

## Xiaoxian Zhang

### Data Quality Report

#### Description: -

- The data provided is real estate property valuation and assessment records. It provides address, owner, property characteristics, assessed tax values, etc.
- Where it came from: The data is obtained from NYC Open Data and provided by Department of Finance.
- Time period it covers: This data covers applications between 2010 and 2011.
- Number of fields: There are 32 columns/fields.
- Number of records: There are 1,070,994 records.

#### Summary Tables: -

**Table for Numeric or Date Time Fields: -**

Field Name	% Populated	Min	Max	Mean	Std dev	% Zero
LOT	100	1	9,978	364.72	853.22	0
LTFRONT	100	0	9,999	36.64	74.03	15.79
LTDEPTH	100	0	9,999	88.86	76.40	15.89
STORIES	94.75	1	119	5.01	8.37	0
FULLVAL	100	0	6,150,000,000	874,264.51	11,582,430.99	1.21
AVLAND	100	0	2,668,500,000	85,067.92	4,057,260.06	1.21
AVTOT	100	0	4,668,308,947	227,238.17	6,877,529.31	1.21
EXLAND	100	0	2,668,500,000	36,423.89	3,981,575.79	45.91
EXTOT	100	0	4,668,308,947	91,186.98	6,508,402.82	40.39
BLDFRONT	100	0	7,575	23.04	35.58	21.36
BLDDEPTH	100	0	9,393	39.92	42.71	21.37
AVLAND2	26.40	3	2,371,005,000	246,235.72	6,178,962.56	0
AVTOT2	26.40	3	4,501,180,002	713,911.44	11,652,528.95	0
EXLAND2	8.17	1	2,371,005,000	351,235.68	10,802,212.67	0
EXTOT2	12.22	7	4,501,180,002	656,768.28	16,072,510.17	0

**Table for Categorical Fields: -**

Field Name	% Populated	# Unique Values	Most Common Value
RECORD	100	1,070,994	Unique value for each record
BBLE	100	1,070,994	Unique value for each record
BORO	100	5	4
BLOCK	100	13,984	3944

EASEMENT	43.29	13	E
OWNER	97.04	863,348	PARKCHESTER PRESERVAT
BLDGCL	100	200	R4
TAXCLASS	100	11	1
EXT	33.08	3	G
EXCD1	59.62	130	1017
STADDR	99.94	839,281	501 SURF AVENUE
ZIP	97.20	197	10314
EXMPTCL	1.45	15	X1
EXCD2	8.68	61	1017
PERIOD	100	1	FINAL
YEAR	100	1	2010/11
VALTYPE	100	1	AC-TR

## Distribution of Fields: -

### 1. RECORD

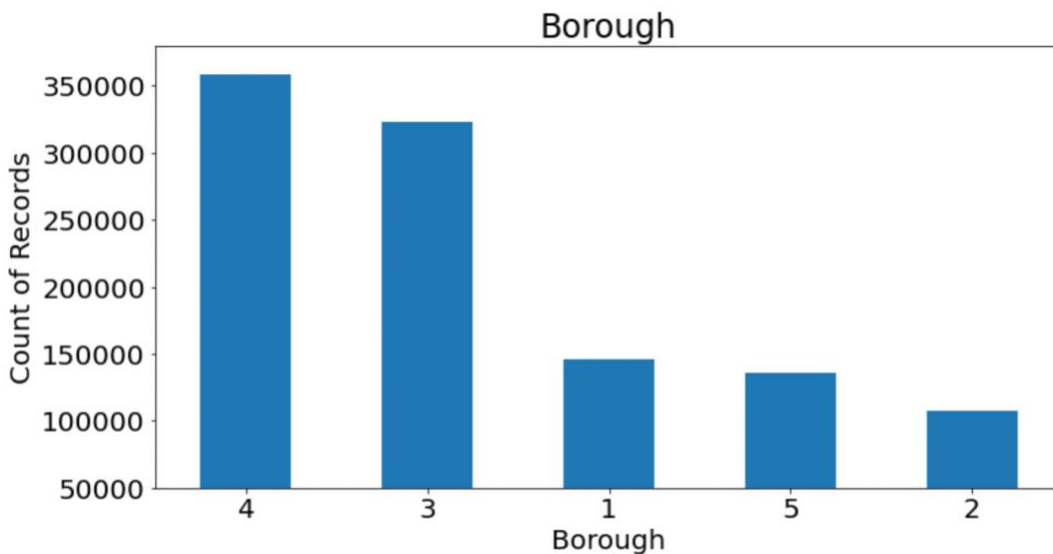
- The field is a unique identifier for each record, essentially index from 1 to 1070994.
- Each row here has a unique non null value.

### 2. BBLE

- The field is also a unique identifier for each record, essentially a serial number of file key.
- Each row here has a unique non null value.

### 3. BORO

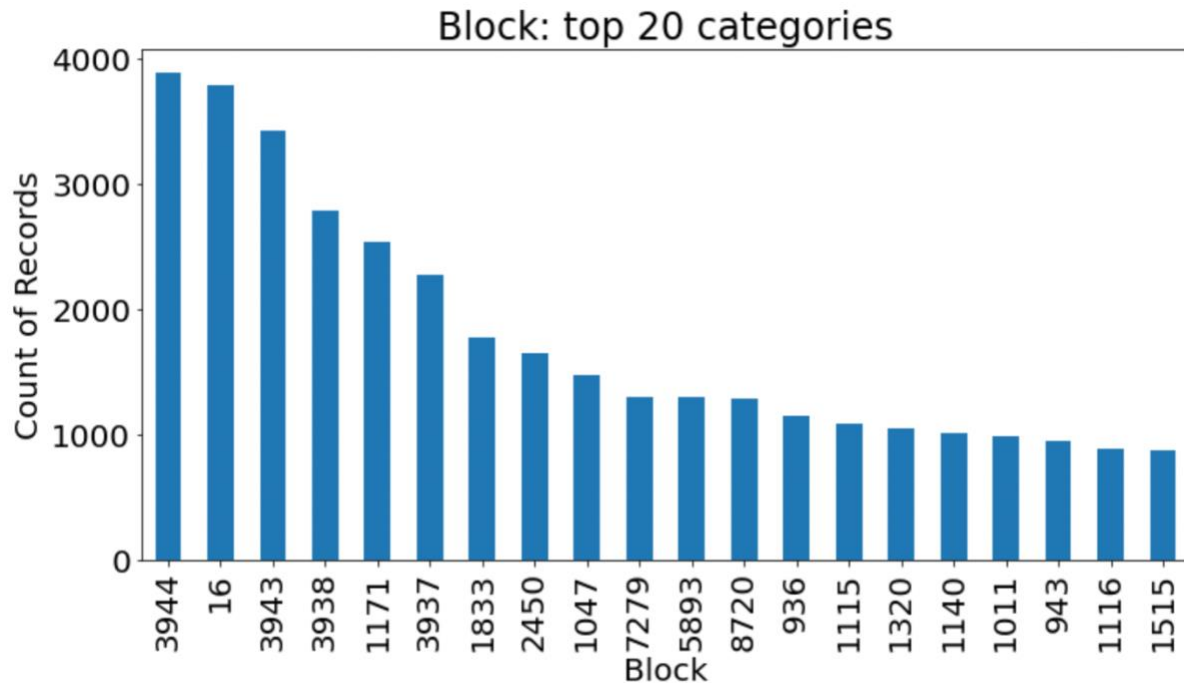
- This field represents the borough codes of the property.
- 1 = Manhattan, 2 = Bronx, 3 = Brooklyn, 4 = Queens, 5 = Staten Island



#### 4. BLOCK

-This field represents the block code of the property.

