Assignment 04

Amol Gote

2/10/2020

# Question 4

(Extra credit) Try to think about how you can use the zip code data for location information and map price to zipcodes

zipcodeShapeData <- st\_read('data/tl\_2019\_us\_zcta510.shp')

## Reading layer `tl\_2019\_us\_zcta510' from data source `C:\Aamol\GitRepos\msdatascience\DataVisualization\Week4\Assignment4\data\tl\_2019\_us\_zcta510.shp' using driver `ESRI Shapefile'  
## Simple feature collection with 33144 features and 9 fields  
## geometry type: MULTIPOLYGON  
## dimension: XY  
## bbox: xmin: -176.6847 ymin: -14.37378 xmax: 145.8305 ymax: 71.34132  
## epsg (SRID): 4269  
## proj4string: +proj=longlat +datum=NAD83 +no\_defs

houses <- read\_csv("data/KING COUNTY House Data.csv")  
counties <- st\_as\_sf(map("county", plot = FALSE, fill = TRUE))  
counties\_wa <-counties %>%   
 filter(str\_detect(ID, 'washington,'))  
  
counties\_wa\_king <- counties\_wa %>%   
 filter(str\_detect(ID, "king"))  
  
averagePriceByZip <- houses %>% group\_by(zipcode) %>% summarize(averagePrice = mean(price, na.rm = TRUE))  
mergedShapeAndAvgZipCodeData <- merge(zipcodeShapeData,averagePriceByZip,by.x=c("ZCTA5CE10"),by.y=c("zipcode"))

mergedShapeAndAvgZipCodeData %>%  
 ggplot() +   
 geom\_sf(data=counties\_wa\_king) +   
 geom\_sf(aes(fill=averagePrice)) +   
 scale\_fill\_viridis\_c() +  
 labs(x = "Longitude", y = "Latitude",  
 title = "King County house prices by zip codes") +  
 theme\_minimal()

