Census Data Exploration

January 27, 2021

1 Census Data Exploration

Created by Keven Michel   The data used for this project was provided by the website, "Census Reporter," and focuses on the zipcodes within the county of Los Angeles. I decided to use zipcode data for this project so that I can broadly compare different races within the LA county region. Through these visualizations produced, I hope to gain a better understanding of the different percentages of races that reside within the county. I also hope to learn more on this data and how it can work alongside data such as income to provide a narrative on the health disparities in the Los Angeles County.

1.1 Import Data

First, I imported the neccessary libraries so that my code can run. I also uploaded my files using a .geojson.

```
[1]: import pandas as pd
```

```
[2]: import geopandas as gpd
```

```
[3]: import contextily as cx
```

```
[4]: import matplotlib.pyplot as plt
```
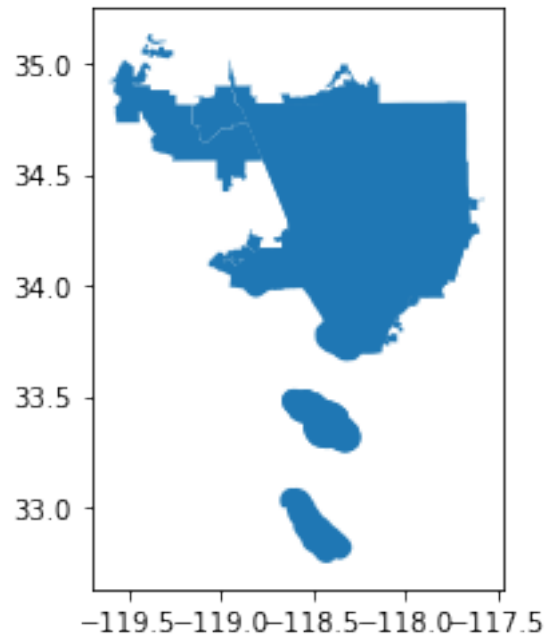
```
[5]: census = gpd.read_file("datasets/zipcodes.geojson")
```

1.1.1 Census Plot

This map shows a general overview of the zipcodes in the Los Angeles County. This provides a general shape.

```
[6]: census.plot()
```

```
[6]: <matplotlib.axes._subplots.AxesSubplot at 0x7fb65c9f3c70>
```

1.1.2 Cleaning The Data

In order for my data to provide an accurate information, I have to delete a row that does not correspond to the rest of the dataset. Below, I dropped the column which featured "Los Angeles" as it had information that was not relevant to the corresponding zipcodes I was using. Additionally, I renamed my columns so that they can be represented when creating my maps.

```
[7]: census = census.drop([0])
```

```
[8]: ctk = ['geoid',
    'name',
    'B03002001',
    'B03002002',
    'B03002003',
    'B03002004',
    'B03002005',
    'B03002006',
    'B03002007',
    'B03002008',
    'B03002009',
    'B03002012',
    'geometry' ]
```

```
[9]: census = census[ctk]
```

```
[10]:  census.columns = [
        'geoid',
        'Zipcode',
        'Total',
        'Non Hispanic',
        'Non Hispanic White',
        'Non Hispanic Black',
        'Non Hispanic American Indian and Alaska Native',
        'Non Hispanic Asian',
        'Non Hispanic Native Hawaiian and Other Pacific Islander',
        'Non Hispanic Some other race',
        'Non Hispanic Two or more races',
        'Hispanic',
        'geometry']
```

1.1.3 Creating Percentages For Each Racial Category

To create the visual maps for my research, I have to convert the census data of each race into percentages. Below, we can see how each race is converted in this way.

```
[11]:  census["Percent Non Hispanic"] = census["Non Hispanic"]/census["Total"]*100
       census['Percent Hispanic'] = census['Hispanic']/census['Total']*100
```

```
[12]:  census['Percent Non Hispanic White'] = census['Non Hispanic White']/
         ↪census['Total']*100
       census['Percent Non Hispanic Black'] = census['Non Hispanic Black']/
         ↪census['Total']*100
       census['Percent Non Hispanic American Indian and Alaska Native'] = census['Non␣
         ↪Hispanic American Indian and Alaska Native']/census['Total']*100
       census['Percent Non Hispanic Asian'] = census['Non Hispanic Asian']/
         ↪census['Total']*100
       census['Percent Non Hispanic Native Hawaiian and Other Pacific Islander'] =␣
         ↪census['Non Hispanic Native Hawaiian and Other Pacific Islander']/
         ↪census['Total']*100
       census['Percent Non Hispanic Some other race'] = census['Non Hispanic Some␣
         ↪other race']/census['Total']*100
       census['Percent Non Hispanic Two or more races'] = census['Non Hispanic Two or␣
         ↪more races']/census['Total']*100
```