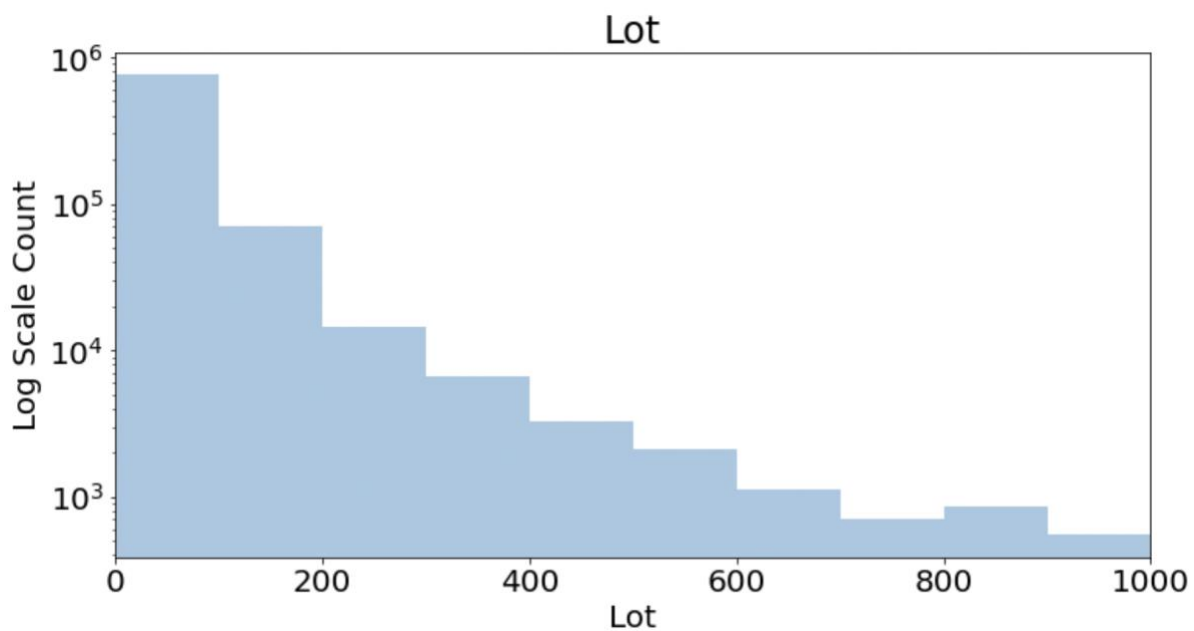
-Manhattan: 1 to 2255, Bronx: 2260 to 5958, Brooklyn: 1 to 8955, Queens: 1 to 16350, Staten Island: 1 to 8050



#### 5. LOT

-This field represents unique number within borough/block.

-The plot covers 80% of records.

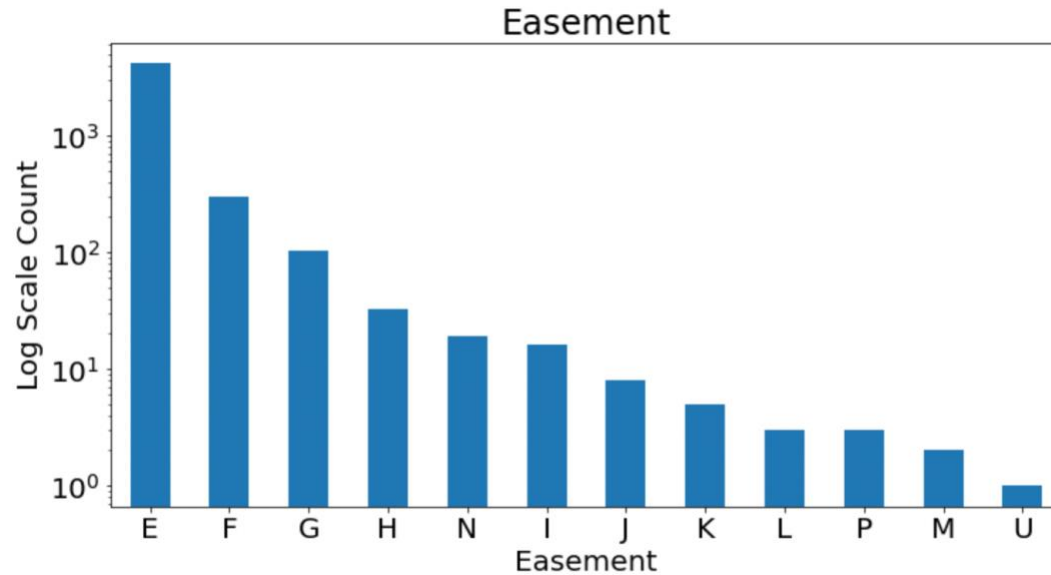


## 6. EASEMENT

-This field represents type of easement for each property.

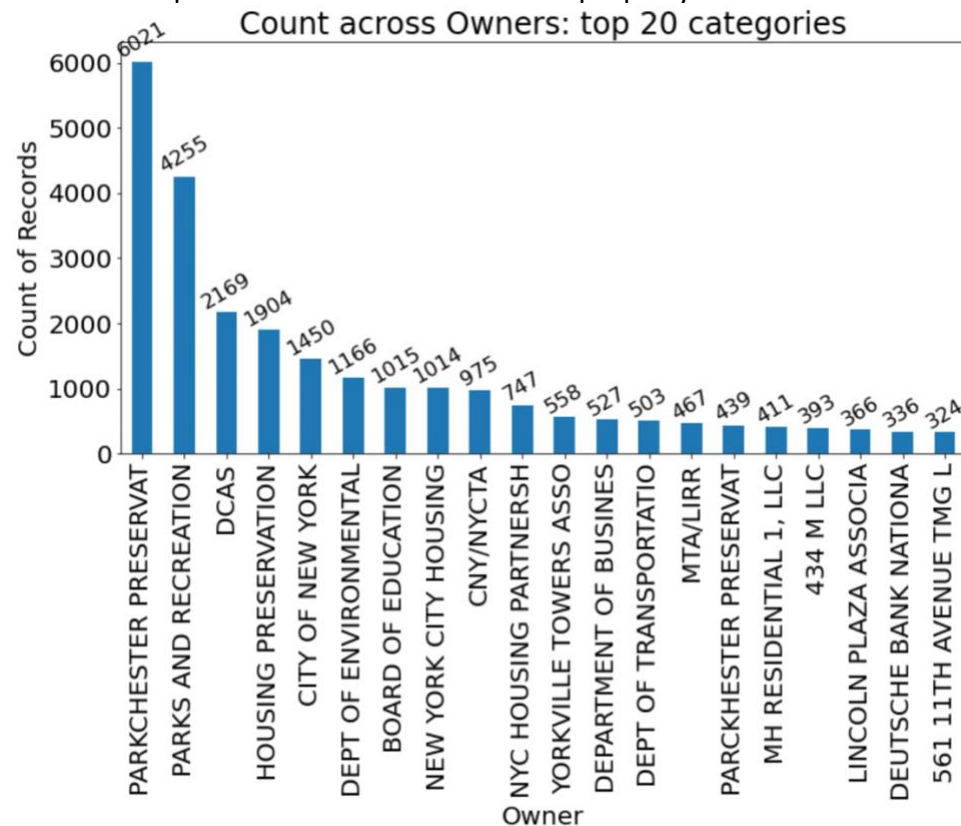
-Space = No Easement, A = Air Easement, B = Non-Air Rights, E = Land Easement, F Thru M are duplicates of E, N = Non-Transit Easement, P = Pier, R = Railroad, S = Street, U = U.S.

Government



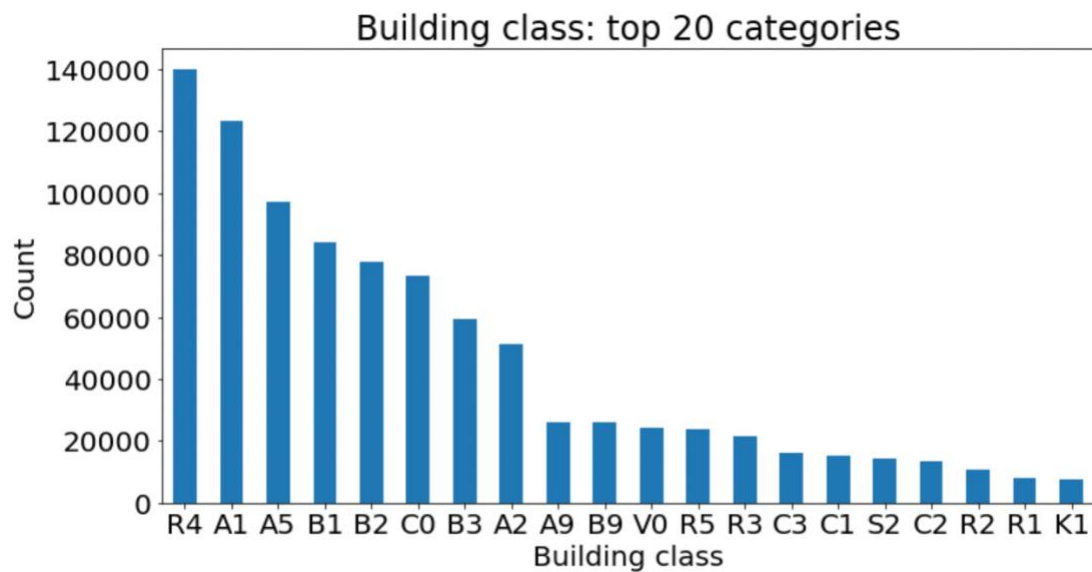
## 7. OWNER

-This field represents owner name for each property.



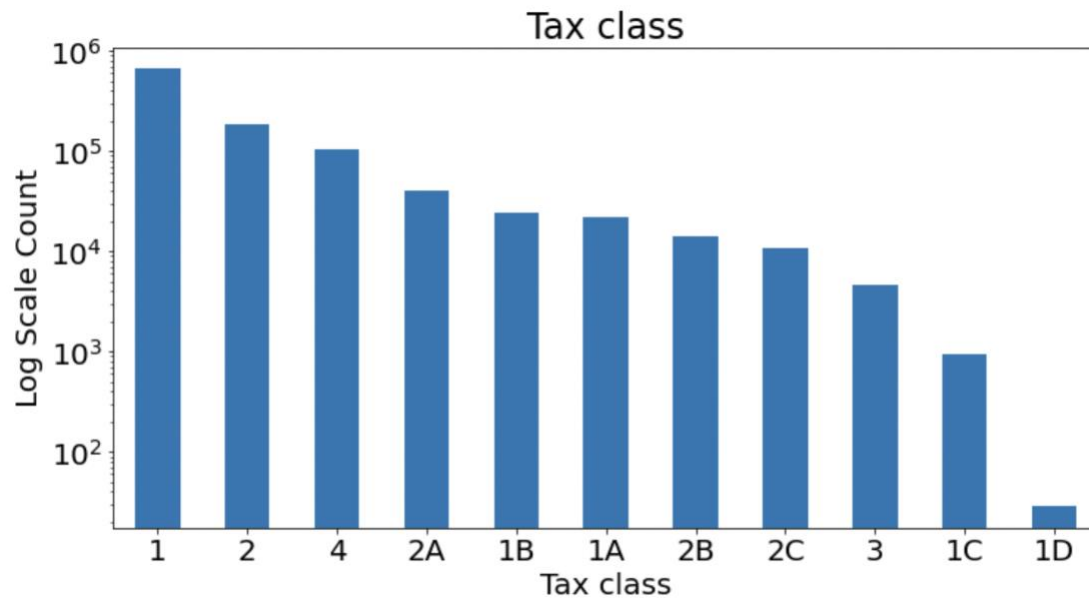
## 8. BLDGCL

-This field represents building class for each property.



## 9. TAXCLASS

-This field represents tax class for each property.

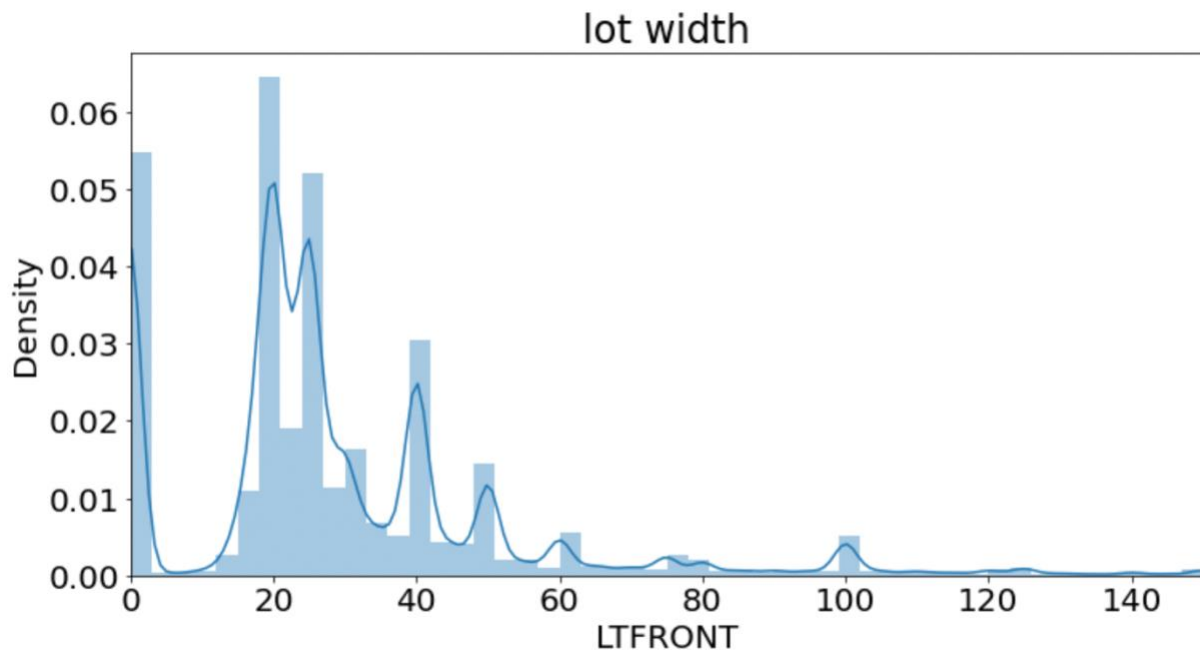


## 10. LTFRONT

-This field represents lot width for each property.

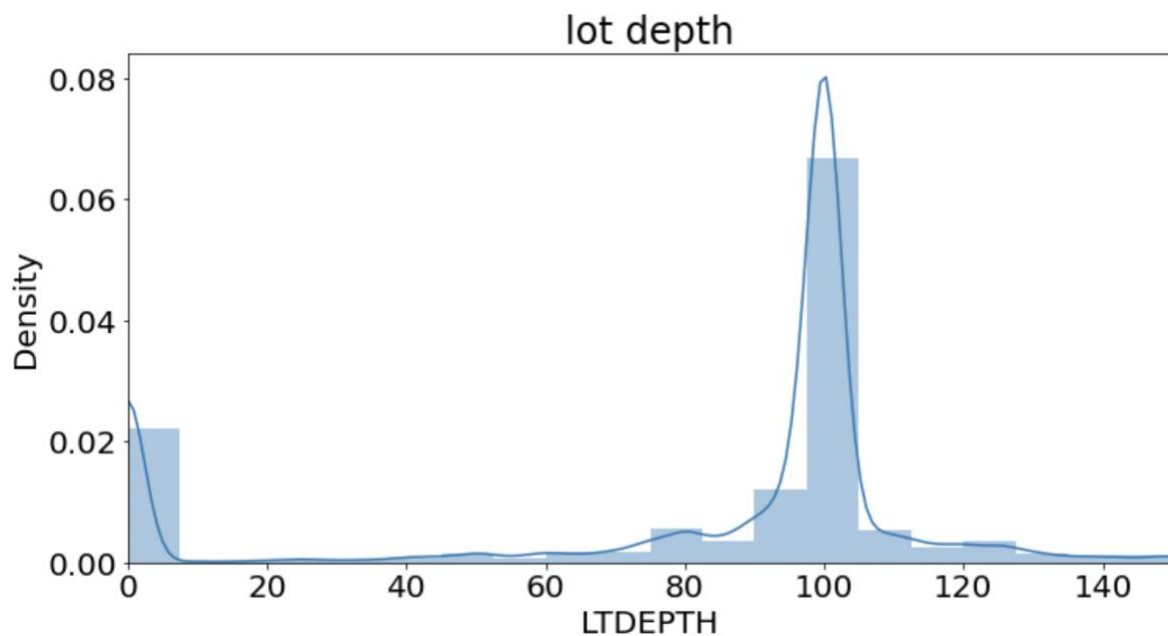
-Also, we see a lot of entries with zero values, which might be representing missing values.

-The plot covers 97% of records.



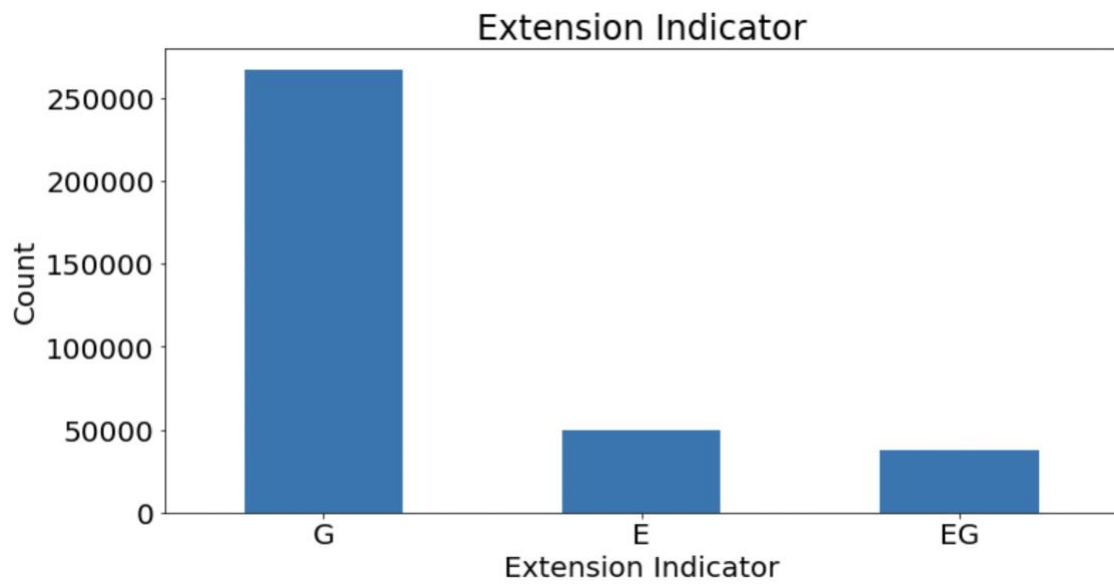
#### 11. LTDEPTH

- This field represents lot depth for each property.
- Also, we see a lot of entries with zero values, which might be representing missing values.
- The plot covers 96% of records.



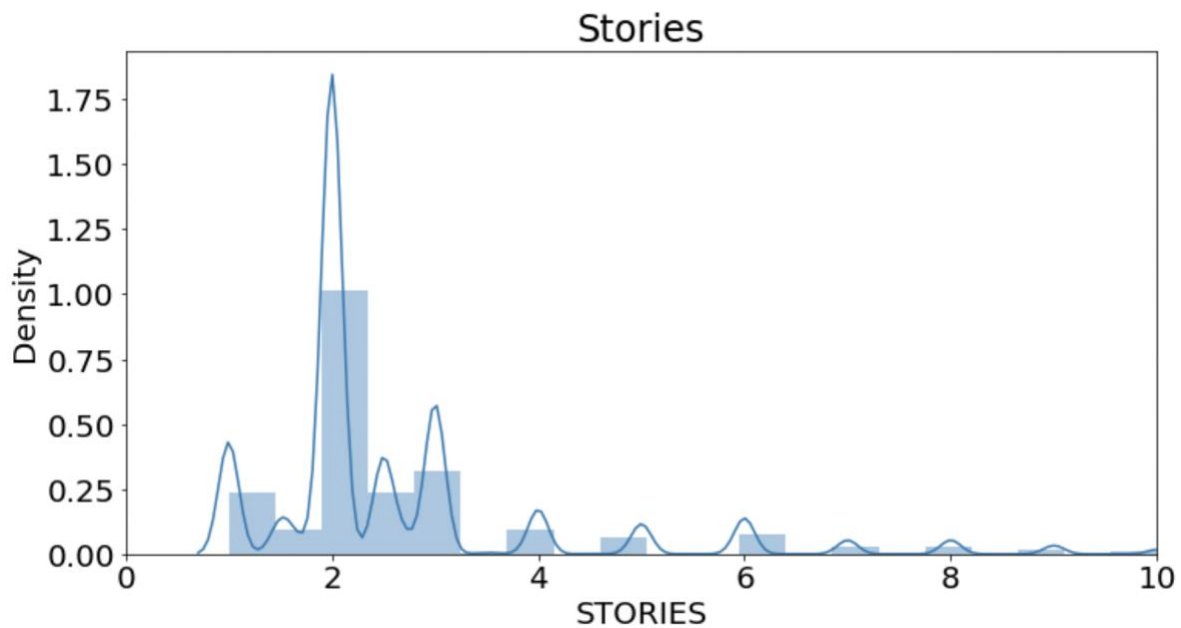
#### 12. EXT

- This field represents extension indicator for each property.



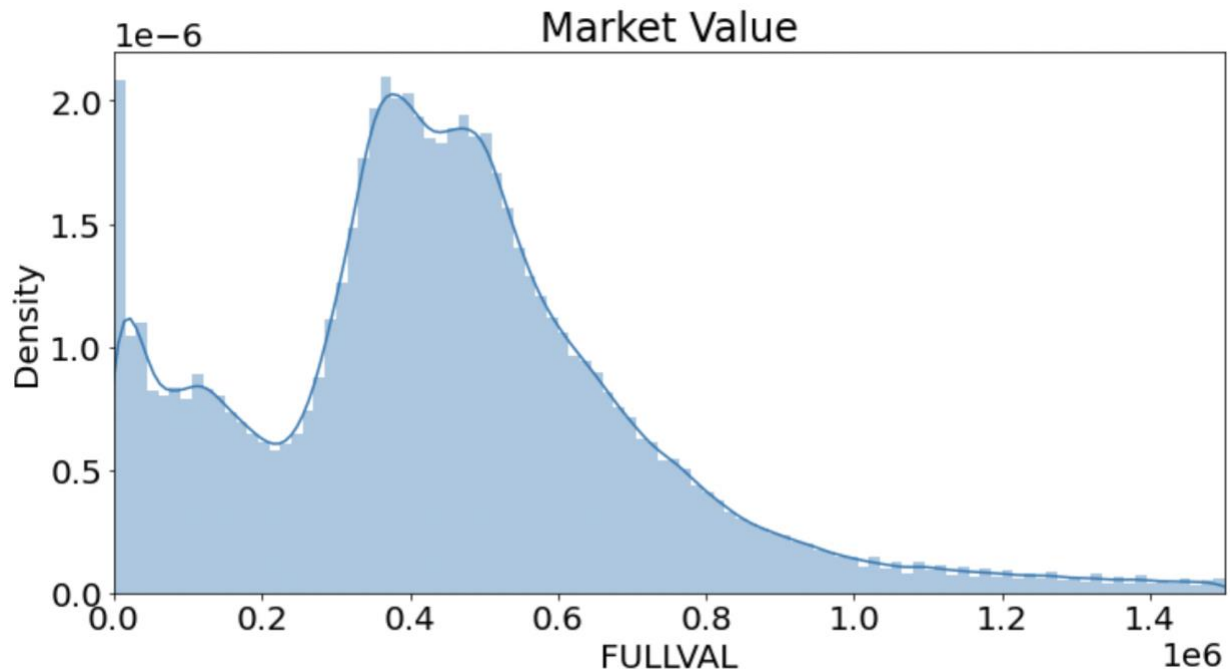
### 13. STORIES

- This field represents number of stories in each property.
- The plot covers 90% of records.



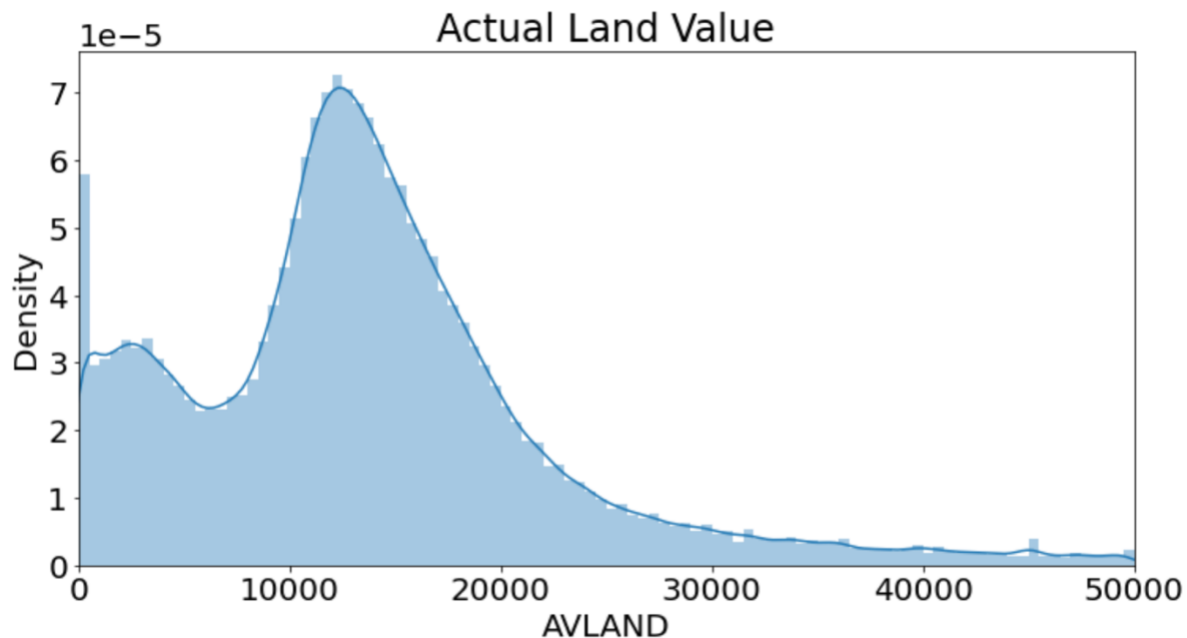
### 14. FULLVAL

- This field represents market value of each property.
- Also, we see a lot of entries with zero values, which might be representing missing values.
- The plot covers 95% of records.



#### 15. AVLAND

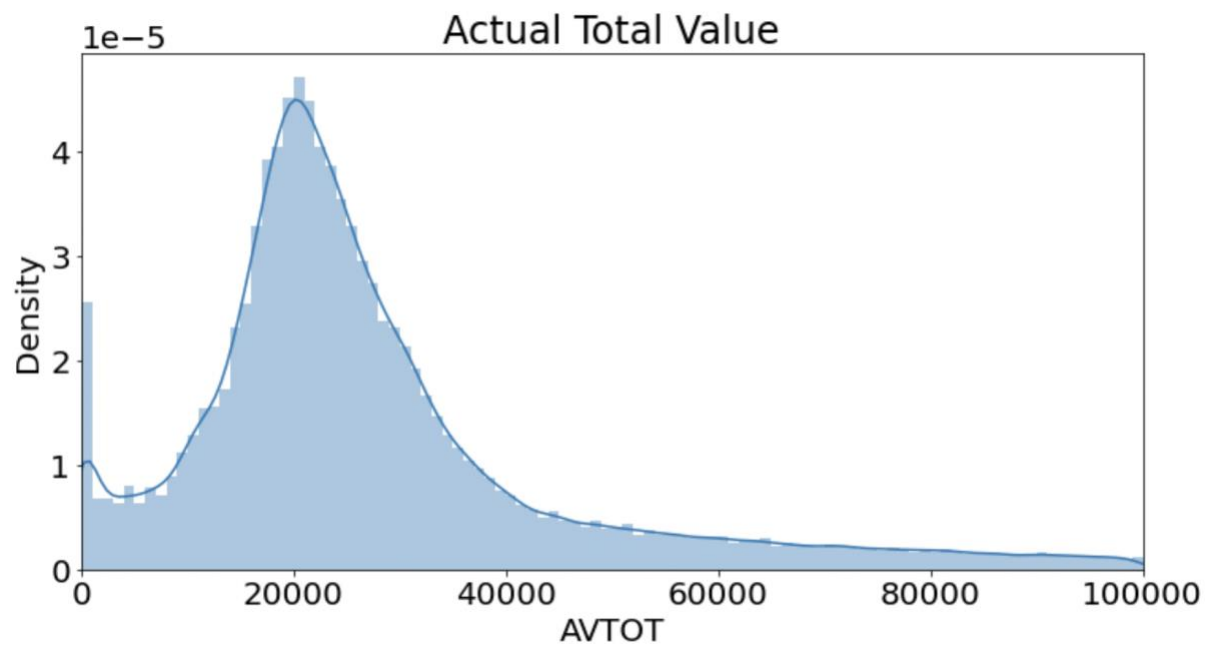
- This field represents actual land value of each property.
- Also, we see a lot of entries with zero values, which might be representing missing values.
- The plot covers 91% of records.



#### 16. AVTOT

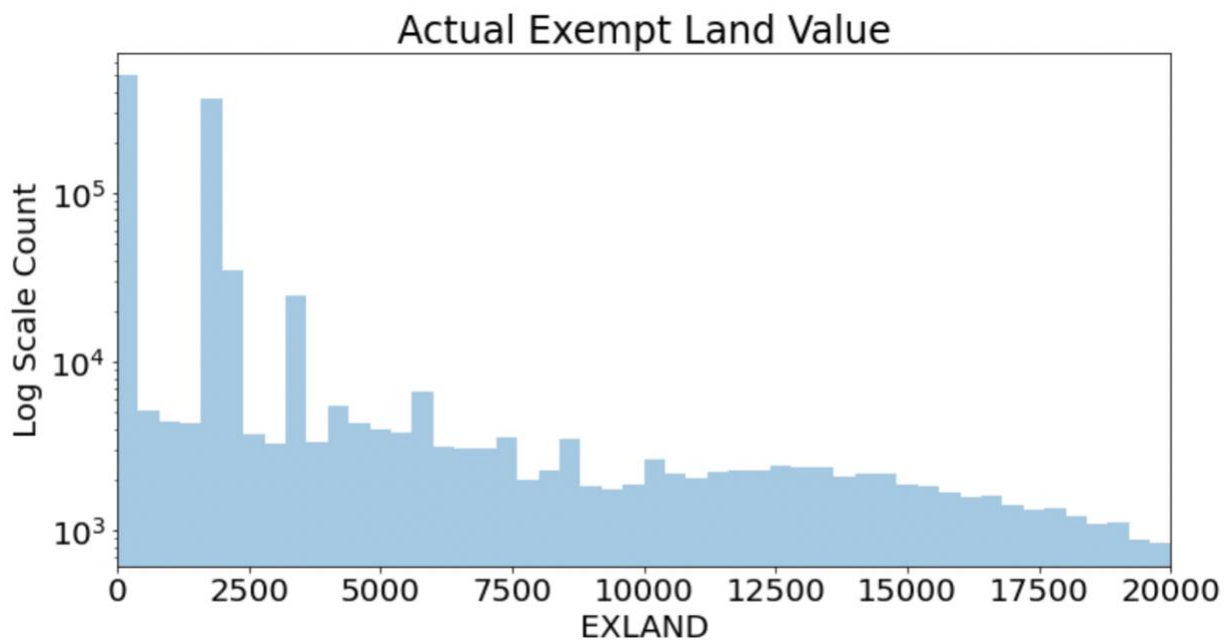
- This field represents actual total value of each property.
- Also, we see a lot of entries with zero values, which might be representing missing values.
- The plot covers 86% of records.





#### 17. EXLAND

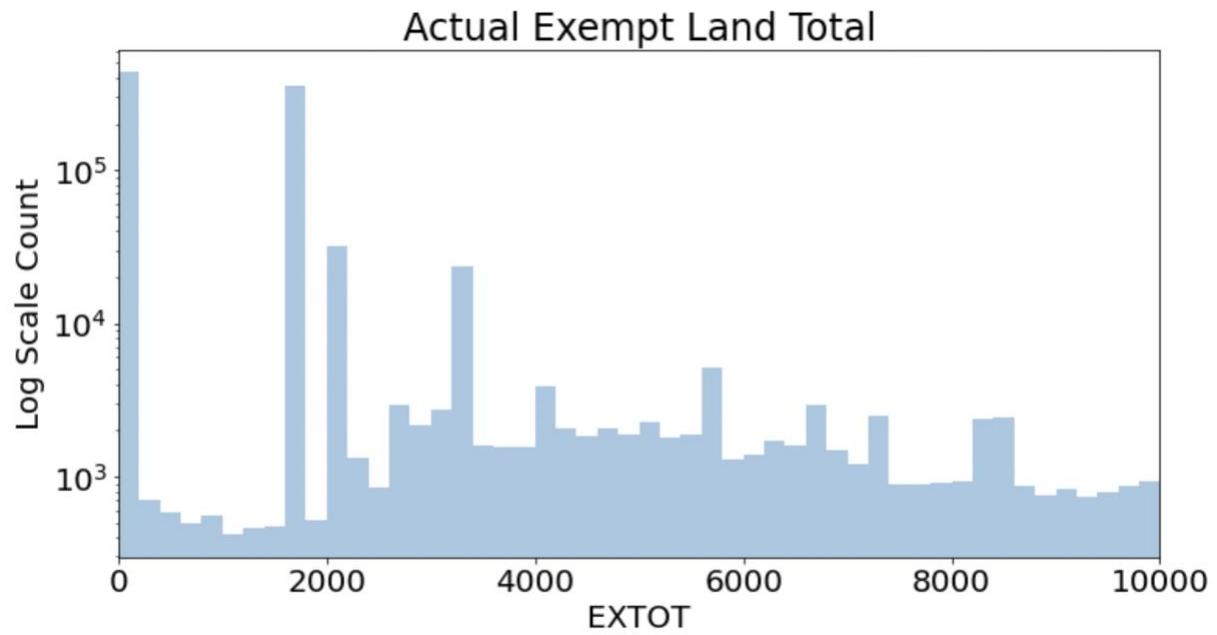
- This field represents actual exempt land value of each property.
- Also, we see a lot of entries with zero values, which might be representing missing values.
- The plot covers 97% of records.



#### 18. EXTOT

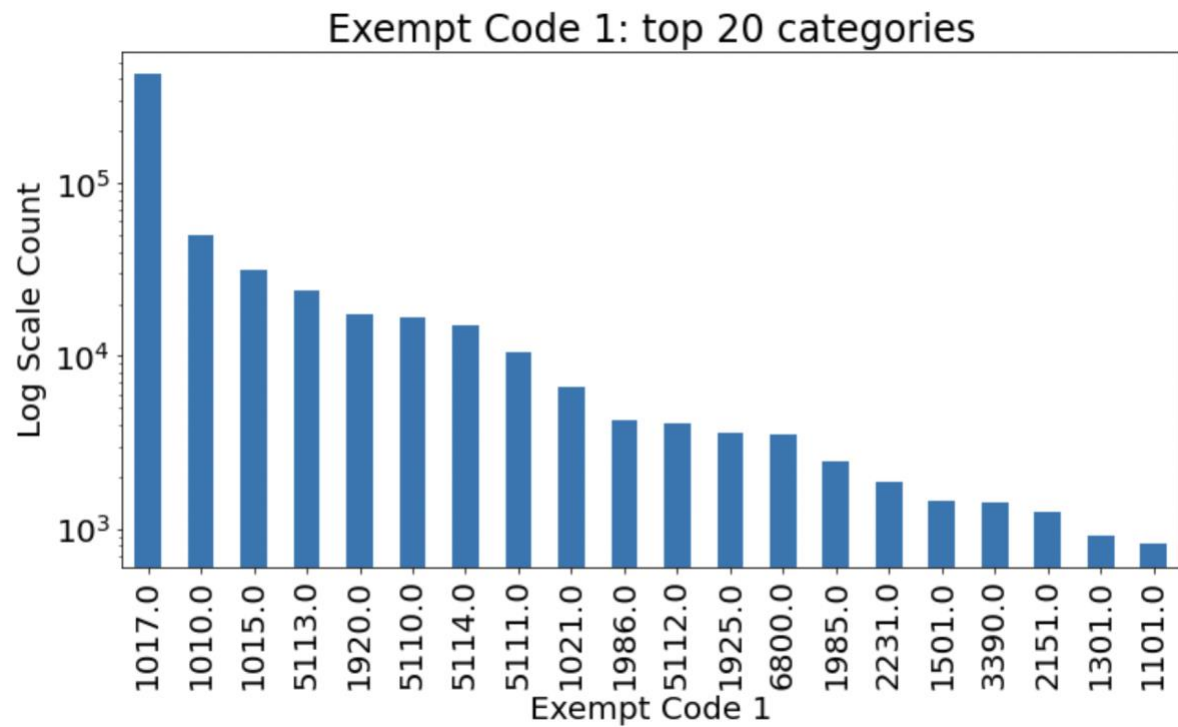
- This field represents actual exempt land total of each property.
- Also, we see a lot of entries with zero values, which might be representing missing values.

-The plot covers 85% of records.



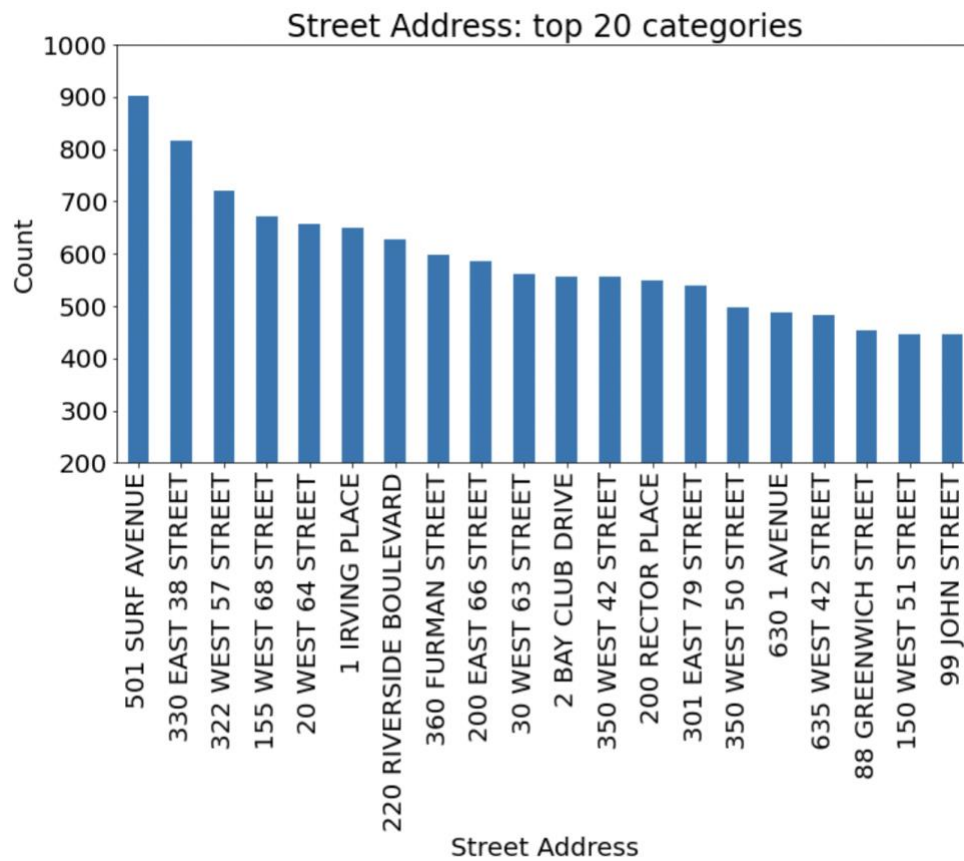
#### 19. EXCD1

-This field represents exempt code 1 of each property.



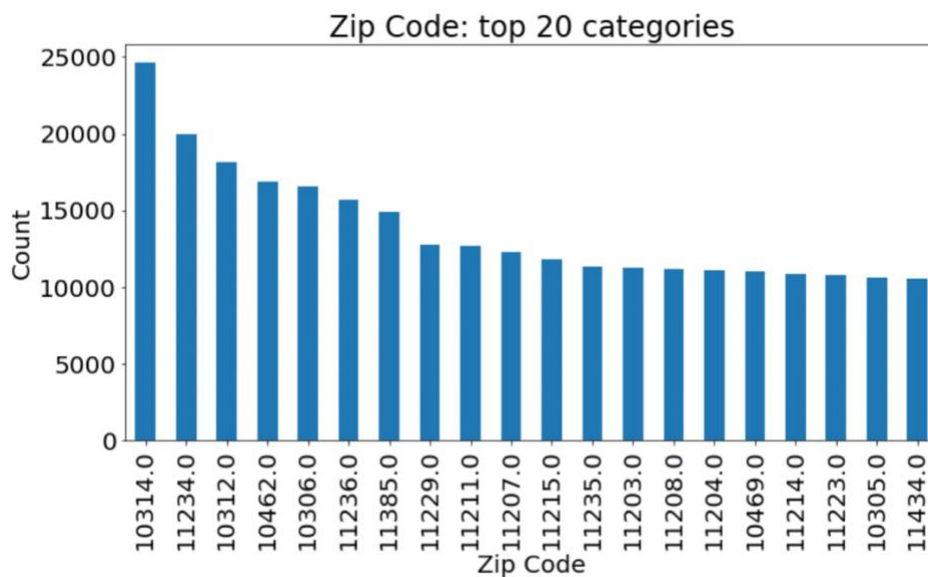
## 20. STADDR

-This field represents street address of each property.



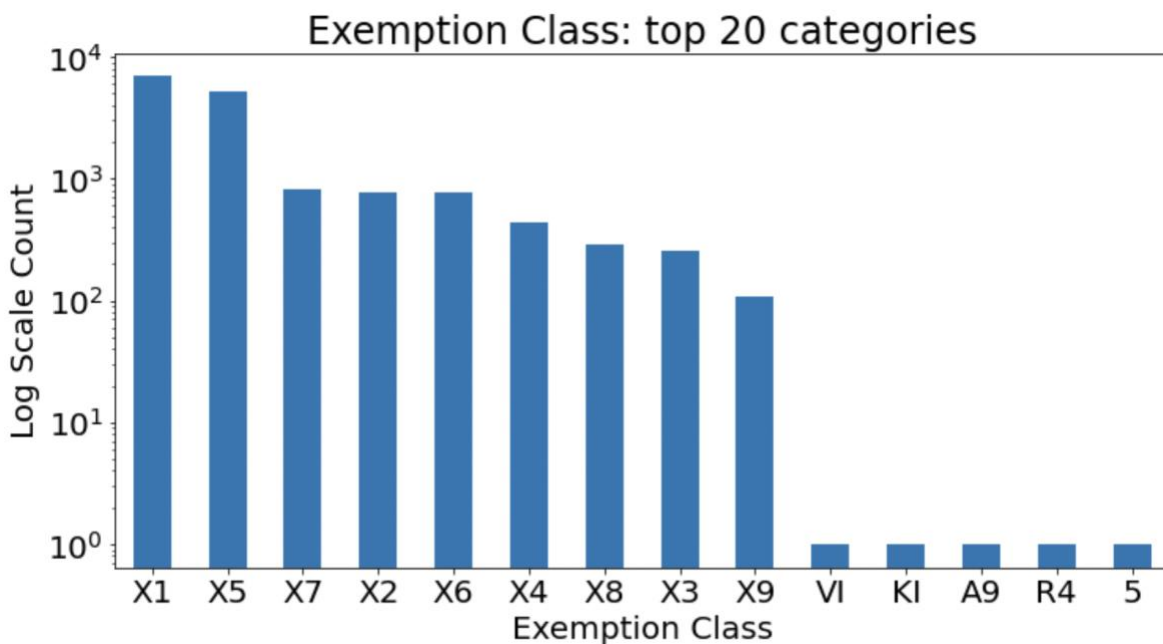
## 21. ZIP

-This field represents zip code of each property.



## 22. EXMPTCL

-This field represents exemption class of each property.

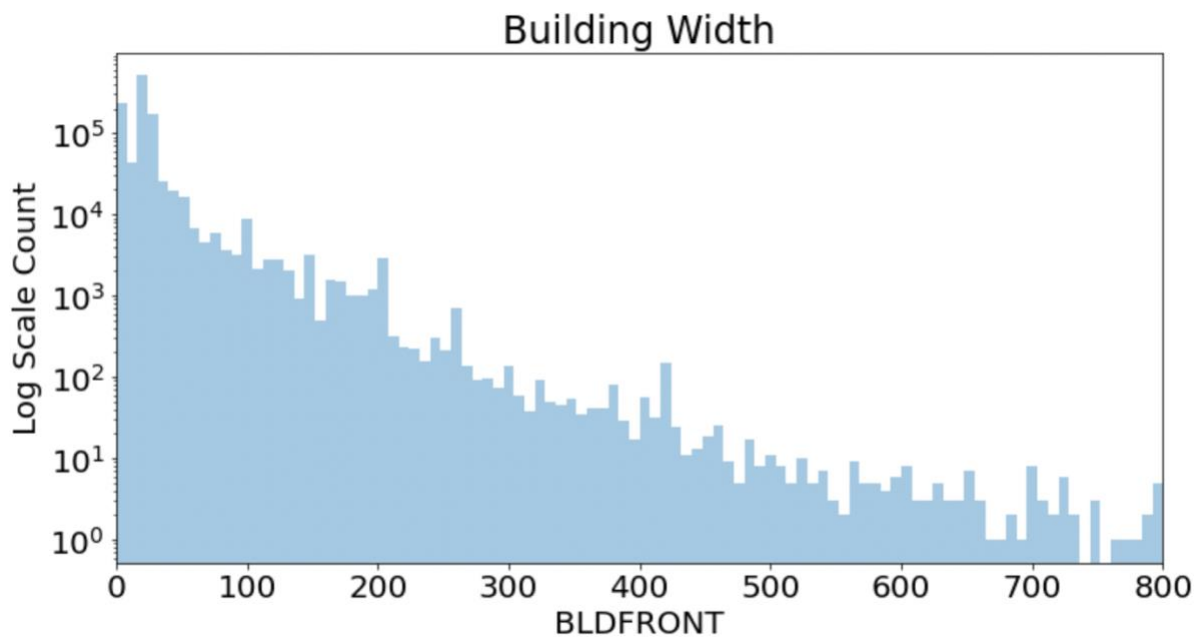


## 23. BLDFRONT

-This field represents building width of each property.

-Also, we see a lot of entries with zero values, which might be representing missing values.

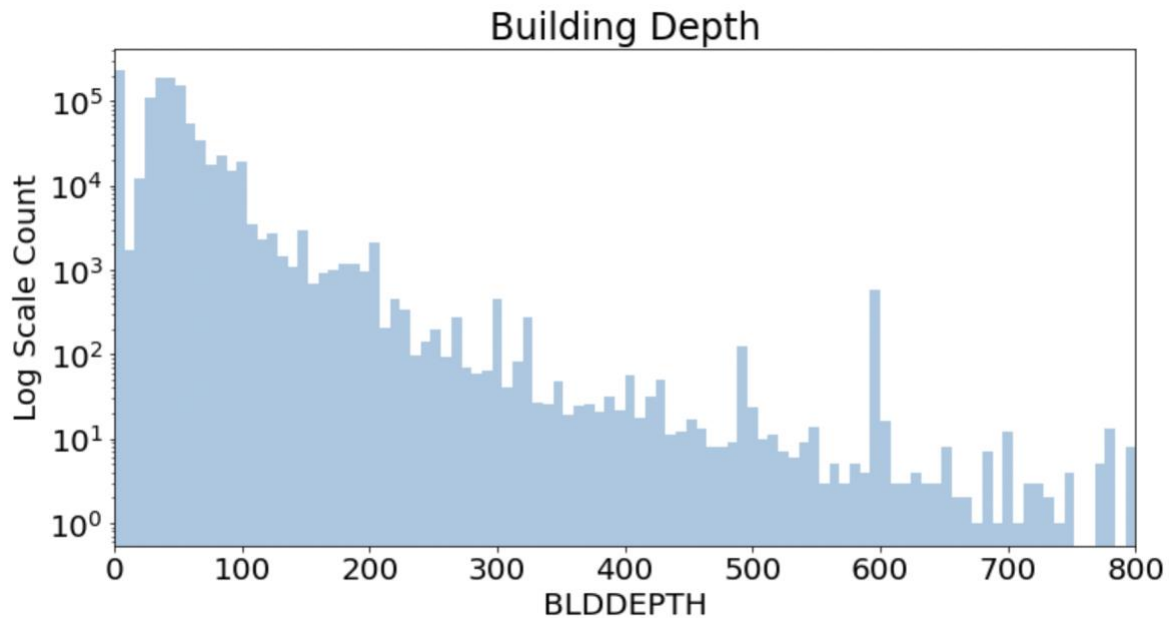
-The plot covers 100% of records.



## 24. BLDDEPTH

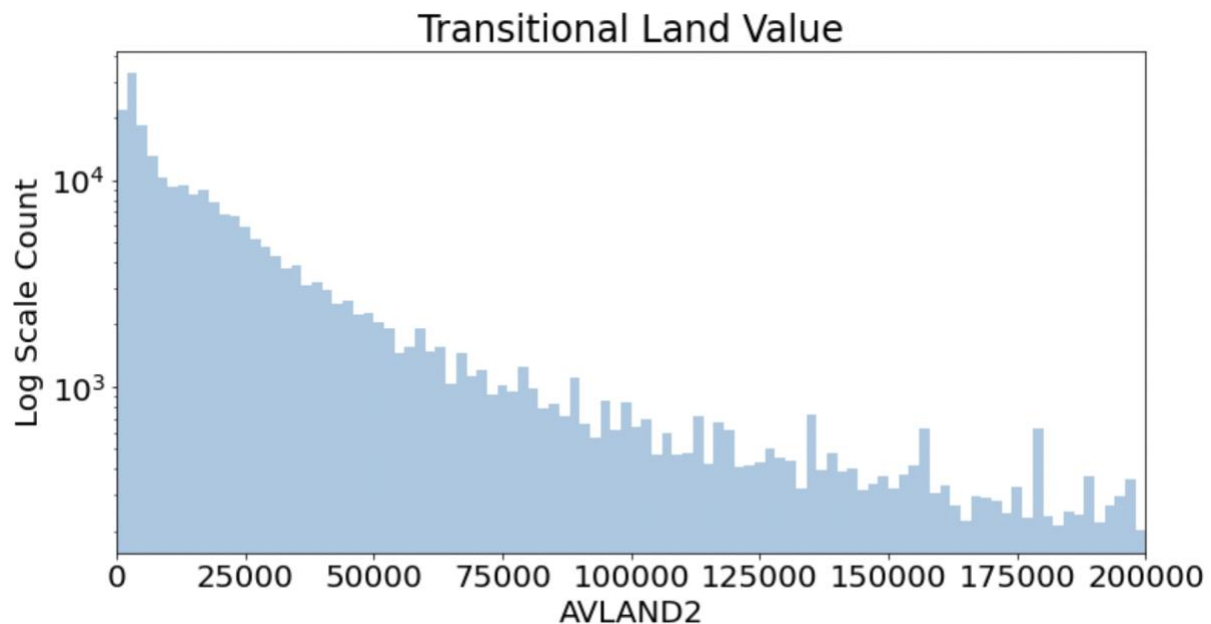
-This field represents building depth of each property.

- Also, we see a lot of entries with zero values, which might be representing missing values.
- The plot covers 100% of records.



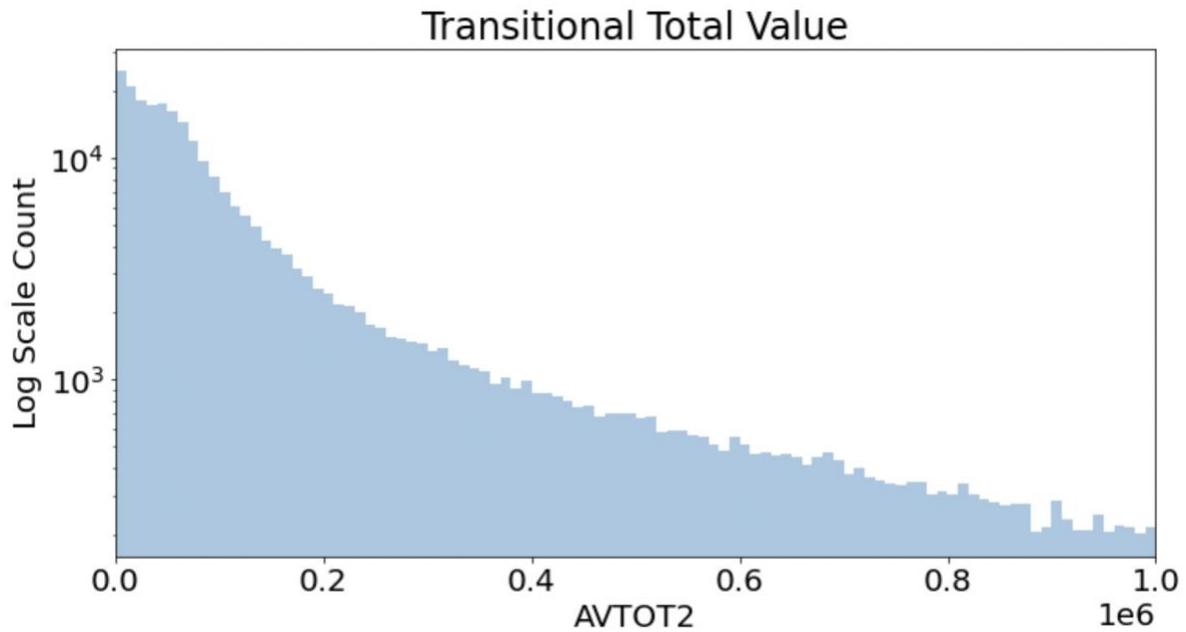
#### 25. AVLAND2

- This field represents transitional land value of each property.
- The plot covers 88% of records.



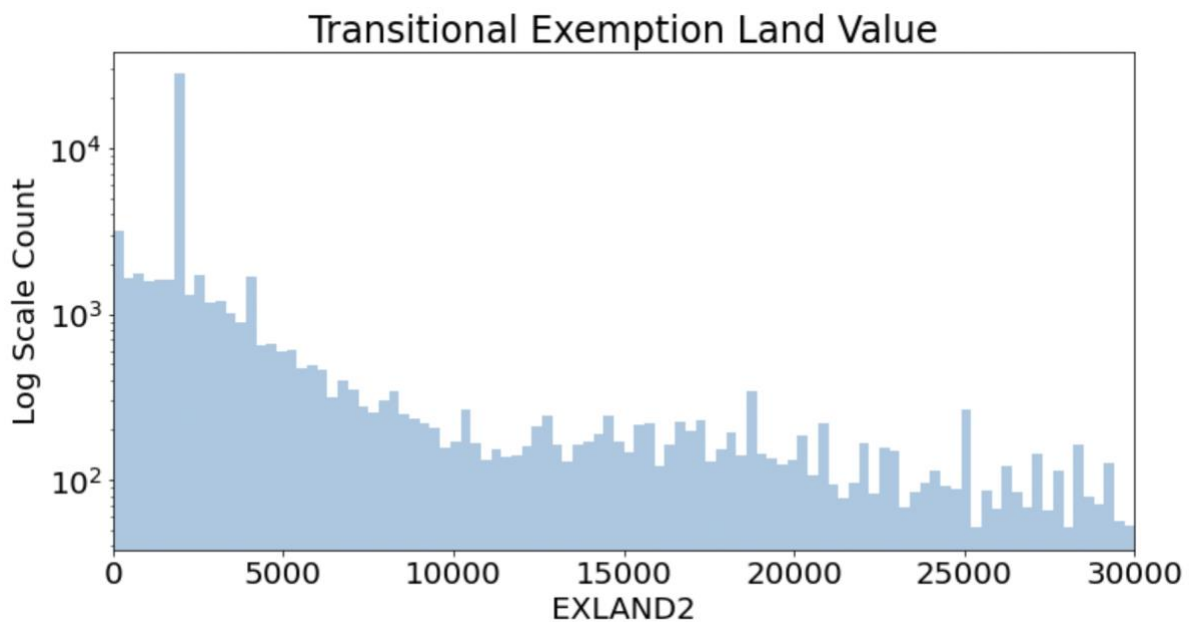
#### 26. AVTOT2

- This field represents transitional total value of each property.
- The plot covers 92% of records.



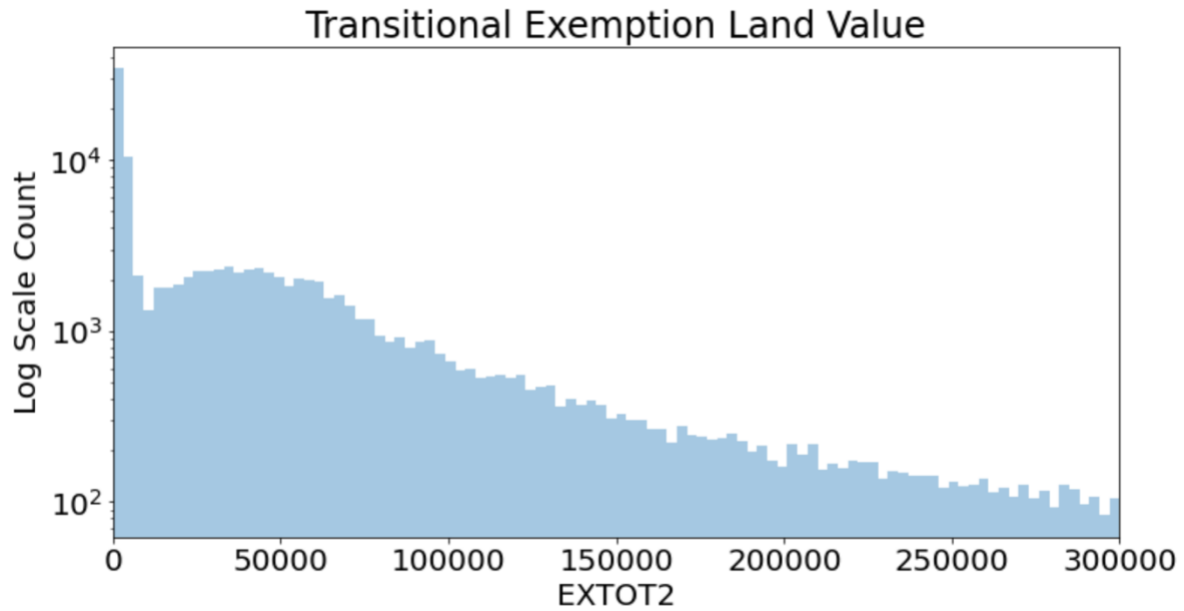
#### 27. EXLAND2

- This field represents transitional exemption land value of each property.
- The plot covers 74% of records.



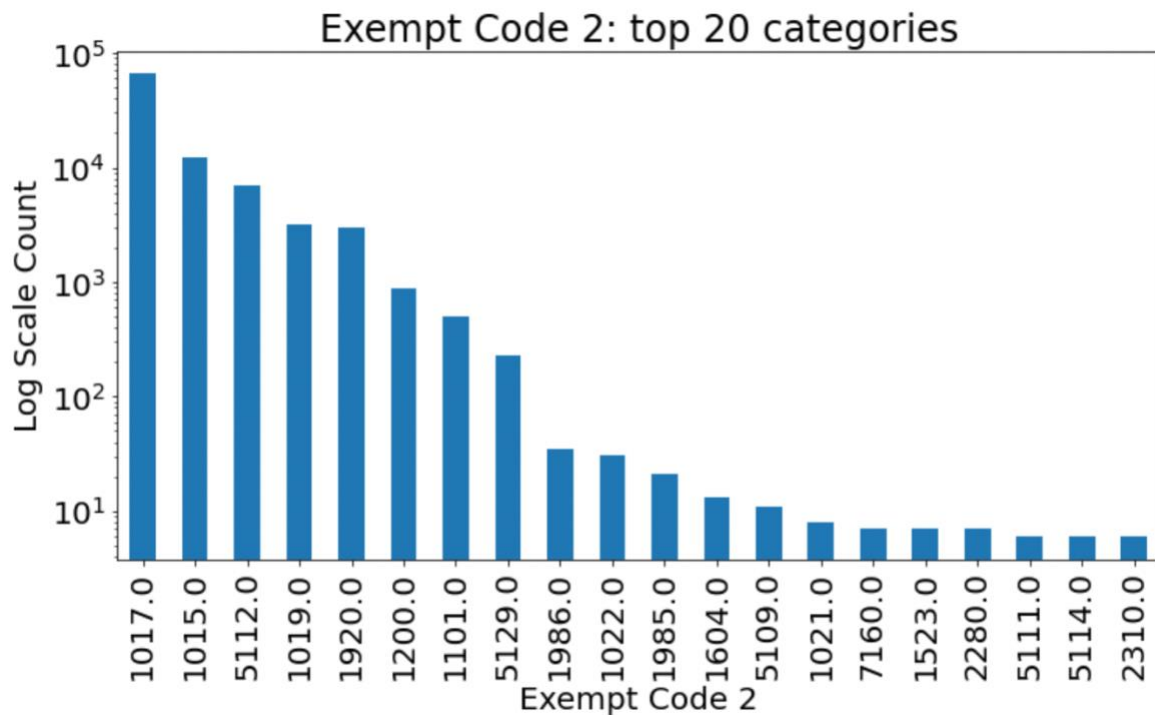
#### 28. EXTOT2

- This field represents transitional exemption land total of each property.
- The plot covers 86% of records.



#### 29. EXCD2

-This field represents Exemption Code 2 of each property.



#### 30. PERIOD

-This field represents assessment period of each property, with only one unique value "FINAL".

### 31. YEAR

-This field represents assessment year of each property, with only one unique value "2010/11".

### 32. VALTYPE

-This field represents validation type of each property, with only one unique value "AC-TR".