

대통령 선거 데이터 분석

10대, 20대, 30대는 어떻게 움직였는가?



TABLE OF CONTENTS

1 프로젝트 배경

2 데이터 소개

3 분석 결과

4 서비스 배포

5 프로젝트 한계

6 참고자료

I. 프로젝트 배경



랭킹순위

등수	아이디	점수
1	사라진초밥십인분	99,999
2	[REDACTED]	15,000
3	[REDACTED]	10,980
4	[REDACTED]	10,620
5	[REDACTED]	10,300
6	[REDACTED]	10,200
7	[REDACTED]	10,100
8	[REDACTED]	10,090
9	[REDACTED]	10,030
10	[REDACTED]	10,000



2022년 20대 대한민국 대통령 선거를 코앞에 두고도 누구를 차기 대통령으로 선택할지 모르겠다는 19세 최원서 씨와 응원하는 후보가 없다는 20세 구태경 씨는 모두 2030세대, 이른바 'MZ 세대'의 일원이다.

특히 이번 대선에서 '양강'인 이재명 더불어민주당 대선후보와 윤석열 국민의힘 대선 후보의 접전이 계속되면서 MZ 세대들이 승패를 좌우할 '캐스팅 보터'로 급부상했다.

10대와 20대를 주목했던 지난 선거



20대 대선 득표율 단위: %

48.6

당선



윤석열
국민의힘

47.8



이재명
더불어민주당

2.4

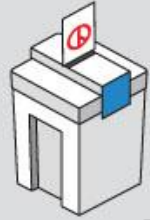


심상정
정의당

자료: 중앙선거관리위원회

그런데, **MZ세대**의 지지도는?

2022대선 성별·연령별 출구조사 결과



방송 3사 KBS · MBC · SBS

신뢰수준 95%, 표본오차 ±0.8%p

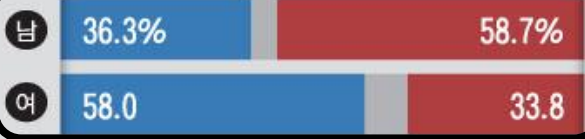


더불어민주당
이재명



국민의힘
윤석열

20대
이하



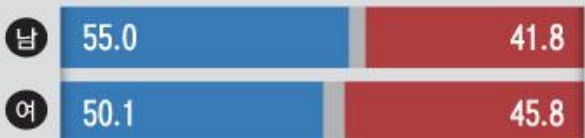
30대



40대



50대



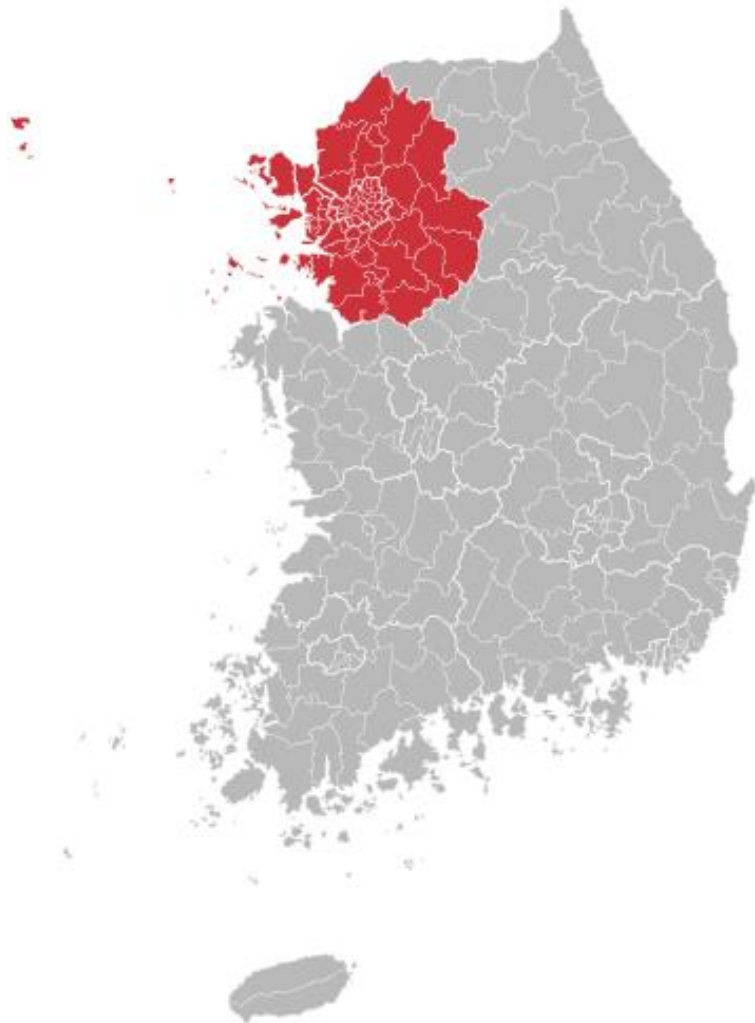
60대

이상



구분	이재명	윤석열
남	36.3	58.7
여	58.0	33.8

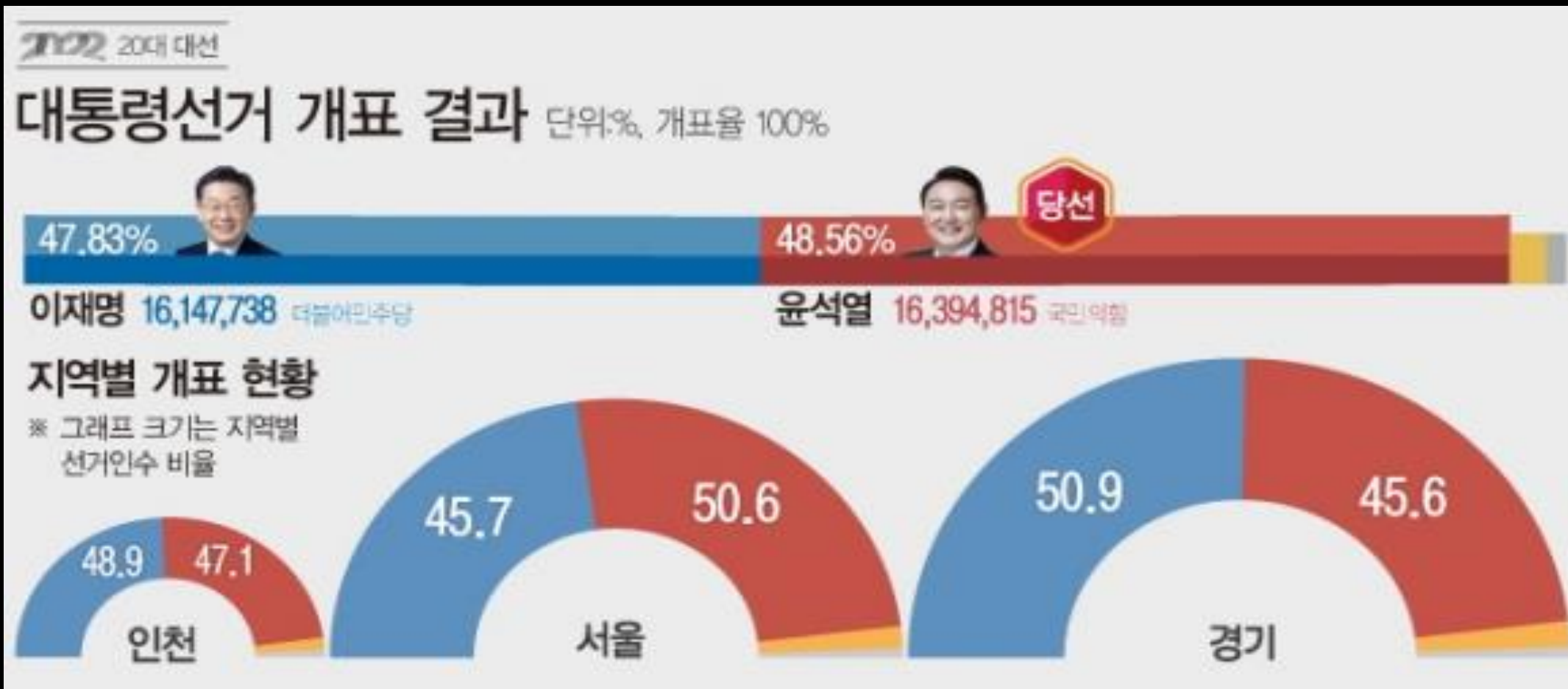
왜 서로 다를까?



시도	인구수	선거인수		
		계	남	여
전체 합계	51,629,720	44,168,510	21,892,189	22,276,321
서울	9,505,765	8,336,646	4,012,626	4324020
경기	13,572,415	11,428,857	5,730,410	5,698,447
인천	2,949,606	2,518,329	1,254,543	1,263,786
수도권 합계	26,027,786	22,283,832	10,997,579	11,286,253

전체 유권자의 **50.4%** 집중

같은 수도권, 그러나 다른 결과



수도권은 왜 다를까요?

