Text Mining

Lecture 01

RPython Installation

Keungoui Kim

awekim@handong.edu

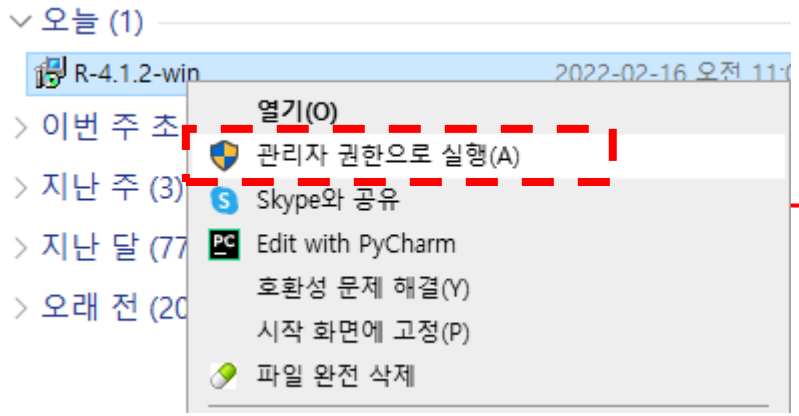
Handong Global University, 2024-2.

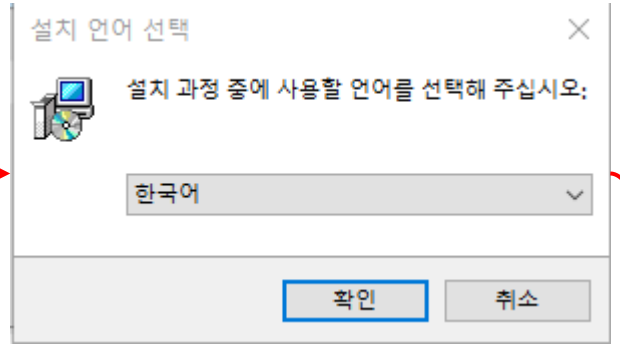
R Installation

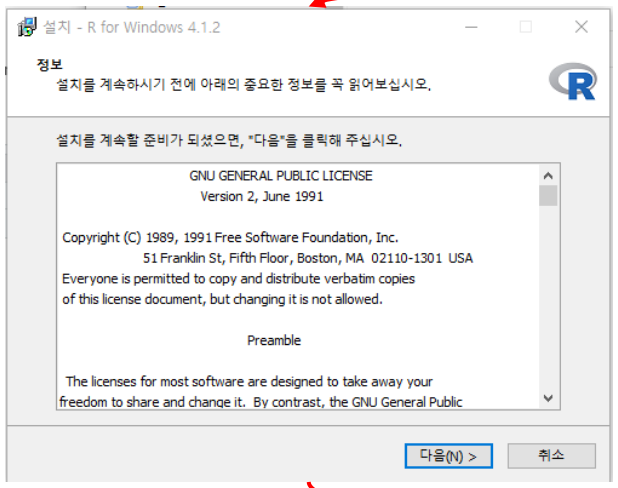
- R
  - Download from https://cran.r-project.org/bin/windows/base/
  - (search “download R”)
- R studio
  - Integrated Developing Environment IDE for R
  - Download from https://www.rstudio.com/products/rstudio/download/
- R studio server
  - R working environment available from a remote server
  - Accounts are already set up for all enrolled students

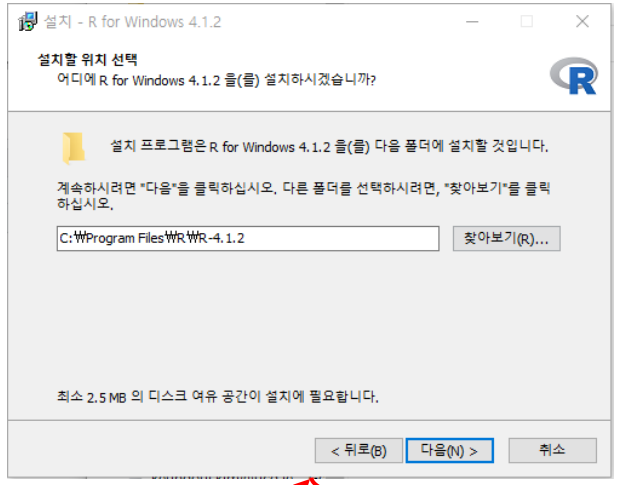
- on Windows
  - https://youtu.be/MFfRQuQKGYg
- on Mac
  - https://youtu.be/Ywj6yNfc5nM
- on Linux
  - (Ubuntu) https://youtu.be/GsuA5ugYqyw

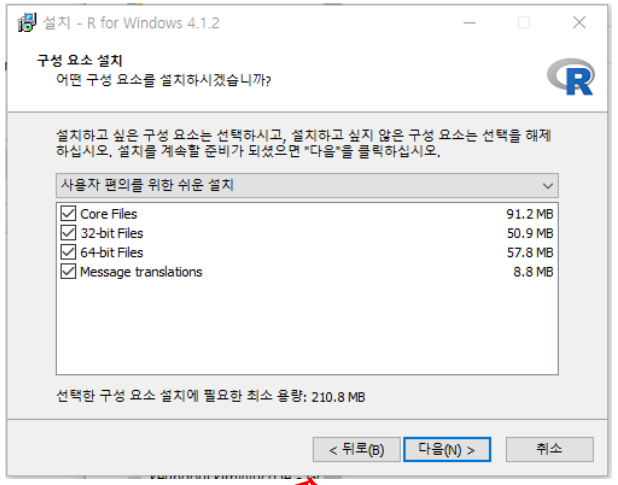
- Install with “Run as administrator”

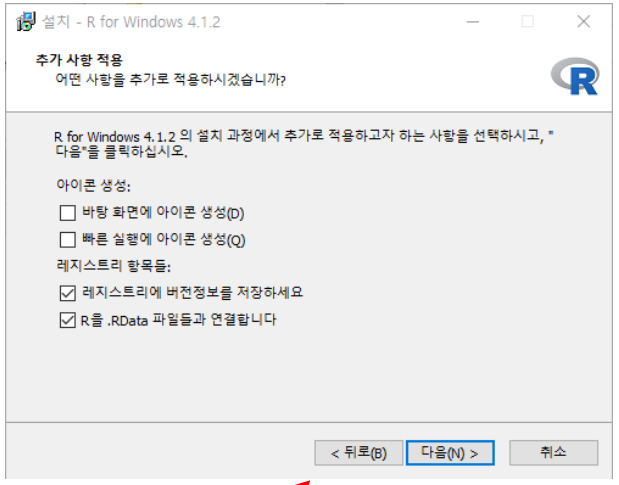






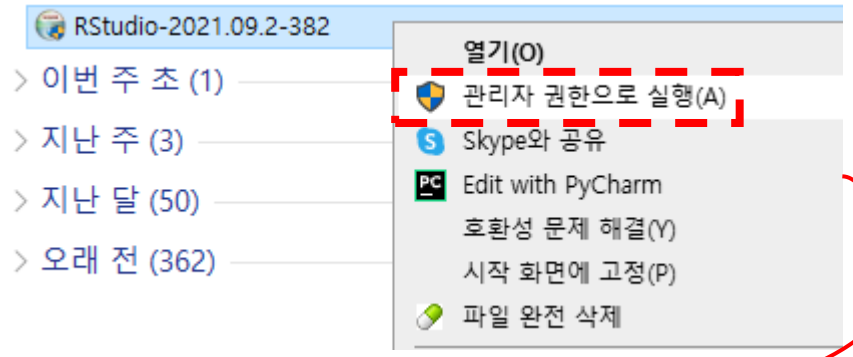


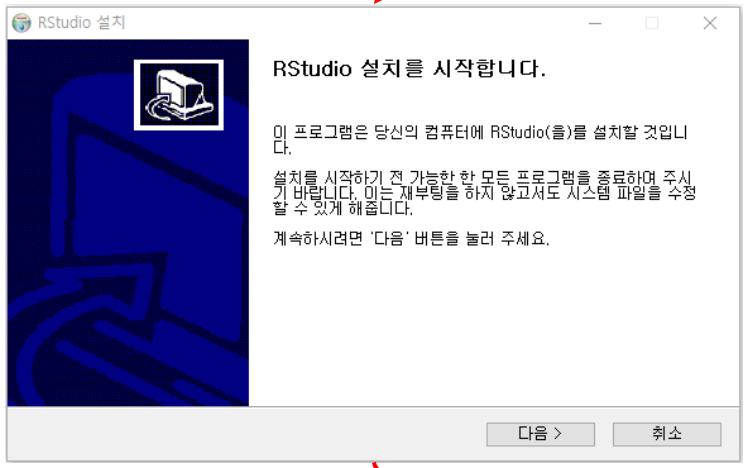


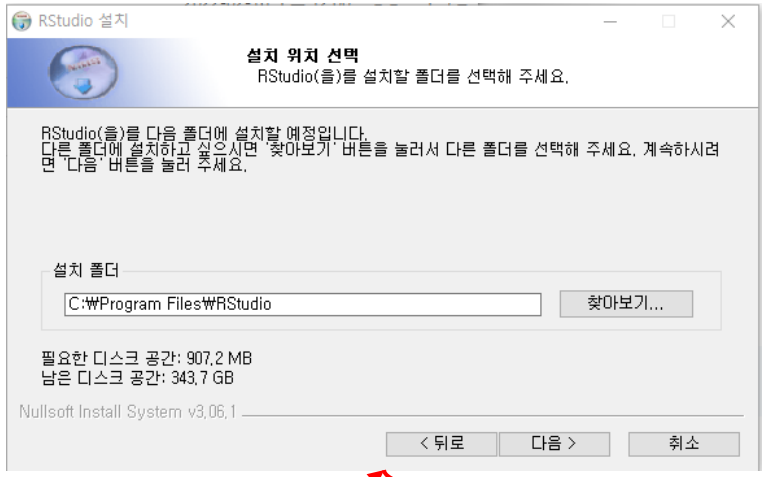


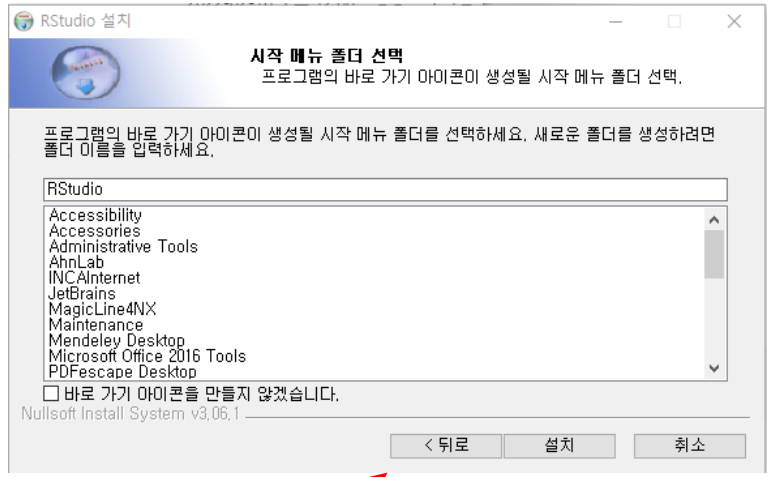
Install R Studio

- Install with “Run as administrator”







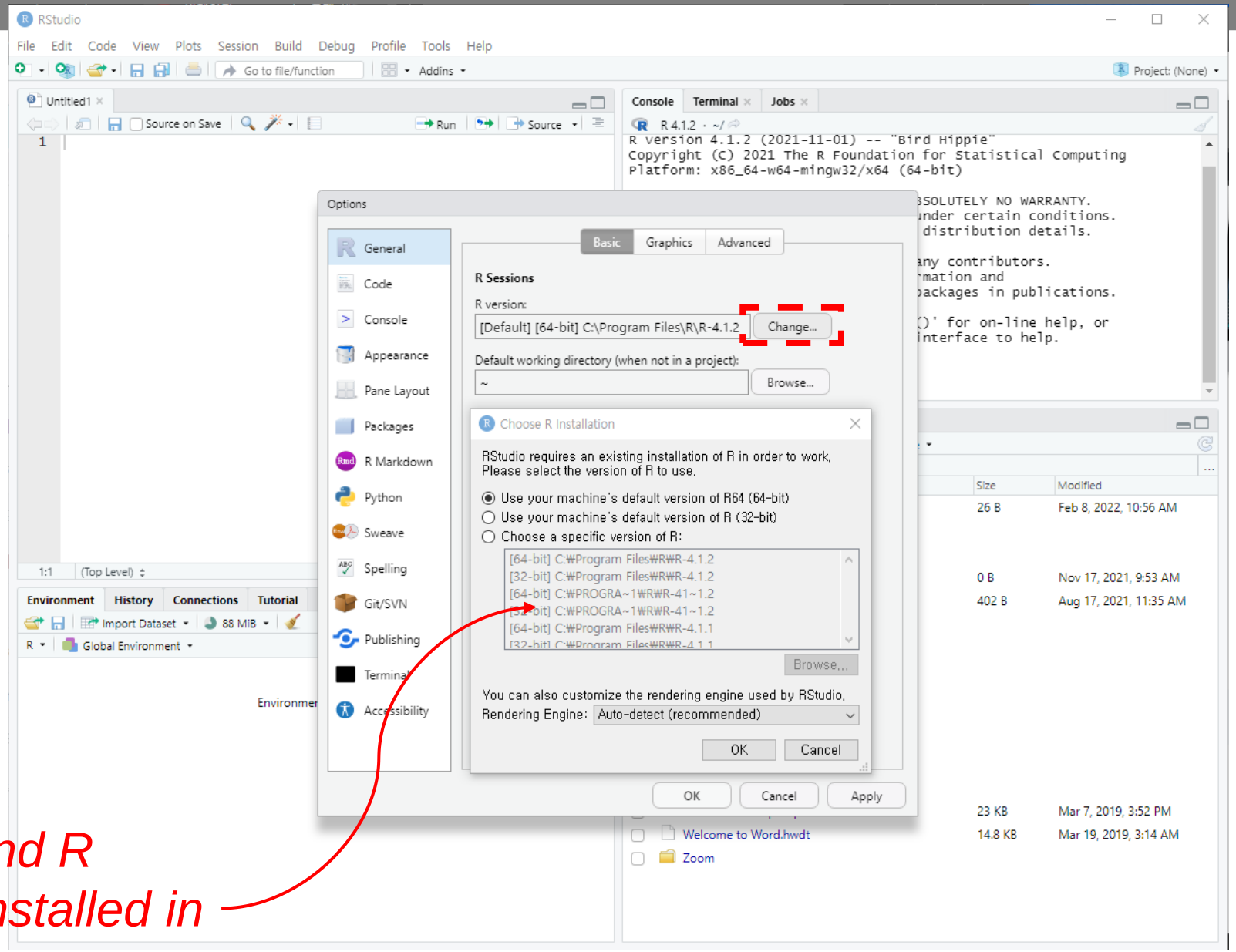


Run R & Version Control

- Run as “Run as administrator”

- Version control
  - Tools → Global Options… → General → R version:

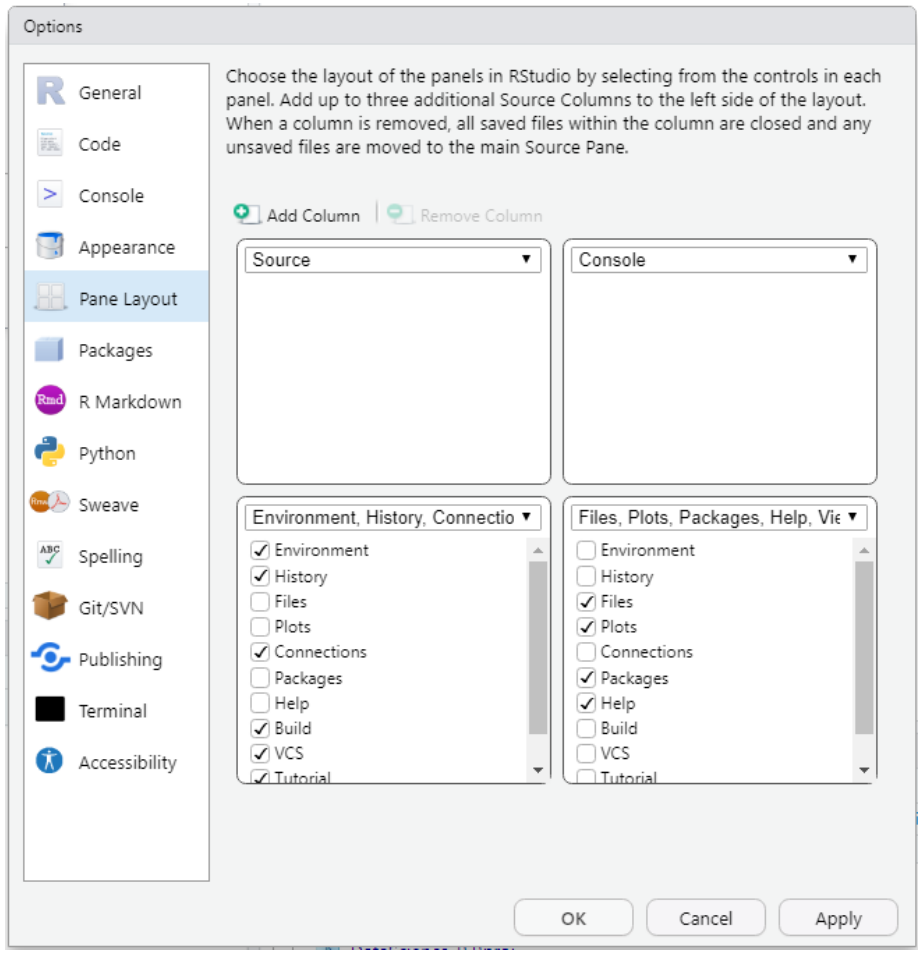
  - When using an R package, different versions of R might be required. → Select R versions you wish to use.

