

DATA QUALITY REPORT

1. Dataset Description

Dataset Name: Property Valuation and Assessment Data

Dataset Source: Open data of NYC government

Dataset URL: <https://data.cityofnewyork.us/Housing-Development/Property-Valuation-and-Assessment-Data/rgy2-tti8>

Time Period: November 2010

No. of Fields: 32

No. of Records: 1,070,994

2. Summary

2.a) Numerical

| Field Name | # records that have a value | % populated | # unique values | # records with value 0 | Min | Max | Mean | Standard Deviation |
|------------|-----------------------------|-------------|-----------------|------------------------|-----|------------|-----------|--------------------|
| LTFRONT | 1070994 | 100.0% | 1297 | 169108 | 0 | 9999 | 36.63 | 74.03 |
| LTDEPTH | 1070994 | 100.0% | 1370 | 170128 | 0 | 9999 | 88.86 | 76.39 |
| STORIES | 1014730 | 94.7% | 112 | 0 | 1 | 119 | 5 | 8.36 |
| FULLVAL | 1070994 | 100.0% | 109324 | 13007 | 0 | 6150000000 | 874264.5 | 11582430.9 |
| AVLAND | 1070994 | 100.0% | 70921 | 13009 | 0 | 2668500000 | 85067.91 | 4057260.05 |
| AVTOT | 1070994 | 100.0% | 112914 | 13007 | 0 | 4668308947 | 227238.16 | 6877529.3 |
| EXLAND | 1070994 | 100.0% | 33419 | 491699 | 0 | 2668500000 | 36423.89 | 3981575.79 |
| EXTOT | 1070994 | 100.0% | 64255 | 432572 | 0 | 4668308947 | 91186.98 | 6508402.81 |
| BLDFRONT | 1070994 | 100.0% | 612 | 228815 | 0 | 7575 | 23.04 | 35.5 |
| BLDDEPTH | 1070994 | 100.0% | 621 | 228853 | 0 | 9393 | 39.92 | 42.7 |
| AVLAND2 | 282726 | 26.4% | 58592 | 0 | 3 | 2371005000 | 246235.71 | 6178962.56 |
| AVTOT2 | 282732 | 26.4% | 111361 | 0 | 3 | 4501180002 | 713911.43 | 11652528.94 |
| EXLAND2 | 87449 | 8.2% | 22196 | 0 | 1 | 2371005000 | 351235.68 | 10802212.67 |
| EXTOT2 | 130828 | 12.2% | 48349 | 0 | 7 | 4501180002 | 656768.28 | 16072510.17 |

2.b) Categorical

| Field Name | # records that have a value | % populated | # unique values | Most common value |
|------------|-----------------------------|-------------|-----------------|--------------------------|
| RECORD | 1070994 | 100.0% | 1070994 | NA |
| BBLE | 1070994 | 100.0% | 1070994 | NA |
| B | 1070994 | 100.0% | 5 | 4 |
| BLOCK | 1070994 | 100.0% | 13984 | 3944 |
| LOT | 1070994 | 100.0% | 6366 | 1 |
| EASEMENT | 4636 | 0.4% | 13 | E |
| OWNER | 1039249 | 97.0% | 863347 | PARKCHESTER PRESERVAT |
| BLDGCL | 1070994 | 100.0% | 200 | R4 |
| TAXCLASS | 1070994 | 100.0% | 11 | 1 |
| EXT | 354305 | 33.1% | 4 | G |
| EXCD1 | 638488 | 59.6% | 130 | 1017 |
| STADDR | 1070318 | 99.9% | 839281 | 501 SURF AVENUE |
| ZIP | 1041104 | 97.2% | 197 | 10314 |
| EXMPTCL | 15579 | 1.5% | 15 | X1 |
| EXCD2 | 92948 | 8.7% | 61 | 1017 |
| PERIOD | 1070994 | 100.0% | 1 | FINAL |
| YEAR | 1070994 | 100.0% | 1 | 2010/11 |
| VALTYPE | 1070994 | 100.0% | 1 | AC-TR |

3. DATA FIELD EXPLORATION

Field 1

Field Name: RECORD

Description: Unique integer for each record

Field 2

Field Name: BBLE

Description: Concatenation of Borough code, block code, unique number

Field 3

Field Name: B

Description: Borough Codes

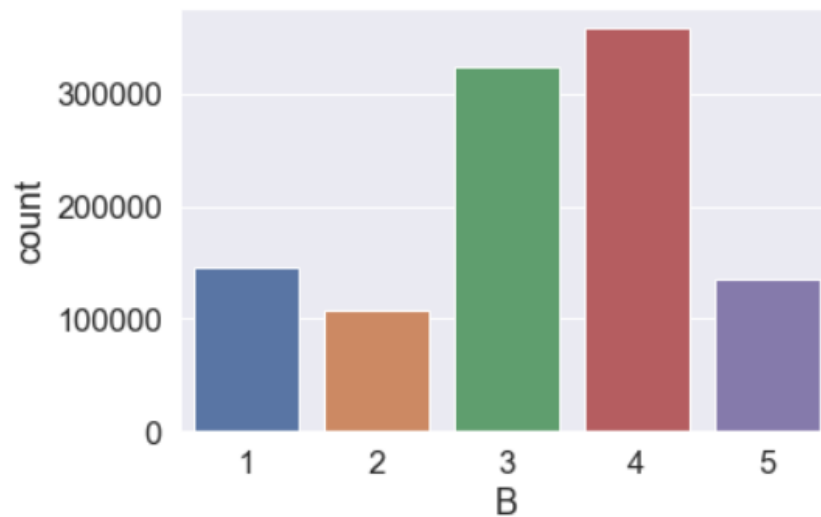


Fig. 1. Histogram plot showing distribution of field 'B'

Field 4**Field Name:** BLOCK**Description:** Valid block ranges by borough codes

| BLOCK | COUNT |
|-------|-------|
| 3944 | 3838 |
| 16 | 3786 |
| 3943 | 3424 |
| 3938 | 2794 |
| 1171 | 2535 |
| 3937 | 2275 |
| 1833 | 1774 |
| 2450 | 1651 |
| 1047 | 1480 |
| 7279 | 1302 |
| 5893 | 1295 |
| 8720 | 1281 |
| 936 | 1151 |
| 1115 | 1090 |
| 1320 | 1049 |
| 1140 | 1017 |

*Table 1. Top 15 Field Values***Field 5****Field Name:** LOT**Description:** Unique # Within Borough/Block

| LOT | COUNT |
|-----|-------|
| 1 | 24367 |
| 20 | 12294 |
| 15 | 12171 |
| 12 | 12143 |
| 14 | 12074 |
| 16 | 12042 |
| 17 | 11982 |
| 18 | 11979 |
| 25 | 11949 |
| 21 | 11840 |
| 23 | 11705 |
| 22 | 11665 |
| 6 | 11646 |
| 19 | 11640 |
| 30 | 11596 |

