

SRI CASPER



HOME VALUE ANALYSIS



OUTLINE

- Introduction
- Data Collection
- Data Wrangling
- Data Analysis Insights
- Data Modeling
- Data Evaluation
- Future Plans





PROBLEM STATEMENTS

1. What are the factors or features that affect the value of the properties?
2. What is the home value changing rate from 2016 to 2017?
3. What is the price range of a certain type of property?



DATA COLLECTION

Resource

kaggle  Zillow®

Dataset

2,985,217 rows
58 columns

Year

2016 - 2017

DATA WRANGLING

Removed
36
columns

Replace
Missing Values
KNN Imputer

Columns
with
80%
of missing values
will be removed

Added
2
columns

DATA WRANGLING

Added Columns

2

columns

`ageinyear`

- The year the property was built to the inspection year

`assessedvalue`

- The sum of structure value and land value

UPDATED DATASET

parcelid	bathroomcnt	bedroomcnt	buildingqualitytypeid	calculatedfinishedsquarefeet	...	roomcnt	unitcnt	yearbuilt	structuretaxvaluedollarcnt
14165040	3.5	6	4	2774.0	...	7.0	4.0	1970	189508.0
12144802	2.0	4	4	1818.0	...	7.0	1.0	1956	377159.0
13101491	3.0	4	4	2509.0	...	8.0	1.0	1981	281659.0
13926960	2.0	3	7	1619.0	...	5.0	2.0	1959	47244.0
13980106	2.5	4	7	1976.0	...	7.0	1.5	1954	102352.0

