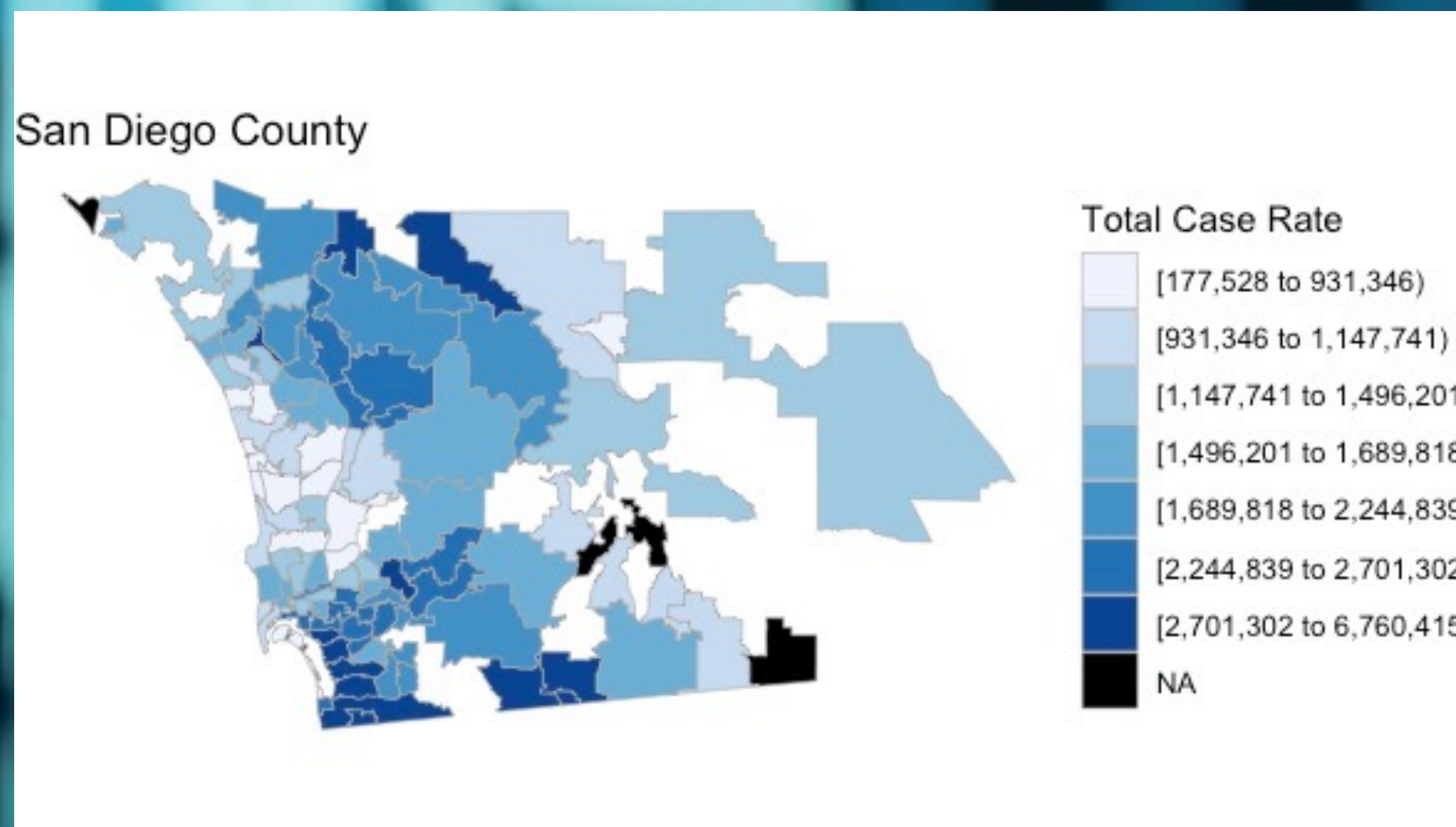# Covid on the Border

Introduction:
- ❖ Research on the Covid cases by zip code in San Diego County.
- ❖ Q: How does Covid affect regions in San Diego County by race and economic indicators?

Data:
- ❖ Data from San Diego Association of Governments (SANDAG).
- ❖ Two data sets:
  - ❖ First data set organized by cases per day in each zip code between 3/30/2020 and 12/28/2021
  - ❖ Second data set organized the demographic data by zip code
- ❖ Analysis in R:
  - ❖ Summarized the demographic data and merged the two data sets
  - ❖ Computed the case rates by population
  - ❖ Created maps and graphs for exploratory data analysis
- ❖ R packages used: Tidyverse, ggplot, choropleth, and choroplethrZip

San Diego County

Total Case Rate
- [177,528 to 931,346)
- [931,346 to 1,147,741)
- [1,147,741 to 1,496,201)
- [1,496,201 to 1,689,818)
- [1,689,818 to 2,244,839)
- [2,244,839 to 2,701,302)
- [2,701,302 to 6,760,415)
- NA

This map shows the distribution of Covid cases in San Diego County Zip Codes.

Methods:
- ❖ Maps: I created choropleth maps of San Diego County zip codes using case rates and demographics.
- ❖ Time Series: I created time series graphs by splitting the demographic data into high and low groups where high is the upper quartile and low is the lower quartile of zip codes.
- ❖ Scatterplot: I created scatterplots with regression and scatter plot smooth lines to study the relationship between case rates per zip code and demographics.
- ❖ Stacked Bar Chart: I created stacked bar charts of categorical demographic variables. The categories are grouped into four quartiles from low to high case rate zip codes.

## Proportion of People with a College Degree
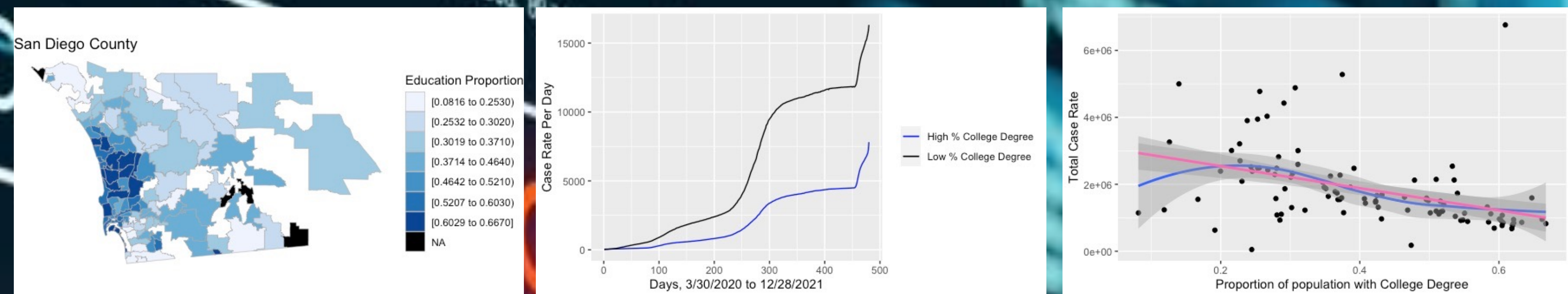
Conclusion: Regions with a higher proportion of educated citizens have lower case rates. Also, the difference between regions with high and low proportions of educated people grew as we moved further into the pandemic.

Proportion of People below the Poverty Level
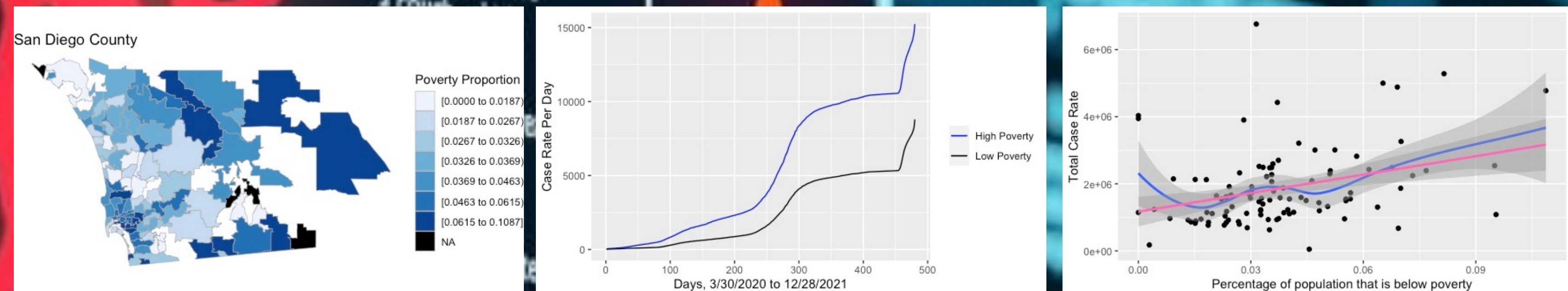
Conclusion: Regions with higher rates of poverty have higher case rates. Also, the difference between regions with high and low proportions of poverty grew as we moved further into the pandemic.

Ethnicity
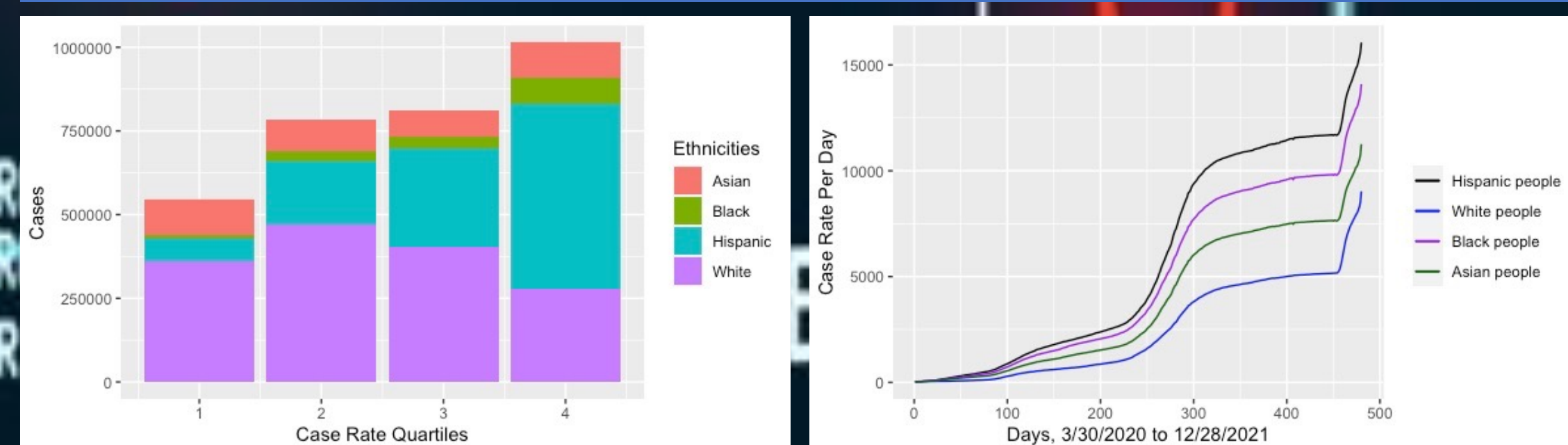
Conclusion: Zip codes with higher Covid rates tend to have a higher proportion of Hispanic and Black people while the proportion of Asian people in these zip codes stays about the same. Also, the difference in cases between these ethnic groups was increasing as we moved farther into the pandemic.

Future Research:
I studied the before and after vaccine data and found the results to be very similar. I wanted to include death rates in my research but I only found death rates at the county level but not at the zip code level. So next time I will include these death rates at the zip code level and add to my before and after vaccine data to further study differential impacts on demographic data.

References:
- ❖ Data obtained from: SANDAG/SanGIS Regional GIS Data Warehouse Open Data Portal
- ❖ Mentor: Professor Joey Lin, SDSU