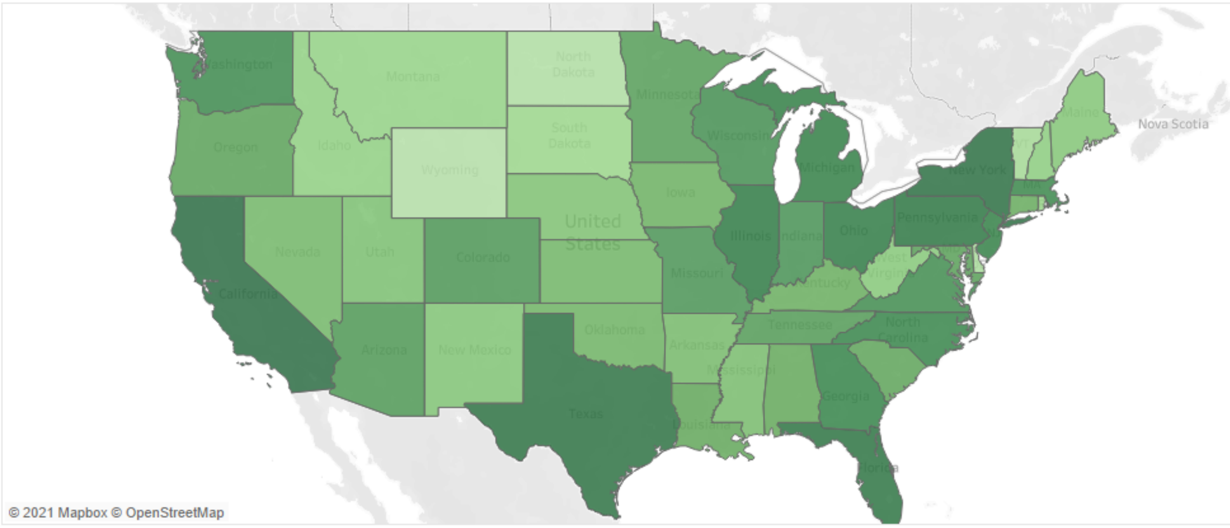
Patrick Keener
Homework 2
DSC 465 – Data Visualization
1/31/2021

Question 1- Food Service by County

1a. Food Services Availability by State (2007)

