CS-302 Design and Analysis of Algorithms Project

Attention

- Make sure that you read and understand each and every instruction. If you have any questions or comments you are encouraged to discuss with your colleagues and instructors.
- Start early otherwise you will struggle with the assignment.
- Project will be checked for plagiarism. If found plagiarized, both the involved parties will be awarded zero marks in the relevant evaluation instrument, all of the instruments or even an F grade in the course.
- In part II, your solution will be graded on correctness and efficiency(complexity). You are required to submit a complete algorithm which you made for each sub part of PART II along with detailed analysis of its complexity in the form of project report.
- This is individual based project.
- You don't change the formatting of given dataset file.

Objective

- How to use algorithm concepts on real life problems.

PART I (DUE 27th April 2020)

My apartment has n computers. My friend's apartment also has n computers. In each apartment, some pairs of computers are connected to each other with AcidNet cables (ignoring the routers). Each connection has a certain bandwidth (in bytes per second). My friend always brags about the speed of his computer network. He always shows me his n-by-n table that lists the bandwidths between each pair of computers. My network is slower, and I want to rebuild it. So, I want to know how I should connect my computers in order to have the same n-by-n bandwidth table.

Since I don't want to buy too many AcidNet cables, you'll need to find a solution with the minimum number of connections. You may use AcidNet cables of any integer bandwidth — they all have the same price at my local Imaginary Hardware Store.

Given a graph, you can compute the all-pairs maximum flow table, right? Now do the opposite: given an n-by-n symmetric table, find a graph with fewest edges that has the given table of all-pairs maximum flows.

Input

The first line of input gives the number of cases, N. N test cases follow. Each one is a line containing n ($0 < n \leq 200$), followed by n lines with n integers each, giving the table T.

- T[u][u] will always be 0.
- T[u][v] will always be positive and equal to T[v][u].
- $T[i][j] \leq 10000$

T[u][v] is the largest possible speed (in bytes per second) for sending information from computer u to computer v, assuming there is no other traffic on the network.

Output

For each test case, output one line containing 'Case #x:' followed by m — the number of cables I have to buy. The next m lines will each contain 3 integers u, v and w meaning that I need to connect computer u to computer v using an AcidNet cable of bandwidth w. Computers are numbered starting at 0.

If there is no solution, print 'Impossible'.

Sample Input

```
4
2
0 10
10 0
3
0 1 1
1 0 2
1 2 0
1
0
4
0 2 2 1
2 0 2 2
2 2 0 2
1 2 2 0
```

Sample Output

```
Case #1: 1
0 1 10
Case #2: 2
0 1 1
1 2 2
Case #3: 0
Case #4: Impossible
```

This question is taken from onlinejudge (link). You may know that **UVa Online Judge** is an online automated judge for programming problems. There is an automated system that evaluates the submissions. You are required to make an online account on onlinejudge and submit your code. The automated system will test the code and provide you with acceptance/rejection report.

You are required to check your code using online judge and submit the evaluation snapshot by the onlinejudge.org along with your code to the submission portal at google classroom.

Part II (Due 18th May 2020)

The **2019–20 coronavirus pandemic** is an ongoing pandemic of coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). As of 13 April 2020, more than 1.84 million cases of COVID-19 have been reported in 210 countries and territories, resulting in more than 114,000 deaths. More than 421,000 people have recovered, although there may be a possibility of reinfection.

In this project, you are required to answer queries regarding the COVID-19 pandemic. You are provided with data regarding the pandemic containing information regarding the daily new cases and cumulative cases for each country.

WHO COVID-19 global data

You must read the given dataset and answer the following queries. For each query, think very carefully about your approach in regards to both the time and space complexity of your solution.

1) On a given day, find the top 20 countries with the most confirmed cases. **(Efficiently)**

2) Find the country(s) with the highest new cases between two given dates. **(Efficiently)**

3) Find the starting and ending days of the longest spread period for a given country. The spread period is defined as the period where daily new cases tend to increase. They may contain days where new confirmed cases were relatively lower or none at all.

   For example, [5, 2, 9, 16, 11, 27, 14, 45, 11] has a longest spread period elapsing 7 days from day 2 (2) to day 8 (45). **(Efficiently)**

4) Find the longest daily death toll decrease period for a given country. They may contain days where new confirmed deaths were relatively higher.

   For example, [9, 5, 1, 16, 11, 23, 8, 3, 27, 14, 45, 11] has a longest daily death toll decrease period of 4 {16, 11, 8, 3} (**Efficiently)**

5) You have decided to help with the relief effort by collaborating with an organization that distributes essential supplies to affected regions. However, they must consider the overhead cost for each country (distance, shipping etc.). They have assigned a score to each country, which is simply the total number of active cases on the latest day (they prefer to help countries with higher cases). A country may or may not be selected for aid. They have gathered their projected costs for each country and have tasked you with finding the highest possible score attainable as well as the countries selected given a budget of 300. **(Efficiently)**

6) We wish to compare the response of any two countries against this virus. To do this, we must compare how similar the change of their daily active cases is. The similarity is measured by the longest number of days their daily active cases share similar values. However, since it is extremely unlikely that two countries have the exact same values for a sequence of days, we define a compare threshold K such that any two countries are said to have 'similar' daily active cases on a given day if their active cases differ <= K.
   For example, C1 = [0, 0, 0, 1, 6, 10, 17, 27, 48, 94] & C2 = [3, 4, 8, 14, 31, 32, 49] and K = 10

similarity (C1, C2) = 5 (C1[3:7] = C2[0:4])

Given two countries and the compare threshold K, find their similarity. **(Efficiently)**

Your project should be a menu driven application that solves the above queries. You may use either C++ or Python. Solutions will be graded on correctness, complexity and coding practices in that order.