Problem 116: Data Lockdown

Difficulty: Easy

Author: Richard Green, Whiteley, Hampshire, United Kingdom

Originally Published: Code Quest Australia 2019

Problem Background

Corporations around the world deal with a lot of information that they need to keep relatively secret. If information about business proposals, active projects, or employee records gets released to the public - either deliberately or accidentally - it can have a serious impact on a company's ability to do business. In extreme cases, it can leave the company subject to serious legal penalties, even if the release of data wasn't directly the fault of the company.

As a result, many companies, including Lockheed Martin, put considerable effort into making sure that their proprietary information is kept secure. IT departments may block access to certain websites, prevent certain types of files from being attached to emails, and always make sure that anti-virus software is kept up to date and running. Despite all these efforts, data breaches remain possible, and constant vigilance is needed to ensure that breaches are prevented, or failing that, contained before they can create a serious problem.

Problem Description

You're working with Lockheed Martin's Enterprise Business Services company to monitor internet traffic by Lockheed employees. The proxy Lockheed uses maintains a log showing what website someone is trying to access, and how much data is being sent in their request. Small requests are normal; internet browsers have to send websites a small amount of data to let them know what they want to see. However, very large requests may be an indication that someone is trying to upload a file to an external website... which could be a data breach.

You need to write a tool that will automatically flag such requests for additional review. The proxy logs will indicate the website being accessed and the amount of data being transferred in the request (in kilobytes). Requests to Lockheed websites - which all end with ".lmco.com" - can be ignored. However, any request to another website that transfers more than a megabyte of data (1000 kilobytes) should be reviewed.

Sample Input

The first line of your program's input, received from the standard input channel, will contain a positive integer representing the number of test cases. Each test case will include:

- A line containing a positive integer, X, indicating the number of records in the log

© 2019, 2020 Lockheed Martin Corporation. All Rights Reserved.

- X lines, each representing an entry in the proxy log. Entries contain:
    - The URL of the website being accessed. This URL may contain lowercase letters, numbers, and periods
    - A space
    - A positive integer representing the amount of data being sent to the website, in kilobytes

```
1
6
www.google.com 128
www.dropbox.com 1424
www.cnn.com 142
codequest.global.lmco.com 212
www.australia.gov.au 4120
filetransfer.us.lmco.com 2186
```

Sample Output

For each test case, your program must print the log entries that require additional review, following the rules described above. Log entries should be printed in the order received in the input.

```
www.dropbox.com 1424
www.australia.gov.au 4120
```

© 2019, 2020 Lockheed Martin Corporation. All Rights Reserved.