Dallas Real Estate Market
MIDS W200 Section 2
Project 2 Data Analysis Proposal
By: Nicholas Cirella, Theodore Fong, Thomas Martinez

Purpose

The purpose of this project is to conduct an analysis on Real Estate trends on the Dallas real estate market and find correlations between housing prices and a variety of housing characteristic variables from Redfin's open-source data.

Context

First time buyers and families looking to purchase homes often are overwhelmed by the countless variables that go into buying a home. Companies like Trulia, Zillow, Realtor.com, Estately, Homes .com, and Redfin seek to make available homes easier to find by future homeowners, but can easily be overwhelmed by the 50+ variables on Zillow or 75+ variables on Redfin. Since the first thing you see when looking at a home is the listed price, the square footage of the home, and the number of days the home was listed, we wanted to examine if these variables, even matter when considering a home offer price. We also look into Home Owner Associations (HOA), because there are varying viewpoints on the pros/cons of buying a home with HOA fees.

Data

We downloaded data from RedFin specifically in the Dallas region, but we had to clean the data up to get a final data set. We conducted this process for both listed homes as of March 30, 2019 and sold homes within the past 90 days from April 9, 2019. All available homes were downloaded from Redfin of up to 350 homes at a time and then compiled into one master .csv file. We then removed all duplicate listings and changed the variable names to lowercase and underscores to replace spaces between words. We deleted irrelevant columns for our data analysis. Please refer to the project_2_data_collection document for the full data collection process.

Data Source: https://www.redfin.com/city/30794/TX/Dallas/filter/max-price=10M

Variable	Description
sale_type	how the home is being sold
property_type	type of property (ie. Multi-family, single-family, vacant land)
address	property address
city	City
zip_or_postal_code	zip code
price	listed price
beds	number of beds
baths	number of baths
location	city district of the home
square_feet	total square feet of living space
lot_size	total size of the property in square feet
year_built	what year the home was built
days_on_market	number of days on the market
price_per_square_feet	price variable divided by square_feet variable
hoa_per_month	amount of monthly HOA fees due
status	property's status as active or not
source	where the information for the listing was found

We also identified the latitude/longitude of each home from the given redfin data labeled as 'clean_data_redfin_30MAR2019_sfinc_final.csv' where we used this to create a heat map overlap on a map of Dallas, Texas .

Focus Questions

Given the context, our research was focused on answering the following questions:

- For individuals that have made a decision to buy, what factors should they consider that may be counterintuitive?
- How does zip code impact price/availability?
- What is the relationship between lot size to overall price of a home?
- What can we understand from the number of days a home is listed on the market?
- What is the impact of HOA-dues on home sales?

Raw Data

Section I: Available Inventory Basics

Before diving into more detailed metrics, it is useful to obtain a general overview of the market. What inventory is currently available and what has sold recently and at what prices.

First, we can examine available properties by price and square footage. The heatmap below lays out the currently available homes.

Currently available homes by size and price
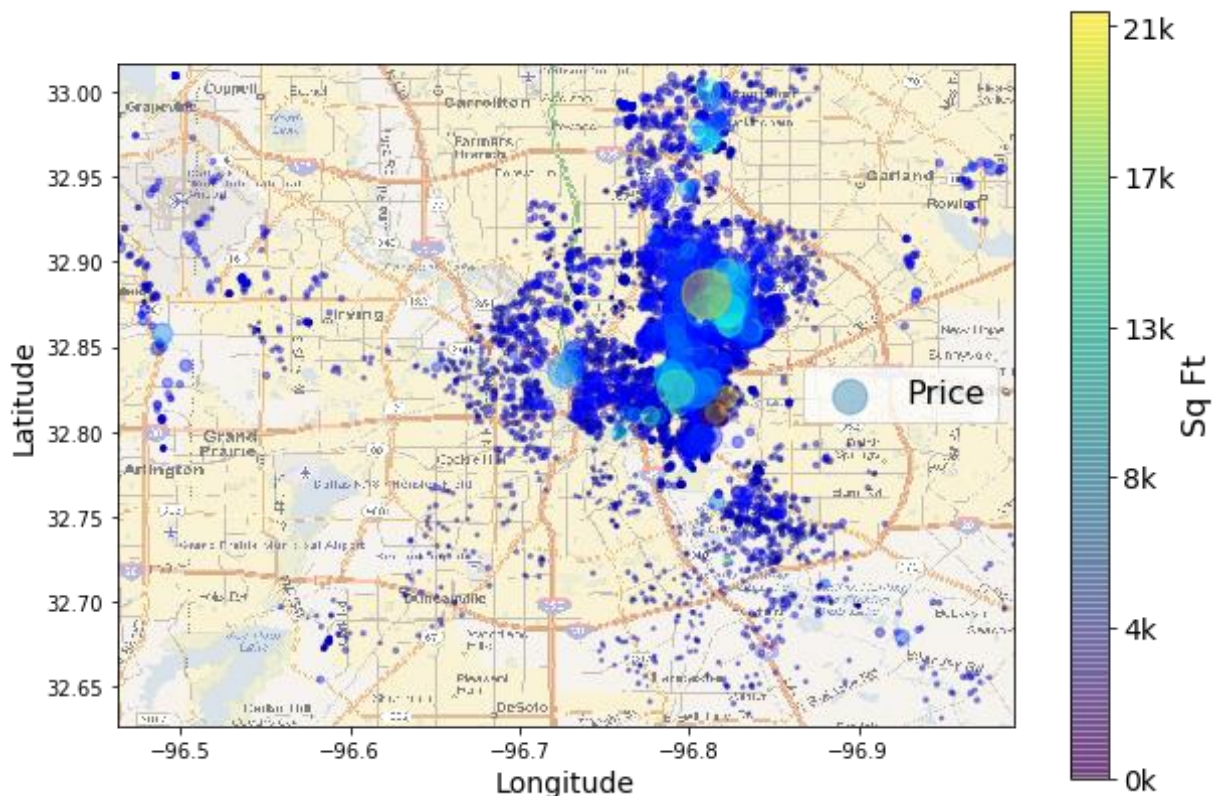
We can see that price seems loosely correlated with size, but also may increase for properties located in more central areas. This is consistent with what one might expect, but knowing what areas are hot and where smaller home are more highlight priced may help a new buyer to understand what neighborhoods are desirable to live in, especially if one does not have experience in the Dallas market.

As a next step, the histograms below outline these metrics and give some useful initial insights into the market.

Currently Available Homes (number currently available by asking price)
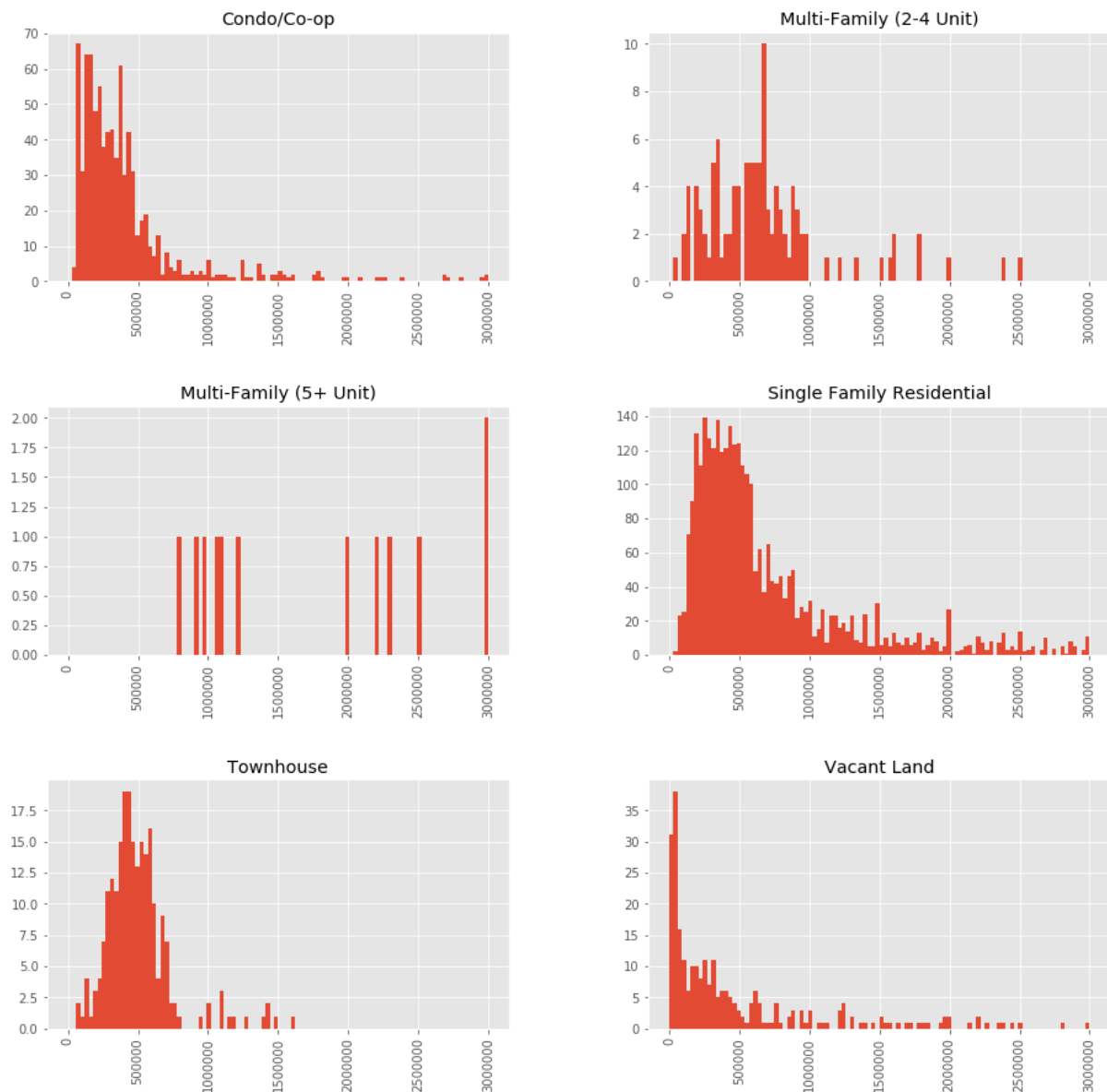
The x-axis for the above graphs show the price, while the y-axis show the number of homes available. One can quickly identify that most homes are listed under a million dollars and are listed around the half a million dollar mark.

*Recently Sold Homes (number sold over the last three months by sale price)*
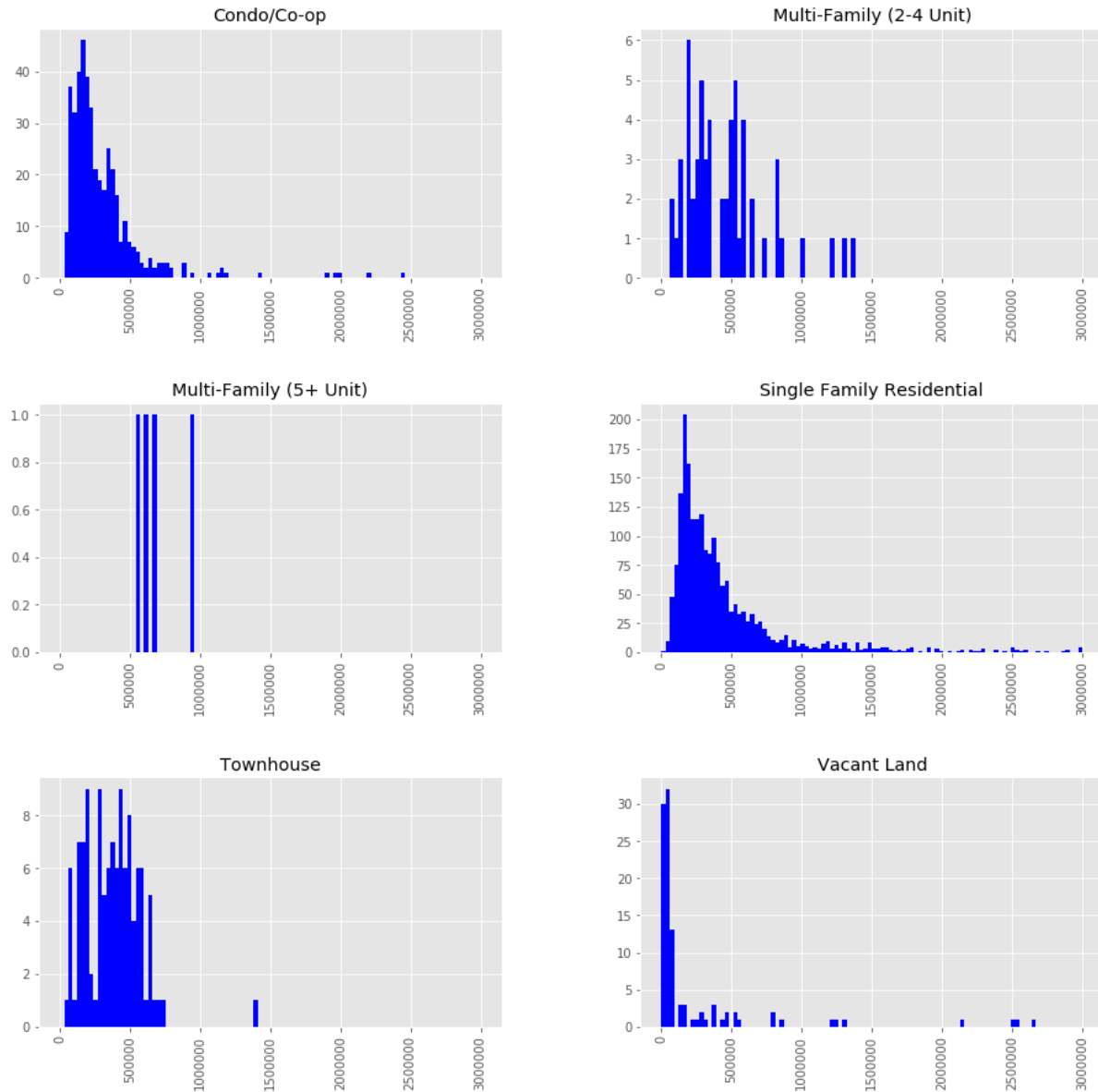
The graphs above are limited to properties sold for up to $3 million. This encompasses both the vast majority of properties and maintains a range that is not completely out of the question for many prospective buyers. What we can see it that sales are grouped tightly in the $500k range for singe family, condo and townhouse residences for recently sold properties.

It is interesting to note that asking prices appear to range higher in terms of frequency than actual sales. This should tell prospective buyers that the market may be ripe for counteroffers below the asking price, particularly for properties priced above $500k. This, of course, is not a full picture, but is a useful starting point.

Finally, the highest inventories appear to be in single family and condo properties. This is another indicator that may inform the thought process of prospective buyers. However, since these are also the highest among the recently sold properties, demand may match supply and offset some of the bargaining leverage.

Section II: Understanding the Market

Before diving into available inventory, it is important to understand the recent market. We have analyzed sales data over the prior three months, which will be vital in analyzing the available inventory.


Sold Homes by Property Type
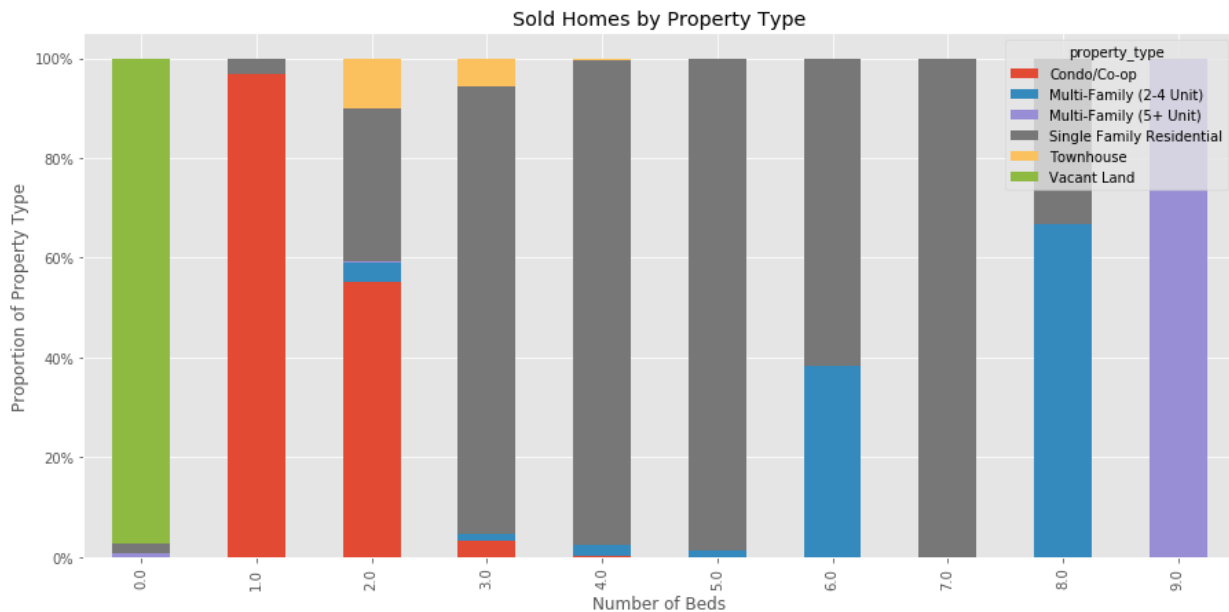
This graph shows the number of beds in the homes sold, colored by property type. Condos dominate the one- and two-bedroom market, while town homes appear to only be relevant in the two to three bed homes. Prospective buyers searching for townhomes will be limited to two- to three-bedroom homes, and are unlikely to find a four-bedroom townhome. Multi-family homes appear to be mainly two to eight bed homes and prospective buyers for real-estate investment are unlikely to find multi-family homes less than six beds.

Examining Price vs Size

There are two important sizing metrics to consider when examining available properties: livable square feet and lot size. Lot size is the acreage of the property, which livable square footage is the size of the home itself (including multiple floors, etc.). Vacant lots would not have any livable square footage, for example. These two primary perspectives are useful to develop a general familiarity with the nature and volume of the inventory population at large. These perspectives provide the added benefit of revealing general trends as well as shifting trends in the data, such as tiered changes in the ratio of price to square foot and in the ratio of price to lot size.

Home Sold prices by Year Built
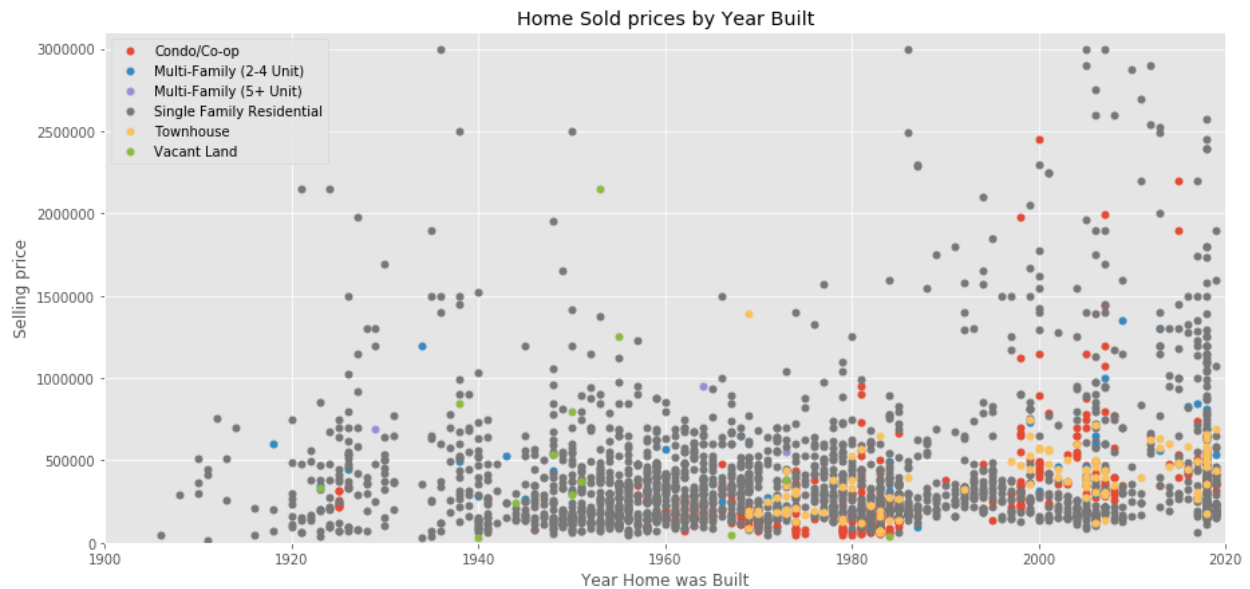
This graph shows the year a home was built and the price that it was sold at for all homes purchased in Dallas within the last three years. Reinforcing common convention, homes that were built before the 1980s tend to be cheaper, or in this case under a million dollars. Homes that go over a million dollars tend to be newer homes since 2000, or 1920-1960s homes that have been preserved well. For buyers seeking investment property from multi-family homes, they may be relieved that multi-family homes tend to sell for under a million dollars indicating a lower upfront cost for long-term growth.

Home Sold prices by Year Built
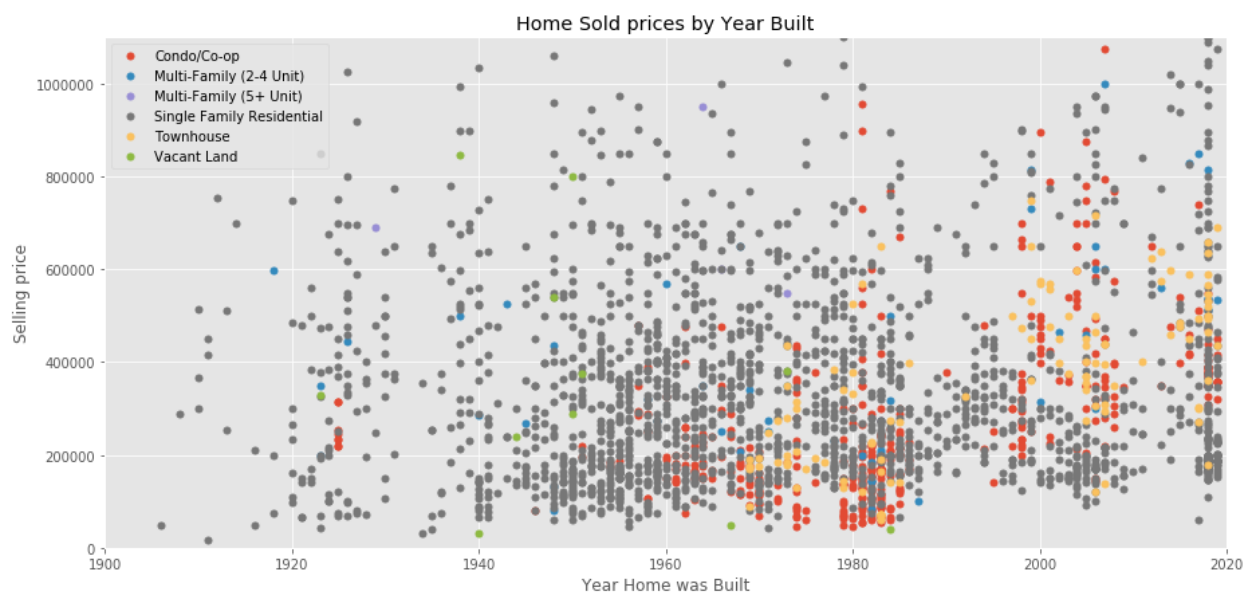
Honing in on the year built for homes being sold under a million dollars, we find that the average home before the 1990s tend to be selling around $100,000 to $400,000, where the graph is most dense. Homes built after 2000 though tend to sell between $150,000 to $600,000. Supporting the idea that newer homes sell better, this graph also indicates that newer townhomes/condos tend to sell at a higher price than older ones within the under a million dollar Dallas market. We believe this discovery

should caution potential townhome and condo buyers in thinking that their homes can be sold at a high price over time. With rising home prices in the past several decades, an apartment/condo may increase in value, but this graph predicts that buyers may prefer new-build as opposed to recently-built when it comes to condo/townhome inventory.


This graph shows the number of days since a home was sold and the number of homes sold by property type. The gaps in the lines indicate dates when no home was sold. Sellers would find this graph interesting because this time-series plot indicates that most homes are close out and sold on Fridays as shown by the weekly spikes in this graph. Right after December 2018 / January 2019, buyers appear to have more confidence and are back in the market to buying homes, which has since tempered a bit. This may be indicative of families moving cities in the new year as hiring ticks up post-holidays.

HOA Dues

For some properties, a homeowners association ("**HOA**") may charge monthly dues. This can be for the upkeep of common areas in communities or building, amenities like pools, taxes and other general management expenses. These monthly dues increase the carrying cost of a property and as a result we could expect an impact on home prices. Below we see an examination of recently sold homes by price per square foot relative to HOA dues.

Actual price per square foot vs HOA Dues
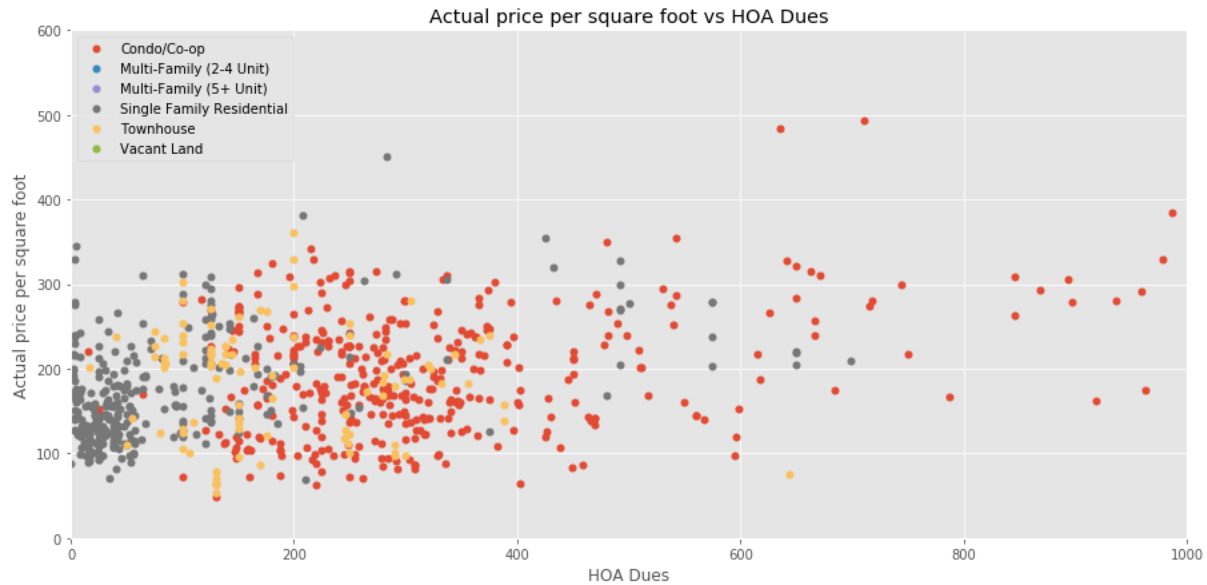
Perhaps counterintuitively we don't seem to see higher dues decreasing the price per square foot. We theorize that higher price point buyers are less sensitive to carrying costs and that the value of the amenities offered may offset the additional costs. This is a determination that a buyer should bear in mind, but must be evaluated on a case by case basis. The fact that condos and townhomes generally appear to have higher dues is consistent with the fact that these properties share commonly owned space that must be managed and paid for.

Build Year

Finally, we can look at the year built for recently sold properties to examine the impact on price.

Actual price per square foot vs Year Built
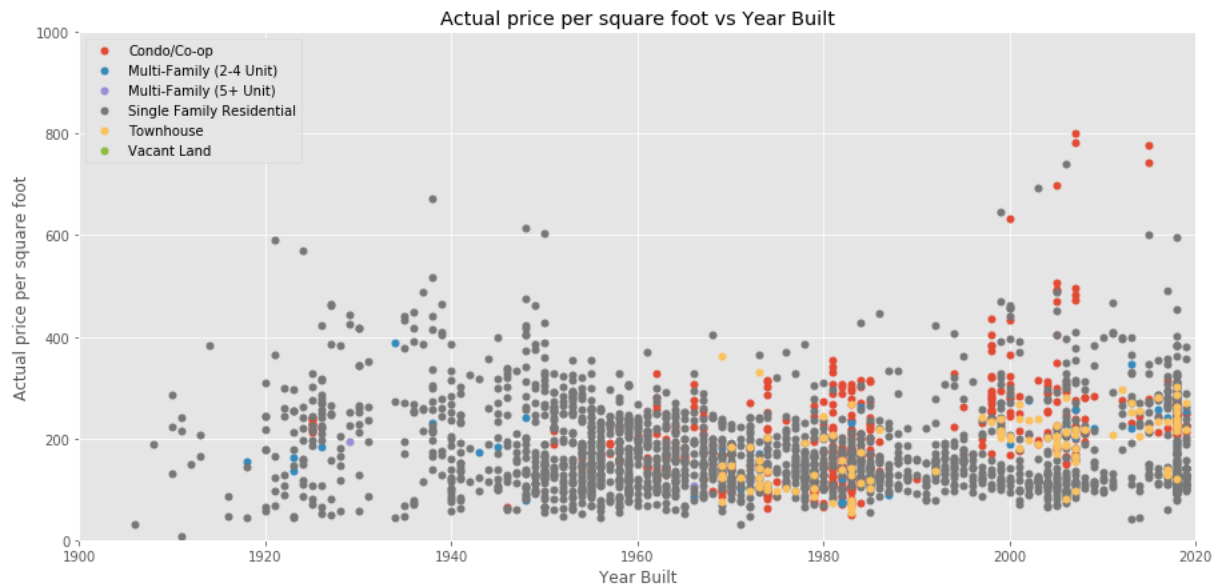
The above shows something interesting. Condos are more likely to be extremely expensive when recently built, whereas single family homes can be more expensive both for more recent builds and older homes. This makes sense. Older homes may have charm and value, especially when remodeled. Condos are often apartment-style and new builds have amenities like central air that would be prized in

a market like Dallas. Deals may be found in the mid-range and older, but buyers should be careful of the need for investment.

Section III: Looking Closer at Inventory

Most buyers agree that it all comes down to price and when a seller has left a property on the market for a while, it may indicate that the listing price is too high. Below is a graph showing listed prices relative to time on the market.


This graph shows listed prices by how long a home is on the market up to a year by property type. Since a listed home price can often be negotiated downwards after a home has been on the market for a while. Our graph shows that the threshold where buyers have that negotiating power is at the 90-day mark. Based on the above, there seems to be a significant drop-off in available properties after approximately 90 days on the market. Prospective buyers can look to properties available for longer than 90 days for potential deals or opportunities to negotiate.

The average townhome appears to be centered around $300,000–$750,000, despite the amount of days a home is listed on the market. Likewise, single family residential homes appear to be mostly under a million dollars throughout with a higher concentration around $400,000, despite the number of days on the market. One may hypothesize that sellers may initially price their homes optimistically (or unrealistically) and over time end up lowering the asking price to the market value. Prospective buyers may also assume that if a seller's home has been listed on the market for a longer period of time, the seller may have unrealistic expectations and there may be an opportunity to negotiate. Of course, this could also mean the seller is recalcitrant and will not provide a realistic opportunity for a market deal.

Overview - List Price to Livable Square Feet

We will first look at some general plots in useful ranges for available properties. By examining these plots and some of their fundamental trends a buyer can get a sense of what properties are available and what a reasonable price point is based on livable square footage.

Listed Price to Square Feet of Home
0 - 3,000 sq ft
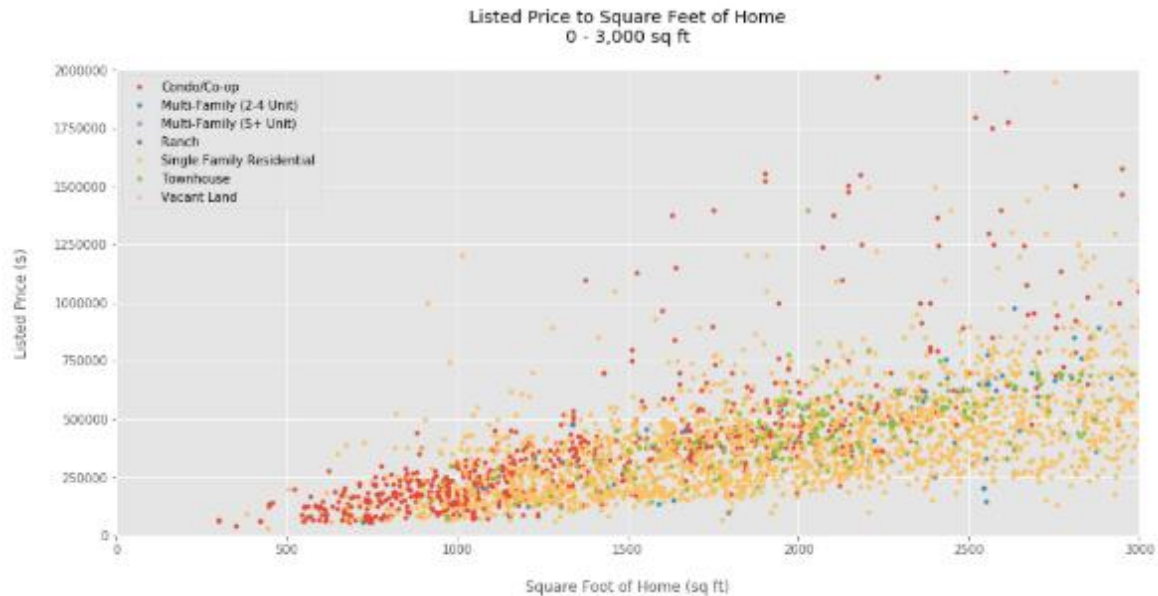
The above graph displays the Dallas inventory no greater than $2 million and up to 3,000 square feet of home interior living space. The 3,000 square feet threshhold was selected because the inventory population density is greater below this thresshold and best displayed and studied within this limited range. Of note are the trendline for this population as well as the trendline for each property type denoted by each of the various colors.

Listed Price to Square Feet of Home
0 - 6,000 sq ft

The square feet range of this graph includes that of the prior graph (0 – 3,000 sq ft) and demonstrates a possible trajectory shift in trendlines above 3,000 square feet. In other words, the price to square foot ratio appears to escalate price more rapidly above 3,000 square feet.

Overview - List Price to Lot Size

Next we will take a closer look at lot size as a function of price. The following two graphs represent an

overview of the same Dallas real estate inventory from the perpsective of the relationship between list price and lot size. For consistency and comparibility, the price limits on this graph are the same as those applied to the prior overview graphs, $0 to $2M.

Listed Price to Lot Size
Up to 218,000 sq ft (5 Acres)

Listed Price to Lot Size
Up to 43,650 sq ft (1 Acre)
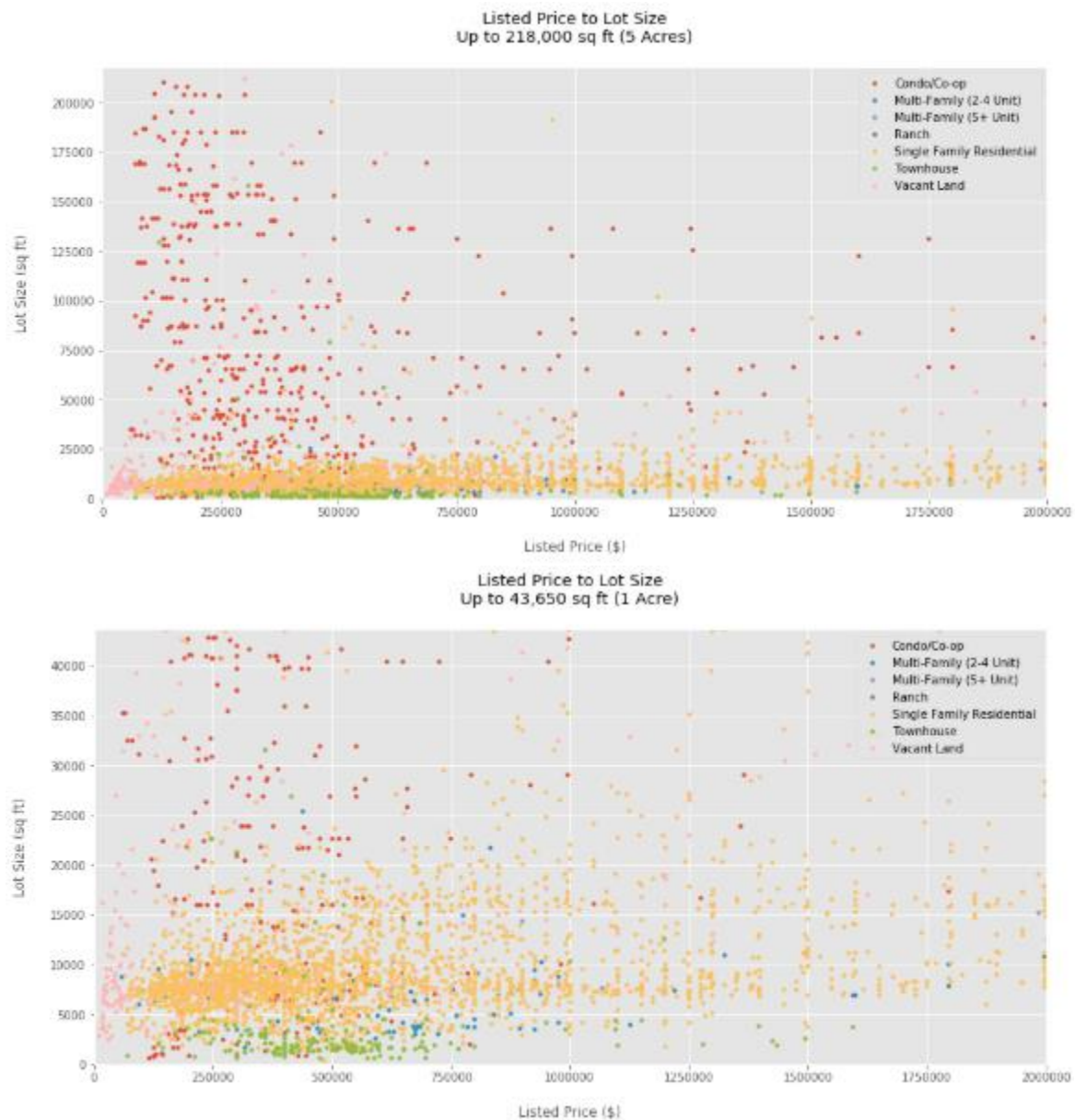
Of note in these graphs is the vertical pattern unique to the condo/co-op property type. This pattern is unique to this entire study and does not appear to demonstrate any discernible information. This pattern may be the result of differences in data entry methods from the various reporting parties providing this data. For simplicity and concission, this data element in this perspecive will be studied limitedly.

With this baseline data above in mind we can take a deeper look into overall price metrics relative to square footage.

Ratio Analysis – Price to Livable Square Feet

Listed Price to Price Per Sq Ft
$50,000 - 500,000
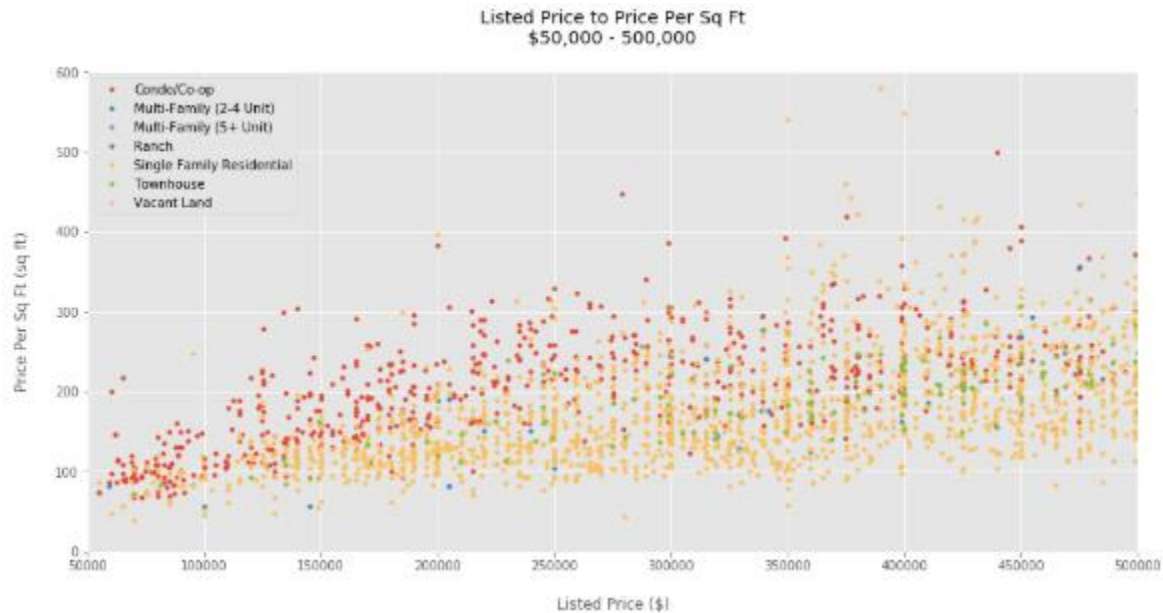
This graph plots the ratio of price per square foot relative to list prices from $50,000 to $500,000, which as we saw above is a primary grouping category for most property types. In noting the trendlines for each property type in this graph, the single family residential type demonstrates a trendline which is relatively flat in comparison to that of condo/co-op. this suggests that, within this price range, the single family residential price per square foot ratio is less susceptible than that of condo/co-op to trajectory shift as price escalates. This characteristic would be of note to a buyer wanting to maximize square footage relative to purchase price. such as buyer is more likely to obtain more square footage in a single family residence than a condo-co-op, all other things being equal.

Extending this same study to a list price range of $500,000 to $2M, the condo/co-op trendline continues above that of single family residence. Additionally, the reduced volume of inventory is noted by the decreased density on this graph.

Listed Price to Price Per Sq Ft
$501,000 - 2,000,000

Finally we can combine the data from the prior two graphs to compare price range characteristics above and below $500,000. Trendlines for both single family residence and condo/co-op increase more steeply below $500,000, then relatively flatten above this threshhold. This indicates that as purchase price increases, square footage is likely decremented in progressively lesser proportions.


Listed Price to Price Per Sq Ft
$50,000 - 2,000,000

Ratio Analysis – Price to Lot Size[1]

In the following graphs, a similar study is applied to demonstrate relationships of price to lot size at various price ranges.

Listed Price to Price Per Lot Size
$50,000 - 500,000

---

[1] Although not critical to this study, it should be noted that such property types at which land is held in common by multiple owner/residents, such as Condo/Co-op and Townhouse, may be less relevant to this element of the study and possibly could be excluded in order to simplify the graphical representation.

This graph demonstrates the ratio of price to lot size for the price points within a range from $0 to $500,000. The trendline for each property type is noted by the color of each property type and denotes a mild upward and consistent trajectory as price increases.

Listed Price to Price Per Lot Size
$501,000 - 2,000,000
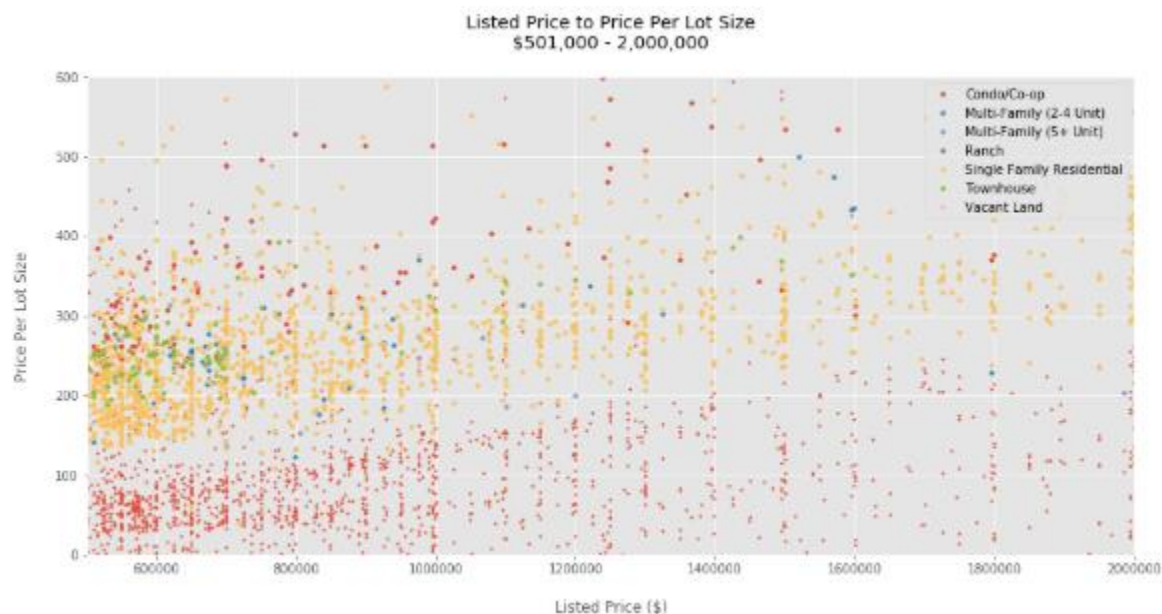
The same study applied to a price range from $501,000 to $2M demonstrates a similar mild, consistent trajectory as price increases. The trajectory angles appear similar to those at the lower price range in the previous graph with no apparent transposition of the trendlines.

Listed Price to Price Per Lot Size
$50,000 - 2,000,000
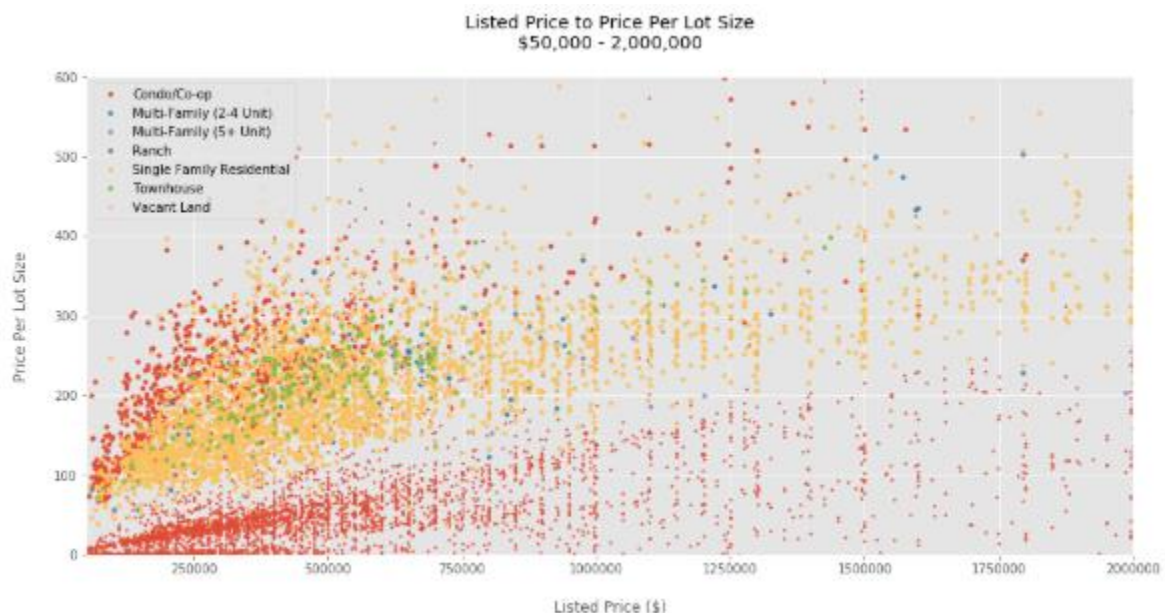
This graph combines the data, including price ranges, of the prior two graphs and confirms that the trendlines and their respective trajectories are relatively consistent throughout the combined price range of $500,000 - $2M.

Of note is that the vacant land holds a very linear and generally lower price to lot size ratio than do the other property types included in this graph. For the potential buyer whose priority is to maximize lot size

at any price level below $2M, this data suggests that the best way to do so is by purchase of vacant land.

HOA Dues

Similar to our examination for recently sold properties, asking price is likely impacted by the monthly expense of HOA dues. We do not opine on the value of the amenities or other deliverables of the HOA, which will vary significantly from property to property, but hope to find some informative trends.
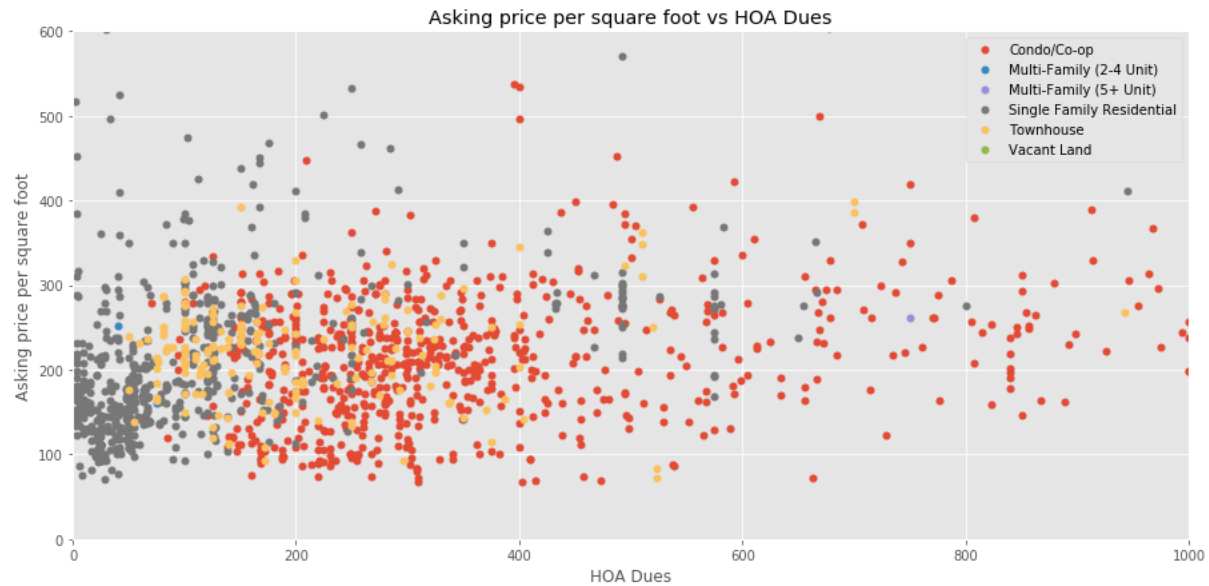
The above graph is consistent with what we learned in the recent market. HOA dues does not have an obvious trend relative to price per square feet and a buyer must evaluate these dues relative to the value of the amenities and services offered at a given property. The one common take away is that condos and townhomes generally can be expected to have higher dues, which is consistent with the shared property model for these homes.

Build Year

As we did with recently sold properties, we examine price per square foot by build year.
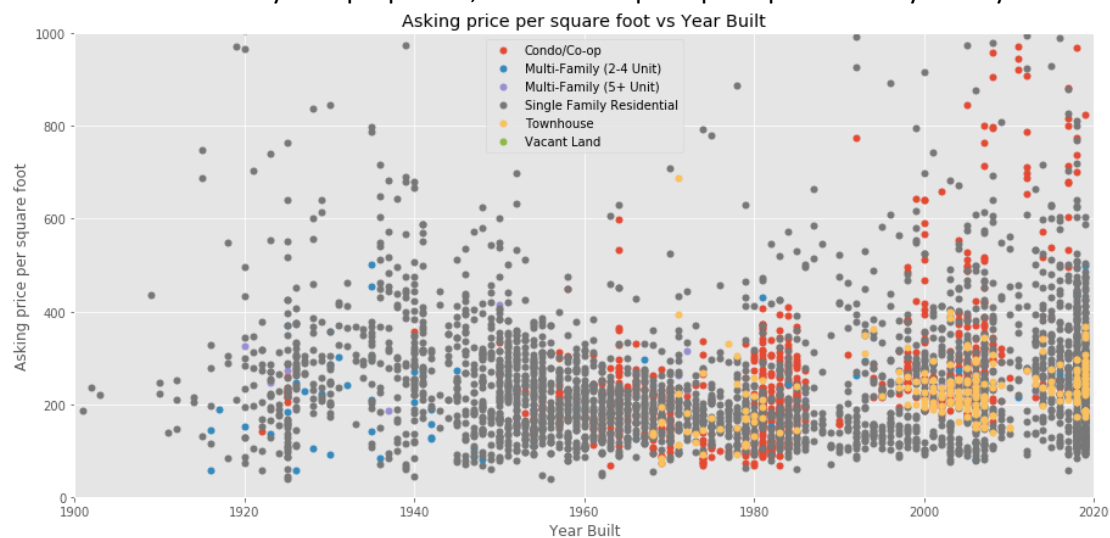
We see that the general the insights gleaned from recently sold properties continue for asking prices. However, asking prices continue to appear elevated relative to recent sales. A buyer should expect to negotiate in an effort to bring these prices down.

Section IV: Combined Inferences (Recently Sold vs Available Properties)

Now that we've had a chance to analyze both the recent market and the currently available properties we can start to evaluate where opportunities may lie and what price points are reasonable to expect to prospective buyers.

To start we will look at price dislocations by geography. What this means are areas where the average asking price is most meaningfully above the recent sales price from the same area. While not determinative, since every property is unique, this may indicate areas that a prospective buyer may want to look for deals.
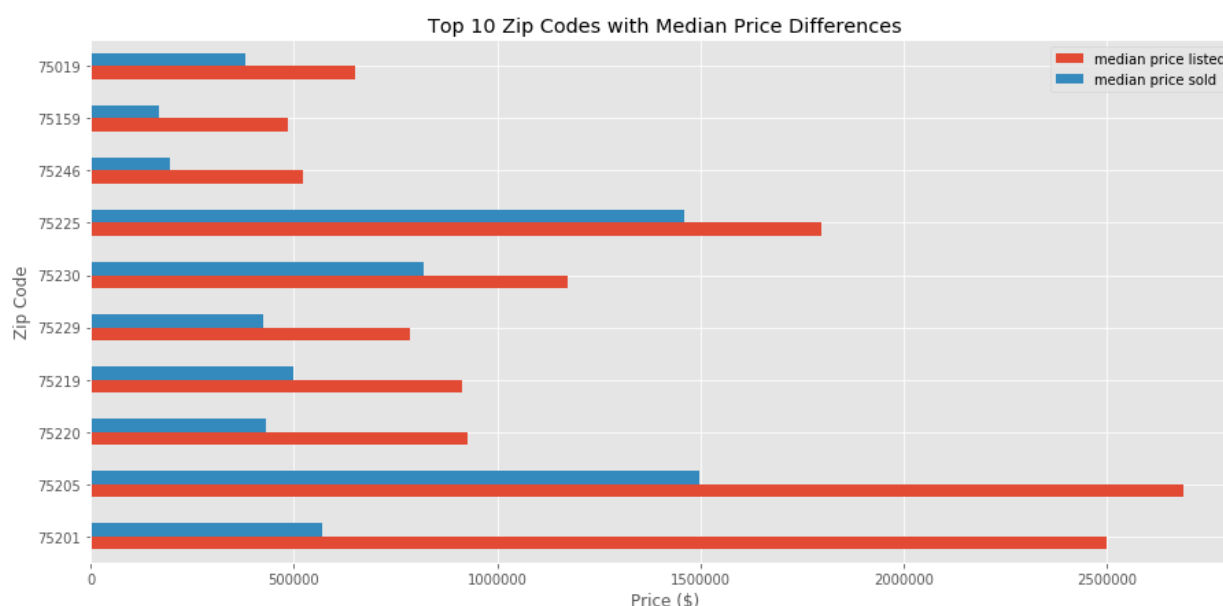
This graph shows the top 10 zip codes where the median Dallas single family residential home difference in price for the first quarter of 2019. We can see that some zip codes have a high median listing price, but low median selling price as in zip codes 75201 and 75205. There are many potential inferences here. If there is a large difference between listing and selling price, that may indicate that demand is lower in that area than anticipated and there may be an opportunity for prospective buyers. Alternately it could indicate an area in decline.

Overall, the above seems to indicate that current listing prices may be overheated, and prospective buyers should negotiate downwards relative to the listing price.

Section V: Conclusions

Ultimately, real estate, especially homes, are very situationally specific, but the trends examined above should be informative to prospective buyers. We see that negotiation is likely the right strategy across the board in Dallas as asking prices seems to exceed recent sale prices generally. Importantly, the overheated asking price market may indicate a future bubble so buyers should be careful if they're looking to realize near-term value increases.

Our investigation in the Dallas real estate data trends allowed us to conclude the following in response to our research questions:

1. Individuals that wish to purchase a home on the market should consider days on the market, year built, lot size, zip code of the property, and type of home as they tend to tell the buyer the most information for a more informed decision.
2. Zip code impacts price/availability by a significant factor. Buyers should be wary of cheaper median home selling zip codes due to the high demand for Dallas' real estate in some particular zip codes.
3. Lot size is important, but that seems to be more of a factor for suburban areas relative to more densely populated areas. Newer condos are generally preferred and a discerning buyer may find deals in the second-owner range, but due diligence is required to ensure that the build-quality and amenities are up to the buyer's standards.
4. Average type of homes listed on the market tend to not decrease in listing price over time indicating that homes that have been listed for longer may not have a lower listing price or may be a market correction from a higher original listing price. About 90 days is the cut-off where homes on the market tend to become more scarce, which means that buyers will have more negotiating power to lower the price of a home after it has been on the market for 90 days.
5. HOA-dues on home sales do not have any obvious trends in relative price per square feet which means that HOA-dues are a significant burden in regards to higher monthly fees to purchase a home. Furthermore, HOA-due homes tend to decrease in value over time, whereas single family residential homes without HOA fees have more variance in listing prices over the home's age.