

# Re: How to cleanse a data file like this in VbScript

**Source:**

<http://www.tech-archive.net/Archive/Scripting/microsoft.public.scripting.vbscript/2005-03/1057.html>

---

**From:** McKirahan (*News\_at\_McKirahan.com*)

**Date:** 03/18/05

Date: Thu, 17 Mar 2005 19:00:58 -0600

"Karen Middleton" <karenmiddleol@yahoo.com> wrote in message news:a5fd468a.0503171502.fbaeccc@posting.google.com...

> *Hello All*

>

> *I get the following data file which is basically a report generated  
> from a legacy system output as a text file it as some page headers and  
> column  
> headers I need to get rid of the page headers and column headers and  
> read only the columns of data can any of you please provide me with a  
> sample VB*

> *script code that can accomplish the cleaning. So basically the script  
> should be able to read the file and skip all lines that do not have a  
> customer in the first field because a page header or footer line will  
> not have a customer # in it.*

>

> *The file structure is as follows:*

>

>

> *[2w10:12 PM 02-2005 XYZ Com Ltd.  
> 34715 ABC CO 15-Mar-05  
> New York :NIGHTM VENDOR  
> SALES ANALYSIS  
> PAGE 1*

>

> *-----Net Sales----- Whse Stock-----Other  
> Sales-----  
> Customer CUR YTD CUR YTD CUR  
> YTD*

>

> -----

> *102509 53 53  
> 102582 16 16 16 16  
> 10770632 157 32 157*

>

>

>

microsoft.public.scripting.vbscript: Re: How to cleanse a data file like this in VbScript

> *In the above file all I need is the output file in the following form:*  
>  
> *102509 53 53*  
> *102582 16 16 16 16*  
> *10770632 157 32 157*  
>  
> *It must take the data in the above report form with header and footers*  
> *and column headers and remove them and generate the above file which*  
> *is devoid of headers, footers and column headers the rest must be*  
> *removed sometimes the report is so big multiple pages could appear*  
> *each with a page header please advise the best way of handling and*  
> *cleansing this kind of data and picking up only things needed and*  
> *discarding the unwanted data.*  
>  
> *Please share the VB script code how to cleanse and generate the above*  
> *output file given the data file which is also described above.*  
>  
> *Thanks*  
> *Karen*

Is the eventual (unstated) goal to get the extracted data into a CSV file, spreadsheet, or database?

A "report file scraping" script will break if unexpected data is encountered.

What you posted wrapped lines so we can't really see what you do.

To do this properly one would have to examine a complete report and preferably several reports to ensure consistency between them.

As it's not trivial, if you'd like to hire someone to do this for you ...