

서울시 숙박업소 밀집도 조사

14 JULY 2023

(산대특)빅데이터 분석 기반 AI 알고리즘 개발 과정

김정훈

목차

01	분석 개요
-	분석의 배경 및 목적
-	분석 범위

02	데이터 구축 및 분석 방법
-	분석 프로세스
-	데이터 수집 방법
-	분석 방법

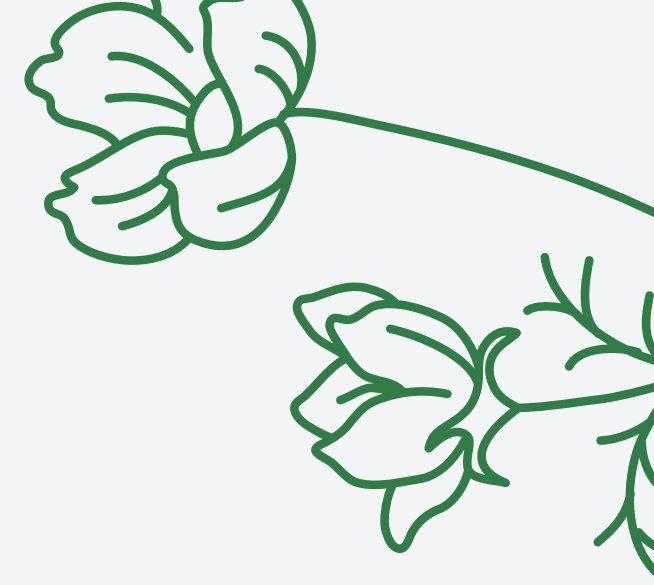
03	분석 결과
04	결론



1. 분석 개요(Analysis Overview)

0%





분석 배경 및 목적

2010년 이후 대한민국 호텔 붐
이 일어나면서 서울에 많은 호텔
들이 건설되었다. 이로 인해 서울
지역권의 호텔 및 호스텔의 밀집
도를 알아보고자 한다.

관련 기사 발췌

MTO 머니투데이 (+) 구독

서울 호텔 건설 '붐'... 5년후엔 2배로?

입력 2012.11.03. 오후 6:39 · 수정 2012.11.03. 오후 6:41 기사원문

공감 10

기사원문

[머니투데이 원종태, 이지혜 기자] 40년간 2만5000실 지었는데 향후 5년간 2만1000실 넘게 건립, "공급과잉" 주의보



서울 종로6가 동대문종합시장 D동상가 옆 주차장 부지는 요즘 별 4개짜리 1급 호텔 'JW 메리어트 동대문'의 신축 공사가 한창이다. 지난달 30일 찾은 이 현장은 이미 터파기 공사를 끝내고, 골조공사를 6층까지 마무리 지었다. 같은 시각 서울 충무로1가 명동 밀리오레 쇼핑몰도 3~17층까지 15개층은 엘리베이터조차 운행하지 않은 채 폐쇄한 상태다. 이곳에도 2급(별 3개) 비즈니스 호텔이 들어설 예정이다.

서울 곳곳에서 호텔 건설이 붐을 이루고 있다. 2일 서울시의 관광숙박시설 건립 추진현황 자료에 따르면 지난 10월15일 기준으로 서울 지역에서 사업승인을 받고 착공을 했거나, 착공을 기다리고 있는 호텔은 총 56곳. 이 호텔들이 모두 완공되는 2015년말이면 8776실의 객실이 더 늘어난다.

◇향후 5년간 신축 호텔이 기존 호텔보다 많을 수도=호텔 러시는 여기서 멈추지 않는다. 지난 10월말 현재 서울시에 호텔을 짓겠다고 사업계획 승인을 신청했거나, 앞으로 신청 예정인 호텔만도 32곳에 달한다. 대부분 2014~2017년까지 집중적으로 완공될 예정으로 추가로 객실 1만 2535실이 늘어난다.

이에 따라 2017년까지 증가하는 예상 객실수만 2만1241실. 현재 서울지역 호텔 객실수의 84% 규모 신규 객실이 추가로 생겨나는 것이다. 지난 40여년간 지어진 서울지역 전체 호텔 객실수에 육박하는 객실이 불과 5년만에 새롭게 탄생한다.

서울지역 호텔 연도별 공급 예정 현황							
자료:개소, 실							
구분	2011.12	2012년	2013년	2014년	2015년	2016년	2017년
증가 호텔 수	-	32	21	19	12	8	5
증가 객실 수	-	4,132	2,762	4,782	3,635	5,008	1,918
고급 객실(저녁)	25,160	26,292	27,054	28,836	30,471	32,479	34,397

연구 범위



서울 정보소통광장
(<https://opengov.seoul.go.kr/>)에
게시되어 있는 “관광숙박업 등록
현황” 2022년 자료를 선정하여 진행

2. 데이터 수집 및 분석 방법 (Data Acquisition and Analysis Methods)

20%



분석 프로세스

파이썬을 활용하여 데이터 가공,
분석에 적합하게 만듦

시각화를 토대로한 종합 결과 확인



데이터 수집



데이터 전처리



탐색적 데이터 분석



데이터 시각화



결과 보고

서울 정보소통광장
([HTTPS://OPENGOV.SEOUL.GO.KR/](https://opengov.seoul.go.kr/))
을 통해 숙박업소 리스트 데이터 수집

전처리 된 데이터를 이용하여
시각화 그래프 제작 및 결과
확인

데이터 수집방법(1)



서울정보 소통광장
([HTTPS://OPENGOV.SEOUL.GO.KR/](https://opengov.seoul.go.kr/))

서울정보소통광장은 결재문서, 사전공개정보 등 법령에서 정한 비공개 사항을 제외한 행정정보를 시민에게 효과적으로 개방하고 공유하기 위한 서울시 정보공개 전용 누리집.

서울시 소관부서에서 결재문서, 업무추진비, 일정공개 등의 사전공개정보와 시정정보를 포함한 다양한 콘텐츠를 언제 어디서나 쉽게 찾아볼 수 있는 서비스를 제공하고 있음.

데이터 수집방법(2)

게 없던 미술 축제 '서울아트위크'가 온다!					서울소
원문정보	사전공개	공개청구	시정정보	통계정보	
결재문서	사전공개정보	정보공개청구 안내	서울시의회 정보	서울의 통계	1
부시장이상 결재문서	업무추진비	비공개 세부기준	건설사업정보	공공데이터	사
정책실명제 문서	일정공개	정보공개관련규정	보도자료	서울의 재정	사
	위원회 회의정보	정보공개청구목록	주요사업백서	서울시조직운영지표	사
	정책연구자료	정보(결재문서)목록	국외훈련보고서	정보공개현황	문
호텔업 등록현황 (22.12.31.기준)					
지역2 (시군구)	업종	호텔명(등록명칭)	영업상태 (22.12.31.기준)	등급	
조리구	관광호텔업	동대문관광호텔	영업중	2선	
	관광호텔업	센추럴	영업중		
	광호텔업	VATICA HOTEL(바티카 호텔)	영업중		

검색창에서 “관광숙박업 등록 현황” 입력.
사전공개>사전공개정보>사전공개문서로 들어가
2022년 4분기 기준 서울시 관광숙박업 등록
(승인)현황 파일 확인 가능

변수 확인

변수명	내용	비고
번호	숙박 업소 순번	서울 정보 소통광장(HTTPS://OPENGOV.SEOUL.GO.KR/) 2022년 4분기 기준 서울시 관광숙박업 등록 (승인)현황파일이용
지역_1(시도)	숙박 업소가 위치한 시	
지역_2(시군구)	숙박 업소가 위치한 행정 구	
업종	해당 숙박 업소의 영업 종류	
호텔명_(등록명칭)	숙박 업소 명칭	
영업상태_(22.12.31)기준	22년 12월 31일 기준 운영 여부	
등급	숙박 업소 등급	
등급부여일	숙박 업소 등급 날짜	
결정기관	숙박 업소 등급 심사 결정 기관	
객실 수	숙박 업소 객실 수	
주소(도로명주소)	숙박 업소 주소	구글맵 URL주소에서 직접 추출 진행
위도	구글 지도상 x좌표	
경도	구글 지도상 y좌표	

분석 방법

✿ PANDAS의 READ_EXCEL함수 이용하여 엑셀 파일 LOAD, 데이터 프레임 생성

✿ 데이터 타입 확인, 결측치 확인, 인덱스 SETTING

✿ 불필요한 행이나 열 제거 후 재배치

✿ 기초통계량 분석

✿ 데이터 시각화(도수분포표, 히스토그램, FOLIUM 지도시각화, 바플롯, 파이차트, SVG 지도시각화)

데이터 전처리

프로젝트를 진행하기 위해 라이브러리 설정 진행

#패키지 참조

```
from pandas import read_excel, DataFrame #표
import numpy # 과학적 계산을 위한 라이브러리
import folium # 지도 시각화를 위한 라이브러리
from matplotlib import pyplot # 데이터 시각
import seaborn # 간단하게 통계적 그래프와 색
from IPython.display import SVG # IPython
from bs4 import BeautifulSoup # 웹 스크래핑
import requests #HTTP 요청을 보내고 응답을 받
```

- Pandas: 데이터 분석과 조작을 위한 라이브러리.
- Numpy: 수치 계산을 위한 라이브러리.
- Folium: 지도 시각화를 위한 라이브러리.
- Matplotlib.pyplot: 데이터 시각화를 위한 라이브러리 중 하나로, PYPLOT 모듈을 사용.
- seaborn: 데이터 시각화를 위한 라이브러리 중 하나로, 통계 기반의 그래프를 생성하는데 도움을 줌.
- IPython.display.SVG: IPYTHON 환경에서 SVG 이미지를 표시하기 위한 모듈.
- bs4 (BEAUTIFULSOUP): 웹 스크래핑과 HTML/XML 파싱을 위한 라이브러리.
- requests: HTTP 요청을 보내고 응답을 받기 위한 라이브러리.

행 번호	지역 1_(시도)	지역 2_(시군구)	업종	호텔명_(등록명칭)
0	1	서울특별시	종로구	관광호텔업 동대문관광호텔
1	2	서울특별시	종로구	관광호텔업 센추럴
2	3	서울특별시	종로구	관광호텔업 VATICA HOTEL(바티카 호텔)
3	4	서울특별시	종로구	관광호텔업 동화킴스관광호텔
4	5	서울특별시	종로구	가족호텔업 오라카이 인사동 스위트

데이터 전처리

- READ_EXCEL 함수 이용하여 서울시 호텔리스트 2022년 12월 기준 파일 LOAD
- 데이터가 400행이 넘기 때문에 일부만 보이도록 HEAD() 함수 이용

데이터 전처리

```
번호          int64
지역1_(시도)  object
지역2_(시군구) object
업종          object
호텔명_(등록명칭) object
영업상태_(22.12.31.기준) object
등급          object
등급부여일    object
결정기관      object
객실수        int64
주소(도로명주소) object
위도          float64
경도          float64
dtype: object
```

**DTYPES 속성 이용하여 각 컬럼
별 데이터 타입 확인**

```
번호          0
지역1_(시도)  0
지역2_(시군구) 0
업종          0
호텔명_(등록명칭) 0
영업상태_(22.12.31.기준) 0
등급          69
등급부여일    139
결정기관      141
객실수        0
주소(도로명주소) 0
위도          0
경도          0
dtype: int64
```

**ISNA() 및 SUM()메서드 이용하여
결측치가 있는지 확인**

데이터 전처리

호텔의 밀집도를 확인할 예정이기 때문에 '영업상태' (22.12.31.기준)', '등급부여일', '등급', '결정기관' 4가지 컬럼은 우선 제외하기로 한다.

	번호	지역1_(시도)	지역2_(시군구)	업종	호텔명_(등록명칭)	객실수	주소(도로명주소)
0	1	서울특별시	종로구	관광호텔업	동대문관광호텔	52	서울특별시 종로구 종로 286 (창신동, 이스턴호텔)
1	2	서울특별시	종로구	관광호텔업	센추럴	78	서울특별시 종로구 청계천로 137 (장사동, 센트럴관광호텔)
2	3	서울특별시	종로구	관광호텔업	VATICA HOTEL(바티카 호텔)	50	서울특별시 종로구 청계천로 341 (창신동, 삼호호텔)
3	4	서울특별시	종로구	관광호텔업	동화킴스관광호텔	51	서울특별시 종로구 평창문화로 136 (평창동, 킴스관광호텔)
		서울특별시	종로구	관광호텔업			서울특별시 종로구 창신동 1가 10 (창신동, 이스턴호텔)

	지역1_(시 도)	지역2_(시 군구)	업종	
번호				
1	서울특별시	종로구	관광호텔업	
2	서울특별시	종로구	관광호텔업	
3	서울특별시	종로구	관광호텔업	VATICA H
4	서울특별시	종로구	관광호텔업	
5	서울특별시	종로구	가족호텔업	오라카
...	
451	서울특별시	송파구	관광호텔업	소피텔 앰배서더 서울
452	서울특별시	강동구	관광호텔업	

데이터 전처리

번호 컬럼을 인덱스로 설정한다.

데이터 전처리

	번호	객실수	위도	경도
count	455.000000	455.000000	455.000000	455.000000
mean	228.000000	130.395604	37.541977	126.982994
std	131.491445	137.429556	0.037363	0.066968
min	1.000000	6.000000	37.449021	126.802516
25%	114.500000	40.000000	37.513178	126.928486
50%	228.000000	73.000000	37.552280	126.988577
75%	341.500000	188.500000	37.564628	127.029343
max	455.000000	1151.000000	37.685143	127.138069

기초통계량 확인 결과 특별히 유의미한 확인은 할 수 없었다.

1) Folium 지도 시각화 분석 방법 및 분석 결과

FOLIUM 지도 시각화 도구로 확인 결과
관광 특구인 강남과 중구에 숙박시설이 많이
밀집되어있는 것을 마커 밀집도로 확인 할 수
있음



```
map_osm=folium.Map(location=[37.566651, 126.978428],zoom_start=11)

html="<font color='green' style='white-space:nowrap'><b>%s</b></font>"

#데이터프레임의 행 수만큼 반복하면서 마커 생성
for i in df1.index:
    #행 우선 접근 방식으로 값 추출하기
    name=df1.loc[i,'호텔명_(등록명칭)']
    lat=df1.loc[i,'위도']
    lng=df1.loc[i,'경도']

    #추출한 정보를 지도에 표시
    popup_html=folium.Popup(html%name,parse_html=False)
    marker=folium.Marker([lat,lng],popup=popup_html)
    marker.add_to(map_osm)

map_osm.save('map_osm2.html') # 지도시각화 저장
```


2) 도수분포표 분석 방법

주소(도로명주소) 라는 컬럼에서 문자열
데이터를 SPLIT(자르기) 후에 공백을 기준
으로 2번째 열만 선택

ASSIGN 함수 이용하여 'GU' 컬럼을 '구'로
변경해서 추가

"구" 컬럼을 활용하여 각 행정구별로 숙박
업소(호텔, 호스텔)의 수를 살펴봄

행정구		count
0	중구	103
1	강남구	63
2	종로구	42

객실 수	주소(도로명주소)	위도	경도	행정 구
52	서울특별시 종로구 종로 286 (창 신동, 이스턴호텔)	37.570765	127.010427	종로 구
78	서울특별시 종로구 청계천로 137 (장사동, 센트럴관광호텔)	37.568812	126.992885	종로 구
50	서울특별시 종로구 청계천로 341 (창신동, 삼호호텔)	37.570083	127.015532	종로 구
51	서울특별시 종로구 평창문화로 136 (평창동, 김스관광호텔)	37.609446	126.975163	종로 구
213	서울특별시 종로구 인사동4길 18 (낙원동)	37.572962	126.987332	종로 구
...

```

hist,bins=numpy.histogram(cdf['count'],bins=5)
hist,bins

info=[]
for i,v in enumerate(hist): #해당 구문 적용안하면 마지막 문구도 마
    if i +1<len(hist):
        tpl="%d이상 %d미만"
    else:
        tpl="%d이상 %d이하"

    cls=tpl%(bins[i],bins[i+1])
    line={"숙박업소_수":cls,"행정구역의_수":v}
    info.append(line)

hist_df=DataFrame(info)
hist_df.set_index('숙박업소_수',inplace=True)
hist_df

