

Homebrewによるソフトウェアの実装(3)

-- Mac OS 10.15 (Catalina)用 --

- ① Anaconda (Spyder) (Python開発環境プラットフォーム)
<https://www.python.jp/install/anaconda/>
- ② pandas (Python用データ処理関数ライブラリ)
<https://pandas.pydata.org/>
- ③ Matplotlib (Python用データ描画関数ライブラリ)
<https://matplotlib.org/>
- ④ IPAexフォント (オープンソースライセンス日本語フォント)
<https://ipafont.ipa.go.jp/>

溝 口 佳 寛

九州大学マス・フォア・インダストリ研究所

2020年5月23日 (Ver.1)

① Anaconda (Spyder) (Python開発環境プラットフォーム)

① ターミナルウィンドウでの作業:

作業を始める前に, Homebrewにアップデートがないか確認し, 必要があれば, アップデートします.

```
brew update -force  
brew upgrade
```

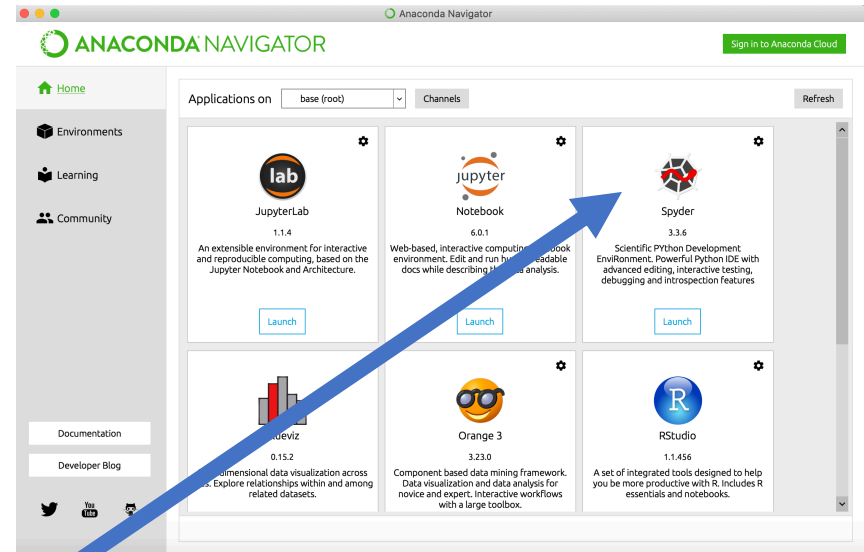
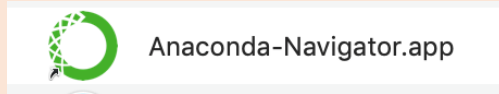
Anacondaのインストールは, まず,

```
brew install pyenv
```

で行います. 次に,

```
pyenv install anaconda3-2019.10
```

で, Anacondaをインストールします.
アプリケーションフォルダに下のアイコンをダブルクリックして起動します.



③ Spyderの起動

Anacondaに入っている開発ツールのなかで, 「Spyder」をクリックして起動します.

注. ターミナル中で以下のスクリプトを実行しても「Spyder」を起動可能です.

```
~/.anaconda/navigator/scripts/spyder.sh
```

② pyenvのインストール後, zshまたは, bashの環境設定のために以下のコマンドを実行します.

OS10.15 (Catalina) の場合, ターミナルはzshで動いていますので, 次のようにします.

```
echo 'eval "$$(pyenv init -)"' >> ~/.zshrc
```

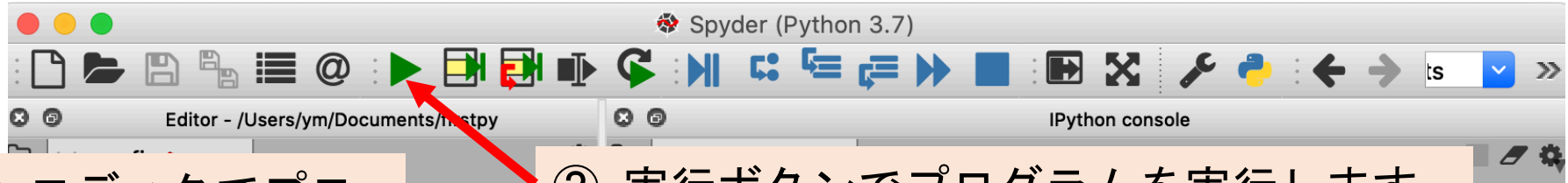
OS10.14 (Majave) 以前の場合, ターミナルはbashで動いていますので, 次のようにします.

```
echo 'eval "$$(pyenv init -)"' >> ~/.bash_profile
```

新しいターミナルを開いて, 新しいpythonコマンドの場所を `which python` で確認します.

`/Users/ym/.pyenv/shims/python` (注意. `/User/ym` の部分は人により異なります.)

①-2. Anaconda (Spyder) の使い方



① エディタでプログラムを書きます。

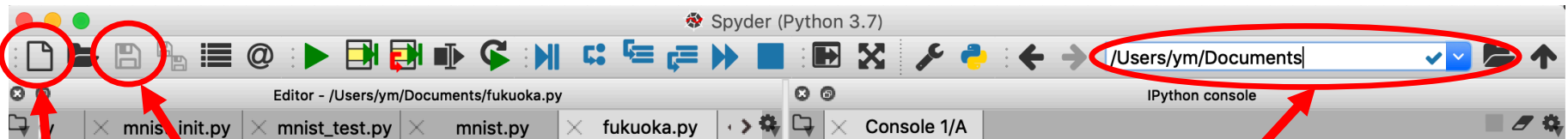
② 実行ボタンでプログラムを実行します。

③ コンソールに実行結果が表示されます。

```
Created on Saturday 23 May 2020 11:14  
@author: ym  
print ("1+1=")  
print (1+1)
```

```
Type "copyright", "credits" or "license" for more information.  
IPython 7.8.0 -- An enhanced interactive Python  
In [1]: runfile('/Users/ym/Documents/fukuoka.py')  
1+1=  
2  
In [2]:
```

メニューバーの **Consoles** メニューから新しいコンソールウィンドウを開けます。コンソールウィンドウに直接Pythonコマンドを入力して実行することも出来ます。各ウィンドウの左上の **×ボタン** でウィンドウを閉じることが出来ます。



プログラムをファイルに保存します。本スライドの説明では、Documents(書類)フォルダに保存することを仮定してします。

新しいエディタ画面を開く

プログラムを保存するフォルダは、本スライドでは、Documents(書類)フォルダを指定して説明します。

② pandas (Python用データ処理関数ライブラリ)

Pandasは、表データを処理したり、入出力する関数群を備えたライブラリです。プログラムの先頭に `import pandas as pd` と入れて利用します。pdの部分は自分で決めることが出来、ライブラリ pandas 内の関数を pd.[関数名]として呼ぶことが出来ます。表データ処理とは、表計算ソフトExcelで行う作業をイメージすると良いでしょう。また、pandasはExcel等で作ったデータを CSV(コンマ区切り) ファイルをPythonの配列に読み込む関数も備えています。

③ Matplotlib (Python用データ描画関数ライブラリ)

Matplotlibは、各種数値データや関数を描画するための関数群を備えたライブラリです。プログラムの先頭に `import matplotlib` と入れて利用します。matplotlibは単体でも様々な機能を持っていますが、今回のスライドでは、pandasライブラリのplot関数の中で呼ばれる関数描画機能のみ紹介します。

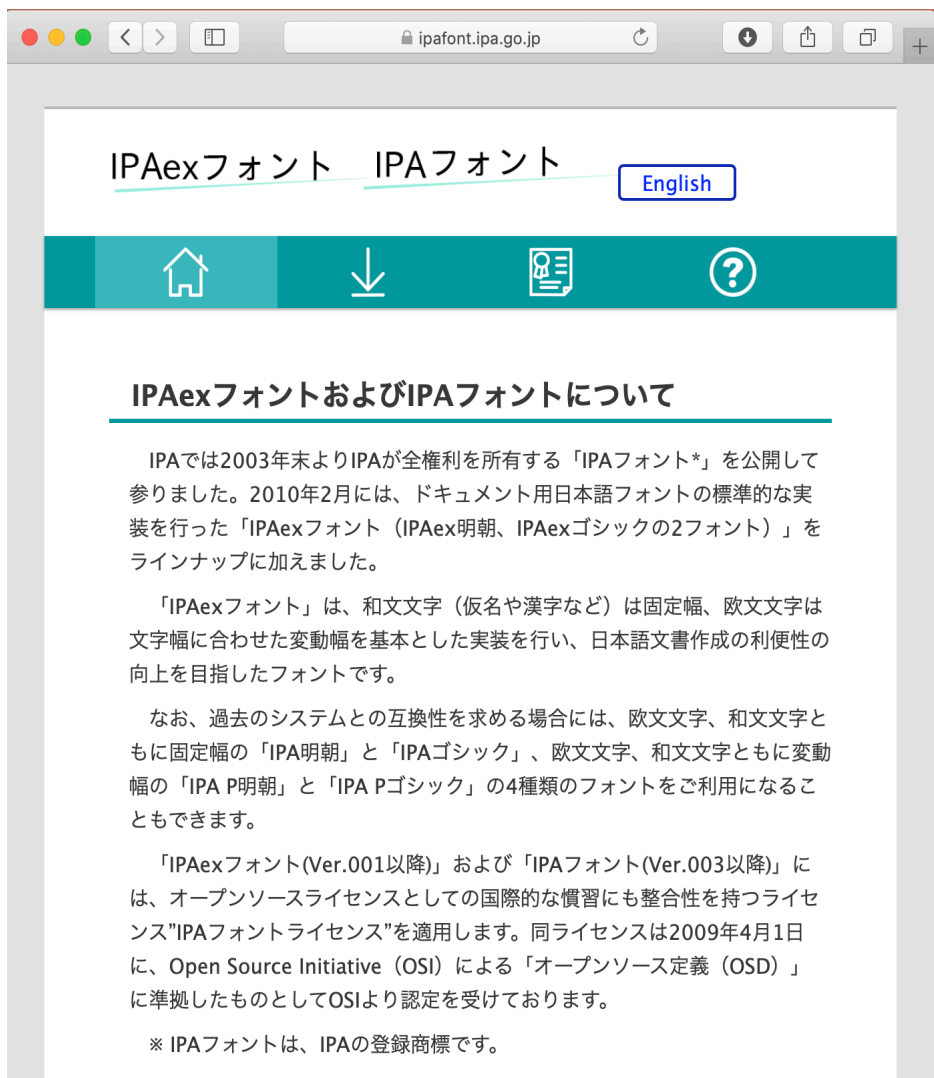
※ Matplotlibは日本語フォント処理が不十分なので、そのまま使うと日本語の入ったデータ描画で文字化けします。そのために、本スライドでは、独自に日本語フォントを入れてMatplotlibで使う方法を紹介します。

補足: pandas も Matplotlib もAnacondaのインストールで実装されていますので、Matplotlibの日本語描画以外の機能には、追加のインストール作業はありません。

④ IPAexフォント(オープンソースライセンス日本語フォント)

IPA (独) 情報処理推進機構

<https://ipafont.ipa.go.jp/node17#jp>

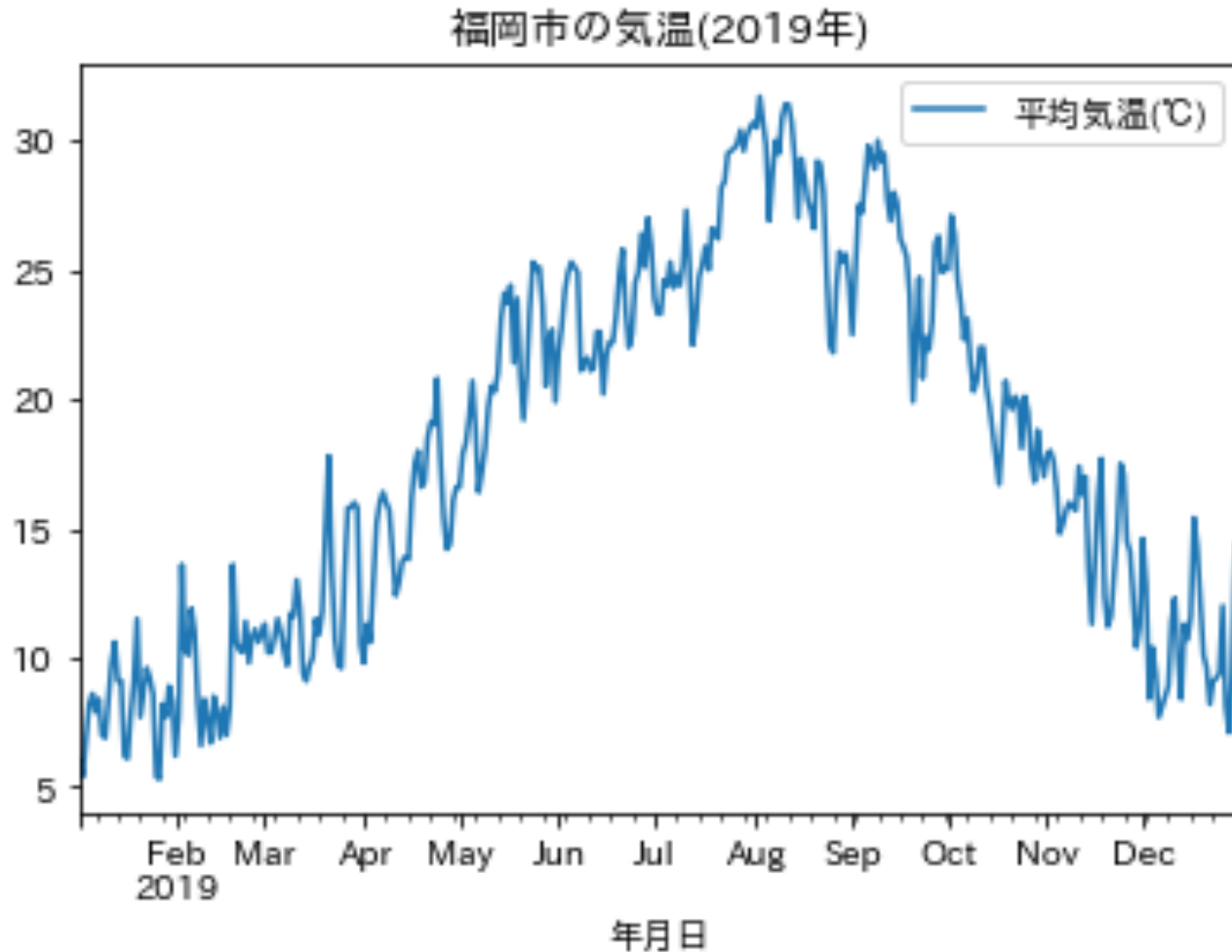


IPAexフォントのダウンロードページより、(IPAexゴシック (Ver. 004. 01)) フォントの zipファイル [ipaexg00401.zip](#) をダウンロードして展開します。その中で必要なファイルは、ipaexg.ttf だけです。

- (1) ipaexg.ttf を ~/Library/Fonts/ フォルダへコピーします。~/ は自分のホームページです。
例えば、~/Downloads/ipaexg00401 フォルダに ipaexg.ttf 展開されている場合、ターミナルで、
cd ~/Downloads/ipaexg00401
cp ipaexg.ttf ~/Library/Fonts
とすればコピー出来ます。
- (2) もし、~/matplotlib フォルダが存在している場合、一度フォルダを消去してcacheを全て消しておきます。
rm -ri ~/matplotlib
次にMatplotlibを実行したときに改めて作成されます。

⑤ pandasとMatplotlibを使ったグラフ描画例

ホームページから、福岡市の1年間の気温データをCSVファイルで入手し、pandasライブラリでPythonに読み込み、必要部分のデータ抽出を行います。そして、Matplotlib を使って折れ線グラフで描画します。



⑤-2. 福岡市の2019年1年間の気温データの入手.

「気象庁ホームページ」 <http://www.jma.go.jp/jma/index.html> の
「過去の地点気象データ・ダウンロードページ」
<http://www.data.jma.go.jp/gmd/risk/obsdl/> から、下記の手順で、
福岡市の2019年の1年間の毎日の平均気温のデータを入手します。

The screenshot shows the JMA data download interface with three main sections:

- 地点を選ぶ (Select Location):** A map of Japan with a red dot indicating Fukuoka City. The text "他の都道府県を選ぶ" (Select other prefectures) is visible.
- 項目を選ぶ (Select Item):** A list of data types under "データの種類の ? 詳細" (Data Type Details). The "日別値" (Daily Value) option is selected. A red box highlights the "最初に選択してください" (Please select first) instruction. Below, the "項目" (Item) list shows "日平均気温" (Daily Average Temperature) selected.
- 期間を選ぶ (Select Period):** A section for selecting the time period. The "連続した期間で表示する" (Display for continuous period) option is selected. The period is set to "2019年1月1日から2019年12月31日まで" (From January 1, 2019, to December 31, 2019).

① 地点を選ぶ.
「福岡県」「福岡市」
をクリック

② 項目を選ぶ.
データの種類は「日別値」
気温は「日平均気温」だけ
を選択します。

③ 期間を選ぶ.
ここでは、2019年1月1日から、
2019年12月31日までを指定します。

④ ダウンロードボタン [CSVファイルをダウンロード ▶](#) を押してCSVファイルをダウンロードします。
ファイル名 **data.csv** で、ダウンロードフォルダに保存されます。
Pythonプログラムで利用するため、**「ドキュメント(書類)」フォルダにコピーします。**

⑤-3. 福岡市の2019年1年間の気温データの処理.

```
import pandas as pd
import matplotlib
font = {'family': 'IPAexGothic'}
matplotlib.rc('font', **font)
df = pd.read_csv("data.csv",
                 skiprows=[0, 1, 2, 4, 5],
                 encoding='shift_jis',
                 index_col=0 )
df = df.drop(df.columns[[1, 2]], axis=1)
df.index = pd.to_datetime(df.index)
df.plot(title='福岡市の気温(2019年)')
```

① 日本語表示のため

② ファイル
からのデータ
入力と不要な
行や列の削除

③ 日付書式変更

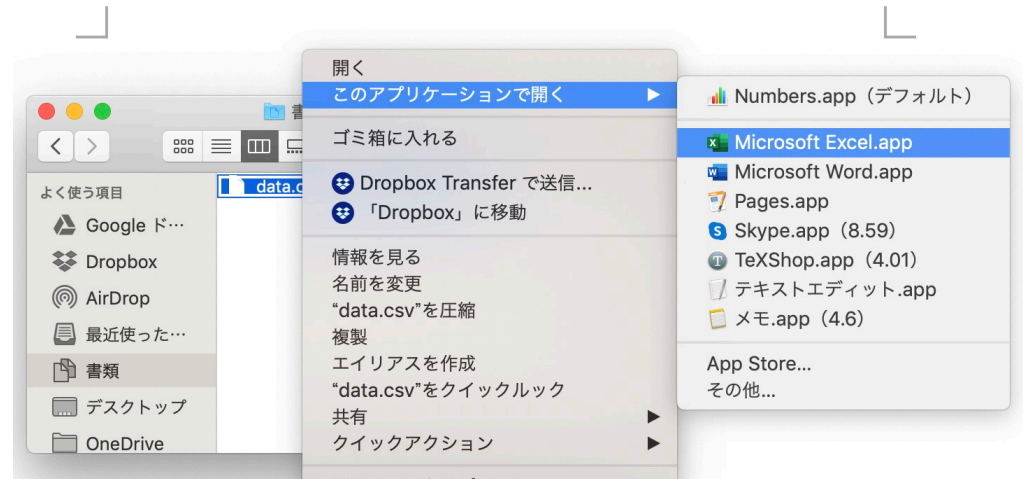
④ グラフ描画

② 補足: 取得CSVファイルは, 4行目に「年月日」と「平均気温」の列の説明がありますが, 1, 2, 3, 5, 6行目の情報はグラフには不要なのでskipして入力します。(注意. Pythonはエクトルの要素を1から数えず, 0から数えます.)

③ 補足: 日付データを文字列ではなく日付として認識させると, Matplotlibでの描画の際に自動的に表示項目を工夫してくれます.

⑤-4. CSVデータをExcelで確認してみる!?

① ダウンロードした data.csv をExcelを使って開いてみます。data.csvのあるフォルダをファインダーで開いて右クリックで開くアプリケーションとして、Excelを指定することで、開くことができます。



	A	B	C	D	E
1	ダウンロードした時刻：2020/05/23 08:43:26				
2					
3		福岡	福岡	福岡	
4	年月日	平均気温(℃)	平均気温(℃)	平均気温(℃)	
5			品質情報	均質番号	
6	2019/1/1	6.8	8	1	
7	2019/1/2	6	8	1	
8	2019/1/3	5.4	8	1	
9	2019/1/4	6.9	8	1	
10	2019/1/5	8.2	8	1	

② 左図を見てわかるように、1, 2, 3, 5行目のデータはグラフ描画には不要です。また、3(C), 4(D)列の情報も不要です。Excelではマウスクリックして行や列を簡単に削除することが出来ます。また、Excelでも折れ線グラフ表示は可能です。ある意味では便利ですが、同じ作業を他の都市データ等、何度も行わないといけない場合には、pandasを使ったプログラムを書いておいた方が便利です。