Applied Demo: Discovering Market Patterns Using Spatial Analysis Workflows

Daniel Fasteen, Ph.D.

Senior Research Scientist

P&R Special Projects

daniel.fasteen@tylertech.com

TYLER CONNECT

EMPOWER | COLLABORATE | IMAGINE

Agenda

Introduction and Problem

Jupyter Notebooks and Python

Data and EDA

Build Models

Visualize Results

Application of Clustering Methodologies

Other Methodologies that Could Be Applied

Summary Stats and Visualizations

tyler technologies

Overview

TYLER CONNECT

tyler technologies

Disclaimer

This session is meant to be a “What you can do Session” given open-source technology with the ability to analyze a data extract from CAMA.

Data Science and technology is advancing quickly and Assessors offices have more tools than just excel these days.

There is a bit of a learning curve, but with AI and YouTube, you can pretty much do anything. ☺

This analysis is just one pipeline that the analyst can use. There are endless possibilities of how one could assemble an analysis.

I will make the cluster code available on Github so you can apply to your dataset!

Problem Scope

Last year you may have seen my session on Clustering markets to derive segments for building groupings for the market approach or Cost Approach. This is kind of a spinoff of that with more detail. **Overall goal for this analysis is to:**

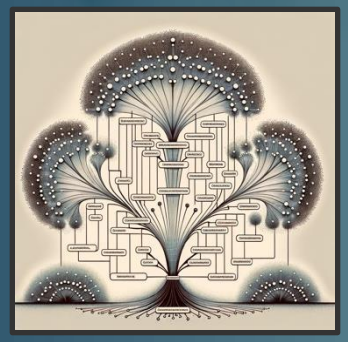
> **"Identify and define meaningful market segments using spatial analysis techniques that improves model accuracy, interpretability, and equity."**

We are going to use real data to analyze this problem and derive a solution. This is a common problem that I see a lot of my clients have and my goal is to provide potential solutions.

Market Segmentation Methods

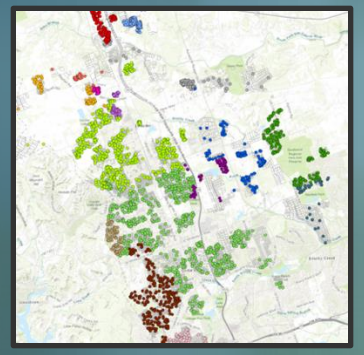
“Eyeballing”

- Easiest and requires less tools or expertise
- Least accurate due to lack of quantitative evidence



Clustering

- Quickly find clusters or groupings
- Most commonly used to find “hot or cold spots”
- Can’t really apply to entire study area

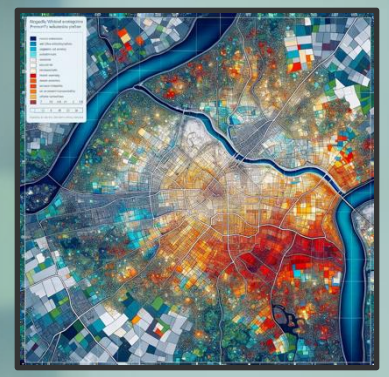


Spatial Modeling (GWR)

- Great for parsing out spatial heterogeneity
- Predictive power
- Steeper learning curve

Location Factor Grouping

- Standardizes the effects of location
- Provides for better visualization
- Can extend analysis with LocF



Interpolation

- More complexity for tuning and understanding
- Provides an estimate for all points in the study area
- Cluster or group all subjects in study area

TYLER CONNECT

tyler technologies

Technology Used for Demo

TYLER CONNECT

tyler technologies

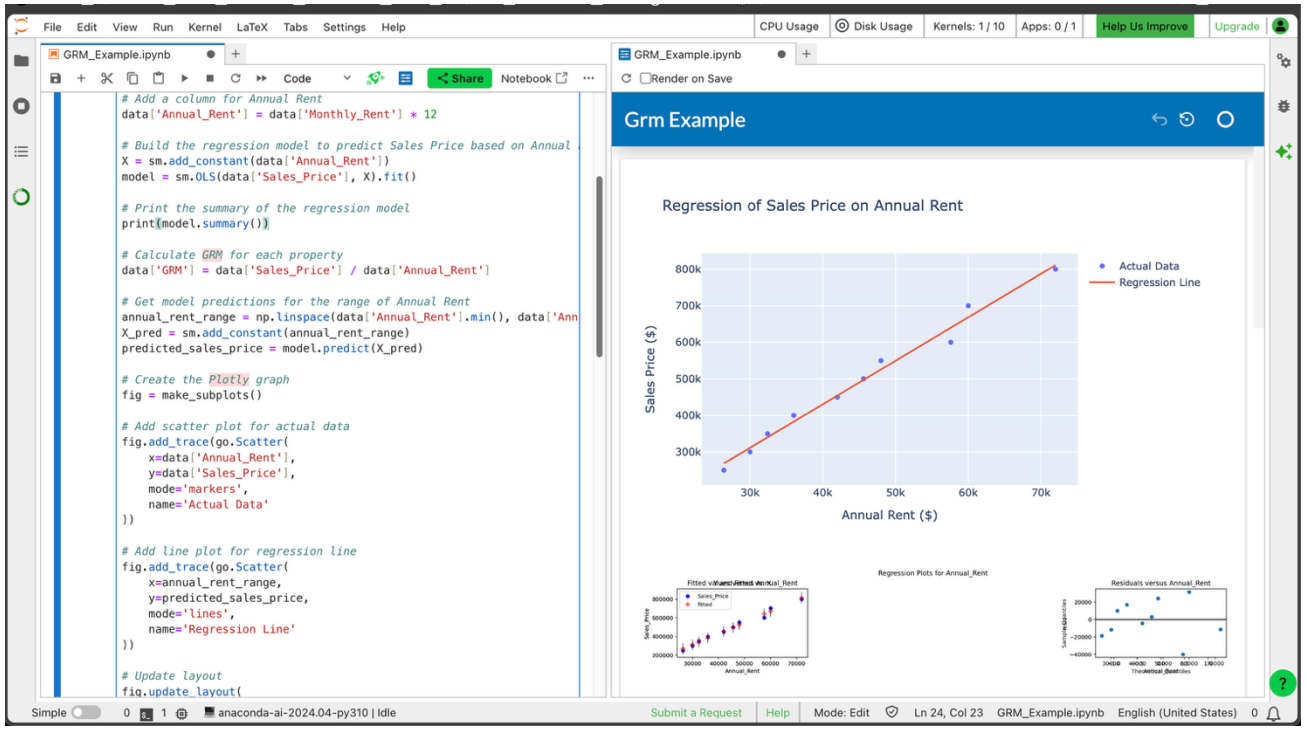
Jupyter Notebook in Positron

Using Jupyter Notebook with Python and Positron IDE (beta),

- One could also use Quarto, Google Collab or other markdown environments.

Not going to show how to set up the environment, but I generally create virtual environments to containerize my analysis.

Also, Anaconda has the cloud-based version of python that works well!

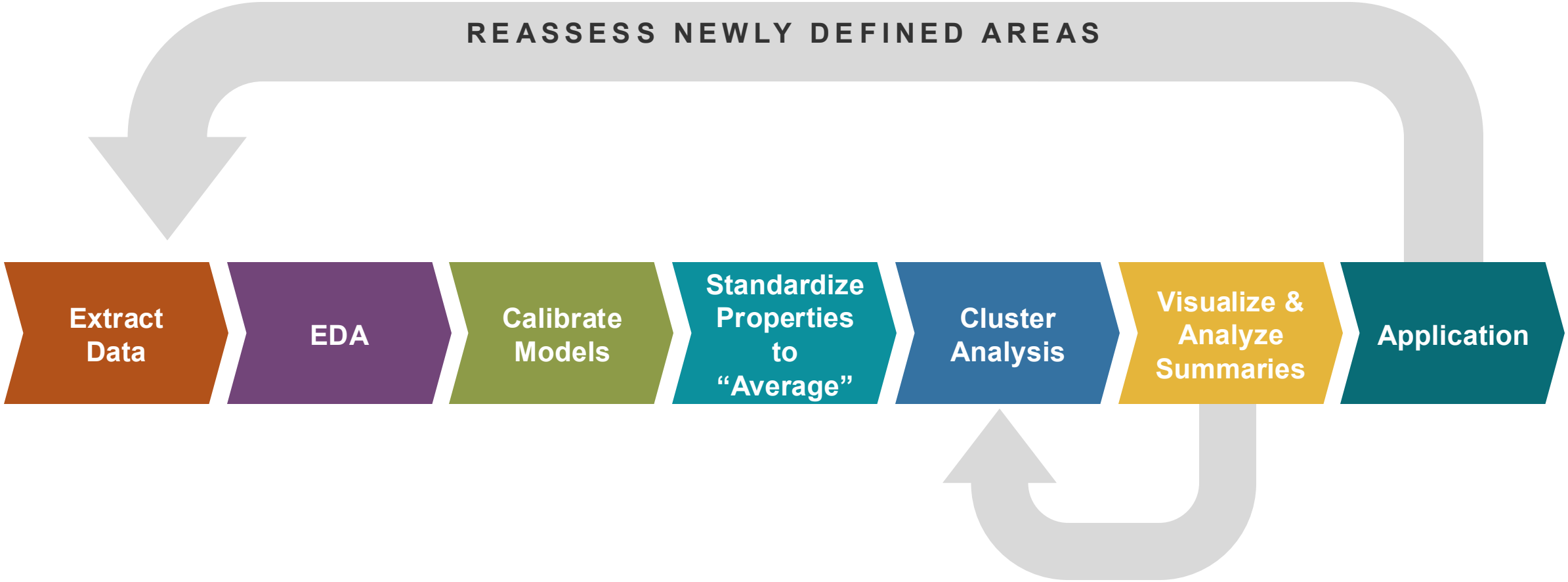


Python Libraries Used

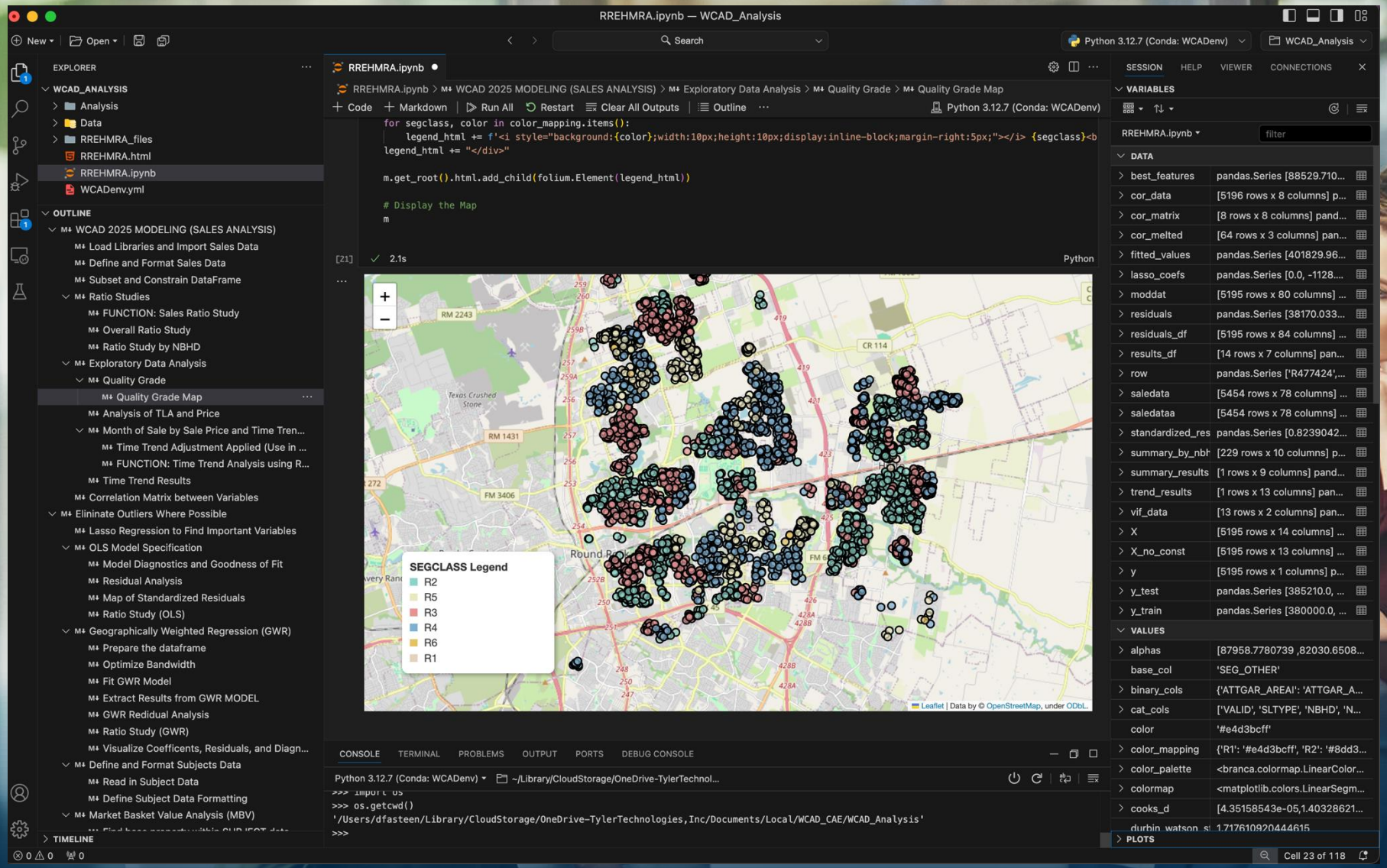
- pandas – Data frames
- numpy – math library
- matplotlib – plots
- statsmodels – statistical modeling
- scipi – advanced modeling
- plotly - interactive visualizations
- leaflet - mapping

- geopandas – spatial data frames
- mgwr – geographically weighted regression
- sklearn – machine learning (clustering)

Analysis Pipeline



Dans GitHub Repo



https://positron.posit.co/start.html

Extract Data and Exploratory Data Analysis

TYLER CONNECT

tyler technologies

Data Extract

- Pull data from the validated sales history pool.
- Need a flattened version of the data (both sales and subjects) such that we can perform analysis.
  - Aggregate Additions and OBY's
  - Add all area by **each** addition code or OBY code.
- Derive a standardized extract either through SQL, Inquire, CRS, etc…
  - Reused and revised as necessary.

Need a Flat File Extract

ATENT

```
-- ADDITIONS AGGREGATE
ADN AS
 (
SELECT SALEKEY,

-- PORCHES, PATIOS, AND DECKS
SUM(CASE WHEN FIRST IN ('11','21') THEN AREA ELSE 0 END) OPNPORCH,
SUM(CASE WHEN FIRST = '41' THEN AREA ELSE 0 END) SCRPORCH,
SUM(CASE WHEN FIRST IN ('12','22') THEN AREA ELSE 0 END) ENCPORCH,
SUM(CASE WHEN FIRST IN ('33','34') THEN AREA ELSE 0 END) PATIO,
SUM(CASE WHEN FIRST IN ('31','40') THEN AREA ELSE 0 END) DECKS,

-- ATTACHED GARAGES AND CARPORTS
SUM(CASE WHEN FIRST IN ('13','23','93') THEN AREA ELSE 0 END) ATTGAR,
SUM(CASE WHEN FIRST = '30' THEN AREA ELSE 0 END) ATTCARPRT,

-- NONCOMPLIENT AND BASMENT AREA
SUM(CASE WHEN FIRST IN ('17','27','97') THEN AREA ELSE 0 END) NONCOMPLNT -- 2nd or 3rd story SF NONCOMPLIANT LIVING AREA (I.E., HALFWALLS)

FROM ADDN_SALE
GROUP BY SALEKEY
 ),

--OBYS AGGREGATE
OBY AS
  (
SELECT SALEKEY,
SUM(CASE WHEN CODE IN ('RG1','RG2') THEN AREA ELSE 0 END)  DETGAR,
SUM(CASE WHEN CODE IN ('RC1') THEN AREA ELSE 0 END)  DETCARPRT,
SUM(CASE WHEN CODE IN ('RS1','RS2','RS3','WS1') THEN AREA ELSE 0 END)  SHEDAREA,
SUM(CASE WHEN CODE IN ('BT2') THEN AREA ELSE 0 END)  BATHHOUSE,
SUM(CASE WHEN CODE IN ('RP1','RP2','RP3','RP4','RP5','RP6') THEN AREA ELSE 0 END) AS POOL,
SUM(CASE WHEN CODE NOT IN('RP1','RP2','RP3','RP4','RP5','RP6','RG1','RG2') THEN ADJRCNLD ELSE 0 END) AS OTHOBYVAL  --OBY COST Value without Pools and Detached Garages

FROM OBY_SALE
WHERE TAXYR = '2024' AND CUR = 'Y'
GROUP BY SALEKEY
  )
```

ATPRO

```
-- ADDITIONS AGGREGATED
WITH SEG AS
(
SELECT propertyid
, SUM(CASE WHEN fsegType = 'G' THEN fArea ELSE 0 END) ATTGAR_AREA
, SUM(CASE WHEN fsegType = 'DG' THEN fArea ELSE 0 END) DETGARAGE_AREA
, SUM(CASE WHEN fsegType = 'OP' THEN fArea ELSE 0 END) OPNPRCH_AREA
, SUM(CASE WHEN fsegType = 'CSP' THEN fArea ELSE 0 END) POOL_AREA
, SUM(CASE WHEN fsegType = 'P' THEN fArea ELSE 0 END) PATIO_AREA
, SUM(CASE WHEN fsegType = 'EG' THEN fArea ELSE 0 END) ENCL_GARAGE
, SUM(CASE WHEN fsegType = 'CP'THEN fArea ELSE 0 END) CARPRT
, SUM(CASE WHEN fsegType = 'U' THEN fArea ELSE 0 END) UTIL
, SUM(CASE WHEN fsegType = 'SI' THEN fArea ELSE 0 END) SITEIMP
, SUM(CASE WHEN fsegType = 'DK' THEN fArea ELSE 0 END) DECK_AREA
, SUM(CASE WHEN fsegType = 'OK1' THEN fArea ELSE 0 END) OUTDRKIT_AREA
, SUM(CASE WHEN fsegType = 'MA' THEN fArea ELSE 0 END) FIRST_AREA
, SUM(CASE WHEN fsegType = 'MA2' THEN fArea ELSE 0 END) SECOND_AREA
, SUM(CASE WHEN fsegType = 'MA3' THEN fArea ELSE 0 END) THIRD_AREA
, SUM(CASE WHEN fsegType = 'MA4' THEN fArea ELSE 0 END) FOURTH_AREA
, SUM(CASE WHEN fsegType = 'FP' THEN 1 ELSE 0 END) FIREPLACE
, SUM(CASE WHEN fsegType IN('EG','WS','GU','GR','GA') THEN vTSGRSEG_RCN ELSE 0 END) SEG_OTHER
FROM dbo.cama_TSGRSeg
WHERE AdHocTaxYear = '2020'
GROUP BY propertyid
),


--IMPROVMENTS 1:1
SIMP AS
(
SELECT PropertyID
, AdHocTaxYear
, ISNULL(fActYear,0) ACTYRBLT
, ISNULL(fEffYear,0) EFFYRBLT
, fSegClass SEGCLASS
, fConstStyle STYLE
--, fIntFinish INTFINISH
--, fExtFinish EXTFINISH
--, fFoundation FOUNDATION
, fRoof ROOFTY
, fHeatAC HVAC
, vTSGRSeg_AdjFactor ADJFACT
, fAreaFactor SEGFACTOR
, fUnitPrice CLASSUNITFACT
, vTSGRSeg_UnitPrice UNITPRICE

FROM dbo.cama_TSGRSeg
WHERE fSegType = 'MA' AND AdHocTaxYear = '2020' AND FirstPage = '1'
)

-- EXTRACT SALE DATA --

SELECT PG.quickrefid QUICKREFID
, PG.PropertyID PARID
, S.SaleDate SLDATE

, S.price PRICE
, S.TransferValidityCode VALID
, PG.PropertyTypeCode PROPTY
, ISNULL(I.fImpType, 'NA') IMPTYPE
```

Data Extract

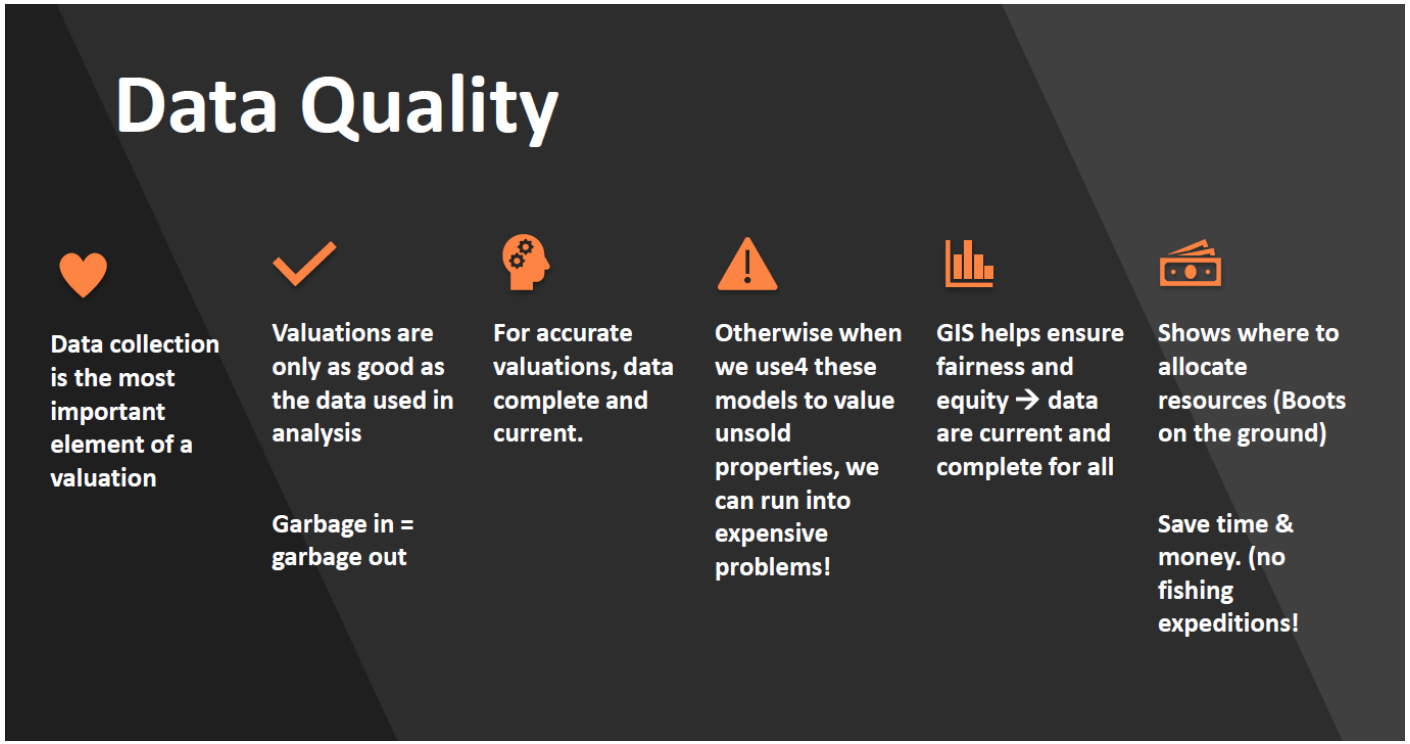
PARID	SLDATE	PRICE	VALID	PROPTY	SLTYPE	IMPTYPE	TOTALVAL	LANDVAL	IMPVAL	RCN	MRAVAL	MKTVAL	SCHTAXUN	NBHD	NGROUP	CLUST	DEFMOD	ACRES	ACTYRBLT	EFFYRBLT	SEGCLASS	SEGFACTO	CLASSUNIT	UNITPRICE	STYLE	ADJFACT	NBHDFACT	TLA
202958	6/17/22	440000.00	VALID	RES	ALML	R	405935.00	77000.00	352427.00	196316.00	407290.00	405935.00	SRR	R435499H	ERRGRP8	RREH-C	RREHMRA	0	2000	2000	R2	0.00	45.90	94.31	NA	1.80	204.00	1859.00
117946	7/7/22	517000.00	VALID	RES	ALML	R	388876.00	68000.00	316011.00	190782.00	391176.00	388876.00	SRR	R462475H	ERRGRP8	RREH-C	RREHMRA	0	1976	1990	R2	0.00	48.00	101.06	NA	1.66	202.00	1557.00
216632	4/15/22	715000.00	VALID	RES	ALML	R	707556.00	151000.00	559507.00	392097.00	729641.00	707556.00	SRR	R566600B	ERRGRP9	RREH-C	RREHMRA	0	2002	2002	R5	0.00	62.20	117.73	NA	1.43	155.00	2955.00
202978	8/10/22	374500.00	VALID	RES	ALML	R	387991.00	77000.00	311186.00	175337.00	386486.00	387991.00	SRR	R435499H	ERRGRP8	RREH-C	RREHMRA	0	1999	1999	R2	0.00	48.00	101.96	NA	1.77	204.00	1510.00
433260	2/17/22	309464.00	VALID	RES	ALML	R	242450.00	57000.00	185450.00	217818.00	386325.00	346541.00	SJA	J208321D	JSFM	EJAR-C	RREHMRA	0	2021	2021	R3	0.00	0.00	99.46	NA	1.29	129.00	1983.00
229751	3/9/22	799500.00	VALID	RES	ALML	R	831873.00	151000.00	693885.00	476243.00	850628.00	831873.00	SRR	R566600B	ERRGRP9	RREH-C	RREHMRA	0	2005	2005	R5	0.00	0.00	110.70	NA	1.46	155.00	3898.00
433300	2/25/22	393000.00	VALID	RES	ALML	R	162061.00	57000.00	105061.00	262720.00	451020.00	396444.00	SJA	J208321D	JSFM	EJAR-C	RREHMRA	0	2021	2021	R3	0.00	0.00	92.15	NA	1.29	129.00	2583.00
433392	7/29/22	352500.00	VALID	RES	ALML	R	271278.00	57000.00	214278.00	215723.00	383423.00	347332.00	SJA	J208321D	JSFM	EJAR-C	RREHMRA	0	2021	2021	R3	0.00	0.00	100.15	NA	1.29	129.00	1939.00
118024	1/26/22	325000.00	VALID	RES	ALML	R	367847.00	68000.00	285635.00	191086.00	369425.00	367847.00	SRR	R462475H	ERRGRP8	RREH-C	RREHMRA	0	1977	1977	R2	0.00	48.00	101.06	NA	1.49	202.00	1575.00
433309	3/17/22	384999.00	VALID	RES	ALML	R	160716.00	57000.00	103716.00	259356.00	446824.00	392260.00	SJA	J208321D	JSFM	EJAR-C	RREHMRA	0	2021	2021	R3	0.00	0.00	92.66	NA	1.29	129.00	2540.00
433306	3/4/22	369999.00	VALID	RES	ALML	R	154367.00	57000.00	9736																			
433283	2/24/22	290590.00	VALID	RES	ALML	R	209852.00	57000.00	15285																			
434557	5/5/22	343390.00	VALID	RES	ALSS	R	145342.00	74000.00	7134																			
117944	7/13/22	530000.00	VALID	RES	ALML	R	439155.00	68000.00	35063																			
117901	10/27/22	330000.00	VALID	RES	ALML	R	461518.00	68000.00	37073																			
202953	2/10/22	405000.00	VALID	RES	ALML	R	429754.00	77000.00	40355																			
117881	5/19/22	520000.00	VALID	RES	ALML	R	501556.00	68000.00	41268																			
433318	1/21/22	386510.00	VALID	RES	ALML	R	230230.00	57000.00	17323																			
433275	2/17/22	307590.00	VALID	RES	ALQ	R	259309.00	57000.00	20230																			
118010	4/6/22	440350.00	VALID	RES	ALML	R	382966.00	68000.00	31080																			
216616	12/16/22	670000.00	VALID	RES	ALML	R	714192.00	151000.00	56767																			
216719	3/4/22	775501.00	VALID	RES	ALML	R	808448.00	151000.00	64406																			
434547	6/30/22	394219.00	VALID	RES	ALQ	R	94000.00	74000.00	2000																			
216697	12/29/22	610000.00	VALID	RES	ALML	R	721018.00	151000.00	58413																			
433328	1/6/22	369999.00	VALID	RES	ALML	R	270370.00	57000.00	21337																			

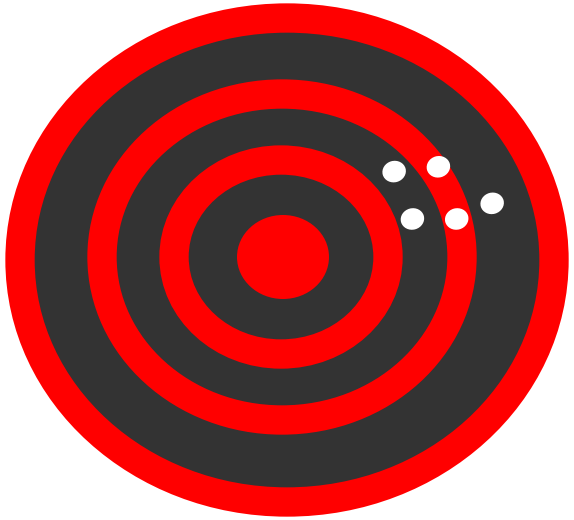
ATTGAR_AREA	DETGARAGE_AREA	OPNPRCH_AREA	POOL_AREA	PATIO_AREA	OUTDRKIT_AREA	FIRST_AREA	SECOND_AREA	THIRD_AREA	FOURTH_AREA	CARPRT_AREA	UTIL_AREA	SITEIMP_AREA	DECK_ARE	SEG_OTHER	COOR
399.00	0.00	49.00	0.00	100.00	0.00	1859.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
441.00	0.00	368.00	0.00	0.00	0.00	1557.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
480.00	0.00	440.00	0.00	0.00	0.00	1838.00	1117.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
368.00	0.00	30.00	0.00	0.00	0.00	1510.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
394.00	0.00	40.00	0.00	0.00	0.00	794.00	1189.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.8
692.00	0.00	123.00	0.00	0.00	0.00	2140.00	1758.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.4
453.00	0.00	163.00	0.00	0.00	0.00	1058.00	1525.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.8
399.00	0.00	60.00	0.00	0.00	0.00	1131.00	808.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.8
437.00	0.00	316.00	0.00	0.00	0.00	1575.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
432.00	0.00	168.00	0.00	0.00	0.00	1052.00	1488.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.8
399.00	0.00	135.00	0.00	0.00	0.00	928.00	1402.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.8
380.00	0.00	40.00	0.00	0.00	0.00	1635.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.8
408.00	0.00	184.00	0.00	0.00	0.00	1644.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.6
420.00	0.00	342.00	383.00	0.00	0.00	1707.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
578.00	0.00	404.00	0.00	0.00	0.00	2112.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
399.00	0.00	40.00	0.00	0.00	0.00	1111.00	1510.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
306.00	0.00	642.00	0.00	0.00	0.00	2042.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
432.00	0.00	122.00	0.00	0.00	0.00	1052.00	1488.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.8
394.00	0.00	40.00	0.00	0.00	0.00	794.00	1189.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.8
552.00	0.00	212.00	0.00	0.00	0.00	1470.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
433.00	0.00	208.00	0.00	82.00	0.00	2048.00	1088.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
542.00	0.00	132.00	0.00	0.00	0.00	2061.00	1594.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.6
702.00	0.00	93.00	0.00	378.00	0.00	1842.00	1283.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.5
475.00	0.00	95.00	0.00	0.00	0.00	852.00	1327.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.8
445.00	0.00	26.00	0.00	0.00	0.00	2020.00	1394.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	30.5
399.00	0.00	135.00	0.00	0.00	0.00	928.00	1402.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	30.8

Model Building is Both an Art and a Science

Modeling is an iterative process – combination of the right variables, variable combinations, and the type of sample that the modeler is trying generalize over the population.

- Good data is essential to extracting good adjustments from the market.
- Garbage in = Garbage Out

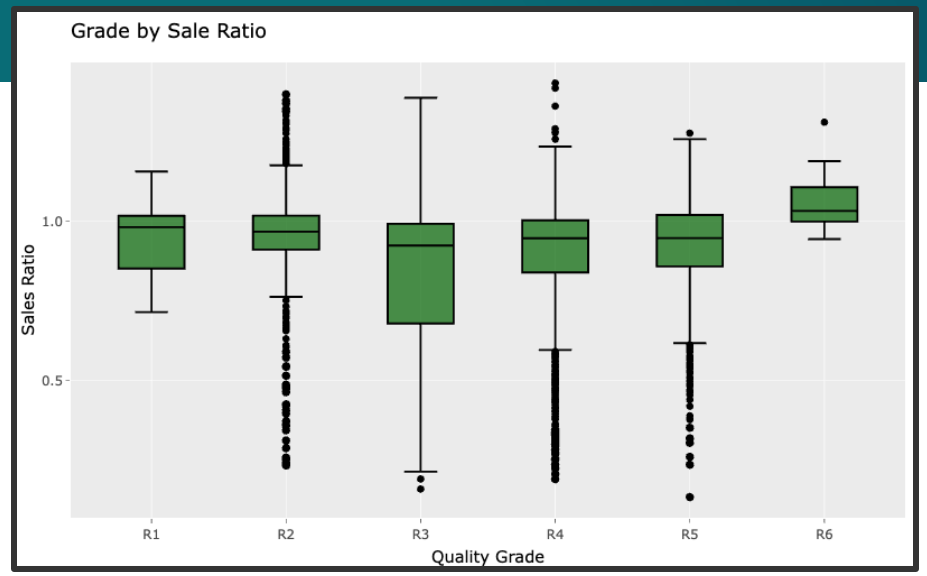


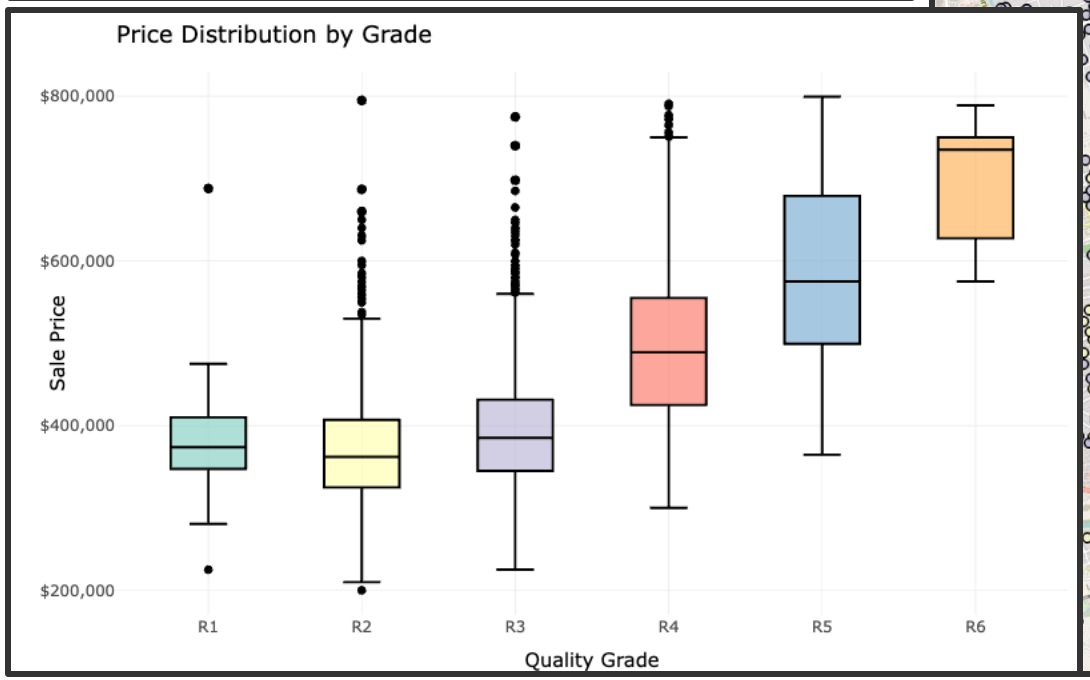


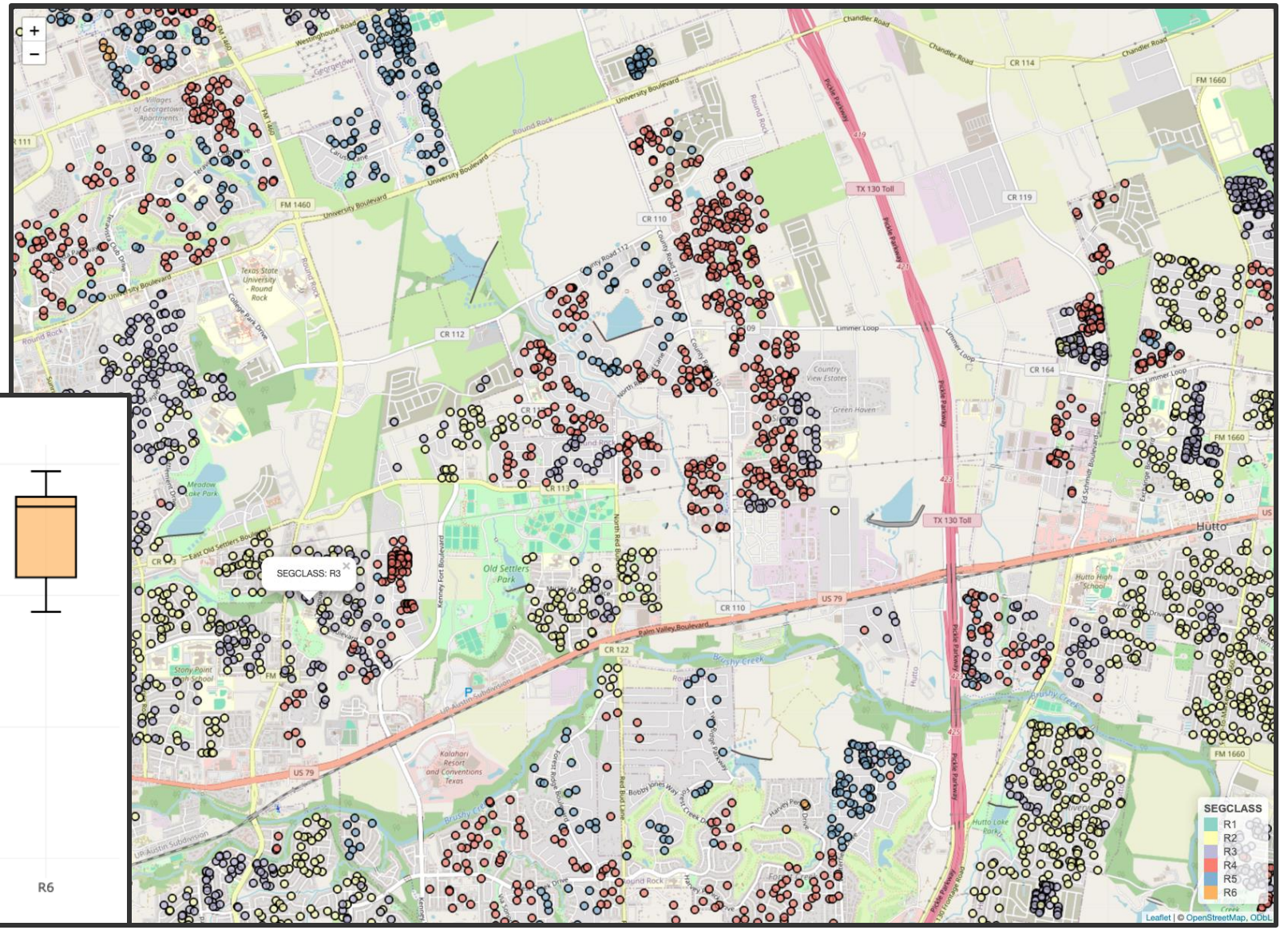
TYLER CONNECT

tyler technologies

Quality Grade Analysis







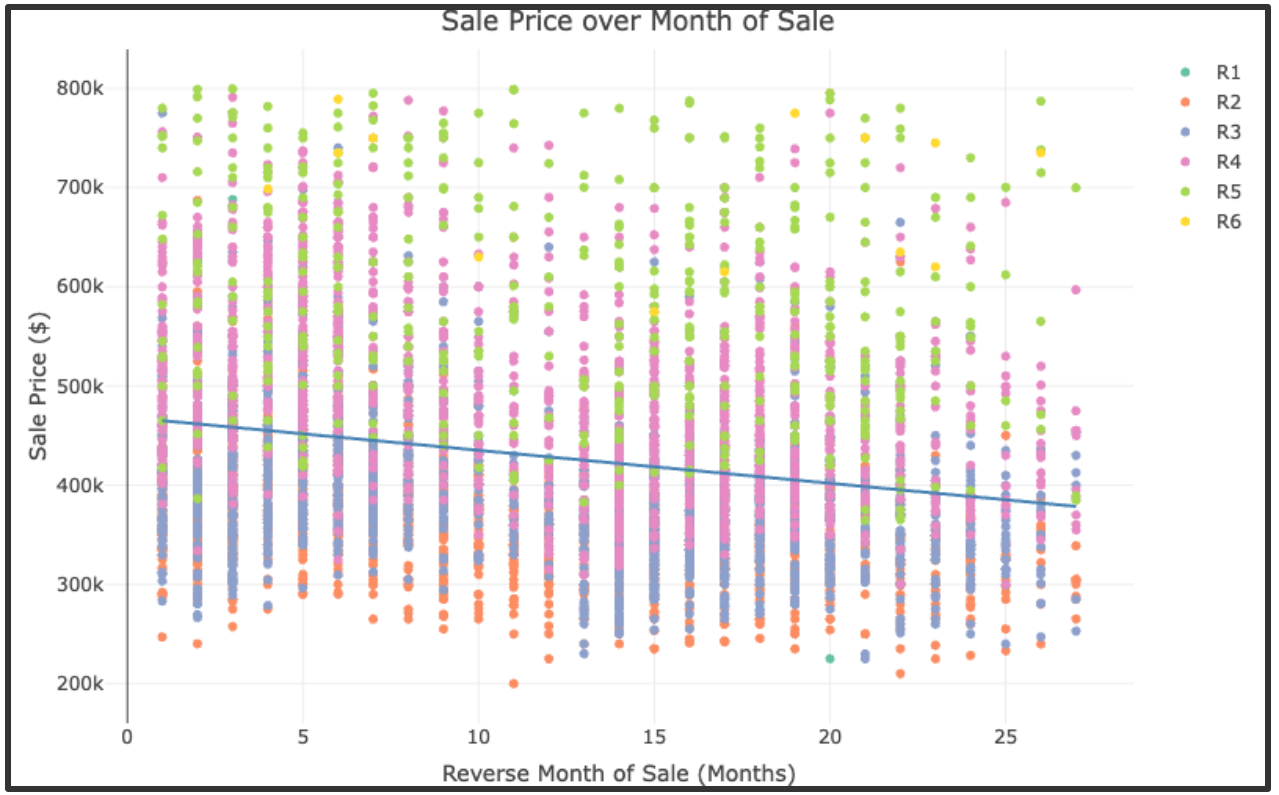
Analysis of Living Area and Price

Time Trend Analysis

Before we do any kind of analysis, make sure to apply a time trend to price to ensure that prices are of the same date (e.g. Jan 1st of a given year)

- Use reverse month of sale (RMOS) to determine if a trend exists
- Use the ratio trend methodology or other adjustment to quantify the trend and apply to sale price

$$TASP = PRICE \times (1 + \text{Monthly Trend})^{rmos}$$



Sale Count	Upper Bound	Lower Bound	Trimmed Sale Count	Median Inverse Ratio	Slope	P-Value	Significance	Monthly Trend	Standard Error	T-Value	Annual Trend Info	Annual Trend
4457	1.225401	0.636071	4457	1.055485	-0.001344	0.0	1.0	-0.001344	0.000232	-5.804373	-0.016007	-0.016007

EDA and Ratio Study Demo

TYLER CONNECT

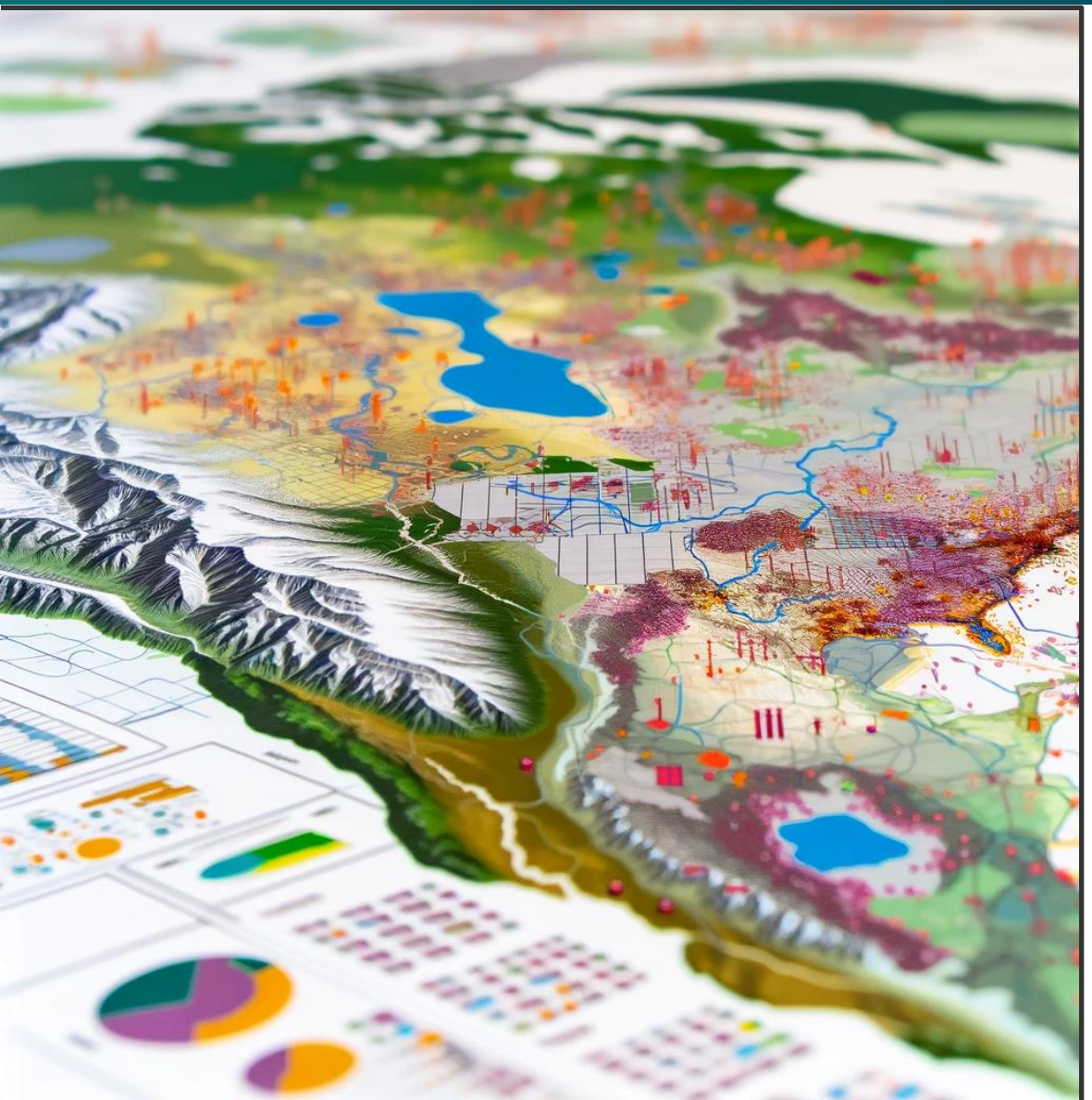
tyler technologies

Model Building

TYLER CONNECT

tyler technologies

Use Spatial Models to parse out Dependence and Heterogeneity



Advantages:

- Improved understanding of spatial relationships.
- Enhanced decision making.
- Predictive power
- Integration of spatial data.

Disadvantages:

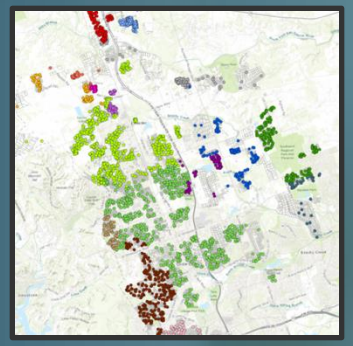
- Complexity and expertise required.
- Data quality and availability
- Computational requirements

Common Spatial Models

Regression with Spatial Variables

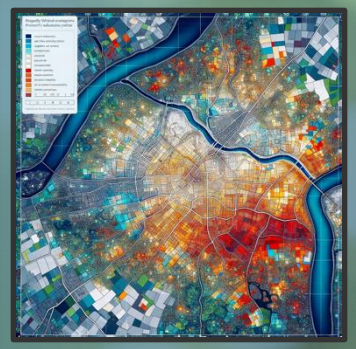
Location indicator variables

Distance/proximity variables



Locally Weighted Regression

Geographically Weighted Regression



Spatial Interpolation Models

Kriging, Co-Kriging, EBK

Response Surface Analysis

Spatial Regression Models

Spatial Lag Model (SLM)

Spatial Error Model (SEM)

TYLER CONNECT

tyler technologies

Geographically Weighted Regression (GWR)

- Utilize the specification of an OLS model but incorporate XY coordinates to analyze locational effects
- Create a Kernel and bandwidth specification to utilize
- Run GWR
- Run Ratio Stats
- Analyze residual graphs

n <int>	MEANRAT <dbl>	MEDRAT <dbl>	COD <dbl>	PRD <dbl>
5188	1.004269	1.000314	6.426022	0.9930016

```
***********************************************************************
*          Results of Geographically Weighted Regression              *
***********************************************************************

*******************Model calibration information*******************
Kernel function: exponential
Adaptive bandwidth: 110 (number of nearest neighbours)
Regression points: the same locations as observations are used.
Distance metric: Great Circle distance metric is used.

***************Summary of GWR coefficient estimates:*****************
                         Min.       1st Qu.       Median       3rd Qu.       Max.
Intercept      -261098.06699 -127111.29133  -89259.67979  -39601.09498 273811.5932
GRDFACT         -49753.71983  111934.77336  161364.95249  196185.22600 275541.2147
LANDVAL             -0.86823       0.68046       0.86344       1.20220      2.6353
TLA                 59.87628      81.85310      95.25893     105.92194    129.7210
RMOS              1329.10823    3739.63232    4370.59682    4869.89179   6169.3632
EFFAGE           -3449.95582   -1278.16493    -317.30015     -32.30707   1696.5209
POOL_AREA           -61.41306     103.05238     154.96474     189.85494    325.9196
ATTGAR_AREA          -8.92250      75.52996     103.86284     126.77925    269.4772
DETGARAGE_AREA      -24.51916      77.61641     106.43430     134.73900    454.0555
TOTPORCH            -32.55936      20.95587      38.59254      50.24101    137.0074
STHT2           -35612.64751  -20114.63130  -15807.41437  -10611.58246   6252.4686
AGEMAX60         -2539.52288     105.13430     645.63239    1382.36703   3638.6577
SEG_OTHER           -1.81875       1.29164       1.99579       2.82623     13.8492
CARPRT_AREA       -121.82492      51.34540     124.44693     302.46107    657.1691
**********************Diagnostic information**********************
Number of data points: 5188
Effective number of parameters (2trace(S) - trace(S'S)): 481.0902
Effective degrees of freedom (n-2trace(S) + trace(S'S)): 4706.91
AICc (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): 124537.6
AIC (GWR book, Fotheringham, et al. 2002,GWR p. 96, eq. 4.22): 124176.4
BIC (GWR book, Fotheringham, et al. 2002,GWR p. 61, eq. 2.34): 121386.4
Residual sum of squares: 7094553153047
R-square value:  0.8808935
Adjusted R-square value:  0.8687171

***********************************************************************
```

Modeling Demo

TYLER CONNECT

tyler technologies

Build a Market Basket Standard

TYLER CONNECT

tyler technologies

What is a Market Basket?

A market basket is an econometric concept that provides a baseline for comparison.

Compares the same property across all properties in the study area.

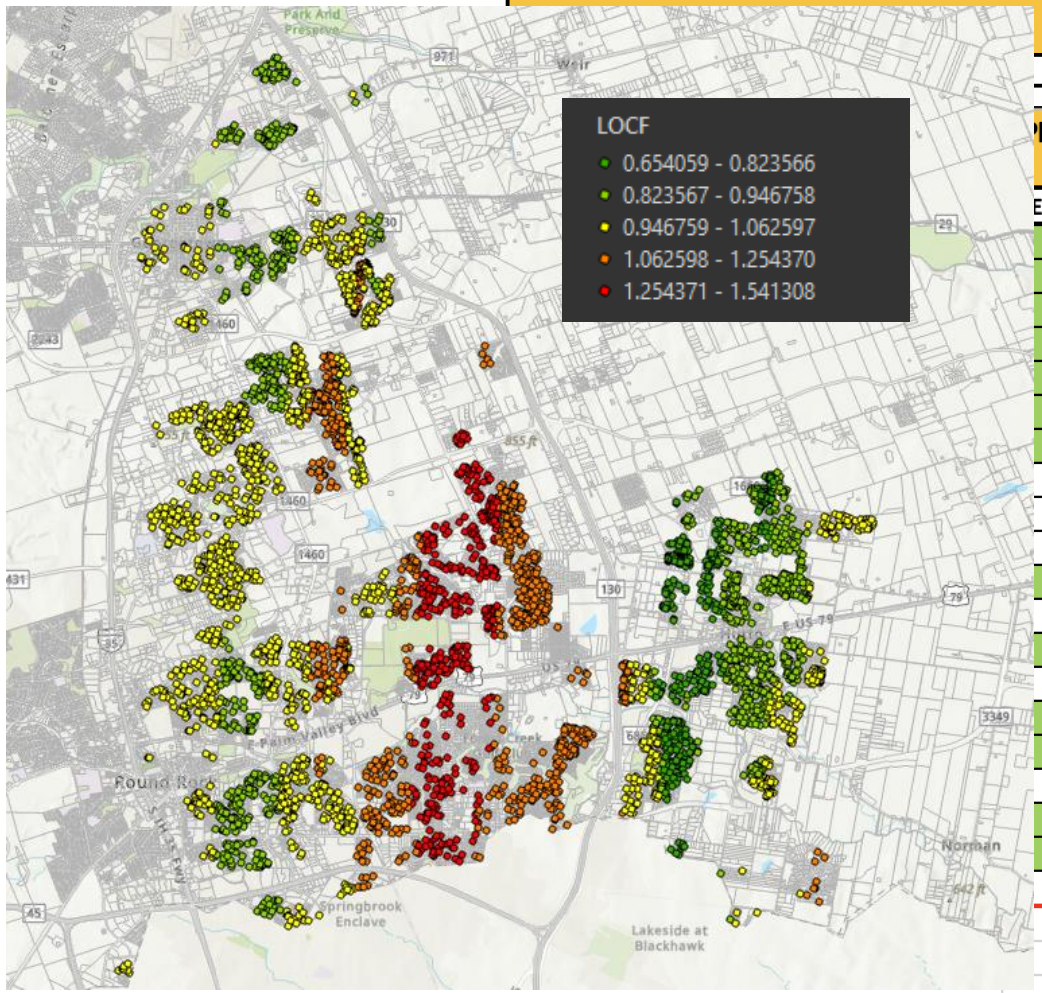
- Finds properties that are inferior or superior to the ”average” property and quantifies it.
- Evaluate GWR coefficients at the mean or median value across the study area.
- MBV Represents the “net effects of the spatially varying parameters”

Advantages of building a market basket

- Reduces reliance on single sale comparisons
- Enhances Modeling Location Adjustments
- Increases Equity and Consistency
- Adapts to market changes
- Bridges the gap between traditional valuation methods and modern data science

tyler technologies

Build a Market Basket and LOCF using GWR Coefficients



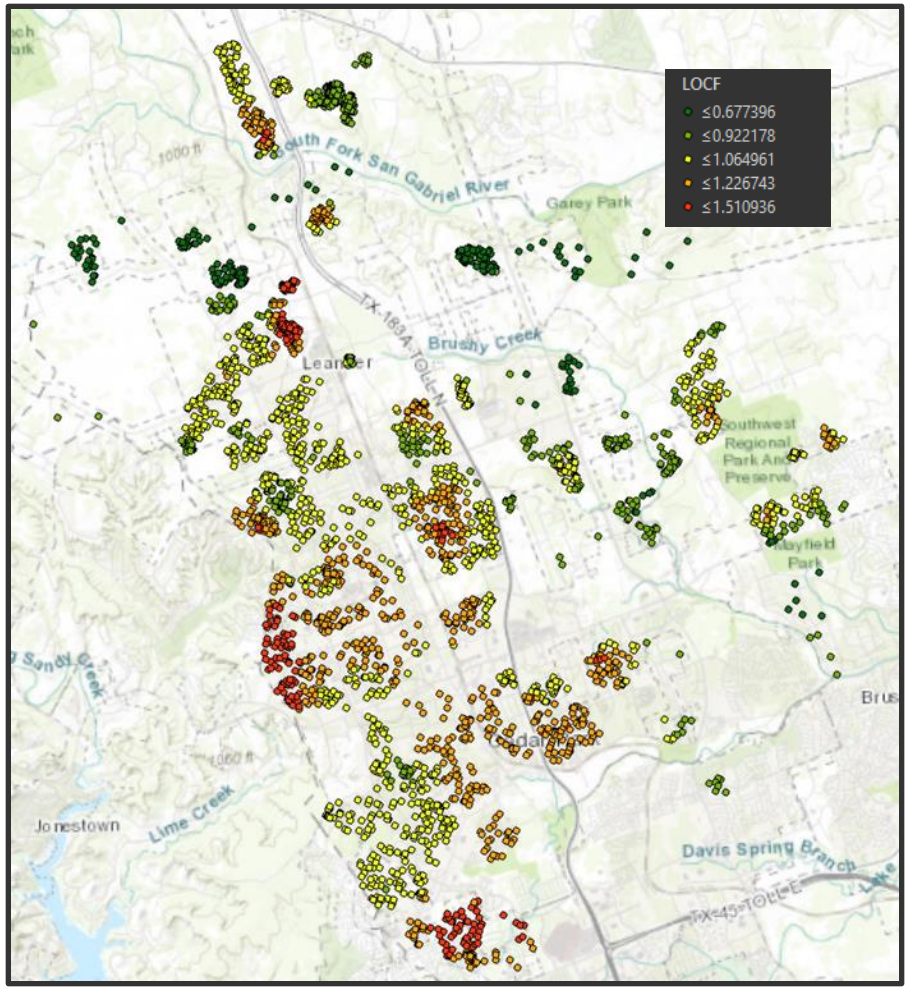
AVERAGE PROPERY IN STUDY AREA			
	Median Sale	Mean Population	Median Population
1,142	1,142	9,835	9,835

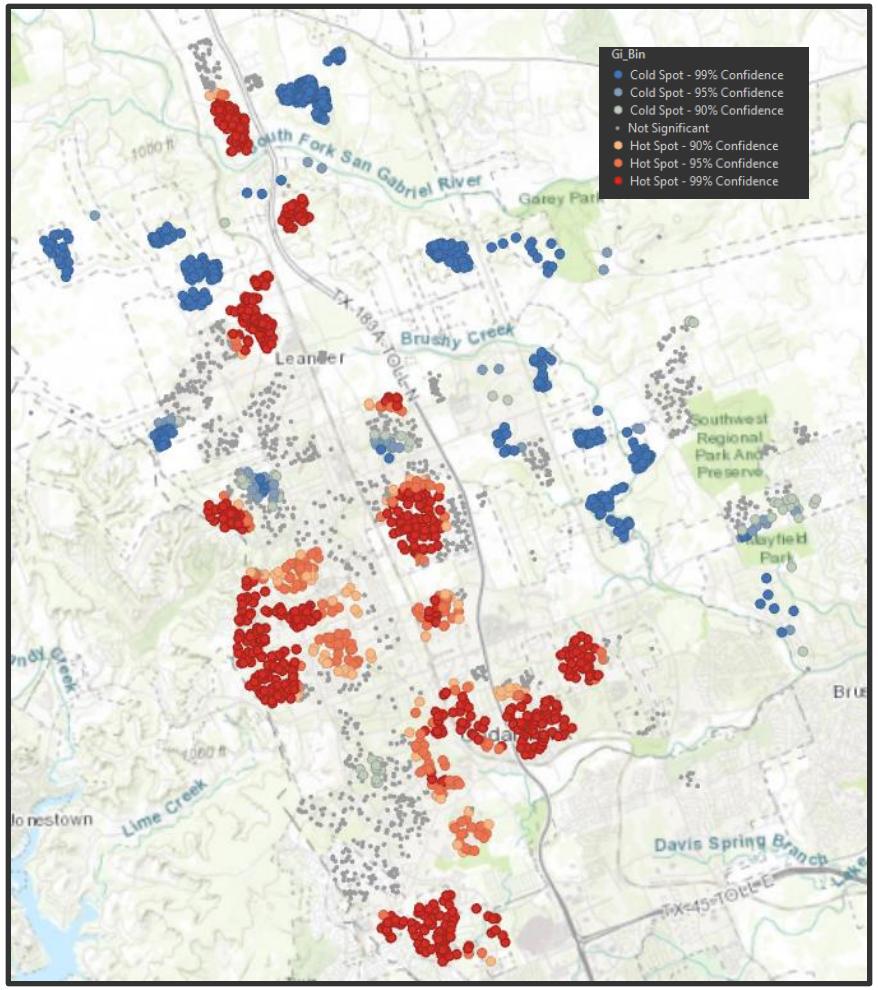
PROPERTY		EXAMPLE SALE		
		SALEKEY	447809	
EAN VALUE	COL	VARIABLE	COEFFICIENT	MEANVAL *COEFF
1	C	INTERCEPT	409256.4536	409,256.45
100000	D	LANDVAL	0.32969552	32,969.55
1.56	E	GRDFACT	28.71730713	44.80
19.447	F	RMOS	-1.054164116	(20.50)
70	G	ACTAGE	1.212372792	84.87
0.0859	H	WALLFACT	0.76625717	0.07
1.77	I	CDUFACT	14.80868096	26.21
0	J	POOLSF	68.59099835	-
0	K	TOTPORCHSF	57.37219988	-
0	L	OTHOBYVAL	8.707475466	-
2.3	M	BATHS	13590.25664	31,257.59
0	N	MSCTOTAL	5.792193832	-
403	O	FINBSMTAREA	78.45224606	31,616.26
0	P	HEATNOAC	-40.53201289	-
1.385	Q	STRYHT	-15241.10312	(21,108.93)
3.48	R	BATHFIXADDN	13641.54134	47,472.56
0	S	DETGARSF	27.27935513	-
16.94	T	DS6	0.962643199	16.31
18.79	U	DS12	0.298317709	5.61
0	V	SRA	-52578.85352	-
			MBVSUM	$531,621
			MBVAVG	$408,672
			LOCF	1.301

TYLER CONNECT

tyler technologies

Analyze LocF Across the Study Area





Analysis of Location Factors (LocF)

Use LocF as an indicator itself for grouping (should show adequate spatial patterns)

Can use cluster algorithms per method #2

Use classification algorithms

- Spatially constrained multivariate clustering - SKATER
- Forest-based model for classification
- Gradient-boosted model for classification

Advantages

- Standardizes the effects of location based on an average property
- Easy for comparison and to visually see patterns or run confirmatory analysis

Disadvantages

- Need to specify a proper OLS and GWR model
- Requires basic math skills to calculate MBV and derive LOCF
- Extending analysis requires a more advanced skillset

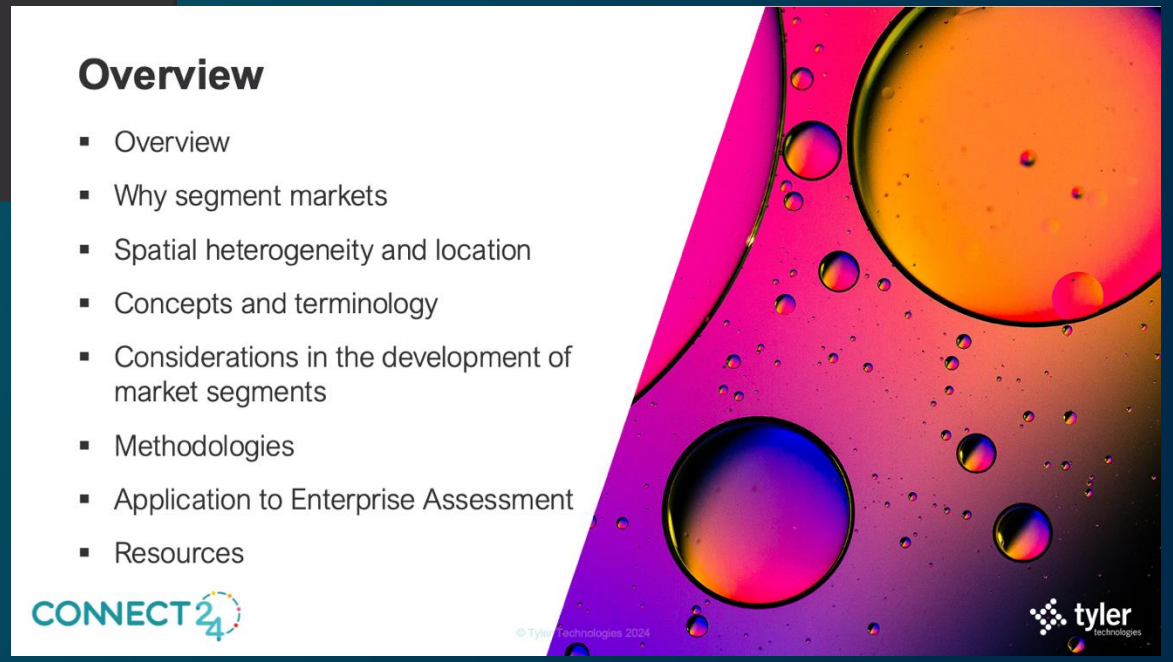
Cluster Analysis

TYLER CONNECT

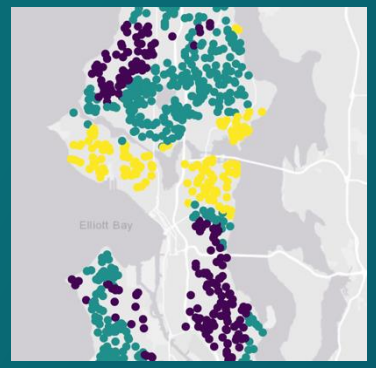
tyler technologies

CONNECT 24

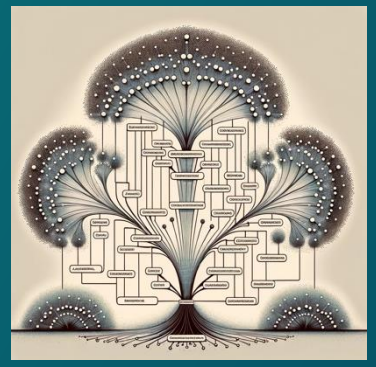
The Truth Behind Valuation Neighborhood Boundaries



Types of Clustering

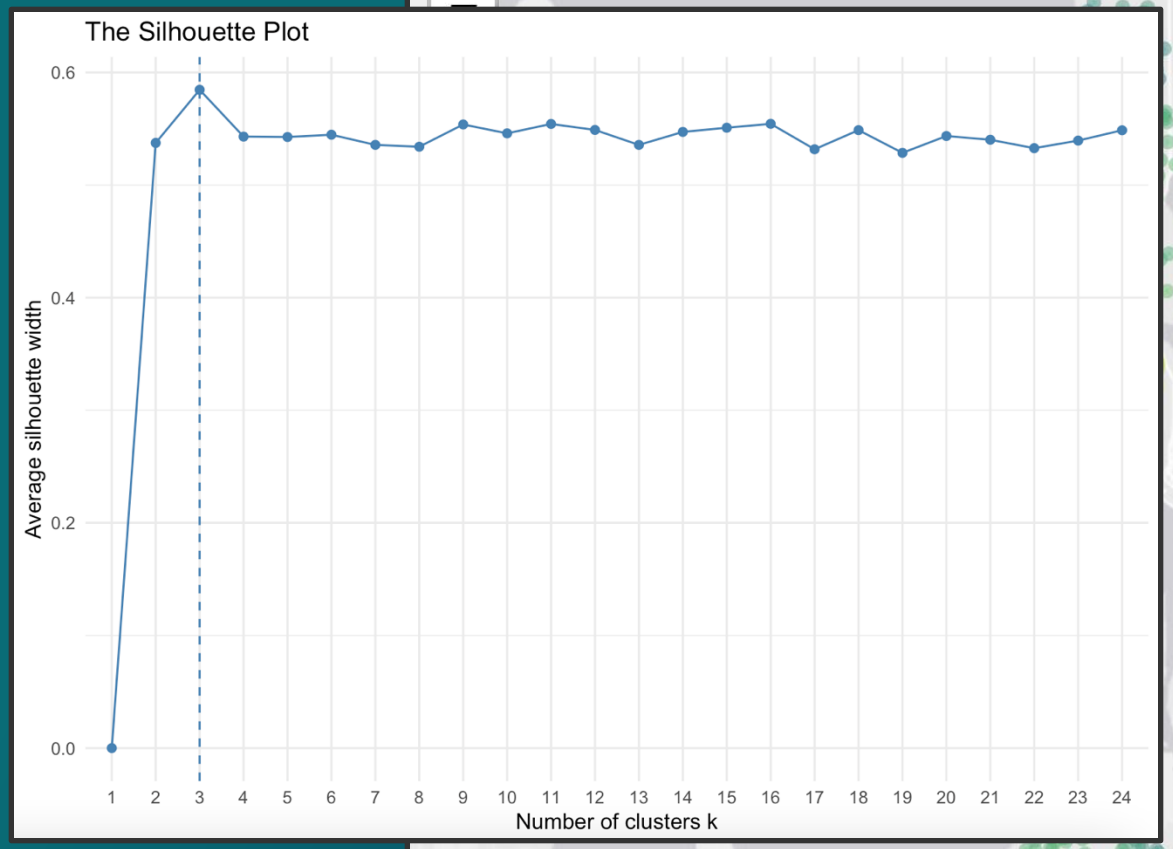


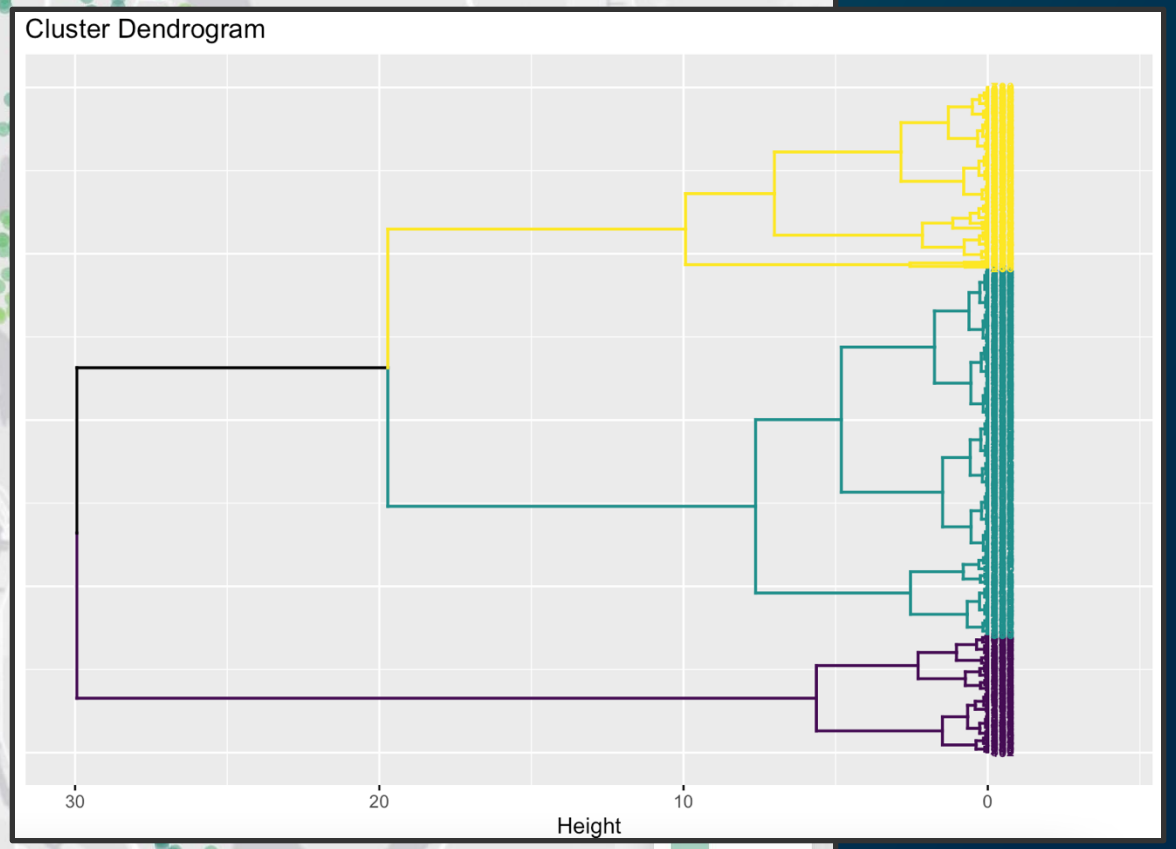
K-Means Clustering – Partitions the data into K distinct clusters based on distance to the centroid of a cluster. Assigns phenomena to clusters based on the nearest mean distance from a cluster's centroid.

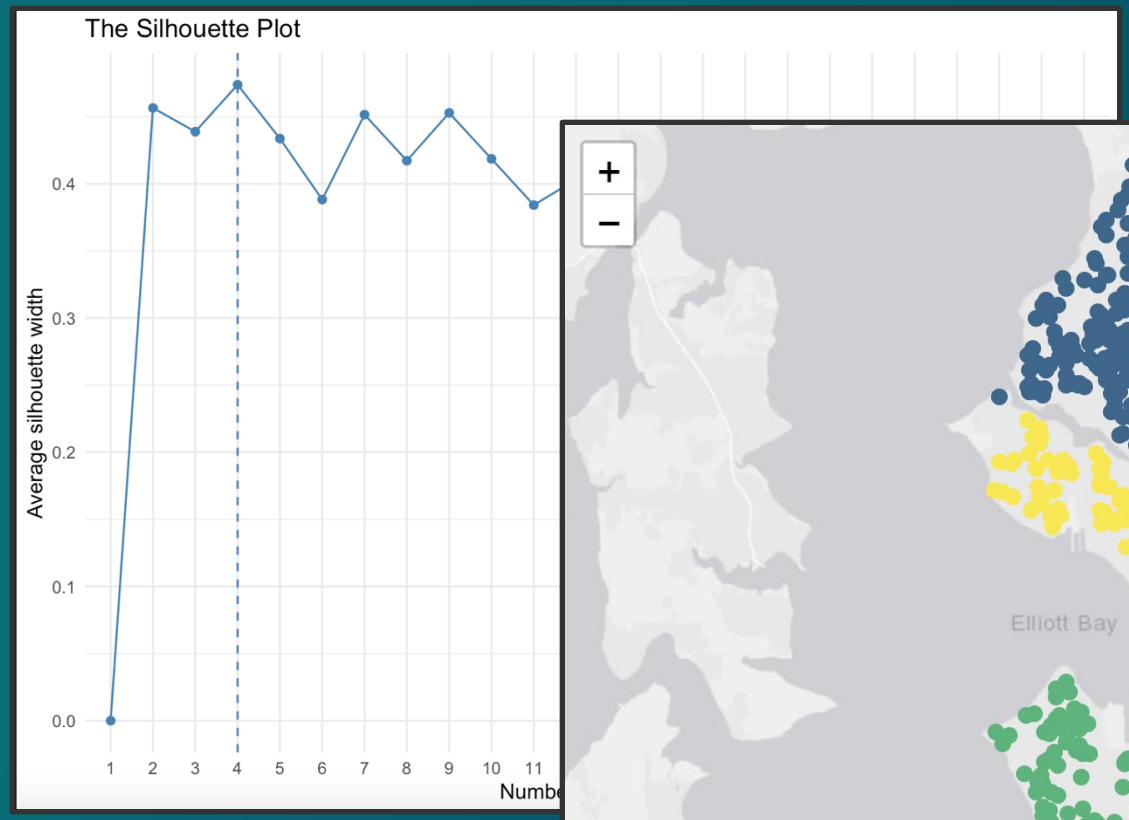


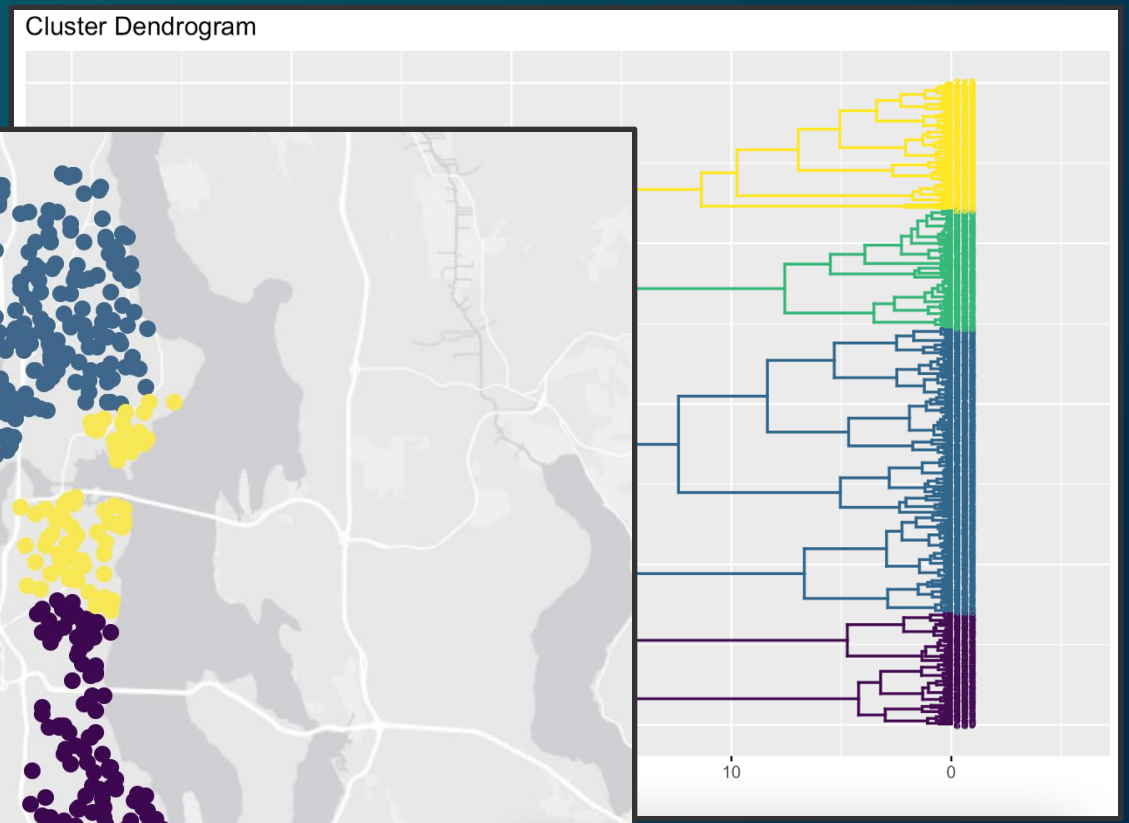
Hierarchical Clustering – Builds a hierarchy of clusters through:

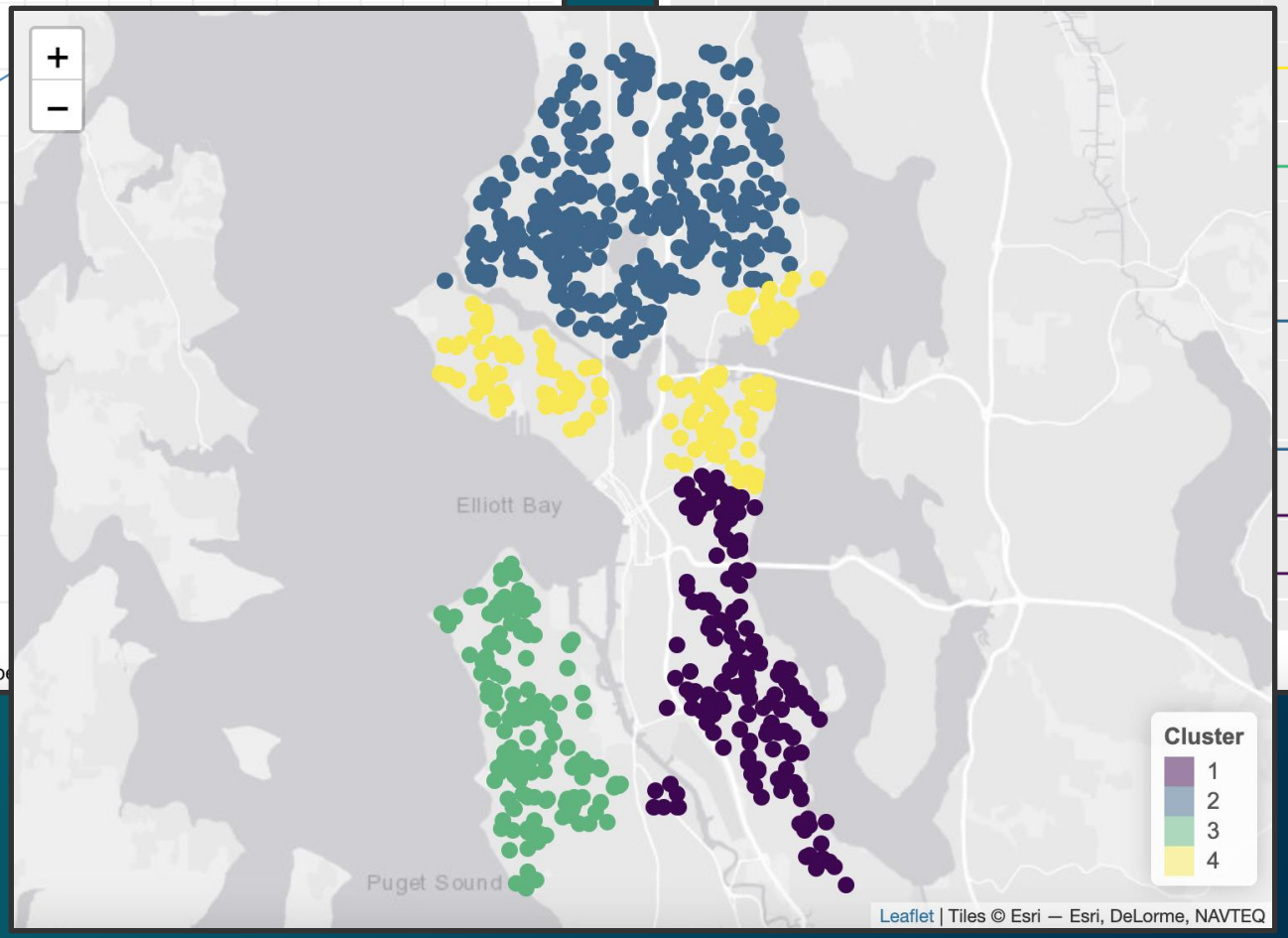
- **Agglomerative** – bottom-up approach where each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy
- **Divisive** – top-down approach where all observations start in one cluster, and splits are performed recursively as one moves down the hierarchy











Clustering Demo

TYLER CONNECT

tyler technologies

Build Summaries

TYLER CONNECT

tyler technologies

Maximize Between Groups and Minimize Within Groups

Counts		Land Value					SEGCLASS				Class Unit Price				ASMT Value				ASMT Per SqFt			MASMT				MASMT PerSqFt		
CLUSTER	n	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	Min	Max	Mean	Median	Min	Max	Mean	Median	Mean	Median	SD	Min	Max	Mean	Median	Mean	Median	SD
1	205	$52,800	$85,850	$71,494	$70,850	$3,494	3	5	4	4	$70	$99	$87	$88	$220,354	$436,142	$278,711	$271,959	$133	$132	$15	$216,101	$452,011	$276,862	$261,603	$133	$131	$17
2	34	$30,000	$70,850	$67,111	$70,850	$8,424	2	5	5	5	$75	$109	$95	$96	$231,482	$612,741	$321,119	$317,757	$140	$138	$19	$234,379	$529,767	$329,563	$327,712	$144	$143	$19
3	746	$39,285	$243,620	$62,707	$56,100	$18,532	1	6	3	3	$51	$109	$80	$80	$179,391	$722,939	$300,220	$265,329	$138	$136	$24	$179,076	$747,079	$297,524	$261,938	$137	$134	$25
4	56	$25,508	$101,996	$72,079	$59,950	$24,897	3	6	4	4	$74	$102	$89	$91	$199,995	$704,820	$391,927	$306,786	$145	$149	$25	$201,380	$659,059	$382,071	$307,598	$142	$143	$21
5	105	$49,050	$59,950	$55,217	$57,343	$4,097	3	4	3	3	$64	$98	$82	$82	$204,555	$409,634	$264,167	$256,452	$144	$142	$23	$210,736	$365,601	$259,483	$254,438	$142	$141	$25
6	72	$57,334	$73,800	$59,191	$57,900	$3,713	2	6	4	4	$58	$103	$83	$84	$195,777	$507,161	$308,293	$292,417	$133	$132	$19	$211,048	$501,070	$303,193	$296,776	$132	$129	$22
7	89	$52,600	$92,050	$63,147	$60,931	$6,223	2	4	3	3	$56	$90	$77	$77	$206,681	$573,088	$345,228	$326,441	$154	$149	$30	$214,103	$531,370	$337,803	$317,930	$150	$148	$27
8	81	$35,000	$180,069	$86,644	$75,800	$28,957	2	7	5	5	$68	$122	$90	$90	$227,524	$859,511	$444,446	$424,944	$153	$141	$36	$222,836	$807,497	$430,798	$419,207	$148	$144	$27
9	119	$46,946	$225,949	$81,407	$84,039	$18,180	2	6	5	5	$75	$105	$90	$91	$261,597	$699,942	$429,332	$424,234	$140	$136	$24	$259,879	$682,067	$425,734	$421,028	$139	$134	$21
10	155	$52,800	$53,000	$52,839	$52,800	$79	3	3	3	3	$65	$92	$80	$79	$203,157	$330,521	$240,536	$236,860	$126	$127	$17	$203,839	$302,002	$235,013	$230,803	$124	$125	$19
11	172	$56,211	$85,892	$73,848	$73,030	$9,113	3	5	4	4	$66	$98	$81	$81	$246,856	$510,622	$355,647	$346,473	$129	$126	$12	$257,424	$543,892	$353,341	$347,093	$128	$125	$15
12	66	$51,300	$75,313	$63,058	$65,400	$6,812	2	4	4	4	$63	$92	$81	$81	$211,901	$522,753	$308,342	$312,301	$135	$134	$19	$209,847	$481,039	$301,990	$305,766	$133	$132	$17
13	624	$17,136	$154,791	$58,318	$54,391	$12,483	1	6	3	3	$57	$102	$81	$82	$163,224	$646,538	$284,159	$260,814	$133	$133	$21	$149,223	$633,993	$280,619	$262,667	$132	$132	$22
14	264	$51,400	$107,400	$64,398	$60,931	$10,316	2	6	3	3	$53	$101	$78	$78	$213,181	$698,361	$329,949	$312,888	$140	$138	$20	$204,000	$676,146	$325,161	$308,947	$138	$136	$20
15	33	$56,159	$118,961	$71,241	$62,457	$16,274	3	5	4	3	$61	$96	$77	$77	$258,784	$770,887	$376,213	$324,115	$143	$126	$46	$265,538	$736,694	$355,385	$332,474	$134	$125	$31
16	36	$59,950	$59,950	$59,950	$59,950	$0	4	4	4	4	$78	$91	$87	$88	$236,843	$330,661	$267,868	$259,443	$129	$133	$12	$222,028	$385,708	$266,883	$253,773	$128	$132	$11

Counts		Sale Price				Sale Price per SqFt			Age of Property				Total Living Area (TLA)				Asmt Ratio Statistics						MASMT Ratio Statistics						Residual Values		
CLUSTER	n	Min	Max	Mean	Median	Mean	Median	SD	Min	Max	Mean	Median	Min	Max	Mean	Median	MEANRAT	MEDRAT_CI_LB	MEDRAT	MEDRAT_CI_UB	COD	PRD	MEANRAT	MEDRAT_CI_LB	MEDRAT	MEDRAT_CI_UB	COD	PRD	MEAN_RESID	MED_RESID	MEANPCTCHG
1	205	$205,231	$467,350	$277,879	$265,52	$133	$132	$18	1	4	1	1	1,316	4,522	2,145	2,082	1.01	1.00	1.01	1.02	4.88	1.00	1.00	1.00	1.01	1.01	3.54	1.00	$1,016	-$615	0.1%
2	34	$227,000	$600,000	$328,973	$321,450	$143	$140	$23	1	42	4	2	1,352	3,942	2,348	2,309	0.98	0.99	1.01	1.03	4.69	1.01	1.01	1.00	1.03	1.05	5.41	1.01	-$590	-$6,695	-1.1%
3	746	$182,500	$775,000	$296,735	$261,75	$137	$134	$26	1	69	18	17	1,052	5,086	2,243	2,104	1.01	1.01	1.01	1.01	4.04	1.00	1.01	1.00	1.00	1.01	4.65	1.00	-$789	-$937	-0.5%
4	56	$203,000	$691,270	$383,610	$309,995	$143	$142	$25	1	15	5	3	1,319	4,156	2,649	2,662	1.02	1.00	1.01	1.03	4.48	1.00	1.00	0.99	1.00	1.02	4.15	1.01	$1,539	-$1,502	-0.4%
5	105	$209,938	$359,117	$259,950	$252,99	$142	$141	$26	1	5	2	2	1,024	3,748	1,921	1,747	1.02	1.00	1.01	1.02	3.66	1.00	1.00	0.99	1.00	1.00	2.42	1.00	$466	$1,179	0.1%
6	72	$198,500	$560,000	$304,345	$289,200	$132	$133	$22	3	33	14	12	1,158	3,912	2,381	2,502	1.01	1.00	1.01	1.02	4.72	1.00	1.00	0.99	1.00	1.01	4.12	1.00	$1,152	$893	0.0%
7	89	$195,000	$560,000	$337,992	$320,000	$151	$148	$32	1	58	23	22	1,131	4,235	2,339	2,096	1.02	1.00	1.01	1.03	4.43	1.00	1.01	0.98	1.00	1.02	5.80	1.01	$189	-$4	-0.5%
8	81	$228,000	$777,106	$438,753	$414,902	$152	$143	$39	1	44	4	2	1,087	4,934	2,920	2,893	1.01	1.00	1.01	1.02	4.45	1.00	0.99	0.99	1.00	1.01	5.21	1.00	$7,955	$375	1.3%
9	119	$260,000	$745,000	$428,680	$415,000	$140	$135	$25	1	36	5	3	1,613	5,180	3,129	3,170	1.01	1.00	1.01	1.02	3.99	1.00	1.00	0.99	1.00	1.01	4.83	1.01	$2,947	$143	0.2%
10	155	$200,000	$302,000	$235,603	$232,000	$124	$125	$20	1	14	6	6	1,275	3,698	1,955	1,864	1.02	1.00	1.01	1.02	3.89	1.00	1.00	1.00	1.00	1.01	2.70	1.00	$590	-$54	0.1%
11	172	$253,800	$570,223	$353,047	$347,000	$128	$125	$16	1	30	9	6	1,684	4,672	2,804	2,724	1.01	1.00	1.01	1.02	4.71	1.00	1.00	1.00	1.00	1.01	4.04	1.00	-$293	-$1,976	-0.4%
12	66	$208,840	$515,000	$302,315	$301,963	$133	$132	$18	2	35	8	5	1,335	3,999	2,325	2,195	1.02	1.01	1.02	1.03	3.52	1.00	1.00	1.00	1.01	1.01	3.07	1.00	$325	-$1,477	-0.2%
13	624	$145,000	$640,000	$280,923	$257,250	$132	$132	$23	1	69	11	7	822	4,972	2,196	2,126	1.01	1.01	1.01	1.01	4.15	1.00	1.00	1.00	1.00	1.00	4.03	1.00	$304	$284	-0.3%
14	264	$210,000	$752,545	$326,476	$312,150	$138	$137	$22	1	44	21	22	1,221	4,744	2,423	2,363	1.01	1.00	1.01	1.02	4.52	1.00	1.00	0.99	1.00	1.01	5.02	1.00	$1,315	-$699	0.0%
15	33	$265,000	$790,000	$372,969	$329,642	$142	$128	$49	1	45	25	21	1,794	4,458	2,733	2,556	1.02	0.99	1.01	1.02	4.71	1.01	0.98	0.95	0.99	1.02	7.37	1.02	$17,583	$5,591	2.5%
16	36	$225,000	$412,248	$264,919	$249,990	$127	$130	$13	1	6	2	2	1,670	3,173	2,107	1,951	1.02	1.02	1.03	1.04	4.27	1.01	1.01	0.99	1.01	1.02	3.28	1.00	-$1,964	-$1,418	-0.9%

TYLER CONNECT

tyler technologies

Ratio Study Iterations on New Clusters

Look at ratio study statistics overall and over each grouping.
Visualize and group by NBHD, Grade, AGE, STHT, Wall Codes, etc…

n	ASMT_MEANRAT	ASMT_MEDRAT	ASMT_COD	ASMT_PRD	PRB	PRB_CI_Lower	PRB_CI_Upper	Gini_Coefficient
5195	88.96%	94.7%	14.348062	1.002240	-0.048202	-0.071670	-0.024733	0.137678

Initial Iteration

	n	ASMT_MEANRAT	ASMT_MEDRAT	ASMT_COD	ASMT_PRD	PRB	PRB_CI_Lower	PRB_CI_Upper	Gini_Coefficient
0	5195	101.0%	100.3%	8.292442	1.009989	-0.167745	-0.179031	-0.156459	0.147157

Final Iteration

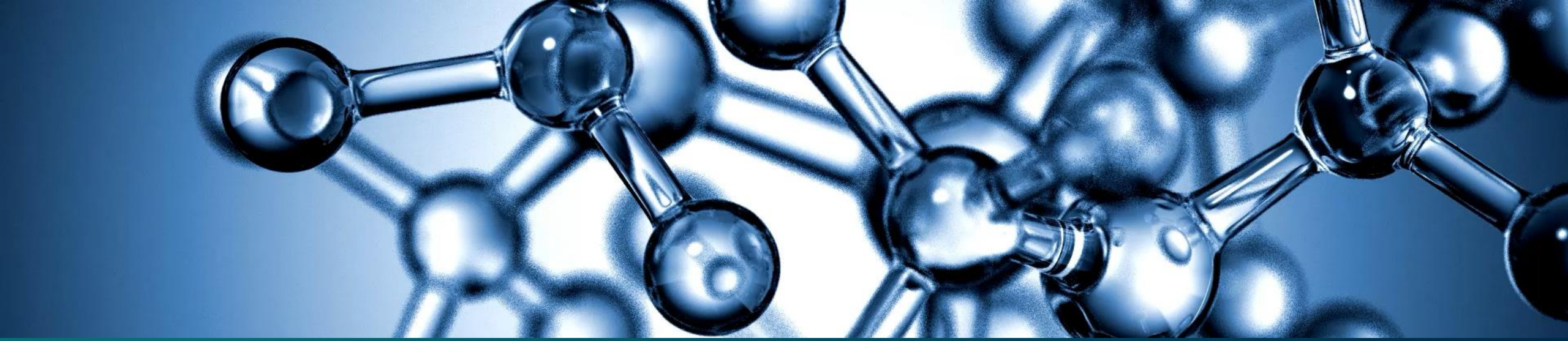
Measure	IAAO Standards	Description
Coefficient of Dispersion (COD)	Between 5 and 15	Measures spread around the median
Price Related Differential (PRD)	0.98 – 1.03	Compares High-value vs. low-value property treatments
Price Related Bias (PRB)	-0.05 and 0.05	Regression based tes of bias across values. + = regressivity, - = progressivity
Gini Coefficient	None, but lower is better!	Measures overall inequality in ratio distribution.

Final Thoughts

Develop a Pipeline That Works for You!

It may be beneficial to try a combination of methods to provide a clear picture of the market forces driving value in the area. Example:

1. Build out an OLS Model
2. Build a GWR model to explore spatial variations of coefficients, errors, and point estimates
3. Use local Morans I, K-means, etc. to cluster any of the spatially varying coefficients, residuals, or diagnostics
4. Develop LOCF
5. Interpolate LOCF over study area
6. Use either K-means or SCMVC to derive optimal segments
7. Create a web application with results of analysis and present to the local market experts
8. Manually draw based on results of analysis and knowledge of market



- Market segmenting is not only a science but also a subjective art
- Analyze various variables that may contribute to value (PEGS)
- Try various methods, if not all methods, for initial segment development
- Continuously analyze segment summaries for both sales and subjects including ratio statistics to ensure proper homogeneity within groups and heterogeneity between groups
- Refine every year to keep up with changing market forces
- Provide analysis to appraisers or analysts in the form of web apps or other web-based interactive tools and have them derive the segments and rerun segment stats

TYLER CONNECT

EMPOWER | COLLABORATE | IMAGINE

tylertech.com

TYLER CONNECT

Thank you!

Daniel Fasteen

DANIEL.FASTEEN@TYLERTECH.COM

Empower to serve the public

Collaborate with industry peers

Imagine the possibilities

TYLER CONNECT

EMPOWER | COLLABORATE | IMAGINE

TYLER CONNECT

EMPOWER | COLLABORATE | IMAGINE

tylertech.com