Table 2. Top 15 Field Values

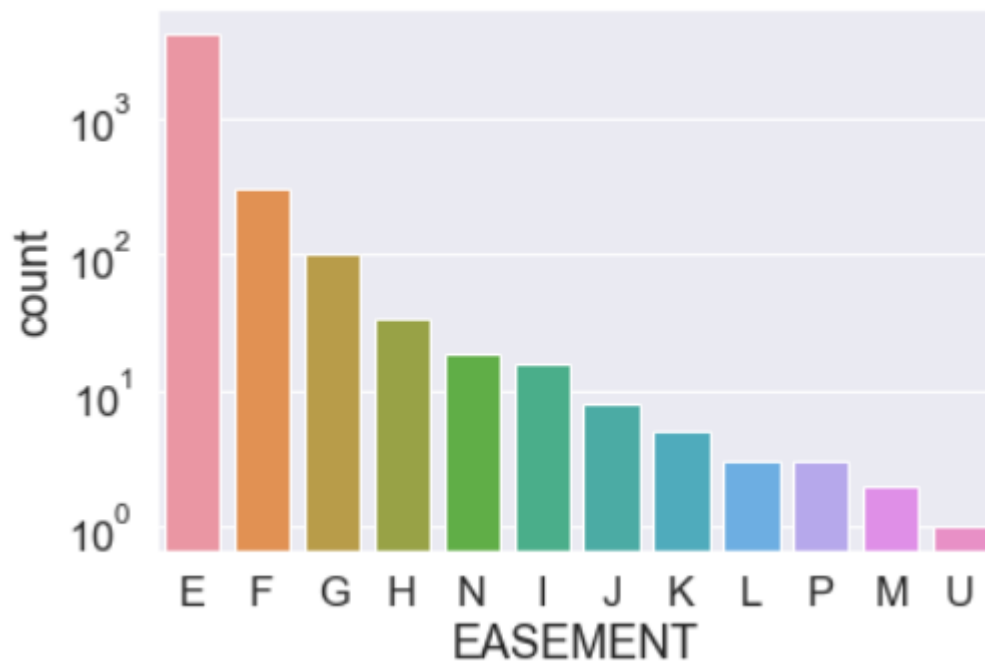
Field 6**Field Name:** EASEMENT**Description:** A field that is used to describe easement

Fig. 2. Histogram plot showing distribution of field 'EASEMENT'

Field 7**Field Name:** OWNER**Description:** Owner's Name

| OWNER | COUNT |
|-----------------------|-------|
| PARKCHESTER PRESERVAT | 6020 |
| PARKS AND RECREATION | 4255 |
| DCAS | 2169 |
| HOUSING PRESERVATION | 1904 |
| CITY OF NEW YORK | 1450 |
| DEPT OF ENVIRONMENTAL | 1166 |
| BOARD OF EDUCATION | 1015 |
| NEW YORK CITY HOUSING | 1014 |
| CNY/NYCTA | 975 |
| NYC HOUSING PARTNERSH | 747 |

Table 3. Top 10 Field Values

Field 8

Field Name: BLDGCL

Description: Building class where position 1 = ALPHA & position 2 = NUMERIC. There is a direct correlation between the Building Class and the Tax Class.

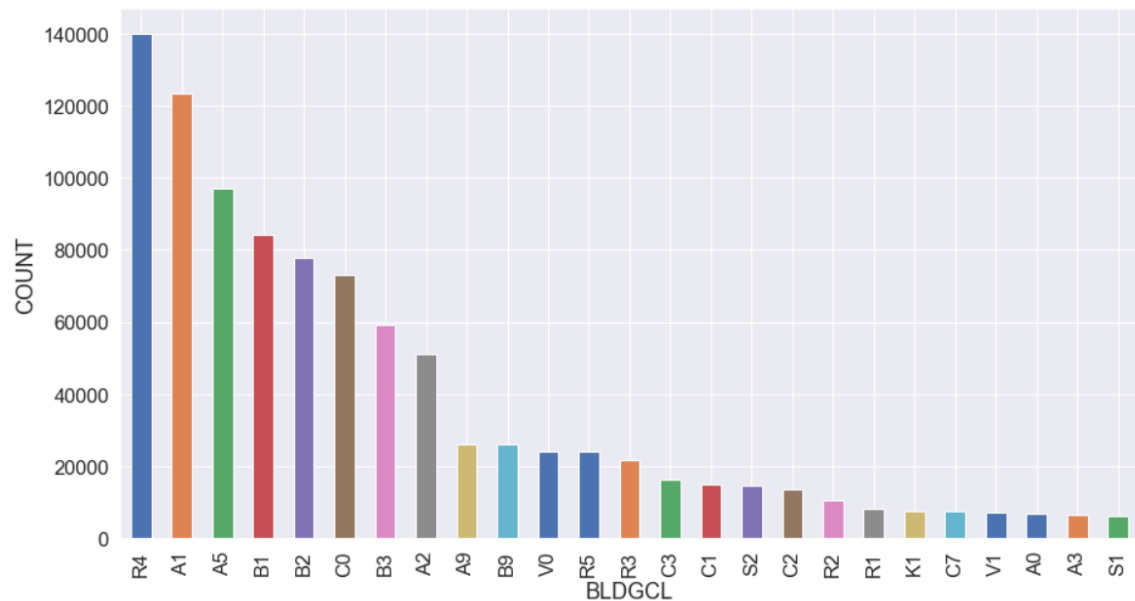


Fig. 3. Histogram plot showing distribution of field 'BLDGCL'

Field 9

Field Name: TAXCLASS

Description: Current Property Tax Class Code (NYS Classification)

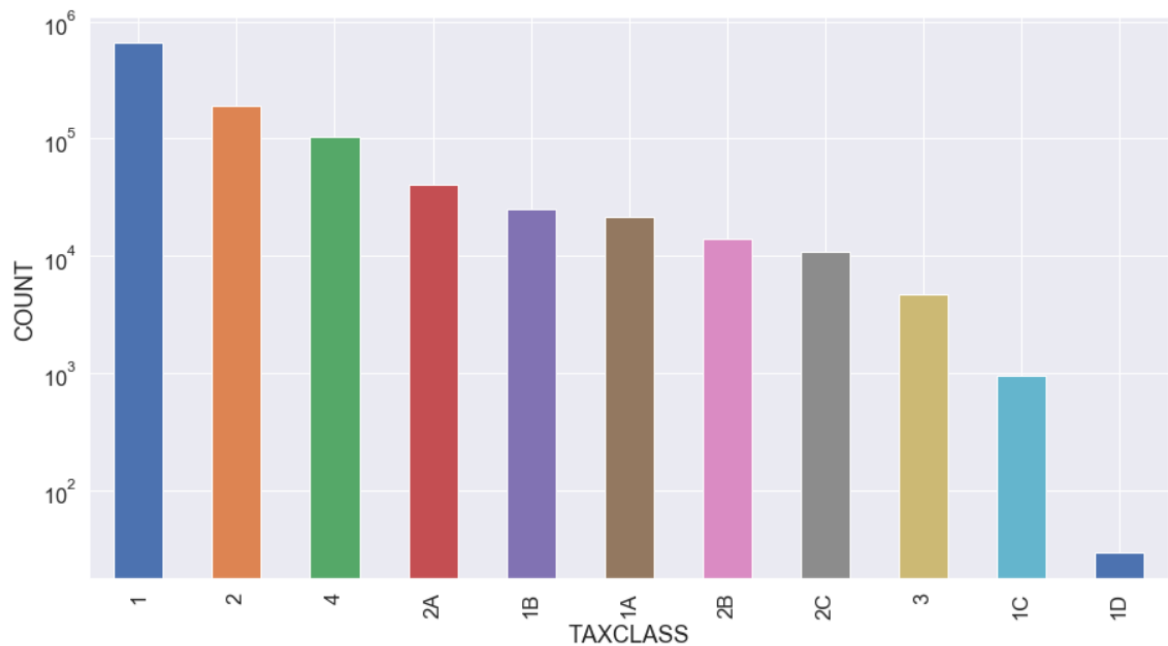
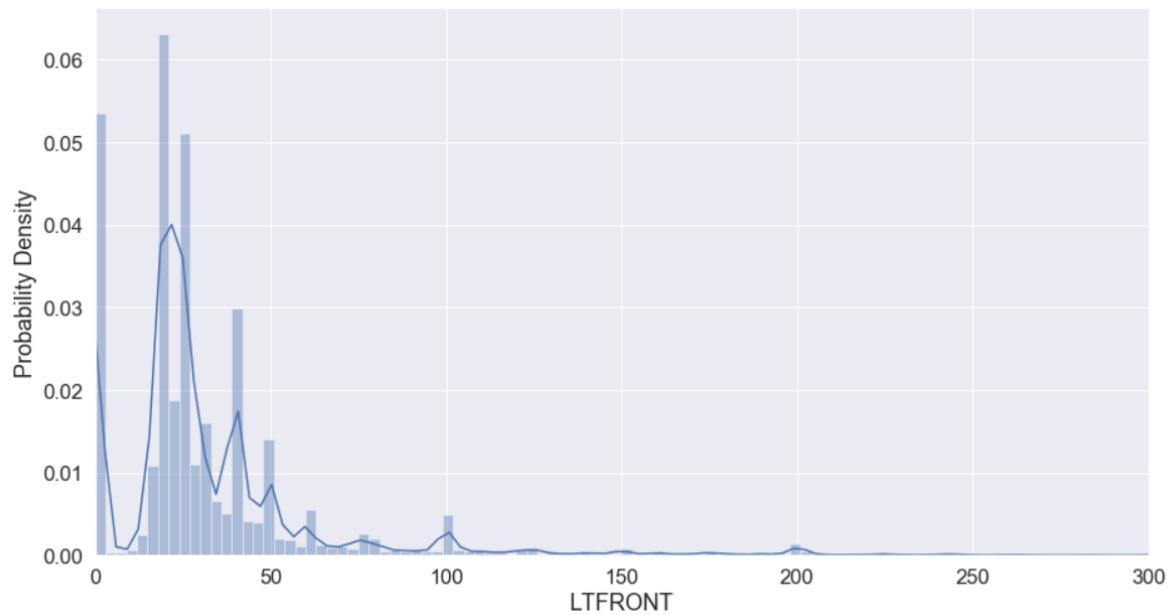


Fig. 4. Histogram plot showing distribution of field 'TAXCLASS'

Field 10

Field Name: LTFRONT

Description: Lot Frontage in feet

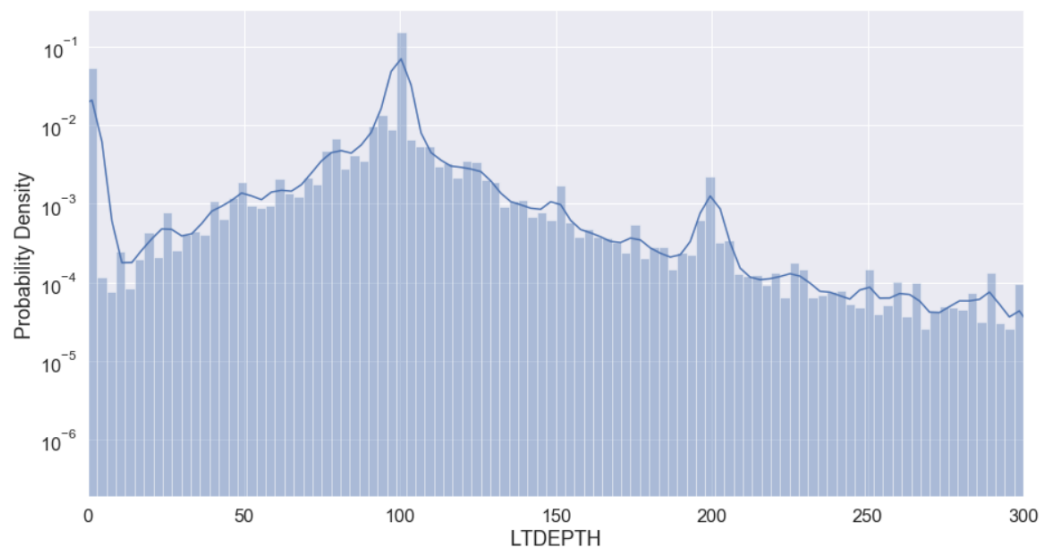


*Fig. 5. Kernel Density Distribution plot showing distribution of field 'LTFRONT'
(Outliers > 300 removed)*

Field 11

Field Name: LTDEPTH

Description: Lot Depth in feet



*Fig. 6. Kernel Density Distribution plot showing distribution of field 'LTDEPTH'
(Outliers > 300 removed)*

Field 12

Field Name: EXT

Description: E-Extension, G- Garage, EG- Extension and Garage

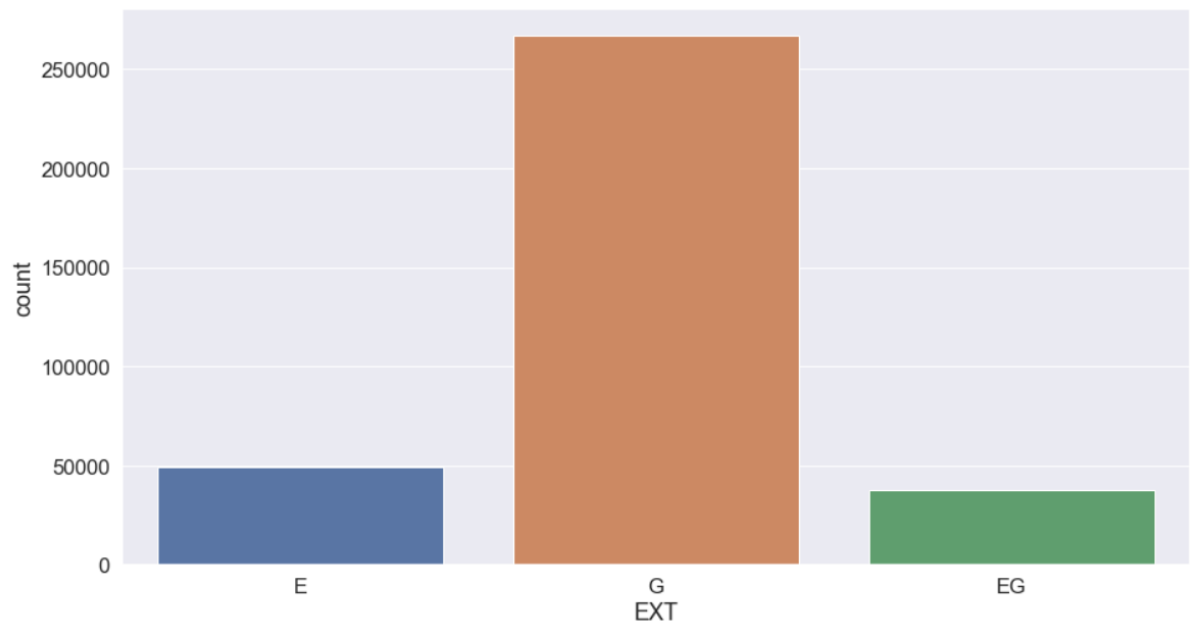


Fig. 7. Histogram plot showing distribution of field 'EXT'

Field 13

Field Name: STORIES

Description: Number of floors in the building

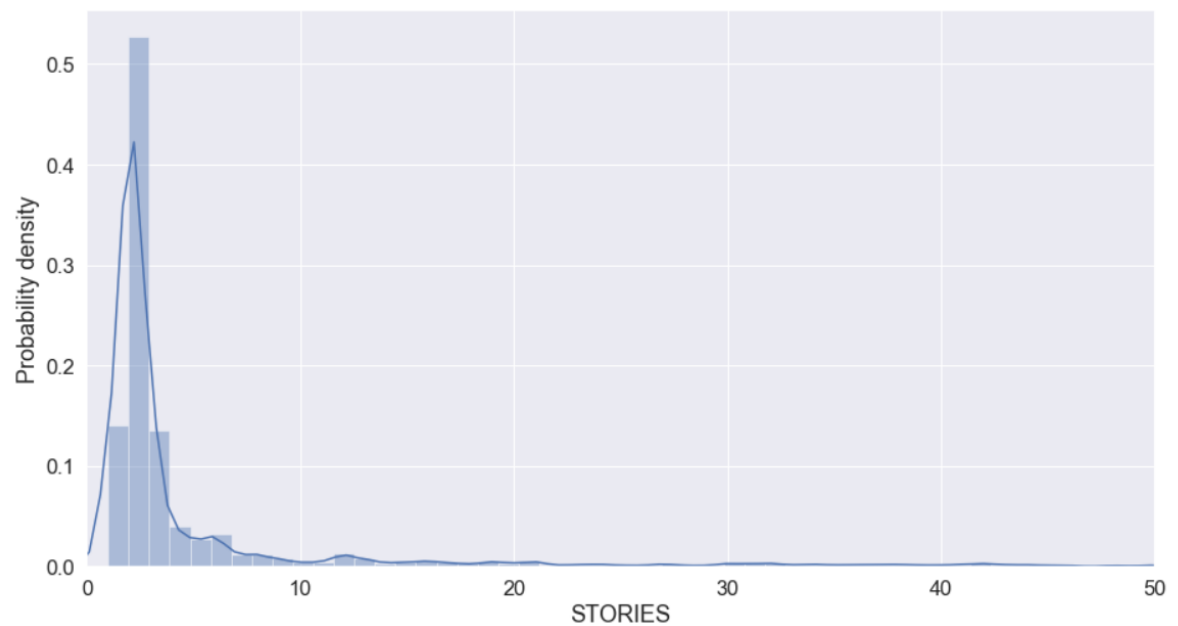


Fig. 8. Kernel Density Distribution plot showing distribution of field 'STORIES'

(Outliers > 50 removed)

Field 14

Field Name: FULLVAL

Description: Total Market Value

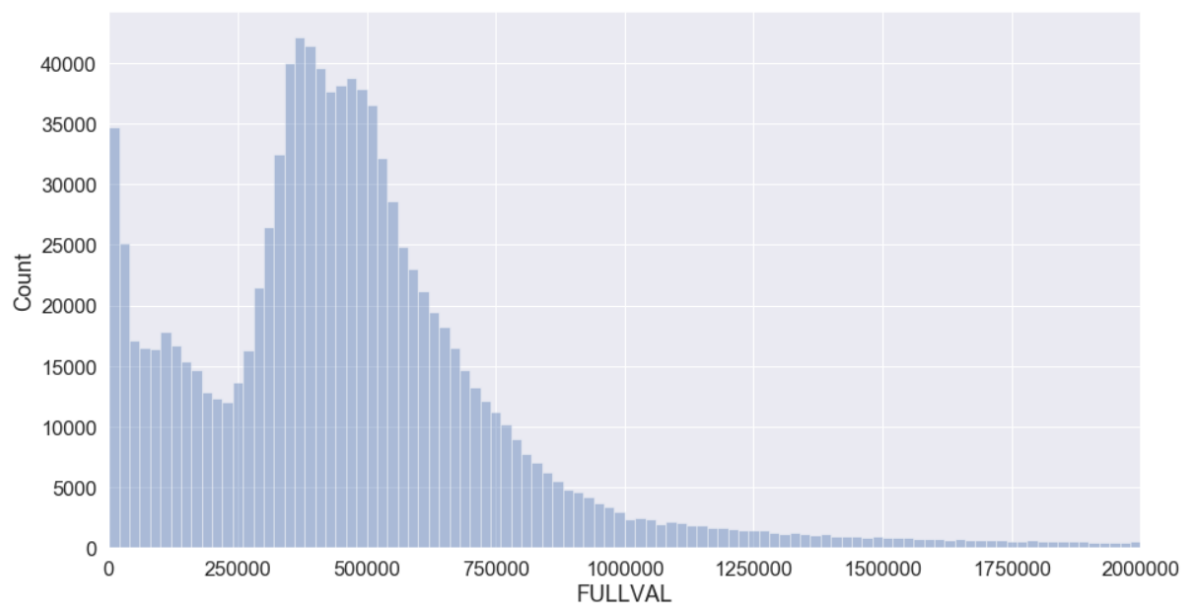


Fig. 9. Distribution plot showing distribution of field 'FULLVAL'

(Outliers > 2000000 removed)

Field 15

Field Name: AVLAND

Description: Market value of land

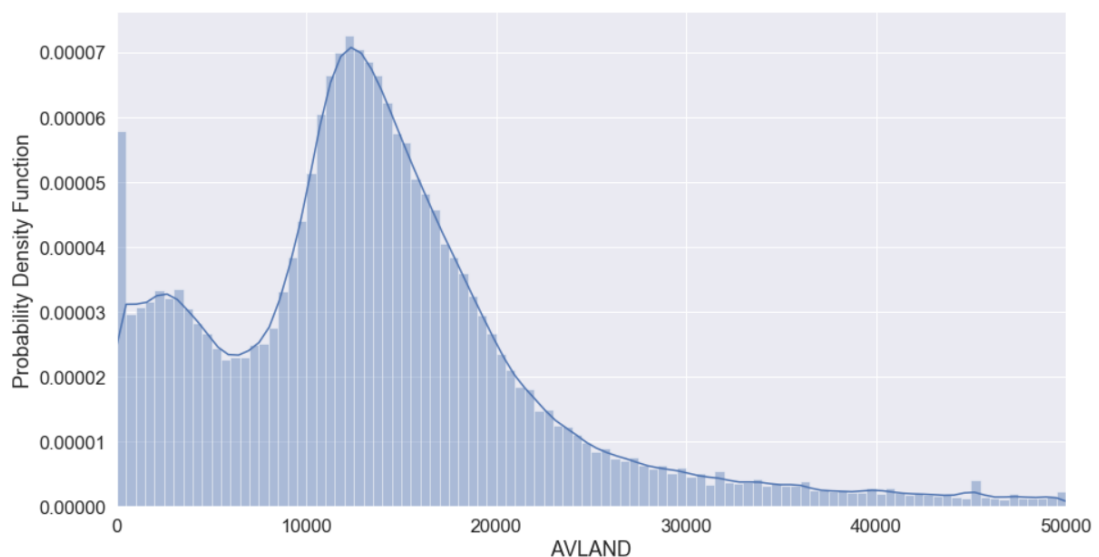


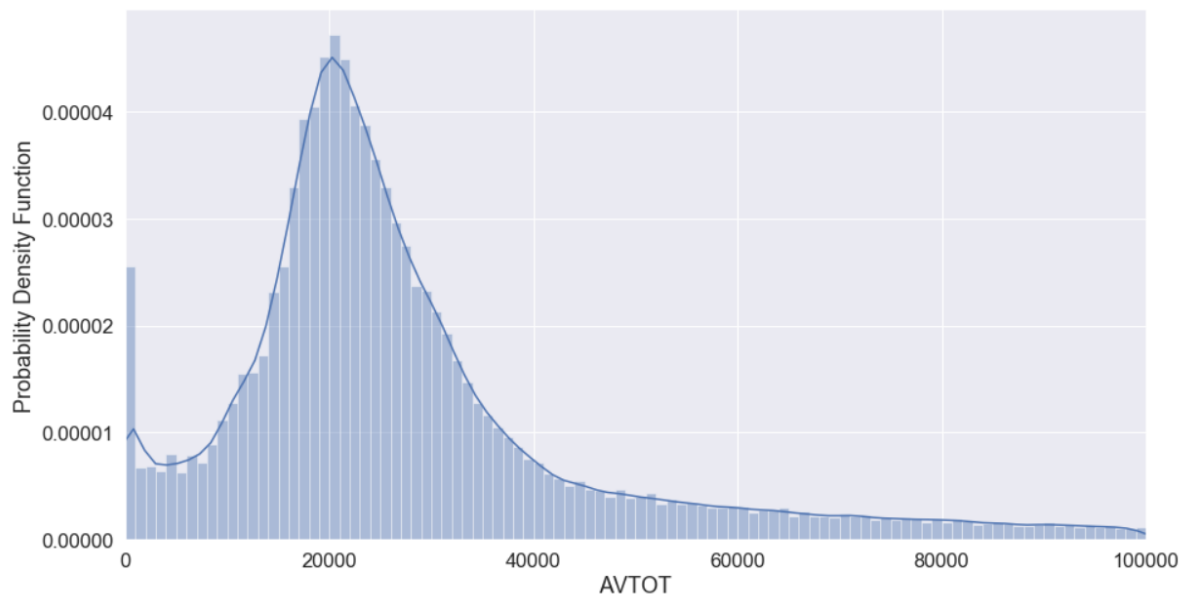
Fig. 10. Kernel Density Distribution plot showing distribution of field 'AVLAND'

(Outliers > 50000 removed)

Field 16

Field Name: AVTOT

Description: Total market value

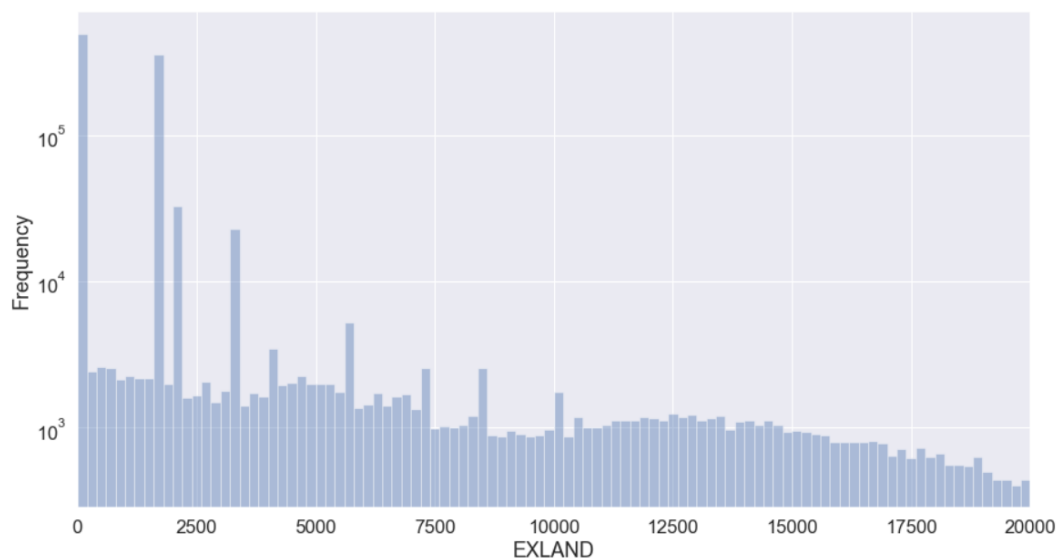


*Fig. 11. Kernel Density Distribution plot showing distribution of field 'AVTOT'
(Outliers > 100,000 removed)*

Field 17

Field Name: EXLAND

Description: Tentative transitional exempt land value



*Fig. 12. Distribution plot showing distribution of field 'EXLAND'
(Outliers > 20,000 removed)*

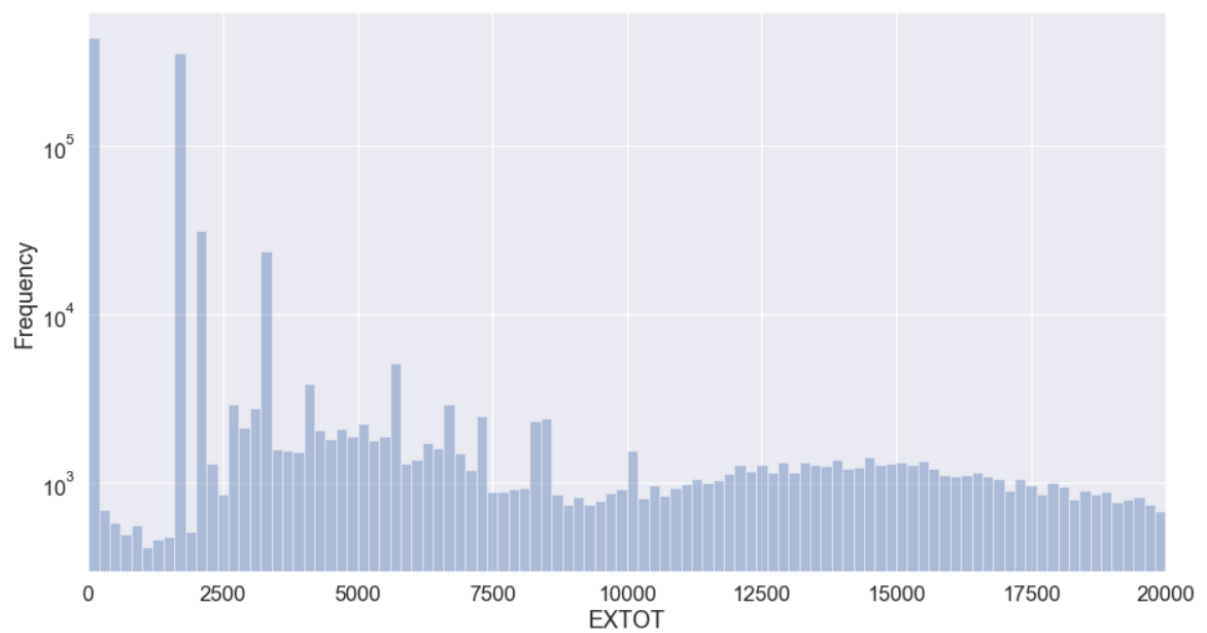
Field 18**Field Name:** EXTOT**Description:** Tentative transitional exempt total value

Fig. 13. Distribution plot showing distribution of field 'EXTOT'
(Outliers > 20,000 removed)

Field 19**Field Name:** EXCD1**Description:**

| EXCD1 | COUNT |
|-------|--------|
| 1017 | 425348 |
| 1010 | 49756 |
| 1015 | 31323 |
| 5113 | 23858 |
| 1920 | 17594 |
| 5110 | 16834 |
| 5114 | 14984 |
| 5111 | 10609 |
| 1021 | 6613 |
| 1986 | 4231 |

Table 4. Top 10 Field Values

Field 20**Field Name:** STADDR**Description:** Street name of property

| STREET ADDRESS | COUNT |
|--------------------|-------|
| 501 SURF AVENUE | 902 |
| 330 EAST 38 STREET | 817 |
| 322 WEST 57 STREET | 720 |
| 155 WEST 68 STREET | 671 |
| 20 WEST 64 STREET | 657 |
| 1 IRVING PLACE | 650 |
| 628 | |
| 599 | |
| 585 | |
| 562 | |

*Table 5. Top 10 Field Values***Field 21****Field Name:** ZIP**Description:** Postal ZIP code of the property

| ZIP | COUNT |
|-------|-------|
| 10314 | 24606 |
| 11234 | 20001 |
| 10312 | 18127 |
| 10462 | 16905 |
| 10306 | 16578 |
| 11236 | 15678 |
| 11385 | 14921 |
| 11229 | 12793 |
| 11211 | 12710 |
| 11207 | 12293 |

Table 6. Top 10 Field Values

Field 22

Field Name: EXMPTCL

Description: Exempt class used for fully exempt properties only

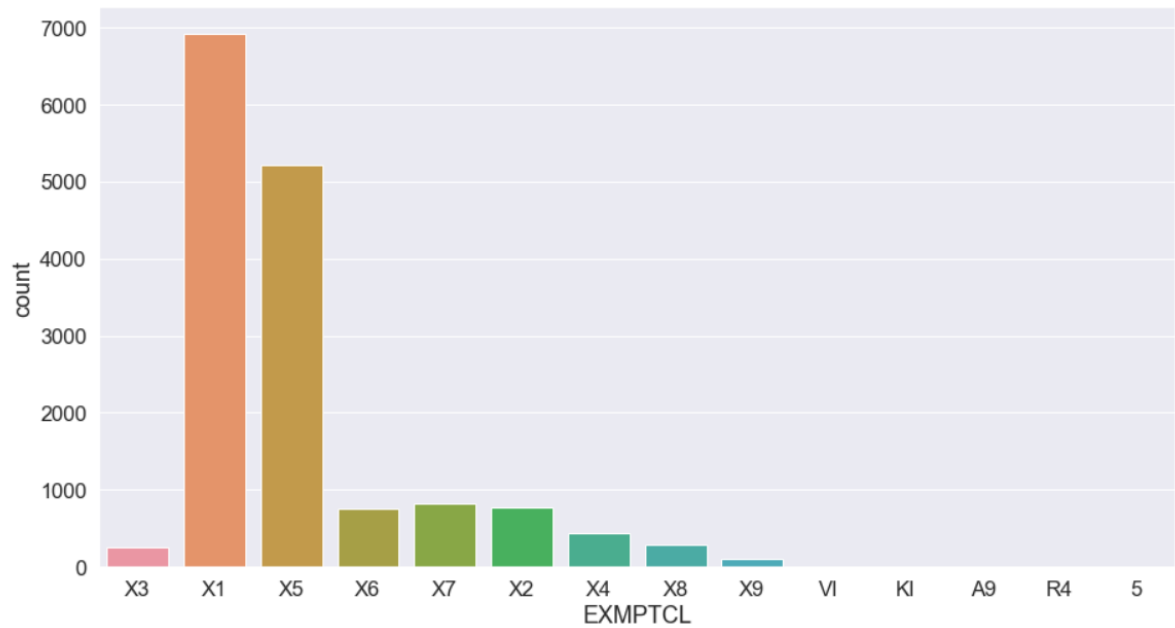


Fig. 14. Histogram plot showing distribution of field 'EXMPTCL'

Field 23

Field Name: BLDFRONT

Description: Building frontage in feet

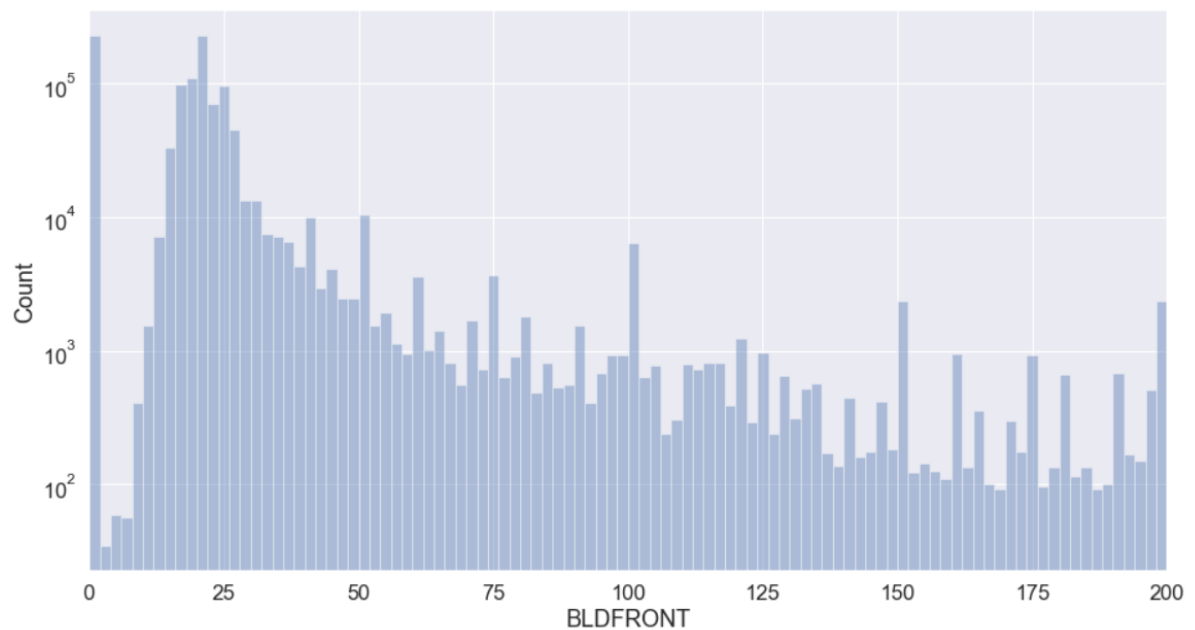


Fig. 15. Distribution plot showing distribution of field 'BLDFRONT'

(Outliers > 200 removed)

Field 24

Field Name: BLDDEPTH

Description: Lot depth in feet

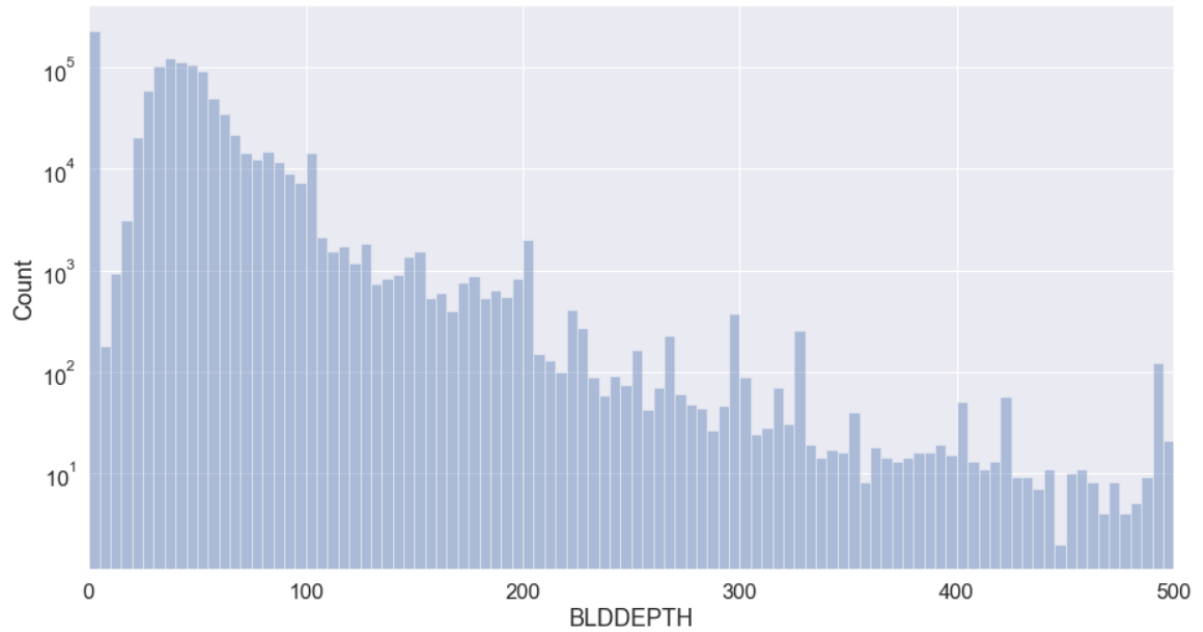


Fig. 16. Distribution plot showing distribution of field 'BLDDEPTH'

(Outliers > 500 removed)

Field 25

Field Name: AVLAND2

Description: New market value of land

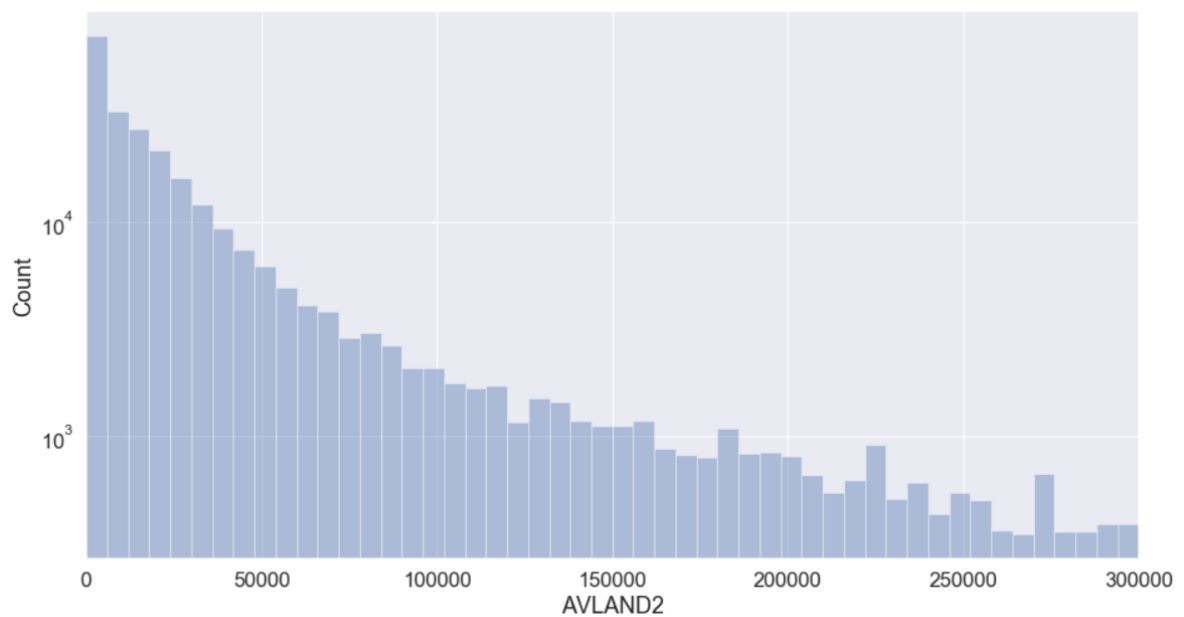


Fig. 17. Histogram plot showing distribution of field 'AVLAND2'

(Outliers > 300,000 removed)

