## fetch-data.sh logs explanation

These logs include timestamps to mark the **start and end**, aiding in debugging and tracking progress. Messages also detail the **downloading of data files, clearing directories, successful or unsuccessful file downloads, and what stage program is in right now**, enabling users to monitor script actions and detect issues promptly. In summary, the logs enhance script management and troubleshooting by documenting essential actions.

```
[Info] Sat Oct  7 20:52:46 BST 2023
[Info] Downloading data from https://cs1007web.teaching.cs.st-andrews.ac.uk/Coursework/A02/data/filelist.txt into files
[Info] Clearing contents of 'data' directory.
[Info] Clearing contents of 'out' directory.
[Info] Downloading file from https://cs1007web.teaching.cs.st...
[Info] Downloaded file from https://cs1007web.teaching.cs.st...
...
[Info] Program finished
[Info] Sat Oct  7 20:52:57 BST 2023
```

## Evaluation of part 2 script

The primary objective of this script is to **identify the five largest files**. The 'if' condition efficiently verifies the correct number of arguments; however, it does not validate the directory path, as this responsibility is delegated to the 'du' command. Let's delve deeper into the last line:

| | |
|---|---|
| `du -h $1/data/*` | This code segment **performs a space occupancy analysis** of files located in the data folder. These files were downloaded through the fetch-data script and are presented in a human-readable format (-h argument) using the disk usage command. |
| `sed 's/\.\d*//g'` | This operation involves the **elimination of decimal points**, which may result in potential data loss. |
| `sed 's/K/000/g'` | **Substitute 'K' with '000'**. This transformation is specific to kilobytes, and it may not function correctly for larger file sizes, such as megabytes. |
| `sort` | It organizes file sizes alphabetically, which may not accurately reflect their numerical values, making it **unsuitable for numerical comparisons**. |
| `head -n5` | Retrieves the initial **five lines of data**. |