

2021

빅데이터

분석을 통한
중국 시장 국내
중소기업 살아남기!

금경용

2021년 04월 02일

목요일 3시 30분

빅데이터 BI 전문가 양성(파이썬, R)

목차

▶ 서론 및 이론적 배경

일정관리 및 개발환경

연구 방법 및 목적

▶ 실증 분석

빅데이터 텍스트 마이닝을 이용한 “한국” 키워드 분석

중국 인구 소득 및 지출

중국 수입품목 조사 및 연령 조사

국내 거주하고 있는 중국 유학생 분석

2

▶ 결론

연구의 결과 및 시사점

연구 한계 및 향후 연구 방향

※ 개발환경

기간

2021.03.29 ~ 2021.04.02

참여 인원

1명

OS

Windows 10 Pro

Language

Python 3.8.5

IDE

Anacomda jupyter notebook

Open
Source

LAC/ selenium/ requests/ BeautifulSoup³/ WordCloud/ Collections / warnings
folium/ Collections/ pandas/ numpy/ matplotlib.pyplot/ seaborn/

※수집 데이터

※웹 크롤링을 통한 데이터



-신문 400개의 데이터 수집

※CSV 데이터



중국 국가 통계국



중국 국가 통계국



교육통계서비스

-2020 고등교육기관 외국인 유학생 현황

-대학교 위도 경도 자료

-중국 국민 소득과 지출데이터

-중국 수입물품 데이터

-중국 소비자 한국 온라인쇼핑 판매 품목 데이터

※실행 함수 이용

-중국 홈페이지를 웹 크롤링 및
워드클라우드-

Python 실행어	사용도
from selenium import webdriver	웹 브라우저 호출
import requests	HTTP link 가져오기
from bs4 import BeautifulSoup	URL-HTTP 정제 및 추출(어근)
from LAC import LAC	중국어 형태소 분석
from wordcloud import WordCloud	워드 클라우드
from PIL import Image	이미지 불러오기
import matplotlib.pyplot as plt	그래프 시각화
import collections	글자 수

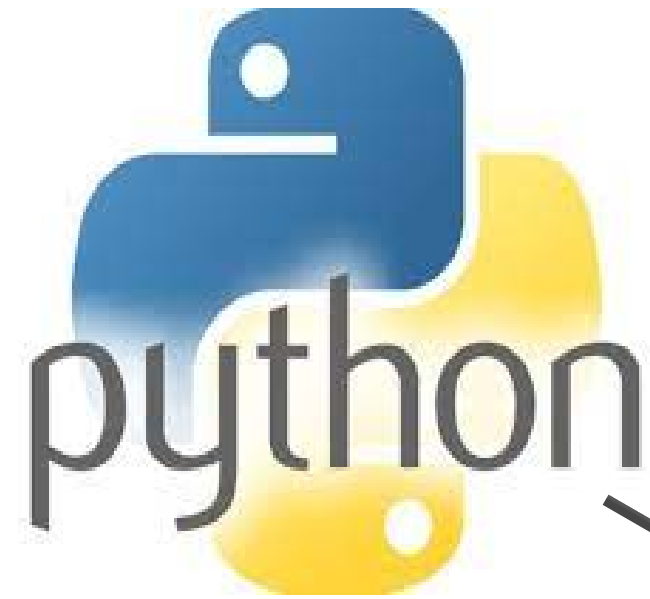
- 데이터 전처리 및 시각화-

Python 실행어	사용도
Import pandas	데이터 전처리
Import warnings	시각화시 에러 미발생
import matplotlib.pyplot	그래프 시각화
Import seaborn	그래프 시각화
Import folium	지도 시각화

※일정관리

활동내용	일정표																			
	3/29.오전		3/29.오후		3/30.오전		3/30.오후		3/31.오전		3/31.오후		4/1.오전		4/1.오후		4/2.오전		4/2.오후	
기획 및 설계																				
주제선정 및 요구사항, 일정관리	계획																			
	완료		완료																	
설계, 기능정의, 데이터수집	계획																			
	완료		완료		완료		완료													
서론 및 이론적 배경																				
연구 배경 및 목적	계획																			
					완료															
연구 내용 및 방법	계획																			
							완료													
프로젝트의 구성	계획																			
							완료													
실증 분석 및 결론																				
각자에 맞는 데이터분석 및 기술통계분석	계획																			
	완료		완료		완료		완료													
추가 분석 및 데이터 시각화	계획																			
	완료		완료		완료		완료													
테스트 및 수정	계획																			
																	완료		완료	

연구방법



Python 3.8.5

텍스트 마이닝
빅데이터 분석



통계청, 한국무역협회
중국국가통계국(国家
统计局)

결과 도출



* 목적

2017년 한반도의 사드 배치로 중국의 사드 보복으로 인해 상당한 경제 피해손실을 받을 뿐만 아니라 앞으로의 중국 시장의 진출의 상당한 걸림돌과 높은 진입장벽을 구축하였다.

2020년 한중관계가 개선되는 조짐이 보였으나 현재, 2021년 북미, 미-중 관계가 나날이 악화와 북한의 미사일 기술이 발전함에 따라 대한민국의 사드 미사일 업그레이드가 불가피해진 상황에서 “百度” 웹 크롤링 키워드 분석과 빅데이터 분석을 통해 최소한의 투자로 효율을 낼 수 있는 방법을 모색해 볼 것이다.

01.

百度 “한국”
키워드 맵 크롤링

"한국"
키워드 검색



| 韩国



百度一下

百度热榜

换一换

1 江苏致11死车祸原因:货车轮胎脱落 **热**

2 BBC涉疆报道画面和解说词完全相反

3 云南新增本土确诊病例2例

4 泰国考虑为外国游客提供免费机票

5 黎智英等3人认罪

6 清华北大30名学霸参军入伍

韩国“遗憾”,日本“晴天霹雳”



中国青年网 33分钟前

韩国“感到遗憾” 朝鲜参加2018年在韩国举办的平昌冬奥会,是此前令人瞩目的拉近朝韩外交关系中的一个决定性因素。朝鲜领导人金正恩的妹妹金与正当时出席了平昌冬奥会,而且朝鲜的参加就是一大广告。平壤6日的宣布终结了首尔的希望,后者原... 百度快照

中国女足“生死战”对阵韩国,我们的铿锵玫瑰从未如此神秘



澎湃新闻 40分钟前

4月8日15时,中国女足就将迎来客场比赛,两队将争夺仅有的一张东京女足,一边是从未拿到过奥运会入空气中弥漫。相比以往的多次...

韩国单日新增病例创近3个月新高

中国新闻网 45分钟前

中新社首尔4月7日电(记者 曾燕)韩国非首都圈疫情恶化,全据韩国中央防疫对策本部7日通报,过去24小时,新增确诊病例今年1月8日以来单日新增病例最高值。 韩国... 百度快照

前瞻-神秘之师低调出征韩国 中国女足首战



网易体育 21分钟前

Python 실행어	사용도
from selenium import webdriver	웹 브라우저 호출
import requests	HTTP link 가져오기
from bs4 import BeautifulSoup	URL-HTTP 정제 및 추출(어근)
from LAC import LAC	중국어 형태소 분석
from wordcloud import WordCloud	워드 클라우드
from PIL import Image	이미지 불러오기
import matplotlib.pyplot as plt	그래프 시각화
import collections	글자 수

* 시각화



* 도표

중국어	한국어	횟수
韩国	한국	287
身材	몸매	106
中国	중국	101
手机	핸드폰	72
人	사람	62
文化	문화	54
LG	LG	48
电影	영화	43
运动	운동	41
日本	일본	39
健康	건강	36
市场	시장	34
业务	비즈니스	33
行业	업종, 산업	32
价格	가격	31

2021 / 04 / 02

(중국) 도시 인구

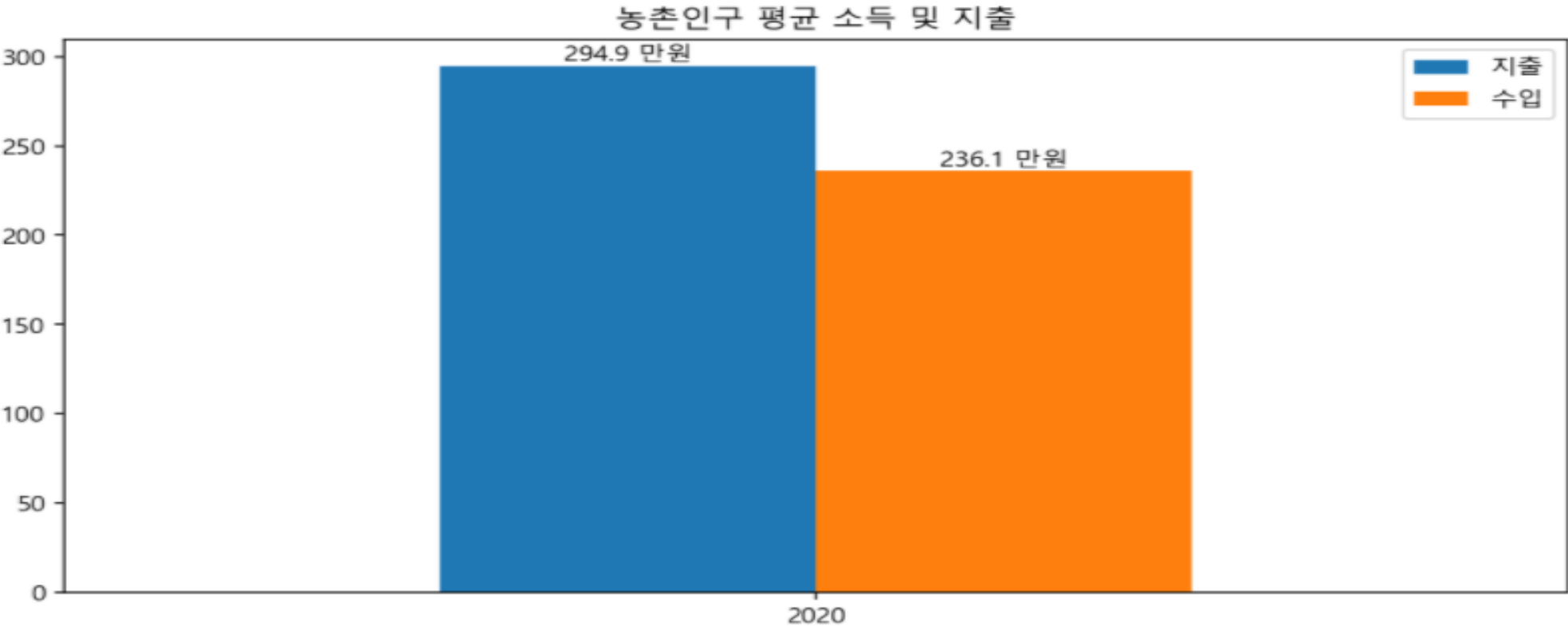
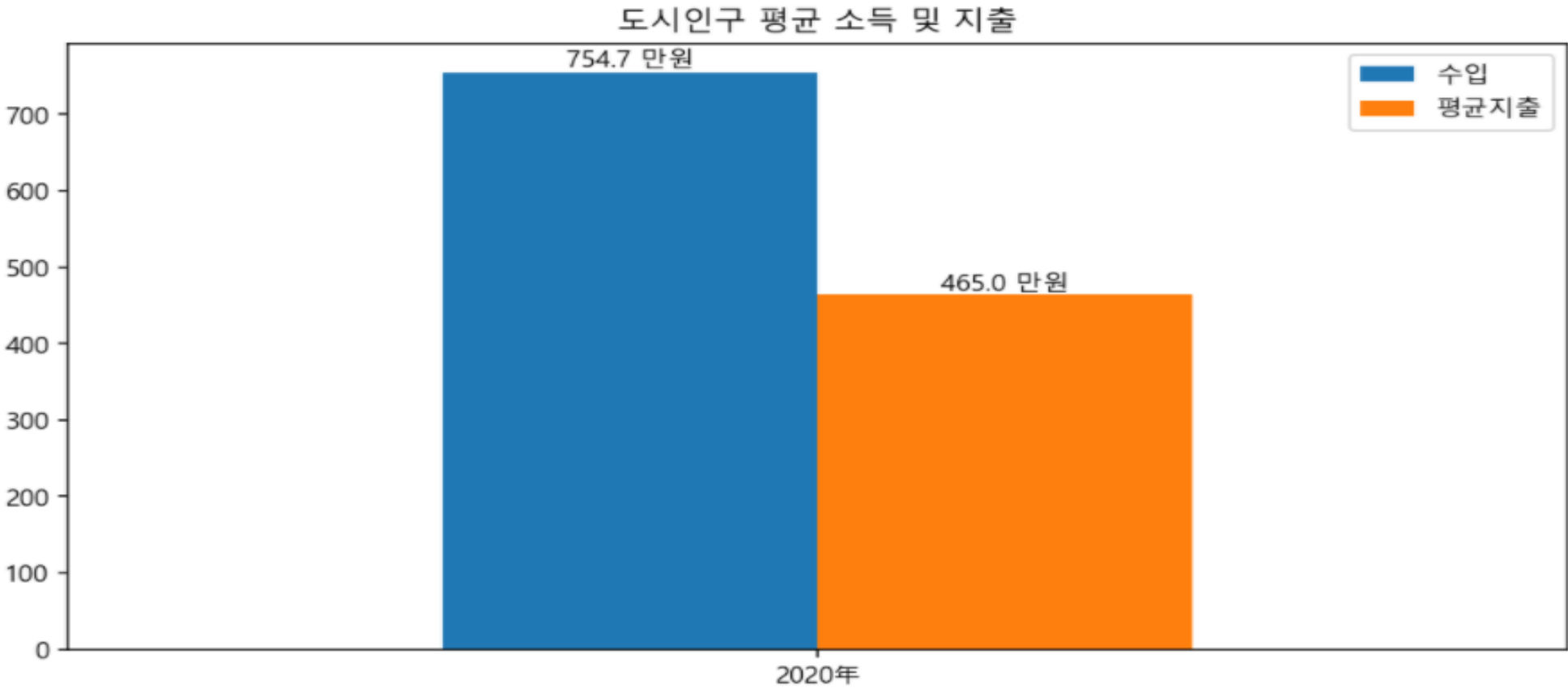
소득 및 지출

분석

금 경 용

도시 인구 2020년 평균
수입: 754,7만원
지출: 465.0만원

18



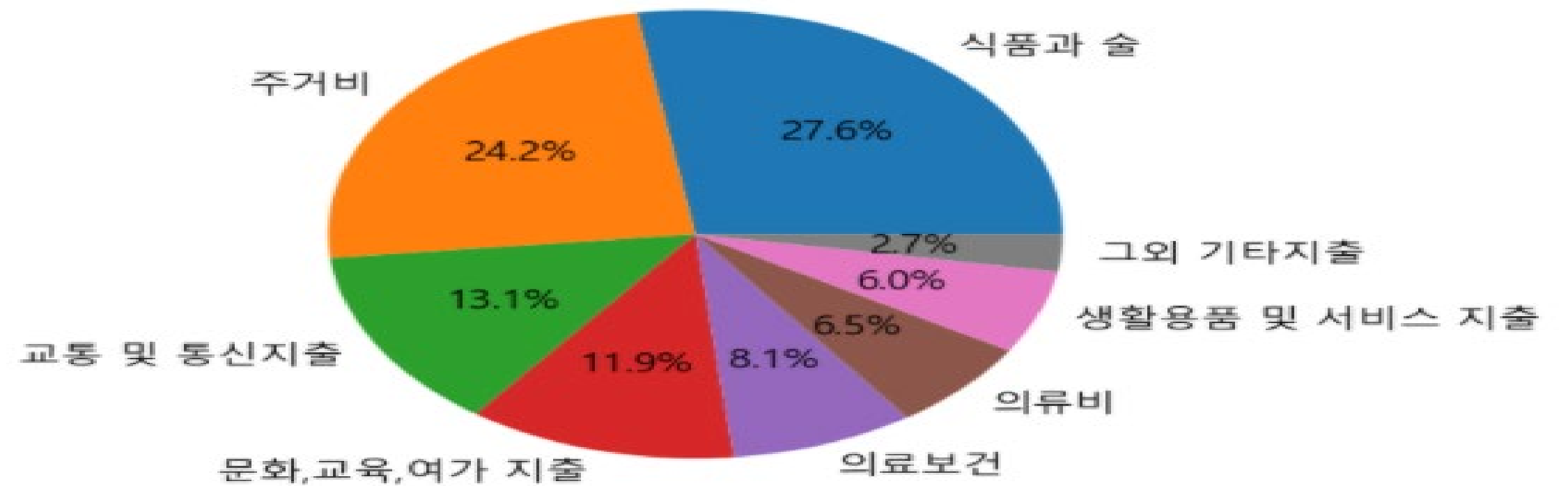
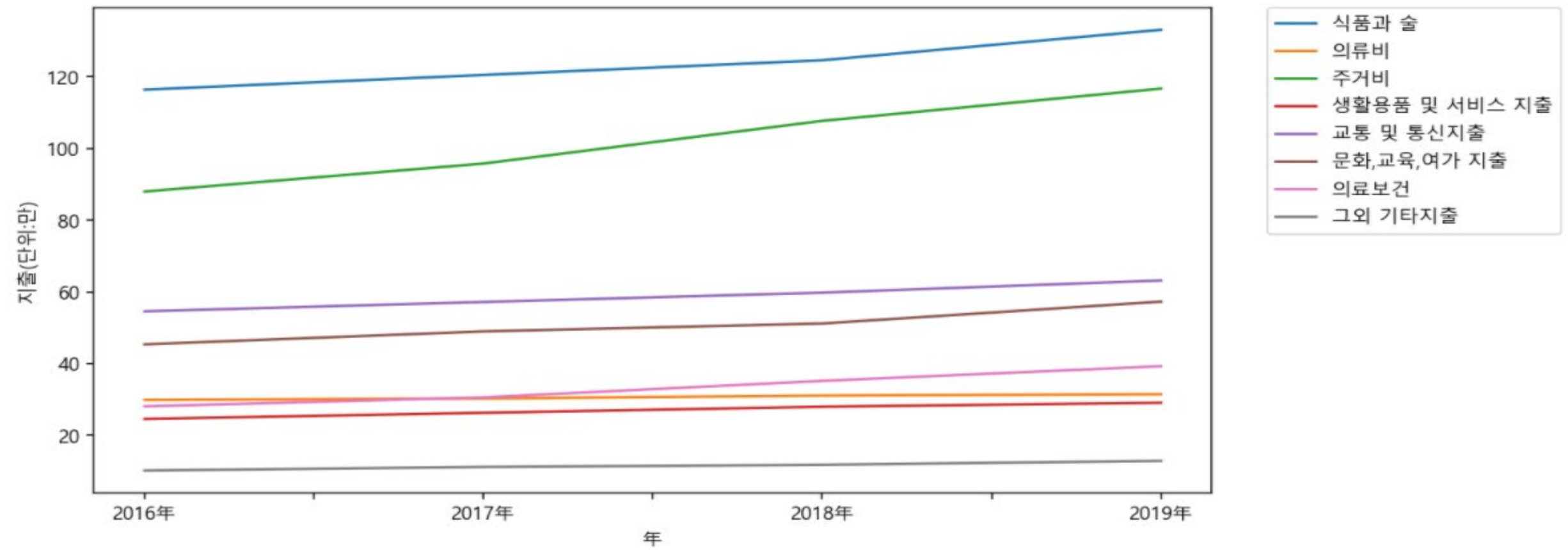
농촌 인구 2020년 평균
수입: 294,9만원
지출: 236.1만원

02

2020 년 지출

식품 & 술	:27.6%
주거비	:24.2%
교통 및 통신지출	:13.3%
문화,교육,여가,지출	:11.9%
의료보건	:8.1%
의류 비	:6.5%
생활용품 및 서비스 지출	:6.0%
그 외 기타지출	:2.7%

2021 / 04/ 02



금 경 용

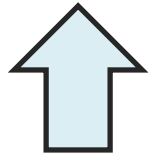
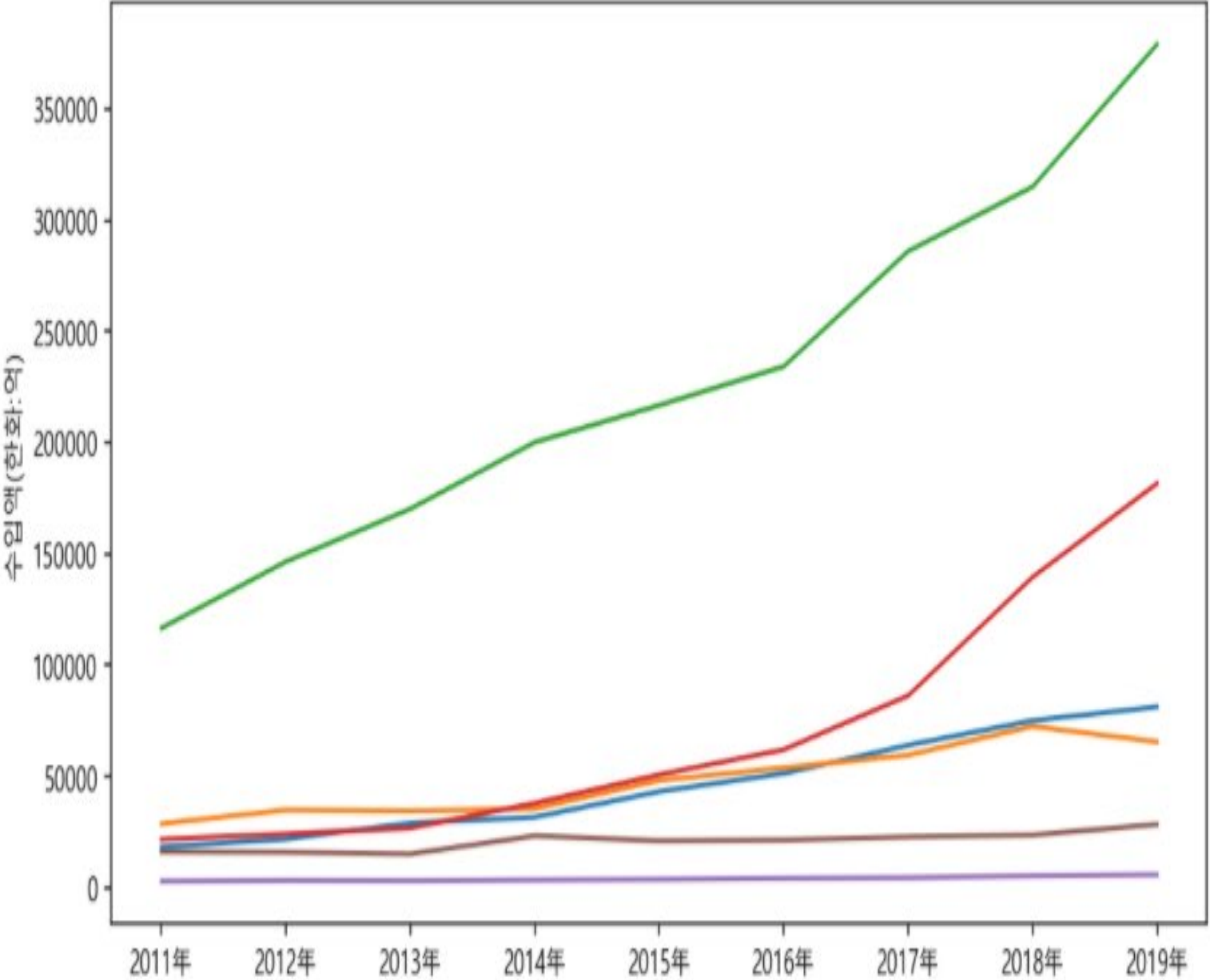
2 0 2 1 / 04/ 02

03. ■

중국의 “수입”

품목 조사

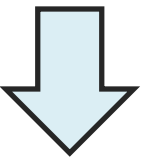
금 경 용



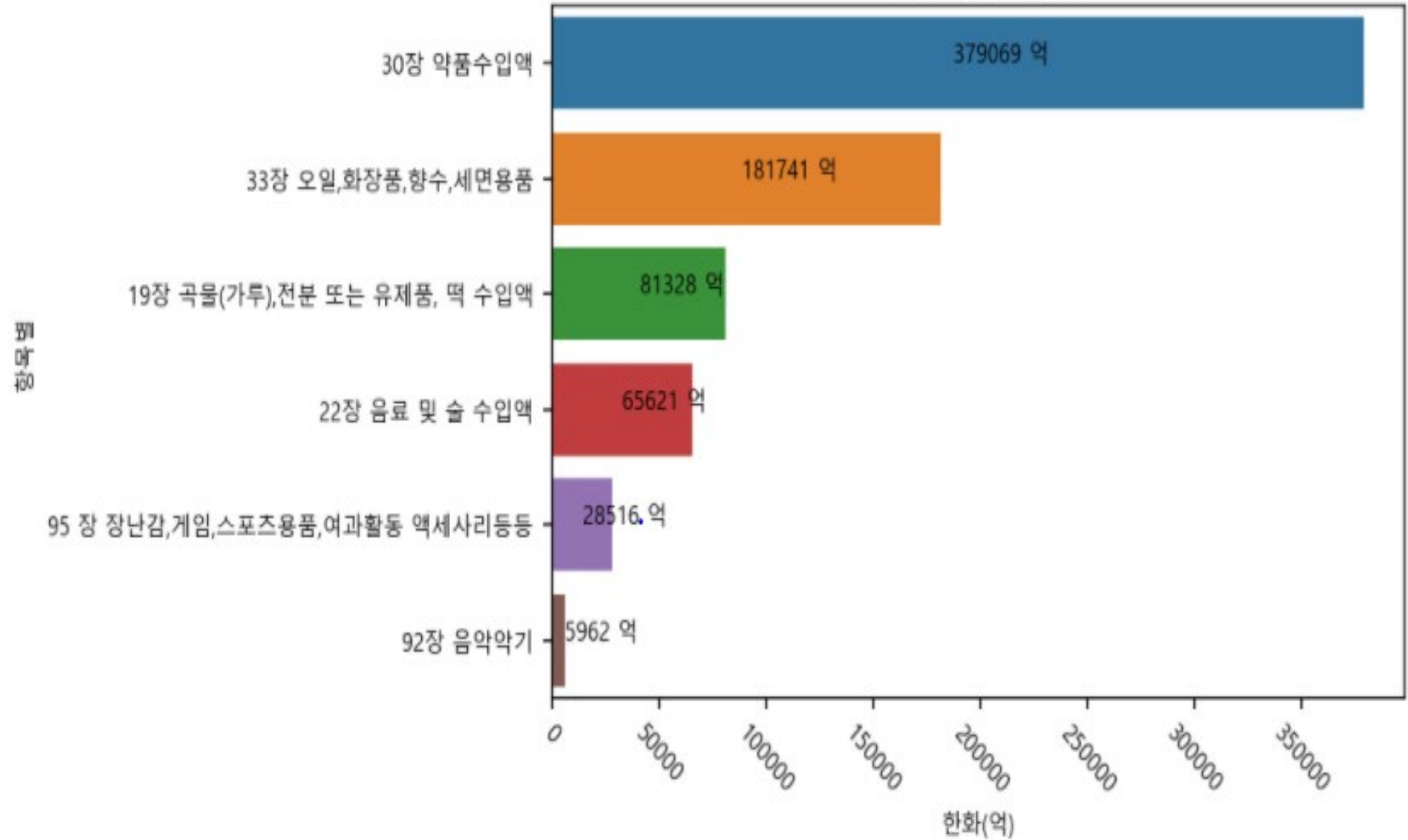
2011~2019 수입추세

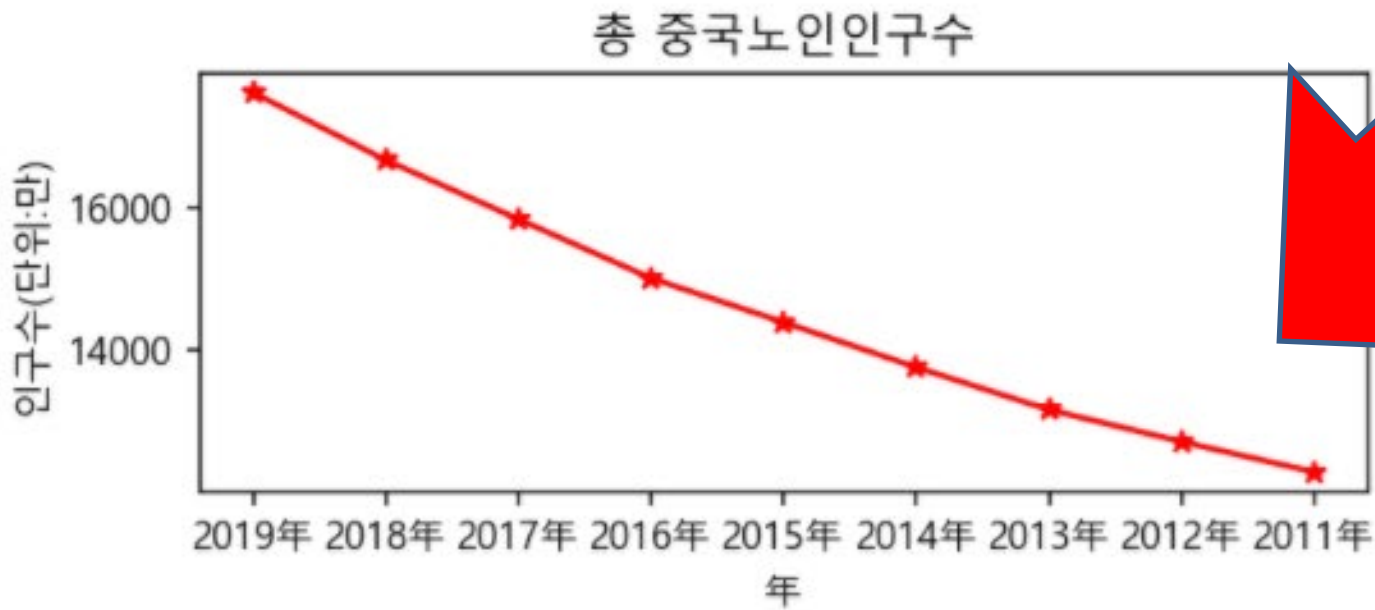
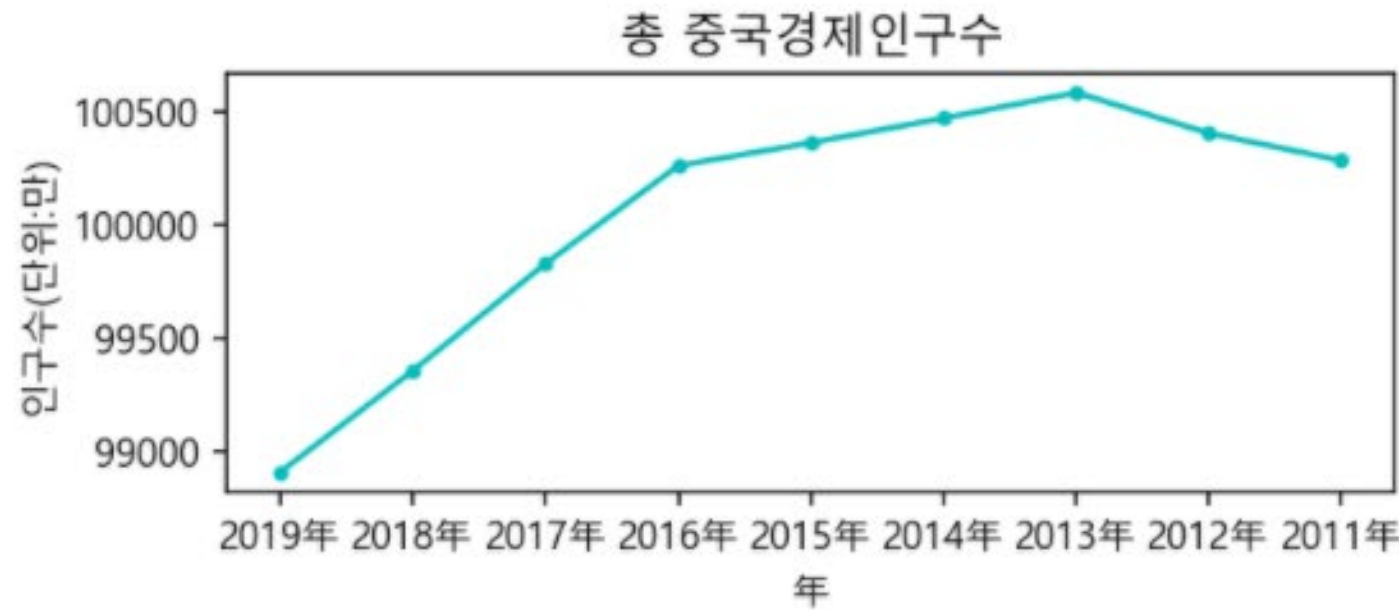
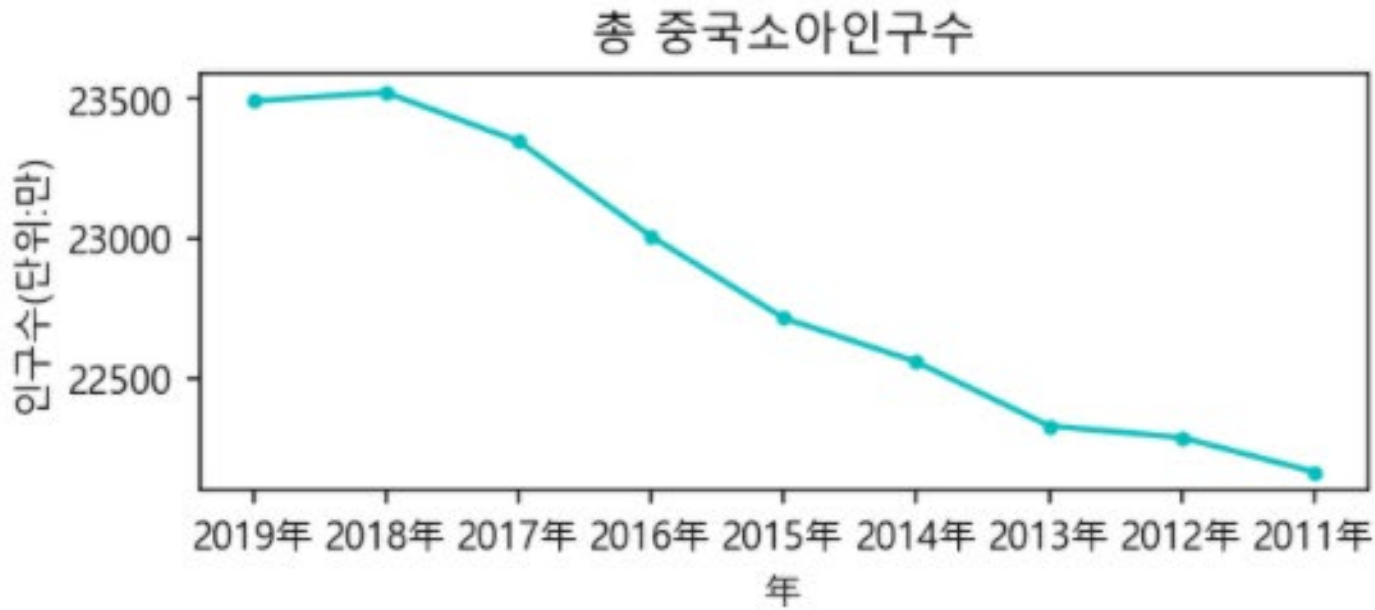
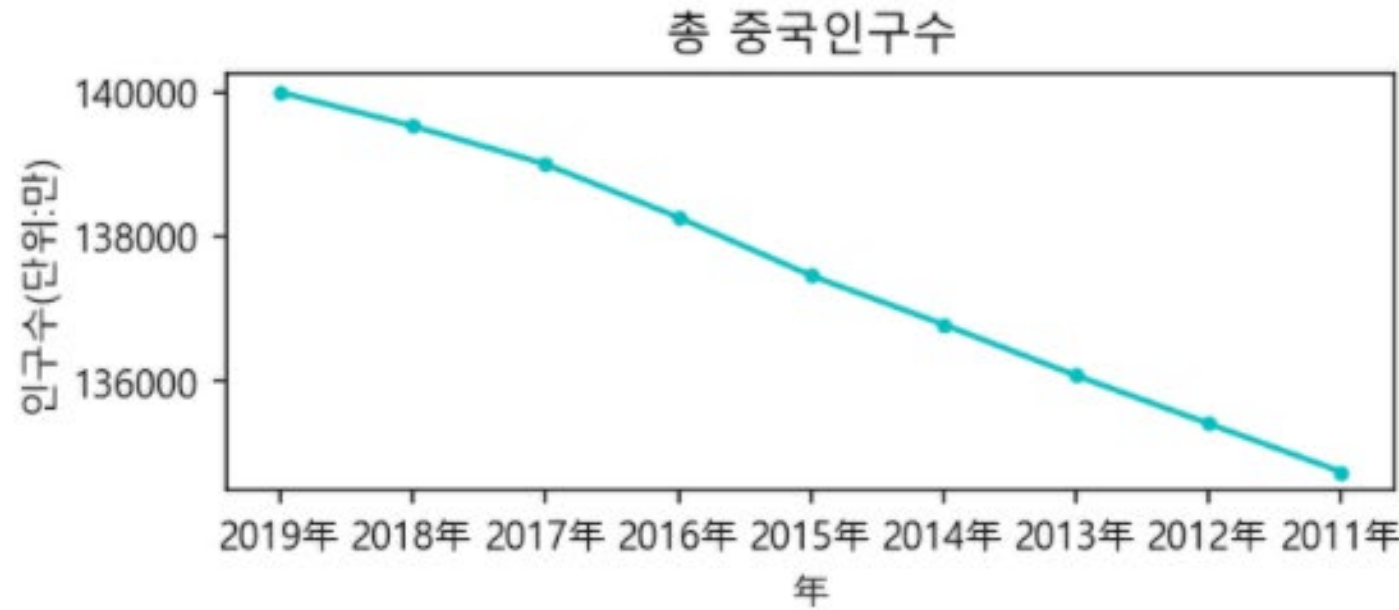
- 19장 곡물(가루),전분 또는 유제품, 떡 수입액
- 22장 음료 및 술 수입액
- 30장 약품수입액
- 33장 오일,화장품,향수,세면용품
- 92장 음악악기
- 95 장 장난감,게임,스포츠용품,여과활동 액세서리등등

2019

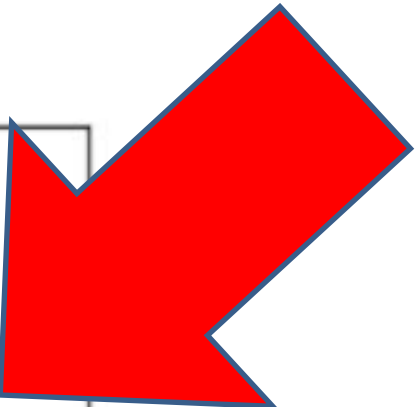


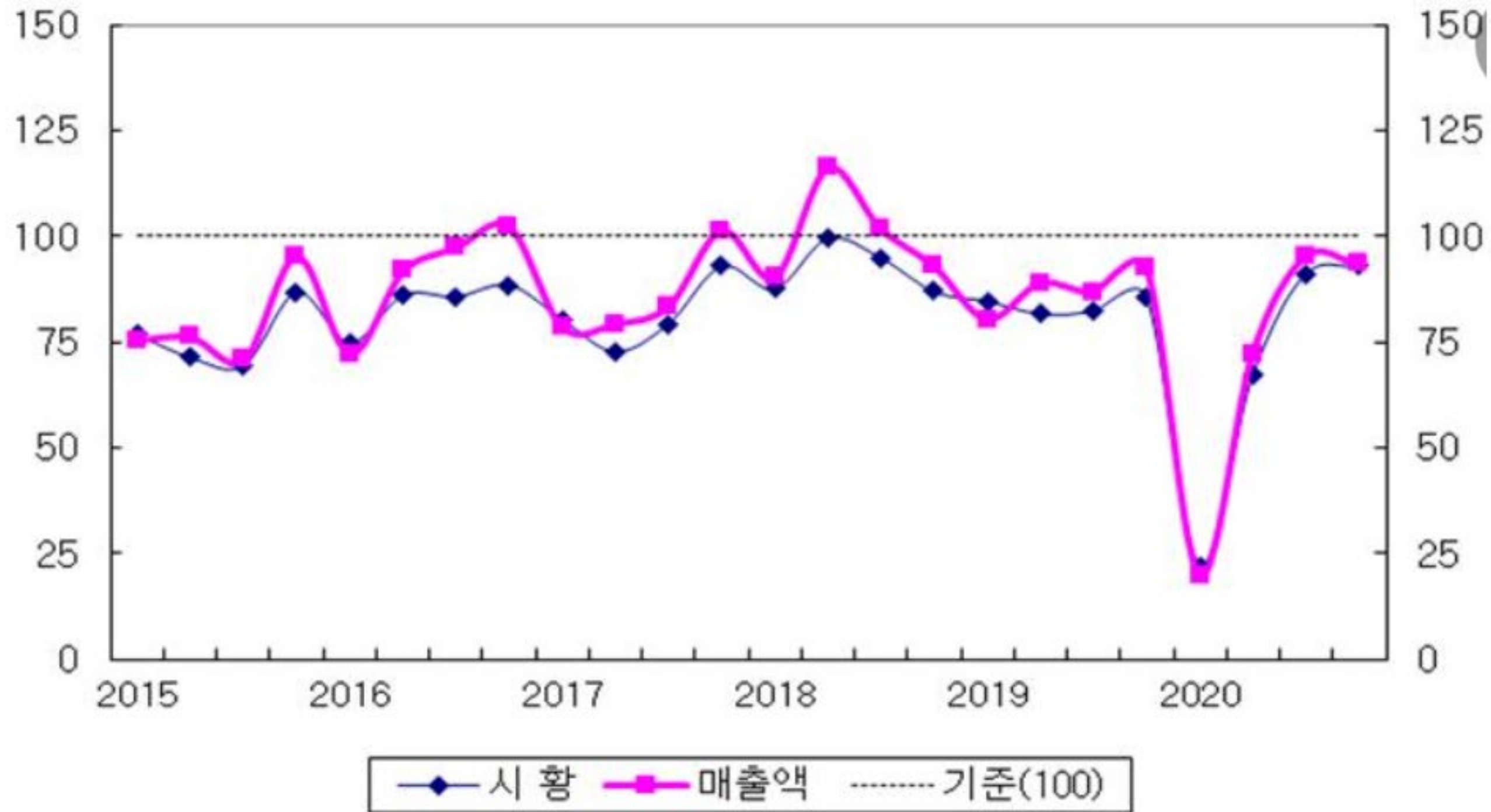
6개 항목 수입품목





매년
1000만명





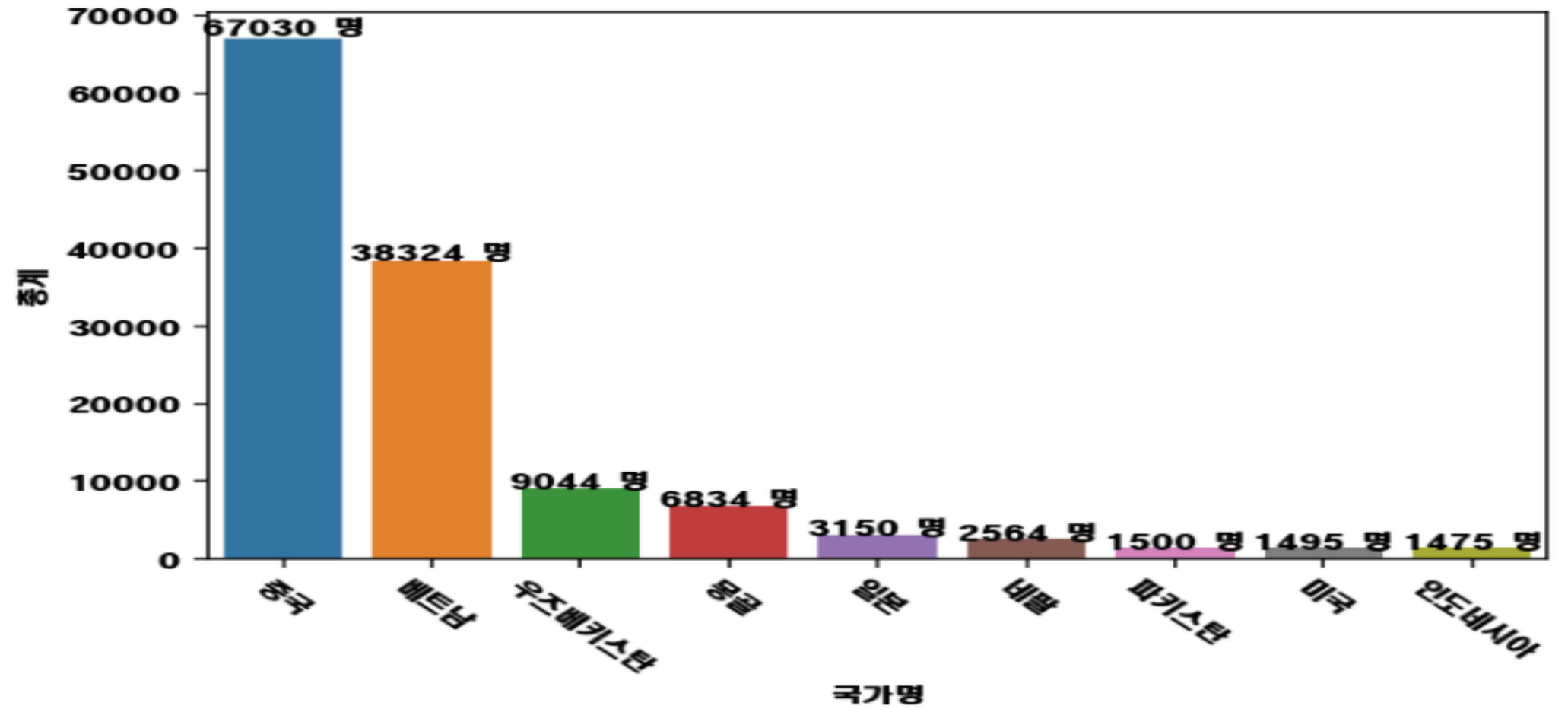
출처:산업연구원(KIET)

2 0 2 1 / 04/ 02

04.

국내 거주하고 있는
중국 유학생

나라별 국내 유학생

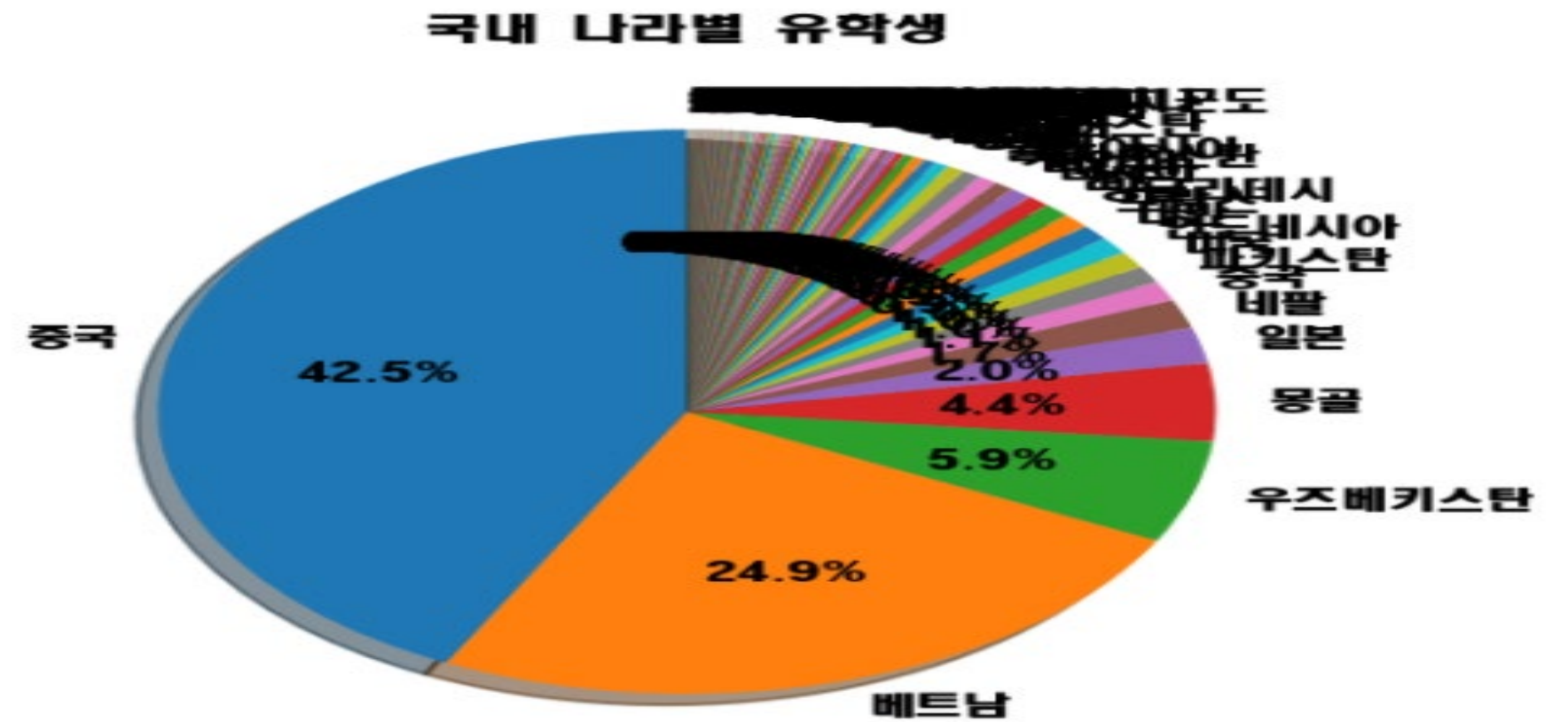


01

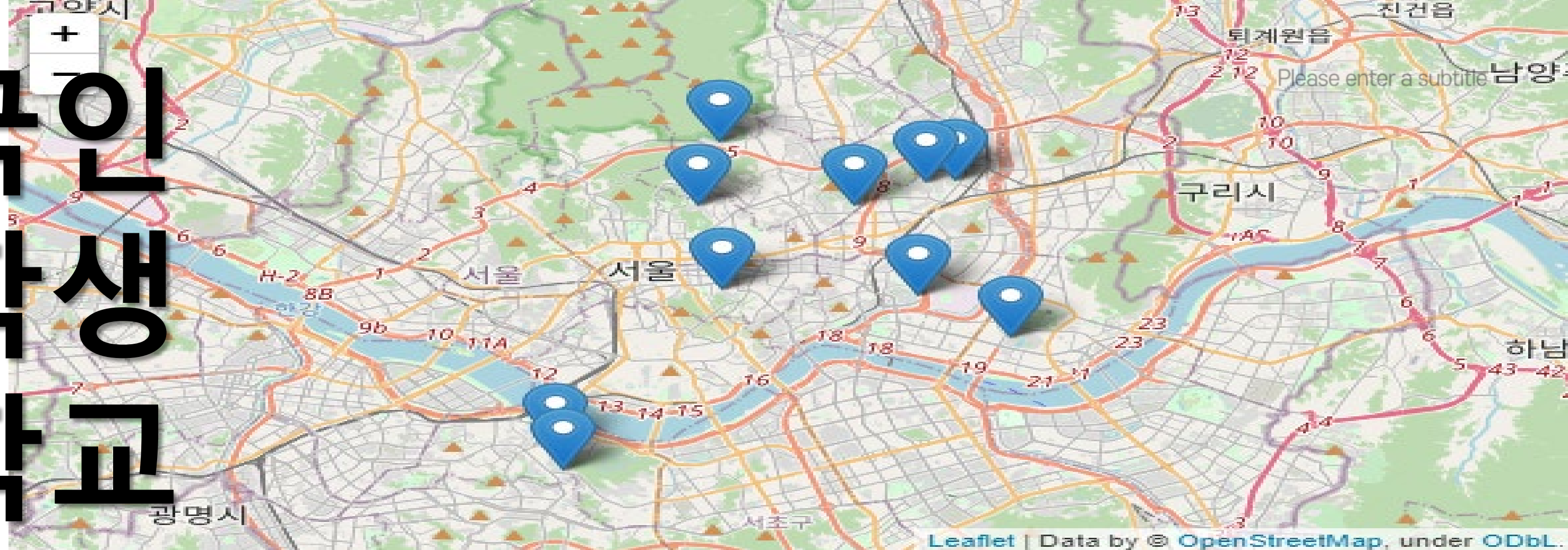
Top 10

02

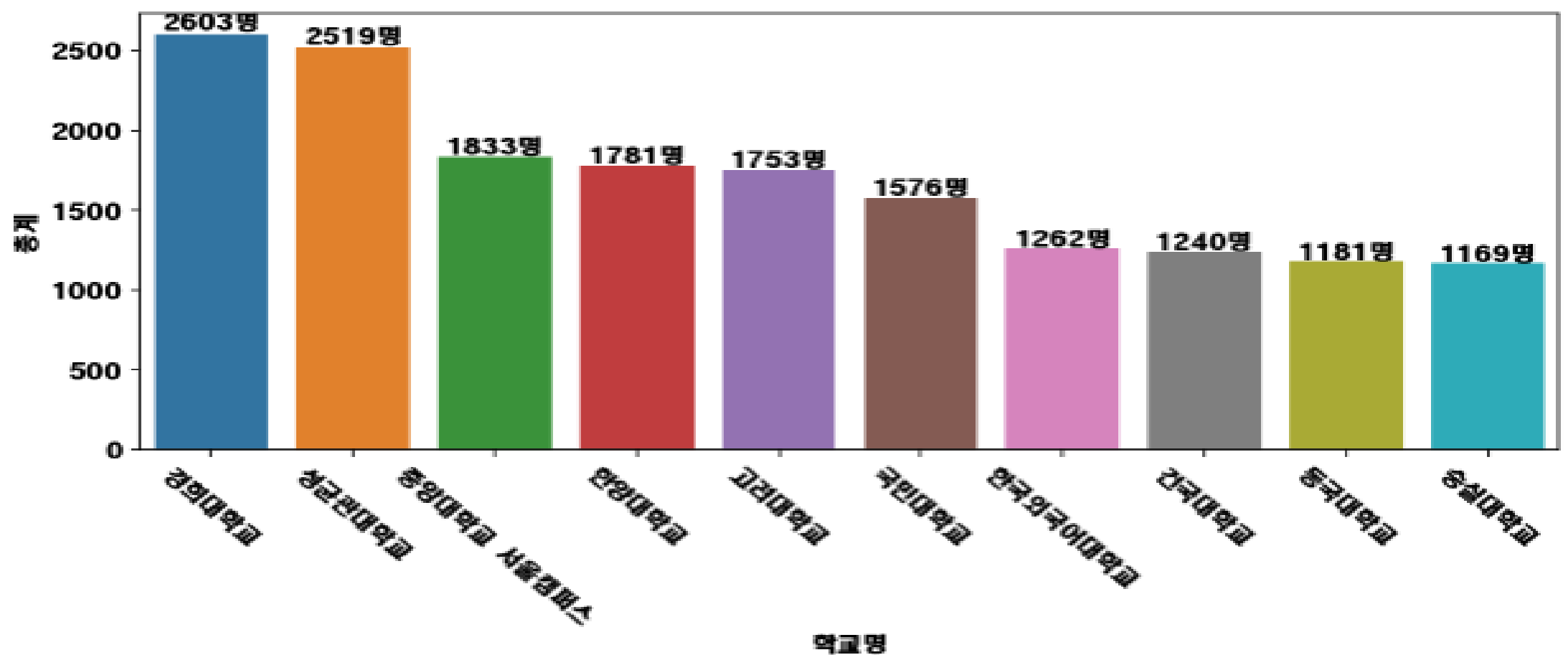
중국
유학생: 42.5%

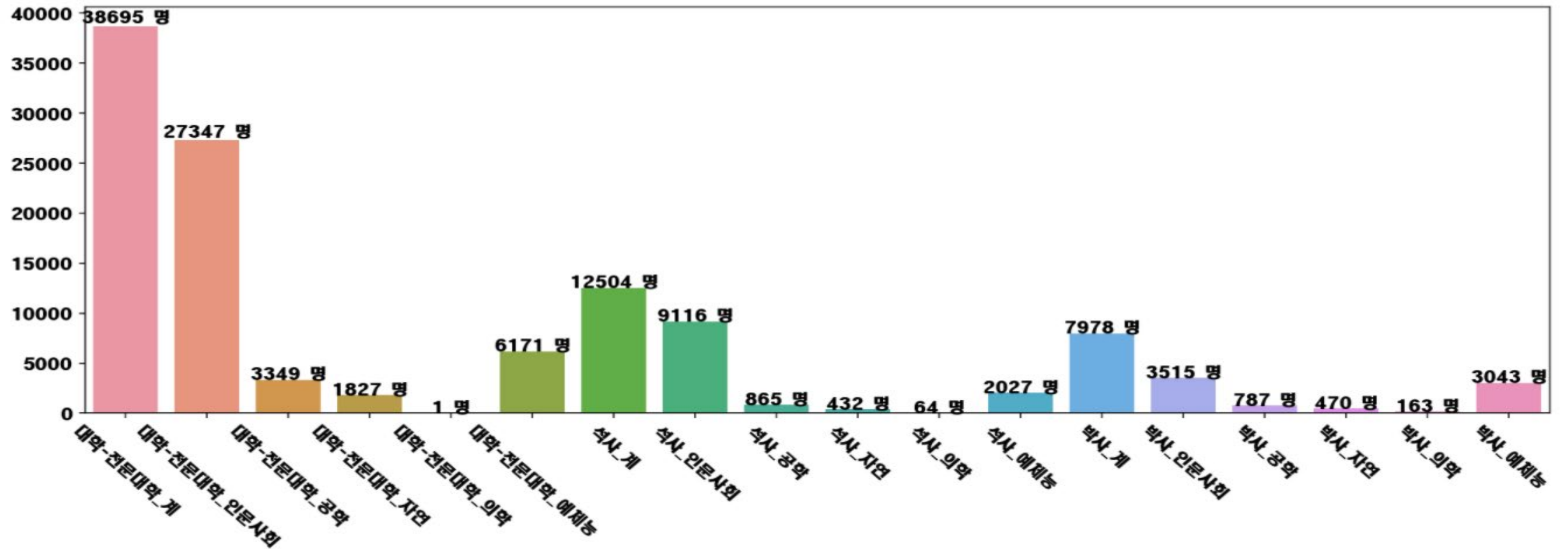


중국인 유학생 대학교



정원 순위
top10





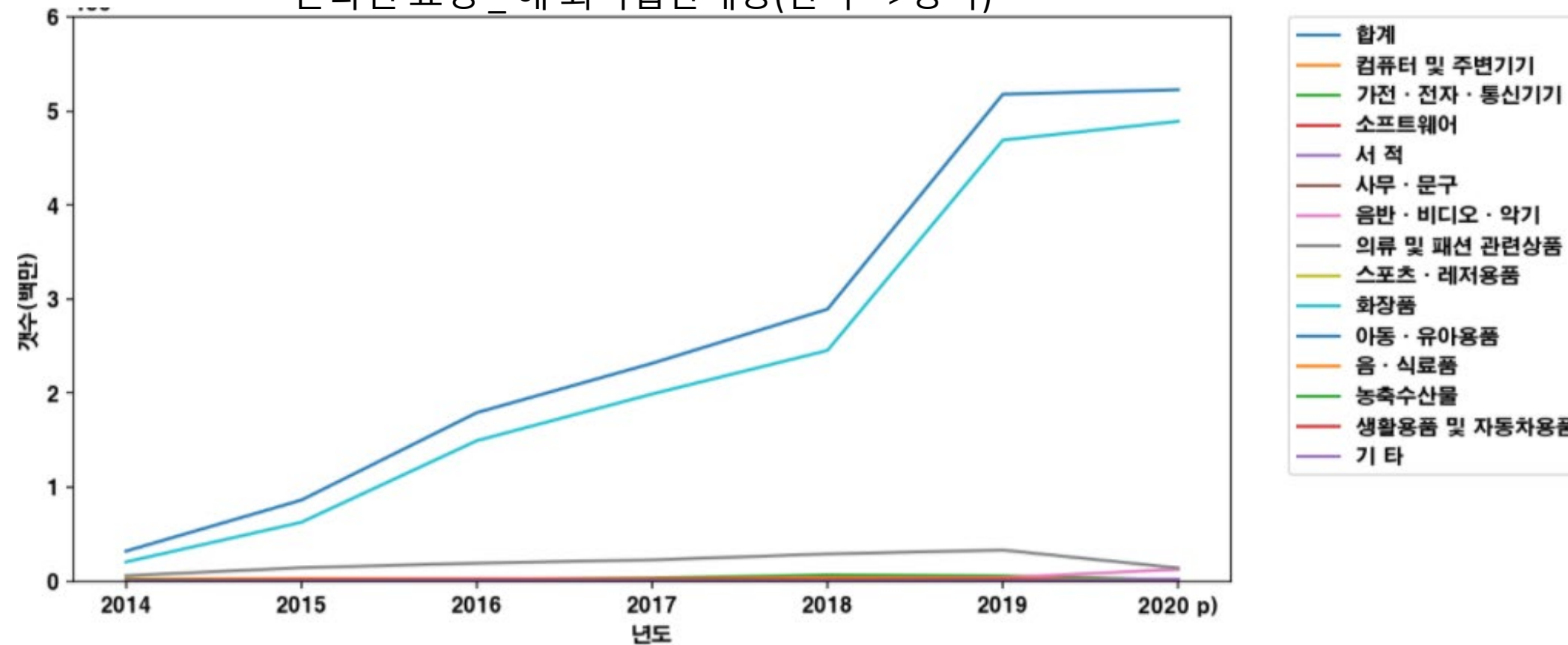
: 본론

대다수 학생이 인문사회 및 예체능

결론.

(유학생) 서포터즈

온라인 쇼핑 _ 해외직접판매량(한국 -> 중국)



2021

소셜 미디어 플랫폼을 이용한
중국 유학생 서포터즈

또한, 중국 화장품 소비시장은 온라인 채널이 가장 큰 비중을 차지하고 있어 소셜미디어 플랫폼을 적극적으로 활용하는 기업들의 성과가 두드러진다는 특징이 나타난다.

현지 언론, 웨이신(중국판KAKAO),더우인(중국판 틱톡), 웨이보(중국판 트위터)
등을통해 소비자에게 제품을 지속적으로 노출시켜 브랜드의 인지도 UP ↑

2 0 2 1 / 04/ 02

THANK
YOU

금 경 용

이름 금경용

연락처 long_3014@naver.com

Contact me