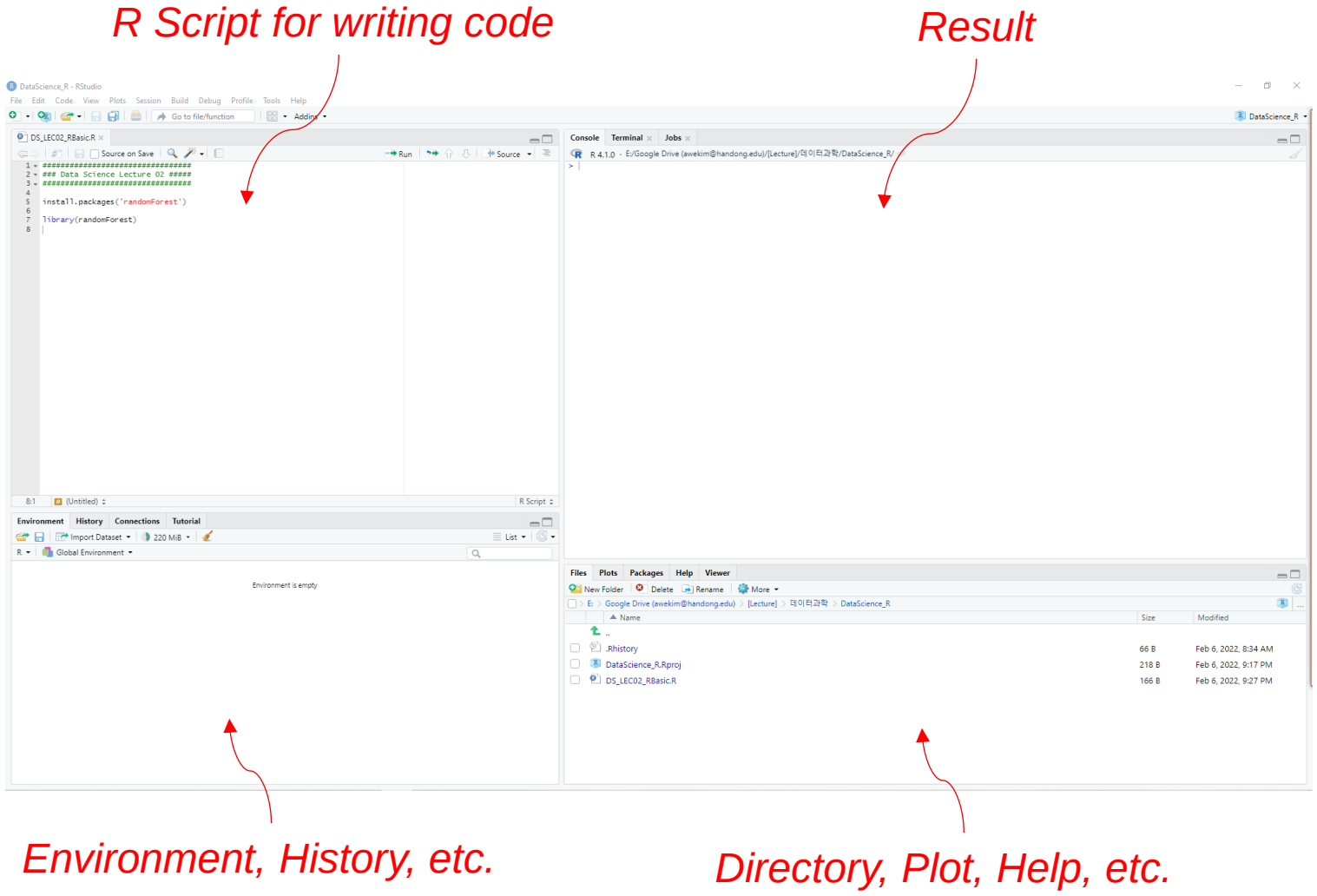


You can find R versions installed in your computer

Display Setting

- Tools → Global Option → Pane Layout



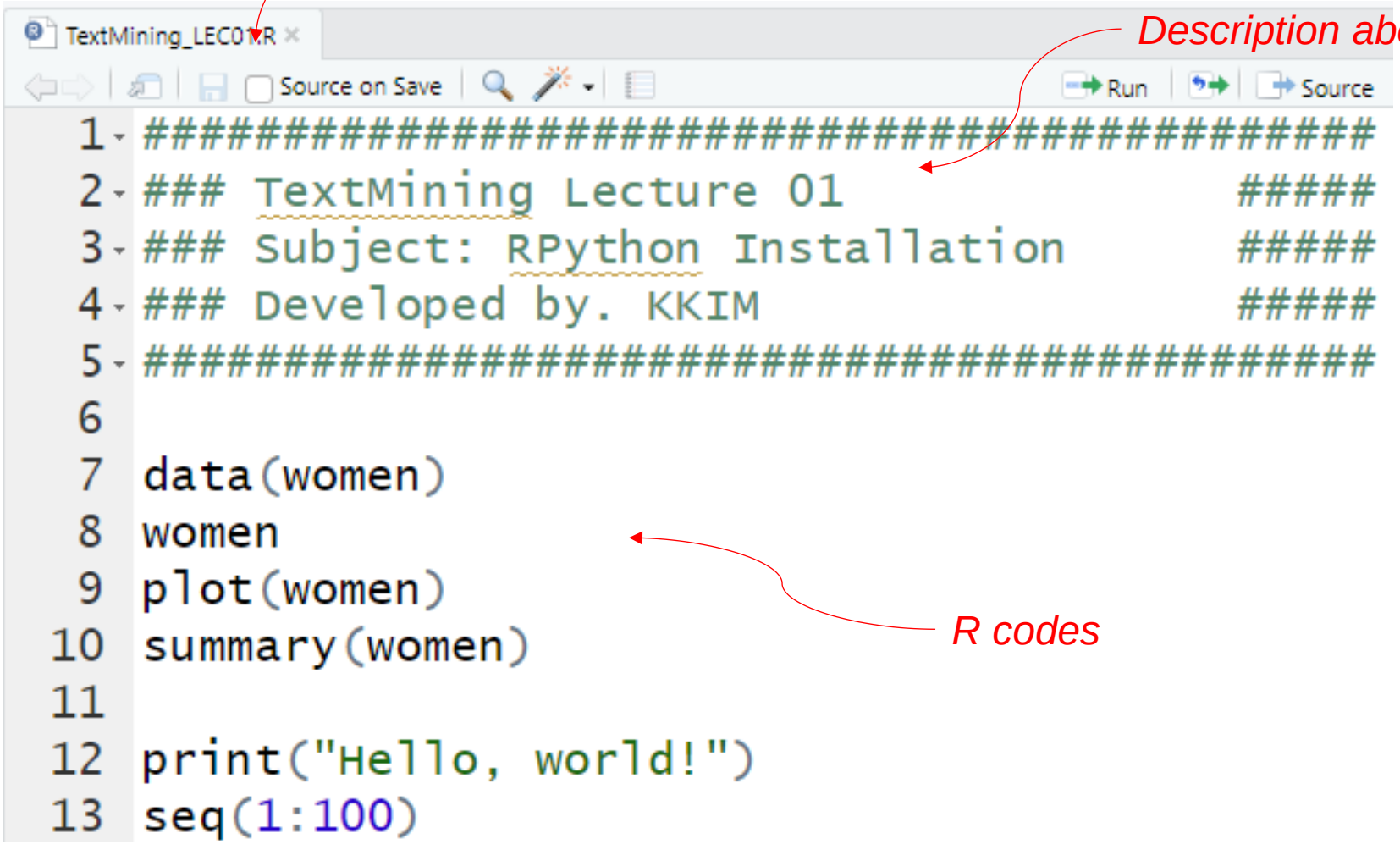


First R exercise

- Try

Rename File name

Description about the code



R codes

First R exercise

- Try

TextMining_LEC01.R

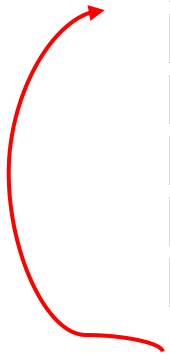
Source on Save | Run | Source

```
1  ##########################################
2  ### TextMining Lecture 01             #####
3  ### Subject: RPython Installation    #####
4  ### Developed by. KKIM               #####
5  ##########################################
6
7  data(women)
8  women
9  plot(women)
10 summary(women)
11
12 print("Hello, world!")
13 seq(1:100)
```

```
> data(women)
> women
   height weight
1      58    115
2      59    117
3      60    120
4      61    123
5      62    126
6      63    129
7      64    132
8      65    135
9      66    139
10     67    142
11     68    146
12     69    150
13     70    154
14     71    159
15     72    164
> plot(women)
> summary(women)
      height          weight
 Min.   :58.0   Min.   :115.0
 1st Qu.:61.5   1st Qu.:124.5
 Median :65.0   Median :135.0
 Mean   :65.0   Mean   :136.7
 3rd Qu.:68.5   3rd Qu.:148.0
 Max.   :72.0   Max.   :164.0
> print("Hello, world!")
[1] "Hello, world!"
> 1+1
[1] 2
```

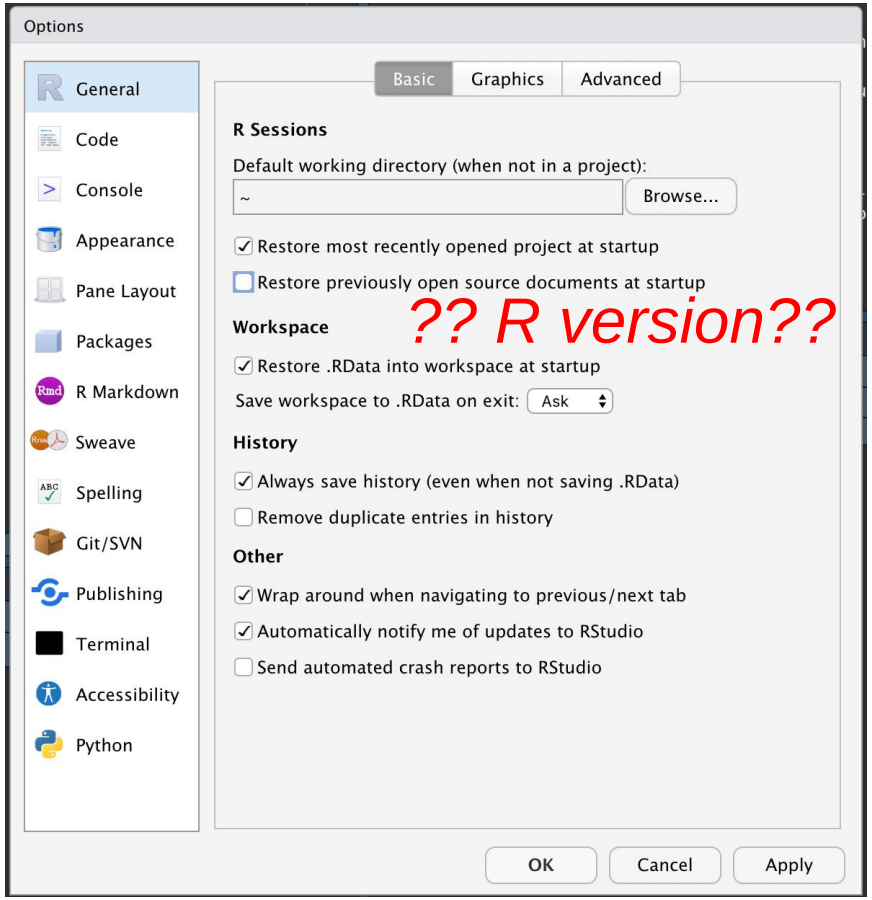
- Try

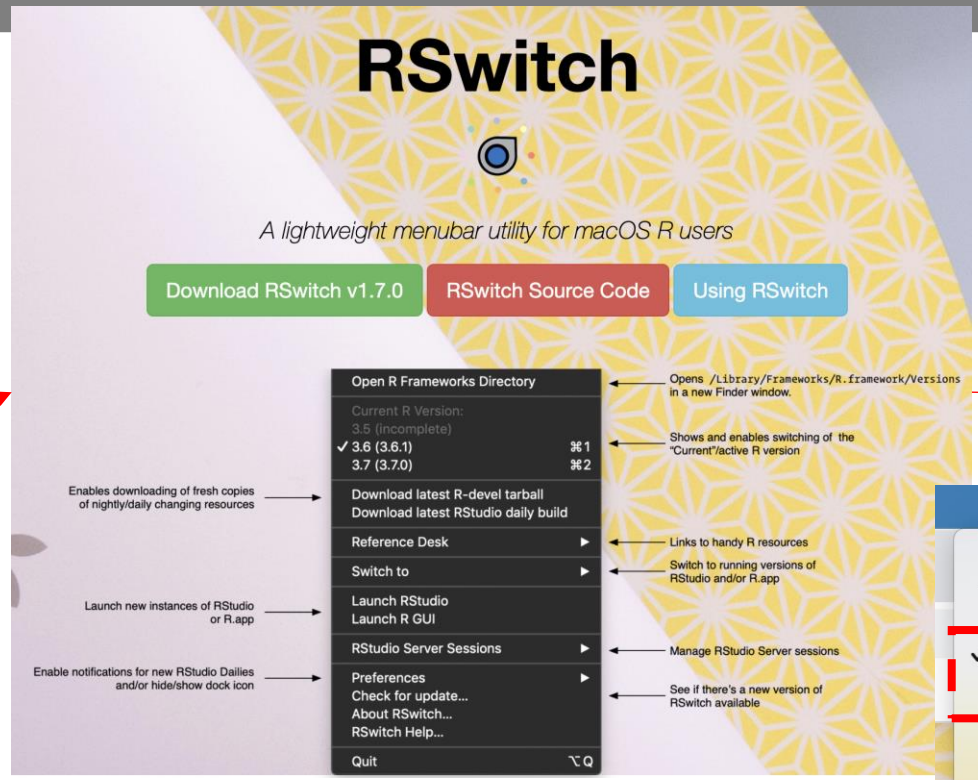
```
> print("Hello, world!")
[1] "Hello, world!"
> seq(1:100)
  [1]   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15
 [16]  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30
 [31]  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45
 [46]  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
 [61]  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75
 [76]  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90
 [91]  91  92  93  94  95  96  97  98  99 100
```

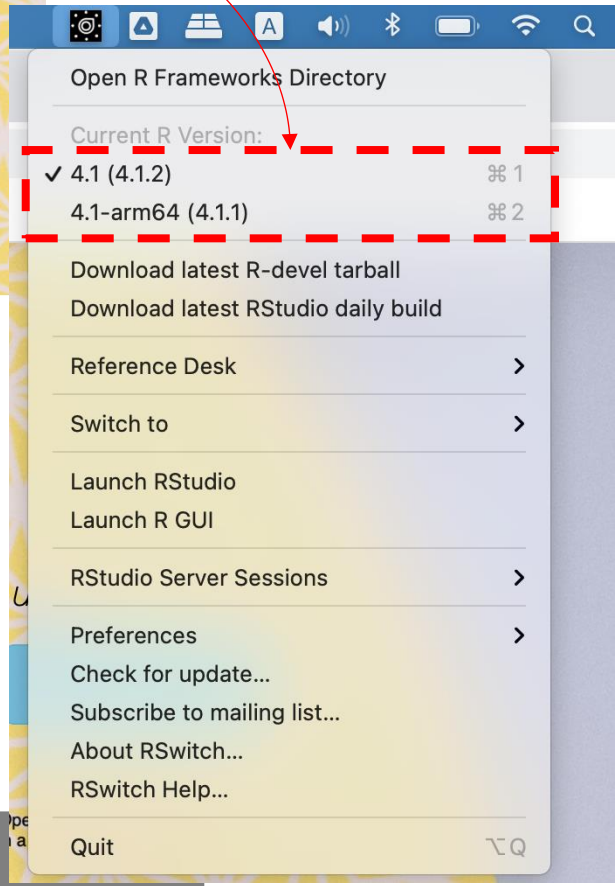


Index of result

- Changing R version in Mac
  - https://rud.is/rswitch/

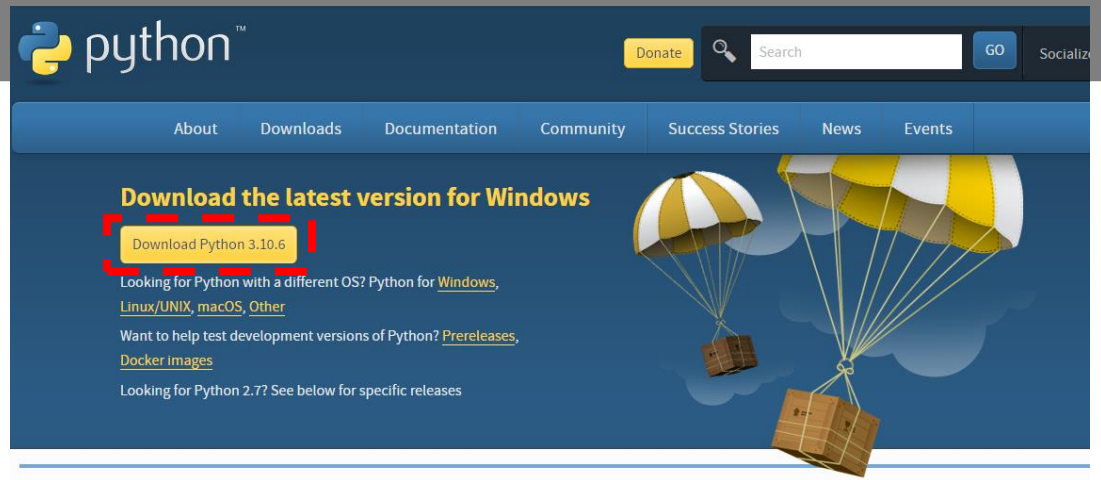


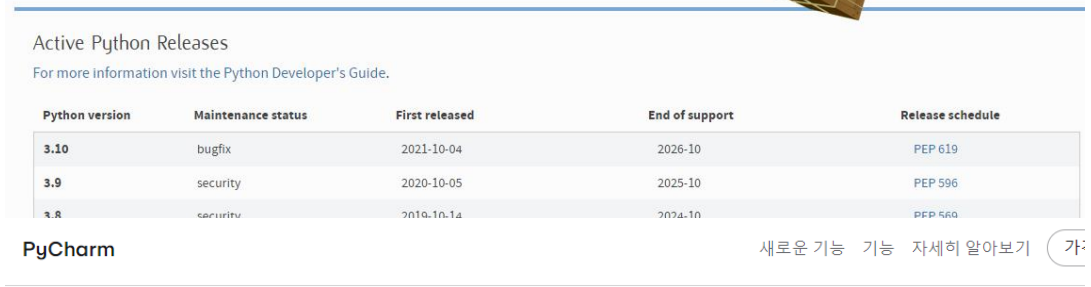




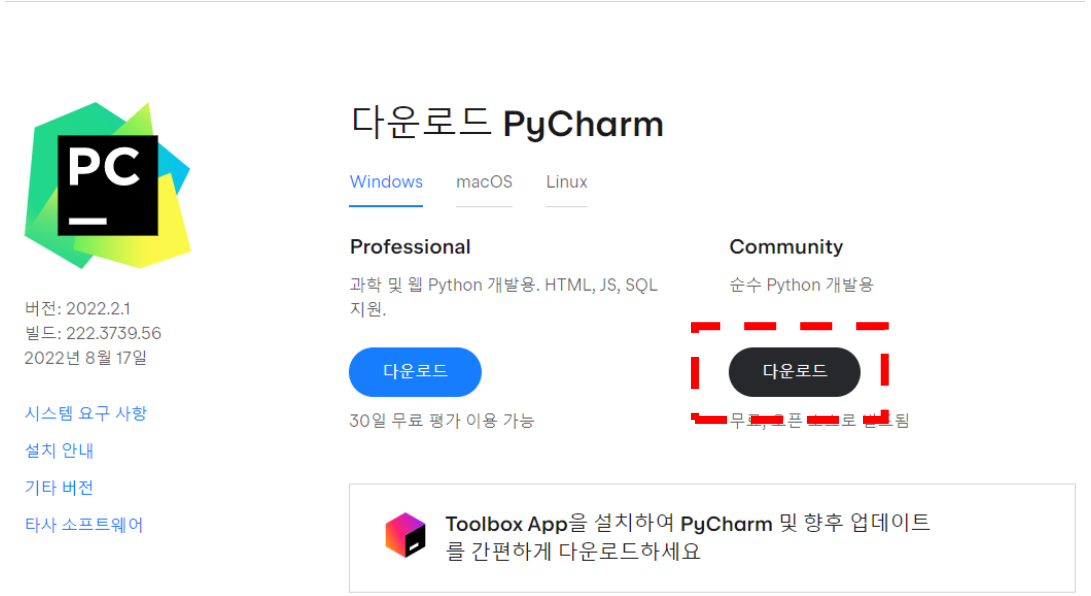
Python Installation

- Python
  - Download from https://www.python.org/downloads/
  - (search “download Python”)



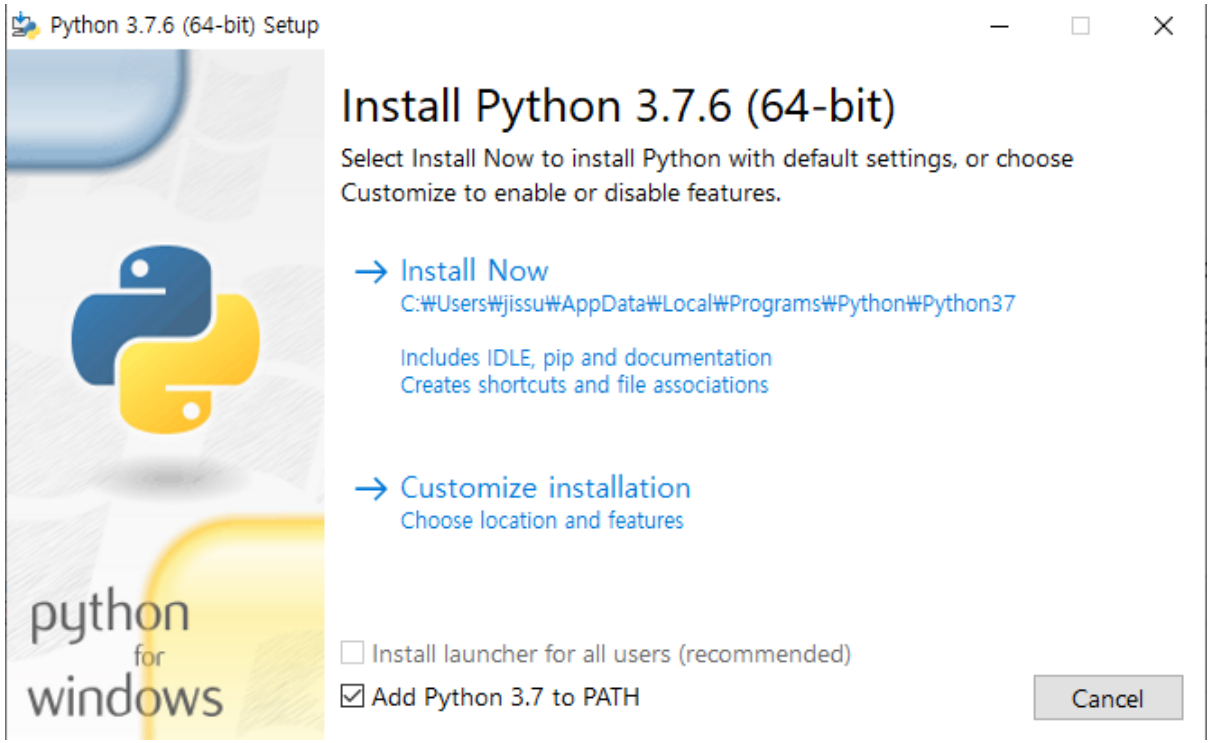


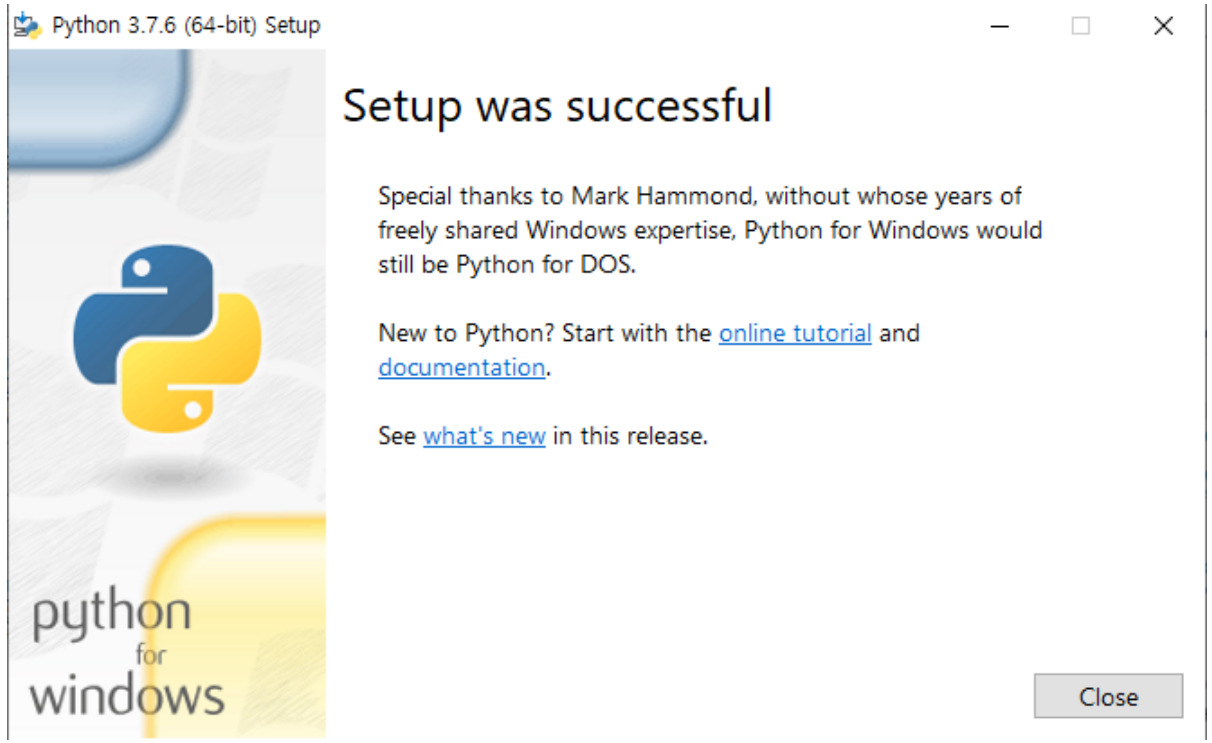
- Pycharm
  - Community version
  - Download from https://www.jetbrains.com/ko-kr/pycharm/download/#section=windows



Install Python

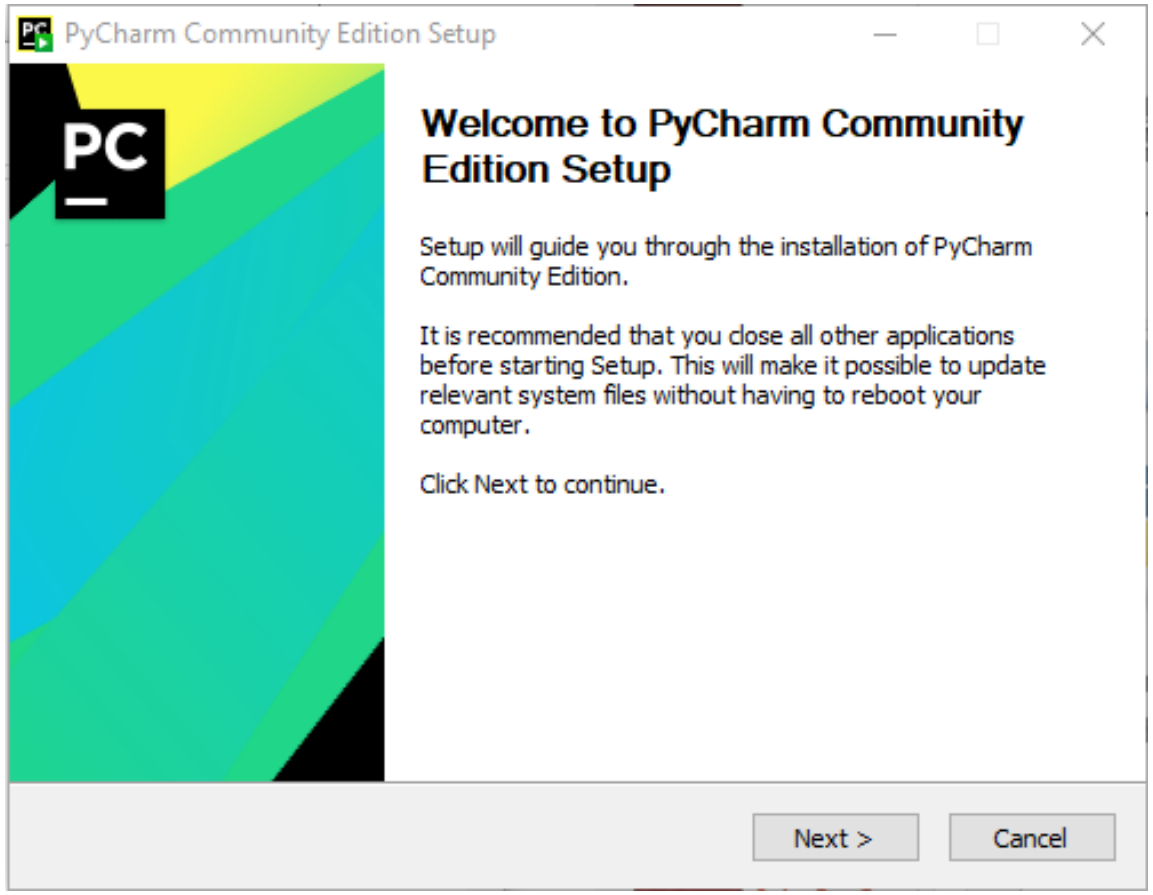
- Install with “Run as administrator”

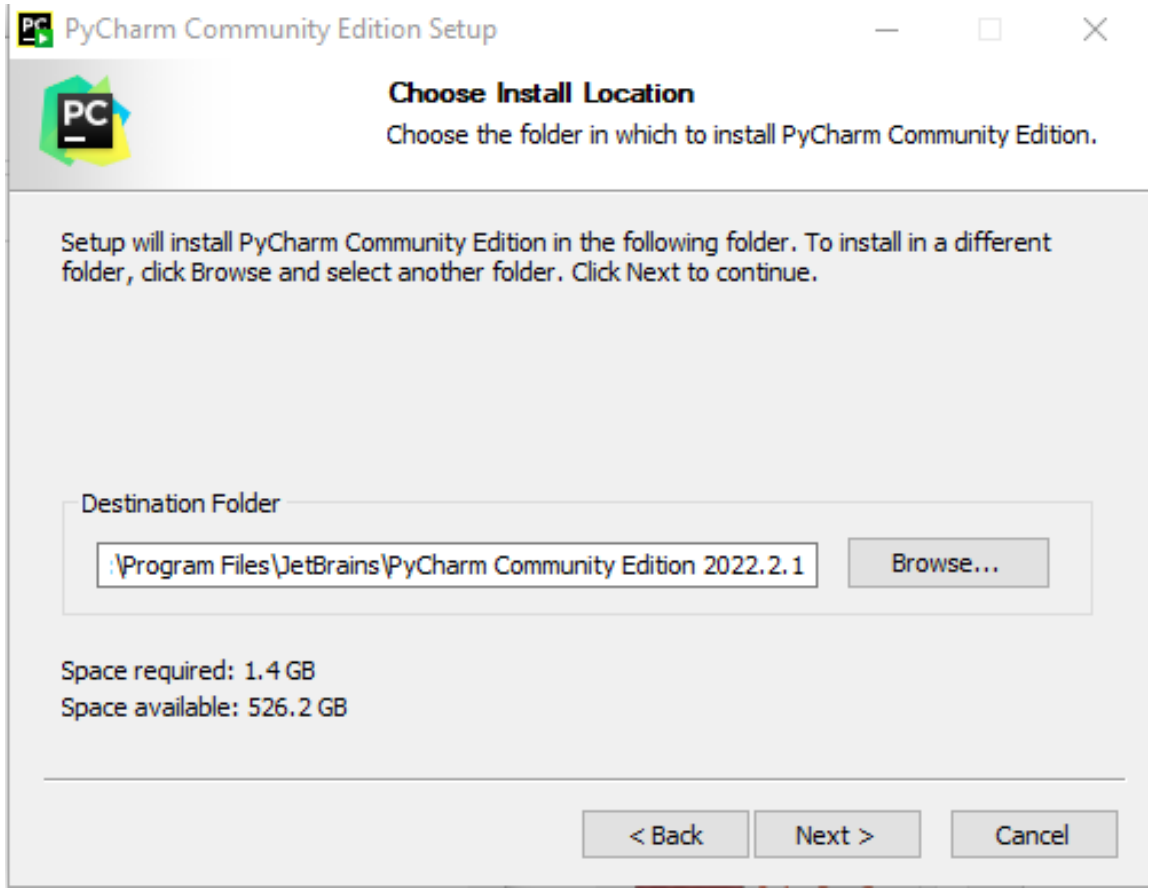




Install Pycharm

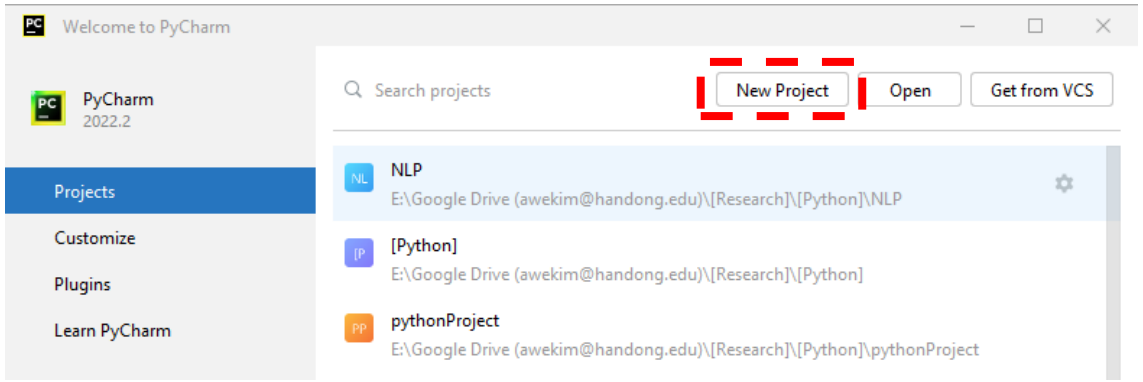
- Install with “Run as administrator”

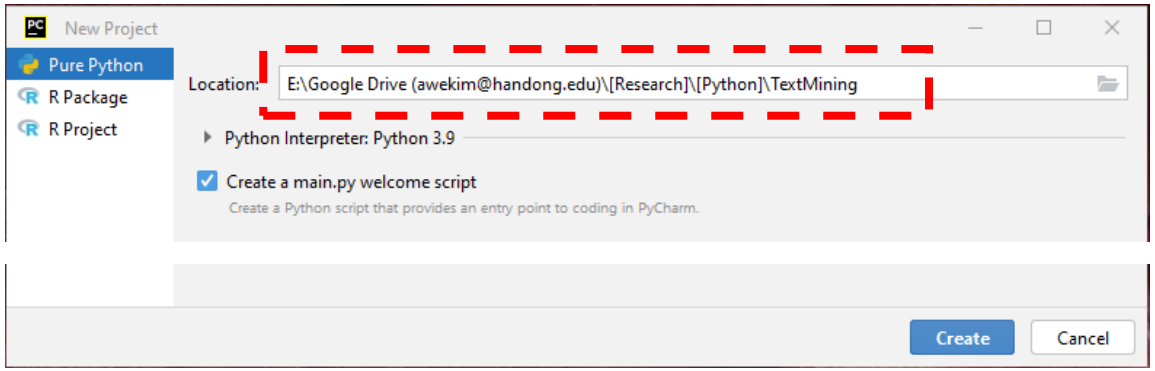


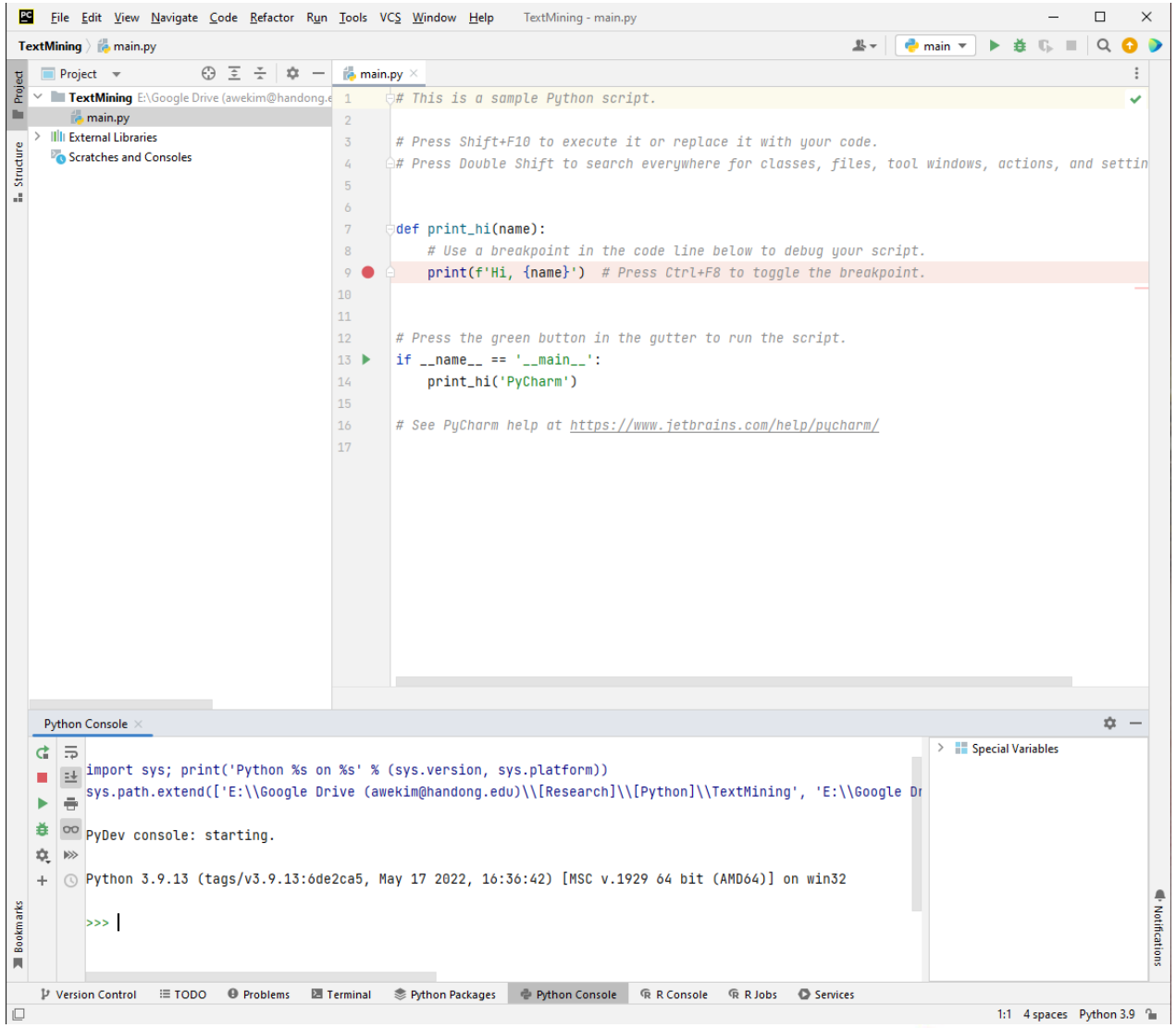


Run Python & Version Control

- Create new project

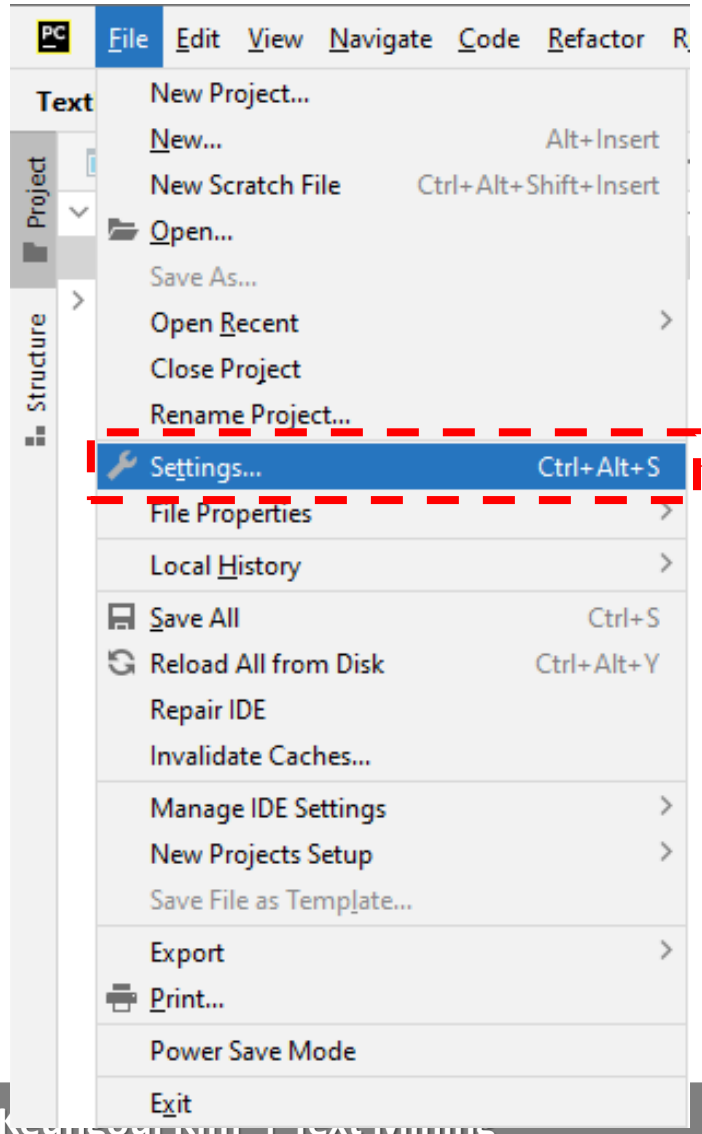


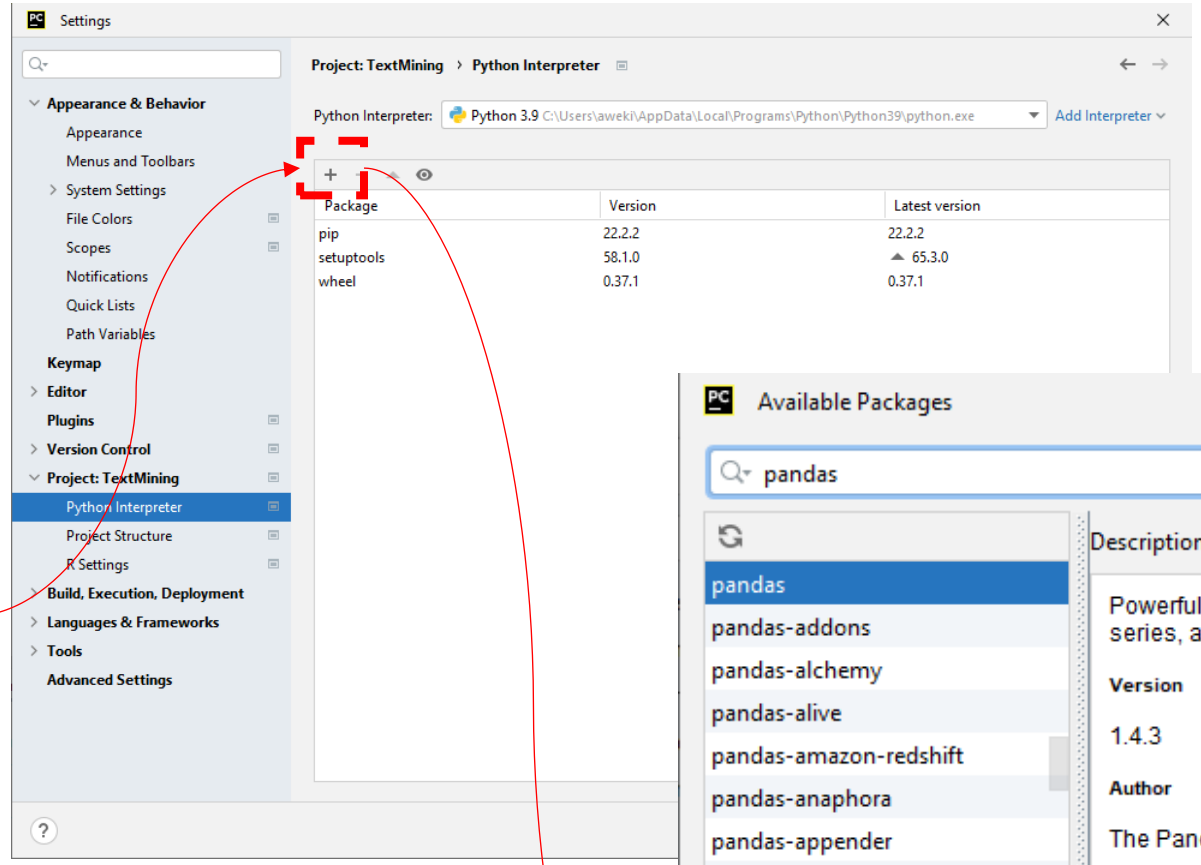


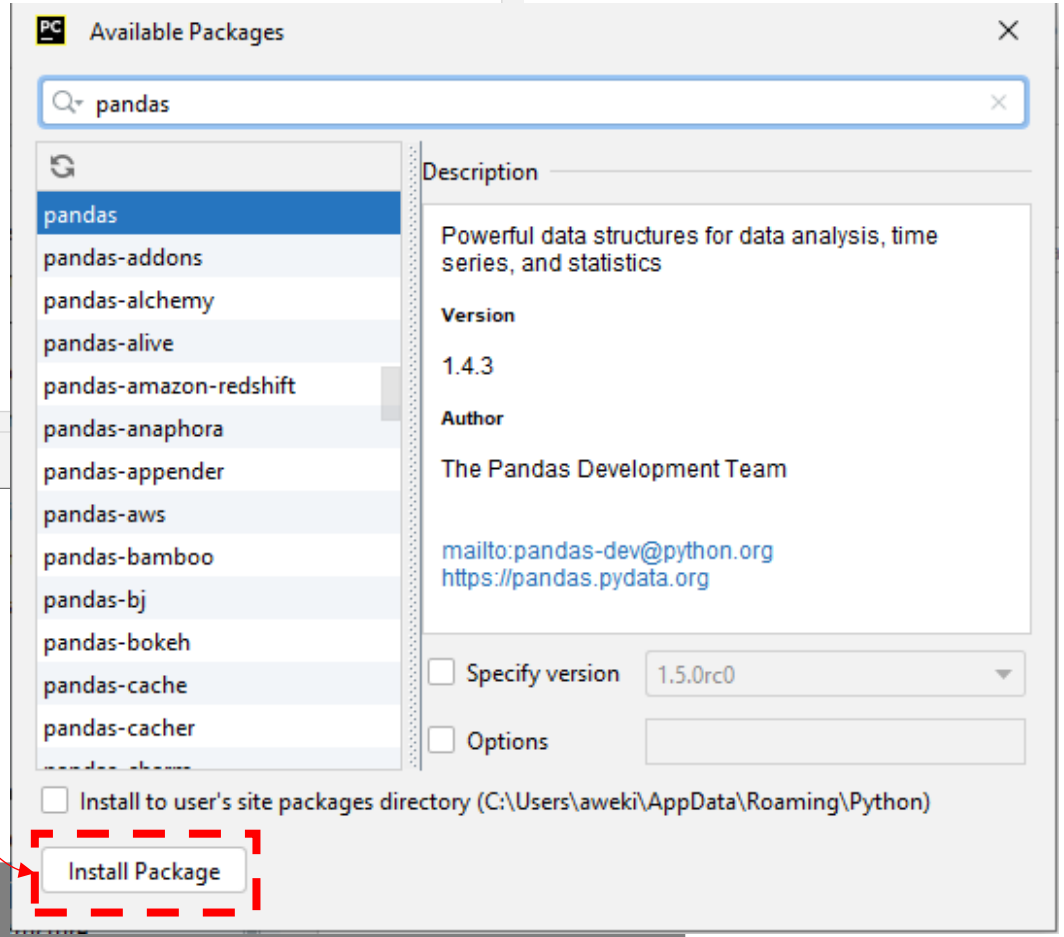


Run Python & Version Control

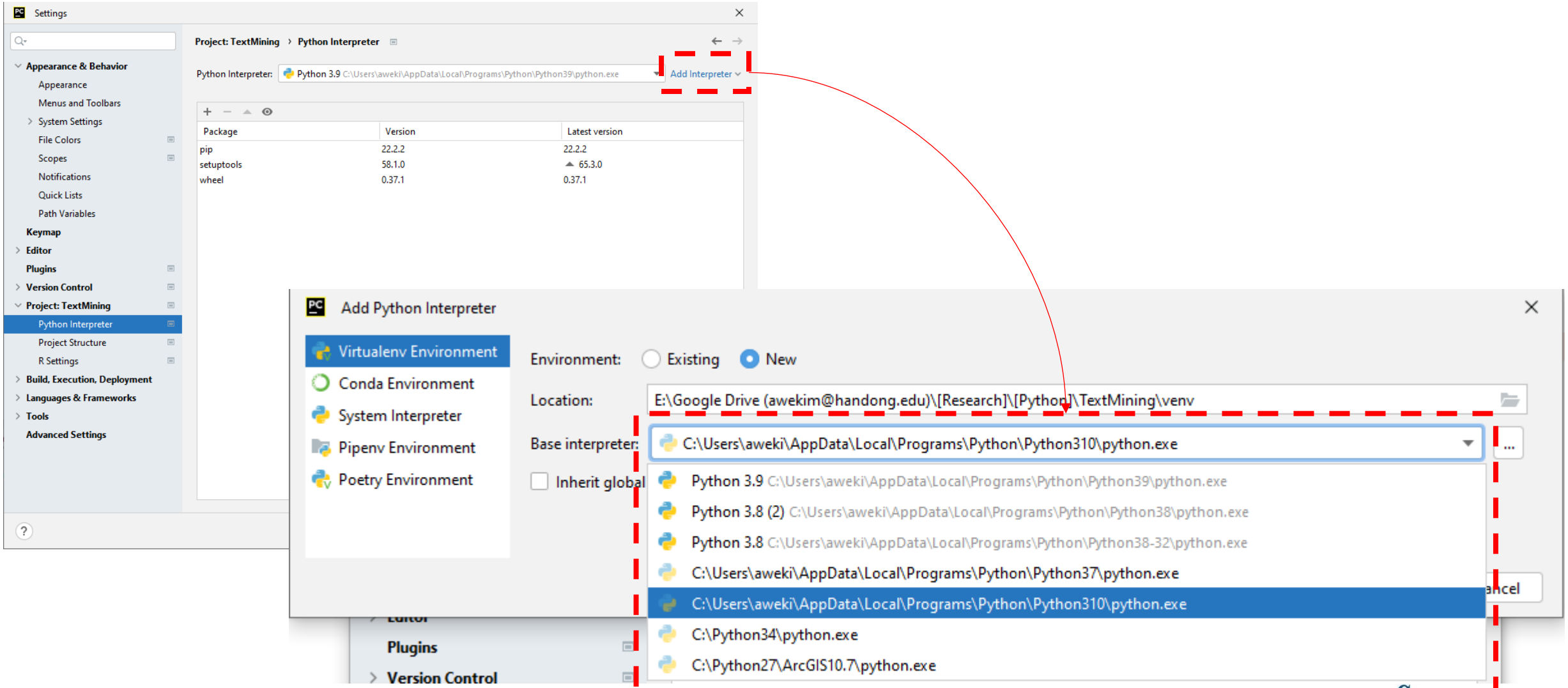
- Add new package







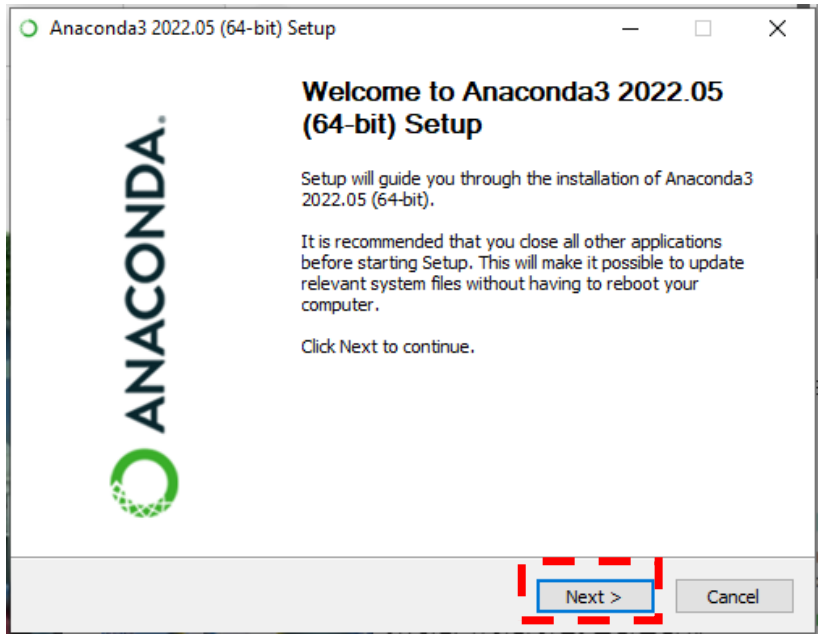
- Change Python version

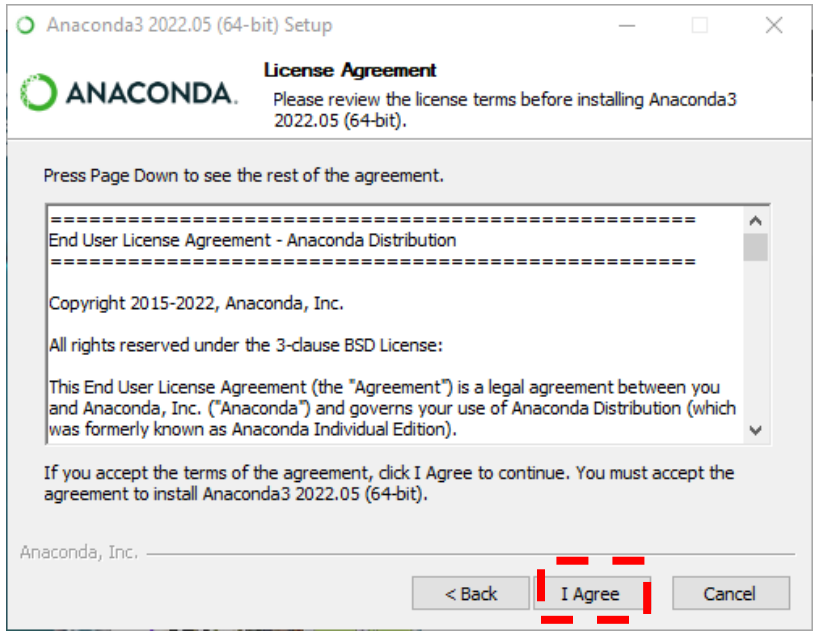


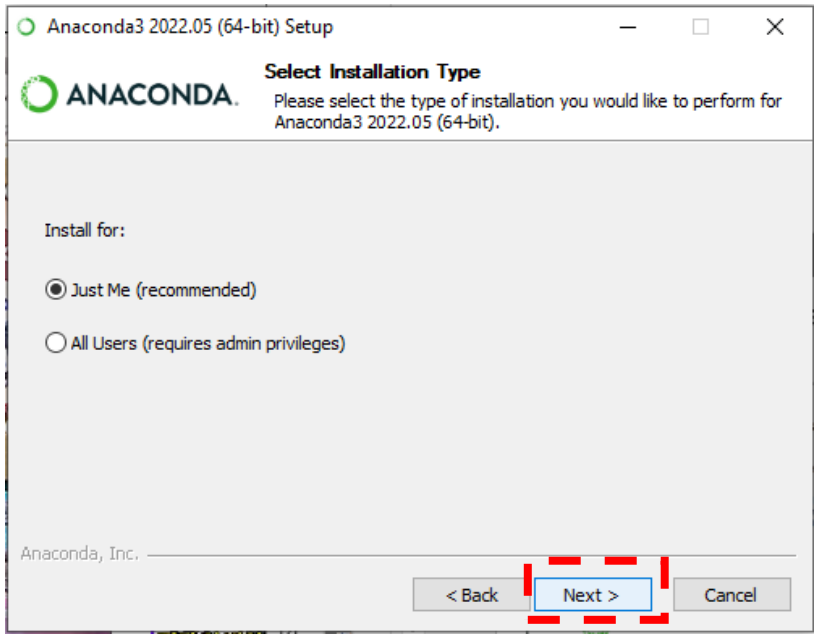
- Anaconda
  - Download from https://www.anaconda.com/products/distribution
  - Includes some useful packages as a default setting
  - Allows us to run a separate virtual environment
  - Often used for deep learning
  - If data storage is limited, you can use “miniconda”

Install Anaconda

- Install with “Run as administrator”

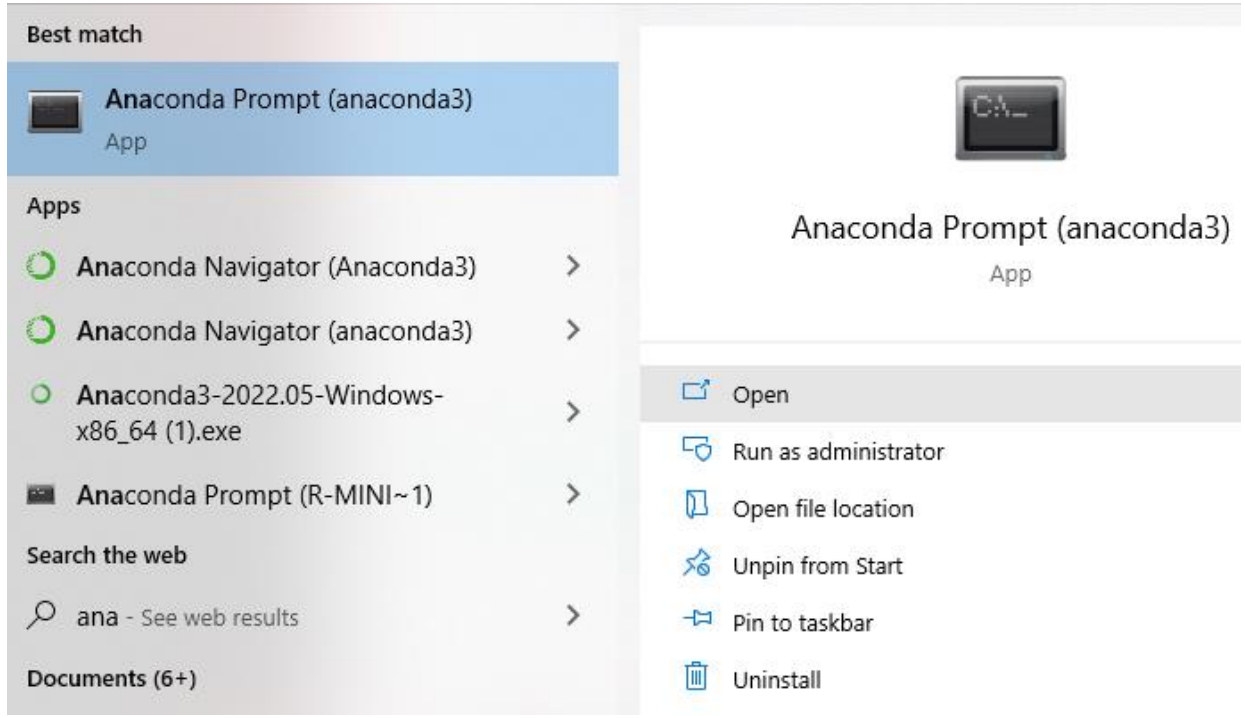






Install Anaconda

- Create and activate virtual environment



Name of virtual environment

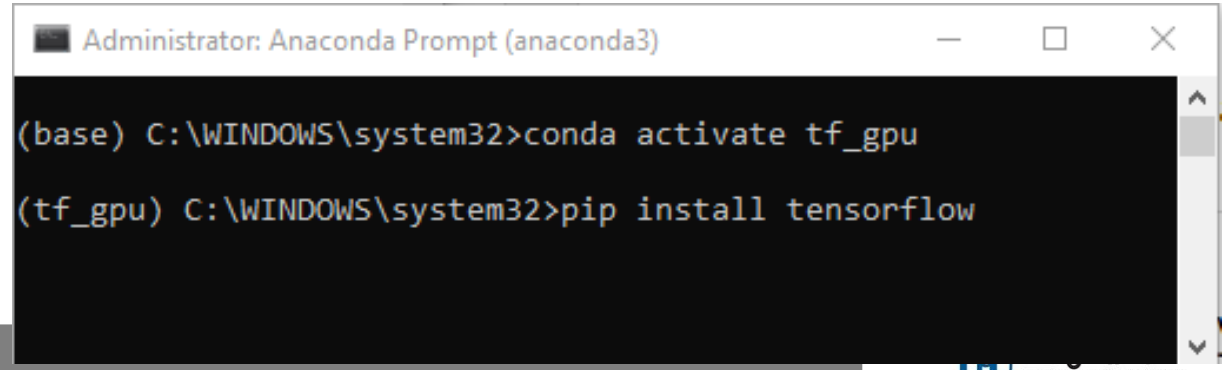
> conda activate –n *virtualenv* python=3.8

> conda env list *Returns list of conda environment*

> conda create –clone *virtualenv -n newvirtualenv*

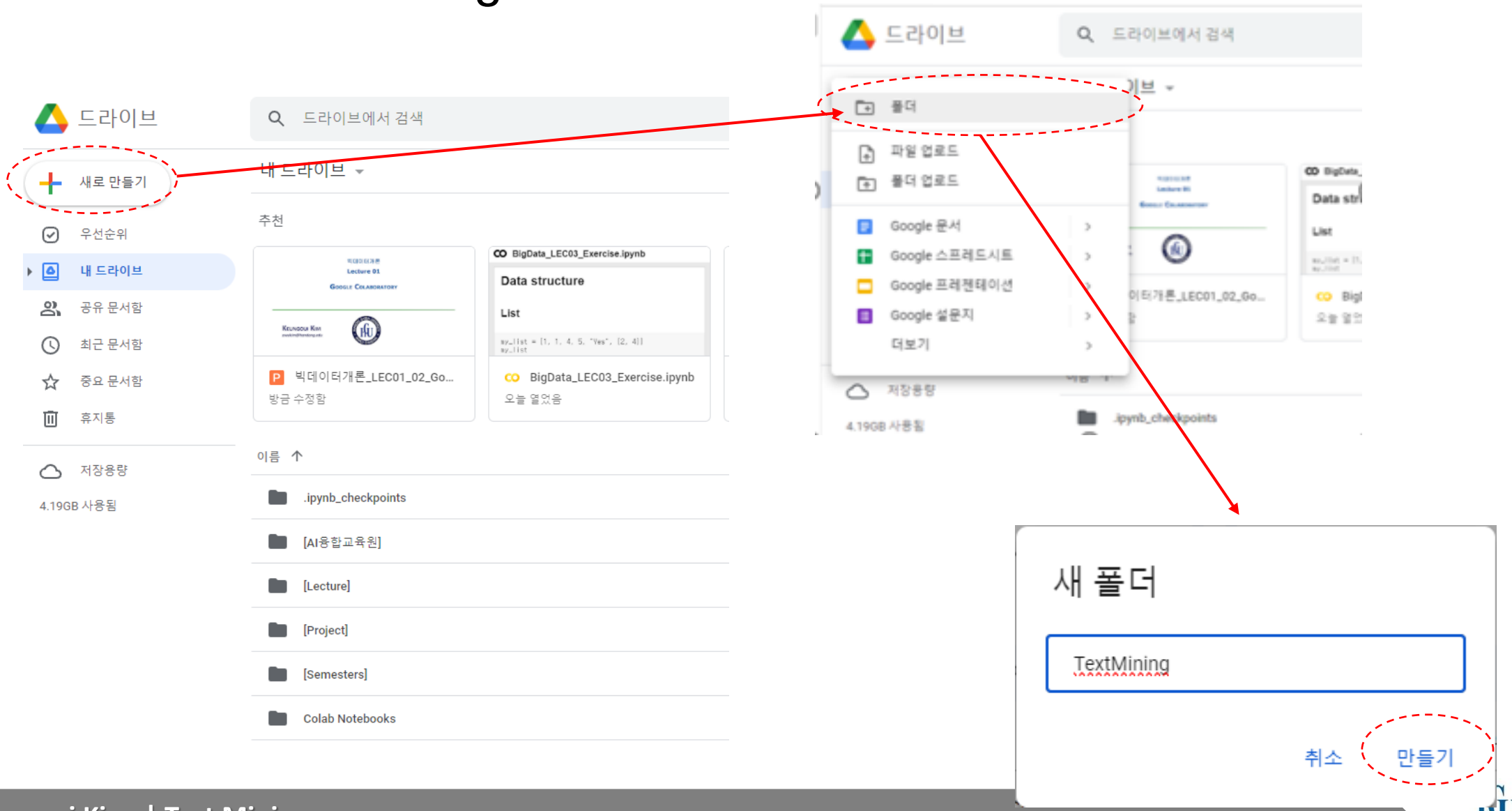
> conda activate *virtualenv*
> conda deactivate

> conda env remove –n *virtualenv* *Remove conda environment*

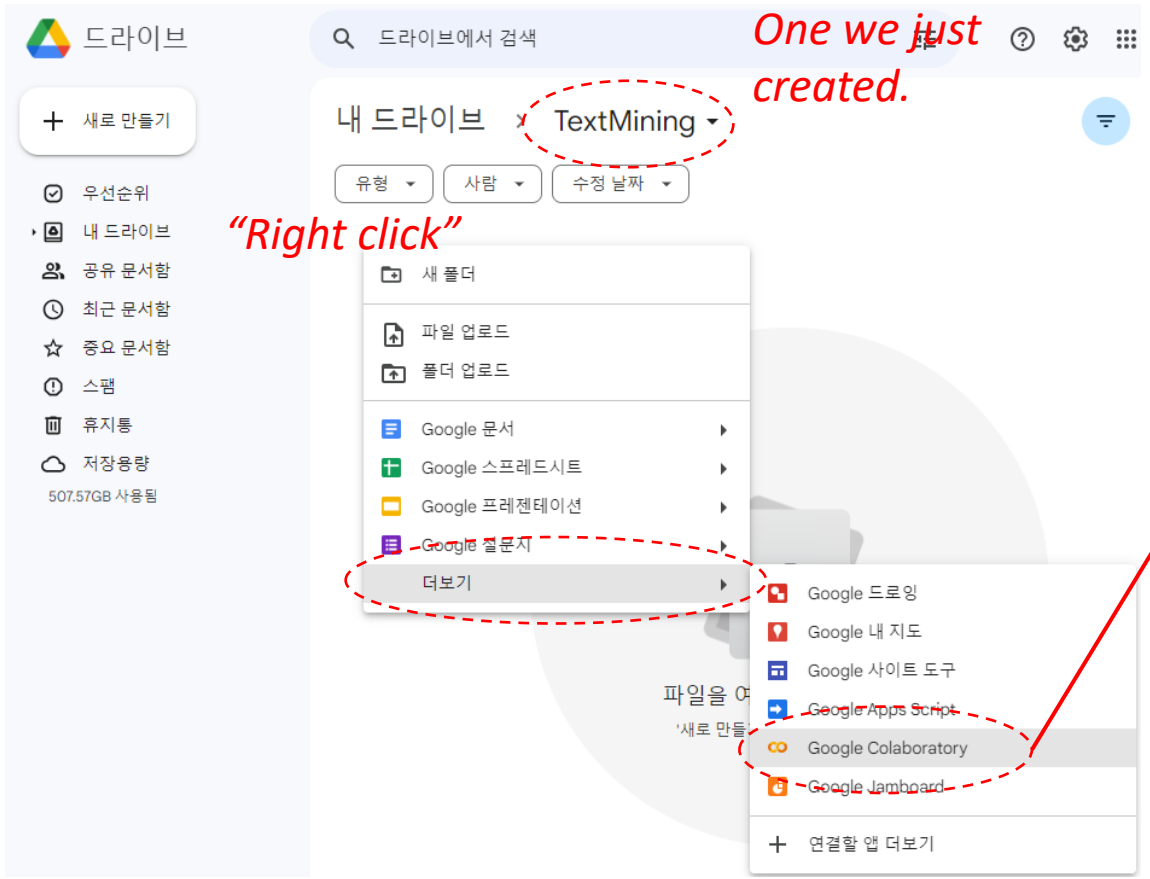


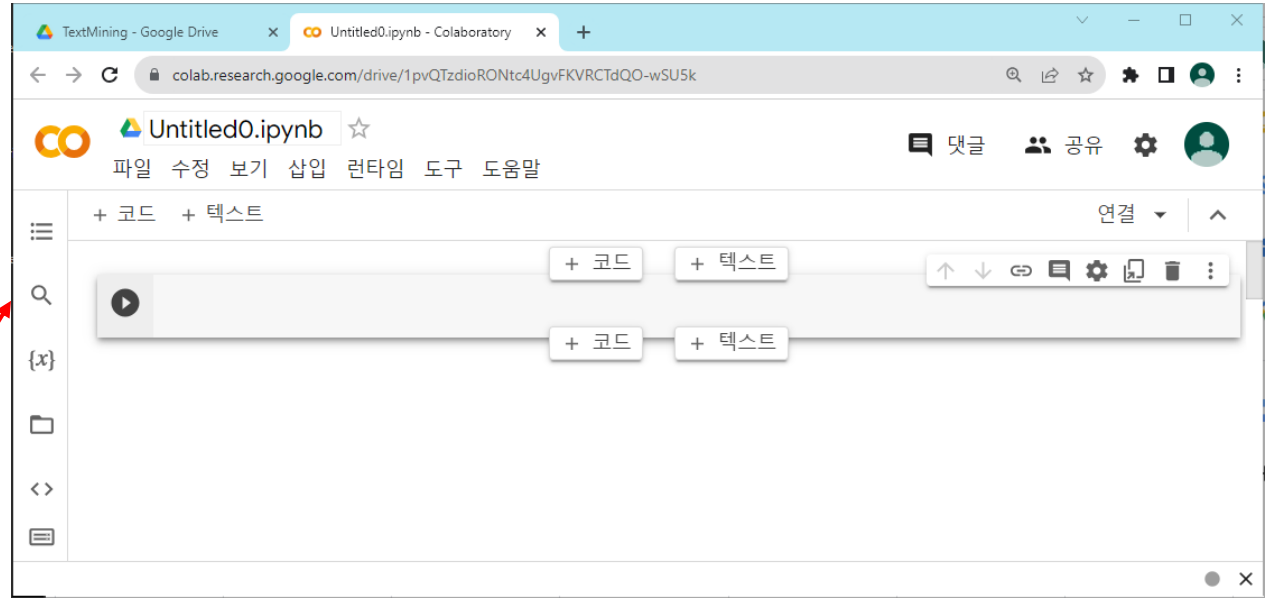
RPython in Google Colab

- Create folder for Google Colab



- Create Google Colab document





- Change Runtime type