Pre-Class Batch Processing Instructions for CJL

Benjamin Blemings[*]

This Version: September 15, 2021

Updates at: Github Link

Abstract

The goal of this is session is to discuss data management and **demonstrate how I automate the execution of scripts across multiple software types** which greatly increases chances for replication and reduces time doing mundane data management tasks. This document helps setup prerequisites to make the session more interactive.

[*] Benjamin Blemings, Ph.D. Dyson School of Applied Economics and Management, Cornell University, Ithaca NY 14850. E-mail: benblemings@gmail.com

1 Introduction

I use **Stata, R, and Python**. I know Stata is not free and understand if you do not have it, although I think you should buy it if you plan to remain in academic economics.

R and Python are free, so this is still worth learning.

Even if you only use 1 of these programs, I am going to make the case for you to use all of them so I encourage you to keep an open mind and follow this guide.

This **session will run smoothest if you have these programs installed and know where their executable (.exe)** is in your computer's file system.

This guide shows how to install the programs and find their executable files.

Then, I show how to add the executables to path.

I have also never tested this system on MacOS, I know you can run "virtual Windows" on a Mac and it might be worth it for this.

2 One Click Execution of Common Software for Economists

Sooner or later, you'll run into something difficult in your preferred software

This gives you 4 options:

1. Spend a lot of time re-inventing the wheel in your preferred language

2. Limit the research questions you can ask and answer, because it's not in your preferred language

3. Try to find coauthors who use the other softwares

4. Do it in another language, but that might break your scripting flow*

*Notice, I'm already assuming you have a "makefile" that runs all scripts in order

There's a really easy way create a multi-language makefile that I will show you

Batch processing is a .txt file saved with the extension .bat

It’s contents are lines of code with this general recipe:

```
“software path” “ManageBuildDirectories.py”
::“software path” “script path”
::Old script that is commented out and not used anymore
“software path” “script path”
“software path” “script path”
“software path” “D: Dropbox \ GregBenSean \ NYCTough \ Work \ Build \ Code \ MergeData.R”
“C:\Program Files\R\R-4.1.0\bin\Rscript.exe” “D: Dropbox . . . Work \ Build \ Code \ CreateVariables.R”
“C:\Program Files\Stata16\StataMP-64.exe” /e do “D: \ Dropbox  . . . Work \ Analysis \ Code \ AnalyzeData.do”
```

The slides are mostly about where script paths go.

There’s an actual example batch file in the slides.

This document is more about the software aspect.

3 R

Stata people, install R. You don’t have to replace Stata with it, but you might need something it does at some point.

R people, please don’t sneer at my passing knowledge of R.

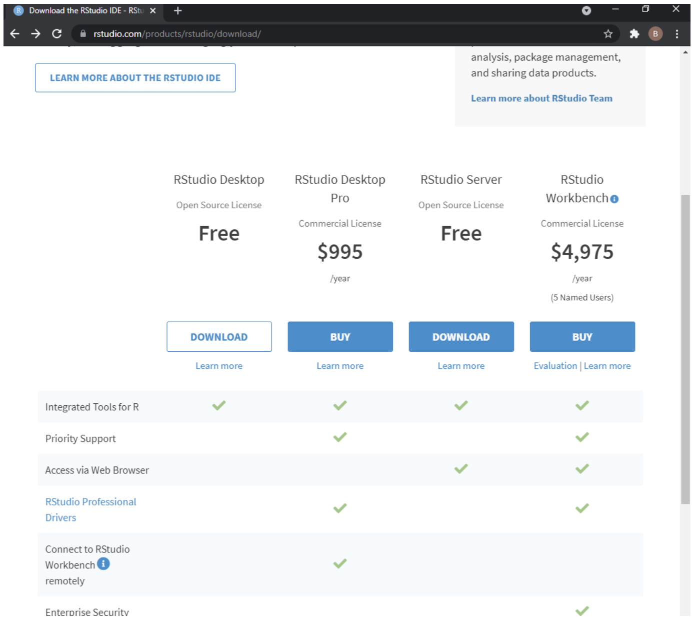
For R, I use R Studio as my IDE.

Install it from: https://www.rstudio.com/products/rstudio/download/

Here’s 2 pictures:

Figure 1

(a) R Studio Home Page



Download the RStudio IDE - RStu
rstudio.com/products/rstudio/download/

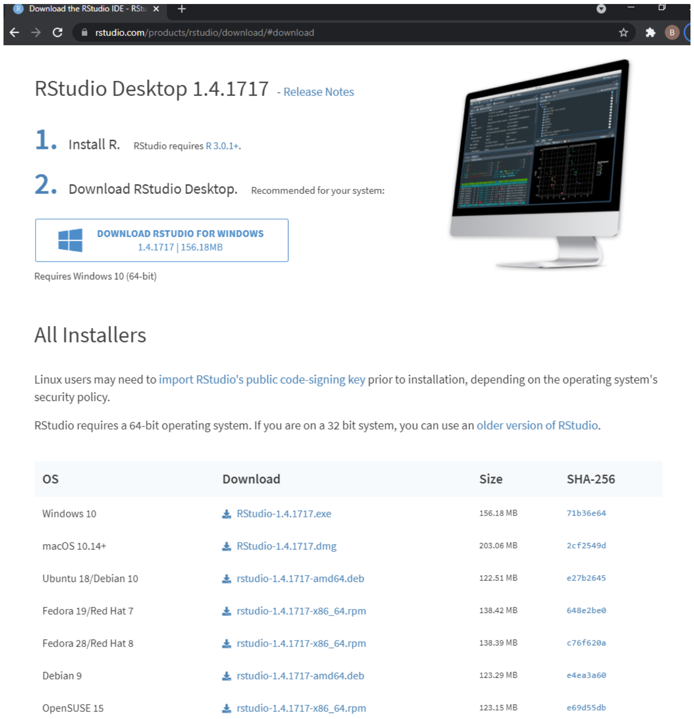
analysis, package management, and sharing data products.

Learn more about RStudio Team

LEARN MORE ABOUT THE RSTUDIO IDE

	RStudio Desktop	RStudio Desktop Pro	RStudio Server	RStudio Workbench
	Open Source License	Commercial License	Open Source License	Commercial License
	Free	$995 /year	Free	$4,975 /year (5 Named Users)
	DOWNLOAD Learn more	BUY Learn more	DOWNLOAD Learn more	BUY Evaluation | Learn more
Integrated Tools for R	✓	✓	✓	✓
Priority Support		✓		✓
Access via Web Browser			✓	✓
RStudio Professional Drivers		✓		✓
Connect to RStudio Workbench remotely		✓		
Enterprise Security				✓

(b)



Download the RStudio IDE - RStu
rstudio.com/products/rstudio/download/#download

RStudio Desktop 1.4.1717 - Release Notes

1. Install R. RStudio requires R 3.0.1+.

2. Download RStudio Desktop. Recommended for your system:

DOWNLOAD RSTUDIO FOR WINDOWS
1.4.1717 | 156.18MB

Requires Windows 10 (64-bit)

All Installers

Linux users may need to import RStudio's public code-signing key prior to installation, depending on the operating system's security policy.

RStudio requires a 64-bit operating system. If you are on a 32 bit system, you can use an older version of RStudio.

OS	Download	Size	SHA-256
Windows 10	RStudio-1.4.1717.exe	156.18 MB	71b36e64
macOS 10.14+	RStudio-1.4.1717.dmg	203.06 MB	2cf2549d
Ubuntu 18/Debian 10	rstudio-1.4.1717-amd64.deb	122.51 MB	e27b2645
Fedora 19/Red Hat 7	rstudio-1.4.1717-x86_64.rpm	138.42 MB	648e2be0
Fedora 28/Red Hat 8	rstudio-1.4.1717-x86_64.rpm	138.39 MB	c76f620a
Debian 9	rstudio-1.4.1717-amd64.deb	123.29 MB	e4ea3a60
OpenSUSE 15	rstudio-1.4.1717-x86_64.rpm	123.15 MB	e69d55db

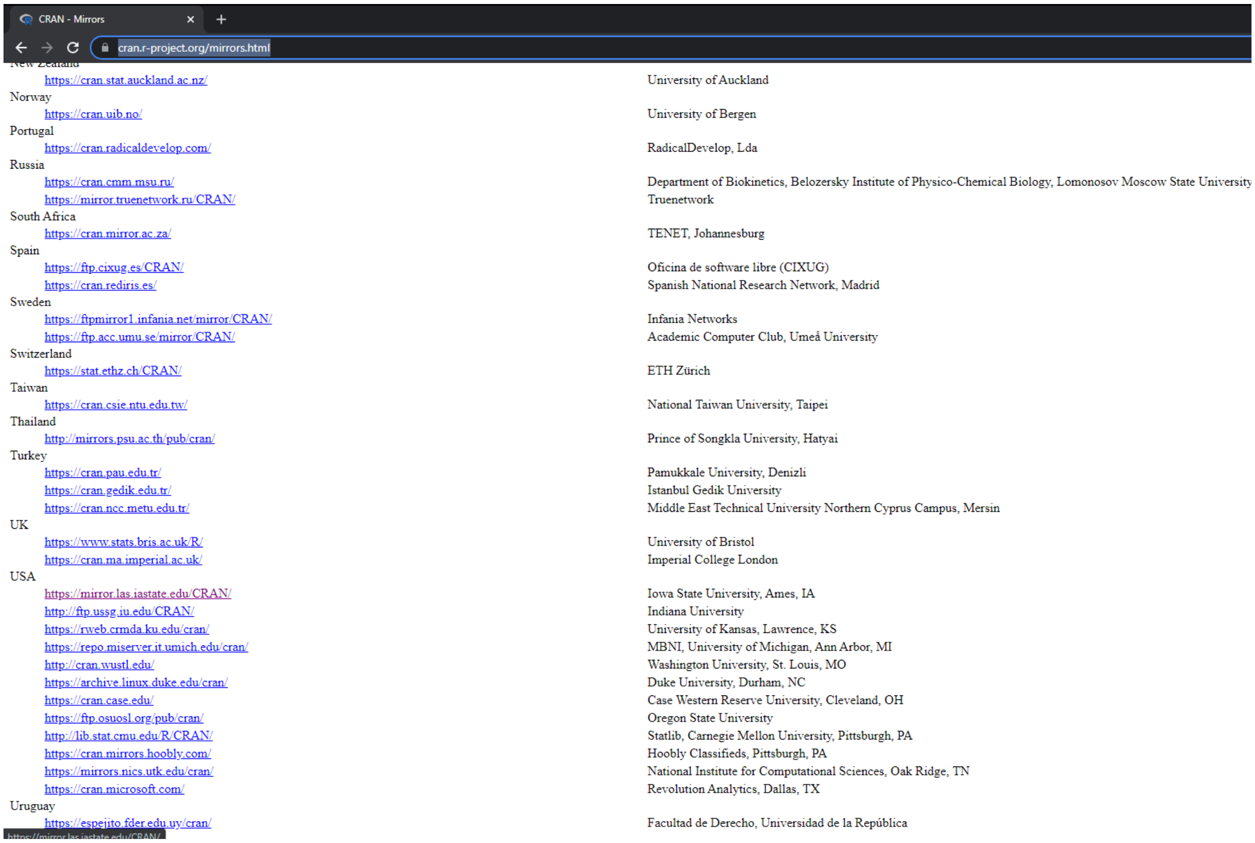
Note: Choose free in (a), choose you OS in (b).

