**Analysis and Conclusions for Further Paths of Inquiry**

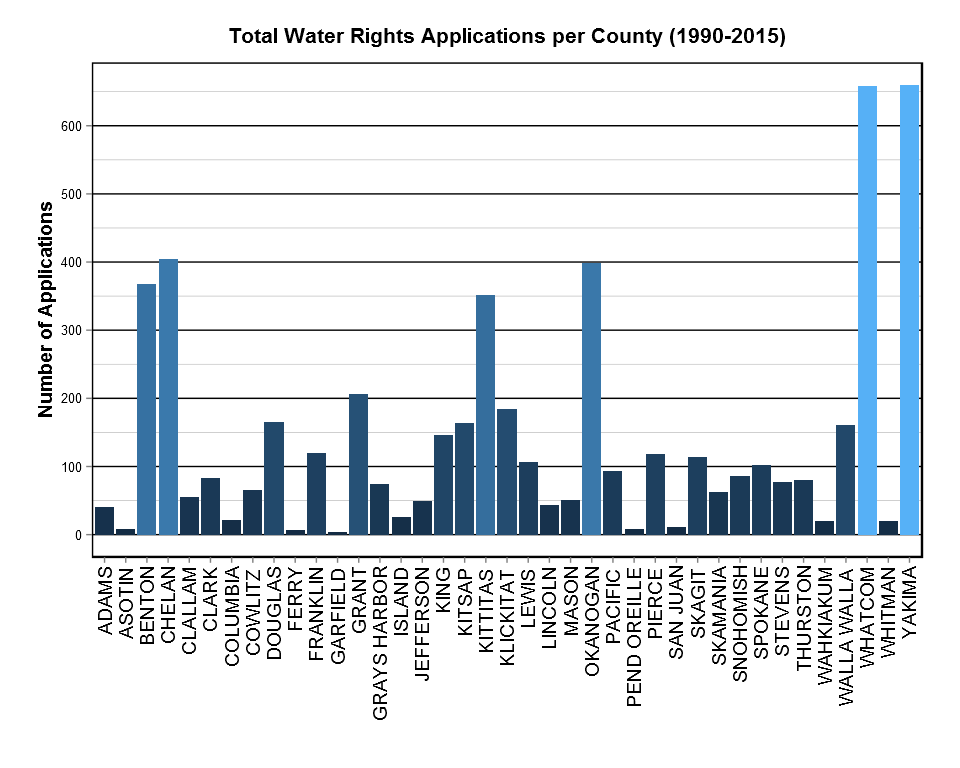
For this project I pretended that I was doing a large scale exploration of Water Right Applications in Washington State and what drives them. Water Rights are a major issue in the Western U.S., and a complete model of underlying inputs and hidden drivers of water right applications could be key to solving and/or preparing for future conflicts among a variety of users. The following analysis represents the initial exploration and analysis of historical water right applications as a means to begin building a more complete and robust model.

As a note, packages I used that I have no prior experience with include: ggplot2, ggthemes, reshape2, plyr, stats, and ggmap. The majority of my time on this project was spent learning the details of melting data (in some cases by hand) and manipulating the many feature of the plots in ggplot2. Specific decisions regarding how the packages were applied, and how the data was cleaned and manipulated and for what purpose, are included in the code of the script itself.

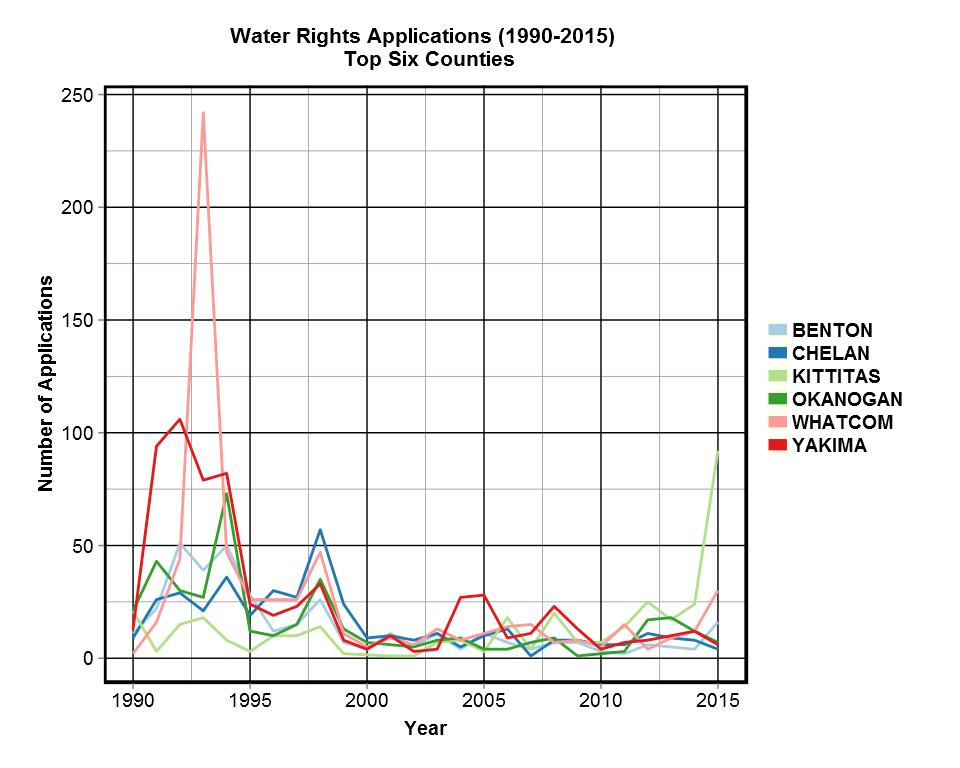
-------------------------------------------------------------------------------------------------------------------------------

My initial visualization was a simple bar graph to give me a feel for the total number of applications in each county since 1990. I decided to use 1990 as a cutoff because much open data in Washington State is complete from this date forward.

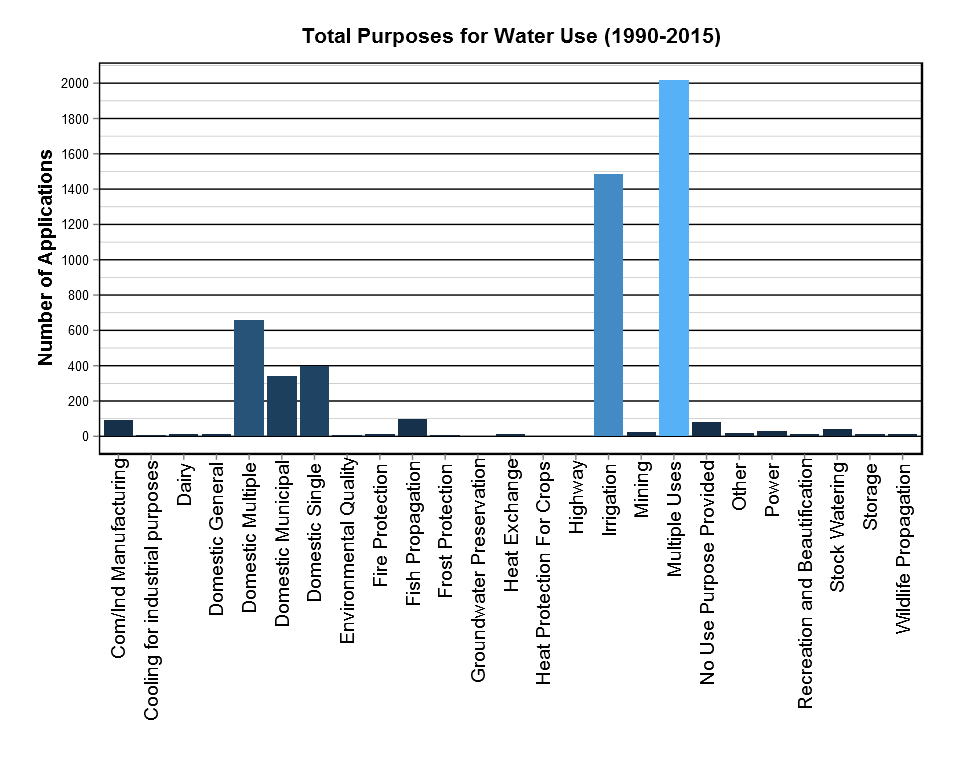
What immediately jumps out here is that there are six counties that have significantly more applications that the others: Benton, Chelan, Kittitas, Okanogan, Whatcom, and Yakima.



