

Capstone Project – MD Presentation

Present analysis to help residential home builder locate in the
Houston, Texas metro area

Houston is a large city with many different areas and neighborhoods. A builder looking to locate could use assistance in understanding where to focus attention to pick a final location.

- Try and group neighborhoods into distinct groups
- Use a mix of data to open up more variables
- Provide observations and some insight so builder can make final decision based on their criteria

Data acquisition and cleaning

Data consist of Houston metro area neighborhoods with relevant real estate data over a period ending in 2016.

This is combined with zip code data to get geospatial coordinates for mapping

Further, venue data for the neighborhoods is gathered via Foursquare to add more detail to the analysis

The cleaned and grouped data consist of 130 neighborhoods and 16 features with the first 5 shown below

	Neighborhood	ZIP Code	2016 Median Home Price	% Growth 2010-2016	% Growth 2015-2016	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	1960/Cypress	77065	179000	0.455	0.085	29.9319	-95.6106	5.0	Pizza Place	Mexican Restaurant	Intersection	Donut Shop	Yoga Studio	Doctor's Office	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant
1	Aldine Area	77039	133500	0.571	0.077	29.9067	-95.3334	5.0	Pizza Place	Pharmacy	Mexican Restaurant	Fast Food Restaurant	Chinese Restaurant	Taco Place	Gas Station	Dive Bar	Flower Shop	Factory
2	Alief	77072	164000	0.802	0.147	29.6990	-95.5862	5.0	Lounge	Park	Gas Station	Yoga Studio	Doctor's Office	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory
3	Alvin North	77511	227000	0.434	0.056	29.4120	-95.2515	5.0	Insurance Office	Restaurant	Baseball Field	Dessert Shop	Yoga Studio	Doctor's Office	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant
4	Alvin South	77511	163900	0.463	0.062	29.4120	-95.2515	5.0	Insurance Office	Restaurant	Baseball Field	Dessert Shop	Yoga Studio	Doctor's Office	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant

A map of the Houston, Texas area, showing the distribution of 100 restaurants. The map includes major highways (Interstates 10, 27, 29, 610, and State Routes 281, 241, 290, 151, 160, 101, 51, 28, 24, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000). The dots are color-coded: purple (25), yellow (30), orange (15), red (5), blue (10), and cyan (15). The map shows a high concentration of restaurants in the central urban area, with fewer restaurants in the surrounding suburban and rural areas. The map also shows major highways, water bodies, and city names.

Observational analysis of each cluster with the first five neighborhoods from each

Cluster 0

	Neighborhood	2016 Median Home Price	% Growth 2010-2016	% Growth 2015-2016	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
13	Braeswood Place	715000	0.585	0.014	29.6889	-95.4341	0.0	Recreation Center	Park	Yoga Studio	Dive Bar	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory	Eye Doctor
72	Knollwood /Woodside Area	430000	0.352	-0.075	29.6889	-95.4341	0.0	Recreation Center	Park	Yoga Studio	Dive Bar	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory	Eye Doctor
90	Mission Bend Area	171000	0.513	0.097	29.6947	-95.6511	0.0	Park	Yoga Studio	Dive Bar	Food Service	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory	Eye Doctor

Cluster 0: These neighborhoods are represented by the Red dots on the map. The most popular venue data for these is park or recreation center. The remaining venues are oriented toward residential customers. The home value appreciation for these neighborhoods were good, with all over 30% for the 2010 - 2016 period.

Cluster 1

	Neighborhood	2016 Median Home Price	% Growth 2010-2016	% Growth 2015-2016	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
6	Atascocita South	199000	0.327	0.008	29.9507	-95.2622	1.0	Boat or Ferry	Home Service	Construction & Landscaping	Auto Workshop	Electronics Store	Business Service	Doctor's Office	Food Service	Food	Fondue Restaurant
14	Brays Oaks	225000	0.772	0.100	29.6581	-95.5413	1.0	Construction & Landscaping	Historic Site	Dive Bar	Food Service	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory	Eye Doctor
20	Chambers County West	260000	0.396	0.036	29.7700	-94.8608	1.0	Moving Target	Business Service	Yoga Studio	Doctor's Office	Food Service	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory
21	Champions Area	234195	0.338	0.055	29.9863	-95.5208	1.0	Home Service	Locksmith	Boutique	Playground	Historic Site	Dive Bar	Fondue Restaurant	Hobby Shop	Flower Shop	Fast Food Restaurant
33	Cypress North	230995	0.320	-0.006	29.9766	-95.6358	1.0	Yoga Studio	Home Service	Video Store	Gym / Fitness Center	Dive Bar	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory

Cluster1: These neighborhoods are represented by Purple dots on the map. They are grouped due the the first or second most popular venue is oriented toward industrial customers. This indicates that many of the neighborhoods are located in or near an industrial area. The home appreciation for the period 2010 - 2016 ranged anywhere from just over 10% to over 100%.

Cluster 2

	Neighborhood	2016 Median Home Price	% Growth 2010-2016	% Growth 2015-2016	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
12	Bellaire	933000	0.333	0.009	29.7023	-95.4611	2.0	Pharmacy	Yoga Studio	Dive Bar	Food Service	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory	Eye Doctor
83	Memorial Close In	2348200	1.539	0.269	29.7696	-95.5201	2.0	Clothing Store	Business Service	Yoga Studio	Doctor's Office	Food Service	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory
85	Memorial Villages	1368088	0.410	-0.148	29.7696	-95.5201	2.0	Clothing Store	Business Service	Yoga Studio	Doctor's Office	Food Service	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory
107	Rice/Museum District	819200	0.366	0.100	29.7179	-95.4263	2.0	Boutique	Pet Store	Clothing Store	Beach	Bakery	Yoga Studio	Electronics Store	Food Service	Food	Fondue Restaurant
108	River Oaks Area	2194000	0.809	0.139	29.7517	-95.4054	2.0	Men's Store	Gym	Cosmetics Shop	Spa	Clothing Store	Kids Store	French Restaurant	Italian Restaurant	Café	Gym / Fitness Center

Cluster 2: These neighborhoods are represented by dark Blue dots on the map. These are most of the luxury neighborhoods in the area with the average 2016 home price over \$800,000 and all but 2 neighborhoods well over 1 million. The venue data confirms this.

Cluster 3

	Neighborhood	2016 Median Home Price	% Growth 2010-2016	% Growth 2015-2016	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
37	Denver Harbor	105000	1.543	0.235	29.7758	-95.3121	3.0	Fast Food Restaurant	Sandwich Place	Pizza Place	Mexican Restaurant	Pharmacy	Gas Station	Discount Store	Bus Station	Grocery Store	Intersection
39	East End-Galveston	176900	0.474	0.109	29.2983	-94.7930	3.0	Museum	Pharmacy	College Football Field	Library	Mexican Restaurant	Fried Chicken Joint	Donut Shop	Bar	Monument / Landmark	Taco Place
51	Greenway Plaza	1005000	1.781	0.951	29.7330	-95.4306	3.0	Mexican Restaurant	Bank	Hotel	Thai Restaurant	Seafood Restaurant	Spa	Sandwich Place	Burger Joint	Bar	Café
61	Humble Area South	108300	0.740	0.610	30.0041	-95.2825	3.0	Fried Chicken Joint	Sandwich Place	Fast Food Restaurant	Department Store	Mexican Restaurant	Sporting Goods Shop	Smoke Shop	Shipping Store	Seafood Restaurant	Buffet
73	La Marque	137000	0.730	0.489	29.3676	-94.9742	3.0	Burger Joint	Discount Store	Yoga Studio	Food Service	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory	Eye Doctor

Cluster 3: These neighborhoods are represented on the map by light Blue dots. These neighborhoods show strong growth over the 2010 - 2016 period. All are over 60% growth with the exception of one at 47%. Several had growth rates over 100%. The venue data shows that most of the popular are oriented toward residential customers.

Cluster 4

	Neighbor hood	2016 Median Home Price	% Growth 2010- 2016	% Growth 2015- 2016	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
8	Bayou Vista	240000	0.601	0.160	29.3398	-94.9926	4.0	Bar	Yoga Studio	Doctor's Office	Food Service	Food	Fondue Restauran t	Flower Shop	Fast Food Restauran t	Factory	Eye Doctor
56	Hitchcock	114900	0.017	-0.123	29.3398	-94.9926	4.0	Bar	Yoga Studio	Doctor's Office	Food Service	Food	Fondue Restauran t	Flower Shop	Fast Food Restauran t	Factory	Eye Doctor
100	Omega Bay	289900	0.734	0.035	29.3398	-94.9926	4.0	Bar	Yoga Studio	Doctor's Office	Food Service	Food	Fondue Restauran t	Flower Shop	Fast Food Restauran t	Factory	Eye Doctor

Cluster4: These neighborhoods are represented by Teal dots on the map. There are only three neighborhoods in this group and they all share the same venue data. The most popular is a bar followed by Yoga studio and Doctor's office. This indicates that these neighborhoods are close to some type of retail center. The growth ranged from just under 2% to over 70% for period 2010 - 2016.

Cluster 5

	Neighborhood	2016 Median Home Price	% Growth 2010-2016	% Growth 2015-2016	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	1960/Cypress	179000	0.455	0.085	29.9319	-95.6106	5.0	Pizza Place	Mexican Restaurant	Intersection	Donut Shop	Yoga Studio	Doctor's Office	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant
1	Aldine Area	133500	0.571	0.077	29.9067	-95.3334	5.0	Pizza Place	Pharmacy	Mexican Restaurant	Fast Food Restaurant	Chinese Restaurant	Taco Place	Gas Station	Dive Bar	Flower Shop	Factory
2	Alief	164000	0.802	0.147	29.6990	-95.5862	5.0	Lounge	Park	Gas Station	Yoga Studio	Doctor's Office	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory
3	Alvin North	227000	0.434	0.056	29.4120	-95.2515	5.0	Insurance Office	Restaurant	Baseball Field	Dessert Shop	Yoga Studio	Doctor's Office	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant
4	Alvin South	163900	0.463	0.062	29.4120	-95.2515	5.0	Insurance Office	Restaurant	Baseball Field	Dessert Shop	Yoga Studio	Doctor's Office	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant

Cluster 5: These neighborhoods are represented by light Brown on the map. This is the largest group of neighborhoods. The growth for these has a large range from little growth to nearly 100%. The venue data, while diverse, does indicate that they are located near retail centers. Most of the these neighborhoods had moderate home values in 2016.

Cluster 6

	Neighborhood	2016 Median Home Price	% Growth 2010-2016	% Growth 2015-2016	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
11	Bear Creek	162500	0.548	0.062	29.8323	-95.7360	6.0	Home Service	Yoga Studio	Dive Bar	Food Service	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory	Eye Doctor
28	Conroe Southwest	325000	0.383	-0.044	30.3217	-95.5285	6.0	Home Service	Locksmith	Yoga Studio	Doctor's Office	Food Service	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory
29	Copperfield Area	204000	0.333	0.046	29.8941	-95.6481	6.0	Home Service	Yoga Studio	Dive Bar	Food Service	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory	Eye Doctor
64	Katy-North	182000	0.517	0.058	29.8323	-95.7360	6.0	Home Service	Yoga Studio	Dive Bar	Food Service	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory	Eye Doctor
74	La Porte/Shoreacres	159900	0.355	0.088	29.6884	-95.0513	6.0	Home Service	Food	Yoga Studio	Dive Bar	Food Service	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory	Eye Doctor
104	Porter/New Caney East	227000	1.536	-0.025	30.1579	-95.1980	6.0	Home Service	Yoga Studio	Dive Bar	Food Service	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory	Eye Doctor
112	Santa Fe	181250	0.353	0.076	29.4032	-95.0734	6.0	Home Service	Garden Center	Yoga Studio	Dive Bar	Food	Fondue Restaurant	Flower Shop	Fast Food Restaurant	Factory	Eye Doctor

Cluster 6: These neighborhoods are represented by Orange dots on the map. Growth for all these neighborhoods was good with all over 30% and one over 100%. The venue data for all has home service as the most popular. All are located further out from the center of the metro area. From the data, I assume that many venues are oriented toward residential services.

Recommendations

My recommendation to the residential builder would be as follows. If they want to build luxury homes, start looking in the neighborhoods from cluster 2 as a start point. If they are looking to build moderate customer or planned neighborhoods, clusters 0, 3, 5, and 6 might be good areas to start looking. Cluster 3 are all located close to center of metro area and may be a good place to look for urban style living while clusters 0, 5 and 6 are generally further out and might be better for suburban living. Clusters 1 and 4 seem to be located near industrial or major retail area and should be investigated for further insight.

Conclusions

We were able to pull in data from several sources and run an analysis to try and determine a good location for a residential home builder to locate. We were able to analyze most of the neighborhoods by some real estate metrics and using Foursquare to present the top 10 venues in each neighborhood. Based on this analysis, we have provided the home builder with several neighborhoods that could be targets for further study, depending on type and style of homes they want to build.