

In [1]:

```
#!/usr/bin/env python
#-*-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 [3]:

```

# 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='indigo')
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('아파트 평균 가격', fontsize=12)
ax.legend(['평균가격'], fontsize=12)

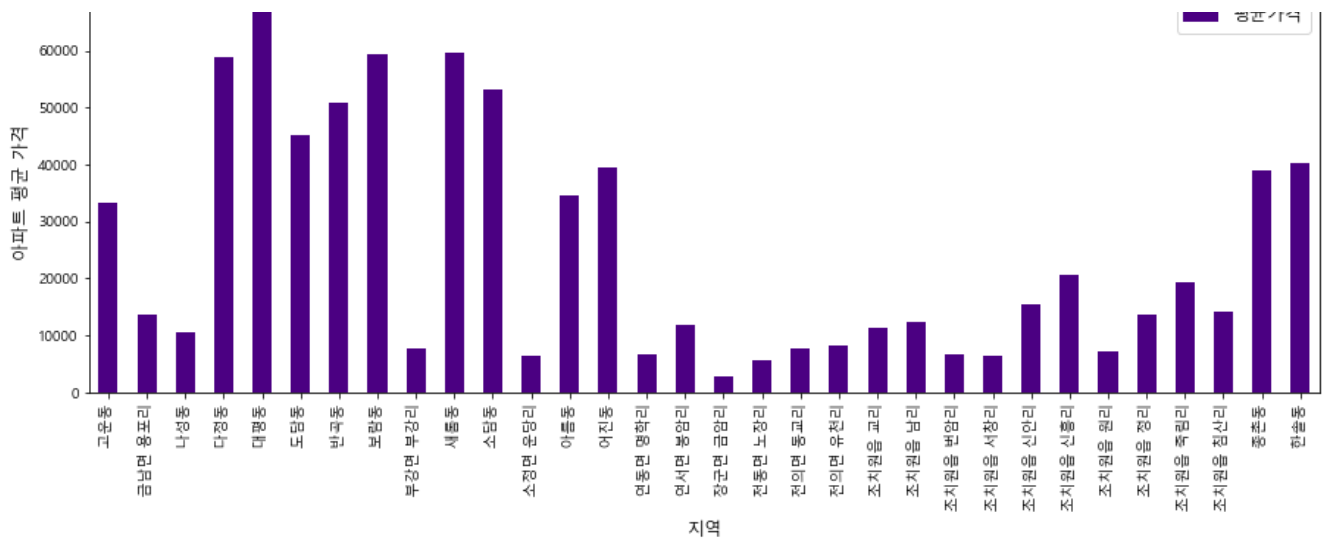
```

/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[3]:

<matplotlib.legend.Legend at 0x7f9deaf2f940>



In [4]:

```
# 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

# 그래프 한글깨짐 방지
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,
e, fontsize=10, color='deepskyblue')
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('연립다세대 평균 가격', fontsize=12)
ax.legend(['평균가격'], fontsize=12)
```

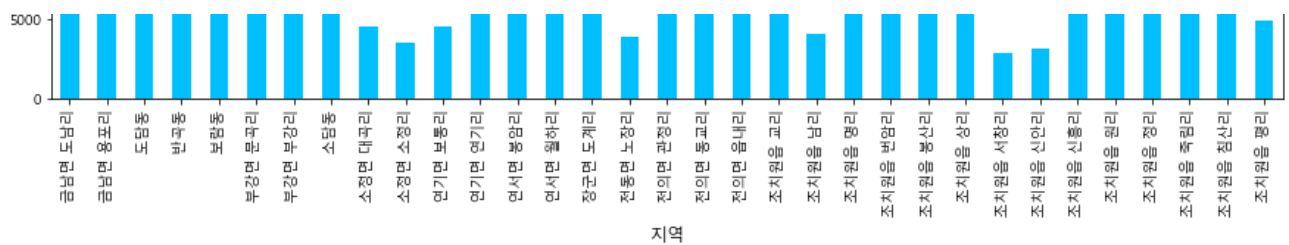
/opt/app-root/lib/python3.6/site-packages/ipykernel_launcher.py:7: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

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[4]:

<matplotlib.legend.Legend at 0x7f9d4464aa58>





In [5]:

