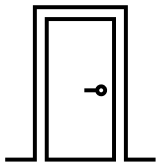
Hackathon Rules & Submission Process

Rules

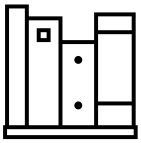
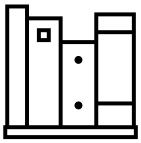
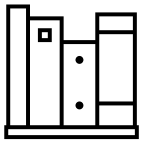
1. No AI Assistance is allowed. Your code will be reviewed for originality.
2. Any programming language is accepted, sample code is provided in python to process an unknown number of input files from the input directory.
   2 input and output files will be visible for each part, your solutions will be tested on more input files than are available to you.
3. You are responsible for submitting your code correctly using the below guidelines.

Submission Guidelines

1. Your final submission must be a single ZIP file containing all your code for Parts 1, 2 and 3.
2. Once your zip is ready, email it to the following email addresses:
   kevin.bradley@mycit.ie
   kevin.bradley@netapp.com
   Ignacio.Castineiras@mtu.ie
   Dylan.Smyth@mtu.ie



WELCOME to the NetApp Hackathon 2025!

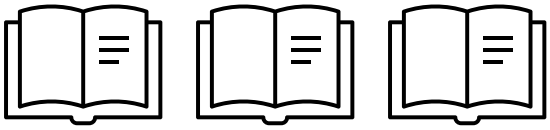




Hey! You, between the shelves!

Welcome, welcome, fellow wanderer! Come closer!

You have been selected to go on this exhilarating journey of *library management!*

Oh?



Doesn't sound fun? *???*

Nonsense!

Here at the **NetApp Library**, we pride ourselves on our innovative technologies and the most efficient library system in the lands far and wide. *What is it, you might wonder?* Well, that is a fabulous question! I am thrilled to hear of your growing interest!

Have I mentioned that you are now the official **Librarian**?

What happened to our previous librarian? He disappeared under mysterious circumstances, and we desperately need a replacement!

That's you!

Congratulations!

But wait, there's more! This isn't just any ordinary library. The **NetApp Library** has a mind of its own, and it loves storage and order. If you don't finish the tasks in time, well... let's just say the library has a way of expressing its displeasure. How do you think we lost... eh, misplaced the last librarian?

Books might start flying, shelves could rearrange themselves, and you might find yourself in a maze...

...in the endless corridors...

...organising books for all eternity.

We don't want you to disappear as well, *do we?*

Your mission, should you choose to accept it (*and you really should, because the library has already chosen* ***you***), is to reorganize the library before it gets too... *cranky*. Each task in the library holds a unique challenge, and only the bravest and most clever will prevail.

So, gear up, brave **Librarian**!

Your first task awaits you.

Remember, the clock is ticking, and the library is always watching.

Tick, tick, tick! Don't get nervous now!

Go on then, don't get any paper cuts!

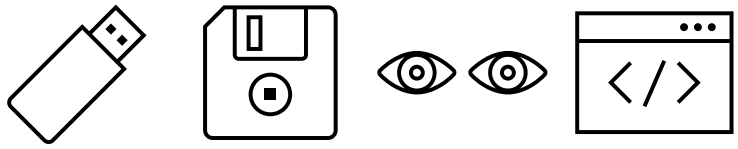
Good luck!

You'll need it...

1

Part 1: Ranking Books in a Digital Library

I told you about our fantastical digitisation efforts, haven't I?



We love technology at the **NetApp Library**.

Let's talk you through your first task, brave **Librarian**:

You are designing a system to rank books in a digital library based on how often they are borrowed. The library has a hot shelf that can store only a limited number of books represented by X.

Books that are borrowed the most and spend the least amount of time sitting idle (waiting between loans) should be prioritised for the hot shelf.

Your Task

1. Count how many times each book has been borrowed (check-out/check-in pairs).
2. Calculate idle time (time between a book being returned and its next checkout).
3. Sort the books:
    a. First, by most borrows.
    b. If the number of borrows is the same, rank by least idle time.
    c. If both are the same, sort alphabetically.
4. Additional metrics are to be tracked, these are explained below the output format

Input Format

1. First line → The number of books to keep on the hot shelf.
2. Each book starts with its title.
3. Dates are listed in pairs (checkout, return).
4. Books are separated by a blank line.

Example Input File

```
2
Book_1
```

```
12/02/2024
14/02/2024
17/02/2024
18/02/2024
28/02/2024
04/03/2024

Book_2
20/02/2024
25/02/2024
28/02/2024

Book_3
11/02/2024
15/02/2024
18/02/2024
19/02/2024
22/02/2024
26/02/2024
```

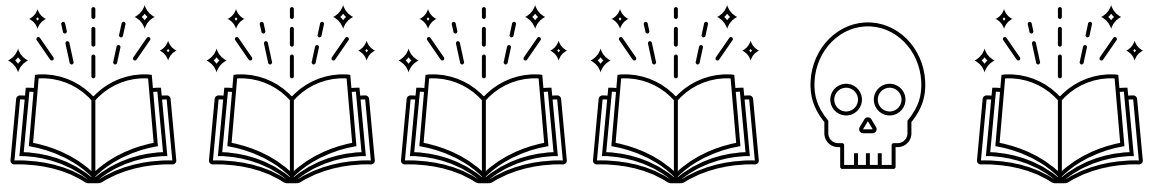
Expected Output Format

```
Book Name:In/Out Pairs,Idle Time,Total Borrowed,Longest
Borrow,Shortest Idle
Book_3:3,6,9,4,3
Book_1:3,13,8,5,3
```

Explanation of the Calculations

- Total Borrowed Time = sum of all borrow durations.
- Longest Single Borrow = max single borrowing period.
- Shortest Idle Time = shortest time between return & next checkout.

Good luck, brave **Librarian**! The library is counting on you!

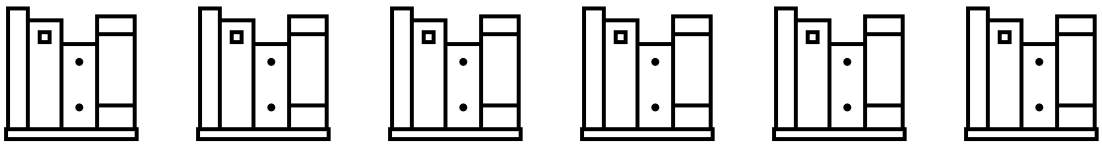


2

Part 2: Ranking Books in a 2D Library Layout

Hm.

Hmmmm... **Interesting**...



I like the way your mind is thinking. I knew you were the right person for this task, that's why our library chose **you**.

Are you ready for your second task, brave **Librarian**?

The library now organises books in a 2D matrix (rows & columns).

Your Task

1. Rank books based on popularity (number of borrows) and idle time.
2. Extract the top-ranked books into a hot shelf (also a matrix).
3. Print the hot shelf in a matrix format.

Input Format

1. First line → Library's current matrix size (rows & columns).
2. Second line → Hot shelf dimensions (rows & columns).
3. The remaining lines contain book borrowing data, following the same format as Part 1.

Example Input File

```
2 2
1 2

Book1
23/01/2024
01/02/2024
24/02/2024
02/03/2024
```

```
25/03/2024
04/04/2024
07/05/2024

Book2
04/01/2024
12/01/2024
27/01/2024
04/02/2024
22/02/2024
23/02/2024
24/03/2024

Book3
06/01/2024
08/01/2024
14/02/2024

Book4
17/01/2024
26/01/2024
01/03/2024
02/03/2024
21/03/2024
28/03/2024
18/04/2024
```

Expected Output Format:

```
Hot Shelf:
Book2     Book4
```

Don't let the **NetApp Library** wait too long...

3

Part 3: Hybrid Storage Optimization

Ah, there you are, brave **Librarian**!

I see you have successfully completed the second task. The library is pleased with your progress. You truly have the mind of a **master organiser**.

But there's no time to rest on your laurels;

the library has more challenges in store for you.

You might have noticed that the **NetApp Library** has many similarities to data storage. Maybe our library is more digitised than you thought?

There might be certain similarities to **NetApp**... Let's look at some more details for your third task, brave **Librarian**:

Large-scale data solution companies like NetApp frequently use two types of storage:

1. NVMe Drives – High-speed, low-latency storage for frequently accessed data.
2. Spinning Drives – Slower but high-capacity storage for less frequently accessed data.

Your task is to design a storage management system that optimally places files across these storage types to minimize read latency, while ensuring that the NVMe drive does not exceed its capacity.

<u>Your Task</u>

1. Rank files based on their importance using:
   - Access Frequency – Files accessed more frequently are prioritized.
   - Recency – More recently accessed files are prioritized.
   - Size Efficiency – Larger files are deprioritized unless frequently accessed.
2. Assign files to storage tiers:
   a. Store the highest-ranked files on NVMe storage until its capacity is full.

    b. Store remaining files on Spinning Drives.
3. Simulate file access latency:
    a. If a file is in NVMe → Use NVMe Latency.
    b. If a file is in Spinning Drive → Use Spinning Drive Latency.
    c. Compute the average read latency based on the provided Access Pattern.

Input Format

1. First line → NVMe storage capacity in MB.
2. Second line → NVMe access latency (in milliseconds).
3. Third line → Spinning drive latency (in milliseconds).
4. Each file entry:
    a. File name
    b. File size (MB)
    c. List of access timestamps (epoch time in `<seconds>.<milliseconds>` format).
5. Access Pattern:
    a. A list of file names representing a sequence of file requests.

Example Input File

```
500  # NVMe capacity in MB
5 ms  # NVMe latency
50 ms  # Spinning drive latency

FileA
200
1704137428.410
1704141120.620

FileB
300
1704069856.847
1704096289.467


FileC
100
```

```
1704074677.906
1704078441.952
1704137751.333

Access Pattern:
FileA
FileB
FileC
FileA
```

Expected Output Format

The output should list:

1. Which files are stored in NVMe vs Spinning Drives.
2. The average read latency for the given access pattern.

Example Output

```
NVMe Storage:
FileC (100 MB)
FileA (200 MB)

Spinning Drive Storage:
FileB (300 MB)

Average Read Latency: 16.25 ms
```

How the Ranking Works

Each file is assigned a priority score based on the following formula:

```
Score = (3 * Access Frequency) - (2 * File Size in MB) + (1 / (1 +
Seconds Since Most Recent Access))
```

Where:

- Access Frequency = Number of access timestamps of a file.

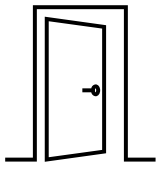
- File Size (MB) = Larger files are deprioritized unless frequently accessed.
- Recency Factor:
    - For each file, find its last recorded access timestamp. Then, subtract it from the most recent access timestamp across all files to compute the time difference.
    - A smaller value means the file has been accessed more recently.

Files are ranked by score, and the highest-scoring files are placed in NVMe first until capacity is full. If two files have the same score, they are ranked alphabetically.

Average read latency should be to 2 decimal places.

The library's – and NetApp's – efficiency depends on your keen judgment and swift action.

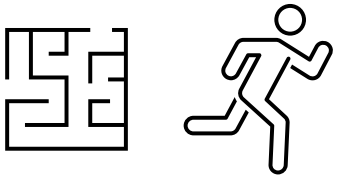
And don't get it wrong, you don't want to anger the **NetApp Library**!



Conclusion: Escaping the Maze!

Brave Librarian!

You have navigated the labyrinth of tasks with unparalleled skill and precision. The **NetApp Library**'s shelves are now in perfect order, and its data storage is optimised to perfection. Your coding and organisational prowess have not only impressed the library but have also



ensured your escape from its ever-shifting maze.

Congratulations! You have escaped the maze!

Your amazing coding and organizational skills have saved you from an eternity of wandering the endless corridors. The books and data are now in perfect harmony, a testament to your dedication and expertise.

As you step out of the library, you might realise that this journey was more than just a series of tasks. It was a connection between the physical world of books and the digital realm of data storage. Just as the library relies on efficient organization and quick access to information, so does **NetApp's** data storage solutions. Both require a keen mind, a strategic approach, and the ability to adapt to ever-changing demands.

Thank you for your hard work and determination. The **NetApp Library** is now a model of efficiency, and you, brave **Librarian**, are its hero.

Good luck in your future endeavours, and always remember: the skills you've honed here will serve you well in any challenge you face.

Farewell, and may your paths always be clear and well-organised!

Maybe there will even be prizes in your near future?