# MIS503 - Final Project

## Lugo, Alex

### Zillow Home Value Index Analysis

### Wake County Home Sales

1. What have been the overall trends in Wake County Home Values? Home values have generally been increasing from 1996 to 2018.
2. There were dips in home values in the past 20 years. What years did these occur? There are 2 ranges where home values dipped based on the charts below. First, from 1996 to 1998 and again from 2007 to 2012.
3. Based on the analysis, where would be the least expensive area to purchase home? Most expensive area? Based on the trends and last years home values, 2018, the least expensive area to buy a home is in Zebulon.The most expensive would be Raleigh.
4. Are any area home values trending down? Is there one area that stands out compared to others? There doesn’t seem to be any area where home values are trending down, however, Raleigh has experienced growth in home values at a very rapid pace compared to all other areas in recent years.

library(tidyverse)

## -- Attaching packages -------------------------------------------------------------------------------- tidyverse 1.2.1 --

## v ggplot2 3.2.1 v purrr 0.3.3  
## v tibble 2.1.3 v dplyr 0.8.3  
## v tidyr 1.0.0 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.4.0

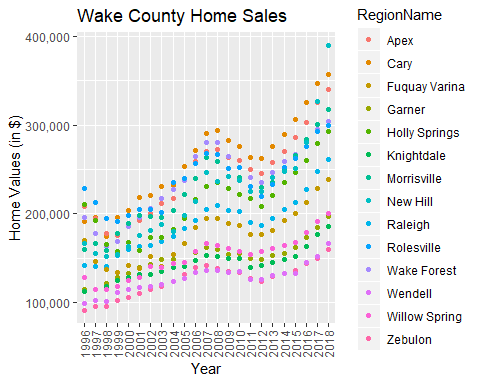
## -- Conflicts ----------------------------------------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

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

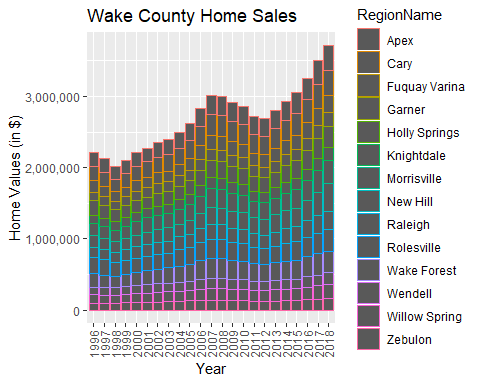
## Parsed with column specification:  
## cols(  
## .default = col\_double(),  
## RegionName = col\_character(),  
## State = col\_character(),  
## Metro = col\_character(),  
## CountyName = col\_character()  
## )

## See spec(...) for full column specifications.

WakeCountySales <- filter(SingleFamilyResidenceSales, State== "NC", CountyName== "Wake County")  
WakeCountySales <- select(WakeCountySales, "RegionName", "State", "CountyName", "Metro", "1996-05", "1997-05", "1998-05", "1999-05", "2000-05", "2001-05", "2002-05", "2003-05", "2004-05", "2005-05", "2006-05", "2007-05", "2008-05", "2009-05", "2010-05", "2011-05", "2012-05", "2013-05", "2014-05", "2015-05", "2016-05", "2017-05", "2018-05")  
WakeCountySales <- rename(WakeCountySales, "1996"="1996-05", "1997"="1997-05", "1998"="1998-05", "1999"="1999-05", "2000"="2000-05", "2001"="2001-05", "2002"="2002-05", "2003"="2003-05", "2004"="2004-05", "2005"="2005-05", "2006"="2006-05", "2007"="2007-05", "2008"="2008-05", "2009"="2009-05", "2010"="2010-05", "2011"="2011-05", "2012"="2012-05", "2013"="2013-05", "2014"="2014-05", "2015"="2015-05", "2016"="2016-05", "2017"="2017-05", "2018"="2018-05")  
WakeCountySales <- gather(WakeCountySales, "1996", "1997", "1998", "1999", "2000", "2001", "2002", "2003", "2004", "2005", "2006", "2007", "2008", "2009", "2010", "2011", "2012", "2013", "2014", "2015", "2016", "2017", "2018", key='YR', value='ZHVI')  
ggplot(WakeCountySales, aes(YR, ZHVI, color= RegionName))+  
 geom\_point()+  
 labs(title= "Wake County Home Sales",  
 x= "Year",  
 y= "ZHVI")+  
 theme(axis.text.x = element\_text(angle = 90, vjust=0.5))+  
 scale\_y\_continuous(name="Home Values (in $)", labels = scales::comma)



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



### NC Rental Market

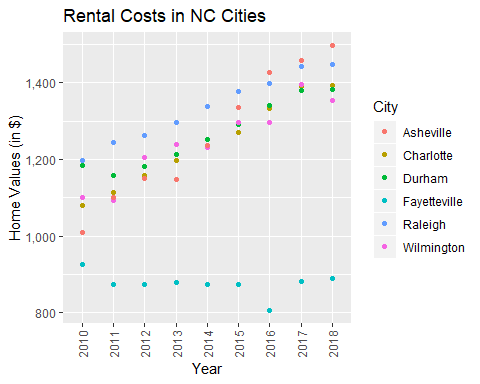
1. What has been the overall trend in the rental market around the state? Are there any cities that have not followed this trend? The trend has been an overall increase in rental cost. Fayetteville has not followed the trend as it has remained fairly constant aside from a drop in 2016.
2. Where is the most expensive city to rent in? Least expensive? The most expensive place to rent is Asheville based on the below chart. The least expensive city is Fayetteville.
3. You are trying decide between Wilmington and Asheville. Which market has the lowest rent? Wilmington has lower rent as Asheville has past the average rental cost in Wilmington in recent years.

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

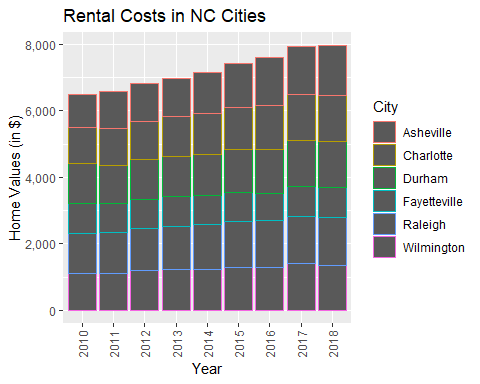
## Parsed with column specification:  
## cols(  
## .default = col\_double(),  
## RegionName = col\_character(),  
## State = col\_character(),  
## Metro = col\_character(),  
## CountyName = col\_character()  
## )

## See spec(...) for full column specifications.

Rentals <- filter(SingleFamilyResidenceRental, RegionName %in% c("Asheville", "Charlotte", "Durham", "Fayetteville", "Raleigh", "Wilmington"))  
Rentals <- filter(Rentals, State== "NC")  
Rentals <- mutate(Rentals, City = RegionName)  
Rentals <- select(Rentals, "City", "State", "2010-11", "2011-11", "2012-11", "2013-11", "2014-11", "2015-11", "2016-11", "2017-11", "2018-10")  
Rentals <- rename(Rentals, "2010"="2010-11", "2011"="2011-11", "2012"="2012-11", "2013"="2013-11", "2014"="2014-11", "2015"="2015-11", "2016"="2016-11", "2017"="2017-11", "2018"="2018-10")  
Rentals <- gather(Rentals, "2010", "2011", "2012", "2013", "2014", "2015", "2016", "2017", "2018", key='YR', value='ZHVI')  
ggplot(Rentals, aes(YR, ZHVI, color= City))+  
 geom\_point()+  
 labs(title= "Rental Costs in NC Cities",  
 x= "Year",  
 y= "ZHVI")+  
 theme(axis.text.x = element\_text(angle = 90, vjust=0.5))+  
 scale\_y\_continuous(name="Home Values (in $)", labels = scales::comma)



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



### Home Values in Select Rental Markets

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

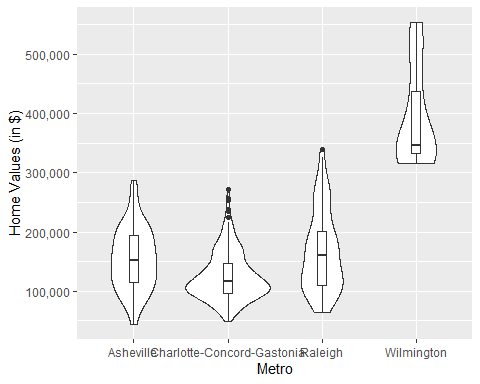
NCHomeSales <- filter(SingleFamilyResidenceSales, State %in% "NC", Metro==c("Asheville", "Charlotte-Concord-Gastonia", "Raleigh", "Wilmington"))

## Warning in Metro == c("Asheville", "Charlotte-Concord-Gastonia",  
## "Raleigh", : longer object length is not a multiple of shorter object  
## length

NCHomeSales <- select(NCHomeSales, "RegionName","State", "Metro", "1996-05", "1997-05", "1998-05", "1999-05", "2000-05", "2001-05", "2002-05", "2003-05", "2004-05", "2005-05", "2006-05", "2007-05", "2008-05", "2009-05", "2010-05", "2011-05", "2012-05", "2013-05", "2014-05", "2015-05", "2016-05", "2017-05", "2018-05")  
NCHomeSales <- rename(NCHomeSales, "1996"="1996-05", "1997"="1997-05", "1998"="1998-05", "1999"="1999-05", "2000"="2000-05", "2001"="2001-05", "2002"="2002-05", "2003"="2003-05", "2004"="2004-05", "2005"="2005-05", "2006"="2006-05", "2007"="2007-05", "2008"="2008-05", "2009"="2009-05", "2010"="2010-05", "2011"="2011-05", "2012"="2012-05", "2013"="2013-05", "2014"="2014-05", "2015"="2015-05", "2016"="2016-05", "2017"="2017-05", "2018"="2018-05")  
NCHomeSales <- gather(NCHomeSales, "1996", "1997", "1998", "1999", "2000", "2001", "2002", "2003", "2004", "2005", "2006", "2007", "2008", "2009", "2010", "2011", "2012", "2013", "2014", "2015", "2016", "2017", "2018", key='YR', value='ZHVI')  
NCHomeSales <- group\_by(NCHomeSales, Metro)  
ggplot(NCHomeSales, aes(Metro, ZHVI,))+  
 geom\_violin()+  
 geom\_boxplot(width=0.1)+  
 scale\_y\_continuous(name="Home Values (in $)", labels = scales::comma)

## Warning: Removed 27 rows containing non-finite values (stat\_ydensity).

## Warning: Removed 27 rows containing non-finite values (stat\_boxplot).



### Relocation Home Value Comparison

1. Based on your analysis, which city’s housing is most affordable? Least affordable? Houston is the most affordable. New York is the least affordable.
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)? In the last 5 years, Denver and New York have experienced the largest change in prices. Houston has remained the most consistent.
3. During the market downturn in 2012, which cities were most impacted? Which cities have recovered? Chicago and New York were most impacted by the market downturn in 2012. Denver, Houston, and New York have all recovered from their marks before the downturn. Only Chicago hasn’t returned to its previous high before it.

NationalHomeSales <- filter(SingleFamilyResidenceSales, State %in% c("IL", "TX", "CO", "NY"), RegionName %in% c("Chicago", "Denver", "Houston", "New York"))  
NationalHomeSales <- select(NationalHomeSales, "RegionName", "State", "Metro", "1996-05", "1997-05", "1998-05", "1999-05", "2000-05", "2001-05", "2002-05", "2003-05", "2004-05", "2005-05", "2006-05", "2007-05", "2008-05", "2009-05", "2010-05", "2011-05", "2012-05", "2013-05", "2014-05", "2015-05", "2016-05", "2017-05", "2018-05")  
NationalHomeSales <- rename(NationalHomeSales, "1996"="1996-05", "1997"="1997-05", "1998"="1998-05", "1999"="1999-05", "2000"="2000-05", "2001"="2001-05", "2002"="2002-05", "2003"="2003-05", "2004"="2004-05", "2005"="2005-05", "2006"="2006-05", "2007"="2007-05", "2008"="2008-05", "2009"="2009-05", "2010"="2010-05", "2011"="2011-05", "2012"="2012-05", "2013"="2013-05", "2014"="2014-05", "2015"="2015-05", "2016"="2016-05", "2017"="2017-05", "2018"="2018-05")  
NationalHomeSales <- gather(NationalHomeSales, "1996", "1997", "1998", "1999", "2000", "2001", "2002", "2003", "2004", "2005", "2006", "2007", "2008", "2009", "2010", "2011", "2012", "2013", "2014", "2015", "2016", "2017", "2018", key='YR', value='ZHVI')  
ggplot(NationalHomeSales, aes(YR, ZHVI))+  
 facet\_wrap(~Metro)+  
 geom\_point()+  
 labs(title= "City Home Value Comparison",  
 x= "Year",  
 y= "ZHVI")+  
 theme(axis.text.x = element\_text(angle = 90, vjust=0.5, size=5))+  
 scale\_y\_continuous(name="Home Values (in $)", labels = scales::comma)

