

Challenge Statement

Imagine a file in the following fixed format:

`<unique record identifier><white_space><numeric value>`

e.g.

1426828011	9
1426828028	350
1426828037	25
1426828056	231
1426828058	109
1426828066	111

.
.
.

Write a program that reads from 'stdin' the contents of a file, and optionally accepts the absolute path of a file from the command line. The file/stdin stream is expected to be in the above format. The output should be a list of the unique ids associated with the X-largest values in the rightmost column, where X is specified by an input parameter. For example, given the input data above and X=3, the following would be valid output:

1426828028
1426828066
1426828056

Note that the output does not need to be in any particular order. Multiple instances of the same numeric value count as distinct records of the total X. So if we have 4 records with values: 200, 200, 115, 110 and X=2 then the result must consist of the two IDs that point to 200 and 200 and no more.

Your solution should take into account extremely large files.

What to return back to us

1. Your code in the language of your preference. If it's in Python, Java or C++ it's preferable for us.
2. Include in your code comments about your solution's algorithmic complexity for both time and memory.
3. Include instructions on how to build and run your code in a README file. Please include operating system information if necessary.
4. Provide tests for your code to illustrate it works and it's robust.
5. Please zip everything in a directory named `yourfirst.lastname/` and return via email.
6. In your email response please let us know roughly how many hours you spent for this exercise.