**Volunteer Nurses in North Carolina during COVID-19**

Feiya Suo

Client: Jill Forcina

Date: March 2022

**Executive Summary**

Health care workers are in huge demand during COVID-19, and there was a huge number of volunteers participated in to help. To collect information about available volunteer nurses in each county of North Carolina, we designed an online questionnaire survey about the willingness of volunteer nursing during the pandemic. By using quantitative data from the survey with 3,089 registered nurses in NC, we describe the geographic distribution of available and volunteered nurses across states, and analyze the correlation between available volunteer nurses and COVID-19 vaccination coverage rate. We found that urban states have more nurses who are willing to volunteer and who actually volunteered. The percentages of available and volunteered nurses over county population are also high in those urban counties; but it is also the highest in some rural counties with small population.

In the correlation analysis, we found that the numbers of available and volunteered nurses are significantly correlated with vaccination rate. The proportion of volunteered nurses over the number of available nurses does not have statistically significant correlation with vaccination coverage rate. After aggregating the data into region level, we found the same trend.

**Introduction**

During the COVID-19 pandemic, there is a shortage of health care workers because of the rapidly increasing demand. Considering the negative effect of worker shortage, it is important to determine the problems in supply, distribution, and retention of health care professionals, and facilitate the process. The North Carolina aims at providing and supporting educational activities and services inside the state, particularly focusing on primary care in rural communities and those with less access to resources to recruit, train, and retain the workforce.

AHEC worked with the NC Board of Nursing to design and distribute a questionnaire about nurses’ willingness of volunteer work during the pandemic. By collecting the survey data, we have better understand of how many volunteer nurses are available in each county, and how many of them actually volunteered. Combined with county-level data on vaccination, we study the variation in available nurses, and its correlation with vaccination coverage rates.

**Data**

We designed two rounds of questionnaire surveys to ask about the volunteer information in North Carolina during the COVID-19 pandemic. In the first survey, respondents were invited to participate in the voluntary survey when they accessed the website of NC Board of Nursing. If they chose the option that they are willing to be a volunteer nurse, they would be redirected to the survey, asking for their primary county, other countries of volunteer willingness, nurse license type, and employment status. We also recorded their personal contact information and the county health care provide may contact them for volunteer work. We received 3,089 responses in the first round.

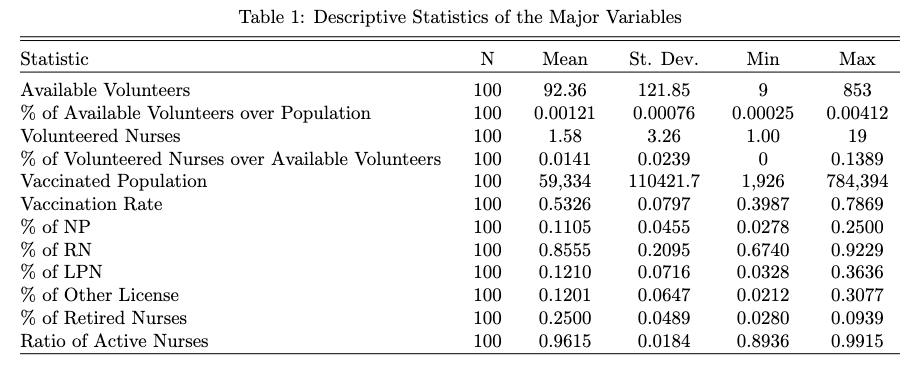
In the second survey, we ask if the respondent actually volunteered since the first survey. If the answer is yes, we collect the county where they volunteered. We collected 562 responses in the second round. The two surveys are separate and cannot be connected with each other.

We also collected information on country-level population, the number of people who are fully vaccinated, and the vaccination coverage rates. We first transferred the survey data from individual level to county level, and then merge the data together to generate the final dataset for future analysis. The final data has 100 rows and each represents a county inside North Carolina.

**Analysis**

1. Data visualization

There are many variables of interest in this project. Table 1 summarizes the descriptive statistics of the major variable.

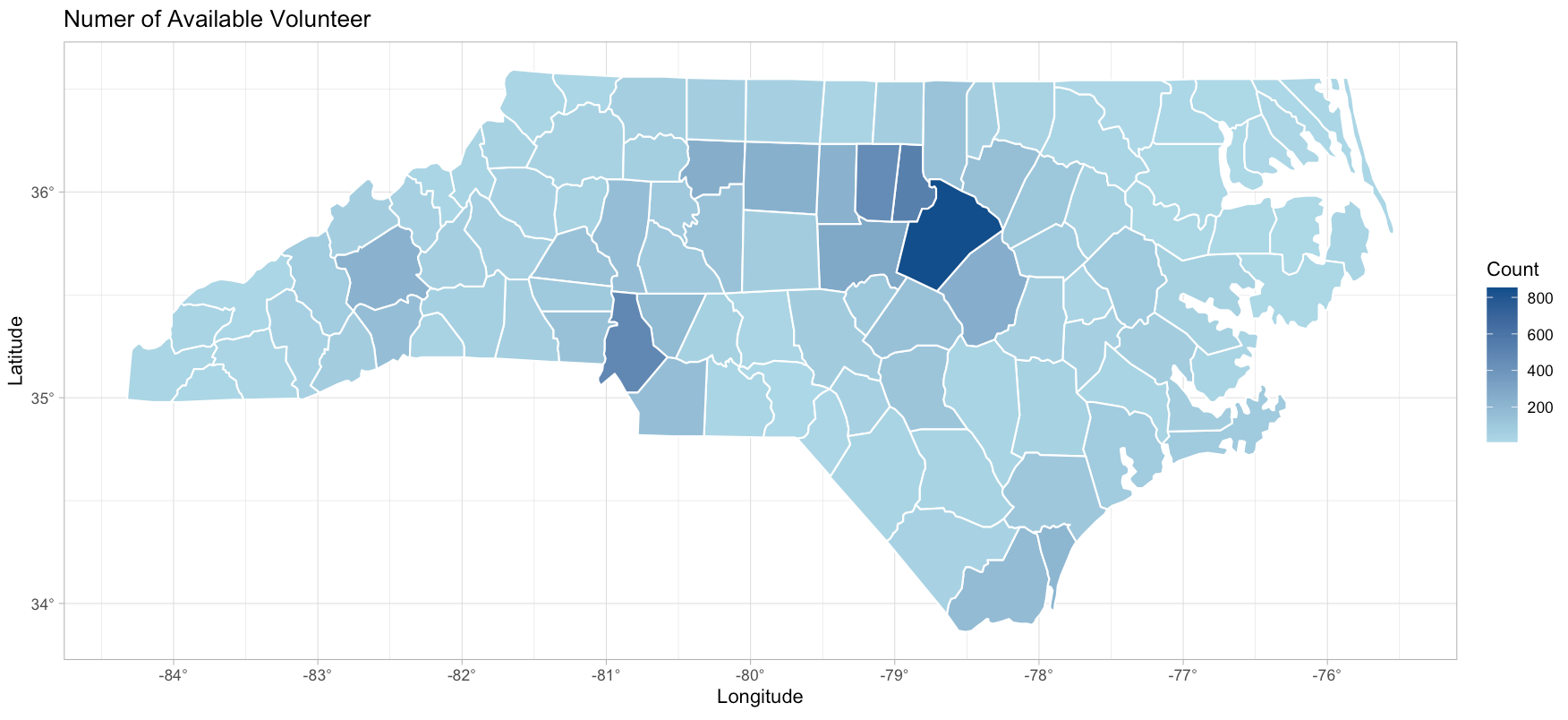


We first use the Shiny app in R to visualize the survey data at country level. The create interactive map application can help the client to get the map visualization easily; they can choose the metric of interest from the interactive menu, and also obtain the detailed county level data by clicking the county in the map.

First of all, we show the number of available volunteers in each county. Figure 1 show the number of available nurse and Figure 2 shows the percentage based on the county population. In the two figures, darker blue represents higher counts/percentages of available nurses during the pandemic.

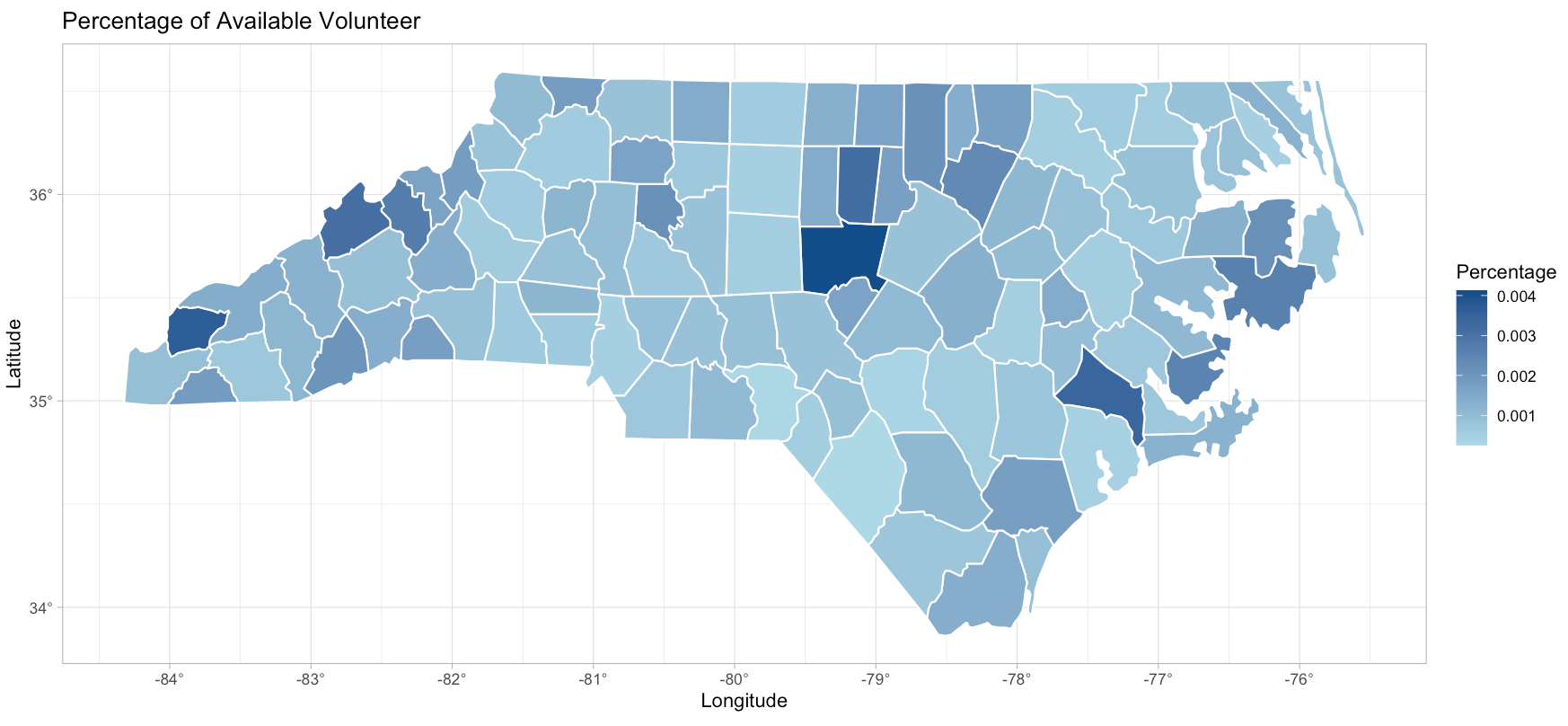
From Figure 1, the highest count is in Wake County. The nearby counties, Durham and orange, also have a higher number of available nurses. The other high ones are more urban[[1]](#footnote-1) and populous including Mecklenberg and Buncombe. This is consistent with our hypothesis that more urban counties have more available nurses.

Figure 1: Number of Available Nurse Volunteer, County Level



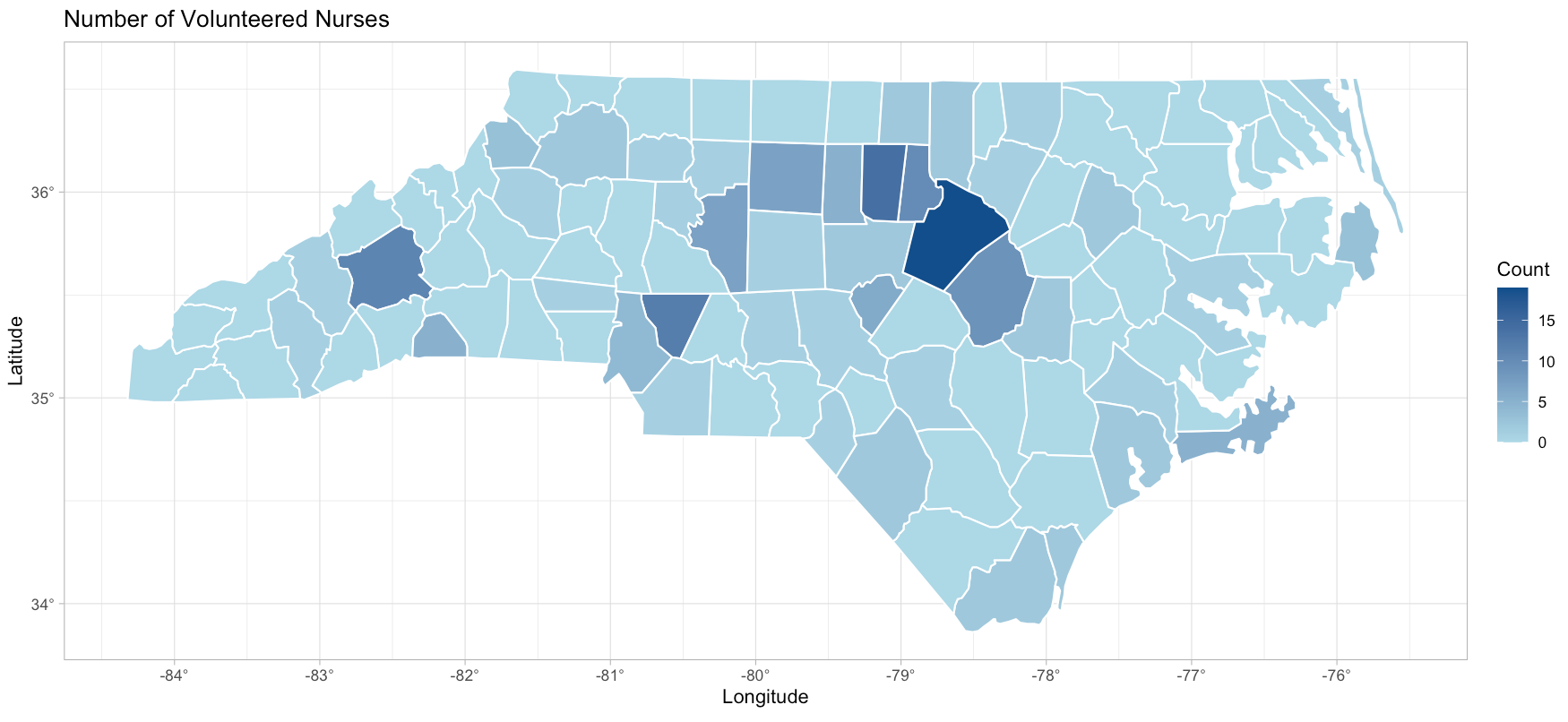
In Figure 2, the counties with the higher percentage are different. Although some urban counties, such as Durham and Orange, still have high percentage of available volunteers, some rural counties have a higher percentage. For example, Chatham, Graham, Madison, Yancey, and Jones have more available nurse based on county population. They all rank relatively low in state GDP, but also have smaller population size.

Figure 2: Percentage of Available Nurse Volunteer on Population, County Level



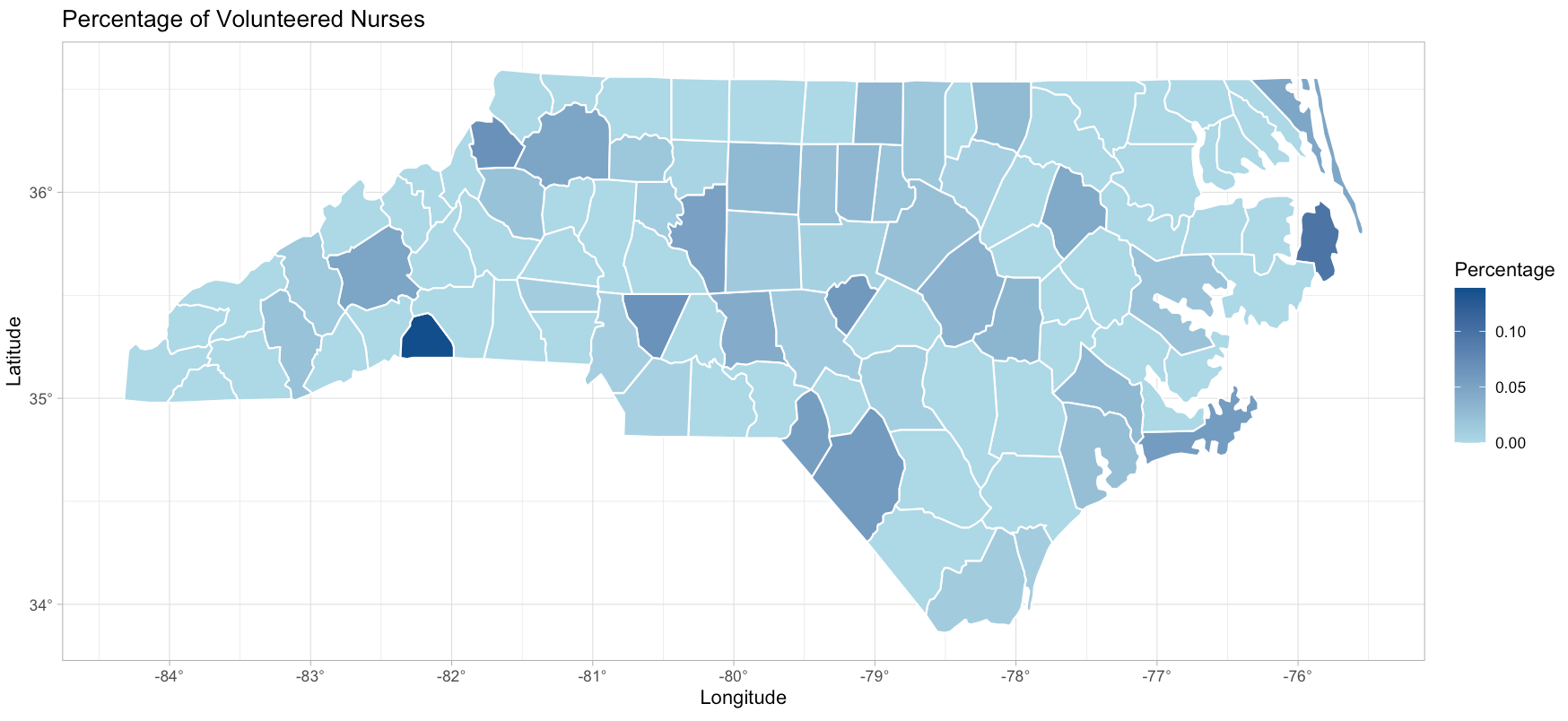
We then visualize the number of nurses who actually volunteered in each country. Figure 3 and Figure 4 show the counts and percentages respectively. In Figure 3, we have similar findings as Figure 1: urban counties, such as Wake, Durham, Orange, and Buncombe, have more nurses who volunteered during COVID 19.

Figure 3: Number of Volunteered Nurse, County Level



In Figure 4, Polk County has the highest percentage of volunteered nurse based on county population. We can still find that urban counties mentioned above (Wake, Durham, Orange) as well as some regional cities and suburban (Buncombe, Cabarrus, and Davidson) have a higher percentage of volunteered nurses, even though their population sizes are big. Meanwhile, rural counties with fewer population, including Scotland, Wilkes, Moore, and Watauga, also have large percentage of volunteered nurses. Thus, larger urban counties have a higher percentage of nurses who actually volunteered during the pandemic, but some rural counties also have higher percentages due to smaller population size.

Figure 4: Percentage of Volunteered Nurses on Population, County Level



We are interested in the correlation between volunteer nurse and vaccination coverage at county level. We first show the number of vaccinated population and the coverage rates in Figure 5 & 6. Again, darker blue represents higher counts/rates. In Figure 5, urban counties like Wake, Mecklenburg, Guilford, Forsyth, Durham, and Buncombe have the largest vaccinated population. It is consistent to the number of available nurse volunteers.

Figure 5: Number of Vaccinated Population, County Level

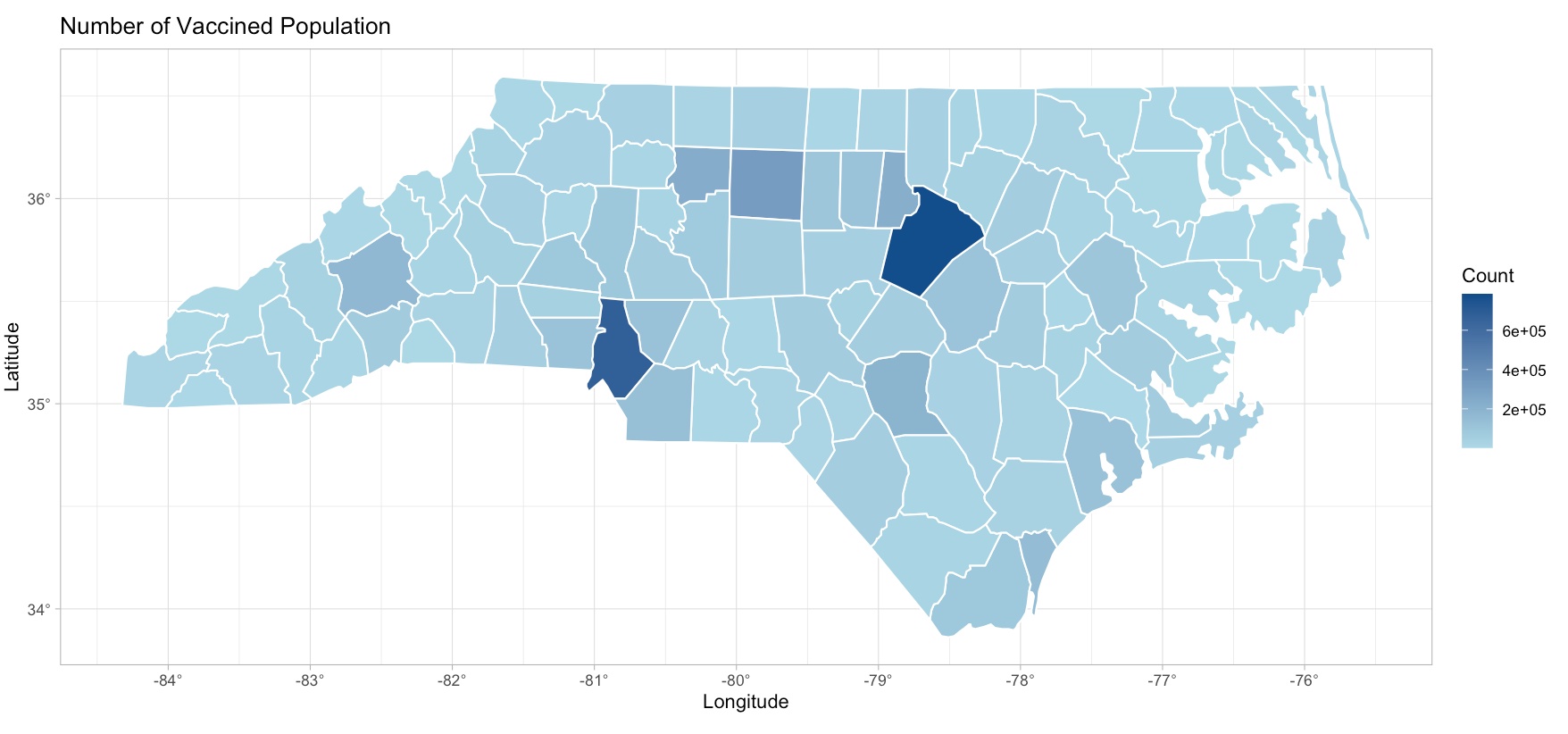
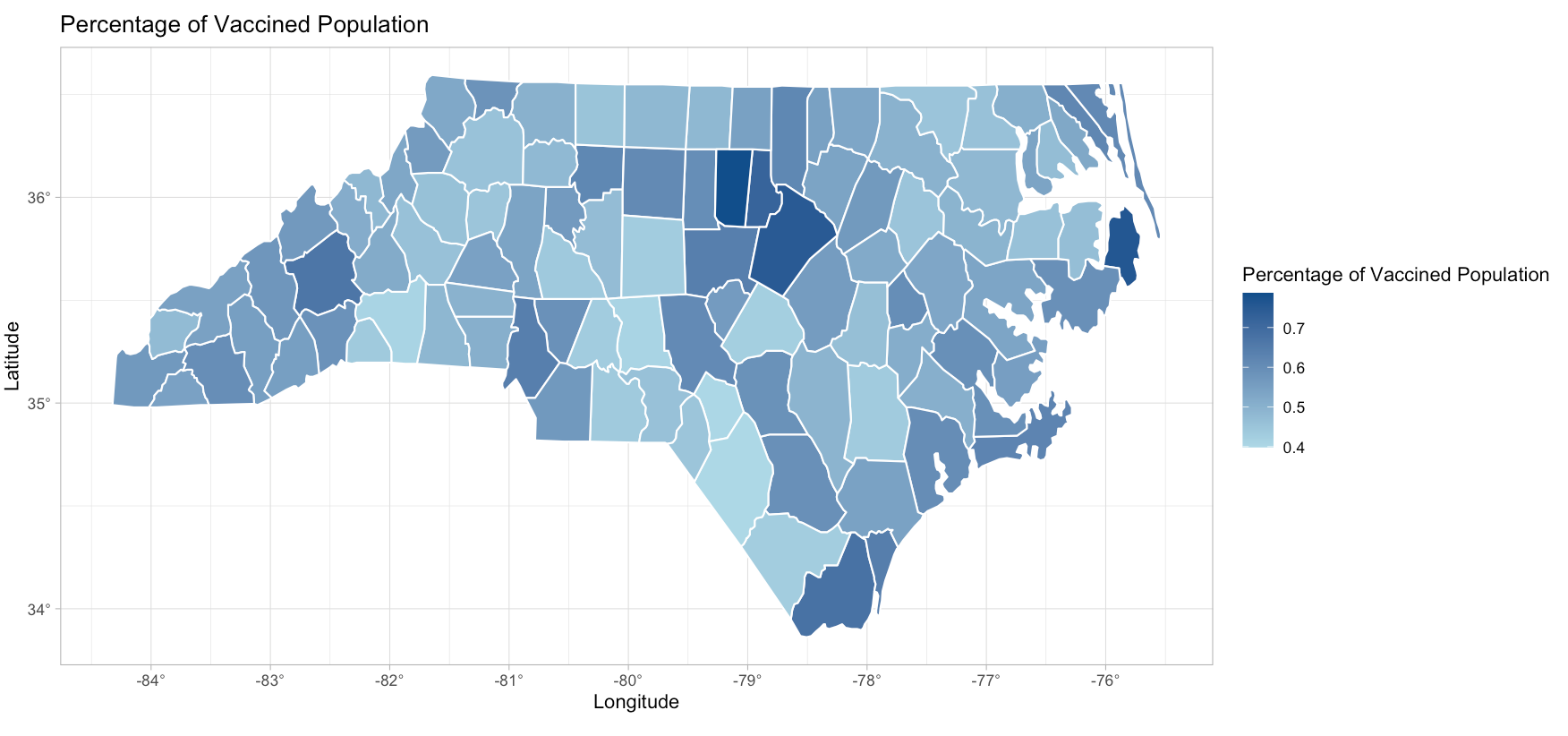


Figure 6 shows that most urban counties mentioned above have high vaccination rates. However, the vaccination coverage rates in rural counties are also high. Similar to the findings on volunteer percentage, the smaller population size can be one potential explanation. The statistical analysis between volunteer nurse and vaccines can be found in the next section.

Figure 6: Vaccination Coverage Rates, County Level



We also use the same visualization method to show the percentage of available nurses based on their license types (NP, RN, LPN, NAII, and other), license status (active or inactive), and retirement status (retired or not). More additional figures can be found in the appendix.

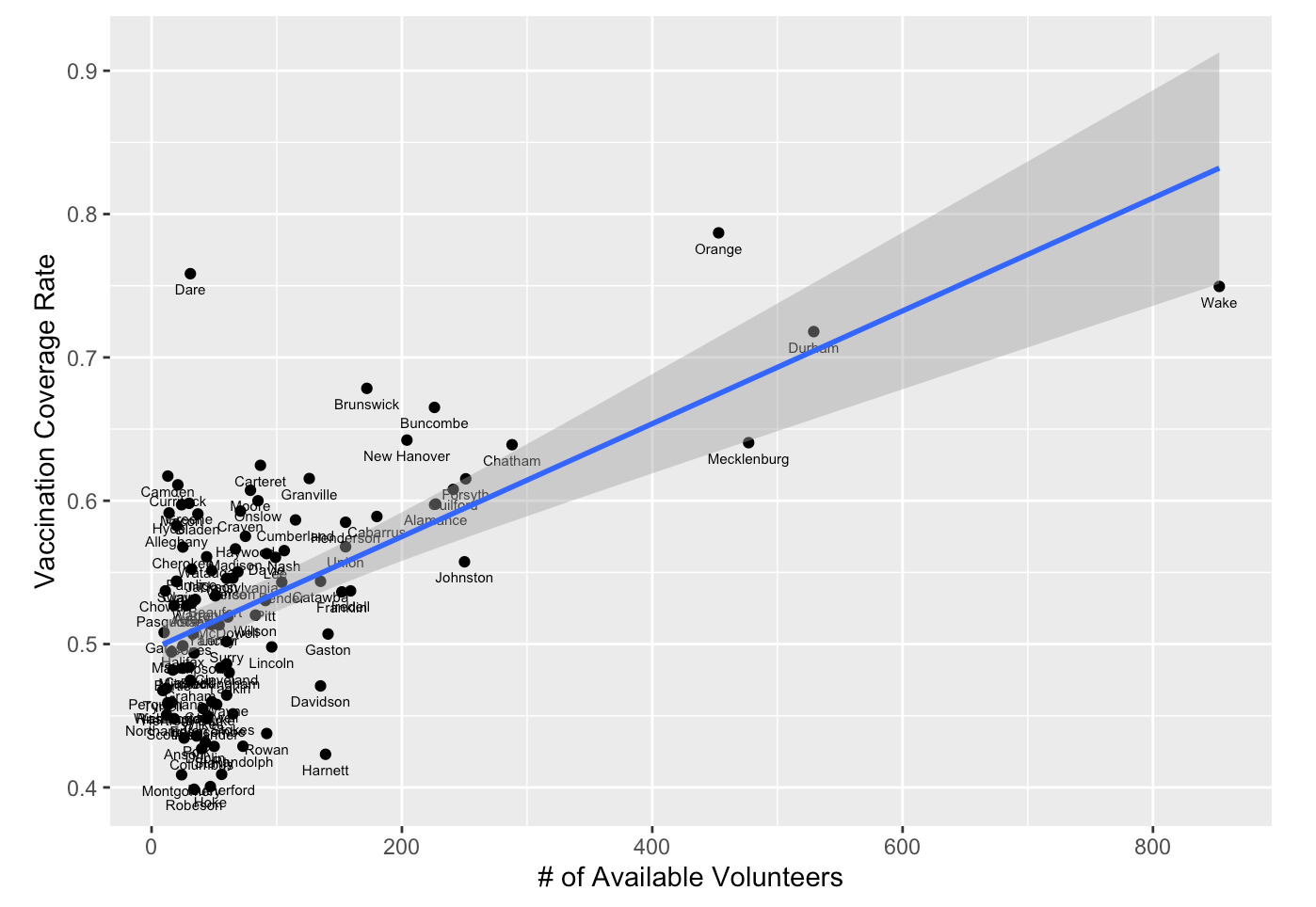
1. Correlation Analysis

We are also interested in the correlations between volunteer nurse and COVID-19 vaccination rates. The hypothesis is that there is significant correlation between the number of available nurses who are willing to volunteer, or more nurses who actually volunteered, and the COVID-19 vaccination coverage rate at county level in NC. We first show the scatter plots between the variables of interests.

Figure 7 shows the scatter plot between vaccination coverage rate & the number of available volunteers. Each point represents a county. There are many points in the left side of the figure, indicating that many counties have a very low number of available volunteers. The counties on the right the urban counties mentioned above, which have more available volunteers. They also have relatively high vaccination coverage rate. The blue line represents the regression line, showing that there is a positive correlation between the two variables. The grey area indicates the 95% confidence interval of the regression line, which shows the uncertainty inside the estimation.

We also use hypothesis testing to further examine the correlation. We use the vaccination coverage rate as the outcome variable and the number of available volunteers as the explanatory variable in a linear regression model. The coefficient is 3.937e-04 and is also statistically significant at the 5% significance level. The p-value is 3.56e-11, so we have a strong evidence to reject the null hypothesis that the correlation between the two variables is zero. It shows there is a significant positive correlation between the two variables.

Figure 7: Scatter Plot between Vaccination Coverage Rate & the Number of Available Volunteers, County Level



From the regression model, we can also use the R-squared value to explain to what extent the variance of one variable explains the variance of the second variable. R-squared value is a statistical measure that represents the proportion of the variance for a dependent variable that is explained by an independent variable or variables in a regression model. It ranges from 0 to 1 and are commonly stated as percentages from 0% to 100%. An R-squared of 100% means that all movements of the dependent variable are completely explained by movements in the independent variable(s).The R-squared value in this model is 0.3621, which means the explanatory variable can interpret 36.21% of the variations inside the outcome variable.

Similarly, Figure 8 shows the correlation between vaccination coverage rate and the number of nurses who actually volunteered. We can still observe the urban counties as outliers in the scatter plot, as well as the positive correlation. From the regression model, the coefficient is 0.01318 and the p-value is 6.94e-09. We have strong evidence that the two variables are significantly correlated, and the correlation is positive. The R-squared value is 0.2911, which means the explanatory variable can interpret 29.11% of the variations inside the outcome variable.

Figure 8: Scatter Plot between Vaccination Coverage Rate & the Number of Nurses who Actually Volunteered, County Level

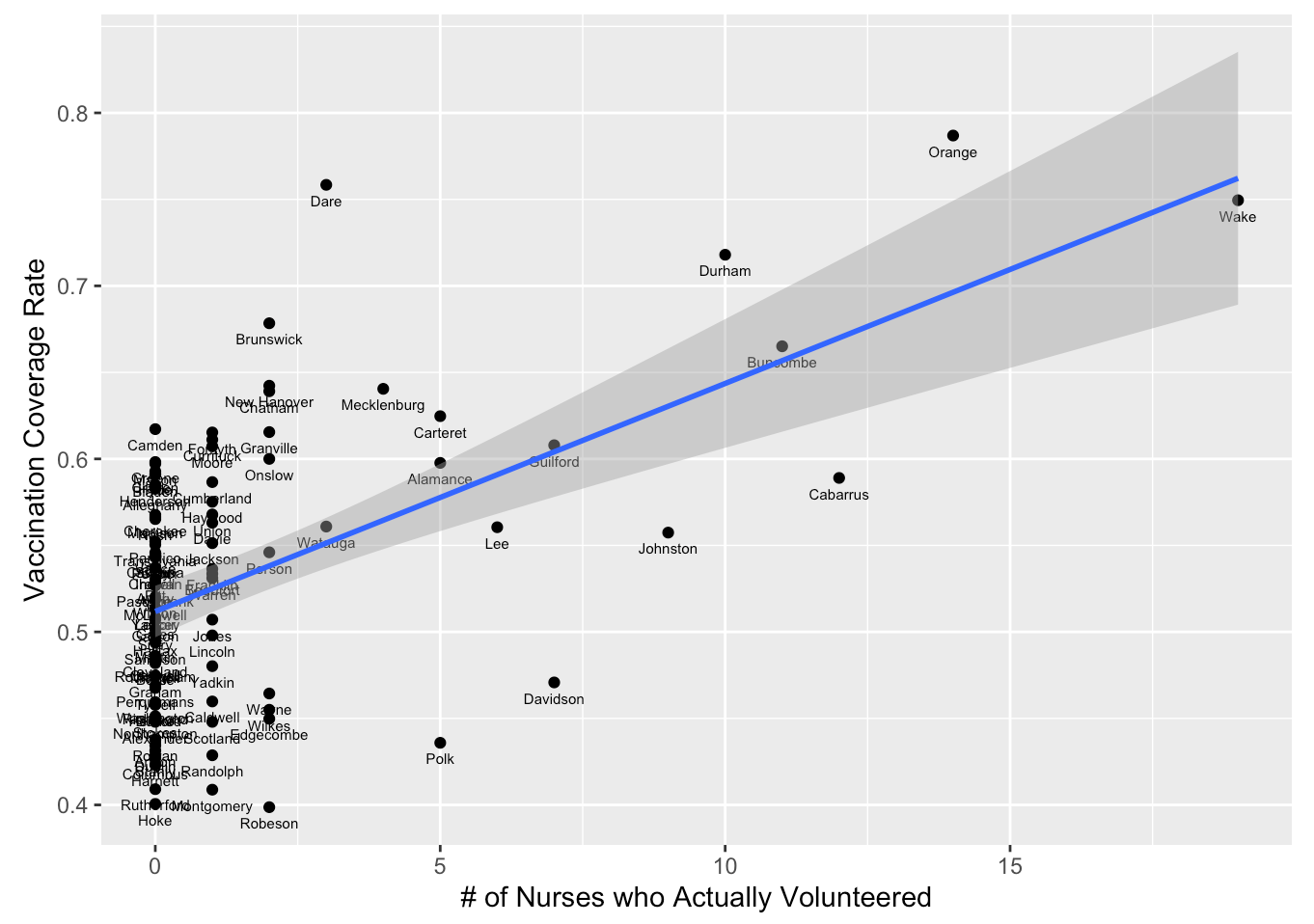
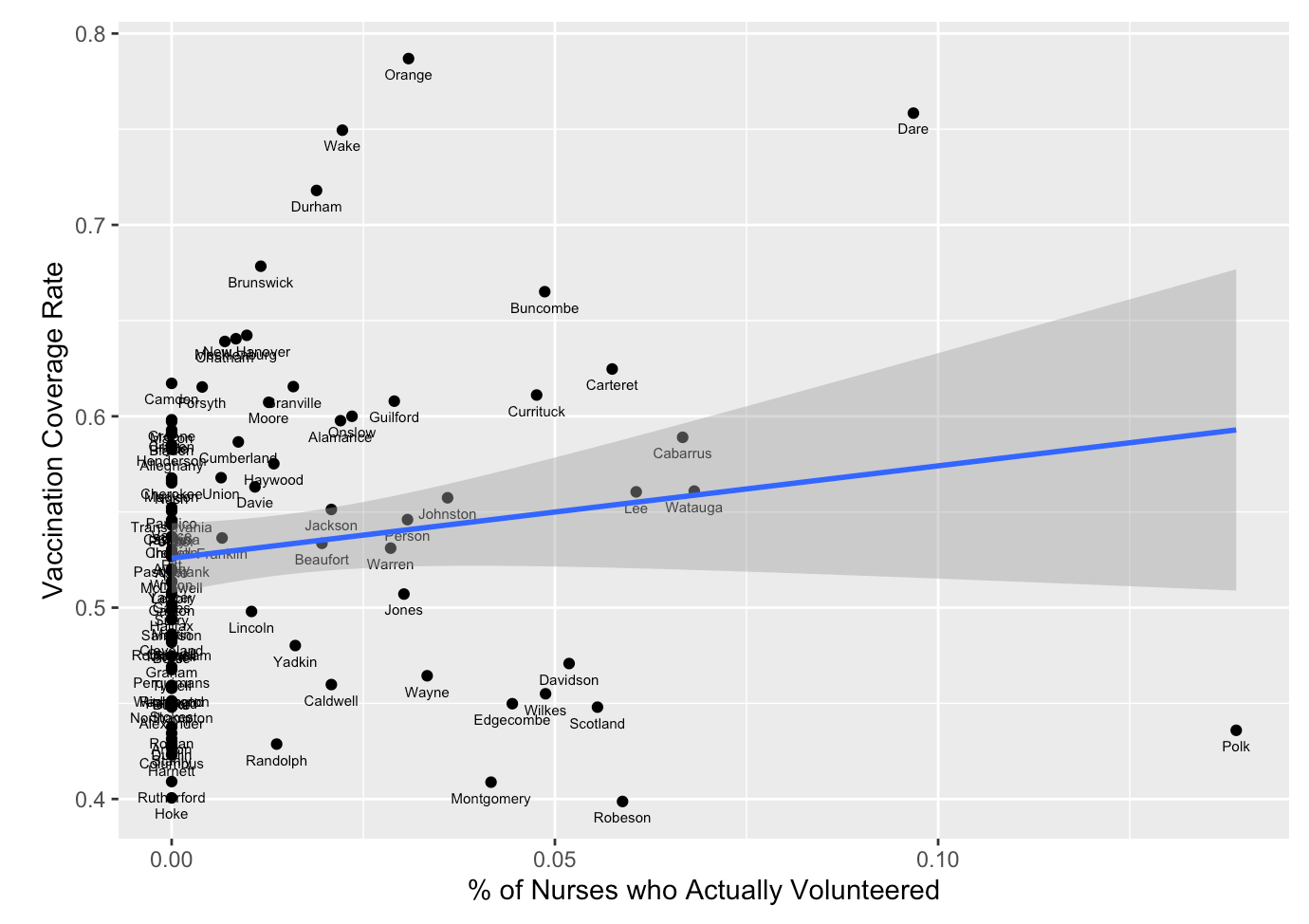


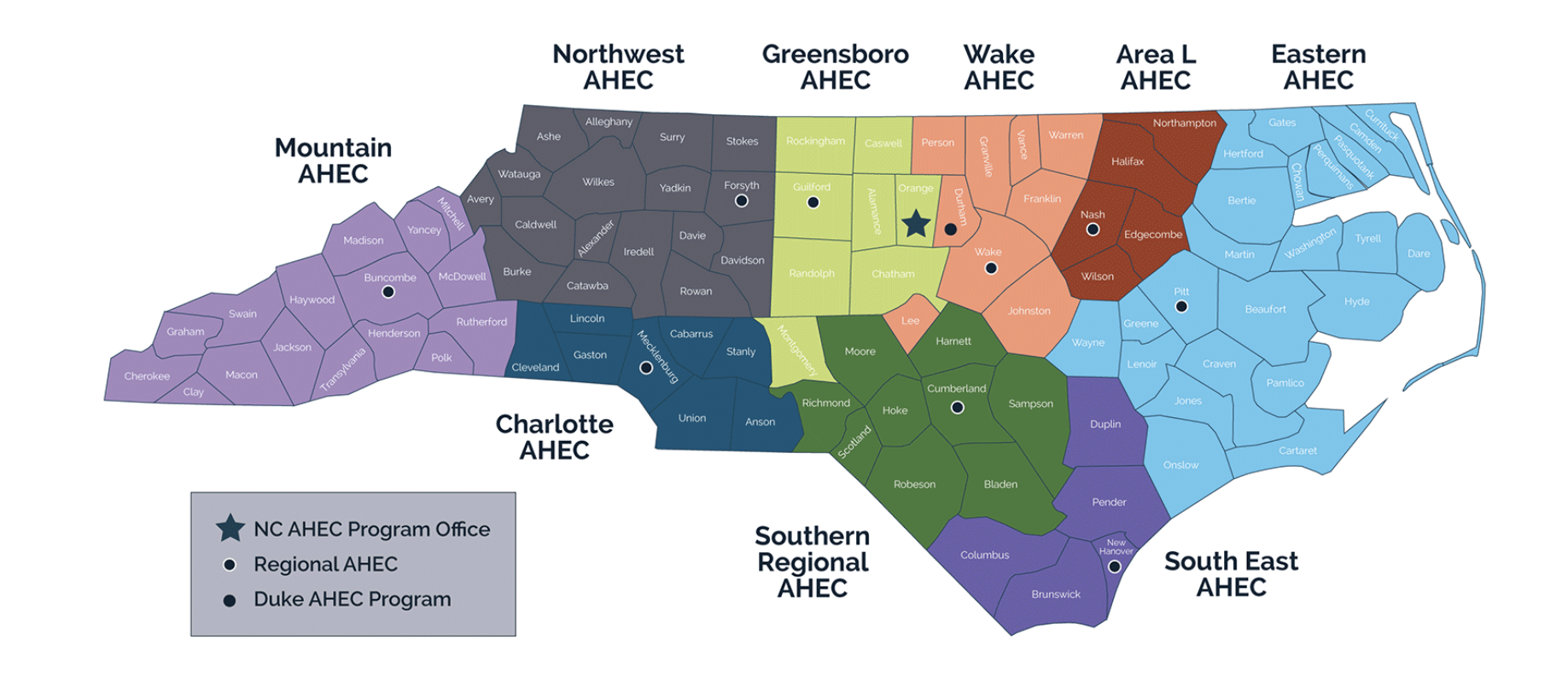
Figure 9 is the scatter plot between vaccination coverage rate and the proportion of nurses who actually volunteered over the total number of available volunteers. We can find a slight positive regression line in this figure, and many counties have zero inside the data. From the regression model, the coefficient is 0.4826, but the p-value is 0.15. We cannot reject the null hypothesis that the correlation between the two variables is zero. Thus, the relationship between vaccination rate and the volunteered nurse proportion is not significantly correlated. The R-squared value is also low, which is 0.021. It indicates that the explanatory variable can only interpret 2.1% of the variations inside the outcome variable.

Figure 9: Scatter Plot between Vaccination Coverage Rate & the Proportion of the Number of Nurses who Actually Volunteered over the Number of Available Volunteers, County Level



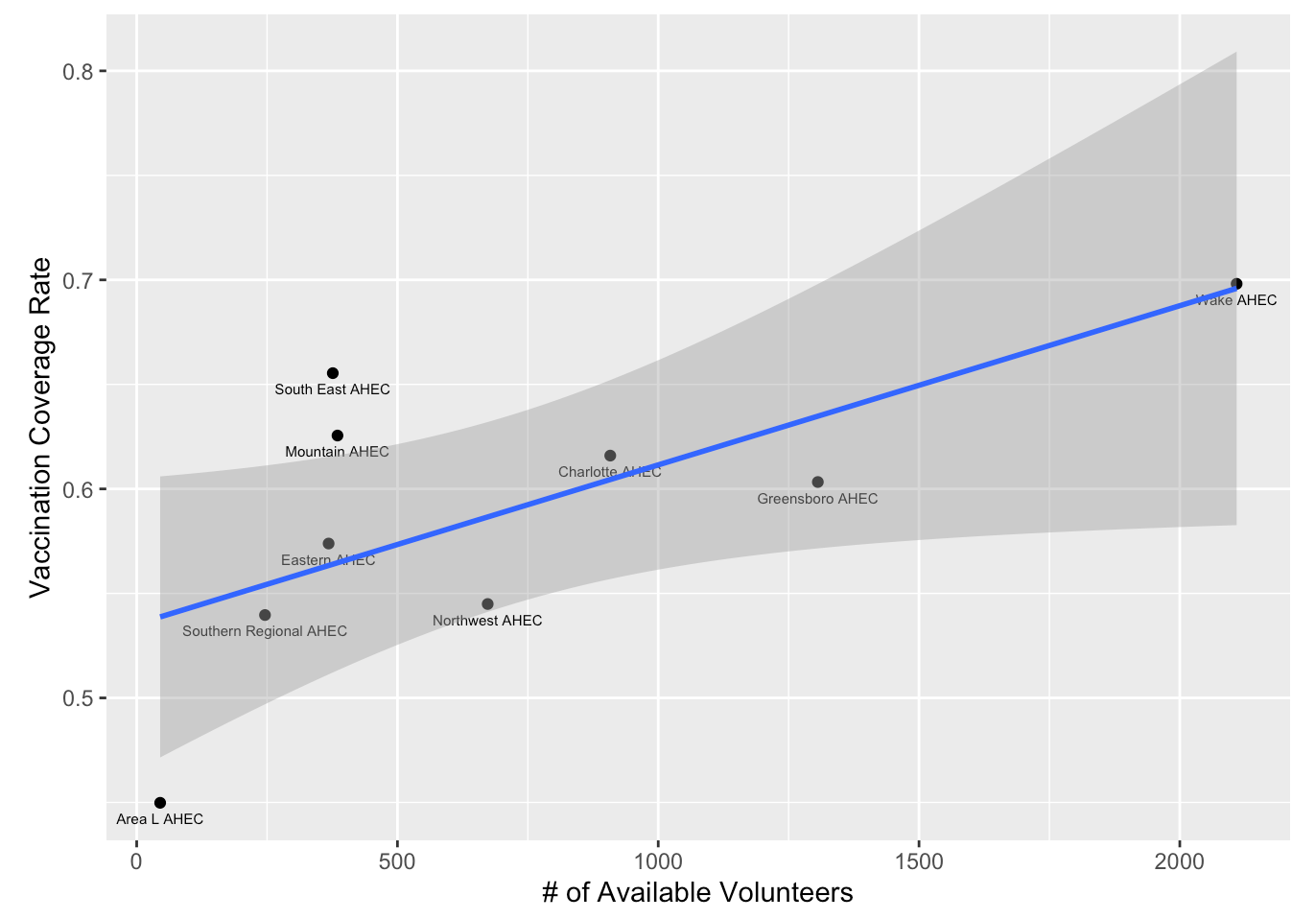
As we can observe, there are many zeros in the three figures above due to the small sample size. That is, the number of volunteer nurses, especially the number of nurses who actually volunteered, is low in many states. It may influence the correlation analysis results. Thus, we use the NC AHEC region to further aggregate the data and then analyze the correlations. The 100 counties in NC are divided into nine regions. We first aggregate the raw data into region level, and then calculate the metrics of interest, such as the proportions of volunteered nurses over available nurses and the vaccination coverage rates. We use the aggregated data to analyze the correlation by the same method. Figure 10 shows the AHEC regions.

Figure 10: AHEC Regions



We first analyze the correlation between vaccination rate and the number of available volunteers. From Figure 11, we can still observe the positive correlation between the two variables, which is consistent with the findings at county level. Area L, South East, and Mountain are more far away from the regression line and its 95% confidence interval, and they are all regions including more rural counties; but there are fewer outliers in the aggregated data. In the hypothesis testing, the correlation coefficient is 7.616e-05, with a p-value of 0.046. It shows that we can reject the null hypothesis that the correlation between the two variables is zero. The R-squared value is 0.4559, which indicates that the explanatory variable can only interpret 45.59% of the variations inside the outcome variable.

Figure 11: Scatter Plot between Vaccination Coverage Rate & the Number of Available Volunteers, Region Level



In Figure 12, there is a positive correlation between the vaccination rate and the number of volunteered nurses. In the regression model, the coefficient is 0.003188; but the p-value is 0.0552. We do not have strong evidence to reject the null hypothesis of zero correlation at the 5% significance level. The value of R-squared is still high, which is 0.4299.

Figure 12: Scatter Plot between Vaccination Coverage Rate & the Number of Nurses who Actually Volunteered, Region Level

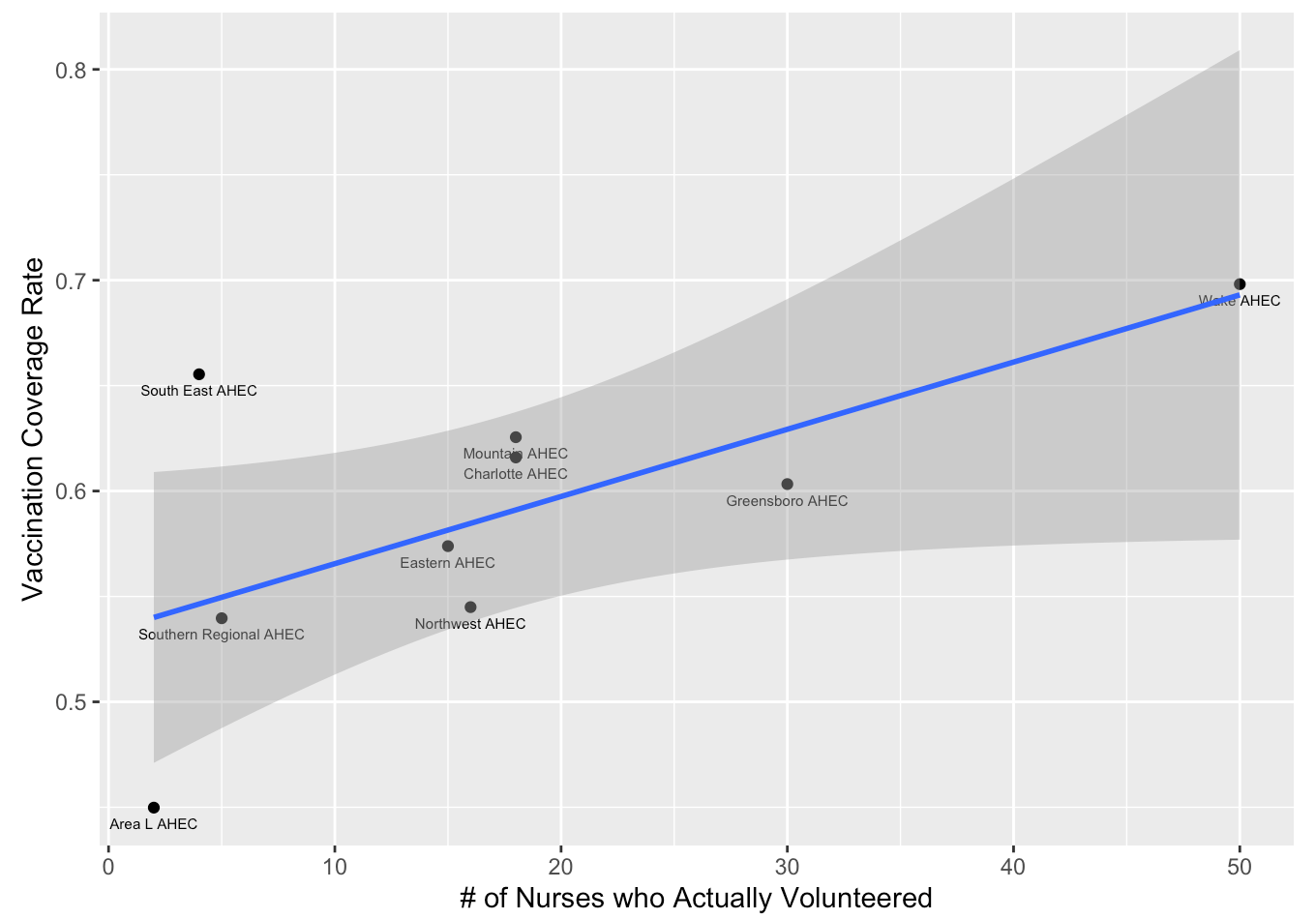
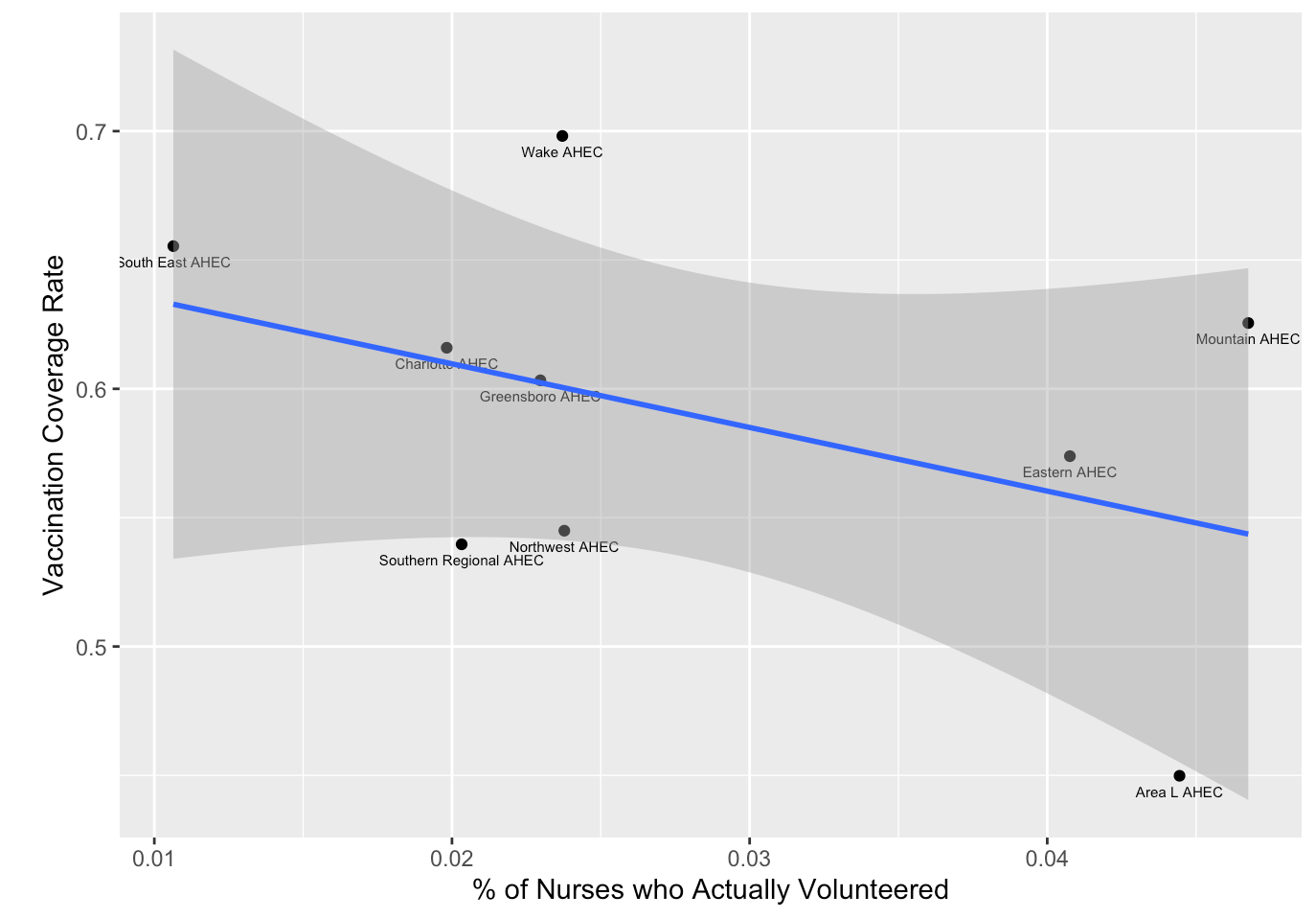


Figure 13 shows different results. The correlation between the vaccination coverage rate and the proportion of the volunteered nurses is negative. It seems that a lower proportion of volunteered nurses over available nurses is associated with a lower vaccination rate, which is different from our hypothesis. However, the regression coefficient is not statistically significant. The estimated coefficient is -2.470 and the p-value is 0.251. Thus, we cannot reject the null hypothesis of zero correlation. The R-squared is also lower (0.1826) compared to the two models above.

Figure 13: Scatter Plot between Vaccination Coverage Rate & the Proportion of the Number of Nurses who Actually Volunteered over the Number of Available Volunteers, Region Level



**Conclusion**

In this project, we first visualize the data to explore the county-level difference on volunteer nurses. From the county maps, we find that urban, regional cities and suburban counties have a higher number of available nurses and nurses who actually volunteered. However, urban counties also have more population, so the percentage of nurses over county population is the highest in both urban and rural counties. The number of vaccinated population is also higher in urban counties than rural counties, but the vaccination coverage rate in NC is high in general.

We also analyze the correlation between volunteer nurses and vaccination rates. Out hypothesis is that a higher number or proportion of volunteer nurses is associated with a higher vaccinate rate. When using the number of available and volunteered nurses as the explanatory variables, we find statistically significant correlation. We have strong evidence to reject the null hypothesis of zero correlation. When using the proportion of volunteered nurses over the available nurses as the independent variable, the correlation coefficient is not statistically significant. We find the same trend when aggregating the data at AHEC region level.

It seems that in counties with more available volunteer nurses, the healthcare system might be more developed and have more health care workers in general. It is associated with a higher vaccination rates. However, regardless of the proportion of volunteered nurses, the number of nurses is the most important. It is important for local government to facilitate the health care system and relevant program to recruit more health workers. It is also essential for AHEC to provide and support educational activities and services with a focus on primary care in rural communities and those with less access to resources to recruit, train, and retain the workforce.

1. The definition of rural/urban county comes from the NC Rural Center (www.ncruralcenter.org). There are 78 rural counties in NC, with an average population density of 250 people per squared miles or less. Regional cities and suburban counties refer to those have 250-750 people per squared miles; there are 16 in NC. There are 6 urban counties in NC and they have more than 750 people per squared miles on average. [↑](#footnote-ref-1)