

In [1]:

```
 -*-coding:utf-8 -*-

import pathlib
import random
from functools import reduce
from collections import defaultdict

import pandas as pd
import geopandas as gpd
import folium
import shapely
import numpy as np
from IPython.display import display
import matplotlib.pyplot as plt
from tqdm.notebook import tqdm
import xgboost
import sklearn.cluster
import tensorflow as tf
import matplotlib as mpl

from pandas import DataFrame
from geoband import API
```

In [2]:

```
input_path = pathlib.Path('./input')
if not input_path.is_dir():
    input_path.mkdir()

from geoband.API import *
GetCompasData('SBJ_2102_001', '1', '1.세종시 전유부.csv')
GetCompasData('SBJ_2102_001', '2', '2.세종시 표제부.csv')
GetCompasData('SBJ_2102_001', '3', '3.세종시 아파트(매매)_실거래가.csv')
GetCompasData('SBJ_2102_001', '4', '4.세종시 연립다세대(매매)_실거래가.csv')
GetCompasData('SBJ_2102_001', '5', '5.세종시 단독다가구(매매)_실거래가.csv')
GetCompasData('SBJ_2102_001', '6', '6.세종시 오피스텔(매매)_실거래가.csv')
GetCompasData('SBJ_2102_001', '7', '7.세종시 아파트(전월세)_실거래가.csv')
GetCompasData('SBJ_2102_001', '8', '8.세종시 연립다세대(전월세)_실거래가.csv')
GetCompasData('SBJ_2102_001', '9', '9.세종시 단독다가구(전월세)_실거래가.csv')
GetCompasData('SBJ_2102_001', '10', '10.세종시 오피스텔(전월세)_실거래가.csv')
GetCompasData('SBJ_2102_001', '11', '11.세종시 상업업무용(매매)_실거래가.csv')
GetCompasData('SBJ_2102_001', '12', '12.세종시 토지(매매)_실거래가.csv')
GetCompasData('SBJ_2102_001', '13', '13.세종시 분양권_실거래가.csv')
GetCompasData('SBJ_2102_001', '14', '14.세종시 상권정보.csv')
GetCompasData('SBJ_2102_001', '15', '15.세종시 상권정보_업종코드.csv')
GetCompasData('SBJ_2102_001', '16', '16.세종시 사업체_매출정보.geojson')
GetCompasData('SBJ_2102_001', '17', '17.세종시 사업체_매출정보.csv')
GetCompasData('SBJ_2102_001', '18', '18.세종시 개별공시지가(2017~2020).csv')
GetCompasData('SBJ_2102_001', '19', '19.세종시 연령별 거주인구정보_격자.geojson')
GetCompasData('SBJ_2102_001', '20', '20.세종시 전입자수.csv')
GetCompasData('SBJ_2102_001', '21', '21.세종시 전출자수.csv')
GetCompasData('SBJ_2102_001', '22', '22.세종시 연령별 인구현황.csv')
GetCompasData('SBJ_2102_001', '23', '23.세종시 도로명주소 건물.geojson')
GetCompasData('SBJ_2102_001', '24', '24.세종시 건축물연면적_격자.geojson')
GetCompasData('SBJ_2102_001', '25', '25.세종시 가구 월평균 소비지출액.csv')
GetCompasData('SBJ_2102_001', '26', '26.세종시 거주기간.csv')
GetCompasData('SBJ_2102_001', '27', '27.세종시 월평균 가구소득.csv')
GetCompasData('SBJ_2102_001', '28', '28.세종시 지역별 세대원수별_세대수.csv')
GetCompasData('SBJ_2102_001', '29', '29.세종시 거주 의사(향후).csv')
GetCompasData('SBJ_2102_001', '30', '30.세종시 법정경계(시군구).geojson')
GetCompasData('SBJ_2102_001', '31', '31.세종시 법정경계(읍면동).geojson')
GetCompasData('SBJ_2102_001', '32', '32.세종시 행정경계(읍면동).geojson')
GetCompasData('SBJ_2102_001', '33', '33.세종시 지적도.geojson')
GetCompasData('SBJ_2102_001', '34', '34.세종시 건물분포_이미지.zip')

for path in list(input_path.glob('*.csv')) + list(input_path.glob('*.geojson')):
    print(path)
```

