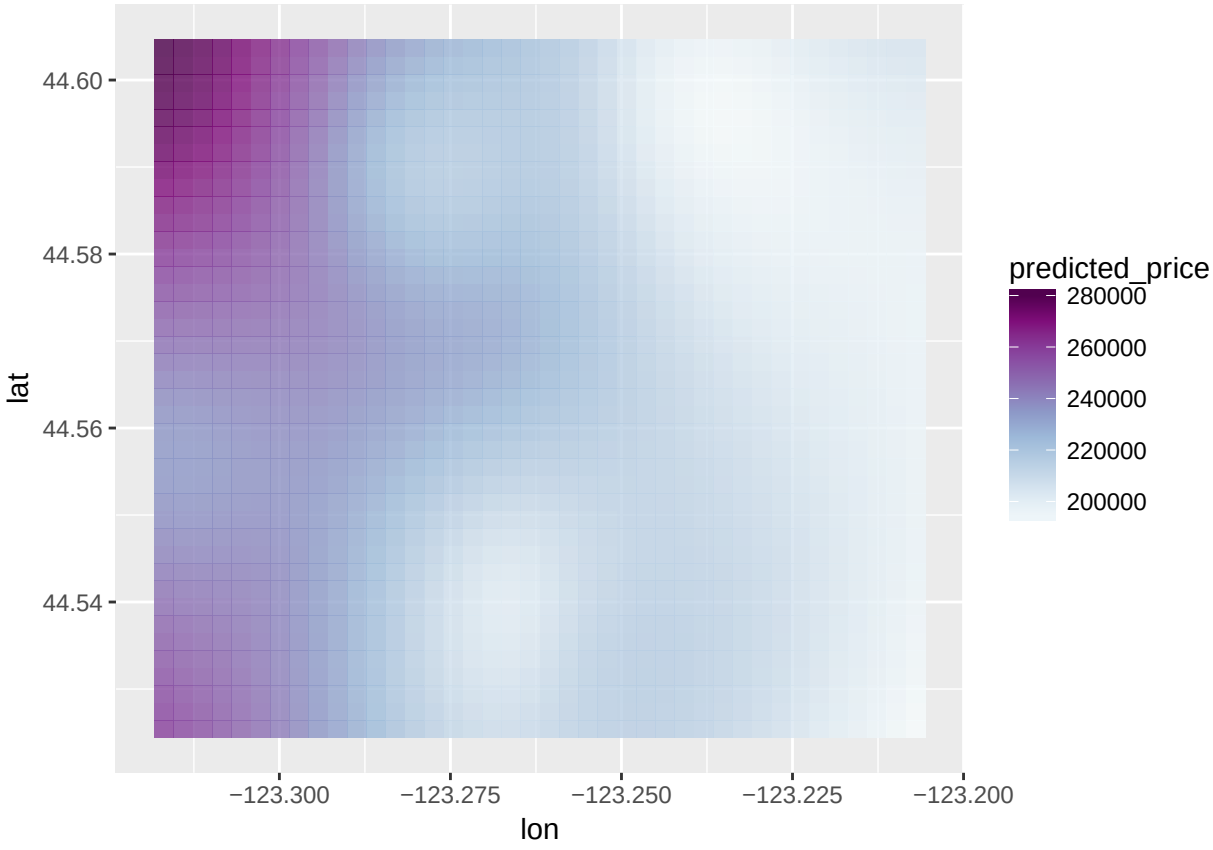
R for Spatial Analysis

Aaron Kleyn

May 17, 2018

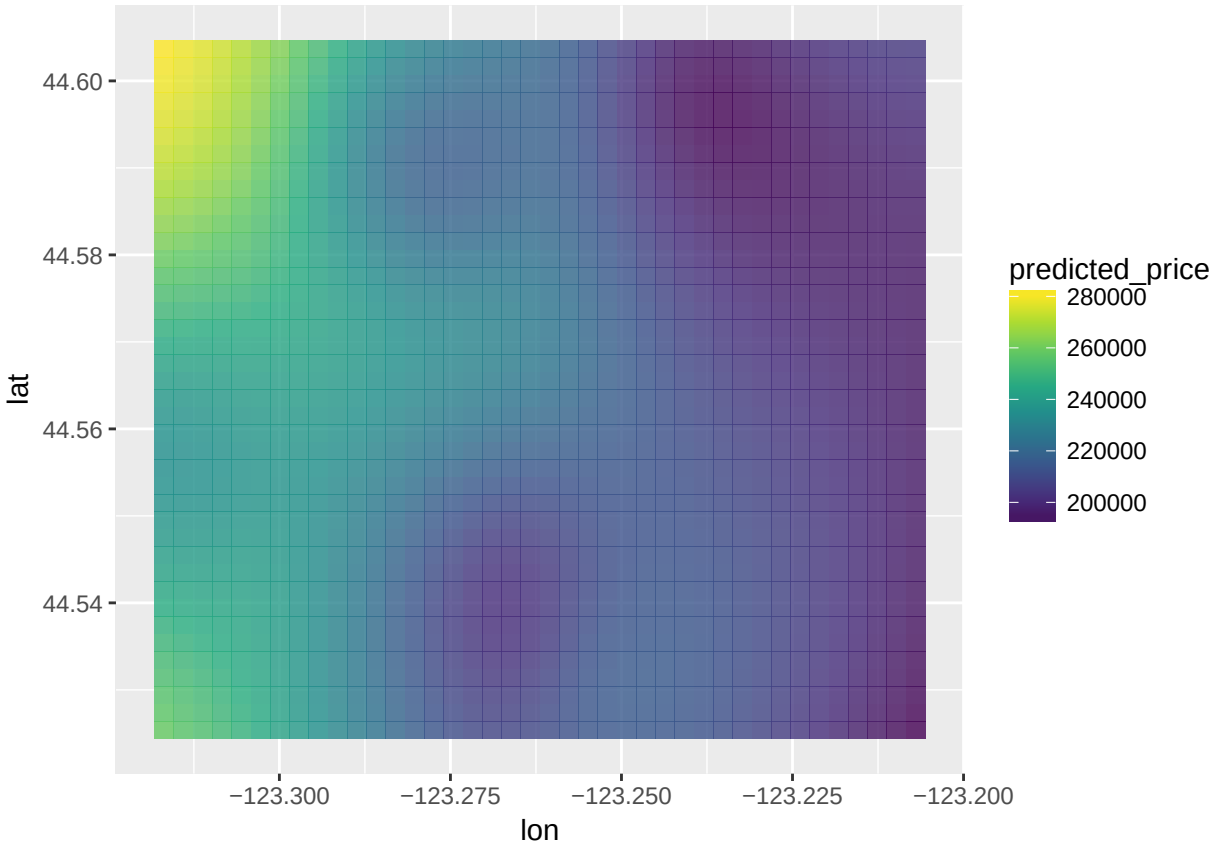
The following three plots will be used to construct a demographic plot New York City and the surrounding regions.

```
library(ggplot2)
library(sp)
library(RColorBrewer)
library(raster)
library(tmap)


# 9 steps on the RColorBrewer "BuPu" palette: blups
blups <- brewer.pal(9, "BuPu")

# Add scale_fill_gradientn() with the blups palette
ggplot(preds) +
  geom_tile(aes(lon, lat, fill = predicted_price), alpha = 0.8) +
  scale_fill_gradientn(colors = blups)
```




```
library(viridisLite)
# viridisLite viridis palette with 9 steps: vir
```

```
vir <- viridis(9)

# Add scale_fill_gradientn() with the vir palette
ggplot(preds) +
  geom_tile(aes(lon, lat, fill = predicted_price), alpha = 0.8) +
  scale_fill_gradientn(colors = vir)
```




```
# mag: a viridisLite magma palette with 9 steps
mag <- viridisLite::magma(7)
```

```
library(classInt)
```

```
## Loading required package: spData
```

```
## To access larger datasets in this package, install the spDataLarge
## package with: `install.packages('spDataLarge',
## repos='https://nowosad.github.io/drat/', type='source'))`
```

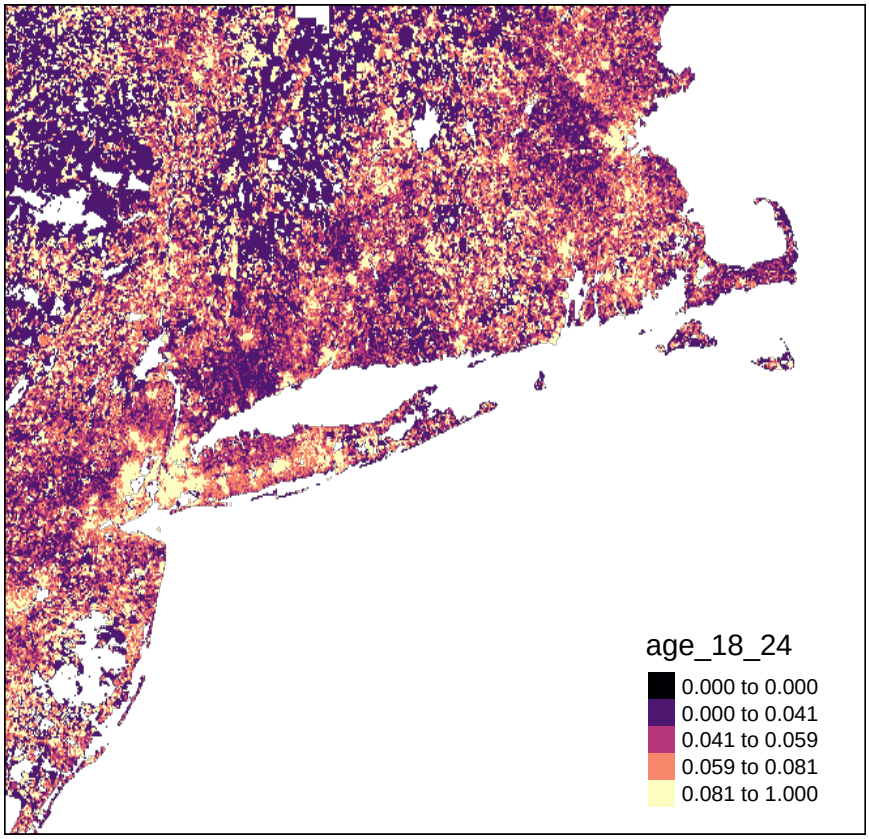
```
library(RColorBrewer)

# Create 5 "pretty" breaks with classIntervals()
classIntervals(values(prop_by_age[["age_18_24"]]),
               n = 5, style = "pretty")
```

```
## style: pretty
##   [0,0.2) [0.2,0.4) [0.4,0.6) [0.6,0.8)   [0.8,1]
##    130770      1775       302       170       138
```

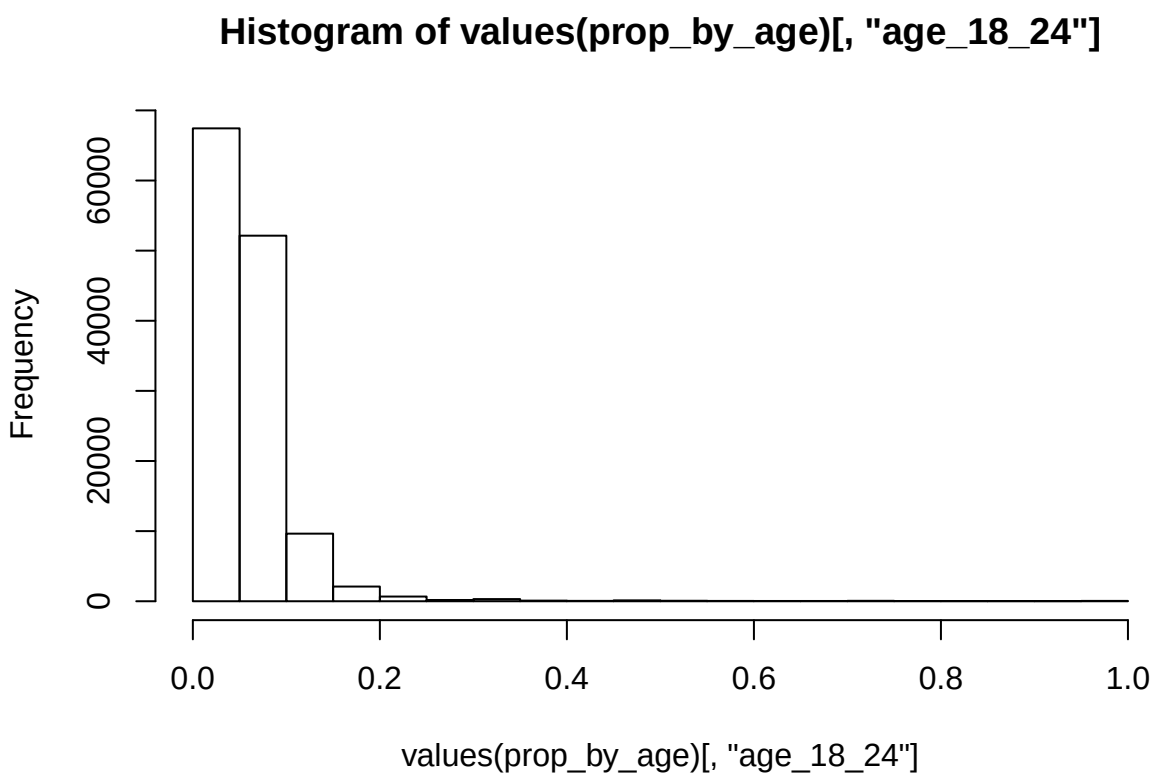
```
# Create 5 "quantile" breaks with classIntervals()
classIntervals(values(prop_by_age[["age_18_24"]]),
               n = 5, style = "quantile")
```

```
## style: quantile
##                   [0,0)          [0,0.04054054) [0.04054054,0.05882353)
##                       0                   53218                   25925
## [0.05882353,0.08108108)          [0.08108108,1]
##                   27217                   26795
```

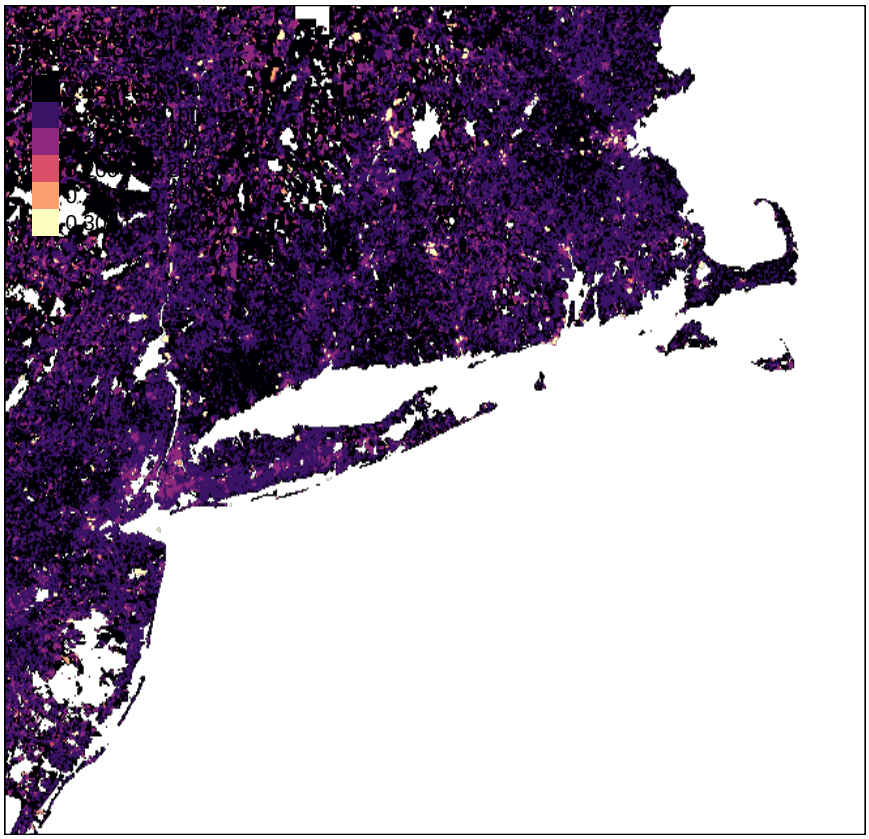
```
# Use 5 "quantile" breaks in tm_raster()
tm_shape(prop_by_age) +
  tm_raster("age_18_24", palette = mag, style = "quantile") +
  tm_legend(position = c("right", "bottom"))
```




```
# Create histogram of proportions
hist(values(prop_by_age)[, "age_18_24"])
```




```
# Use fixed breaks in tm_raster()
tm_shape(prop_by_age) +
  tm_raster("age_18_24", palette = mag,
    style = "fixed", breaks = c(0.025, 0.05, 0.1, 0.2, 0.25, 0.3, 1))
```



```
# Save your plot to "prop_18-24.html"
save_tmap(filename = "prop_18-24.html")
```

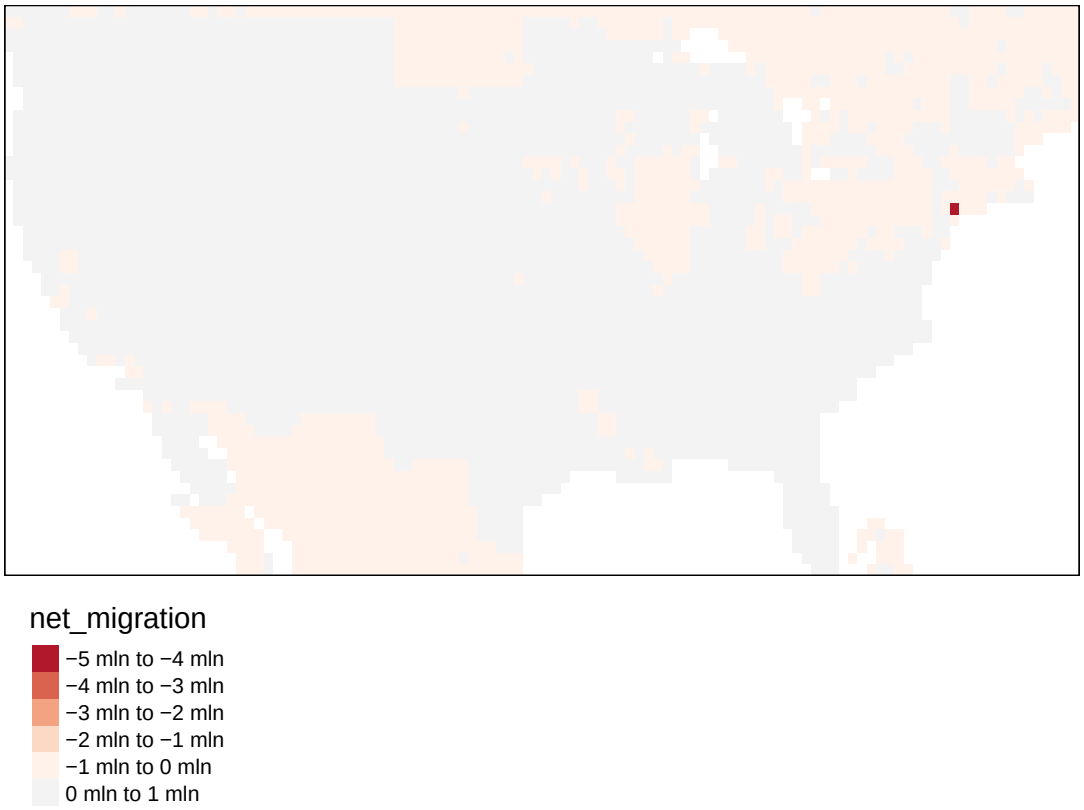
```
## Interactive map saved to C:\Users\Aaron\Documents\prop_18-24.html
```

The following code will derive a plot of net migration in the US from 1990 - 2000

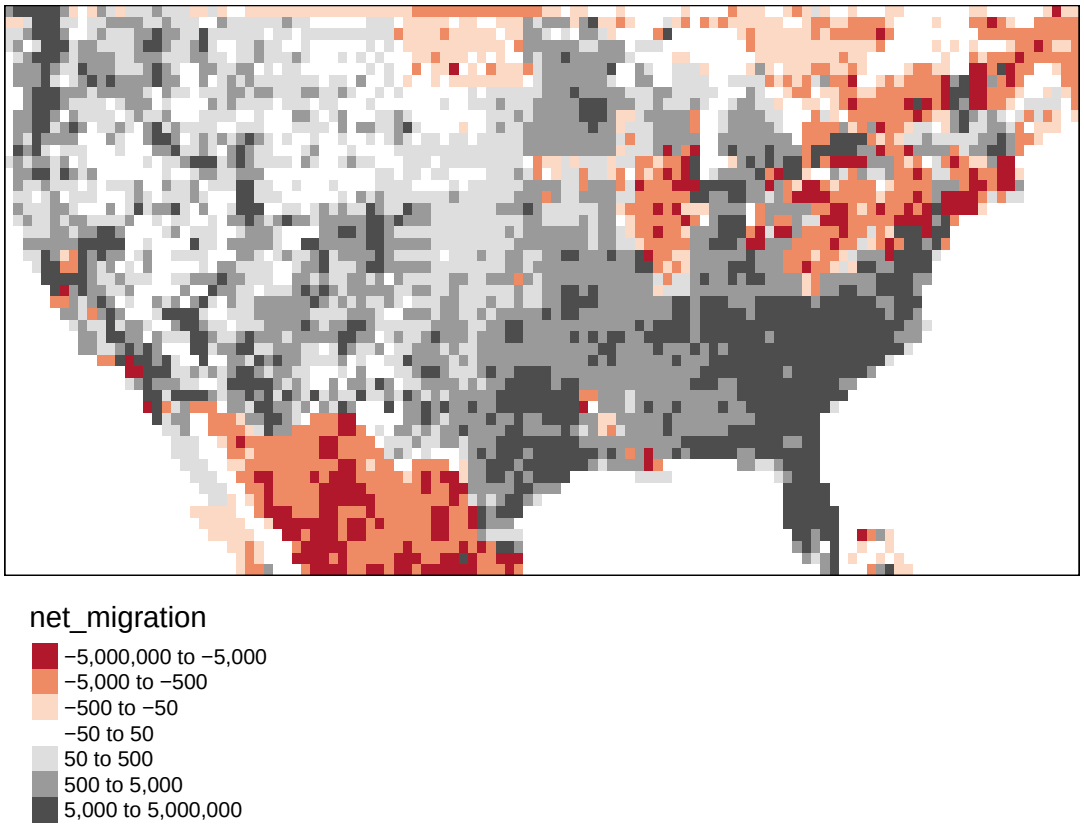
```
# Print migration

migration
```

```
## class       : RasterLayer
## dimensions  : 49, 116, 5684  (nrow, ncol, ncell)
## resolution  : 0.5, 0.5  (x, y)
## extent      : -125, -67, 25, 49.5  (xmin, xmax, ymin, ymax)
## coord. ref. : +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0
## data source : in memory
## names       : net_migration
## values      : -4560234, 806052.2  (min, max)
```

```
# Diverging "RdGy" palette
red_gray <- brewer.pal(7, "RdGy")

# Use red_gray as the palette
tm_shape(migration) +
  tm_raster(palette = red_gray) +
  tm_legend(outside = TRUE, outside.position = c("bottom"))
```

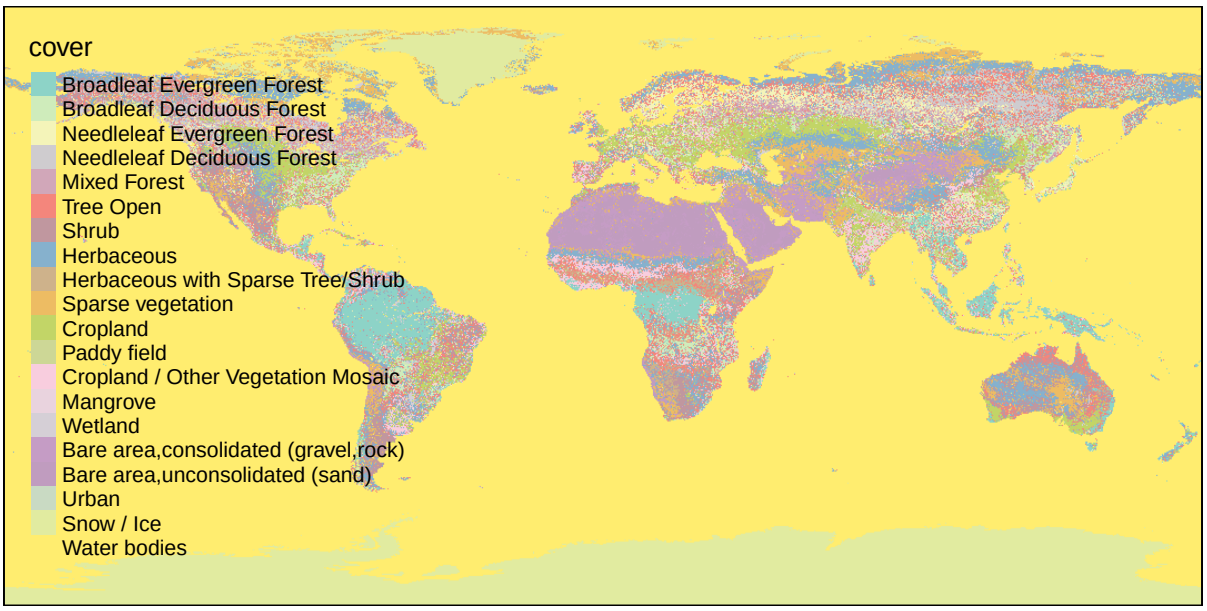



```
# Add fixed breaks
tm_shape(migration) +
  tm_raster(palette = red_gray, style = "fixed",
     breaks = c(-5e6, -5e3, -5e2, -5e1, 5e1, 5e2, 5e3, 5e6)) +
  tm_legend(outside = TRUE, outside.position = c("bottom"))
```

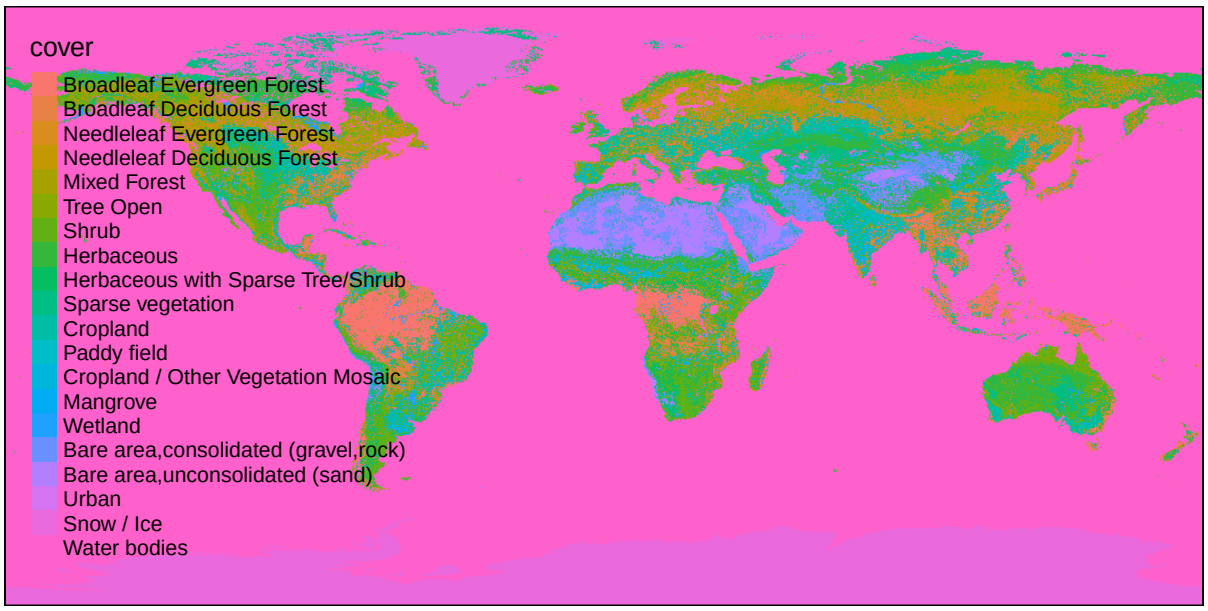


The following R code will derive a plot of vegetation/climate cover based on longitudinal and latitudinal coordinates.

```
library(raster)

# Plot land_cover
tm_shape(land_cover) +
  tm_raster()
```



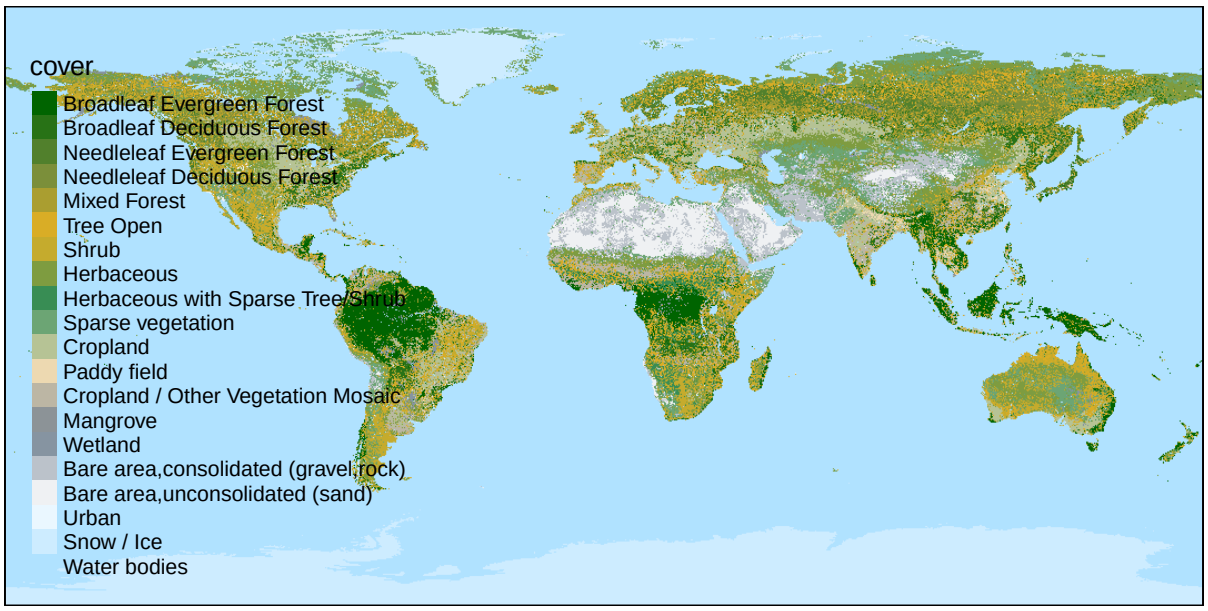

```
# Palette like the ggplot2 default
hcl_cols <- hcl(h = seq(15, 375, length = 9),
                c = 100, l = 65)[-9]

# Use hcl_cols as the palette
tm_shape(land_cover) +
  tm_raster(palette = hcl_cols)
```




```
# Examine levels of land_cover
levels(land_cover)
```

```
## NULL
```

```
# A set of intuitive colors
intuitive_cols <- c(
  "darkgreen",
  "darkolivegreen4",
  "goldenrod2",
  "seagreen",
  "wheat",
  "slategrey",
  "white",
  "lightskyblue1"
)

# Use intuitive_cols as palette
tm_shape(land_cover) +
  tm_raster(palette = intuitive_cols) +
  tm_legend(position = c("left", "bottom"))
```



The following R code will be used to develop a plot of average annual income in Manhattan.

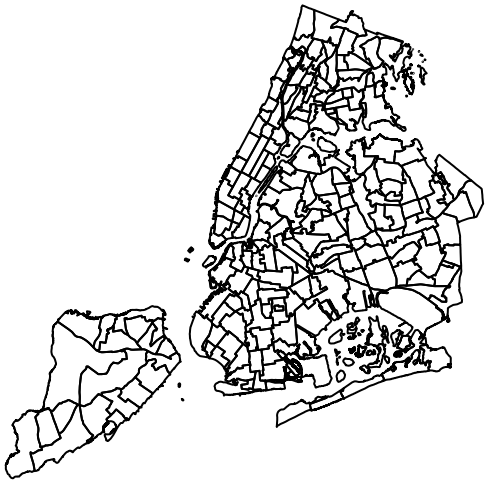
```
library(sp)
library(rgdal)

# Use dir() to find directory name
dir()

# Call dir() with directory name
dir('nynta_16c')

# Read in shapefile with readOGR(): neighborhoods
neighborhoods <- readOGR('nynta_16c', 'nynta')
```

```
# summary() of neighborhoods
summary(neighborhoods)
```

```
## Object of class SpatialPolygonsDataFrame
## Coordinates:
##        min       max
## x 913175.1 1067382.5
## y 120121.9  272844.3
## Is projected: TRUE
## proj4string :
## [+proj=lcc +lat_1=40.66666666666666 +lat_2=41.03333333333333
## +lat_0=40.16666666666666 +lon_0=-74 +x_0=300000 +y_0=0
## +datum=NAD83 +units=us-ft +no_defs +ellps=GRS80 +towgs84=0,0,0]
```

```
## Data attributes:
##     BoroCode          BoroName  CountyFIPS    NTACode
##  Min.   :1   Bronx        :38   005:38     BK09   :  1
##  1st Qu.:2   Brooklyn     :51   047:51     BK17   :  1
##  Median :3   Manhattan    :29   061:29     BK19   :  1
##  Mean   :3   Queens       :58   081:58     BK21   :  1
##  3rd Qu.:4   Staten Island:19   085:19     BK23   :  1
##  Max.   :5                                BK25   :  1
##                                           (Other):189
##                                       NTAName      Shape_Leng
##  Airport                                  :  1   Min.   : 11000
##  Allerton-Pelham Gardens                  :  1   1st Qu.: 23824
##  Annadale-Huguenot-Prince's Bay-Eltingville:  1   Median : 30550
##  Arden Heights                            :  1   Mean   : 42003
##  Astoria                                  :  1   3rd Qu.: 41877
##  Auburndale                               :  1   Max.   :490474
##  (Other)                                  :189
##    Shape_Area
##  Min.   :  5573902
##  1st Qu.: 19392084
##  Median : 32629789
##  Mean   : 43227674
##  3rd Qu.: 50237449
##  Max.   :327759719
##
```

```r
# Plot neighboorhoods
plot(neighborhoods)
```



