How to Convert HTML to PDF in Python

Learn how you can convert HTML pages to PDF files from an HTML file, URL or even HTML content string using wkhtmltopdf tool and its pdfkit wrapper in Python.

Abdou Rockikz · 7 min read · Updated Jul 2022 · PDF File Handling

__Disclosure__: This post may contain affiliate links, meaning when you click the links and make a purchase, we receive a commission.

There are a lot of online tools that provide converting HTML to PDF documents, and most of them are free. In this tutorial, you will learn how you can do that with Python.

We will use the wkhtmltopdf **tool, an open-source command-line utility that renders HTML into PDF using the Qt WebKit rendering engine.**

Here is the table of contents of this tutorial:

To get started, we have to install wkhtmltopdf **tool and its** `pdfkit` **wrapper in Python.**

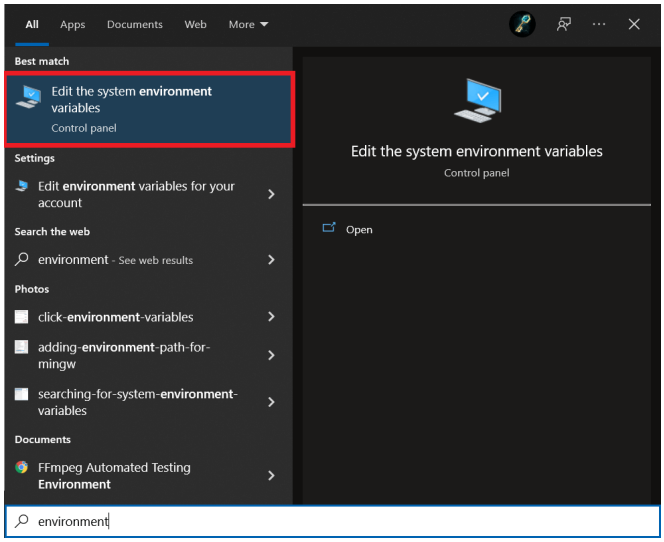
Installing wkhtmltopdf

On Windows

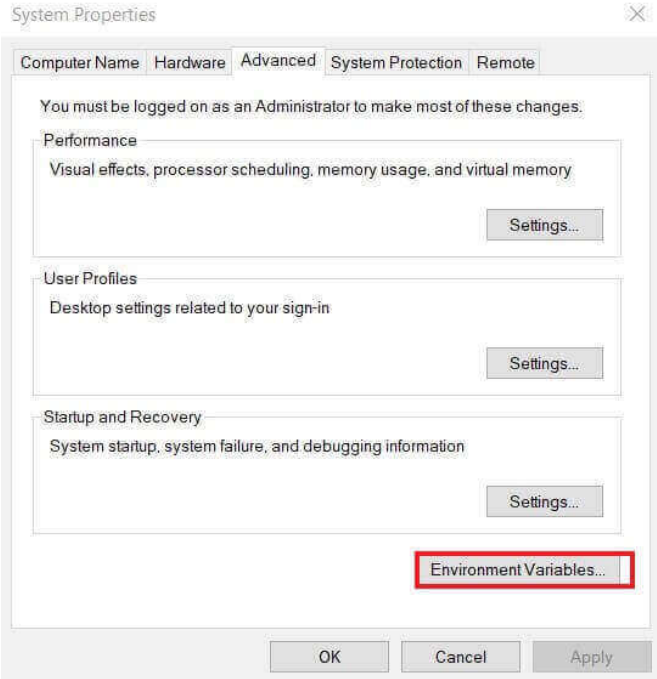
Go to the wkhtmltopdf **official downloads page**, and download the Windows installer for your Windows architecture. In my case, I downloaded the 64-bit architecture one that is supported on Vista or later since I have Windows 10.

After you have downloaded the installer and successfully installed the wkhtmltopdf tool, now you need to add it to the PATH environment variable.

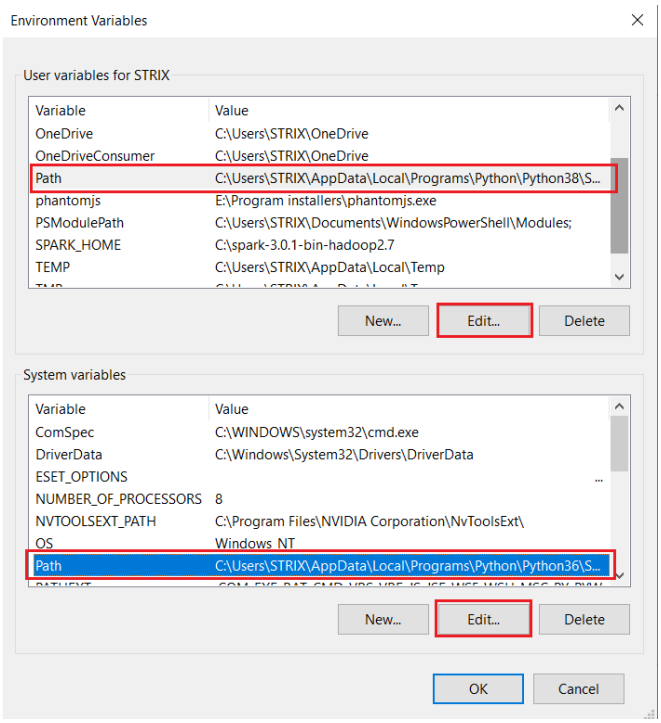
To do it, you must go to Windows search and write "environment", you'll see "Edit the system environment variables", click on it:



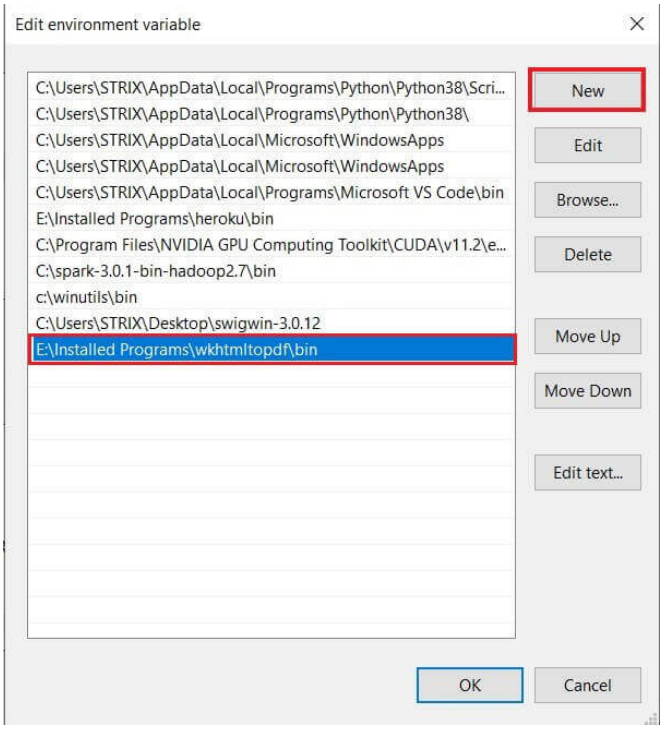
A new window will appear, and click on "Environment Variables...":



In the new window, you're free to choose the system or user variables and find the PATH variable to edit:



Once you click on Edit on either variables, go on and add the path of where you've installed wkhtmltopdf **to the PATH variable:**



After you've done that, click the OK button and close the previous windows, and you're good to go.

On Linux

If you're on Linux, it's much simpler as it'll be added to PATH automatically using your package manager.

Below is the command for Ubuntu/Debian:

```
$ apt update
$ apt install wkhtmltopdf
```

And below is for Debian/CentOS:

```
$ sudo yum makecache --refresh
$ sudo yum -y install wkhtmltopdf
```

On macOS

You can simply install it using brew:

```
$ brew install Caskroom/cask/wkhtmltopdf
```

Converting HTML from URL to PDF

`pdfkit` **did a great job wrapping** `wkhtmltopdf` **in Python; we use effortless methods to do such complicated tasks. Let's install it:**

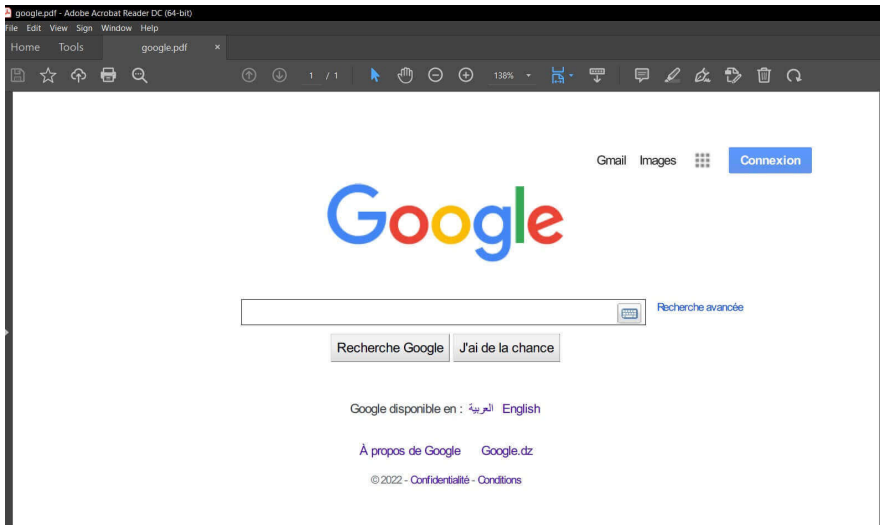
```
$ pip install pdfkit
```

For instance, let's convert the Google search page to a PDF document:

```
import pdfkit

# directly from url
pdfkit.from_url("https://google.com", "google.pdf", verbose=True)
print("="*50)
```

The first argument to the `from_url()` **function is the URL you want to convert, and the second argument is the PDF document name you wish to generate. Here's the output PDF document:**



Converting Local HTML File to PDF

You can also convert a local HTML file in your machine to a PDF document; here's how:

```
# from file
pdfkit.from_file("webapp/index.html", "index.pdf", verbose=True, options={"enable-local-file-access": True})
print("="*50)
```

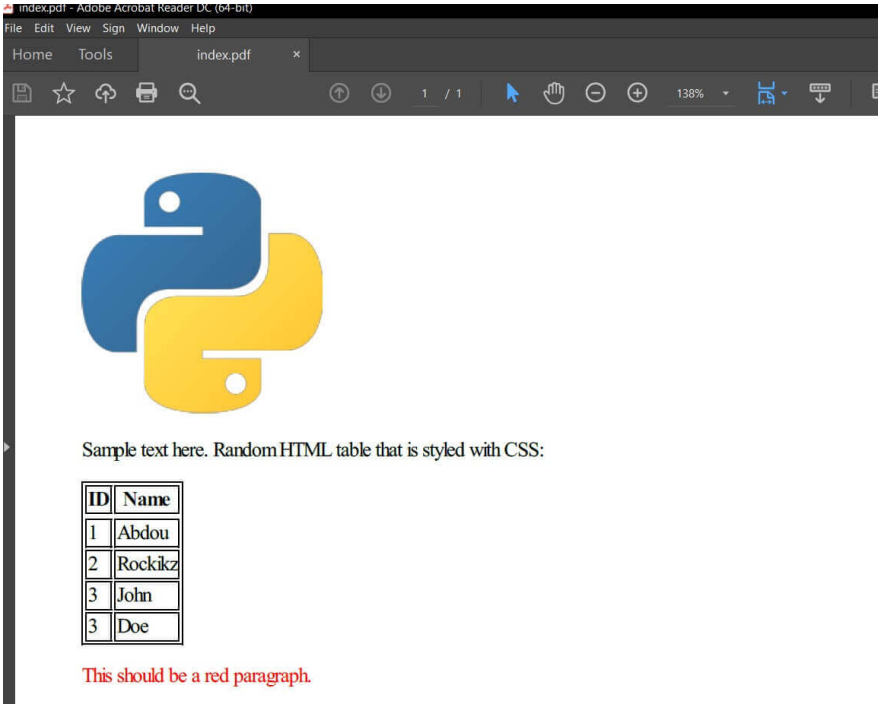
The `webapp/` folder (in which you can view it here) contains the `index.html`, its `style.css` CSS file, and a sample image `image.png`.

Here's the content of `index.html`:

```
<!DOCTYPE html>
<!--[if lt IE 7]>      <html class="no-js lt-ie9 lt-ie8 lt-ie7"> <![endif]-->
<!--[if IE 7]>         <html class="no-js lt-ie9 lt-ie8"> <![endif]-->
<!--[if IE 8]>         <html class="no-js lt-ie9"> <![endif]-->
<!--[if gt IE 8]>      <html class="no-js"> <!--<![endif]-->
<html>
    <head>
        <meta charset="utf-8">
        <meta http-equiv="X-UA-Compatible" content="IE=edge">
        <title></title>
        <meta name="description" content="">
        <meta name="viewport" content="width=device-width, initial-scale=1">
        <link rel="stylesheet" href="style.css">
        <style>
            table, th, td {
                border: 1px solid black;
            }
        </style>
    </head>
    <body>
        <!--[if lt IE 7]>
            <p class="browsehappy">You are using an <strong>outdated</strong> browser. Please <a href="#">upgrade your browser</a> to
        <![endif]-->
        <img src="image.png" alt="Python logo">
        <p>Sample text here. Random HTML table that is styled with CSS:</p>
        <table bordered>
            <thead>
                <th>ID</th>
                <th>Name</th>
            </thead>
            <tbody>
                <tr>
                    <td>1</td>
                    <td>Abdou</td>
                </tr>
                <tr>
                    <td>2</td>
                    <td>Rockikz</td>
                </tr>
                <tr>
                    <td>3</td>
                    <td>John</td>
                </tr>
                <tr>
                    <td>3</td>
                    <td>Doe</td>
                </tr>
            </tbody>
        </table>
        <p class="red-text">This should be a red paragraph.</p>
    </body>
</html>
```

We use the `from_file()` function, the first argument is the location of the HTML file, and the second is the resulting PDF document path, we set the `enable-local-file-access` to `True` in the `options` parameter to allow local file access from this HTML file to images and CSS/JS files.

Here's the content of `index.pdf`:

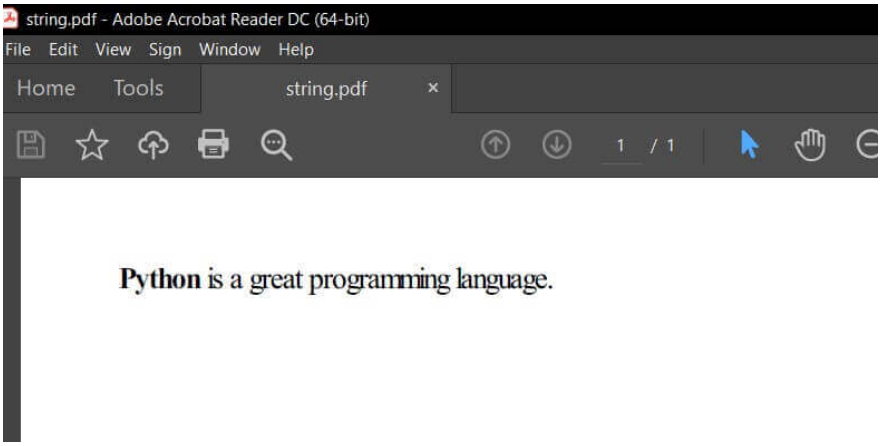


Converting HTML String to PDF

Finally, you can also convert HTML content from a Python string to a PDF document:

```
# from HTML content
pdfkit.from_string("<p><b>Python</b> is a great programming language.</p>", "string.pdf", verbose=True)
print("="*50)
```

Here's the content of `string.pdf`:



Conclusion

Awesome, I hope this tutorial was helpful to get you started with the wkhtmltopdf tool that helps convert HTML from either a URL, local file, or string to a PDF document in Python with the help of pdfkit wrapper library.

You can get the complete code here.

Finally, if you're a beginner and want to learn Python, I suggest you take the Python For Everybody Coursera course, in which you'll learn a lot about Python. You can also check our resources and courses page to see the Python resources I recommend on various topics!

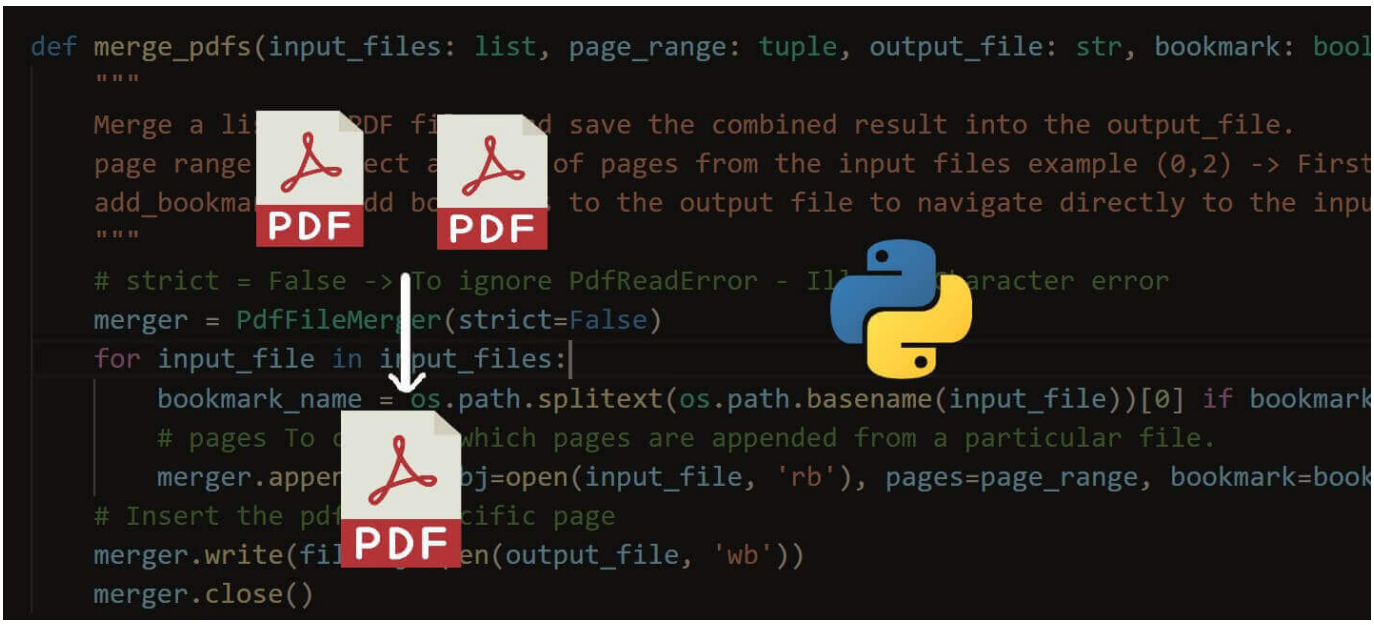
Learn also: How to Convert PDF to Docx in Python

Happy coding ♥

VIEW FULL CODE

Sharing is caring!

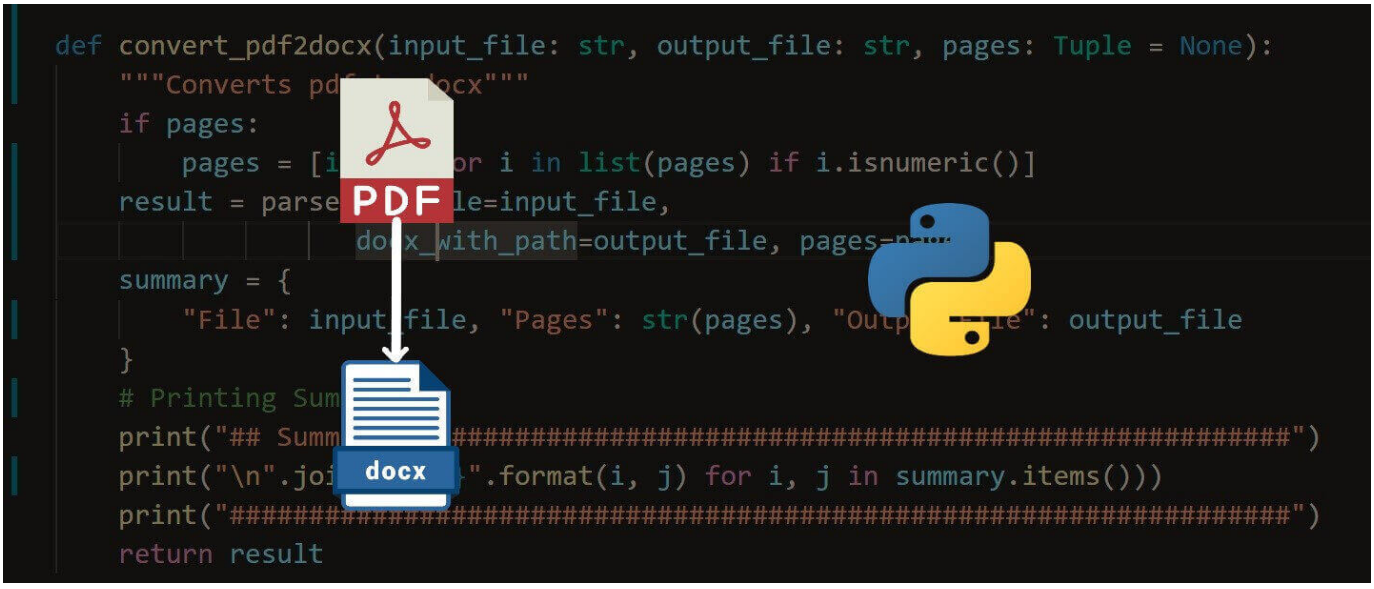
Read Also



How to Merge PDF Files in Python

Learn how to merge two or multiple PDF files into a single PDF file using PyPDF4 library in Python

VISIT



How to Convert PDF to Docx in Python

Learn how you can use pdf2docx library to convert PDF files to docx word files in Python

VISIT

How to Convert PDF to Images in Python

Learn how to use PyMuPDF library to convert PDF files into individual images per page in Python.

VISIT

Comment panel

Your email address will not be published.

Enter your name

Enter Email

Write a comment...

☑ Subscribe for our newsletter

COMMENT

Tags

- Machine Learning
- Ethical Hacking
- General Python Tutorials
- Web Scraping
- Computer Vision
- Python Standard Library
- Application Programming Interfaces
- Database
- Finance
- Packet Manipulation Using Scapy
- Natural Language Processing

New Tutorials

Popular Tutorials