목차

- <Step1.> : 문제정의 및 가설설정
 - [문제정의]
 - [가설설정]
- <Step2.> : 데이터 수집
 - [데이터 불러오기]
 - [데이터 결합]
- <Step3.> : 데이터 전처리
 - [데이터 결측값 제거]
 - [데이터 구조 변경]
- <Step4.>: 데이터 모델링 및 시각화
 - [요인분석 1:지역구]
 - [1.1 유종 간 상관관계 _ (Pairplot Chart)]
 - [1.2 지역구별 유가 _ (Boxplot)]
 - [1.3 지역구별 유가 _ (Barplot)]
 - [1.4 지역구별 월별 유가 (Heatmap)]
 - [요인분석 2 : 인건비-셀프여부]
 - [2.1 인건비와 유가의 상관관계 _ (Violinplot)]
 - [2.2 지역구별 인건비와 유가의 상관관계 _ (Violinplot)]

pr01

- [요인분석 3 : 대리점 공급가]
 - [3.1 지역별 대리점의 수 (Wordcloud)]
 - [3.2 지역구별 대리점의 수와 유가 _ (Plot)]

<Step1.> : 문제정의 및 가설설정

[문제정의]

유가에 영향을 끼치는 요인들이 무엇인지 알아보고자 한다.

[가설설정]

유가에 영향을 미치는 요인을 **3**가지로 설정하고 이 **3**가지요인이 각각 정말 유가에 영향을 미치는지 확인한다.

요인 1) 지역구

요인 2) 인건비 (셀프여부)

요인 3) 대리점 공급가

[참고자료]

전략: 같은 지역인데 주유소마다 기름값이 다른 이유는 무엇일까?

시각화: 데이터 사이언스 스쿨 Seaborn gallery

<Step2.> : 데이터 수집

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

In [2]: import platform
    from matplotlib import font_manager, rc
    plt.rcParams['axes.unicode_minus'] = False
    if platform.system() == 'Darwin':
        rc('font', family='AppleGothic')
    elif platform.system() == 'Windows':
        path = "c:/Windows/Fonts/malgun.ttf"
        font_name = font_manager.FontProperties(fname=path).get_name()
        rc('font', family=font_name)
    else:
        print('Unknown system... sorry~~~~')
```

[데이터 불러오기]

```
# 상반기 주유소 판매 데이터 불러오기 first = pd.read_csv('../data/2020년 상반기 주유소 판매가격.csv', encoding='cp949') first
```

	번호	지역	상호	주소	기간	상표	셀프 여부	고급휘 발유	휘발유	경유	실내 등유	
0	기준 : 월간 (202001~202006)	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
1	A0006039	서울 강남 구	(유)동 하석유 힐탑셀 프주유 소	서울 강남 구 논 현로 640	2020 년 01월	SK에 너지	셀프	1802.00	1657.00	1495.00	0.0	
2	A0006039	서울 강남 구	(유)동 하석유 힐탑셀 프주유 소	서울 강남 구 논 현로 640	2020 년 02월	SK에 너지	셀프	1795.38	1637.97	1483.97	0.0	
3	A0006039	서울 강남 구	(유)동 하석유 힐탑셀 프주유 소	서울 강남 구 논 현로 640	2020 년 03월	SK에 너지	셀프	1741.26	1570.61	1431.13	0.0	
4	A0006039	서울 강남 구	(유)동 하석유 힐탑셀 프주유 소	서울 강남 구 논 현로 640	2020 년 04월	SK에 너지	셀프	1617.00	1425.33	1290.00	0.0	
	2	o 기준 : 월간 (202001~202006) 1 A0006039 2 A0006039 3 A0006039	0 기준 : 월간 (202001~202006) NaN 1 A0006039 서울 강남구 2 A0006039 서울 강남구 3 A0006039 상울 강남구 4 A0006039 상울 강남	1 기준 : 월간 (202001~202006) NaN NaN 1 A0006039 서울 하석유 힐탑셀 프주유소 2 A0006039 서울 하석유 힐탑셀 프주유소 3 A0006039 서울 하석유 힐탑셀 프주유소 4 A0006039 서울 하석유 힐탑셀 프주유소 4 A0006039 사용 하석유 힐탑셀 프주유소	1 기준 : 월간 (202001~202006) NaN NaN NaN 1 A0006039 사물 하석유 강남 질탑셀 구 논 현로 소 640 2 A0006039 사물 하석유 강남 질탑셀 구 논 현로 640 3 A0006039 사물 하석유 강남 필탑셀 구 논 현로 640 4 A0006039 사물 하석유 강남 필탑셀 구 논 현로 640 4 A0006039 사물 하석유 강남 필탑셀 구 논 현로 640	1 기준 : 월간 (202001~202006) NaN NaN NaN NaN 1 A0006039 서울 하석유 강남 열탑셀 구 논 년 연현로 01월 연로 01월 연로 01월 연로 01월 연로 01월 연로 01월 연료로 01월 연료로 02월 연료로 03월 연료로 04월 3 A0006039 서울 하석유 강남 열탑셀 구 논 년 03월 연료로 03월 연료로 04월	1 기준 : 월간 (202001~202006) NaN NaN NaN NaN NaN NaN NaN 1 A0006039 서울 하석유 강남 힐탑셀 구노 변경 이1월 사물 하석유 강남 일탑셀 구노 변경 이1월 SK에 어디원 보지 2 A0006039 서울 하석유 강남 필타셀 구노 연로 640 지금 SK에 너지 3 A0006039 사물 하석유 강남 필타셀 구노 연로 640 지금 SK에 너지 4 A0006039 사물 하석유 강남 필타셀 구노 연로 640 지금 SK에 너지 4 A0006039 사물 하석유 강남 필타셀 구노 연로 640 지물 32020 보기 기본 14지	1 A0006039	1 A0006039	1 A0006039	1 A0006039 서울 아석유 강남 일반생 구는 서울 2020 년 이 2월 SK에 설프 시대 설프 1741.26 1570.61 1431.13 대원 시대	1

	번호	지역	상호	주소	기간	상표	셀프 여부	고급휘 발유	휘발유	경유	실내 등유
2990	A0009197	서울 중랑 구	현대의 영의 교육 중 성의 유주 소의 교육 주 소의 교육 수 있는데 무슨이 되었다.	서울 중랑 구 망 우로 229 (중화 동)	2020 년 02월	SK에 너지	셀프	1872.93	1542.14	1365.55	0.0
2991	A0009197	서울 중랑 구	현대 의명 이 의 의 의 의 의 의 의 의 의 의 의 의 의 의 의 의 의	서울 중랑 구 망 우로 229 (중화 동)	2020 년 03월	SK에 너지	셀프	1798.10	1479.32	1278.87	0.0
2992	A0009197	서울 중랑 구	현대 의명 이 의명	서울 중랑 구 망 우로 229 (중화 동)	2020 년 04월	SK에 너지	셀프	1652.93	1322.80	1128.67	0.0
2993	A0009197	서울 중랑 구	현 왕 교 영 교 유 양 왕 교 유 양 왕 교 유	서울 중랑 구 망 우로 229 (중화 동)	2020 년 05월	SK에 너지	셀프	1593.19	1252.55	1058.48	0.0
2994	A0009197	서울 중랑 구	현 왕 의 의 의 의 의 의 의 의 의 의 의 의 의 의 의 의 의 의	서울 중랑 구 망 우로 229 (중화 동)	2020 년 06월	현대 오일 뱅크	셀프	1618.07	1328.73	1137.67	0.0

2995 rows × 11 columns

second = pd.read_csv('../data/2020년 하반기 주유소 판매가격.csv', encoding='cp949') second

Out[4]: 고급 실내 휘발 번호 지역 상호 주소 기간 기준 : 월간 NaN (202007~202012) (유)동 하석유 서울 2020 강남구 SK에 힐탑셀 A0006039 강남 년 셀프 1635.0 1435.00 1265.00 0.0 논현로 너지 구 07월 640 소

	번호	지역	상호	주소	기간	상표	셀프 여부	고급 휘발 유	휘발유	경유	실내 등유
2	A0006039	서울 강남 구	(유)동 하석유 힐탑셀 프주유 소	서울 강남구 논현로 640	2020 년 08월	SK에 너지	셀프	1635.0	1435.00	1265.00	0.0
3	A0006039	서울 강남 구	(유)동 하석유 힐탑셀 프주유 소	서울 강남구 논현로 640	2020 년 09월	SK에 너지	셀프	1635.0	1435.00	1265.00	0.0
4	A0006039	서울 강남 구	(유)동 하석유 힐탑셀 프주유 소	서울 강남구 논현로 640	2020 년 10월	SK에 너지	셀프	1635.0	1435.00	1265.00	0.0
•••											
2946	A0009197	서울 중랑 구	현대오 일뱅크 ㈜직영 중랑교 셀프주 유소	서울 중랑구 망우로 229 (중화 동)	2020 년 08월	현대 오일 뱅크	셀프	1598.0	1402.65	1211.10	0.0
2947	A0009197	서울 중랑 구	현대오 일뱅크 ㈜직영 중랑교 셀프주 유소	서울 중랑구 망우로 229 (중화 동)	2020 년 09월	현대 오일 뱅크	셀프	1598.0	1393.00	1203.00	0.0
2948	A0009197	서울 중랑 구	현대오 일뱅크 ㈜직영 중랑교 셀프주 유소	서울 중랑구 망우로 229 (중화 동)	2020 년 10월	현대 오일 뱅크	셀프	1598.0	1338.42	1148.42	0.0
2949	A0009197	서울 중랑 구	현대오 일뱅크 ㈜직영 중랑교 셀프주 유소	서울 중랑구 망우로 229 (중화 동)	2020 년 11월	현대 오일 뱅크	셀프	1598.0	1292.00	1092.37	0.0
2950	A0009197	서울 중랑 구	현대오 일뱅크 ㈜직영 중랑교 셀프주 유소	서울 중랑구 망우로 229 (중화 동)	2020 년 12월	현대 오일 뱅크	셀프	1598.0	1367.58	1167.58	0.0

2951 rows × 11 columns

[데이터 결합]

In [5]:

상반기와 하반기 데이터를 상하로 결합

df = pd.concat([first, second])
df

Out[5]:

t[5]:		번호	지역	상호	주소	기간	상표	셀프 여부	고급휘 발유	휘발유	경유	실내 등유
	0	기준 : 월간 (202001~202006)	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
	1	A0006039	서울 강남 구	(유)동 하석유 힐탑셀 프주유 소	서울 강남 구 논 현로 640	2020 년 01월	SK에 너지	셀프	1802.00	1657.00	1495.00	0.0
	2	A0006039	서울 강남 구	(유)동 하석유 힐탑셀 프주유 소	서울 강남 구 논 현로 640	2020 년 02월	SK에 너지	셀프	1795.38	1637.97	1483.97	0.0
	3	A0006039	서울 강남 구	(유)동 하석유 힐탑셀 프주유 소	서울 강남 구 논 현로 640	2020 년 03월	SK에 너지	셀프	1741.26	1570.61	1431.13	0.0
	4	A0006039	서울 강남 구	(유)동 하석유 힐탑셀 프주유 소	서울 강남 구 논 현로 640	2020 년 04월	SK에 너지	셀프	1617.00	1425.33	1290.00	0.0
	•••											
	2946	A0009197	서울 중랑 구	현대오 일뱅크영 ㈜직망교 중랑교 유소	서울 중랑 구 망 우로 229 (중화 동)	2020 년 08월	현대 오일 뱅크	셀프	1598.00	1402.65	1211.10	0.0
	2947	A0009197	서울 중랑 구	현대오 일뱅크 영국의 중랑교 중랑교 유소	서울 중랑 구 망 우로 229 (중화 동)	2020 년 09월	현대 오일 뱅크	셀프	1598.00	1393.00	1203.00	0.0
	2948	A0009197	서울 중랑 구	현대오 일뱅크 영국 중랑교 중랑교 유소	서울 중랑 구 망 우로 229 (중화 동)	2020 년 10월	현대 오일 뱅크	셀프	1598.00	1338.42	1148.42	0.0

	번호	지역	상호	주소	기간	상표	셀프 여부	고급휘 발유	휘발유	경유	실내 등유
2949	A0009197	서울 중랑 구	현대오 일뱅직영 중랑교 주소 일반주주선 유소	서울 중랑 구 망 우로 229 (중화 동)	2020 년 11월	현대 오일 뱅크	셀프	1598.00	1292.00	1092.37	0.0
2950	A0009197	서울 중랑 구	현대의 의명 이 의명 이 의명 이 의명 이 의명 이 되는데 이 의명 이 되는데 이 되는데 이 되는데 이 되었다.	서울 중랑 구 망 우로 229 (중화 동)	2020 년 12월	현대 오일 뱅크	셀프	1598.00	1367.58	1167.58	0.0

5946 rows × 11 columns

```
# 지역 컬럼의 값을 이용하여 시와 구를 분리

df['시'] = df['지역'].str[:2]
df['구'] = df['지역'].str[3:]
df.head()
```

Out[6]:		번호	지역	상호	주소	기간	상표	셀프 여부	고급휘 발유	휘발유	경유	실내 등유	시	Ŧ
	0	기준 : 월간 (202001~202006)	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
	1	A0006039	서울 강남 구	(유) 동 석 유 일 합 프 유 소	서울 강남 구 논현 로 640	2020 년 01월	SK 에너 지	셀프	1802.00	1657.00	1495.00	0.0	서울	강닏
	2	A0006039	서울 강남 구	(유) 동 수 유 합 된 유 오 수 소	서울 강남 구 논현 로 640	2020 년 02월	SK 에너 지	셀프	1795.38	1637.97	1483.97	0.0	서울	강닏
	3	A0006039	서울 강남 구	(유) 동하 석유 힐탑 젤 주 소	서울 강남 구 논현 로 640	2020 년 03월	SK 에너 지	셀프	1741.26	1570.61	1431.13	0.0	서울	강닏 구

	번호	지역	상호	주소	기간	상표	셀프 여부	고급휘 발유	휘발유	경유	실내 등유	시	7
4	A0006039	서울 강남 구	(유) 동하 석유 힐탑 셀프 주 소	서울 강남 구 논현 로 640	2020 년 04월	SK 에너 지	셀프	1617.00	1425.33	1290.00	0.0	서울	강닏

<**Step3.>** : 데이터 전처리

[결측값 제거]

df.dropna(inplace=True)

Out[7]:		번호	지 역	상호	주소	기간	상 표	셀 프 여 부	고급휘 발유	휘발유	경유	실 내 등 유	시	구
	1	A0006039	서 울 강 남 구	(유)동하 석유 힐탑 셀프주유 소	서울 강 남구 논 현로 640	2020 년 01월	SK 에 너 지	셀 프	1802.00	1657.00	1495.00	0.0	서 울	강 남 구
	2	A0006039	서 울 강 남 구	(유)동하 석유 힐탑 셀프주유 소	서울 강 남구 논 현로 640	2020 년 02월	SK 에 너 지	셀 프	1795.38	1637.97	1483.97	0.0	서 울	강 남 구
	3	A0006039	서 울 강 남 구	(유)동하 석유 힐탑 셀프주유 소	서울 강 남구 논 현로 640	2020 년 03월	SK 에 너 지	셀 프	1741.26	1570.61	1431.13	0.0	서 울	강 남 구
	4	A0006039	서 울 남 구	(유)동하 석유 힐탑 셀프주유 소	서울 강 남구 논 현로 640	2020 년 04월	SK 에 너 지	셀 프	1617.00	1425.33	1290.00	0.0	서 울	강 남 구
	5	A0006039	서 울 강 남 구	(유)동하 석유 힐탑 셀프주유 소	서울 강 남구 논 현로 640	2020 년 05월	SK 에 너 지	셀 프	1565.00	1350.81	1197.74	0.0	서 울	강 남 구

	번호	지 역	상호	주소	기간	상 표	셀 프 여 부	고급휘 발유	휘발유	경유	실 내 등 유	시	구
2946	A0009197	서 울 중 랑 구	현대오일 뱅크㈜직 영 중랑교 셀프주유 소	서울 중 랑구 망 우로 229 (중 화동)	2020 년 08월	현대오일뱅크	셀 프	1598.00	1402.65	1211.10	0.0	서 울	중 랑 구
2947	A0009197	서 울 중 랑 구	현대오일 뱅크㈜직 영 중랑교 셀프주유 소	서울 중 랑구 망 우로 229 (중 화동)	2020 년 09월	현 대 오 일 뱅 크	셀 프	1598.00	1393.00	1203.00	0.0	서 울	중 랑 구
2948	A0009197	서 울 중 랑 구	현대오일 뱅크㈜직 영 중랑교 셀프주유 소	서울 중 랑구 망 우로 229 (중 화동)	2020 년 10월	현 대 오 일 뱅 크	셀 프	1598.00	1338.42	1148.42	0.0	서 울	중 랑 구
2949	A0009197	서 울 중 랑 구	현대오일 뱅크㈜직 영 중랑교 셀프주유 소	서울 중 랑구 망 우로 229 (중 화동)	2020 년 11월	현 대 오 일 뱅 크	셀 프	1598.00	1292.00	1092.37	0.0	서 울	중 랑 구
2950	A0009197	서 울 중 랑 구	현대오일 뱅크㈜직 영 중랑교 셀프주유 소	서울 중 랑구 망 우로 229 (중 화동)	2020 년 12월	현 대 오 일 뱅 크	셀 프	1598.00	1367.58	1167.58	0.0	서 울	중 랑 구

5944 rows × 13 columns

[데이터 구조 변경]

In [8]: # 시 컬럼의 서울을 서울특별시로 변환

df['시'] += '특별시'
df.head()

Out[8]: 셀 실내등유 상 표 프 고급휘 번호 상호 주소 기간 경유 시 구 여 부 서 서 SK · 울 강 강 울 특 별 (유)동하석 서울 강 2020 에 셀 남 **1** A0006039 유 힐탑셀 남구 논 년 01 1802.00 1657.00 1495.00 0.0 프 너 남 구 프주유소 현로 640 지 시

	번호	지 역	상호	주소	기간	상 표	셀 프 여 부	고급휘 발유	휘발유	경유	실 내 등 유	시	구
2	A0006039	서 울 강 남 구	(유)동하석 유 힐탑셀 프주유소	서울 강 남구 논 현로 640	2020 년 02 월	SK 에 너 지	셀 프	1795.38	1637.97	1483.97	0.0	서 울 특 별 시	강 남 구
3	A0006039	서 울 남 구	(유)동하석 유 힐탑셀 프주유소	서울 강 남구 논 현로 640	2020 년 03 월	SK 에 너 지	셀 프	1741.26	1570.61	1431.13	0.0	서 울 특 별 시	강 남 구
4	A0006039	서 울 남 구	(유)동하석 유 힐탑셀 프주유소	서울 강 남구 논 현로 640	2020 년 04 월	SK 에 너 지	셀 프	1617.00	1425.33	1290.00	0.0	서 울 특 별 시	강 남 구
5	A0006039	서 울 남 구	(유)동하석 유 힐탑셀 프주유소	서울 강 남구 논 현로 640	2020 년 05 월	SK 에 너 지	셀 프	1565.00	1350.81	1197.74	0.0	서 울 특 별 시	강 남 구

<Step4.>: 데이터 모델링 및 시각화

[요인분석1:지역구]

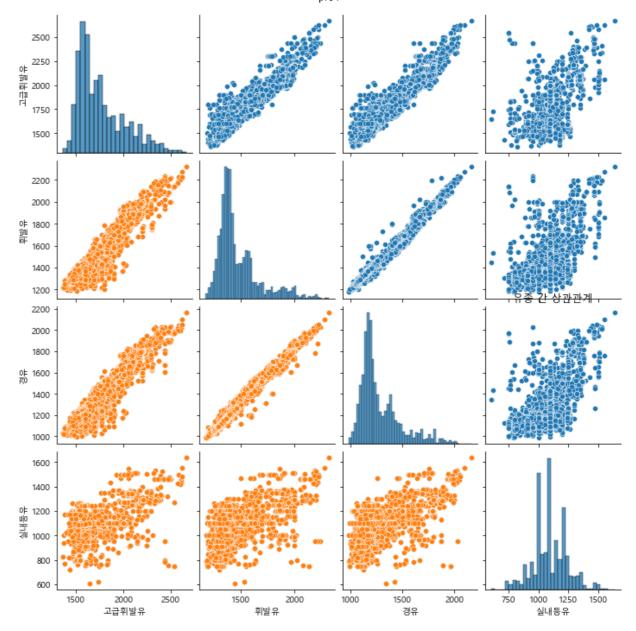
1.1 유종 간 상관관계 _ (Pairplot Chart)

```
df_nonull= df.replace(0, np.NaN)
df_nonull

import matplotlib.pyplot as plt
import matplotlib.font_manager as fm

fm.get_fontconfig_fonts()
font_location = 'C:\text{WWindows\text{WFonts\text{Wmalgun.ttf'}}}
font_name = fm.FontProperties(fname=font_location).get_name()
plt.rc('font', family=font_name)

graph = sns.pairplot(df_nonull)
graph.map_lower(sns.scatterplot)
graph.map_upper(sns.regplot, scatter=False, truncate=False, ci=False)
plt.title("유종 간 상관관계")
plt.show()
```



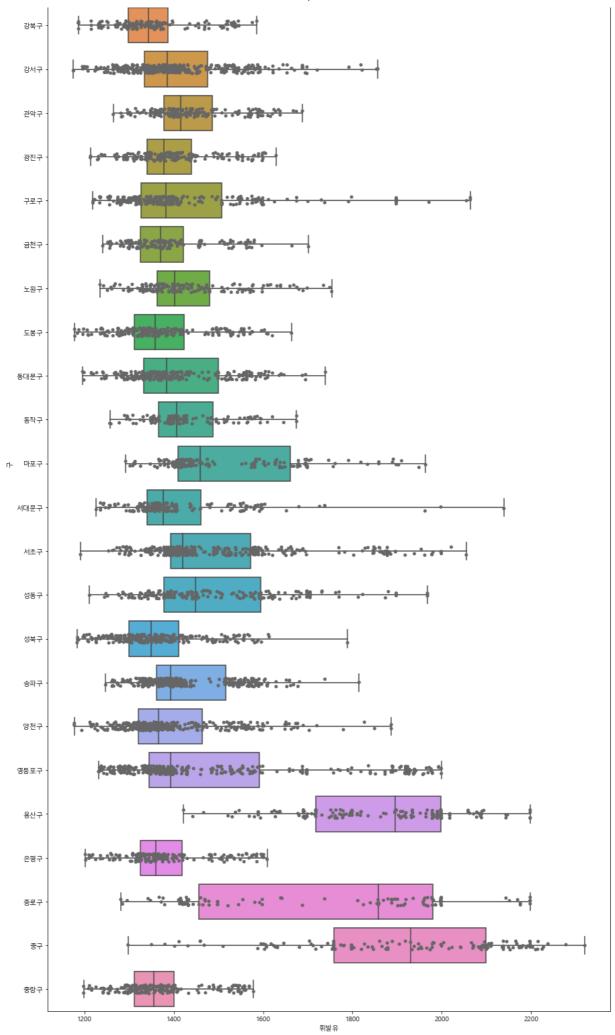
분석결과

지역구 별 유가를 비교하기 전에 유가별 상관관계를 파악하여 가장 대표적인 유종이 무엇인지 알아보았다.

> 분석결과 : 모든 유종은 양의 상관관계를 가지고있으며 그 중에서도 휘발유와 경유의 상관관계가 제일 높다. 따라서, 편의상 휘발유를 대표적 유종으로 선정하였다.

1.2 지역구별 유가 _ (Boxplot)

```
plt.figure(figsize=(15, 30))
plt.title("지역구별 휘발유가")
sns.boxplot(x="휘발유", y="구", data=df, whis=np.inf)
sns.stripplot(x="휘발유", y="구", data=df, jitter=True, color="0.4")
plt.show()
```



분석결과

대표 유종인 휘발유의 지역구별 가격을 알아보았다.

> 분석결과 : Boxplot으로 알 수 있는 것은 중앙값, 산포도, 최대값, 최소값, 이상치 등이 있다.

위 그래프를 해석하면, 휘발유의 평균가격이 가장 높은 지역구는 중구, 용산구, 종로구 순이고 가장 낮은 지역구는 강북구,성북구, 중랑구임을 알 수 있다.

특히, 중구의 그래프를 보면 2,3,4분위수 범위의 간격이 비교적 일정한 것으로보아 중구의 높은 휘발유 평균가격은 이상치에 의해 좌우된 결과가 아님을 알 수있다.

따라서, 해당 그래프는 유의미하다.

1.3 지역구별 유가 _ (Barplot)

df_a = df.replace(0, np.NaN) # 0을 제외하고 평균을 내기 위해 nan처리 df_a = df_a.groupby('구')[['고급휘발유', '휘발유', '경유', '실내등유']].mean() # df의 df_a

Out[11]:		고급휘발유	휘발유	경유	실내등유
	구				
	강남구	1820.577020	1601.450148	1426.829409	1250.738244
	강동구	1742.760476	1483.317188	1297.847500	1082.447983
	강북구	1740.838125	1358.062500	1168.330962	942.135278
	강서구	1616.750072	1407.981343	1218.795721	1068.254754
	관악구	1678.794375	1440.764404	1266.050363	1087.529294
	광진구	1633.337292	1397.081373	1202.645196	1065.067746
	구로구	1588.155789	1429.092710	1226.033435	1070.564000
	금천구	1533.020625	1390.482484	1196.913248	1065.275529
	노원구	1661.301000	1434.355371	1246.633600	1066.436341
	도봉구	1669.422292	1378.404167	1190.656389	993.692963
	동대문구	1616.753056	1408.674144	1215.942824	1088.621518
	동작구	1600.552453	1430.096083	1264.556583	1060.882778
	마포구	1835.779714	1533.331408	1356.522676	1161.473958
	서대문구	1708.995488	1413.302099	1223.311547	1140.450896
	서초구	1712.043105	1490.589930	1324.340634	1155.363154
	성동구	1758.545816	1492.690099	1307.588218	1165.003548
	성북구	1666.202018	1362.384425	1173.994390	1034.432033
	송파구	1595.915313	1423.959394	1242.611439	1125.365122
	양천구	1637.812301	1401.761962	1209.962310	1083.694105
	영등포구	1804.245648	1497.842619	1313.364603	1122.859420

고급휘발유 휘발유 경유 실내등유 구 2133.165179 1858.926071 1709.826726 1154.484000 은평구 1560.113553 1382.617371 1192.846031 966.631096 종로구 2036.235000 1746.083853 1571.764495 1146.844706 중구 2194.344646 1904.910216 1718.084029 1197.343226 중랑구 1618.713793 1370.667396 1180.817396 1011.434426

In [12]

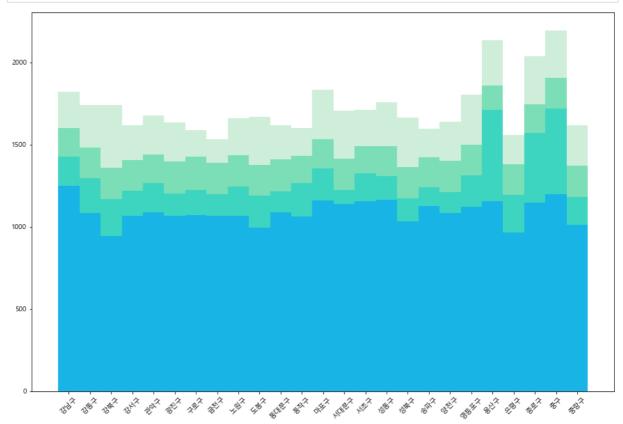
df_a.reset_index(inplace=True)
df_a

구 고급휘발유 휘발유 경유 실내등유 강남구 1820.577020 1601.450148 1426.829409 1250.738244 0 강동구 1742.760476 1483.317188 1297.847500 1082.447983 1 강북구 1740.838125 1358.062500 1168.330962 2 942.135278 3 강서구 1616.750072 1407.981343 1218.795721 1068.254754 관악구 1678.794375 1440.764404 1266.050363 1087.529294 4 광진구 1633.337292 1397.081373 1202.645196 1065.067746 5 구로구 1588.155789 1429.092710 1226.033435 1070.564000 6 금천구 1533.020625 1390.482484 1196.913248 1065.275529 8 노원구 1661.301000 1434.355371 1246.633600 1066.436341 도봉구 1669.422292 1378.404167 1190.656389 993.692963 **10** 동대문구 1616.753056 1408.674144 1215.942824 1088.621518 11 동작구 1600.552453 1430.096083 1264.556583 1060.882778 마포구 1835.779714 1533.331408 1356.522676 1161.473958 12 서대문구 1708.995488 1413.302099 1223.311547 1140.450896 13 서초구 1712.043105 1490.589930 1324.340634 1155.363154 14 15 성동구 1758.545816 1492.690099 1307.588218 1165.003548 16 성북구 1666.202018 1362.384425 1173.994390 1034.432033 17 송파구 1595.915313 1423.959394 1242.611439 1125.365122 양천구 1637.812301 1401.761962 1209.962310 1083.694105 18 19 영등포구 1804.245648 1497.842619 1313.364603 1122.859420 20 용산구 2133.165179 1858.926071 1709.826726 1154.484000 21 은평구 1560.113553 1382.617371 1192.846031 966.631096 종로구 2036.235000 1746.083853 1571.764495 1146.844706 22 23 중구 2194.344646 1904.910216 1718.084029 1197.343226 중랑구 1618.713793 1370.667396 1180.817396 1011.434426 24

```
plt.figure(figsize=(15, 10))
plt.xticks(rotation=45)

plt.bar(df_a['구'], df_a['고급휘발유'], alpha = 0.3, width=1, color='#62c983')
plt.bar(df_a['구'], df_a['휘발유'], alpha = 0.4, width=1, color="#00c786")
plt.bar(df_a['구'], df_a['경유'], alpha = 0.5, width=1, color="#00cec9")
plt.bar(df_a['구'], df_a['실내등유'], alpha = 0.6, width=1, color="#009fff")

plt.show()
```



분석결과

지역별 전체 유가 평균을 살펴보았다.

> 분석결과 : Boxplot으로 봤을 때와 같이 평균가격이 가장 높은 지역구는 중구, 용산구, 종로구 순이고 가장 낮은 지역구는 강북구,성북구, 중랑구임을 알 수 있 다.또한 지역구별로 유종간에 상관관계가 있음도 확인할 수 있었다.

1.4 지역구별 월별 유가 (Heatmap)

```
ln [21]:

df1=df.replace(0,np.NaN)
 df1=df1.groupby(['구','기간'])[['고급휘발유','휘발유','경유','실내등유']].mean()
 df1
```

Out[21]:			고급휘발유	휘발유	경유	실내등유
	구	기간				
	강남구	2020년 01월	1983.367059	1771.252750	1620.019000	1332.228462
		2020년 02월	1968.100294	1754.086750	1600.688750	1332.799167
		2020년 03월	1928.760588	1703.356500	1545.263750	1313.131818
		2020년 04월	1815.537941	1580.068250	1418.166250	1230.497692

		고급휘말유	휘말유	경유	실내능유
구	기간				
	2020년 05월	1740.169118	1504.057250	1335.514000	1161.004545
•••	•••				
중랑구	2020년 08월	1570.570000	1364.416875	1165.096875	995.800000
	2020년 09월	1564.000000	1349.686250	1151.018125	995.000000
	2020년 10월	1560.130000	1326.126250	1126.568750	982.742000
	2020년 11월	1555.600000	1306.229375	1107.904375	966.500000
	2020년 12월	1575.270000	1371.651875	1171.588750	987.452000

300 rows × 4 columns

```
ln [22]:

df2=pd.pivot_table(df1,index="구",columns='기간', values="고급휘발유")

df3=pd.pivot_table(df1,index="구",columns='기간', values="휘발유")

df4=pd.pivot_table(df1,index="구",columns='기간', values="경유")

df5=pd.pivot_table(df1,index="구",columns='기간', values="경유")

plt.figure(figsize=(12, 12))

sns.heatmap(df2,annot=True,fmt=".Of",linewidths=1,cmap='Blues')

plt.title('지역구 월별 고급휘발유 가격 비교')

plt.show()
```

지역구 월별 고급휘발유 가격 비교

지역구 월별 고급휘발유 가격 비교														
강남구 -	1983	1968	1929	1816	1740	1765	1786	1792	1765	1764	1761	1776		
강동구 -	1930	1896	1826	1693	1631	1698	1732	1711	1709	1696	1690	1701		
강북구 -	1822	1806	1784	1724	1686	1730	1734	1734	1734	1734	1734	1665		
강서구 -	1747	1730	1679	1588	1541	1576	1602	1600	1589	1576	1560	1598		- 2200
관악구 -	1842	1832	1790	1671	1567	1596	1638	1648	1652	1646	1634	1631		
광진구 -	1795	1774	1713	1607	1565	1621	1635	1617	1568	1556	1557	1592		
구로구 -	1740	1723	1672	1563	1499	1519	1555	1572	1566	1561	1543	1568		
금천구 -	1709	1669	1605	1482	1433	1489	1530	1542	1536	1491	1482	1525		
노원구 -	1797	1774	1722	1640	1601	1642	1650	1640	1621	1610	1608	1630		- 2000
도봉구 -	1839	1818	1731	1572	1534	1623	1652	1658	1658	1645	1639	1664		- 2000
동대문구 -	1790	1766	1692	1560	1512	1583	1625	1600	1572	1554	1552	1594		
동작구 -	1781	1763	1711	1592	1516	1522	1557	1584	1571	1552	1541	1564		
다 마포구 -	2103	2090	1975	1799	1708	1754	1783	1792	1790	1769	1766	1786		
서대문구 -	1933	1893	1848	1710	1616	1647	1671	1671	1659	1612	1598	1625		
서초구 -	1868	1858	1819	1704	1630	1646	1671	1681	1677	1661	1652	1668		- 1800
성동구 -	1956	1945	1873	1734	1669	1715	1733	1731	1729	1715	1686	1706		
성북구 -	1810	1796	1743	1633	1590	1603	1618	1641	1646	1643	1635	1645		
송파구 -	1760	1755	1718	1577	1489	1529	1558	1561	1559	1556	1554	1567		
양천구 -	1841	1815	1743	1597	1523	1560	1588	1612	1616	1606	1582	1608		
영등포구 -	1904	1900	1861	1772	1729	1770	1795	1802	1784	1775	1765	1795	,	- 1600
용산구 -	2254	2257	2236	2118	2059	2095	2109	2097	2098	2096	2086	2092		
은평구 -	1735	1708	1629	1470	1406	1522	1557	1568	1546	1527	1515	1555		
종로구 -	2168	2158	2130	1994	1939	1997	2012	2014	2016	2016	1992	1998		
중구 -	2330	2309	2269	2148	2085	2121	2150	2176	2197	2192	2184	2165		
중랑구 -	1848	1814	1725	1581	1536	1605	1624	1571	1564	1560	1556	1575		
	2020년 01월 -	2020년 02월 -	2020년 03월 -	2020년 04월	2020년 05월 -	- 890 년 06월 -	다 2020년 07월 -	2020년 08월 -	2020년 09월 -	2020년 10월 -	2020년 11월 -	2020년 12월 -		

```
plt.figure(figsize=(12, 12))
sns.heatmap(df3,annot=True,fmt=".0f",linewidths=1,cmap='Blues')
plt.title('지역구 월별 휘발유 가격 비교')
plt.show()
```

지역구 월별 휘발유 가격 비교

강남구 -	1771	1754	4700										시역구 월열 위말유 가격 비교												
		1/54	1703	1580	1504	1541	1570	1574	1554	1550	1543	1564													
강동구 -	1647	1619	1556	1419	1356	1423	1466	1473	1475	1456	1443	1467		- 2000											
강북구 ·	1538	1514	1438	1271	1216	1312	1355	1357	1346	1313	1283	1354		2000											
강서구 -	1590	1561	1483	1345	1289	1361	1389	1391	1383	1360	1341	1396													
관악구 -	1608	1587	1534	1406	1315	1379	1419	1420	1418	1403	1386	1416		- 1900											
광진구 -	1572	1542	1464	1331	1266	1347	1385	1390	1379	1353	1341	1393													
구로구 -	1616	1595	1521	1383	1311	1381	1411	1409	1394	1371	1361	1409													
금천구 -	1576	1539	1475	1341	1271	1345	1383	1391	1374	1320	1304	1371		- 1800											
노원구 -	1619	1591	1521	1380	1314	1380	1415	1408	1400	1386	1372	1411													
도봉구 -	1557	1529	1446	1289	1231	1336	1379	1381	1368	1335	1317	1374													
동대문구 -	1588	1563	1495	1355	1268	1359	1402	1401	1383	1354	1339	1395		- 1700											
동작구 -	1600	1574	1508	1381	1307	1368	1413	1421	1414	1390	1377	1409													
다 마포구	1752	1746	1691	1531	1408	1449	1479	1484	1486	1471	1460	1480													
서대문구 -	1623	1585	1515	1358	1272	1363	1400	1401	1383	1354	1336	1376		- 1600											
서초구 -	1653	1636	1580	1456	1381	1423	1458	1465	1468	1456	1441	1468													
성동구 -	1659	1644	1583	1433	1368	1435	1475	1477	1476	1454	1432	1473													
성북구 -	1562	1512	1425	1267	1217	1316	1364	1365	1351	1319	1293	1360		- 1500											
송파구 -	1584	1572	1529	1383	1302	1364	1405	1404	1390	1378	1372	1405													
양천구 -	1593	1561	1478	1338	1273	1350	1379	1390	1384	1357	1337	1390													
영등포구 -	1651	1626	1554	1440	1401	1464	1492	1492	1471	1453	1438	1487		- 1400											
용산구 -	1988	1986	1963	1835	1766	1807	1827	1833	1835	1828	1816	1823													
은평구 -	1563	1531	1450	1293	1243	1339	1373	1377	1360	1342	1328	1369													
종로구 -	1898	1875	1830	1685	1619	1710	1724	1730	1728	1727	1709	1721		- 1300											
중구 -	2063	2048	1999	1849	1803	1830	1861	1887	1893	1882	1867	1864													
중랑구 -	1548	1522	1437	1291	1238	1327	1366	1364	1350	1326	1306	1372													
	2020년 01월 -	2020년 02월 -	2020년 03월 -	2020년 04월 -	2020년 05월 -	- 890 의 주 기	- 룡20 위0207 가	2020년 08월 -	2020년 09월 -	2020년 10월 -	2020년 11월 -	2020년 12월 -													

```
plt.figure(figsize=(12, 12))
sns.heatmap(df4,annot=True,fmt=".0f",linewidths=1,cmap='Blues')
plt.title('지역구 월별 경유 가격 비교')
plt.show()
```

지역구 월별 경유 가격 비교

					시억-		경유 가격	0176					
강남구 -	1620	1601	1545	1418	1336	1362	1386	1392	1369	1362	1352	1370	
강동구 -	1494	1458	1380	1238	1168	1233	1273	1277	1280	1259	1245	1269	
강북구 -	1381	1351	1256	1075	1024	1119	1159	1159	1146	1111	1083	1154	
강서구 -	1417	1389	1298	1156	1099	1167	1196	1199	1187	1163	1149	1200	
관악구 -	1450	1425	1366	1236	1152	1205	1237	1239	1235	1220	1201	1230	
광진구 -	1400	1367	1286	1144	1073	1145	1184	1187	1176	1149	1134	1187	
구로구 -	1435	1406	1323	1184	1104	1173	1203	1202	1187	1163	1150	1199	
금천구 -	1402	1364	1288	1145	1076	1145	1182	1193	1176	1121	1105	1172	
노원구 -	1457	1419	1342	1196	1136	1191	1218	1210	1202	1187	1172	1212	
도봉구 -	1395	1369	1271	1101	1043	1138	1179	1184	1173	1139	1122	1176	
동대문구 -	1407	1382	1311	1161	1080	1162	1202	1203	1186	1155	1140	1194	
동작구 -	1457	1425	1351	1217	1149	1198	1238	1250	1242	1216	1200	1232	
마 마포구 -	1603	1598	1543	1381	1244	1259	1284	1290	1291	1277	1265	1284	
서대문구 -	1455	1422	1339	1169	1084	1165	1200	1201	1184	1154	1136	1176	
서초구 -	1514	1499	1441	1302	1214	1249	1280	1286	1286	1273	1256	1285	
성동구 -	1491	1472	1406	1259	1188	1246	1282	1284	1281	1261	1237	1278	
성북구 -	1398	1339	1241	1076	1035	1127	1173	1176	1155	1115	1091	1164	
송파구 -	1427	1410	1356	1197	1126	1178	1214	1214	1204	1192	1184	1211	
양천구 -	1422	1386	1289	1149	1085	1157	1181	1190	1184	1158	1138	1190	
영등포구 -	1491	1462	1382	1260	1208	1274	1301	1302	1280	1259	1243	1293	
용산구 -	1844	1838	1810	1699	1628	1653	1675	1681	1681	1673	1664	1672	
은평구 -	1398	1364	1269	1110	1058	1142	1173	1178	1160	1141	1127	1169	
종로구 -	1743	1723	1679	1563	1480	1519	1532	1536	1533	1531	1511	1518	
중구 -	1880	1846	1789	1667	1612	1650	1680	1704	1708	1697	1687	1685	
중랑구 -	1380	1347	1256	1109	1053	1135	1167	1165	1151	1127	1108	1172	
	2020년 01월 -	2020년 02월 -	2020년 03월 -	2020년 04월 -	2020년 05월 -	2020년 06월 -	- അ.40 리020Z 간	2020년 08월 -	2020년 09월 -	2020년 10월 -	2020년 11월 -	2020년 12월 -	

```
plt.figure(figsize=(12, 12))
sns.heatmap(df5,annot=True,fmt=".0f",linewidths=1,cmap='Blues')
plt.title('지역구 월별 실내등유 가격 비교')
plt.show()
```

지역구 월별 실내등유 가격 비교

시역구 월일 실내용유 가격 비교														
강남구 -	1332	1333	1313	1230	1161	1208	1238	1238	1203	1227	1235	1254		
강동구 -	1190	1182	1163	1105	1052	1039	1034	1043	1043	1043	1043	1047		
강북구 -	1067	1067	1063	985	883	883	881	883	884	900	900	910		
강서구 -	1130	1140	1124	1083	1042	1032	1034	1031	1031	1036	1040	1057		
관악구 -	1160	1160	1148	1107	1071	1069	1071	1071	1061	1048	1043	1046		- 1300
광진구 -	1133	1131	1104	1103	1055	1022	1022	1022	1026	1029	1053	1047		
구로구 -	1165	1159	1124	1060	1018	1037	1048	1044	1044	1051	1044	1056		
금천구 -	1104	1104	1098	1088	1049	1016	1036	1038	1038	1064	1039	1060		
노원구 -	1156	1153	1125	1060	1030	1044	1037	1037	1033	1032	1025	1039		
도봉구 -	1077	1066	1047	984	965	984	984	984	982	965	940	946		- 1200
동대문구 -	1172	1172	1131	1108	1051	1046	1063	1091	1090	1046	1028	1066		
동작구 -	1090	1090	1073	1030	1020	1054	1063	1063	1063	1061	1060	1063		
마 마포구 -	1394	1337	1257	1083	1025	1106	1118	1127	1109	1112	1121	1153		
서대문구 -	1150	1150	1150	1164	1125	1130	1136	1134	1136	1130	1143	1144		
서초구 -	1198	1193	1177	1138	1113	1129	1149	1149	1186	1149	1136	1139		- 1100
성동구 -	1229	1220	1207	1138	1108	1141	1158	1144	1144	1157	1143	1155		
성북구 -	1169	1135	1097	1001	952	998	998	998	996	994	980	1026		
송파구 -	1180	1175	1162	1130	1104	1100	1098	1115	1107	1093	1109	1114		
양천구 -	1122	1123	1120	1084	1071	1089	1062	1071	1071	1066	1054	1063		
영등포구 -	1186	1164	1167	1122	1089	1094	1092	1110	1108	1111	1107	1117	,	- 1000
용산구 -	1301	1283	1219	1090	1072	1113	1129	1131	1128	1122	1123	1140		
은평구 -	1037	1047	1001	938	890	952	952	952	949	946	944	979		
종로구 -	1276	1259	1225	1148	1080	1085	1108	1108	1108	1108	1108	1108		
중구 -	1310	1297	1297	1224	1148	1143	1116	1164	1128	1122	1129	1196		
중랑구 -	1053	1058	1046	1032	1006	994	1005	996	995	983	966	987	,	- 900
	2020년 01월 -	2020년 02월 -	2020년 03월 -	2020년 04월 -	2020년 05월 -	- 890 의 5020기	- 룡20 위0207	2020년 08월 -	2020년 09월 -	2020년 10월 -	2020년 11월 -	2020년 12월 -		

분석결과

지역과 유종, 유가간의 관계를 알아보았으니 유가에 월별 영향이 미치는지 알아보았다.

> 분석결과 : 월별 유가를 비교해보니 각 월이 유가에 미치는 영향은 적은 것을 확인 할 수 있었다. 따라서 위 그래프는 무의미하다고 봤다.

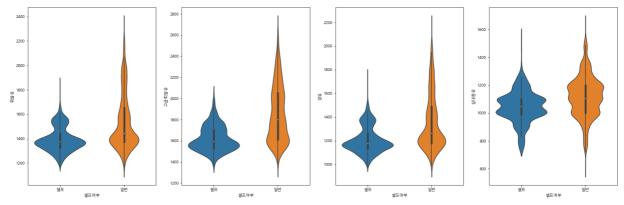
[요인분석2:인건비]

2.1 인건비와 유가의 상관관계 _ (Violinplot)

```
# 한글처리 import matplotlib.pyplot as plt import matplotlib.font_manager as fm fm.get_fontconfig_fonts() font_location = 'C:\text{WWindows\text{WFonts\text{Wmalgun.ttf'}}}
```

```
font_name = fm.FontProperties(fname=font_location).get_name()
plt.rc('font', family=font_name)
```

df0 = df.replace(0, np.NaN)
fig, ax = plt.subplots(figsize=(26, 8), ncols=4)
sns.violinplot(data=df0,x='셀프여부', y='휘발유', ax=ax[0])
sns.violinplot(data=df0,x='셀프여부', y='고급휘발유', ax=ax[1])
sns.violinplot(data=df0,x='셀프여부', y='경유', ax=ax[2])
sns.violinplot(data=df0,x='셀프여부', y='실내등유', ax=ax[3])
plt.show()



분석결과

인건비가 유가에 미치는 영향을 알아보았다.

> 분석결과 : 대부분의 지역에서는 유종에 상관없이 셀프주유소의 가격이 일반 주유소의 가격보다 저렴하다는 것을 알 수 있다.

2.2 지역구별 인건비와 유가의 상관관계 (Violinplot)

fig, ax = plt.subplots(figsize=(120, 8),ncols=4)
sns.violinplot(data=df, hue='셀프여부',x='구', y='휘발유', split=True, ax=ax[0])
sns.violinplot(data=df, hue='셀프여부',x='구', y='고급휘발유', split=True, ax=ax[1])
sns.violinplot(data=df, hue='셀프여부',x='구', y='경유', split=True, ax=ax[2])
sns.violinplot(data=df, hue='셀프여부',x='구', y='실내등유', split=True, ax=ax[3])
plt.show()

분석결과

지역별로 셀프 주유소와 일반 주요소의 가격 차이를 알아보았다.

> 분석결과 : 대부분의 지역에서는 셀프 주유소의 가격이 일반 주유소의 가격보다 저렴했다.

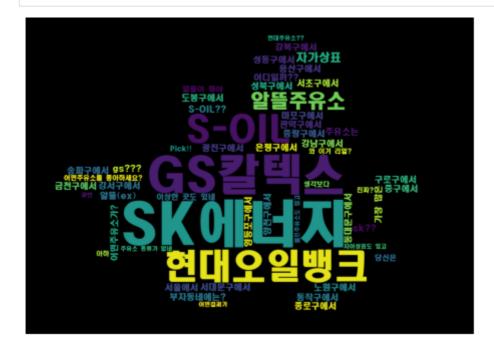
하지만, 일부 지역 (ex: 강북구, 광진구, 금천구 등)에서는 인건비의 영향이 무의미하기도 했다.

따라서, 인건비는 유가에 영향을 미치는 결정적인 요소는 아니라고 판단된다.

[요인분석3: 대리점 공급가]

3.1 지역별 대리점의 수 _ (Wordcloud)

```
In [37]:
         # 서울시 전체 상호 중복
         part = df
         part.dropna(inplace=True)
         seoul=part.drop_duplicates(['상호'], keep='first')
         #서울 상표 카운트
         s_freq = seoul.groupby(['상표']).count()
         s_count = s_freq.loc[:,['지역']]
         s_count
                    지역
               상표
            GS칼텍스
                     133
              S-OIL
                     76
            SK에너지
                     190
            알뜰(ex)
                      1
          알뜰주유소
                      11
            자가상표
                      2
         현대오일뱅크
                      80
         sdict0 = s count['지역'].to dict()
         추가 = {"가장 많은":1, "주유소는":1, "어디일까??":1, "부자동네에는?":1, "어떤주유소가?
                   'gs???":1, "sk??":1, "S-OIL??":1, "서울에서":1, "강남구에서":1, "강북구에서":
               "강서구에서":1, "관악구에서":1, "광진구에서":1, "구로구에서":1, "금천구에서":1, "도봉구에서":1, "동대문구에서":1, "동작구에서":1, "마포구에서":1, "서대문구에서"
"성동구에서":1, "성북구에서":1, "송파구에서":1, "양천구에서":1, "영등포구에서":1
               "은평구에서":1, "종로구에서":1, "중구에서":1, "중랑구에서":1, "이상한 곳도 있네"
                  "당신은":1, "어떤주유소를 좋아하세요?":1, "Pick!!":1, "와 이거 리얼?":1, "진기
                  "생각보다 ":1. "주유소 종류가 없네":1, "현대주유소??":1, "셀프주유소도 많고";
                  "과연":1, "어떤결과가":1}
         sdict0.update(추가)
In [40]:
         from wordcloud import WordCloud
         from PIL import Image
         import numpy as np
         imgmask = np.array(Image.open("img\seoul.png"))
         keywords = sdict0
         wc = WordCloud(font_path="C:\Windows\Fonts\HMKMRHD.ttf",
                                 width=800.
                                 height=600.
                                 background_color='black',
                                 mask = imgmask
         cloud=wc.generate_from_frequencies(keywords)
         plt.figure(figsize=(8, 8))
         plt.imshow(cloud)
         plt.axis("off")
         plt.show()
```



분석결과

지역별 대리점의 수를 알아보았다.

3.2 지역구별 대리점의 수와 유가 _ (Plot)

```
| df_m = df.replace(0, np.NaN) # 0을 제외하고 평균을 내기 위해 nan처리 df_m = df_m.groupby('상표')[['휘발유']].mean() # df의 mean()은 nan 제외한 평균을 내줌 df_m
```

```
Out [14]: 휘발유
```

상표

GS칼텍스 1480.077478

S-OIL 1434.653561

SK에너지 1517.494476

알뜰(ex) 1330.236667

알뜰주유소 1355.216043

자가상표 1414.056800

현대오일뱅크 1427.582881

In [15]:

추가됨

df_m.sort_values(by='휘발유').drop(['알뜰(ex)', '알뜰주유소', '자가상표'])

Out [15]: **휘발유**

상표

현대오일뱅크 1427.582881

S-OIL 1434.653561

GS칼텍스 1480.077478

휘발유

상표

SK에너지 1517.494476

```
ln [16]:

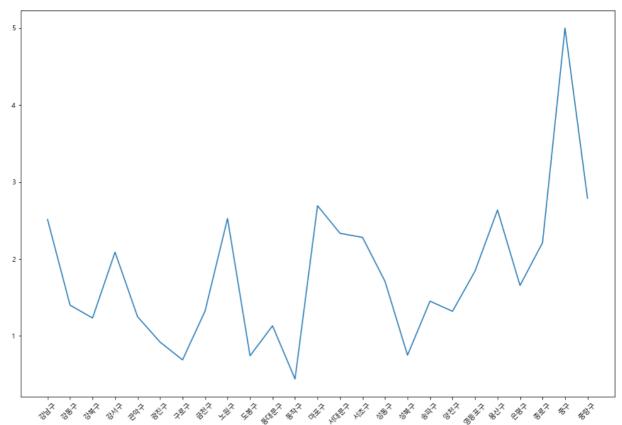
df_k = df.pivot_table('휘발유', index='구', columns='상표', aggfunc='count').dropna(t df_k.fillna(0, inplace=True) df_k
```

```
상표 GS칼텍스 S-OIL SK에너지 현대오일뱅크
     구
 강남구
             143.0
                     71.0
                                            74.0
                              186.0
 강동구
              60.0
                     24.0
                               63.0
                                            45.0
 강북구
              36.0
                     24.0
                               53.0
                                            43.0
 강서구
              84.0
                                            68.0
                     72.0
                              142.0
 관악구
              45.0
                     24.0
                               60.0
                                            48.0
 광진구
              60.0
                     36.0
                                            50.0
                               46.0
 구로구
              48.0
                     72.0
                               53.0
                                            77.0
 금천구
              34.0
                     36.0
                               41.0
                                            31.0
 노원구
              72.0
                                            19.0
                     36.0
                               48.0
 도봉구
              36.0
                     72.0
                               46.0
                                            62.0
동대문구
              68.0
                     33.0
                               86.0
                                            76.0
 동작구
              36.0
                     12.0
                               22.0
                                            50.0
 마포구
              12.0
                     34.0
                               70.0
                                            26.0
서대문구
              35.0
                     36.0
                               77.0
                                            33.0
 서초구
             167.0
                     60.0
                              130.0
                                            57.0
 성동구
              36.0
                     24.0
                               82.0
                                            48.0
 성북구
              84.0
                     58.0
                                            76.0
                               57.0
 송파구
             120.0
                     48.0
                                            93.0
                              135.0
 양천구
              60.0
                     12.0
                              132.0
                                           100.0
영등포구
             119.0
                     31.0
                              140.0
                                            76.0
 용산구
              48.0
                     0.0
                               87.0
                                            33.0
 은평구
              72.0
                     24.0
                               53.0
                                            32.0
 종로구
              24.0
                     24.0
                               42.0
                                            19.0
   중구
              43.0
                     12.0
                               60.0
                                            12.0
 중랑구
              60.0
                     60.0
                               53.0
                                            19.0
```

```
| # df_k = (df_k['SK에너지'] + df_k['GS칼텍스']) / (df_k['S-0|L'] + df_k['현대오일뱅크'] | df_k = df_k['SK에너지'] / df_k['현대오일뱅크'] | df_k
```

```
강남구
                   2.513514
        강동구
                   1.400000
        강북구
                   1.232558
        강서구
                   2.088235
        관악구
                   1.250000
        광진구
                   0.920000
        구로구
                   0.688312
        금천구
                   1.322581
        노원구
                   2.526316
        도봉구
                   0.741935
        동대문구
                    1.131579
        동작구
                   0.440000
        마포구
                   2.692308
        서대문구
                   2.333333
        서초구
                   2.280702
        성동구
                   1.708333
        성북구
                   0.750000
        송파구
                   1.451613
        양천구
                   1.320000
        영등포구
                    1.842105
        용산구
                   2.636364
        은평구
                   1.656250
        종로구
                   2.210526
        중구
                  5.000000
        중랑구
                   2.789474
        dtype: float64
In [18]:
         df_k.sort_values()
        동작구
구로구
                   0.440000
                   0.688312
        도봉구
                   0.741935
        성북구
                   0.750000
        광진구
                   0.920000
        동대문구
                    1.131579
        강북구
                   1.232558
        관악구
                   1.250000
        양천구
금천구
                   1.320000
                   1.322581
        강동구
                   1.400000
        송파구
                   1.451613
        은평구
                   1.656250
        성동구
                   1.708333
        영등포구
                    1.842105
        강서구
                   2.088235
        종로구
                   2.210526
        서초구
                   2.280702
        서대문구
                   2.333333
        강남구
                   2.513514
        노원구
                   2.526316
        용산구
                   2.636364
        마포구
                   2.692308
        중랑구
                   2.789474
        중구
                  5.000000
        dtype: float64
         plt.figure(figsize=(15, 10))
         plt.xticks(rotation=45)
         plt.plot(df_k)
```

Out[19]: [<matplotlib.lines.Line2D at 0x16e3967dee0>]

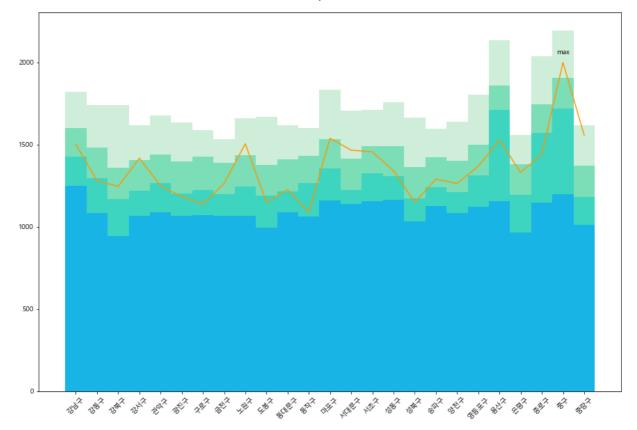


```
In [30]: df_k = df_k * 200 + 1000

In [31]: plt.figure(figsize=(15, 10))
    plt.xticks(rotation=45)
    plt.text(22.7, 2050, f"max")

plt.bar(df_a['구'], df_a['고급휘발유'], alpha = 0.3, width=1, color='#62c983')
    plt.bar(df_a['구'], df_a['취발유'], alpha = 0.4, width=1, color="#00c786")
    plt.bar(df_a['구'], df_a['경유'], alpha = 0.5, width=1, color="#00cec9")
    plt.bar(df_a['구'], df_a['실내등유'], alpha = 0.6, width=1, color="#009fff")
    plt.plot(df_k, color='#ff9900')

plt.show()
```



분석결과

지역별로 공급업체가 유가에 영향을 끼치는지 확인했다.

분석결과> 휘발유 가격이 가장 높은 업체의 수와 가장 낮은 업체의 수로 비율을 구해서 지역별 전체 유가 평균 그래프와 함께 봤다. 유가 평균이 높은 지역구에는 해당 비율이 높게 나오고,유가 평균이 낮은 지역구에는 해당 비율이 낮게 나온다.

따라서, 공급업체가 유가에 영향을 끼친다고 할 수 있다.

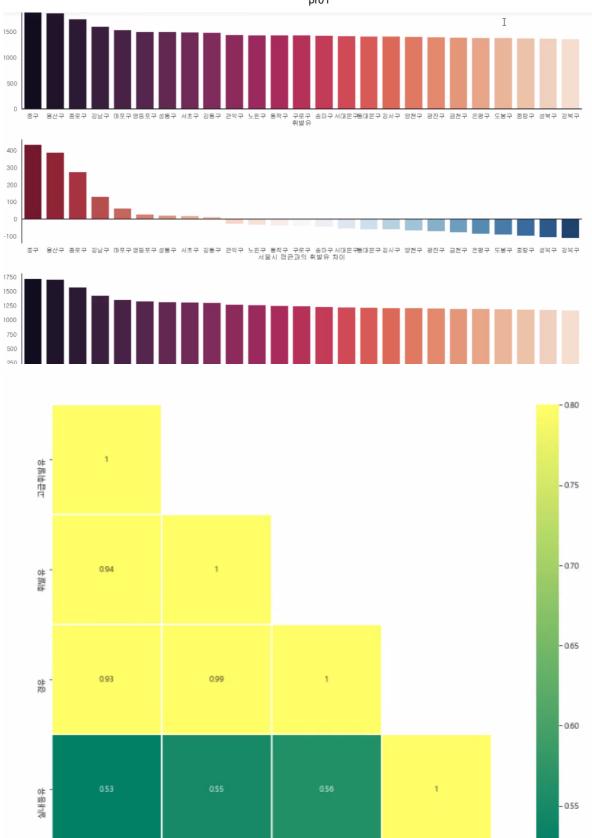
Feedback

- 데이터 전략 잘짰다.
- 데이터 수집: 기업, 공공, 크롤링
- 데이터 전략은 전처리에 많은 영향을 끼친다.
- 강남구가 있는데 중구가 왜 비쌀까 같은 거? 아쉽
- 시각화 여러 그래프로 해서 좋았다.

아쉬운 점

- 우리는 주어진 데이터에 국한되어 분석하는 걸 당연하게 생각했다. 하지만 외부 데이터를 가지고 와서 사용한 팀을 보고, 갇힌 사고를 한 게 아닌가 생각했다.
- 다른 팀 작업방식 중 질문에 대한 답을 찾아낼 때까지 과정을 담아낸 게 좋았던 것 같다.
- 다른 팀들의 발표를 보고 좋았던 시각화 자료들

다른 팀 잘한 거



고급휘발유

휘발유

경유

실내등유