Rows

29,852

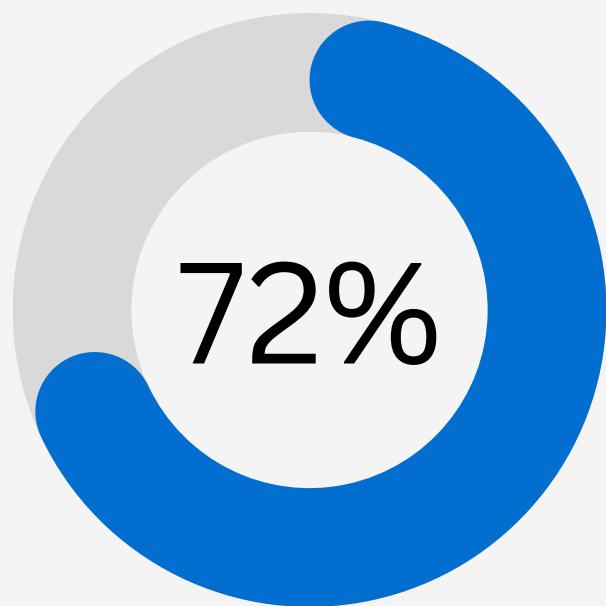
Columns

24

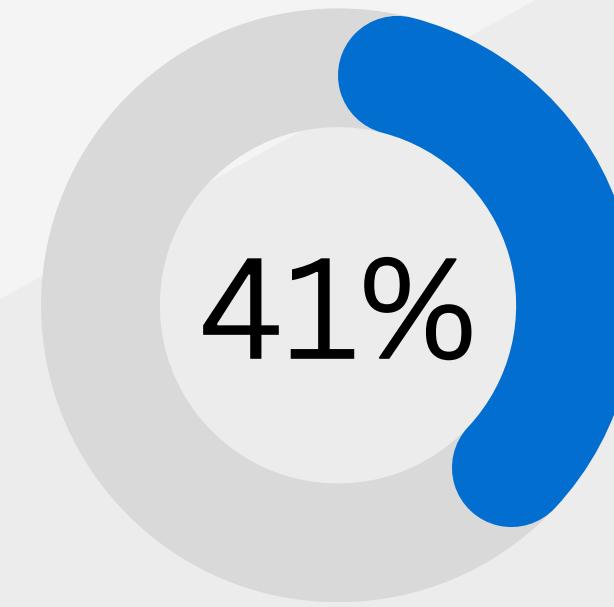
Year

2016 - 2017

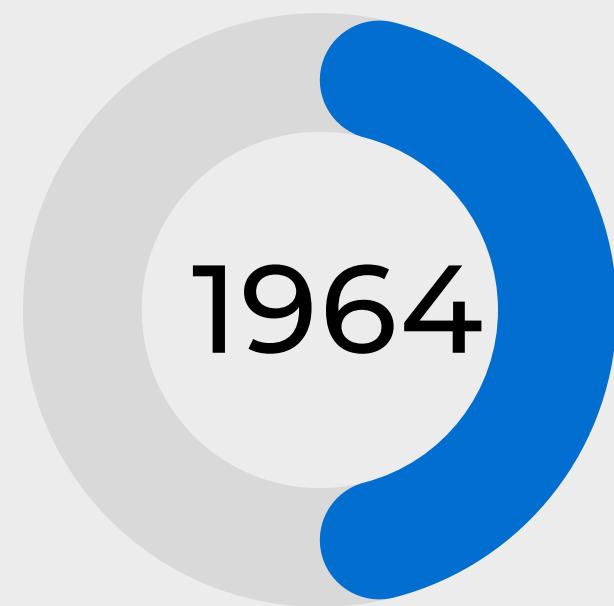
DATA ANALYSIS INSIGHTS



- 72 % of the properties are Single Family Residential



- 2 bedrooms & 2 bathrooms



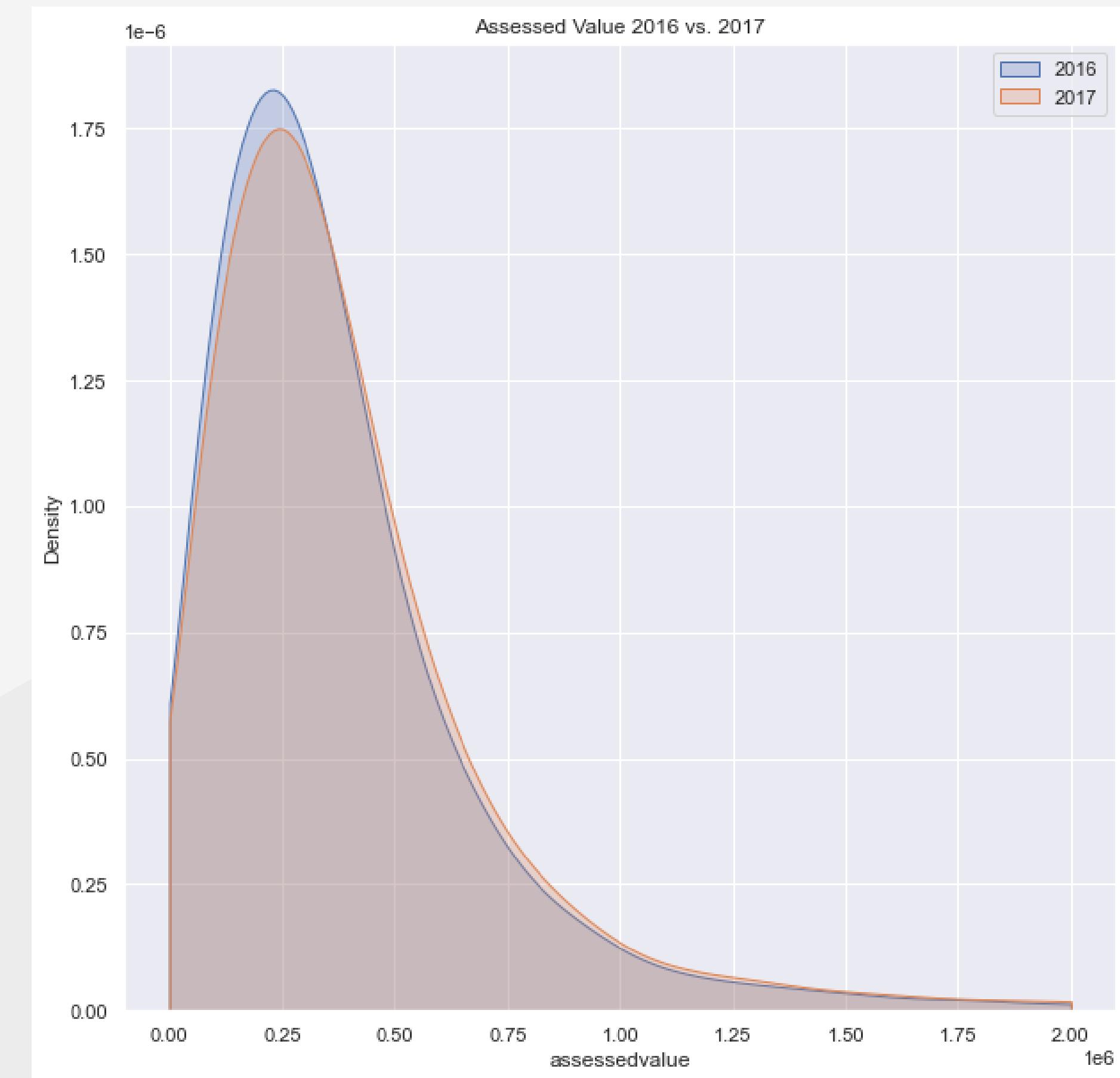
- Average of the year the properties were built

