

How to remove empty pages in VB.NET with ByteScout PDF Extractor SDK

How to code in VB.NET to remove empty pages with this step-by-step tutorial

The coding tutorials are designed to help you test the features without need to write your own code. ByteScout PDF Extractor SDK: the SDK that helps developers to extract data from unstructured documents, pdf, images, scanned and electronic forms. Includes AI functions like automatic table detection, automatic table extraction and restructuring, text recognition and text restoration from pdf and scanned documents. Includes PDF to CSV, PDF to XML, PDF to JSON, PDF to searchable PDF functions as well as methods for low level data extraction. It can remove empty pages in VB.NET.

You will save a lot of time on writing and testing code as you may just take the VB.NET code from ByteScout PDF Extractor SDK for remove empty pages below and use it in your application. In order to implement the functionality, you should copy and paste this code for VB.NET below into your code editor with your app, compile and run your application. Implementing VB.NET application typically includes multiple stages of the software development so even if the functionality works please test it with your data and the production environment.

Our website provides trial version of ByteScout PDF Extractor SDK for free. It also includes documentation and source code samples.

FOR MORE INFORMATION AND FREE TRIAL:

[Download Free Trial SDK \(on-premise version\)](#)

[Read more about ByteScout PDF Extractor SDK](#)

[Explore API Documentation](#)

[Get Free Training for ByteScout PDF Extractor SDK](#)

[Get Free API key for Web API](#)

[visit www.Bytescout.com](http://www.Bytescout.com)

Source Code Files:

```
Imports System.IO
Imports Bytescout.PDFExtractor

''' <summary>
''' The example demonstrates detection of empty pages, splitting the document to separate
''' pages excluding empty ones, then combine parts back to a single document.
''' </summary>
Module Module1

    Dim InputFile = ".\sample.pdf"
    Dim OutputFile = ".\result.pdf"
    Dim TempFolder = ".\temp"

    Sub Main()

        ' Create and setup Bytescout.PDFExtractor.TextExtractor instance
        Dim extractor As New TextExtractor("demo", "demo")

        ' Load PDF document
        extractor.LoadDocumentFromFile(InputFile)

        ' List to keep non-empty page numbers
        Dim nonEmptyPages = New List(Of String)()

        ' Iterate through pages
        For pageIndex = 0 To extractor.GetPageCount() - 1
            ' Extract page text
            Dim pageText = extractor.GetTextFromPage(pageIndex)
            ' If extracted text is not empty keep the page number
            If pageText.Length > 0 Then
                nonEmptyPages.Add((pageIndex + 1).ToString())
            End If
        Next

        ' Cleanup
        extractor.Dispose()

        ' Form comma-separated list of page numbers to split ("1,3,5")
        Dim ranges As String = String.Join(",", nonEmptyPages)

        ' Create Bytescout.PDFExtractor.DocumentSplitter instance
        Dim splitter = new DocumentSplitter("demo", "demo")
        splitter.OptimizeSplittedDocuments = true

        ' Split document by non-empty in temp folder
        Dim parts = splitter.Split(InputFile, ranges, TempFolder)

        ' Cleanup
        splitter.Dispose()
```

```
' Create ByteScout.PDFExtractor.DocumentMerger instance
Dim merger = New DocumentMerger("demo", "demo")

' Merge parts
merger.Merge(parts, OutputFile)

' Cleanup
merger.Dispose()

' Delete temp folder
Directory.Delete(TempFolder, true)

' Open the result file in default PDF viewer (for demo purposes)
Process.Start(OutputFile)
```

End Sub

End Module

VIDEO

https://www.youtube.com/watch?v=s28W3_KMraU

ON-PREMISE OFFLINE SDK

[60 Day Free Trial](#) or [Visit ByteScout PDF Extractor SDK Home Page](#)
[Explore ByteScout PDF Extractor SDK Documentation](#)
[Explore Samples](#)
[Sign Up for ByteScout PDF Extractor SDK Online Training](#)

ON-DEMAND REST WEB API

[Get Your API Key](#)
[Explore Web API Docs](#)
[Explore Web API Samples](#)

[visit www.ByteScout.com](http://www.ByteScout.com)

[visit www.PDF.co](http://www.PDF.co)