```

숙박업소_수

10이상 21미만 19

21이상 41미만 3

41이상 62미만 1

62이상 82미만 1

82이상 103이하 1

2) 도수분포표 분석 방법 및 분석 결과

CDF['COUNT'] 열의 데이터를 이용하여 히스토그램을 생성. BINS=5로 설정하여 5개의 구간으로 나눔

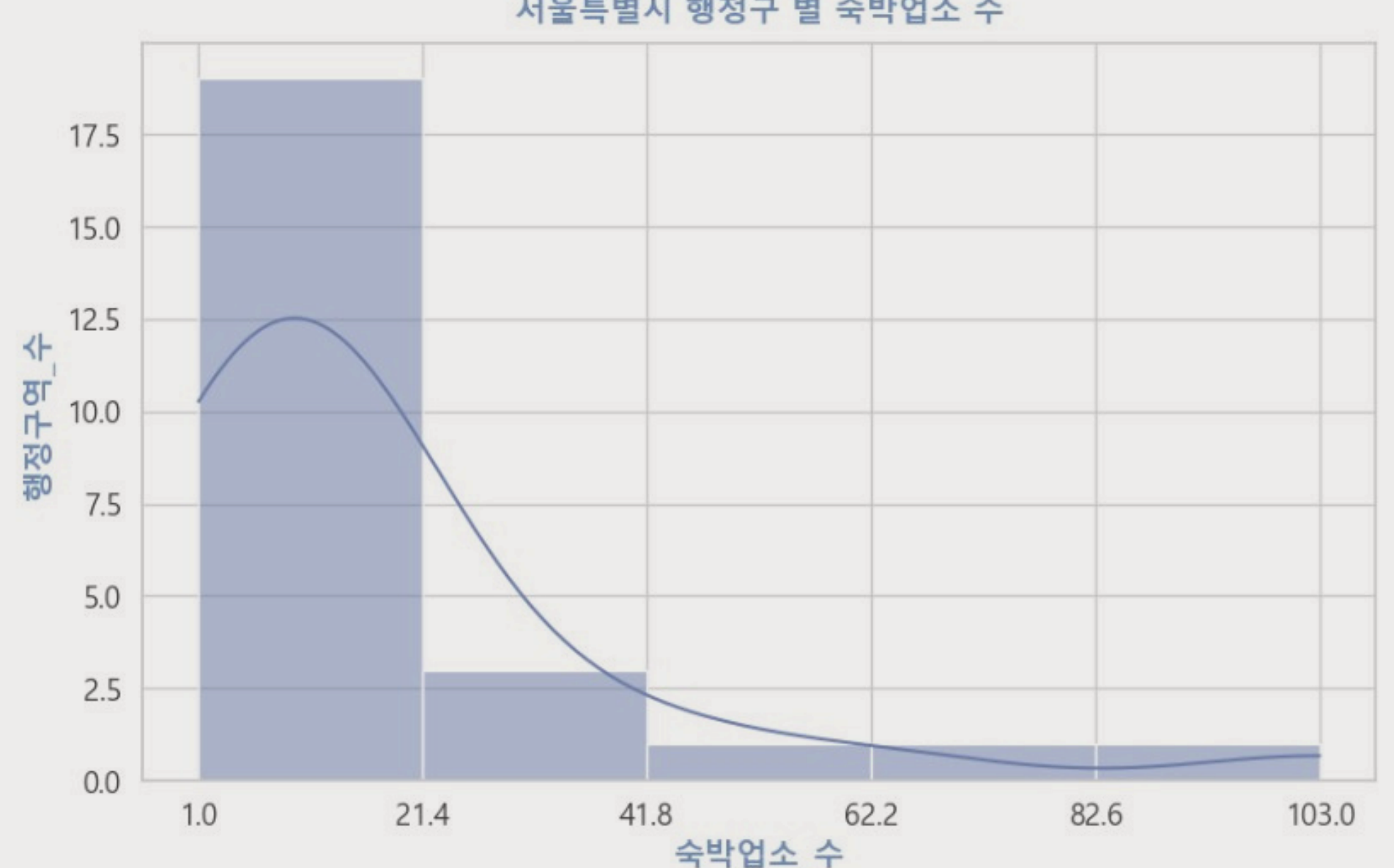
HIST: 각 구간(BIN)에 속하는 데이터 개수를 담은 배열
BINS: 구간(BIN)의 경계값을 담은 배열

확인 결과 1 이상 21 미만의 숙박업소 수를 가진 행정구는 총 19개로 압도적으로 집중되어 있음을 확인할 수 있음. 이를 바탕으로 추론해 볼 수 있는 점은, 해당 19개 행정구는 숙박시설이 상대적으로 부족한 지역임을 암시하고 있으며, 반면에 나머지 상위 2개 구간인 62 이상 부터 82 미만 및 82 이상부터 103 이하에 해당하는 행정구(총 2곳)에서는 숙박업소가 주로 분포되어 있는 것을 관찰할 수 있음.

3) 히스토그램 분석 방법 및 분석 결과

히스토그램 시각화를 하기 위해
그래프 세팅(한글폰트, 폰트사이
즈, 차트사이즈, 음수표시설정)
후 히스토그램 시각화 진행

도수분포표와 같은 결론을 확인
할 수 있다.



```
pyplot.rcParams[ 'font.family' ] = 'AppleGothic' if sys.platform == 'darwin'
pyplot.rcParams[ "font.size" ] = 12
pyplot.rcParams[ "figure.figsize" ] = (10, 6)
pyplot.rcParams[ "axes.unicode_minus" ] = False
```

```
pyplot.figure()
seaborn.set_theme(style='whitegrid', font='Malgun Gothic', font_scale=1.2)
seaborn.histplot(data=cdf,x='count',bins=5,kde=True)
# for문을 사용하여 bins에 있는 값을 하나씩 읽어와서 "{:.1f}".format(b)와 같이
# 그 결과를 리스트로 만들어 눈금 레이블로 사용.
pyplot.xticks(bins, ["{:.1f}".format(b) for b in bins])
pyplot.title("서울특별시 행정구 별 숙박업소 수", pad=10,color='steelblue',fontweight='bold')
pyplot.xlabel("숙박업소_수",color='steelblue',fontweight='bold')
pyplot.ylabel("행정구역_수",color='steelblue',fontweight='bold')
pyplot.show()
pyplot.close()
```


4) 막대 그래프 시각화 (1)

분석 방법 및 분석 결과

SEABORN 막대그래프 활용하여 시각화 진행

- 확인 결과, 중구에 가장 많은 숙박업소가 있으며 그 다음으로 강남구, 종로구 순으로 보여진다.
- 숙박업소 수가 가장 적은 곳은 양천구와 노원구로 확인된다.



```
pyplot.rcParams["figure.figsize"]=(16,4)
pyplot.figure()
pyplot.title("서울 지역별 숙박업소 수",pad=20,color='steelblue',fontsize=20,font
seaborn.barplot(data=cdf,x=cdf['행정구'],y=cdf['count'])
pyplot.xlabel("행정구",color='steelblue',fontweight='bold')
pyplot.ylabel("숙박업소 수",color='steelblue',fontweight='bold')
pyplot.xticks(rotation=45) # x축 레이블(텍스트)각도 조절
pyplot.grid()
pyplot.show()
pyplot.close()
```

4) 막대그래프 시각화 (2)

분석 방법 및 분석 결과

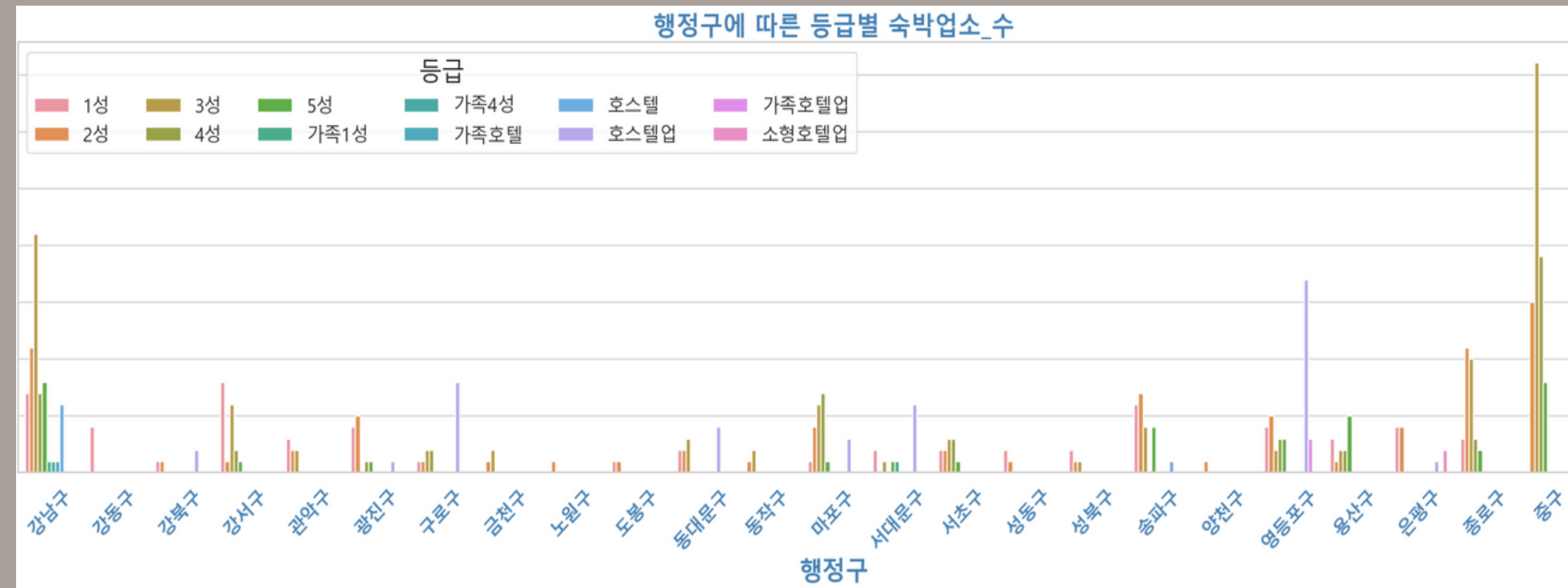
SEABORN 막대그래프 활용하여
행정구에 따른 등급별 숙박업소_수 시각화

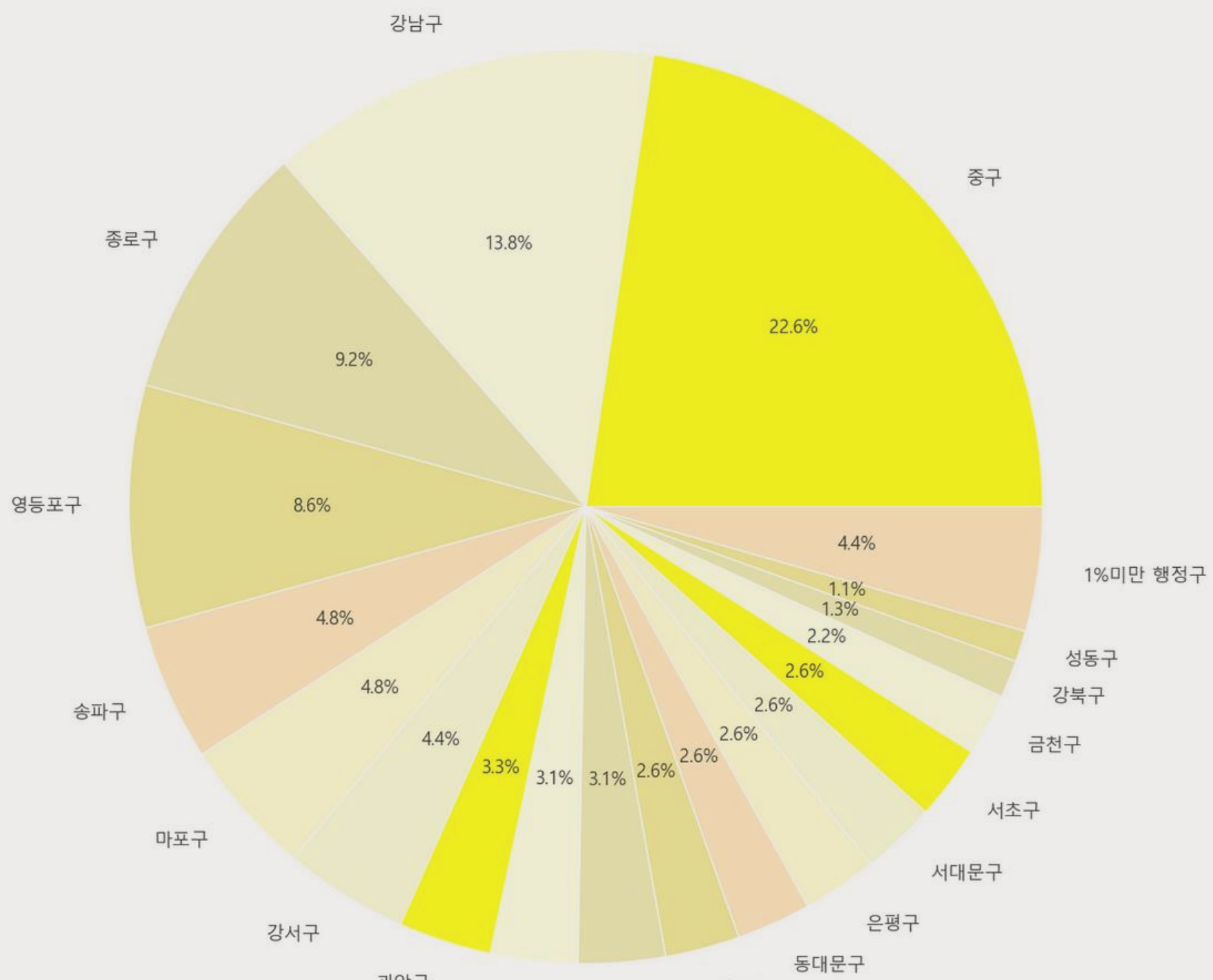
- 행정구에 따른 등급별 숙박업소_수를 확인할 수 있는 그래프이다.
- 여기서 알 수 있는 사실은 3성급 호텔은 중구에 가장 많이 집중되어있으며 영등포구에는 호텔업이 압도적으로 많은 것을 알 수 있다.
- 5성급 호텔의 경우 숙박업소 밀집도가 가장 높은 중구, 강남구에 많이 있는것으로 확인된다.

```
df_cleaned = df.dropna(subset=['등급'])

# 지역과 등급에 따른 호텔 수 계산
df_grouped = df_cleaned.groupby(['지역2_(시군구)', '등급']).size().reset_index(name='호텔수')

# 막대그래프 그리기
pyplot.figure(figsize=(25, 6))
seaborn.barplot(x='지역2_(시군구)', y='호텔수', hue='등급', data=df_grouped)
pyplot.title('행정구에 따른 등급별 숙박업소_수', color='steelblue', fontsize=20, fontweight='bold')
pyplot.xlabel('행정구', color='steelblue', fontsize=20, fontweight='bold')
pyplot.ylabel('숙박업소_수', color='steelblue', fontsize=20, fontweight='bold')
pyplot.xticks(rotation=45, fontsize=15, color='steelblue', fontweight='bold') # rotation: x축 레이블
# 범례 사이즈 조절
pyplot.legend(title='등급', title_fontsize='20', fontsize='15', loc='upper left', ncol=6) #loc: 범례
pyplot.show()
pyplot.close()
```





5) 파이차트 시각화(2) 분석방법 및 분석 결과

파이차트를 활용하여 시각화 진행 2번째

1 % 미만 행정구인 노원구, 양천구, 중랑구, 도봉구, 강동구, 성북구 총 6곳의 데이터를 하나로 묶어 재시각화 진행. 가독성이 훨씬 나아진 것을 확인할 수 있다.

```
# 작은 비율 찾기
threshold = 5 # 작은 비율 임계값 (1% 미만)
small_values = cdf[cdf['count'] < threshold] # 작은 값 필터링
```

```
# 작은 값들의 합 구하기
small_sum = small_values['count'].sum()
```

```
# 기존 데이터프레임에서 작은 값들 제거
cdf_filtered = cdf[~cdf['행정구'].isin(small_values['행정구'])]
```

```
# 새로운 '기타' 항목 추가
cdf_filtered.loc[len(cdf_filtered)] = ['1%미만 행정구', small_sum]
```

```
# 그래프 출력
colors = ['yellow', 'lightyellow', 'palegoldenrod', 'khaki', 'moccasin', 'lemonchiffon', 'lightgoldenrodyellow']
pyplot.title("서울 지역별 숙박업소 수", loc='left', pad=20, fontweight='bold')
pyplot.pie(cdf_filtered['count'], labels=cdf_filtered['행정구'], colors=colors, autopct='%0.1f%%')
pyplot.axis('equal')
```

6) SVG 지도시각화 분석 방법

```
cdf2 = cdf.reset_index()

조건 = [
    cdf2['행정구'] == "은평구",
    cdf2['행정구'] == "영등포구",
    cdf2['행정구'] == "동대문구",
    cdf2['행정구'] == "강동구",
    cdf2['행정구'] == "종로구",
    cdf2['행정구'] == "강남구",
    cdf2['행정구'] == "중구",
    cdf2['행정구'] == "송파구",
    cdf2['행정구'] == "성동구",
    cdf2['행정구'] == "서초구",
    cdf2['행정구'] == "중랑구",
    cdf2['행정구'] == "노원구",
    cdf2['행정구'] == "구로구",
    cdf2['행정구'] == "서대문구",
    ]

값 = ["Eunpyeong-gu", "Yeongdeungpo-gu", "Jongno-gu", "Gangnam-gu", "Jung-gu", "Seocho-gu", "Jungnang-gu", "Nowon-gu", "Dobong-gu", "Seongbuk-gu", "Geumcheon-gu", "Gwanak-gu", "Gangseo-gu", "Gangbuk-gu", "Dongjak-gu"]

cdf2['Gu'] = numpy.select(조건, 값, default="")

# 영문으로 생성된 Gu 컬럼을 인덱스로 설정
cdf2.set_index('Gu', inplace=True)

# 기존에 있던 구 컬럼은 삭제
cdf2.drop('행정구', axis=1, inplace=True)

# count 컬럼을 "숙박업소_수"로 변경
cdf2.rename(columns={'count': '숙박업소_수'})
cdf2
```

SVG 파일 ID값이 영문이므로
값을 통일해야하만 싱크를 맞출 수
있음. 엑셀파일 행정 구를 영문
으로 변환하기위한 절차 진행

영문으로 변환하기 위해 인덱스를
리셋 후 리스트 생성하여 값을 집
어넣음

NUMPY.SELECT 함수를 사용
하여 조건에 따라 값을 선택 하여
'GU' 컬럼에 할당

분석 방법

데이터 확인 결과 영문으로
잘 변경된 것을 확인 할 수
있다.

	index	숙박업소_수
Gu		
Jung-gu	0	103
Gangnam-gu	1	63
Jongno-gu	2	42
Yeongdeungpo-gu	3	39
Songpa-gu	4	22
Mapo-gu	5	22
Gangseo-gu	6	20
Yongsan-gu	7	15
Guro-gu	8	14
Gwanjin-gu	9	14

분석 방법

1. request를 사용하여 SVG 파일 load

2. requests.get() 함수를 통해 서울 SVG파일을 불러와 주소에서 SVG 파일을 요청하고, 가져온 데이터를 변수 map_svg에 저장.

3. 그 후 BeautifulSoup 라이브러리를 사용, SVG 파일을 파싱하고, 'path' 태그 중에 'id' 속성이 있는 요소들을 추출하여 리스트로 저장.

```
#지도 이미지 가져오기
```

```
user_agent = "Mozilla/5.0 (Windows NT 10.0; Win64; x64) Ap  
header_info = {'User-agent': user_agent, 'referer': None}
```

```
r = requests.get('https://data.hossam.kr/D04/map_seoul.svg')
```

```
# 결과 검사
```

```
if r.status_code != 200:
```

```
    # 에러코드와 에러메시지를 문자열로 구성
```

```
map_svg = r.text
```

```
# SVG(map_svg)
```

```
#지도 이미지 파싱
```

```
soup = BeautifulSoup(map_svg)
```

```
# soup
```

```
#구 단위로 추출
```

```
path_list = soup.select('path[id]')
```

```
# 출력 결과는 path 단위를 저장한 리스트
```

```
# path_list
```

```
# 지도에서 확인한 지명만큼 반복
```

```
for p in path_list:
```

```
    지역명 = p['id']
```

```
    # if 지역명 not in cdf2.index: # 지도와 행정구 갯수가 매칭이 안되므로 if구문 및  
    # continue # 이용해서 없는 행정구가 있으면 건너된다.
```

```
    호텔수 = cdf2.loc[지역명, '숙박업소_수']
```

```
    # print(지역명, " -->", 숙박업소_수)
```

```
    # 히스토그램 분포도수를 참고
```

```
    if 호텔수 > 82.6: color_index = 4
```

```
    elif 호텔수 > 62.2: color_index = 3
```

```
    elif 호텔수 > 41.8: color_index = 2
```

```
    elif 호텔수 > 21.4: color_index = 1
```

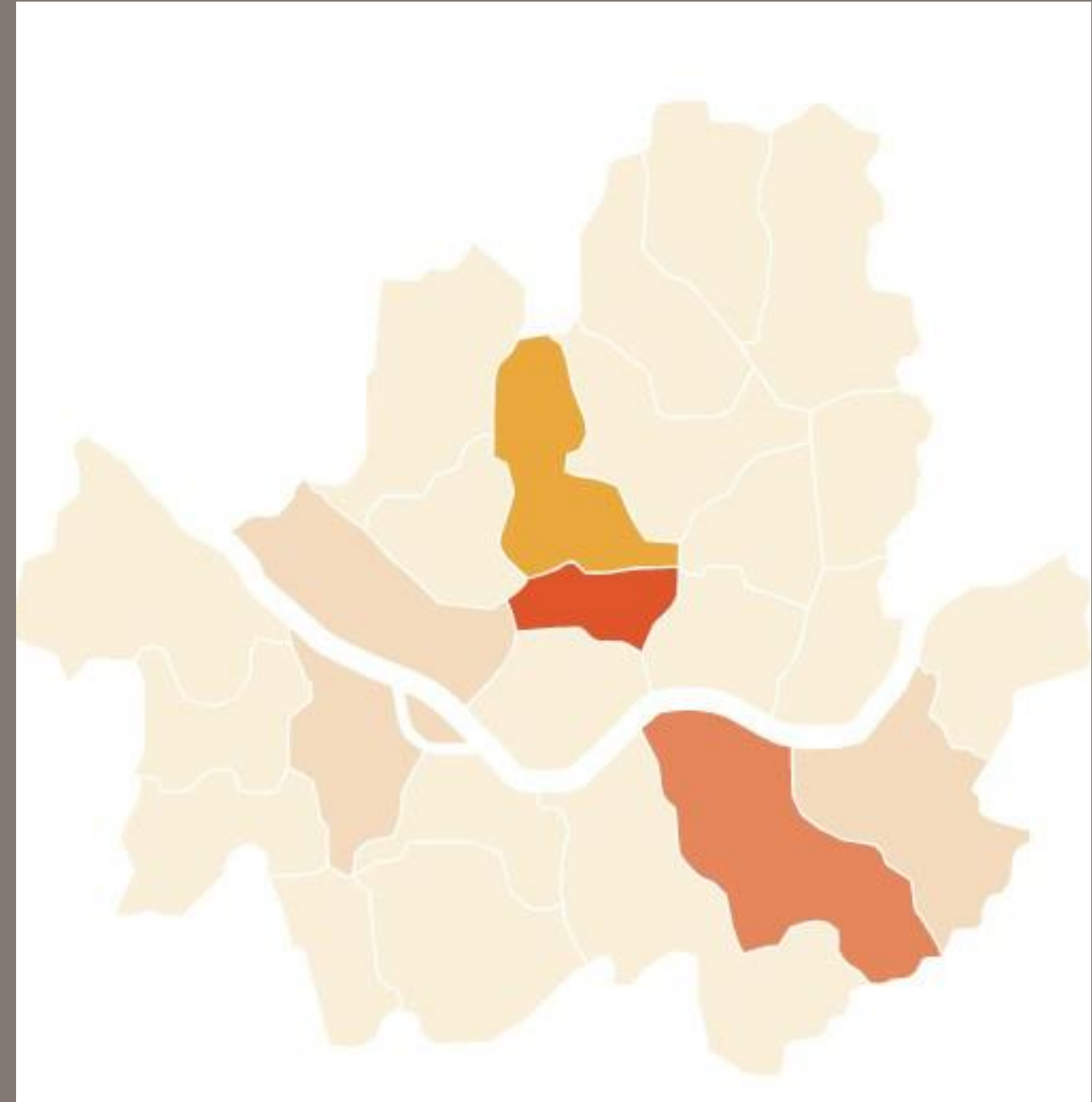
```
    else: color_index = 0
```

```
    # 각 지역별 색상값을 데이터 프레임에 변환
```

```
    cdf2.loc[지역명, '색상값'] = colors[color_index]
```

분석 결과

COLOR가 들어간 지도 확인 결과
중구에 숙박업소가 가장 많이 분
포되어 있으며 강남구, 종로구 순
으로 나타나는 걸 알 수 있다.



3. 분석 결과(Analysis Results)

80%



RESULTS



해당 데이터를 토대로 데이터 시각화한 결과, 서울 내 숙박업소의 약 절반(45%)이 중구 및 강남구에 밀집되어 있는 것으로 확인된다.

4. 결론(conclusion)

100%



CONCLUSION

- 숙박 시설은 여행객이나 장거리 출장자들에게 안락한 휴식 공간을 제공함으로써, 체력 회복과 개인 위생을 유지하는 데 필수적이다. 또한 이는 지역 경제 활성화에 기여하며, 관광 산업의 중요한 부분을 차지한다.
- 특정 행정 구에 숙박 시설이 집중되면 그 지역의 인프라스트럭처에 과부하를 주고, 주거 공간 부족과 임대료 상승 등의 문제를 초래할 수 있다. 또한 일부 지역만 관광객이 몰리게 되어 지역 간 경제적 불균형이 심화될 수 있다.
- 숙박 시설이 거의 없는 행정 구에서는 해당 지역의 문화적, 역사적 자원을 홍보하고, 적절한 인프라와 교통 네트워크를 구축하여 여행객들이 방문하기 용이하도록 만드는 것이 중요하다고 볼 수 있다. 또한 현지 정부가 유인책(세금 감면 등)을 마련하여 사업자들로 하여금 새로운 숙박 시설을 설립하는 것을 독려할 수 있을 것이다.

THANK YOU!

F O R Y O U R A T T E N T I O N