

Certainly! Data deduplication involves identifying and removing duplicate entries from a dataset. In this dataset, we can identify duplicates by comparing all columns or a subset of columns that uniquely identify a record. Here, I'll assume that a combination of `brokered_by`, `street`, `city`, `state`, and `zip_code` can uniquely identify a property listing. Let's go through the dataset and remove duplicates based on these columns:

1. **Original Dataset:**

- Total records: 100

2. **Identified Duplicates:**

- 48807;for_sale;375000;3;3;0.2;88 East Ave;San Antonio;Texas;78254;2346;2004-05-14 appears twice.
- 22671;sold;90000;3;2;0.09;553 Third Ct;Saint Louis;Missouri;63115;1152;2022-03-24 appears three times.
- 16829;sold;350000;4;2;0.18;219 Field St;Houston;Texas;77021;1626;2022-03-31 appears twice.

3. **Deduplicated Dataset:**

- Remove the duplicate entries, keeping only one instance of each duplicate.

Here is the deduplicated dataset:

```
brokered_by;status;price;bed;bath;acre_lot;street;city;state;zip_code;house_size;prev_sold_date
81316;sold;749000;4;2;0.27;285 Center Ln;Fairfax;Virginia;22030;2948;2021-12-10
96817;sold;425000;4;3;0.06;33 Pine Ave;Alexandria;Virginia;22309;2075;2022-02-04
53673;for_sale;112900;3;2;0.5;169 Meadow St;Mount Pleasant;Michigan;48858;1524;2012-08-13
10437;sold;350000;3;2;1.07;70 Cedar Dr;Cumming;Georgia;30041;1681;2022-02-23
19903;sold;849500;3;3;0.18;249 South Pl;Houston;Texas;77030;2150;2021-12-31
20385;for_sale;599500;5;3;0.37;217 Church Ln;Calabash;North Carolina;28467;3058;2020-03-24
22792;sold;525000;3;2;0.59;977 Adams Blvd;Westford;Massachusetts;1886;2211;2021-11-19
104873;for_sale;888000;4;4;0.21;210 Church Ln;Houston;Texas;77082;4156;2013-09-26
104876;sold;398000;3;2;0.11;981 Harbor Ave;Chicago;Illinois;60659;1171;2021-11-30
58970;sold;949900;3;4;0.23;828 Spring St;Kensington;Maryland;20895;2500;2021-12-30
94681;sold;409900;2;1;2.12;655 First St;Stockbridge;Massachusetts;1262;1015;2021-11-29
78139;for_sale;425000;1;1;0.58;949 Franklin Blvd;Palisade;Colorado;81526;987;2006-07-06
53138;sold;479000;3;3;0.22;47 Lincoln Rd;Bowie;Maryland;20720;2375;2022-01-04
52946;for_sale;150000;4;2;10.0;675 Circle Dr;Hulbert;Oklahoma;74441;1896;2018-04-19
687;sold;355000;3;2;0.19;272 Terrace Ln;Orlando;Florida;32828;1864;2021-12-08
103967;for_sale;635000;3;3;0.23;885 Washington Ave;Iselin;New Jersey;8830;2000;2019-02-28
10987;for_sale;500000;2;3;0.15;229 Broadway Ave;Corpuschristi;Texss;7841;241;0219-4-3
2765;sold;585000;3;3;4.44;699 Wilson Dr;Crossville;Tennessee;38555;3352;2022-04-15
78460;for_sale;278900;3;2;0.7;95 Meadow Ave;Somerset;Kentucky;42503;1923;2016-07-01
53232;for_sale;330000;2;1;0.3;225 Broadway Ct;Kingston;New York;12401;1062;2014-07-01
22916;sold;376000;4;3;0.28;73 Franklin Rd;Jacksonville;North Carolina;28546;3162;2022-04-25
57424;for_sale;999000;3;2;0.15;277 Princess Dr;Livermore;California;94550;1785;2022-04-22
81112;for_sale;115900;4;1;0.15;380 Valley Ln;Fairborn;Ohio;45324;1116;2020-08-03
92736;for_sale;175000;3;2;0.26;862 King Blvd;Owasso;Oklahoma;74055;1144;2017-12-29
68269;sold;389854;3;2;0.15;144 Terrace Blvd;Caldwell;Idaho;83605;1574;2021-12-17
32769;for_sale;1300000;2;3;27.59;550 Second Pl;Julian;California;92036;986;2020-05-15
84529;for_sale;1075000;3;3;0.06;162 Mill St;Tustin;California;92782;1928;2011-11-14
22217;sold;898000;3;2;0.21;52 North Blvd;Alhambra;California;91803;1465;2022-04-08
65293;for_sale;725000;3;3;1.04;715 Terrace Dr;Sparks;Nevada;89441;1876;2004-09-10
7689;sold;98500;4;1;9.0;272 Broadway Ln;Fairview;Michigan;48621;2064;2022-04-14
78247;sold;339900;3;2;0.28;676 South Blvd;Troy;Illinois;62294;1613;2022-01-26
105798;for_sale;155990;4;3;0.21;195 Spring Ave;Albany;Indiana;47320;2009;2020-06-19
2177;sold;479900;3;2;0.13;842 Cedar Dr;Windsor;Colorado;80550;2638;2022-03-18
26012;sold;259900;4;2;0.13;391 Bridge Rd;Great Falls;Montana;59405;1632;2021-11-09
96014;sold;370000;4;2;0.12;224 Lincoln St;Spokane;Washington;99207;2084;2022-04-05
48807;for_sale;375000;3;3;0.2;88 East Ave;San Antonio;Texas;78254;2346;2004-05-14
55214;sold;509900;3;3;0.2;603 Princess St;San Jacinto;California;92583;2383;2022-05-03
53556;for_sale;375000;5;3;0.3;611 Hill Pl;Wichita;Kansas;67207;2696;2021-11-12
84534;sold;689900;3;4;0.07;920 Field Dr;Vienna;Virginia;22180;2075;2022-04-22
24268;for_sale;550000;3;2;0.22;81 Roosevelt Dr;Charlottesville;Virginia;22903;1512;2011-09-29
845289;for_sale;107500;4;2;0.05;163 Mill St;Tustine;CALIFORNIA;93782;1920;2111-11-14
26543;sold;600000;4;3;1.08;763 Adams Pl;Melbourne;Florida;32934;3056;2022-01-21
22671;sold;90000;3;2;0.09;553 Third Ct;Saint Louis;Missouri;63115;1152;2022-03-24
```

78075;sold;280000;2;1;0.06;686 East Ln;Reno;Nevada;89506;851;2021-11-19
81824;sold;292677;5;3;0.29;63 Field Dr;Sumter;South Carolina;29154;3040;2021-12-22
3479;sold;419000;5;4;0.3;864 Lake Ln;Irving;Texas;75060;3110;2021-12-09
29538;for_sale;1150000;3;2;0.37;223 Lake St;Long Beach;California;90802;1409;2016-06-10
109950;sold;225000;3;2;0.23;414 Second Ln;Fort Worth;Texas;76133;1552;2022-01-19
75016;sold;317900;2;2;0.22;168 Bridge Pl;Kirkwood;Missouri;63122;2032;2021-11-23
53016;sold;239000;3;2;0.56;829 Main Blvd;Oak Ridge;Tennessee;37830;1724;2022-02-02
21986;sold;269900;4;3;0.17;614 Garden Ct;Pooler;Georgia;31322;2368;2021-12-30
102016;for_sale;374900;3;2;0.14;368 Shore Rd;Zephyrhills;Florida;33542;1694;2019-04-25
53377;sold;189900;4;1;0.28;286 Bridge Ln;Saint Louis;Missouri;63114;1624;2022-01-18
8;for_sale;517300;3;2;35.0;624 River Ln;Cheyenne;Wyoming;82007;3018;2022-10-31
53173;sold;350000;3;2;0.9;693 Ridge Pl;Harpers Ferry;West Virginia;25425;3220;2021-12-10
31355;sold;2365000;5;5;0.19;592 Meadow Ave;Edina;Minnesota;55424;4905;2021-12-30
16829;for_sale;1979000;6;6;0.29;400 Adams Ct;Port Washington;New York;11050;5828;2017-05-05
101497;for_sale;129900;4;2;1.95;575 King Pl;Eastover;South Carolina;29044;2128;2021-09-20
53555;for_sale;475000;6;4;0.32;611 Hill;Wicita;Kansaa;6727;2700;2021-11-1
10649;sold;219900;3;2;0.2;430 Church Rd;Allison Park;Pennsylvania;15101;1188;2022-04-13
5145;sold;750000;3;4;0.19;59 Park Blvd;Meridian;Idaho;83642;2796;2022-01-10
53592;for_sale;619900;4;6;2.1;779 Third Blvd;Mchenry;Illinois;60050;4145;2020-04-02
107955;for_sale;2199000;3;3;0.3;20 First Dr;Hingham;Massachusetts;2043;3650;1992-10-22
78200;sold;149000;3;2;0.15;688 Garfield Ct;Akron;Ohio;44314;1560;2022-02-09
15757;for_sale;69900;3;1;0.17;520 View Rd;Enid;Oklahoma;73701;1333;2016-06-09
16829;sold;350000;4;2;0.18;219 Field St;Houston;Texas;77021;1626;2022-03-31
86329;for_sale;699000;4;5;14.25;347 Oak Ln;Wetumpka;Alabama;36092;4758;2015-02-19
22611;for_sale;215000;3;2;0.12;995 Maple Ave;Calumet City;Illinois;60409;1304;2013-08-15
57424;sold;935000;5;5;0.16;409 Wood Ct;Manteca;California;95337;3789;2022-02-18
56084;sold;279900;2;2;0.16;593 Elm Ave;Madison;Wisconsin;53704;1699;2021-12-17
22611;sold;280000;4;2;0.19;952 King Rd;Monroeville;Pennsylvania;15146;1874;2022-04-19
108243;sold;499000;3;3;0.18;197 Meadow Ct;Kennewick;Washington;99337;2289;2022-03-30
4630;for_sale;519950;4;3;0.13;906 Mill Blvd;Huntsville;Alabama;35801;2778;2020-03-09
2792;sold;524000;2;1;0.55;977 Adams Boulevard;westford;Masachusets;1886;2210;2021-12-19
81031;sold;219000;3;2;0.31;73 Lake Ln;Chambersburg;Pennsylvania;17202;1189;2021-11-24
45913;sold;250000;3;2;0.86;94 Washington Ave;Queensbury;New York;12804;1883;2021-11-23
106177;for_sale;399500;3;3;0.27;411 Wilson Ave;Savannah;Georgia;31405;1760;2002-06-03
22562;for_sale;210000;3;2;0.15;272 Farm Blvd;Dearborn Heights;Michigan;48125;1960;2010-10-14
109978;sold;760000;3;2;0.36;760 Madison Ln;Escondido;California;92026;1888;2021-12-22
76215;sold;339000;3;3;0.34;60 Franklin Ave;Bella Vista;Arkansas;72714;2035;2021-12-07
81671;sold;325000;3;2;8.36;651 Vista Blvd;Plato;Missouri;65552;1784;2021-12-17
109987;for_sale;490000;3;4;0.16;229 Broadway Ave;Corpus Christi;Texas;78418;2416;2019-04-03
28222;sold;2450000;4;4;0.19;687 Wilson Ave;Los Angeles;California;90064;2751;2022-02-22
85655;sold;249900;3;1;0.22;159 Meadow Rd;Scituate;Rhode Island;2831;1310;2022-02-04
78167;sold;350000;3;2;0.15;324 Sunset St;Saint Louis Park;Minnesota;55426;1908;2022-04-01
19415;for_sale;309000;3;2;3.0;820 View Ln;Ojibwa;Wisconsin;54862;2025;2020-02-04
45936;for_sale;539900;3;2;0.21;293 Terrace Ct;Skaneateles;New York;13152;1918;2015-06-15
33901;for_sale;225000;2;2;0.33;417 East Ct;Greenfield;Indiana;46140;1456;2017-07-26
100987;for_sale;499000;2;3;0.1;299 Broadway Ave;Corpus Cristi;texas;78428;2420;2019-4-3
78184;sold;385000;3;3;0.23;390 Madison Ct;North Port;Florida;34286;1708;2022-03-18
79245;for_sale;599900;5;5;0.42;254 Princess Ave;Bettendorf;Iowa;52722;3765;2016-11-21
53177;for_sale;105000;3;2;0.24;434 Court Ct;Woodward;Oklahoma;73801;2036;2007-06-29
81311;sold;575000;5;4;0.22;827 Elm Ct;Leland;North Carolina;28451;3057;2022-04-22
82978;for_sale;335000;3;2;0.18;71 Hill Ln;Temple;Texas;76502;2344;2013-05-01
22721;sold;799999;4;4;0.15;900 Baker Ln;Natick;Massachusetts;1760;3422;2022-04-05
5294;for_sale;155000;5;3;10.05;675 Circle;Hubert;Oklaoma;74442;1900;2018-4-18

- **Total records after deduplication: 95**

This deduplicated dataset removes the duplicate entries while retaining one instance of each.