



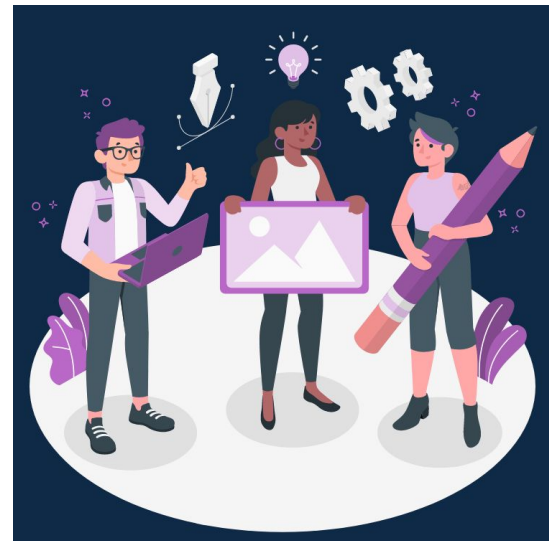
Land Use Data

Assuring quality and consistency in MTC's collection
of Bay Area jurisdictions' land use attributes

Danh Nguyen
Okeefe Niemann
University of San Francisco Practicum

Motivation

- What are our **challenges**?
 - Limited staff, limited time
 - “Mountain” of data
- What are our **current resources**?
 - Domain experts
 - Data Science Tools from past interns
 - Data (BASIS, publicly available jurisdiction policies)
- **How** can we take advantage of these resources to overcome these challenges?



Goals

1. Implement **tools** that further address the limitation of staff and time.
2. Provide **insights** into the zoning policy changes, evolution of land use over time.
3. **Bridge** the gap between current technologies and city planning.





Beyond Web Scrapping

Challenges

Cupertino

Table 19.28.060 Site Development Regulations			
	R1-5	R1-6, 7.5, 8, 10, 20, etc., and R1-6e	R1-a
A. Minimum net lot area ¹	5,000 square feet	the number multiplied by 1,000 square feet	10,000 square feet
B. Minimum lot width (at the front setback line)	50 feet	60 feet	75 feet
C. Landscaping	See Chapter 14.15 , Landscape Ordinance		Landscaping plans are required for all additions or new homes. The purpose of the landscaping is to beautify the property and to achieve partial screening of building

Scrapping tables in jurisdiction policies is an unexpectedly nuanced process

HEIGHT, AREA & SETBACK REGULATIONS FOR PRIMARY STRUCTURE											
Zone	Maxim um Height Feet ¹	Minimum Building Site Sq. Ft.	Minimum Lot Width in Feet		Maximu m Lot Coverag e	Minimum Density Allowed (Units per Gross Developable Acre)	Maximum Density Allowed Units Per Gross Developable Acre ⁶	Front Yard Minimum	Minimum Side Yard Required in Feet ²		Minimum Rear Yard Required in Feet
			Corner	Interior					Corner	Interior	
RE	TO BE DETERMINED BY CITY COUNCIL THROUGH PLANNED DEVELOPMENT PROCESS										
RR	TO BE DETERMINED BY CITY COUNCIL THROUGH PLANNED DEVELOPMENT PROCESS										
R-4	35	6,000	65	60	40%	NA	4 du/acre	f	f	5 ft.	20 ft.
R-6	35	6,000	65	60	40%	NA	6 du/acre	f	f	5 ft.	20 ft.
R-10	45	6,000	65	60	40%	NA	10 du/acre	f	f	5 ft.	10 ft.
R-20	45	20,000	70	70	40%	NA	20 du/acre	f	f	5 ft.	10 ft.
R-25	45	20,000	70	70	50%	20 du/acre	25 du/acre	f	f	5 ft.	10 ft. ^m
R-35	45	20,000	70	70	50%	30 du/acre	35 du/acre	f	f	5 ft.	10 ft. ^m
PBC	35	20,000	65	60	35%	NA	0	f	f	0 ft.	0 ft.

Antioch

Process Overview

1. Retrieve scraped documents from DB

- scraped policies from jurisdictions in the bay area

- Format: raw text files (limited table extraction)

*Scrape Date: 4/21/2020

2. Extract information by attribute

- scan through policy contents, identifying lines referring to building attributes

- Info pulled via pre-determined word pattern/value range for each attribute

3. Organize found attributes

- Populate a single dataframe for all cities with the following columns:

- City
- Attribute
- Values
- Context
- Policy Subsection
- Line No.

4. Identify potential zone context

- Find any potential zone candidates for found attributes

- Finalize the DataFrame with this information under "Zone Candidates"

Use Case

VERSION: JAN 21, 2021 (CURRENT) ▼
AND DEFINITIONS

➤ Chapter 14.04 - ZONING DISTRICTS
DESIGNATED

▼ Chapter 14.06 - R1-10 SINGLE-FAMILY DISTRICT

14.06.010 - R1-10 District.

14.06.020 - Permitted uses (R1-10).

14.06.030 - Conditional uses (R1-10).

14.06.040 - Site area (R1-10).

14.06.050 - Site frontage, width and depth (R1-10).

14.06.060 - Coverage (R1-10).

14.06.070 - Floor area ratio (R1-10).

14.06.080 - Setbacks (R1-10).

14.06.090 - Height of structures (R1-10).

14.06.100 - Daylight plane (R1-10).

14.06.110 - Basements (R1-10).

14.06.120 - Outdoor kitchens,
barbeques, fireplaces and swimming
pools.

14.06.130 - Design control (R1-10).

14.06.140 - Off-street parking (R1-10).

14.06.150 - Signs (R1-10).

14.06.020 - Permitted uses (R1-10).

The following uses shall be permitted in the R1-10 District:

- A. Single-family residences, including ancillary accessory structures;
- B. Accessory dwelling units as provided in [Chapter 14.14](#) of this title;
- C. Home occupations;
- D. Agriculture and horticulture;
- E. Animals as provided in [Chapter 5.10](#) of this code; and
- F. Small family day care.

(Ord. 05-285 § 2 (part); Ord. 04-267 § 2 (part))

(Ord. No. [2018-448](#), § 3, 7-10-2018)

14.06.030 - Conditional uses (R1-10).

Upon the granting of a use permit in accordance with the provisions of [Chapter 14.80](#) of this title, the following uses shall be permitted in the R1-10 District:

- A. Flag lots;
- B. Large family day care; and
- C. Pre-existing community facilities as provided in [Chapter 14.70](#) of this title. New community facilities are prohibited.

(Ord. 05-285 § 2 (part); Ord. 05-271 § 2 (part); Ord. 04-267 § 2 (part))

14.06.040 - Site area (R1-10).

The minimum site area shall be ten thousand (10,000) square feet, except that on corner lots the minimum site area shall be eleven thousand (11,000) square feet and on flag lots the minimum site area shall be fifteen thousand (15,000) square feet.

(Ord. 04-267 § 2 (part))

14.06.050 - Site frontage, width and depth (R1-10).

- A. The minimum site frontage and width shall be eighty (80) feet and the minimum site depth shall be one hundred (100) feet, except that the minimum site width for a corner lot shall be ninety (90) feet and the minimum site frontage on a cul-de-sac turnaround shall be sixty (60) feet.
- B. The minimum width of the access corridor for each flag lot shall be twenty (20) feet, and shall access directly to a public or private street. The access corridor shall not connect to any portion of the turnaround space of a cul-de-sac.

(Ord. 04-267 § 2 (part))

Context - In Corpus

```
print_Context(df=df_all_cities, line_num=13392, context_window=10)
```

City: los_altos

Attribute: minimum_lot_sqft

Policy Section: Title 14 - ZONING

Line Number: 343

Line: The minimum site area shall be ten thousand (10,000) square feet, except that on corner lots the minimum site area shall be eleven thousand (11,000) square feet and on flag lots the minimum site area shall be fifteen thousand (15,000) square feet.

Context:

334: A.

335: Flag lots;

336: B.

337: Large family day care; and

338: C.

339: Pre-existing community facilities as provided in Chapter 14.70 of this title. New community facilities are prohibited.

340: (Ord. 05-285 § 2 (part): Ord. 05-271 § 2 (part): Ord. 04-267 § 2 (part))

341: 14.06.040 - Site area (R1-10).

342: COMPARE VERSIONS

343: The minimum site area shall be ten thousand (10,000) square feet, except that on corner lots the minimum site area shall be eleven thousand (11,000) square feet and on flag lots the minimum site area shall be fifteen thousand (15,000) square feet.

344: (Ord. 04-267 § 2 (part))

345: 14.06.050 - Site frontage, width and depth (R1-10).

346: COMPARE VERSIONS

347: A.

348: The minimum site frontage and width shall be eighty (80) feet and the minimum site depth shall be one hundred (100) feet, except that the minimum site width for a corner lot shall be ninety (90) feet and the minimum site frontage on a cul-de-sac turnaround shall be sixty (60) feet.

349: B.

350: The minimum width of the access corridor for each flag lot shall be twenty (20) feet, and shall access directly to a public or private street. The access corridor shall not connect to any portion of the turnaround space of a cul-de-sac.

351: (Ord. 04-267 § 2 (part))

352: 14.06.060 - Coverage (R1-10).

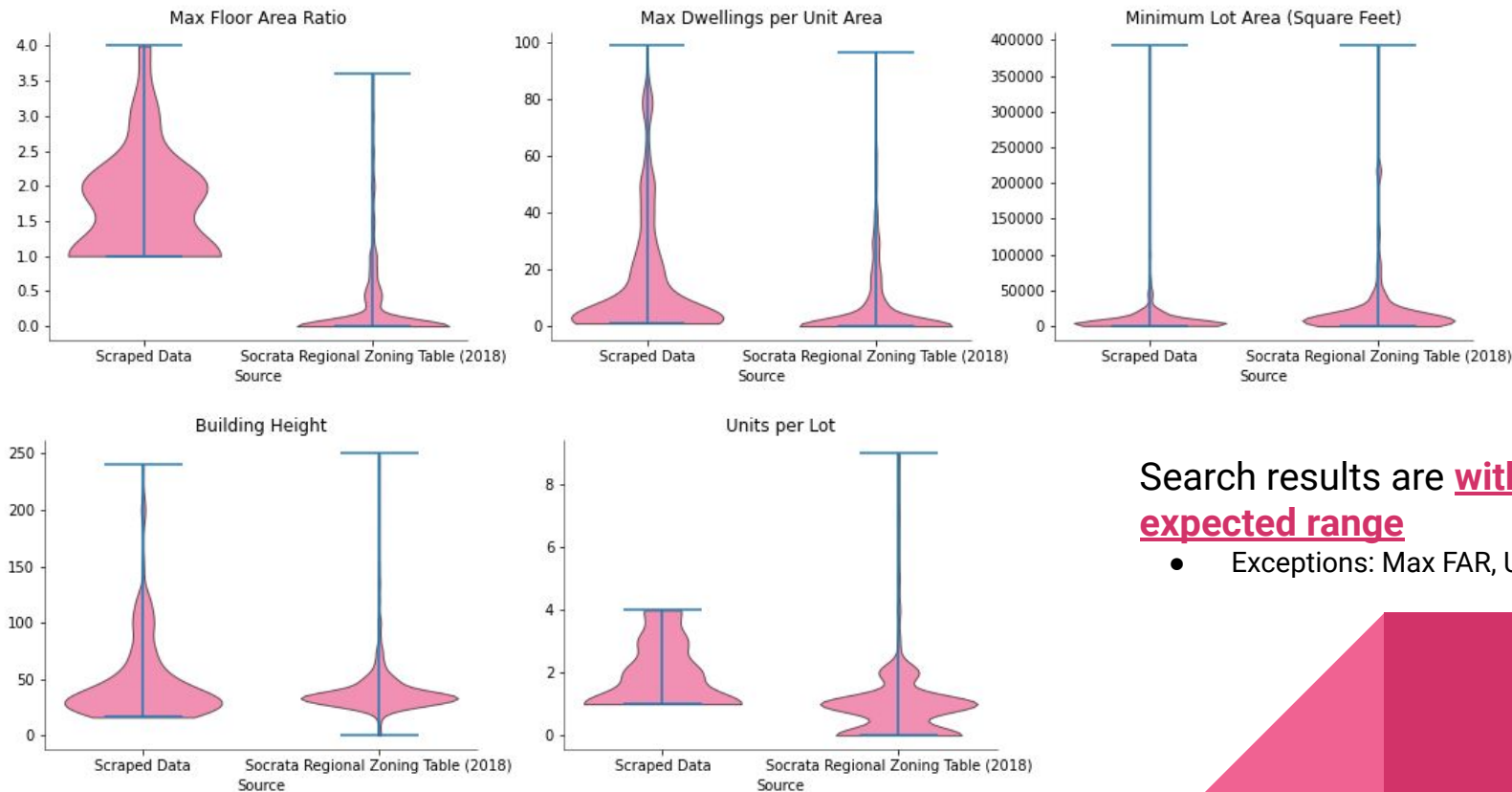
Identify Potential Zone Codes

	City	Attribute	Values	Context	Policy Subsection	Line No.	Zone Candidates
13381	los_altos	minimum_lot_sqft	2019.0, 750.0	R313.1 Townhouse automatic fire sprinkler systems. An automatic residential fire sprinkler system shall be installed in accordance with National Fire Protection Association's (NFPA) Standard 13D in all new townhouses and in existing townhouses, when additions are made that increase the building area to more than the allowable Fire-Flow Appendix B, Tables B105.1(1) and B105.1(2) of the 2019 California Fire Code, and/or additions exceeding fifty (50) percent of the existing living area (existing square foot calculations shall not include existing basement) and/or additions exceeding seven hundred and fifty (750) square feet. When automatic fire sprinkler systems are required by this section, all associated attached garages shall be included. Additions over fifty (50) percent and/or seven hundred and fifty (750) square feet as referenced above, shall be treated as a new structure regarding installation of fire sprinkler systems. For the purpose of this section, removal of roof framing with associated exterior walls down to, or below the subfloor/slab shall be included in the above calculations. Therefore, the following shall apply:	Title 12 - BUILDINGS AND CONSTRUCTION	86	D2, R301, R313, D0, B105, D1
13382	los_altos	minimum_lot_sqft	2019.0, 750.0	R313.2 One and two-family dwellings automatic fire sprinklers systems. An automatic residential fire sprinkler system shall be installed in accordance with National Fire Protection Association's (NFPA) Standard 13D in all new one and two-family dwellings and in existing dwellings, when additions are made that increase the building area to more than the allowable Fire-Flow Appendix Tables B105.1(1) and B105.1(2) of the 2019 California Fire Code, and/or additions exceeding fifty (50) percent of the existing living area (existing square foot calculations shall not include existing basement) and/or additions exceeding seven hundred and fifty (750) square feet. When automatic fire sprinkler systems are required by this section, all associated garages shall be included. Additions over fifty (50) percent and/or seven hundred and (750) square feet as referenced