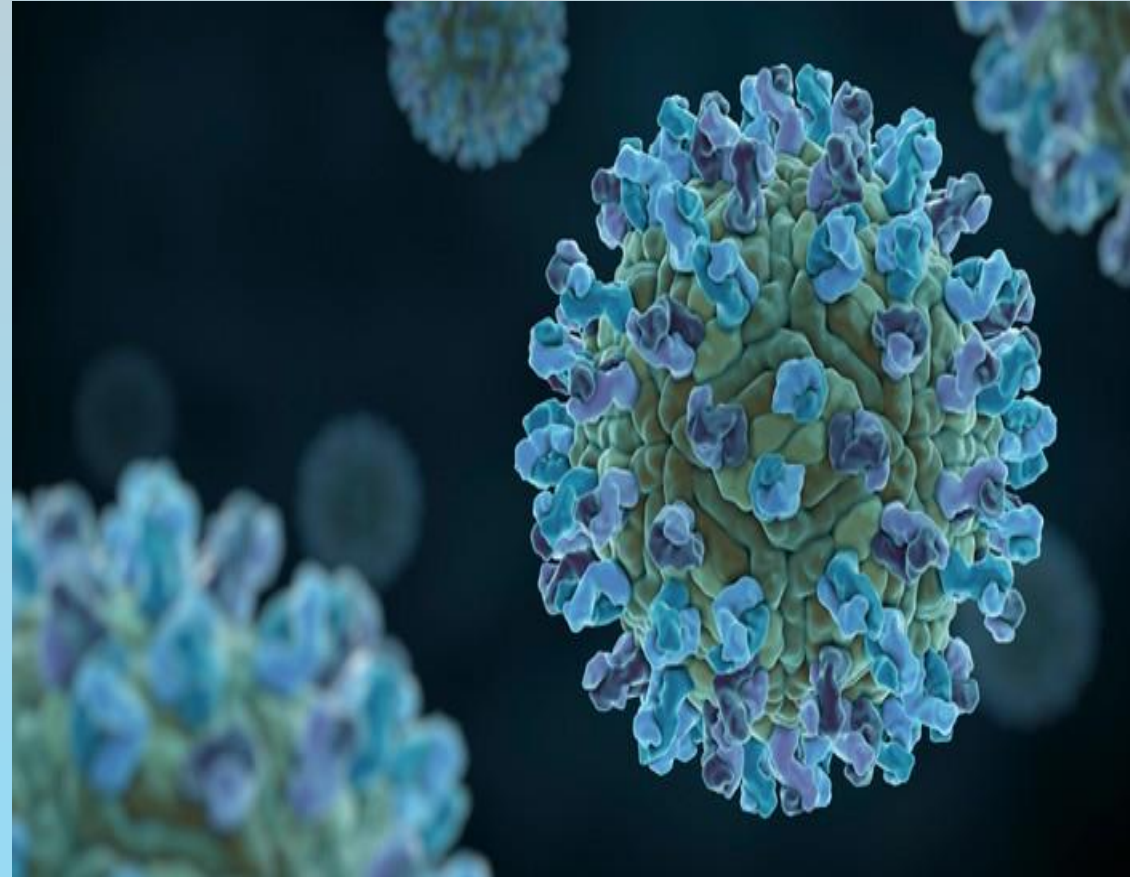


# DATA VIZUALIATION

## West Nile Virus Project

<https://public.tableau.com/profile/ashish.gandhi#!/vizhome/West-Nile-Virusproject/Dashboard1?publish=yes>



