

[www.bytescout.com](http://www.bytescout.com)

## How to extract line items from tables on multiple pages in VB.NET and ByteScout Document Parser SDK

ByteScout Document Parser SDK: the template based data extraction engine can process thousands and millions of documents per day based on templates. Can work online and offline for privacy focused tasks. Templates can be supported and updated without any programming or technical knowledge required. Output is generated in JSON, CSV, XML and custom format if required.

FOR MORE INFORMATION AND FREE TRIAL:

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

[Read more about ByteScout Document Parser SDK](#)

[Explore API Documentation](#)

[Get Free Training for ByteScout Document Parser SDK](#)

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

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

## Source Code Files:

ExtractLineItemFromTableOnMultiplePages.sln

```
Microsoft Visual Studio Solution File, Format Version 12.00
# Visual Studio 15
VisualStudioVersion = 15.0.27703.2026
MinimumVisualStudioVersion = 10.0.40219.1
Project("{F184B08F-C81C-45F6-A57F-5ABD9991F28F}") = "ExtractLineItemFromTableOnMultiplePages", "ExtractLineItemFromTableOnMultiplePages.csproj", "{F184B08F-C81C-45F6-A57F-5ABD9991F28F}"
EndProject
Global
    GlobalSection(SolutionConfigurationPlatforms) = preSolution
        Debug|Any CPU = Debug|Any CPU
        Release|Any CPU = Release|Any CPU
    EndGlobalSection
    GlobalSection(ProjectConfigurationPlatforms) = postSolution
```

```

    {A5FA1293-C15B-4249-A771-ED4301658454}.Debug!Any CPU.ActiveCfg = Debug!Any CPU
    {A5FA1293-C15B-4249-A771-ED4301658454}.Debug!Any CPU.Build.0 = Debug!Any CPU
    {A5FA1293-C15B-4249-A771-ED4301658454}.Release!Any CPU.ActiveCfg = Release!Any CPU
    {A5FA1293-C15B-4249-A771-ED4301658454}.Release!Any CPU.Build.0 = Release!Any CPU
EndGlobalSection
GlobalSection(SolutionProperties) = preSolution
    HideSolutionNode = FALSE
EndGlobalSection
GlobalSection(ExtensibilityGlobals) = postSolution
    SolutionGuid = {DCA85560-1523-40E0-917A-5BB3AA7E031F}
EndGlobalSection
EndGlobal

```

## ExtractLineItemFromTableOnMultiplePages.vbproj

```

<?xml version="1.0" encoding="utf-8"?>
<Project ToolsVersion="15.0" xmlns="http://schemas.microsoft.com/developer/msbuild/2003">
  <Import Project="$(MSBuildExtensionsPath)\$(MSBuildToolsVersion)\Microsoft.Common.props" Condition="Exists('$(MSBuildExtensionsPath)\$(MSBuildToolsVersion)\Microsoft.Common.props')" />
  <PropertyGroup>
    <Configuration Condition="'$(Configuration)' == ''">Debug</Configuration>
    <Platform Condition="'$(Platform)' == ''">AnyCPU</Platform>
    <ProjectGuid>{A5FA1293-C15B-4249-A771-ED4301658454}</ProjectGuid>
    <OutputType>Exe</OutputType>
    <StartupObject>ExtractLineItemFromTableOnMultiplePages.Module1</StartupObject>
    <RootNamespace>ExtractLineItemFromTableOnMultiplePages</RootNamespace>
    <AssemblyName>ExtractLineItemFromTableOnMultiplePages</AssemblyName>
    <FileAlignment>512</FileAlignment>
    <MyType>Console</MyType>
    <TargetFrameworkVersion>v4.0</TargetFrameworkVersion>
  </PropertyGroup>
  <PropertyGroup Condition="'$(Configuration)|$(Platform)' == 'Debug|AnyCPU' ">
    <PlatformTarget>AnyCPU</PlatformTarget>
    <DebugSymbols>>true</DebugSymbols>
    <DebugType>full</DebugType>
    <DefineDebug>>true</DefineDebug>
    <DefineTrace>>true</DefineTrace>
    <OutputPath>bin\Debug</OutputPath>
    <NoWarn>42016,41999,42017,42018,42019,42032,42036,42020,42021,42022</NoWarn>
  </PropertyGroup>
  <PropertyGroup Condition="'$(Configuration)|$(Platform)' == 'Release|AnyCPU' ">
    <PlatformTarget>AnyCPU</PlatformTarget>
    <DebugType>pdbonly</DebugType>
    <DefineDebug>>false</DefineDebug>
    <DefineTrace>>true</DefineTrace>
    <Optimize>>true</Optimize>
    <OutputPath>bin\Release</OutputPath>
    <NoWarn>42016,41999,42017,42018,42019,42032,42036,42020,42021,42022</NoWarn>
  </PropertyGroup>
  <PropertyGroup>
    <OptionExplicit>On</OptionExplicit>
  </PropertyGroup>
  <PropertyGroup>
    <OptionCompare>Binary</OptionCompare>
  </PropertyGroup>
  <PropertyGroup>
    <OptionStrict>Off</OptionStrict>
  </PropertyGroup>
  <PropertyGroup>
    <OptionInfer>On</OptionInfer>
  </PropertyGroup>
  <ItemGroup>
    <Reference Include="ByteScout.DocumentParser, Version=1.0.0.100, Culture=neutral, PublicKeyToken=f7dd1bd9" />
    <SpecificVersion>False</SpecificVersion>
    <HintPath>c:\Program Files\ByteScout Document Parser SDK\net40\ByteScout.DocumentParser.dll</HintPath>
    </Reference>
    <Reference Include="System" />
  </ItemGroup>

```

```

<Reference Include="System.Core" />
</ItemGroup>
<ItemGroup>
  <Import Include="Microsoft.VisualBasic" />
  <Import Include="System" />
  <Import Include="System.Collections" />
  <Import Include="System.Collections.Generic" />
  <Import Include="System.Data" />
  <Import Include="System.Diagnostics" />
  <Import Include="System.Linq" />
  <Import Include="System.Xml.Linq" />
</ItemGroup>
<ItemGroup>
  <Compile Include="Module1.vb" />
</ItemGroup>
<ItemGroup>
  <None Include="..\..\MultiPageTable.pdf">
    <Link>MultiPageTable.pdf</Link>
    <CopyToOutputDirectory>Always</CopyToOutputDirectory>
  </None>
  <None Include="..\..\_Sample Templates\MultiPageTable-template1.yml">
    <Link>MultiPageTable-template1.yml</Link>
    <CopyToOutputDirectory>Always</CopyToOutputDirectory>
  </None>
  <None Include="..\..\_Sample Templates\MultiPageTable-template2.yml">
    <Link>MultiPageTable-template2.yml</Link>
    <CopyToOutputDirectory>Always</CopyToOutputDirectory>
  </None>
</ItemGroup>
<Import Project="$(MSBuildToolsPath)\Microsoft.VisualBasic.targets" />
</Project>

```

## Module1.vb

```
Imports ByteScout.DocumentParser
```

```
' This example demonstrates extracting line items from tables on multiple pages with two different approaches.
' See comments in the code of templates.
```

```
Module Module1
```

```
  Sub Main()
```

```
    Dim inputDocument As String = ".\MultiPageTable.pdf"
    Dim template1 As String = ".\MultiPageTable-template1.yml"
    Dim template2 As String = ".\MultiPageTable-template2.yml"
```

```
    ' Process using template-1
```

```
    Using documentParser As New DocumentParser("demo", "demo")
```

```
        Console.WriteLine("Loading template 1...")
        documentParser.AddTemplate(template1)
```

```
        Console.WriteLine("Template 1 loaded.")
        Console.WriteLine()
```

```
        Console.WriteLine($"{code}Parsing ""{inputDocument}""...")
        Console.WriteLine()
```

```
        ' Parse document data in JSON format
```

```
        documentParser.ParseDocument(inputDocument, "result1.json", OutputFormat.JSON)
```

```
        Console.WriteLine("Parsing results saved to `result1.json`.")
        Console.WriteLine()
```

```
    End Using
```

```
' Process using template-2
Using documentParser As New DocumentParser("demo", "demo")

    Console.WriteLine("Loading template 2...")
    documentParser.AddTemplate(template2)

    Console.WriteLine("Template 2 loaded.")
    Console.WriteLine()

    Console.WriteLine({code}quot;Parsing ""{inputDocument}""...")
    Console.WriteLine()

' Parse document data in JSON format
documentParser.ParseDocument(inputDocument, "result2.json", OutputFormat.JSON)

Console.WriteLine("Parsing results saved to `result2.json`.")
Console.WriteLine()

End Using

Console.WriteLine()
Console.WriteLine("Press any key to continue...")
Console.ReadLine()

End Sub

End Module
```

---

## VIDEO

<https://www.youtube.com/watch?v=MG5FftWWSVE>

## ON-PREMISE OFFLINE SDK

[60 Day Free Trial](#) or [Visit ByteScout Document Parser SDK Home Page](#)  
[Explore ByteScout Document Parser SDK Documentation](#)  
[Explore Samples](#)  
[Sign Up for ByteScout Document Parser 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)

