Notes on Data Science

This file contain my notes on various data science techniques.

Google for Education

Google for Education: Python is a website providing basic training on Python. I first worked on it on 30 April 2023.

The resources have been download to the "C:\Users\andyw\google-python-exercises" at the ASUS Zenbook.

There was some mistake in the `hello.py` file downloaded. Parentheses are missed. After adding them the code worked.

Command line, Python interpreter and IPython

This topic is based on Section "Python Setup" in Google for Education: Python, which I worked on it on 30 April 2023.

There should be some difference between running Python program file at **command line** and **run Python code interactively at Python interpreter**.

Command line

Enter **`cmd`** at the Start menu to enter the terminal.

At the **command line**, you type **`python filename.py`** to run the Python programme file **`filename.py`**.

For example, if you are at the directory which it contain the Python file, you can type the following to execute the file.

```
C:\Users\andyw\google-python-exercises>python hello.py
Hello World
```

Or you can enter the full path of the Python file if you are not at the directory which contain the file.

```
C:\Users\andyw>python C:/Users/andyw/google-python-exercises/hello.py
Hello World
```

You can execute Python programme file at the command line but you cannot run Python code directly at the command line.

Python interpreter

To run **Python code interactively** at a Python interpreter, select the Run... command from the Start menu, and type "Python".

When you type **`python`** or **`py`** at the command line, you enter the **Python interpreter** and can run Python code directly by entering code at the prompt.

For example:

```
>>> print("Hello World")
Hello World
```

On Windows, use **`Ctrl-Z`** and then press **`Enter`** to exit (on all other operating systems it's **`Ctrl-D`** to exit).

IPython

You type **`ipython`** at the start menu or at the command line launch the IPython shell.

```
C:\Users\andyw>ipython
... some message ...

In [1]:
```

You can execute arbitrary Python statements by typing them and pressing Return (or Enter).

```
In [1]: a = 5

In [2]: b = 9

In [3]: a + b
Out[3]: 14
```

You can execute a Python programme file by **`%run`** command and entering the full path of the Python programme file.

```
In [4]: %run C:/Users/andyw/google-python-exercises/hello.py
Hello World
```

JupyterLab and Github

Launch Jupyter Notebook and JupyterLab

Type **`jupyter notebook`** or **`jupyter lab`** at the command line. These work on both ASUS ZenBook and Dell desktop.

```
# Launch Jupyter Notebook
C:\Users\andyw>jupyter notebook

# Launch JupyterLab
C:\Users\andyw>jupyter lab
```

Set up GitHub linkage with the JupyterLab

In order to enable JupyterLab to interact with GitHub account, you have to install **git extension** in JupyterLab.

To install git extension, you should follow the steps below:

1. You have to install JupyterLab.
2. You have to install **git**. See https://github.com/git-guides/install-git
3. To install git extension perform the following steps, with pip at the command line:

   ```
   C:\Users\andyw>pip install --upgrade jupyterlab jupyterlab-git
   ```

   (See https://github.com/jupyterlab/jupyterlab-git for reference.)

Connection to GitHub using JupyterLab

Only some function is working in the Dell desktop. The ASUS ZenBook has not yet set up link with the GitHub account.

Everytime you use JupyterLab, you better synchronize the GitHub account repository by go to the tap `Git` and select `Pull from Remote` first.

Add a new notebook

1. Create a new notebook
2. On the left panel under the `GitHub` icon and under `Untracked`, select the new notebook and hit `+` (i.e. to change to track this file)
3. Under `Staged` the number should become from 0 to 1.
4. At the lower part of the left panel at come comment in the box provided and then hit `COMMIT`
5. At the tap `Git` on the top, select `Push to Remote`

Existing file amended

1. On the left panel under the `GitHub` icon and under `Changed`, hit `+` (i.e. to stage all changes)
2. At the lower part of the left panel at come comment in the box provided and then hit `COMMIT`
3. At the tap `Git` on the top, select `Push to Remote`

Deleting an existing file

1. On the left panel under the `File Folder` right click the file you want to delete and then select `Delete`
2. On the left panel under the `GitHub` icon and under `Changed`, hit `+` (i.e. to stage all changes)
3. At the lower part of the left panel at come comment in the box provided and then hit `COMMIT`
4. At the tap `Git` on the top, select `Push to Remote`

Useful reference

You may find following tutorial on youtube useful:

1. Using the JupyterLab Git Extension in a Data Science Notebook Session
2. Introduction to JupyterLab

YAML Headers

Please refer to the following link for introduction to the YAML Headers:

https://zsmith27.github.io/rmarkdown_crash-course/lesson-4-yaml-headers.html

Some templates have been prepared in the `YAML Headers` sub-folder.