```
# 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

# 그래프 한글깨짐 방지
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='midnightblue')
ax.set_xlabel('지역', fontsize=30)
ax.set_ylabel('단독다가구 평균 가격', fontsize=30)
ax.legend(['평균 가격'], fontsize=30)
```

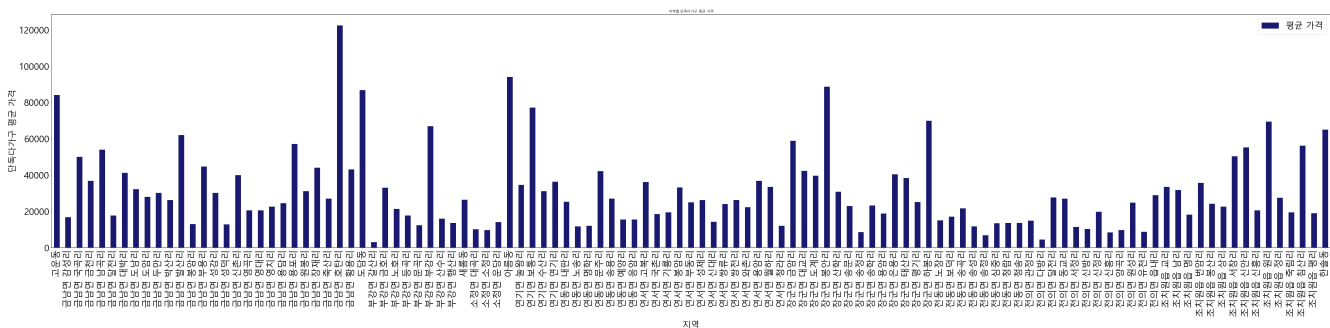
/opt/app-root/lib/python3.6/site-packages/ipykernel_launcher.py:7: SettingWithCopyWarning: A value is trying to be set on a copy of a slice from a DataFrame

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[5]:

<matplotlib.legend.Legend at 0x7f9d453a90f0>



In [6]:

```
# 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

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

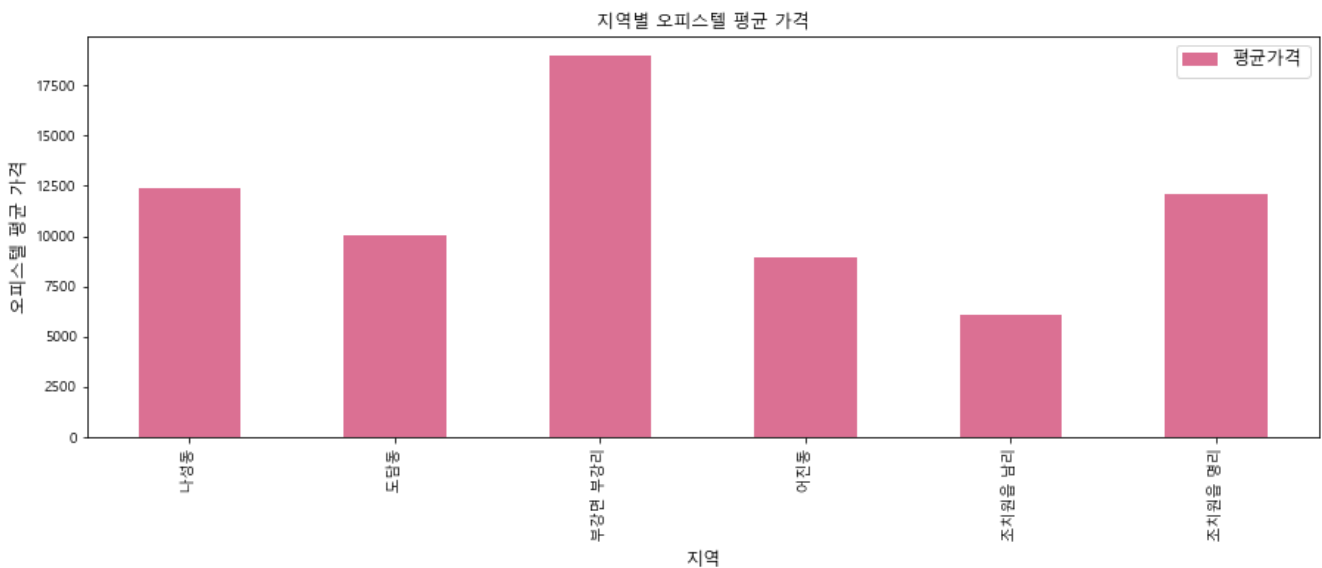
# 지역 별 평균 가격 그래프
ax = data_오피스텔.plot(kind='bar', title='지역별 오피스텔 평균 가격', figsize=(15, 5), legend=True,
fontsize=10, color='palevioletred')
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('오피스텔 평균 가격', fontsize=12)
ax.legend(['평균가격'], fontsize=12)
```

/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[6]:

<matplotlib.legend.Legend at 0x7f9d445d7748>



In []:

In []:

In [15]:

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

# 행 확인
아파트_전월세.columns

# 세종특별자치시 없애기
for i in range(len(아파트_전월세['시군구'])):
```

```
아파트_전월세['시군구'][i] = 아파트_전월세['시군구'][i][7:]
```

```
# 지역마다 전, 월세 평균
```

```
data_아파트_전월세 = 아파트_전월세.groupby(['시군구', '전월세구분'])[['보증금 (만원)', '월세 (만원)']].mean()
data_아파트_전월세 = data_아파트_전월세.unstack().fillna(0)
data_아파트_전월세
```

```
data_아파트_전월세_월세보증금 = data_아파트_전월세['보증금 (만원)', '월세']
data_아파트_전월세_월세 = data_아파트_전월세['월세 (만원)', '월세']
data_아파트_전세 = data_아파트_전월세['보증금 (만원)', '전세']
```

```
#data_아파트_전월세['보증금 (만원)', '전세'] = data_아파트_전세
```

```
del (data_아파트_전월세['월세 (만원)', '전세'])
```

```
del (data_아파트_전월세['보증금 (만원)', '전세'])
```

```
round(data_아파트_전월세, 2)
```

```
# 그래프 한글깨짐 방지
```

```
matplotlib.font_manager._rebuild()
```

```
plt.rc("font", family="Malgun Gothic")
```

```
# 지역 별 평균 가격 그래프
```

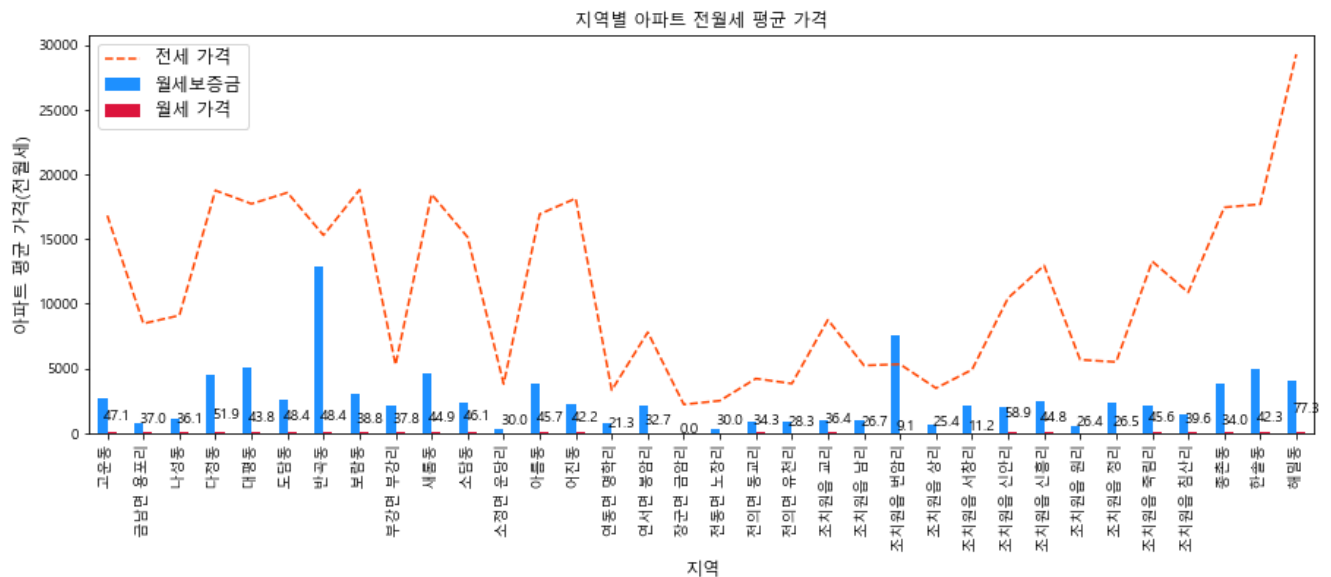
```
ax = data_아파트_전월세.plot(kind='bar', title='지역별 아파트 전월세 평균 가격', figsize=(15, 5),
legend=True, fontsize=10, color=['dodgerblue', 'crimson'])
plt.plot(data_아파트_전세, linestyle='--', color='orangered')
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('아파트 평균 가격(전월세)', fontsize=12)
ax.legend(['전세 가격', '월세보증금', '월세 가격'], fontsize=12)
```

```
for i, v in enumerate(data_아파트_전월세['월세 (만원)', '월세']):
    ax.text(i-0.1, v+v*20, str(round(v, 1)))
```

/opt/app-root/lib/python3.6/site-packages/ipykernel_launcher.py:10: 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

```
# Remove the CWD from sys.path while we load stuff.
```



```
In [16]:
```

```
# 자료 불러오기
```

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

```
# 행 확인
```

```
연립다세대_전월세.columns
```

```
# 세종특별자치시 없애기
```

```
for i in range(len(연립다세대_전월세['시군구'])):
    연립다세대_전월세['시군구'][i] = 연립다세대_전월세['시군구'][i][7:]
```

```
# 지역마다 전, 월세 평균
data_연립다세대_전월세 = 연립다세대_전월세.groupby(['시군구', '전월세구분'])[['보증금 (만원)', '월세 (만원)']].mean()
data_연립다세대_전월세 = data_연립다세대_전월세.unstack().fillna(0)
data_연립다세대_전월세

data_연립다세대_전월세_월세보증금 = data_연립다세대_전월세[['보증금 (만원)', '월세']]
data_연립다세대_전월세_월세 = data_연립다세대_전월세[['월세 (만원)', '월세']]
data_연립다세대_전세 = data_연립다세대_전월세[['보증금 (만원)', '전세']]

#data_연립다세대_전월세[['보증금 (만원)', '전세']] = data_연립다세대_전세
del(data_연립다세대_전월세[['월세 (만원)', '전세']])
del(data_연립다세대_전월세[['보증금 (만원)', '전세']])

round(data_연립다세대_전월세, 2)

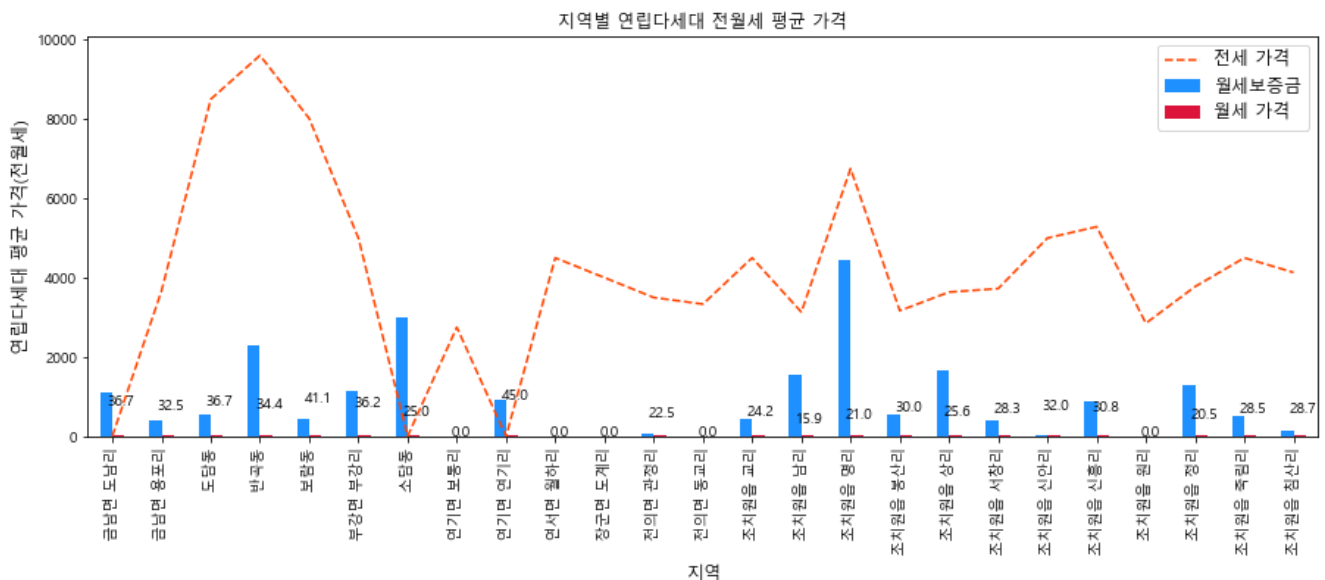
# 그래프 한글깨짐 방지
matplotlib.font_manager._rebuild()
plt.rc("font", family="Malgun Gothic")

# 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=['dodgerblue', 'crimson'])
plt.plot(data_연립다세대_전세, linestyle='--', color='orangered')
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('연립다세대 평균 가격 (전월세)', fontsize=12)
ax.legend(['전세 가격', '월세보증금', '월세 가격'], fontsize=12)

for i, v in enumerate(data_연립다세대_전월세[['월세 (만원)', '월세']]):
    ax.text(i-0.1, v+v*20, str(round(v, 1)))
```

/opt/app-root/lib/python3.6/site-packages/ipykernel_launcher.py:10: 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
Remove the CWD from sys.path while we load stuff.



In [32]:

```
# 자료 불러오기
단독다가구_전월세 = pd.read_csv(input_path.joinpath('9.세종시_단독다가구 (전월세)_실거래가.csv'), thousands =
    ',', encoding='cp949')
단독다가구_전월세

# 행 확인
단독다가구_전월세.columns

# 세종특별자치시 없애기
for i in range(len(단독다가구_전월세[['시군구']])):
    단독다가구_전월세[['시군구']][i] = 단독다가구_전월세[['시군구']][i][7:]
```

```

# 지역마다 전,월세 평균
data_단독다가구_전월세 = 단독다가구_전월세.groupby(['시군구', '전월세구분'])[['보증금(만원)', '월세(만원)']].mean()
data_단독다가구_전월세 = data_단독다가구_전월세.unstack().fillna(0)
data_단독다가구_전월세

data_단독다가구_전월세_월세보증금 = data_단독다가구_전월세['보증금(만원)', '월세']
data_단독다가구_전월세_월세 = data_단독다가구_전월세['월세(만원)', '월세']
data_단독다가구_전세 = data_단독다가구_전월세['보증금(만원)', '전세']

#data_단독다가구_전월세['보증금(만원)', '전세'] = data_단독다가구_전세
del(data_단독다가구_전월세['월세(만원)', '전세'])
del(data_단독다가구_전월세['보증금(만원)', '전세'])

round(data_단독다가구_전월세,2)

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

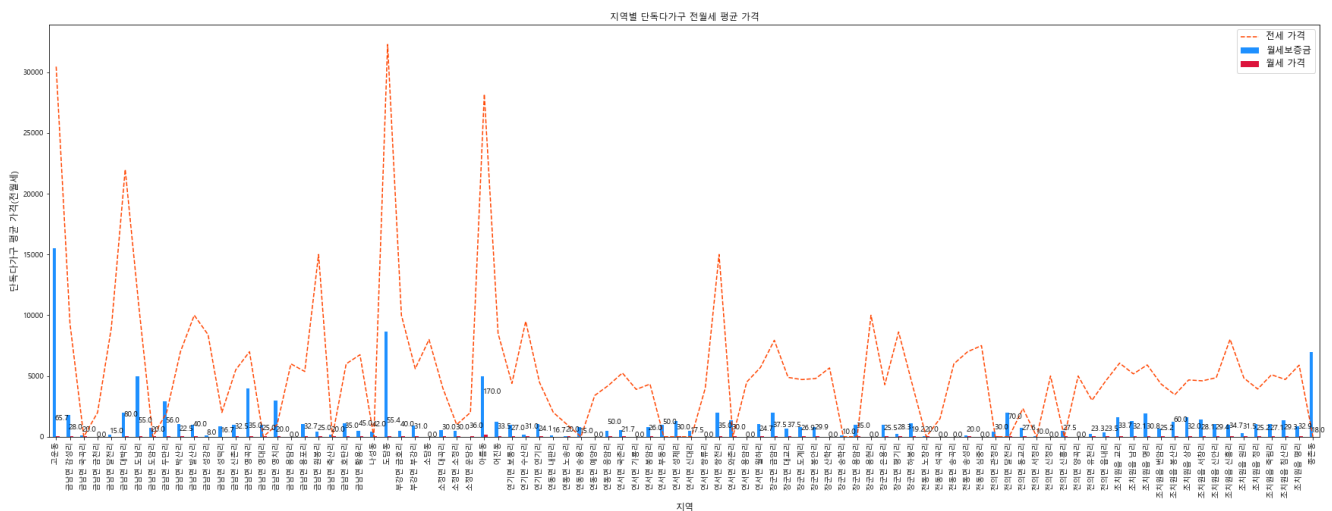
# 지역 별 평균 가격 그래프
ax = data_단독다가구_전월세.plot(kind='bar', title='지역별 단독다가구 전월세 평균 가격', figsize=(30, 10),
    legend=True, fontsize=10, color=['dodgerblue', 'crimson'])
plt.plot(data_단독다가구_전세, linestyle='--', color='orangered')
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('단독다가구 평균 가격(전월세)', fontsize=12)
ax.legend(['전세 가격', '월세보증금', '월세 가격'], fontsize=12)

for i, v in enumerate(data_단독다가구_전월세['월세(만원)', '월세']):
    ax.text(i-0.1, v+v*20, str(round(v, 1)))

```

/opt/app-root/lib/python3.6/site-packages/ipykernel_launcher.py:10: 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
Remove the CWD from sys.path while we load stuff.



In [34]:

```

# 자료 불러오기
오피스텔_전월세 = pd.read_csv(input_path.joinpath('10.세종시_오피스텔(전월세)_실거래가.csv'), thousands = ', ', encoding='cp949')
오피스텔_전월세

# 행 확인
오피스텔_전월세.columns

# 세종특별자치시 없애기
for i in range(len(오피스텔_전월세['시군구'])):
    오피스텔_전월세['시군구'][i] = 오피스텔_전월세['시군구'][i][7:]

# 지역마다 전,월세 평균
data_오피스텔_전월세 = 오피스텔_전월세.groupby(['시군구', '전월세구분'])[['보증금(만원)', '월세(만원)']].mean()

```



```

data_오피스텔_전월세 = data_오피스텔_전월세.unstack().fillna(0)
data_오피스텔_전월세

data_오피스텔_전월세_월세보증금 = data_오피스텔_전월세['보증금 (만원)', '월세']
data_오피스텔_전월세_월세 = data_오피스텔_전월세['월세 (만원)', '월세']
data_오피스텔_전세 = data_오피스텔_전월세['보증금 (만원)', '전세']

#data_오피스텔_전월세['보증금 (만원)', '전세'] = data_오피스텔_전세
del (data_오피스텔_전월세['월세 (만원)', '전세'])
del (data_오피스텔_전월세['보증금 (만원)', '전세'])

round(data_오피스텔_전월세, 2)

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

# 지역 별 평균 가격 그래프
ax = data_오피스텔_전월세.plot(kind='bar', title='지역별 오피스텔 전월세 평균 가격', figsize=(15, 5), legend=True, fontsize=10, color=['dodgerblue', 'crimson'])
plt.plot(data_오피스텔_전세, linestyle='--', color='orangered')
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('오피스텔 평균 가격(전월세)', fontsize=12)
ax.legend(['전세 가격', '월세보증금', '월세 가격'], fontsize=12)

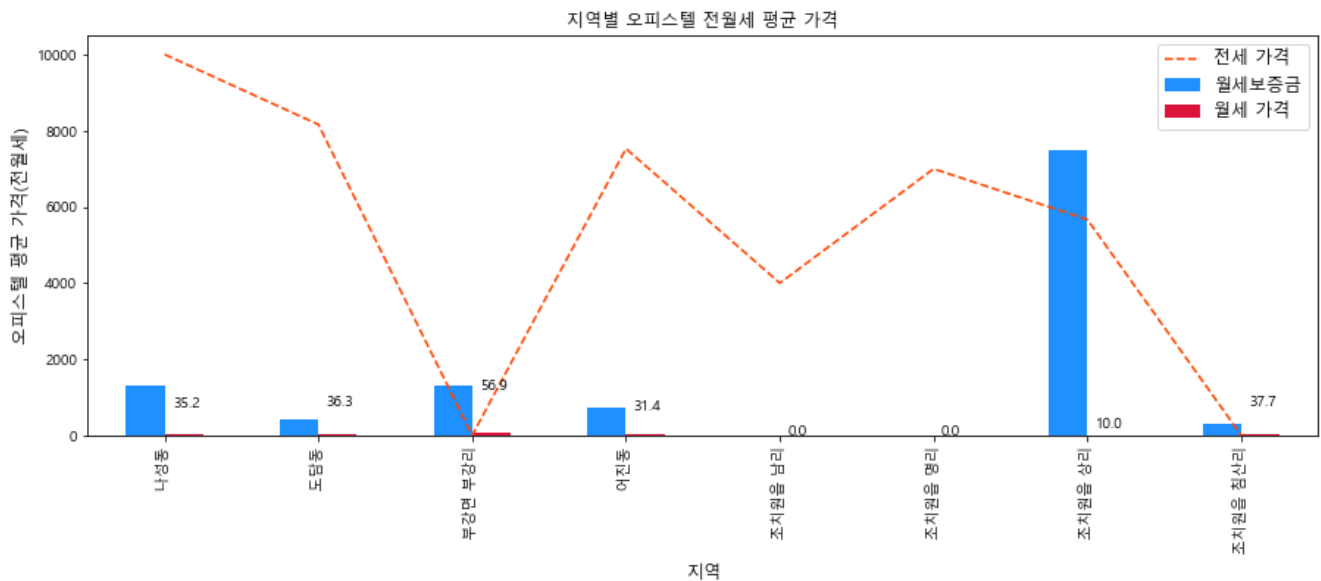
for i, v in enumerate(data_오피스텔_전월세['월세 (만원)', '월세']):
    ax.text(i+0.05, v+v*20, str(round(v, 1)))

```

/opt/app-root/lib/python3.6/site-packages/ipykernel_launcher.py:10: 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

Remove the CWD from sys.path while we load stuff.



