

Job Recruitment Posts Research Project

2 조

✓ 개요 및 핵심목표

극심해지는 청년실업이 과열되는 시대에 코로나 19 라는 천재지변이 닥쳐서 20 대 청년실업 80%이상을 돌파하는 작금의 대한민국의 시작되었습니다. 그로 인해, 넘쳐나는 구직자 및 취업준비생들은 취업을 포기하는 경지에 치달았는데요. 저희 2 조는 그 문제점을 완화하고자 많은 취업준비생들이 겪는 불편함이 무엇인가 토의해본 결과, 취업사이트 여러개를 번갈아가면서 봐야하고 각각의 취업공고 플랫폼의 장점만 살릴수 있는 방법을 모색하여 진행하였습니다. 해당 타겟을 데이터베이스 모델링하여 MySQL 과 Python 으로 데이터베이스를 구축하여 좀 더 편리한 구직활동을 제공하고자 이번 프로젝트를 시작했습니다. 본 조사보고서는 잡코리아와 잡플레닛의 장점을 뽑아서 스크래핑 및 문서화 작업을 한 내용을 담고 있습니다. 첫째, 잡코리아를 선택한 이유는 다른 타사이트들에 비해 많은 양의 구인공고를 가지고 있는 플랫폼이었습니다. 둘째, 잡플레닛은 다른 많은 기업리뷰 웹사이트들에 비해 신뢰도가 높은 기업평가 리뷰점수 데이터 및 신입 연봉 데이터를 가지고 있었습니다. 그리하여, 저희는 이 두 개의 플랫폼을 번갈아가면서 보지않고 한 눈 문서화되어 볼 수 있는 데이터베이스를 구축하게 되었습니다.

✓ 핵심코드와 모듈

```
import selenium
from selenium import webdriver
from selenium.webdriver.common.keys import Keys
import time
import csv
import pandas as pd
--사용한 모듈들--
Company_name = [] #기업 리스트 선언
for i in range(1,101):
    jobkorea_0 =
browser.find_element_by_xpath("/html/body/div[5]/div[1]/div/form/div[2]/div/div[2]/div[1]/ol/li[{0}]/div[2]/div[1]/a/b".format(i))
    Company_name.append(jobkorea_0.text) #기업 100 위 목록 추출

-- 기업의 필요한 정보를 추출--
comp_class = pd.read_csv('company_classifications.csv', encoding = 'euc-kr')

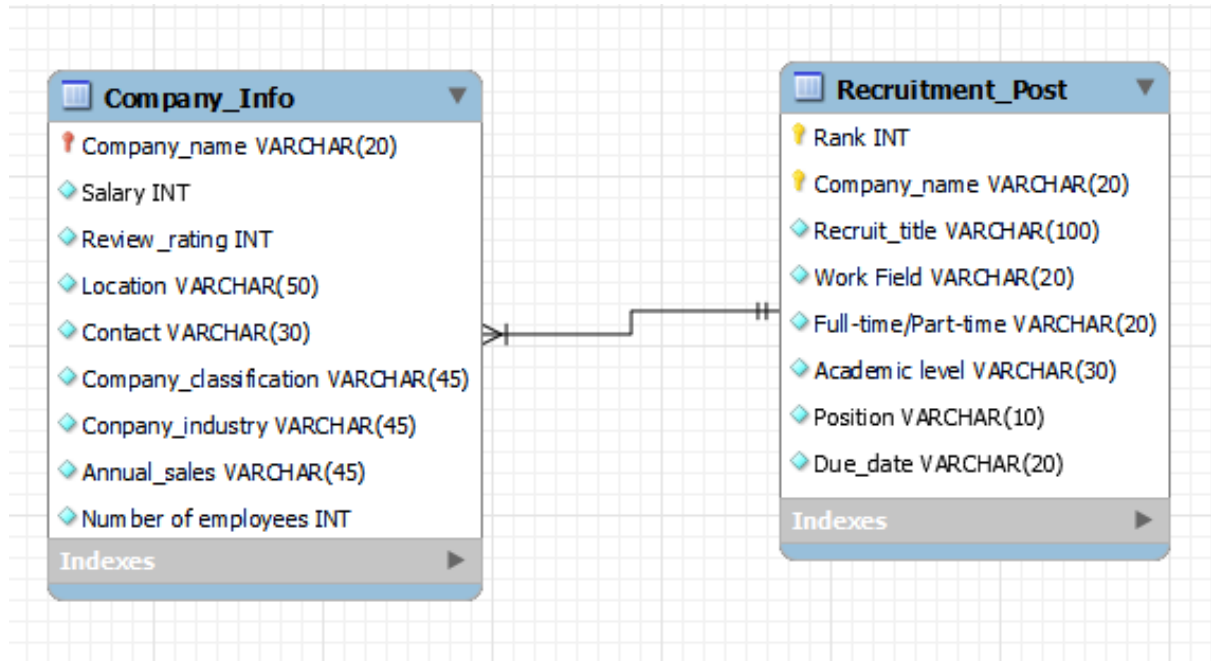
--pandas 를 사용해서 DataFrame 병합 --
jobPost = pd.concat([pd.DataFrame(Rank), pd.DataFrame(Company_name), pd.DataFrame(Recruit_title),
pd.DataFrame(Work_field), pd.DataFrame(Fulltime_Parttime), pd.DataFrame(Academic_level),
pd.DataFrame(Position), pd.DataFrame(Due_date)],axis=1)

jobPost.to_csv('jobPost.csv',index = False, encoding = 'euc-kr')
```

--MY SQL 에서 두개의 테이블 INNER JOIN --

SELECT * FROM JOBPOST A INNER JOIN COMPANY_INFO B ON (A.COMPANY_NAME = B.COMPANY_NAME);

✓ 데이터 베이스 모델링



**데이터베이스 구현:

Rank	Company_name	Recruit_title	Work_Field	Academic_Level	Position	Due_date	Company_name	Salary	Review_rating	Location	Company_classification	Annual_sales	No_of_employees
1	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어-하드웨어 소프트웨어 응용프로그램...	초대졸+	정규직	~07/26(수)	SKD&D	2,852	2.3	서울, 경기, 대전 외	그 외 기타 기계 및 장비 도매업	1180억9900만	350명
2	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대졸+	정규직	~08/31(목)	SKD&D	4,018	3.2	서울, 경기, 인천 외	통신 및 위성 통신업	22,349명	11억9천만원
3	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 시스템소프트웨어 DBA-데이터...	대학원	정규직	~07/31(목)	SKD&D	2,839	3.1	경기	항공 산업업	5480만명	574명
4	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 HTML-개발직...	대학원	정규직	~07/22(목)	SKD&D	3,588	2.1	서울, 충북	기타 종이 및 판지 제조업	5480만명	574명
5	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 DBA-데이터...	대학원	정규직	상시채용	기타	4,835	4	서울, 경기, 인천 외	자동차	5480만명	574명
6	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~08/01(수)	SKD&D	2,964	3	서울, 경기, 대전 외	소프트웨어-소프트웨어 개발업	1,614억7천	282명
7	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~08/01(수)	SKD&D	4,007	2.6	서울, 경기, 대전 외	일반도매 및 소매업	2,507억7천	3,529명
8	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	2,928	3.1	서울, 경기, 부산 외	소프트웨어-소프트웨어 개발 및 공급업	151억2천만원	200명
9	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	3,861	2.8	서울, 경기	컴퓨터시스템 통합 자문 및 구축 서비스업	3041억	974명
10	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	2,602	1.8	경기	산업설비 제조업	2313억	100명
11	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	3,096	3.6	서울, 경기	소프트웨어-소프트웨어 개발업	460억	150명
12	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	3,653	3.2	서울, 경기, 대전 외	항공 산업업	5,611억3천	1,050명
13	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	3,750	2.9	서울, 경기, 대전 외	항공 산업업	530억1천만원	93명
14	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	3,229	2.5	서울, 경기, 대전 외	소프트웨어-소프트웨어 개발업	253억3천만원	38명
15	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	3,518	2.2	서울, 경기	자동차-자동차 부품 제조업	1,650억	430명
16	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~08/08(목)	SKD&D	3,822	3.1	인천	소프트웨어-소프트웨어 개발업	2044억	1,014명
17	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	2,490	1.5	서울, 경기, 인천	소프트웨어-소프트웨어 개발업	191억4천만원	100명
18	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	3,391	3	경기	자동차-자동차 부품 제조업	3,198억9천만원	656명
19	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	2,672	2.2	서울	소프트웨어-소프트웨어 개발업	598억	201명
20	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	3,299	3.8	서울	항공-항공기 제조업	417억8527만	290명
21	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	2,441	2.8	서울, 경기	소프트웨어-소프트웨어 개발업	8억7,281만	18명
22	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	2,445	4.4	서울	교육업	1200억	750명
23	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	3,044	2.6	서울, 경기, 인천 외	차고산업	5650억4260만	1885명
24	SKD&D	2021년 하반기 신입/경력 채용	소프트웨어 응용프로그램 시스템소프트웨어...	대학원	정규직	~07/22(목)	SKD&D	3,044	2.6	서울, 경기, 인천 외	차고산업	5650억4260만	1885명

✓ 느낀점

- 크롤링도 어렵지만 크롤링된 데이터를 sql 에 import 하는 것이 생각보다 쉬운 일이 아니었다. 행별로 데이터를 추출했다면 쉽게 sql 쿼리를 사용하여 테이블을 구축했을텐데, 크롤링 방향을 칼럼별로 추출하다보니 sql 쿼리를 잘 활용하지 못했다는 생각이 든다.
- 잡코리아의 일간 Top100 채용공고와 잡플래닛의 리뷰와 칼럼이 매치가 되지 않는 경우가 있었다.
- 웹 스크래핑으로 정보를 다수의 페이지에서 가지고 오게 될때, 각 기업만의 양식을 가지고 있는 기업정보 같은 경우, 스크래핑을 하는데 있어서 어려움이 크다.
- 데이터베이스의 크기와 양이 커지면 커질수록 수집시 만한 다양성을 가만하고 수집형 알고리즘을 구현해야한다.
- 시간차에 따른 해당 직업순위 스크래핑은 해당 데이터가 실시간으로 변동되므로 가능한 큰 변동폭이 있기전에 스크래핑을 마치는 것이 좋다.
- 한 눈에 보기 힘든 취업공고를 크롤링으로 문서화 했을 때, 한번 코드와 데이터베이스화 해놓으면 그 용도에 있어서 편리하다.
- utf-8 은 인코딩이 잘 안돼서 euc-kr 로 인코딩하는게 낫다

✓ 개선점

- 시간에 따라서 끊임없이 변동되는 순위에 따라서 계속해서 업데이트가 될 수 있는 자동화 코드가 업데이트되면 더 편리할 것으로 보인다.
- 해당 프로젝트에서는 데이터베이스 스택킹을 열 별로 했는데, 행별로 하는 것이 SQL 을 사용하기에 더 적합한 방법으로 보인다..

✓ 마무리

해당 프로젝트를 진행하면서 여러 취업공고 사이트들을 접해보고 장점만을 추려 한 가지 데이터베이스로 구축을 해보았을 때, 일일이 번갈가며 볼 필요가 없어졌고, 한눈의 기업의 규모 및 재무사항까지 고려하여 입사지원을 할 수 있다는 큰 장점이 있었다. 무엇보다도 경제위기 상황으로 한 걸음씩 나아가고 있는 작금의 대한민국은 충분히 기업의 상황도 고려해서 지원을 해야한다고 생각한다. 해당 프로젝트를 서비스해서 플랫폼 위에 플랫폼을 비교하는 플랫폼을 구축할 수 있는 날이 왔으면 좋겠다고 생각하는 바이다.

****해당 데이터베이스 확대본 1 & 2**

	Rank	Company_Name	Recruit_Title	Work_Field	Academic_Level	Position	Due_date
▶	1	㈜카스	2021년 7월 수시채용	소프트웨어·하드웨어 웹프로그래머 응용프로그...	초대졸↑	정규직	~07/26(월)
	2	㈜케이티	2021년 KT 신입 / 경력 채용	웹프로그래머 응용프로그래머 시스템프로그래...	대졸↑	정규직	~08/31(화)
	3	와디즈㈜	와디즈 대규모 인제 영입	응용프로그래머 시스템프로그래머 DBA·데이터...	학력무관	정규직	~07/31(토)
	4	깨끗한나라㈜	2021년 깨끗한나라㈜ 수시채용	컨텐츠·사이트운영 경리·회계·결산 재무·세무·I...	대졸↑	정규직	~07/22(목)
	5	기아㈜	각 부문별 신입/경력 채용	웹프로그래머 응용프로그래머 HTML·퍼블리싱·...	학력무관	정규직	상시채용
	6	㈜메쉬코리아	2021년 부문별 경력/신입 채용	웹프로그래머 응용프로그래머 시스템프로그래...	학력무관	정규직	~08/01(월)
	7	㈜케이씨씨	2021년 경력/신입 수시채용	웹디자인 기획·전략·경영 경리·회계·결산 재무·...	학력무관	정규직, 계약직	내일마감
	8	엔아이티서비스㈜	엔아이티서비스 하반기 채용	응용프로그래머 네트워크·서버·보안 웹프로그...	학력무관	정규직	~07/23(화)
	9	㈜신세계아이앤씨	하반기 경력사원(시스템개발/운영, SAP, AI, Cl...	웹프로그래머 응용프로그래머 시스템프로그래...	학력무관	정규직	~07/27(금)
	10	SK D&D㈜	2021년 하반기 각 부문별 전환할 인턴 및 경력...	응용프로그래머 ERP·시스템분석·설계 시스템...	대졸↑	정규직, 인턴	~10/07(목)
	11	㈜한국생활건강	2021년 하반기 각 부문 신입/경력사원 모집	응용프로그래머 DBA·데이터베이스 웹디자인 E...	고졸↑	정규직	채용 시 마감
	12	㈜우아한형제들	각 부문별 경력/신입 인재영입	응용프로그래머 빅데이터·AI(인공지능) 웹프로...	학력무관	정규직, 계약직	~10/15(금)
	13	유안타증권㈜	2021년 유안타증권 채용	웹프로그래머 응용프로그래머 DBA·데이터베이...	학력무관	계약직	~07/28(수)
	14	에스케이임업㈜	SK임업 신입/경력직 모집	웹마케팅 웹디자인 기획·전략·경영 인사·노무·...	초대졸↑	정규직, 계약직	~07/23(금)
	15	㈜세아네트웍스	각 부문 신입 및 경력 채용	웹기획·PM·ERP·시스템분석·설계 DBA·데이터베...	대졸↑	정규직, 인턴	~07/25(월)
	16	니덱모빌리티코...	니덱모빌리티코리아주식회사 신입/경력 채용	소프트웨어·하드웨어 시스템프로그래머 응용프...	초대졸↑	정규직	~07/25(월)
	17	하나금융티아이	2021년 하나금융TI 신입사원 채용	응용프로그래머 시스템프로그래머 네트워크·서...	대졸↑	정규직	~08/08(월)
	18	㈜두산 디지털이...	채용연계할 인턴 모집	ERP·시스템분석·설계 소프트웨어·하드웨어 시...	학력무관	인턴	~07/25(월)
	19	디만트코리아㈜	2021년 하반기 각 부문별 인재 모집	컨텐츠·사이트운영 동영상제작·편집 사무·출무...	대졸↑	정규직, 계약직	~07/30(금)
	20	㈜케이씨텍	(주)케이씨텍 각 분야 신입/경력 채용	웹프로그래머 응용프로그래머 시스템프로그래...	초대졸↑	정규직	~07/25(월)
	21	㈜탑코	2021 탑코 하반기 각 부문 채용	웹프로그래머 DBA·데이터베이스 웹기획·PM HT...	학력무관	정규직, 계약직	~07/31(토)
	22	㈜대학내일	2021 통합 대학내일 상시채용	빅데이터·AI(인공지능) 기획·전략·경영 마케팅...	학력무관	정규직	상시채용
	23	㈜이루다	2021년 하반기 각 부문별 인재 모집	시스템프로그래머 응용프로그래머 소프트웨어·...	학력무관	정규직	~07/30(금)
	24	㈜에듀윌	에듀윌 부문별(출판/콘텐츠마케팅/경영지원 등...	웹프로그래머 응용프로그래머 웹기획·PM 웹마...	학력무관	정규직, 계약...	~07/31(토)
	38	오스텀임플란트㈜	IT/SW, R&D 부문 채용	웹프로그래머 응용프로그래머 시스템프로그래...	대졸↑	정규직, 계약직	~07/25(월)
	25	오스텀임플란트㈜	연구지원 및 영업/마케팅, 경영지원 등 부문 채용	웹디자인 ERP·시스템분석·설계 재무·세무·IR ...	대졸↑	정규직, 계약직	~07/25(월)

<확대본. 1>

Company_name	Salary	Review_rating	Location	Company_classification	Annual_sales	No_of_emp
㈜카스	2,852	2.3	서울, 경기, 대전 외	그 외 기타 기계 및 장비 도매업	1180억6900만	350명
㈜케이티	4,018	3.2	서울, 경기, 인천 외	무선 및 위성 통신업	17조 8천억원	22,349명
와디즈㈜	2,839	3.1	경기	경영 컨설팅업	11억 9천만원	37명
깨끗한나라㈜	3,588	2.1	서울, 충북	기타 종이 및 판지 제조업	3,480만원	574명
기아㈜	4,835	4	서울, 경기, 인천 외	자동차	54조 1,698억원	35,665명
㈜메쉬코리아	2,964	3	서울, 경기, 대전 외	소방용·오픈마켓·소셜커머스	1,614억 7천	282명
㈜케이씨씨	4,007	2.6	서울, 경기, 광주 외	일반용 도로 및 관련제품 제조업	2조 718억원	3,528명
엔아이티서비스㈜	2,928	3.1	서울, 경기, 부산 외	시스템 소프트웨어 개발 및 공급업	151억 2천만원	200명
㈜신세계아이앤씨	3,861	2.8	서울, 경기	컴퓨터시스템 통합 자문 및 구축 서비스업	3041억	974명
SK D&D㈜	NULL	NULL	경기	산업건설 시공 토목 조경	2313억	100명
㈜한국생활건강	2,602	1.8	서울, 경기	소방용·오픈마켓·소셜커머스	450억	150명
㈜우아한형제들	3,096	3.6	서울	모바일·APP	5,611억 3천	1,050명
유안타증권㈜	3,653	3.2	서울, 충북, 대구 외	증권 중개업	1조 515억천	1,717명
에스케이임업㈜	3,750	2.9	서울, 충북, 전국	조경 건설업	530억 1천만원	93명
㈜세아네트웍스	3,229	2.5	서울, 전북, 경남	네트워크·통신서비스	253억 3천만원	38명
니덱모빌리티코...	3,518	2.2	서울, 경기	자동차·조선·철강·항공	1600억원	430명
하나금융티아이	3,822	3.1	인천	솔루션·SI·CRM·ERP	2044억	1,014명
㈜두산 디지털이...	NULL	NULL	서울, 경기, 인천	솔루션·SI·CRM·ERP	NULL	600명
디만트코리아㈜	2,490	1.5	서울	제약·보건·바이오	191억 4천만원	100명
㈜케이씨텍	3,391	3	경기	반도체 제조용 기계 제조업	3,198억 9천만원	656명
㈜탑코	2,672	2.2	서울	솔루션·SI·CRM·ERP	588억	201명
㈜대학내일	3,299	3.8	서울	광고·홍보·전시	417억 8527만	290명
㈜이루다	2,441	2.8	서울, 경기	호텔업	8억 7,281만원	18명
㈜에듀윌	2,445	4.4	서울	교육업	1200억	750명
오스텀임플란트㈜	3,044	2.6	서울, 경기, 인천 외	치과산업	5650억 4260만	1885명
오스텀임플란트㈜	3,044	2.6	서울, 경기, 인천 외	치과산업	5650억 4260만	1885명

<이어진 테이블> <확대본. 2>