

Pandas 사용하기

cvs 파일 시각화

백석대학교 강윤희

데이터 얻기

초단기실황

■ 자료설명

초단기실황자료는 전국의 읍,면,동 단위별 상세한 날씨를 매시각 제공하는 실황 관측자료입니다.

※ 2018년 7월11일 이전 제공값에 대해서는 AWS 관측값이 아닌 분석값을 제공

※ 시간은 협정세계시(UTC, 한국표준시 -9) 사용

※ 자료제공기간은 2010년 6월부터 조회일 기준 전월까지 입니다.

데이터셋

OPEN-API

■ 검색조건

기간 2019 년 04 ~ 2020 년 03

☒ 보령시
☒ 아산시
☒ 서산시
☒ 논산시
☒ 계룡시
☒ 당진시
☒ 금산군
☒ 부여군
☒ 서천군
☒ 청양군
☒ 홍성군

☒ 전체
☒ 요소
☐ 강수형태
☐ 습도
☐ 강수
☐ 하늘상태
☒ 기온
☐ 뇌전
☐ 풍향
☐ 풍속

<https://data.kma.go.kr/>

데이터 얻기

csv 파일읽기

```
In [1]: import pandas as pd
```

```
In [2]: try:
        df = pd.read_csv('400_20200411221745769.csv')
    except OSError as err:
        print("OS error: {}".format(err))
    df.head()
```

Out[2]:

	format: day	hour	value	location:63_108	Start : 20190401
0	1	0.0			8.1
1	1	100.0			8.8
2	1	200.0			8.3
3	1	300.0			9.8
4	1	400.0			9.5

데이터 얻기

자료 구조 이해하기

```
In [8]: df.columns
```

```
Out[8]: Index(['format: day', 'hour', 'value location:63_108 Start : 20190401'], dtype='object')
```

```
In [10]: df = df.rename(columns={'value location:63_108 Start : 20190401 ': 'temp'})
```

```
In [19]: df.head()
```

```
Out[19]:
```

	format: day	hour	temp
0	1	0.0	8.1
1	1	100.0	8.8
2	1	200.0	8.3
3	1	300.0	9.8
4	1	400.0	9.5

데이터 얻기

자료 요약하기

```
In [12]: df_temp = df['temp']  
df_temp.describe()
```

```
Out[12]: count      8784.000000  
mean         12.549966  
std          10.702835  
min          -50.000000  
25%           3.800000  
50%          12.900000  
75%          21.500000  
max          36.700001  
Name: temp, dtype: float64
```

```
In [13]: df_temp.shape
```

```
Out[13]: (8795,)
```

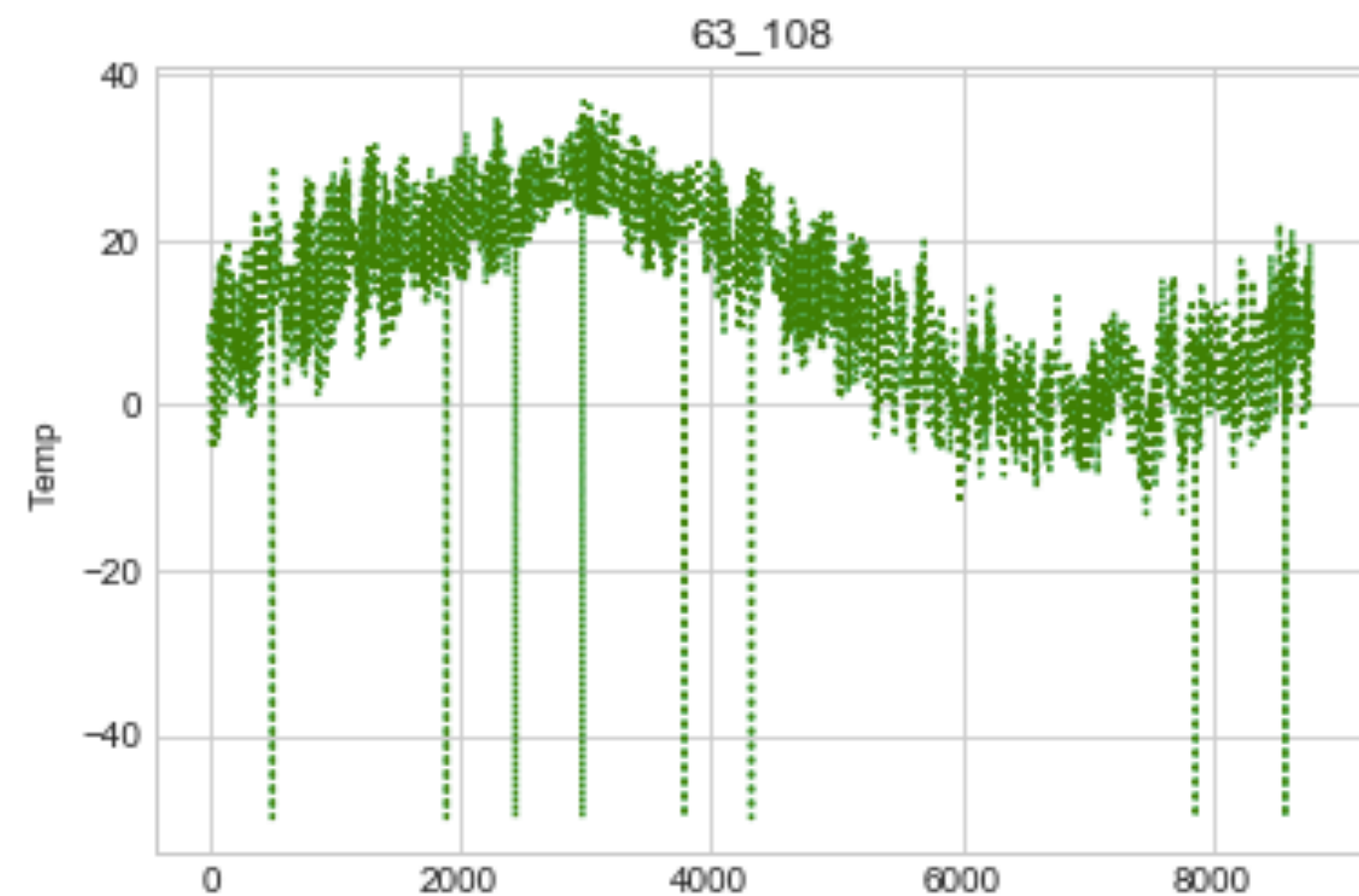
데이터 얻기

자료 가시화하기

```
In [14]: import matplotlib.pyplot as plt  
plt.style.use('seaborn-whitegrid')
```

```
In [15]: x = list(range(len(df_temp)))
```

```
In [18]: plt.plot(x, df_temp, ':b', color='g')  
plt.xlim(0, 5)  
plt.ylim(15, 50);  
plt.axis('tight');  
plt.title("63_108")  
plt.ylabel("Temp");  
  
plt.show()
```





Branch: master ▾ [DataAnalysis](#) / Pandas CSV Read.ipynb

 iceman67 Add files via upload

1 contributor

347 lines (347 sloc) | 67.6 KB

  Raw

지역별 온도데이터

<https://data.kma.go.kr/>

In [1]: `import pandas as pd`

In [2]:

```
try:
    df = pd.read_csv('400_20200411221745769.csv')
except OSError as err:
    print("OS error: {0}".format(err))
df.head()
```

Out[2]:

	format: day	hour	value location:63_108 Start : 20190401
0	1	0.0	8.1
1	1	100.0	8.8
2	1	200.0	8.3
3	1	300.0	9.8
4	1	400.0	9.5

<https://github.com/iceman67/DataAnalysis>