In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In [9]:

```
# 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
```

Out [9]:

	계약년월	계약일	시군구	본번	부번	단지명	전용면적(㎡)	거래금액(만원)	층	건축년도	평수
0	201701	1	보람동	628	2	스타힐타워I(C2-3)	16.9849	8600	7	2016	0.0
1	201701	1	중촌동	683	0	가재마을2단지(호반베르디움3차)	59.8652	27300	7	2014	10.0
2	201701	2	고운동	1393	0	가락마을15단지(중흥파크뷰)	84.9894	32000	3	2015	20.0
3	201701	2	고운동	1706	0	가락마을8단지(고운뜰아파트)	84.9800	30500	20	2015	20.0
4	201701	2	아름동	1284	0	범지기마을7단지 호반에코시티	84.9280	31300	6	2014	20.0
...
21135	202010	27	조치원읍 신흥리	399	0	조치원 신흥 e편한세상(아)	176.3430	50000	10	2008	50.0
21136	202010	27	조치원읍 침산리	89	0	육일아파트	84.9900	23500	20	1996	20.0
21137	202010	28	도담동	639	0	도렘마을17단지(모아미래도)	84.9690	68000	9	2015	20.0
21138	202010	28	부강면 부강리	401	48	주연	59.8200	6500	11	1992	10.0
21139	202010	28	조치원읍 죽림리	396	0	조치원죽림자이	84.6641	37000	6	2008	20.0

21140 rows × 11 columns

In []:

In [10]:

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

Out [10]:

평수	0.0	10.0	20.0	30.0	40.0	50.0	60.0	All
시군구								
고운동	0	880	1317	256	0	0	1	2454
금남면 용포리	0	452	89	0	0	0	0	541
나성동	288	58	0	0	0	0	0	346
다정동	0	179	434	137	0	0	0	750
대평동	0	42	144	58	0	0	0	244
도담동	165	517	1262	278	1	0	0	2223
반곡동	1	74	78	16	0	0	0	169
보람동	25	149	293	215	0	1	0	683
부강면 부강리	0	101	11	0	0	0	0	112
새롬동	0	421	509	219	1	0	0	1150
소담동	0	342	593	33	0	0	0	968
소정면 운당리	0	28	1	0	0	0	0	29
아름동	0	97	2262	179	0	0	0	2538
임지동	155	27	187	20	0	0	0	400

연동면 명학리 시군구	0.0	10.0	20.0	30.0	40.0	50.0	60.0	All
연동면 봉암리	0	54	69	0	0	0	0	123
장군면 금암리	69	1	0	0	0	0	0	70
전동면 노장리	0	21	0	0	0	0	0	21
전의면 동교리	0	64	0	0	0	0	0	64
전의면 유천리	0	36	0	0	0	0	0	36
조치원읍 교리	0	35	65	3	0	0	0	103
조치원읍 남리	4	11	104	0	0	0	0	119
조치원읍 번암리	0	204	0	0	0	0	0	204
조치원읍 서창리	0	27	0	0	0	0	0	27
조치원읍 신안리	146	140	111	236	46	0	0	679
조치원읍 신흥리	0	208	341	167	51	11	0	778
조치원읍 원리	17	0	2	0	0	0	0	19
조치원읍 정리	0	3	7	0	0	0	0	10
조치원읍 죽림리	0	117	724	294	11	0	0	1146
조치원읍 침산리	0	226	109	15	22	0	0	372
종촌동	0	682	1770	294	0	0	0	2746
한솔동	0	255	901	671	12	0	0	1839
All	899	5600	11383	3101	144	12	1	21140

In []:

In [11]:

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

Out[11]:

	계약일	1	2	3	4	5	6	7	8	9	10	...	23	24	25	26	27	28	29	30	31	All
시군구	평수																					
고운동	10.0	33	26	26	34	24	31	37	31	36	28	...	20	20	27	22	29	32	34	28	16	880
	20.0	38	35	30	50	42	37	42	37	44	42	...	47	38	41	45	44	37	51	36	27	1317
	30.0	9	8	7	5	4	7	9	8	4	9	...	9	12	5	11	9	9	6	15	6	256
	60.0	0	1	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	1
금남면 용포리	10.0	16	8	10	16	18	16	19	22	13	12	...	13	21	9	11	16	17	16	16	6	452
...
한솔동	10.0	9	6	14	10	8	8	6	11	5	7	...	14	13	8	11	8	6	3	4	3	255
	20.0	34	32	25	21	20	21	25	40	32	37	...	17	32	28	35	20	46	29	21	19	901
	30.0	29	23	14	30	26	23	18	26	16	22	...	20	23	20	10	23	19	26	21	14	671
	40.0	2	0	1	0	1	0	0	1	0	0	...	0	0	1	0	0	0	1	0	0	12
All		699	626	628	666	1228	643	706	712	655	710	...	668	633	642	586	671	653	677	605	388	21140

93 rows × 32 columns

In []:

In [12]:

```
# 전의면의 전세 평균구하기
```

```

# 지역마다 전세 평균 파악하기

# 시군구, 전월세구분 행을 index에서 해제시킴
data_아파트_전세=data_아파트_전월세.reset_index()
data_아파트_전세

# 행 이름 체크
data_아파트_전세.columns

# 전월세 구분 행에서 '월세' 열 삭제
data = data_아파트_전세[data_아파트_전세['전월세구분'] == '월세'].index
data_아파트_전세 = data_아파트_전세.drop(data)
data_아파트_전세

# 시군구 행을 index로 다시 설정 (그래프x축에 나타내기 위함)
data_아파트_전세.set_index('시군구', inplace = True)
data_아파트_전세

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

# 지역 별 평균 가격 그래프
ax = data_아파트_전세.plot(kind='bar', title='지역별 아파트 전세 평균 가격', figsize=(15, 5), legend=True,
                           fontsize=10, color='lightslategray')

ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('아파트 평균 가격', fontsize=12)
ax.legend(['평균가격'], fontsize=12)

```

```

-----
KeyError                                Traceback (most recent call last)
/opt/app-root/lib/python3.6/site-packages/pandas/core/indexes/base.py in get_loc(self, key,
method, tolerance)
    2897         try:
-> 2898             return self._engine.get_loc(casted_key)
    2899         except KeyError as err:

```

```

pandas/_libs/index.pyx in pandas._libs.index.IndexEngine.get_loc()

pandas/_libs/index.pyx in pandas._libs.index.IndexEngine.get_loc()

pandas/_libs/hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHashTable.get_item()

pandas/_libs/hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHashTable.get_item()

KeyError: '전월세구분'

```

The above exception was the direct cause of the following exception:

```

KeyError                                Traceback (most recent call last)
<ipython-input-12-9bfbd17b7cce> in <module>
      9
     10 # 전월세 구분 행에서 '월세' 열 삭제
--> 11 data = data_아파트_전세[data_아파트_전세['전월세구분'] == '월세'].index
     12 data_아파트_전세 = data_아파트_전세.drop(data)
     13 data_아파트_전세

/opt/app-root/lib/python3.6/site-packages/pandas/core/frame.py in __getitem__(self, key)
    2903         if is_single_key:
    2904             if self.columns.nlevels > 1:
-> 2905                 return self._getitem_multilevel(key)
    2906             indexer = self.columns.get_loc(key)
    2907             if is_integer(indexer):

```

```

/opt/app-root/lib/python3.6/site-packages/pandas/core/frame.py in _getitem_multilevel(self, key)
    2953     def _getitem_multilevel(self, key):
    2954         # self.columns is a MultiIndex
-> 2955         loc = self.columns.get_loc(key)
    2956         if isinstance(loc, (slice, np.ndarray)):
    2957             new_columns = self.columns[loc]

```

```

/opt/app-root/lib/python3.6/site-packages/pandas/core/indexes/multi.py in get_loc(self, key,
method)
    2702
    2703     if not isinstance(key, tuple):

```

```

2704         loc = self._get_level_indexer(key, level=0)
2705         return _maybe_to_slice(loc)
2706
/opt/app-root/lib/python3.6/site-packages/pandas/core/indexes/multi.py in _get_level_indexer(self,
key, level, indexer)
2967     else:
2968
-> 2969         code = self._get_loc_single_level_index(level_index, key)
2970
2971         if level > 0 or self.lexsort_depth == 0:

/opt/app-root/lib/python3.6/site-packages/pandas/core/indexes/multi.py in
_get_loc_single_level_index(self, level_index, key)
2635         return -1
2636     else:
-> 2637         return level_index.get_loc(key)
2638
2639     def get_loc(self, key, method=None):

/opt/app-root/lib/python3.6/site-packages/pandas/core/indexes/base.py in get_loc(self, key,
method, tolerance)
2898         return self._engine.get_loc(casted_key)
2899     except KeyError as err:
-> 2900         raise KeyError(key) from err
2901
2902         if tolerance is not None:

```

KeyError: '전월세구분'

In []:

```

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

# 지역 별 평균 가격 그래프
colors = [ 'dodgerblue', 'orangered']
ax = data_아파트_전월세.plot(kind='bar', title='지역별 아파트 월세 평균 가격', figsize=(15, 5), legend
=True, fontsize=10, color=colors)
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('아파트 평균 가격(월세)', fontsize=12)
ax.legend(['월세보증금', '월세 가격'], fontsize=12)

for i, v in enumerate(data_아파트_전월세['월세(만원)', '월세']):
    ax.text(i-0.1, v*40, str(round(v, 1)))

```

In []:

```

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

# 지역 별 평균 가격 그래프
ax = data_아파트_전세.plot(kind='bar', title='지역별 아파트 전세 평균 가격', figsize=(15, 5), legend=T
rue, fontsize=10)
ax.set_xlabel('지역', fontsize=12)
ax.set_ylabel('아파트 평균 가격(전세)', fontsize=12)
ax.legend(['전세 가격'], fontsize=12)

```

In []: