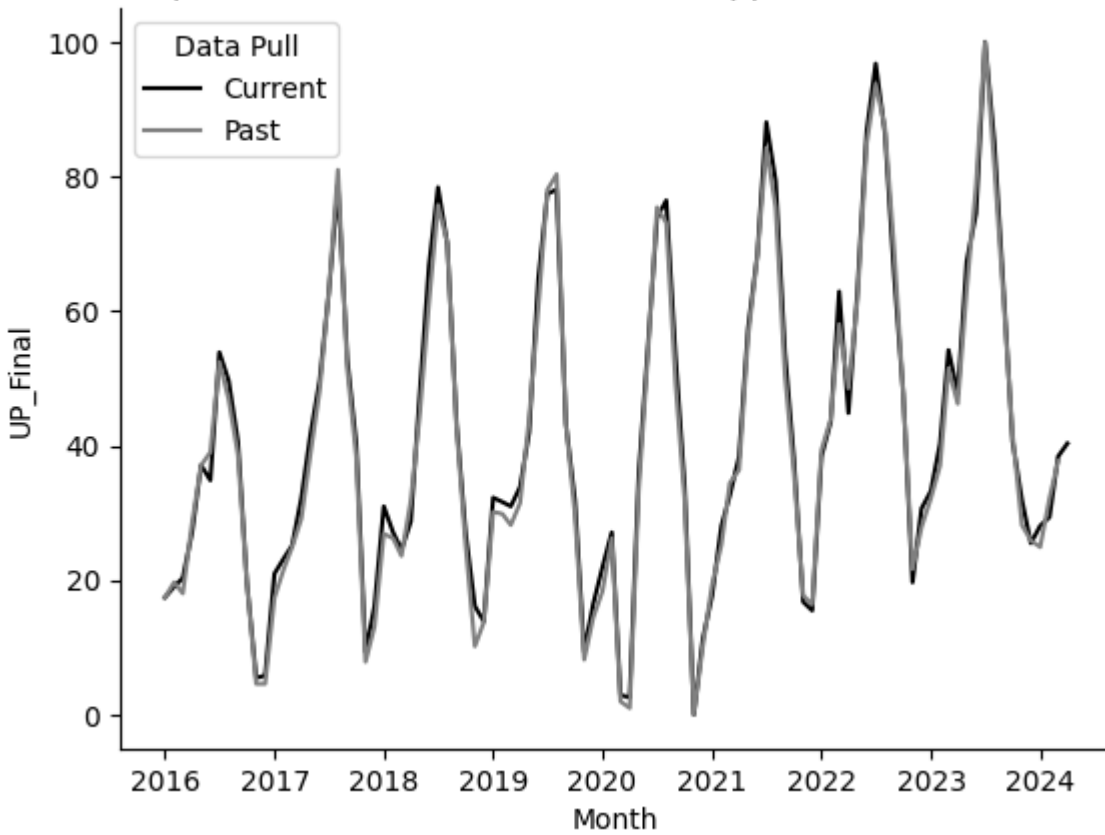
One last factor to consider is my source for weather forecasts, with a focus on temperature. All of my historical weather data used to train my models is from weather.gov. However, I source temperature forecasts from both weather.gov (average of averages) and accuweather (average maximums and average minimums). To some degree, it's difficult to compare these sources because average, maximum, and minimum temperatures are different statistics. Nonetheless, throughout this project, I've found that weather.gov usually forecasts above normal temperatures. To some extent I wonder if forecasts from a government agency are more likely to be high given politics surrounding issues like global warming.

Therefore, going forward I will roughly track the one month forward predictive performance of the weather.gov and accuweather forecasts. First, I will look at the previous month's average temperature and determine if it actually was above historical normals. Next, I will compare the previous month's forecasted accuweather average max and min temperatures to actual values.

For April, the normal average temperature is 42.2 degrees for Houghton Lake and 37.3 degrees for Sault St Marie. The actual average temperature values were Houghton Lake 45.9 and Sault St Marie 42.5. So actual temperatures were, indeed, above normals.

Based on the accuweather forecasts, we see that these were also quite accurate for April. The maximum temperature forecasts were almost exactly the same and the minimum temperature forecasts were slightly higher than actuals.



Comparison of Average Maximum Temperatures for Previous Month
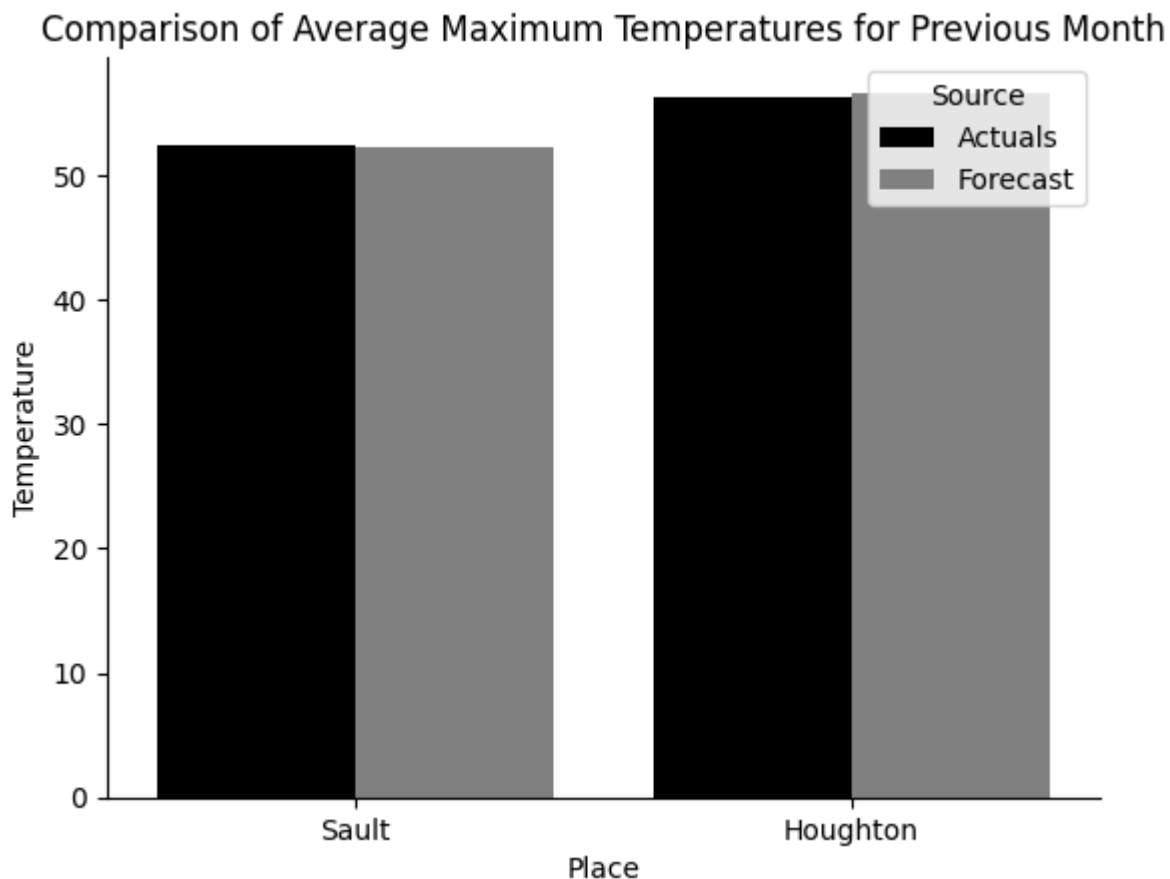
Comparison of Average Minimum Temperatures for Previous Month