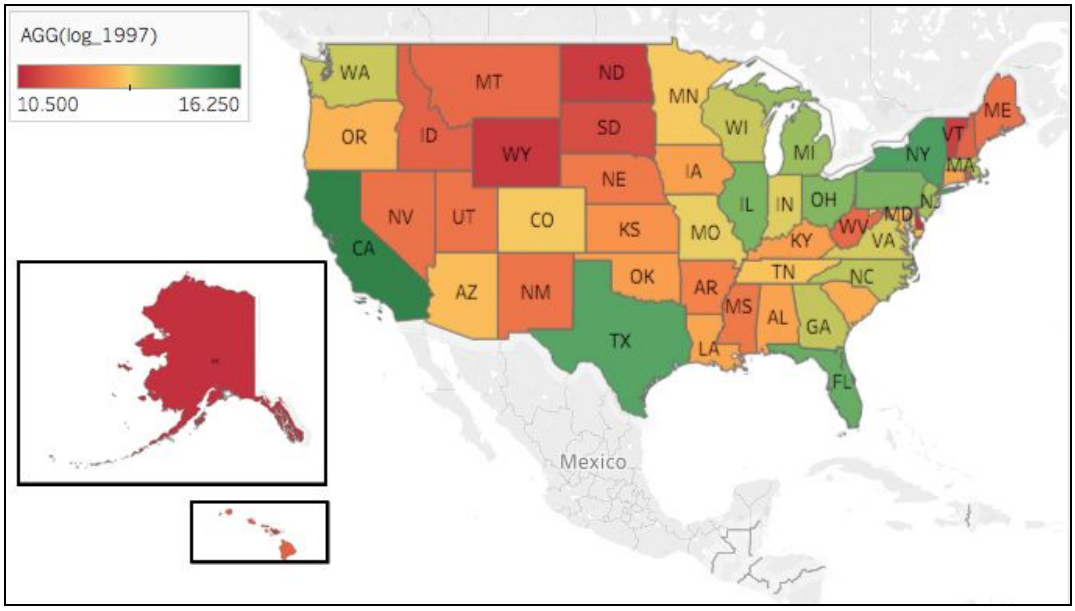
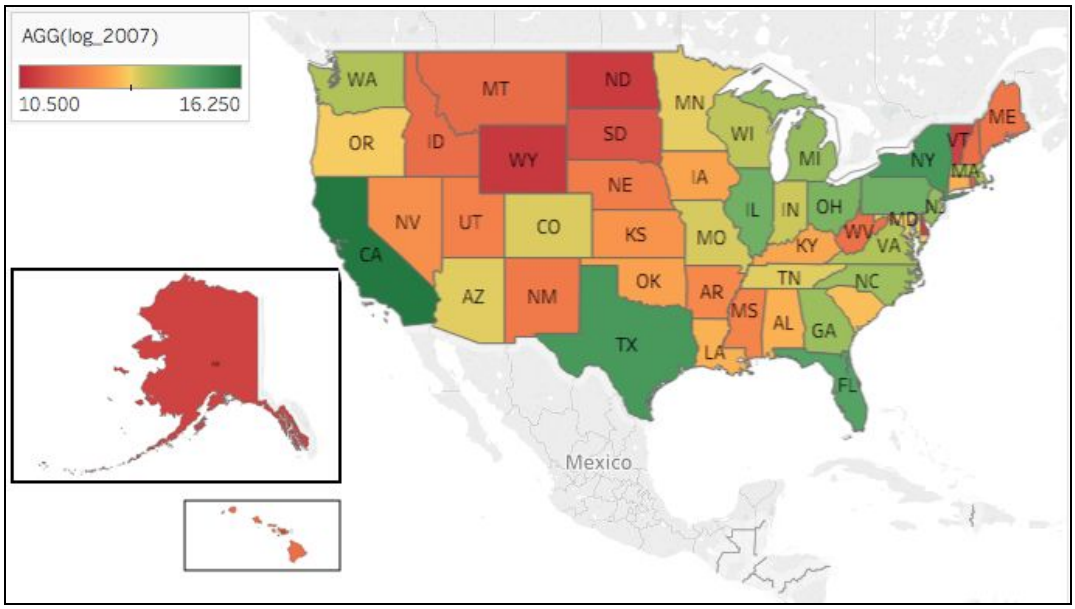
Alex Teboul **DSC 465 - Homework 4**

1) (20 points) Download the FoodSrvcByCounty.txt file and create the following visualizations for this geographical data.

a. Graph food services by state as a choropleth for two of the years. Note any patterns that arise. Your visualization should clearly display states that have high levels or low levels of food service availability.



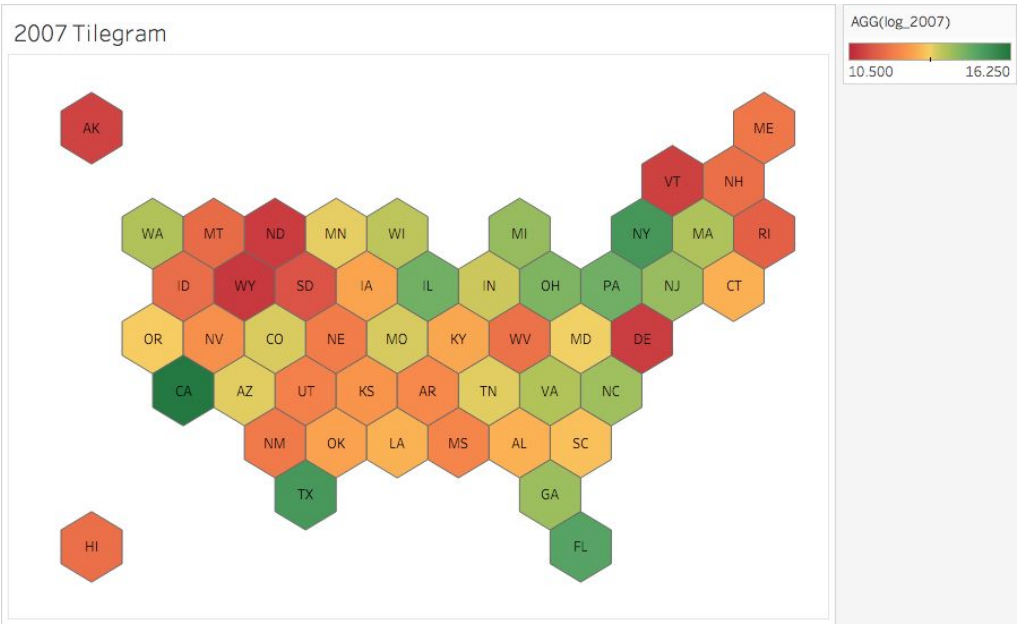
1997



2007

Explanation: The first choropleth shows the Sum of Food Services by State in 1997 and below for 2007. To highlight the high vs low Food Services offered states, I chose a diverging Red-Gold-Green color scheme. Note the insets for Alaska and Hawaii are not perfectly to scale. The scale I used was log base 2 to further differentiate given the scale present. It does not appear that there is a large change in Food Services offered in States between 1997 and 2007.

b. Build a tile cartogram (this is possible with either R or Tableau) and create a cartogram/tileogram of the state data from this dataset for two years. Analyze the difference in visual communication between your graphs in a) and b). What differences can you see? Is there anything that you can see easily in one graph that is more difficult to see in the other?



Explanation: The first Tilegram shows the Sum of Food Services by State in 1997 and below for 2007. There isn't much of a difference between the two charts visually. Instead it tells us that between 1997 and 2007 not much has changed in terms of the food services offered in each state. Those states with low availability of food services still have relatively low availability and the high availability states remain so as well. For example, California added a lot as did New

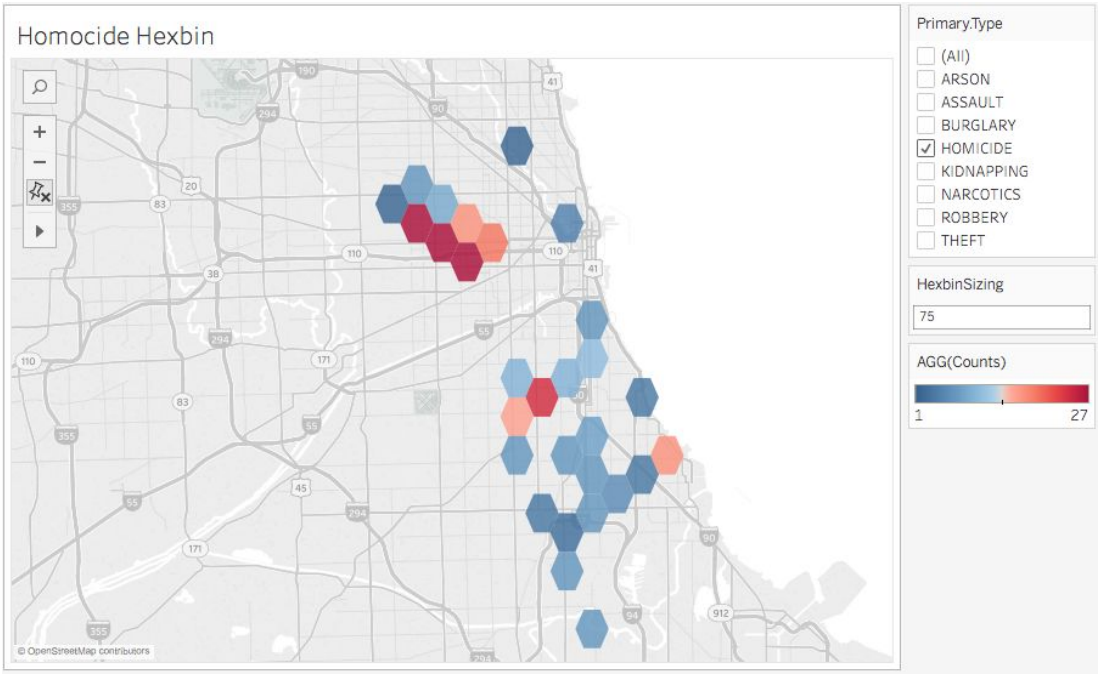
York (upwards of 10k each), but relative to other states that also added availability they remained much higher and so the color difference was not particularly noticeable. If the goal of this chart was for comparison, I think that a between option would be to calculate the difference in food services offered in 1997 vs 2007 and graph that. In that way you could see which states saw losses and gains.

I also tried different color schemes, and scales but the story the data tells remains the same. That said, there are some key differences between the choropleth in 'a' and the tilegram in 'b'. One such difference is that 'a' preserves the general geographic representation of the data with color values assigned to a map, at the cost of requiring insets for Alaska and Hawaii to be seen and not absurdly area-distorted by a mercator projection. The tilegram in 'b' on the other hand loses the geographic/spatial nature of the data, but comes at the benefit of making comparison between states easier and getting Alaska and Hawaii easily in the chart without area-distortion being a distraction.

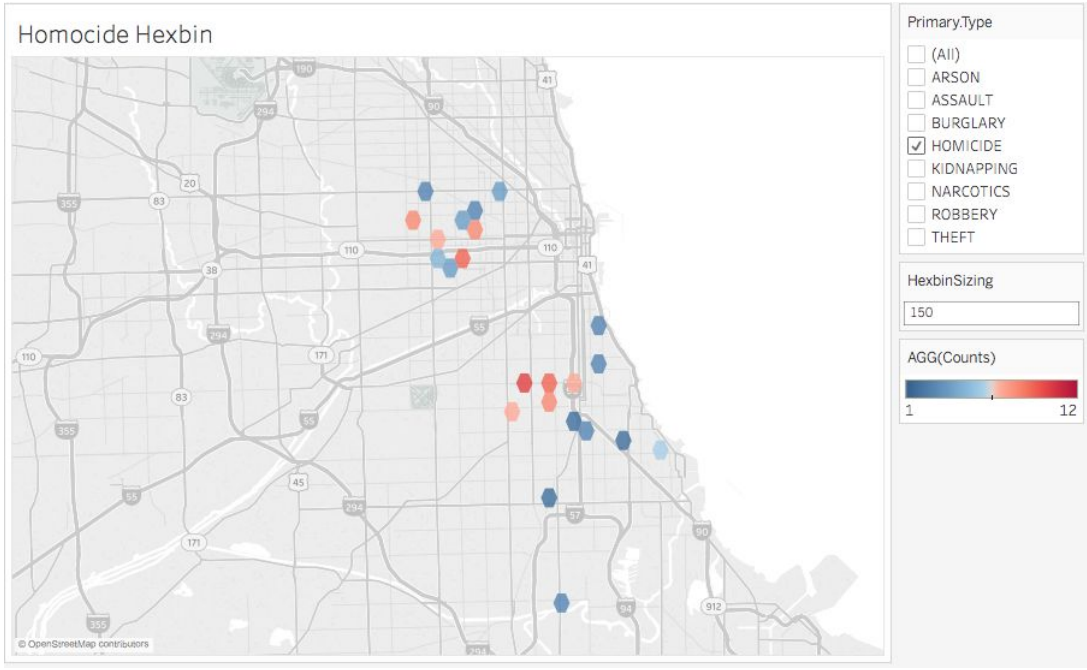
---

2) (20 pts) With the Chicago crime dataset from last week create the following graphs

a. Begin with your hex-bin plot from last week and add a map of Chicago underneath that is correctly scaled to the data.

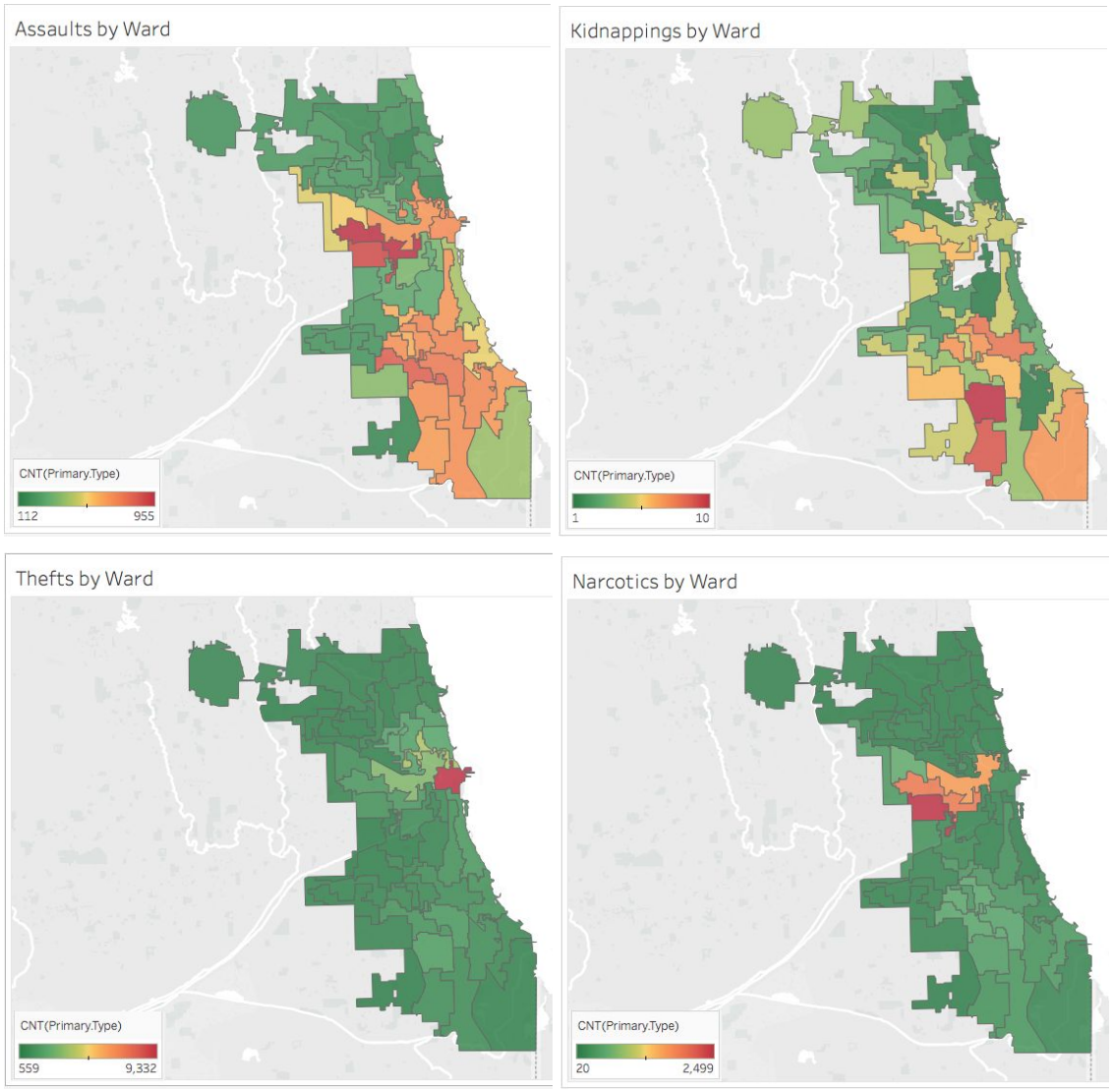


Explanation: For this hexbin I use a HexbinSizing parameter and auto-scaling with padding for 6-sided polygons. Setting the HexbinSizing parameter to 75, the plot shows areas of high homicide rate in red and low in blue. I chose this red-blue diverging color scheme to highlight those high homicide rates. Note that the hexbins are mapped to Chicago underneath.

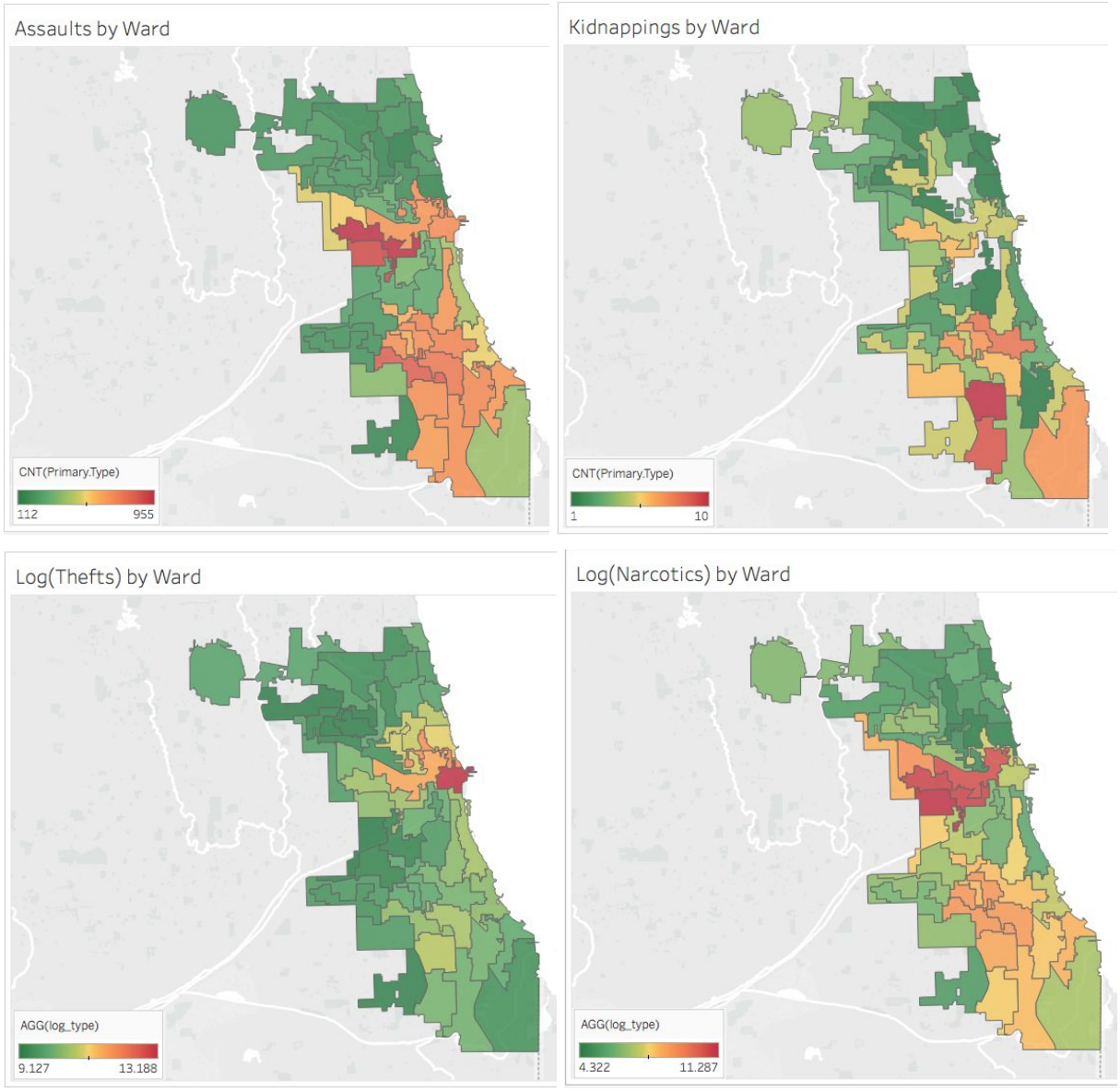


Explanation: This method also allows us to get as granular as needed. Here I doubled the HexbinSizing parameter to get more hexbins and therefore more specific. I personally prefer a slightly larger hexagon size (smaller HexbinSizing parameter) to see broader patterns but we can get down to more refined areas if necessary.

b. Create a small-multiples table of choropleths (like you did with the rose-plots in the last homework) with at least 4 different crime types. The color schemes for the four crimes should be unified.



Explanation: I present this first small-multiples plot for Assault, Kidnapping, Thefts, and Narcotics Incidents. There are some clear higher-crime rate wards on towards the city center and on the South and West sides. Looking at Theft and Narcotics incidents, we lose a lot of meaning in that sea of green due to the wide range in values. In order to fix this, I take the log of the count of Crime-Incident Type below.

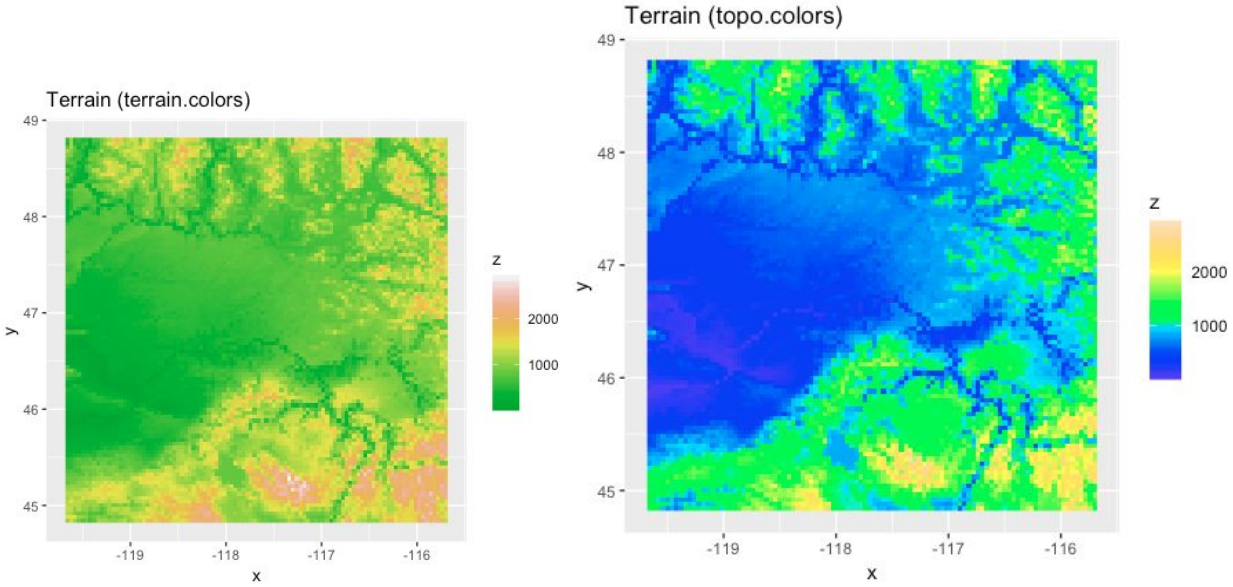


Explanation: Taking it a step further, I fixed the scales to better visualize the differences between wards. In order to do this, I took the log_2 of counts for Thefts and Narcotics Incidents. Now we can better visualize some of the other wards with moderately high crime rates. I did not fix the scales to set start-end values because it would cause incidents like Kindappings (which have very low incidence) to essentially disappear on the maps. What matters here is identifying those high crime wards for each type.

---

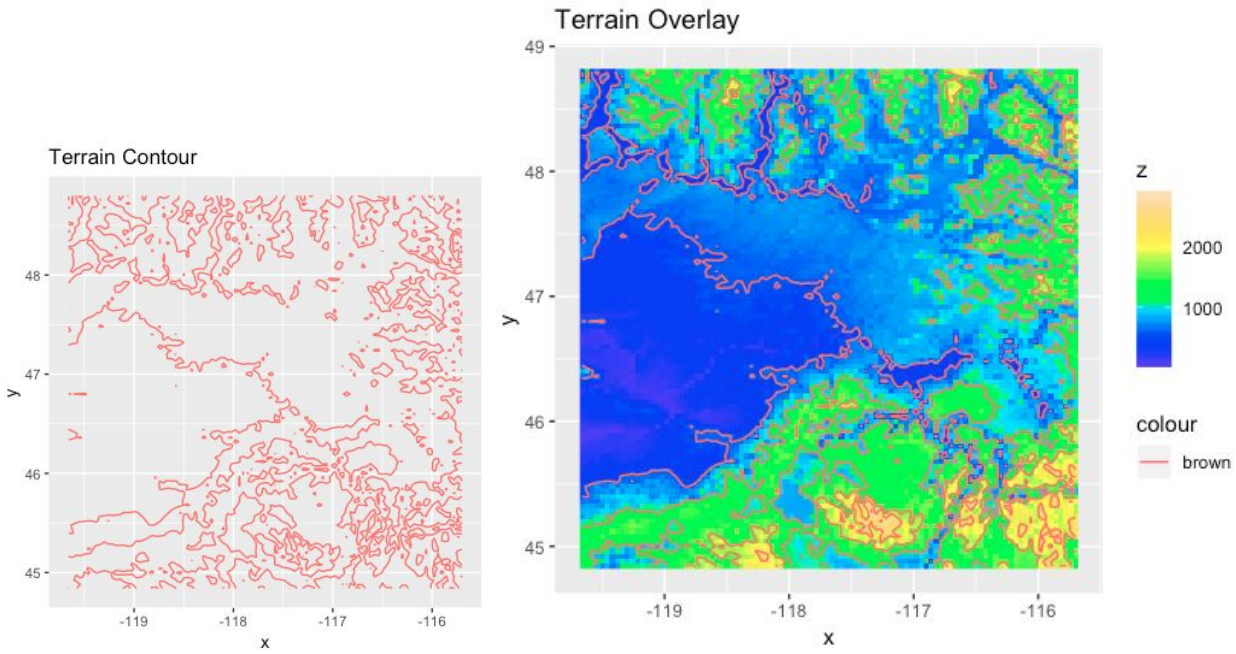
3) (20 pts) Download the terrain data from D2L (will be posted during week 9) and create the following maps of the data

a. A heat map of the terrain data mapped by color. You can use the "terrain colors" that are built into R that are good for such applications. The color map goes through blues to yellows to browns to whites for mountain tops.



Explanation: The heatmap on the left uses the "terrain colors" color palette, but that palette does not include the blues to yellows to browns to whites as described in the problem statement, so I made a second plot on the right using the topo.colors palette which more closely matches that description. This gives a reasonable visual of the terrain.

b. A contour plot of the same data. The contours should not be too dense to be unreadable, but they should be dense enough to be able to see the details and judge the steepness of the terrain. Overlay this plot over the heat map.



Explanation: On the left is the Terrain Contour map that was generated, and on the right is the contour map overlayed on top of the heat map. This the terrain boundaries without overly cluttering the plot.

c. Compare the two plots for what they communicate.

Explanation: The first plot in part ‘a’ shows the heatmap view of terrain with smooth transitions between layers. The second plot in part ‘b’ adds a countour overlay to help bring out the distinct terrain layers and give boundaries to the plot. The color choice for the overlay was a red-brown so that it could be easily seen over the topo-colors palette from the heatmap. The problem statement mentions steepness of terrain, but the overlay doesn’t add much in this respect. Overlaying with shadow might do more towards that end.