

# File - 01\_extract2csv

```

1 "C:\Program Files\Anaconda3\python.exe" "D:/Program Files/JetBrains/Local anacondapy3/hai kou_pandas/
  01_extract2csv.py"
2 image directory is datasets
3 file_list is ['dwv_order_make_haikou_1.txt', 'dwv_order_make_haikou_2.txt']
4 num of file_list = 2
5 datasets\dwv_order_make_haikou_1.txt has been find!
6 -----数据概览-----
7 df.shape (99, 24)
8 df.shape after dropna (72, 24)
9 df.index.values
10 [ 0  1  2  3  4  5  6  7  9 10 11 12 14 15 16 17 18 19 21 22 23 24 25 26
11 28 29 30 31 32 33 34 35 36 37 38 39 40 43 45 46 50 51 53 55 56 57 58 61
12 63 64 67 68 69 70 71 72 73 74 78 79 81 82 84 85 87 89 90 91 94 95 96 98]
13 df.columns.values
14 ['dwv_order_make_haikou_1.order_id' 'dwv_order_make_haikou_1.product_id'
15 'dwv_order_make_haikou_1.city_id' 'dwv_order_make_haikou_1.district'
16 'dwv_order_make_haikou_1.county' 'dwv_order_make_haikou_1.type'
17 'dwv_order_make_haikou_1.combo_type'
18 'dwv_order_make_haikou_1.traffic_type'
19 'dwv_order_make_haikou_1.passenger_count'
20 'dwv_order_make_haikou_1.driver_product_id'
21 'dwv_order_make_haikou_1.start_dest_distance'
22 'dwv_order_make_haikou_1.arrive_time'
23 'dwv_order_make_haikou_1.departure_time'
24 'dwv_order_make_haikou_1.pre_total_fee'
25 'dwv_order_make_haikou_1.normal_time'
26 'dwv_order_make_haikou_1.bubble_trace_id'
27 'dwv_order_make_haikou_1.product_level'
28 'dwv_order_make_haikou_1.dest_lng' 'dwv_order_make_haikou_1.dest_lat'
29 'dwv_order_make_haikou_1.starting_lng'
30 'dwv_order_make_haikou_1.starting_lat' 'dwv_order_make_haikou_1.year'
31 'dwv_order_make_haikou_1.month' 'dwv_order_make_haikou_1.day']
32 df.head(10)
33   dwv_order_make_haikou_1.order_id  ... dwv_order_make_haikou_1.day
34 0                        17592719043682  ...                19
35 1                        17592719302995  ...                19
36 2                        17592719330238  ...                19
37 3                        17592720943629  ...                19
38 4                        17592721794216  ...                19
39 5                        17592721831669  ...                19
40 6                        17592722325070  ...                19
41 7                        17592722363412  ...                19
42 9                        17592722883936  ...                19
43 10                       17592723562993  ...                19
44
45 [10 rows x 24 columns]
46 -----经纬度最值-----
47 starting_lng_max: 110.4506
48 starting_lng_min: 110.2543
49 starting_lat_max: 20.0655
50 starting_lat_min: 19.9418
51 dest_lng_max: 110.4395
52 dest_lng_min: 110.2663
53 dest_lat_max: 20.0671
54 dest_lat_min: 19.9644
55 dest_lat_mean: 20.01860694444445
56 dest_lat_max: 20.0671
57 dest_lat_min: 19.9644
58 start_df.shape (72, 3)
59 dest_df.shape (72, 3)
60 start_df.shape after inside area (69, 3)

```

File - 01\_extract2csv

```
61 dest_df.shape after inside area (72, 3)
62 datasets\dwv_order_make_haikou_2.txt has been find!
63 -----数据概览-----
64 df.shape (99, 24)
65 df.shape after dropna (84, 24)
66 df.index.values
67 [ 0  1  2  3  5  7  8  9 10 14 15 18 19 20 21 22 23 26 27 28 29 30 31 32
68  33 34 35 36 37 38 39 40 41 42 43 44 46 47 48 49 50 51 53 54 55 56 57 58
69  60 61 62 63 64 65 66 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84
70  85 86 88 89 90 91 92 93 94 95 96 97]
71 df.columns.values
72 ['dwv_order_make_haikou_2.order_id' 'dwv_order_make_haikou_2.product_id'
73  'dwv_order_make_haikou_2.city_id' 'dwv_order_make_haikou_2.district'
74  'dwv_order_make_haikou_2.county' 'dwv_order_make_haikou_2.type'
75  'dwv_order_make_haikou_2.combo_type'
76  'dwv_order_make_haikou_2.traffic_type'
77  'dwv_order_make_haikou_2.passenger_count'
78  'dwv_order_make_haikou_2.driver_product_id'
79  'dwv_order_make_haikou_2.start_dest_distance'
80  'dwv_order_make_haikou_2.arrive_time'
81  'dwv_order_make_haikou_2.departure_time'
82  'dwv_order_make_haikou_2.pre_total_fee'
83  'dwv_order_make_haikou_2.normal_time'
84  'dwv_order_make_haikou_2.bubble_trace_id'
85  'dwv_order_make_haikou_2.product_level'
86  'dwv_order_make_haikou_2.dest_lng' 'dwv_order_make_haikou_2.dest_lat'
87  'dwv_order_make_haikou_2.starting_lng'
88  'dwv_order_make_haikou_2.starting_lat' 'dwv_order_make_haikou_2.year'
89  'dwv_order_make_haikou_2.month' 'dwv_order_make_haikou_2.day' ]
90 df.head(10)
91   dwv_order_make_haikou_2.order_id  ...  dwv_order_make_haikou_2.day
92  0                               17592880231474  ...                26
93  1                               17592880435172  ...                26
94  2                               17592880622846  ...                26
95  3                               17592880665344  ...                26
96  5                               17592880885186  ...                26
97  7                               17592881199105  ...                26
98  8                               17592881962918  ...                26
99  9                               17592882308885  ...                26
100 10                               17592882846802  ...                26
101 14                               17592883268444  ...                26
102
103 [10 rows x 24 columns]
104 -----经纬度最值-----
105 starting_lng_max: 110.4836
106 starting_lng_min: 110.1979
107 starting_lat_max: 20.0697
108 starting_lat_min: 19.9418
109 dest_lng_max: 110.4634
110 dest_lng_min: 110.1954
111 dest_lat_max: 20.0679
112 dest_lat_min: 19.9369
113 dest_lat_mean: 20.01437380952381
114 dest_lat_max: 20.0679
115 dest_lat_min: 19.9369
116 start_df.shape (84, 3)
117 dest_df.shape (84, 3)
118 start_df.shape after inside area (84, 3)
119 dest_df.shape after inside area (84, 3)
120 time use: 0.32599973678588867 s
121
```

File - 01\_extract2csv

122 Process finished with exit code 0

123