

Jupyter Notebook 入門

はんなり Python #3

2018/02/16



自己紹介

森崎雅之

@masayuki14



Follow me !!

自己紹介

- コミュニティ
 - はんなりPython
 - OSS Gate
- 主夫
 - 兼 パートタイムプログラマー
 - 兼 スーパーキーズアンバサダー

スプーキーズ@京都

ソーシャルゲーム × ボードゲーム



スプーキーズ@京都

社内勉強会やってます

- WebRTCを活用する
- Docker勉強会
- ゲームAIを作って競い合う構想

スプーキーズ@京都

エンジニア勉強会 #1 3/23(金)

connpass 79323

スプーキーズ@京都

Webエンジニア募集中

今日する話

Jupyter Notebook 入門

今日する話

- 1. Jupyter Notebook とは**
- 2. Use Jupyter Notebook**
- 3. Use Pandas**
4. 身近なデータを可視化

Jupyter Notebook とは

Jupyter Notebook とは

Project Jupyter

複数のプログラミング言語にまたがるインタラクティブ
ブコンピューティングのためのサービスを開発する

Jupyter Notebook とは

Project Jupyter のサービスの1つ

ライブコード、方程式、可視化、テキストを含むドキュメントを作成して共有できるオープンソースの
Webアプリケーション

Use Jupyter Notebook



Use Jupyter Notebook by Docker

```
FROM python:latest

# Install miniconda to /miniconda
RUN curl -LO 'https://repo.continuum.io/miniconda/Miniconda3-late
RUN bash Miniconda3-latest-Linux-x86_64.sh -b -p /miniconda
ENV PATH=/miniconda/bin:${PATH}
RUN conda update -y conda
# install for jupyter notebook
RUN conda install -y pandas matplotlib nb_conda
RUN conda install -y pyyaml
RUN mkdir -p /root/notebook
WORKDIR /root/notebook
CMD jupyter notebook --ip=0.0.0.0 --allow-root
```

Use Jupyter Notebook

Build image

```
$ docker build -t jupyter .
```

Run docker

```
$ docker run -it --rm -v $(pwd)/notebook:/root/notebook -p 80:888
```

Use Jupyter Notebook

Copy/paste this URL **into** your browser when you connect **for the fi**
to login **with a token**:

`http://0.0.0.0:8888/?token=ba4fc6de0d99161f5e144ad4c1167ebf07`

起動ログに初回アクセスのURLが表示される

Use Jupyter Notebook

`http://localhost/?
token=ba4fc6de0d99161f5e144ad4c1167ebf074ddc29b916`

にアクセス！！

Use Jupyter Notebook



Use Jupyter Notebook

Hello world

```
def hello():  
    return 'Hello Jupyter.'  
  
hello()
```

```
'Hello Jupyter.'
```

Use Jupyter Notebook

グラフを表示

```
# グラフ表示を有効化
%matplotlib inline
import pandas as pd

df = pd.DataFrame([1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987])
df.plot()
```



Use Jupyter Notebook

データの特徴

```
fibo_pd.describe()
```

	0
count	20.000000
mean	52428.750000
std	128047.770566
min	1.000000
25%	28.000000
50%	768.000000
75%	20480.000000
max	524288.000000

Use Jupyter Notebook

Markdown

- Markdownでメモなどを残せる
- LaTeXで数式を書ける

Use Jupyter Notebook

Notebook URL

保存した .ipynb ファイルをGitHubにPushすると表示できる

notebook

Use Pandas

Use Pandas

- 1次元データ構造のシリーズ (Series)
- 2次元データ構造のデータフレーム (DataFrame)

Use Pandas

Pandasをつかおう

```
import pandas as pd
```

Use Pandas

データフレームの基本的な使い方

```
In [1]: import pandas as pd
...:
...: # columnsオプションで列名を指定
...: df = pd.DataFrame([1,2,3],
...:                    columns=['value'])
...: df
```

```
Out[1]:
   value
0      1
1      2
2      3
```

Use Pandas

タプルで配列を渡す

```
In [2]: df = pd.DataFrame([
...:     ('apple', 100), ('orange', 230), ('grape', 290), ('ba
...:     columns=['name', 'price']
...: ])
...: df
```

```
Out[2]:
```

	name	price
0	apple	100
1	orange	230
2	grape	290
3	banana	100

Use Pandas

ディクショナリで配列を渡す

```
In [3]: df = pd.DataFrame({  
...:     'name': ['apple', 'orange', 'pear', 'peach'],  
...:     'price': [120, 150, 230, 360],  
...:     'order': [3, 8, 4, 5]  
...: })  
...: df
```

```
Out[3]:
```

	name	order	price
0	apple	3	120
1	orange	8	150
2	pear	4	230
3	peach	5	360

Use Pandas

列の追加

```
In [4]: df['color'] = ['red', 'orange', 'green', 'pink']
...: df['total'] = df['order'] * df['price']
...: df
```

Out[4]:

```
name
order
price
color
total
0
apple
```

3

Use Pandas

インデックスの追加

```
In [5]: df.index = ['Apple', 'Orange', 'Pear', 'Peach']  
...: df
```

Out[5]:

	name	order	price	color	total
Apple	apple	3	120	red	360
Orange	orange	8	150	orange	1200
Pear	pear	4	230	green	920
Peach	peach	5	360	pink	1800

Use Pandas

列の選択

```
In [6]: df['price']
```

```
Out[6]:  
Apple      120  
Orange     150  
Pear       230  
Peach      360  
Name: price, dtype: int64
```


Use Pandas

列の選択（複数）

```
In [7]: df[['price', 'color']]
```

```
Out[7]:
```

	price	color
Apple	120	red
Orange	150	orange
Pear	230	green
Peach	360	pink

Use Pandas

head, tail

```
In [8]: df.head(2)
```

```
Out[8]:
```

	name	order	price	color	total
Apple	apple	3	120	red	360
Orange	orange	8	150	orange	1200

Use Pandas

インデックス指定

```
In [9]: df.loc[['Apple', 'Pear']]
```

```
In [9]: df[1:3]
```

```
Out[9]:
```

	name	order	price	color	total
Orange	orange	8	150	orange	1200
Pear	pear	4	230	green	920

Use Pandas

条件による指定

```
In [10]: df[df.price > 200]
```

```
Out[10]:
```

	name	order	price	color	total
Pear	pear	4	230	green	920
Peach	peach	5	360	pink	1800

Use Pandas

- 対話的、探索的にデータを操作できる

<https://git.io/vA8Uj>

身近なデータを可視化

身近なデータを可視化

Pandas でデータ操作だけじゃつ
まらない

身近なデータを可視化

そうだと可視化しよう

身近なデータを可視化 グラフを書いてみよう

身近なデータを可視化

今年は雪が多い

福井県出身なんです



身近なデータを可視化

昭和**38**年、昭和**56**年と比較する

身近なデータを可視化

気象庁の観測データを**DL**

<http://www.data.jma.go.jp/gmd/risk/obsdl/index.php>

身近なデータを可視化

最深積雪をグラフにする

身近なデータを可視化

年別の最深積雪



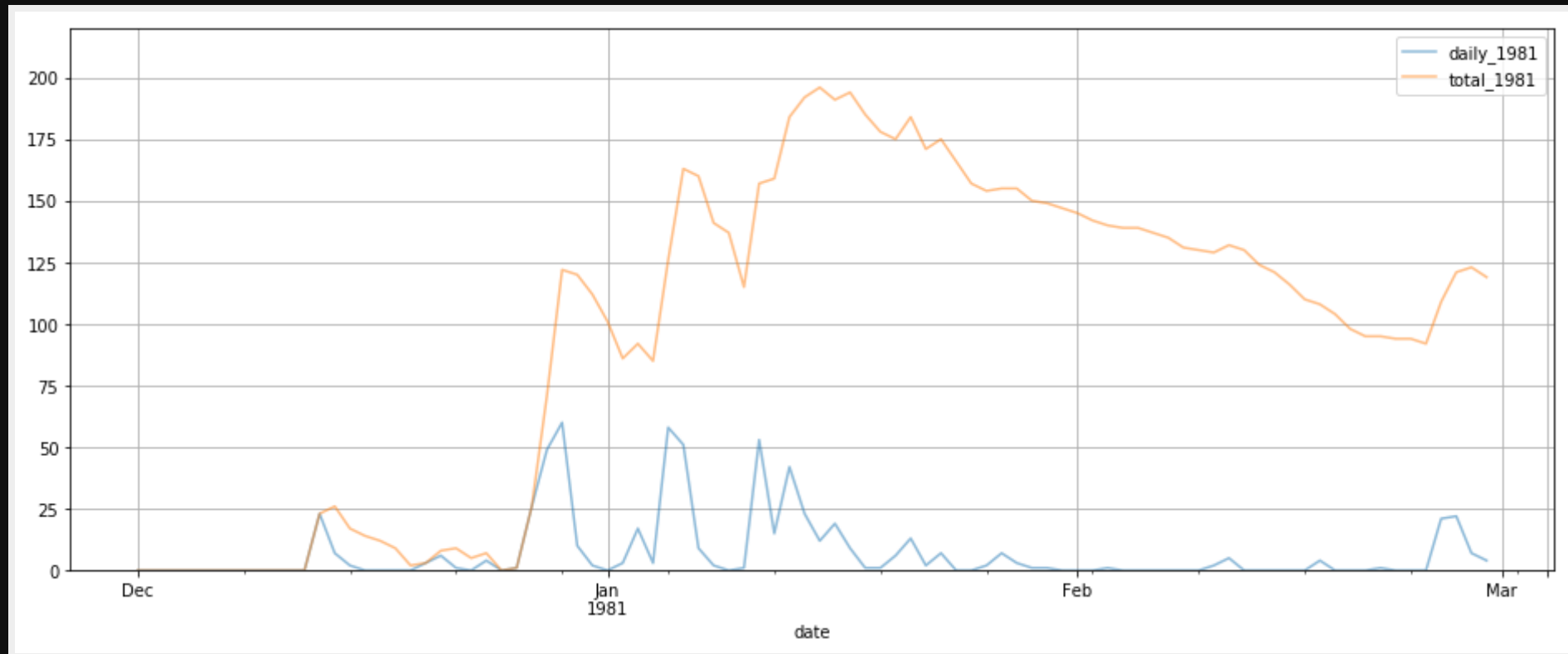
身近なデータを可視化

昭和**38**年 (1963年)



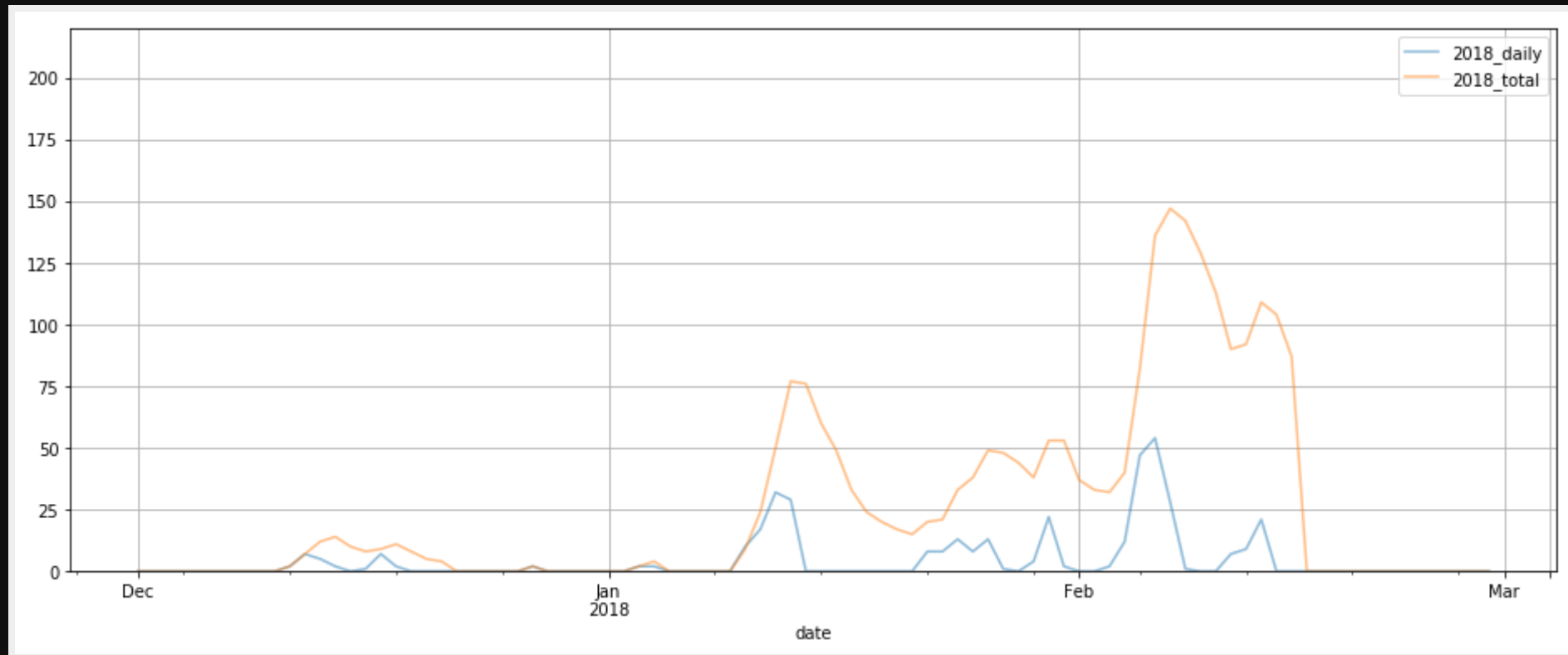
身近なデータを可視化

昭和56年 (1981年)



身近なデータを可視化

平成30年 (2018年)



まとめ

Jupyter Notebook を使うことは
データサイエンスのはじめの一步

ありがとうございました

発表資料 <https://git.io/vAl8>