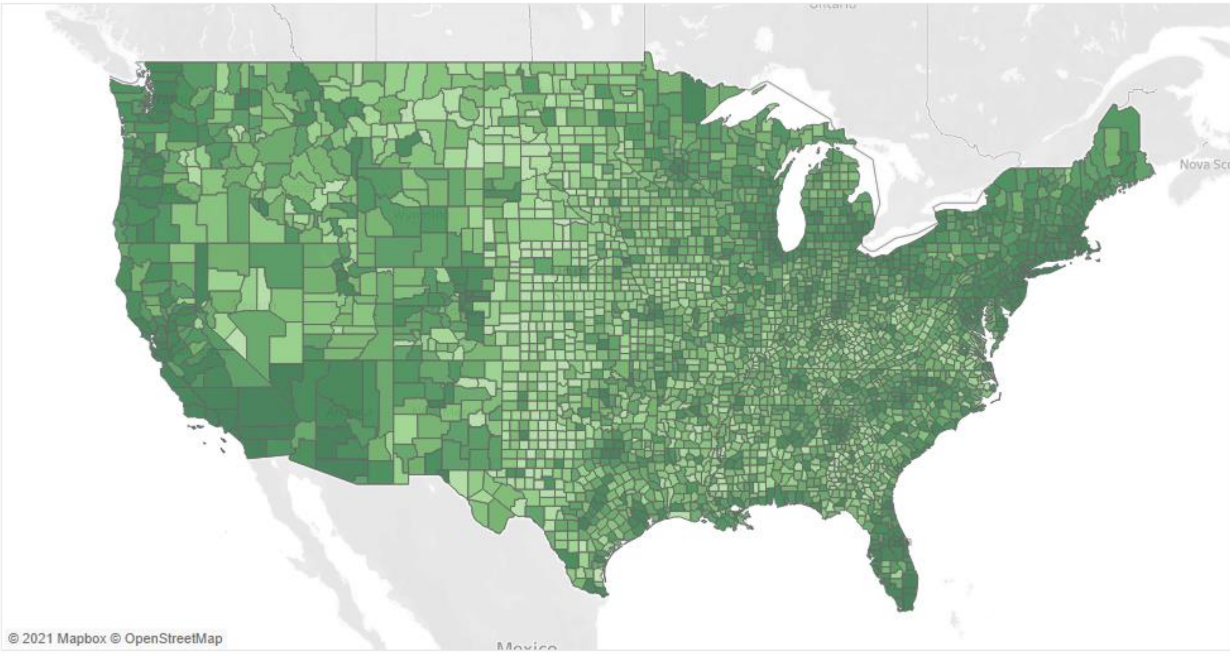


Food Services availability by State was created using the map feature in Tableau. Once state and food availability was added, there was a null value that needed to be filtered (it was the aggregate value of the dataset). I then set the value as the percentile, since we're interested in a relative comparison. Finally, I selected a sequential Green palette because green is the color of life, and we are comparing levels of the same thing rather than categroical or different outcomes.

It's interesting to note that the mid-northwest and central US ranks lowest, given that most food is grown here. This is likely (untested) explained by low population density.

Patrick Keener
Homework 2
DSC 465 – Data Visualization
1/31/2021

1b. Food Services Availability by County (2007)



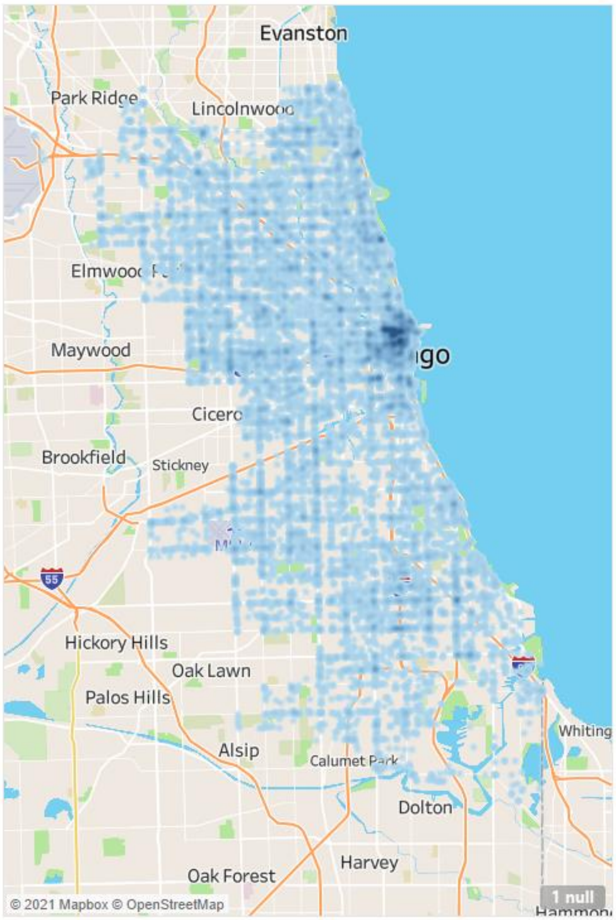
Food Services availability by County was created using the map feature in Tableau. Once state and food availability was added, there was a null value that needed to be filtered (it was the aggregate value of the dataset). I then set the value as the percentile, since we're interested in a relative comparison. Finally, I selected a sequential Green palette because we are comparing levels of the same thing rather than categroical or different outcomes, and green is the color of growth/plants and life (food).

It's interesting to note that the mid-northwest and central US ranks lowest, given that most food is grown here. This is likely (untested) explained by low population density, which is nominally supported by the green "hotspots" in otherwise barren states within the mid-northwest and central US.

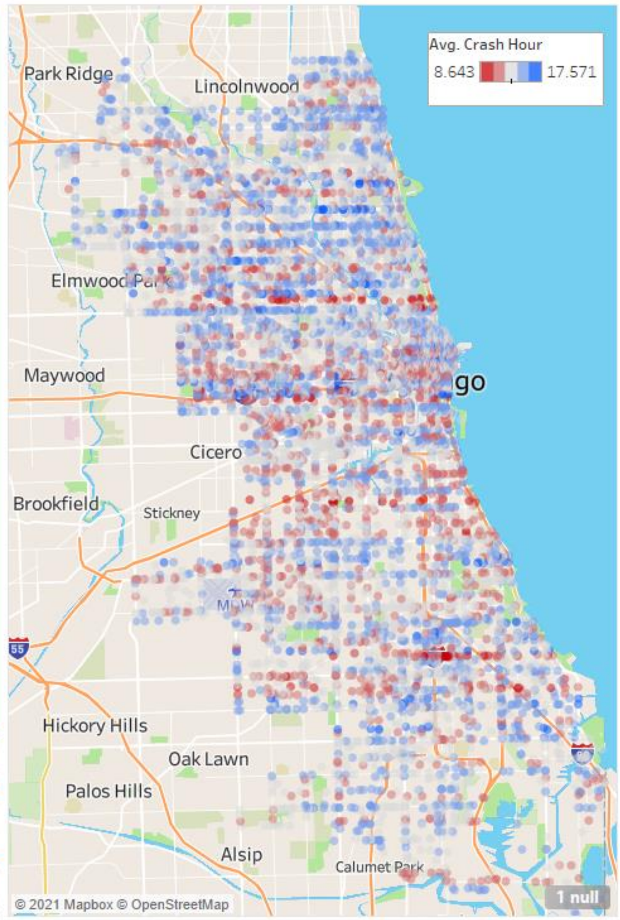
Patrick Keener
Homework 2
DSC 465 – Data Visualization
1/31/2021