input/1.세종시 전유부.csv

input/10.세종시 오피스텔(전월세)_실거래가.csv

```

input/10.세종시_오피스텔(전월세)_실거래가.csv
input/11.세종시_상업업무용(매매)_실거래가.csv
input/12.세종시_토지(매매)_실거래가.csv
input/13.세종시_분양권_실거래가.csv
input/14.세종시_상권정보.csv
input/15.세종시_상권정보_업종코드.csv
input/17.세종시_사업체_매출정보.csv
input/18.세종시_개별공시지가(2017~2020).csv
input/2.세종시_표제부.csv
input/20.세종시_전입자수.csv
input/21.세종시_전출자수.csv
input/22.세종시_연령별_인구현황.csv
input/28.세종시_지역별_세대원수별_세대수.csv
input/25.세종시_가구_월평균_소비지출액.csv
input/26.세종시_거주기간.csv
input/27.세종시_월평균_가구소득.csv
input/29.세종시_거주의사(향후).csv
input/3.세종시_아파트(매매)_실거래가.csv
input/4.세종시_연립다세대(매매)_실거래가.csv
input/5.세종시_단독다가구(매매)_실거래가.csv
input/6.세종시_오피스텔(매매)_실거래가.csv
input/7.세종시_아파트(전월세)_실거래가.csv
input/8.세종시_연립다세대(전월세)_실거래가.csv
input/9.세종시_단독다가구(전월세)_실거래가.csv
input/16.세종시_사업체_매출정보.geojson
input/19.세종시_연령별_거주인구정보_격자.geojson
input/23.세종시_도로명주소_건물.geojson
input/24.세종시_건축물연면적_격자.geojson
input/30.세종시_법정경계(시군구).geojson
input/31.세종시_법정경계(읍면동).geojson
input/32.세종시_행정경계(읍면동).geojson
input/33.세종시_지적도.geojson

```

In [40]:

```

# 자료 불러오기
아파트_매매 = pd.read_csv(input_path.joinpath('3.세종시_아파트(매매)_실거래가.csv'), encoding='cp949')
아파트_매매

# 세종 특별자치시 없애기
for i in range(len(아파트_매매['시군구'])):
    아파트_매매['시군구'][i] = 아파트_매매['시군구'][i][7:]

# col 확인
아파트_매매.columns

# 주소별 거래금액 평균 구하기
data_아파트 = 아파트_매매.groupby('시군구')['거래금액(만원)'].mean()
data_아파트

# 그래프 한글깨짐 방지
import matplotlib
matplotlib.font_manager._rebuild()
plt.rc("font", family="Malgun Gothic")

# 지역 별 평균 가격 그래프
ax = data_아파트.plot(kind='bar', title='지역별 아파트 평균 가격', figsize=(15, 5), legend=True,
fontsize=10, color='#9467bd')
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('아파트 평균 가격', fontsize=12)
ax.legend(['평균가격'], fontsize=12)

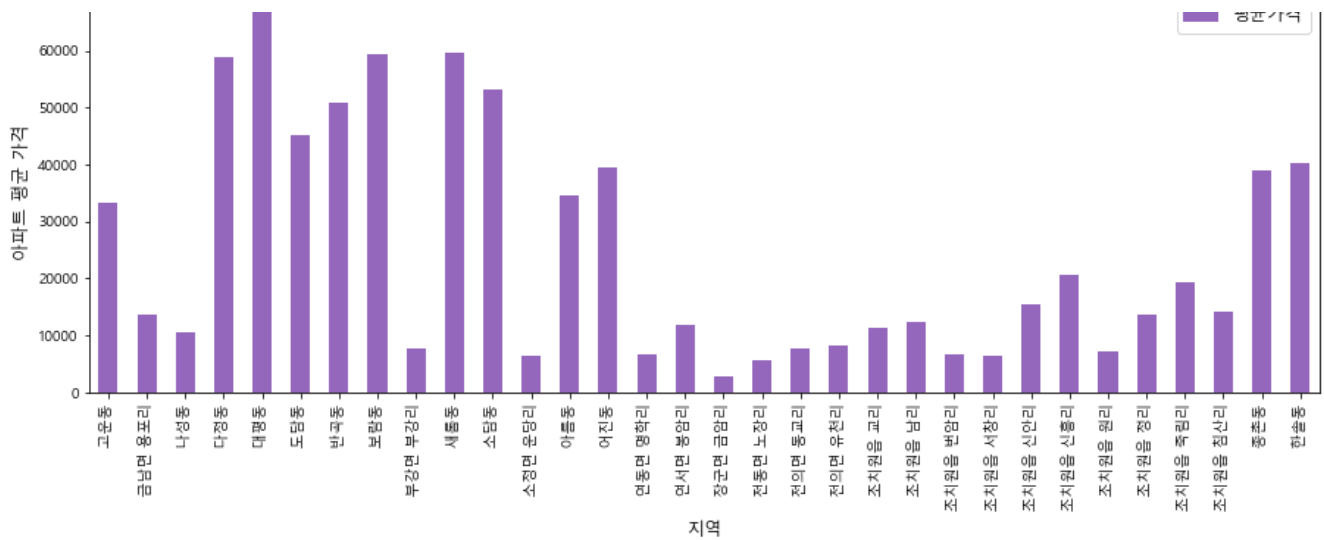
```

/opt/app-root/lib/python3.6/site-packages/ipykernel_launcher.py:7: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
import sys

Out[40]:

<matplotlib.legend.Legend at 0x7f730d494ac8>



In [41]:

```
# 자료 불러오기
연립다세대_매매 = pd.read_csv(input_path.joinpath('4.세종시_연립다세대(매매)_실거래가.csv'), thousands =
',', encoding='cp949')
연립다세대_매매

# 세종 특별자치시 없애기
for i in range(len(연립다세대_매매['시군구'])):
    연립다세대_매매['시군구'][i] = 연립다세대_매매['시군구'][i][7:]

# col 확인
연립다세대_매매.columns

# 주소별 거래금액 평균 구하기
data_연립다세대 = 연립다세대_매매.groupby('시군구')['거래금액(만원)'].mean()
data_연립다세대

import matplotlib
matplotlib.font_manager._rebuild()
plt.rc("font", family="Malgun Gothic")

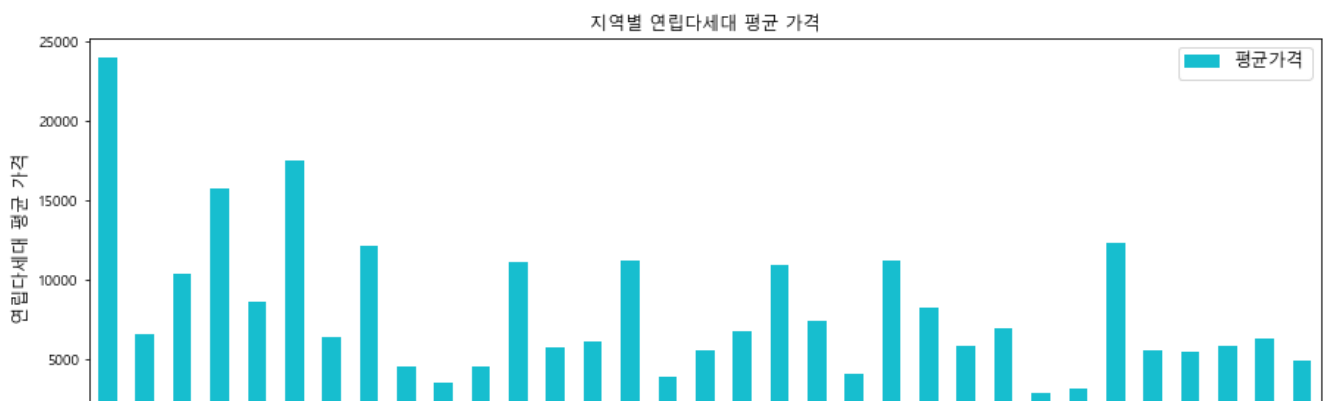
# 지역 별 평균 가격 그래프
ax = data_연립다세대.plot(kind='bar', title='지역별 연립다세대 평균 가격', figsize=(15, 5), legend=True,
    fontsize=10, color='#17becf')
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('연립다세대 평균 가격', fontsize=12)
ax.legend(['평균가격'], fontsize=12)
```

/opt/app-root/lib/python3.6/site-packages/ipykernel_launcher.py:7: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
import sys

Out[41]:

<matplotlib.legend.Legend at 0x7f730d48ae80>





In [42]:

```
# 자료 불러오기
단독다가구_매매 = pd.read_csv(input_path.joinpath('5.세종시_단독다가구(매매)_실거래가.csv'), encoding='cp949')
단독다가구_매매

# 세종 특별자치시 없애기
for i in range(len(단독다가구_매매['시군구'])):
    단독다가구_매매['시군구'][i] = 단독다가구_매매['시군구'][i][7:]

# col 확인
단독다가구_매매.columns

# 주소별 거래금액 평균 구하기
data_단독다가구 = 단독다가구_매매.groupby('시군구')['거래금액(만원)'].mean()
data_단독다가구
import matplotlib
matplotlib.font_manager._rebuild()
plt.rc("font", family="Malgun Gothic")

# 지역 별 평균 가격 그래프
ax = data_단독다가구.plot(kind='bar', title='지역별 단독다가구 평균 가격', figsize=(80,15), legend=True,
fontsize=32, color='#8c564b')
ax.set_xlabel('지역', fontsize=30)
ax.set_ylabel('단독다가구 평균 가격', fontsize=30)
ax.legend(['평균 가격'], fontsize=30)
```

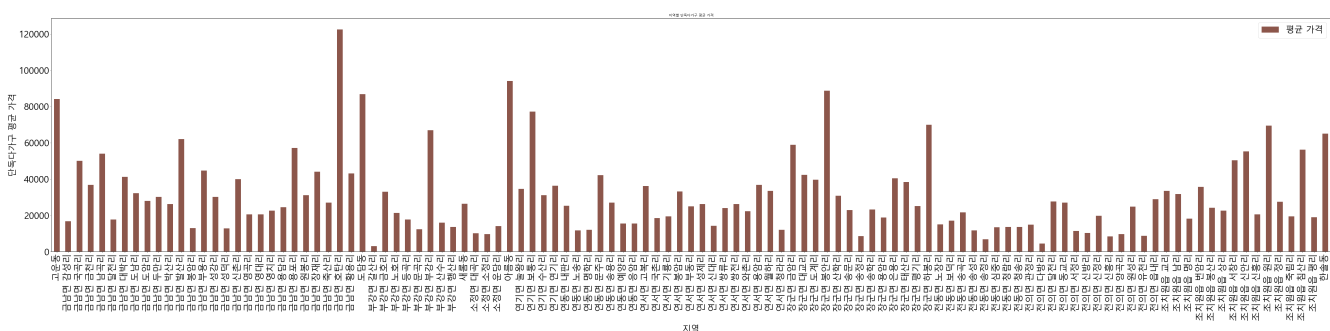
/opt/app-root/lib/python3.6/site-packages/ipykernel_launcher.py:7: SettingWithCopyWarning: A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
import sys
```

Out[42]:

<matplotlib.legend.Legend at 0x7f730ca6f2b0>



In [44]:

```
# 자료 불러오기
오피스텔_매매 = pd.read_csv(input_path.joinpath('6.세종시_오피스텔(매매)_실거래가.csv'), encoding='cp949')
오피스텔_매매

# 세종 특별자치시 없애기
for i in range(len(오피스텔_매매['시군구'])):
    오피스텔_매매['시군구'][i] = 오피스텔_매매['시군구'][i][7:]

# col 확인
오피스텔_매매.columns

# 주소별 거래금액 평균 구하기
```

```
data_오피스텔 = 오피스텔_매매.groupby('시군구')['거래금액(만원)'].mean()
data_오피스텔
import matplotlib
matplotlib.font_manager._rebuild()
plt.rc("font", family="Malgun Gothic")

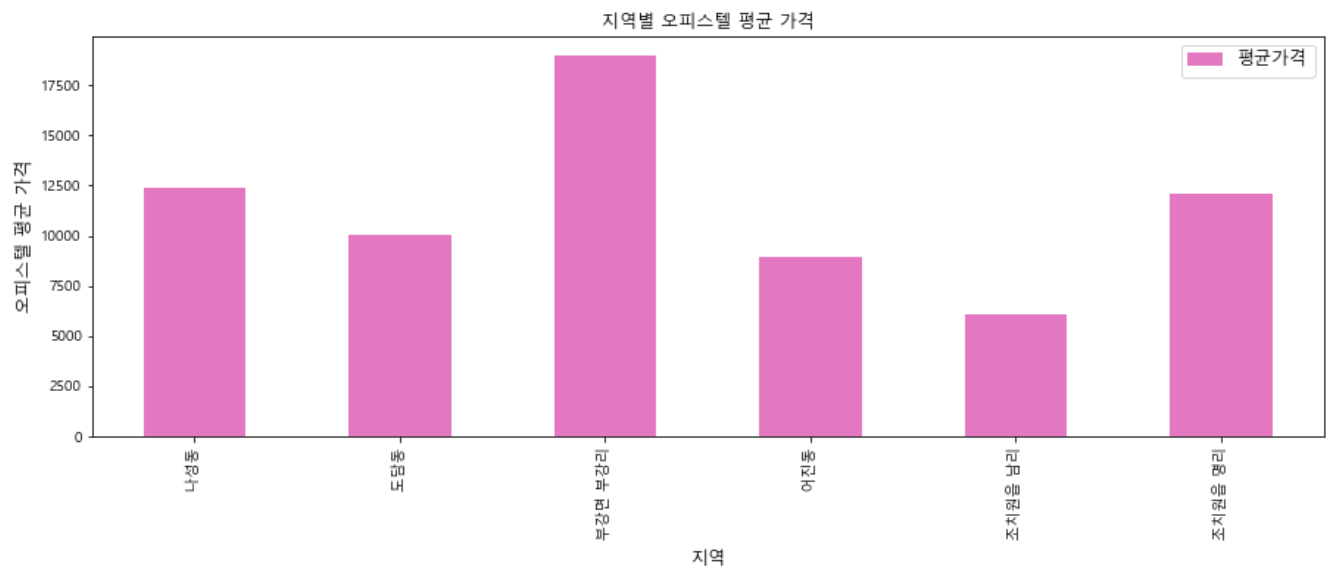
# 지역 별 평균 가격 그래프
ax = data_오피스텔.plot(kind='bar', title='지역별 오피스텔 평균 가격', figsize=(15, 5), legend=True,
fontsize=10, color='#e377c2')
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('오피스텔 평균 가격', fontsize=12)
ax.legend(['평균가격'], fontsize=12)
```

/opt/app-root/lib/python3.6/site-packages/ipykernel_launcher.py:7: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
import sys

Out[44]:

<matplotlib.legend.Legend at 0x7f730c0a42e8>



In [49]:

```
# 자료 불러오기
아파트_전월세 = pd.read_csv(input_path.joinpath('7.세종시_아파트(전월세)_실거래가.csv'), thousands = ',', encoding='cp949')
아파트_전월세

# 세종 특별자치시 없애기
for i in range(len(아파트_전월세['시군구'])):
    아파트_전월세['시군구'][i] = 아파트_전월세['시군구'][i][7:]

# col 확인
아파트_전월세.columns

# 주소별 거래금액 평균 구하기
data_아파트_전월세 = 아파트_전월세.groupby('시군구')['보증금(만원)'].mean()
data_아파트_전월세
import matplotlib
matplotlib.font_manager._rebuild()
plt.rc("font", family="Malgun Gothic")

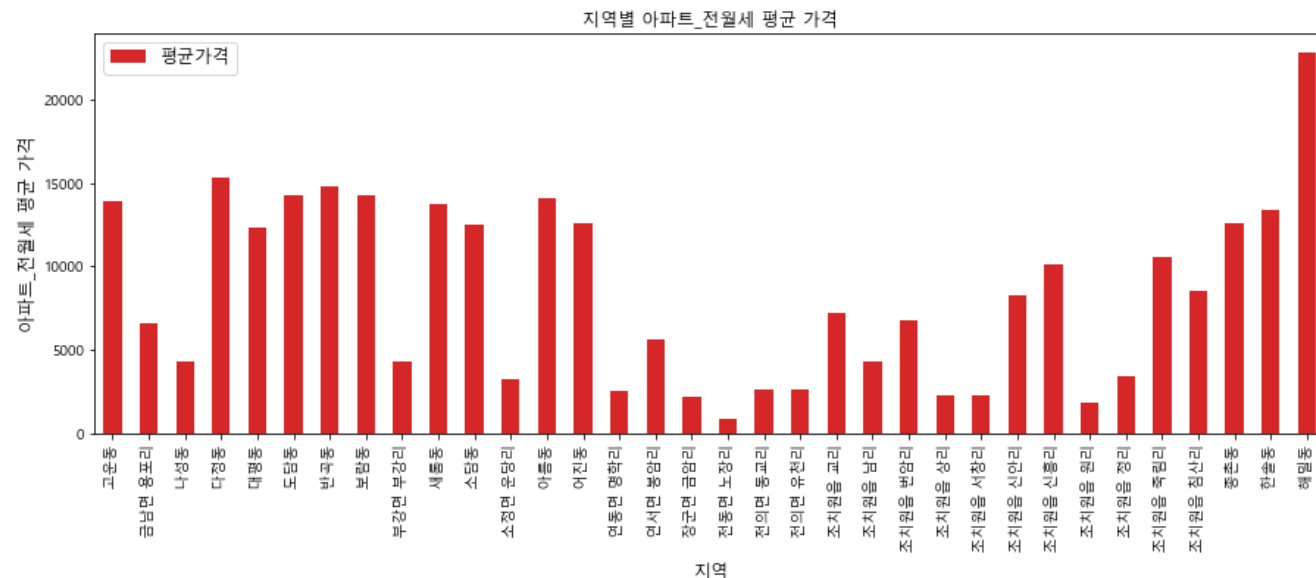
# 지역 별 평균 가격 그래프
ax = data_아파트_전월세.plot(kind='bar', title='지역별 아파트_전월세 평균 가격', figsize=(15, 5),
legend=True, fontsize=10, color='#d62728')
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('아파트_전월세 평균 가격', fontsize=12)
ax.legend(['평균가격'], fontsize=12)
```

```
/opt/app-root/lib/python3.6/site-packages/ipykernel_launcher.py:7: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
import sys
```

Out[49]:

<matplotlib.legend.Legend at 0x7f730c03ea58>



In [52]:

```
# 자료 불러오기
연립다세대_전월세= pd.read_csv(input_path.joinpath('8.세종시_연립다세대 (전월세)_실거래가.csv'),thousands =
',',encoding='cp949' )
연립다세대_전월세

# 세종 특별자치시 없애기
for i in range(len(연립다세대_전월세['시군구'])):
    연립다세대_전월세['시군구'][i] = 연립다세대_전월세['시군구'][i][7:]

# col 확인
연립다세대_전월세.columns

# 주소별 거래금액 평균 구하기
data_연립다세대_전월세 = 연립다세대_전월세.groupby('시군구')['보증금(만원)'].mean()
data_연립다세대_전월세
import matplotlib
matplotlib.font_manager._rebuild()
plt.rc("font", family="Malgun Gothic")

# 지역 별 평균 가격 그래프
ax = data_연립다세대_전월세.plot(kind='bar', title='지역별 연립다세대_전월세 평균 가격', figsize=(15, 5)
, legend=True, fontsize=10, color='#2ca02c')
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('연립다세대_전월세 평균 가격', fontsize=12)
ax.legend(['평균가격'], fontsize=12)
```

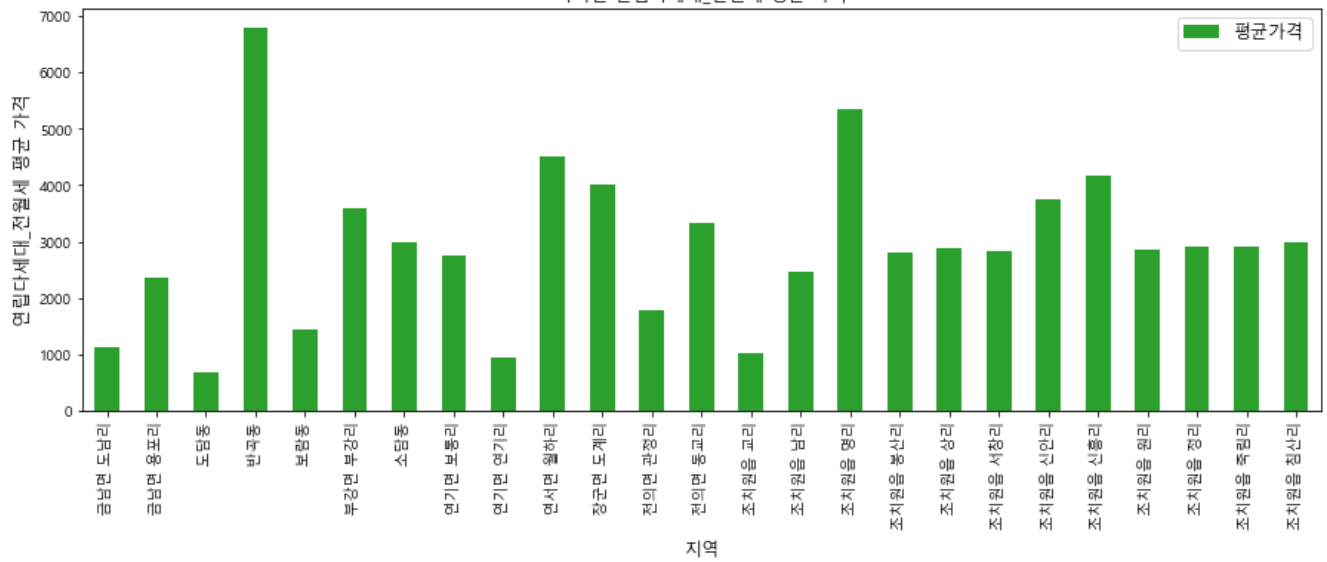
```
/opt/app-root/lib/python3.6/site-packages/ipykernel_launcher.py:7: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
import sys
```

Out[52]:

<matplotlib.legend.Legend at 0x7f730e36ccc0>

지역별 근원시설별_근원시설 평균 가격



In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

```
# 아파트 전용면적으로 평수 구하기
아파트_매매['평수'] = 아파트_매매['전용면적 (㎡)']/3.3

# 평수를 소수점 2자리까지 남기기
아파트_매매['평수'] = round(아파트_매매['평수'], 2)

아파트_매매['평수'] = (아파트_매매['평수']/10)*10
아파트_매매
```

In []:

In []:

```
data1 = pd.crosstab(아파트_매매.시군구, 아파트_매매.평수, margins=True)
data1
```

In []:

```
data2 = pd.crosstab(index=[아파트_매매.시군구, 아파트_매매.평수], columns=아파트_매매['계약일'],
margins=True)
data2
```

In []:

In []:

In []:

In []:

