Absorption Analysis

Absorption phenomenon

```
# 新建站点信息
New Station ID : 3461
New station appear time: 2017-05-18 12:15:
# 温度 [before, after]
temperature[4.58354135064, 11.8816189938]
# 临近站点的demand变化 [before, after], 最后一项是与新建station的距离
nearStation id :309, workday[300.538461538, 281.5], holiday[286.5, 278.0],
151.42 m
nearStation id :152, workday[251.076923077, 216.8], holiday[174.5, 164.5],
178.86 m
nearStation id :319, workday[240.384615385, 287.9], holiday[198.0, 218.5],
227.37 m
```

```
# 新建站点信息
New Station ID : 3249
New station appear time: 2016-04-05 18:47:58
# 温度 [before, after]
temperature[-2.29718230487, -2.58800067035]
# 临近站点的demand变化 [before, after], 最后一项是与新建station的距离
nearStation id :437, workday[53.0, 57.4166666667], holiday[44.8, 55.0],
211.58 m
```

```
# 新建站点信息
New Station ID : 3469
New station appear time: 2017-06-08 15:03:17
# 温度 [before, after]
temperature[8.21299317513, 13.8245949074]
# 临近站点的demand变化 [before, after], 最后一项是与新建station的距离
nearStation id :3116, workday[108.33, 111.15], holiday[143.5, 127.0],
170.68 m
```

```
# 新建站点信息
New Station ID : 3436
New station appear time: 2016-09-30 17:01:38
# 温度 [before, after]
temperature[12.612604338, 5.21078793825]
# 临近站点的demand变化 [before, after], 最后一项是与新建station的距离
nearStation id :248, workday[115.9, 118.7], holiday[66.1666666667, 47.5],
206.55 m
nearStation id :328, workday[242.67, 183.14], holiday[161.5, 182.25],
306.19 m
```

```
# 新建站点信息
New Station ID : 3474
New station appear time: 2017-06-29 16:44:35
# 温度 [before, after]
temperature[12.0737418301, 17.4493122202]
# 临近站点的demand变化 [before, after], 最后一项是与新建station的距离
nearStation id :128, workday[566.75, 537.75], holiday[449.0, 409.0], 227.07
m
```

新建站点的选择方法：人工选择

其他发现：一些距离新建站点很近的旧站点，没有in / out记录，应该是旧站点被新站点取代了，这个因素需要考虑吗?

Weather Factor

Data Source : NOAA

download link : https://www.ncei.noaa.gov/orders/cdo/1135040.csv

All the days with abnormal weathre are removed.

Overall, the bad weather days account for 42.9% in 4 years (2013 - 2017).

Holiday and Workday

Holiday include : Saturday, Sunday and public holidays.

The public holidays includes: (Data Source : https://publicholidays.us/new-york/)

Date	Day	Holiday
1 Jan	Sun	New Year's Day
2 Jan	Mon	New Year Holiday
16 Jan	Mon	Martin Luther King Jr. Birthday
12 Feb	Sun	Lincoln's Birthday
13 Feb	Mon	Lincoln's Birthday Holiday
20 Feb	Mon	President's Day
29 May	Mon	Memorial Day
4 Jul	Tue	Independence Day
4 Sep	Mon	Labor Day
9 Oct	Mon	Columbus Day
10 Nov	Fri	Veterans Day Holiday
11 Nov	Sat	Veterans Day
23 Nov	Thu	Thanksgiving Day
25 Dec	Mon	Christmas Day

Method to find S* and near-Stations (未实践)

time threshold : T = 30 days

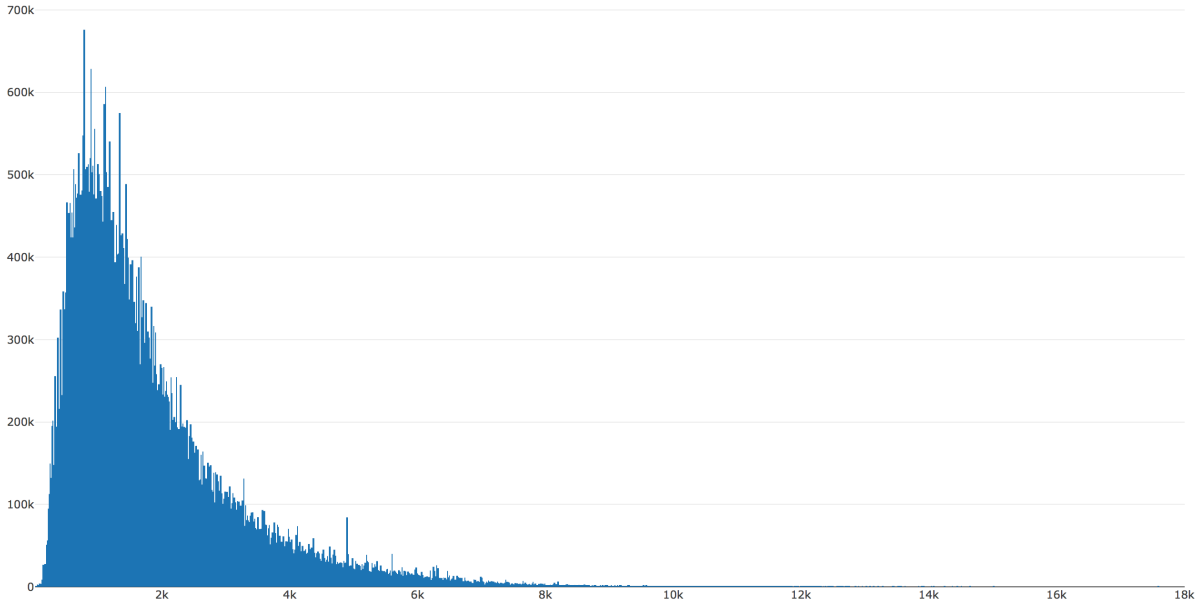
distance threshold : D = 150m (for absorption), 5000m (for stimulation)

For a station A, get the near stations set A_NN (within 150m of A)

Rule 1 : The stations in A_NN are all built at least one month earlier than A

Rule 2 : After A was built, no station is built within 150m + 5000m of A. (+5000m is to ensure the near stations are not influenced by stimulation)

Method to find satisfied new station and effected station list(未实践)



上图为骑行距离的统计图，可以用t分布做一下假设检验，来寻找距离的阈值，假设置信区间是 [150米, 6000米]，那就认为两个车站如果在150米以内，就存在竞争关系，就在这个距离阈值内寻找Absorption的现象。

在[150米, 6000米]距离之间的stations寻找stimulation的现象