1.1.4 Maps and Comparison

The maps below show the areas where the high percentages of each race is more present compared to others. The purpose of this is to gain new insights on the different racial percentages between zipcodes. We can also understand more on where certain races are more concentrated than others within the County.

```
[13]: fig, axs = plt.subplots(1,2, figsize= (15,12))
      ax1, ax2 = axs
      census.plot(column = "Percent Hispanic",
                  cmap = "plasma",
                  scheme = "quantiles",
                  k = 5,
                  edgecolor = "white",
                  linewidth = 0,
                  alpha = 0.75,
                  ax = ax1,
                  legend = True)
      ax1.axis("off")
      ax1.set_title("Percent Hispanic")

      census.plot(column = "Percent Non Hispanic White",
                  cmap = "plasma",
                  scheme = "quantiles",
                  k=5,
                  edgecolor = "white",
                  linewidth = 0,
                  alpha = 0.75,
                  ax = ax2,
                  legend = True)
      ax2.axis("off")
      ax2.set_title("Percent Non Hispanic White")
```
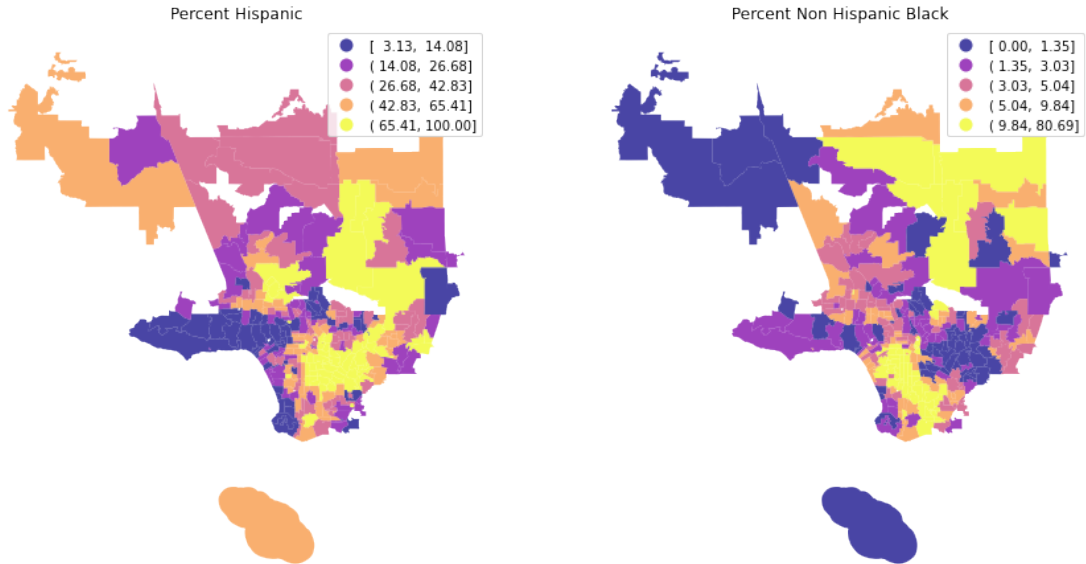
[13]: Text(0.5, 1.0, 'Percent Non Hispanic White')
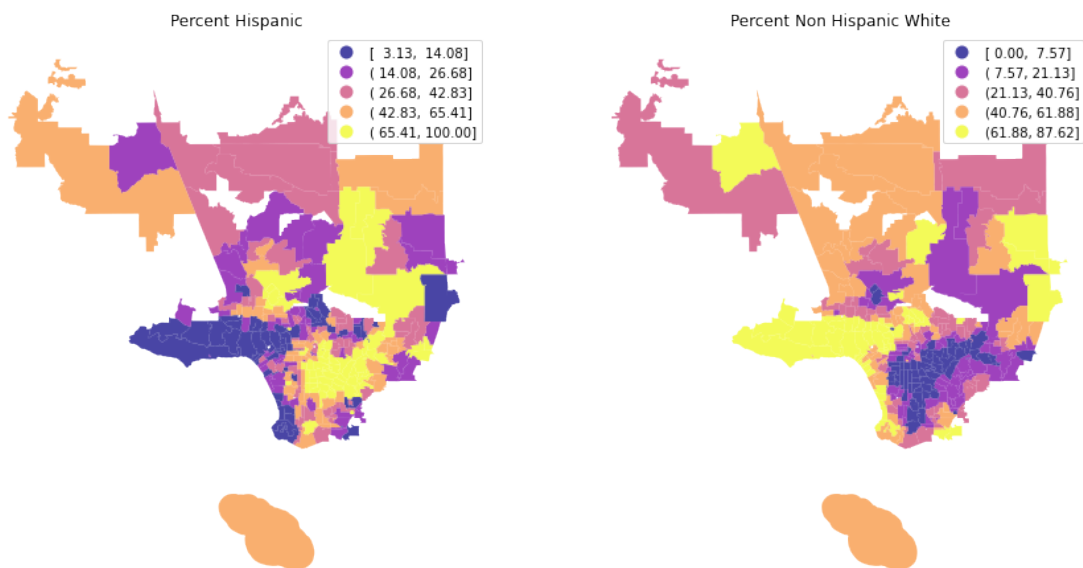
Percent Hispanic v. Percent Non Hispanic White   These two maps allow us to identify where each racial group is located more in the County according to thier zipcodes provided by the Census. As we can see, the yellow on the "percent hispanic" graph identifies how 65% of more of Hispanics live within that area. Compared to the map on the right, titled "percent non hispanic white," it shows areas mapped in yellow are where this percentage is higher compared to the darker regions.

```
[14]: fig, axs = plt.subplots(1,2, figsize= (15,12))
      ax1, ax2 = axs
      census.plot(column = "Percent Hispanic",
                  cmap = "plasma",
                  scheme = "quantiles",
                  k = 5,
                  edgecolor = "white",
                  linewidth = 0,
                  alpha = 0.75,
                  ax = ax1,
                  legend = True)
      ax1.axis("off")
      ax1.set_title("Percent Hispanic")

      census.plot(column = "Percent Non Hispanic Black",
                  cmap = "plasma",
                  scheme = "quantiles",
                  k=5,
                  edgecolor = "white",
                  linewidth = 0,
                  alpha = 0.75,
                  ax = ax2,
                  legend = True)
      ax2.axis("off")
      ax2.set_title("Percent Non Hispanic Black")
```

```
[14]: Text(0.5, 1.0, 'Percent Non Hispanic Black')
```

Percent Hispanic | Percent Non Hispanic Black

Percent Hispanic legend:
- [3.13, 14.08]
- (14.08, 26.68]
- (26.68, 42.83]
- (42.83, 65.41]
- (65.41, 100.00]

Percent Non Hispanic Black legend:
- [0.00, 1.35]
- (1.35, 3.03]
- (3.03, 5.04]
- (5.04, 9.84]
- (9.84, 80.69]

Percent Hispanic v. Percent Non Hispanic Black In these two maps, we can see a comparison between the Hispanic and Non Hispanic Black racial percentages and where they are more concentrated within the County. As mentioned before, the yellow color represent where each race has a 65% or higher percentage concentrated within that area. Eacg block represents a zipcode within the county.

[15]:
```
fig, axs = plt.subplots(1,2, figsize= (15,12))
ax1, ax2 = axs
census.plot(column = "Percent Hispanic",
            cmap = "plasma",
            scheme = "quantiles",
            k = 5,
            edgecolor = "white",
            linewidth = 0,
            alpha = 0.75,
            ax = ax1,
            legend = True)
ax1.axis("off")
ax1.set_title("Percent Hispanic")

census.plot(column = "Percent Non Hispanic Asian",
            cmap = "plasma",
            scheme = "quantiles",
            k=5,
            edgecolor = "white",
            linewidth = 0,
            alpha = 0.75,
```
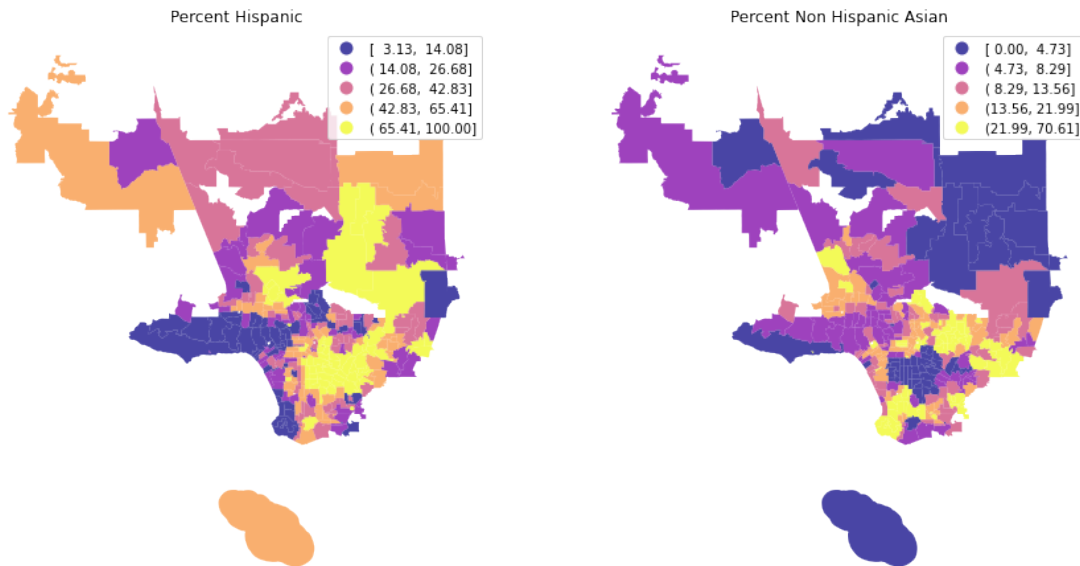
```
            ax = ax2,
            legend = True)
ax2.axis("off")
