



# 군산사랑상품권 가맹점 리스트 DB/XML 프로젝트

최민서, 한동엽

본 저작물은 군산시청 소상공인지원과에서 '22년 작성하여 공공누리 제4유형으로 개방한 저작물을 이용하였습니다.

# 팀 역할 분담

---



## 최민서

---

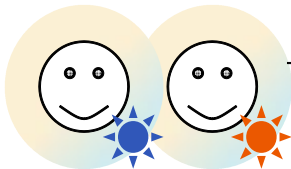
- XML/XSD 파일 관리
- 데이터 수집 / 전처리
- 웹 시각화 표현



## 한동엽

---

- 프로젝트(csv 등) 파일 관리
- DB 테이블 관련 관리
- 데이터 수집/크롤링/전처리



## 공동역할

---

- XSD 설계
- 트리 설계
- 공통어휘 찾기
- DB 구조 설계
- 정규화

# 목차 Content

---

1 주제 소개

2 DB 설계

3 XML 설계

4 마무리

---

# 1

# 주제 소개

---

- 군산사랑상품권이란?
- 고객 니즈
- 프로젝트 목적
- 데이터 현황 및 수집
- 데이터 전처리
- 데이터 크롤링

# 군산사랑상품권이란?

- 군산시의 지역 상품권
  - 어려운 재래시장과 골목상권을 보호,  
지역 자금의 타지역 유출을 방지함
  - 도내 자영업 및 소상공인들의 소득 증  
대로 선순환 지역경제구조를 유도
    - 상품권을 10% 가격 감면으로 소비자들  
도 애용하고 있는 상품권임




# 고객 니즈분석



대상	- 군산사랑상품권 이용자
상황	- 가게에 들어가 군산사랑상품권으로 결제하려 하지만 가맹점이 아니라 결제를 못하여 낭패
고민	- 가맹점이 아닌 곳을 찾고 싶음. - 가맹점은 어디에 있는지 알고 싶음.
충족	군산사랑상품권 DB를 구현하여 가맹점 여부를 파악할 수 있음

# 프로젝트의 목적

---

- 공공데이터포털(data.go.kr)에 있는 **군산시 상가 리스트**
    - (병원, 식품제조업체, 의원, 제과점, 안경점 등)
  - 군산시청(gunsan.go.kr)에 있는 **군산사랑상품권 가맹점 리스트**
- 
- XML 설계를 통해 **교집합** 부분을 찾음
  - 만든 상가 리스트로 **가맹점 해당여부** 태그를 만듦
  - 상가 리스트에 **시설 편의여부**를 조사함
    - 예) 병원 리스트 -> 병실수, 병동수 / 음식점 리스트 -> 모범음식점 여부

# 데이터 수집 현황

전라북도 군산시_관광호텔현황_202111...	2022-10-29 오전 4:13	한컴오피스 한글 ...	1KB
전라북도 군산시_노래연습장현황_20220...	2022-10-29 오전 4:14	한컴오피스 한글 ...	18KB
전라북도 군산시_대형마트 및 SSM현황_...	2022-10-29 오전 4:15	한컴오피스 한글 ...	1KB
전라북도 군산시_모범음식점_20220811...	2022-10-29 오전 4:13	한컴오피스 한글 ...	5KB
전라북도 군산시_문화시설현황_202111...	2022-10-29 오전 4:14	한컴오피스 한글 ...	1KB
전라북도 군산시_병원현황_20220609.csv	2022-10-29 오전 4:12	한컴오피스 한글 ...	3KB
전라북도 군산시_식품제조업체현황_202...	2022-10-29 오전 4:14	한컴오피스 한글 ...	18KB
전라북도 군산시_안경업소 현황_202205...	2022-10-29 오전 4:14	한컴오피스 한글 ...	5KB
전라북도 군산시_약국 현황_20220531.csv	2022-10-29 오전 4:12	한컴오피스 한글 ...	10KB
전라북도 군산시_의원현황_20220503.csv	2022-10-29 오전 4:12	한컴오피스 한글 ...	12KB
전라북도 군산시_정신건강의학과현황_2...	2022-10-29 오전 4:14	한컴오피스 한글 ...	1KB
전라북도 군산시_저과점현황_20220630...	2022-10-29 오전 4:14	한컴오피스 한글 ...	11KB
전라북도 군산시_제조업공장현황_20211...	2022-10-29 오전 4:14	한컴오피스 한글 ...	167KB
전라북도 군산시_죽석판매제조가공업 현...	2022-10-29 오전 4:14	한컴오피스 한글 ...	55KB
전라북도 군산시_지역화폐 발행현황_20...	2022-10-29 오전 4:14	한컴오피스 한글 ...	1KB
전라북도 군산시_지역화폐 이용현황_20...	2022-10-29 오전 4:14	한컴오피스 한글 ...	3KB
전라북도 군산시_직판가격업소 현황_20...	2022-10-29 오전 4:12	한컴오피스 한글 ...	2KB
전라북도 군산시_지과의원현황_202205...	2022-10-29 오전 4:13	한컴오피스 한글 ...	6KB
전라북도 군산시_한의원현황_20220503...	2022-10-29 오전 4:11	한컴오피스 한글 ...	6KB
전라북도_나들가게 개점 현황_2020102...	2022-10-29 오전 4:17	한컴오피스 한글 ...	1KB



	A	B	C	D	E	F	G	H
1	의료기관종별	의료기관명	의료기관주소(도로명)	의료기관전화번호	병실수	병상수	비고(건물층수)	
2	종합병원	의료법인오성	전라북도 군산시 조촌	063-440-0570	94	283	지상8층 지하1층	
3	종합병원	전라북도 군산	전라북도 군산시 의료	063-472-5000	113	413	본관 지상8층 지하1층	
4	병원	의료법인 관주	전라북도 군산시 대학	063-460-3000	49	203	지상6층 지하1층	

전라북도 군산시\_병원현황\_20220609.csv

R=(의료기관종별, 의료기관명, 의료기관주소(도로명),  
의료기관전화번호, 병실수, 병상수, 비고(건물층수))



# 데이터 수집

The image shows a web browser window displaying a list of restaurants in Gunsan City. Overlaid on the browser is an Excel spreadsheet titled '군산맛집 목록 안내표' (Gunsan Restaurant List Guide). The spreadsheet has columns for '번호' (Number), '집 이름' (Restaurant Name), '소재지 (도로명)' (Address (Road Name)), and '연락처' (Contact Information). The list includes various restaurants such as '2. 경원대학교', '2. 경원대학교', '2. 경원대학교', etc., with their respective addresses and phone numbers. The browser window shows the website 'www.gunsang.kr' with a search bar and a list of restaurants.

엑셀 내 사이트 표  
가져오기 기능을 활용하여  
크롤링  
(군산시청 맛집 리스트)

# 데이터 전처리

	A	B	C	D	E	F
1	제목	단위	기준	출처		
2	상가정보		2021-08-01	소상공인시장진흥공단		
3						
4	상호명	상권업종중분류명	행정동명	지번주소	도로명주소	중정보
5	접속노래방	무도/유흥/가무	조촌동	전라북도 군산전라북도 군산		
6	준베이커리	제과제빵떡케익	옥서면	전라북도 군산전라북도 군산		
7	태화각	중식	조촌동	전라북도 군산전라북도 군산		

상가현황.xlsx

업종에 요소 두 세개가 있기도 하며, 같은 업종이라도 다른 방식으로 써져 있는 경우가 많음.  
그러므로 간략한 제약을 둘 예정

업종 - 소매업, 보건업, 숙박업, 개인서비스업, 음식점업, 교육서비스업, 제조업, 스포츠여가관련서비스업, 기타

## 데이터 수집 특이사항

1. 노래방, 제과점 등의 파일에는 전화번호가 적혀있지 않음.  
-> 사업장전화번호는 minoccurs="0"으로 할 예정
2. 모범음식점과 제조업공장, 병원 등은 건물층수, 주차대수 등 세부적인 인스턴스 사항들이 있음

※ 업종분류: 소매업, 보건업, 숙박업, 개인서비스업, 음식점업, 교육서비스업, 제조업, 스포츠여가관련서비스업, 기타

# 데이터 전처리











- csv / xlsx -> xml 변경작업 (python) csv to xsd

```
1 import pandas as pd
2
3 xlsx = pd.read_excel("xmlldb_dataset\\상가정보.xlsx")
4 xlsx.to_csv("xmlldb_dataset\\상가정보.csv", encoding="utf-8-sig")

1 import csv
2 reader = csv.reader(open('xmlldb_dataset\\군산시별\\전라북도 군산시_모범음식점_20220811.csv', 'r'), delimiter=",")
3
4 f=open('xmlldb_dataset\\군산시별\\전라북도 군산시_모범음식점.xml', 'w+', encoding="utf-8-sig")
5 f.write('<?xml version="1.0" encoding="utf-8"?>\n\n')
6 f.write('<goodRestaurant>\n')
7
8 for row in reader:
9     f.write('\t<Restaurant>\n')
10    f.write('\t\t<legalName>' + row[1] + '</legalName>\n')
11    f.write('\t\t<EastAsianAddress>' + row[2] + '</EastAsianAddress>\n')
12    f.write('\t\t<telephone>' + row[3] + '</telephone>\n')
13    f.write('\t\t<servesCuisine>' + row[4] + '</servesCuisine>\n')
14    f.write('\t</Restaurant>\n')
15 f.write('</goodRestaurant>')
```

# 데이터 전처리

- 데이터 활용을 위해 파이썬으로 다양한 전처리

 CSVtoXML.py	2022-12-02 오후 7:01
 datatype.py	2022-12-02 오후 11:27
 duplicatecheck.py	2022-12-02 오후 11:34
 finalresultCSVtoXML.py	2022-12-03 오전 12:54
 isThisChackanStore.py	2022-12-03 오전 12:40
 isThisGoodRestaurant.py	2022-12-03 오전 12:35
 isThisKunsancredit.py	2022-12-02 오후 10:19
 isThisKunsancreditMK2.py	2022-12-03 오전 12:28
 kunsancredit.py	2022-12-02 오후 6:47
 silhum.py	2022-12-02 오후 10:28

# 데이터 크롤링

---

- 가맹점 리스트 파일이 제공되지 않음
  - 군산시청 내 사이트 가맹점 데이터 크롤링 필요.
    - (군산시청 소상공인지원과 연락 후 허락받음)

# 데이터 크롤링

## 가맹점 조회

상호
검색

※ 업종분류: 소매업, 보건업, 숙박업, 개인서비스업, 음식점업, 교육서비스업, 제조업, 스포츠여가관련서비스업, 기타



번호	상호	업종	소재지	행정동	사업장전화번호
12607	아람솔루션(테스트)	제조업	전라북도 군산시 전주시 완산구 흥산북로 60-10(효자동3가), 2	효자동3가	
12606	회현떡방앗간	제조업	전라북도 군산시 회현면 회현초교2길 28-9	회현면	0634665036
12605	논	음식점업	전라북도 군산시 회현면 회현초교1길 47, 1층	회현면	

번호	상호	업종	소재지	행정동	사업장전화번호
12607	아람솔루션(테스트)	제조업	전라북도 군산시 전주시 완산구 흥산북로 60-10(효자동3가), 2	효자동3가	
12606	회현떡방앗간	제조업	전라북도 군산시 회현면 회현초교2길 28-9	회현면	0634665036
12605	논	음식점업	전라북도 군산시 회현면 회현초교1길 47, 1층	회현면	
12604	금강이용원	개인서비스업	전라북도 군산시 회현면 회현로 198	회현면	
12603	금강농약사	소매업	전라북도 군산시 회현면 회현로 190	회현면	0634666095
12602	카메라실	음식점업	전라북도 군산시 회현면 회현로 182(1층)	회현면	0634653555
12601	해인머리방	개인서비스업	전라북도 군산시 회현면 회현로 182	회현면	0634654956

<tr><td> 부분 크롤링

# 데이터 크롤링

---

## - Python selenium 패키지 사용, pandas

```
from selenium import webdriver
from selenium.webdriver.chrome.service import Service
from webdriver_manager.chrome import ChromeDriverManager
from selenium.webdriver.common.by import By
import time

def set_chrome_driver():
    chrome_options = webdriver.ChromeOptions()
    driver = webdriver.Chrome(service=Service(ChromeDriverManager().install()), options=chrome_options)
    return driver

url = 'https://www.gunsan.go.kr/main/m308'
driver = set_chrome_driver()
driver.get(url)
```

```
# 릴레이션 이름 가져오기
```

```
gunsan_cert_column = (driver.find_elements(By.XPATH, '//*[@id="mapmap"]/div/div[1]')[0].text).split('\n')  
gunsan_cert_column
```

```
['번호', '상호', '업종', '소재지', '행정동', '사업장전화번호']
```

```
import time  
from tqdm import tqdm_notebook
```

```
# 마지막 페이지 번호 구하기
```

```
driver.find_element(By.XPATH, '//*[@id="containertab"]/section/div[6]/div/ul/li[last()]').click()  
fn = driver.find_element(By.XPATH, '//*[@id="containertab"]/section/div[6]/div/ul/li[last()]/a').text  
time.sleep(2)  
driver.find_element(By.XPATH, '//*[@id="containertab"]/section/div[6]/div/ul/li[1]').click()
```

```
print(fn)
```



# 마지막 페이지 번호 구하기

```
driver.find_element(By.XPATH, '//*[@id="containertab"]/section/div[6]/div/ul/li[last()]').click()
fn = driver.find_element(By.XPATH, '//*[@id="containertab"]/section/div[6]/div/ul/li[last()]/a').text
time.sleep(2)
driver.find_element(By.XPATH, '//*[@id="containertab"]/section/div[6]/div/ul/li[1]').click()
```

```
print(fn)
```

1261

```
gunsan_cert = []
```

```
for i in tqdm_notebook(range(int(fn))):
    for i in range(2, 12):
        gunsan_cert.append((driver.find_elements(By.XPATH, '//*[@id="mapmap"]/div/div['+str(i)+']')[0].text).split('\n'))
    driver.get('https://www.gunsan.go.kr/main/m308/list?s_idx='+str(i))
```

```
gunsan_cert
```

Output exceeds the [size limit](#). Open the full output data [in a text editor](#)

```
[['12507', '정우수산', '소매업', '전라북도 군산시 내항2길 297(해망동)', '해망동'],  
 ['12506', '일출물산', '소매업', '전라북도 군산시 내항2길 293(해망동)', '해망동'],  
 ['12505', '옥봉수산', '소매업', '전라북도 군산시 내항2길 289(해망동)', '해망동'],  
 ['12504', '더돌풍수산', '소매업', '전라북도 군산시 내항2길 289', '해망동'],  
 ['12503', '광민수산', '소매업', '전라북도 군산시 내항2길 283(해망동)', '해망동'],
```

```
import pandas as pd  
import csv  
gunsan_cert  
with open('C:\\\\Users\\\\우리집\\\\Downloads\\\\test.csv', 'w', newline='') as f:  
    writer = csv.writer(f)  
    writer.writerows(gunsan_cert)
```

# 데이터 이슈

---

- 수집한 업소, 도로명 주소 **적힌게 다름.** 가맹점 여부를 구하기 어려움
  - 정제에 힘드므로 9개 업종별 중 음식점업에 집중
- 음식점업의 모범음식점, 군산시청 맛집 등의 **데이터 활용**

# 데이터 사용

- 따라서 사용할 데이터를 **재수집 및 재정리**



hospital.csv  
전라북도 군산시\_관광호텔현황\_202111...  
전라북도 군산시\_노래연습장현황\_20220...  
전라북도 군산시\_대형마트 및 SSM현황...  
전라북도 군산시\_모범음식점\_20220811...  
전라북도 군산시\_문화시설현황\_202111...  
전라북도 군산시\_병원현황\_20220609.csv  
전라북도 군산시\_식용제조업재현황\_202...  
전라북도 군산시\_안경업소 현황\_202205...  
전라북도 군산시\_약국 현황\_20220531.csv  
전라북도 군산시\_의원현황\_20220503.csv  
전라북도 군산시\_청신건강의학과현황\_2...  
전라북도 군산시\_제과점현황\_20220630...  
전라북도 군산시\_제조업공장현황\_20211...  
전라북도 군산시\_죽식판매제조가공업 현...  
전라북도 군산시\_지역화폐 발행현황\_20...  
전라북도 군산시\_지역화폐 이용현황\_20...  
전라북도 군산시\_착한가격업소 현황\_20...  
전라북도 군산시\_치과의원현황\_202205...  
전라북도 군산시\_한의원현황\_20220503...  
전라북도\_나들가게 개점 현황\_2020102...  
통합.csv

Gunsan\_food(gift).csv  
Gunsan\_food.csv  
QualityRestaurant.csv  
군산\_백년가게.csv  
군산시\_모범음식점.csv  
군산시\_착한가격업소.csv  
군산시청\_맛집선정리스트.csv

# 데이터 전처리

Custom text transform on column RoadNameAddress

Expression Language: General Refine Expression Language (GREL) ▼

value.replace(/,([^\s])+/,"").replace(/#([^\s])+/,"") No syntax error.

Preview History Starred Help

row	value	value.replace(/,([^\s])+/,"").re ...
1.	전라북도 군산시 회현면 회현초교2길 28-9	전라북도 군산시 회현면 회현초교2길 28-9
2.	전라북도 군산시 회현면 회현초교1길 47, 1층	전라북도 군산시 회현면 회현초교1길 47
3.	전라북도 군산시 회현면 회현로 198	전라북도 군산시 회현면 회현로 198
4.	전라북도 군산시 회현면 회현로 190	전라북도 군산시 회현면 회현로 190
5.	전라북도 군산시 회현면 회현로 182(1층)	전라북도 군산시 회현면 회현로 182

On error: ☒ keep original ☐ set to blank ☐ store error ☐ Re-transform up to 10 times until no change

OK Cancel

도로명주소에  
괄호 안에 둘러진 심표, 문자 등을 제거

# 데이터 전처리

리스트를 활용하여  
군산상가 리스트에 있는  
모범음식점 찾기

```
kunsanstore=open('새 폴더\\finalresult.csv', 'w', encoding='utf-8')
goodRestaurant=open('새 폴더\\QualityRestaurant.csv', 'w', encoding='utf-8')
finalresult=open('새 폴더\\finalresult2.csv', 'w', encoding='utf-8')
finalresult.write("id, legalName, category, RoadNameAddress, AdministrativeNeighborhood, isThisKunsancredit, isThisgoodrestaurant, isThisCheckanStore\n")
strings=""
a=0
kunsanstorelines=kunsanstore.readlines()
goodRestaurantlines=goodRestaurant.readlines()

for kunsanstoreline in kunsanstorelines:    #2중 반복, 시간복잡도는 n^2
    kunsanstoresplit = kunsanstoreline.split(',')
    for goodRestaurantline in goodRestaurantlines:
        goodRestaurantsplit=goodRestaurantline.split(',')
        if goodRestaurantsplit[2]==kunsanstoresplit[1]:
            strings=kunsanstoresplit[0]+","+kunsanstoresplit[1]+","+kunsanstoresplit[2]+","+kunsanstoresplit[3]+","+kunsanstoresplit[4]+","+kunsanstoresplit[5].replace("\n", "")+",true,\n"    #만약 이름이 같을 경우 여부에 true
            finalresult.write(strings)
        a=1
    if a==0:
        strings =kunsanstoresplit[0]+","+kunsanstoresplit[1]+","+kunsanstoresplit[2]+","+kunsanstoresplit[3]+","+kunsanstoresplit[4]+","+kunsanstoresplit[5].replace("\n", "")+",false,\n"    #여부에 false
        finalresult.write(strings)
    a=0
```

# 데이터 전처리

```
kunsanstore=open('새 품다\\gunsanStore-csv.csv','r',encoding='utf-8')
kunsancredit=open('새 품다\\secondmodified.csv','r',encoding='utf-8')
finalresult=open('새 품다\\finalresult.csv','w',encoding='utf-8')
finalresult.write("id,legalName,category,RoadNameAddress,AdministrativeNeighborhood,isThisKunsancredit")
strings=""
a=1
kunsanstorelines=kunsanstore.readlines()
kunsancreditlines=kunsancredit.readlines()

for kunsanstoreline in kunsanstorelines:    #2중 반복, 시간복잡도는 n^2
    kunsanstoresplit = kunsanstoreline.split(',')
    for kunsancreditline in kunsancreditlines:
        kunsancreditsplit=kunsancreditline.split(',')
        if kunsancreditsplit[2]==kunsanstoresplit[1]:
            strings=kunsanstoresplit[0]+","+kunsanstoresplit[1]+","+kunsanstoresplit[2]+","+kunsanstoresplit[3]+","+kunsanstoresplit[4]+",true\n"    #민약 이름이 같을 경우 여부에 true
            finalresult.write(strings)
            a=1#그리고 true라는 것을 증명하는 a를 1로
        if a==0: #반복문이 끝날 때 까지 a가 1이 되지 않았다는 것은 군산사랑상품권 csv에 이름이 없다는 것이므로
            strings =kunsanstoresplit[0]+","+kunsanstoresplit[1]+","+kunsanstoresplit[2]+","+kunsanstoresplit[3]+","+kunsanstoresplit[4]+",false\n"    #여부에 false
            finalresult.write(strings)
        a=0
#약 200만 번 연산함.
#1분 30초를 걸리는 것 같음. 이렇게 오래 걸리는 코딩은 처음 해 봄
```

문자열을 계속 비교, 같으면 참 반환  
다르면 거짓 반환

# 데이터 전처리

```
kunsanstore=open('새 폴더\\finalresult2.csv','w',encoding='utf-8')
chackanstore=open('새 폴더\\전라북도 군산시 착한가격업소 현황_20220601.csv','w',encoding='cp949')
finalresult=open('새 폴더\\finalresult3.csv','w',encoding='utf-8')
finalresult.write("id,legalName,category,RoadNameAddress,AdministrativeNeighborhood,isThisKunsancredit,isThisgoodRestaurant,isThisChackanStore\n")
strings=""
a=0
kunsanstorelines=kunsanstore.readlines()
chackanstorelines=chackanstore.readlines()

for kunsanstoreline in kunsanstorelines:    #2 중 반복, 시간복잡도는 n^2
    kunsanstoresplit = kunsanstoreline.split(',')
    for chackanstoreline in chackanstorelines:
        chackanstoresplit=chackanstoreline.split(',')
        if chackanstoresplit[2]==kunsanstoresplit[1]:
            strings=kunsanstoresplit[0]+","+kunsanstoresplit[1]+","+kunsanstoresplit[2]+","+kunsanstoresplit[3]+","+kunsanstoresplit[4]+","+kunsanstoresplit[5]+","+kunsanstoresplit[6].replace("\n", "")+",true\n"    # 만약 이름이 같을 경우 여기에 true
            finalresult.write(strings)
            a+=1
        if a==0:
            strings=kunsanstoresplit[0]+","+kunsanstoresplit[1]+","+kunsanstoresplit[2]+","+kunsanstoresplit[3]+","+kunsanstoresplit[4]+","+kunsanstoresplit[5]+","+kunsanstoresplit[6].replace("\n", "")+",false\n"    # 어쨌든 false
            finalresult.write(strings)
            a+=1
```

이하 착한가격업소



# 데이터 전처리

```
import csv
reader = csv.reader(open('새 폴더\\finalresult2.csv', 'r', encoding="utf-8-sig"), delimiter=",")

f=open('새 폴더\\finalresult.xml', 'w', encoding="utf-8-sig")
i=0
f.write('<?xml version="1.0" encoding="utf-8"?>\n\n')

f.write('<kunsancredit>\n')

for row in reader:
    f.write('\t<store id="'+str(i)+'">\n')
    f.write('\t\t<legalName>' + row[1] + '</legalName>\n')
    f.write('\t\t<category>' + row[2] + '</category>\n')
    f.write('\t\t<RoadNameAddress>' + row[3] + '</RoadNameAddress>\n')
    f.write('\t\t<AdministrativeNeighborhood>' + row[4] + '</AdministrativeNeighborhood>\n')
    f.write('\t\t<isThisKunsancreditid>' + row[5] + '</isThisKunsancreditid>\n')

    f.write('\t</store>\n')
    i=i+1
f.write('</kunsancredit>')
```

## 전처리한 CSV를 이하 xml로 변경

# 데이터 전처리

- 리스트마다 주소가 각기  
다름, -> 카카오맵 API  
를 통하여 주소 통일 작업

```
def lon2(name): # 주소로 검색한다.  
    try:  
        url = url = 'https://dapi.kakao.com/v2/local/search/address.json?query={address}'\br/>            .format(address=name)  
        headers = {"Authorization": "KakaoAK " + api_key}  
        result = json.loads(str(requests.get(url, headers=headers).text))  
        return result['documents'][0]['road_address']['address_name'], result['documents'][0]  
    except: # 가게가 안나왔다면?  
        return "X", "X"
```

✓ 0.5s

Python

```
# 건물이름을 담기 위한 데이터프레임  
templist2 = pd.DataFrame({'buliding_name':[]})
```

✓ 0.4s

Python

```
for i in tqdm_notebook(range(len(g))):  
    # 카테고리 중 음식점이 아니거나 X 표시인것, 전복 군산시가 아닌것들을 다시 검색  
    if(templist.iloc[i,1].split('>')!="음식점 " or templist.iloc[i, 1]=="X" \br/>        or templist.iloc[i, 0][:6] == "전복 군산시"):  
        t2 = lon2(g.iloc[i, 3]) # 도로명주소 검색  
        t = lon(g.iloc[2, 3]) # 카테고리 가져오기  
        templist.loc[i] = t2[0], t[1]  
        templist2.loc[i] = t2[1]
```

7m 2.5s

Python

C:\Users\우리집\AppData\Local\Temp\ipykernel\_20896\340300863.py:1: TqdmDeprecationWarning: This function will be removed in tqdm==5.0.0  
Please use `tqdm.notebook.tqdm` instead of `tqdm.tqdm\_notebook`  
for i in tqdm\_notebook(range(len(g))):

55%

2878/5251 [07:02<05:59, 6.60it/s]

# 데이터 전처리

	road_address	category
0	전북 군산시 옥서면 옥구저수지로 230	음식점 > 간식 > 제과,베이커리
1	인천 남동구 만수로75번길 32	음식점 > 중식 > 중국요리
2	전북 군산시 미장안7길 34	음식점 > 한식 > 국밥 > 전주현대옥
3	전북 군산시 동수송안길 25	음식점 > 술집 > 일본식주점
4	전북 군산시 월명로 241	음식점 > 한식 > 육류,고기 > 족발,보쌈 > 원할머니보쌈족발
...	...	...
5246	서울 동작구 양녕로 167	여행 > 관광,명소 > 문화유적 >릉,묘,총
5247	전북 군산시 해망로 79-1	음식점 > 한식 > 육류,고기 > 갈비
5248	X	X
5249	전북 익산시 무왕로15길 19	음식점 > 술집 > 와인바
5250	경북 문경시 영신로 3	가정,생활 > 유흥시설 > 유흥주점

5251 rows × 2 columns

수집 완료 및  
카테고리도 불러옴

카테고리는 split을 통해  
두번째 항목을 가져옴

# 데이터 전처리

208	208	11948	도서관	전북 군산시 계산로 87-6	술집	
243	243	11773	군산고래밥	전북 군산시 중앙로 196	건어울	서울유리
255	255	11704	에이앤디(AND)	전북 군산시 구영7길 95	카페	
262	262	11693	나들목 칠명점	전북 군산시 구영6길 83	한식	
263	263	11690	수목원	전북 군산시 구영5길 81	카페	수목원
268	268	11673	서래포구마을	전북 군산시 서래안2길 22	카페	서래포구마을
288	288	11581	뽕이네 뼈다귀탕	전북 군산시 진포로 177	한식	
289	289	11580	우리들 해장국	전북 군산시 진포로 177	한식	
300	300	11560	대청마루 초촌점	전북 군산시 초촌안5길 68	한식	가마솔우렁쌈
317	318	11494	고우담본점	전북 군산시 초촌로 82	한식	
327	328	11466	머시기수산회포차	전북 군산시 초촌로 168	술집	지우빌딩
330	331	11450	(유)다음 던킨도너츠	전북 군산시 초촌로 149	X	동군산병원
351	352	11378	제이피케이푸드	전북 군산시 초촌4길 34	X	
361	362	11364	암샘김밥 군산시	전북 군산시 초촌4길 3	X	하준빌딩
382	383	11320	부흥부페식당	전북 군산시 양촌1길 38	X	한빛산기
383	384	11317	바다애항기	전북 군산시 양안로 82	X	저섬오리
384	385	11316	가연	전북 군산시 양안로 51	결혼	한상구이
391	392	11297	햇살마루청동오리	전북 군산시 양안로 76	X	햇살마루
395	396	11288	다올바지락죽	전북 군산시 양안로 32	X	천남백화점
396	397	11287	예술냉면칼국수(	전북 군산시 양안로 32	X	천남백화점
405	406	11258	해바라기가맥	전북 군산시 양안로 107	X	아차울

불러들여온  
카테고리 수작업

# 데이터 전처리

```
import pandas as pd

kunsanfood = pd.read_csv('새 폴더\\사용 데이터\\Gunsan_food_1.csv')
kunsangift = pd.read_csv('새 폴더\\사용 데이터\\Gunsan-food(gift).csv')

a = 0
print(kunsanfood.iloc[1]['RoadNameAddress'])

for i in range(0, kunsanfood.shape[0]):
    irow = kunsanfood.iloc[i]
    for j in range(0, kunsangift.shape[0]):
        jrow = kunsangift.iloc[j]
        if (irow['RoadNameAddress'] == jrow['RoadNameAddress'] and irow['LegalName'] == jrow['LegalName']):
            kunsanfood.loc[i, 'isThisGunsanGift'] = '0'
            kunsanfood.loc[i, 'category'] = jrow['category']
            kunsanfood.loc[i, 'telephone'] = jrow['telephone']
            print(irow['isThisGunsanGift'])
            print(irow['category'], jrow['category'])
            a = 1
    if (a != 1):
        kunsanfood.loc[i, 'isThisGunsanGift'] = 'X'
    a = 0
    print(i)

kunsanfood.to_csv('새 폴더\\사용 데이터\\여부.csv')
```

```
import pandas as pd

kunsanfood = pd.read_csv('새 폴더\\사용 데이터\\Gunsan_food_1.csv')
kunsangood = pd.read_csv('새 폴더\\사용 데이터\\군산시장_맛집선정리스트.csv')

a = 0
print(kunsanfood.iloc[1]['RoadNameAddress'])

for i in range(0, kunsanfood.shape[0]):
    irow = kunsanfood.iloc[i]
    for j in range(0, kunsangood.shape[0]):
        jrow = kunsangood.iloc[j]
        if (irow['RoadNameAddress'] == jrow['RoadNameAddress'] and irow['LegalName'] == jrow['LegalName']):
            kunsanfood.loc[i, 'isThisHoleInTheWall'] = '0'
            kunsanfood.loc[i, 'category'] = jrow['category']
            a = 1
    if (a != 1):
        kunsanfood.loc[i, 'isThisHoleInTheWall'] = 'X'
    a = 0
    print(i)

kunsanfood.to_csv('새 폴더\\사용 데이터\\여부1.csv')
```

```
kunsanfood = pd.read_csv('새 폴더\\사용 데이터\\Gunsan_food_1.csv')
kunsangood = pd.read_csv('새 폴더\\사용 데이터\\군산시장_맛집선정리스트.csv')

a = 0

for i in range(0, kunsanfood.shape[0]):
    irow = kunsanfood.iloc[i]
    for j in range(0, kunsangood.shape[0]):
        jrow = kunsangood.iloc[j]
        if (irow['RoadNameAddress'] == jrow['RoadNameAddress'] and irow['LegalName'] == jrow['LegalName']):
            kunsanfood.loc[i, 'isThisGoodRestaurant'] = '0'
            kunsanfood.loc[i, 'category'] = jrow['category']
            a = 1
    if (a != 1):
        kunsanfood.loc[i, 'isThisGoodRestaurant'] = 'X'
    a = 0
    print(i)

kunsanfood.to_csv('새 폴더\\사용 데이터\\여부2.csv')
```

군산 음식점(모범음식점, 착한가격) 등  
여러가지가 적힌 csv를 비교 후  
주소 and 가게명이 일치하면 0 요소 추가

# 데이터 전처리

---

, 한식, 이가매생이수제비칼국수, 전북 군산시 조촌안3

29, 한식, 일해옥, 전북 군산시 구영7길 19

8,, 일해옥콩나물국밥, 전북 군산시 구영7길 19,

각기의 csv 파일에는 동일 가게이지만  
이름이 다른 가게들이다.

# 데이터 전처리

```
df.loc[df['RoadNameAddress'].str.contains("옥구|옥정|상평|미국|수신|오국|5269", 'na')].str.contains("옥구|남내|당복|향동|금성"), 'AdministrativeNeighborhood'] = '향동면'
df.loc[df['RoadNameAddress'].str.contains("화면|월연|금광|대정|고서|세정|5270", 'na')].str.contains("화면|월연|금광|대정|고서|세정|5270", 'na')].str.contains("임피|옥정|상평|미국|수신|오국|5271", 'na')].str.contains("임피|옥정|상평|미국|수신|오국|5271", 'na')].str.contains("서수|옥동|관원|마룡|화동|금암|5edit", 'na')].str.contains("서수|옥동|관원|마룡|화동|금암|5edit", 'na')].str.contains("대야|산울|보덕|지경|북고|광교|5edit", 'na')].str.contains("대야|산울|보덕|지경|북고|광교|5edit", 'na')].str.contains("아동|은회|아산|통서|발산|옥석|5edit", 'na')].str.contains("아동|은회|아산|통서|발산|옥석|5edit", 'na')].str.contains("성산|성덕|동덕|고동|도암|여암|여천|연곡|영동사", 'na')].str.contains("성산|성덕|동덕|고동|도암|여암|여천|연곡|영동사", 'na')].str.contains("나로|장상|육군|부곡|주곡|서포리"), 'AdministrativeNeighborhood'] = '나로면'
df.loc[df['RoadNameAddress'].str.contains("옥도|개야도|연도|여흥도|선유도|무녀도|신서도|여미도|장자도|관리도|달도|비안도"), 'AdministrativeNeighborhood'] = '옥도면'
df.loc[df['RoadNameAddress'].str.contains("옥서|옥동|선연리"), 'AdministrativeNeighborhood'] = '옥서면'
df.loc[df['RoadNameAddress'].str.contains("해암|신룡|금동"), 'AdministrativeNeighborhood'] = '해신동'
df.loc[df['RoadNameAddress'].str.contains("장미|월당|신정|중앙로1|영곡|장미|선암|돈골|장성|당산|송정|가죽"), 'AdministrativeNeighborhood'] = '월당동'
df.loc[df['RoadNameAddress'].str.contains("오룡|금광|삼학"), 'AdministrativeNeighborhood'] = '삼학동'
df.loc[df['RoadNameAddress'].str.contains("신룡|송정|문화"), 'AdministrativeNeighborhood'] = '신룡동'
df.loc[df['RoadNameAddress'].str.contains("중앙로2|영|신영|옥성|평화|영|금암동"), 'AdministrativeNeighborhood'] = '중앙동'
df.loc[df['RoadNameAddress'].str.contains("중앙로3|대영|장제|미원|흥남|서동남동"), 'AdministrativeNeighborhood'] = '동남동'
df.loc[df['RoadNameAddress'].str.contains("조촌|경장"), 'AdministrativeNeighborhood'] = '조촌동'
df.loc[df['RoadNameAddress'].str.contains("경암동"), 'AdministrativeNeighborhood'] = '경암동'
df.loc[df['RoadNameAddress'].str.contains("구암|나룡"), 'AdministrativeNeighborhood'] = '구암동'
df.loc[df['RoadNameAddress'].str.contains("가정|사정"), 'AdministrativeNeighborhood'] = '가정동'
df.loc[df['RoadNameAddress'].str.contains("수송|미정|지곡"), 'AdministrativeNeighborhood'] = '수송동'
df.loc[df['RoadNameAddress'].str.contains("나운1|나운 1"), 'AdministrativeNeighborhood'] = '나운 1동'
df.loc[df['RoadNameAddress'].str.contains("나운|개사"), 'AdministrativeNeighborhood'] = '나운동'
df.loc[df['RoadNameAddress'].str.contains("나운2|나운 2|나운5길"), 'AdministrativeNeighborhood'] = '나운 2동'
df.loc[df['RoadNameAddress'].str.contains("나운3|나운 3|미룡|신관"), 'AdministrativeNeighborhood'] = '나운 3동'
df.loc[df['RoadNameAddress'].str.contains("소룡|오식도|비룡도"), 'AdministrativeNeighborhood'] = '소룡동'
df.loc[df['RoadNameAddress'].str.contains("신복|나초"), 'AdministrativeNeighborhood'] = '미정동'
```

많이 전처리를 하였지만  
상단 표처럼 빠진 부분도 있었음.  
(5000개 중 20개)

# 데이터 전처리

Custom text transform on column telephone

Expression Language Python / Jython

return "0"+value No syntax error.

**Preview** History Starred Help

5.	null	Error: Traceback (most recent call last): File "<string>", line 2, in __temp_1285185590__ TypeError: cannot concatenate 'str' and 'NoneType' objects
6.	634660958	0634660958
7.	634665284	0634665284
8.	634665556	0634665556
9.	634663858	0634663858
10.	634468172	0634468172

On error ☒ keep original ☐ Re-transform up to  times until no change  
☐ set to blank  
☐ store error

OK Cancel

사용하는 프레임워크를  
번갈아가며 사용하니  
전화번호도 0이 없어졌다.

▼ telephone
634653555
634660958
634665284
634665556
634663858
634468172



---

## 2

# DB 설계

---

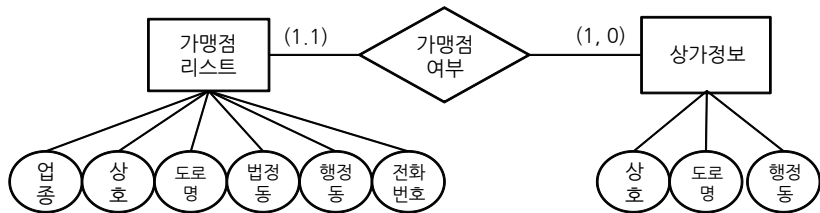
- 테이블 구조
- ER, IE 다이어그램
- 정규화 과정

# 요구사항

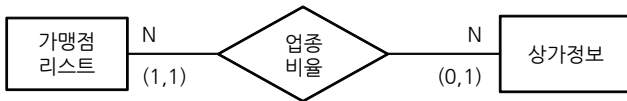
- 데이터 열이름은 변동사항이 있을 수 있기에 기재하지 않음.
- 데이터 모델 수정사항으로 요구사항 변동이 있을 수 있음.

- 1. 상품권 가맹점()은 번호(), 업종() 상호명(), 도로명주소(), 법정동(), 행정동(), 전화번호()를 가진다
- 2. 군산시 상가정보()는 번호(), 업종(), 상호명(), 도로명주소(), 법정동(), 행정동() 전화번호()를 가진다
- 3. 군산시 상가정보()와 상품권 가맹점()의 전화번호는 NULL 값이 허용된다.
- 4. 군산시 상호리스트 중 상품권 가맹점 리스트()에 소속될 수 있는데, 이때 각 업종별로 가맹점비율()이 기록된다.
- 5. 군산시 상가정보에는 업종()이 무엇인지 소매업, 보건업, 숙박업 등으로만 기록하게 되어있다.
- 6. 군산시 상가정보()와 상품권 가맹점()은 도로명주소()를 참고해 법정동을 참고한다.
- 7. 군산시 상가정보()와 상품권 가맹점()의 법정동()을 관할하는 행정동()을 기재한다.

# 데이터 모델링



# 데이터 모델링



가맹점 리스트 (가맹점번호, 업종, 상호명, 도로명주소, 법정동, 행정동, 전화번호)  
상가정보 (상가번호, 업종, 상호명, 도로명주소, 법정동, 행정동, 전화번호)  
업종비율 (업종번호, 업종(FK), 상품권업종비율)

예) 상가정보의 업종별로 군산사랑상품권 가맹점인 비율이 얼마나 되는가?

## 전라북도 군산시 업체별 리스트

제조업공장현황		
U	회사명	VARCHAR(30)
	대표자명	VARCHAR(10)
FK	공장명	VARCHAR(150)
	주소	VARCHAR(15)
N	전화번호	INTEGER
N	팩스번호	INTEGER
	생년월일	VARCHAR(50)

관광호텔		
U	업체명	VARCHAR(30)
FK	도로명주소	VARCHAR(150)
	전화번호	INTEGER
N	영업권	INTEGER
	객실수	INTEGER

제과점		
PK	연번 업종명 업소명	INTEGER VARCHAR(15) VARCHAR(30)
FK	소재지_도로명주소	VARCHAR(150)
	소재지_지번주소	VARCHAR(150)

약국		
U	약국명	VARCHAR(150)
	약국전화번호	INTEGER
FK	약국소재지	VARCHAR(150)
	약국구분	VARCHAR(10)

안경업소		
U	안경업소명	VARCHAR(30)
FK	사업장소재지 도로명	VARCHAR(150)
N	사업장전화번호	INTEGER

의원		
PK	순번 의분기 의분기 의분기	INTEGER VARCHAR(30) VARCHAR(150)
N	의분기 의분기 의분기	INTEGER

치과의원현황		
PK	순번	INTEGER
	의료기관명	VARCHAR(30)
FK	의료기관주소 도로명	VARCHAR(150)
	의료기관전화번호	INTEGER

대형마트 및 SSM		
PK	승인번호	INTEGER
	상호명	VARCHAR(15)
	주소	VARCHAR(30)
FK	전화번호	VARCHAR(150)
		INTEGER

정신건강의학과		
PK	순번	INTEGER
	의뢰기관명	VARCHAR(15)
	의뢰기관주소	VARCHAR(30)
FK	의뢰자명	VARCHAR(150)
N	의뢰자전화번호	INTEGER

노래연습장		
PK	습비 업종명	INTEGER VARCHAR(15)
FK	상호명	VARCHAR(30)
	소재지 도로명주소	VARCHAR(150)
	소재지 지번주소	VARCHAR(150)

문화시설		
PK	구분	INTEGER
	시설명	VARCHAR(30)
FK	도로명주소	VARCHAR(150)
	전화번호	INTEGER

식품제조업체		
U	업소명	VARCHAR(30)
FK	소재지 도로	VARCHAR(150)
	소재지 지번주소	VARCHAR(150)
N	소재지 전화	INTEGER

병원		
PK	의뢰기관명	INTEGER
	의뢰기관명	VARCHAR(30)
FK	의뢰기관주소	VARCHAR(150)
N	의뢰기관전화번호	INTEGER
	병상수	INTEGER
	병상수	INTEGER
	비고	INTEGER

죽석판매회사가공업		
PK	연번	INTEGER
	업소명	VARCHAR(30)
FK	소재지_도로명	VARCHAR(150)
	소재지_지번	VARCHAR(150)

The diagram illustrates the relationships between four tables in a database:

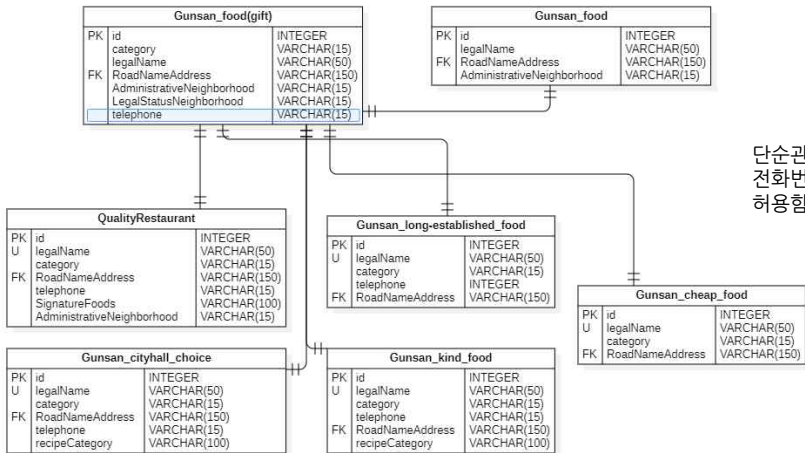
- 전라북도 군산시 상가현황** (Jeollabuk-do Gunsan-si Sanggahyeonhwang):
  - PK: 상가번호 (Sanggabomobu)
  - FK: 도로명주소 (Doromyeongjuso)
  - Attributes: INTEGER, VARCHAR(30), VARCHAR(15), VARCHAR(10), VARCHAR(150), VARCHAR(150), INTEGER.
- 군산사랑상품권 가맹점** (Gunsan Sarang Sangpumgwin Gamaengjeom):
  - PK: 점번호 (Jeombomobu)
  - FK: 상가번호 (Sanggabomobu)
  - N: 상가번호 (Sanggabomobu)
  - Attributes: INTEGER, VARCHAR(30), VARCHAR(15), VARCHAR(10), VARCHAR(150), VARCHAR(20).
- 착한가격업소** (Chahan-gagagyeopso):
  - PK: 업소번호 (Eopso-bomobu)
  - FK: 도로명주소 (Doromyeongjuso)
  - Attributes: INTEGER, VARCHAR(15), VARCHAR(30), VARCHAR(150).
- 모범음식점** (Mobeom-eumsikjeom):
  - PK: 연번 (Eonbeon)
  - FK: 업소번호 (Eopso-bomobu)
  - Attributes: INTEGER, VARCHAR(30), VARCHAR(150), INTEGER, VARCHAR(50).

Relationships are indicated by lines with crow's foot notation:

- 전라북도 군산시 상가현황 (PK) to 군산사랑상품권 가맹점 (FK): 1:1 relationship.
- 전라북도 군산시 상가현황 (FK) to 군산사랑상품권 가맹점 (N): 1:N relationship.
- 전라북도 군산시 상가현황 (FK) to 착한가격업소 (FK): 1:1 relationship.
- 군산사랑상품권 가맹점 (FK) to 모범음식점 (FK): 1:1 relationship.
- 착한가격업소 (FK) to 모범음식점 (FK): 1:1 relationship.

DB(군산사랑상품권 가맹점)





단순관계를 1:1로 그림.  
전화번호 부분은 NULL값을 허용함.



# 데이터 삽입

Table Data Import

**Import Data**

The following tasks will now be performed. Please monitor

☒ Prepare Import  
☐ Import data file

Row import failed with error: ("Incorrect integer value: " for column 'telephone' at row 1", 1366)

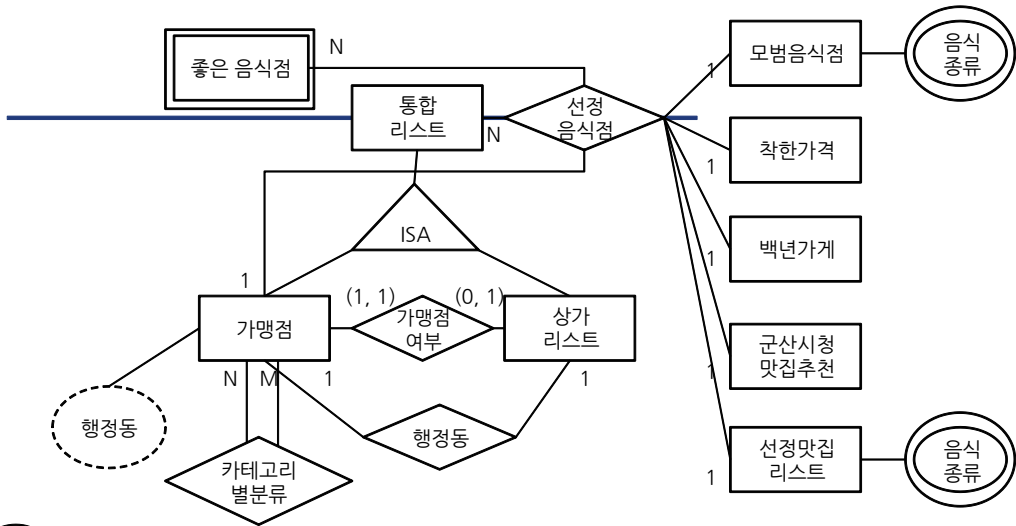
Message Log

- Data import
- Data import
- Row import failed with error: ("Incorrect integer value: " for column 'telephone' at row 1", 1366)
- Data import
- Row import failed with error: ("Incorrect integer value: " for column 'telephone' at row 1", 1366)
- Row import failed with error: ("Incorrect integer value: " for column 'telephone' at row 1", 1366)
- Data import
- Row import failed with error: ("Incorrect integer value: " for column 'telephone' at row 1", 1366)
- Data import
- Row import failed with error: ("Incorrect integer value: " for column 'telephone' at row 1", 1366)
- Data import
- Data import

	id	category	LegalName	RoadNameAddress	AdministrativeNeighborhood	LegalStatusNeighborhood	telephone
	5107	일식	불맛계대로냉면스50	전북 군산시 대학로 533	나운3동	미용동	
	5109	한식	한복수육국밥	전북 군산시 대학로 535	나운3동	미용동	
	5114	카페	커피면회소(군산대점)	전북 군산시 대학로 549	나운3동	미용동	
	5144	간식	파리바게트	전북 군산시 향릉로 12	나운3동	미용동	
	5156	치킨	조선치킨 군산대점	전북 군산시 미룡로 15	나운3동	미용동	
	5157	술집	변학도	전북 군산시 미룡로 45	나운3동	미용동	
	5172	카페	아트	전북 군산시 미룡로 12	나운3동	미용동	
	5176	한식	가람영태	전북 군산시 미룡로 12	나운3동	미용동	
	5218	분식	또또꼬마김밥 미용점	전북 군산시 미룡로 46	나운3동	미용동	
	5246	치킨	맥스칸 미용점	전북 군산시 미계1길 10	나운3동	미용동	
	5247	중식	야미마라탕	전북 군산시 미계1길 11	나운3동	미용동	
	5248	치킨	고촌치킨 군산대점	전북 군산시 미계1길 12	나운3동	미용동	

앞서 전처리한 csv 데이터를 불러옴





○ 속성은 특이값 제외 생략(앞면 참고)

# 조인 테이블

SQL File 5\* x qualityrestaurant gunsan\_food\_gift gunsan\_cityhall\_ch

Limit to 1000 rows

```
1 • select distinct o.legalName, o.RoadNameAddress
2 from gunsan_food o, gunsan_food_gift t
3 where o.RoadNameAddress = t.RoadNameAddress;
```

Result Grid Filter Rows: Export: Wrap Cell C

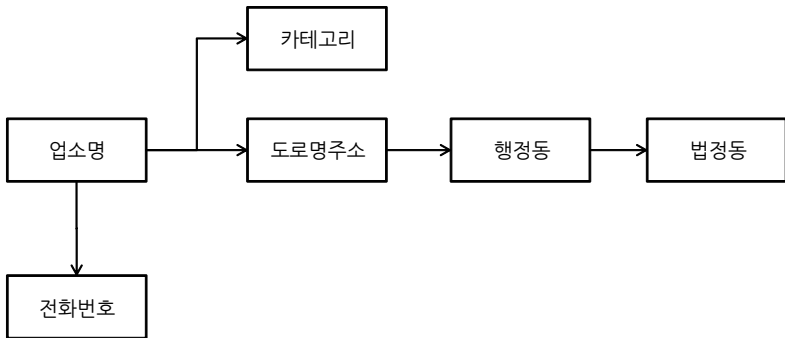
legalName	RoadNameAddress
현대옥	전북 군산시 상나운3길 29-3
청담이상	전북 군산시 동수송안길 25
시루향기	전북 군산시 수송안길 23
꽃보다소	전북 군산시 칠성로 172
더갈비	전북 군산시 공단대로 601
코코호도군산나우점	전북 군산시 하나우로 45

테이블 조인으로  
군산사랑상품권 가맹점 찾기가  
가능해짐

# 정규화

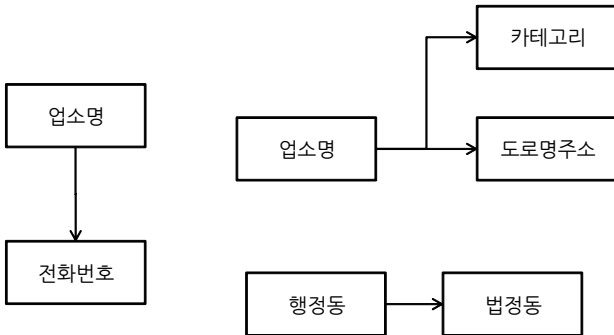
---

- 군산사랑상품권 테이블



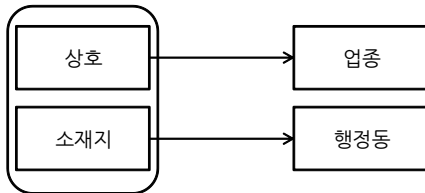
# 정규화

---



# 정규화 진행(제2정규화)

군산사랑상품권 가맹점		
PK	번호	INTEGER
	상호	VARCHAR(30)
	업종	VARCHAR(15)
FK	소재지	VARCHAR(150)
U	행정동	VARCHAR(10)
N	사업장전화번호	INTEGER



업종

번호	상호	카테고리
1	라라코스트	양식
2	명륜진사갈비 군산점	한식
3	롯데리아	패스트푸드

행정동

소재지	행정동	법정동
군산시 나운2동 별별로	나운2동	나운2동
군산시 미룡동 대학로~	나운3동	미룡동
군산시 미장동 좋은도로	수송동	미장동

# 정규화 진행

행정동

상호	소재지	행정동
CU 군산00점	군산시 별별로 (나운2동)	나운2동
명륜진사갈비 군산점	군산시 대학로~ (미룡동)	나운3동
군산사랑게스트하우스	군산시 좋은도로 (미장동)	수송동

행정동이란?

시의 하위 행정구역으로서  
여러개의 법정동을 하나의 행정동 기관을  
설치해 관할하는 행정구역이다.

예) 관할 법정동

(미장동, 수송동, 지곡동) -> 수송동

(나운3동, 미룡동, 개사동, 신관동) -> 나운3동

(비응도동, 소룡동, 오식도동) -> 소룡동



# 구현예정

업종

번호	상호	업종
1	CU 군산OO점	소매업
2	명륜진사갈비 군산점	음식점업
3	군산사랑게스트하우스	숙박업

행정동

소재지	행정동
군산시 별별로 (나운2동)	나운2동
군산시 대학로~ (미룡동)	나운3동
군산시 좋은도로 (미장동)	수송동

군산사랑상품권 업종별 비율

업종	비율
소매업	0.4
음식점업	0.3
숙박업	0.1

군산사랑상품권 행정동별 비율

행정동	비율
나운2동	0.4
나운3동	0.3
수송동	0.1

# 제 1정규화

- 맛집선정 리스트에 속성값이 원자성을 충족하지 않음  
따라서 두 차례로 나눈 투플을 만듦

	id	legalName	category	RoadNameAddress	telephone	recipeCategory	
	16	새만금횃집	한식	전북 군산시 비응안7길 13	063-464-1001	박대탕	
	17	서진해물곱돌솔밥	한식	전북 군산시 구영6길 72	063-442-2282	해물곱돌솔밥	
	18	수원성	한식	전북 군산시 월명로 281	063-453-0002	돼지갈비, 갈비탕	
	19	스위트인디아	아시아...	전북 군산시 거석길 36	063-445-8522	인도커리, 탄두라...	
	20	신돈쪽갈비	한식	전북 군산시 경암로 48	063-442-9903	갈비	
	--	--	--	--	--	--	
16		새만금횃집	전북 군산시 비응안7길 13	063-464-1001	박대탕	한식	
17		서진해물곱돌솔밥	전북 군산시 구영6길 72	063-442-2282	해물곱돌솔밥	한식	
18		수원성	전북 군산시 월명로 281	063-453-0002	돼지갈비	한식	
19		수원성	전북 군산시 월명로 281	063-453-0002	갈비탕	한식	
20		스위트인디아	전북 군산시 거석길 36	063-445-8522	인도커리	아시아...	
21		스위트인디아	전북 군산시 거석길 36	063-445-8522	탄두라음식	아시아...	
22		신돈쪽갈비	전북 군산시 경암로 48	063-442-9903	갈비	한식	
--		--	--	--	--	--	



# 제 2정규화

id	category	legalName	RoadNameAddress	AdministrativeNeighborhood	LegalStatusNeighborhood	telephone
9	한식	힐링과메기전문점	전북 군산시 중정길 11	월명동	개복동	634456878
10	카페	인디무드	전북 군산시 중정길 14	월명동	개복동	
17	한식	운정식당	전북 군산시 중정길 8-2	월명동	개복동	634460891

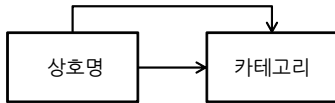
상호명 테이블(카테고리)

id	category	legalName
12605	카페	논
12602	카페	카페마실
12600	한식	자매식당
12598	술집	여그여 동무야
12596	분식	해바라기분식&식당
12595	카페	회현커피

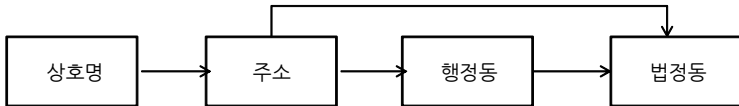
RoadNameAddress	AdministrativeNeighborhood	LegalStatusNeighborhood
전북 군산시 회현면 회현초교1길 47	회현면	회현면
전북 군산시 회현면 회현로 182	회현면	회현면
전북 군산시 회현면 회현로 180-1	회현면	회현면
전북 군산시 회현면 회현로 178	회현면	회현면
전북 군산시 회현면 회현로 175	회현면	회현면

## 제 2정규화

---

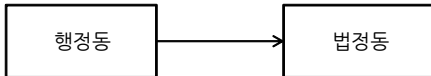


상호명 테이블(카테고리)



주소 테이블(행정동)

# 제3정규화

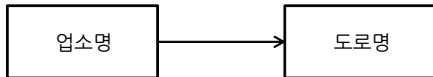


행정동	법정동
월명동	개북동
개정동	개정동
개정동	개정면
경암동	경암동
조촌동	경장동
구암동	구암동
삼학동	금광동
해신동	금동
중앙동	금암동

RoadNameAddress	AdministrativeNeighborhood
전북 군산시 회현면 회현초교1길 47	회현면
전북 군산시 회현면 회현로 182	회현면
전북 군산시 회현면 회현로 180-1	회현면
전북 군산시 회현면 회현로 178	회현면
전북 군산시 회현면 회현로 175	회현면

# 제3정규화

LegalName	RoadNameAddress
힐링과메기전문점	전북 군산시 중정길 11
인디무드	전북 군산시 중정길 14
운정식당	전북 군산시 중정길 8-2
꿀밥	전북 군산시 동개정길 20
더블미	전북 군산시 동개정길 20
밥머꼬달머꼬	전북 군산시 번영로 345-4
개정동커피집	전북 군산시 번영로 407
개정황금가든	전북 군산시 번영로 411
군산커피마을	전북 군산시 개정면 개정들4길 63
오래들	전북 군산시 개정면 개정문화길 61-19
(농)(유)가소앤...	전북 군산시 개정면 구암로 206



# 기타

	행정동	count(*)/3614 * 100
	수송동	23.2983
	나운동	18.7327
	조촌동	8.4671
	소룡동	7.9137
	월명동	5.7277
	중앙동	5.2297
	미성동	4.7593

	category	count(*)/3614 * 100
	한식	43.9679
	카페	14.2501
	음식점	1.9092
	양식	2.9054
	중식	5.1467
	술집	8.4947
	치킨	5.8661

---

# 3

# XML 설계

---

- XML 트리
- 공통어휘
- 통합 스키마

# 공통어휘

## 군산사랑상품권 가맹점

필드명	설명	통합스키마 적용여부	매핑어휘
legalName	상호명	O	<a href="#">schema:legalName</a>
address	소재지	O	<a href="#">schema:address</a>
telephone	전화번호	O	<a href="#">schema:telephone</a>
RoadNameAddress	도로명주소	O	<a href="#">juso:RoadNameAddress</a>
EastAsianAddress	지번주소	X	<a href="#">juso:EastAsianAddress</a>
AdministrativeNeighborhood	행정동	O	<a href="#">juso:AdministrativeNeighborhood</a>

# 공통어휘

## 군산상가 리스트

필드명	설명	통합스키마 적용여부	매핑어휘
legalName	상호명	O	<a href="#">schema:legalName</a>
address	소재지	O	<a href="#">schema:address</a>
telephone	전화번호	O	<a href="#">schema:telephone</a>
RoadNameAddress	도로명주소	O	<a href="#">juso:RoadNameAddress</a>
EastAsianAddress	지번주소	O	<a href="#">juso:EastAsianAddress</a>



# 공통어휘

기타 통합 리스트(백년가게, 착한가격 등)

필드명	설명	통합스키마 적용여부	매핑어휘
category	음식 카테고리	O	schema:category
SignatureFoods	식당 대표메뉴	X	schema:SignatureFoods
recipeCategory	식당 음식종류	X	schema:Category

# 공통어휘

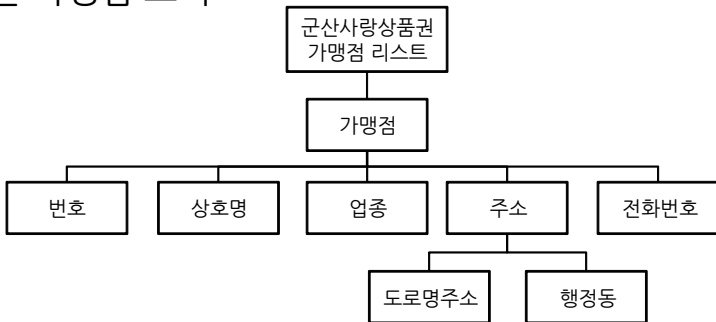
---

- <https://schema.org/Store> (legalName, address, telephone)
- <http://rdfs.co/juso/kr/RoadNameAddress> (도로명주소)
- <http://rdfs.co/juso/kr/EastAsianAddress> (지번주소)
- <http://rdfs.co/juso/kr/AdministrativeNeighborhood> (행정동)
- <https://schema.org/Hotel>에 numberOfRooms, starRating
- <https://schema.org/Organization>에 faxNumber(팩스 번호), founder(대표자), owns(생산품)
- <https://schema.org/Hospital> (대형병원)
- <https://schema.org/MedicalClinic> (의원)
- <https://schema.org/Dentist> (치과)
- <https://schema.org/Pharmacy> (약국)

# XML 트리 (前)

---

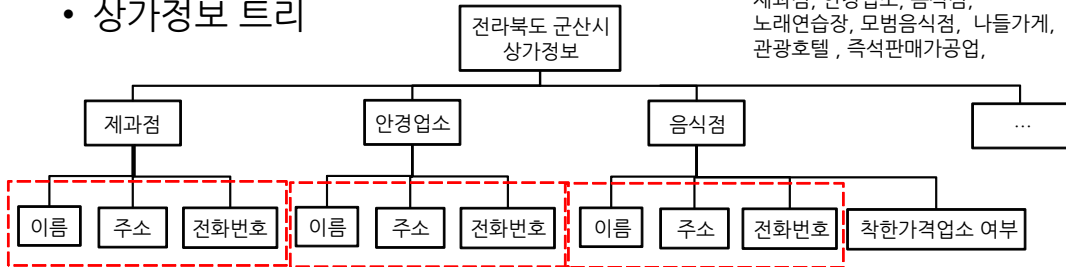
- 상품권 가맹점 트리



# XML 트리 (前)

- 상가정보 트리

제과점, 안경업소, 음식점,  
노래연습장, 모범음식점, 나들가게,  
관광호텔, 즉석판매가공업,



그룹으로 묶을 예정

# XML 트리 (前)

- 업종별 XML 트리

전라북도 군산시  
의료시설 현황

병원, 약국, 의원, 정신건강의학과, 치과의원,  
한의원

병원

한의원

...

번호

이름

주소

전화번호

병상수

병동수

번호

이름

주소

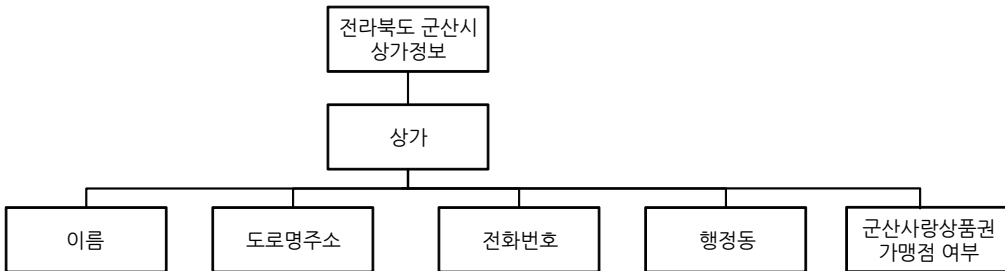
전화번호

이는 상가정보 XML과도 공통부분임

# 적용예정

---

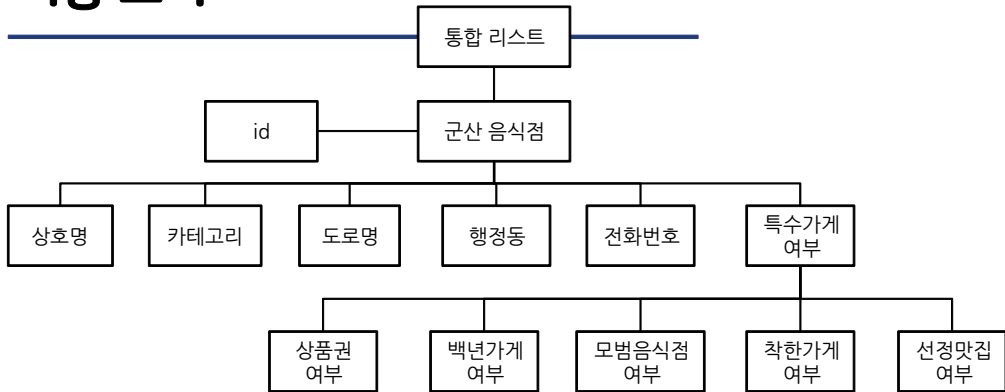
- 통합 스키마 예상(군산상가 리스트)



# 최종 트리



# 최종 트리





# 통합 스키마

---

```
xsi:noNamespaceSchemaLocation='finalresult.xsd'>
  <store id="1">
    <legalName>준베이커리</legalName>
    <category>미분류됨</category>
    <RoadNameAddress>전북 군산시 옥서면 옥구저수지로 230</RoadNameAddress>
    <AdministrativeNeighborhood>옥서면</AdministrativeNeighborhood>
    <telephone>-</telephone>
    <IsThisSpecial>
      <IsThisGunsanGift>X</IsThisGunsanGift>
      <IsThisMillenniumStore>X</IsThisMillenniumStore>
      <IsThisHoleInTheWall>X</IsThisHoleInTheWall>
      <IsThisChackanStore>X</IsThisChackanStore>
      <IsThisGoodRestaurant>X</IsThisGoodRestaurant>
    </IsThisSpecial>
  </store>
```

```

<xs:element name="kunsanGift">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="store" type="storeType" maxOccurs="unbounded"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>

<xs:complexType name="storeType">
  <xs:sequence>
    <xs:element name="legalName" type="xs:string"/> <!--이름-->
    <xs:element name="category" type="categoryType" minOccurs="0"/> <!--업종-->
    <xs:element name="RoadNameAddress" type="xs:string"/> <!--도로명주소-->
    <xs:element name="AdministrativeNeighborhood" type="AdministrativeNeighborhoodType"/> <!--행정동-->
    <xs:element name="telephone" type="telephoneType"/> <!--전화번호-->
    <xs:element name="IsThisSpecial" type="IsThisSpecialType"/>
  </xs:sequence>
  <xs:attribute name="id" type="xs:integer"/> <!--아이디-->
</xs:complexType>

<xs:complexType name="IsThisSpecialType">
  <xs:sequence>
    <xs:element name="IsThisGunsanGift" type="OXType"/> <!--군산사랑상품권 가맹점 여부-->
    <xs:element name="IsThisMillenniumStore" type="OXType"/> <!--백년가게 여부-->
    <xs:element name="IsThisHoleIntheWall" type="OXType"/> <!--맛집 여부-->
    <xs:element name="IsThisChackanStore" type="OXType"/> <!--착한가격업소 여부-->
    <xs:element name="IsThisGoodRestaurant" type="OXType"/> <!--모범식당-->
  </xs:sequence>
</xs:complexType>

<xs:simpleType name="OXType">
  <xs:restriction base="xs:string">
    <xs:enumeration value="0"/>
    <xs:enumeration value="X"/>
  </xs:restriction>
</xs:simpleType>

```

## 기본적인 XML 스키마

```

<xs:simpleType name="categoryType">
  <xs:restriction base="xs:string">
    <xs:enumeration value="간식"/>
    <xs:enumeration value="도시락"/>
    <xs:enumeration value="분식"/>
    <xs:enumeration value="튀김"/>
    <xs:enumeration value="샤브샤브"/>
    <xs:enumeration value="술집"/>
    <xs:enumeration value="슈퍼마켓"/>
    <xs:enumeration value="식품"/>
    <xs:enumeration value="식품판매"/>
    <xs:enumeration value="아시아음식"/>
    <xs:enumeration value="야식"/>
    <xs:enumeration value="양식"/>
    <xs:enumeration value="유흥시설"/>
    <xs:enumeration value="음식점"/>
    <xs:enumeration value="일식"/>
    <xs:enumeration value="중식"/>
    <xs:enumeration value="치킨"/>
    <xs:enumeration value="카페"/>
    <xs:enumeration value="패스트푸드"/>
    <xs:enumeration value="퓨전 요리"/>
    <xs:enumeration value="한식"/>
    <xs:enumeration value="패스트푸드"/>
    <xs:enumeration value="패스트푸드"/>
    <xs:enumeration value="미분류됨"/>
  </xs:restriction>
</xs:simpleType>

```

<- 카테고리  
행정동 ->

```

<xs:simpleType name="AdministrativeNeighborhoodType">
  <xs:restriction base="xs:string">
    <xs:enumeration value="옥구읍"/>
    <xs:enumeration value="옥산면"/>
    <xs:enumeration value="회현면"/>
    <xs:enumeration value="임피면"/>
    <xs:enumeration value="서수면"/>
    <xs:enumeration value="대야면"/>
    <xs:enumeration value="개정면"/>
    <xs:enumeration value="성산면"/>
    <xs:enumeration value="나포면"/>
    <xs:enumeration value="옥도면"/>
    <xs:enumeration value="옥서면"/>
    <xs:enumeration value="해신동"/>
    <xs:enumeration value="월명동"/>
    <xs:enumeration value="삼학동"/>
    <xs:enumeration value="신풍동"/>
    <xs:enumeration value="중앙동"/>
    <xs:enumeration value="홍남동"/>
    <xs:enumeration value="조촌동"/>
    <xs:enumeration value="경암동"/>
    <xs:enumeration value="구암동"/>
    <xs:enumeration value="개정동"/>
    <xs:enumeration value="수송동"/>
    <xs:enumeration value="나운동"/>
    <xs:enumeration value="소룡동"/>
    <xs:enumeration value="미성동"/>
  </xs:restriction>
</xs:simpleType>

```

# 제약조건

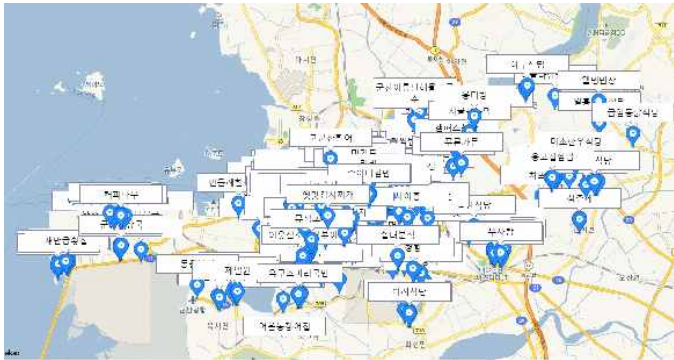
---

```
<xs:simpleType name="telephoneType">
  <xs:restriction base="xs:string">
    <xs:pattern value="063\d{3,4}\d{4}"/>
    <xs:pattern value="-"/>
  </xs:restriction>
</xs:simpleType>
```

정규 표현식 전화번호

# 웹 시각화

카카오맵 API 활용한 지도 마커 표시,  
군산사랑상품권 가맹점 수 등 그래프 시각화  
가맹점 등의 필터 여부 등을 확인 가능



# 웹 시각화

군산 전역에 상가를 볼 수 있음.  
(노란색 음영은 군산사랑상품권  
가맹점을 뜻함)

## 군산의 음식점 현황

- 식당가게주소 보기
- 음식점 보기
- 군산사랑상품권 가맹점 보기
- 맛집 보기
- 배설가게 보기
- 찜통 보기

주소



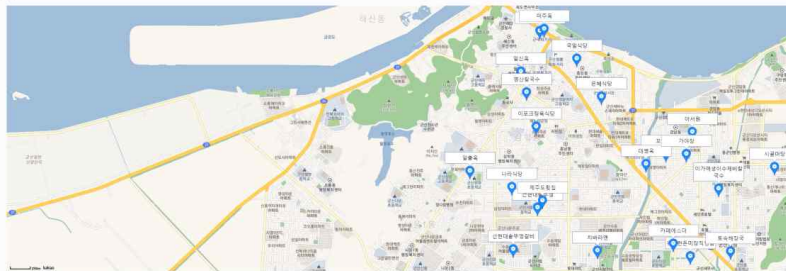
114	한양	음식	한양 군산사랑상품권	주요점	-	0
115	한양	음식	한양 군산사랑상품권	주요점	-	1
116	한양	음식	한양 군산사랑상품권	주요점	-	1
117	한양	음식	한양 군산사랑상품권	주요점	3040190	0
118	한양	음식	한양 군산사랑상품권	주요점	-	1
119	한양	음식	한양 군산사랑상품권	주요점	3040190	0
120	한양	음식	한양 군산사랑상품권	주요점	-	1
121	한양	음식	한양 군산사랑상품권	주요점	3040190	0

구분	주소명	주소	도로명주소	영양분	한도번호	군산사랑상품권
1	한양가게	의정부동	한양 군산 사랑 상품권 210	주요점	-	1

# 웹 시각화

## 군산의 착한가격업소 현황

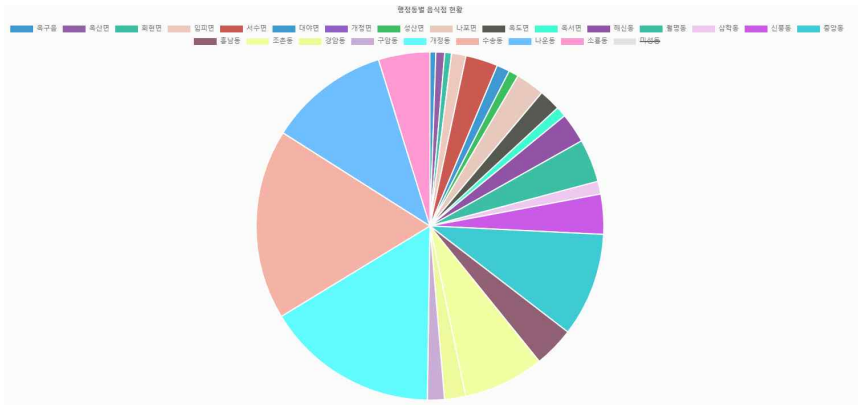
- 착한가격업소 보기
- 로컬식당 보기
- 군산사랑상품권 가맹점 보기
- 맛집 보기
- 백년가게 보기
- 천재 보기
- 초점



구분	업소명	업종	도로명주소	명칭	전화번호	군산사랑상품권
43%	봉산국	한식	전북 군산시 구경1길 26	송정동	0634111904	O

착한가격 업소,  
백년가게 등 볼수가 있음

# 웹 시각화





# 웹 시각화

```
tr:has(td[IsThisGunsanGift="0"]){  
  background-color: rgb(255, 251, 124) !important;  
  font-weight: bolder;  
}
```

군산사랑상품권 가맹점이 맞을 경우  
tr 부분의 배경을 노랗게 만들기

```
<script type="text/javascript">  
var gunsanStoreCount = <?> $gunsanStoreCount ?>;  
var gunsanGiftCount = <?> $gunsanGiftCount ?>;  
var nonGunsanGiftCount = <?> $nonGunsanGiftCount ?>;  
var MillenniumStoreCount = <?> $MillenniumStoreCount ?>;  
var HoleInTheWallCount = <?> $HoleInTheWallCount ?>;  
var ChackanStoreCount = <?> $ChackanStoreCount ?>;  
var GoodRestaurantCount = <?> $GoodRestaurantCount ?>;  
  
new Chart(document.getElementById("bar-chart"), {  
  type: 'bar',  
  data: {  
    labels: ["전체 음식점", "군산사랑상품권 가맹점", "대기업점", "백년가게", "최한가죽점소", "맛집", "모범음식점"],  
    datasets: [  
      {  
        label: "장모수",  
        backgroundColor: ["#3e95cd", "#8e5ea2", "#3cba9f", "#e8c3b9", "#c45850", "#c45850", "#c45850"],  
        data: [gunsanStoreCount, gunsanGiftCount, nonGunsanGiftCount, MillenniumStoreCount, HoleInTheWallCount, ChackanStoreCount, GoodRestaurantCount]  
      }  
    ]  
  },  
  options: {  
    legend: { display: false },  
    title: {  
      display: true,  
      text: "군산 음식점 현황 (2022)"  
    }  
  }  
});  
</script>
```

그래프 그리기

# 웹 시각화

```
tr:nth-child(even) {background-color: rgb(236, 236, 236)}
```

짝수 tr 색을 바꿔주는 코드

```
<div id="splited">
  <a href="http://localhost/splited/ChackanStore.php">착한가격업소 필터</a>
  <a href="http://localhost/splited/GoodRestaurant.php">모범식당 필터</a>
  <a href="http://localhost/splited/GunsanGift.php">군산사랑상품권 가맹점 필터</a>
  <a href="http://localhost/splited/HoleIntheWall.php">맛집 필터</a>
  <a href="http://localhost/splited/MillenniumStore.php">백년가게 필터</a>
</div>
```

특정 가게 목록을  
보러 가게하기

---

# 4

# 마무리

---

- 숫자로 보는 우리의 프로젝트
- 로드맵 / 간트차트
- 마무리

# 숫자로 보는 우리의 프로젝트

군산사랑상품권  
가맹점 리스트

12507 16243

군산상가 수

군산사랑상품권  
가맹점 (음식점업)

3646 5251

군산상가 수  
(음식점업)

13

군산시 내  
백년가게 음식점

57

군산시  
모범음식점

21

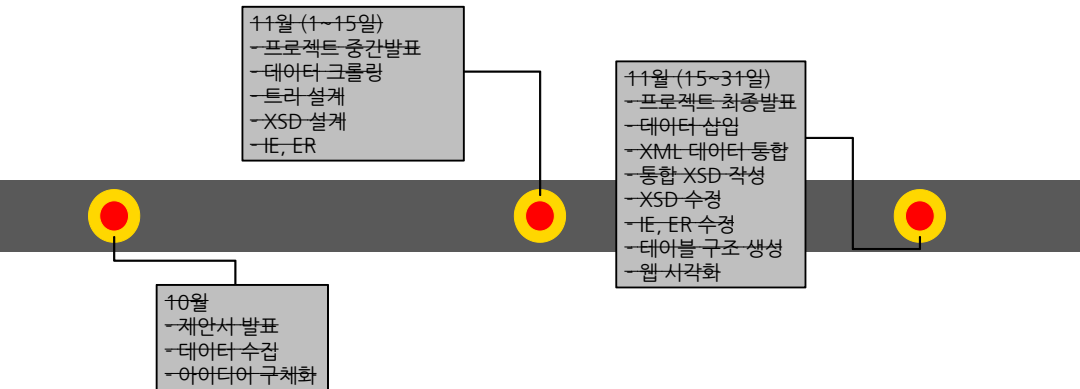
군산시  
착한가격업소

37

군산시청  
선정맛집

※ 등록된 업소 기준으로 해당 수와 다를 수 있음.

# 로드맵



# 간트 차트

내용/주	4 (초안)	5	6	7	8	9	10	11 (중간)	12	13	14 (최종)
요구사항 수집 및 분석											
데이터 전처리											
XML 설계											
구현 및 수정											

# 세부 일정표

담당	내용	시작일	종료일	진행율
최민서	데이터 수집 / 전처리	2022-11-03	2022-11-12	100%
	XSD / XML 파일 관리	2022-11-10	2022-12-01	95%
한동엽	데이터 수집(크롤링) / 전처리	2022-10-01	2022-11-06	100%
	DB 정규화 / 삽입 / 관리	2022-11-15	2022-12-01	90%
공통	XML 트리 설계	2022-10-15	2022-11-07	100%
	공통어휘 찾기	2022-11-01	2022-11-06	100%
	IE, ER 설계 / 데이터 정규화	2022-11-03	2022-11-10	90%
	데이터 통합 과정 / 통합 XSD 작성	2022-11-15	2022-12-01	95%

# 지금까지의 이슈사항

---

- 주소 / 이름 통일성
- 팀 변동사항 및 일정
- 카테고리 등의 부재
- 기타 등등...



# 팀 소감

---

- 2학년 2학기 들어서 갑자기 학교생활이 갑자기 **어려워진 기분**이다. 엄청 하기 싫지만 마음을 먹었고 행동했기에 결과가 나왔다. 뭐든지 "일단 해보면 어떻게든 된다"라는 지극히 평범한 말에 대해 다시 한번 생각해보았다. 결과적으로는 문제를 멀리서 보는 능력, 멀티태스킹 능력이 약간이나마 **발전**한 것 같음을 느꼈다. 내가 만약 코드를 잘 짰으면 30분이라는 시간을 단축할 수 있었을 것이다. 정말 **알고리즘** 하나 짜는 게 중요하구나.
- 시간이 모자랐고 병행한 프로젝트에 의해 체력이 너무 많이 소모됐다. 진행하며 계획대로 될 것 같았던 일들이 앞서 발생한 이슈사항이 **진행을 더디게 했고**, 경험했습니다. 팀원 최민서 분과 같이 협동하였으며, 좋은 파트너였다. 하면서 **전처리 작업이 가장 중요**하다고 꼽았다. 이 경험으로 다음 트랙에 나아갈 수 있을 거 같다. 부족한 프로젝트였지만 성장 방향을 잡아준 프로젝트였다.



# 군산사랑상품권 가맹점 리스트 DB/XML 프로젝트

최민서, 한동엽

## 감사합니다!

'22.11.10.

본 저작물은 군산시청 소상공인지원과에서 '22년 작성하여 공공누리 제4유형으로 개방한 저작물을 이용하였습니다.