

Michigan Outdoor Recreation Search Interest: First Installment

=====

There are many outdoor recreation activities popular in Michigan. Many people websearch these activities on Google and other search engines. The goal of this project is to analyze and forecast the trends of Google Search interest within Michigan for 10 different forms of outdoor recreation: atving, boating, camping, fishing, hiking, kayaking, rving, hunting, skiing, and snowmobiling. Search trends for these activities are likely related to the popularity and interest in these activities in general.

In these project, I use Google trends/pytrends to pull daily data from Jan 2021 to April 2024. For this first installment, I will discuss steps required to pull and clean data, show trends for all ten activities, and investigate correlation between them.

=====

Data

Data is pulled from Google trends using the pytrends python package for each search term from Michigan only. All data trends are normalized between 0 and 100 representing the Google search interest for the search term on that day relative to other days in the data pull. The highest period value for each search term is 100.

For each of the outdoor activities, both the base and present particpal form are used, i.e camp and camping. (Depending on the activity, one or the other is generally more common.) For each activity, both word forms are added together and renormalized between 0 and 100.

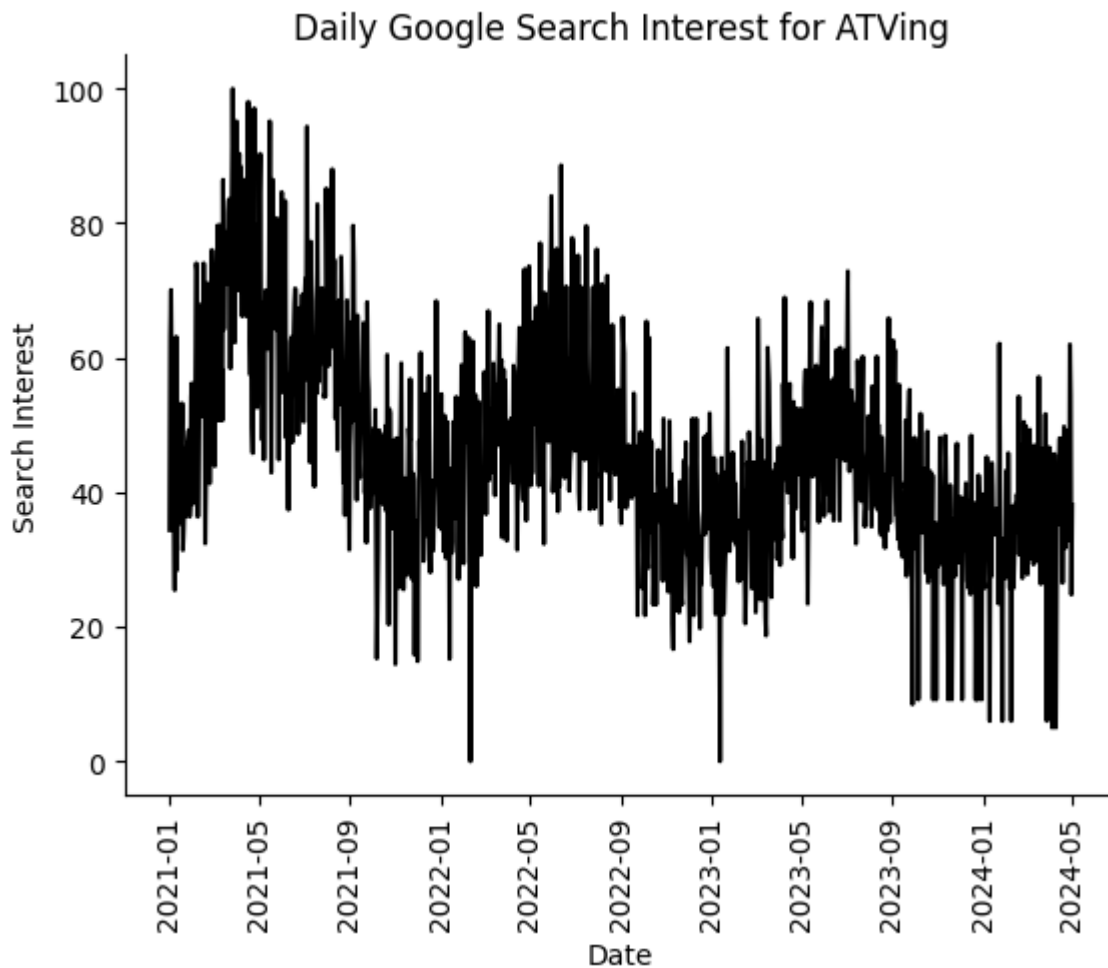
Due to limitations when pulling data with pytrends, daily data is only available when pulling data less than 6 months at a time. Since each one of these data pulls is internally normalized from 0 to 100, different data pulls are not directly comparable. Therefore, I pull many six month intervals that overlap by three months. This three month overlap allows me to combine two series and re-normalize them to the same scale. Whichever series is lower during the overlapping period is the series that has the highest value. Therefore, the average ratio between the two series during the overlapping period is used to adjust the other series downward. This puts both series on the same level. Subsequently, I rescale the data series by standard deviation. Once two time periods are combined, the combination is iteratively combined with the other periods.

Finally, there are many extremely high search interest values that I identify as outliers. These outliers aren't completely at random in that they appear more likely to occur during times of high search interest. However, these are single day values that are several times higher than adjacent days. I determined that these values are either measurement error or, in the least, values I'm unlikely to explain using forecasting and analysis. Therefore, I remove these by setting values after large one period changes to missing and linearly interpolating the values using the adjacent days' search interest values.

While the resulting dataset is patched together and imperfect, it is what we have and I believe the data will contain a lot of information, insights, and predictive value.

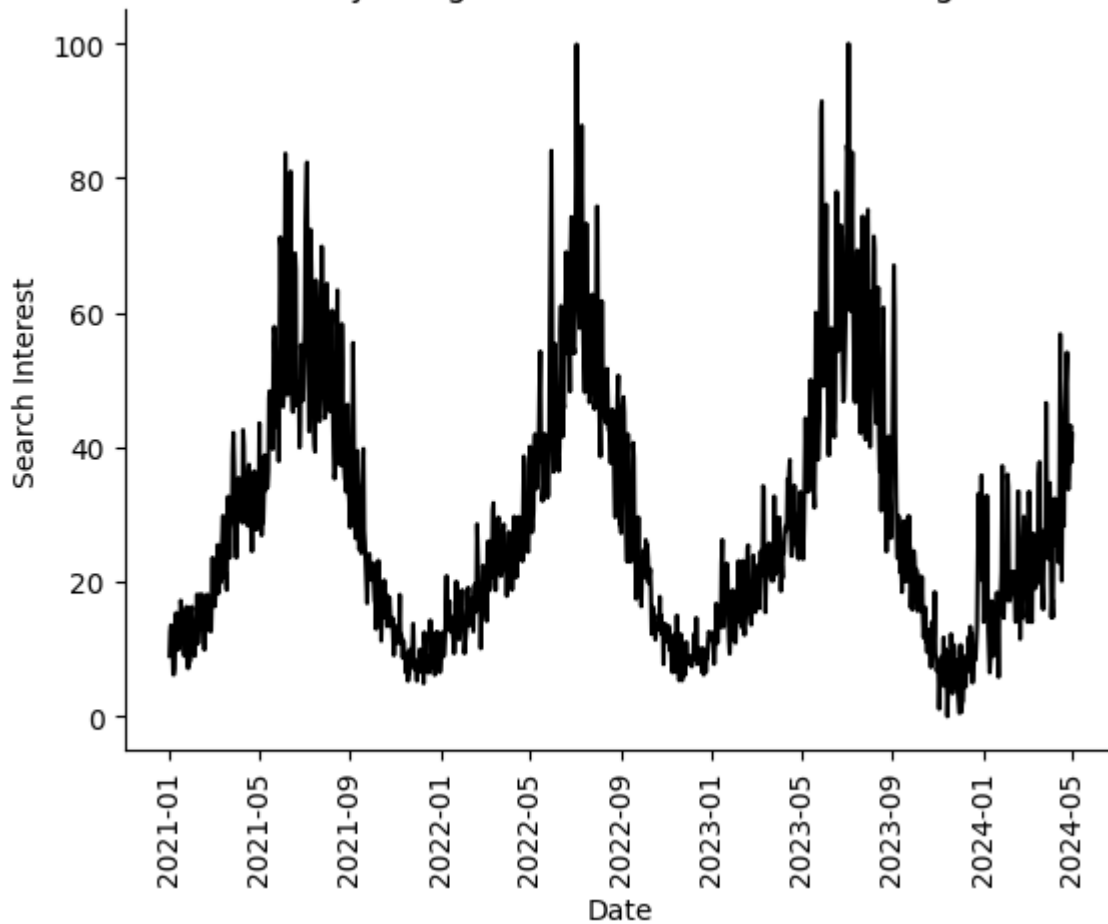
Search Interest Trend Plots

Now, I will plot the 10 resulting data series. First I plot the series for atving. As with most of these series, the plot for atving is cyclical or seasonal with most search interest in the summer months. In this case, search interest appears to be trending downward. To some extent, some forms of outdoor recreation peaked during or after the COVID-19 pandemic. This was especially the case for forms of outdoor recreation that involved buying expensive items due to the influx of government money during the same time period. ATVs were in short supply during 2020.



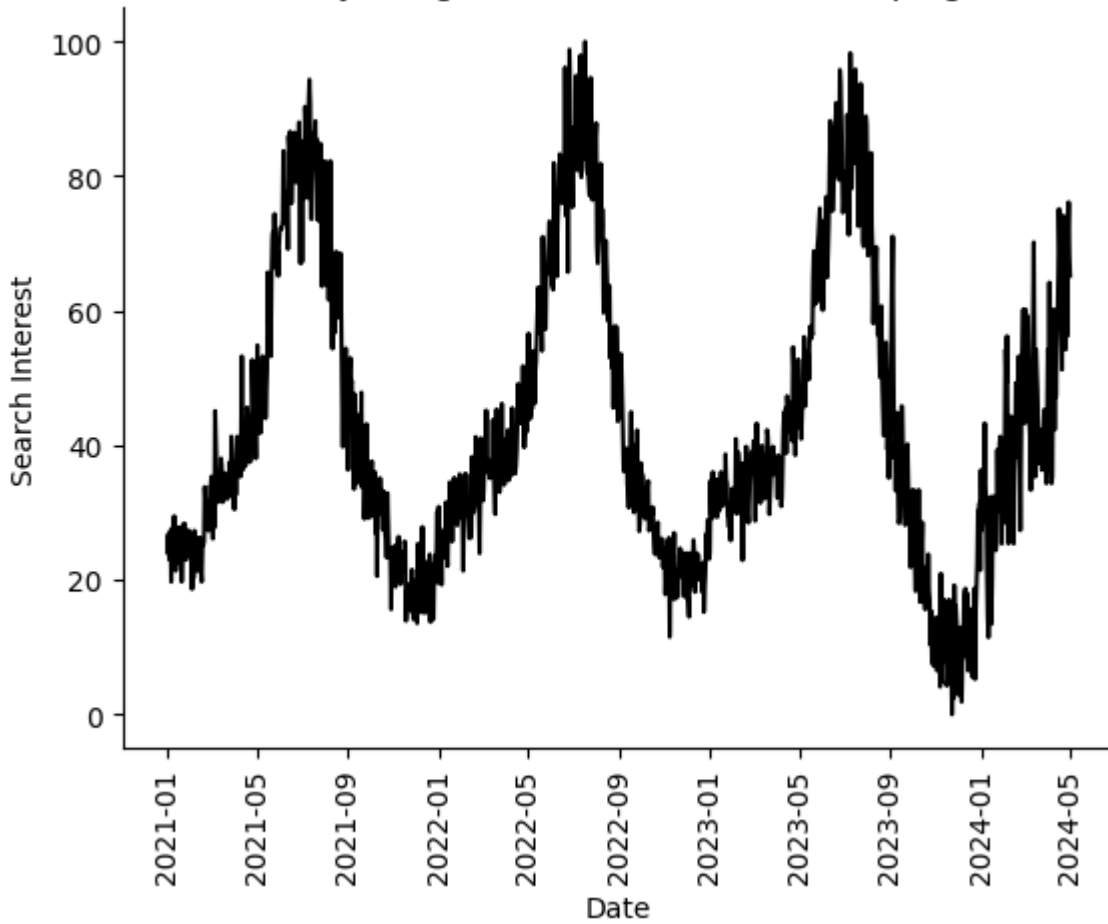
Interest for boating is even more peaked during the summer and is not trending downward unlike ATVing.

Daily Google Search Interest for Boating



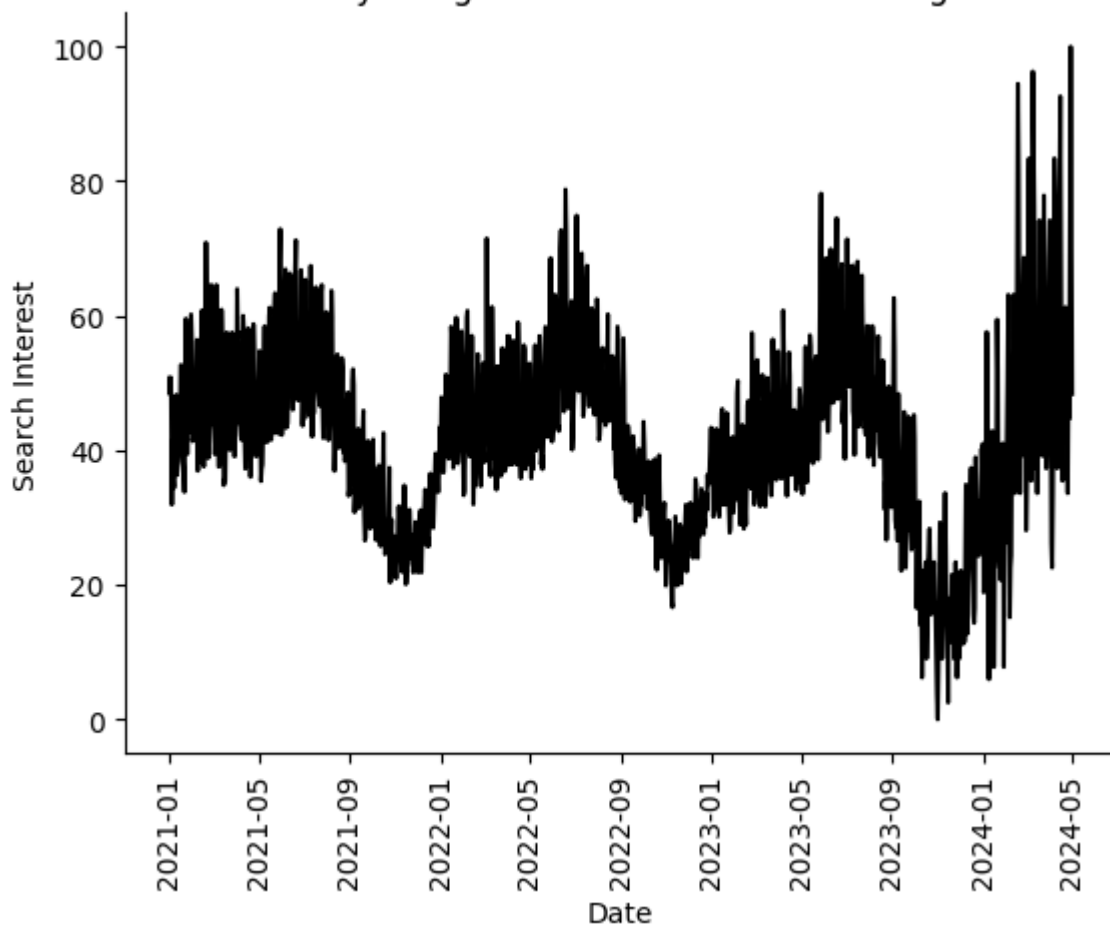
The search interest for camping is similar to boating. Camping search interest had a very low value during the winter of 2023-2024. This may be tied to the warm winter leading the loss of winter recreation opportunities. (Camping is often done in the pursuit of other activities.)

Daily Google Search Interest for Camping



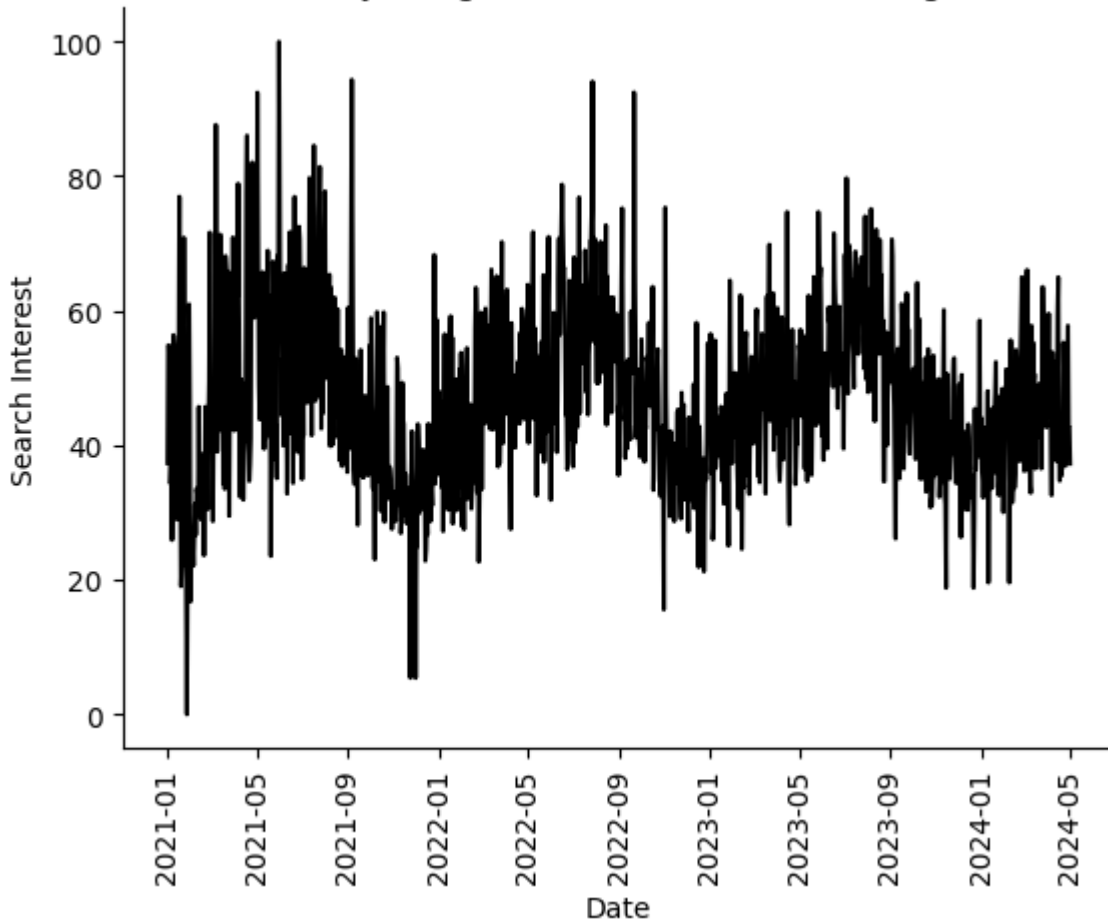
Fishing is less seasonal than other activities and has its low point in late fall. Many fishermen participate in ice-fishing during the winters. The search interest for fishing was very low during the winter of 2023-2024. The aforementioned warm weather reduced the amount of safe ice in much of the state.

Daily Google Search Interest for Fishing



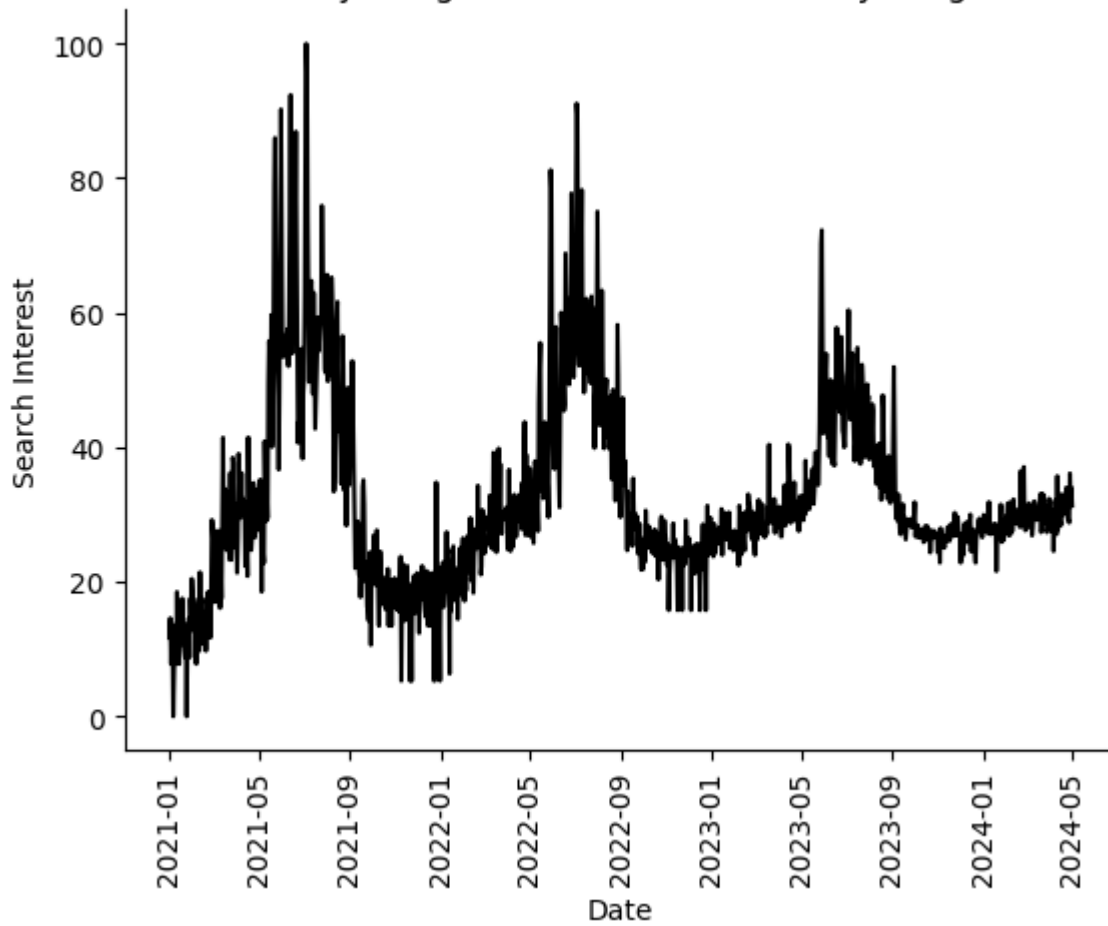
Like fishing, hiking is most popular during the summer; but hiking generates a fair amount of interest year around. Hiking is least popular in the late fall, again like fishing.

Daily Google Search Interest for Hiking



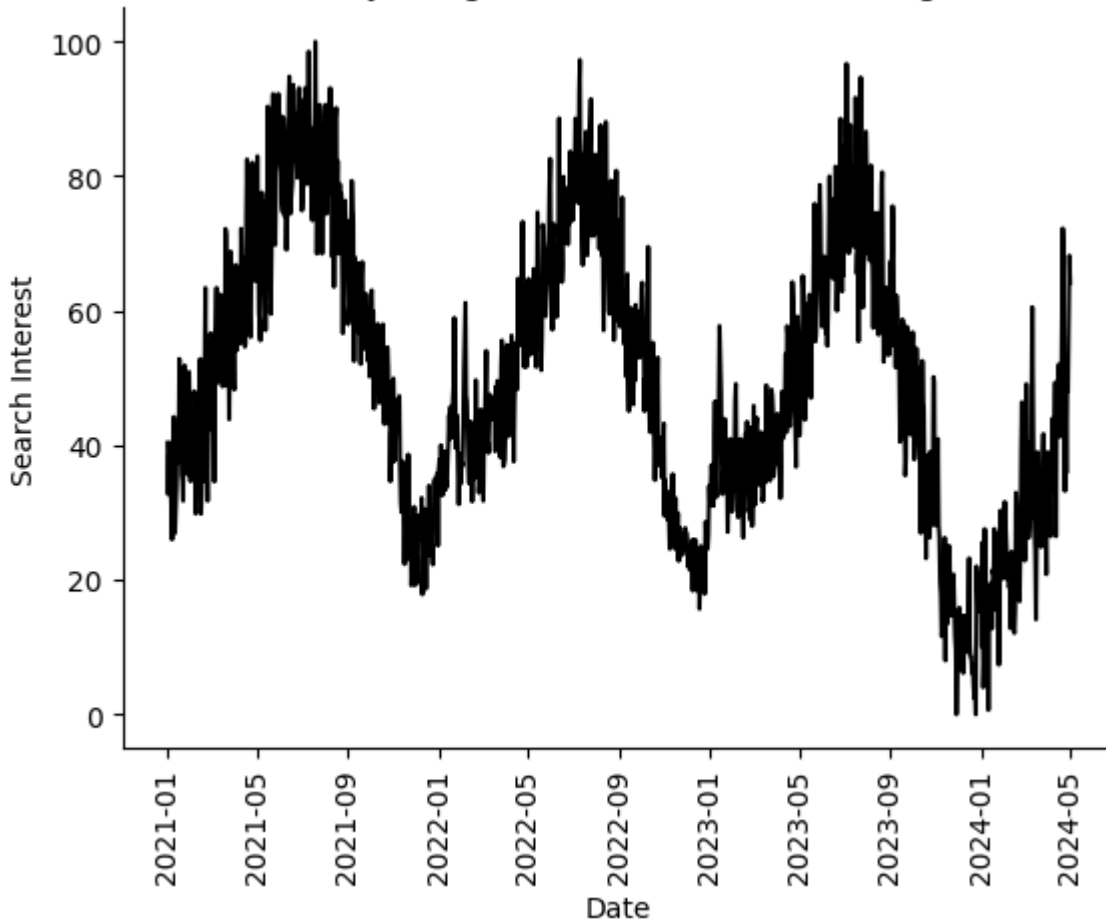
Kayaking shows very interesting trends. Interest is highly seasonal; however, the peak summers are declining in interest over time while the winters are increasing in interest. Kayaking may show a silver lining from warmer winters.

Daily Google Search Interest for Kayaking



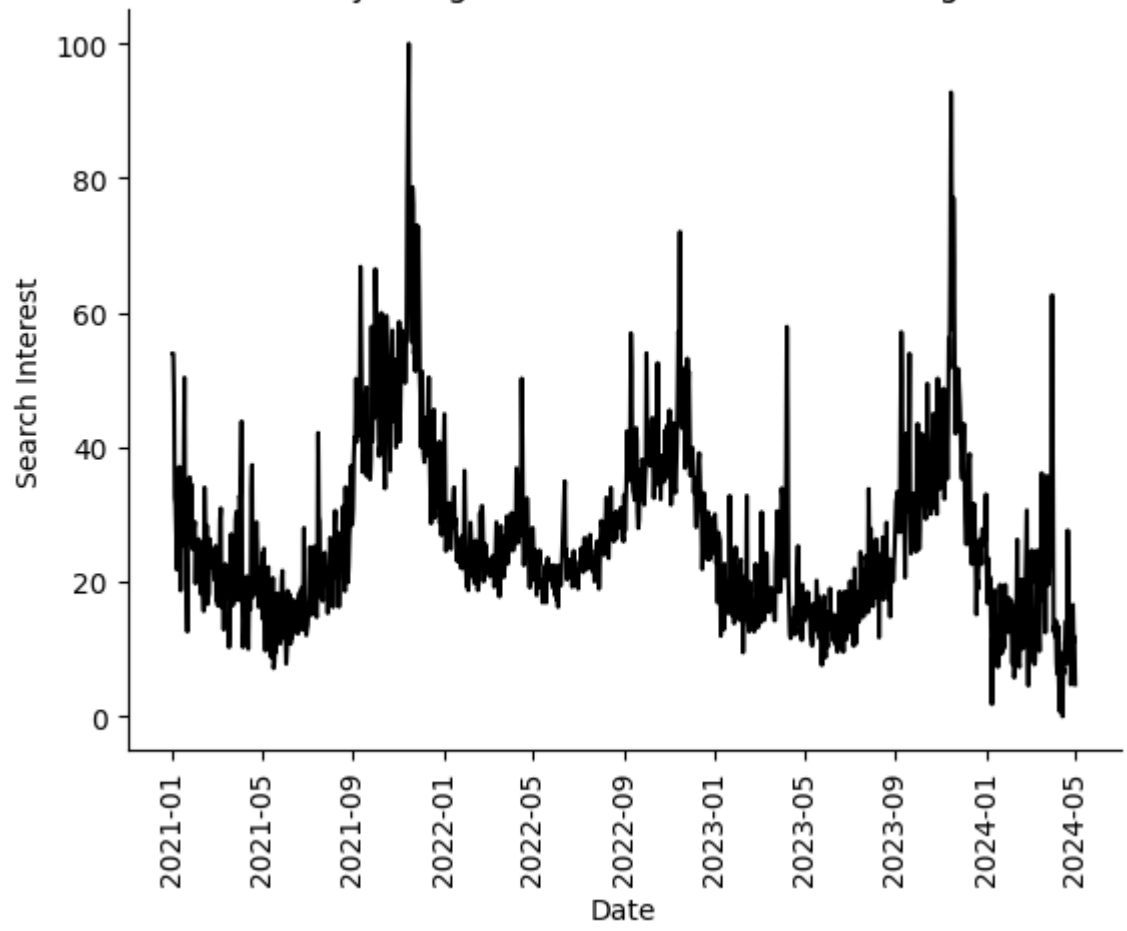
RVing is highly seasonal. While it is more seasonal than ATVing, it is similar to ATVing in terms of declining search interest. RVs were also a popular pandemic era purchase.

Daily Google Search Interest for RVing



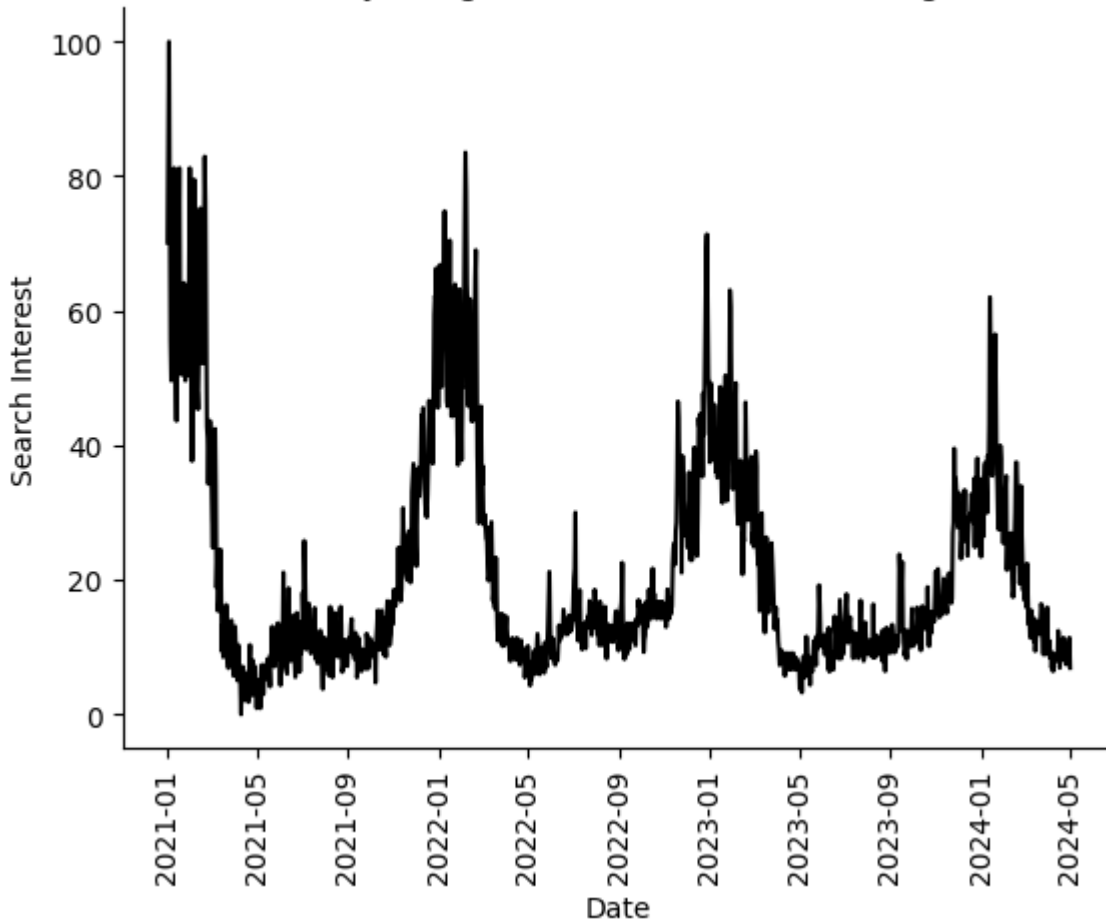
Hunting was a popular activity during the pandemic and appears to be trending downward now. Hunting search interest is seasonal in ways different from any of the other activities. Hunting is most popular during the fall, and secondarily during the spring. The very large peaks in search interest during November are on November 15th which is the first day of firearms deer season. Peaks in April are due to the start of turkey season.

Daily Google Search Interest for Hunting



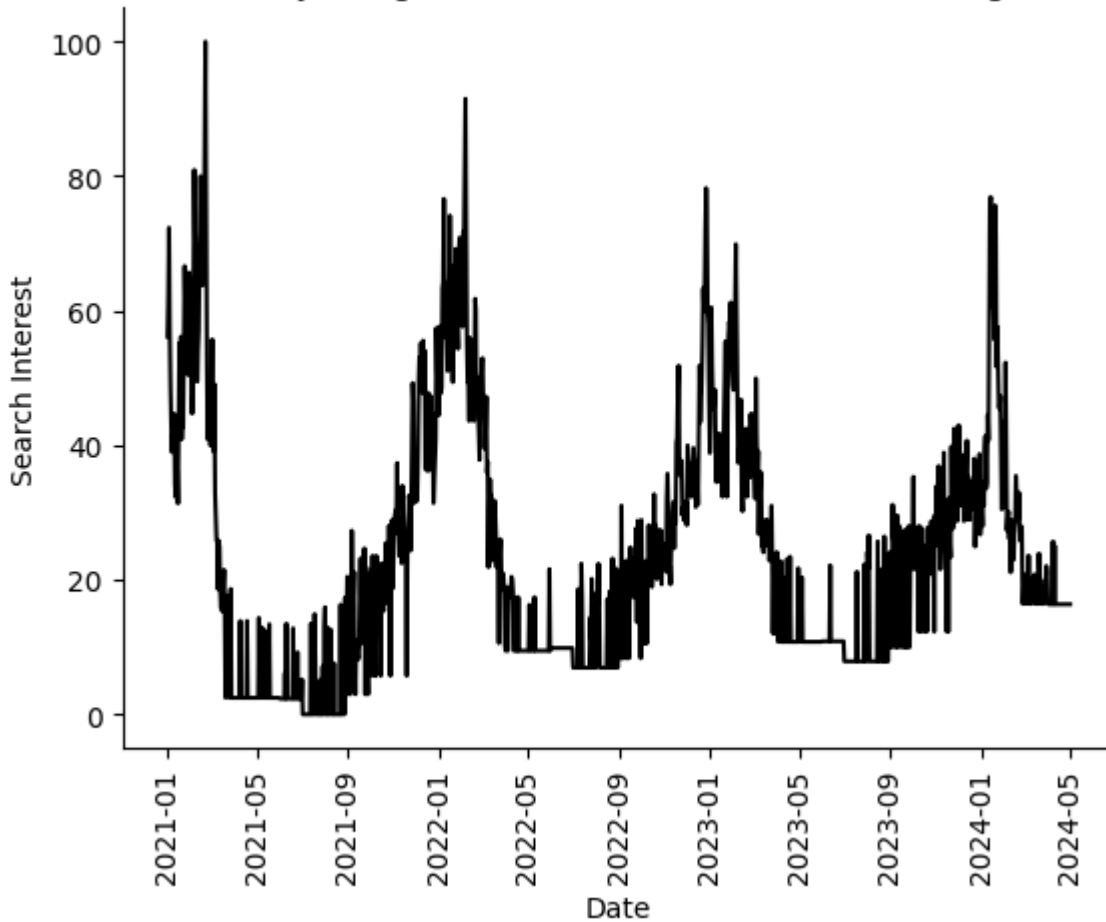
Unlike summer activities, skiing interest is extremely seasonal with most interest in the first few months of winter. Skiing interest is trending downward possibly due to warmer winters.

Daily Google Search Interest for Skiing



The final activity I analyze is snowmobiling. Snowmobiling is seasonal with interest in early winters. Snowmobiling interest is trending downward, likely at least partly due to warmer recent winters. However, search interest in some of the more recent summers seems, ironically higher. I'm not sure what explains this. If I was to make a random guess, it's possible that people are ordering snowmobiles during this time period for the following winters. It's also possible that there have been promotions or sales in recent years given the lower interest in snowmobiling in general and attempts to push snowmobile sales.

Daily Google Search Interest for Snowmobiling



Search Interest Correlation Plots

Search interest between various forms of outdoor recreation is correlated. For instance, camping and boating have a correlation of over 0.9. Skiing and snowmobiling have search interest correlation of almost 0.9.

However, most of this correlation is spurious and due to seasonality or, in some cases, trends over time. For instance, all of the summer activities are positively correlated: atving, boating, camping, fishing, hiking, kayaking, and rving. Both of the winter activities (skiing and snowmobiling) are highly positively correlated. Hunting correlations are generally low, negative with summer activities, and slightly positive with winter activities. This is because hunting occurs in fall and spring when both summer and winter activities are less popular.

While this spurious seasonal correlation is interesting, it's not unexpected or all that useful for forecasting or prediction. The reason that hiking and kayaking appear to be related isn't due to people actually choosing these activities together but rather due to both happening to be popular during the summer due to warm weather. Changes due to seasonality can easily be anticipated by appropriate time series techniques without considering the values or correlations with other variables.

What would be more interesting and useful is to find that two activities are related even after accounting for seasonality. For instance, it's possible that many people combine activities such as camping and hiking. However, even then, correlation may be due to other factors such as responses to warmer weather or too much rain.

In the end, statistical analysis is necessary to determine more interesting results from the data. This will be the subject of the next installment.

Correlations Between Michigan Search Interest for Outdoor Recreation