INSIGHTS



ASSESSED VALUE

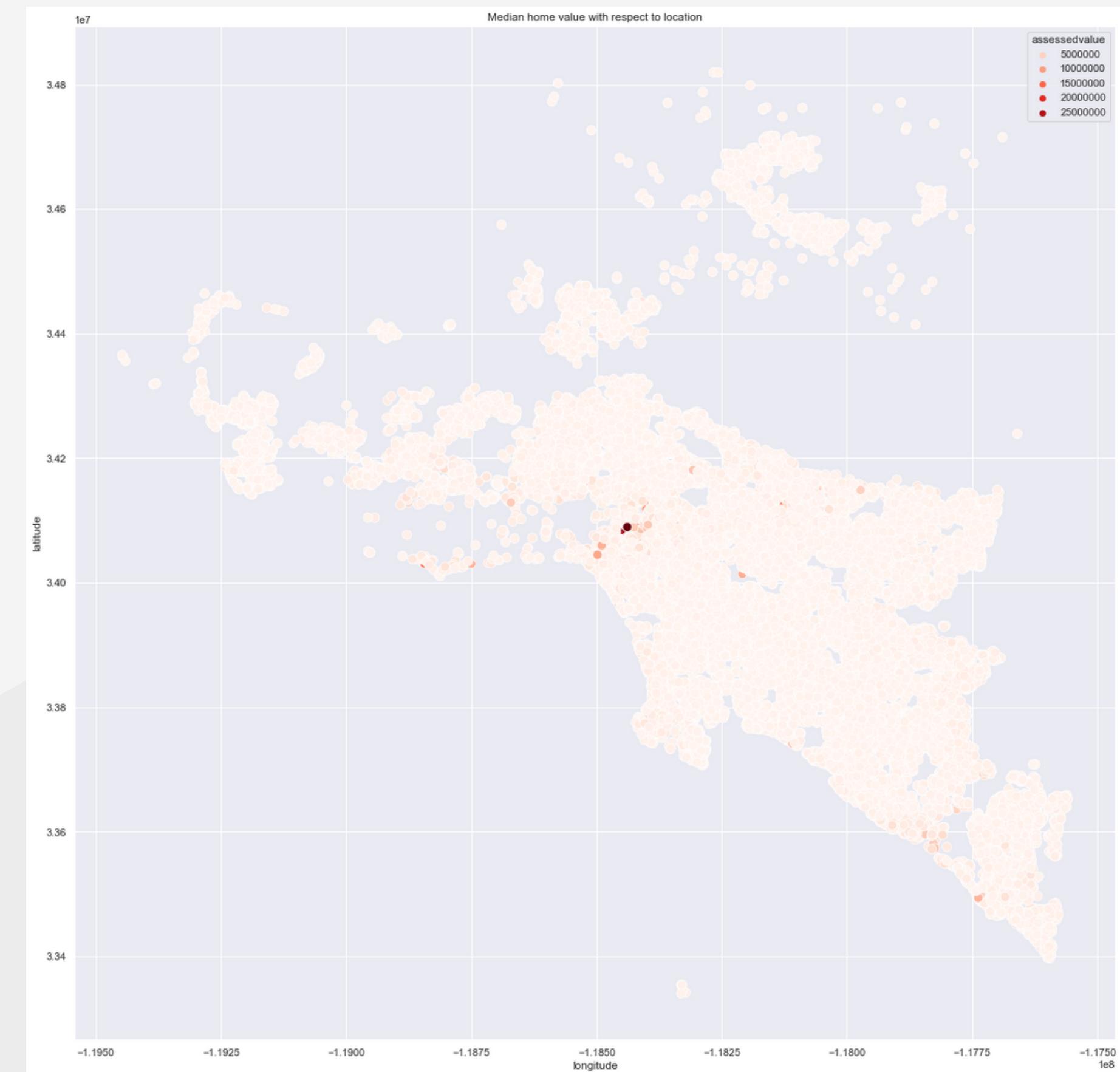
- The average assessed value: \$424,220.
- The median assessed value: \$303,250.
- The highest assessed value: \$29,502,930.



INSIGHTS



ASSESSED VALUE & LOCATION



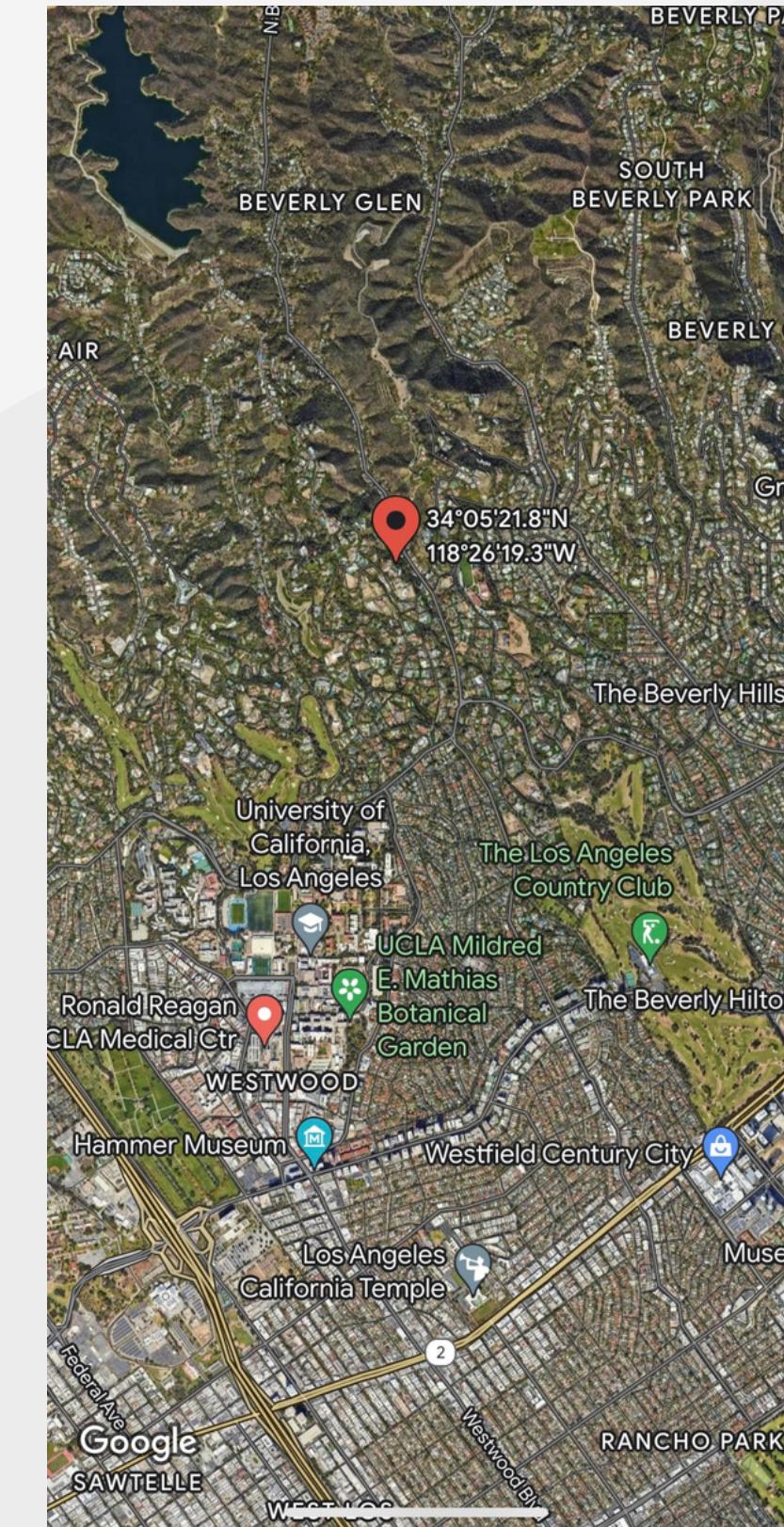
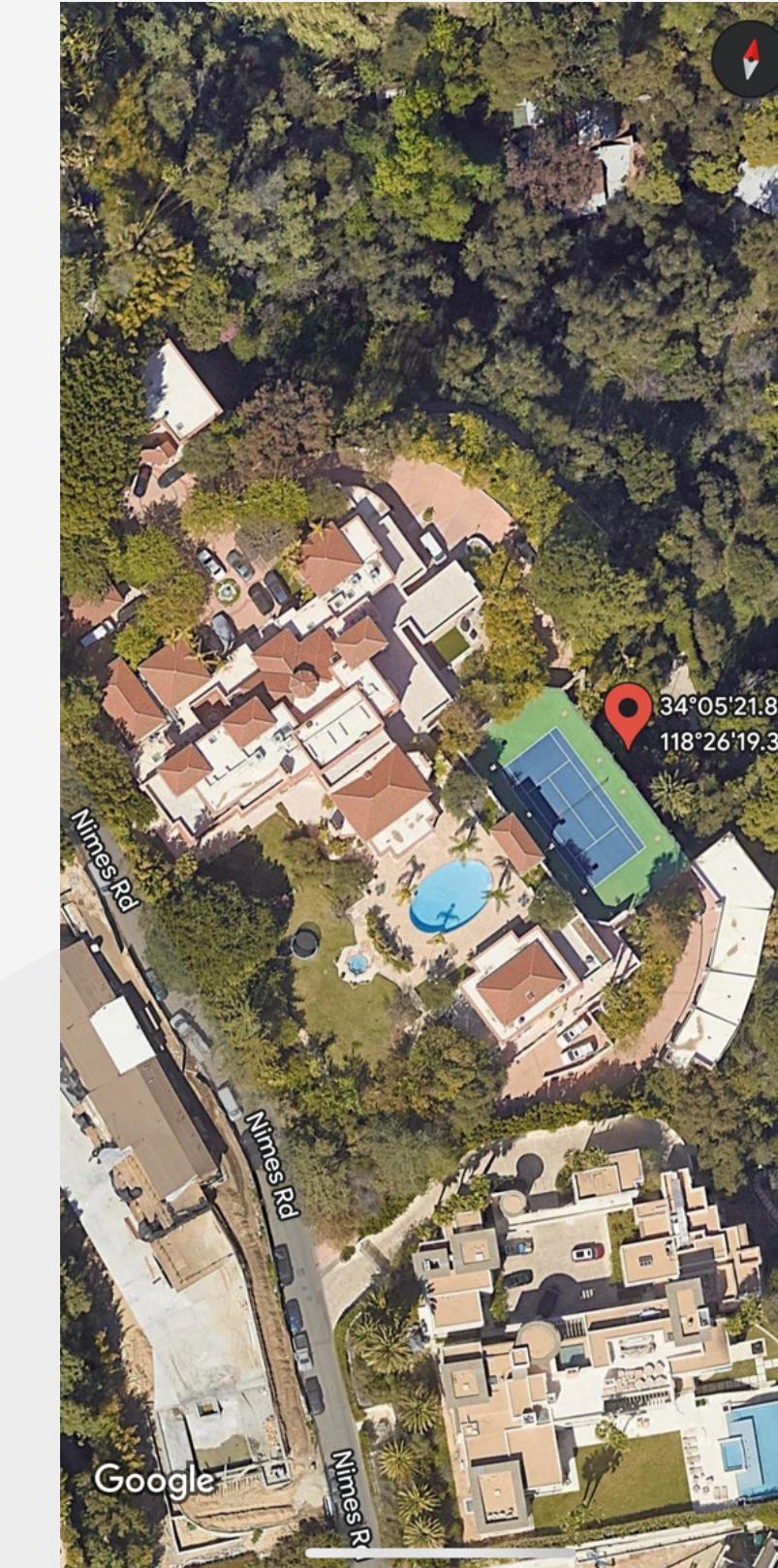
EDA

0 0 0 0

MOST EXPENSIVE PROPERTY IN 2016

- 6 bedrooms
- 15 bathrooms
- 18,417 sqft. of living area
- Built in 1985
- Assessed value: \$29,502,930

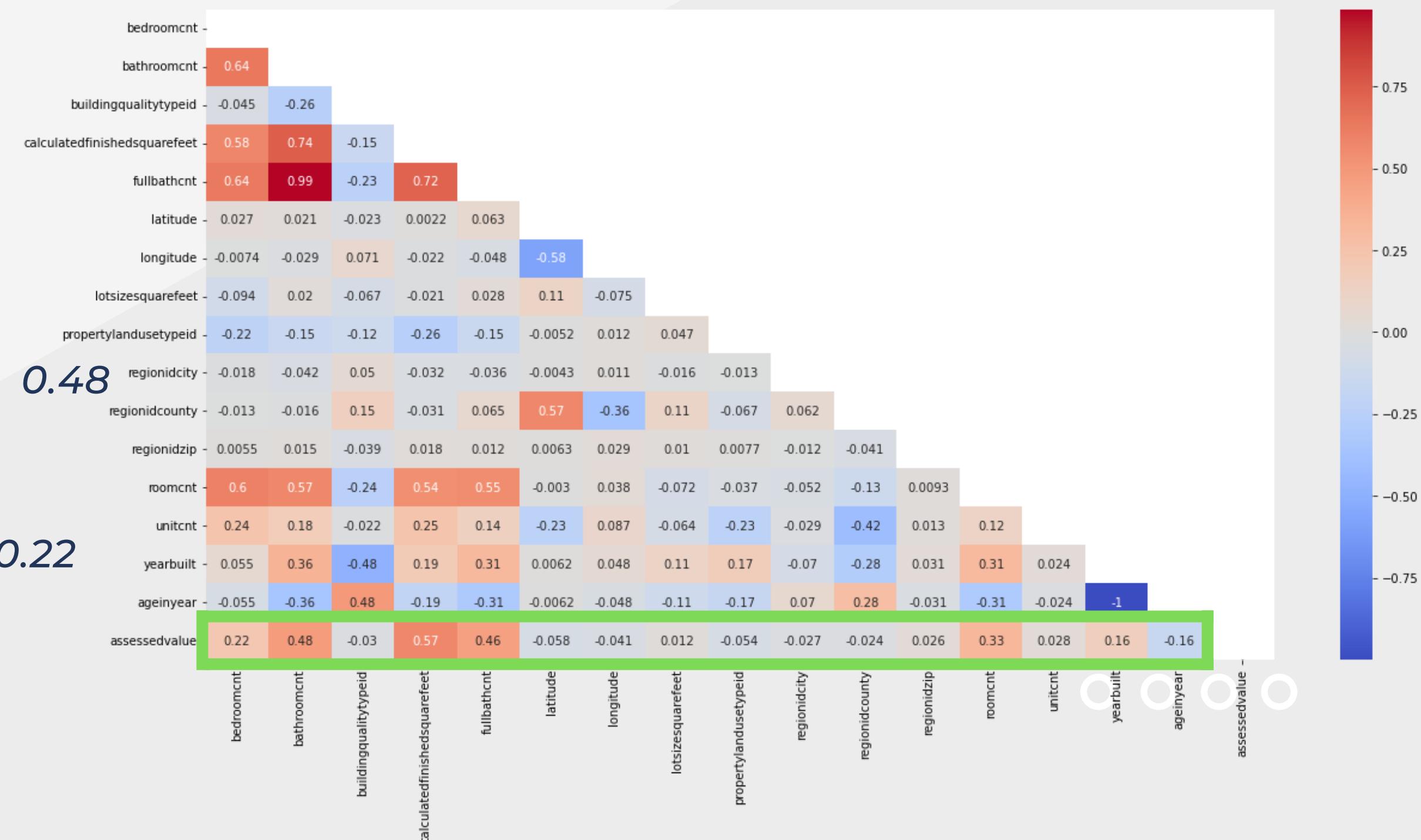
0 0 0 0



○ ○ ○ ○

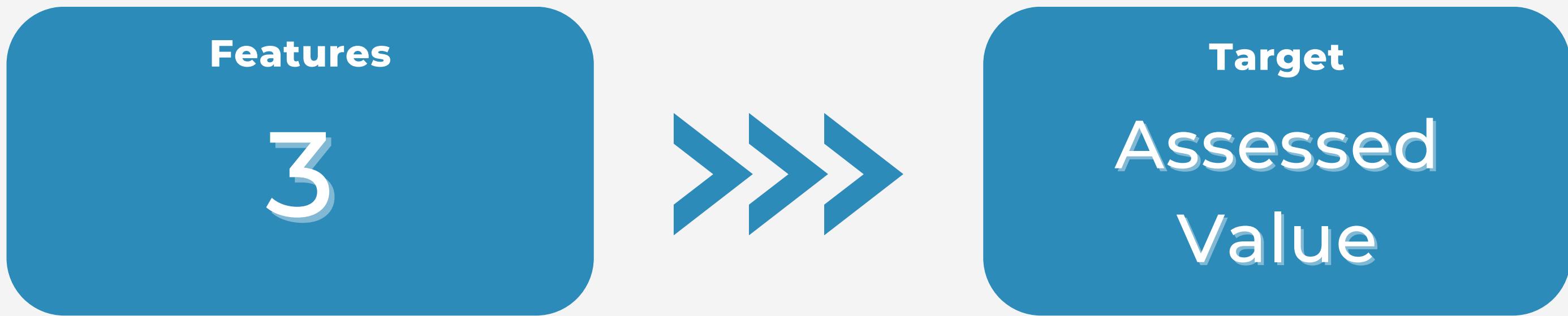
CORRELATION COEFFICIENT

- Assessed value & Living area: 0.57
- Assessed value & Bathroom count: 0.48
- Assessed value & Bedroom count: 0.22





DATA MODELING





MODEL EVALUATION

3 features

- Bedroom Count

- Bathroom Count

- Amount of Living Area

Model	Test Score
Linear Regression	37.18%
Lasso Regression	37.18%
Decision Tree	45.59%
Random Forests	45.95%
Gradient Boost	41.42%
XG Boost	44.37%
KNN	36.22%



MODEL TESTING

Bedroom

2

Bathroom

2

Living Area (sqft.)

1500

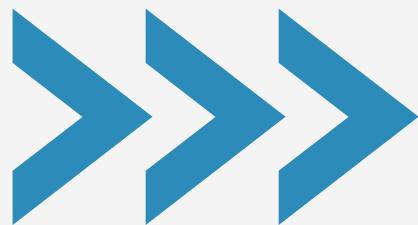
Assessed Value

\$159,602



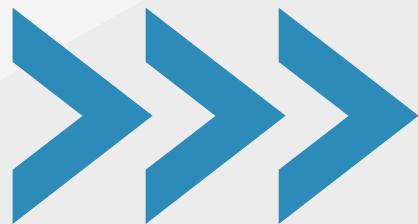
SUMMARY

1. What are the factors or features that affect the value of the properties?



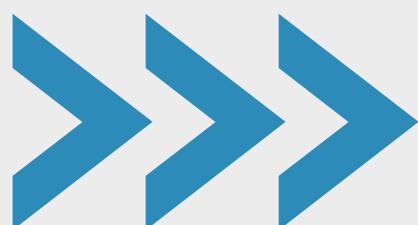
Bedroom count, Bathroom count, Amount of living area

2. What is the home value changing rate from 2016 to 2017?



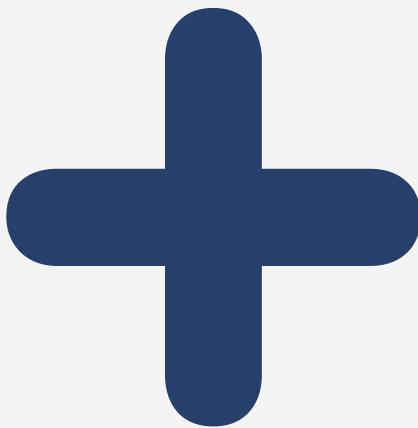
2.97%

3. What is the price range of a certain type of property?



Single Family Residential:
\$437,793 - \$25,605,020

FUTURE PLANS



**Low features,
more accurate**

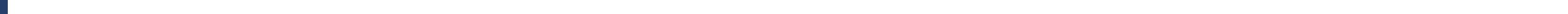
**Time Series &
Forecasting**



**Check models
for errors**

**Data with
more features
and factors**

**Create a
website or
an application**



Look forward to working with you

SRI CASPER



Austin, Texas



linkedin.com/in/sricasper/



srisakul.casper@gmail.com