# PROJECT OVERVIEW

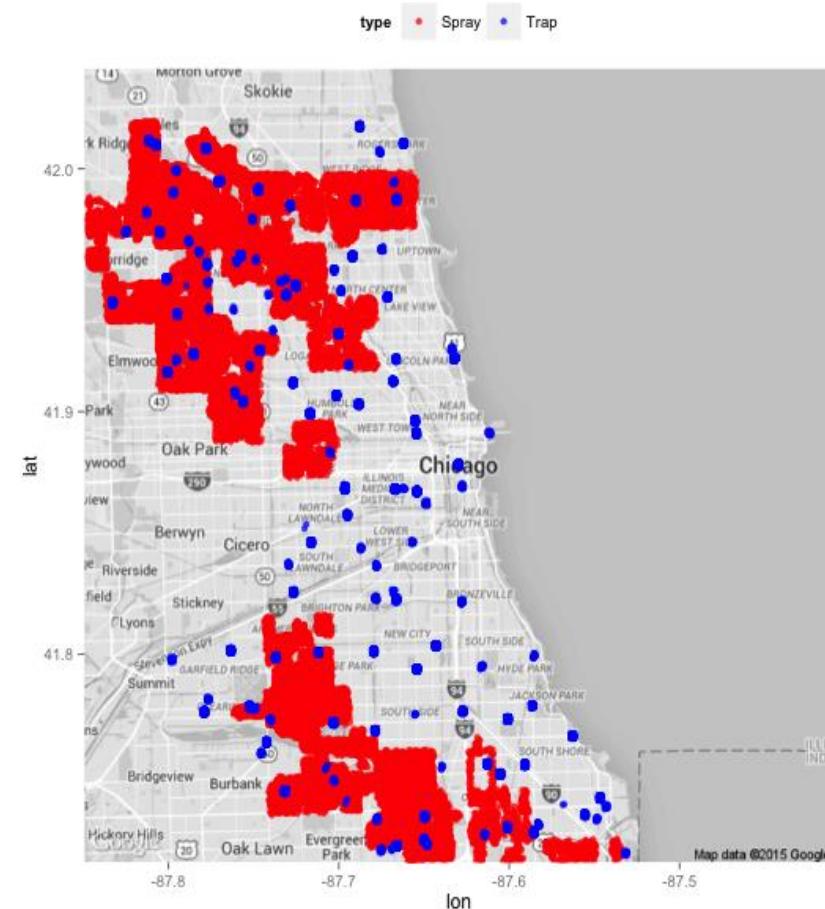
In 2002, the first human cases of West Nile virus were reported in Chicago. By 2004 the City of Chicago and the Chicago Department of Public Health (CDPH) had established a comprehensive surveillance and control program that is still in effect today.

Every week from late spring through the fall, mosquitos in traps across the city are tested for the virus. The results of these tests influence when and where the city will spray airborne pesticides to control adult mosquito populations.

Given weather, location, testing, and spraying data, this competition asks you to predict when and where different species of mosquitos will test positive for West Nile virus. A more accurate method of predicting outbreaks of West Nile virus in mosquitos will help the City of Chicago and CPHD more efficiently and effectively allocate resources towards preventing transmission of this potentially deadly virus.

In this competition, you will be analyzing weather data and GIS data and predicting whether or not West Nile virus is present, for a given time, location, and species.

All locations and all sprays in 2007-2014



# ANALYSIS APPROACH INSIGHTS

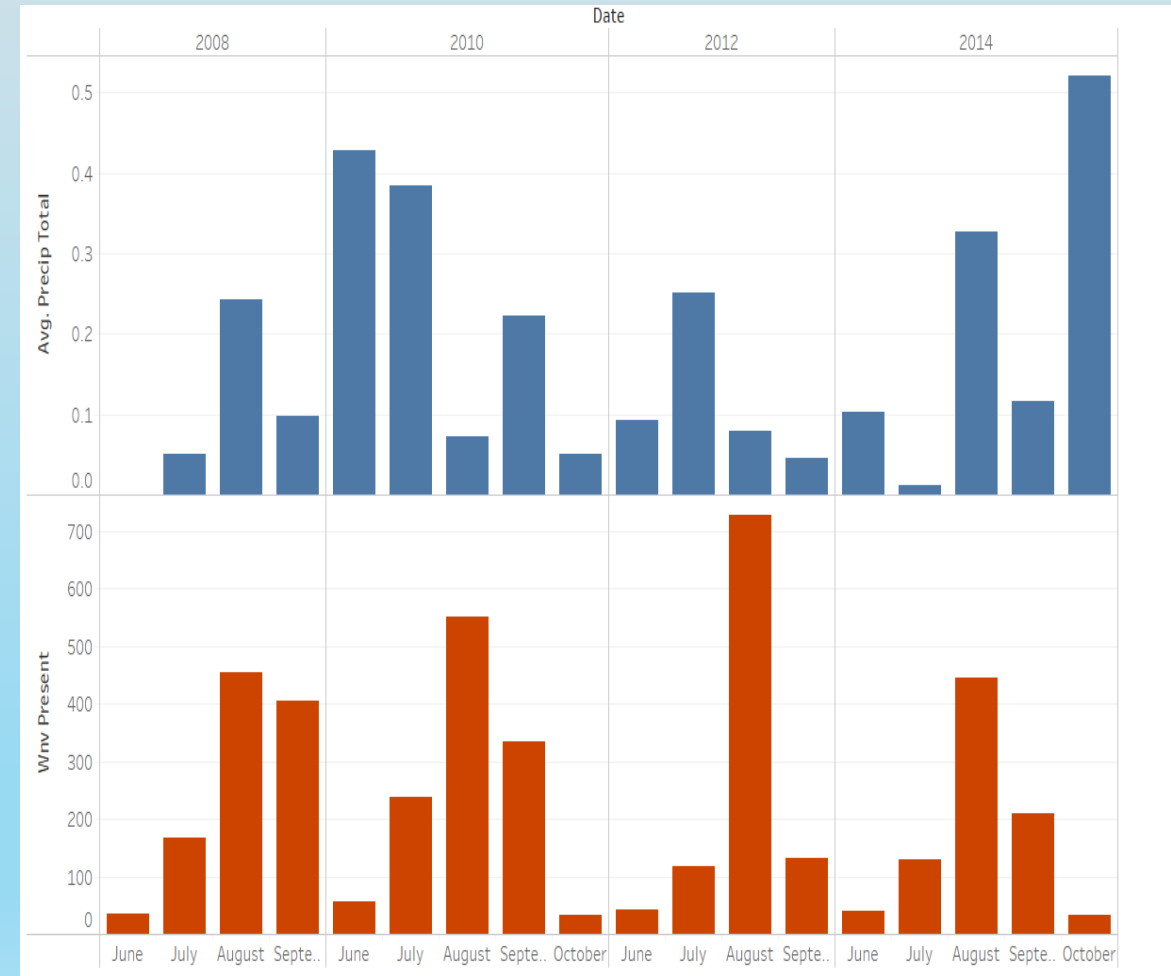
With the help of this insight, we tried to determine if the avg. precipitation rate has anything to do with the outgoing West Nile Virus.

**For this insight, we used the weather and test.csv file.**

**Insight – We came to the conclusion with this insight that Avg. Precipitation rate does not affect the West- Nile-Virus as much as it may seem.**

**In 2010, the avg. precipitation rate is highest in June, but the most number of positive WNV cases were found in August despite having the avg. precip. Rate less than rest of the months in the same year.**

**Same is the case with rest of years where Avg. Precip. Rate does not seem to affect the Wnv Rate.**



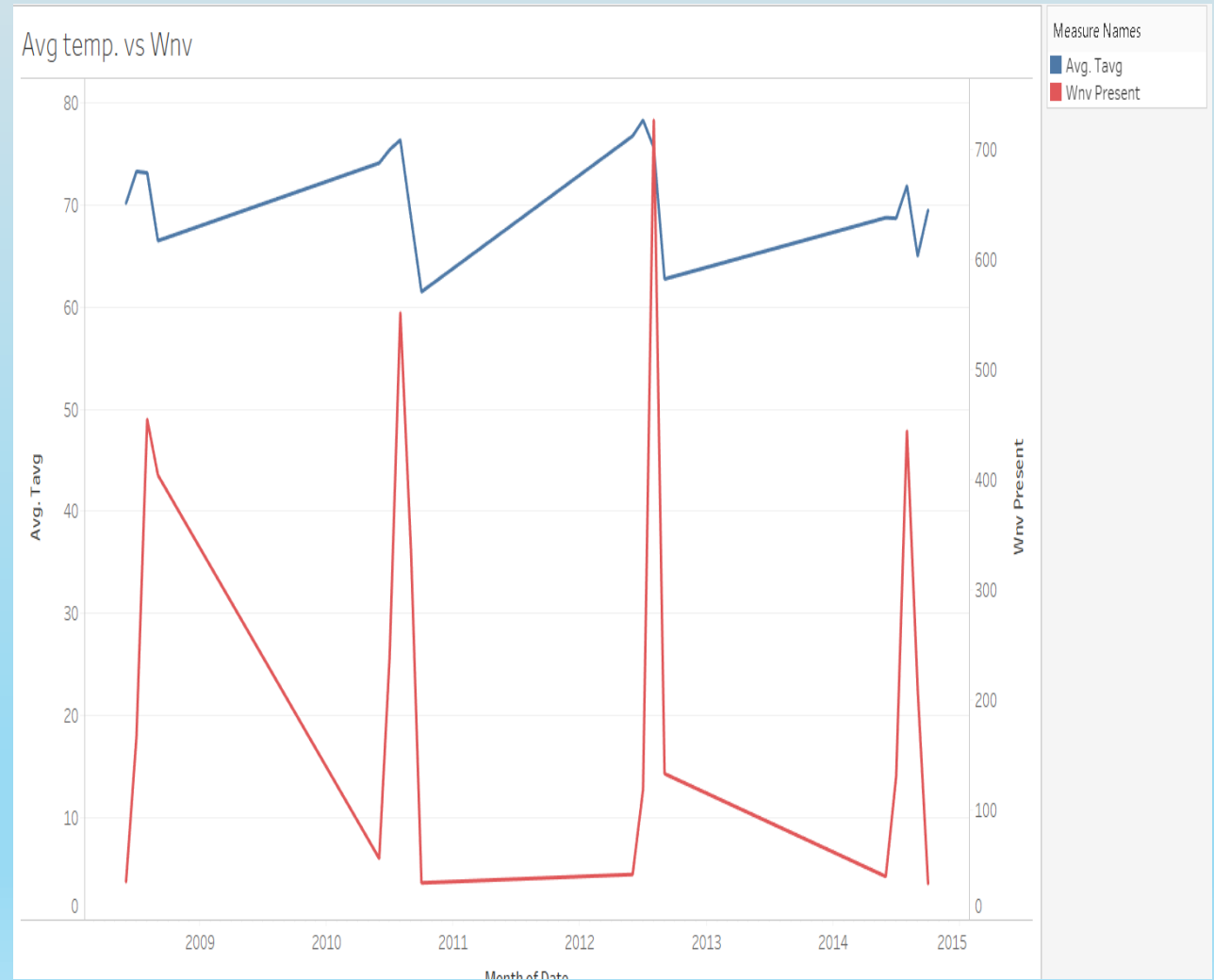
# ANALYSIS APPROACH INSIGHTS

Our second observation helps us analyze the relationship between Avg Temp. and Wnv.

Since we are using the Date type variable to analyze this observation, Line chart is considered as the best approach to do so.

**Insight –** With the help of this observation, we can easily say that Avg. temp has an impact over the Wnv. You can see that the Wnv variable goes up to maximum when the Avg. Temp is high. In 2010, the avg. Temp is highest in August and likewise the number of Wnv is highest as well in that case.

Sometimes the Avg. Temp is not the highest for the highest Wnv Value, but close to the highest like in 2012.





# ANALYSIS APPROACH INSIGHTS

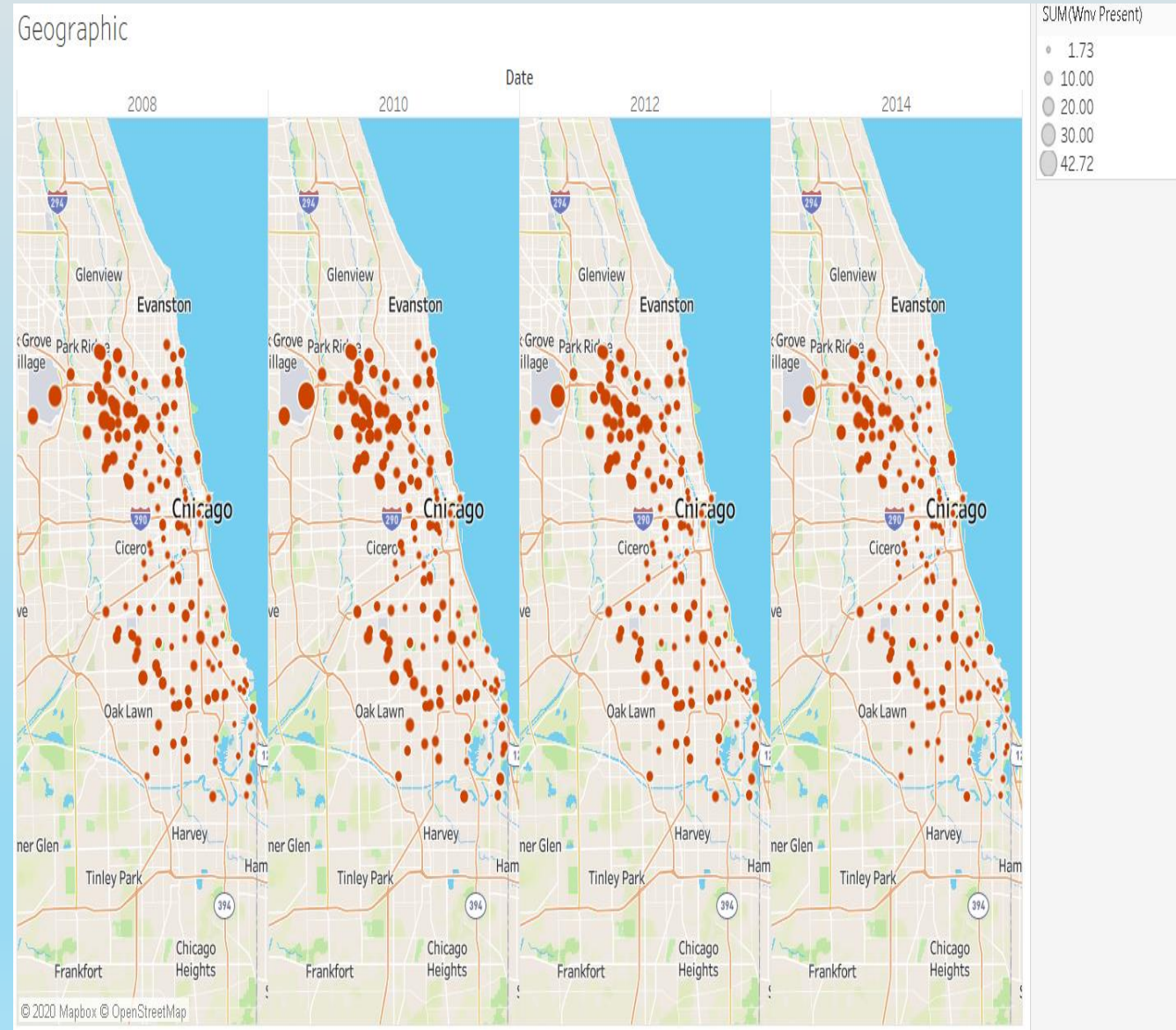
For this insight we used the longitude and Latitude dimension present in the test.csv file to plot this symbol map.

We used the discrete time-series data with Wnv count foreach location to determine which location saw the most cases.

**Insight:** In this insight we analyzed which year saw the most cases in which location. The size of the bubble determines the number of positive cases in that location.

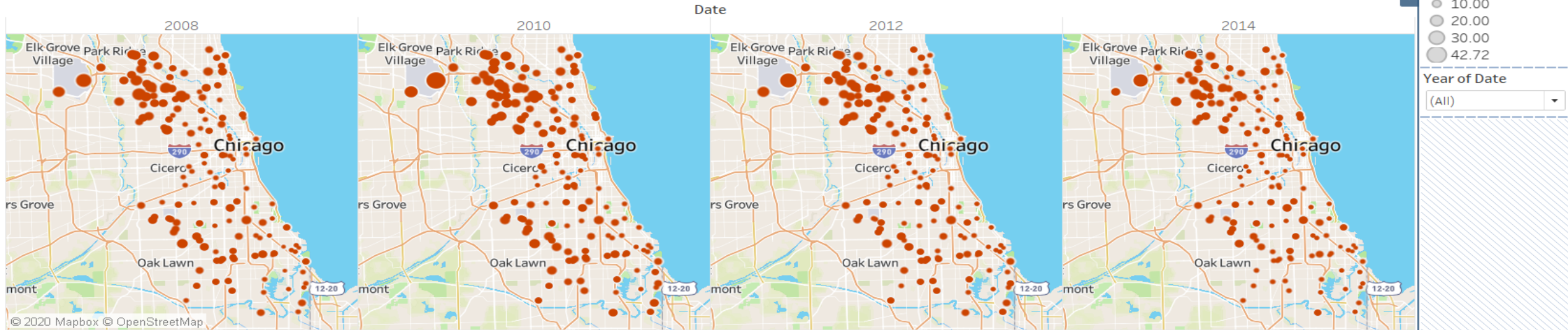
On inspecting the data, we found that 2010 and 2012 witnessed more cases than 2008 and 2014. It is also seen that the north region is the hotspot for the virus in each year .

In 2014, we could see a small dip in the Wnv cases from the previous two years in the picture.

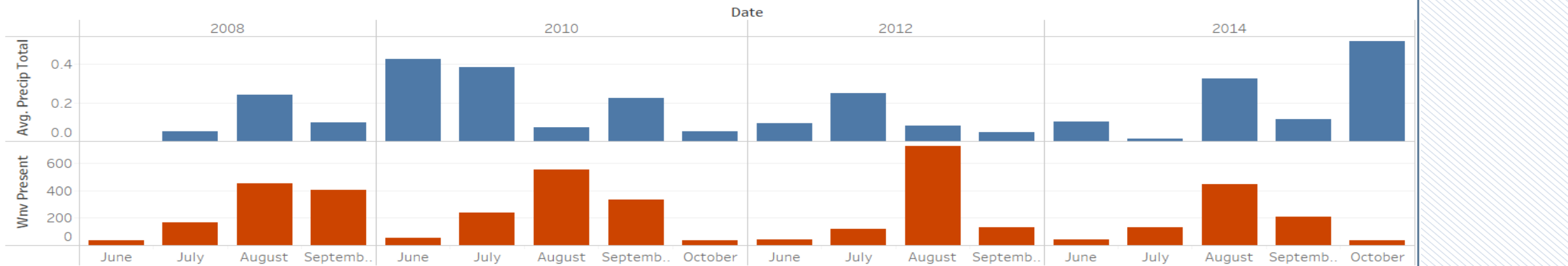


# DASHBOARD #1

## Geographic

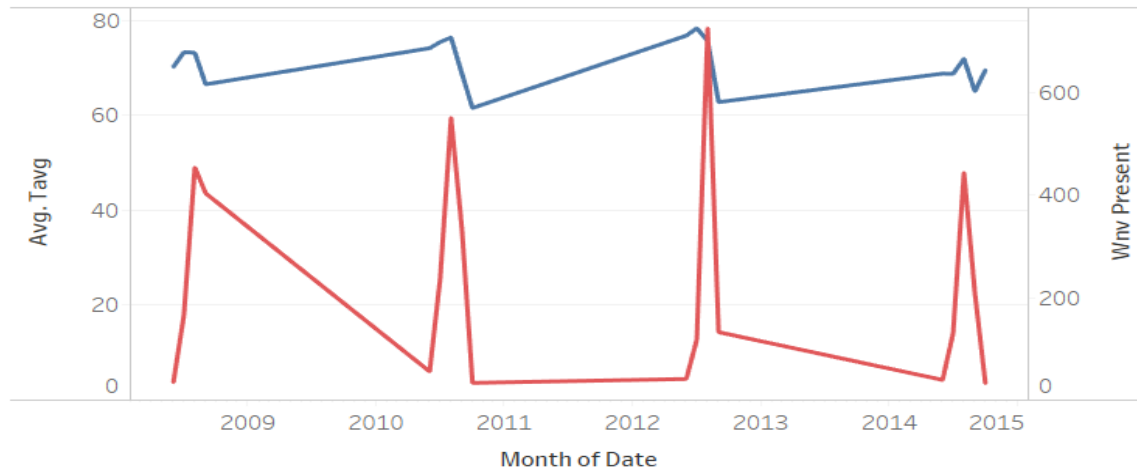


## Avg Precip. vs Wnv

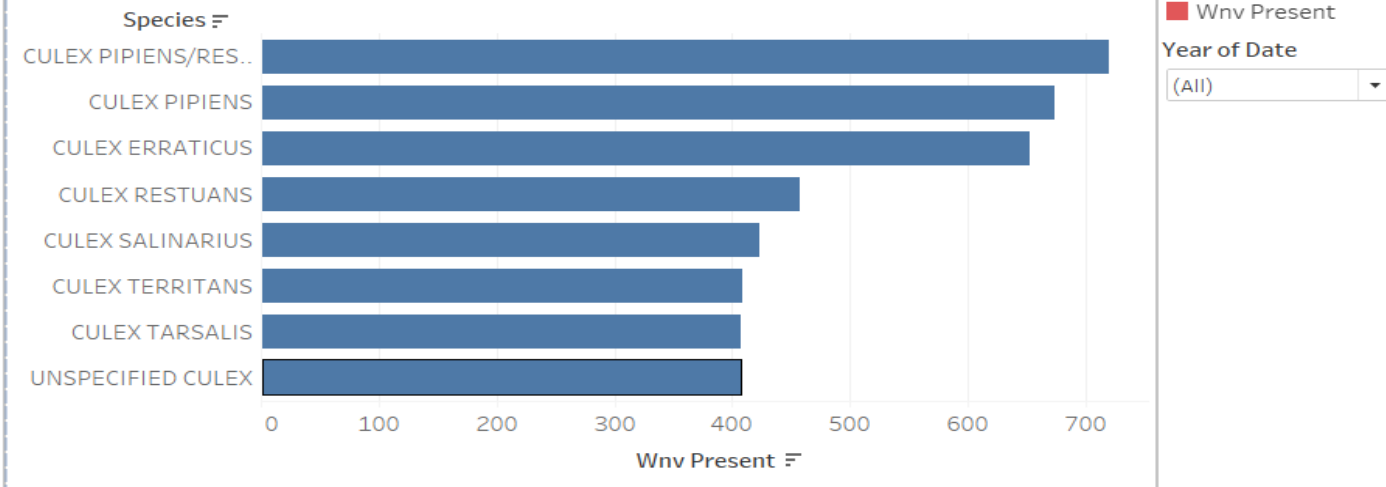


# DASHBOARD #2

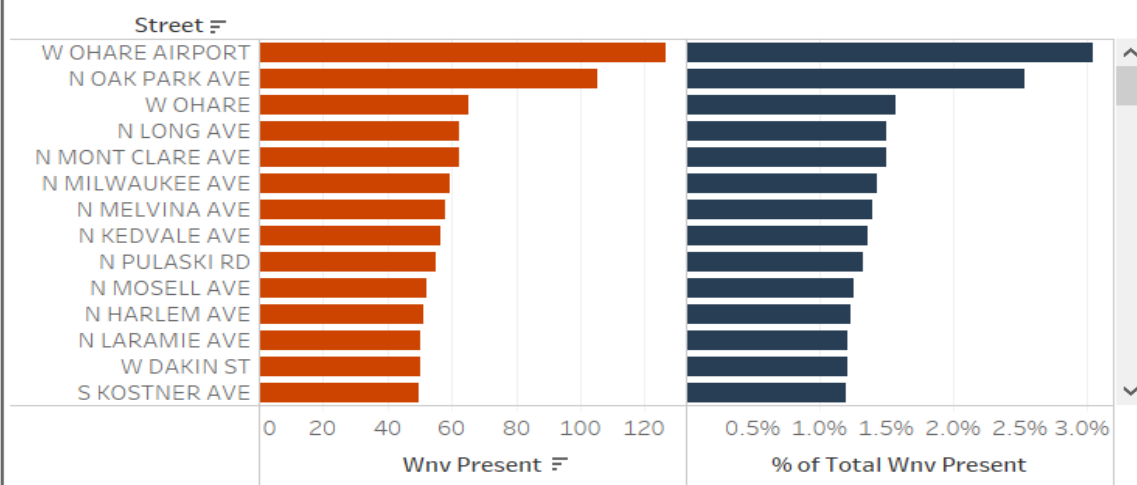
Avg temp. vs Wnv



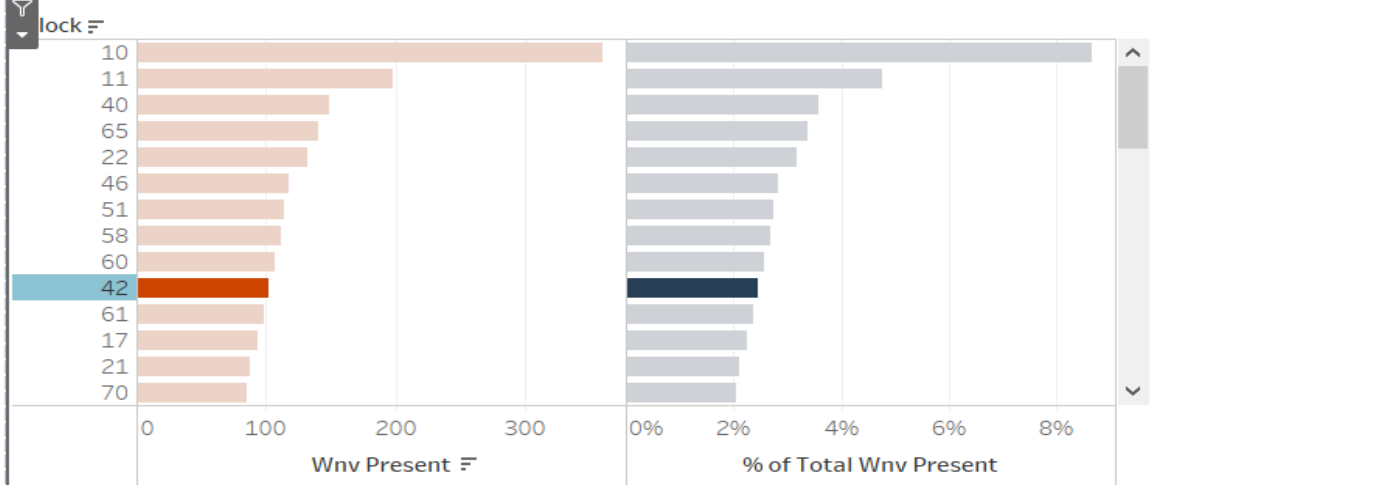
Species



Streets



Blocks





**THANK YOU**