ax2.set_title("Percent Non Hispanic Asian")
```

[15]: Text(0.5, 1.0, 'Percent Non Hispanic Asian')

Percent Hispanic v. Percent Non Hispanic Asian In these two maps, we compare the percentages between Hispanics and Non Hispanic Asians within the county. As mentioned before, the yellow color represents where each raical percentage is more concentrated compared to their counterparts.

1.1.5 Building A Detailed Map

In this process, I created a detailed map that clearly shows the user/audience where different cities are located within the radius of the L.A. County. The data is still represented through zipcode blocks.

[16]: ```
census_web_mercator = census.to_crs(epsg = 3857)
```

[17]: ```
fig, ax = plt.subplots(figsize = (15,15))

census_web_mercator[census_web_mercator["Percent Hispanic"]>50].plot(ax=ax,␣
 ↪alpha=0.8)

ax.axis("off")
```
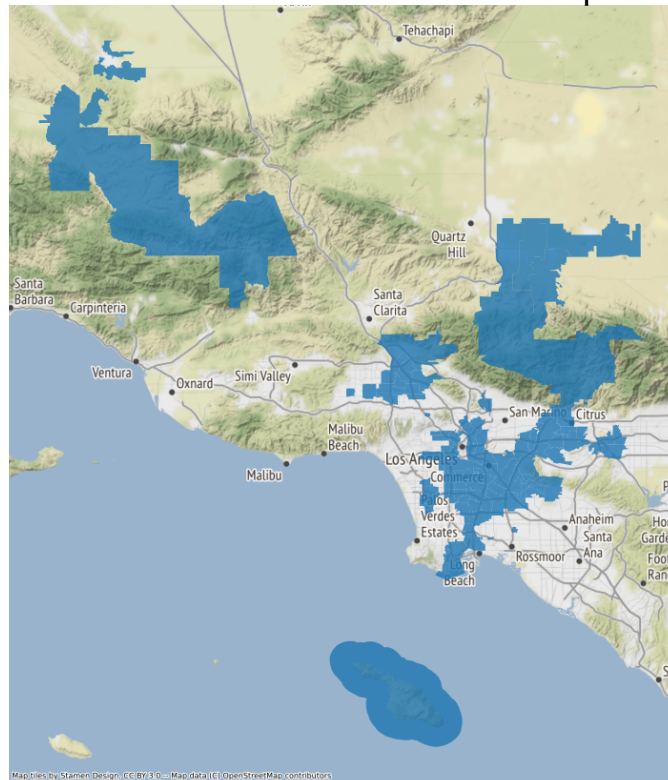
7

```
ax.set_title("Census Zipcodes with more than 50% Hispanic Population", fontsize␣
 ↪= 40)

cx.add_basemap(ax)
```

Census Zipcodes with more than 50% Hispanic Population

Census Zipcodes Map   In this map, the blue area is where more than 50% of this Hispanic Population is concentrated. Using this map, it gives a more detailed visualization of cities in and outside of Los Angeles County. It also allows me to gain a better insights on where these populations are located within their respective zipcodes.

[ ]: