Jack New

[Email address]

Guide to easily merge, and translate MD to proper markdown standards

LTIO

Table of Contents

[**Anaconda Navigator** 2](#_Toc24013816)

[Download 2](#_Toc24013817)

[Setting up new environment 2](#_Toc24013818)

[**MarkdownTools2** 3](#_Toc24013819)

[Downloading package 3](#_Toc24013820)

[**Convert PDF to Markdown (MD)** 5](#_Toc24013821)

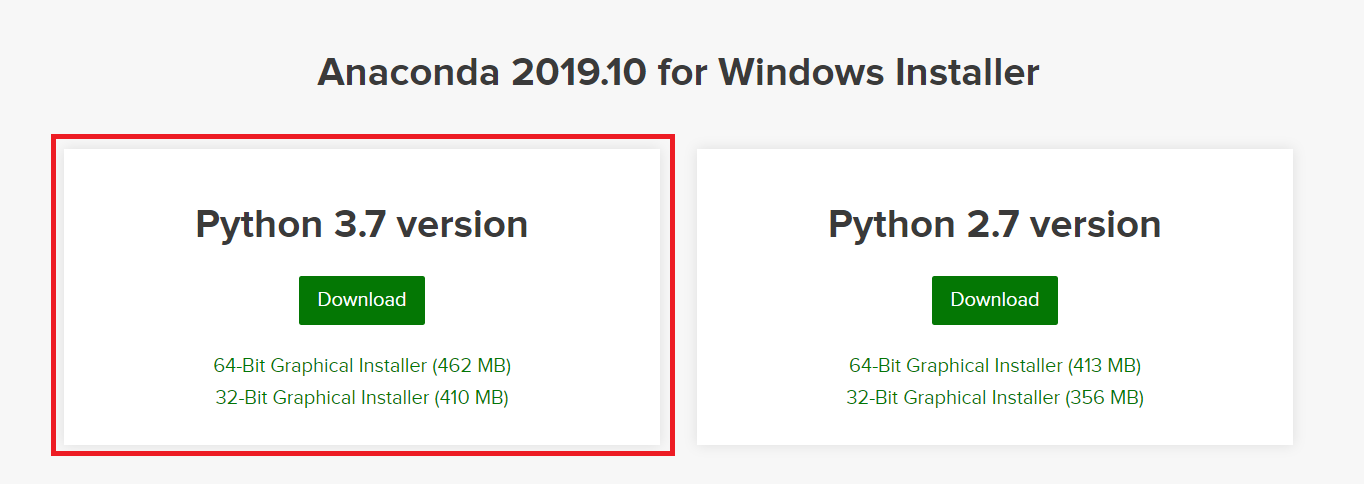
[**Merging MD files** 7](#_Toc24013822)

[**Executing script to clean up MD** 8](#_Toc24013823)

# **Anaconda Navigator**

## Download

Download Anaconda Navigator from here: <https://www.anaconda.com/distribution/>

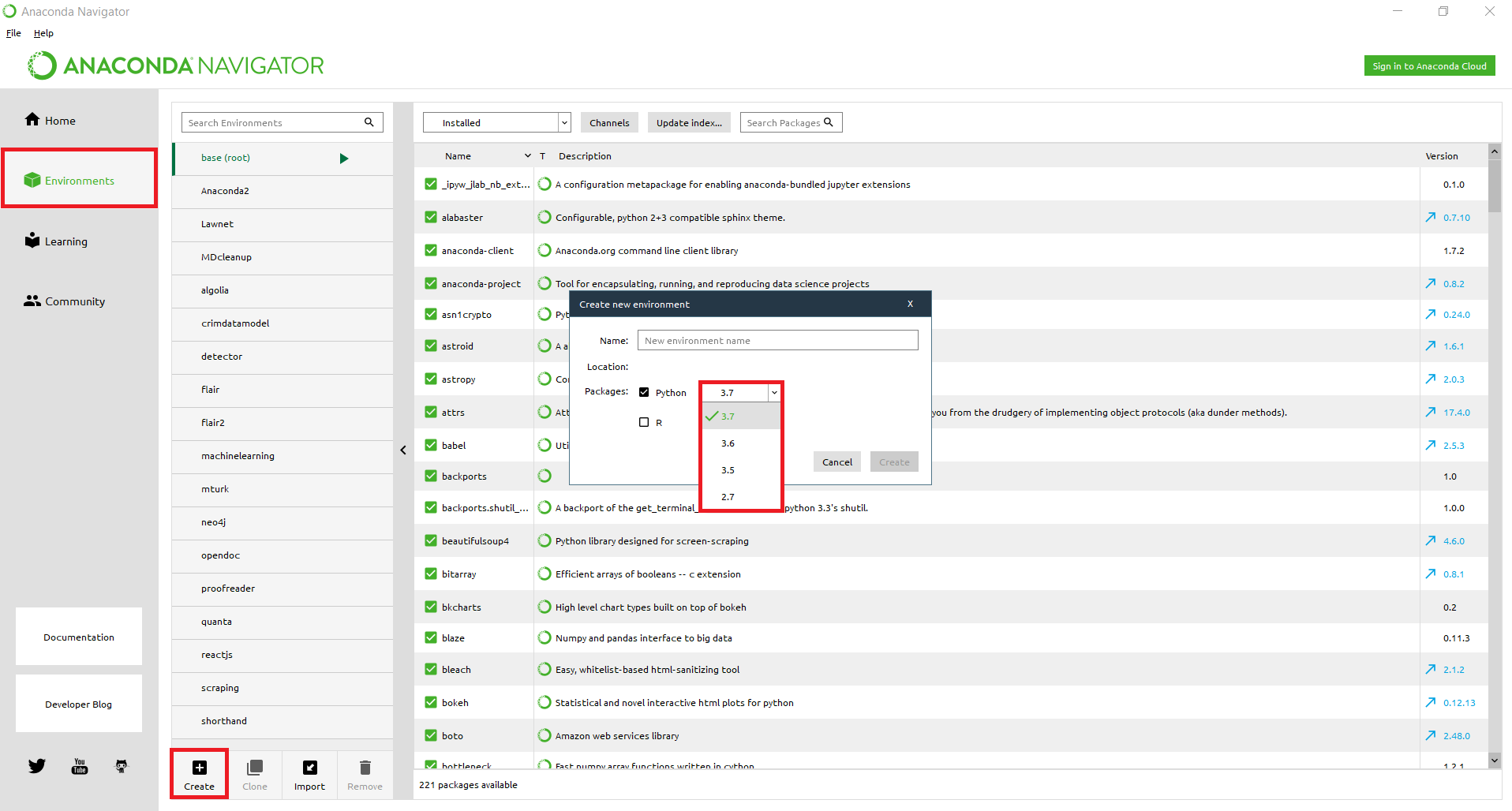


Download the latest version

Use the default settings and complete the installation.

## Setting up new environment

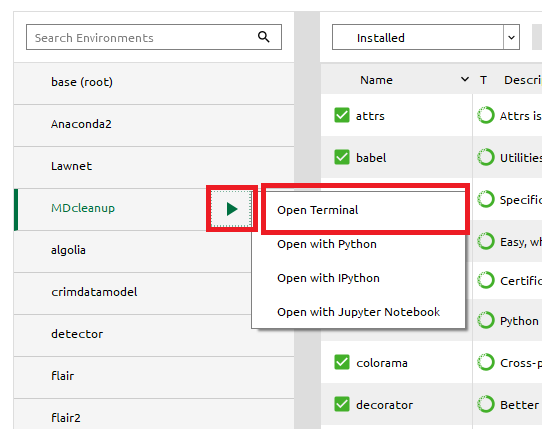
To set up a new environment, click on environment > create. You may input any name for your environment and select the latest python version which should be 3.7.



