



1조 발표

[Chicken Store Data]

I love chicken, you love chicken, We love chicken.

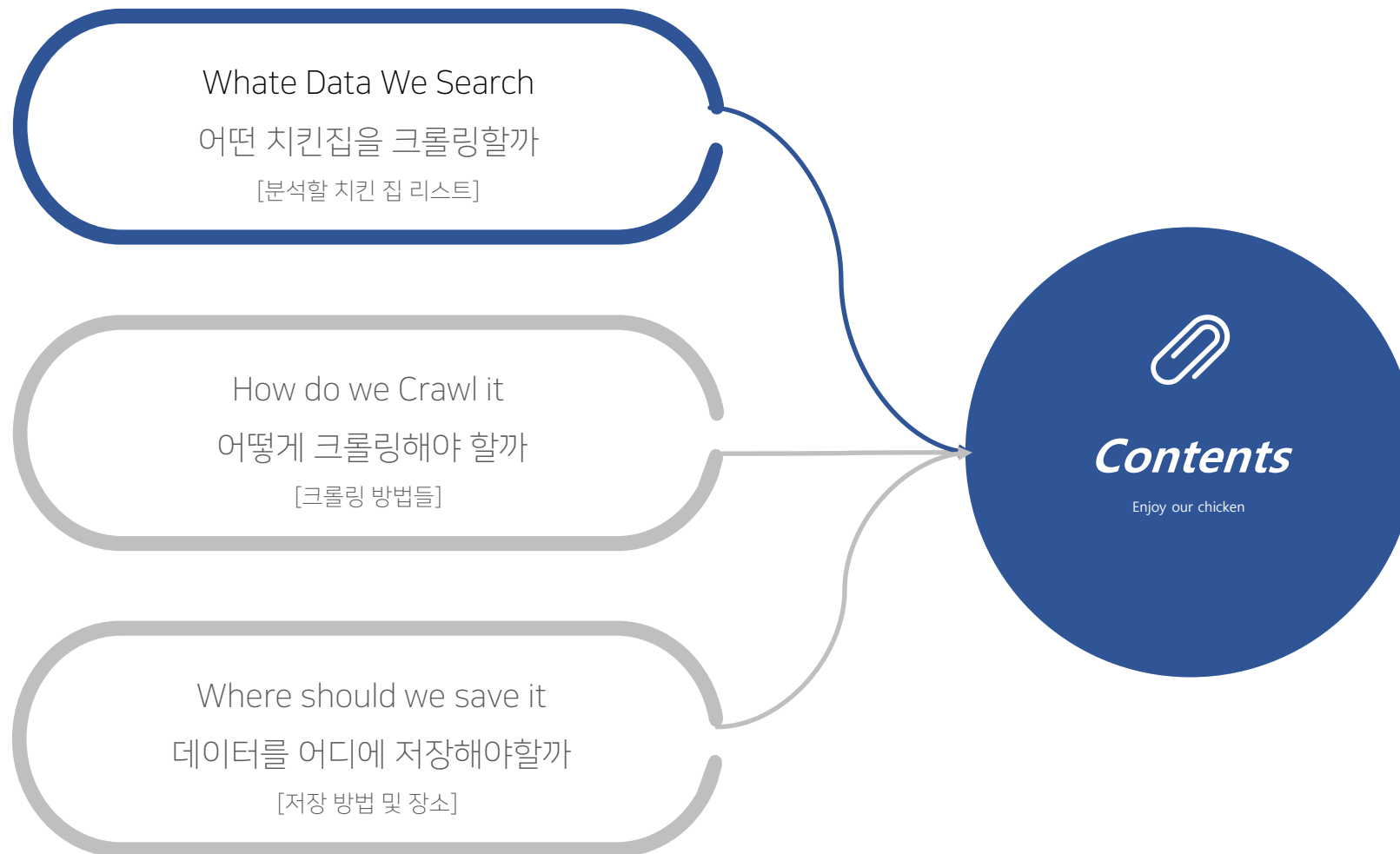
A+



01

CONTENTS

I love chicken, you love chicken, We love chicken

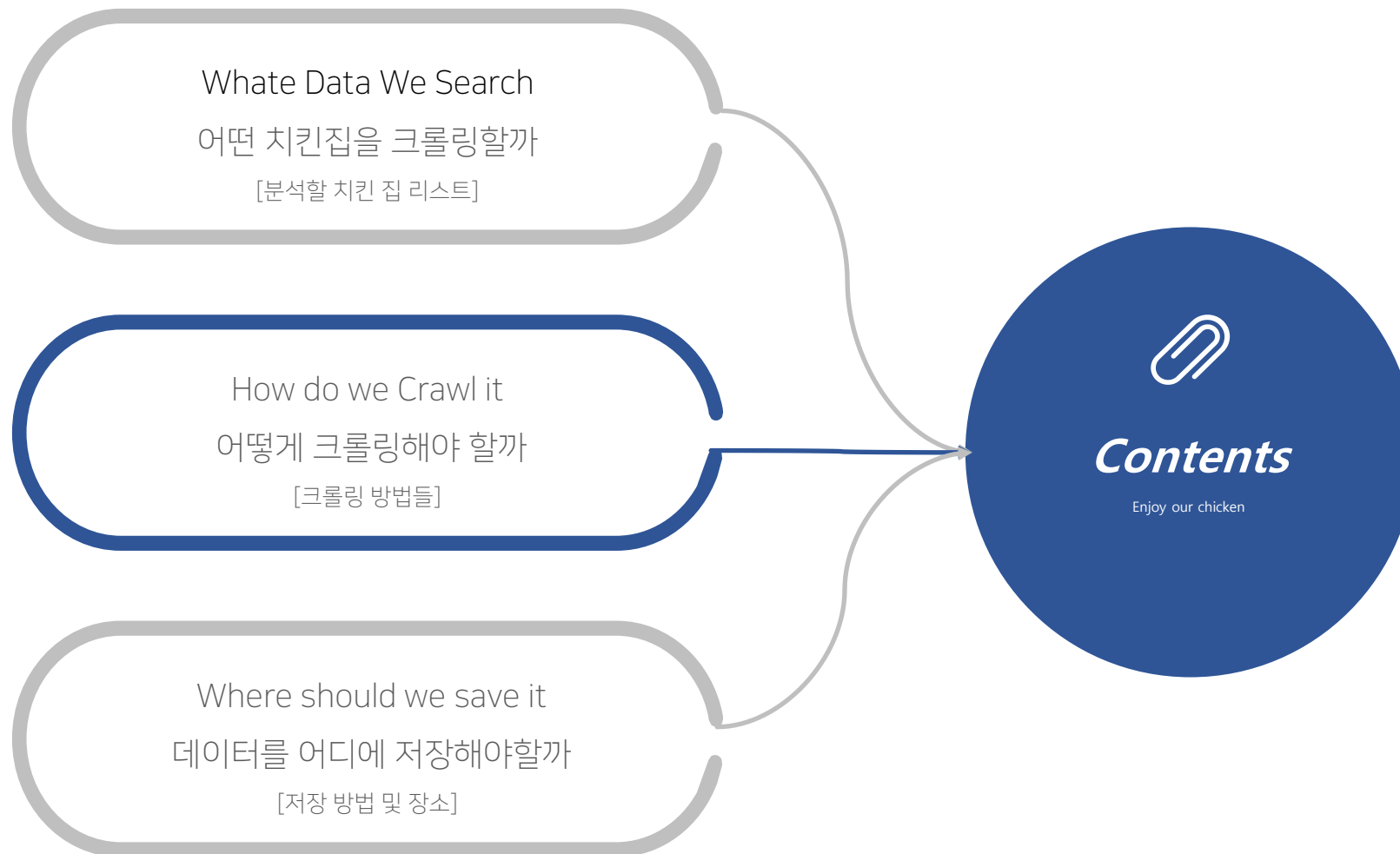




01

CONTENTS

I love chicken, you love chicken, We love chicken

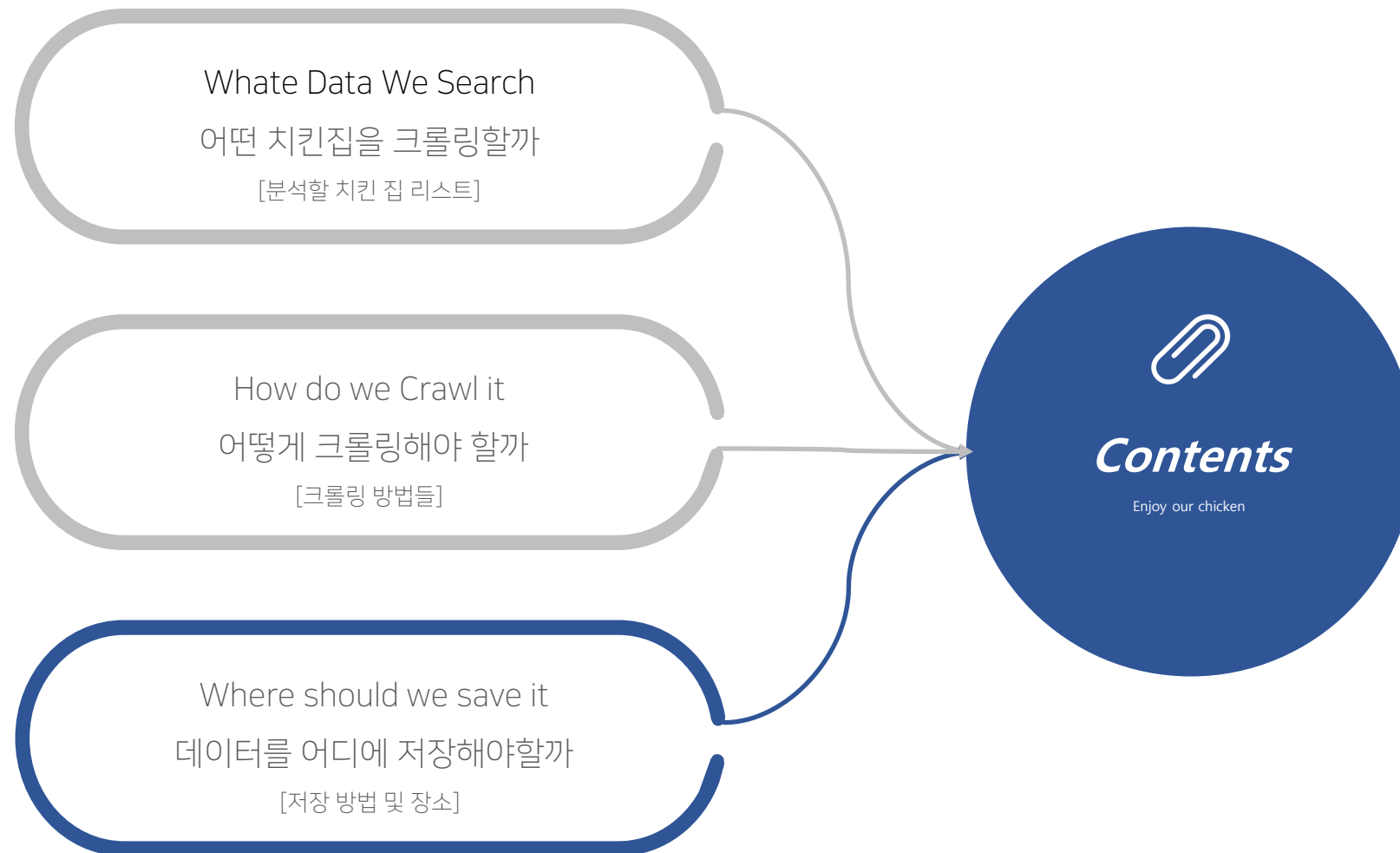




01

CONTENTS

I love chicken, you love chicken, We love chicken





02

What Data We Search

I love chicken, you love chicken, We love chicken

어디를 해야할까..

좋아하는 치킨집이 다 다르다

