ZHVI_project

Ben Holden

12/18/2022

```
library(tidycensus)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.7
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1
```

```
## Warning: package 'tidyr' was built under R version 4.1.2
```

```
## Warning: package 'readr' was built under R version 4.1.2
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(ggplot2)
library(choroplethrZip)
library(RColorBrewer)
```

PART 1: Data Wrangling: ZCTA land area, population, demographics and income.

Data frame no. 1: ZCTA Land Area & 2010 Census Population

```
#Loading ZCTA census package by "jjchern/zcta" GitHub repository:
devtools::install_github("jjchern/zcta")
```

```
## Skipping install of 'zcta' from a github remote, the SHA1 (763a2a86) has not changed since last insta
##   Use `force = TRUE` to force installation
```

```
# Selecting all ZCTAs into dataframe matching ZCTAs to counties:
zcta_area <- zcta::zcta_county_rel_10
```

```
#Removing Duplicates, converting area from square meters to square miles, selecting relevant variables
zcta_area <- zcta_area%>%
  distinct(zcta5, .keep_all = TRUE)%>%
  mutate(zcta_sq_mi = zarealand/2590000)%>%
  select(zcta5, zpop, zcta_sq_mi)
```

Data frame no. 2: 2019 ACS population

```
#Loading 2019 American Community Survey (ACS) estimates of population by ZCTA-5
zcta_population <-  read.csv("https://github.com/ben-luc-hol/Zillow-ZHVI-Pandemic-Home-Values-Project/r

zcta_population <-  zcta_population%>%
                        rename("acs2019_population" = B01003_001E,   #renaming variables for population
                               "zcta5" = NAME)%>%
                            select(acs2019_population, zcta5)%>%       #selecting variables
                            relocate(zcta5, .before = acs2019_population)

zcta_population <- zcta_population[-1,] #removing first observation containing description


#Tidying unique identifier (ZCTA ID)
zcta_population <- zcta_population%>%
                        separate(zcta5, c("key", "zcta5"), " ")%>%
                        select(zcta5, acs2019_population)%>%
                        mutate(acs2019_population = as.numeric(acs2019_population, na.rm = TRUE))
```

Data frame no. 3: ZCTA Income, ACS 2019

```
#Income dataset
zcta_income <- read.csv("https://github.com/ben-luc-hol/Zillow-ZHVI-Pandemic-Home-Values-Project/raw/ma

zcta_income <- zcta_income%>%
                rename("median_income" = B19013_001E,   #renaming variables for income and zcta
                       "zcta5" = NAME)

zcta_income <- zcta_income[-1,]


#Tidying income dataset
zcta_income <-  zcta_income%>%
                    separate(zcta5, c("key", "zcta5"), " ")%>%   #splitting variable to only contain zct
                    select(zcta5, median_income)%>%
                    arrange(zcta5, .before = median_income)%>%
                    mutate(median_income = as.numeric(median_income))
```

```
## Warning in mask$eval_all_mutate(quo): NAs introduced by coercion
```

Data frame no. 4: ZCTA Demographics, 2019 ACS

```
#Race dataset, same procedure as other ACS datasets
zcta_race <- read.csv("https://github.com/ben-luc-hol/Zillow-ZHVI-Pandemic-Home-Values-Project/raw/main,

zcta_race <- zcta_race[-1,]

zcta_race <- zcta_race%>%
  rename("population" = B02001_001E,
         "white_total" = B02001_002E,
         "black_total" = B02001_003E,
```

```
        "zcta5" = NAME)%>%
    select(zcta5,population,white_total,black_total)%>%
    mutate(perc_white = (as.numeric(white_total)/as.numeric(population))*100,
        perc_black = (as.numeric(black_total)/as.numeric(population))*100,
        perc_other = 100-(perc_black + perc_white))%>%
    separate(zcta5, c("key", "zcta5"), " ")%>%    #splitting variable to only contain zcta code)%>%
    select(zcta5, perc_white, perc_black, perc_other)%>%
    arrange(zcta5, .before = perc_white)
```

Data frame no. 5: All ZCTA characteristics

```
#joining population and area datasets into ZCTA main
zctas <- inner_join(zcta_population,
                    zcta_area,
                    by = "zcta5")


#adding income dataset
zctas <- inner_join(zctas,
                    zcta_income,
                    by = "zcta5")


zctas <- inner_join(zctas,
                    zcta_race,
                    by = "zcta5")



#creating population density variables
zctas <-  zctas%>%
  mutate(zcta_density = (zpop/zcta_sq_mi),
        zcta_density19 = (acs2019_population/zcta_sq_mi))%>%
  rename("zpop19" = acs2019_population)

 #Prepare df to join into main
zctas_main <- zctas%>%
        select( zcta5, median_income, zcta_density, perc_white, perc_black, perc_other, zpop)%>%
        rename("RegionName" = zcta5)
```

PART 2: DATA WRANGLING — CREATING MAIN DATAFRAME

```
#Loading dataset
zips <- read_csv("https://github.com/ben-luc-hol/Zillow-ZHVI-Pandemic-Home-Values-Project/raw/main/data,

## Rows: 30502 Columns: 270
## -- Column specification ----------------------------------------------------
## Delimiter: ","
## chr   (7): RegionName, RegionType, StateName, State, City, Metro, CountyName
## dbl (263): RegionID, SizeRank, 2000-01-31, 2000-02-29, 2000-03-31, 2000-04-3...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.

zips <- zips%>%
  select(RegionName, City, CountyName, Metro, State, `2018-06-30`:`2021-09-30`)%>%    #Selecting only t
```

```r
  mutate(beforeprice = ((`2019-06-30`+ `2019-07-31`+`2019-08-31`+`2019-09-30`)/4),     #Calculating "b
         afterprice = ((`2021-06-30`+ `2021-07-31`+`2021-08-31`+`2021-09-30`)/4),
         perc_chg = (as.numeric(afterprice/beforeprice)-1),                            #Percent change (
         perc_chg_rounded = perc_chg*100)                                              # *100


#joining dataframes
zips_main <- inner_join(zips,
                        zctas_main,
                        by = "RegionName")


#shuffling around columns in dataframe
zips_main <- zips_main%>%
              relocate((`2018-06-30`:`2021-09-30`), .after = perc_other)%>%
              relocate(beforeprice, .after = perc_other)%>%
              relocate(afterprice, .after = beforeprice)%>%
              relocate(perc_chg, .after = perc_other)%>%
              relocate(perc_chg_rounded, .after = perc_chg)

#Creating scaled income and density variables for regressions
zips_main <- zips_main%>%
              mutate(income_k = median_income/1000,
                     density_k = zcta_density/1000)
```

```r
glimpse(zips_main)
```

```
## Rows: 29,924
## Columns: 57
## $ RegionName       <chr> "10025", "60657", "10023", "77494", "60614", "77449",~
## $ City             <chr> "New York", "Chicago", "New York", "Katy", "Chicago",~
## $ CountyName       <chr> "New York County", "Cook County", "New York County", ~
## $ Metro            <chr> "New York-Newark-Jersey City", "Chicago-Naperville-El~
## $ State            <chr> "NY", "IL", "NY", "TX", "IL", "TX", "NY", "TX", "TX",~
## $ median_income    <dbl> 91624, 95173, 132605, 135943, 117138, 78112, 36982, 6~
## $ zcta_density     <dbl> 125860.102, 29028.563, 124357.545, 1510.660, 21040.37~
## $ perc_white       <dbl> 66.19332, 86.01567, 78.81710, 64.71921, 83.96337, 57.~
## $ perc_black       <dbl> 11.5651863, 3.0243242, 4.6513487, 8.5856067, 4.284681~
## $ perc_other       <dbl> 22.241493, 10.960005, 16.531555, 26.695184, 11.751953~
## $ perc_chg         <dbl> 0.058616188, 0.061720278, 0.039003304, 0.203309631, 0~
## $ perc_chg_rounded <dbl> 5.8616188, 6.1720278, 3.9003304, 20.3309631, 4.278790~
## $ beforeprice      <dbl> 1046016.5, 483265.0, 1401721.0, 337560.2, 627683.2, 1~
## $ afterprice       <dbl> 1107330.0, 513092.2, 1456392.8, 406189.5, 654540.5, 2~
## $ `2018-06-30`     <dbl> 1049497, 489327, 1482564, 336827, 635667, 180200, 888~
## $ `2018-07-31`     <dbl> 1044842, 488444, 1484431, 336660, 634543, 180534, 887~
## $ `2018-08-31`     <dbl> 1048917, 487529, 1501195, 337035, 633456, 181247, 881~
## $ `2018-09-30`     <dbl> 1043957, 487929, 1515711, 338123, 634215, 182060, 900~
## $ `2018-10-31`     <dbl> 1052782, 488448, 1529230, 339019, 634929, 182992, 925~
## $ `2018-11-30`     <dbl> 1060405, 490339, 1529444, 339949, 637329, 183714, 957~
## $ `2018-12-31`     <dbl> 1065955, 491396, 1505561, 339699, 638664, 184411, 967~
## $ `2019-01-31`     <dbl> 1084131, 489902, 1468424, 339983, 635965, 185300, 977~
## $ `2019-02-28`     <dbl> 1089543, 488093, 1430422, 339517, 633871, 185966, 995~
## $ `2019-03-31`     <dbl> 1100527, 486245, 1428218, 339374, 631674, 186393, 101~
## $ `2019-04-30`     <dbl> 1091757, 486150, 1417918, 338367, 632395, 186328, 992~
```

```
## $ `2019-05-31`    <dbl> 1087597, 485167, 1415310, 337679, 630941, 186495, 984~
## $ `2019-06-30`    <dbl> 1076631, 484122, 1409515, 337119, 629117, 186933, 989~
## $ `2019-07-31`    <dbl> 1057004, 483884, 1413724, 337365, 628107, 187585, 100~
## $ `2019-08-31`    <dbl> 1026366, 482921, 1407107, 337483, 627114, 188175, 100~
## $ `2019-09-30`    <dbl> 1024065, 482133, 1376538, 338274, 626395, 188493, 994~
## $ `2019-10-31`    <dbl> 1039973, 480666, 1363832, 338682, 625342, 189224, 100~
## $ `2019-11-30`    <dbl> 1074291, 480019, 1378399, 339019, 624598, 190469, 999~
## $ `2019-12-31`    <dbl> 1086290, 479544, 1417384, 338735, 623769, 191610, 102~
## $ `2020-01-31`    <dbl> 1080434, 479689, 1443785, 338366, 623647, 191823, 106~
## $ `2020-02-29`    <dbl> 1055334, 481297, 1450511, 337973, 624085, 191852, 109~
## $ `2020-03-31`    <dbl> 1021225, 482659, 1446545, 338265, 623859, 192958, 110~
## $ `2020-04-30`    <dbl> 1005297, 483786, 1455405, 339134, 623069, 194617, 106~
## $ `2020-05-31`    <dbl> 995100, 482886, 1456083, 340090, 621611, 195744, 1034~
## $ `2020-06-30`    <dbl> 994083, 482619, 1455080, 340531, 620433, 195970, 1000~
## $ `2020-07-31`    <dbl> 1003734, 484184, 1457939, 340670, 621882, 196327, 988~
## $ `2020-08-31`    <dbl> 1011176, 486312, 1467214, 341971, 623431, 197121, 983~
## $ `2020-09-30`    <dbl> 1023517, 490482, 1479629, 343561, 628612, 199246, 972~
## $ `2020-10-31`    <dbl> 1041639, 493571, 1477100, 346209, 632476, 201532, 966~
## $ `2020-11-30`    <dbl> 1075896, 497476, 1476888, 349471, 637548, 204124, 949~
## $ `2020-12-31`    <dbl> 1109692, 500087, 1475579, 353965, 640982, 206456, 919~
## $ `2021-01-31`    <dbl> 1103483, 501122, 1457248, 358010, 641971, 208686, 878~
## $ `2021-02-28`    <dbl> 1093575, 502434, 1446666, 362614, 643528, 211565, 850~
## $ `2021-03-31`    <dbl> 1082559, 502468, 1432506, 368069, 643668, 212969, 841~
## $ `2021-04-30`    <dbl> 1090164, 504373, 1443528, 374704, 645972, 215139, 852~
## $ `2021-05-31`    <dbl> 1096777, 505035, 1444495, 381647, 645922, 217765, 857~
## $ `2021-06-30`    <dbl> 1097715, 507623, 1445244, 390291, 648377, 223264, 875~
## $ `2021-07-31`    <dbl> 1102062, 511422, 1448051, 401064, 652306, 229263, 888~
## $ `2021-08-31`    <dbl> 1109991, 515354, 1459086, 412065, 657159, 235681, 913~
## $ `2021-09-30`    <dbl> 1119552, 517970, 1473190, 421338, 660320, 240759, 921~
## $ zpop           <dbl> 94600, 65996, 60998, 61600, 66617, 94382, 81410, 9513~
## $ income_k       <dbl> 91.624, 95.173, 132.605, 135.943, 117.138, 78.112, 36~
## $ density_k      <dbl> 125.860102, 29.028563, 124.357545, 1.510660, 21.04037~
```

```r
#creating separate dataframes for each Metropolitan Statistical Area

#New York
zips_newyork <- zips_main%>%
                filter(Metro %in% "New York-Newark-Jersey City")
#DC
zips_dc      <- zips_main%>%
                filter(Metro %in% "Washington-Arlington-Alexandria")
#Boston
zips_boston  <- zips_main%>%
                filter(Metro %in% "Boston-Cambridge-Newton")


#San Francisco Bay Area
sfbaymsas <- c("San Francisco-Oakland-Hayward", "San Jose-Sunnyvale-Santa Clara")
sfbaymsas01 <-  c("San Francisco-Oakland-Hayward, CA","San Jose-Sunnyvale-Santa Clara, CA")


zips_sfbay   <-  zips_main%>%
                filter(Metro %in% sfbaymsas)-> zip_sf

#Denver and surrounding counties
```

```
centralco <- c("Adams County", "Arapahoe County", "Broomfield County", "Clear Creek County", "Denver Cou
                "Jefferson County", "Park County", "Gilpin County", "Boulder County", "Teller Coun
                "Larimer County", "Weld County", "Morgan County")


zips_centralco <- zips_main%>%
                   filter(CountyName %in% centralco,
                          State == "CO")
```

PART 3: CHOROPLETHS

```
#custom color scale for ZHVI choros
m20 <- "#8e0152"
m15 <-  "#c51b7d"
m10 <-  "#de77ae"
m5 <- "#f1b6da"
m0 <- "#fde0ef"
p0 <- "#f7f7f7"
p5 <- "#e6f5d0"
p10 <- "#b8e186"
p15 <-  "#7fbc41"
p20 <- "#4d9221"
p30 <-  "#276419"
```

#NEW YORK

```
zips_newyork%>%
  select(RegionName, City, CountyName, Metro, State, perc_chg, zcta_density)%>%
  rename(region = RegionName,
         value = perc_chg)%>%
  mutate(value, original_value = value) -> ny_ch


ny_ch$value <- cut(ny_ch$value,
                   breaks = c(-Inf, -0.20, -0.15, -0.1, -0.05, 0, 0.05, 0.10, 0.15, 0.20, 0.30, Inf),
                   labels = c( "(20+)", "(15-20)", "(10-15)","(5-10)", "(0-5)", "0-5", "5-10", "10-15",
                   ordered_result = TRUE,
                   include.lowest = FALSE)

choro <-  ZipChoropleth$new(ny_ch)
choro$title <- "% change in typical home value"
choro$ggplot_scale <- scale_fill_manual(values = c(m20,m15,m10,m5,m0,p0,p5,p10,p15,p20,p30),
                                          drop = FALSE,
                                          na.value = "#d2d2d2")
choro$set_zoom_zip(state_zoom = NULL, county_zoom = NULL, msa_zoom = "New York-Newark-Jersey City, NY-N.
choro$render()

## Warning in self$bind(): The following regions were missing and are being set to
## NA: 08540, 07970, 07939, 11042, 08691, 07495, 10174, 10597, 10119, 10545, 11371,
## 10110, 07840, 07865, 07882, 10271, 18405, 18457, 18460, 10171, 10162, 10177,
## 10152, 11549, 18302, 10279, 07870, 08006, 10115, 11697, 10111, 10112, 10167,
## 10474, 11351, 11359, 11424, 11425, 11451, 07881, 18445, 18325, 08525, 10169,
## 10103, 10311, 10153, 10154, 07961, 10199, 11794, 10165, 10168, 11973, 10278,
```
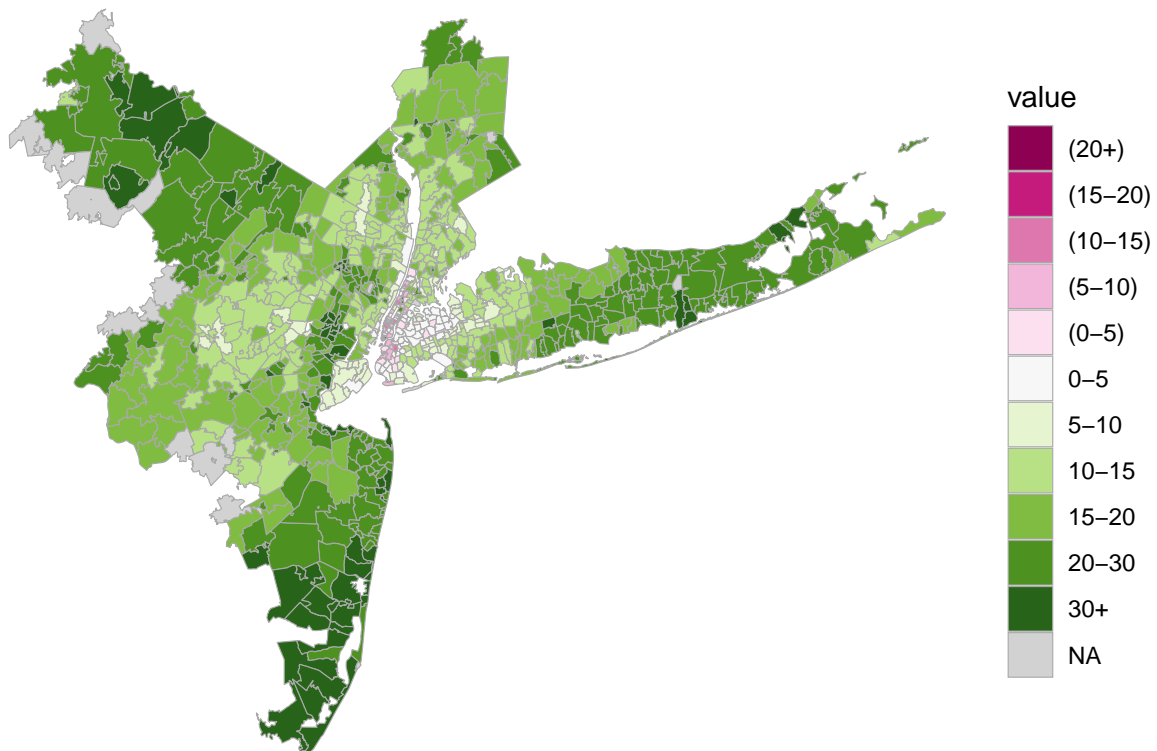
```
## 07311, 07608, 10173, 11109, 11556, 10170, 08890, 10172, 11005
```

% change in typical home value


```
zips_dc%>%
  select(RegionName, City, CountyName, Metro, State, perc_chg, zcta_density)%>%
  rename(region = RegionName,
         value = perc_chg)%>%
  mutate(value, original_value = value) -> dc_ch


dc_ch$value <- cut(dc_ch$value,
                   breaks = c(-Inf, -0.20, -0.15, -0.1, -0.05, 0, 0.05, 0.10, 0.15, 0.20, 0.30, Inf),
                   labels = c( "(20+)", "(15-20)", "(10-15)","(5-10)", "(0-5)", "0-5", "5-10", "10-15",
                   ordered_result = TRUE,
                   include.lowest = FALSE)


choro2 <-  ZipChoropleth$new(dc_ch)
choro2$title <- "% change in typical home value"
choro2$ggplot_scale <- scale_fill_manual(values = c(m20,m15,m10,m5,m0,p0,p5,p10,p15,p20,p30),
                                         drop = FALSE,
                                         na.value = "#d2d2d2")
choro2$set_zoom_zip(state_zoom = NULL, county_zoom = NULL, msa_zoom = "Washington-Arlington-Alexandria,
choro2$render()

## Warning in self$bind(): The following regions were missing and are being set to
## NA: 20777, 20053, 20418, 22611, 20230, 22620, 22185, 21787, 20593, 20307, 20540,
## 22655, 20659, 21719, 20405, 22645, 22663, 20506, 21705, 21776, 21783, 22211,
```
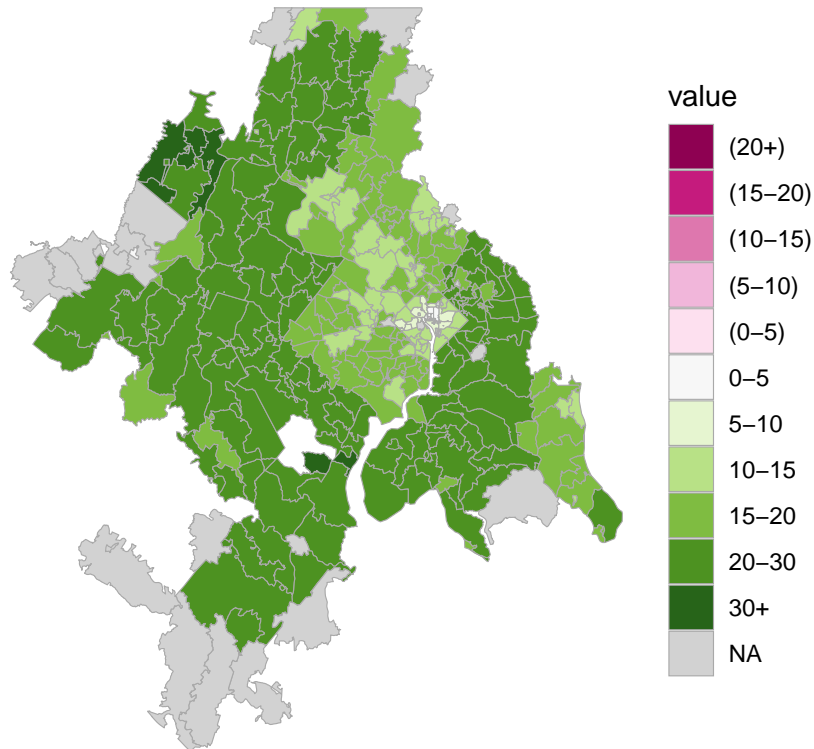
```
## 22401, 20427, 22508, 20240, 20317, 20510, 20560, 20762, 22580, 20118, 23015,
## 23024, 22035, 20057, 20245, 22657, 20045, 20373, 20535, 20565, 20064, 20553,
## 20228, 20899, 22046, 20202, 23117, 20204, 22214, 20566, 22960, 20319, 20742,
## 20551, 20260, 20520, 20390
```

% change in typical home value


```
zips_sfbay%>%
  select(RegionName, City, CountyName, Metro, State, perc_chg, zcta_density)%>%
  rename(region = RegionName,
         value = perc_chg)%>%
       mutate(value, original_value = value) -> sf_ch

sf_ch$value <- cut(sf_ch$value,
                 breaks = c(-Inf, -0.20, -0.15, -0.1, -0.05, 0, 0.05, 0.10, 0.15, 0.20, 0.30, Inf),
                 labels = c( "(20+)", "(15-20)", "(10-15)","(5-10)", "(0-5)", "0-5", "5-10", "10-15",
                 ordered_result = TRUE,
                 include.lowest = FALSE)

choro3 <-  ZipChoropleth$new(sf_ch)
choro3$title <- "Change in Home Value"
choro3$ggplot_scale <- scale_fill_manual(values = c(m20,m15,m10,m5,m0,p0,p5,p10,p15,p20,p30),
                                         drop = FALSE,
                                         na.value = "#d2d2d2")

choro3$set_zoom_zip(state_zoom = NULL, county_zoom = NULL, msa_zoom = sfbaymsas01, zip_zoom = NULL)
choro3$render()
```
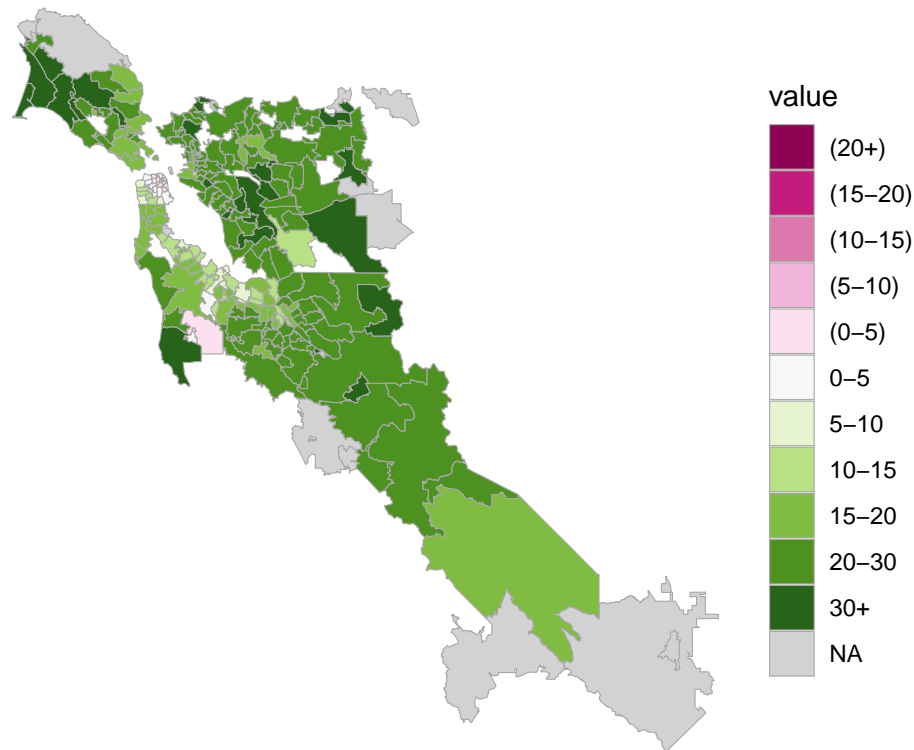
```
## Warning in self$bind(): The following regions were missing and are being set to
```

```
## NA: 94720, 94569, 94575, 94613, 95391, 94952, 95013, 94130, 95219, 95004, 95377,
## 94972, 94128, 95053, 93210, 93930, 95076, 94850
```

Change in Home Value


```r
zips_boston%>%
  select(RegionName, City, CountyName, Metro, State, perc_chg, zcta_density)%>%
  rename(region = RegionName,
         value = perc_chg)%>%
  mutate(value, original_value = value) -> bos_ch

bos_ch$value <- cut(bos_ch$value,
                    breaks = c(-Inf, -0.20, -0.15, -0.1, -0.05, 0, 0.05, 0.10, 0.15, 0.20, 0.30, Inf),
                    labels = c( "(20+)", "(15-20)", "(10-15)","(5-10)", "(0-5)", "0-5", "5-10", "10-15"
                    ordered_result = TRUE,
                    include.lowest = FALSE)
#construct choropleth
choro4 <-  ZipChoropleth$new(bos_ch)
choro4$title <- "Change in Home Value"
choro4$ggplot_scale <- scale_fill_manual(values = c(m20,m15,m10,m5,m0,p0,p5,p10,p15,p20,p30),
                                         drop = FALSE,
                                         na.value = "#d2d2d2")
choro4$set_zoom_zip(state_zoom = NULL, county_zoom = NULL, msa_zoom = "Boston-Cambridge-Newton, MA-NH",
choro4$render()
```
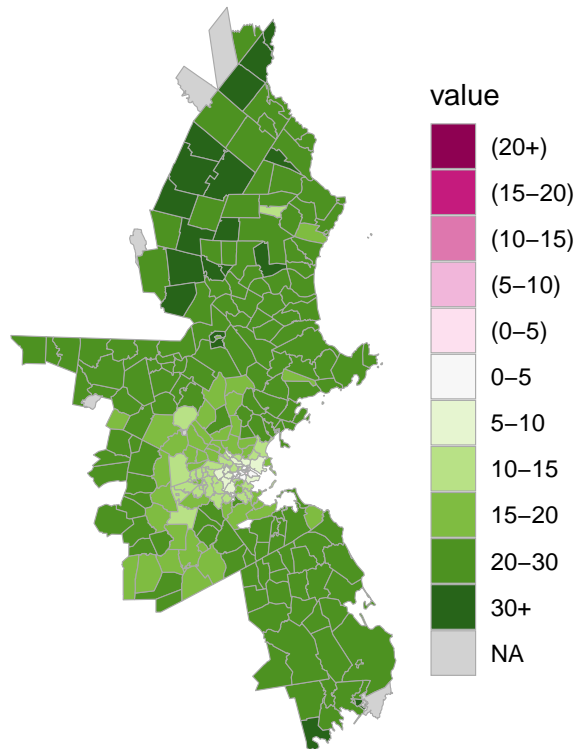
```
## Warning in self$bind(): The following regions were missing and are being set to
## NA: 03103, 02203, 02366, 03855, 03225, 02163, 02457, 02532, 01731, 01937, 01434
```

Change in Home Value


```
#CENTRAL COLORADO
centralcofips <- c(8001, 8005, 8014, 8019, 8031, 8035,  8039, 8059, 8093, 8047, 8013, 8119, 8041, 8117,

zips_centralco%>%
  select(RegionName, City, CountyName, Metro, State, perc_chg, zcta_density)%>%
  rename(region = RegionName,
         value = perc_chg)%>%
  mutate(value, original_value = value) -> co_ch

co_ch$value <- cut(co_ch$value,
                   breaks = c(-Inf, -0.20, -0.15, -0.1, -0.05, 0, 0.05, 0.10, 0.15, 0.20, 0.30, Inf),
                   labels = c( "(20+)", "(15-20)", "(10-15)","(5-10)", "(0-5)", "0-5", "5-10", "10-15",
                   ordered_result = TRUE,
                   include.lowest = FALSE)


#choropleth
choro5 <-  ZipChoropleth$new(co_ch)
choro5$title <- "Change in Mean Home Value"
choro5$ggplot_scale <- scale_fill_manual(values = c(m20,m15,m10,m5,m0,p0,p5,p10,p15,p20,p30),
                                         drop = FALSE,
                                         na.value = "#d2d2d2")
choro5$set_zoom_zip(state_zoom = NULL, county_zoom = centralcofips, msa_zoom = NULL, zip_zoom = NULL)
choro5$render()

## Warning in self$bind(): The following regions were missing and are being set to
## NA: 80544, 80733, 80432, 80427, 80723, 80476, 81008, 80456, 80448, 80045, 80750,
## 80742, 80828, 80840, 80902, 80444, 80438, 80107, 80741, 80654, 80475, 80653,
```
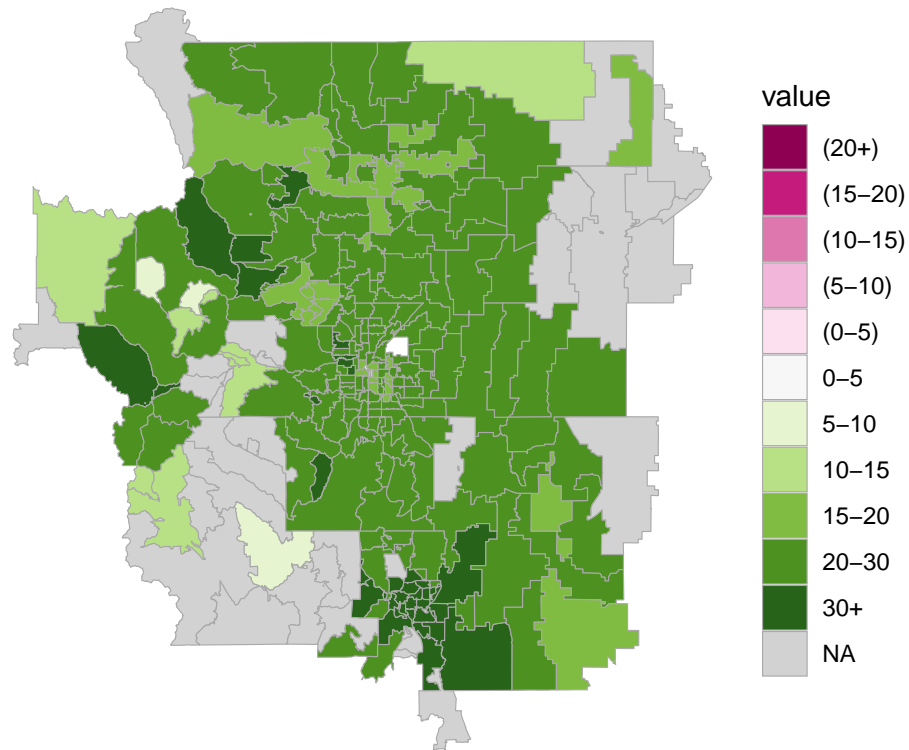
```
## 80310, 80913, 80701, 80705, 80814, 80816, 80820, 80290, 80419, 80421, 80422,
## 80914, 80449, 80939, 82063, 80293, 80863, 80813, 80294, 80423, 80264
```

Change in Mean Home Value


```
#choropleths for Population Density

#NYC:
zips_newyork%>%
  select(RegionName, City, CountyName, Metro, State, perc_chg, zcta_density)%>%
  rename(region = RegionName)%>%
  mutate(zcta_density_2010, "value" = zcta_density_2010) -> ny_ch1



choro6 <-  ZipChoropleth$new(ny_ch1)
choro6$title <- "Population Density"
choro6$legend_title <- "%"
choro6$ggplot_scale <- scale_fill_brewer(palette = "GnBu")
choro6$set_zoom_zip(state_zoom = NULL, county_zoom = NULL, msa_zoom = "New York-Newark-Jersey City, NY-I
choro6$render()


#DC

zips_dc%>%
  select(RegionName, City, CountyName, Metro, State, perc_chg, zcta_density)%>%
  rename(region = RegionName)%>%
  mutate(zcta_density_2010, "value" = zcta_density_2010) -> dc_ch1
```

```r
choro7 <-  ZipChoropleth$new(dc_ch1)
choro7$title <- "Population Density"
choro7$ggplot_scale <- scale_fill_brewer(palette = "GnBu")
choro7$set_zoom_zip(state_zoom = NULL, county_zoom = NULL, msa_zoom = "Washington-Arlington-Alexandria,
choro7$render()


#SF

zips_sfbay%>%
  select(RegionName, City, CountyName, Metro, State, perc_chg, zcta_density)%>%
  rename(region = RegionName)%>%
  mutate(zcta_density_2010, "value" = zcta_density_2010) -> sf_ch1


#choropleth
choro8 <-  ZipChoropleth$new(sf_ch1)
choro8$title <- "Population Density"
choro8$ggplot_scale <- scale_fill_brewer(palette = "GnBu")
choro8$set_zoom_zip(state_zoom = NULL, county_zoom = NULL, msa_zoom = sfbaymsas01, zip_zoom = NULL)
choro8$render()


#Boston

zips_boston%>%
  select(RegionName, City, CountyName, Metro, State, perc_chg, zcta_density)%>%
  rename(region = RegionName)%>%
  mutate(zcta_density_2010, "value" = zcta_density_2010) -> bos_ch1

#construct choropleth
choro9 <-  ZipChoropleth$new(bos_ch1)
choro9$title <- "Population Density"
choro9$ggplot_scale <- scale_fill_brewer(palette = "GnBu")
choro9$set_zoom_zip(state_zoom = NULL, county_zoom = NULL, msa_zoom = "Boston-Cambridge-Newton, MA-NH",
choro9$render()


#Colorado

zips_centralco%>%
  select(RegionName, City, CountyName, Metro, State, perc_chg, zcta_density)%>%
  rename(region = RegionName)%>%
  mutate(zcta_density_2010, "value" = zcta_density_2010) -> co_ch1


#choropleth
choro10 <-  ZipChoropleth$new(co_ch1)
choro10$title <- "Population Density"
choro10$ggplot_scale <- scale_fill_brewer(palette = "GnBu")
choro10$set_zoom_zip(state_zoom = NULL, county_zoom = centralcofips, msa_zoom = NULL, zip_zoom = NULL)
choro10$render()
```

PART 4: Data Analysis & Regressions

```
weighted.mean(zips_newyork$perc_chg_rounded, zips_newyork$zpop, na.rm = TRUE)
```

```
## [1] 13.11669
```

```
weighted.mean(zips_boston$perc_chg_rounded, zips_boston$zpop, na.rm = TRUE)
```

```
## [1] 20.24371
```

```
weighted.mean(zips_sfbay$perc_chg_rounded, zips_sfbay$zpop, na.rm = TRUE)
```

```
## [1] 20.54377
```

```
weighted.mean(zips_dc$perc_chg_rounded, zips_dc$zpop, na.rm = TRUE)
```

```
## [1] 18.286
```

```
weighted.mean(zips_centralco$perc_chg_rounded, zips_centralco$zpop, na.rm = TRUE)
```

```
## [1] 25.80753
```