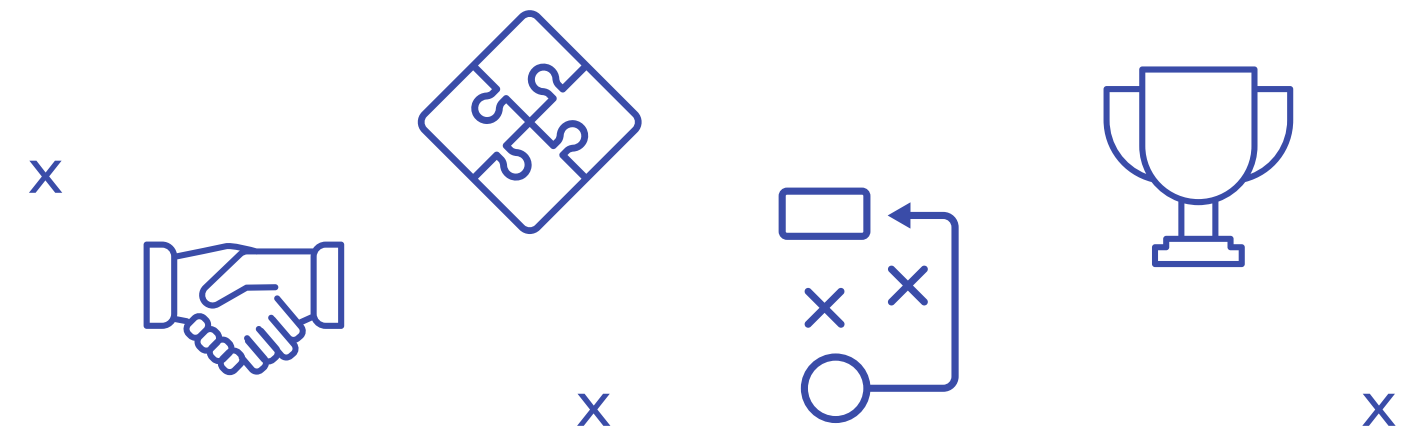


2024-1

합격 자기소개서를 통해 취업 준비하기



Contents



01

주제 선정배경

02

데이터 수집 및 전처리

03

분석 과정

04

분석 결과

05

개선 방안 및 한계점

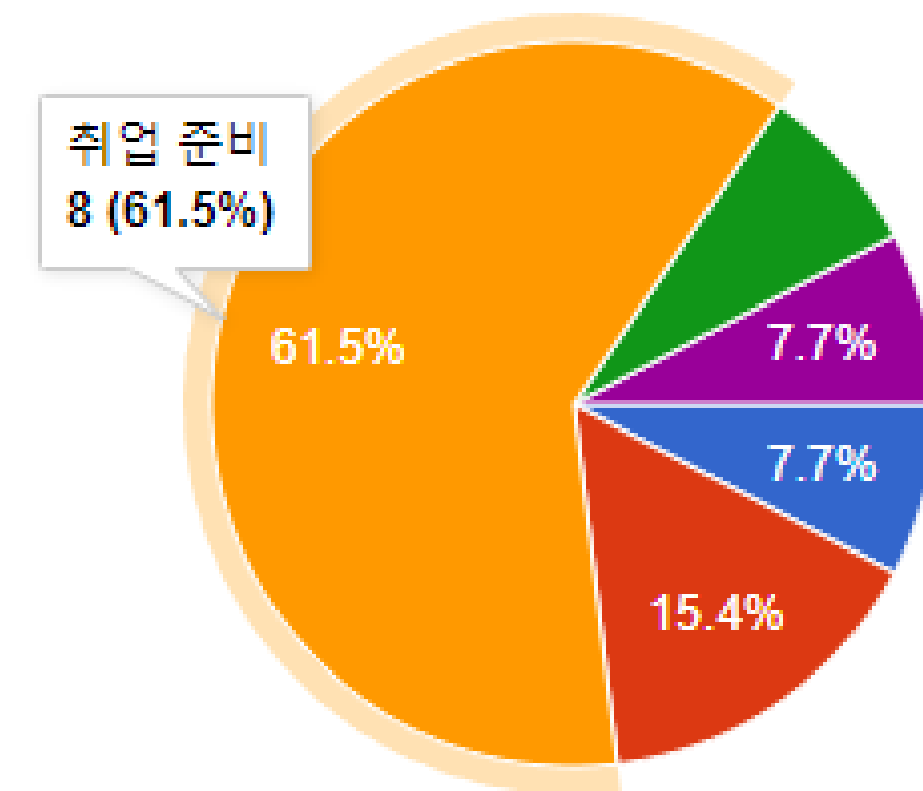
06

활용방안

주제 선정배경

[선정배경]

- 빅데이터응용학과 졸업전 취업률 ↓
- 학과 내 구두조사결과 취업분야, 진로 방향 모름 대다수
- 취업 시장에 공부하지 않으면 취업 어려움



01

주제 선정배경

[주제 선정]

링커리어
LINKareer

LINKareer

채널 커뮤니티 실시간 공고 공고 달력 자소서 만능 검색기 자소서 작성 멘토 게시판

자소서 만능 검색기

무엇이든 검색해보세요. 딱 맞는 합격 자소서가 3초 만에 나옵니다.
스펙업 X 링커리어가 검증한 18,904개 2024 최신 합격자소서 대방출

전체 | 대기업 | 공기업 | 은행권 | 인턴

기업명 직무 자소서 항목, 내용 등 키워드 검색

추천검색어: 롯데, 주택도시보증공사, 카카오, 지원동기, 성장과정, 성격의 장단점, 포부

합격연도	기업명	직무명	분류	합격스펙내용	스크랩
2024 상반기	삼성애스원	경영지원	대기업, 신입	인서울4년제 / 영어 / 학점 4.23 / 토익: 930, 토익스피킹: 레벨6 / 공기업 인턴 ... 1. 지원한 이유와 입사 후 회사에서 이루고 싶은 꿈을 기술하십시오. 학부시절 지역아동센터에서 멘토링 활동을 하며 공간의 중요성을 실감하였습니다. 문화체험 활동으로 멘티와 에버랜드에 가면서 아이의 새로운 면을 볼 수 있었습니다. 내...	77



- 링커리어 합격 자소서를 통해 텍스트마이닝

데이터수집 및 전처리

[데이터 수집]



- 13,018개 데이터 수집
- 기업, 직무, 지원연도를 info열에 저장
- 출신학교, 학점, 스펙 등을 specification열에 저장
- 자기소개서를 self_intro에 저장

검색결과 13028건

합격 자소서

삼성에스원 / 경영지원 / 2024 상반기

인서울4년제 / 영어 / 학점 4.23 / 토익: 930, 토익스피킹: 레벨6 / 공기업 인턴 6개월 / 사조사2급

자소서 작성 스크랩하기 공유하기 새창 목록

마음에 드는 문장을 스크랩 할 수 있어요!
최고 품질의 상품들을 지금보다 더 많은 소비자들이 여러 유통 채널에서 더욱 폭 넓고 쉽게...
[다시보기 않기](#)

1. 지원한 이유와 입사 후 회사에서 이루고 싶은 꿈을 기술하십시오.

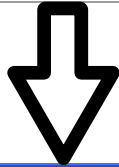
학부시절 지역아동센터에서 멘토링 활동을 하며 공간의 중요성을 실감하였습니다. 문화체험 활동으로 멘티와 에버랜드에 가면서 아이의 새로운 면을 볼 수 있었습니다. 내성적이었던 멘티는 평소 에 감정 표현을 잘 하지 않았기 때문에 멘토링에 어려움이 있었습니다. 하지만 에버랜드 카니발 퍼레이드를 함께 보면서 멘티가 활짝 웃는 모습을 처음 볼 수 있었습니다. 이를 계기로 멘티가 느꼈던 감정과 속 이야기를 하면서 급속도로 가까워질 수 있었습니다.

02

데이터수집 및 전처리

[데이터 전처리]

index	info	specification	self_intro
0	한국농수산물유통공사 / 회계직 / 2020 상반기	부산진여자상업고등학교 / 금융경영 / 학점 X / 한국사검정 시험: 고급, 컴퓨터활용...	1. 공사의 업무 중 본인이 가장 관심 있는 업무 한 가지를 제시하고, 해당 업무와...
1	한국콜마 / 생산 / 2018 하반기	지방대 / 기계과 / 학점 3.4 / 토익 785, 토스 130 / 봉사활동 1회,...	한국콜마의 인재상인 4성 5행 中 가장 중요하다고 생각하는 것을 택 1 하여 그 정...



index	기업	직무	합격년도	학교	전공	학점	스펙_1	스펙_2	스펙_3	자기소개서	질문_1	답변_1	...	질문_1	답변_1
0	한국농수산물유통공사	회계직	2020_상반기	부산진여자상업고등학교	금융경영	학점 X	한국사검정시험: 고급, 컴퓨터활용능력: ...	NaN	NaN	1. 공사의 업무 중 본인이 가장 관심 있는 업무 한 가지를 제시하고...	1. 공사의 업무 중 본...	<30년 후의 회계 전문가를 위한 활동>Wn가장	...	None	None
1	한국콜마	생산	2018_상반기	지방대	기계과	학점 3.4	토익 785, 토스 130	봉사활동 1회,인턴 1회	NaN	한국콜마의 인재상인 4성 5행 中 가장 중요하다고 생각하...	None	None	...	None	None

- info열에 저장된 기업, 직무, 지원연도를 '/'를 기준으로 분리하여 각각 저장
- specification열에 저장된 출신학교, 학점 , 스펙 등을 '/'를 기준으로 spec_n으로 저장
- self_intro열에 저장된 자기소개서의 패턴을 파악하여 자기소개서, 질문_n, 답변_n으로 저장

데이터수집 및 전처리

[데이처 전처리]

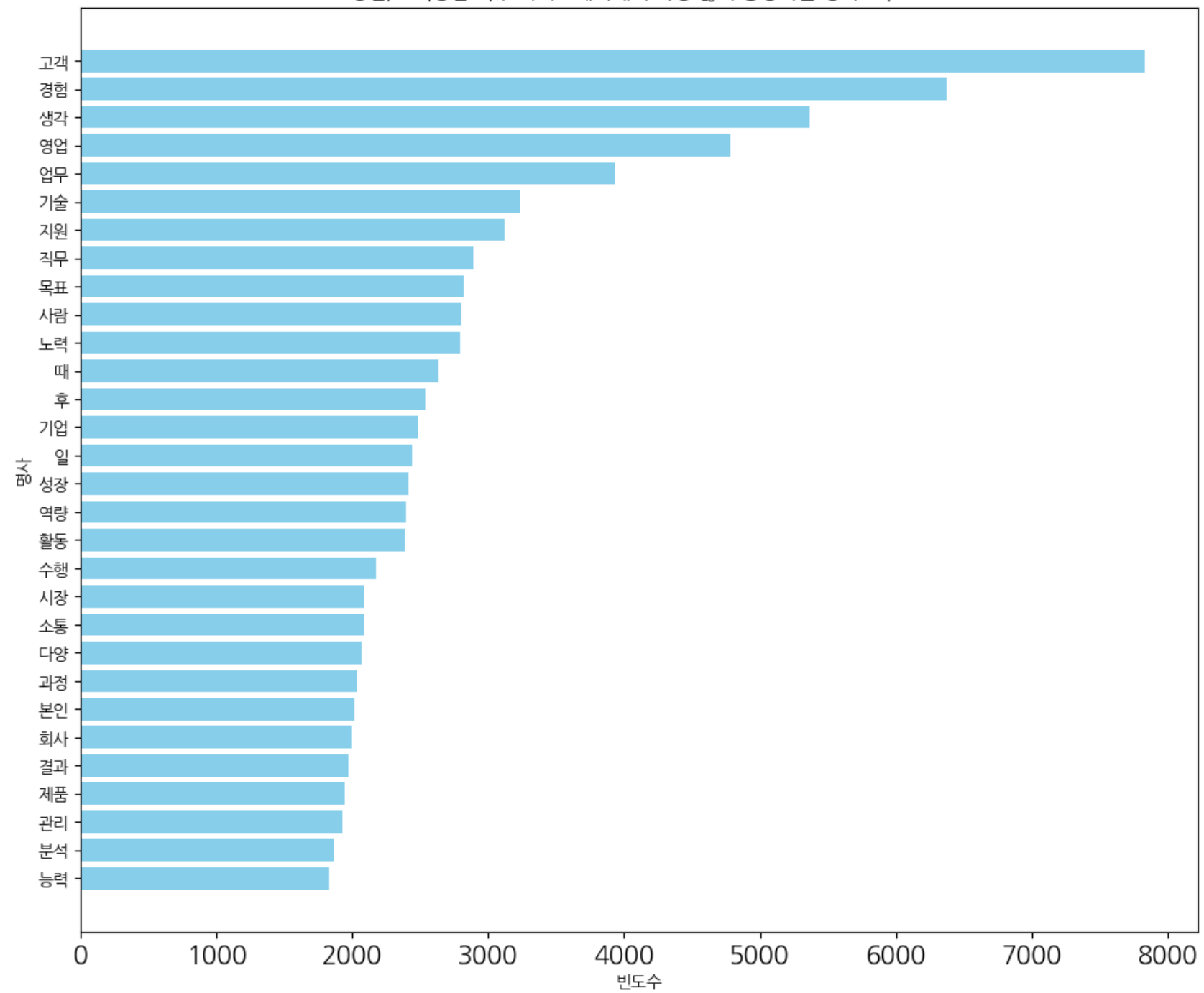
직무를 6개의 범주로 변환

- `CS,cs,고객,응대,영업` 이 포함된 행을 `영업/고객상담` 범주로 변환
- `경영,지원,인사,노무,채용,면접,전략,ERP,재무,세무,회계` 이 포함된 행을 `경영/사무` 범주로 변환
- `마케팅,광고,홍보,기획` 를 `마케팅/홍보/광고` 이 포함된 행을 범주로 변환
- `생산,관리,제조,가공,설비` 를 `생산/제조` 이 포함된 행을 범주로 변환
- `연구,R&D,설계` 이 포함된 행을 `연구개발/설계` 범주로 변환
- `개발,SW,sw,웹,안드로이드,보안,네트워크,데이터,data,Data,AI` 이 포함된 행을 `IT` 범주로 변환

1. KoNLPy 패키지 사용
2. 꼬꼬마 분석기 kkma 사용
3. 직무별로 자기소개서 형태소 분석
4. 'NNG'(보통명사), 'NNP'(고유명사) 추출 후 단어 빈도 확인
5. 단어빈도가 가장 많은 30개 단어를 시각화

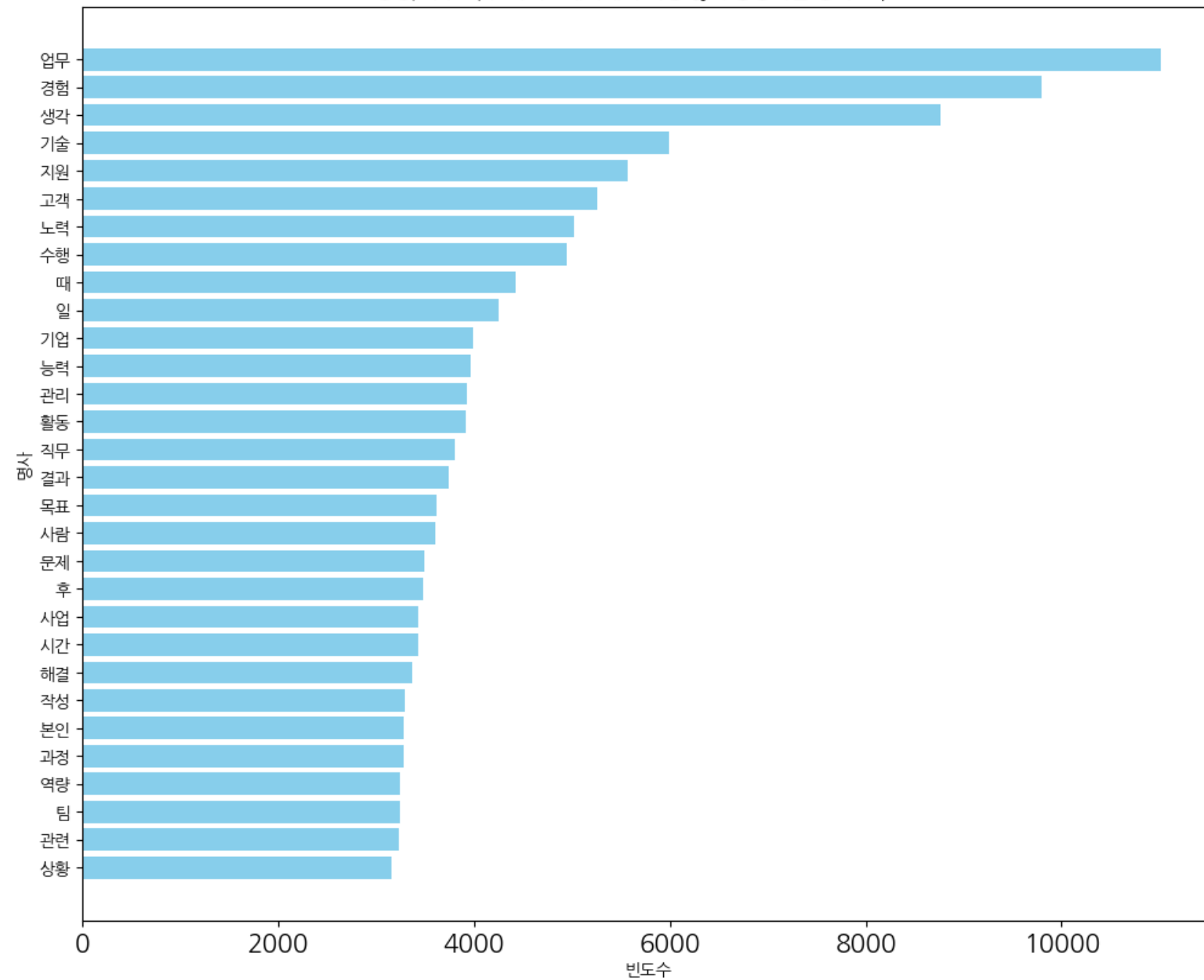


영업/고객상담 직무 자기소개서에서 가장 많이 등장하는 명사 Top 30

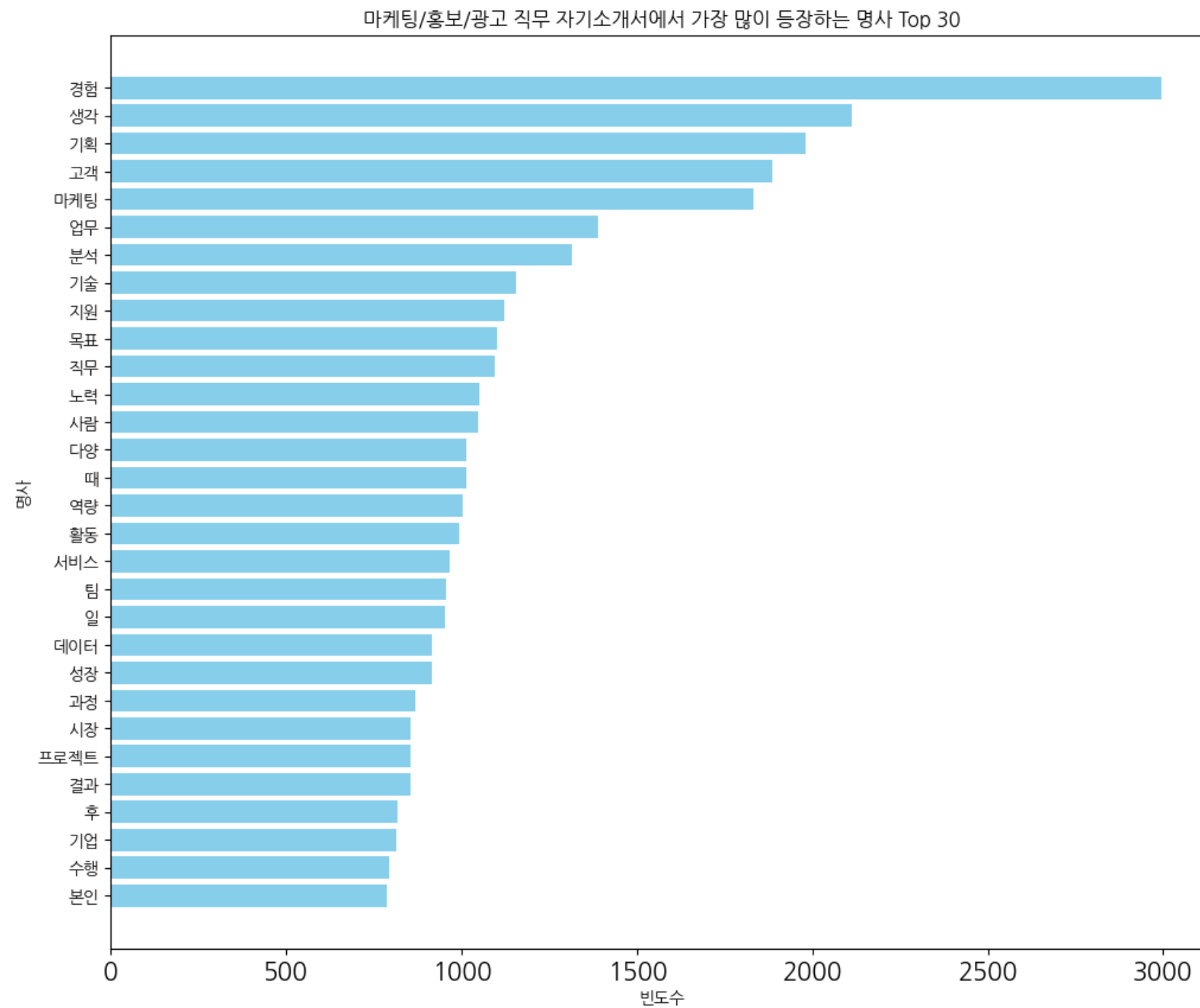


- 고객, 경험, 생각, 영업 등 상위권

경영/사무 직무 자기소개서에서 가장 많이 등장하는 명사 Top 30

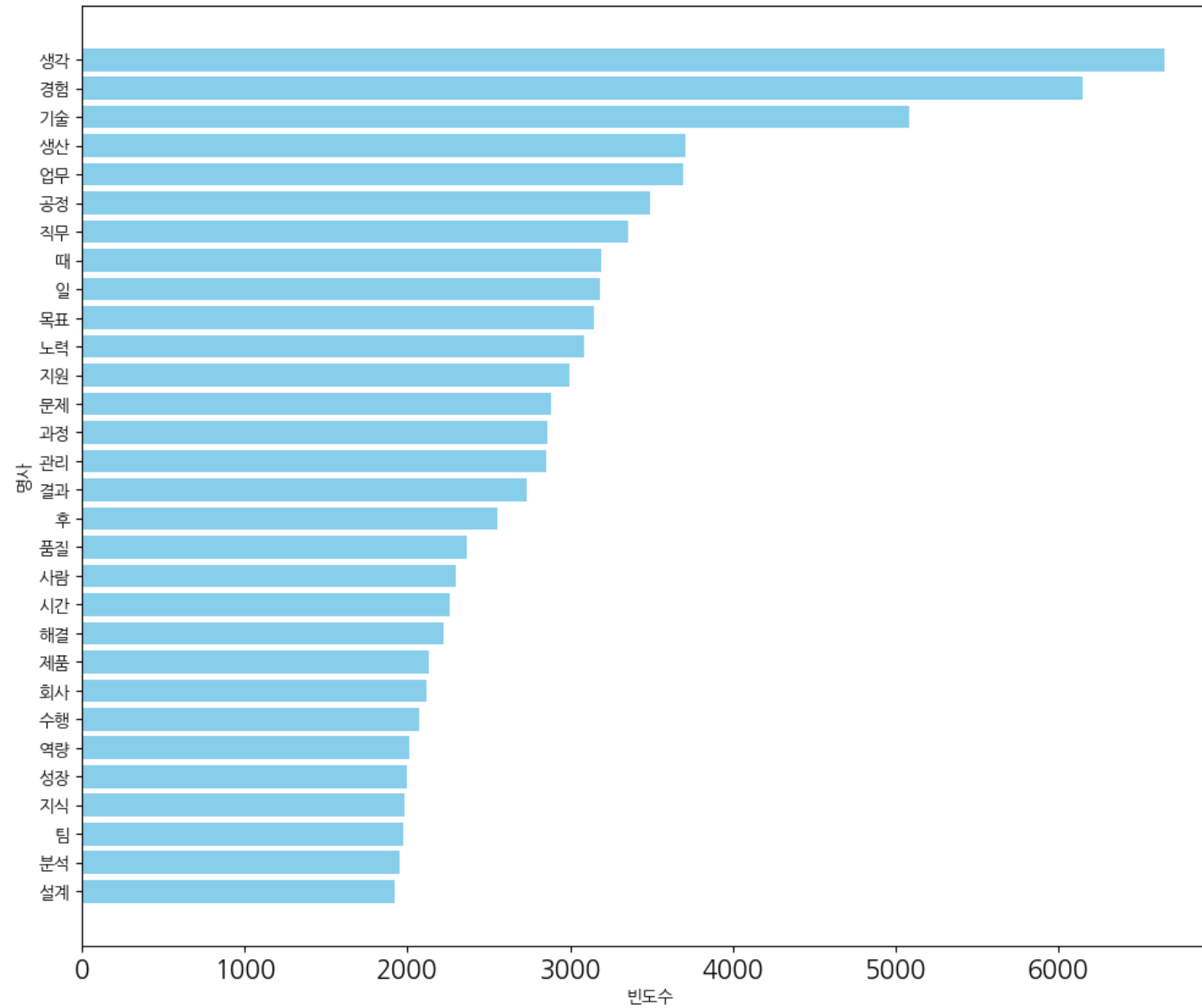


- 업무, 경험, 기술, 지원 등 상위권



- 경험, 생각, 기획, 마케팅 등 상위권

생산/제조 직무 자기소개서에서 가장 많이 등장하는 명사 Top 30

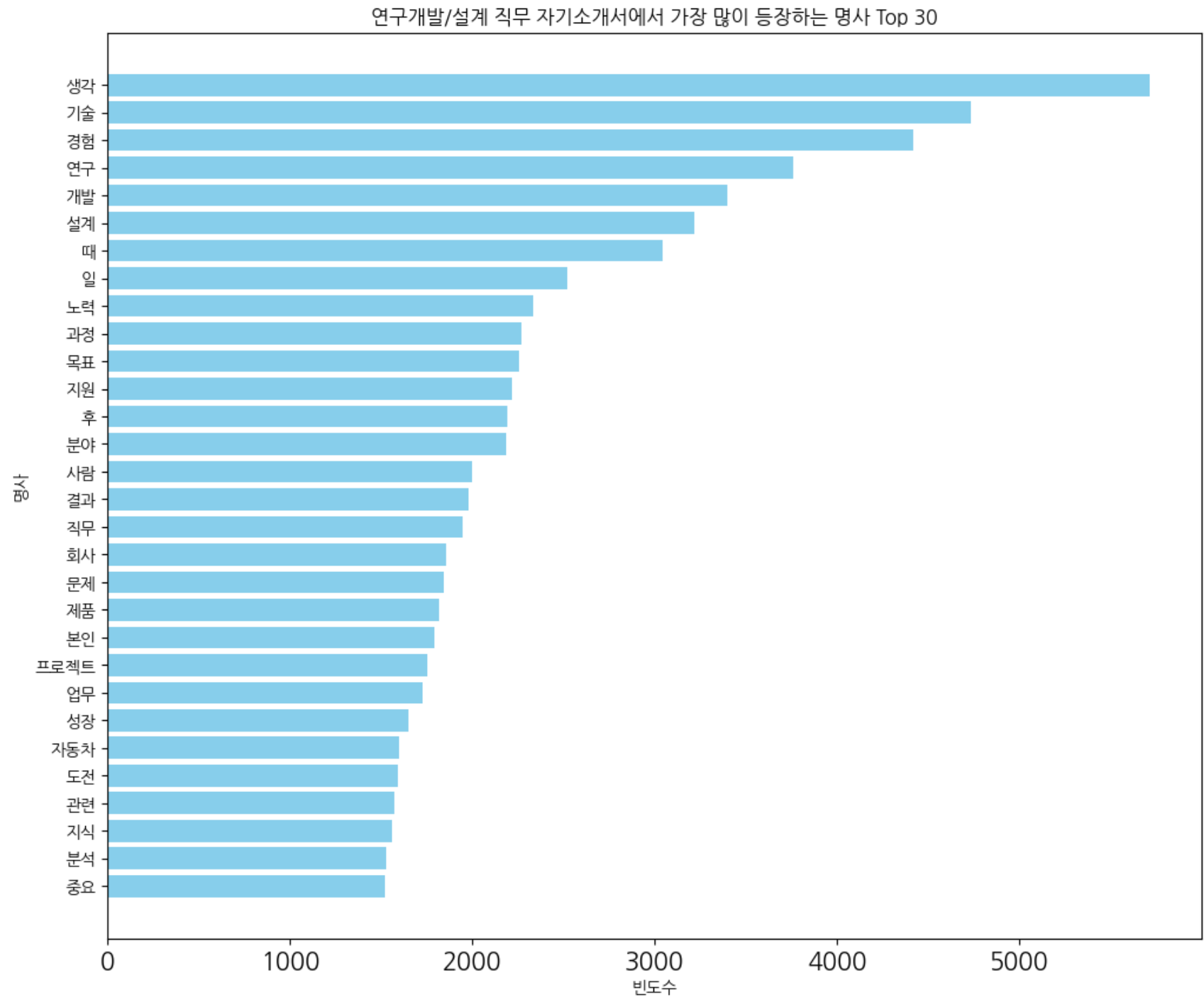


- 생각, 경험, 기술, 생산, 공정 등 상위권

04

분석 결과

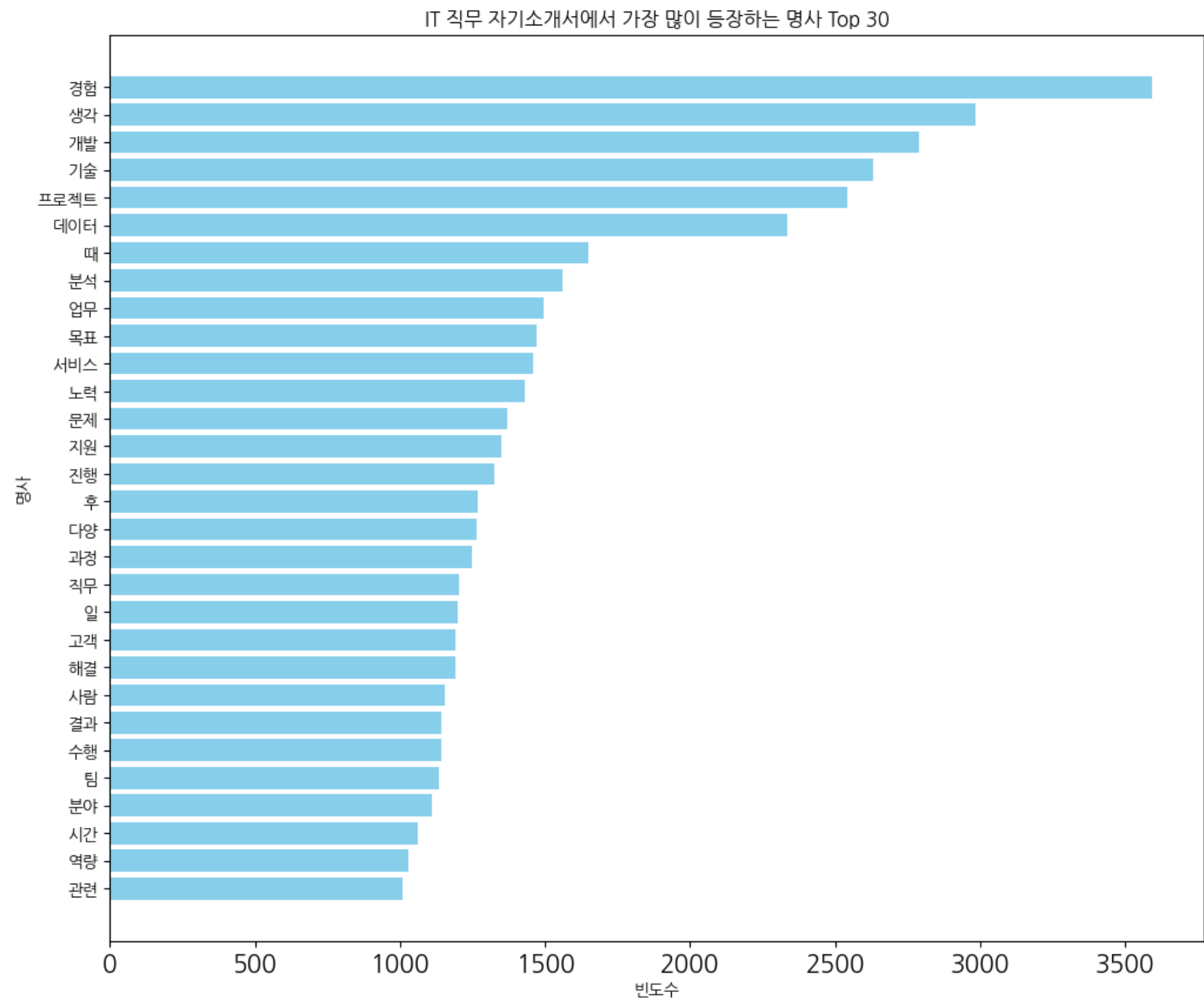
[연구개발/설계 직무]



- 생각, 기술, 경험, 연구, 개발, 설계 등 상위권

분석 결과

[IT 직무]



- 경험, 생각, 개발, 기술, 프로젝트, 데이터 등 상위권

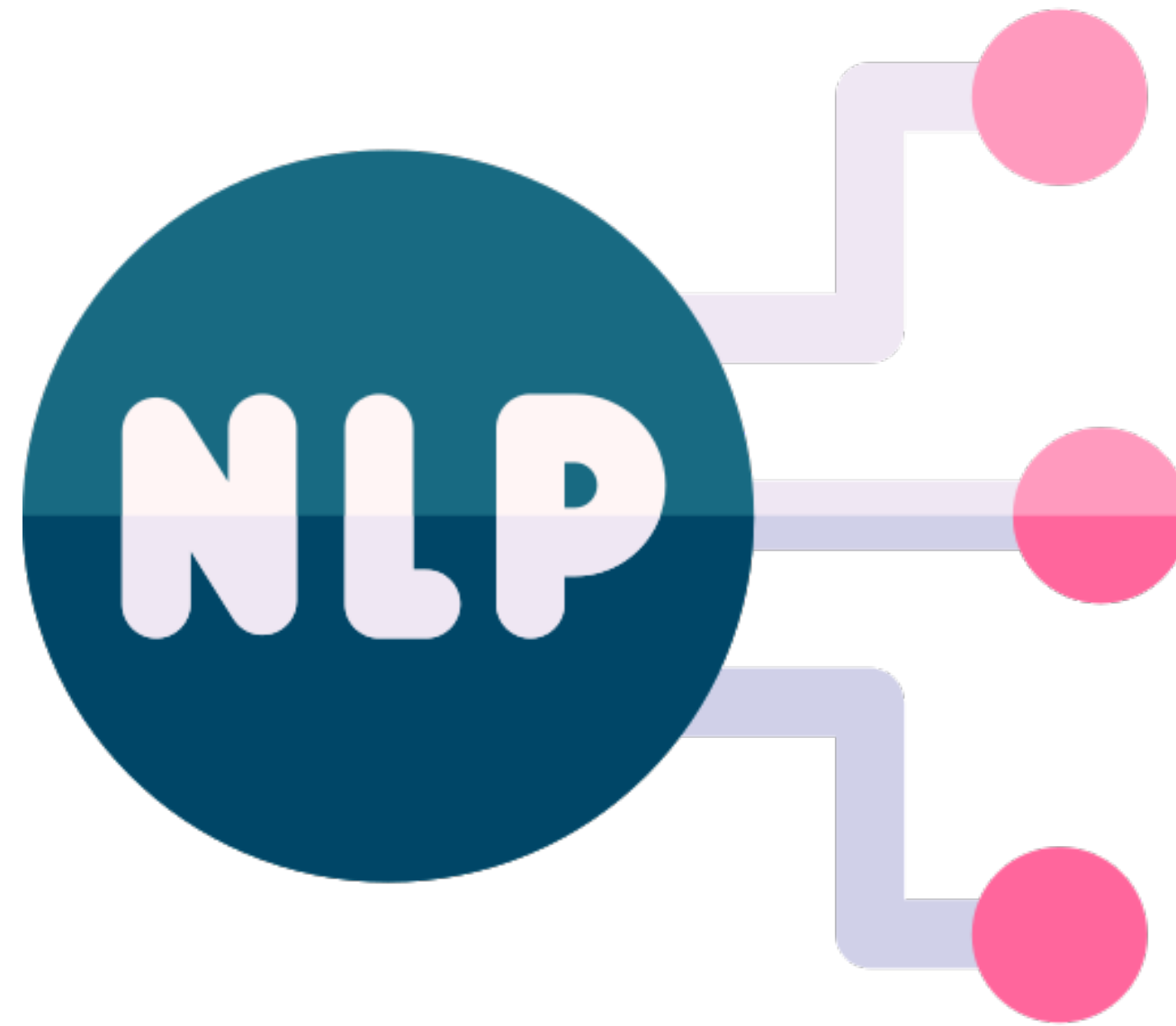
분석 결과

04

[최종 분석 결과]

- 모든 직무에서 경험을 가장 중요하게 생각
- 영업/고객상담 직무에서는 지원, 제품, 관리, 소통
- 경영/사무 직무에서는 업무, 지원, 해결, 관리, 상황
- 마케팅/홍보/광고 직무에서는 기획, 마케팅, 서비스, 데이터, 시장
- 생산/제조 직무에서는 문제, 관리, 품질, 분석
- 연구개발/설계 직무에서는 지식, 분석
- IT 직무에서는 서비스, 문제, 팀, 역량, 시간

1. 데이터 전처리에서 누락된 데이터를 포함하여 토큰화
2. 불용언 제거 등 세부적인 데이터 전처리 필요
3. 기업, 연도 별 단어 트렌드 분석
4. 기업, 직무 별 다중 분류 분석
5. 연관단어 분석 및 키워드 주변 단어 분석
6. 불합격 데이터를 추가하여 합격,불합격 분류 모델



NLP에 대한 이해도 부족

- 사회 취약 계층에게 취업 지원 서비스
- 학부 저학년들의 취업 진로 방향 도움
- 취업 자기소개서 작성 시 가이드라인 제공

자료 출처

- 합격 자기소개서 데이터:
 - 링커리어:<https://linkareer.com/cover-letter/search>
- KoNLPy:
 - (박은정, 조성준, “KoNLPy: 쉽고 간결한 한국어 정보처리 파이썬 패키지”, 제 26회 한글 및 한국어 정보처리 학술대회 논문집, 2014.)

**THANK YOU FOR
ATTENTION**