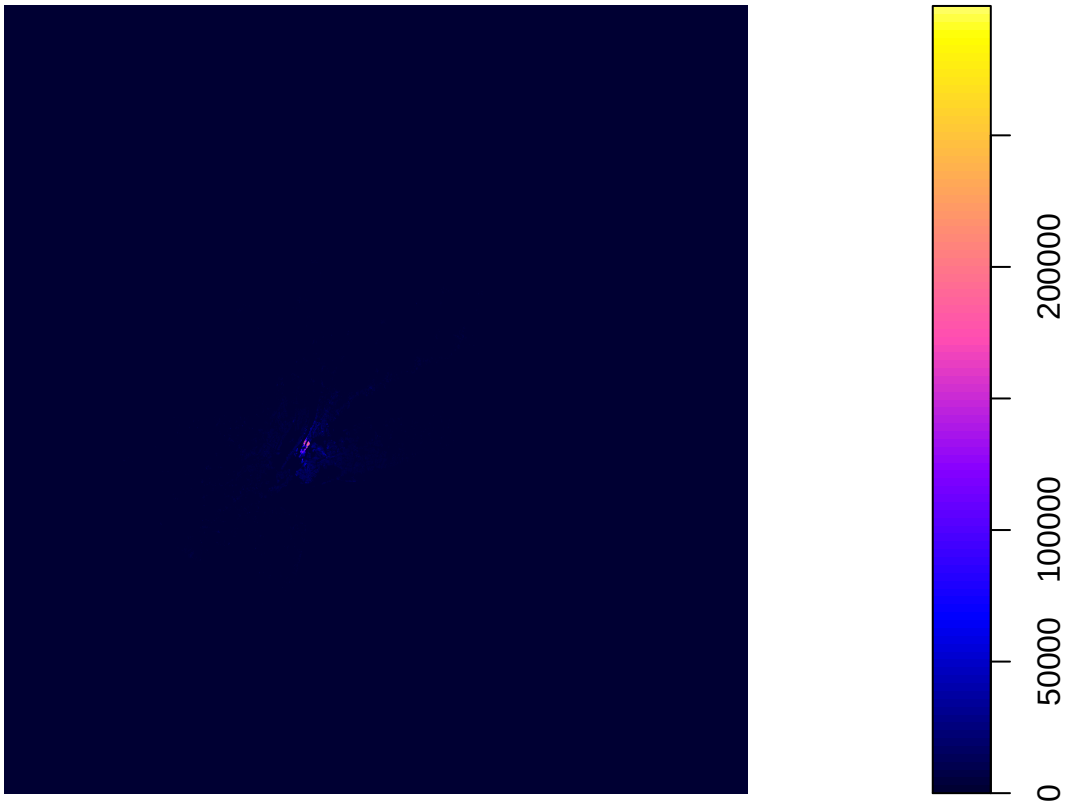
```
library(raster)

# Call dir()
dir()

# Call dir() on the directory
dir('nyc_grid_data')

# Use raster() with file path: income_grid
income_grid <- raster('nyc_grid_data/m5602ahhi00.tif')
```

```
# Call summary() on income_grid
summary(income_grid)
```

```
## Object of class SpatialGridDataFrame
## Coordinates:
##   min max
## x -76 -71
## y  39  43
## Is projected: FALSE
## proj4string :
## [+proj=longlat +datum=NAD83 +no_defs +ellps=GRS80 +towgs84=0,0,0]
## Grid attributes:
##   cellcentre.offset    cellsize cells.dim
## x         -75.99896 0.002083333      2400
## y          39.00104 0.002083333      1920
## Data attributes:
```

```
##      band1
##  Min.   :     0.0
##  1st Qu.:     0.0
##  Median :     0.0
##  Mean   :   120.9
##  3rd Qu.:     0.0
##  Max.   :299176.0
```

```
# Call plot() on income_grid
plot(income_grid)
```




```
library(sp)
library(tigris)
# Call tracts(): nyc_tracts
nyc_tracts <- tracts(state = 'NY', county = 'New York', cb = TRUE)
```

```
##
  |
  |                                                                      |   0%
  |
  |=                                                                     |   1%
  |
  |=                                                                     |   2%
  |
  |==                                                                    |   3%
  |
  |===                                                                   |   4%
```

```
  |
  |================================================                 |  78%
  |
  |=================================================                |  79%
  |
  |==================================================               |  80%
  |
  |===================================================              |  81%
  |
  |====================================================             |  82%
  |
  |=====================================================            |  83%
  |
  |======================================================           |  84%
  |
  |=======================================================          |  85%
  |
  |========================================================         |  86%
  |
  |=========================================================        |  87%
  |
  |==========================================================       |  87%
  |
  |===========================================================      |  88%
  |
  |============================================================     |  89%
  |
  |=============================================================    |  90%
  |
  |==============================================================   |  91%
  |
  |===============================================================  |  92%
  |
  |================================================================ |  93%
  |
  |================================================================ |  94%
  |
  |================================================================= |  95%
  |
  |================================================================= |  96%
  |
  |================================================================== |  97%
  |
  |==================================================================== |  98%
  |
  |===================================================================== |  99%
  |
  |======================================================================| 100%
```

```
# proj4string() on nyc_tracts and neighborhoods
proj4string(nyc_tracts)
```

```
## [1] "+proj=longlat +datum=NAD83 +no_defs +ellps=GRS80 +towgs84=0,0,0"
```

```r
proj4string(neighborhoods)
```

```
## [1] "+proj=lcc +lat_1=40.66666666666666 +lat_2=41.03333333333333 +lat_0=40.16666666666666 +lon_0=-74
```

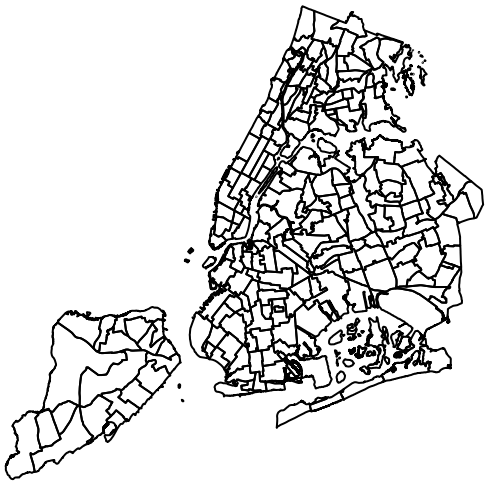
```r
# coordinates() on nyc_tracts and neighborhoods
head(coordinates(nyc_tracts))
```

```
##         [,1]     [,2]
## 37 -74.01297 40.70905
## 38 -73.98114 40.72126
## 39 -73.98146 40.72901
## 40 -73.99708 40.72762
## 41 -74.00411 40.73046
## 42 -73.97467 40.74632
```

```r
head(coordinates(neighborhoods))
```

```
##        [,1]     [,2]
## 0  987397.5 169148.4
## 1 1037640.2 214077.6
## 2 1043002.7 212969.8
## 3 1037005.2 219265.3
## 4 1020706.7 217413.9
## 5 1011274.1 240777.4
```

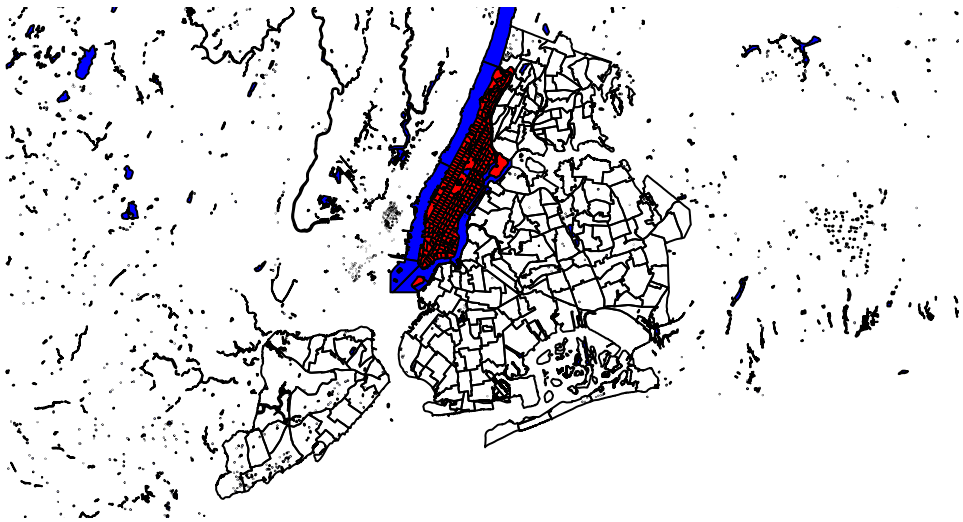
```r
# plot() neighborhoods and nyc_tracts
plot(neighborhoods)
plot(nyc_tracts, add = TRUE, col = 'red')
```



```
library(sp)
library(raster)
library(tigris)
# Call tracts(): nyc_tracts
nyc_tracts <- tracts(state = 'NY', county = 'New York', cb = TRUE)
# Use spTransform on neighborhoods: neighborhoods
neighborhoods <- spTransform(neighborhoods,
                             proj4string(nyc_tracts))

# head() on coordinates() of neighborhoods
head(coordinates(neighborhoods))
```

```
##        [,1]     [,2]
## 0 -73.98866 40.63095
## 1 -73.80729 40.75411
## 2 -73.78795 40.75104
## 3 -73.80955 40.76835
## 4 -73.86840 40.76335
## 5 -73.90235 40.82751
```

```
# Plot neighborhoods, nyc_tracts and water
plot(neighborhoods)
plot(nyc_tracts, add = TRUE, col = "red")
plot(water, add = TRUE, col = "blue")
```



```
library(sp)
library(tigris)
```

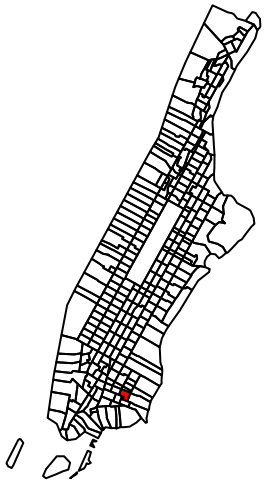
```r
# Call tracts(): nyc_tracts
nyc_tracts <- tracts(state = 'NY', county = 'New York', cb = TRUE)
# Use str() on nyc_income and nyc_tracts@data
str(nyc_income)
```

```
## 'data.frame':    288 obs. of  6 variables:
##  $ name    : chr  "Census Tract 1, New York County, New York" "Census Tract 2.01, New York County, Ne
##  $ state   : int  36 36 36 36 36 36 36 36 36 36 ...
##  $ county  : int  61 61 61 61 61 61 61 61 61 61 ...
##  $ tract   : chr  "000100" "000201" "000202" "000500" ...
##  $ estimate: num  NA 23036 29418 NA 18944 ...
##  $ se      : num  NA 3083 1877 NA 1442 ...
```

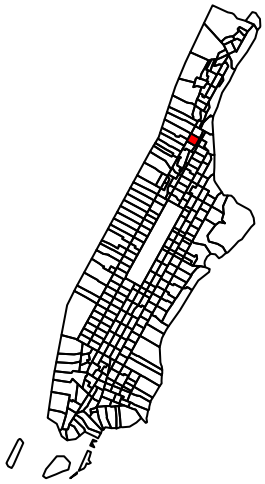
```r
str(nyc_tracts@data)
```

```
## 'data.frame':    288 obs. of  9 variables:
##  $ STATEFP : chr  "36" "36" "36" "36" ...
##  $ COUNTYFP: chr  "061" "061" "061" "061" ...
##  $ TRACTCE : chr  "001300" "002202" "003400" "005501" ...
##  $ AFFGEOID: chr  "1400000US36061001300" "1400000US36061002202" "1400000US36061003400" "1400000US360
##  $ GEOID   : chr  "36061001300" "36061002202" "36061003400" "36061005501" ...
##  $ NAME    : chr  "13" "22.02" "34" "55.01" ...
##  $ LSAD    : chr  "CT" "CT" "CT" "CT" ...
##  $ ALAND   : chr  "312732" "56112" "160257" "124447" ...
##  $ AWATER  : chr  "0" "0" "0" "0" ...
```

```r
# Highlight tract 002201 in nyc_tracts
plot(nyc_tracts)
plot(nyc_tracts[nyc_tracts$TRACTCE == "002201", ],
     col = "red", add = TRUE)
```



```r
# Set nyc_tracts@data to nyc_income
nyc_tracts@data <- nyc_income

# Highlight tract 002201 in nyc_tracts
plot(nyc_tracts)
plot(nyc_tracts[nyc_tracts$tract == "002201", ],
     col = "red", add = TRUE)
```



```
library(tigris)
# Call tracts(): nyc_tracts
nyc_tracts <- tracts(state = 'NY', county = 'New York', cb = TRUE)
# Check for duplicates in nyc_income
any(duplicated(nyc_income$tract))
```

```
## [1] FALSE
```

```
# Check for duplicates in nyc_tracts
any(duplicated(nyc_tracts$TRACTCE))
```

```
## [1] FALSE
```

```
# Check nyc_tracts in nyc_income
all((nyc_tracts$TRACTCE %in% nyc_income$tract))
```

```
## [1] TRUE
```

```
# Check nyc_income in nyc_tracts
all((nyc_income$tract %in% nyc_tracts$TRACTCE))
```

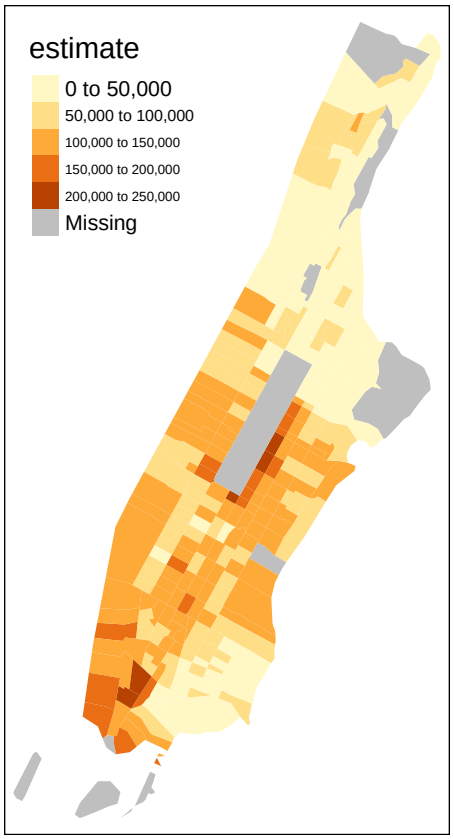
```
## [1] TRUE
```

```
library(sp)
library(tmap)
library(tigris)
# Call tracts(): nyc_tracts
nyc_tracts <- tracts(state = 'NY', county = 'New York', cb = TRUE)
# Merge nyc_tracts and nyc_income: nyc_tracts_merge
```

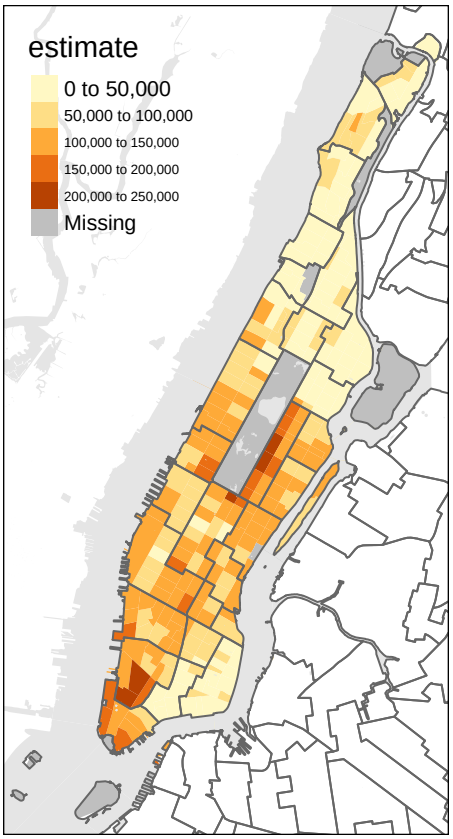
```r
nyc_tracts_merge <- sp::merge(nyc_tracts, nyc_income, by.x = "TRACTCE", by.y = "tract")

# Call summary() on nyc_tracts_merge
summary(nyc_tracts_merge)
```

```
## Object of class SpatialPolygonsDataFrame
## Coordinates:
##         min       max
## x -74.04731 -73.90700
## y  40.68419  40.88207
## Is projected: FALSE
## proj4string :
## [+proj=longlat +datum=NAD83 +no_defs +ellps=GRS80 +towgs84=0,0,0]
## Data attributes:
##    TRACTCE            STATEFP            COUNTYFP
##  Length:288         Length:288         Length:288
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##    AFFGEOID            GEOID               NAME
##  Length:288         Length:288         Length:288
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##      LSAD              ALAND              AWATER
##  Length:288         Length:288         Length:288
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##      name               state        county      estimate
##  Length:288         Min.   :36   Min.   :61   Min.   : 12479
##  Class :character   1st Qu.:36   1st Qu.:61   1st Qu.: 39038
##  Mode  :character   Median :36   Median :61   Median : 81786
##                     Mean   :36   Mean   :61   Mean   : 82405
##                     3rd Qu.:36   3rd Qu.:61   3rd Qu.:112561
##                     Max.   :36   Max.   :61   Max.   :232266
##                                               NA's   :9
##        se
##  Min.   :  1117
##  1st Qu.:  5107
##  Median :  8998
##  Mean   : 11678
##  3rd Qu.: 14281
##  Max.   :132737
##  NA's   :9
```

```
# Choropleth with col mapped to estimate
tm_shape(nyc_tracts_merge) +
  tm_fill(col = "estimate")
```



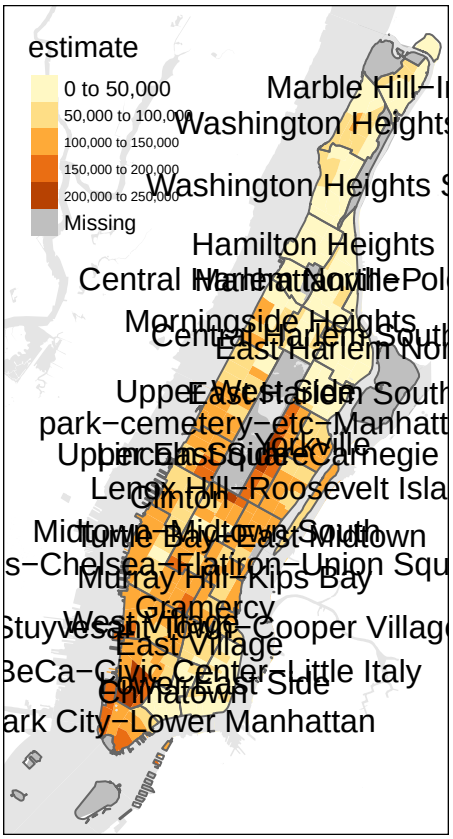

```
library(tmap)
library(tigris)
tm_shape(nyc_tracts_merge) +
  tm_fill(col = "estimate") +
  # Add a water layer, tm_fill() with col = "grey90"
  tm_shape(water) +
  tm_fill(col = "grey90") +
  # Add a neighborhood layer, tm_borders()
  tm_shape(neighborhoods) +
  tm_borders()
```



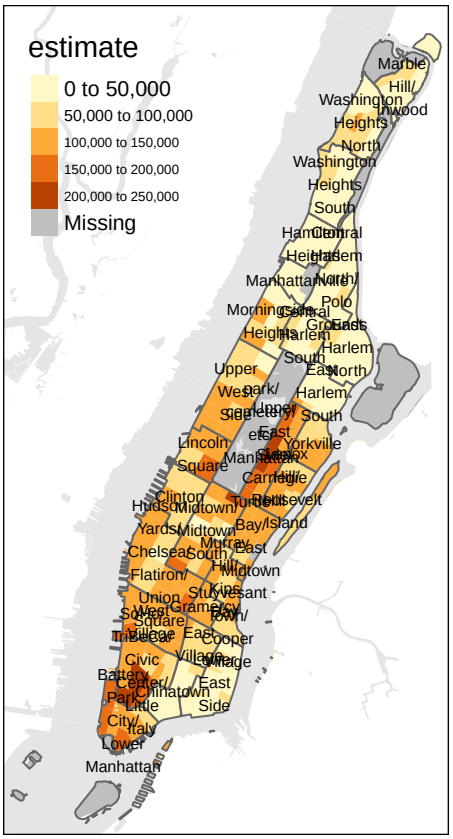

```
library(tmap)

# Find unique() nyc_tracts_merge$COUNTYFP
unique(nyc_tracts_merge$COUNTYFP)
```

```
## [1] "061"
```

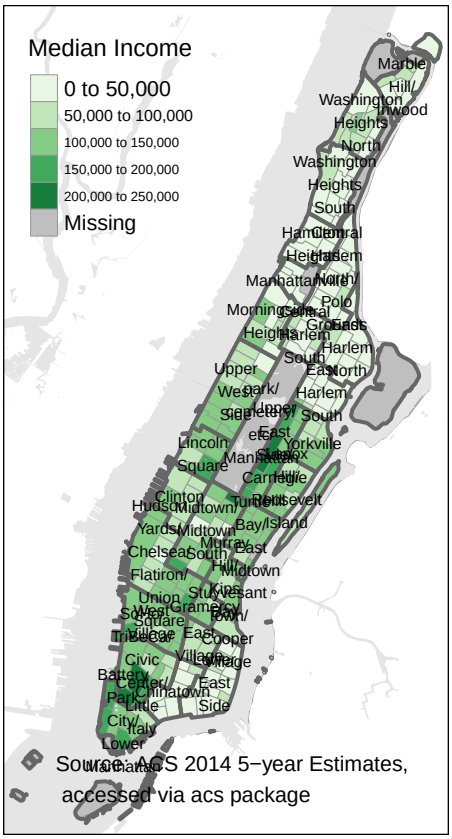
```
# Add logical expression to pull out New York County
manhat_hoods <- neighborhoods[neighborhoods$CountyFIPS == "061", ]

tm_shape(nyc_tracts_merge) +
  tm_fill(col = "estimate") +
  tm_shape(water) +
  tm_fill(col = "grey90") +
  # Edit to use manhat_hoods instead
  tm_shape(manhat_hoods) +
  tm_borders() +
  # Add a tm_text() layer
  tm_text(text = "NTAName")
```




```
library(tmap)

# gsub() to replace " " with "\n"
manhat_hoods$name <- gsub(" ", "\n", manhat_hoods$NTAName)

# gsub() to replace "-" with "/\n"
manhat_hoods$name <- gsub("-", "/\n", manhat_hoods$name)

# Edit to map text to name, set size to 0.5
tm_shape(nyc_tracts_merge) +
    tm_fill(col = "estimate") +
  tm_shape(water) +
    tm_fill(col = "grey90") +
  tm_shape(manhat_hoods) +
    tm_borders() +
    tm_text(text = "name", size = 0.5)
```




```
library(tmap)

tm_shape(nyc_tracts_merge) +
  # Add title and change palette
  tm_fill(col = "estimate",
          title = "Median Income",
          palette = "Greens") +
  # Add tm_borders()
  tm_borders(col = "grey60", lwd = 0.5) +
  tm_shape(water) +
  tm_fill(col = "grey90") +
  tm_shape(manhat_hoods) +
  # Change col and lwd of neighborhood boundaries
  tm_borders(col = "grey40", lwd = 2) +
  tm_text(text = "name", size = 0.5) +
  # Add tm_credits()
  tm_credits("Source: ACS 2014 5-year Estimates, \n accessed via acs package",
             position = c("right", "bottom"))
```




```
# Save map as "nyc_income_map.png"
save_tmap(filename = "nyc_income_map.png", width = 4, height = 7)
```