At this point your group should have completed your inferential model of your data. As we transition to the final scaffold activity your task is to make meaning OF YOUR MODEL.

So while there are many interesting aspects about your data, focus your attention on being able to visualize your model.

In your group articulate

**1. What is the main conclusion from our model in relation to our variables?**

[**https://colab.research.google.com/drive/1R2UD0Jszwqn\_vXB-sDRo\_IPmttEPEpJ2#scrollTo=LVdVp43F8gm\_**](https://colab.research.google.com/drive/1R2UD0Jszwqn_vXB-sDRo_IPmttEPEpJ2#scrollTo=LVdVp43F8gm_)

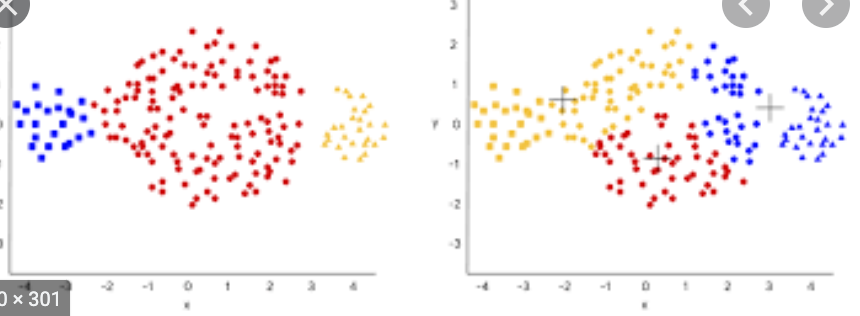
Looking at the table which shows cities and the clusters they are in for all three years, overall we noticed that from 1975 to 1995 either cities remained constant (low to low or mid to mid like Los Angeles) or they decreased in homicides per 100k (Cleveland for example). For 1995 to 2014 we noticed that most cities, even those that saw an uptake in homicide rates from 1975 to 1995, shifted and decreased (Salt Lake City). Many of the cities have changed from a higher crime group to a lower crime group over the three periods, 1975, 1995, and 2014. While there are certainly outliers that saw an increase, New Orleans for example, the majority did not.

2. **What would be an appropriate way to convey this information through a written description?** **(this involves "translating" statistical results into everyday language. For example, if there is a significant difference in the number of times Newfoundland dogs eat off the counter and the number of times a poodle eats off the counter, I could say, the t test was significant at .05 and I reject the null hypothesis. But that means nothing to most people. What I might want to write is something like, my analysis found that Newfoundlands eat off the counter much more, and to a statistically significant amount more, than poodles. The mean number of times for newfoundlands was 10 compared to the mean number of times for poodles of 5.**

Many of the cities have changed from a higher crime group to a lower crime group over the three periods, 1975, 1995, and 2014. While there are certainly outliers that saw an increase, New Orleans for example, the majority did not.

**3. What kind of graph would illustrate what you found in your MODEL (again, limit this to your model)**

(not a graph) but a map visualization could be effective in communicating our data



These graphs are useful because they can give a quick understanding of the changes between the two groups. For instance, in graph 1, blue is quite small; this indicates there are less cities in the blue cluster, but in graph 2, the blue fills up more of the graph indicating there are more cities in cluster blue. Each color represents a different cluster (ie. low, mid, and high crime cluster groups) and we would have a legend showing which cluster is which color (for example. blue = low, red = high, and yellow = mid).

4.  **This is really important. Examine your responses and review them to make sure you are not drawing conclusions that extend beyond your data and your results. In doing so, consolidate the decisions you have been making about your project throughout the semester. Be aware of using technical terms inappropriately or making broad statements that are not supported by your data!**

We think that our conclusions do match with our data. We will make sure to keep an eye out for missued technical terms in our conclusions and future visualizations.

One potential error in our conclusions would be if we said that we fully examined data for a period of 40 years (like it appears in our original data). For our project we analyzed cities in the years: 1975, 1995, and 2014. This is because for 2015 (our 40th year) we had too many missing values and decided to analyze 2014 instead.

We should be mindful that we excluded data that was not reported over a 12 month period. We did this to standardize our data yet we now cannot claim that we’ve analyzed *all* data from major cities.

Due to time, we decided to only analyze three years instead of 40 years.

We chose to work with homicides per 100k instead of total homicide as when we tested the magnitude of r-correlation on both groups, it was higher for homicides per 100k compared to the sum of homicides.