Question 2: Chicago Crashes

2a. Density Map of Crashes



2b. Average Hour of Crash in Vicinity of Point



Patrick Keener
Homework 2
DSC 465 – Data Visualization
1/31/2021

2a.

This chart was created using the Map (symbol) chart-type in Tableau. After latitude and longitude were added (as dimensions) to Row & Column, one null point that was excluded from the visualization (it was at location 0,0 off the coast of Africa). The points were set as density and set to 55% opacity. The size of the points was adjusted downwards from default in an effort to make it easier to distinguish them.

The data itself yields only a few patterns, and few unexpected. The Loop, which is the dark spot along the lake in the middle of the graph is one of the busiest spots in the city, so more accitdents are to be expected. It is interesting to note that there are a few spots on the west side of the city- presumably busy streets - that have additional accidents. There also seems to be a heightened density along the coast going north of the city (following Lake Shore Drive, a primary thoroughfaire). Another cluster can be found along the highway to the south of the city, near where 94 and 80 converge.

2b.

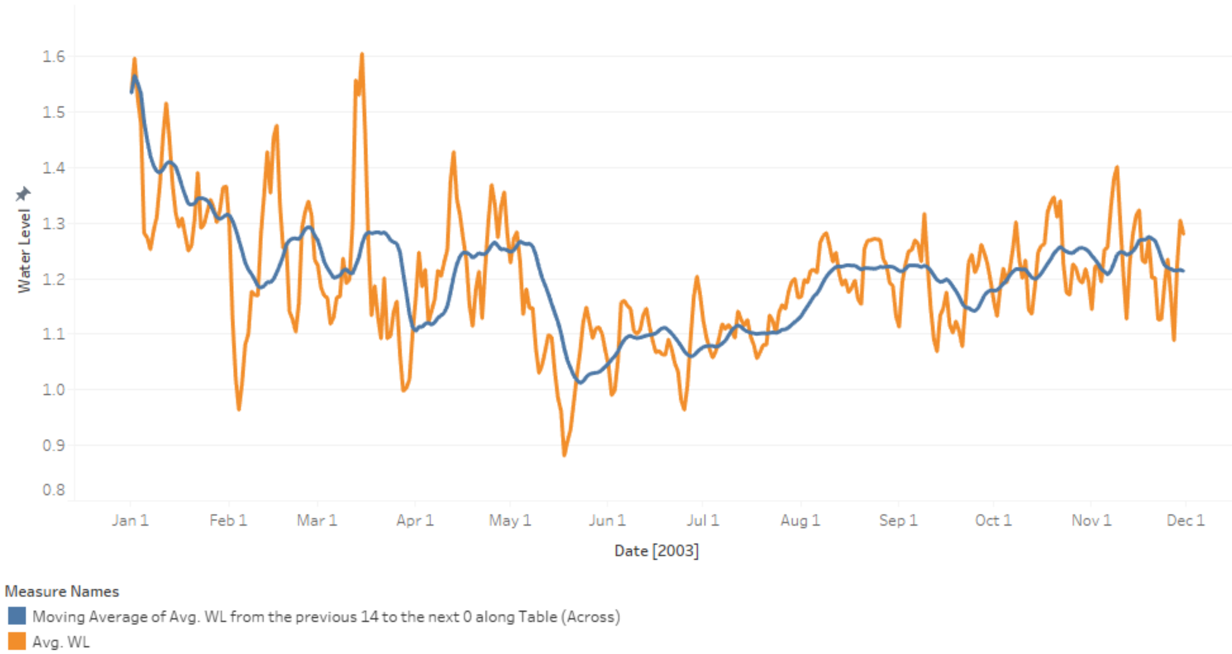
Similar to the chart from 2a, this chart was created using the Map (symbol) chart-type in Tableau. After latitude and longitude were added (as dimensions) to Row & Column, one null point that was excluded from the visualization (it was at location 0,0 off the coast of Africa).

The goal was to visualize the points using colors, with red representing morning (sunrise) and blue representing evening (sundown). At first application, with each point coded to its respective hour, the chart was impossible to read due to clutter, and no helpful patterns emerged. To facilitate this, smoothing was applied via a simple average: Each point had the crash hour from points +/- 10 places of lattitude and longitude from it averaged, which allowed certain patterns to emerge, such as specific roads coloring more blue or red (which can be explained by the prevalance of one-way streets in Chicago; these patterns were mostly seen along lattitude, so one possible explanation is that red roads are eastbound, meaning heavier traffic in the mornings as people enter the city for work, and blue roads are westbound. This theory was not tested). One potential issue (distortion) with this approach is that, since it is average records nearby, if there is an area with few crashes it may average in data from a far geographical distance. This was determined to be a minor risk due to the density of the data itself.

Patrick Keener
Homework 2
DSC 465 – Data Visualization
1/31/2021

Question 3: Water Levels

3a. Water Levels Over Time - Smoothed

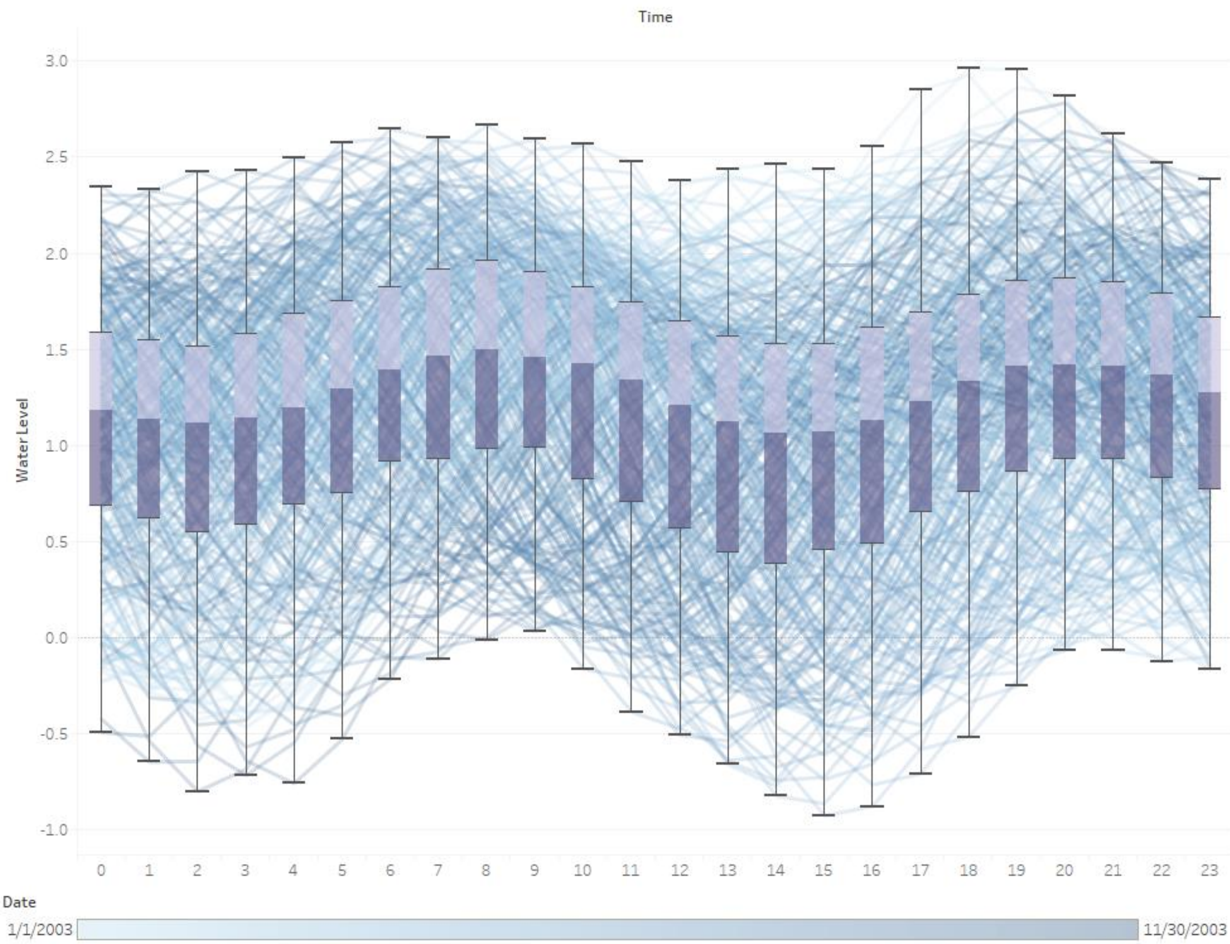


Note: Both Water Levels graphs were made in Tableau due to time constraints (it's busy season at work).

The first Water Level graph was created in Tableau by using Date in the columns and average water level in the rows. There are two lines on the graph: a daily average and a 2-week average. The purpose of the daily average is to give a baseline, while the 2-week average is to help see the bigger picture- specifically, cycles over the months and year. The month of december has been excluded from this graph because it is only a partial month.

Patrick Keener
Homework 2
DSC 465 – Data Visualization
1/31/2021

3b. Water Levels by Hour



This water level graph was also made in Tableau. Each line represents a date and displays the water level over a 24 hour period. The oscillations over the month is clear based on the chart itself (opacity turned down to help show more data). Additionally, I added a box plot which emphasizes the daily average values as well as the overall highs. It indicates that water level reliably changes over the course of the day, and allows us to judge the relative change. Again, December was left out of this chart due to incomplete data and to be consistent with the other charts.

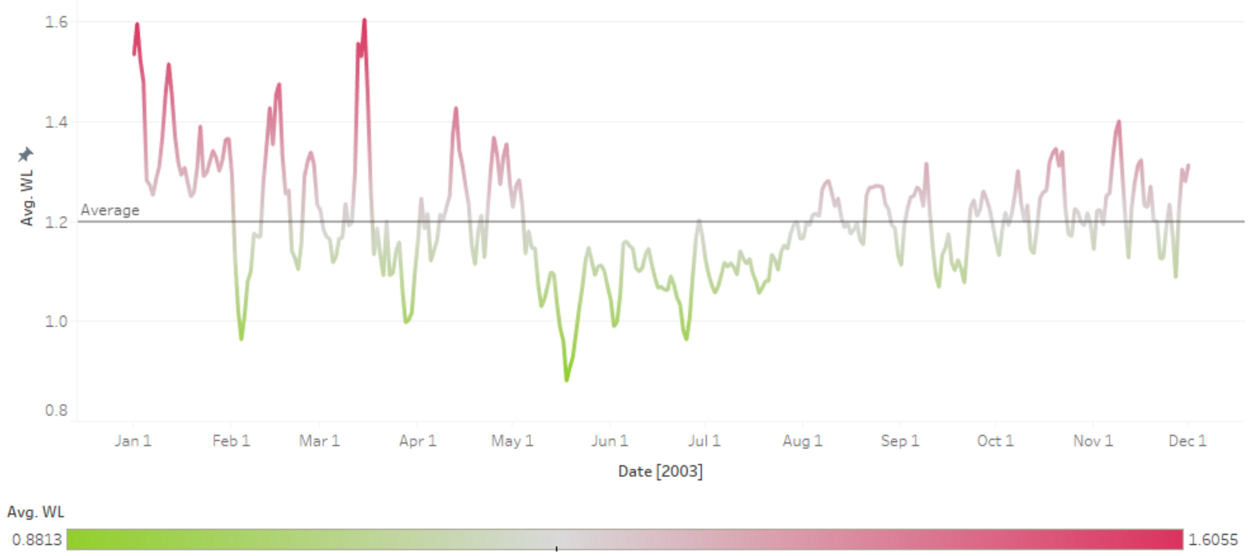
3c.

Each graph communicates different information. The first graph gives context over the entire year. It provides a way to see both intra-month patterns (but no more precise) as well as patterns over the course of the year (such as a decrease in water levels from jan to jun before an increase from jun to jan). The second graph communicates information at the daily level: how the water levels change throughout each day as well as hourly-level summary data over the course of the year (average, range).

Patrick Keener
Homework 2
DSC 465 – Data Visualization
1/31/2021

Question 4: Color Scales (Water Levels Part 2)

4. Color Scales (Water Levels Part 2)



This chart was made using a divergent color scale. The higher and lower values cut off at approximately 1.19 (an 'average' line has been added to aid the reader). I chose green and magenta for two colors because Green is 'good' (lower water levels indicating less climate change) and 'red' is bad, however bright red was too contrasted so I went with about a 150 degree shift for the counter-color. I also reduce saturation slightly to make the color less contrasting.