Darrell_Nelson_HW07

Darrell Nelson II

March 4, 2019

Step 01: Load the data

```r
# Darrell Nelson II
# HW 07
# VizMap HW: Median Income

# Step 01:
# Read Excel document into R and upload it
# install.packages("readxl")
library("readxl")

incomedata <- read_excel("MedianZIP_2_2.xlsx")
```

```
## New names:
## * `` -> `..2`
## * `` -> `..3`
## * `` -> `..4`
```

```r
# Replace column names with names in row 1 then remove row 1
cnames <- incomedata[ 1, ]
cnames[1] <- "zip"
colnames(incomedata) <- cnames
incomedata <- incomedata[-1,]


# Load the zipcode package
# install.packages("zipcode")
library("zipcode")
data("zipcode")

# Ensure there are 5 numbers in incomedata df to match zipcode df
# install.packages("stringr")
library("stringr")
incomedata$zip <- str_pad(incomedata$zip, width=5, side="left", pad="0")

# Merge the zip code information from the two dfs
Income2 <- merge(x = incomedata, y = zipcode, by = "zip", all.x = TRUE)


# Remove Hawaii, Alaska, and Washington D.C.
Income2 <- Income2[Income2$state != "HI" , ]
Income2 <- Income2[Income2$state != "AK" , ]
Income2 <- Income2[Income2$state != "DC" , ]
```

Step 02: Show the income & population per state

```
# Step 02: Show the income & population per state
# Create df with average median income, pop per state, and state names (must be lowercase)
Income2$Median <- as.numeric(Income2$Median)
Income2$Pop <- as.numeric(Income2$Pop)

PerState <- tapply(Income2$Pop, Income2$state, sum)
PerState <- data.frame(PerState)
PerState$Avg_Median <- tapply(Income2$Median, Income2$state, mean)
colnames(PerState) <- c("Pop", "Avg_Median")


# Building a map using GGPLOT
# Load required libraries
require(ggplot2)
```

```
## Loading required package: ggplot2
```

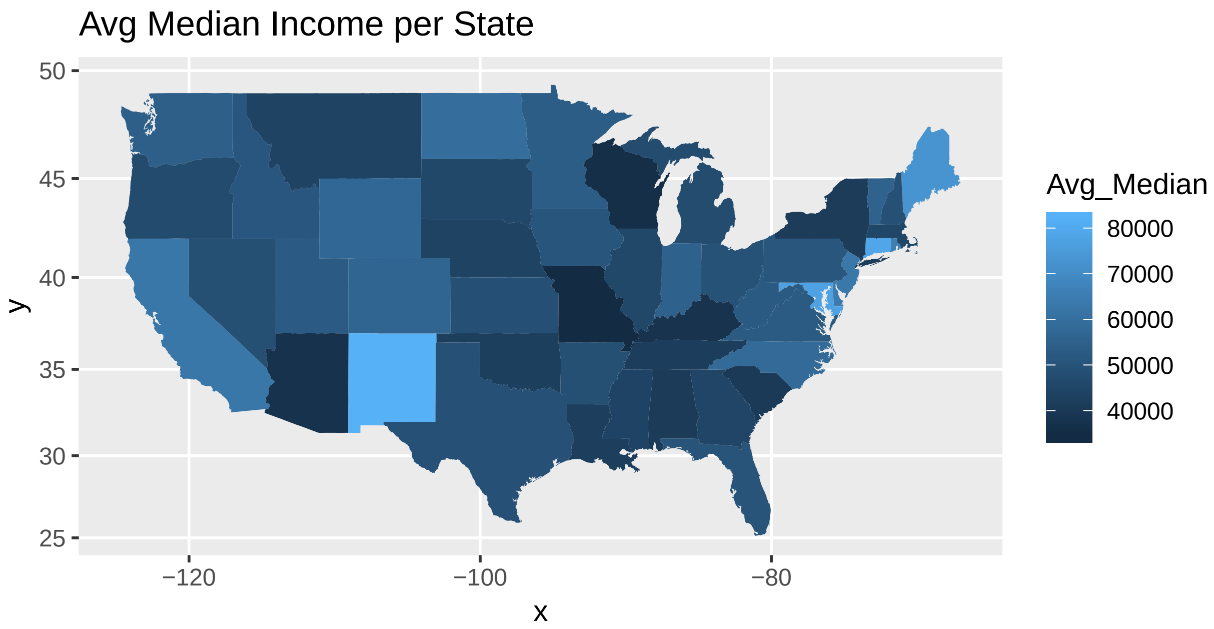
```
require(maps)
```

```
## Loading required package: maps
```

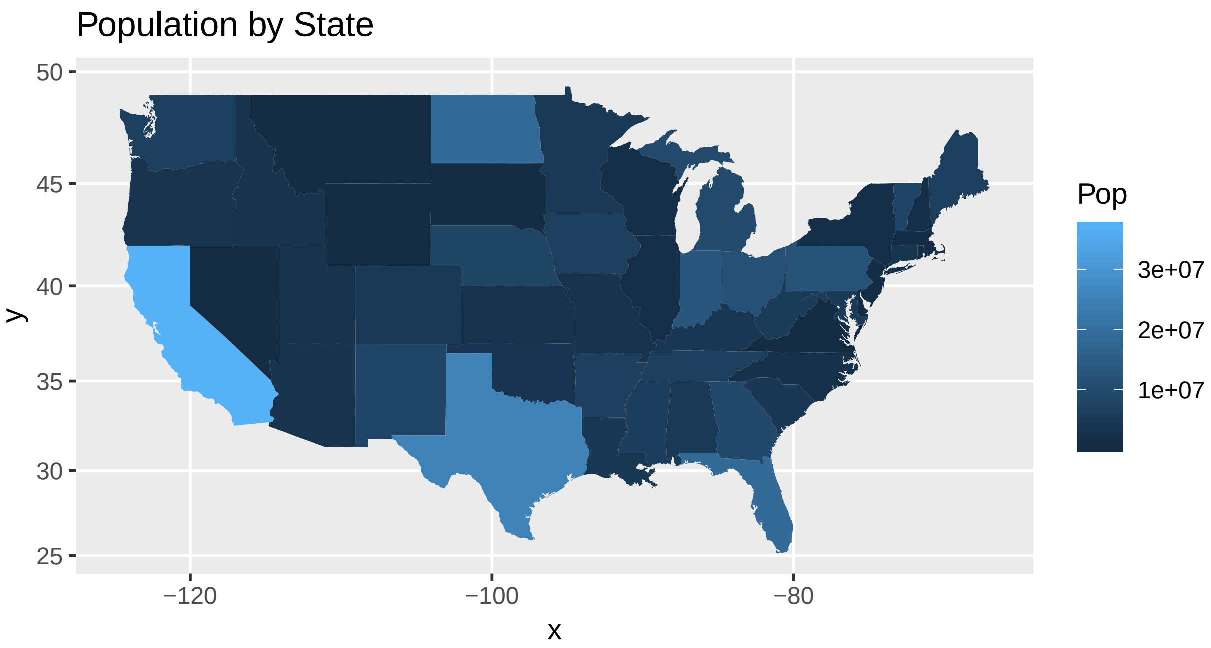
```
# install.packages("ggmap")
library("ggmap")
```

```
## Google's Terms of Service: https://cloud.google.com/maps-platform/terms/.

## Please cite ggmap if you use it! See citation("ggmap") for details.
```

```
# install.packages("mapproj")
library("mapproj")
# Get map of US
us <- map_data("state")
statenames <- data.frame(state.name, stringsAsFactors = FALSE)

# Remove Hawaii and Alaska from statenames
statenames <- statenames[statenames != "Hawaii"]
statenames <- data.frame(statenames, stringsAsFactors = FALSE)
statenames <- statenames[statenames != "Alaska"]
statenames <- tolower(statenames)

# Add statenames and abbreviatoins to PerState df
PerState$state <- statenames
abb <- row.names(PerState)
PerState$Abb <- abb

# Show US map w/ avg median income as color
map.popColor <- ggplot(PerState, aes(map_id=state))
map.popColor <- map.popColor + geom_map(map=us, aes(fill=Avg_Median))
map.popColor <- map.popColor + expand_limits(x=us$long, y=us$lat)
map.popColor <- map.popColor + coord_map() + ggtitle("Avg Median Income per State")
map.popColor
```

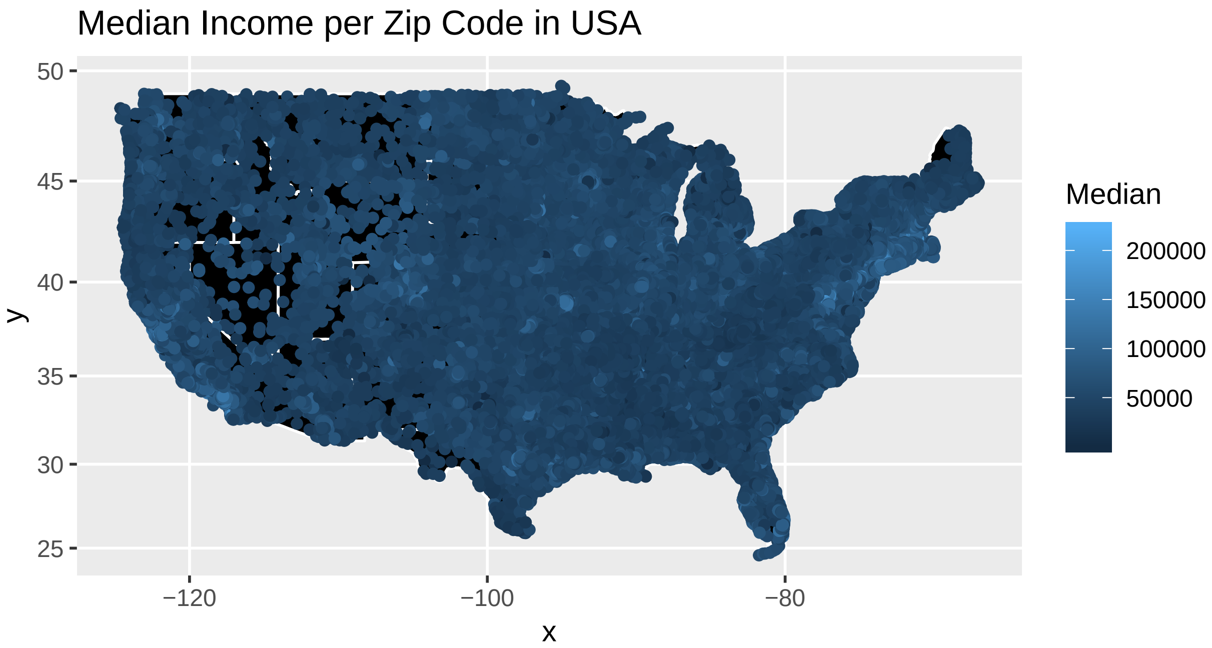


```
# Show US map w/ population of state as color
map.popColor2 <- ggplot(PerState, aes(map_id=state))
map.popColor2 <- map.popColor2 + geom_map(map=us, aes(fill=Pop))
map.popColor2 <- map.popColor2 + expand_limits(x=us$long, y=us$lat)
map.popColor2 <- map.popColor2 + coord_map() + ggtitle("Population by State")
map.popColor2
```



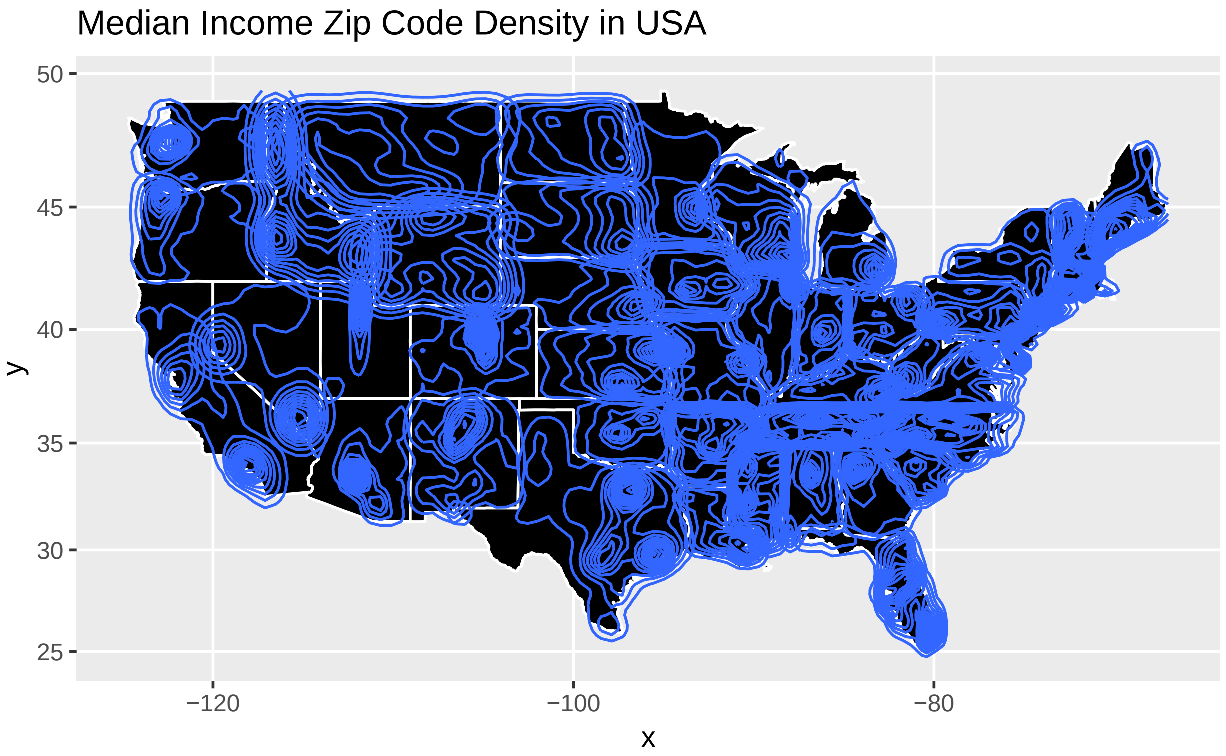
# Step 03: Show the income per zip code

```
# Step 03: Show the income per zip code
# Get map of US
us <- map_data("state")
dummyDF <- data.frame(state.name, stringsAsFactors = FALSE)
dummyDF$state <- tolower(dummyDF$state.name)
map.simple <- ggplot(dummyDF, aes(map_id=state))
map.simple <- map.simple + geom_map(map=us, fill="black", color="white")
map.simple <- map.simple + expand_limits(x=us$long, y=us$lat)
map.simple <- map.simple + coord_map() + ggtitle("Median Income per Zip Code in USA")
map.simple <- map.simple + geom_point(data = Income2, aes(x=Income2$longitude, y=Income2$latitude, colo
map.simple
```



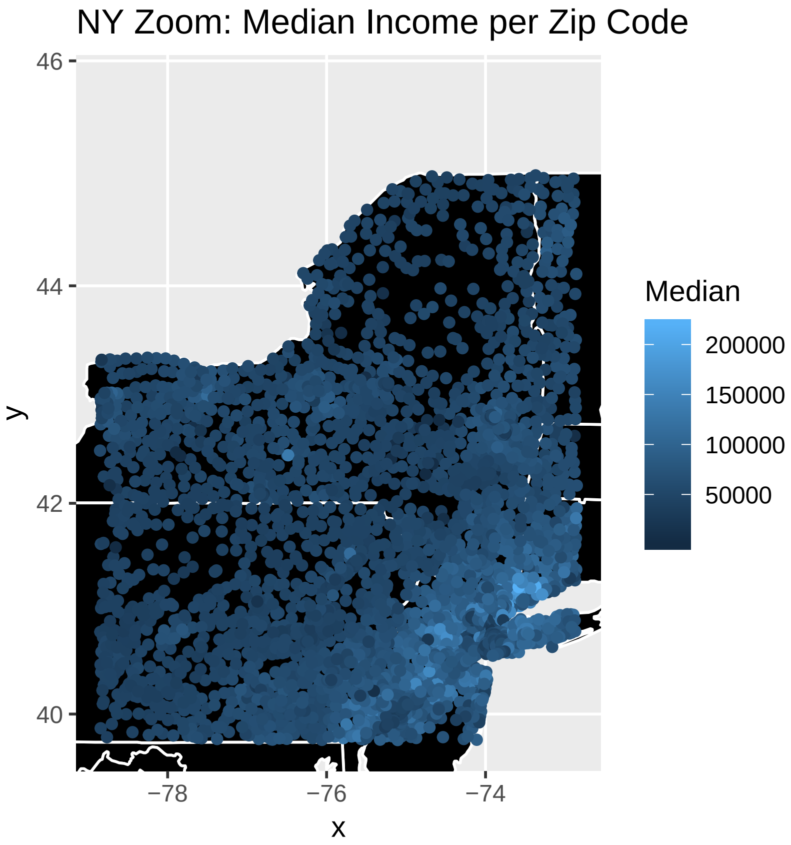
# Step 04: Show Zip Code density

```
# Step 04: Show Zip Code density
us <- map_data("state")
dummyDF <- data.frame(state.name, stringsAsFactors = FALSE)
dummyDF$state <- tolower(dummyDF$state.name)
map.simple2 <- ggplot(dummyDF, aes(map_id=state))
map.simple2 <- map.simple2 + geom_map(map=us, fill="black", color="white")
map.simple2 <- map.simple2 + expand_limits(x=us$long, y=us$lat)
map.simple2 <- map.simple2 + coord_map() + ggtitle("Median Income Zip Code Density in USA")
map.simple2 <- map.simple2 + stat_density_2d(data = Income2, aes(x=Income2$longitude, y=Income2$latitude
map.simple2
```

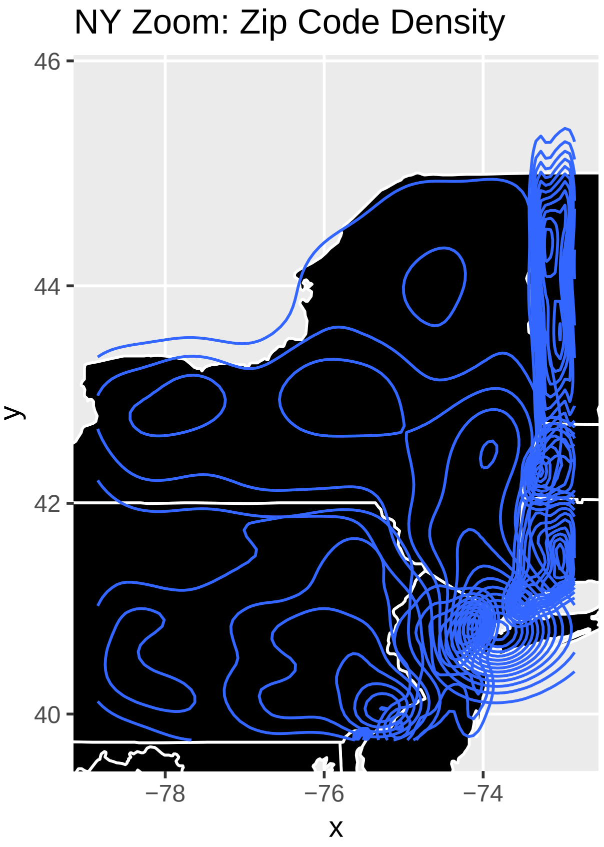


# Step 05: Zoom in on NYC region

```
# Step 05: Zoom in on NYC region
# Find the middle/center of NY

NY <- Income2[Income2$state == "NY" , ]
midlat <- mean(range(NY$latitude))
midlong <- mean(range(NY$longitude))
Window = 3
xlimit <- c(midlong - Window, midlong + Window)
ylimit <- c(midlat - Window, midlat + Window)

us <- map_data("state")
dummyDF <- data.frame(state.name, stringsAsFactors = FALSE)
dummyDF$state <- tolower(dummyDF$state.name)
map.zoom <- ggplot(dummyDF, aes(map_id=state))
map.zoom <- map.zoom + geom_map(map=us, fill="black", color="white")
map.zoom <- map.zoom + expand_limits(x=xlimit, y=ylimit)
map.zoom <- map.zoom + coord_map() + ggtitle(("NY Zoom: Median Income per Zip Code"))

zoom.zip <- Income2[Income2$longitude > xlimit[1] , ]
zoom.zip <- zoom.zip[zoom.zip$longitude < xlimit[2] , ]
zoom.zip <- zoom.zip[zoom.zip$latitude > ylimit[1] , ]
zoom.zip <- zoom.zip[zoom.zip$latitude < ylimit[2] , ]

map.zoom <- map.zoom + geom_point(data = zoom.zip, aes(x=zoom.zip$longitude, y=zoom.zip$latitude, color
map.zoom
```



```
us <- map_data("state")
dummyDF <- data.frame(state.name, stringsAsFactors = FALSE)
dummyDF$state <- tolower(dummyDF$state.name)
map.zoom <- ggplot(dummyDF, aes(map_id=state))
map.zoom <- map.zoom + geom_map(map=us, fill="black", color="white")
map.zoom <- map.zoom + expand_limits(x=xlimit, y=ylimit)
map.zoom <- map.zoom + coord_map() + ggtitle("NY Zoom: Zip Code Density")
map.zoom2 <- map.zoom + stat_density_2d(data = zoom.zip, aes(x=zoom.zip$longitude, y=zoom.zip$latitude)
map.zoom2
```


NY Zoom: Zip Code Density
46
44
42
40
y
−78
−76
−74
x