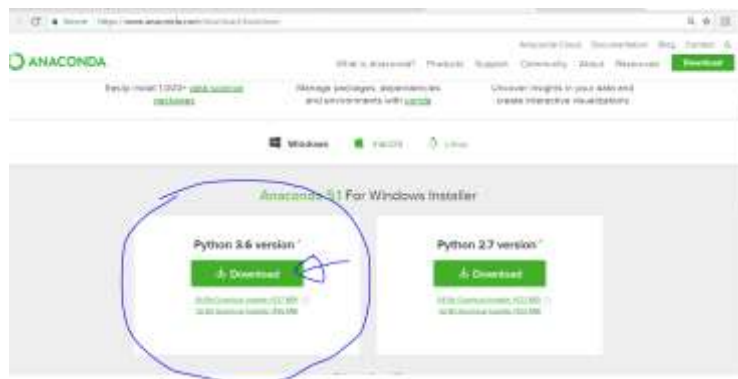


## Introduction to Data Mining – Practical 1

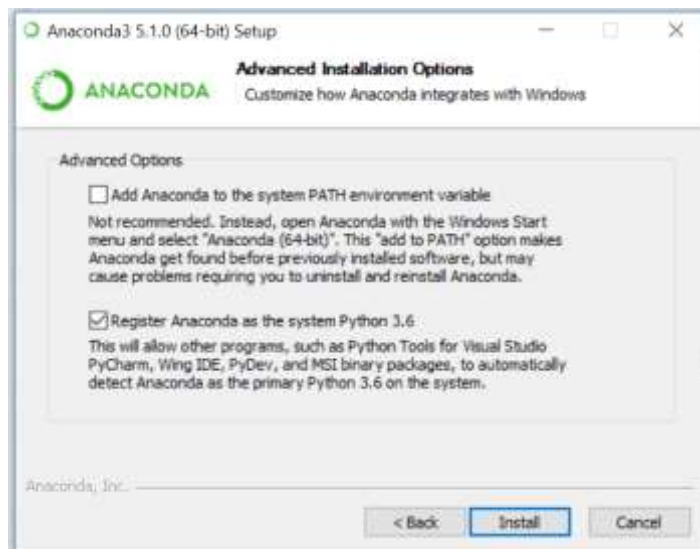
### Installing Software

#### Python

1. There are 2 main variations of Python in use (2.7 and 3.6). Python 2.7 is mainly useful for back-compatibility with certain libraries as Python 3 is more recent. Here we will use Python 3.6.
2. Depending on your operating system and hardware, install the Anaconda distribution (e.g. Windows 64 bit version) from <https://www.anaconda.com/download>



3. Use the default installation options:

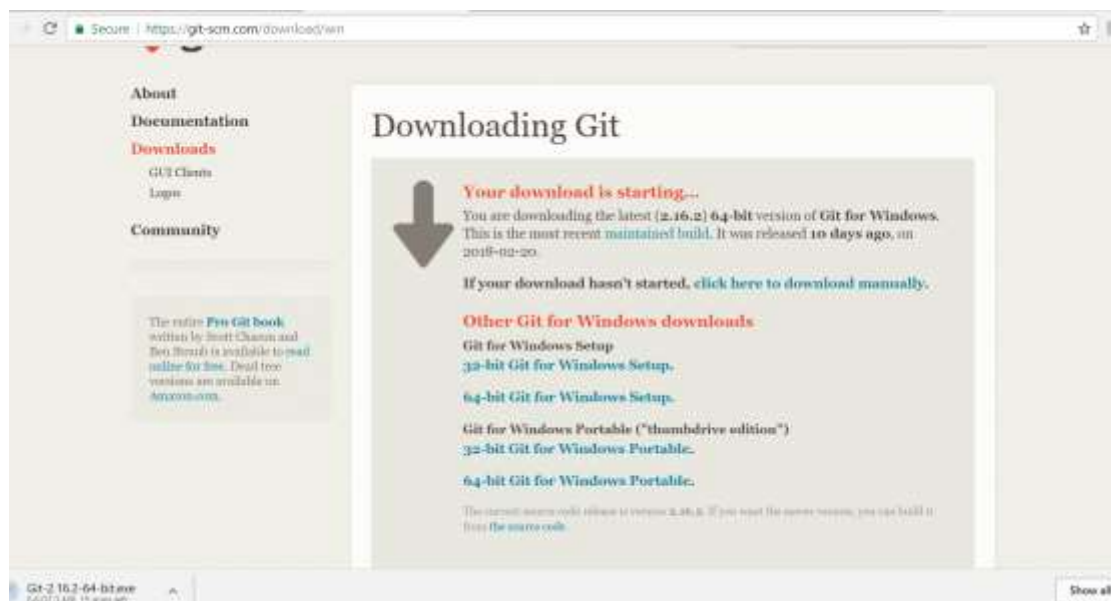


4. This will install both Python and Jupyter (IPython)
5. See also <https://jupyter.readthedocs.io/en/latest/install.html#new-to-python-and-jupyter>

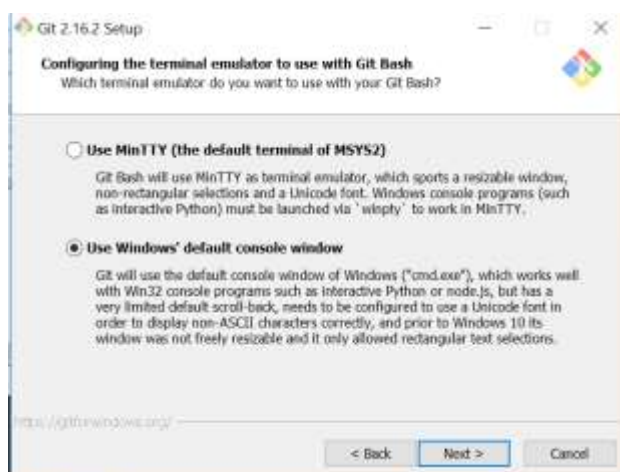
## Git

Git is a version control system, it is used in this tutorial to access the assignment description and data. Again, don't worry if you don't fully understand it yet or have not seen it before, more information will be provided as we progress in the course.

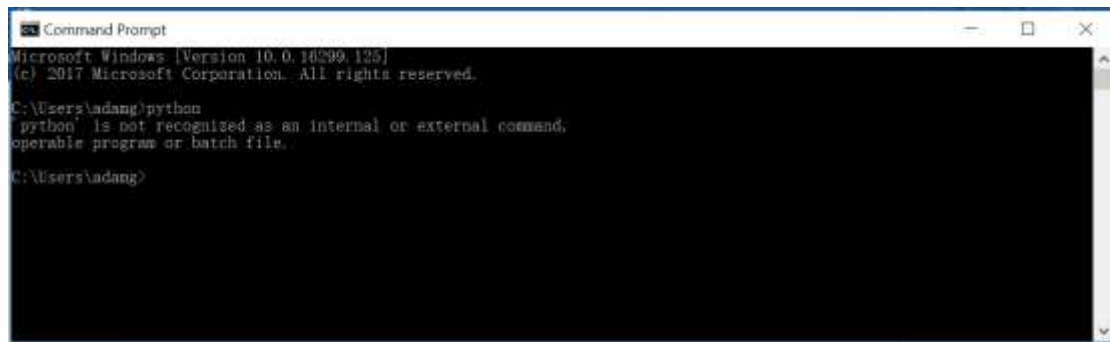
1. Download the appropriate release for your operating system from <https://git-scm.com/downloads>
2. Run the installer



3. Use the default options for the installer, you can use the following option for the terminal emulator to make it easier to run IPython (but then will not have resizable terminal windows):



4. Open a windows command prompt and type "python":



```
Microsoft Windows [Version 10.0.16299.125]
(c) 2017 Microsoft Corporation. All rights reserved.

C:\Users\adang>python
'python' is not recognized as an internal or external command,
operable program or batch file.

C:\Users\adang>
```

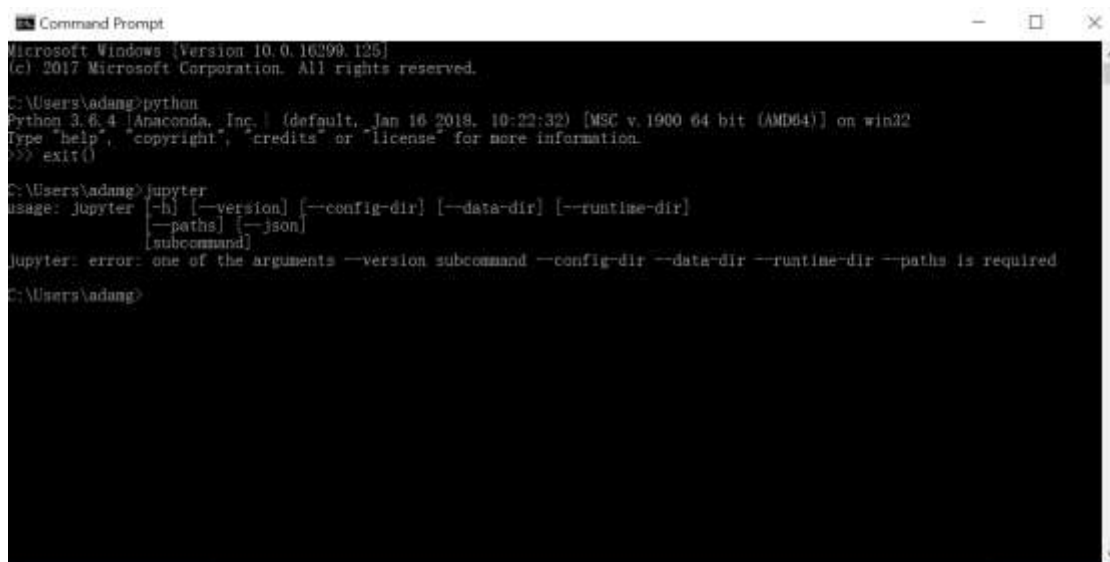
5. Add python and Jupyter to the path variable:

➤ SETX PATH

"%PATH%;C:\ProgramData\Anaconda3\Scripts;C:\ProgramData\Anaconda3"

(note: your path might be different if you performed the installation for a user – check under c:/users/your\_username)

6. Close and reopen the windows command prompt you are able to run Python:



```
Microsoft Windows [Version 10.0.16299.125]
(c) 2017 Microsoft Corporation. All rights reserved.

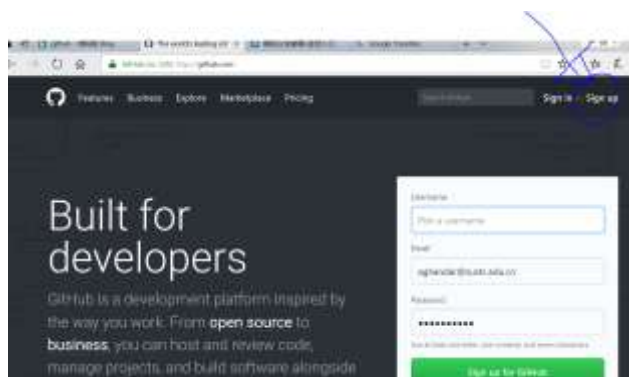
C:\Users\adang>python
Python 3.6.4 [Anaconda, Inc.] (default, Jan 16 2018, 10:22:32) [MSC v.1900 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>> exit()

C:\Users\adang>jupyter
usage: jupyter [-h] [--version] [--config-dir] [--data-dir] [--runtime-dir]
              [--paths] [--json]
              [subcommand]
jupyter: error: one of the arguments --version subcommand --config-dir --data-dir --runtime-dir --paths is required

C:\Users\adang>
```

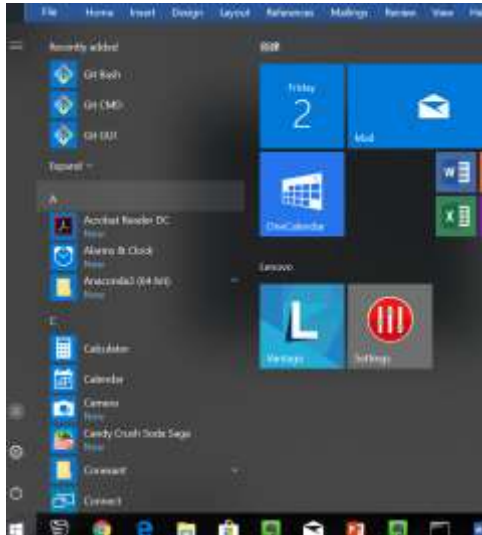
## Create a Github account

Navigate to <http://github.com> and create an account (sign up)

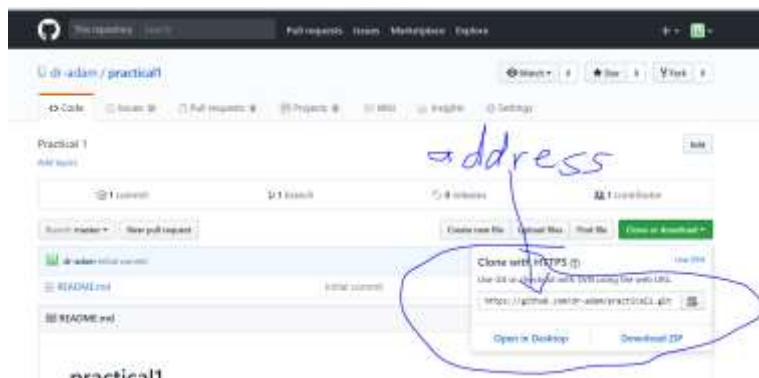


## Clone the practical 1 repo to your local machine

Open a gitbash terminal:



Navigate to the assignment 1 repo at <https://github.com/dr-adam/practical1> and create a clone on your machine (click on “Clone or Download to obtain the address”):



Now in the bash terminal create a directory for the assignment in your machine and then clone the assignment from the course repo to there:

```
MINGW64~/Users/adamg/intro_to_dm
commit      Record changes to the repository
diff        Show changes between commits, commit and Working Tree, etc
merge       Join two or more development histories together
rebase      Reapply commits on top of another base tip
tag         Create, list, delete or verify a tag object signed with GPG

collaborate (see also: git help workflows)
fetch       Download objects and refs from another repository
pull        Fetch from and integrate with another repository or a local branch
push        Update remote refs along with associated objects

'git help -a' and 'git help -g' list available subcommands and some
concept guides. See 'git help <command>' or 'git help <concept>'
to read about a specific subcommand or concept.

adam@LAPTOP-HPF8U8HF MINGW64 ~
$ mkdir intro_to_dm

adam@LAPTOP-HPF8U8HF MINGW64 ~
$ cd intro_to_dm

adam@LAPTOP-HPF8U8HF MINGW64 ~/intro_to_dm
$ git clone https://github.com/dr-adam/practical1.git
Cloning into 'practical1'...
remote: Counting objects: 3, done.
remote: Total 3 (delta 0), reused 0 (delta 0), pack-reused 0
Unpacking objects: 100% (3/3), done.

adam@LAPTOP-HPF8U8HF MINGW64 ~/intro_to_dm
$
```

\*\* unix commands:

- Change directory: ls
- Make new directory: mkdir <directory name>
- Change directory: cd