# **MarkdownTools2**

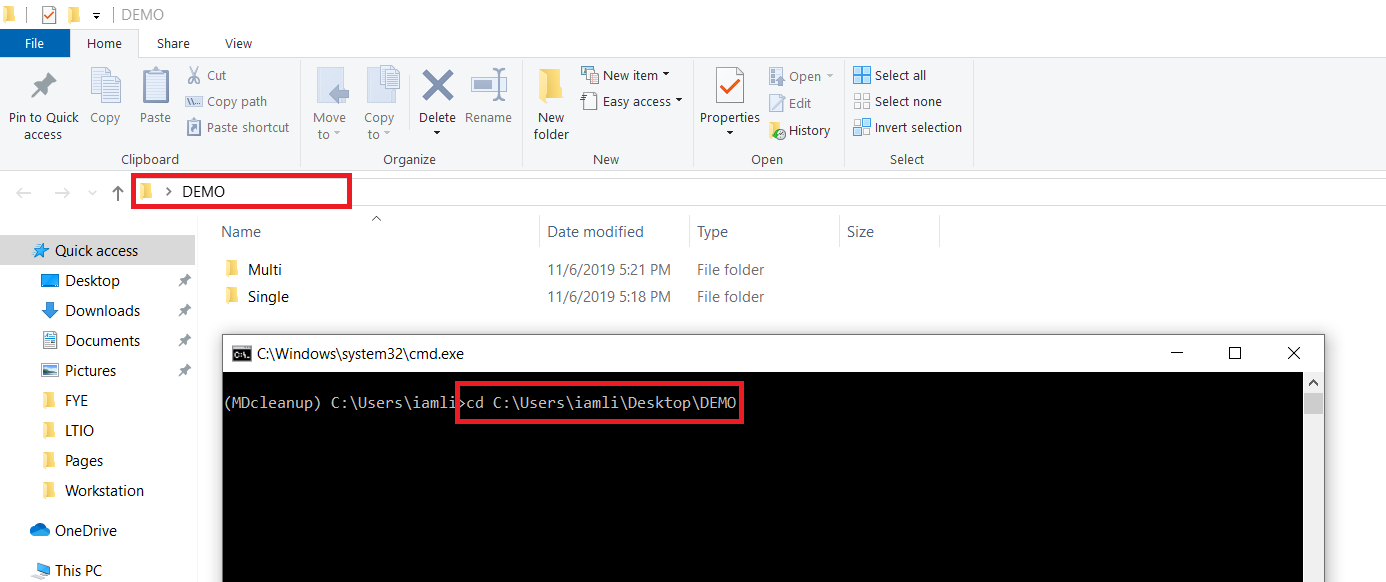
## Downloading package

MarkdownTools2 will be used to merge all MD files into one single MD file. To install, look for the arrow key beside your environment and select “Open Terminal”.



From the terminal, you will want to go to your working directory “MDCleanup” by typing in *cd …* where the path can be found by click on the working directory highlighted below.

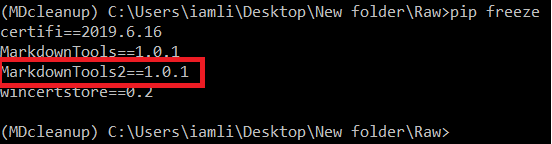
**Note**: Do note that you won’t have the exact same naming convention.



Then in the terminal enter the following:



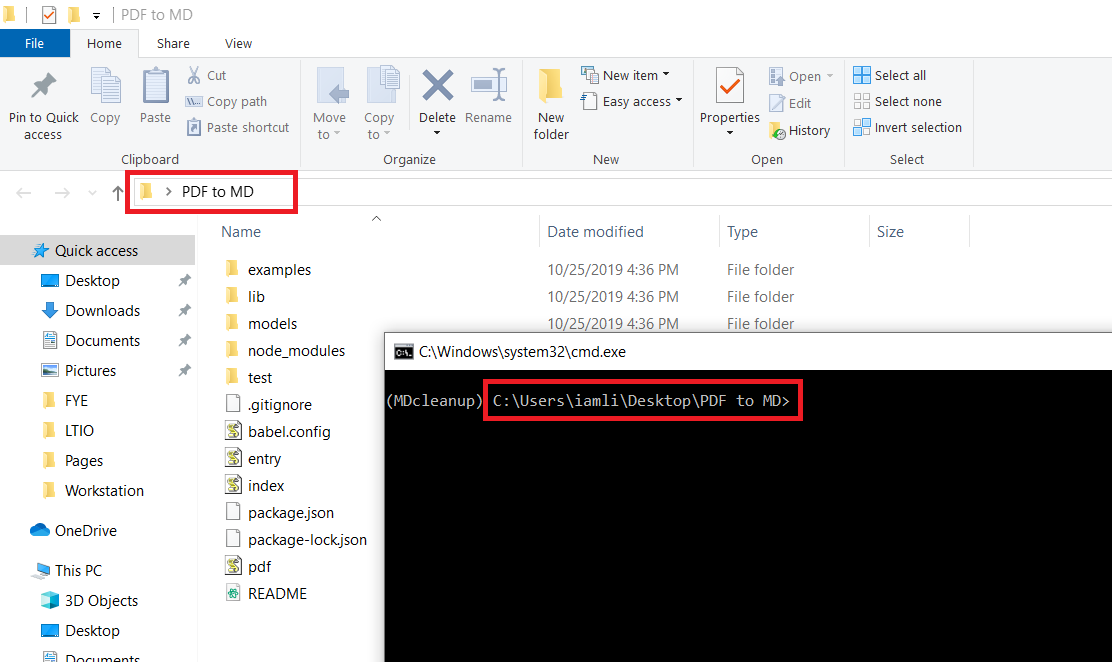
To check if you have successfully downloaded the package in the environment, type in pip freeze and you should see the following:



For the full guide and documentation please visit: <https://pypi.org/project/MarkdownTools2/>

# **Convert PDF to Markdown (MD)**

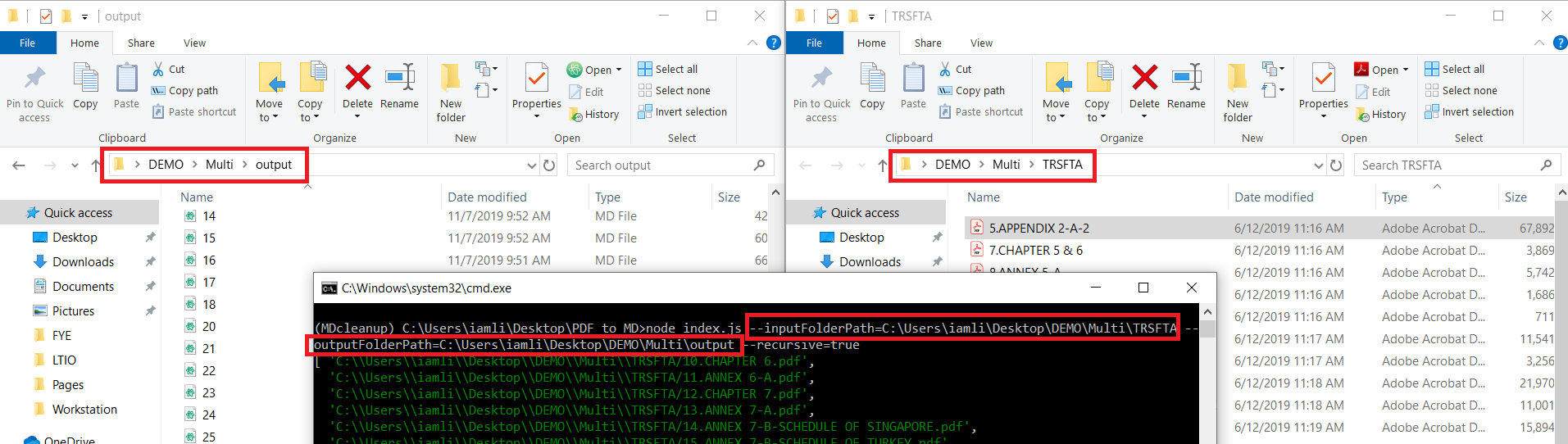
From the terminal, go to your working directory “PDF to MD” by typing in *cd …* where the path can be found by click on the working directory highlighted below.

****

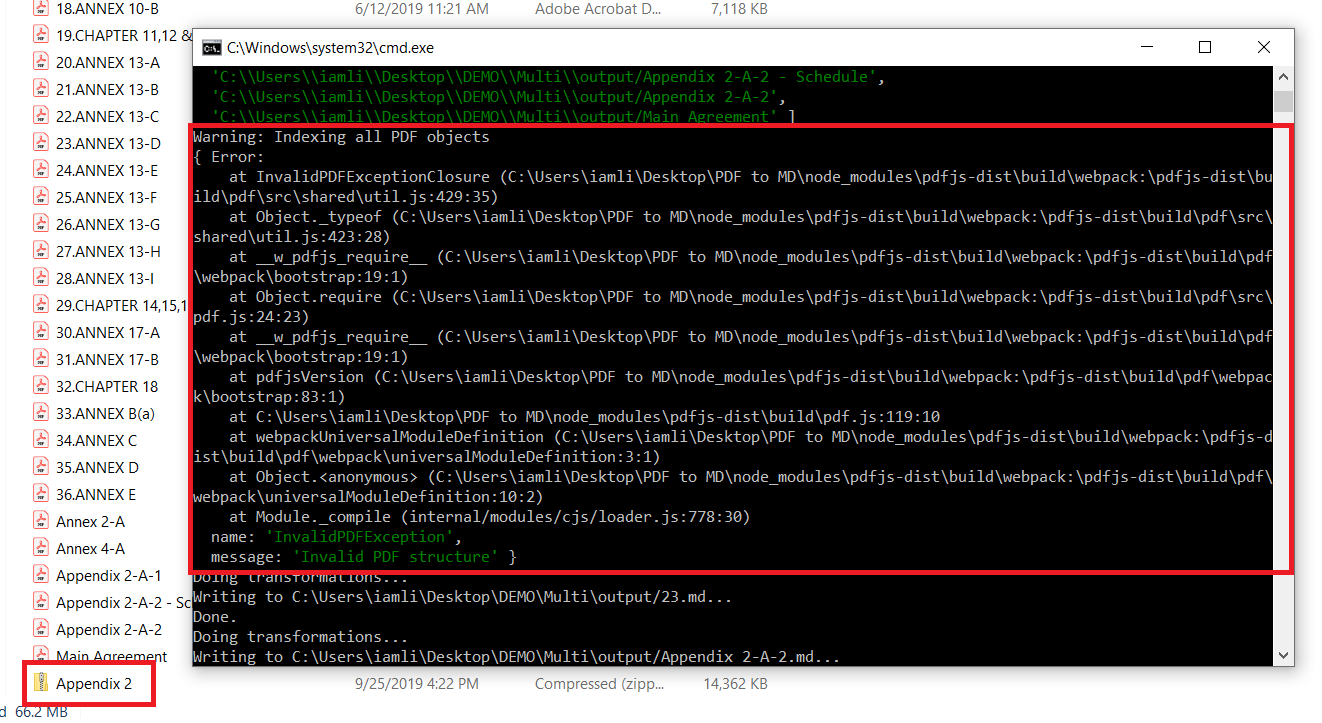
To convert the PDFs to MDs, enter the following codes:

*node index.js --inputFolderPath=… --outputFolderPath=… --recursive=true*

Input folder path will be the path to the folder where your PDFs are located in and output folder path is where you want to store the files in. You will want to do recursive to ensure all files in the folder are converted. Below is an example of how the code would look like.

**

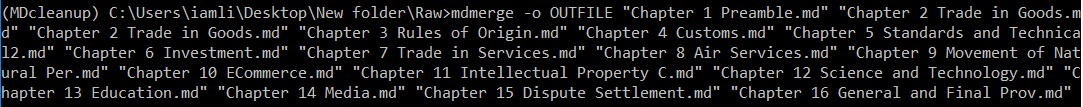
Conversion of the files may take a long time so do wait patiently for the files to be converted.

**

If you encounter errors, it means that there are non-PDF files within the folder. You may ignore the error message and continue with the conversion.

# **Merging MD files**

In your terminal, go to the folder where the MD files are located. Then enter the following code *mdmerge –o OUTFILE [name of MD files in the correct sequence]*:



The OUTFILE name can be changed but for the following script we have hardcoded the script to look for and run OUTFILE. Therefore, please remain the naming convention as it is.

# **Executing script to clean up MD**

Ensure that you have the script (cleanup.py) and OUTFILE in the same folder. Then run the file by entering the following code: *python cleanup.py*

If the script has finished running (should take about 1-3 seconds) you should see a file called finish appearing in the same folder.

