# MIS503 – Final Project

## Tate, Levi

### Zillow Home Value Index Analysis

### Wake County Home Sales

You are considering a move to the Raleigh area and are interested in understanding trends in home values. Run the analysis in the instructions below and come back to this section of the R Markdown document and address the following questions:

1. What have been the overall trends in Wake County Home Values?  
   Overall trend in Wake Count Home Values: appreciation
2. There were dips in home values in the past 20 years. What years did these occur?  
   From 1997-1998 and from 2008-2012
3. Based on the analysis, where would be the least expensive area to purchase home? Most expensive area?  
   Least Expensive: Zebulon | overall Most Expensive: Cary
4. Are any area home values trending down? Is there one area that stands out compared to others?  
   No home values trending down. New Hill home values are increasing at alarming rate.

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.0 ──

## ✓ ggplot2 3.3.2 ✓ purrr 0.3.4  
## ✓ tibble 3.0.4 ✓ dplyr 1.0.2  
## ✓ tidyr 1.1.2 ✓ stringr 1.4.0  
## ✓ readr 1.4.0 ✓ forcats 0.5.0

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

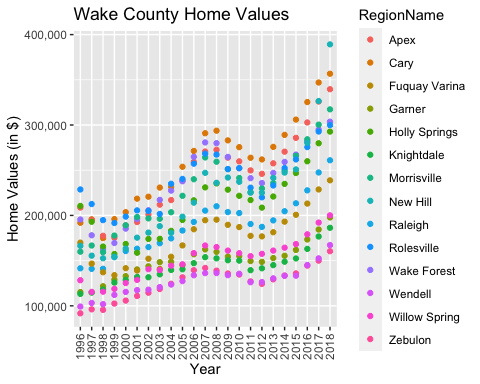
SingleFamilyResidenceSales <- read\_csv("SingleFamilyResidenceSales.csv")

##   
## ── Column specification ────────────────────────────────────────────────────────  
## cols(  
## .default = col\_double(),  
## RegionName = col\_character(),  
## State = col\_character(),  
## Metro = col\_character(),  
## CountyName = col\_character()  
## )  
## ℹ Use `spec()` for the full column specifications.

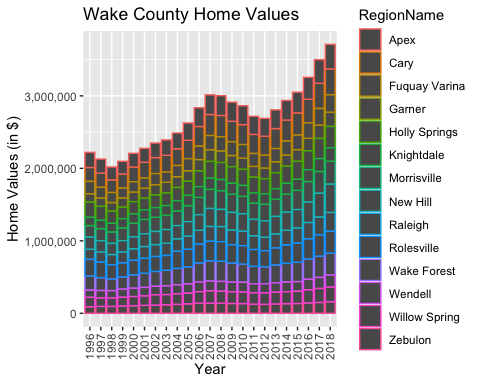
WakeCountySales <- SingleFamilyResidenceSales %>%  
 filter(State == "NC", CountyName == "Wake County") %>%  
   
 select(RegionName, State, CountyName, Metro, ends\_with("-05")) %>%  
   
 rename\_at(vars(ends\_with("-05")), funs(str\_replace(., "-05", ""))) %>%  
   
 pivot\_longer(c(5:27), names\_to = "YR", values\_to = "ZHVI")

## Warning: `funs()` is deprecated as of dplyr 0.8.0.  
## Please use a list of either functions or lambdas:   
##   
## # Simple named list:   
## list(mean = mean, median = median)  
##   
## # Auto named with `tibble::lst()`:   
## tibble::lst(mean, median)  
##   
## # Using lambdas  
## list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last\_warnings()` to see where this warning was generated.

ggplot(data = WakeCountySales, mapping = aes(x = YR, y = ZHVI, color = RegionName)) +  
 geom\_point() +  
 labs(x = "Year",  
 y = "Home Value",  
 title = "Wake County Home Values") +  
 theme(axis.text.x = element\_text(angle = 90, vjust=0.5)) +  
 scale\_y\_continuous(name="Home Values (in $)", labels = scales::comma)



ggplot(data = WakeCountySales, mapping = aes(x = YR, y = ZHVI, color = RegionName)) +  
 geom\_col() +  
 labs(x = "Year",  
 y = "Home Value",  
 title = "Wake County Home Values") +  
 theme(axis.text.x = element\_text(angle = 90, vjust=0.5)) +  
 scale\_y\_continuous(name="Home Values (in $)", labels = scales::comma)



### NC Rental Market

You are considering relocating to a new city and want to rent. You are looking at some of the larger cities within the state including Asheville, Charlotte, Durham, Fayetteville, Raleigh and Wilmington. Run the analysis in the instructions below and come back to this section of the R Markdown document and address the following questions:

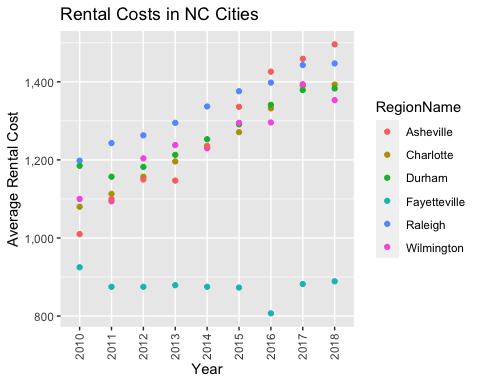
1. What has been the overall trend in the rental market around the state? Are there any cities that have not followed this trend?  
   Overall trend in rental market: Appreciation. Exception: Fayetteville
2. Where is the most expensive city to rent in? Least expensive?  
   Most expensive city: From 2010-2015 Raleigh, From 2016-2018 Asheville  
   Least expensive city: Fayetteville
3. You are trying decide between Wilmington and Asheville. Which market has the lowest rent?  
   Asheville’s market surpassed Wilmington and has steadily increased in rent. Wilmington is cheaper as of October 2018.

SingleFamilyResidenceRental <- read\_csv("SingleFamilyResidenceRental.csv")

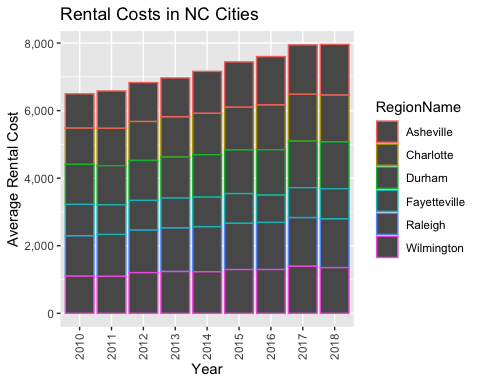
##   
## ── Column specification ────────────────────────────────────────────────────────  
## cols(  
## .default = col\_double(),  
## RegionName = col\_character(),  
## State = col\_character(),  
## Metro = col\_character(),  
## CountyName = col\_character()  
## )  
## ℹ Use `spec()` for the full column specifications.

Rentals <- SingleFamilyResidenceRental %>%  
 filter(State == "NC", RegionName %in% c("Asheville", "Charlotte", "Durham", "Fayetteville", "Raleigh", "Wilmington")) %>%  
   
 select(RegionName, State, ends\_with("-11"), "2018-10") %>%  
   
 rename\_at(vars(ends\_with("-11")), funs(str\_replace(., "-11", ""))) %>%  
   
 rename("2018" = "2018-10") %>%  
   
 pivot\_longer(c(3:11), names\_to = "YR", values\_to = "ZHVI")

ggplot(data = Rentals, mapping = aes(x = YR, y = ZHVI, color = RegionName)) +  
 geom\_point() +  
 labs(x = "Year",  
 y = "Home Value",  
 title = "Rental Costs in NC Cities") +  
 theme(axis.text.x = element\_text(angle = 90, vjust=0.5)) +  
 scale\_y\_continuous(name="Average Rental Cost", labels = scales::comma)



ggplot(data = Rentals, mapping = aes(x = YR, y = ZHVI, color = RegionName)) +  
 geom\_col() +  
 labs(x = "Year",  
 y = "Home Value",  
 title = "Rental Costs in NC Cities") +  
 theme(axis.text.x = element\_text(angle = 90, vjust=0.5)) +  
 scale\_y\_continuous(name="Average Rental Cost", labels = scales::comma)



### Home Values in Select Rental Markets

You have made the choice that you want to focus on 4 regions (Asheville, Charlotte-Concord-Gastonia, Raleigh and Wilmington) and instead of renting, you would like to purchase a home. Run the analysis in the instructions below and come back to this section of the R Markdown document and address the following questions (you may need to do some research on reading a violin plot <https://blog.modeanalytics.com/violin-plot-examples/>):

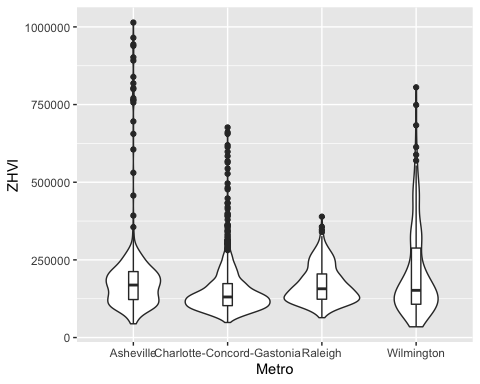
1. According to the results, which market has the lowest median price (represented as horizontal bar in box plot)?  
   Charlotte-Concord-Gastonia
2. The violin plot will show density meaning the wider the plot is, the more observations occur within that area. Which market has the most density around the median value of homes?  
   Charlotte-Concord-Gastonia

NCHomeSales <- SingleFamilyResidenceSales %>%  
   
 filter(State == "NC", Metro %in% c("Asheville","Charlotte-Concord-Gastonia", "Raleigh", "Wilmington")) %>%  
   
 select(RegionName, State, Metro, ends\_with("-05")) %>%  
   
 rename\_at(vars(ends\_with("-05")), funs(str\_replace(., "-05", ""))) %>%  
   
 pivot\_longer(c(4:26), names\_to = "YR", values\_to = "ZHVI") %>%  
   
 group\_by(Metro)

ggplot(data = NCHomeSales, mapping = aes(x = Metro, y = ZHVI)) +  
 geom\_violin() +  
 geom\_boxplot(width=0.1)

## Warning: Removed 90 rows containing non-finite values (stat\_ydensity).

## Warning: Removed 90 rows containing non-finite values (stat\_boxplot).



### Relocation Home Value Comparison

You have been given a new opportunity to relocate for a new position within your company. They have given you the option of 4 different areas in the country (Chicago, Denver, Houston or New York). Run the analysis in the instructions below and come back to this section of the R Markdown document and address the following questions:

1. Based on your analysis, which city’s housing is most affordable? Least affordable?  
   Most affordable: Houston. Least affordable: New York
2. Which cities saw the largest change in prices over the past 5 years? Which city has remained more consistent (i.e., no huge swings up or down in home values)?  
   Largest change in prices over the past 5 years: Denver and New York. Most consistent prices: Houston.
3. During the market downturn in 2012, which cities were most impacted? Which cities have recovered?  
   New York and Chicago were most impacted during the downturn in 2012. All four cities recovered after the market downturn in 2012, and home values have steadily increased in prices.

NationalHomeSales <- SingleFamilyResidenceSales %>%  
   
 filter(RegionName %in% c("Chicago", "Denver", "Houston", "New York"),  
 State %in% c("IL", "CO", "TX", "NY")) %>%  
   
 select(RegionName, State, Metro, ends\_with("-05")) %>%  
   
 rename\_at(vars(ends\_with("-05")), funs(str\_replace(., "-05", ""))) %>%  
   
 pivot\_longer(c(4:26), names\_to = "YR", values\_to = "ZHVI") %>%  
   
 group\_by(Metro)

ggplot(data = NationalHomeSales, mapping = aes(x = YR, y = ZHVI)) +  
geom\_point() +  
 facet\_wrap(~Metro) +  
 labs(x = "Year",  
 y = "Home Values (in $)",  
 title = "City Home Value Comparison") +  
 theme(axis.text.x = element\_text(angle = 90, vjust=0.5, size=5)) +  
 scale\_y\_continuous(name="Home Values (in $)", labels = scales::comma)

