

# **TECHNICAL REPORT 2011-9**

Blanks and Zeroes in the SCAG Parcel Database

Yizhen Gu and Richard Arnott

November 2011

## Blanks and Zeroes in the SCAG Parcel Database

Many fields in the SCAG parcel database have a large number of cells whose datum is or may be missing. If a cell is blank, the datum is obviously missing. If a cell contains a zero, it may either be that the zero represents the recorded value (for example, if a parcel has been misclassified as having a developed land use but is in fact vacant) or that the zero represents "datum unknown" or "datum not recorded" or "datum not yet transcribed from paper to electronic form". The patterns of blanks and zeroes different across counties and across aggregated land use category. Appendix 1 indicates how the aggregated land use categories were obtained from the SCAG land use categories.

This report simply records the percentage of blanks and zeroes by data field, county, and aggregated land use category. It does not report on the results of Yizhen Gu's visits to the various assessment offices, where he enquired into the possible reasons for the blanks and the zeroes. Yizhen Gu plans to write up a technical reporting on the results of his visits.

Table 1: Sales price

land use type\County		Imperial		Los Angeles		Orange	
		% of parcels entered as		% of parcels entered as		% of parcels entered as	
		zero	blank	zero	blank	zero	blank
1	single family residential	sales price unavailable		44.1	0.0	33.8	0.0
2	multi-family residential			38.4	0.0	40.5	0.0
3	mixed residential			45.9	0.0	47.1	0.0
4	office			38.0	0.0	59.4	0.0
5	retail			40.2	0.0	63.2	0.0
6	other commercial			43.8	0.0	70.2	0.0
7	public			55.4	0.0	76.3	0.0
8	warehousing			45.0	0.0	85.2	0.0
9	other industrial			41.1	0.0	57.7	0.0
10	transportation/communication/utilities			68.6	0.0	92.2	0.0
11	mixed			41.9	0.0	62.9	0.0
12	vacant			41.9	0.0	85.0	0.0
13	other			40.4	0.0	85.2	0.0
land use type\County		Riverside		San Bernardino		Ventura	
		% of parcels entered as		% of parcels entered as		% of parcels entered as	
		zero	blank	zero	blank	zero	blank
1	single family residential	59.4	0.0	sales price unavailable	sales price unavailable	sales price unavailable	sales price unavailable
2	multi-family residential	64.1	0.0				
3	mixed residential	67.5	0.0				
4	office	71.4	0.0				
5	retail	75.2	0.0				
6	other commercial	72.6	0.0				
7	public	88.9	0.0				
8	warehousing	78.8	0.0				
9	other industrial	74.6	0.0				
10	transportation/communication/utilities	92.2	0.0				
11	mixed	67.7	0.0				
12	vacant	73.1	0.0				
13	other	88.7	0.0				

Note 1: Parcels entered as "X" refer to parcels with sales price entered as "X".

Note 2: The percent of parcels entered as "X" is calculated as Number of parcels entered as "X" in that land-use type divided by Number of parcels in that land-use type.

Note 3: The current calculation is based on information exported from LA.gdb. We will try to update this table using information exported from the raw datasets later.

Table 2: Sales date

land use type\County		Imperial		Los Angeles		Orange	
		% of parcels entered as		% of parcels entered as		% of parcels entered as	
		zero	blank	zero	blank	zero	blank
1	single family residential	sales date unavailable		8.7	0.0	5.6	0.0
2	multi-family residential		7.4	0.0	7.9	0.0	
3	mixed residential		9.1	0.0	12.2	0.0	
4	office		9.3	0.0	11.3	0.0	
5	retail		12.5	0.0	12.9	0.0	
6	other commercial		16.0	0.0	22.8	0.0	
7	public		37.6	0.0	17.9	0.0	
8	warehousing		27.5	0.0	31.0	0.0	
9	other industrial		15.5	0.0	11.9	0.0	
10	transportation/communication/utilities		57.6	0.0	46.5	0.0	
11	mixed		14.3	0.0	12.4	0.0	
12	vacant		23.8	0.0	31.0	0.0	
13	other		13.5	0.0	20.7	0.0	
land use type\County		Riverside		San Bernardino		Ventura	
		% of parcels entered as		% of parcels entered as		% of parcels entered as	
		zero	blank	zero	blank	zero	blank
1	single family residential	45.4	0.0	71.6	0.0	73.6	0.0
2	multi-family residential	50.9	0.0	65.8	0.0	72.6	0.0
3	mixed residential	52.7	0.0	82.1	0.0	80.2	0.0
4	office	53.9	0.0	77.5	0.0	84.3	0.0
5	retail	58.5	0.0	82.3	0.0	73.4	0.0
6	other commercial	50.1	0.0	89.9	0.0	71.2	0.0
7	public	79.1	0.0	85.7	0.0	79.5	0.0
8	warehousing	66.5	0.0	85.0	0.0	62.9	0.0
9	other industrial	60.1	0.0	75.9	0.0	74.0	0.0
10	transportation/communication/utilities	84.3	0.0	90.9	0.0	92.2	0.0
11	mixed	52.9	0.0	83.7	0.0	76.5	0.0
12	vacant	56.9	0.0	89.3	0.0	82.7	0.0
13	other	74.1	0.0	82.3	0.0	99.9	0.0

Note 1: Parcels entered as "X" refer to parcels with sales date entered as "X".

Note 2: The percent of parcels entered as "X" is calculated as Number of parcels entered as "X" in that land-use type divided by Number of parcels in that land-use type.

Note 3: The current calculation is based on information exported from LA.gdb. We will try to update this table using information exported from the raw datasets later.

Table 3: floor area

land use type\County		Imperial		Los Angeles		Orange	
		% of parcels entered as		% of parcels entered as		% of parcels entered as	
		zero	blank	zero	blank	zero	blank
1	single family residential	8.5	0.1	0.0	2.9	15.1	4.6
2	multi-family residential	14.0	0.4	0.0	4.6	14.0	5.3
3	mixed residential	27.3	0.2	0.1	6.7	19.7	11.8
4	office	14.1	0.8	0.0	6.2	27.2	9.3
5	retail	25.8	0.3	0.2	13.5	25.4	11.9
6	other commercial	24.7	0.0	0.2	14.5	32.2	21.9
7	public	64.7	0.0	0.2	29.9	49.7	16.6
8	warehousing	25.0	2.6	0.3	25.2	27.3	30.3
9	other industrial	40.8	1.4	0.8	19.2	21.7	10.9
10	transportation/communication/utilities	78.0	1.6	0.6	68.2	45.7	45.9
11	mixed	18.7	0.0	0.2	14.3	46.8	11.4
12	vacant	75.0	0.2	0.2	83.2	66.8	26.8
13	other	83.7	15.2	4.2	20.4	24.8	68.3
land use type\County		Riverside		San Bernardino		Ventura	
		% of parcels entered as		% of parcels entered as		% of parcels entered as	
		zero	blank	zero	blank	zero	blank
1	single family residential	2.9	10.5	8.5	1.3	4.7	0.0
2	multi-family residential	1.8	22.3	7.3	8.4	11.6	0.0
3	mixed residential	18.9	7.4	22.9	3.0	36.1	0.0
4	office	3.9	59.5	43.5	3.4	83.5	0.0
5	retail	5.7	84.8	67.9	3.3	89.9	0.0
6	other commercial	7.0	73.9	55.8	2.9	94.1	0.0
7	public	9.2	77.0	79.1	2.2	77.1	0.0
8	warehousing	5.9	92.5	85.0	7.5	96.8	0.0
9	other industrial	10.6	61.5	68.1	2.8	85.7	0.0
10	transportation/communication/utilities	15.8	78.6	82.2	7.8	96.0	0.0
11	mixed	6.8	85.9	59.1	0.9	97.0	0.0
12	vacant	40.7	42.9	87.2	4.2	92.1	0.0
13	other	24.4	63.8	48.6	34.9	99.7	0.0

Note 1: Parcels entered as "X" refer to parcels with FLOOR AREA entered as "X".  
 Note 2: The percent of parcels entered as "X" is calculated as Number of parcels entered as "X" in that land-use type divided by Number of parcels in that land-use type.

Table 4: Year built

land use type\County		Imperial		Los Angeles		Orange	
		% of parcels entered as		% of parcels entered as		% of parcels entered as	
		zero	blank	zero	blank	zero	blank
1	single family residential	0.0	0.1	0.0	2.9	15.0	4.6
2	multi-family residential	0.0	0.4	0.0	4.6	13.6	5.3
3	mixed residential	0.0	0.2	0.0	6.7	17.8	11.8
4	office	0.0	0.8	0.0	6.2	25.0	9.3
5	retail	0.0	0.3	0.0	13.5	20.2	11.9
6	other commercial	0.0	0.0	0.0	14.5	29.5	21.9
7	public	0.0	0.0	0.0	29.9	39.1	16.6
8	warehousing	0.0	2.6	0.0	25.2	23.9	30.3
9	other industrial	0.0	1.4	0.0	19.2	18.9	10.9
10	transportation/communication/utilities	0.0	1.6	0.0	68.2	45.3	45.9
11	mixed	0.0	0.0	0.0	14.3	45.3	11.4
12	vacant	0.0	0.2	0.0	83.2	66.7	26.8
13	other	0.0	15.2	0.0	20.4	24.8	68.3
land use type\County		Riverside		San Bernardino		Ventura	
		% of parcels entered as		% of parcels entered as		% of parcels entered as	
		zero	blank	zero	blank	zero	blank
1	single family residential	2.8	10.5	0.0	0.0	4.6	0.0
2	multi-family residential	1.8	22.3	0.0	0.0	11.6	0.0
3	mixed residential	17.2	7.4	0.0	0.0	35.9	0.0
4	office	3.8	59.5	0.0	0.0	63.0	0.0
5	retail	5.4	84.8	0.0	0.0	31.0	0.0
6	other commercial	6.8	73.9	0.0	0.0	34.5	0.0
7	public	8.7	77.0	0.0	0.0	67.3	0.0
8	warehousing	5.9	92.5	0.0	0.0	51.6	0.0
9	other industrial	10.1	61.5	0.0	0.0	28.0	0.0
10	transportation/communication/utilities	15.4	78.6	0.0	0.0	92.6	0.0
11	mixed	6.8	85.9	0.0	0.0	27.1	0.0
12	vacant	38.9	42.9	0.0	0.0	91.7	0.0
13	other	22.5	63.8	0.0	8.3	99.6	0.0

Note 1: Parcels entered as "X" refer to parcels with Year built entered as "X".

Note 2: The percent of parcels entered as "X" is calculated as Number of parcels entered as "X" in that land-use type divided by Number of parcels in that land-use type.

Table 5: Assessed value (totvalue07)

land use type\County		Imperial		Los Angeles		Orange	
		% of parcels entered as		% of parcels entered as		% of parcels entered as	
		zero	blank	zero	blank	zero	blank
1	single family residential	0.8	0.1	0.1	0.2	2.3	4.6
2	multi-family residential	3.3	0.0	0.2	1.7	8.4	5.3
3	mixed residential	2.4	0.2	0.2	0.7	4.1	11.8
4	office	4.0	0.8	0.4	3.4	5.3	9.3
5	retail	3.7	0.3	0.8	1.6	1.9	11.9
6	other commercial	6.4	0.0	2.3	4.0	6.9	21.9
7	public	54.8	0.0	4.2	5.3	4.4	16.6
8	warehousing	1.7	2.6	10.1	1.6	6.1	30.3
9	other industrial	8.1	1.4	2.4	3.2	2.4	10.9
10	transportation/communication/utilities	49.2	1.6	26.7	1.3	23.8	45.9
11	mixed	1.1	0.0	0.5	2.0	5.6	11.4
12	vacant	9.6	0.2	2.9	2.7	27.4	26.8
13	other	53.6	15.2	2.3	8.6	18.2	68.3
land use type\County		Riverside		San Bernardino		Ventura	
		% of parcels entered as		% of parcels entered as		% of parcels entered as	
		zero	blank	zero	blank	zero	blank
1	single family residential	0.4	7.3	1.2	1.3	1.7	0.0
2	multi-family residential	2.5	4.3	3.4	8.4	9.1	0.0
3	mixed residential	0.6	2.4	1.7	3.0	8.7	0.0
4	office	2.4	6.6	3.9	3.4	50.1	0.0
5	retail	3.0	9.4	4.6	3.3	7.0	0.0
6	other commercial	4.3	6.9	8.5	2.9	7.1	0.0
7	public	12.6	29.0	62.0	2.2	42.9	0.0
8	warehousing	2.4	6.4	14.1	7.5	8.1	0.0
9	other industrial	2.1	8.1	7.1	2.8	8.2	0.0
10	transportation/communication/utilities	22.2	28.7	57.2	7.8	79.4	0.0
11	mixed	1.0	3.4	3.8	0/9	9.8	0.0
12	vacant	5.9	11.8	23.1	4.2	42.5	0.0
13	other	32.7	20.3	26.7	34.9	99.4	0.1

Note 1: Parcels entered as "X" refer to parcels with totvalue07 entered as "X".

Note 2: The percent of parcels entered as "X" is calculated as Number of parcels entered as "X" in that land-use type divided by Number of parcels in that land-use type.

Table 6: Planned land use

land use type\County		Imperial		Los Angeles		Orange	
		% of parcels entered as		% of parcels entered as		% of parcels entered as	
		zero	blank	zero	blank	zero	blank
1	single family residential	0.0	0.1	0.0	0.7	0.1	0.5
2	multi-family residential	0.0	0.2	0.0	1.5	0.2	0.5
3	mixed residential	0.0	0.5	0.0	0.1	0.3	0.1
4	office	0.0	0.8	0.0	2.3	0.1	1.8
5	retail	0.0	0.0	0.0	2.4	0.1	1.6
6	other commercial	0.0	0.3	0.1	1.3	0.1	5.0
7	public	0.0	0.1	0.1	2.0	0.7	1.4
8	warehousing	0.0	0.0	0.1	0.8	0.0	5.1
9	other industrial	0.1	1.3	0.1	1.1	0.1	2.2
10	transportation/communication/utilities	0.3	2.2	0.2	2.2	0.5	15.9
11	mixed	0.0	0.0	0.0	1.3	0.0	0.1
12	vacant	0.0	0.2	0.0	2.2	1.0	2.6
13	other	0.0	15.8	0.0	10.4	0.0	61.9
land use type\County		Riverside		San Bernardino		Ventura	
		% of parcels entered as		% of parcels entered as		% of parcels entered as	
		zero	blank	zero	blank	zero	blank
1	single family residential	1.4	0.2	0.0	1.9	0.2	0.0
2	multi-family residential	1.4	1.0	0.0	1.9	1.0	0.0
3	mixed residential	0.4	0.0	0.0	1.9	0.1	0.0
4	office	0.9	2.9	0.1	0.8	3.6	0.0
5	retail	1.1	2.3	0.0	1.3	0.1	0.0
6	other commercial	1.5	1.6	0.1	0.9	0.6	0.0
7	public	1.0	3.8	1.0	1.6	1.2	0.0
8	warehousing	0.8	0.3	0.0	0.0	3.2	0.0
9	other industrial	0.7	0.4	0.0	1.4	1.1	0.0
10	transportation/communication/utilities	4.3	2.7	1.0	8.3	13.2	0.0
11	mixed	0.5	0.0	0.0	0.4	0.6	0.0
12	vacant	3.1	0.3	0.0	2.0	4.2	0.0
13	other	1.5	14.0	0.0	12.7	58.7	0.0

Note 1: Parcels entered as "X" refer to parcels with planned land use entered as "X".

Note 2: The percent of parcels entered as "X" is calculated as Number of parcels entered as "X" in that land-use type divided by Number of parcels in that land-use type.

Note 3: The accuracy of planned land use is highly related to that of current land use. Thus, as expected, the missing rate of planned land use is higher for parcels of "other" type.



## APPENDIX 1

The correspondence between LA project land use code and SCAG 1993 code

Table 1: Basic

LA project land use code			SCAG 1993 code
Numerical	Alphabetic	Name	3-digit
1	RS-SF	Single-family residential	111
2	RS-MF	Multi-family residential	112
3	RS-MX	Mixed residential	110, 113, 114, 115
4	OF	Office	121
5	RT	Retail	122
6	OC	Other commercial	120, 123
7	P	Public	124, 125, 126, 127
8	W	Warehousing	134
9	I	Industrial	130, 131, 132, 133
10	TCU	Transportation/communication /utilities	140, 141, 142, 143, 144, 145, 146
11	M	Mixed	150, 160
12	V	Vacant	170; 180, 181, 182, 183, 184, 185, 186, 187, 188; 200, 210, 220, 230, 240, 250, 260, 270; 300, 310, 320
13	O	Other	400, 410, 420, 430, 440, 450; 128, 129, 190, 888, 999; 0, missing

Note 1: The original SCAG 1993 code is 4-digit. Here the 3-digit SCAG code refers to that by dropping off the last digit of the original code.

Note 2: The 13<sup>th</sup> land use code O refers to lu08 listed as “4000-4500” (water), and “0”, “1900”, “8888”, “9999”, or missing in the database.

Table 2: Extended with explanations on SCAG code

LA project land use code			SCAG 1993 code		
Numerical	Alphabetic	Name	3-digit	Name	Note
1	RS-SF	Single-family residential	111	Single family residential	
2	RS-MF	Multi-family residential	112	Multi-family residential	
3	RS-MX	Mixed residential	110	Residential	For parcels lacking of more detailed classification
			113	Mobile homes and trailer parks	
			114	Mixed residential	
			115	Rural residential	
4	OF	Office	121	General office use	
5	RT	Retail	122	Retail stores and commercial stores	
6	OC	Other commercial	120	Commercial and Services	For parcels lacking of more detailed classification
			123	Other commercial	
7	P	Public	124	Public facilities	
			125	Special Use Facilities	
			126	Educational institutions	
			127	Military installations	
8	W	Warehousing	134	Wholesaling and warehousing	
9	I	Industrial	130	Industrial	For parcels lacking of more detailed classification
			131	Light industrial	
			132	Heavy industrial	
			133	Extraction	

10	TCU	Transportation/ communication /utilities	14x	Transportation, communications, and utilities	Including 140, 141, 142, 143, 144, 145, 146
11	M	Mixed	150	Mixed commercial and industrial	
			160	Mixed urban	
12	V	Vacant	170	Under construction	
			18x	Open space and recreation	Including 180, 181, 182, 183, 184, 185, 186, 187, 188
			2x0	Agriculture	Including 200, 210, 220, 230, 240, 250, 260, 270
			3x0	Vacant	Including 300, 310, 320
13	O	Other	4x0	Water	Including 400, 410, 420, 430, 440, 450
			0, 1900, 1280, 1290, 8888, 9999, or missing		Not included in SCAG code

Note: The original SCAG 1993 code is 4-digit. Here the 3-digit SCAG code refers to that by dropping off the last digit of the original code.