above, shall be treated as a new structure regarding installation of fire sprinkler systems. For the purpose of this section, removal of roof framing with associated exterior walls down to, or below the subfloor/slab shall be included in the above calculations. Therefore, the following shall apply:	Title 12 - BUILDINGS AND CONSTRUCTION	89	D2, R602, D0, R313, B105, D1
13389	los_altos	minimum_lot_sqft	5000.0, 3999.0	Apartments, two family units or more per building with not more than one dwelling unit for each five thousand (5,000) square feet of lot area; provided, however, if after dividing the area of the site by five thousand (5,000), a remainder of less than five thousand (5,000) square feet but more than three thousand nine hundred ninety-nine (3,999) square feet is obtained, one additional dwelling unit may be located on the site;	Title 14 - ZONING	1385	R3-5
13390	los_altos	minimum_lot_sqft	20000.0, 1000.0, 43560.0, 10000.0, 15000.0	Density. The maximum number of permitted dwelling units shall be calculated by subtracting twenty (20) percent of the gross area of the parcel and dividing the remainder by the required lot area per dwelling unit in the appropriate R1 District (ten thousand (10,000) square feet in the R1-10 District, fifteen thousand (15,000) square feet in the R1-H District, twenty thousand (20,000) square feet in the R1-20 District, and forty-three thousand five hundred sixty (43,560) square feet in the R1-40 District).	Title 14 - ZONING	4781	R1-40, R1-, R1-10, R1-20, R1
13392	los_altos	minimum_lot_sqft	10000.0, 11000.0, 15000.0	The minimum site area shall be ten thousand (10,000) square feet, except that on corner lots the minimum site area shall be eleven thousand (11,000) square feet and on flag lots the minimum site area shall be fifteen thousand (15,000) square feet.	Title 14 - ZONING	344	R1-10
13393	los_altos	minimum_lot_sqft	20000.0, 21000.0	The minimum site area shall be twenty thousand (20,000) square feet, except that on corner lots the minimum site area shall be twenty-one thousand (21,000) square feet.	Title 14 - ZONING	510	R1-



Data At A Glance

Do extracted values align with previous data?

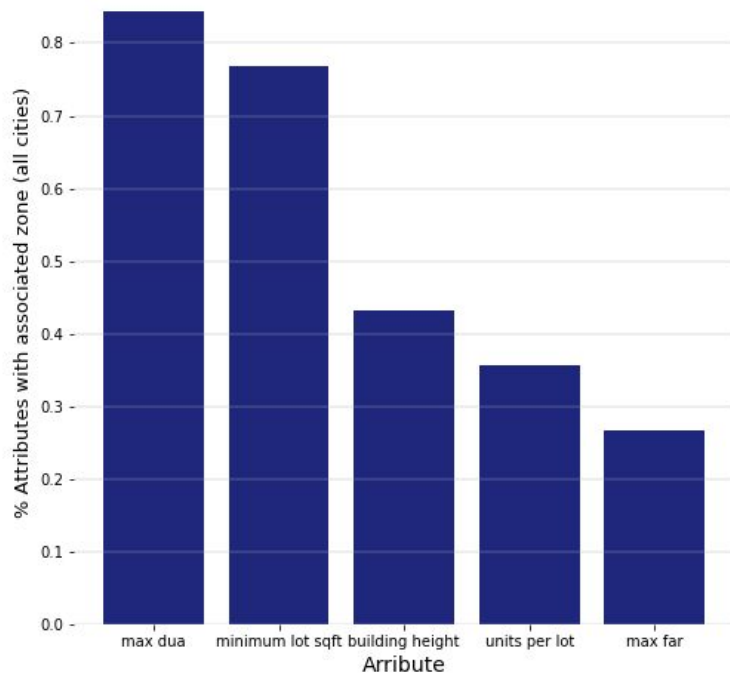


Search results are within expected range

- Exceptions: Max FAR, Units per Lot

Completion Rates

How many zones were found?

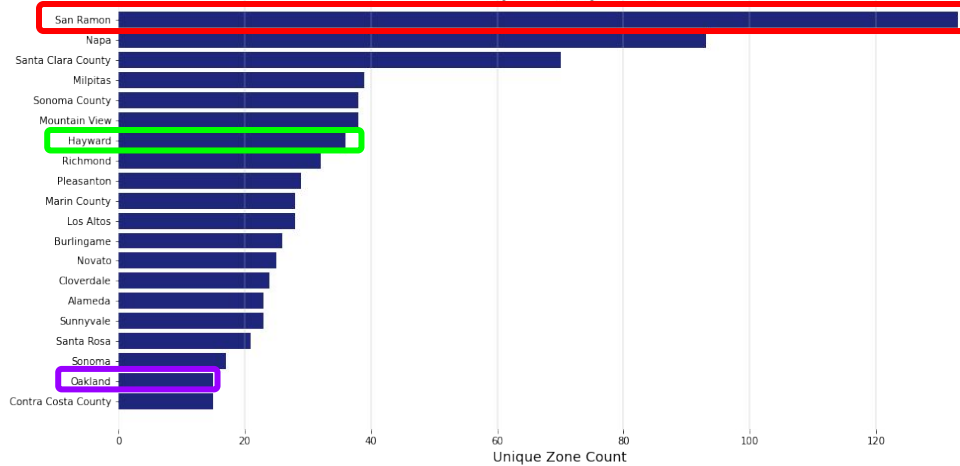


* Derived / hidden values requires more human interventions

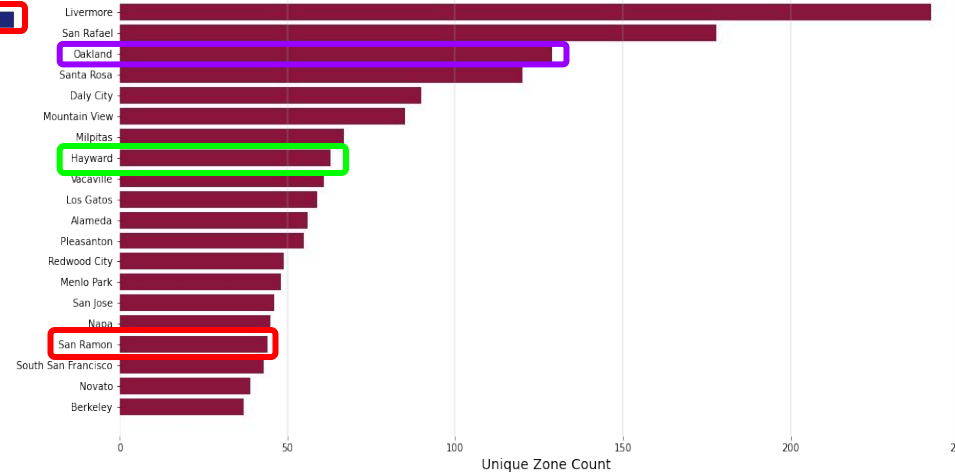
- Max dua vs max far

A consistency problem

Attribute Scraper (Top 20)

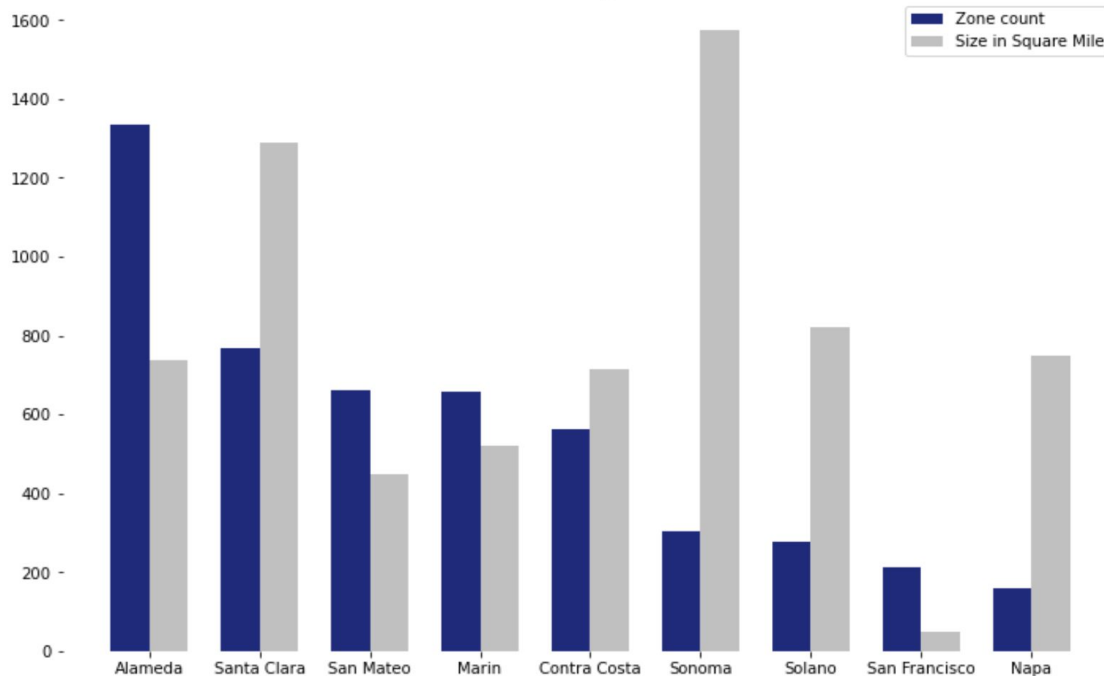


Socrata Regional Zoning Table 2018 (Top 20)



Where do we start?

Zone Count and County Size



Key Points:

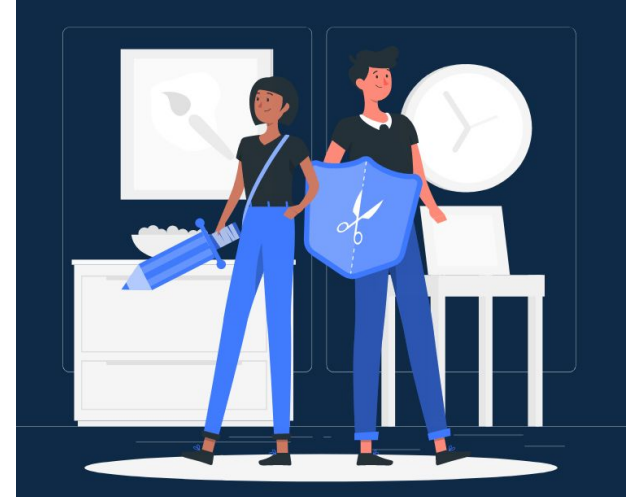
- Not all counties are created equal
- Need standardization of policies



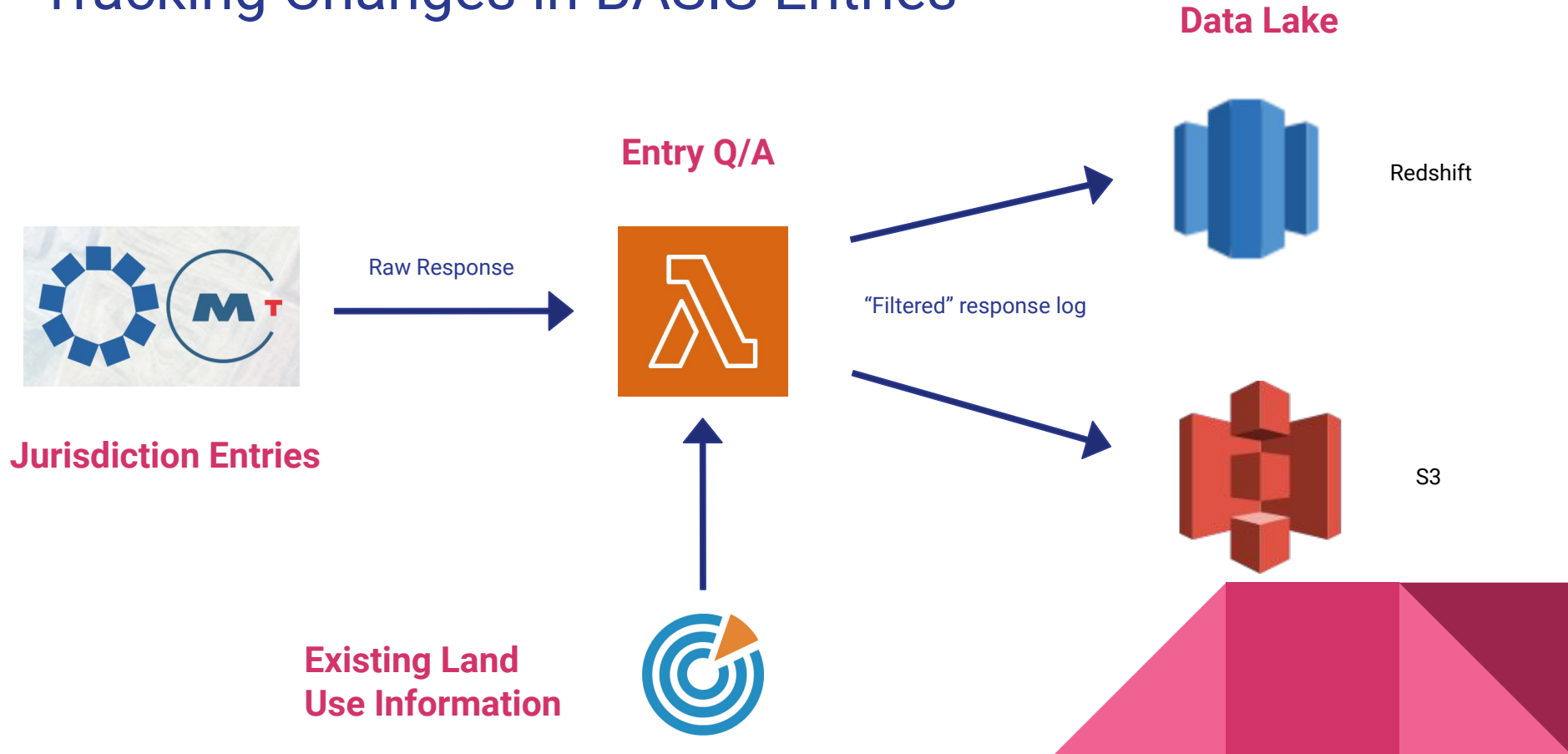
Maintaining Data Integrity

Motivation:

- Jurisdictions given window of time to update or add entries for parcel information with BASIS without much rollback capability
- Limited quality assurance of manual entries in existing pipelines
- Risking clerical errors or category misunderstanding when completing forms



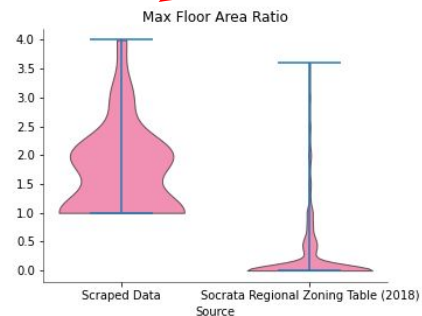
Tracking Changes in BASIS Entries



Use Case

recid	city_name	cols_updated	updated_from_NaN	changed_vals	out_of_range	warnings	editor	edit_timestamp	edit_type (C/U/D)	main_version	old_vals	new_vals
0	thiswasadded1-f595-4fa8-8da2-975dfae46dc4	Cloverdale	row created	NaN	NaN	0		Joshua Croff	2019-10-20 17:00:00	C	1	[thiswasadded1-f595-4fa8-8da2-975dfae46dc4, Cloverdale, M-1, General Industrial, None, 5.0, 0.0, 0.0, nan, nan, Joshua Croff, 2019-10-20 17:00:00, Sonoma, #75187C, nan, nan]
0	dc4e140a-ed57-42c5-b4d0-89444919e88f	Dublin	[zn_area_overlay, regional_lu_class, max_dua, edit_date]	-1	5	0		Michael Cass	2019-10-29 00:00:00	U	1	[dc4e140a-ed57-42c5-b4d0-89444919e88f, Dublin, C-N, Neighborhood Commercial, nan, 3.0, nan, nan, 35.0, nan, Michael Cass, 2019-10-29 00:00:00, Alameda, #FF6B6B, 5000.0, nan]
0	e63d2e0d-2baf-4d4d-a6f8-5eae9bdbc33c	Fremont	[zn_area_overlay, regional_lu_class, building_height, edit_date]	1	4	2	[regional_lu_class value is out of range (val = 123.0), max_dua value is out of range (val = 5000.0)]	Joshua Croff	2019-10-24 17:00:00	U	1	[e63d2e0d-2baf-4d4d-a6f8-5eae9bdbc33c, Fremont, P-87-2, Planned District, nan, 11.0, 0.0, 0.0, nan, nan, Joshua Croff, 2019-05-24 00:00:00, Alameda, #556B2F, nan, nan]
0	f063d0c3-d09f-47e7-9381-f71b954ca0fe	Fremont	[regional_lu_class, max_far, max_dua, edit_date]	0	4	0		Marc Cleveland	2019-10-31 00:00:00	U	1	[f063d0c3-d09f-47e7-9381-f71b954ca0fe, Fremont, OS(Q), Open Space, Quarry Overlay, 7.0, 0.0, 1.0, 30.0, 1.0, Marc Cleveland, 2019-10-31 00:00:00, Alameda, #006A00, 43560.0, nan]
0	62db00d9-c9ff-4764-8d84-59413fe5f0b6	Oakland	[max_far, edit_date, source]	1	2	1	[max_far value is out of range (val = 16.0)]	Joshua Croff	2019-10-24 17:00:00	U	1	[62db00d9-c9ff-4764-8d84-59413fe5f0b6, Oakland, CN-2 - D-KP-3, Commercial - Neighborhood Center, Kaiser Permanente Oakland Medical Center Zone, 3.0, 0.0, 0.0, nan, nan, Joshua Croff, 2019-05-22 00:00:00, Alameda, #FF0000, nan, nan]
0	c1a59692-ee49-4677-95d0-f64e580e5754	Oakland	[zn_area_overlay, units_per_lot, edit_date, minimum_lot_sqft, source]	3	2	1	[units_per_lot value is out of range (val = 12.0)]	Joshua Croff	2019-10-24 17:00:00	U	1	[c1a59692-ee49-4677-95d0-f64e580e5754, Oakland, D-KP-1, Special and Combined - Kaiser Permanente Oakland Medical, nan, 4.0, 0.0, 0.0, nan, nan, Joshua Croff, 2019-05-22 00:00:00, Alameda, #990000, nan, nan]

	city	editor	edit_timestamp	recid	warning
0	Cloverdale	Joshua Croff	2019-10-21 00:00:00	thiswasadded1-f595-4fa8-8da2-975dfae46dc4	max_far value is out of range (val = 0.0)
1	Cloverdale	Joshua Croff	2019-10-21 00:00:00	thiswasadded1-f595-4fa8-8da2-975dfae46dc4	max_dua value is out of range (val = 0.0)
2	Cloverdale	Joshua Croff	2019-10-21 00:00:00	thiswasadded1-f595-4fa8-8da2-975dfae46dc4	0 Joshua Croff\nName: editor, dtype: object not in other row records for editor
3	Fremont	Joshua Croff	2019-10-25 00:00:00	e63d2e0d-2baf-4d4d-a6f8-5eae9bdbc33c	regional_lu_class value is out of range (val = 123.0)
4	Fremont	Joshua Croff	2019-10-25 00:00:00	e63d2e0d-2baf-4d4d-a6f8-5eae9bdbc33c	max_far value is out of range (val = 0.0)
5	Fremont	Joshua Croff	2019-10-25 00:00:00	e63d2e0d-2baf-4d4d-a6f8-5eae9bdbc33c	max_dua value is out of range (val = 5000.0)
6	Fremont	Joshua Croff	2019-10-25 00:00:00	e63d2e0d-2baf-4d4d-a6f8-5eae9bdbc33c	Joshua Croff not in other row records for editor
7	Oakland	Joshua Croff	2019-10-25 00:00:00	62db00d9-c9ff-4764-8d84-59413fe5f0b6	max_far value is out of range (val = 16.0)
8	Oakland	Joshua Croff	2019-10-25 00:00:00	62db00d9-c9ff-4764-8d84-59413fe5f0b6	max_dua value is out of range (val = 0.0)
9	Oakland	Joshua Croff	2019-10-25 00:00:00	62db00d9-c9ff-4764-8d84-59413fe5f0b6	3756 CN-2 - D-KP-3\nName: zn_code, dtype: object not in other row records for zn_code
10	Oakland	Joshua Croff	2019-10-25 00:00:00	62db00d9-c9ff-4764-8d84-59413fe5f0b6	3756 Kaiser Permanente Oakland Medical Center Zone\nName: zn_area_overlay, dtype: object not in other row records for zn_area_overlay
11	Oakland	Joshua Croff	2019-10-25 00:00:00	62db00d9-c9ff-4764-8d84-59413fe5f0b6	3756 25.0\nName: source, dtype: object not in other row records for source
12	Oakland	Joshua Croff	2019-10-25 00:00:00	62db00d9-c9ff-4764-8d84-59413fe5f0b6	3756 Joshua Croff\nName: editor, dtype: object not in other row records for editor
13	Oakland	Joshua Croff	2019-10-25 00:00:00	c1a59692-ee49-4677-95d0-f64e580e5754	max_far value is out of range (val = 0.0)
14	Oakland	Joshua Croff	2019-10-25 00:00:00	c1a59692-ee49-4677-95d0-f64e580e5754	max_dua value is out of range (val = 0.0)
15	Oakland	Joshua Croff	2019-10-25 00:00:00	c1a59692-ee49-4677-95d0-f64e580e5754	units_per_lot value is out of range (val = 12.0)



Next Steps

- **Collaborate** with subject matter experts to improve search results via refining text pattern matching, attribute ranges (ongoing)
- Use machine learning models to fill in missing values in BASIS tables
 - **Break down the barriers** inherent in communication when deciphering ordinance documents
- **Maintain** scrapers and continue to pull new data from policy websites





Thank you!