

情報システム工学特論 (Advanced Information System Engineering)

概要 Abstract

この授業では、データベースシステムに関する種々の演習を実施する。演習では、各自、Windows ノート PC を準備し、python などのソフトウェアをインストールしてきて欲しい。今日の授業では、Python 言語からのリレーショナルデータベースの使用について説明する。リレーショナルデータベースは、複数のテーブルから構成される。テーブルは、数値、テキスト、日時などの情報を、テーブル形式で格納している。リレーショナルデータベースの各種操作を行う言語の世界標準が SQL である。SQL プログラムは、他のプログラミング言語と組み合わせて使用されることが多い。この授業では Python 言語を使う。

In this class, we do several exercises on advance database system topics. Students are required to prepare and bring their Windows note PC, and to install the python software and other sorfwares if possible. Today, we learn how to use a relational database from Python language. Relational database contains relational tables. They stores numeric, text, and datetime information in table format. It is easy to handle. SQL is the standard relational database language to handle relational database. SQL program is often embedded in other programming languages to connect to database management systems, to process relational tables further. Python is one of famous programming language. It is easy to use.

■ テーブル (Relational Table)

リレーショナルデータベースでは、データベースをテーブルの集まり (collection of tables) として記述する。各テーブルには、識別に使うテーブル名があり、テーブルの各列には、識別に使う列名がある。

A relational database is a set of relational tables. Each table has a table name. Each column of a table has a unique attribute name.

【テーブルの例】

id	sepal_length	sepal_width	petal_length	petal_width	species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa

■ Python の繰り返し処理 Python loop

次のプログラムでは、変数 t はデータの集まりであるとする。row は t の各要素について処理を繰り返すための変数になる

In the following program, variable t is a data collection. A variable row is used to loop over each item of t.

```
for row in t:
    print (row)
```

■ Python の SQL プレースホルダー Python SQL place holder

次のプログラムでは、SQL 文の中に、プレースホルダーを示す記号「?」が使われている。「?」の部分が「4」で置き換わり、"select * from iris where id = 4" が評価される。これは Python の機能である。

In the following program, '?' is used to indicate a place holder in a SQL sentence embedded in a Python program. The '?' is replaced by '4', and "select * from iris where id = 4" is evaluated. It is a function of Python.

```
In [27]: cur.execute(u"select * from iris where id = ?", [4])
Out[27]: <sqlite3.Cursor at 0x66afa60>

In [28]: for t in cur:
...:     print t
...:
(4, 4.6, 3.1, 1.5, 0.2, u'setosa')
```

■ CVS ファイル CSV File

カンマで区切られたデータファイル

A CSV file a comma seperated data file

演習 1 (Exercises 1)

(1) ステップ 1 (Step 1)

Windows で Anaconda をインストールしなさい。

Install the Anaconda software on Windows

Anaconda は、Continuum Analytics 社が提供している **Python 処理系と管理用ツールである conda と主要な Python パッケージ**を1つにまとめたソフトウェア。次のアプリケーションも同封されている。

- **jupyter**: Web ペースで動くデータ解析環境。Python, Julia, Ruby, R, Lua, LuaJIT, Haskell, Scala, Go, JavaScript, node.js, bash などに対応。
- **qtconsole**: コンソール。図などにも対応。
- **spyder**: Python 開発環境
- **Anaconda Navigator**: アプリケーションの起動や管理などができるアプリケーション

※ OpenAI を使う予定がある場合には、古い版である Python 3.5 系列を使うこと (2017/03 時点では)。

Anaconda is a software that includes Python processing system, the management tool 'conda' and Python's packages. They are provided by Continuum Analytics. following applications are also included.

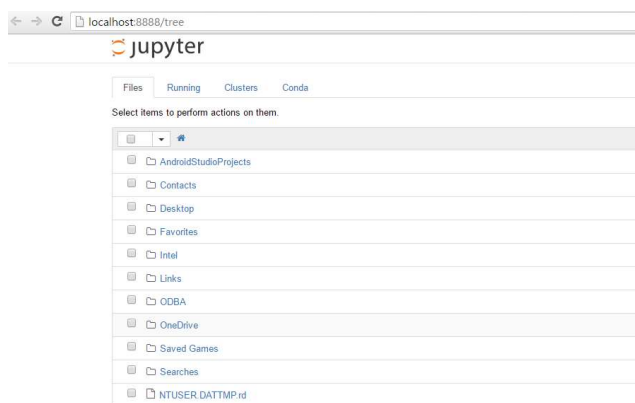
* jupyter: A Web-based analytic data analysis environment that supports Python, Julia, Ruby, R, Lua, LuaJIT, Haskell, Scala, Go, JavaScript, node.js, bash, etc.

* qtconsole: console softwares. It can display figures.

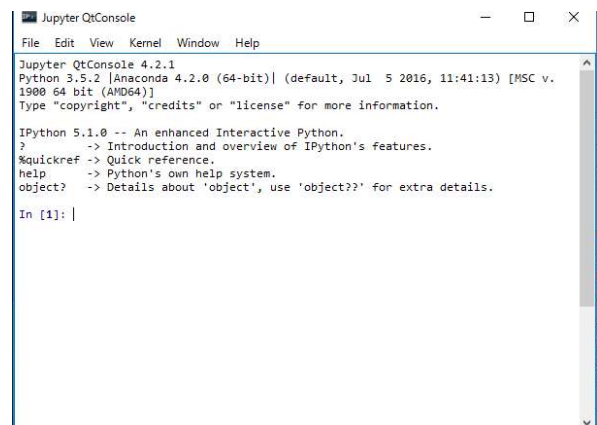
* spyder: Python development environment

* Anaconda Navigator: An application that can launch and manage applications

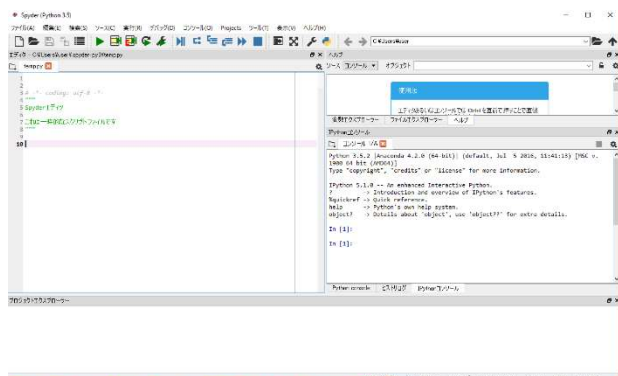
Use Python 3.5 series for compatibility with OpenAI.



jupyter



qtconsole



spyder

◆ ダウンロードとインストール Download and Install

- ① ダウンロード用ウェブページを開く Open the Web page to download

<https://www.continuum.io/downloads>

- ② Python の最新版を使いたい. For Windows の右のボタンをクリック.

Click "64-BIT INSTALLER" to download the latest version of Anaconda.



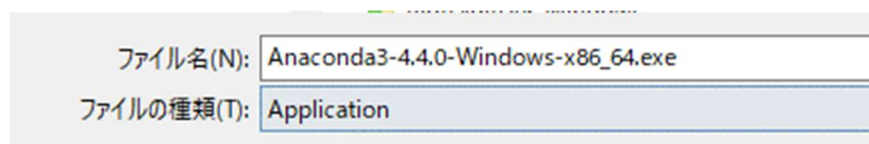
- ② Windows 64 ビット版をダウンロードしたい. 次のファイルを選ぶ.

Choose the following file to download Windows 64-bit version.

- ③ ファイルのダウンロードが始まる.

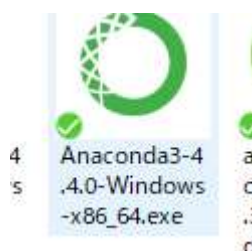
インストールに使うファイル: **Anaconda3-4.4.0-Windows-x86_64.exe**

File download starts. The file to be used to install: ANaconda3-4.4.0-Windows-x86_64.exe



- ④ ダウンロードした .exe ファイルを**実行**.

Launch the .exe file.



