

Name: Yi Liu; USC ID: 8666552780

Record 917,942

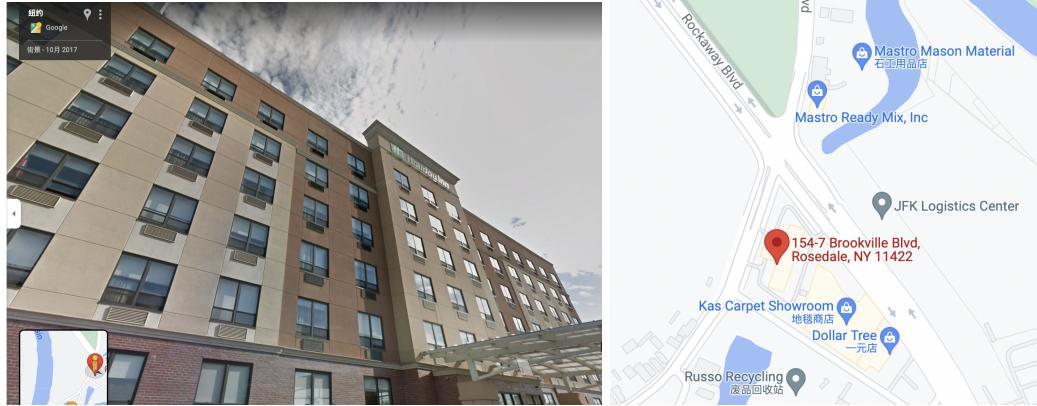


Table 1: Original data record

RECORD	917,942	BLDFRONT	62.08
BORO	4	BLDDEPTH	86.3
ZIP	11,422	LTFRONT	4,910.00
STADDR	154-68 BROOKVILLE BOULEVARD	LTDEPTH	124.44
OWNER	LOGAN PROPERTY, INC.	AVLAND	1,792,808,947.00
TAXCLASS		AVTOT	4,668,308,947.00
STORIES	3	FULLVAL	374,019,883.00

Table 2: Highest scoring records

	r	r_zip3	r_boro
r5	884.03649	900.242759	901.945542
r6	625.448873	606.715597	591.075199
r8	934.683271	971.12882	968.469316
r9	888.115885	890.753853	878.64411

According to the street picture from google map, the property has six floors, instead of 3 in the profile, so I suspect fake building volume will lead to fraudulence.

As for z scores, r5, r6, r8, r9 and the grouped average numbers of these four variables by location are really high. Those 12 variables represent land value or total value of properties per building size or volume. I think the fraudulence results from either the building size/volume being large or the property value being far too low. Since they are highly related with location patterns, I only extract observations of borough 4 or whose zip code is 11422. The 75th percentile of

AVLAND is 15,392.75, less than 1,792,808,947.00; and the 75th percentile of AVTOT is 24,480, less than 4,668,308,947.00. Therefore, property values are less likely to be fraud signs. The 75th percentile of BLDFRONT and BLDDEPTH are 26 and 46, which are quite less than 62.08 and 86.3. The building's stories, length and width are overestimated and need further investigations.

RECORD 996,722



Table 1: Original data record

RECORD	996,722	BLDFRONT	39.5
BORO	5	BLDDEPTH	83.67
ZIP	10,304	LTFRONT	300.00
STADDR	OSGOOD AVENUE	LTDEPTH	400.00
OWNER	IMPERIAL COURT HOMEOW.	AVLAND	1.00
TAXCLASS	1B	AVTOT	1.00
STORIES	4	FULLVAL	250.00

Table 2: Highest scoring records

	inv	inv_zip5	inv_zip3
r1	\	287.705998	\
r4	977.15178	653.938632	971.093118
r5	\	204.866464	213.223181
r7	1015.79693	\	947.266512
r8	368.539683	\	\
r9	221.69821	\	\

For the record 996,722, there isn't any zero or null value needed to be filled in. According to the original profile, the property has four floors, so its land value or total value is not reasonable to be one. Since 9 of 10 highest z scores are related to land value or total value of the property, I believe the property values might be the reason for ramping up to a high combined score. In addition, according to the google map, I can't find any 4-floors buildings. There is only a six-floor orange building on the west side, and a two-floors white house on the east side. Thus, the exact name and property values require further investigation.

RECORD: 956,520

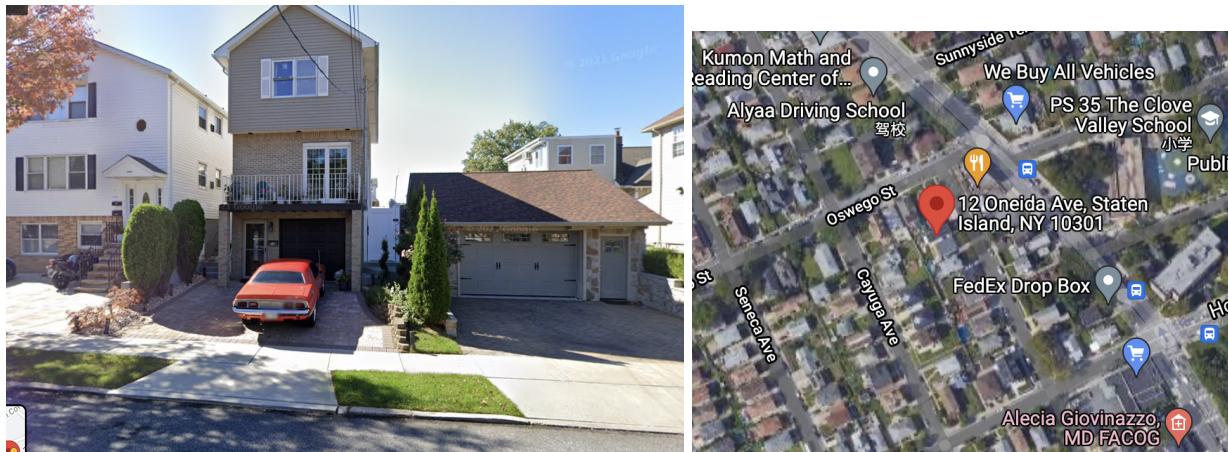


Table 1: Original data record

RECORD	956,520	BLDFRONT	1,812.00
BORO	5	BLDDEPTH	5,020.00
ZIP	10,301	LTFRONT	25
STADDR	12 ONEIDA AVENUE	LTDEPTH	91
OWNER	TROMPETTA RIZALINA.	AVLAND	15,600.00
TAXCLASS		AVTOT	20,892.00
STORIES	3	FULLVAL	348,200.00

Table 2: Highest scoring records

	inv_taxclass		inv_taxclass
r2	648.455279	r6	657.33347
r3	753.597262	r8	625.53622
r5	586.352291	r9	736.753067

Firstly, in the data set of tax class 1, the BLDFRONT and BLDDEPTH of record 956520 is in high range (exceeding 75th percentile number 23 and 49). Since variables with high z scores are all interrelated with building size and volume, and the street picture shows exactly a 3-floor building, probably data of length or width of building results in fraudulence. Apart from that, all highest scoring records focus on inverse value grouped by tax class, which further confirms that building size is suspicious. And whether the property belongs to tax class 1 worths further investigations.

RECORD: 7,040

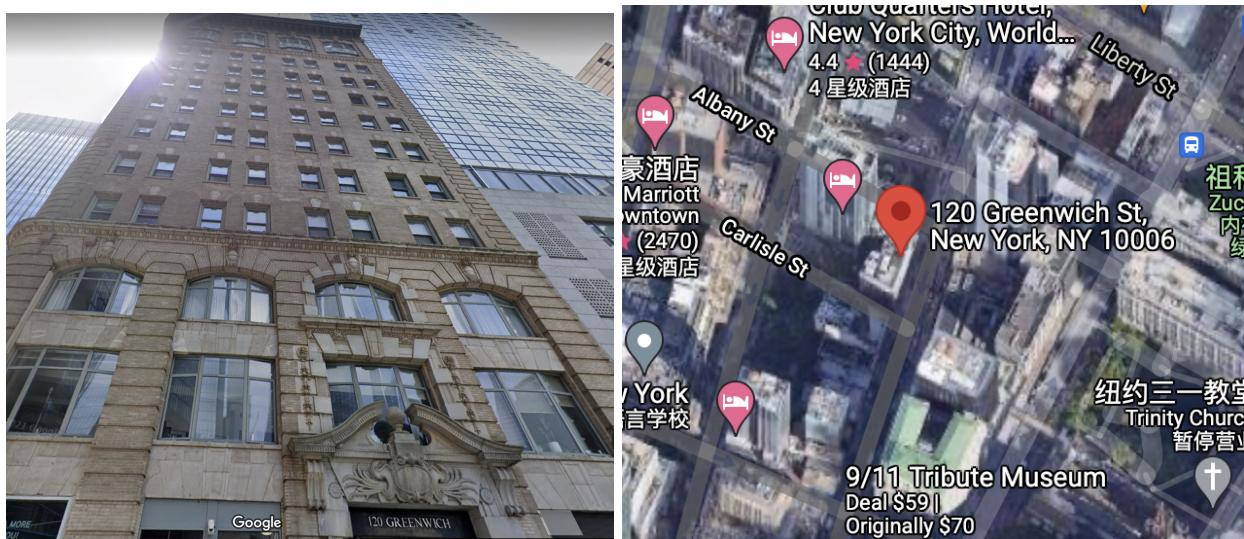


Table 1: Original data record

RECORD	7,040	BLDFRONT	91.05
BORO	1	BLDDEPTH	104.28
ZIP	10,006	LTFRONT	111.58
STADDR	120 GREENWICH STREET	LTDEPTH	130.82
OWNER	BILINKAS, EDWARD	AVLAND	2
TAXCLASS	2	AVTOT	10
STORIES	13	FULLVAL	18

Table 2: Highest scoring records

r1inv	r2inv	r3inv	r4inv	r5inv	r6inv	r7inv	r8inv	r9inv
139.6464	173.8566	233.0032	59.41443	210.8624	160.4865	12.35101	107.6608	210.5360

Record 7,040 represents 10 properties that get the highest z scores. For variables r1-r9 and grouped average of them, z scores of value are not significant, but z scaled inverse values are

mostly obvious. This phenomenon results from wrong property value: AVLAND, AVTOT and FULLVAL. From the picture, 120 GREENWICH has 13 floors and great building volume, those three values are not reasonable. Professional should further estimate its market value, land value and total value.

RECORD: 1,059,883



Table 1: Original data record

RECORD	1,059,883	BLDFRONT	62.08
BORO	5	BLDEPTH	86.30
ZIP	10,309	LTFRONT	5.00
STADDR	SAGONA COURT	LTDEPTH	5.00
OWNER	BILINKAS, EDWARD	AVLAND	444,996.05
TAXCLASS	4	AVTOT	1,294,236.95
STORIES	5.52	FULLVAL	2,772,746.70

Table 2: Highest scoring records

	r	r_zip5	r_boro
r1	207.855436	345.185081	302.57426
r4	254.22162	478.812708	311.814622
r7	253.797324	557.068633	329.134985

Since z scores of r1, r4 and r7 are high, instead of those inverse variables, the fraudulency is more likely to be caused by overestimation of property value and underestimated size.

Extracting the data set in which the zip code is 10309, I find that all of the property values are within the 75th percentile number and maximum value. eg. FULLVAL 602,000<2,772,746.7<290,174,600. Apart from that, I didn't get street photos of this property, so I doubt the number of stories, especially because it's hard to understand why it's a float and not an integer. LTFRONT and LTDEPTH are both below the 25th percentile number (5<25 and 5<90). Researchers should further examine property's value, measure lot's and building's size and take pictures of the exterior.