Problem 1  
  
Given lists of x and y, find interpolated (within range) or extrapolated (not within range) y values for x\_input. You need to handle some special cases when there are multiple points on the same x.  
  
Problem 2   
  
Given temperatures of N towns and NYC  
  
(1) Find town with largest variation. check 1point3acres for more.  
(2) Median temperature of NYC with condition on town2 temperature  
(3) Linear regression  
(4) Which one has smallest MSE  
(5) Find best pair with smallest MSE  
  
Problem 3  
  
Same dataset as P2, approximate best subset of variables.

还是两道原题，第三道题没写，一共三小时。  
Q1 implement a linear interpolate. 有几种特殊情况列的很清楚。 我是先排序，按x排， x一样按y排，然后用binary search去找区间。   
Q2 数据处理，好像是5个小问。找median，standard divination，least square。主要用了pandas和sklearn里的linear regression

第一题是linear interpolation，我之前在地里看到过面经也没多想，但是还是花了不少时间在duplicate的情况上，要求是如果有多个相同x坐标，输入点在左边，那就按最小的y来算，输入点在右边，那就按最大的y来算，这个有点tricky，其他都还好，我审题不清还写了好多corner情况，结果回头看，根本不会出现这样的情况，还是不应该一大早做，头脑不转  
第二题是load数据分析，每个column是一个town的温度，好几个小问题，最后要fit一个linear model看哪个town的温度predict NYC的MSE最小，pandas，sklearn都可以

还是90分钟两道题 不过不是Friend Cycle 和 Longest Chain  
1. Missing Words   
    给两个字符串s和t， t是s的子序列， 要求从s中去掉t 如果t中的单词在s中出现多次 只去掉第一次出现的  
    s = 'I like eating cheese do you like cheese'  
    t = 'like cheese'  
    返回 'I eating do you like cheese'  即只去掉第一个cheese  
2. Substrings  
    给定一个字符串s 在n(n + 1) / 2 个子串中 加上筛选条件 要求以元音开始 辅音结尾 然后返回字典序最小和最大的两个串  
    元音是aeiou  
    如 'aab'返回最小是'aab' 最大是'ab'  
  
    诶 跪在这道题上 全程超时 求大神解答

2016(4-6月) **码农类General** **硕士** **全职**@**TwoSigma** - 内推 - **在线笔试** | **Other** | 在职跳槽

|  |
| --- |
| 第一个是longestchain，这个帖子里的做法就能过 http://www.1point3acres.com/bbs/thread-131978-1-1.html 第二个是一个叫什么zombie的题，就是找联通 比如下面的输入，你要输出2. check 1point3acres for more.  "1100",  "1110",  "0110",  "0001" |

friend circle  
longest chain.