Will Robinson   Springboard Capstone 2   March 17, 2021

Jacksonville Housing Analysis

Problem Statement

House sales have soared since the onset of the pandemic. Demand surged as homes became offices and schools and as many apartment renters sought out space and more control over their environment. Buyers also benefitted from low interest rates and stimulus funds.

Inventory has not kept pace, however, since those who did not have to sell their homes hunkered down. As a result, prices have surged, pricing out many first-time buyers, which creates a business opportunity for homebuilders. In seeking to serve this market, builders need to know what features and locations Jacksonville's first-time buyers are seeking and what the price range is.

I used Jacksonville's county tax roll to answer these questions. I merged parcel and sales files to create a database that detailed each home's age, square footage, lot size, features, subfeatures, sales history, key characteristics, zoning and more. I used this data to explore construction and sales trends and better understand what home traits contribute most to sales price.

Data Wrangling

The parcel and sales data maintained by the Duval County Property Appraiser are available as txt files on the appraiser's website. The parcel file contained 17 tables with pipe delimited rows, and the sales file contained four tables with fixed widths. However, the parcel file contained a table for recent sales, and the sales file contained tables that were also in the parcel file. I put both files into a single, 17-table SQLite file by replacing the parcel file's sales table with the data from the sales file and dropping the duplicative tables from the parcel file.

I added an 18th table which contained a dictionary of qualification codes, which allowed me to limit my analysis to only arms-length transactions.

This SQLite file contained data for every parcel in Duval County, so I wrote queries to limit results to only detached, single-family homes and to combine data from 12 tables into a single table with each row representing one house. Five of the tables were not useful for my analysis.

Getting the data to one row per house required creating dummy columns for features, subfeatures, zoning rights and characteristics. Because each of these have dozens of unique values, I created dummy columns for only the 10 most common of each.

The merged data needed several cleaning steps. This included using string comparison to standardize texts in mailing addresses, cities, states, zip codes and owners. This step aided in exploratory data analysis, but many of these features were not used by my models. Another step involved creating an adjusted valuation that creates distinct values for each house on parcels that contain more than one house.

I used two separate tables for exploratory data analysis, one showing the entire housing inventory and one showing about 30 years of sales. The inventory table consisted of 87 features for about 253,000 homes. The sales file contained eight rows pertaining to more than 340,000 sales. At the modeling stage, I reduced this dataset to the about 49,000 sales completed since 2017.

Exploratory Data Analysis

Jacksonville has nearly 253,000 detached single-family homes with a total valuation of more than $49 billion, by the Duval County Property Appraiser's estimation. Condominiums, townhomes, quadruplexes, duplexes and mobile homes have been excluded from this analysis.
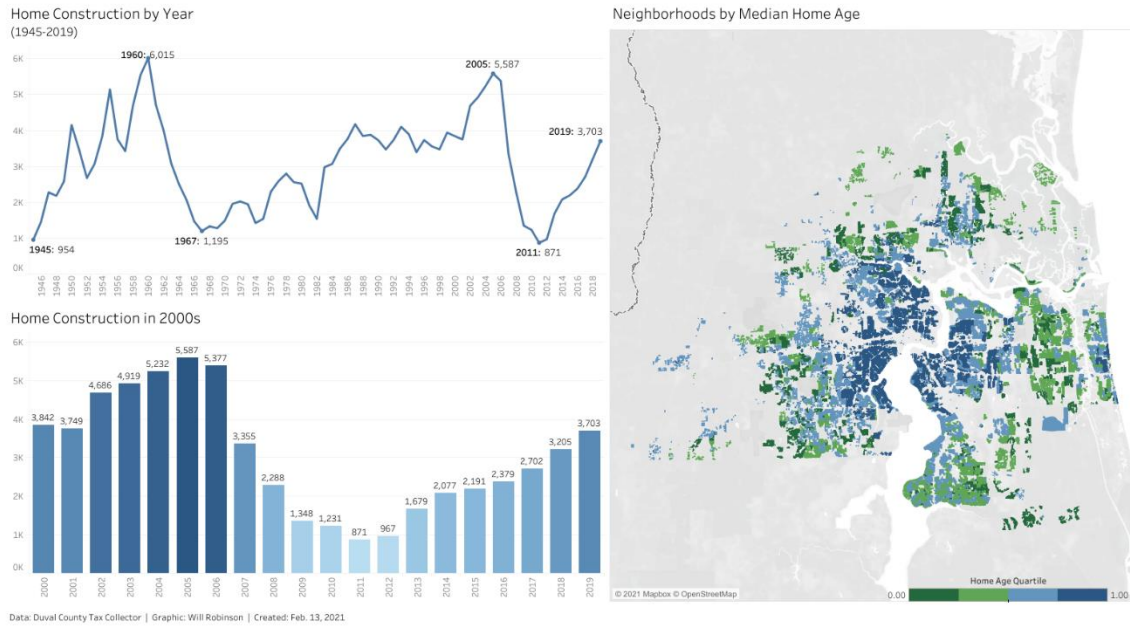
The median home in Jacksonville is 39 years old, with five years knocked off by renovations and updates. It spans a total of 2,142 square feet, 1,644-square-feet of which are heated. It has a value of about $163,000.

Construction

Home construction in Jacksonville has had multiple boom and bust periods. After reaching a peak of more than 6,000 homes constructed in 1960, home construction plummeted to about 1,200 by 1967, a total drop of 80% and an annualized decrease of about 11.5%. The market hit another peak of almost 5,600 in 2005 before dropping to 871 in 2011, a total decrease of 84.5% and an annualized decrease of 14.1%. Home construction has increased every year since 2011, totaling more than 3,700 in 2019.

Historically, home construction in Jacksonville began along the beaches and around downtown at an elbow of the St. Johns River. In the map below, the oldest 50% of homes are colored blue, with the darker shade being the oldest quartile. Some of the oldest districts include San Marco, Avondale, Springfield, the Beaches and Arlington.

More recent construction has concentrated along the Intracoastal Waterway (located between downtown and the beach), in Fleming Island and along the border with St. Johns County to the south, near the airport and waterways to the north and in the fringes of Westside, including near Cecil Commerce Center. In the map below, the youngest 50% of homes are colored green, with the darker shade being the youngest quartile.

Home Construction by Year (1945-2019)

1960: 6,015
2005: 5,587
2019: 3,703
1945: 954
1967: 1,195
2011: 871

Home Construction in 2000s

2000: 3,842, 2001: 3,749, 2002: 4,686, 2003: 4,919, 2004: 5,232, 2005: 5,587, 2006: 5,377, 2007: 3,355, 2008: 2,288, 2009: 1,348, 2010: 1,231, 2011: 871, 2012: 967, 2013: 1,679, 2014: 2,077, 2015: 2,191, 2016: 2,379, 2017: 2,702, 2018: 3,205, 2019: 3,703

Data: Duval County Tax Collector | Graphic: Will Robinson | Created: Feb. 13, 2021

Neighborhoods by Median Home Age

Home Age Quartile   0.00 — 1.00

© 2021 Mapbox © OpenStreetMap

Valuations by location

Jacksonville's most valuable land, irrespective of the buildings on them, are the beaches, the zip code encompassing the Avondale and Ortega historic districts along the river and in the Southern zip codes closest to St. Johns County. The least valuable land is located between downtown and the airport, followed closely by the westernmost zipcodes.

Jacksonville's most valuable homes, irrespective of the land beneath them, are located in the city's southernmost and Northeast Jacksonville, while the zip codes with the least valuable buildings mostly emanate outwardly from downtown in the center of Jacksonville.

This makes intuitive sense, as Jacksonville's buildings are least valuable where people are most concentrated and most valuable where people are least concentrated - that is, buildings are most valuable where land is cheap and therefore buildings tend to be larger. Houses built farthest from downtown, the St. Johns River and the ocean also tend to be younger than those built close to these areas, another reason for their higher building values.

Jacksonville's existing inventory is largest in Northside, Arlington, along the Intracoastal Waterway and to the southwest of downtown. Construction in recent years, 2017 to 2019, has been highest in Northside and Westside. This mostly inverse relationship is unsurprising; builders are going where land is cheap and readily available, but still proximal to existing concentrations of homes. Only Northside has both a high number of homes and a high number of recently built homes.
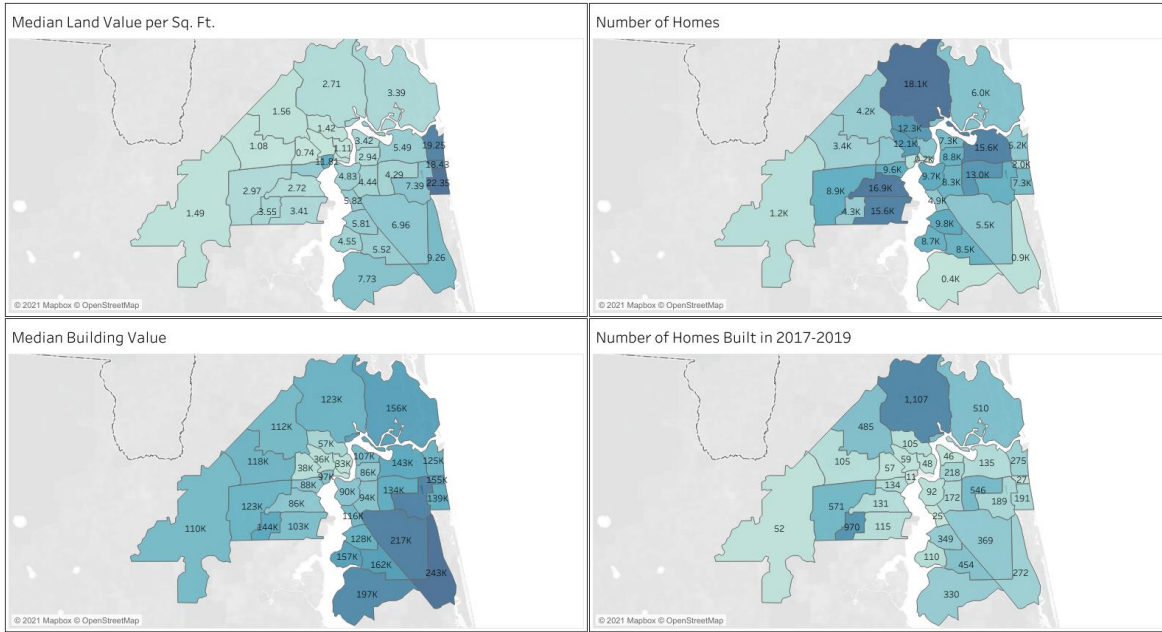
Data: Duval County Property Appraiser | Graphic: Will Robinson | Created: Feb. 13, 2021

The Duval County Property Appraiser, since taxes are levied by parcel, assesses the just value of entire parcels, regardless of the number of buildings on them. To determine the values of each home on a parcel with multiple homes, I therefore calculated an adjusted value. This adjustment takes the ratio of a home's gross square feet to the total gross square feet of buildings in the parcel, applies this ratio to the land value of the parcel, then adds the building value and feature value of each particular building.

The 10 most valuable homes by this calculation are listed below and range in value from $4.7 million to $9.1 million. Each has access to water - the majority to the ocean and the minority to the St. Johns River.
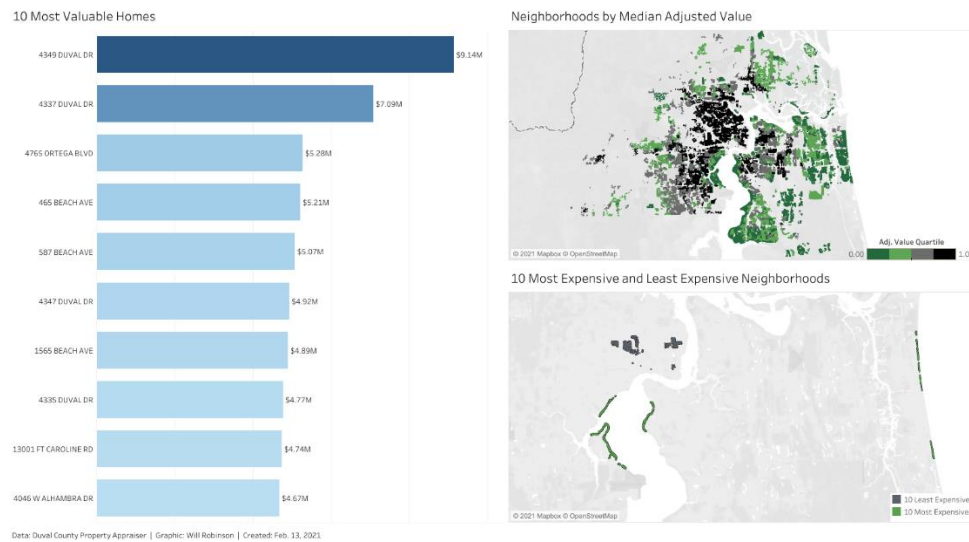
The 10 most valuable neighborhoods are those in the Beaches, San Marco, Avondale and Ortega. Each of these neighborhoods is located along water and has historic stock. This likely reflects not only the value of being near water but also the fact that those who can afford to buy waterfront property can typically afford to invest heavily into the homes as well. Indeed, buildings exposed to the effects of salt water and in historic districts require extensive upkeep and renovations.

The 10 least valuable neighborhoods are landlocked, adjacent to interstates and industrial areas and are closer to areas of downtown blight than areas of downtown revitalization. The neighborhoods are composed of old homes that have not received the significant renovation dollars of the Avondale, Ortega and San Marco historic districts.

That age plays a major role in home valuations is clear from the map of neighborhood median adjusted values below. The most valuable 50% of homes are in green, with dark green depicting the most valuable quartile. The least valuable 50% of homes are in gray, with the darkest shade depicting the least valuable quartile.

The most recently constructed homes are located along the Intracoastal Waterway to the east, in Fleming Island and near St. Johns County to the south, in the far west of Westside and towards Nassau County to the north are all green, while the oldest stock of homes in Arlington, between downtown and the airport and in Westside are all black.

Two other features stand out: homes along the ocean and the river are dark green, as are designated historic districts.

10 Most Valuable Homes

Address	Value
4349 DUVAL DR	$9.14M
4337 DUVAL DR	$7.09M
4765 ORTEGA BLVD	$5.28M
465 BEACH AVE	$5.21M
587 BEACH AVE	$5.07M
4347 DUVAL DR	$4.92M
1565 BEACH AVE	$4.89M
4335 DUVAL DR	$4.77M
13001 FT CAROLINE RD	$4.74M
4046 W ALHAMBRA DR	$4.67M

Data: Duval County Property Appraiser | Graphic: Will Robinson | Created: Feb. 13, 2021

Neighborhoods by Median Adjusted Value

Adj. Value Quartile

© 2021 Mapbox © OpenStreetMap

10 Most Expensive and Least Expensive Neighborhoods

© 2021 Mapbox © OpenStreetMap

10 Least Expensive
10 Most Expensive

Some of the characteristics that we've seen contribute to home value, including whether a home is in an historic district or is in close proximity to a body of water, are demarcated in the county property appraiser's database.

Below are maps of homes with at least one of the 10 most common traits listed in the database. The parcels are heat mapped, with darker hues indicating a higher just value, as estimated by the property appraiser.

Maps of Top 10 Designated Characteristics

Avondale — Just Value $35,113 $4,168,252
Canal — Just Value $63,726 $2,844,505
Golf Course — Just Value $88,787 $1,105,182
Lake — Just Value $24,283 $2,020,399
Misc. Waterway — Just Value $72,844 $1,293,601
Ocean Front — Just Value $594,727 $9,138,427
Ortega — Just Value $92,497 $3,483,574
Springfield — Just Value $9,308 $593,364
St. Johns River — Just Value $49,819 $5,638,809
Trout River — Just Value $36,019 $1,823,510

Data: Duval County Property Appraiser | Graphic: Will Robinson | Created: Feb. 13, 2021

Of the 10 most common characteristics, six pertain to water access, three pertain to historic districts, and the 10th designates homes in golf course neighborhoods.

The oceanfront characteristic is associated with the most valuable homes. This makes sense as oceanfront land is the most valuable land in Jacksonville, the supply of homes (317) is smaller than any of the other top characteristics and the proximity to the ocean requires extensive upkeep, as shown by the fact that the median effective age of oceanfront homes is 42 years younger than the median age of oceanfront homes.
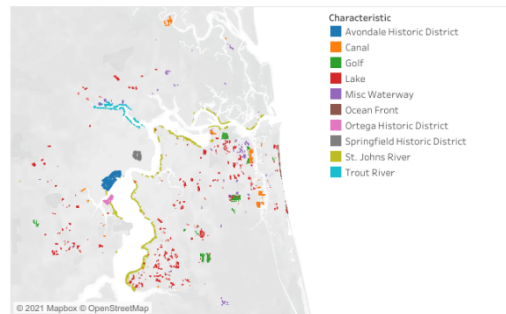
Oceanfront homes are also large. These homes are tied with homes along the St. Johns River for the most median bedrooms and bathrooms with four of each, and the median heated square feet of oceanfront homes are 2nd only to homes along the St. Johns River. Proximity to the St. Johns River is unsurprisingly the 2nd most valuable of the 10 most common characteristics. Ocean frontage still outweighs river frontage, however, as both the median adjusted value per gross square foot and the median adjusted value per heated square foot of oceanfront homes vastly exceed those of homes along the river - or along canals, lakes and miscellaneous waterways.

As previously noted, homes in historic districts have been extensively renovated, as shown by a significant difference between their respective median ages and median effective ages.

Springfield is associated with the least valuable homes of the 10 most common characteristics. That is not to say it is associated with homes that are not valuable. In fact, Springfield is the exception to the rule within its surroundings. Springfield is located immediately north of downtown, in close proximity to the neighborhoods shown in a previous map to be Jacksonville's 10 least valuable.

Unlike its neighbors, Springfield has benefited from a period of revitalization helped by historic tax credits, low interest or interest-free loans from nonprofits and other incentivized funding. Its affordability compared to downtown's other satellite neighborhoods and a multi-year period of downtown redevelopment have added wind to its sails, though whether that momentum continues after the financial ramifications of the Covid-19 pandemic remains to be seen.

10 Most Common Home Characteristics
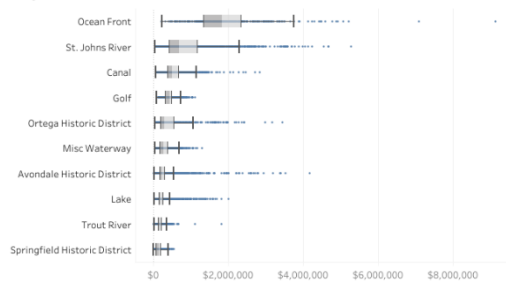
Overview of Market Statistics
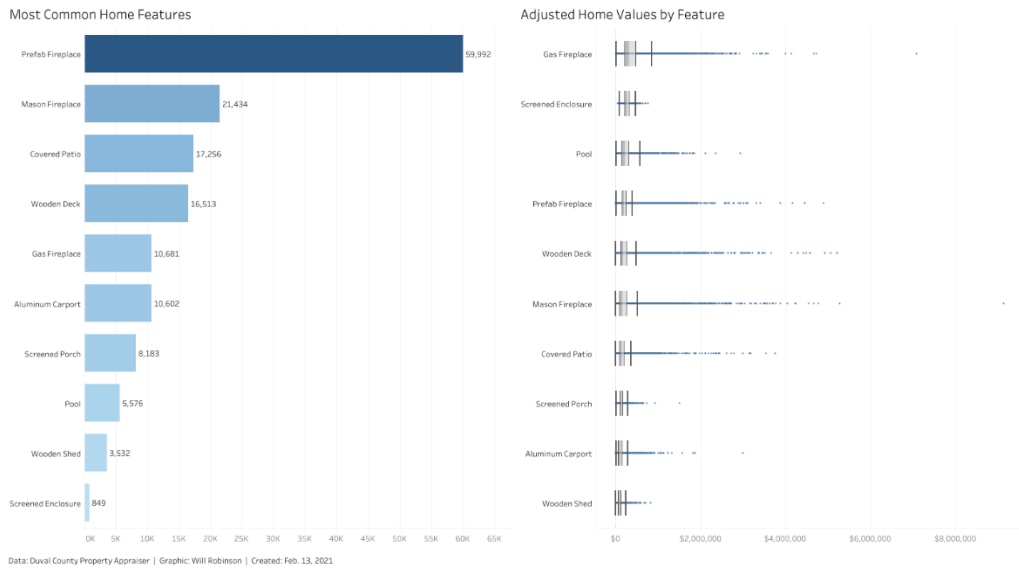(Highest to lowest median adj. value per gross sq. ft.)

	Ocean Front	Canal	St. Johns River	Ortega Historic District	Avondale Historic District	Golf	Misc Waterway	Lake	Trout River	Springfield Historic District
Median Age	63	33	40	81	95	28	18	32	61	107
Median Effec Age	21	27	31	51	51	27	18	31	48	51
Median Bedrooms	4	4	4	3	3	4	4	3	3	3
Median Baths	4	3	4	2	2	3	3	2	2	2
Median Adjusted Value	$1,827,010	$503,244	$696,529	$289,803	$225,514	$430,115	$276,654	$193,990	$175,775	$152,085
Median Building Value	$403,881	$245,385	$295,151	$141,165	$131,985	$290,577	$199,720	$133,764	$103,628	$119,296
Median Gross Sf	4,285	3,443	4,998	2,488	1,924	4,042	3,341	2,316	2,580	2,336
Median Effec Sf	3,449	2,902	4,122	2,139	1,681	3,471	2,870	2,048	2,153	2,028
Median Heated Sf	3,184	2,560	3,564	1,977	1,552	3,069	2,485	1,809	1,904	1,958
Houses	317	1,189	1,577	446	3,813	1,223	744	4,416	531	1,221
Median Adj. Value per Gross SF	440	152	144	124	122	102	87	85	69	64
Median Adj. Value per Heated SF	574	204	204	158	149	135	116	110	95	78

Adjusted Home Values

Data: Duval County Property Appraiser | Graphic: Will Robinson | Created: Feb. 13, 2021

Valuations by home type and feature

In addition to the characteristics discussed above, the county property appraiser database designates features and subfeatures included in each home. The 10 most common features include prefabricated fireplaces, mason fireplaces, covered patios, wooden decks, gas fireplaces, aluminum carports, screened porches, pools, wooden sheds, and screened enclosures.

They are sorted from most common to least common in the graph on the left and from highest to lowest median adjusted home value on the right. Note that in this case, most common does not mean the count of each feature, but rather the count of homes that have at least one of that feature. For example, there are not 59,992 prefabricated fireplaces in Jacksonville; rather, there are 59,992 homes with at least one prefabricated fireplace.
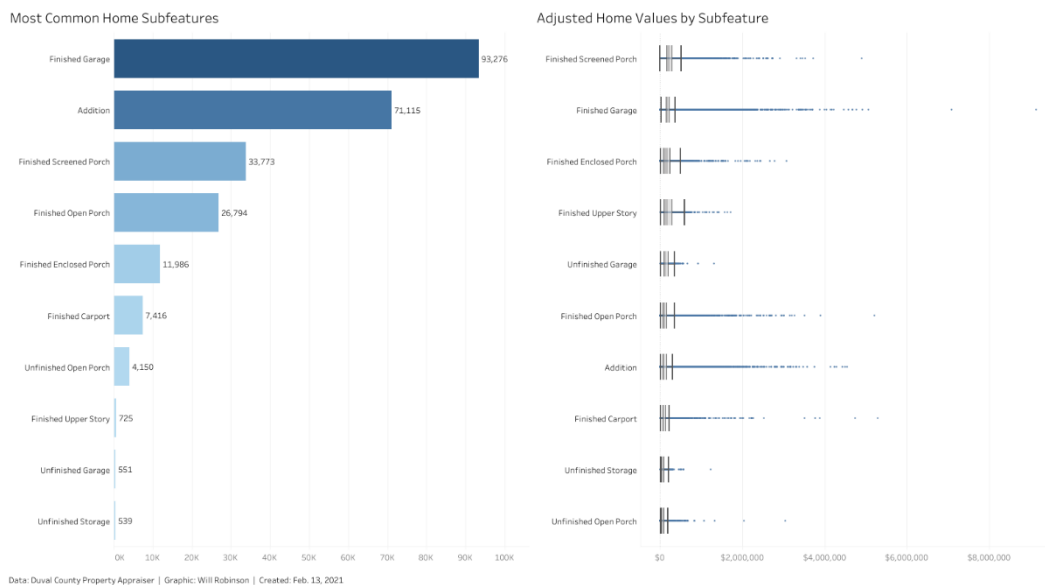
A prefabricated fireplace is the most common feature, while a gas fireplace is associated with the most valuable homes.

Most Common Home Features | Adjusted Home Values by Feature

Data: Duval County Property Appraiser | Graphic: Will Robinson | Created: Feb. 13, 2021

The 10 most common subfeatures include finished garages, additions, finished screened porches, finished open porches, finished enclosed porches, finished carports, unfinished open porches, finished upper stories, unfinished garages and unfinshed storage.

They are sorted from most common to least common in the graph on the left and from highest to lowest median adjusted home value on the right. As with features, most common does not mean the count of each subfeature, but rather the count of homes that have at least one of that subfeature.
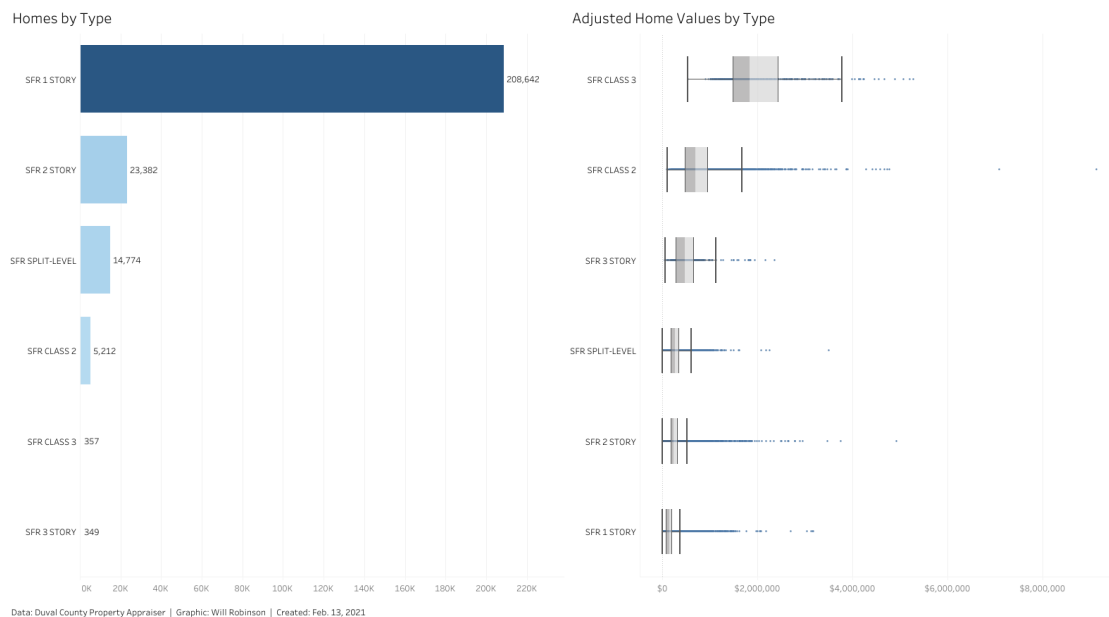
A finished garage is the most common subfeature, while a finished screened porch is associated with the most valuable homes.

Most Common Home Subfeatures | Adjusted Home Values by Subfeature

Data: Duval County Property Appraiser | Graphic: Will Robinson | Created: Feb. 13, 2021

The vast majority of Jacksonville homes, almost 207,000, are single-story. This makes sense as Jacksonville is the largest city by land mass in the continental U.S., meaning there is

ample land to build sprawling single-story homes and little incentive to build vertically. There are fewer than 350 three-story homes in Jacksonville.

However, single-story homes have the smallest median home value. This may be a result of the type's commonality - every other kind of home type is comparatively rare. Having more than one story may also be an indicator of a home's proximity to water, downtown and other valuable areas, since they have higher concentrations of buildings and would be more likely to justify vertical construction. It also may be an indication of a younger home, since more than one story can presumably be associated with later stages of an area's development, newer construction methods and/or more recent aesthetic preferences.
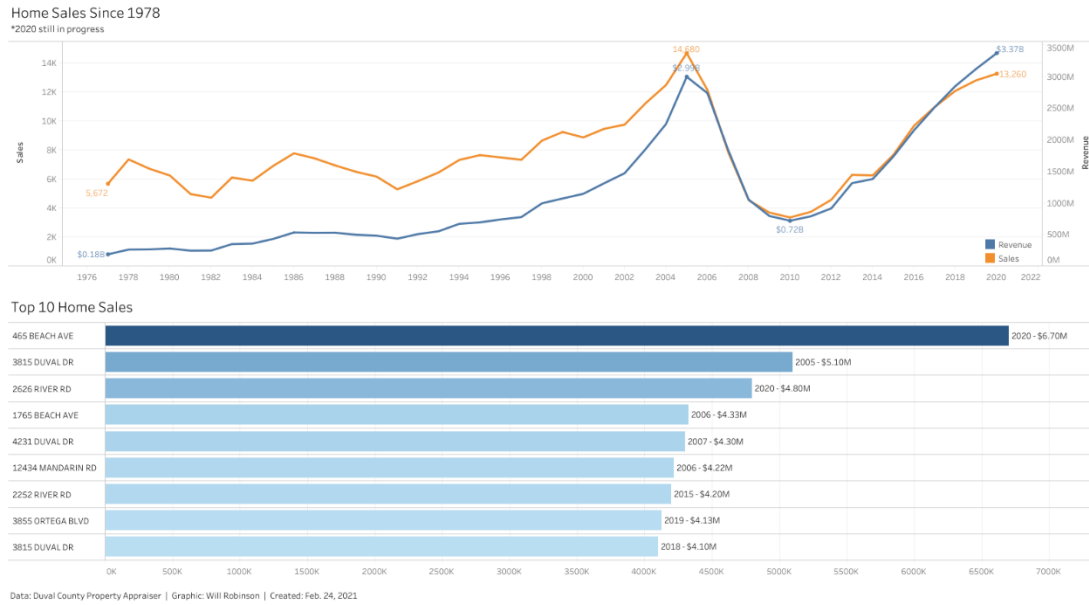

Homes by Type

SFR 1 STORY	208,642
SFR 2 STORY	23,382
SFR SPLIT-LEVEL	14,774
SFR CLASS 2	5,212
SFR CLASS 3	357
SFR 3 STORY	349

Adjusted Home Values by Type (SFR CLASS 3, SFR CLASS 2, SFR 3 STORY, SFR SPLIT-LEVEL, SFR 2 STORY, SFR 1 STORY)

Data: Duval County Property Appraiser | Graphic: Will Robinson | Created: Feb. 13, 2021
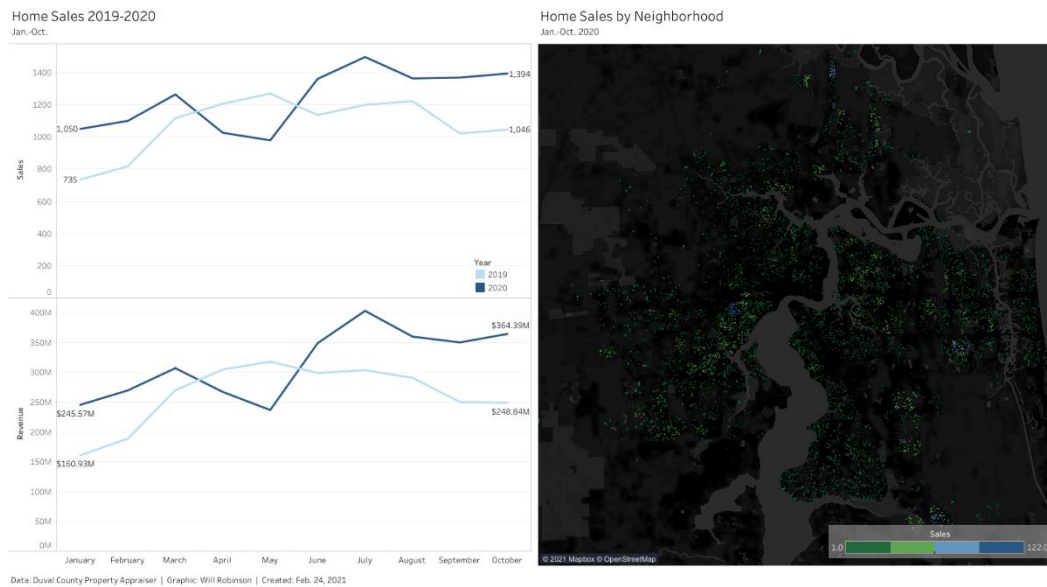
Home sales

As in most U.S. cities, Jacksonville saw a major surge in home sales up to 2005 followed by a very steep decline. Jacksonville hit a peak of more than 14,500 sales in 2005, totaling almost $3 billion in revenue. Four of Jacksonville's 10 largest home sales occurred between 2005 and 2007.

The city finally returned to these record levels last year. The data used for this analysis consists only of qualified sales, meaning sales verified as arms-length transactions with no special circumstances. Because of this, the 2020 data is current only through October.

In the first 10 months of 2020, Jacksonville had its highest number of sales in any year since 2005 and its highest sale revenue of all time.

Home Sales Since 1978
*2020 still in progress

Top 10 Home Sales

465 BEACH AVE	2020 - $6.70M
3815 DUVAL DR	2005 - $5.10M
2626 RIVER RD	2020 - $4.80M
1765 BEACH AVE	2006 - $4.33M
4231 DUVAL DR	2007 - $4.30M
12434 MANDARIN RD	2006 - $4.22M
2252 RIVER RD	2015 - $4.20M
3855 ORTEGA BLVD	2019 - $4.13M
3815 DUVAL DR	2018 - $4.10M

Data: Duval County Property Appraiser | Graphic: Will Robinson | Created: Feb. 24, 2021

Despite initial fears of a slowdown after the onset of the Covid-19 pandemic, home sales continued to surge through 2020. Sales dipped in March and April when the city issued lockdown mandates, but since June both sales and sale revenue have been well above 2019 levels.

Home Sales 2019-2020
Jan.-Oct.

Home Sales by Neighborhood
Jan.-Oct. 2020

Data: Duval County Property Appraiser | Graphic: Will Robinson | Created: Feb. 24, 2021

Model Selection

The goal of my classification models is to predict whether or not a house is classified as a starter home. I also want to better understand what features determine that classification. In labeling my data, I used the same standard as Business Insider in an article it published in partnership with Zillow, in which it defined starter homes as "homes valued in the lower one-

third of all homes" in each market. Therefore, my target metric was the percentile rank of home prices.

My first model was a Random Forest, which allowed me to efficiently make predictions despite having dozens of features. This model, which did not consider price or price percentile, was able to predict whether a home was a starter home with 93% accuracy. More specifically, 95% of its starter home predictions were correct, and 89% of its non-starter home predictions were correct. This model used 300 estimators and a maximum depth of 15.

A review of feature importance found that valuation was the most important predictor, followed by the number of bathrooms, the age of the home, the size of tax exemptions, the quality of added home features and lot size. These factors were more important than a home's number of bedrooms, number of floors and whether it had any one feature or subfeature. Interestingly, the characteristic columns, which showed whether a home was adjacent to a waterway, within an historic district, etc., was not among the 30 most important features.

I then created a K-Nearest Neighbors classifier using only the 10 most important features from the Random Forest classifier. I performed a randomized search to identify the best K, which was 28. This performed almost exactly as well as the Random Forest classifier. It was 93% accurate with 95% accuracy when predicting starter home and 88% accuracy when predicting non-starter home.

Future Research

This project focused only on homes in Jacksonville, Florida. It could easily be replicated on other asset types in the same county or on homes in other areas. Furthermore, one could explore if there is any correlation between asset types; for example, are neighborhoods with high home values the same neighborhoods with high retail property values?

My classification looked at the 10 most common characterizations, most of which related to whether a home was near a certain body of water (ocean, St. Johns River, canals, Ortega River, etc.). It could be instructive to create a single binary feature of whether a property is adjacent to any body of water.

Similarly, one could explore any value changes caused by proximity to other features, including the airport, landfills, major retail centers, etc., rather than relying on only those locales noted by the county property appraiser.