I wanted to get a better look at how the applications were distributed across the time frame, and I focused on the top six counties for visual clarity. We see a significant increase in the early 1990s in Whatcom and Yakima Counties. There is also a secondary bump in later 1996 for many of the counties. We also see a big climb in Kittitas County so far to date in 2015. This naturally begs the question: what was going on in those counties during these years that resulted in such big increases—particularly for Whatcom and Yakima in 1993?

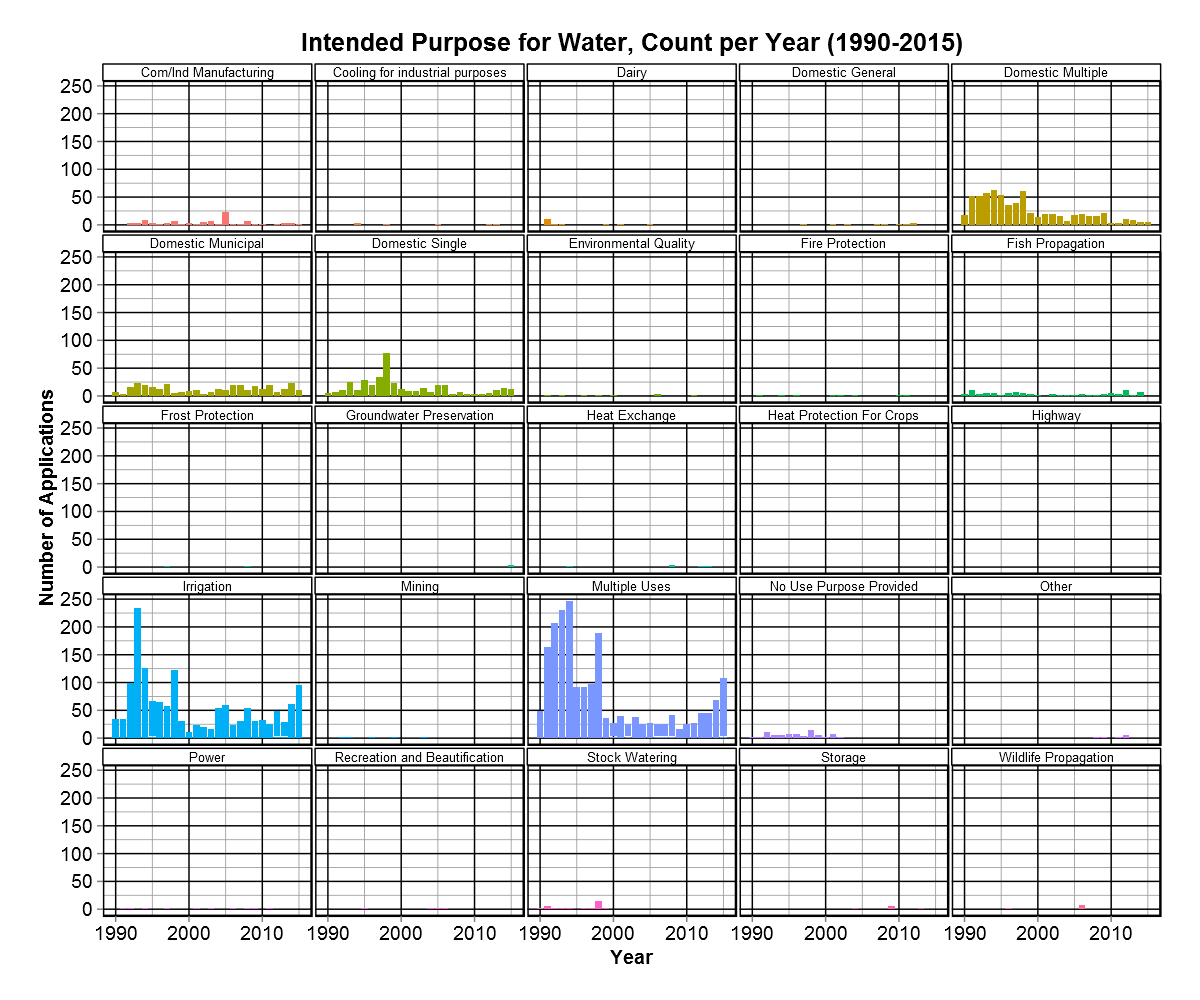


Before going to track down more data from different sources – such as population, agricultural acreage, mining expansion or other possible reasons to apply for a water right – I wanted to delve further into this resource and see if I could narrow down possible directions to go for future exploration. With this in mind, I wanted to see *why* the water right was being applied for. Here is the total number of applications for each purpose:

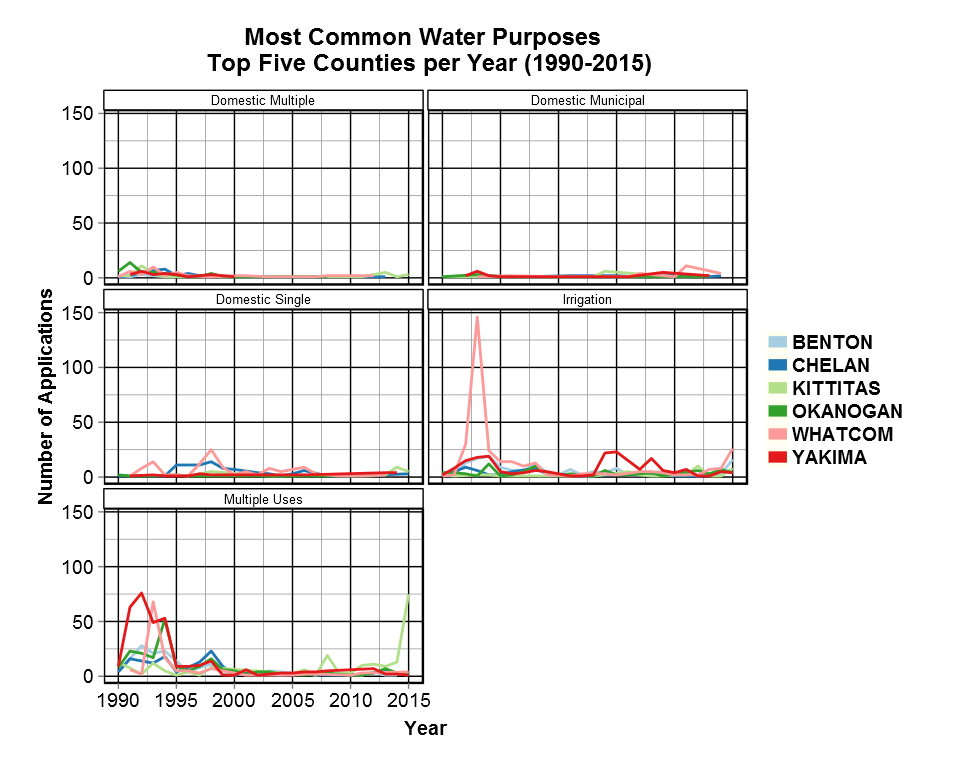


In the above plot we see that *Irrigation* and *Multiple Uses* are by far the most common reasons for a water right request. Though *Multiple Uses* isn’t particularly useful because of lack of detail, *Irrigation* does suggest agricultural reasons. Were I to expand on this project I would begin to pull data from the USDA Census of Agriculture to examine popular crops, particularly in Whatcom and Yakima Counties.

The plot also shows very clearly a “second-tier” of purposes (*Domestic Multiple, Domestic Municipal, Domestic Single*), all of which are related to the construction of housing. The next step for this piece of the puzzle might be to access both population and housing data from the general US census. I also found it helpful to break down the number of applications for each purpose by year.

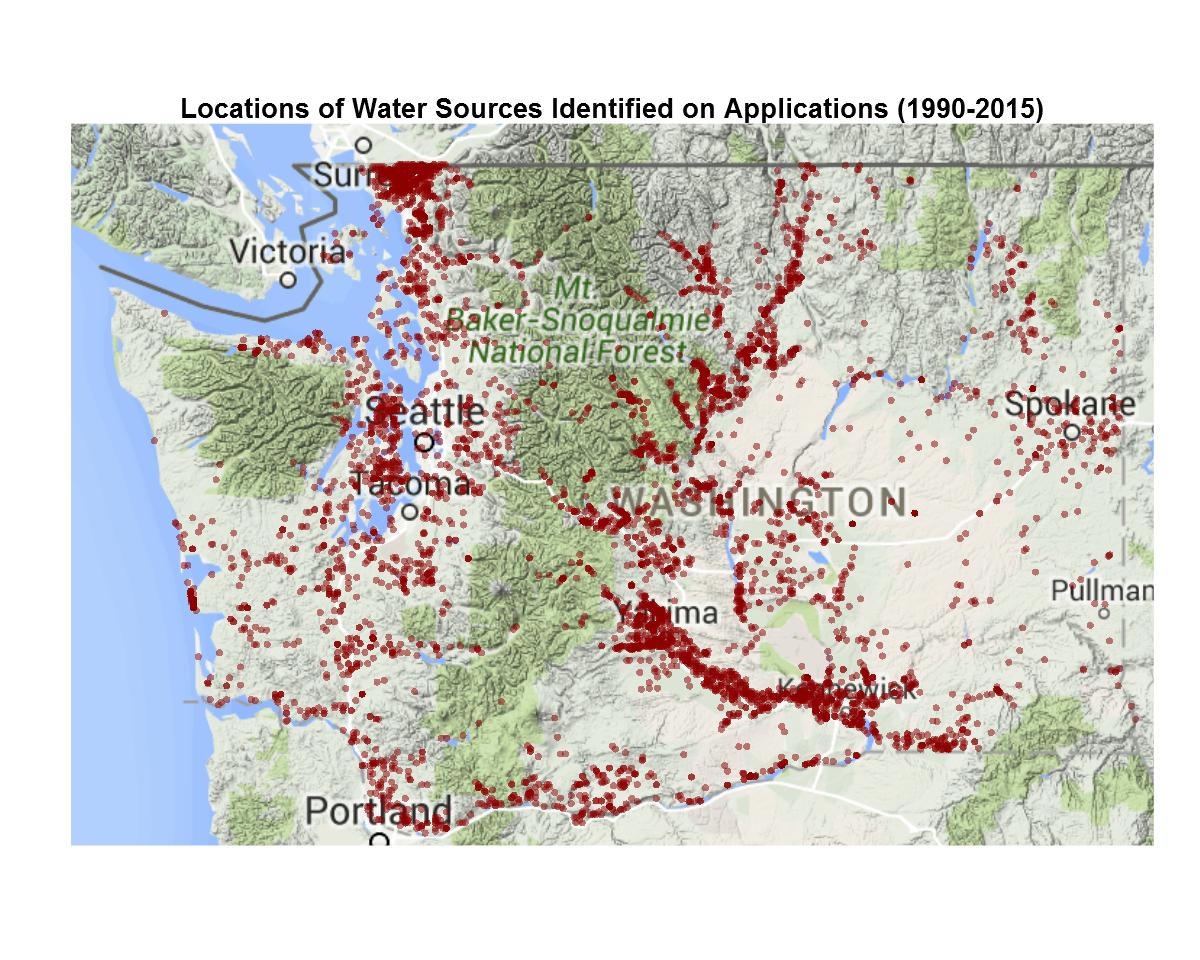


This again indicates that there is a lot of activity in the 1990s across all counties, particularly for *Irrigation* and *Multiple Uses.* But I really want to see if these spikes account for the spikes we saw in the overall number of applications for those top six counties from earlier. Here is a graph which compares the five most common reasons for an application for the six counties with the most total applications, as compared over time:

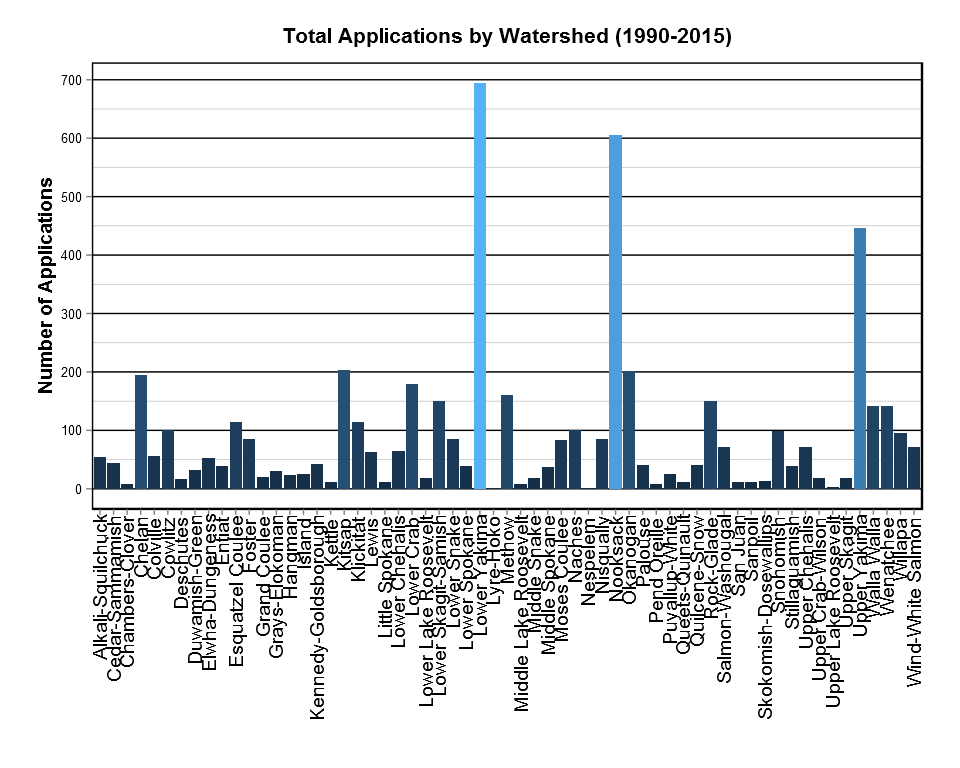


The above plot seems to indicate that we would should be looking at agricultural data for Whatcom County in the early 1990s, but might need to go back into the original data set to determine the individual drivers for the spike in Yakima County since *Multiple Uses* is not very specific. Admittedly, before I broke the data down further, I would have bet that the applications in Yakima, Okanogan and Chelan Counties were primarily for agricultural applications for purposes of growing hops and wine grapes, as in general these have been burgeoning sectors over the last several decades. This is a very good example of early data analysis saving too much work in the wrong direction for latter parts of a project.

Comfortable with the fact that I would need pull in agricultural, housing and population data for future work on this issue, I decided to switch gears to a more geological/geographic analysis. First, I wanted to see a straightforward projection of where the locations of the sources for the water right applications. Here is a Google Map upon which the source for every application has been plotted:



The geographic distribution is unsurprising, as we see a concentration in Whatcom County on the north end of Puget Sound, and along the Snake River going into Yakima County. Okanogan County contains most of the applications stretching northward in the middle of the state. Because many of the source locations lie obviously on rivers and their tributaries, I decided that my next plot should focus on the watersheds in which the sources lie. Again, I started with a bar graph showing the totals for each watershed.



This graph suggests that the next step might be looking into the overall volumes of water available in each watershed compared to the draws upon the watershed at the sources. In fact, this summer multiple sources went dry due to drought, and this could be an important component of an overall water right analysis.

***Final Statement***

It is at this point that I ran out of time for further inquiry. Again, my goal for this project was to build visualization skills for large datasets. I also wanted to push my ability to use visualization to ***guide further study,*** rather than as a fancy final product. If this really was the initial stage of the large-scale project, I would assume that I would probably want to look at similar graphs for most of the fields in the Water Right Applications dataset.