Field 26

Field Name: AVTOT2

Description: New total market value

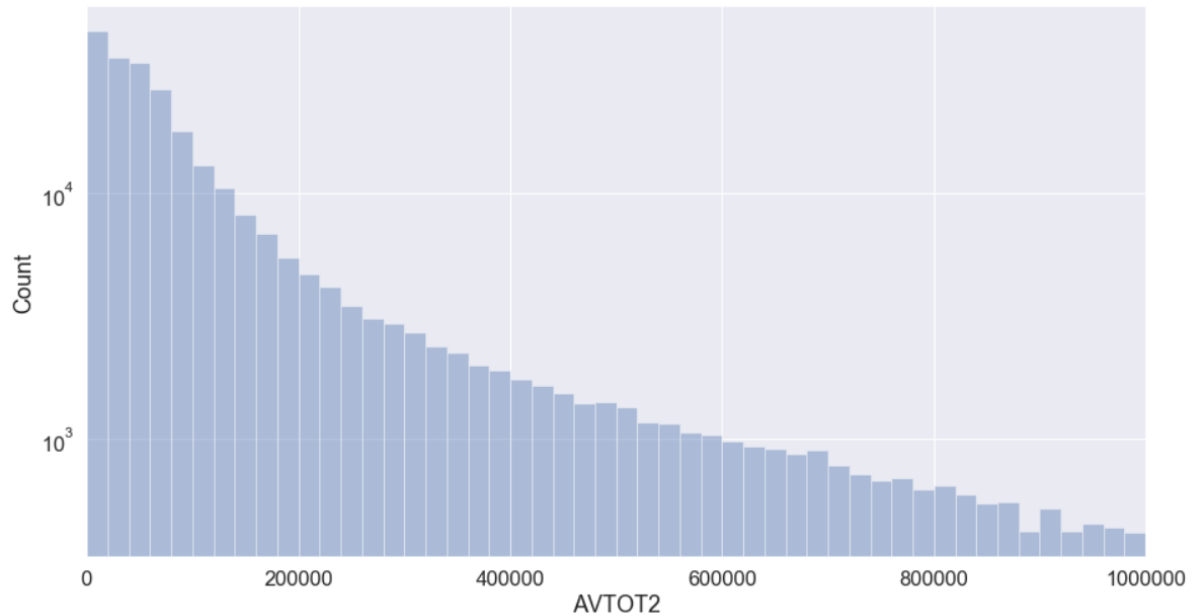


Fig. 18. Histogram plot showing distribution of field 'AVTOT2'

(Outliers > 10,000,000 removed)

Field 27

Field Name: EXLAND2

Description: New Tentative transitional exempt land value

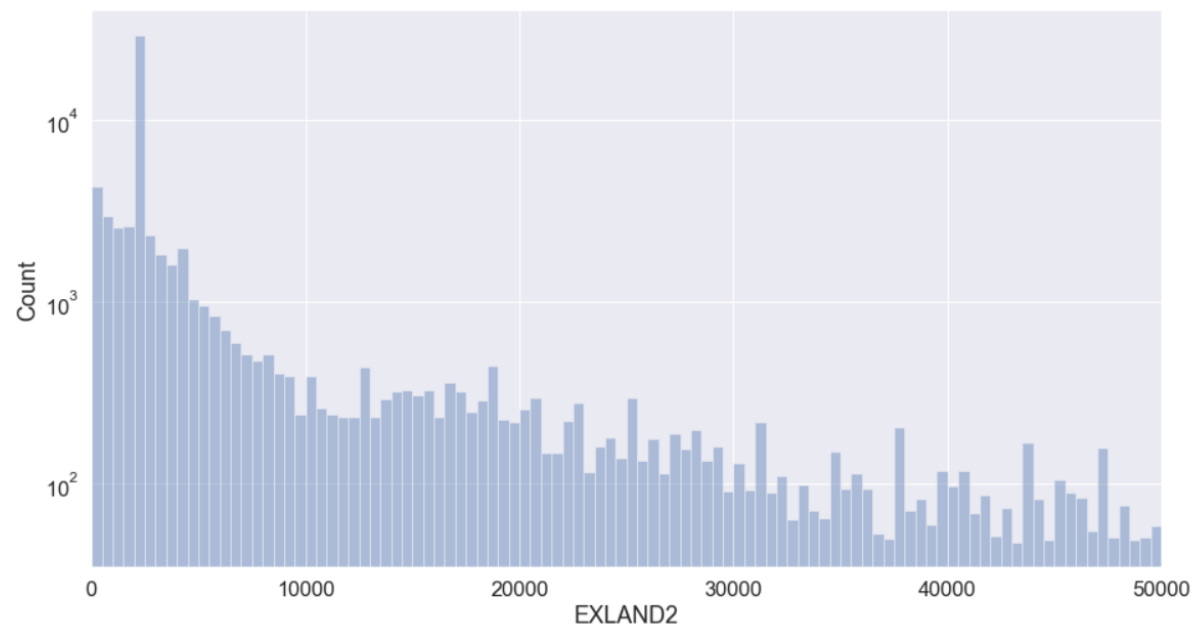


Fig. 19. Histogram plot showing distribution of field 'EXLAND2'

Field 28

Field Name: EXTOT2

Description: New Tentative transitional exempt total value

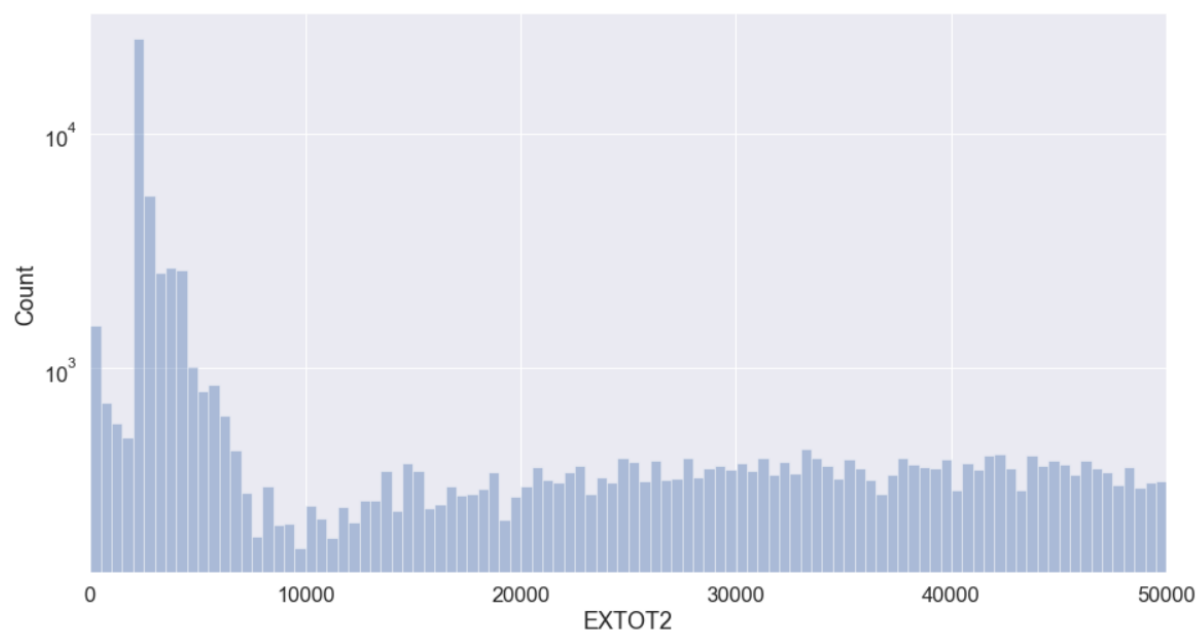


Fig. 20. Distribution plot showing distribution of field 'EXTOT2'

(Outliers > 50,000 removed)

Field 29

Field Name: EXCD2

Description:

| EXCD2 | COUNT |
|-------|-------|
| 1017 | 65777 |
| 1015 | 12337 |
| 5112 | 6867 |
| 1019 | 3178 |
| 1920 | 2961 |
| 1200 | 881 |
| 1101 | 494 |
| 5129 | 227 |
| 1986 | 35 |
| 1022 | 31 |

Table 7. Top 10 Field Values

Field 30**Field Name:** Period**Description:** Final - (Only one unique value)**Field 31****Field Name:** Year**Description:** Four-digit year of the file (2010/11) - (Only one unique value)**Field 32****Field Name:** Valtype**Description:** AC-TR (Only one unique value)