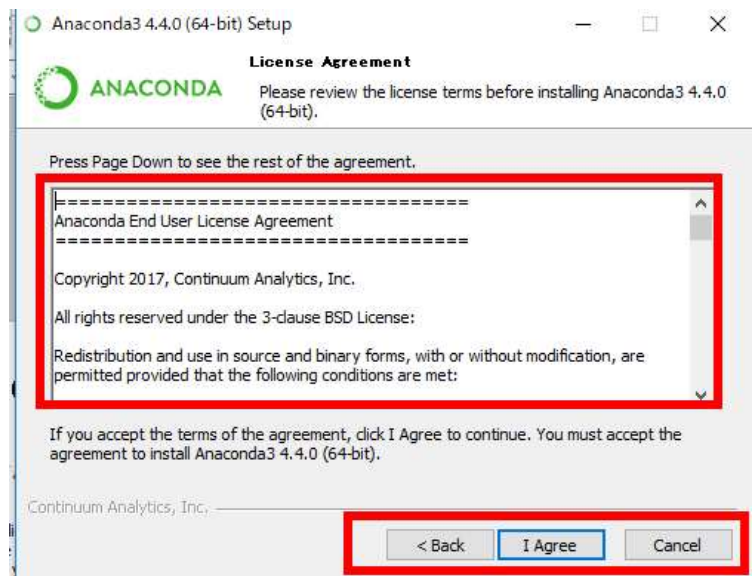
- ⑤ ようこそ画面では、「**Next**」をクリック.

Click "next" in the welcome window.



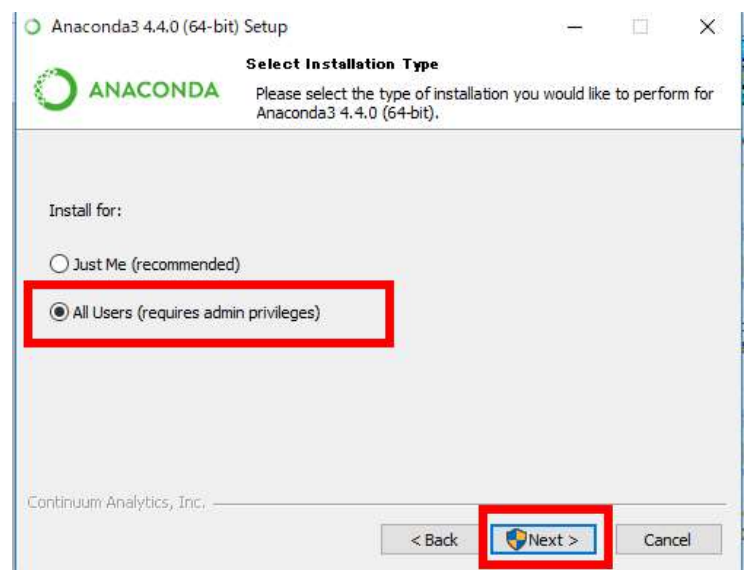
⑥ ライセンス条項の画面

License agreement.

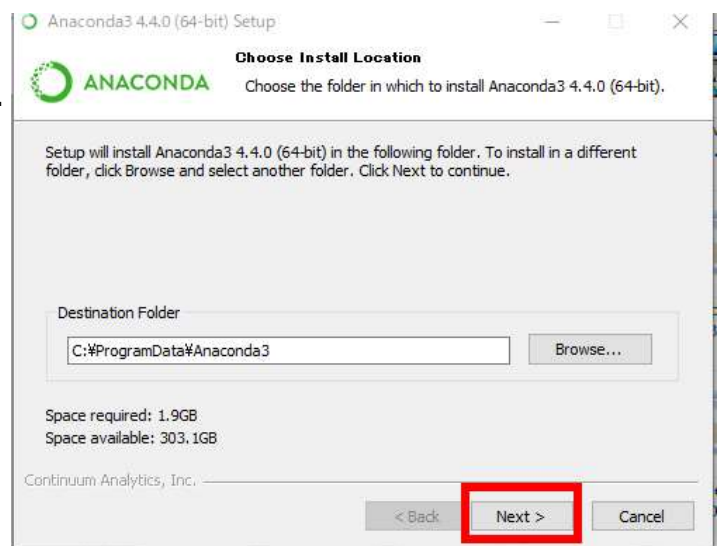


⑦ インストールタイプは「**All Users**」を選び、「**Next**」をクリック。

Choose "All Users" for install type, and click "Next".

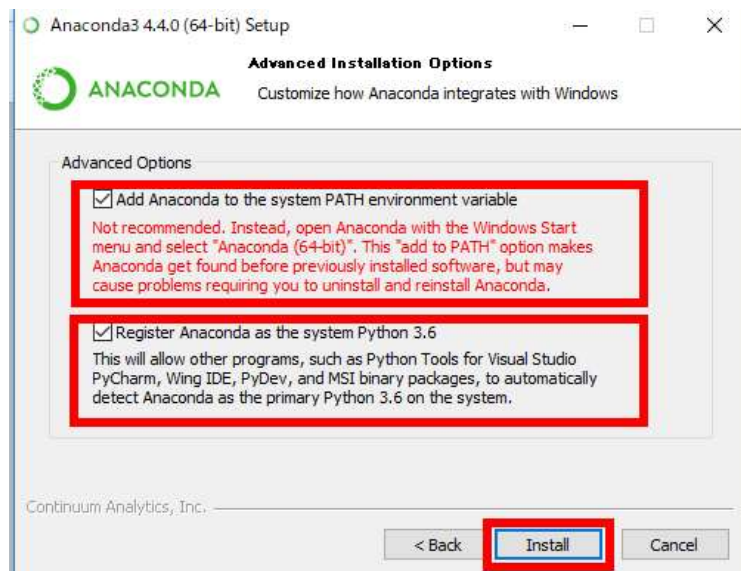


⑧ インストールディレクトリ（フォルダ）は**既定（デフォルト）**のままでよい。「**Next**」をクリック。
Click "Next"

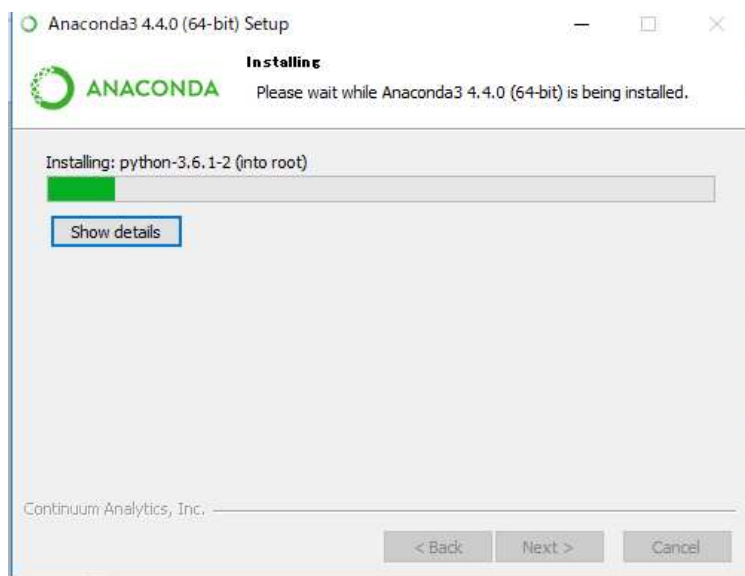


- ⑨ オプションは、**2つともチェックする**。
インストールを開始したいので、「**Install**」をクリック。

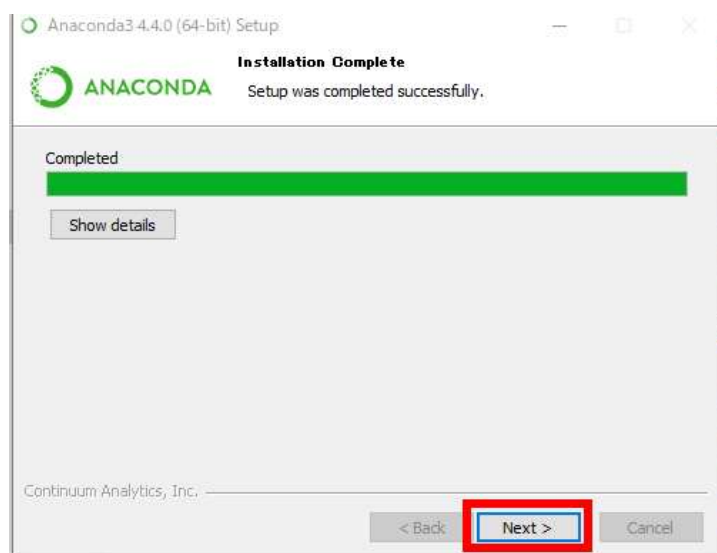
Check the both and click "Install"



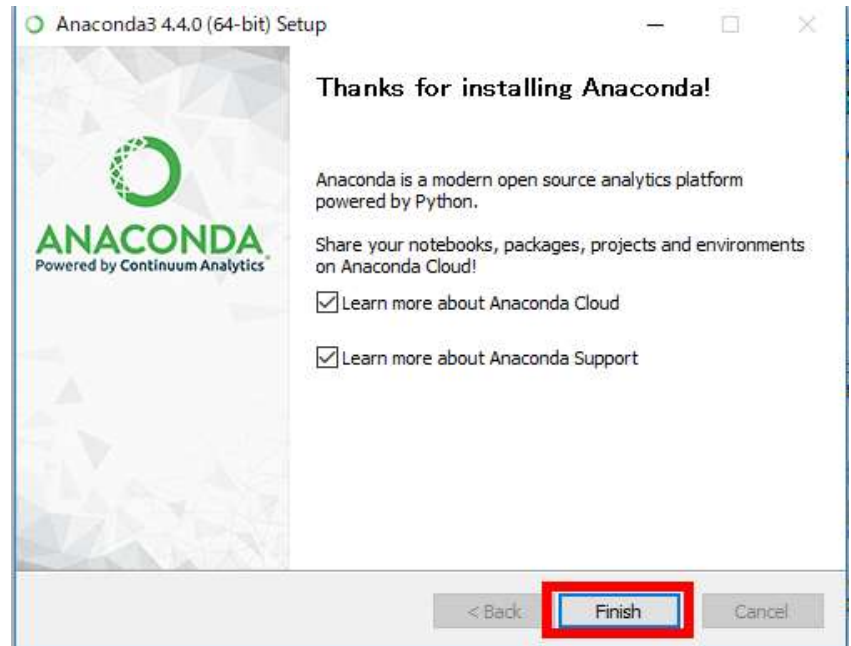
- ⑩ インストールが始まる。
Installation starts.



- ⑪ インストール完了の表示。「**Next**」をクリック。
Click "Next"



⑫ インストール完了の確認
Installation completed.



⑬ Python パッケージの更新と conda-build パッケージのインストール

まず, Anaconda プロンプトを管理者として開くために

スタートメニューの「**Anaconda3 (64bit)**」の下「**Anaconda Prompt**」を右クリックし, メニューで「**管理者として実行**」を選ぶ.

Update python packages and install 'conda-build' package

Right click the "Anaconda Prompt" under "Anaconda3 (64bit)" in a start menu, and select "Run as administrator" in the menu. And close the Anaconda prompt.

Installation of Anaconda is finished.

pip install --upgrade pip

pip install --upgrade -I setuptools

pip install conda-build

※ 途中で質問が出てきたら, y キー + Enter で続行

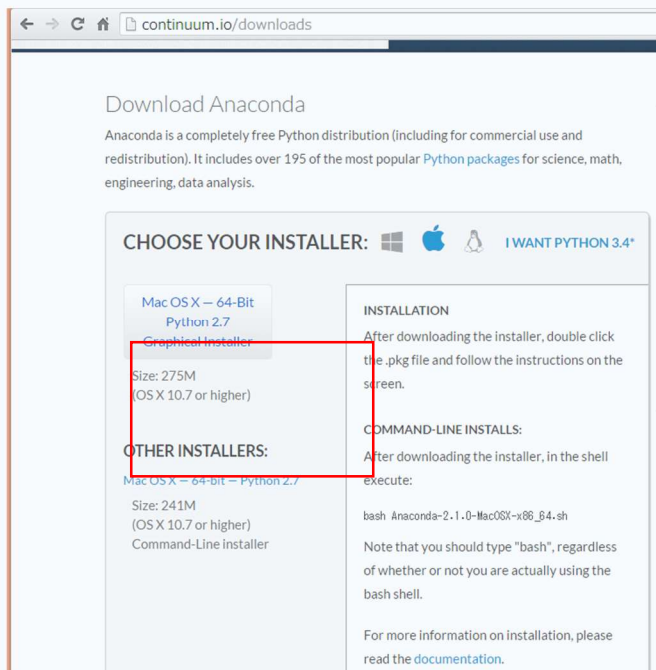
Push 'y' key and push the enter key for any questions in the procedure.

終わったら, Anaconda プロンプトを閉じる. これで Anaconda のインストールは終わり.

※ Mac OSX ユーザ (For Mac OS X users only)

上の説明は無視してください。次の2つの Web ページを参考にして、Python と Spyder をインストールしてください。For Mac OS X users. Please ignore the above steps. Please install the Python and Spyder. Please use the following three Web pages

1. <http://continuum.io/downloads> を開く



※ ダウンロードに数時間かかる場合がある

ダウンロードが終わったら「bash Anaconda-2.1.0-MacOSX-x86_64.sh」を実行

2. <https://bitbucket.org/spyder-ide/spyderlib/downloads> の DMG ファイル

※ <https://bitbucket.org/spyder-ide/spyderlib/> も参考にしてください。

3.python のバージョンの確認

spyder を起動し python のバージョンを確認

4. パッケージの追加

python のバージョンが 2 系列のときは次のコマンドを実行 (execute the following commands for python 2 series)
python のバージョンが 3 系列のときは次のコマンドを実行 (execute the following commands for python 3 series)

```
sudo conda update conda
sudo conda update anaconda
sudo conda install numpy
sudo conda install scipy
sudo conda install pandas
sudo conda install matplotlib
sudo conda install scikit-learn
sudo conda install bottle
sudo conda install numexpr
sudo conda install zodb
```

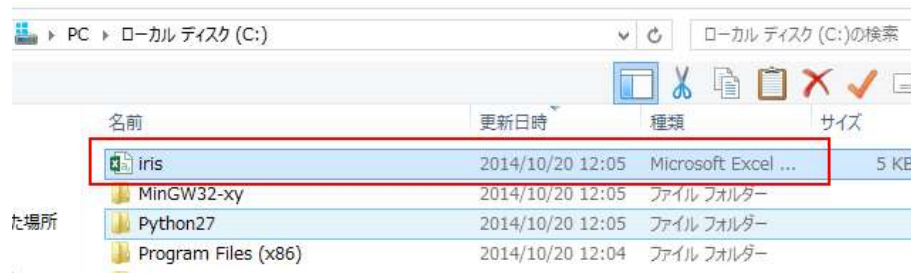
```
sudo pip update conda
sudo pip update anaconda
sudo pip install numpy
sudo pip install scipy
sudo pip install pandas
sudo pip install matplotlib
sudo pip install scikit-learn
sudo pip install bottle
sudo pip install numexpr
sudo pip install zodb
```

(2) ステップ 2 (Step 2)

以下の手順で、テストデータファイル **iris.csv** をダウンロードしなさい

(Download a test data file named 'iris.csv')

- ① Web ブラウザで <http://www.kunihikokaneko.com/free/addb/iris.csv> を開く
(Open <http://www.kunihikokaneko.com/free/addb/iris.csv> using a Web browser)
- ② ダウンロードされたファイル **iris.csv** を **C:¥iris.csv** にコピー
Copy the downloaded file **iris.csv** to **C:¥iris.csv**.



- ③ iris.csv は次のようなファイルである。
iris.csv is a file as follows.

	A	B	C	D	E	F	G
1	id	sepal_len	sepal_wid	petal_len	petal_wid	species	
2	0	5.1	3.5	1.4	0.2	setosa	
3	1	4.9	3	1.4	0.2	setosa	
4	2	4.7	3.2	1.3	0.2	setosa	
5	3	4.6	3.1	1.5	0.2	setosa	
6	4	5	3.6	1.4	0.2	setosa	
7	5	5.4	3.9	1.7	0.4	setosa	
8	6	4.6	3.4	1.4	0.3	setosa	
9	7	5	3.4	1.5	0.2	setosa	
10	8	4.4	2.9	1.4	0.2	setosa	
11	9	4.9	3.1	1.5	0.1	setosa	
12	10	5.4	3.7	1.5	0.2	setosa	
13	11	4.8	3.4	1.6	0.2	setosa	
14	12	4.8	3	1.4	0.1	setosa	
15	13	4.3	3	1.1	0.1	setosa	
16	14	5.8	4	1.2	0.2	setosa	
17	15	5.7	4.4	1.5	0.4	setosa	

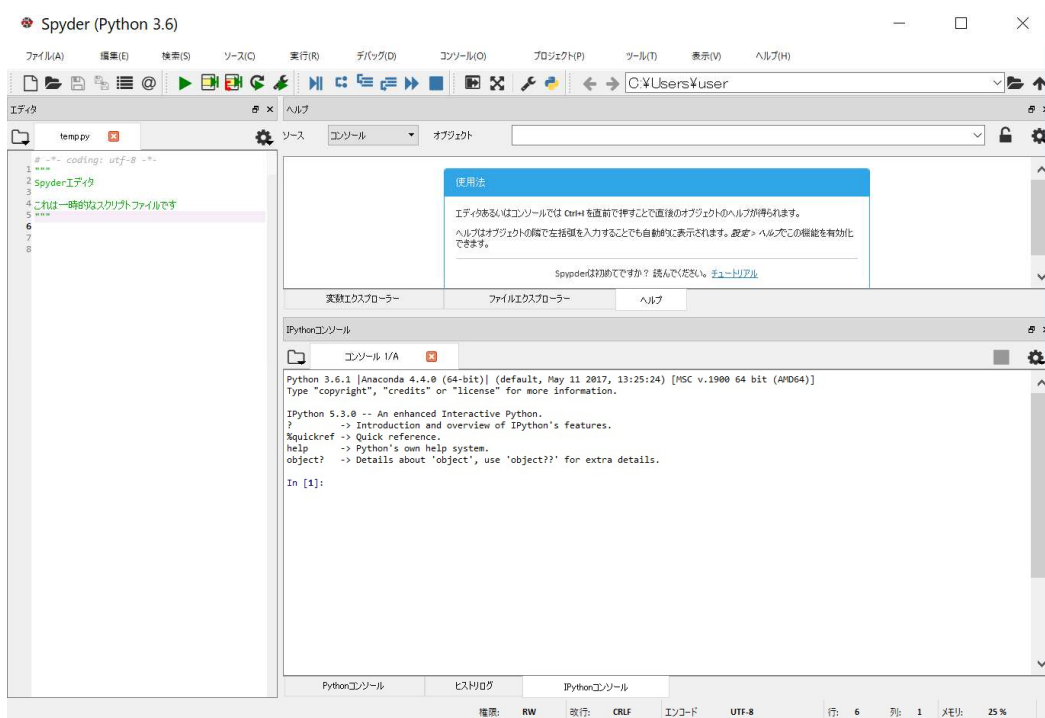
(3) ステップ 3 (Step 3)

以下の手順で、埋め込み SQL を実行しなさい。

(Execute an embedded SQL)

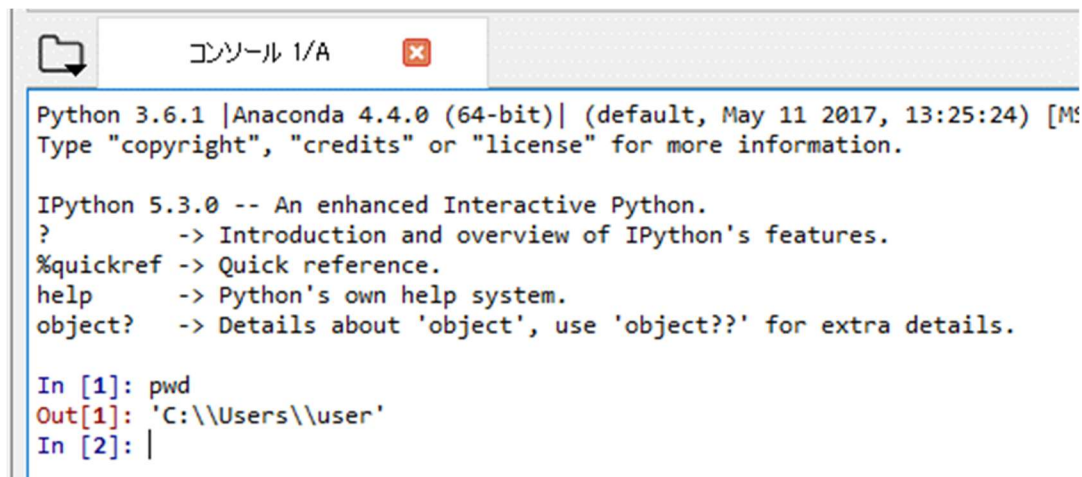
- ① スタートメニューの「**Anaconda3 (64bit)**」の下の「**spyder**」を起動する。右下に IPython コンソール・ウィンドウがある。

Launch "spyder" in the "Anaconda3 (64bit)" in a start menu. There is an IPython console windows in right lower.



- ② IPython コンソール・ウィンドウで、**pwd** コマンドを使い カレントディレクトリを確認。IPython コンソールでは Windows のコマンドも使うことができる。

Get the current directory using the pwd command. Windows commands can be executed in an IPython console window.



③ IPython コンソール・ウインドウで、次の Python プログラムを実行させてみなさい

(Try to execute a Python program using the Spyder's IPython console)

複数行にわたるユニコード文字列を使いたいところでは、「u""" . . . """」のように書く。

```
import pandas as pd
import sqlite3
c = sqlite3.connect('hoge.sqlite')
sql = u"""
create table iris (
    id integer          primary key,
    sepal_length real,
    sepal_width real,
    petal_length real,
    petal_width real,
    species integer );
"""
c.execute(sql)

x = pd.read_csv('c:¥¥iris.csv', header=0)
for index, r in x.iterrows():
    sql = u"insert into iris values (?, ?, ?, ?, ?, ?)"
    c.execute(sql, (r[0], r[1], r[2], r[3], r[4], r[5]))

c.commit()

cur = c.cursor()
cur.execute(u"select * from iris")
for t in cur:
    print (t)

c.close()
```

実行結果の例

```
(0, 5.1, 3.5, 1.4, 0.2, 'setosa')
(1, 4.9, 3.0, 1.4, 0.2, 'setosa')
(2, 4.7, 3.2, 1.3, 0.2, 'setosa')
(3, 4.6, 3.1, 1.5, 0.2, 'setosa')
(4, 5.0, 3.6, 1.4, 0.2, 'setosa')
(5, 5.4, 3.9, 1.7, 0.4, 'setosa')
(6, 4.6, 3.4, 1.4, 0.3, 'setosa')
(7, 5.0, 3.4, 1.5, 0.2, 'setosa')
(8, 4.4, 2.9, 1.4, 0.2, 'setosa')
(9, 4.9, 3.1, 1.5, 0.1, 'setosa')
(10, 5.4, 3.7, 1.5, 0.2, 'setosa')
(11, 4.8, 3.4, 1.6, 0.2, 'setosa')
(12, 4.8, 3.0, 1.4, 0.1, 'setosa')
(13, 4.3, 3.0, 1.1, 0.1, 'setosa')
(14, 5.8, 4.0, 1.2, 0.2, 'setosa')
(15, 5.7, 4.4, 1.5, 0.4, 'setosa')
```

④ さきほど `pwd` で調べたカレントディレクトリに, データベースファイル `hoge.sqlite` ができているので確認する.

カル ディスク (C:) > ユーザー > user

名前	更新日時	種類
hoge.sqlite	2017/06/08 23:55	SQLITE ファ
NTUSER.DAT	2017/06/08 16:46	DAT ファイル