You will also need to download R.

Install it from: https://cran.r-project.org/mirrors.html

Here’s a picture:

Figure 2


CRAN - Mirrors
cran.r-project.org/mirrors.html
https://mirror.las.iastate.edu/CRAN/

New Zealand		
	https://cran.stat.auckland.ac.nz/	University of Auckland
Norway		
	https://cran.uib.no/	University of Bergen
Portugal		
	https://cran.radicaldevelop.com/	RadicalDevelop, Lda
Russia		
	https://cran.cmm.msu.ru/	Department of Biokinetics, Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University
	https://mirror.truenetwork.ru/CRAN/	Truenetwork
South Africa		
	https://cran.mirror.ac.za/	TENET, Johannesburg
Spain		
	https://ftp.cixug.es/CRAN/	Oficina de software libre (CIXUG)
	https://cran.rediris.es/	Spanish National Research Network, Madrid
Sweden		
	https://ftpmirror1.infania.net/mirror/CRAN/	Infania Networks
	https://ftp.acc.umu.se/mirror/CRAN/	Academic Computer Club, Umeå University
Switzerland		
	https://stat.ethz.ch/CRAN/	ETH Zürich
Taiwan		
	https://cran.csie.ntu.edu.tw/	National Taiwan University, Taipei
Thailand		
	http://mirrors.psu.ac.th/pub/cran/	Prince of Songkla University, Hatyai
Turkey		
	https://cran.pau.edu.tr/	Pamukkale University, Denizli
	https://cran.gedik.edu.tr/	Istanbul Gedik University
	https://cran.ncc.metu.edu.tr/	Middle East Technical University Northern Cyprus Campus, Mersin
UK		
	https://www.stats.bris.ac.uk/R/	University of Bristol
	https://cran.ma.imperial.ac.uk/	Imperial College London
USA		
	https://mirror.las.iastate.edu/CRAN/	Iowa State University, Ames, IA
	http://ftp.ussg.iu.edu/CRAN/	Indiana University
	https://rweb.crmda.ku.edu/cran/	University of Kansas, Lawrence, KS
	https://repo.miserver.it.umich.edu/cran/	MBNI, University of Michigan, Ann Arbor, MI
	http://cran.wustl.edu/	Washington University, St. Louis, MO
	https://archive.linux.duke.edu/cran/	Duke University, Durham, NC
	https://cran.case.edu/	Case Western Reserve University, Cleveland, OH
	https://ftp.osuosl.org/pub/cran/	Oregon State University
	http://lib.stat.cmu.edu/R/CRAN/	Statlib, Carnegie Mellon University, Pittsburgh, PA
	https://cran.mirrors.hoobly.com/	Hoobly Classifieds, Pittsburgh, PA
	https://mirrors.nics.utk.edu/cran/	National Institute for Computational Sciences, Oak Ridge, TN
	https://cran.microsoft.com/	Revolution Analytics, Dallas, TX
Uruguay		
	https://espejito.fder.edu.uy/cran/	Facultad de Derecho, Universidad de la República

Note: Pick a mirror in the US. I did the first one which was Iowa State's.

Next you should locate the .exe.

Ben's .exe path is: "C:\Program Files\R\R-4.1.0\bin\Rscript.exe"

My .exe path for R is:

4 Python

Stata and R people, install Python. Do R and Stata do a lot of the same things as Python? Yes.

Do Stata and R do those same things as well as Python? In my opinion, no.

Also, don't forget about the things Python does, that Stata and R do not.

In short, Python is a one-stop shop for Swiss-knife type data analysis/management needs.

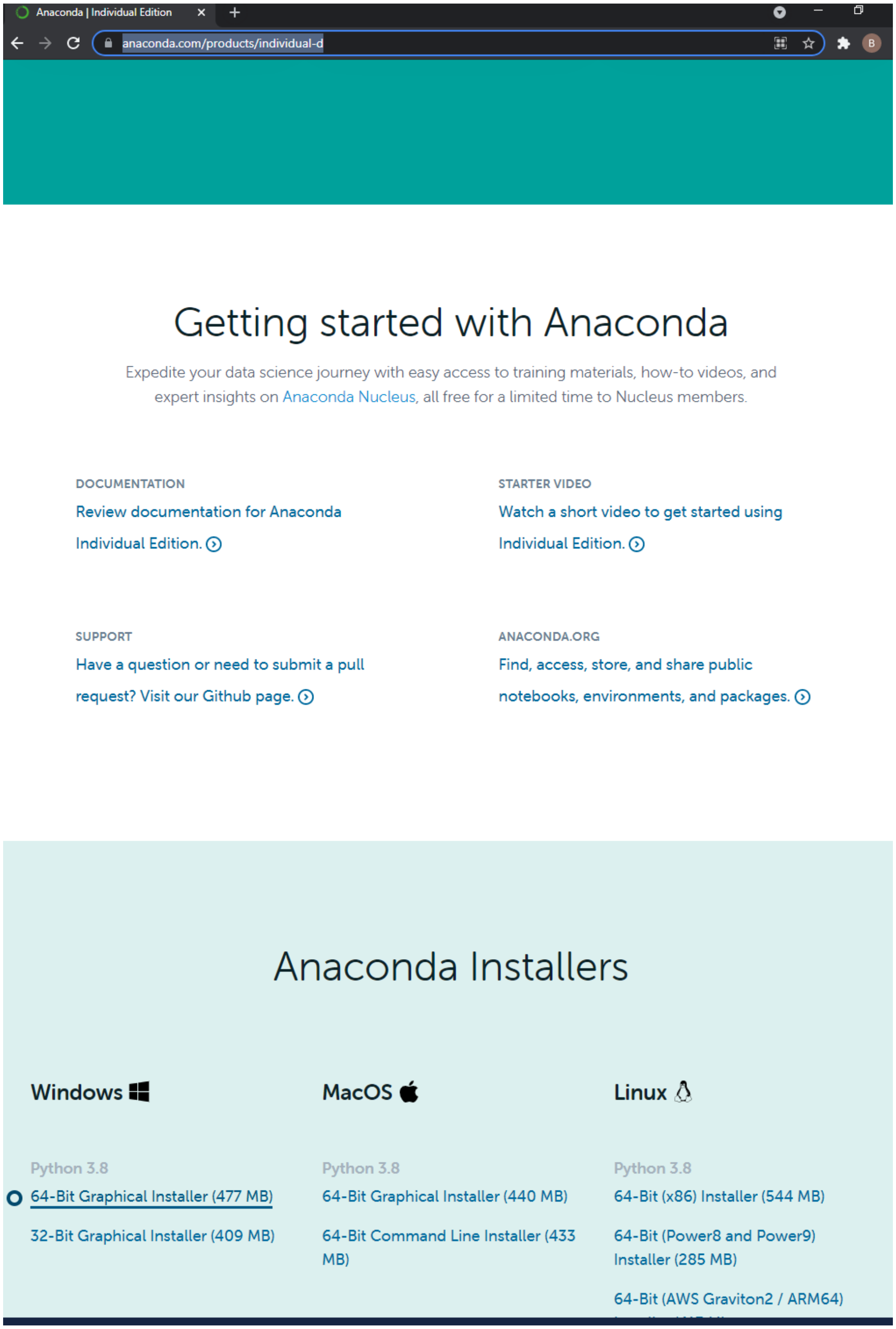
I never took a class in Python, just messed around with and it learned things from Google.

So my default go to GUI is scripting in Anaconda.

With this, you get the Spyder IDE which is what I use.

Website: https://www.anaconda.com/products/individual-d

Figure 3



Anaconda | Individual Edition
anaconda.com/products/individual-d

Getting started with Anaconda

Expedite your data science journey with easy access to training materials, how-to videos, and expert insights on Anaconda Nucleus, all free for a limited time to Nucleus members.

DOCUMENTATION

Review documentation for Anaconda Individual Edition.

STARTER VIDEO

Watch a short video to get started using Individual Edition.

SUPPORT

Have a question or need to submit a pull request? Visit our Github page.

ANACONDA.ORG

Find, access, store, and share public notebooks, environments, and packages.

Anaconda Installers

Windows

Python 3.8

64-Bit Graphical Installer (477 MB)

32-Bit Graphical Installer (409 MB)

MacOS

Python 3.8

64-Bit Graphical Installer (440 MB)

64-Bit Command Line Installer (433 MB)

Linux

Python 3.8

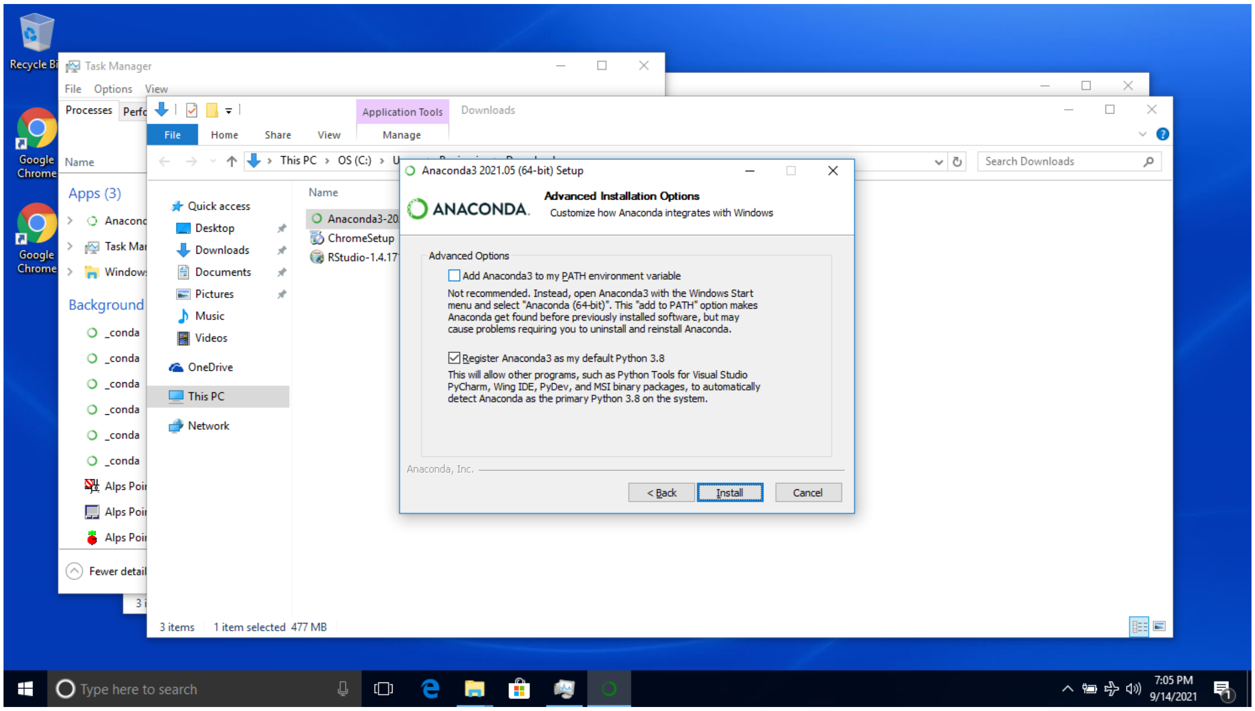
64-Bit (x86) Installer (544 MB)

64-Bit (Power8 and Power9) Installer (285 MB)

64-Bit (AWS Graviton2 / ARM64)

Note: You have to scroll down from https://www.anaconda.com/products/individual-d.

Figure 4: Do Not Add Python/Anaconda to Path, Will Do That Manually. Do Not Check Blue Box



Anaconda3 2021.05 (64-bit) Setup

Advanced Installation Options

Customize how Anaconda integrates with Windows

Advanced Options

□ Add Anaconda3 to my PATH environment variable

Not recommended. Instead, open Anaconda3 with the Windows Start menu and select "Anaconda (64-bit)". This "add to PATH" option makes Anaconda get found before previously installed software, but may cause problems requiring you to uninstall and reinstall Anaconda.

☑ Register Anaconda3 as my default Python 3.8

This will allow other programs, such as Python Tools for Visual Studio PyCharm, Wing IDE, PyDev, and MSI binary packages, to automatically detect Anaconda as the primary Python 3.8 on the system.

Anaconda, Inc.

Note: Do not check the path box. We will do that ourselves.

Important: Do not check the box

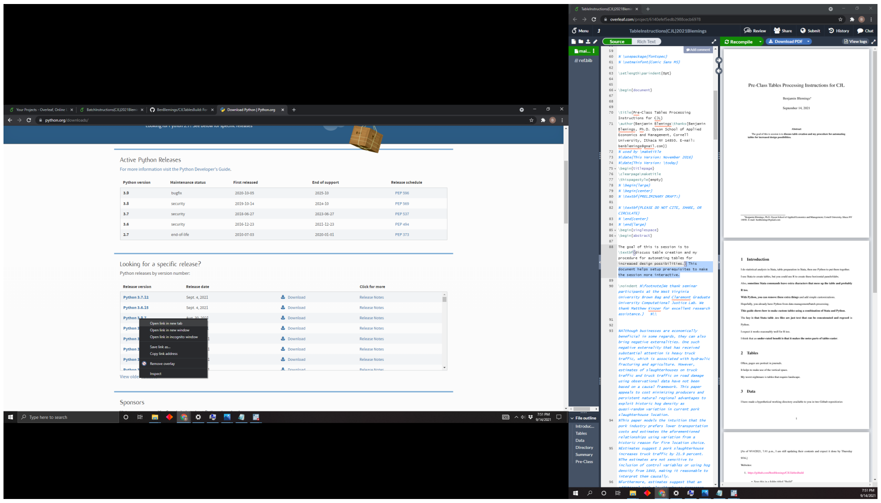
Then I download Python, because my batch hasn't like the Python from Anaconda.

Website: https://www.python.org/downloads/release/python-397/

Picture:

Figure 5: Python Pages

(a) Python Home Page



BatchInstructions(CJL)2021Blemi
BenBlemings/CJLTablesBuild: For
Download Python | Python.org
Open link in new tab
Open link in new window
Open link in incognito window
Save link as...
Copy link address
Remove overlay
Inspect

Active Python Releases

For more information visit the Python Developer's Guide.

Python version	Maintenance status	First released	End of support	Release schedule
3.9	bugfix	2020-10-05	2025-10	PEP 596
3.8	security	2019-10-14	2024-10	PEP 569
3.7	security	2018-06-27	2023-06-27	PEP 537
3.6	security	2016-12-23	2021-12-23	PEP 494
2.7	end-of-life	2010-07-03	2020-01-01	PEP 373

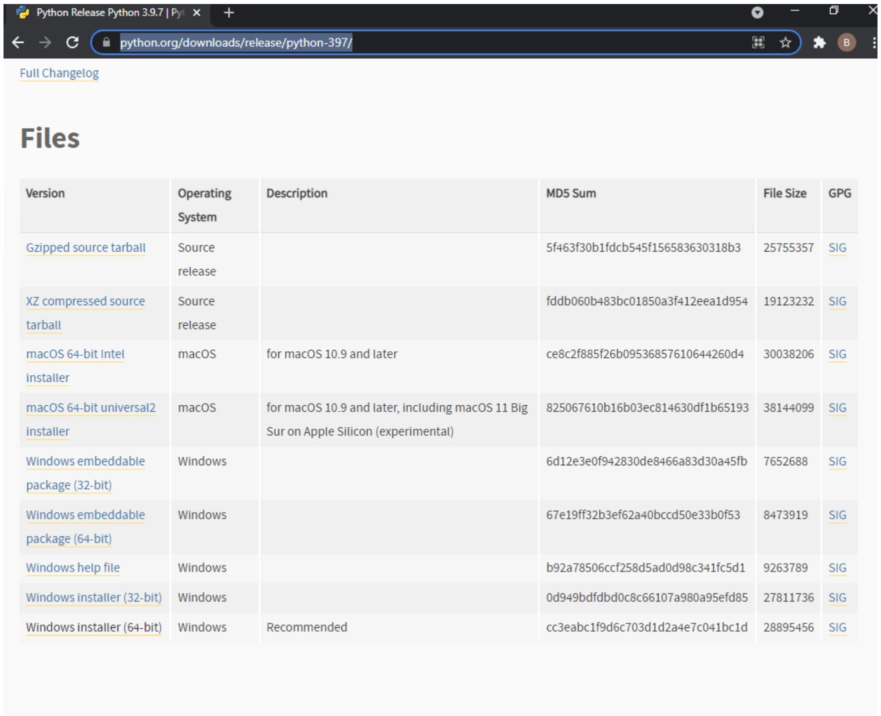
Looking for a specific release?

Python releases by version number:

Release version	Release date		Click for more
Python 3.7.12	Sept. 4, 2021	Download	Release Notes
Python 3.6.15	Sept. 4, 2021	Download	Release Notes
Python 3.9.7	Aug. 30, 2021	Download	Release Notes

Sponsors

(b) Python Choose Installation



Python Release Python 3.9.7 | Pyt
python.org/downloads/release/python-397/

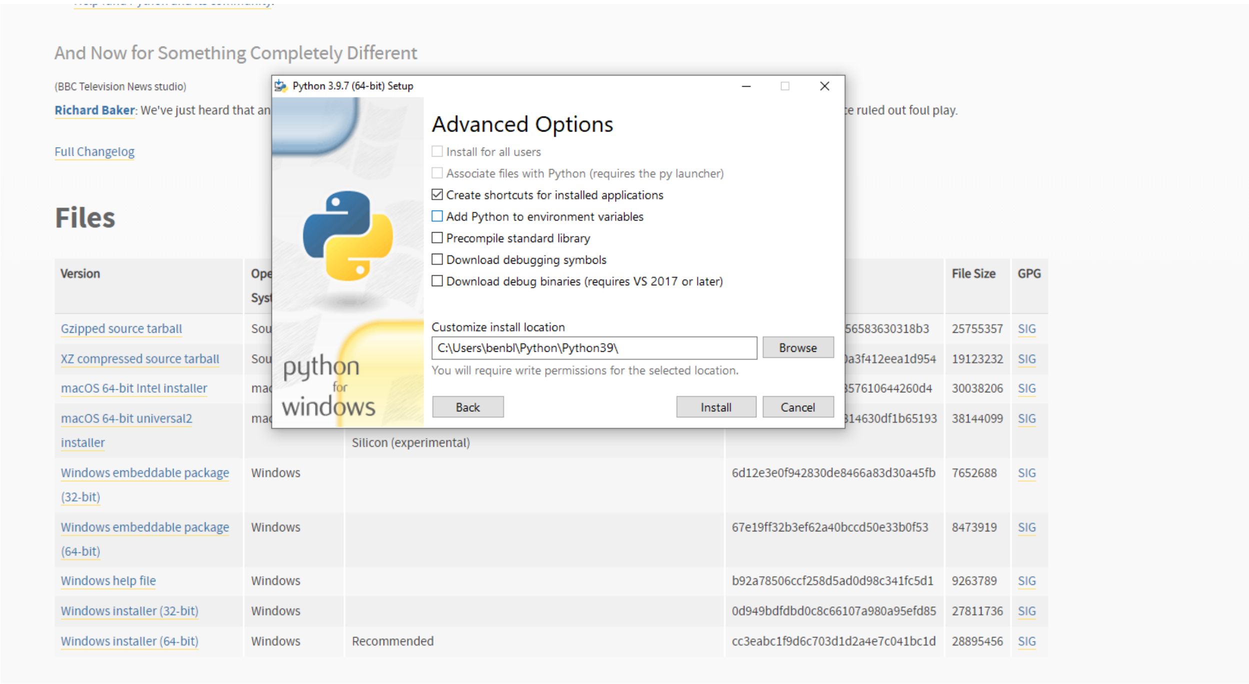
Full Changelog

Files

Version	Operating System	Description	MD5 Sum	File Size	GPG
Gzipped source tarball	Source release		5f463f30b1fdcb545f156583630318b3	25755357	SIG
XZ compressed source tarball	Source release		fddb060b483bc01850a3f412eea1d954	19123232	SIG
macOS 64-bit Intel installer	macOS	for macOS 10.9 and later	ce8c2f885f26b09536857610644260d4	30038206	SIG
macOS 64-bit universal2 installer	macOS	for macOS 10.9 and later, including macOS 11 Big Sur on Apple Silicon (experimental)	825067610b16b03ec814630df1b65193	38144099	SIG
Windows embeddable package (32-bit)	Windows		6d12e3e0f942830de8466a83d30a45fb	7652688	SIG
Windows embeddable package (64-bit)	Windows		67e19ff32b3ef62a40bccd50e33b0f53	8473919	SIG
Windows help file	Windows		b92a78506ccf258d5ad0d98c341fc5d1	9263789	SIG
Windows installer (32-bit)	Windows		0d949bdfdbd0c8c66107a980a95efd85	27811736	SIG
Windows installer (64-bit)	Windows	Recommended	cc3eabc1f9d6c703d1d2a4e7c041bc1d	28895456	SIG

Note: Choose 3.x in (a), choose your OS in (b). Click the left most column link of your OS in (b).

Figure 6: Do Not Add Python to Path, Will Do That Manually. Do Not Check Blue Box



And Now for Something Completely Different

(BBC Television News studio)

Richard Baker: We've just heard that an … e ruled out foul play.

Full Changelog

Python 3.9.7 (64-bit) Setup

Advanced Options

☐ Install for all users

☐ Associate files with Python (requires the py launcher)

☑ Create shortcuts for installed applications

☐ Add Python to environment variables

☐ Precompile standard library

☐ Download debugging symbols

☐ Download debug binaries (requires VS 2017 or later)

Customize install location

C:\Users\benbl\Python\Python39\

You will require write permissions for the selected location.

Files

Version	Ope Syst			File Size	GPG
Gzipped source tarball	Sou		56583630318b3	25755357	SIG
XZ compressed source tarball	Sou		0a3f412eea1d954	19123232	SIG
macOS 64-bit Intel installer	mac		357610644260d4	30038206	SIG
macOS 64-bit universal2 installer	mac	Silicon (experimental)	314630df1b65193	38144099	SIG
Windows embeddable package (32-bit)	Windows		6d12e3e0f942830de8466a83d30a45fb	7652688	SIG
Windows embeddable package (64-bit)	Windows		67e19ff32b3ef62a40bccd50e33b0f53	8473919	SIG
Windows help file	Windows		b92a78506ccf258d5ad0d98c341fc5d1	9263789	SIG
Windows installer (32-bit)	Windows		0d949bdfdbd0c8c66107a980a95efd85	27811736	SIG
Windows installer (64-bit)	Windows	Recommended	cc3eabc1f9d6c703d1d2a4e7c041bc1d	28895456	SIG

Note: Do not check the path box. We will do that ourselves.

Important: Do not check the box

Ben's .exe path for Python is: C:\Users\benbl\Python39

&

C:\Users\benbl\Python39\Scripts

My .exe path for Python is:

5 Stata

If you have this, then great.

Look for your Stata executable.

My .exe path is: "C:\ Program Files\ Stata16 \ StataMP-64.exe"

6 Adding Programs to Path

This cleans up the batch file

It also helps with working across several computers/people

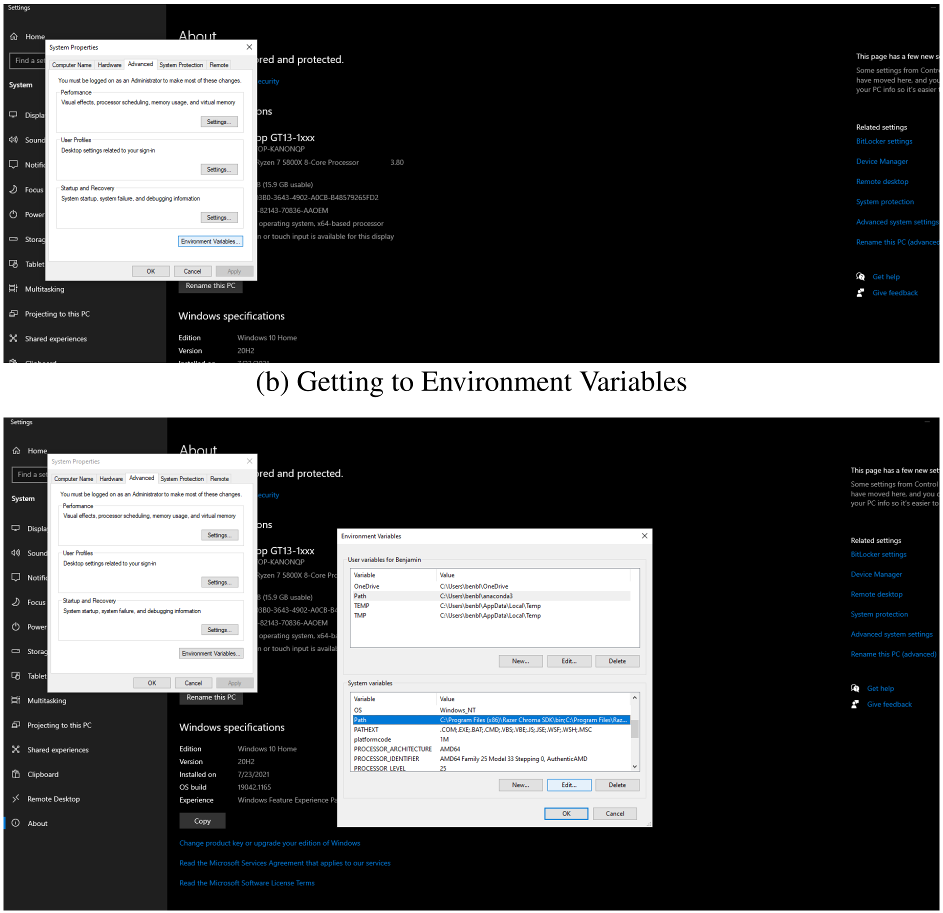
The batch file can look for executables in the “path” so we’ll add them

You need to go to this pc > right click > properties > advanced > environment variables

Here are pictures:

Figure 7: Navigating to System Path

(a) Finding Advanced Settings



System Properties

You must be logged on as an Administrator to make most of these changes.

Performance: Visual effects, processor scheduling, memory usage, and virtual memory

User Profiles: Desktop settings related to your sign-in

Startup and Recovery: System startup, system failure, and debugging information

(b) Getting to Environment Variables

Environment Variables

User variables for Benjamin

Variable	Value
OneDrive	C:\Users\benbl\OneDrive
Path	C:\Users\benbl\anaconda3
TEMP	C:\Users\benbl\AppData\Local\Temp
TMP	C:\Users\benbl\AppData\Local\Temp

System variables

Variable	Value
OS	Windows_NT
Path	C:\Program Files (x86)\Razer Chroma SDK\bin;C:\Program Files\Raz...
PATHEXT	.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;.JSE;.WSF;.WSH;.MSC
platformcode	1M
PROCESSOR_ARCHITECTURE	AMD64
PROCESSOR_IDENTIFIER	AMD64 Family 25 Model 33 Stepping 0, AuthenticAMD
PROCESSOR_LEVEL	25

Note: (a) is after going to properties on this PC. Then clicking the advanced system settings at the right. (b) is after clicking on environment variables in (a).

Add the folders with the .exe's we're using

My path looks like this:

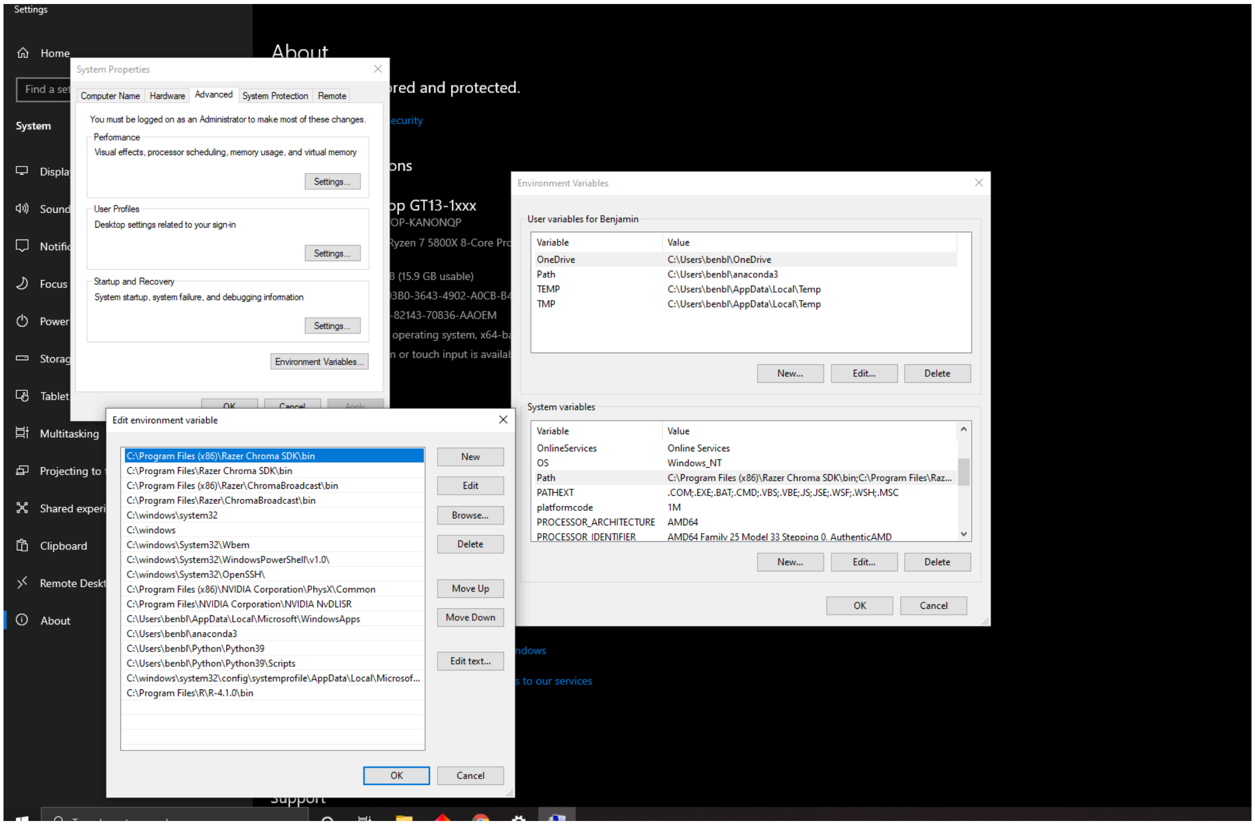
7 Pre-Class Instructions

Install Python and R programs and know the .exe paths for both of them.

Download the class example folder of batch example in different languages from : Github Link

We will add folders to path during class.

Figure 8: Path



System Properties

You must be logged on as an Administrator to make most of these changes.

Performance

Visual effects, processor scheduling, memory usage, and virtual memory

User Profiles

Desktop settings related to your sign-in

Startup and Recovery

System startup, system failure, and debugging information

Environment Variables

User variables for Benjamin

Variable	Value
OneDrive	C:\Users\benbl\OneDrive
Path	C:\Users\benbl\anaconda3
TEMP	C:\Users\benbl\AppData\Local\Temp
TMP	C:\Users\benbl\AppData\Local\Temp

System variables

Variable	Value
OnlineServices	Online Services
OS	Windows_NT
Path	C:\Program Files (x86)\Razer Chroma SDK\bin;C:\Program Files\Raz...
PATHEXT	.COM;.EXE;.BAT;.CMD;.VBS;.VBE;.JS;.JSE;.WSF;.WSH;.MSC
platformcode	1M
PROCESSOR_ARCHITECTURE	AMD64
PROCESSOR_IDENTIFIER	AMD64 Family 25 Model 33 Stepping 0, AuthenticAMD

Edit environment variable

```
C:\Program Files (x86)\Razer Chroma SDK\bin
C:\Program Files\Razer Chroma SDK\bin
C:\Program Files (x86)\Razer\ChromaBroadcast\bin
C:\Program Files\Razer\ChromaBroadcast\bin
C:\windows\system32
C:\windows
C:\windows\System32\Wbem
C:\windows\System32\WindowsPowerShell\v1.0\
C:\windows\System32\OpenSSH\
C:\Program Files (x86)\NVIDIA Corporation\PhysX\Common
C:\Program Files\NVIDIA Corporation\NVIDIA NvDLISR
C:\Users\benbl\AppData\Local\Microsoft\WindowsApps
C:\Users\benbl\anaconda3
C:\Users\benbl\Python\Python39
C:\Users\benbl\Python\Python39\Scripts
C:\windows\system32\config\systemprofile\AppData\Local\Microsof...
C:\Program Files\R\R-4.1.0\bin
```

Note: