

## 공공 데이터 분석 (with folium)

- 전국 전통시장 표준 데이터

### 학습 목표

- folium 등의 라이브러리를 활용하여 전통시장 데이터를 살펴보고 시각해 본다.

### 학습 내용

- 1. 라이브러리 및 데이터 불러오기
- 2. 정보 살펴보기
- 3. 위치 정보를 이용한 시각화

### 목차

[01. 라이브러리 및 데이터 불러오기](#)

[02. 데이터 살펴보기](#)

[03. 위치 정보를 이용한 시각화](#)

- 데이터명 : 전국전통시장 표준데이터
- 데이터 링크 : <https://www.data.go.kr/data/15012894/standard.do>  
(<https://www.data.go.kr/data/15012894/standard.do>)
- 내용 : 전통적 방법으로 거래가 이루어지는 시장 중에 지방 자치 단체가 전통시장으로 인정한 장소 정보
- 정보 제공 : 소상공인시장진흥공단

## 01. 라이브러리 불러오기

[목차로 이동하기](#)

In [1]:

```
import pandas as pd
import folium

# 기본 encoding은 utf-8, 에러 발생시 : euc-kr로 설정
market = pd.read_csv('./data/전국전통시장표준데이터.csv', encoding='utf-8') # encoding='euc-kr'
market.head()
```

Out[1]:

	시장명	시장유형	소재지도로명주소	소재지주소	시장개설주기	위도	경도	점포수	취급품목	사용가능상품권	홈페이지주소	공중화장실보유여부	주차장보유여부	개설연도	전화번호	데이터기준일자
0	도구시장	5일장	NaN	경상북도포항시남구동해면도구리668-1	2일,7일	35.987981	129.439925	10.0	농산물생필품,수산물,잡화등	포항사랑상품권,온누리상품권	NaN	Y	Y	1915	NaN	2020-02-25
1	장기시장	5일장	NaN	경상북도포항시남구장기면임종리333-2	4일,9일	35.891559	129.491439	13.0	농산물생필품,수산물,잡화등	포항사랑상품권,온누리상품권	NaN	Y	N	1914	NaN	2020-02-25
2	신광시장	5일장	NaN	경상북도포항시북구신광면토성리486-5	5일,10일	36.128680	129.261912	22.0	농산물생필품,수산물,잡화등	포항사랑상품권,온누리상품권	NaN	Y	N	1965	NaN	2020-02-25

시 장 명	시 장 유 형	소재 지 로명 주소	소재 지 번 주소	시장 개설 주기	위도	경도	점 포 수	취 급 품 목	사 용 가 능 상 품 권	출 폐 이 지 주소	공 중 화 장 실 보 유 여 부	주 차 장 보 유 여 부	개설 연도	전화 번호	데이 티 기준 일 자	
3	기계 시장	5 일 장	NaN	경상 북도 포항 시 북 구 계면 현내 리 403	1일,6 일	36.069939	129.208468	18.0	농 산 물 생 필 품, 수 산 물, 잡 화 등	포 항 사 랑 상 품 권, 온 누 리 상 품 권	NaN	Y	N	1923	NaN	2020 02-25
4	죽장 시장	5 일 장	NaN	경상 북도 포항 시 북 구 죽 장 면 입암 리 858-34	3일,8 일	36.161583	129.095434	18.0	농 산 물 생 필 품, 수 산 물, 잡 화 등	포 항 사 랑 상 품 권, 온 누 리 상 품 권	NaN	Y	N	1919	NaN	2020 02-25

In [2]:

```
### 데이터 살펴보기
```

```
print(market.columns)
print(market.shape)
print(market.info() )
```

```
Index(['시장명', '시장유형', '소재지도로명주소', '소재지번주소', '시장개설주기', '위
도', '경도', '점포수',
      '취급품목', '사용가능상품권', '홈페이지주소', '공중화장실보유여부', '주차장보유여
부', '개설연도', '전화번호',
      '데이터기준일자', '제공기관코드', '제공기관명'],
      dtype='object')
```

```
(2882, 18)
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 2882 entries, 0 to 2881
```

```
Data columns (total 18 columns):
```

#	Column	Non-Null Count	Dtype
0	시장명	2882 non-null	object
1	시장유형	2882 non-null	object
2	소재지도로명주소	2767 non-null	object
3	소재지번주소	2692 non-null	object
4	시장개설주기	2882 non-null	object
5	위도	2775 non-null	float64
6	경도	2775 non-null	float64
7	점포수	2791 non-null	float64
8	취급품목	2641 non-null	object
9	사용가능상품권	2512 non-null	object
10	홈페이지주소	587 non-null	object
11	공중화장실보유여부	2882 non-null	object
12	주차장보유여부	2882 non-null	object
13	개설연도	2607 non-null	object
14	전화번호	1809 non-null	object
15	데이터기준일자	2882 non-null	object
16	제공기관코드	2882 non-null	object
17	제공기관명	2882 non-null	object

```
dtypes: float64(3), object(15)
```

```
memory usage: 405.4+ KB
```

```
None
```

In [3]:

```
market.isnull().sum()
```

Out[3]:

시장명	0
시장유형	0
소재지도로명주소	115
소재지지번주소	190
시장개설주기	0
위도	107
경도	107
점포수	91
취급품목	241
사용가능상품권	370
홈페이지주소	2295
공중화장실보유여부	0
주차장보유여부	0
개설연도	275
전화번호	1073
데이터기준일자	0
제공기관코드	0
제공기관명	0

dtype: int64

- 전체 2882개 데이터중에 107개 정도가 결측치

In [4]:

```
market_name = market['시장명']
market_lati = market['위도']
market_long = market['경도']
```

```
market.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2882 entries, 0 to 2881
Data columns (total 18 columns):
 #   Column                Non-Null Count  Dtype
---  -
 0   시장명                2882 non-null   object
 1   시장유형              2882 non-null   object
 2   소재지도로명주소      2767 non-null   object
 3   소재지지번주소        2692 non-null   object
 4   시장개설주기          2882 non-null   object
 5   위도                  2775 non-null   float64
 6   경도                  2775 non-null   float64
 7   점포수                2791 non-null   float64
 8   취급품목              2641 non-null   object
 9   사용가능상품권         2512 non-null   object
10   홈페이지주소          587 non-null    object
11   공중화장실보유여부    2882 non-null   object
12   주차장보유여부        2882 non-null   object
13   개설연도              2607 non-null   object
14   전화번호              1809 non-null   object
15   데이터기준일자        2882 non-null   object
16   제공기관코드          2882 non-null   object
17   제공기관명            2882 non-null   object
dtypes: float64(3), object(15)
memory usage: 405.4+ KB
```

## 02. 데이터 살펴보기

[목차로 이동하기](#)

### 결측치를 제외한 행만 가져오기

In [5]:

```
market['위도'].notnull()
```

Out[5]:

```
0      True
1      True
2      True
3      True
4      True
...
2877   True
2878   True
2879   True
2880   True
2881   True
Name: 위도, Length: 2882, dtype: bool
```

In [6]:

```
market_1 = market.loc[ market['위도'].notnull() , :]  
market_1.shape
```

Out[6]:

(2775, 18)

In [7]:

```
market_1.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
Int64Index: 2775 entries, 0 to 2881  
Data columns (total 18 columns):  
#   Column                Non-Null Count  Dtype    
---  -  
0   시장명                2775 non-null   object   
1   시장유형              2775 non-null   object   
2   소재지도로명주소     2668 non-null   object   
3   소재지지번주소       2605 non-null   object   
4   시장개설주기          2775 non-null   object   
5   위도                  2775 non-null   float64  
6   경도                  2775 non-null   float64  
7   점포수                2684 non-null   float64  
8   취급품목              2534 non-null   object   
9   사용가능상품권        2408 non-null   object   
10  홈페이지주소          561 non-null    object   
11  공중화장실보유여부   2775 non-null   object   
12  주차장보유여부       2775 non-null   object   
13  개설연도              2516 non-null   object   
14  전화번호              1752 non-null   object   
15  데이터기준일자       2775 non-null   object   
16  제공기관코드          2775 non-null   object   
17  제공기관명            2775 non-null   object   
dtypes: float64(3), object(15)  
memory usage: 411.9+ KB
```

In [8]:

```
market_1.columns
```

Out[8]:

```
Index(['시장명', '시장유형', '소재지도로명주소', '소재지지번주소', '시장개설주기', '위  
도', '경도', '점포수',  
      '취급품목', '사용가능상품권', '홈페이지주소', '공중화장실보유여부', '주차장보유여  
부', '개설연도', '전화번호',  
      '데이터기준일자', '제공기관코드', '제공기관명'],  
      dtype='object')
```

In [9]:

```
### 필요한 정보(특징)만 가져오기
sel = [ '시장명' , '시장유형' , '시장개설주기' , '위도' , '경도' ,
        '사용가능상품권' , '공중화장실보유여부' , '주차장보유여부' , '개설연도' ,
        '데이터기준일자' , '제공기관명' ]

market_2 = market_1[sel]
market_2
```

Out[9]:

	시장명	시장 유형	시장개 설주기	위도	경도	사용가능상품 권	공중화장 실보유여 부	주차 장보 유여 부	개설 연도	데이터 기준일 자	제 공 기 관 명
0	도구시 장	5일 장	2일,7 일	35.987981	129.439925	포항사랑상품 권, 온누리상 품권	Y	Y	1915	2020- 02-25	경 상 북 도
1	장기시 장	5일 장	4일,9 일	35.891559	129.491439	포항사랑상품 권, 온누리상 품권	Y	N	1914	2020- 02-25	경 상 북 도
2	신광시 장	5일 장	5 일,10 일	36.128680	129.261912	포항사랑상품 권, 온누리상 품권	Y	N	1965	2020- 02-25	경 상 북 도
3	기계시 장	5일 장	1일,6 일	36.069939	129.208468	포항사랑상품 권, 온누리상 품권	Y	N	1923	2020- 02-25	경 상 북 도
4	죽장시 장	5일 장	3일,8 일	36.161583	129.095434	포항사랑상품 권, 온누리상 품권	Y	N	1919	2020- 02-25	경 상 북 도
...	...	...	...	...	...	...	...	...	...	...	...
2877	연지상 점가	상설 장	상점가	35.575348	126.844651	NaN	Y	Y	2000	2019- 06-28	전 라 북 도
2878	새암길 상점가	상설 장	상점가	35.567002	126.853082	NaN	Y	Y	2005	2019- 06-28	전 라 북 도
2879	중앙로 상점가	상설 장	상점가	35.566605	126.853091	NaN	Y	Y	2010	2019- 06-28	전 라 북 도
2880	서부시 장상점 가	상설 장	상점가	35.804433	127.126777	NaN	Y	Y	2008	2019- 06-28	전 라 북 도



	시장명	시장 유형	시장개 설주기	위도	경도	사용가능상품 권	공중화장 실보유여 부	주차 장보 유여 부	개설 연도	데이터 기준일 자	제 공 기 관 명
2881	동문상 점가	상설 장	상점가	35.819008	127.151082	온누리상품권, 전자상품권	Y	Y	2007	2019- 06-28	전 라 북 도

2775 rows × 11 columns

### 03. 위치 정보를 이용한 시각화

[목차로 이동하기](#)

In [10]:

```
market_name = market_2[ '시장명' ]
market_lati = market_2[ '위도' ]
market_long = market_2[ '경도' ]
```

In [11]:

```
import numpy as np

market_lati = list(market_lati)
market_long = list(market_long)
market_loc = np.array([market_lati, market_long]).T
print( market_loc.shape )
print( np.mean( market_lati ) , np.mean(market_long) )
market_loc
```

```
(2775, 2)
36.20336486381247 127.88862370937265
```

Out[11]:

```
array([[ 35.98798126, 129.4399253 ],
       [ 35.891559 , 129.491439 ],
       [ 36.12868046, 129.2619125 ],
       ...,
       [ 35.5666046 , 126.85309052],
       [ 35.804433 , 127.12677653],
       [ 35.81900809, 127.151082 ]])
```

In [12]:

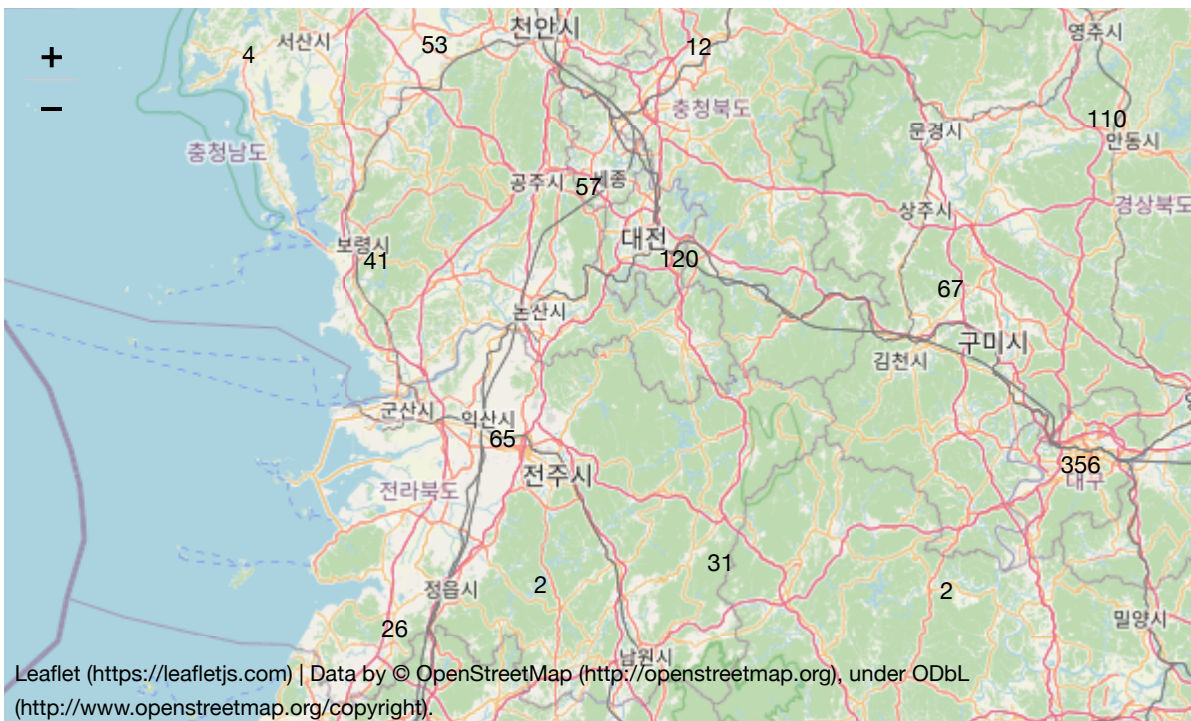
```
from folium import plugins
import os

market_map = folium.Map(location=[ np.mean( market_lati), np.mean(market_long) ],
                           zoom_start=8)

market_name = list(market_name)
plugins.MarkerCluster(market_loc, popups=market_name).add_to(market_map)

market_map.save(os.path.join('.', 'Market_location.html'))
market_map
```

Out[12]:



In [ ]: