

Forum:[Hack Forums](#)

Thread: [Remove everything BUT duplicate values?](#)

Renovatum - February 27, 2018, 12:30 pm

I have a list of about 400k people. I want to delete the people who don't appear twice in this list, is there a way to do this without downloading additional programs?

IE a batch or java script that'll check the list and delete those who don't appear more than once.

Thanks

Gorzoid - February 27, 2018, 5:10 pm

I'd say presort the array, if its names then do it alphabetically, then iterate through and if `arr[n] == are[n-1]`, push it to a new array and make sure to skip subsequent equal values. Will result in better complexity than above answer assuming you're sorting algorithm is at least $O(n \log n)$

m0noxide - March 1, 2018, 12:01 am

(Feb 27, 2018 06:10 PM)Gorzoid Wrote: [CITE](#) I'd say presort the array, if its names then do it alphabetically, then iterate through and if `arr[n] == are[n-1]`, push it to a new array and make sure to skip subsequent equal values. Will result in better complexity than above answer assuming you're sorting algorithm is at least $O(n \log n)$ Your algorithm is $\max(O(n \log n) + O(n))$ but it could be a lot better by being just iterating through the 400k list only twice so $O(2n)$, which most practitioners would reduce to $O(n)$.

You can solve this by iterating through the list with a hashmap(key value pairs), key being the name and the value being the number of times seen.

If the name is not in the set, add it to the hashmap with "name" -> 1.

If it is in the set, increment the value matched to the key, then check if the key value is 2, if it is, then remove it from the hashmap.

Reiterate through hashmap and you now have the names to be deleted.

Here is a javascript implementation, you'll have to pass it an array of names though.

Code function remaining(names){

var remaining = {};

const limit = 2;

for (var i = 0; i < names.length; i++){

if (remaining[names[i]]){

remaining[names[i]]++;

} else {

remaining[names[i]] = 1;

}

}

for (var key in remaining) {

if (remaining.hasOwnProperty(key)){

if (remaining[key] != limit){

```
        delete remaining[key];
    }
}
}

return remaining;
}
```

```
console.log(remaining(["a", "b", "b", "a", "c"]))
```

You may have to modify my above code if names actually appear more than twice; I wasn't sure if that was possible but hopefully the above helps you out OP.

Renovatum - *March 2, 2018, 12:48 pm*

This got real confusing so I just used an addon.

Sorry for wasting you guys's times