후보지도 많고 모두가 재미있게 할 수 있는 방법이 없을까...

긱어오기 편한 곳

검색형 매장 정보 조회
사이트는 피하는게..

그래도 시간 너무 쓰지 않도록

크롤링도 중요하지만 역량에 맞는 곳을 찾아야한다

???

많은 데이터 요구

데이터가 많았으면..

컴퓨터만이 할 수 있다

데이터 양이 적으면 굳이 컴퓨터에게 맡길
필요가 없지 않을까.. 최대한 많이 긱자

공부하기 좋은 곳

각자 공부가 될 수
있으면 좋을 것 같은데..

결국은 실력만이 남는다

모두가 가볍게 넘어가지 않고 모두의
역량을 증진시킬 수 있도록



02

What Data We Search

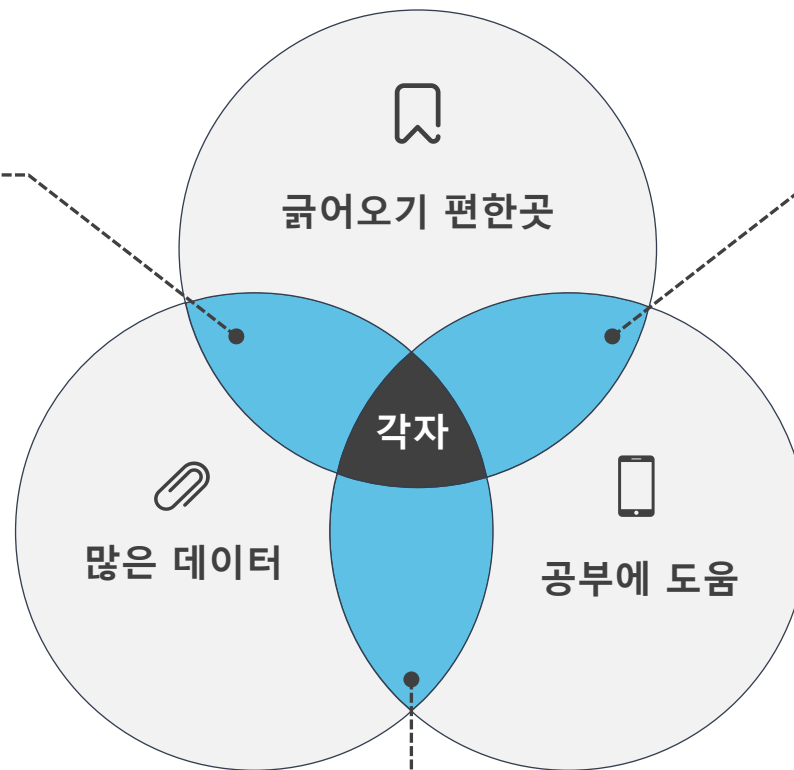
I love chicken, you love chicken, We love chicken

사이트 조사

각자 사이트를 조사해서 적합한 곳을
선별하자

크롤링 한번씩은...

적어도 한번씩은 크롤링 해보고
자기가 만든 데이터를 가공까지 해봐야 과정을
이해하기 쉬울 것이다



많은 데이터, 많은 예외

데이터가 많아질수록 공부하기 좋을 것이다

데이터가 많을 수록 예외처리에 신경 써야한다



02

What Data We Search

I love chicken, you love chicken, We love chicken

사이트 조사

각자 사이트를 조사해서 적합한 곳을
선별하자

긁어오기 편함

크롤링 한번씩은...

각자 한번씩은 크롤링 해보고
저장한 데이터를 가공까지 해봐야 과정을
이해하기 쉬울 것이다

각자

많은 데이터

공부에 도움

많은 데이터, 많은 예외

데이터가 많아질수록 공부하기 좋을 것이다

데이터가 많을 수록 예외처리에 신경 써야한다





Chicken Store Crawling

I love chicken, you love chicken, We love chicken.

페리카나 크롤링 적합성 판단 1

-링크-

https://pelicana.co.kr/store/stroe_search.html?page=1https://pelicana.co.kr/store/stroe_search.html?page=1

get 방식 페이지처리, 아주 좋다.

매장리스트

매장명	주소	전화번호	상세정보
가평역전점	경기도 가평군 가평읍 가화로 54	031-582-5242	상세정보
가평점	경기도 가평군 가평읍 가화로 125-1	031-582-4419	상세정보
가을점	경상북도 영주시 창진로 59	054-634-2500	상세정보
각화주공점	광주광역시 북구 군왕로207번길 16 107호	062-264-8529	상세정보
간성점	강원도 고성군 간성을 간성로 45	033-681-3330	상세정보
간현점	강원도 원주시 지정면 간현로 188	033-731-0887	상세정보
갈곶원동점	경기도 오산시 오산로 68(갈곶동)	031-372-3573 031-372-3574	상세정보
갈마점	대전광역시 서구 갈마로 60(갈마동)	042-531-9962	상세정보
갈산점	인천광역시 부평구 굴포로 42	032-522-9292	상세정보
갈산점(이천)	경기도 이천시 갈산로 81(갈산동, 102)	031-638-8320	상세정보

« < 1 2 3 4 5 6 7 8 9 10 > »

페리카나 크롤링 적합성 판단 2

-태그 구성이 깔끔한가-

```

<tbody>
  <tr>
    <td class="t_center">가평역전점</td> == $0
    <td>경기도 가평군 가평읍 가화로 54</td>
    <td class="t_center"> 031-582-5242</td>
    <td class="t_center">...</td>
  </tr>
  <tr>...</tr>
  <tr>...</tr>
  <tr>...</tr>
  <tr>...</tr>

```

깔끔한 테이블 역시 좋다
페리카나 너로 정했다.



Chicken Store Crawling

I love chicken, you love chicken, We love chicken.

'처갓집양념치킨' 크롤링 적합성 판단 1

-링크-

http://www.cheogajip.co.kr/bbs/board.php?bo_table=store&page=1

http://www.cheogajip.co.kr/bbs/board.php?bo_table=store&page=1

get 방식 페이징처리, 아주 좋다.

지역	매장명	주소	전화번호	자세히
경기도	송정점	경기도 이천시 중신로 273	031-632-4333	more
충청남도	유구점	충청남도 공주시 유구읍 석남리 269-3	041-841-0799	more
전라남도	신기점	전남 여수시 신기남4길 14	061-691-8292	more
경기도	케이준 덕풍점	경기도 하남시 하남대로861번길 5	031-791-7942	more
부산광역시	화명신도시점	부산광역시 북구 금곡대로303번길 69	051-333-5111	more
부산광역시	해운대점	부산광역시 해운대구 해운대해변로209번길26	051-744-9004	more
부산광역시	시청점	부산 연제구 연제로 3	051-851-8292	more
부산광역시	덕포1호점	부산광역시 사상구 사상로 326	051-302-8181	more
경상북도	압량점	경상북도 경산시 청운2로 74	053-802-9985	more
경상북도	축산점	경상북도 영덕군 축산면 축산항길 15	054-734-7377	more

'처갓집양념치킨' 크롤링 적합성 판단 2

-태그 구성이 깔끔한가-

깔끔한 테이블 역시 좋다
처갓집양념치킨 너로 정했다.

```

<div class="tbl_head01 tbl_wrap">
  <table>
    <caption>매장찾기 목록</caption>
    <thead>...</thead>
    <tbody>
      <tr class>
        <td class="td_date">경기도</td>
        <td class="td_date">송정점</td> == $0
        <td class="td_subject">경기도 이천시 중신로 273</td>
        <td class="td_date">031-632-4333</td>
        <td class="td_date" style="width:80px;">...</td>
      </tr>
    </tbody>
  </table>

```



Chicken Store Crawling

I love chicken, you love chicken, We love chicken.

'피자나라치킨공주' 크롤링 적합성 판단 1

-링크-

http://pncg.co.kr/page/sub3_1.html

http://www.pncg.co.kr/page/sub3_1.html?offset=0
http://www.pncg.co.kr/page/sub3_1.html?offset=20

0, 20, 40, 60.. 단위로 증가하는
get 방식 페이징처리

평택고덕점	경기도 평택시 고덕면 고덕국제 3길 64-1, 1층	031-668-8545,6	    
남원시청점	전라북도 남원시 시청로 70, 1층	063-634-8732~3	    
구리수택점	경기도 구리시 원수택로32번길 22	031-554-5052~3	    

1 2 3 4 5 6 7 8 9 10 > >>

'피자나라치킨공주' 크롤링 적합성 판단 2

-태그 구성이 깔끔한가-

```

<table cellpadding="0" cellspacing="0" class="table_style3 mt30">
  <caption>피자나라 치킨공주 매장 검색</caption>
  <colgroup>...</colgroup>
  <thead>...</thead>
  <tbody>
    <tr>
      <td>
        <strong>파주금촌점</strong>
      </td>
      <td class="1">경기도 파주시 새꽃로 62 113호 </td>
    </tr>
    <tr>
      <td>
        <strong>031-949-8683~4</strong>
      </td>
      <td>
      </td>
    </tr>
    <tr>...</tr>
    <tr>...</tr>
    <tr>...</tr>
    <tr>...</tr>
    <tr>...</tr>
    <tr>...</tr>
    <tr>...</tr>
  </tbody>
</table>

```

깔끔한 테이블 역시 좋다
피나치공 너로 정했다.



Chicken Store Crawling

I love chicken, you love chicken, We love chicken.

'60계치킨' 크롤링 적합성 판단 1

-링크-

http://www.60chicken.co.kr/bbs/board.php?bo_table=storehttp://www.60chicken.co.kr/bbs/board.php?bo_table=store&page=2

get 방식 페이징처리, 아주 좋다.

수원서둔탕동점	경기 수원시 권선구 탑동로 5 (탑동) 1층 103호	031-295-6082	VIEW
이천하이닉스점	경기 이천시 부평읍 경춘대로2092번길 23 1동 105호 (이천상성흥푸드센터지)	031-638-6082	VIEW
대구두류남산점	대구 중구 남산로7길 41-1 (남산동) 1층	053-256-0666	VIEW
의왕가꾸단지점	경기 안양시 동안구 호성로 31-5 (호계동) 안하우스 1층	031-427-4936	VIEW
일산마두점	경기 고양시 일산동구 장백로 194 (장항동) 1층 114호(장항동, 위네스21)	031-902-9660	VIEW
양산북정점	경남 양산시 중앙로 312 (북정동) 1층	055-385-3030	VIEW
부산신호점	부산 강서구 신호신당2로49번길 11 (신호동) 1층 104호 (영광빌딩)	051-971-6061	VIEW
서울신림난곡점	서울 관악구 난우길 34 (신림동) 1층	02-862-6082	VIEW
대전도마변동점	대전 서구 도마로 30 (도마동) 1층	042-535-6082	VIEW

1 2 3 4 5 6 7 8 9 10 >

'60계치킨' 크롤링 적합성 판단 2

-태그 구성이 깔끔한가-

```

<!-- 매장명 -->
<a href="http://www.60chicken.co.kr/bbs/board.php?bo_table=store&wr_id=627"> 수원 서둔탕동점
▶<span class="call_btn">...</span>
</td>
<td class="td_location">...</td>
<td class="td_tel">...</td>
<td class="td_view">...</td>
</tr>
<tr class>
<td class="td_subject">

```

깔끔한 테이블 역시 좋다
60계 치킨 너로 정했다.



Chicken Store Crawling

I love chicken, you love chicken, We love chicken.

'바른치킨' 크롤링 적합성 판단 1

-링크-

<https://barunchicken.com/store/><https://barunchicken.com/store/?pageid=2>

get 방식 페이지처리, 아주 좋다.

'바른치킨' 크롤링 적합성 판단 2

-태그 구성이 깔끔한가-

깔끔한 테이블 역시 좋다
바른치킨 너로 정했다.

신길점

주소 : 서울특별시 영등포구 영인로 900
연락처 : 02-2634-5851

신이문점

주소 : 서울특별시 용대문구 이문로 184 바른치킨
연락처 : 02-962-8292

양천향교역점

주소 : 서울 강서구 양천로 344 마곡센트럴대방디점시티 B112호
연락처 : 02-366-5851

여의도점

주소 : 서울특별시 영등포구 63로 45 여의도시범아파트
연락처 : 02-784-5969

오금역점

주소 : 서울특별시 송파구 오금로35길 8 102호
연락처 : 02-443-9292

처음 < 1 2 3 4 5 6 7 8 9 10 > 마지막

```

<!-- 리스트 시작 -->
<ul class="kboard-list">
  <li class="kboard-list-item">...</li>
  <li class="kboard-list-item">
    ::marker
    <a href="/store/?pageid=2&mod=document&uid=9128" title="방학신동아점">...
  </a>
  <div class="kboard-worldmap-franchise-wrap">
    <a href="/store/?pageid=2&mod=document&uid=9128" title="바른치킨">...
    </a>
    <div class="kboard-worldmap-franchise-area">지역 : 서울</div>
    <div class="kboard-worldmap-franchise-branch">지점명 : 방학신동아점
    </div>
    <div class="kboard-worldmap-franchise-address">주소 : 서울특별시 도봉구
    방학로 183</div>
    <div class="kboard-worldmap-franchise-tel"> == $0
      "연락처 : "
      <a href="tel:02-954-6999" title="02-954-6999">02-954-6999</a>
    </div>
    <div class="kboard-worldmap-franchise-homepage">...</div>
  </div>
</li>

```



03

How Do we Crawl it

I love chicken, you love chicken, We love chicken

페리카나

처갓집 양념치킨

60계 치킨

피자나라 치킨공주

바른치킨



03

How Do we Crawl it

I love chicken, you love chicken, We love chicken

나중에 원활하게 데이터를 사용하기 위해서..

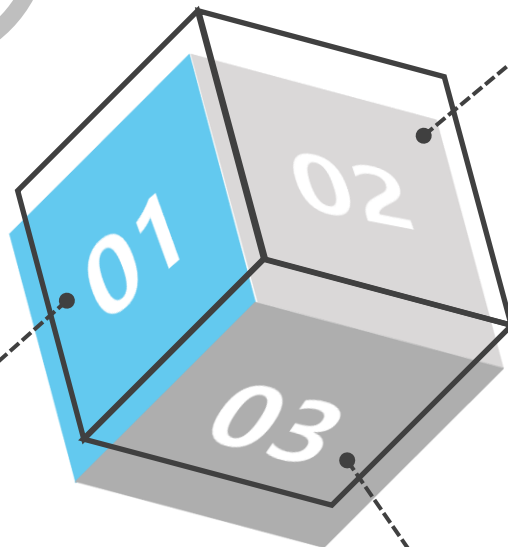
크롤링할 데이터의 속성을 다같이 약속하였다.



01. 가맹점 이름

매장 정보 중 가맹점이름

변수명 : name



02. 가맹점 위치(주소)

가맹점의 위치 정보

변수명 : address



03. 가맹점 전화번호

가맹점의 전화 번호

변수명 : tel



How Do we Crawl it

I love chicken, you love chicken, We love chicken

I love chicken, you love chicken, We love chicken.

```
import multiprocessing
import parmap
import numpy as np
import pericana
import requests
from bs4 import BeautifulSoup
import time
import urllib.request
from selenium.webdriver import Chrome
import re
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.common.keys import Keys
import datetime as dt
import os
```

```
import requests
from bs4 import BeautifulSoup
import time
import urllib.request
from selenium.webdriver import Chrome
import re
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.common.keys import Keys
import datetime as dt
import pandas as pd

def page_make(max_num):
    dial = []
    for i in range(1,max_num+1):
        addr = 'https://pelicana.co.kr/store/stroe_search.html?page='+str(i)
        dial.append(addr)
    return dial

def store_list(lists):
    store_pd2 = []
    path = '../webdriver/chromedriver'
    delay = 2
    driver= Chrome(path)
    driver.implicitly_wait(delay)
    for l in lists:
        link = l
        driver.get(link)
        tr_list = driver.find_elements_by_xpath('//*[@id="contents"]/table/tbody/tr')
        for i in range(1,len(tr_list)+1):
            name = driver.find_element_by_xpath('//*[@id="contents"]/table/tbody/tr['+str(i)+']/td[1]').text
            addr = driver.find_element_by_xpath('//*[@id="contents"]/table/tbody/tr['+str(i)+']/td[2]').text
            if addr == "":
                break
            tel = driver.find_element_by_xpath('//*[@id="contents"]/table/tbody/tr['+str(i)+']/td[3]').text
            tel = tel.split('\n')
            tel = tel[0]
            store_pd2.append([name,addr,tel])
    return store_pd2
```

```
num_cores = multiprocessing.cpu_count() # 12

link_list = pericana.page_make(109)

splited_data = np.array_split(link_list, num_cores)
splited_data = [x.tolist() for x in splited_data]

result = parmap.map(pericana.store_list, splited_data, pm_pbar=True, pm_processes=num_cores)
```

[illegible]



I love chicken, you love chicken, We love chicken.

03

How Do we Crawl it

3. Pandas Dataframe 변환 및 csv 저장

```
import pandas as pd

last = []
for re in result:
    for r in re:
        last.append(r)
list_df = pd.DataFrame(last)
list_df.to_csv("./pericana_data/data.csv", header=True, index=False, encoding='utf-8-sig')
```

[illegible]



Chicken Store Crawling

I love chicken, you love chicken, We love chicken.

```
import time
from selenium import webdriver
from bs4 import BeautifulSoup

#-----1. 웹 페이지 접근
# 웹드라이버 객체 생성
driver = webdriver.Chrome('./webdriver/chromedriver')
driver.implicitly_wait(5)
```

```
result = []
for page_idx in range(1,128):
    # 페이지 접근
    driver.get('http://www.cheogajip.co.kr/bbs/board.php?bo_table=store&page=%d' %page_idx)
    time.sleep(5)
    # 화면의 소스를 가져오기
    html=driver.page_source
    # print(html)

    # BeautifulSoup 이용하여 데이터 출력
    soup=BeautifulSoup(html,'html.parser')
    for store in soup.select('.tbl_head01'):
        name = [name.text.replace(',',' ') for name in store.select('tr > td:nth-child(2)')]
        tel = [tel.text.replace(',',' ') for tel in store.select('tr > td:nth-child(4)')]
        addr = [addr.text.replace(',',' ') for addr in store.select('tr > td:nth-child(3)')]
        for name, tel, addr in zip(name, tel, addr):
            print('{},{,}'.format(name, tel, addr))
            result.append([name, tel, addr])
```

```
import csv
import os
os.makedirs('./data', exist_ok=True)
with open('./data/cheogajip.csv', 'w', encoding='utf-8', newline='') as f:
    writer=csv.writer(f)
    writer.writerows(result)
```

Crawling

1. 웹드라이버 객체 생성
2. 전체 페이지 수만큼 반복문 실행
3. selenium 사용하여 원하는 태그의 텍스트를 각 리스트에 저장
4. 가져온 리스트를 하나의 리스트에 저장(zip() 사용)

CSV 파일 생성

1. 파일을 저장할 디렉토리 생성
 - os.makedirs() 사용 / exist_ok=True 옵션을 사용하여 디렉토리가 이미 존재하더라도 에러 발생 X
2. 파일 생성
 - 파일 생성시 한 라인 생성 후 자동으로 개행이 되어 newline="" 옵션 추가.
 - writerows() 를 사용해 데이터를 입력하면 반복문을 사용하지 않아도 한 레코드씩 입력이 가능



03

How Do we Crawl it

I love chicken, you love chicken, We love chicken

Csv 파일 생성 확인

송정점	031-632-4	경기도 이천시 증신로 273						
유구점	041-841-0	충청남도 공주시 유구읍 석남리 269 -3						
신기점	061-691-8	전남 여수시 신기남4길 14						
케이준 덕점	031-791-7	경기도 하남시 하남대로861번길 5						
화명신도시점	051-333-5	부산광역시 북구 금곡대로303번길 69						
해운대점	051-744-9	부산광역시 해운대구 해운대해변로209번길26						
시청점	051-851-8	부산 연제구 연제로 3						
덕포1호점	051-302-8	부산광역시 사상구 사상로 326						
압량점	053-802-9	경상북도 경산시 청운2로 74						
축산점	054-734-7	경상북도 영덕군 축산면 축산항길 15						
도청점	054-654-2	경상북도 예천군 호명면 양지5길 4						
매호점	053-202-7	대구광역시 수성구 천일로 180						
다산점	054-955-9	경상북도 고령군 다산면 상곡1길 16						
외동점	054-746-2	경북 경주시 외동읍 내외로 229-5						
평리점	053-566-5	대구광역시 서구 통화로 85-2						
임당점	053-802-7	경상북도 경산시 남매공원로2길 35						
장기점	053-292-9	대구시 달서구 장기동 824-3						
범물1점	053-781-3	대구시 수성구 범물동 1369-11						
가산점	054-971-8	경북 칠곡군 가산면 경북대로 1588						
아산원남점	041-544-3	충청남도 아산시 음봉면 음봉면로251번길 48						
유곡점	041-355-8	충남 당진시 송산면 서정9길 12						
함덕점	041-362-7	충남 당진시 함덕읍 미락1안길 31						
삽교점	041-338-1	충청남도 예산군 삽교읍 삽교로 167						
당진CGV점	041-358-3	충남 당진시 복문로1길 41-12						
문화점	042-582-8	대전광역시 중구 송리로 42						
당정점	041-548-9	충남 아산시 당정면 당정면로8번길 93-3						
죽정점	041-933-8	충남 보령시 봉황로 72-11						
사창점	043-903-5	충북 청주시 서원구 창직로 17-1						
성화점	043-234-8	충북 청주시 서원구 신화로 39						
한밭대점	042-822-9	대전 유성구 학하서로121번길 39-11						
태안점	041-675-0	충남 태안군 태안읍 군청7길 31						
청양점	041-944-0	충청남도 청양군 청양읍 칠갑산로 2길 13						
지웰점	070-7622-	충청북도 청주시 흥덕구 대신로73번길 86-1						
지웰백대점	043-236-5	충북 청주시 흥덕구 증안로22번길 13-12						
종촌다정점	044-866-2	세종시 도음1로 106						



Chicken Store Crawling

I love chicken, you love chicken, We love chicken.

```
from selenium import webdriver
from bs4 import BeautifulSoup
import pandas as pd
```

```
driver = webdriver.Chrome('./webdriver/chromedriver')
driver.implicitly_wait(3)
```

```
nameList = []
addrList = []
tellList = []

for page_idx in range(0, 441, 20): # (0, 21, 20) (0, 441, 20) : 0~440 # 1~109
```

```
    # 페이지 접근
    driver.get('http://www.pncq.co.kr/page/sub3_1.html?offset=%d' % page_idx)
```

```
    # 화면의 소스를 가져오기
    html = driver.page_source
```

```
    # BeautifulSoup 을 이용하여 데이터를 출력
    soup = BeautifulSoup(html, 'html.parser')
```

```
    name = soup.select("tbody tr td:nth-child(1)")
    addr = soup.select("tbody tr td:nth-child(2)")
    tel = soup.select("tbody tr td:nth-child(3)")
```

```
    for i in range(len(name)):
        nameList.append(name[i].text)
```

```
    for i in range(1, len(addr)):
        addrList.append(addr[i].text)
```

```
    for i in range(len(tel)):
        tellList.append(tel[i].text)
```

Crawling

1. 웹드라이버 객체 생성
2. 전체 페이지 수만큼 반복문 실행
(20 만큼 점프하며 반복문 실행)
3. selenium 사용하여 원하는 태그의 텍스트를 각 리스트에 저장

```
final = {
    "이름": nameList,
    "주소": addrList,
    "전화번호": tellList
}

print(final)

finalList = pd.DataFrame(final, columns=["이름", "주소", "전화번호"])
print(finalList)

finalList.to_csv("pinachigong.csv", index=False)
```

CSV 파일 생성

1. 가져온 리스트를 하나의 데이터프레임에 저장
2. 판다스 라이브러리의 .to_csv() 이용



03

How Do we Crawl it

I love chicken, you love chicken, We love chicken

Csv 파일 생성 확인

부산서면점,부산광역시 부산진구 부전로 117-7 ,051-805-1885~6
 전주효자점,전라북도 전주시 완산구 용머리로 101 상가1동 106호 1층 피자나라치킨공주,063-237-5378~9
 양주덕계점,경기 양주시 덕계동 696-5 ,031-862-2190~1
 포항오천점,경북 포항시 남구 오천읍 원리 925-4 ,054-293-8579~80
 남양주별내점,경기도 남양주 별내동 960-6 ,031-575-5751
 김포통진점,경기도 김포시 통진을 서암로6번길 5-11,031-996-4519~20
 통영죽림점,경상남도 통영시 광도면 죽림4로 23-96 ,055-643-8676~7
 청송진보점,경북 청송군 진보면 진보로 232 ,054-872-0997~8
 대구평리점,대구광역시 서구 통학로39길 46 ,053-557-9946~7
 칠곡왜관점,경상북도 칠곡군 왜관을 달오3길 26-7 ,054-977-8659~60
 세종다정점,세종특별자치시 다정중앙로 40 1층 피자나라치킨공주,044-868-3489~90
 서울신길점,서울시 영등포구 도신로48길 28-1 ,02-849-8879~80
 전주인후점,전라북도 전주시 덕진구 동가재미1길 24-13 동가재미1길 24-13,063-247-8576~7
 대구국우도남점,대구광역시 북구 국우동 1112-5 ,053-327-2829~30
 서울중화점,"서울 중랑구 봉화산로4길 76, 1층","02-496-8283,4"
 평택고덕점,"경기도 평택시 고덕면 고덕국제3길 64-1, 1층","031-668-8545,6"
 남원시청점,"전라북도 남원시 시청로 70, 1층",063-634-8732~3
 구리수택점,경기도 구리시 원수택로32번길 22,031-554-5052~3
 일산식사점,"경기도 고양시 일산동구 위시티로 55-13, 지하1층 06호",031-969-8577~8
 의정부중앙역점,"경기도 의정부시 태평로 131, 1층 101호(의정부)",031-856-2231~2
 진해경화점,경상남도 창원시 진해구 진해대로 608,055-545-5270~1
 파주운정점,"경기도 파주시 와석순환로252번길 7-21, 101호",031-949-4473~4
 군산조촌점,"전라북도 군산시 양안로 103, 1층상가 102호",063-454-8115~6
 서울내발산점,"서울특별시 강서구 강서로47길 35, 1층 105호",02-2667-0775~6
 안동용상점,경상북도 안동시 용상동 1162-6 ,054-823-9227~8
 통영복신점,경남 통영시 장골로 25 지하 202호,055-643-9533~4
 대구대명점,대구광역시 남구 대명1동 805-14 ,053-656-8568~9
 서울개포점,"서울특별시 강남구 개포로36길 4-4, 1층 103호",02-2058-0818~9
 광양광영점,전라남도 광양시 정수1길 3,061-795-8225~6
 세종소담점,세종특별시 남세종로 357 상가2동 1층 202호,044-868-1650~1
 마포아현점,"서울특별시 마포구 순기정로 42, 1층",02-362-8533~4



Chicken Store Crawling

I love chicken, you love chicken, We love chicken.

```

from selenium import webdriver
import pandas as pd
import csv
from bs4 import BeautifulSoup
driver = webdriver.Chrome('./webdriver/chromedriver')
driver.implicitly_wait(7)
result = []
for page_idx in range(1,69): #68페이지까지
    driver.get("http://www.60chicken.co.kr/bbs/board.php?bo_table=store&page=%d&si=&qu=&dong=" % page_idx)
    html = driver.page_source
    soup = BeautifulSoup(html, 'html.parser')
    for chicken in soup.select("#fboardlist > div > table > tbody > tr"):
        name = [name.text for name in chicken.select("td.td_subject > a")]
        tel = [tel.text for tel in chicken.select("td.td_subject > span > a")]
        address = [address.text for address in chicken.select("td.td_location")]

        for name, tel, address in zip(name, tel, address):
            name = name.replace('\t', ' ').replace('\n', ' ')
            address = address.replace('\t', ' ').replace('\n', ' ').replace(' ', '')
            #print('{0}, {1}, {2}'.format(name, tel, address))
            result.append([name, tel, address])

test = pd.DataFrame(result)
print(test)
test.to_csv('크롤링결과.csv', encoding='utf-8', index=False)

```

Crawling

1. 웹드라이버 객체 생성
2. 전체 페이지 수만큼 반복문 실행
3. selenium 사용하여 원하는 태그의 텍스트를 각 리스트에 저장

CSV 파일 생성

1. 가져온 리스트를 하나의 데이터프레임에 저장
2. 판다스 라이브러리의 .to_csv() 이용



03

How Do we Crawl it
I love chicken, you love chicken, We love chicken

Csv 파일 생성 확인

name	address	tel
수원서둔탑동점	경기수원시권선구탑동로5(탑동)1층103호	031-295-6082
이천하이닉스점	경기이천시부발읍경충대로2092번길231동105호(이천삼성홈프레스티지)	031-638-6082
대구두류남산점	대구중구남산로7길41-1(남산동)1층	053-256-0666
의왕가구단지점	경기안양시동안구호성로31-5(호계동)원하우스1층	031-427-4936
일산마두점	경기고양시일산동구장백로194(장항동)1층114호(장항동,위너스21)	031-902-9660
양산북정점	경남양산시중앙로312(북정동)1층	055-385-3030
부산신호점	부산강서구신호산단2로49번길11(신호동)1층104호(영광빌딩)	051-971-6061
서울신림난곡점	서울관악구난우길34(신림동)1층	02-862-6082
대전도마변동점	대전서구도마로30(도마동)1층	042-535-6082
전주혁신만성이서점	전북완주군이서면오공로11-13105호	063-227-6082
경주안강점	경북경주시안강읍화전길50B동1호	054-761-8000
광명사거리점	경기광명시광명로858(광명동)1층2호	02-2625-6060
진해남문점	경남창원시진해구남영로59(남문동)1층104호	055-552-6077
서울군자점	서울광진구면목로1(군자동)1층	02-462-6082
여수충무점	전남여수시좌수영로39(광무동)1층	061-641-6082
진해충무점	경남창원시진해구벚꽃로78-1(충무동)1층	055-552-6082
군산나운미룡점	전북군산시황룡로10(미룡동)1층103호	063-467-6092
울산명촌점	울산북구명촌15길3-1(명촌동)1층	052-288-8222
여수미평점	전남여수시만성로75-1(미평동)1층	061-653-9282
용인고림점	경기용인시처인구고림로200번길1(고림동)102호,103호	031-322-1800
강원횡성점	강원횡성군횡성읍읍상로22-1	033-344-0060



```
# 페이지 접근
result = []

store_name = []
store_tel = []
store_addr = []

for page_idx in range(1, 15):
    url = driver.get('https://barunchicken.com/store/?pageid=%d&mod=list' % page_idx)

    html = driver.page_source

    soup = BeautifulSoup(html, 'html.parser')

    stores = soup.select('ul.kboard-list > li.kboard-list-item')

    for store in stores:
        name = store.select_one('.kboard-worldmap-franchise-title').text.strip()
        tel = store.select_one('.kboard-worldmap-franchise-tel > a').text.strip()
        addr = store.select_one('.kboard-worldmap-franchise-address').text.strip().replace('주소 : ', '')
        if not name in store_name:
            store_name.append(name)
            store_tel.append(tel)
            store_addr.append(addr)
```

Crawling

1. 웹드라이버 객체 생성
2. 전체 페이지 수만큼 반복문 실행
3. **selenium** 사용하여 원하는 태그의 텍스트를 각 리스트에 저장