Eviction Filings
Part 3

Bliss Cohen

January 04, 2021

Background

The goal of this project is to identify the best way to present data to answer the following questions:

- How are eviction filings trending for a particular region?
- What approach could be used to identify possible thresholds for initiating intervention programs?
- Is one region's filings higher or lower than another?

I revised my approach after Guy and I talked on December 23rd, 2020:

- Explore distributions more thoroughly
- Deprioritize modeling efforts until data is better understood

Data Snapshot

The following file aggregates the number of daily eviction filings from January, 2016 - December, 2020 across 12 regions of interest. Missing days were filled in with a filing value = 0 to ensure each region had a record for every day of the year. Since filings aren't entered on weekends and certain holidays, each record was labeled with a "Yes/No" valid filing day label. (I know I should use Boolean labels but I'm fighting years of habit. . . someday!)

```
## Rows: 21,667
## Columns: 11
## $ New_Date      <date> 2016-01-01, 2016-01-01, 2016-01-01, 2016-01-01, 2016...
## $ Region        <fct> BNGP Counties, Brown, Calumet, Green Lake, Kewaunee, ...
## $ Month         <ord> Jan, Jan, Jan, Jan, Jan, Jan, Jan, Jan, Jan, Jan, Jan...
## $ Year          <dbl> 2016, 2016, 2016, 2016, 2016, 2016, 2016, 2016, 2016,...
## $ Day           <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2,...
## $ Num_evictions <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...
## $ Weekday       <fct> Friday, Friday, Friday, Friday, Friday, Friday, Frida...
## $ Week          <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...
## $ Quarter       <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...
## $ Holiday       <chr> "New Years Day", "New Years Day", "New Years Day", "N...
## $ Valid         <chr> "No", "No", "No", "No", "No", "No", "No", "No", "No",...
```

Exploratory Highlights

Before tackling the questions posed above, I will conduct more exploratory work. I will continue the approach from Evictions_Part2.Rmd and compare a large population to a smaller one.

1. Filtering out non-valid filing days is useful for distribution analysis.

The histograms below compare WI to Kewaunee. Both histograms have a large count near filings = 0 because they contain all the days in a year.
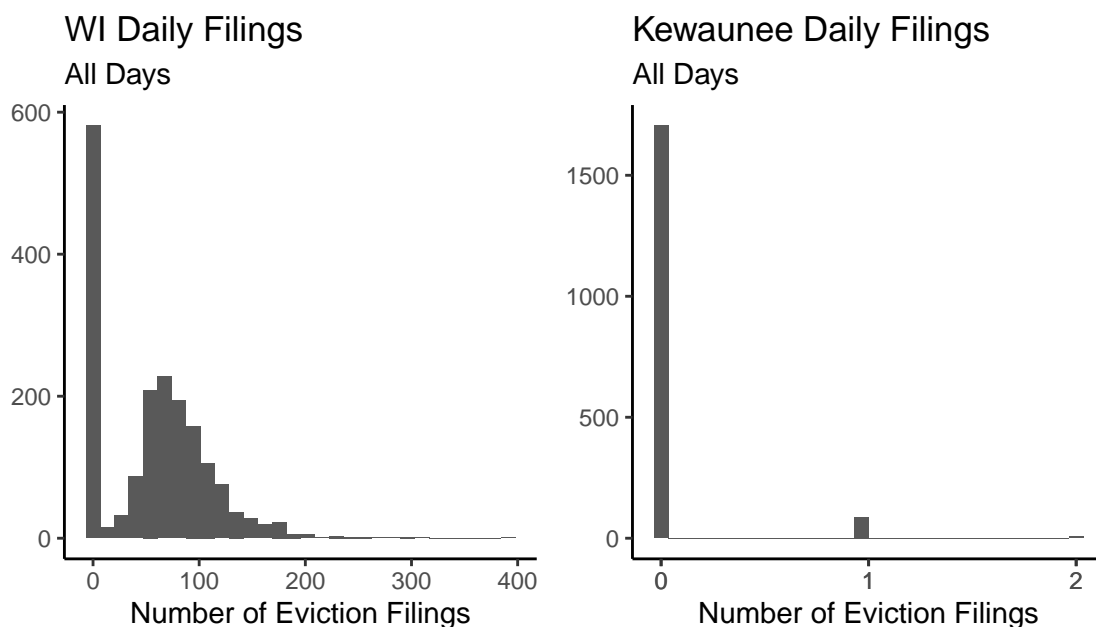

After filtering out non-valid days, WI no longer spikes around 0 but Kewaunee is still dominated by filings = 0. It was surprising to see any 0-day filings for WI, but they appeared legitimate (fell within the eviction moratorium or on a delayed holiday - not shown).
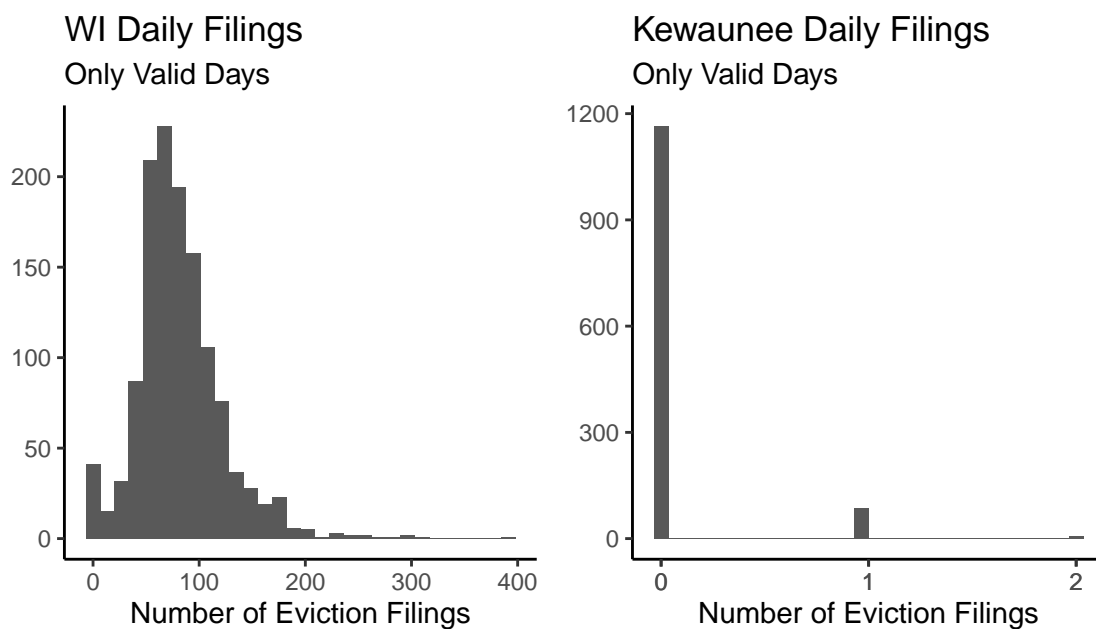
2. Many of the regions' filings are 'outliers'.

The following graphs compare boxplot distributions across all areas ordered by population. The graph on the right zooms in to better view less populated regions.

It is striking to see that even after filtering out non-valid days, many of the regions don't have visible boxplots because of all the filings = 0. Hmmm... comparing regions to each other doesn't seem legit if you only have outliers???
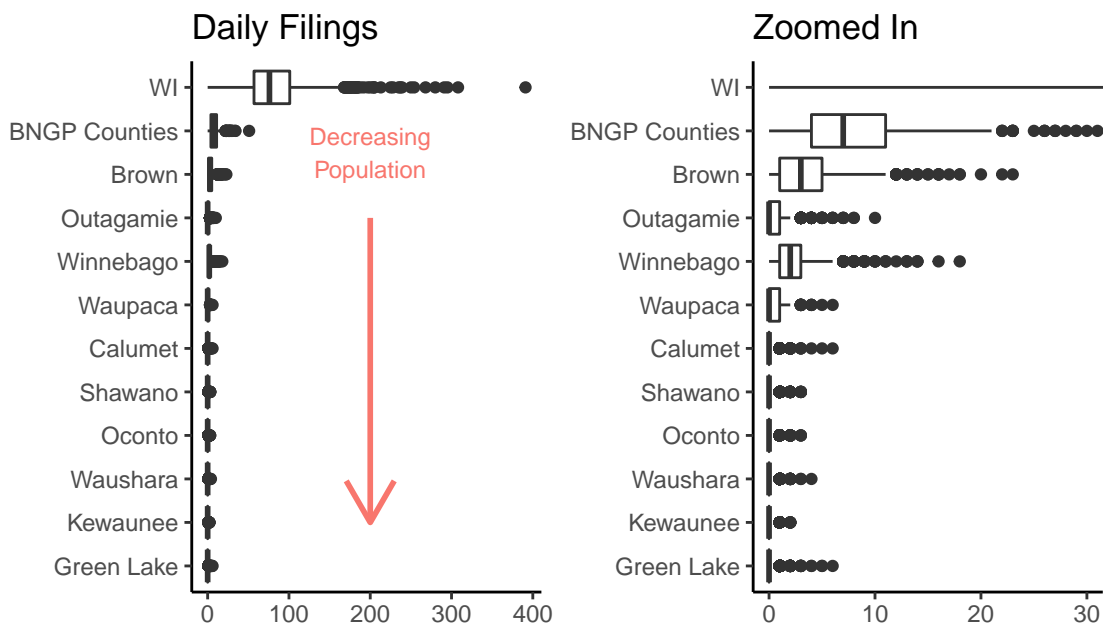
The remaining explorations primarily focus on WI given the state's robust distributions.

3. Filings are consistent across weekdays.

The number of filings is similar by day of week. That said, Monday and Tuesday show higher spikes, possibly driven by backlogs from the weekend. In general, Tuesday tends to be the most productive day of the week; people are either gearing up for the weekend by Wednesday or trying to come out of the weekend fog on Monday. Tuesday does indeed seem to be the most productive filing day.

3
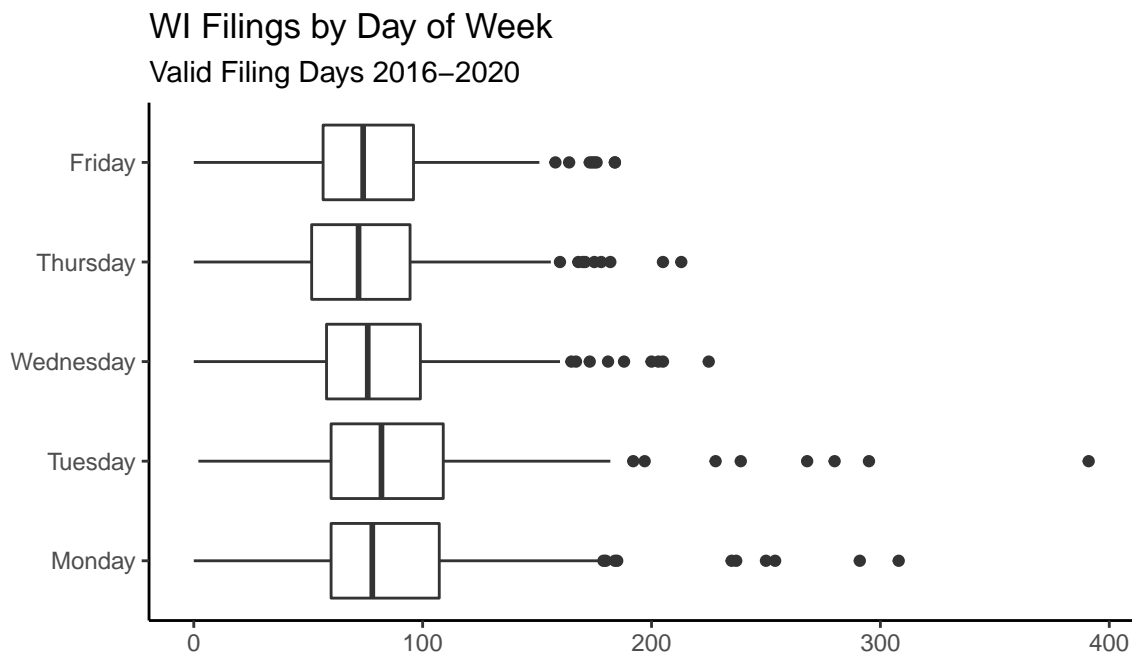
**4. Filings show a wave-like behavior across the year.**

Filings dip in the spring, peak in the summer, and then slowly taper off towards the end of the year, eventually bottoming out again in the spring of the following year. The trend is easier to visualize when plotted by month instead of week given the seasonal patterns.

WI Filings by Month — Valid Filing Days 2016–2020; WI Filings by Week — Valid Filing Days 2016–2020
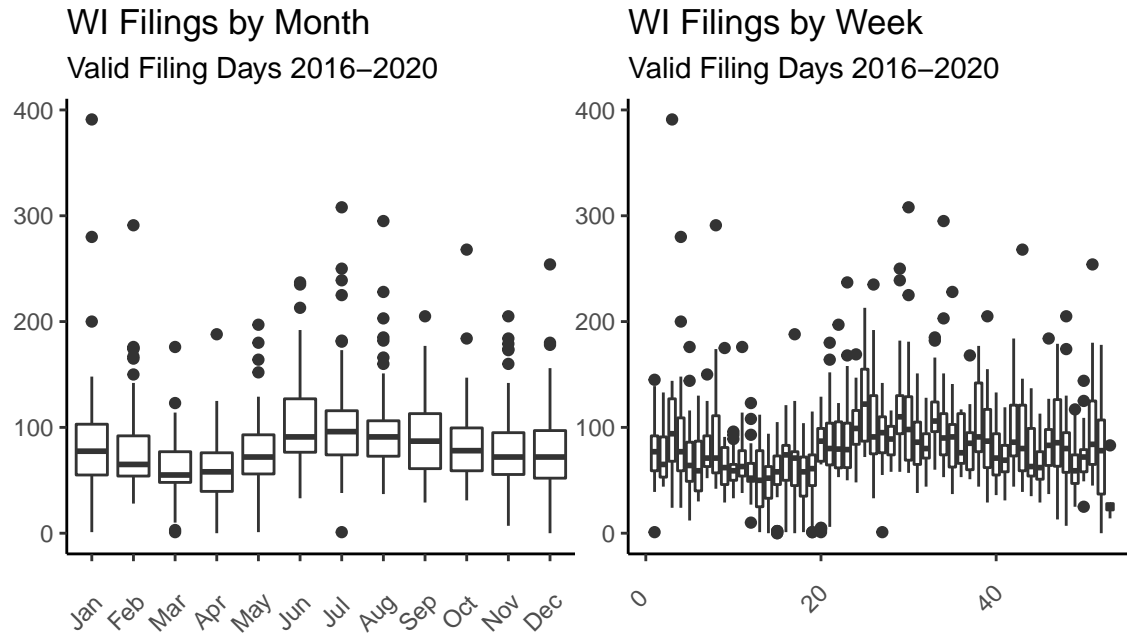
**5. The number of filings changes across the years.**

The following graph aggregates months across the years to yield the following insights:

- Something changed in mid-2018, spiking eviction filings above 2016-2017 levels.
- Filings stayed at the higher levels until early 2020.
- Not surprisingly, 2020 is the most erratic year given the pandemic and eviction moratorium; filings may now be stabilizing at the lower 2016-2017 levels.

4

WI Monthly Filings
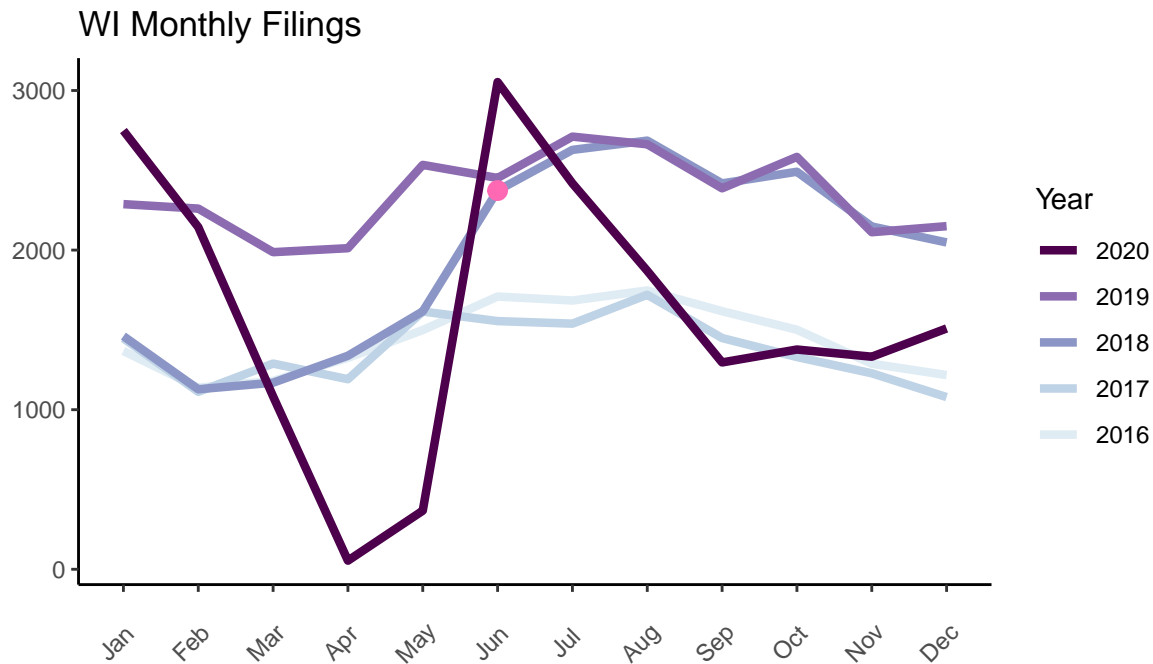
6. Patterns are more difficult to see in small regions.

There is too much noise in small regions, as shown below, to identify clear patterns.

Kewaunee Monthly Filings
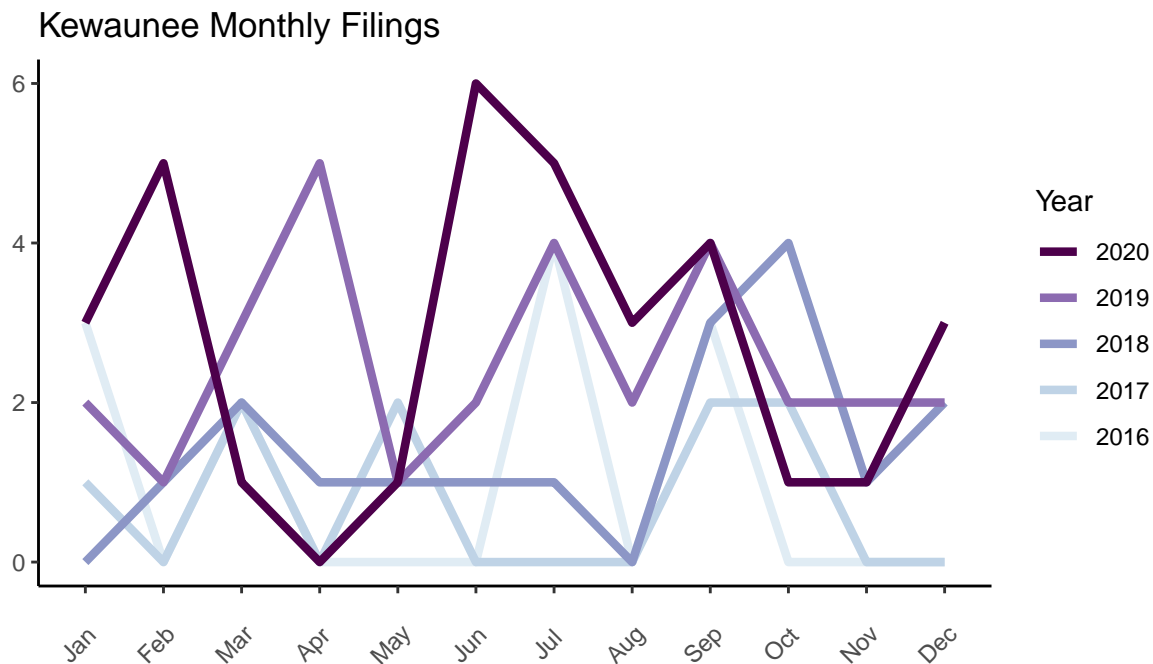
Trends

1. Year-to-Date (YTD)

Identifying thresholds for initiating action is not an easy or well-defined task. Should relief programs be implemented if eviction filings are up 10%? 20%? 50%? To aid decision makers, it may be helpful to implement tools to trial different thresholds.

The table below shows the % change between current vs prior years as of the last update. In a Shiny app,

the user could theoretically enter a threshold % to change 'Value Box' colors for each region (e.g. increasing, decreasing, steady).

Table 1: Snapshot of Most Recent YTD (as of last update)

Region	Start	End	YTD Filings	Pct Change vs Prior Year
WI	2020-01-01	2020-12-31	19254	-32
WI	2019-01-01	2019-12-31	28141	
BNGP Counties	2020-01-01	2020-12-31	2216	-20
BNGP Counties	2019-01-01	2019-12-31	2762	
Brown	2020-01-01	2020-12-31	882	-24
Brown	2019-01-01	2019-12-31	1162	

If something in the snapshot was concerning, the user could explore historical year-on-year changes to see if the most recent trend was anomalous.

In the example below, the end date was changed to February 29, 2020, just to make sure leap year issues were handled correctly. Clearly finding a % threshold is pretty iffy for small regions.

Table 2: WI and Kewaunee YTD as of Feb. 29 (or 28th)

| Region	Start	End	YTD Filings	Pct Change vs Prior Year
WI	2020-01-01	2020-02-29	4894	8
WI	2019-01-01	2019-02-28	4548	76
WI	2018-01-01	2018-02-28	2589	1
WI	2017-01-01	2017-02-28	2559	2
WI	2016-01-01	2016-02-29	2508	
Kewaunee	2020-01-01	2020-02-29	8	167
Kewaunee	2019-01-01	2019-02-28	3	200
Kewaunee	2018-01-01	2018-02-28	1	0
Kewaunee	2017-01-01	2017-02-28	1	-67
Kewaunee	2016-01-01	2016-02-29	3	

2. Monthly

Monthly aggregations provide another way to investigate filings. In a Shiny app, the user could hone in on a region and year.
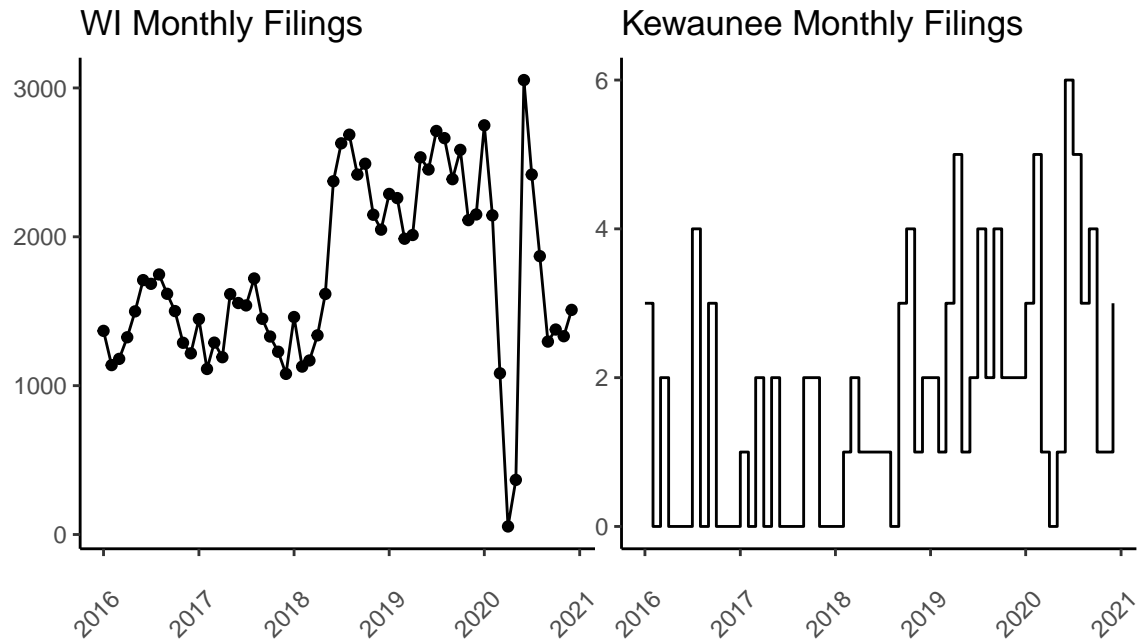
3. Daily

It's not clear if daily aggregations provide additional value. And certainly, smoothing small regions is especially wonky and probably not a good idea.
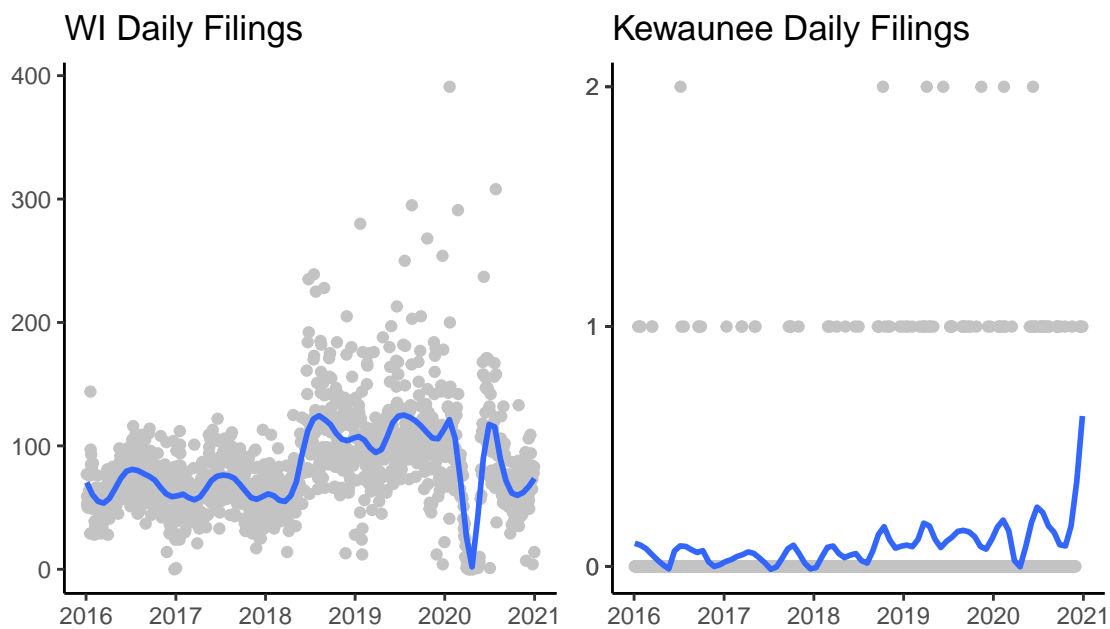

Compare Regions

Given the noise inherent in small regions, it might make more sense to compare different areas using annual aggregations rather than monthly. Filings need to be scaled to an appropriate metric before making comparisons.

The distribution below shows aggregated annual filings overall for the 12 regions.

Aggregated Yearly Filings – All Regions
2016 – 2020
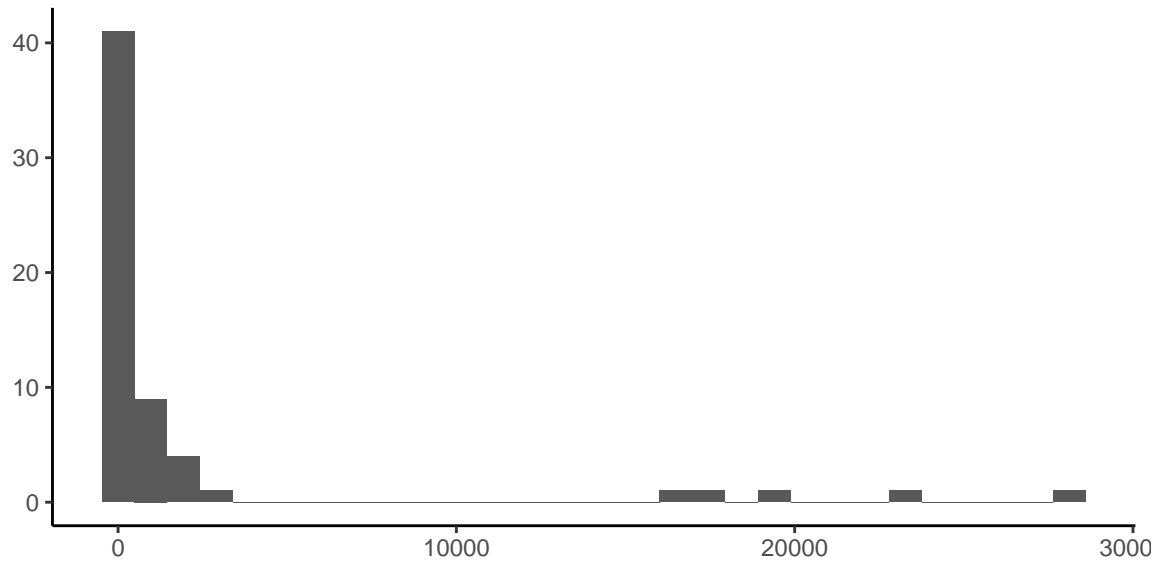

```
##     Min.  1st Qu.   Median    Mean  3rd Qu.     Max.
##     9.00    40.75    62.50  2091.08   768.50  28141.00
```

Each region was joined with the estimated number of occupied rental units, as obtained from the U.S. Census.

Table 3: Estimated Number of Occupied Rental Units

Region	Units
WI	777217
BNGP Counties	104630
Brown	37011
Winnebago	24555
Outagamie	21259
Waupaca	5875
Shawano	3867
Calumet	3834
Oconto	2773
Green Lake	1906
Waushara	1824
Kewaunee	1726

An annual filing rate was calculated for each region by dividing the number of filings by the number of units. The rate was ultimately expressed as the number of filings per 1000 rental units (filings/unit*1000).

The graph below shows the boxplot distributions of filing rates for each region. Visually, the distributions clump into 3 groups: low, medium and high.

Filing Rate: Aggregated Yearly Filings per 1000 Units
2016 − 2020

Filings per 1000 Occupied Rental Units
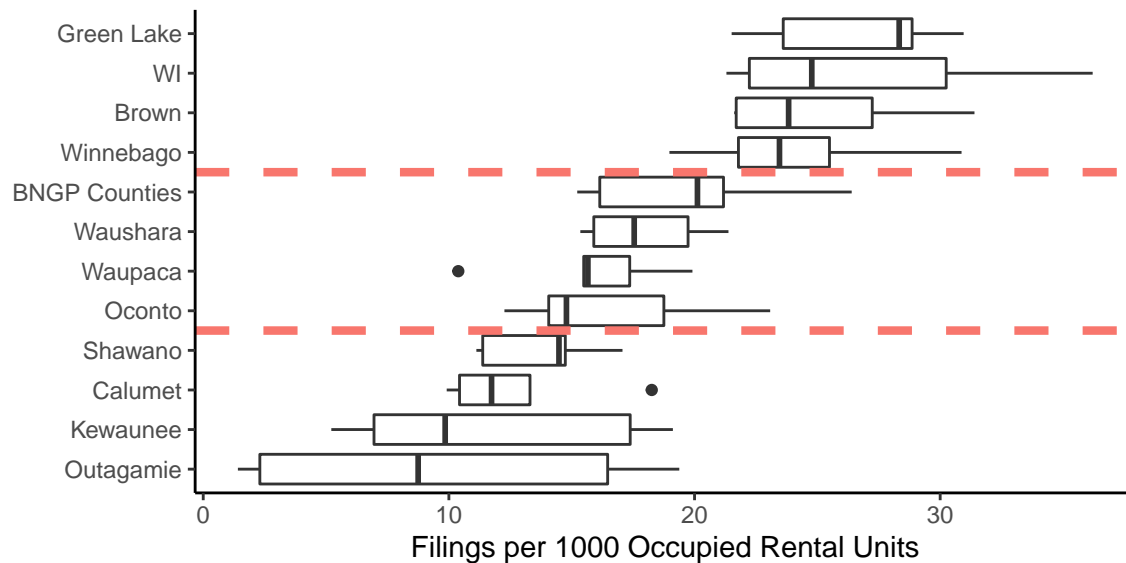
Filing rates were ranked from 1 to 12 with 1 being the lowest and 12 being the highest. The low, medium, and high groups are visualized below across the years:

- Similar to WI, Green Lake, Brown, and Winnebago have the highest rates
- Outagamie, Kewaunee, Calumet, and Shawano have the lowest rates
- The remaining counties, including the BNGP County aggregation, are somewhere in between

Annual Filing Rates: Ranked Lowest (1) to Highest (12)
Filings Per 1000 Units
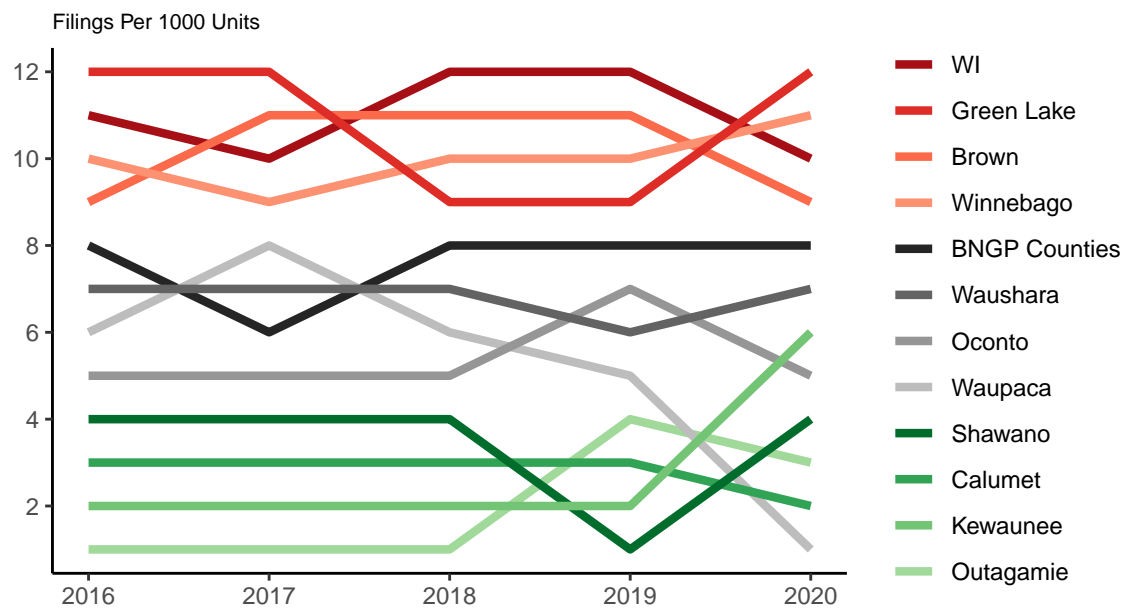
Questions and Next Steps

- Did any analysis strike you as not 'data smart'?

- – Note I did not say 'data dumb' ☻

- What visuals do you prefer and why?
- Any additional explorations?
- It feels like doing stats on this stuff is not a good idea given that many regions only have outliers. Would you agree? Is ranking OK?
- I like the idea of flagging a warning if a particular region is trending up, but how? % change doesn't make sense with small regions... maybe some type of 'control chart' +2-3 sigma?
- In addition to the number of occupied rental units, I joined the features below. Does it make sense to examine possible correlations (multicollinearity will be an issue):

  - – Population
  - – Total number of Households
  - – Number Households under the Official Poverty Level
  - – Household median income
  - – Household mean income
  - – Number ALICE Households (above Official Poverty Level but unable to meet basic needs)

Appendix

This section includes the code to pull and process variables from the U.S. Census and United Way. All U.S. Census data were taken from the 2019 ACS 5-Year Estimates and accessed at the County and State geographies. The United Way ALICE information contained 2018 data.

Process

- Use tidycensus::get_acs() to access Census data; download ALICE csv
- Combine Features into a data frame
- Create a BNGP County Aggregation

  - – Most of the variables could be summed across counties except for Median Household Income and Mean Household Income. The aggregated median was assigned NA since medians aren't calculated. The aggregated mean income was found by taking the average of the county averages.

- Write Feature Files

References

Population: Table B01003

Renter Occupied Units: Table S2502

Total Households, Households Under Poverty Threshold, Median Household Income: Table S2201

Mean Household Income: Table S1901

United Way ALICE for Wisconsin