R기반 데이터 분석 프로젝트



취업 준비생을 위한 전략적인 구직 안내서





5조 윤상원

정하늘



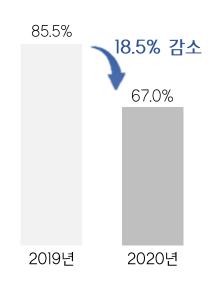
INDEX

- 프로젝트 배경
- 데이터 수집 및 전처리
- 수행 결과(분석 결과 및 한계점, 개선 방향)
- 팀 구성 및 역할



2020, 기업 5곳 중 3곳만 대졸신입 뽑았다

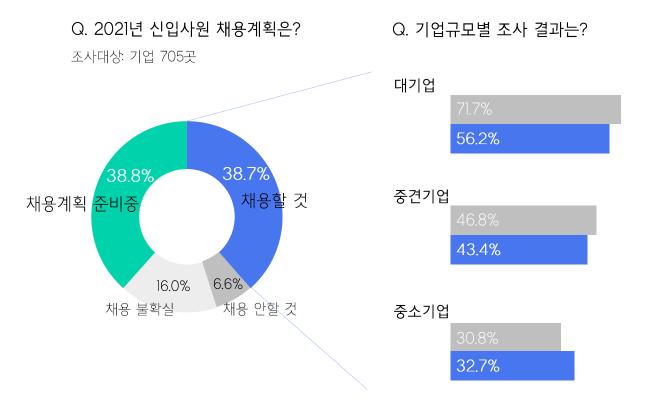
Q. **대졸 신입 '한 명'이라도 채용했습니까?** 조사대상: 기업 536곳



Q. 전년대비 올해 기업규모별 신입채용 비율



2021, 올해 대기업 2곳 중 1곳에서만 신입사원 뽑는다



2021 채용 현황

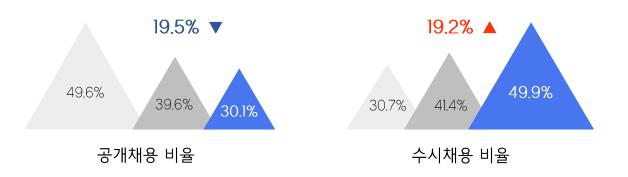


대기업 신입공채 연속 감소세, 수시 모집은 증가

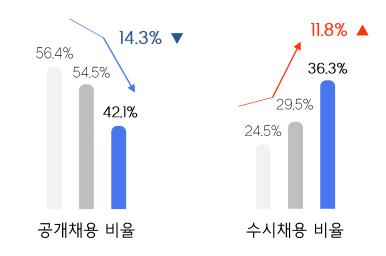
2019년 하반기2020년 하반기

• 2021년 상반기

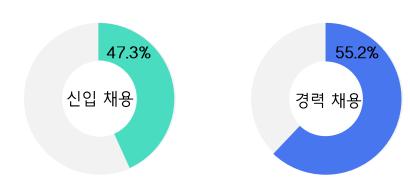
Q. 2021년 대졸 신입사원 채용방식은?



Q. 대기업 채용방식 어떻게 달라졌나?



Q. 대기업, 2021년 상반기 신입보다 경력 많이 뽑는다



2021 채용 현황



<코로나 19로 달라진 채용 트렌드 TOP 5>

화상면접 등 AI면접으로 대체 1위 48.7% 정기공채 사라지고 수시채용 확산 2위 34.5% 온라인 플랫폼 산업 인재 채용 증가 3위 온라인 채용 박람회 확산 4위 언택트 취업상담 서비스 확산 5위

i

전략적인 대응방법 필요

- > 기업들의 달라진 취업 패러다임
- > 미리 준비해야하는 스펙과 자소서





• 데이터 수집 및 전처리

2020년 3월 - 2021년 3월

취업뉴스

총 2,754 건	특수문자를 제외한 키워드를 입력해주세요.	검색 💍
"취준생 절반 이상, 코로나 시대 채용 '자소서 작성' 부담감 커져!" 🔼		TODAY
직장인 10명중 7명, "코로나19로 인해 올해 임금동결"		2021.03.22
"사람인HR, '한국여성과학기술인지원센터 온라인 채용관' 오픈!"		2021.03.18
"현대중공업그룹, C)그룹 등 주요 대기업 신입사원 채용문 활짝!"		2021.03.17
"사람인x시현하다, 개성 담은 '이력서 사진 촬영' 이벤트 진행!"		2021.03.16
"대기업, 올 상반기 신입보다 경력 더 뽑을 것"		2021.03.15
구직자 10명중 9명, "최근 3년간 공채 감소 체감한다"		2021.03.11
"금융계에도 IT 개발자 채용 바람 분다!"		2021.03.10

총 240행

	ntitle	ndate [‡]
1	"취준생 절반 이상, 코로나 시대 채용 '자소서 작성' 부담감 커	TODAY
2	직장인 10명중 7명, "코로나19로 인해 올해 임금동결"	2021.03.22
3	"사람인HR, '한국여성과학기술인지원센터 온라인 채용관' 오	2021.03.18
4	"현대중공업그룹, CJ그룹 등 주요 대기업 신입사원 채용문 활	2021.03.17
5	"사람인x시현하다, 개성 담은 '이력서 사진 촬영' 이벤트 진행!"	2021.03.16
6	"대기업, 올 상반기 신입보다 경력 더 뽑을 것"	2021.03.15
7	구직자 10명중 9명, "최근 3년간 공채 감소 체감한다"	2021.03.11
8	"금융계에도 IT 개발자 채용 바람 분다!"	2021.03.10
9	"IT·공공·디자인 직종, 코로나에도 선방"	2021.03.10
10	"취준생 절반 이상, 코로나 한파에 희망연봉 낮췄다!"	2021.03.08
11	사람인, IT 개발자 채용 플랫폼 '점핏(Jumpit)' 론칭!	2021.03.05
12	수시채용 시대, 1등 스펙봄날 단비 같은 '인턴' 채용 소식	2021.03.05
13	"2020년 주당 평균 취업시간 39시간, 역대 최저치 기록!"	2021.03.05
14	"코로나 시대의 채용, 인재상도 변해!"	2021.03.05
15	"한국철도공사, SK하이닉스 등 주요 기업 신입 채용 시작!"	2021.02.25
16	"기업 5곳 중 3곳, '직무역량' 평가 비중 더 커진다"	2021.02.25
17	"직장인 10명 중 7명, 지난해 이직 미렀다"	2021.02.25
18	코로나19로 "구직자 46% 취업 목표 전환" 업종 변경이 가	2021.02.22
19	"코로나로 빨라진 디지털 트랜스포메이션, 산업 전반 변혁"	2021.02.22
20	"사람인, 웹·앱어워드 4관왕 기념 언택트 축하파티 이벤트 진	2021.02.17

취업시장 뉴스



데이터 수집 및 전처리

2021년 3월 1일 -

기업뉴스

어제의 주요뉴스



'LH서 아파트 싹쓸이 사 퇴' 새만금공사 감사..



위례신도시에 의료복합 타운 조성...미래에셋 컨..



'울며 겨자먹기'?...가덕도 신공항 건설 지원 나..



[게시판] 밀알복지재단, 굿윌스토어 11호점 ..



GS에너지, 한국 기업 최 초 베트남 LNG발전소 ..

총 99,018건

특수문자를 제외한 키워드를 입력해 주세요.

등록일



기사

유명희 "오타와 그룹, WTO 회복 위해 미국과 협력해야"



유명희 "오타와 그룹, WTO 회복 위해 미국과 협력해야"통상장관회의서 WTO 협상분쟁해결 가능 복원 강조[연합뉴스 자료사진] (서울=연합뉴스) 윤보람 기자 = 유명희 산업통상자원부 통상교선본부장은 22일 "세계무역기구(.



LX "구본준 신설지주 상표출원에 법적대응"...LG "디자인 다르다"(종합)

LX "구본준 신설지주 상표출원에 법적대응"...LG "디자인 다르다"(종합)새 사명 "LX" 사용에 대응...LG "상표 디자인 달라 오해 소 지 적어" (세종·서울=연합뉴스) 윤종석 김철선 기자 = LG그룹에...



'충수염 수술' 이재용...부당합병 첫 재판 내달 22일로(종합)

·충수염 수술 이재용…부당합병 첫 재판 내달 22일로(종합)번호인측, 이재용 3주간 안정 필요' 진단서 제출[연합뉴스 자료사진] (서울=연합뉴스) 박형빈 기자 = 이재용 삼성전자 부회장의 계열사 부당합병 첫 공판이 최근 충수염 수술을 _

2021.03.22

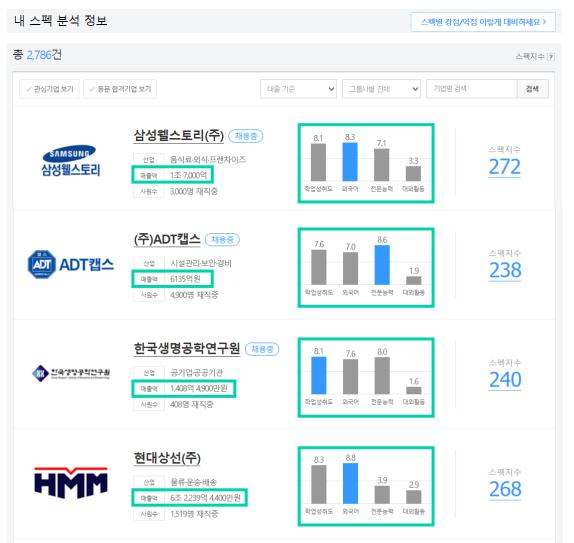
총 3400행

•	title	date [‡]
1	유명희 "오타와 그룹, WTO 회복 위해 미국과 협력해야"	2021.03.22
2	LX "구본준 신설지주 상표출원에 법적대용"LG "디자인 다	2021.03.22
3	'충수염 수술' 이재용부당합병 첫 재판 내달 22일로(종합)	2021.03.22
4	롯데, 바이오 산업 진출 추진코스닥 상장사와 협력 논의	2021.03.22
5	한국타이어家 조양래 회장 작년 보수 40억조현범은 25억	2021.03.22
6	두산그룹 박정원 회장 지난해 연봉 11억원	2021.03.22
7	삼성생명공익재단 이사장에 서정돈 전 성균관대 총장	2021.03.22
8	하청업체 사망사고 난 포스코케미칼 임원·간부 중징계	2021.03.22
9	삼성물산 각 부문 CEO, 자사주 3천200주 매입"책임 경영	2021.03.22
10	'충수염 수술' 이재용계열사 부당합병 첫 재판 연기	2021.03.22
11	LX "구본준 신설 지주사 상표출원에 가처분 등 법적 대응"	2021.03.22
12	금호타이어 사장, 통상임금 패소에 사의신임 대표에 정일택	2021.03.22
13	최태원, SK하이닉스 연봉 30억원 수령작년 보수 총액 63억원	2021.03.22
14	인천도시공사 외국인 전용 임대주택 매각 위법 논란	2021.03.22
15	GL·코파일럿, 가상현실 면접 시스템 개발 협력 MOU	2021.03.22
16	'범현대家 장손' 정의선, 아산 정주영 20주기 추모 사진전 관람	2021.03.22
17	거제시·대우조선 노사, 코로나19 확산 방지 협력	2021.03.22
18	[게시판] SC제일은행 '금융유튜버 박곰희 강의' 웹세미나 개최	2021.03.22
19	[인터뷰] 장인화 부산상의 회장 "일하는 상의 만들겠다"	2021.03.22
20	삼성증권 "유튜브 누적 조회 7천만뷰 돌파업계 최초"	2021.03.22

기업현황 뉴스



● 데이터 수집 및 전처리



총 2494행

•	이름	매출액 💠	스펙지수 👇	학업성취도	외국어	전문능력 🕏	대외활동 🗦
1	㈜다우기술	2.213085e+11	256	7.8	7.6	6.8	3.6
2	대선주조	8.575418e+10	239	7.8	7.4	5.7	1.4
3	㈜한빛소프트	3.298826e+10	233	7.5	8.1	6.0	2.3
4	DB금융투자(주)	8.718253e+11	254	7.8	8.3	7.2	2,4
5	(주)풀무원	7.903302e+10	247	8.0	8.5	5.7	2.4
6	현대자동차(주)	5.066100e+13	302	8.3	8.5	5.2	4.3
7	데상트코리아(주)	6.784000e+11	312	8.2	8.6	5.4	3.9
8	㈜마니커에프앤지	8.570000e+10	236	7.8	8.1	6.7	1.7
9	네이버(주)	4.022600e+12	250	8.0	8.0	5.3	3.6
10	(주)그린웹서비스	3.149666e+10	231	7.8	7.6	6.0	2.4
11	덕일산업	8.700000e+10	259	7.0	7.3	6.4	3.6
12	GS리테일	8.266600e+12	333	8.0	8.4	6.6	4.7
13	두산인프라코어(주)	3.102184e+12	259	8.0	8.1	4.1	2.5
14	세메스(주)	1.125734e+12	266	7.9	8.4	4.8	4.0
15	(주)하이프라자	2.828000e+12	219	7.8	7.9	5.6	2.3
16	SK하이닉스(주)	3.052499e+13	285	8.2	8.4	5.0	4.7
17	삼정케이피엠지㈜	4.170030e+10	209	7.7	8.5	6.0	1.0
18	팅크웨어(주)	1.800000e+03	234	7.7	7.3	6.4	2.2
19	국립암센터	2.898760e+11	238	7.9	7.8	6.8	2.6
20	㈜금화PSC	2.057901e+11	231	7.7	7.0	7.8	1.3



● 데이터 수집 및 전처리



전문가 자소서 총평 5 ★★★★★

본 자기소개서의 장점은 모든 질문에 대해 정확하게 답을 했다는 것입니다. 이 말은 회사와 의사소통이 잘 되는 지원자라는 말입니다. 자신의 생각에만 집중하여 질문에 맞지 않는 내용을 쓰는 지원자들이 많은 상황에서 이런 강점은 분명 회사에서 높이 평가할 것입니다. 그리고 이야기 소재도 개발자와 직접적으로 연관이 있는 경험들이라 면접에서 건설적인 질문이 오갈 수 있는 가능성이 있다는 점이 좋습니다. 잘 썼습니다.

Q1. 성장과정:가족사항, 학창시절, 교우관계, 생활습관, 자신에게 크게 영향을 미친 사건 등을 포함하여 구체적으로 작성

저희 가족은 회사원이신 아버지와 어머니 그리고 누나와 저 4인 가족입니다. 💶 🛣 💵

학창 시절 독서에 관심을 갖고 폭넓은 분야의 책을 지속적으로 읽었으며, 특히 위인전과 고전 소설, 동서양 철학 등 인문 고전을 탐독했습니다. 덕분에 다양한 인간상에 대한 이해와 공감 능력, 그리고 뛰어난 이성과 논리를 습득하는 한편, 인문학적 소양을 쌓으며 위인을 본반으려는 마음가짐을 다지게 되었습니다.

이런 특성으로 인해 독서를 좋아하며 바람직한 도덕 관념과 경손함을 갖춘 친구들과 주로 교우관계를 맺었으며, 성실하고 청렴하신 부모님을 본받아 아침 일찍 등교하고 야간 자율 학습 후 하교하는 성실한 생활 습관을 바탕으로 공부에 임됐습니다. 이지성 작가의 '꿈꾸는 다락방'을 읽으며 범인도 큰 꿈을 가짐으로써 위인으로 거듭나고 결국 그 꿈을 이를 수 있다는 것을 깨달았고, 이것이 저에게 크게 영향을 미쳐 더 나은 사회를 위한 삶을 살고자 하는 평생의 꿈을 품게 되었습니다.

좋은점 2

총 590행

•	star ‡	Q1 ÷	Q2 \$	Q3 ‡	Q4
1	2	"윤활유 같은 인재" 저는 평범한 사람입니다. 초중고 학창시	"다방면의 지식을 갖추는 중입니다" 다양한 현장의 안전을	"혁신을 통한 성취감" 학생회 때 축제부스를 운영했습니다	"ONE TEAM" 00000 공.
2	2	"소통을 통한 점주와의 상생" SC 직무의 역할은 점주와의 동	"분위기를 파악하는 역량" 목표를 달성하기 위해서는 '분위	"책임감을 가진 혁신 전략" 3학년 학생회 때 대학교 축제에서	"CU 편의점 이용 시 장단?
3	2	성장과정에 있어 책은 저에게 삶을 바라보는 시각과 인성을	저는 AI 연구 기관 시큐레이어에서 Meta Feature Extraction	삼성 SDS에 지원한 이유는 MIS 기반 IT 솔루션 개발자로서	NA
4	2	"캠코인의 혁신 노력을 엿보다" 한국자산관리공사의 '신뢰와	"높은 접근성의 투자 지원을 도모하다" 한국자산관리공사의	"협동하는 벌처럼, 다 함께" 동아리 활동으로 2년간 학교 홍	-(상황) 기존 방식의 문제?
5	2	전공선택으로 수강했던 정밀조립 과목으로 생각됩니다. 수	'메카트로닉스' 수업의 팀 프로젝트에서 어려움을 극복하고	단제생활에서 저는 앞장설 수 있는 사람입니다. 모두가 망설	우리 생활에 편리함을 가기
6	2	공식 블로그, 홈페이지 자료 및 발간자료 등을 통해 공사에	'유지관리' 부문에서 효율적인 업무 수행을 위해서는 '안전관	대인관계에 있어서 발생하는 다양한 갈등상황은 자연스러운	농어촌공사가 관리하는 전
7	2	의사소통에서 가장 중요한 것은 '상대방에게 나를 어떻게 이	팀으로 일할 때 가장 중요한 것은 '주인의식을 가지고 책임감	고객 서비스 만족을 위해 가장 중요한 것은 고객의 요구사항	토목 직무에서 안전관리 /
8	2	캡스톤디자인 팀 프로젝트에서 수리 구조물에 대한 프로젝	1. 직무 개요 1) 제가 지원하고자 하는 안전진단본부는 교량,	"정확성과 신뢰성에 주목하다" 공식 SNS인 페이스북, 홈페이	"이론적 지식+문제 해결의
9	2	LH 공사에서 가장 관심 있는 분야는 '도시재생사업'입니다	'단지설계' 부문에서 효율적인 업무 수행을 위해서는 '안전관	의사소통에서 가장 중요한 것은 '상대방에게 나를 어떻게 이	평소 '갈등은 언제나 있으
10	2	일에서 오는 성취감이 저를 달리게 하는 원동력입니다. 학부	유학 시절 남들보다 두세 배 노력해 꼴등에서 1등을 차지했	유학 시절 떡볶이를 팔 때 분석을 통해 초반 부진을 극복하고	의류 전공의 기획 및 디자
11	2	"2% 부족한 점을 100% 만족으로 변화시키기 위해 지원했습	"의사소통 능력, 전공지식, 일 처리의 섬세함" 의사소통을 잘	"끊임없는 성장을 위해 노력하는 CJ제일제당의 엔지니어가	NA
12	2	"기초에서 심화로 가는 여정" 공정 개발직에 지원한 이유는	"미운오리새끼에서 백조로" 고등학교 입학을 시작으로 공부	"바이오 산업이 빠진 아쉬움" 정부가 이번 연도 7월에 발표한	"셔틀콕을 통해 배운 소중
13	2	"미운 오리 새끼에서 백조로" 2학년 때에 아는 누나가 편입에	"과정이 복잡하더라도 원하는 목표를 위해 도전하였고 결국	"처음 시도해보는 논문 내용 발표" 2학년 때 분석화학 강의를	대학 2학년 때 향연이라는
14	2	"소재를 통한 제품의 혁신" 제가 회사를 선택하는 기준은 '내	"개념을 정리하는 습관" 초등학생 때, 어머니께서 요리하시	코로나 사태로 세계 경제 시장은 더욱 약화되고 있지만, 국내	NA
15	2	"긍정적인 결과를 이끌어 내는 의사소통능력" 음식점 아르바	방송부 동아리 활동 중 아침방송 프로그램을 기획한 적이 있	"미흡함을 개선한 후 얻은 팀의 성과" OO공사에서 주최한 U	"적극적인 제안으로 바뀐
16	2	제가 우리은행에 지원한 동기는 제가 평소에 관심있게 공부	저는 연구실의 데모 팀의 리더였습니다. 저와 함께 일한 팀원	NA	NA
17	0	"변화하고 혁신하라" 제 직업관은 '변화를 이끌고 개선하라'	"해상풍력발전의 활성화를 꿈꾸다" 현재 해상풍력발전은 신	"국내 수소산업 선도를 위한 발걸음" 지난 7월 28일, 한국동	"개선점을 찾기 위한 노력
18	2	회사를 선택하는 데 있어 저의 척도는 도전과 혁신입니다. 성	작년 생애 첫 풀코스 마라톤에 도전해서 3시간 59분 4초(SUB	"탄탄한 전문지식" 재무회계팀은 기업의 경영실적과 재무 상	NA
19	2	취미: 축구 특기: 수영 글자수 14자 22Byte	강경원 대위 글자수 6자 11Byte	뛰어난 리더십으로 소대를 이끄는 모습이 인상적이었습니다	"생명보험업의 가치" 생명
20	2	"기업금융 1등 우리은향" 제 목표는 고객에게 통합적인 솔루	"팀워크의 필수전략, 소통" 팀워클를 위해서는 균등하게 바	금융 플랫폼이 발달하고 비대면 거래가 증가하는 상황 속에	NA
21	2	온라인 마케팅 채널 운영 글자수 15자 25Byte	2019.01~2019.06 글자수 15자 15Byte	온라인 채널 운영 담당: 회사 블로그와 SNS 계정, 홈페이지	구글 검색 관리 담당: 구글
22	2	"마케팅에 대한 입체적 이해" 마케팅은 명확한 업무 성과를	"팀워크의 필수전략, 균등한 가치 부여" 팀워크를 높이기 위	"트렌드에 주목" 마케팅 직무에 있어 가장 중요한 것은 트렌	NA

합격자 자소서







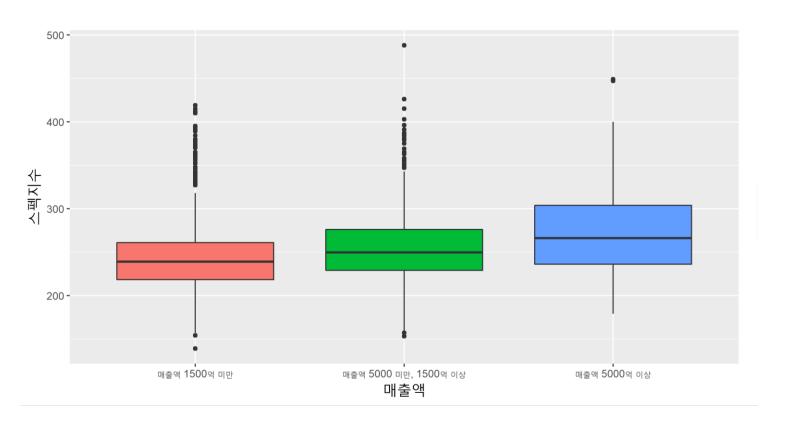
기업 현황 뉴스



최근 뉴스 트렌드



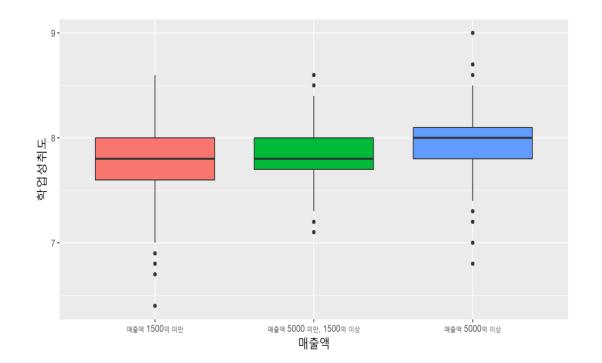
매출액 별 합격자 스펙지수 분포



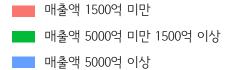
- 매출액 1500억 미만
- 매출액 5000억 미만 1500억 이상
- 매출액 5000억 이상
- > 스펙지수에 포함되는 사항들
- 학업 성취도
- 외국어 능력
- 전문 능력
- 대외 활동

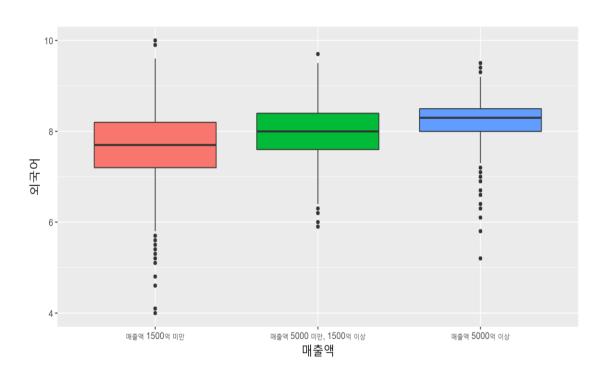


학업 성취도



외국어 능력





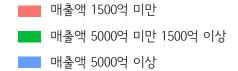


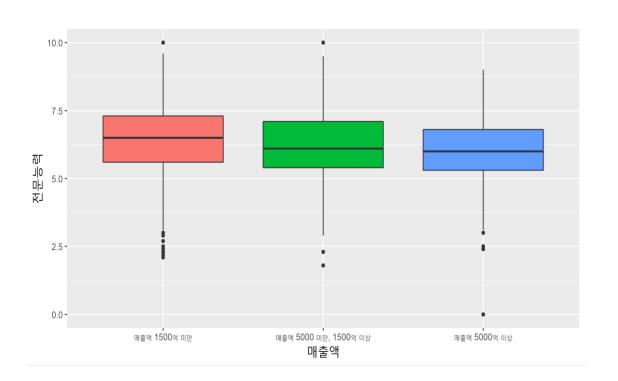


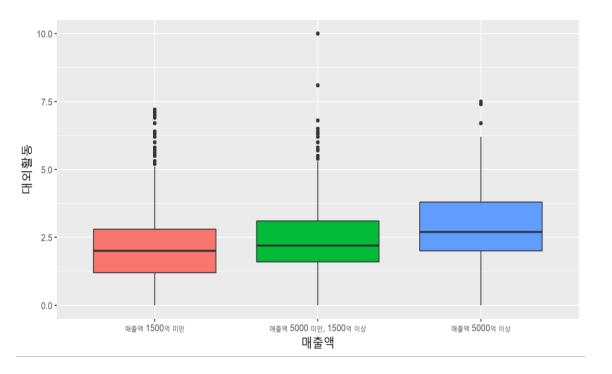
• 분석 결과

전문 능력

대외 활동







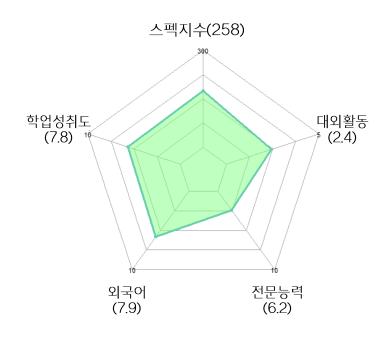




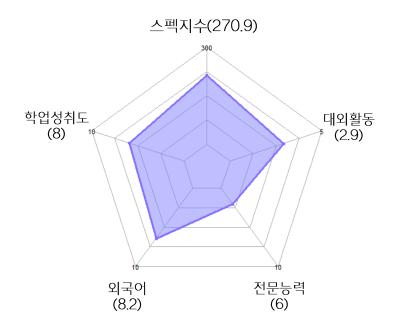
1500억 미만

스펙지수(244.4) 학업성취도 (7.8) 의국어 (7.4) 전문능력 (6.4)

5000억 미만 1500억 이상



5000억 미만

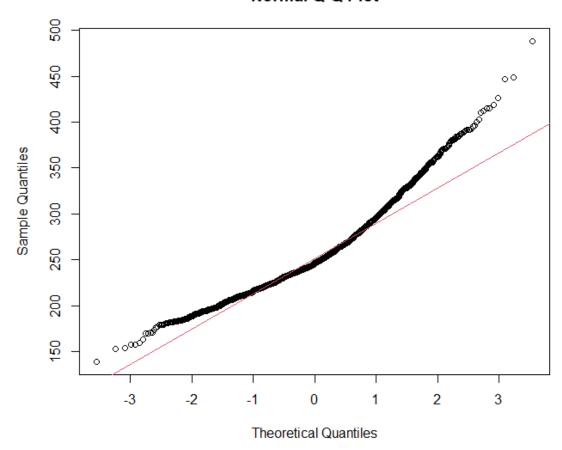


매출액 기준





Normal Q-Q Plot





★★★★ 2점 이하 자소서



★★★★★ 3점 이상 자소서



합격자 자소서



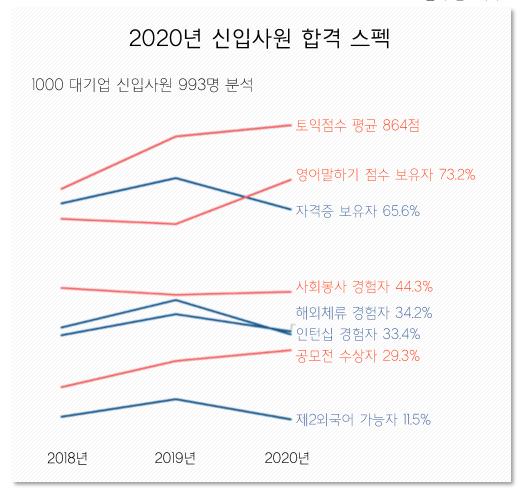




- 합격자 스펙 분석: 매출액 기준
- > But! 매출액 ≠ 좋은 기업
- 분산 가정 성립 X
- > 사후검정 X
- 자소서 워드클라우드
- > 단어로만 되어 있어 자소서 작성에는 도움 X

- \bigcirc
- 기업 분석 시 활용 가능한 기업별 분석
- 합격자 스펙 항목 별 상세 점수 확인
- 자소서 문장 선호도 조사

<출처: 잡코리아>



한계점/개선방향



크롤링이 생각보다 힘들구나 라는 생각이 들었습니다. 특히 동적 크롤링 과정에서는 생각치도 못한 오류가 발생해서 계속 모니터링하고 몇 페이지씩 끊어서 이어서 진행하였습니다. 덕분에 크롤링은 오전에 끝낼 것이라 예상하고 진행 하였는데, 하루나 걸려서 좀 당황스러웠습니다. 그래도 이렇게 가볍게(?) 실전 연습한 것이 도움이 많이 되었습니다.

더 다양하고 자세한 결과를 내기 못해 아쉽지만, 웹 크롤링과 전처리 그리고 분석 방법을 활용해 무엇인가를 구현할 줄 알게 되어 기쁩니다. 짧은 기간동안 프로젝트하느라 수고하셨습니다.



정하늘