12 Bins of Building Types
All Clean Data
N= 6624 & 20 Variables

The FREQ Procedure

bname				
bname	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Education	83	1.27	83	1.27
Entertainment	18	0.28	101	1.55
Facility	242	3.71	343	5.25
Food Service	1	0.02	344	5.27
Gas Station	1	0.02	345	5.28
General Medical	36	0.55	381	5.84
Hotel	142	2.17	523	8.01
Multifamily	5216	79.89	5739	87.90
Office	621	9.51	6360	97.41
Other	74	1.13	6434	98.54
Parking	14	0.21	6448	98.76
Retail	81	1.24	6529	100.00

Input vs. Target
Correlations (Clean)
N= 6624

The CORR Procedure

17 With Variables:	bin1 bin2 bin3 bin4 bin5 bin6 bin7 bin8 bin9 bin10 bin11 bin12 occupy floor_area water_use ws_ngas ws_elec
1 Variables:	site_eui

Simple Statistics							
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
bin1	6529	0.01271	0.11204	83.00000	0	1.00000	D1 Education
bin2	6529	0.00276	0.05244	18.00000	0	1.00000	D2 Entertainment
bin3	6529	0.03707	0.18894	242.00000	0	1.00000	D3 Facility
bin4	6529	0	0	0	0	0	D4 Food Service
bin5	6529	0.0001532	0.01238	1.00000	0	1.00000	D5 Gas Station
bin6	6529	0.00551	0.07406	36.00000	0	1.00000	D6 Medical
bin7	6529	0.02175	0.14587	142.00000	0	1.00000	D7 Hotel
bin8	6529	0.79890	0.40086	5216	0	1.00000	D8 Multifamily
bin9	6529	0.09511	0.29339	621.00000	0	1.00000	D9 Office
bin10	6529	0.01133	0.10586	74.00000	0	1.00000	D10 Other
bin11	6529	0.00214	0.04626	14.00000	0	1.00000	D11 Parking
bin12	6529	0.01241	0.11070	81.00000	0	1.00000	D12 Retail
occupy	6529	99.15224	5.53278	647365	0	100.00000	occupy
floor_area	6529	136860	171146	893557977	500.00000	2376683	floor_area
water_use	6529	16031	157196	104663920	0	6594604	water_use
ws_ngas	6529	1.97718	116.26686	12909	0	9393	ws_ngas
ws_elec	6529	1476976	9362351	9643176395	0	452023582	ws_elec
site_eui	6529	260.57457	10898	1701291	0	869265	site_eui

Pearson Correlation Coefficients, N = 6529 Prob > |r| under H0: Rho=0	
	site_eui
bin1 D1 Education	-0.00063 0.9597
bin2 D2 Entertainment	-0.00061 0.9609
bin3 D3 Facility	-0.00381 0.7585
bin4 D4 Food Service	. .
bin5 D5 Gas Station	-0.00020 0.9870
bin6 D6 Medical	-0.00048 0.9689

bin7 D7 Hotel	0.00468 0.7056	
bin8 D8 Multifamily	0.00442 0.7211	
bin9 D9 Office	-0.00418 0.7354	
bin10 D10 Other	-0.00168 0.8919	
bin11 D11 Parking	-0.00046 0.9706	
bin12 D12 Retail	-0.00150 0.9033	
occupy occupy	-0.02067 0.0948	
floor_area floor_area	-0.00451 0.7155	
water_use water_use	-0.00104 0.9329	
ws_ngas ws_ngas	0.98729 <.0001	
ws_elec ws_elec	0.03153 0.0108	

Categorical vs. Categorical Correlations (Clean)
N= 6624

The CORR Procedure

11 With Variables:	bin1 bin2 bin3 bin5 bin6 bin7 bin8 bin9 bin10 bin11 bin12
11 Variables:	bin1 bin2 bin3 bin5 bin6 bin7 bin8 bin9 bin10 bin11 bin12

Simple Statistics							
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
bin1	6529	0.01271	0.11204	83.00000	0	1.00000	D1 Education
bin2	6529	0.00276	0.05244	18.00000	0	1.00000	D2 Entertainment
bin3	6529	0.03707	0.18894	242.00000	0	1.00000	D3 Facility
bin5	6529	0.0001532	0.01238	1.00000	0	1.00000	D5 Gas Station
bin6	6529	0.00551	0.07406	36.00000	0	1.00000	D6 Medical
bin7	6529	0.02175	0.14587	142.00000	0	1.00000	D7 Hotel
bin8	6529	0.79890	0.40086	5216	0	1.00000	D8 Multifamily
bin9	6529	0.09511	0.29339	621.00000	0	1.00000	D9 Office
bin10	6529	0.01133	0.10586	74.00000	0	1.00000	D10 Other
bin11	6529	0.00214	0.04626	14.00000	0	1.00000	D11 Parking
bin12	6529	0.01241	0.11070	81.00000	0	1.00000	D12 Retail

Pearson Correlation Coefficients, N = 6529 Prob > |r| under H0: Rho=0											
	bin1	bin2	bin3	bin5	bin6	bin7	bin8	bin9	bin10	bin11	bin12
bin1 D1 Education	1.00000	-0.00597 0.6298	-0.02226 0.0721	-0.00140 0.9097	-0.00845 0.4949	-0.01692 0.1716	-0.22617 <.0001	-0.03679 0.0029	-0.01215 0.3263	-0.00526 0.6709	-0.01272 0.3042
bin2 D2 Entertainment	-0.00597 0.6298	1.00000	-0.01032 0.4046	-0.00065 0.9581	-0.00392 0.7518	-0.00784 0.5265	-0.10480 <.0001	-0.01705 0.1684	-0.00563 0.6493	-0.00244 0.8439	-0.00589 0.6340
bin3 D3 Facility	-0.02226 0.0721	-0.01032 0.4046	1.00000	-0.00243 0.8445	-0.01461 0.2379	-0.02925 0.0181	-0.39104 <.0001	-0.06361 <.0001	-0.02101 0.0897	-0.00909 0.4625	-0.02199 0.0756
bin5 D5 Gas Station	-0.00140 0.9097	-0.00065 0.9581	-0.00243 0.8445	1.00000	-0.00092 0.9407	-0.00185 0.8815	-0.02467 0.0462	-0.00401 0.7458	-0.00133 0.9147	-0.00057 0.9630	-0.00139 0.9108
bin6 D6 Medical	-0.00845 0.4949	-0.00392 0.7518	-0.01461 0.2379	-0.00092 0.9407	1.00000	-0.01110 0.3697	-0.14841 <.0001	-0.02414 0.0511	-0.00797 0.5195	-0.00345 0.7804	-0.00835 0.5002
bin7 D7 Hotel	-0.01692 0.1716	-0.00784 0.5265	-0.02925 0.0181	-0.00185 0.8815	-0.01110 0.3697	1.00000	-0.29719 <.0001	-0.04834 <.0001	-0.01596 0.1971	-0.00691 0.5766	-0.01671 0.1770
bin8 D8 Multifamily	-0.22617 <.0001	-0.10480 <.0001	-0.39104 <.0001	-0.02467 0.0462	-0.14841 <.0001	-0.29719 <.0001	1.00000	-0.64619 <.0001	-0.21340 <.0001	-0.09239 <.0001	-0.22339 <.0001
bin9 D9 Office	-0.03679 0.0029	-0.01705 0.1684	-0.06361 <.0001	-0.00401 0.7458	-0.02414 0.0511	-0.04834 <.0001	-0.64619 <.0001	1.00000	-0.03471 0.0050	-0.01503 0.2247	-0.03634 0.0033
bin10 D10 Other	-0.01215 0.3263	-0.00563 0.6493	-0.02101 0.0897	-0.00133 0.9147	-0.00797 0.5195	-0.01596 0.1971	-0.21340 <.0001	-0.03471 0.0050	1.00000	-0.00496 0.6884	-0.01200 0.3323
bin11 D11 Parking	-0.00526 0.6709	-0.00244 0.8439	-0.00909 0.4625	-0.00057 0.9630	-0.00345 0.7804	-0.00691 0.5766	-0.09239 <.0001	-0.01503 0.2247	-0.00496 0.6884	1.00000	-0.00520 0.6747
bin12	-0.01272	-0.00589	-0.02199	-0.00139	-0.00835	-0.01671	-0.22339	-0.03634	-0.01200	-0.00520	1.00000

| D12 Retail | 0.3042 | 0.6340 | 0.0756 | 0.9108 | 0.5002 | 0.1770 | <.0001 | 0.0033 | 0.3323 | 0.6747 | |

Categorical vs. Feature
Correlations (Clean)
N= 6624

The CORR Procedure

5 With Variables:	occupy floor_area water_use ws_ngas ws_elec
11 Variables:	bin1 bin2 bin3 bin5 bin6 bin7 bin8 bin9 bin10 bin11 bin12

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
occupy	6529	99.15224	5.53278	647365	0	100.00000	occupy
floor_area	6529	136860	171146	893557977	500.00000	2376683	floor_area
water_use	6529	16031	157196	104663920	0	6594604	water_use
ws_ngas	6529	1.97718	116.26686	12909	0	9393	ws_ngas
ws_elec	6529	1476976	9362351	9643176395	0	452023582	ws_elec
bin1	6529	0.01271	0.11204	83.00000	0	1.00000	D1 Education
bin2	6529	0.00276	0.05244	18.00000	0	1.00000	D2 Entertainment
bin3	6529	0.03707	0.18894	242.00000	0	1.00000	D3 Facility
bin5	6529	0.0001532	0.01238	1.00000	0	1.00000	D5 Gas Station
bin6	6529	0.00551	0.07406	36.00000	0	1.00000	D6 Medical
bin7	6529	0.02175	0.14587	142.00000	0	1.00000	D7 Hotel
bin8	6529	0.79890	0.40086	5216	0	1.00000	D8 Multifamily
bin9	6529	0.09511	0.29339	621.00000	0	1.00000	D9 Office
bin10	6529	0.01133	0.10586	74.00000	0	1.00000	D10 Other
bin11	6529	0.00214	0.04626	14.00000	0	1.00000	D11 Parking
bin12	6529	0.01241	0.11070	81.00000	0	1.00000	D12 Retail

Pearson Correlation Coefficients, N = 6529
Prob > |r| under H0: Rho=0

	bin1	bin2	bin3	bin5	bin6	bin7	bin8	bin9	bin10	bin11	bin12
occupy occupy	0.01121 0.3651	0.00806 0.5151	-0.09963 <.0001	0.00190 0.8782	-0.04280 0.0005	-0.06541 <.0001	0.15208 <.0001	-0.07347 <.0001	-0.03459 0.0052	-0.11559 <.0001	-0.00534 0.6664
floor_area floor_area	0.00494 0.6899	-0.01299 0.2940	-0.03493 0.0048	-0.00814 0.5109	0.04481 0.0003	0.07820 <.0001	-0.18958 <.0001	0.23142 <.0001	-0.01963 0.1127	-0.00687 0.5790	0.02408 0.0517
water_use water_use	-0.00866 0.4840	-0.00358 0.7722	-0.01460 0.2380	-0.00103 0.9337	0.18134 <.0001	0.01753 0.1566	-0.02563 0.0384	-0.01828 0.1398	0.00476 0.7008	-0.00428 0.7297	0.02971 0.0164
ws_ngas ws_ngas	-0.00049 0.9686	-0.00071 0.9545	-0.00287 0.8169	-0.00020 0.9871	-0.00067 0.9567	-0.00199 0.8723	0.00685 0.5802	-0.00487 0.6938	-0.00139 0.9104	-0.00049 0.9685	-0.00153 0.9019
ws_elec ws_elec	0.00193 0.8760	0.00343 0.7819	-0.01393 0.2603	-0.00008 0.9950	0.03817 0.0020	0.06346 <.0001	-0.12164 <.0001	0.12382 <.0001	-0.00127 0.9181	-0.00018 0.9885	0.02479 0.0452

Feature vs. Feature
Correlations (Clean)
N= 6624

The CORR Procedure

5 With Variables:	occupy floor_area water_use ws_ngas ws_elec
5 Variables:	occupy floor_area water_use ws_ngas ws_elec

Simple Statistics							
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
occupy	6529	99.15224	5.53278	647365	0	100.00000	occupy
floor_area	6529	136860	171146	893557977	500.00000	2376683	floor_area
water_use	6529	16031	157196	104663920	0	6594604	water_use
ws_ngas	6529	1.97718	116.26686	12909	0	9393	ws_ngas
ws_elec	6529	1476976	9362351	9643176395	0	452023582	ws_elec

Pearson Correlation Coefficients, N = 6529 Prob > |r| under H0: Rho=0					
	occupy	floor_area	water_use	ws_ngas	ws_elec
occupy occupy	1.00000	-0.04565 0.0002	-0.05718 <.0001	-0.02018 0.1030	-0.01500 0.2257
floor_area floor_area	-0.04565 0.0002	1.00000	0.06660 <.0001	-0.00624 0.6143	0.29413 <.0001
water_use water_use	-0.05718 <.0001	0.06660 <.0001	1.00000	-0.00100 0.9359	0.02300 0.0631
ws_ngas ws_ngas	-0.02018 0.1030	-0.00624 0.6143	-0.00100 0.9359	1.00000	-0.00122 0.9214
ws_elec ws_elec	-0.01500 0.2257	0.29413 <.0001	0.02300 0.0631	-0.00122 0.9214	1.00000