분석 호기심이 생김

- ① 성별 ~ 연도별 지지도의 차이
- ② 수도권 지지도의 차이

현재는?

WHO ARE WE?

이름	연령	성별	팀내 역할	주요 수행 내용	특징
김미나	10대	여성	팀장	기획, 분석	투표 경험 무
권용석	40대	남성	지원	데이터 수집, 대시보드	반장님
주강희	20대	남성	지원	데이터 수집, 대시보드	코딩 힘들어요!!!

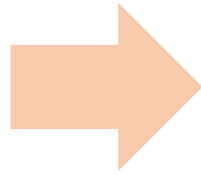
II. 데이터 소개

데이터 수집 및 활용 프로세스

 **Gallup**

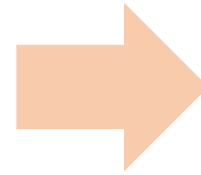
DATA 공공데이터포털
. GO . KR

< 데이터소스 >




python™

< 데이터가공 >




Streamlit

< 대시보드 개발 >

① 갤럽 정당 지지도

② 20대 대선 투표구별 개표 자료

정당 지지도는 모두 엑셀로 재정리함

Trustworthy

한국갤럽 데일리 오피니언

제545호(2023년 6월 1주)

내년 국회의원선거 결과 기대,
장래 정치 지도자 선호도(자유응답)

6월 2주 조사는 쉽니다. 제546호는 16일(금)에 공개합니다



갤럽 정당 지지도 주요 데이터 수집

◎ 정당 지지도: 2020년 1월 2째주 ~ 2022년 12월 3째주

▶ 날짜 : 연도, 월, 주

▶ 정당 : 국민의힘 / 더불어민주당 / 정의당 / 무당층

무당층 -> 기타 정당과 부동층을 포함

▶ 지역 : 서울, 인천/경기, 대전/세종/충청, 광주/전라, 대구/경북, 부산/울산

▶ 성별 : 남자, 여자

▶ 연령 : 10~20대

XLSX 중앙선거관리위원회_제20대 대통령선거 투표구별 개표자료

다운로드

제20대 대통령선거 개표 결과에 대한 정보로, 전국 시도, 구시군, 읍면동, 투표구별 제20대 대통령선거의 개표 결과 데이터를 조회하실 수 있습니다.

0

0

관심

파일데이터 정보

메타데이터 다운로드

데이터 개선요청

오류신고 및 문의

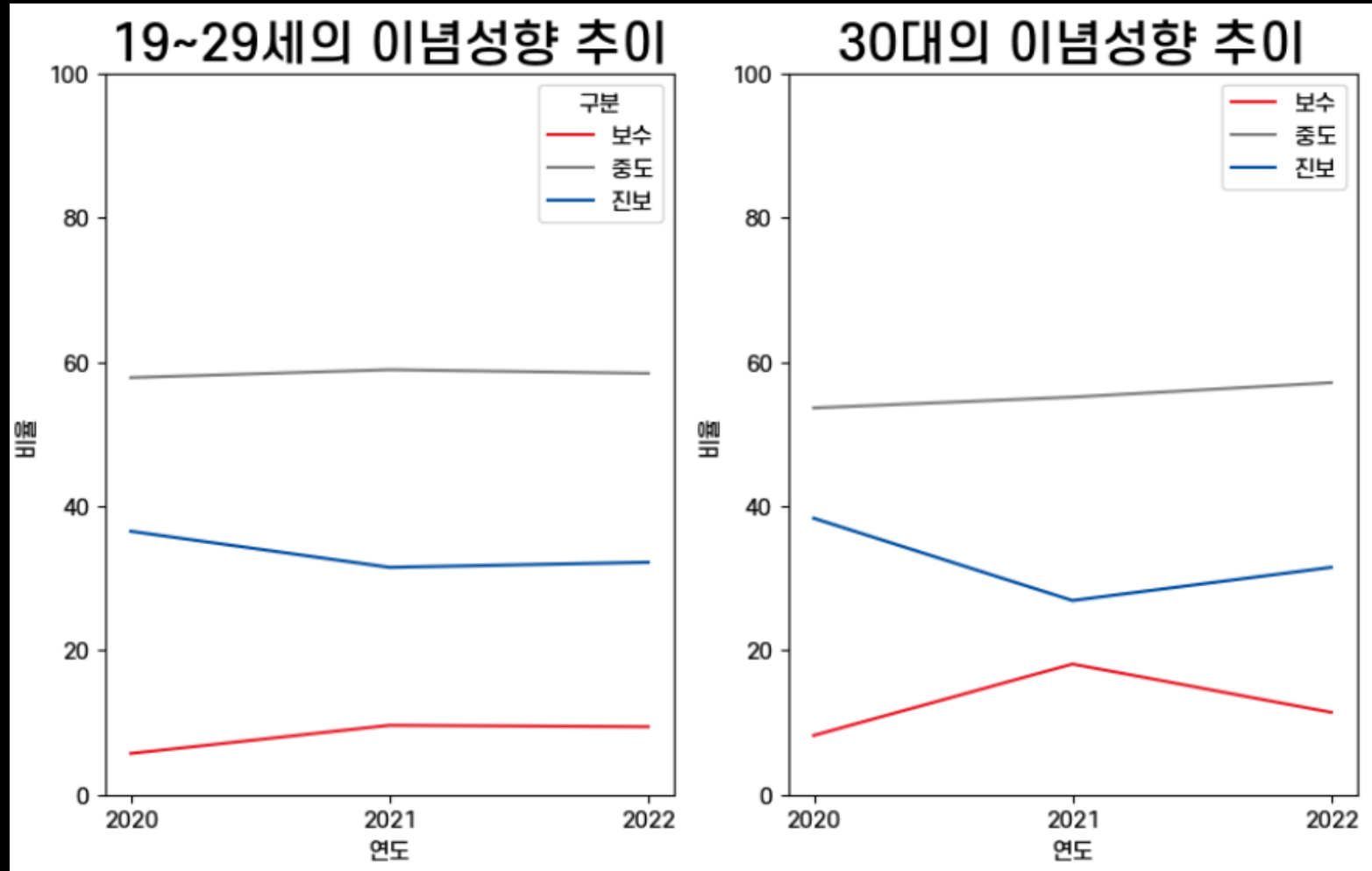
파일데이터명	중앙선거관리위원회_제20대 대통령선거 투표구별 개표자료_20220309		
분류체계	일반공공행정 - 일반행정	제공기관	중앙선거관리위원회
관리부서명	정보화운영과	관리부서 전화번호	02-3294-1153
보유근거	공직선거법	수집방법	선거정보시스템을 통한 선거사무과정의 데이터 수집
업데이트 주기	수시 (1회성 데이터)	차기 등록 예정일	
매체유형	텍스트	전체 행	22753
확장자	XLSX	키워드	대통령선거,선거,개표
데이터 한계	개표데이터 특성상 한계 없음	다운로드(바로가기)	350

III. 분석 결과

분석 기간

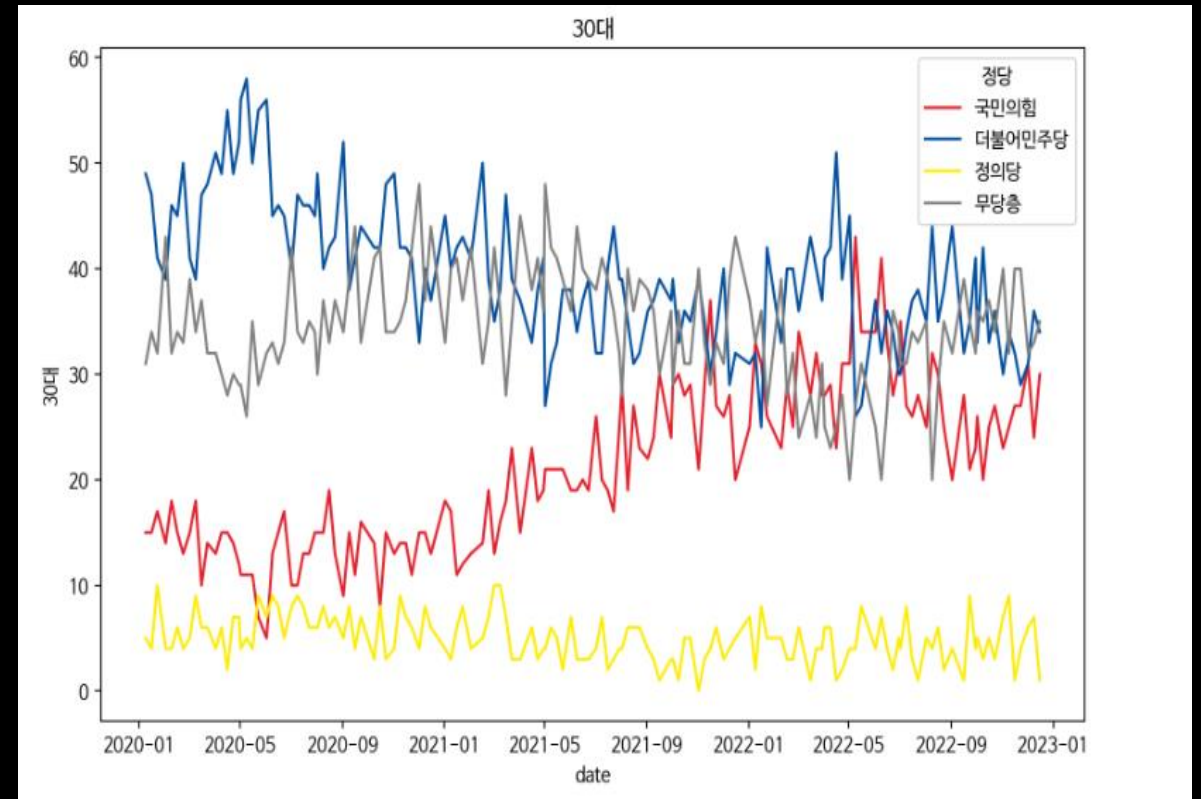
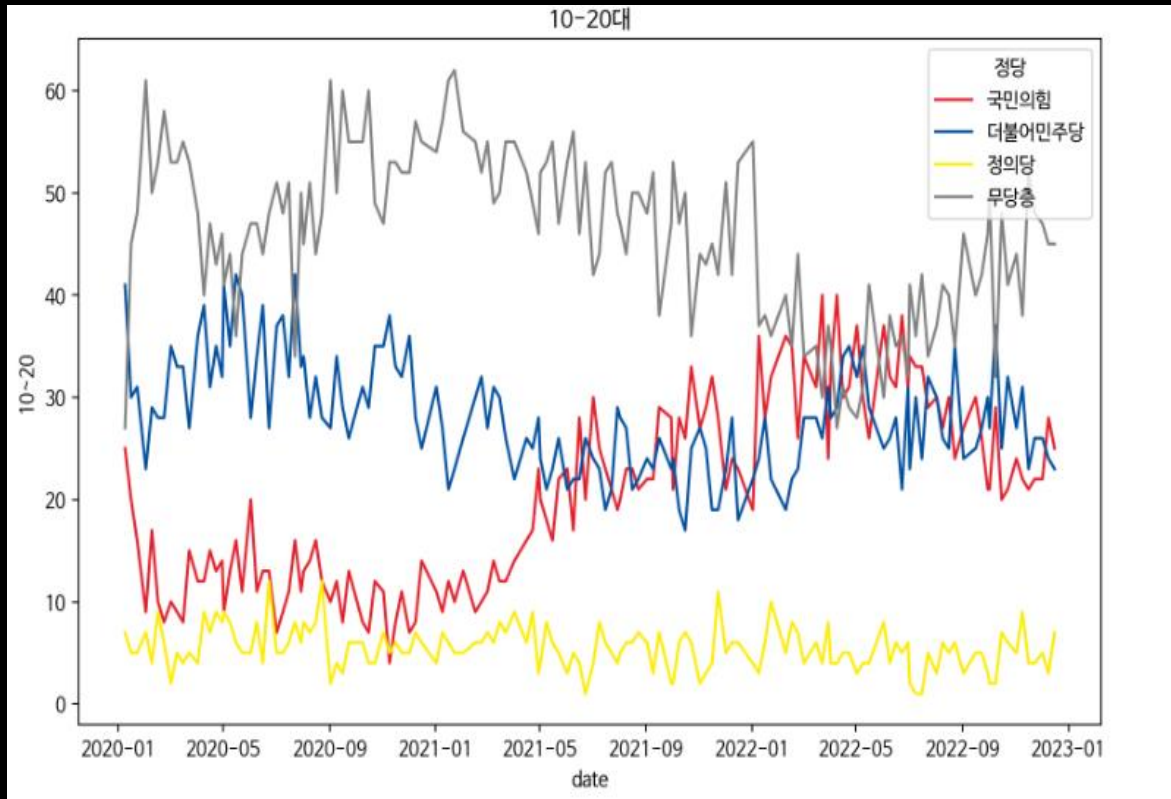
2020.1 – 2022.12

진보의 약세화 ➡ 중도세의 확장, 보수의 대 약진



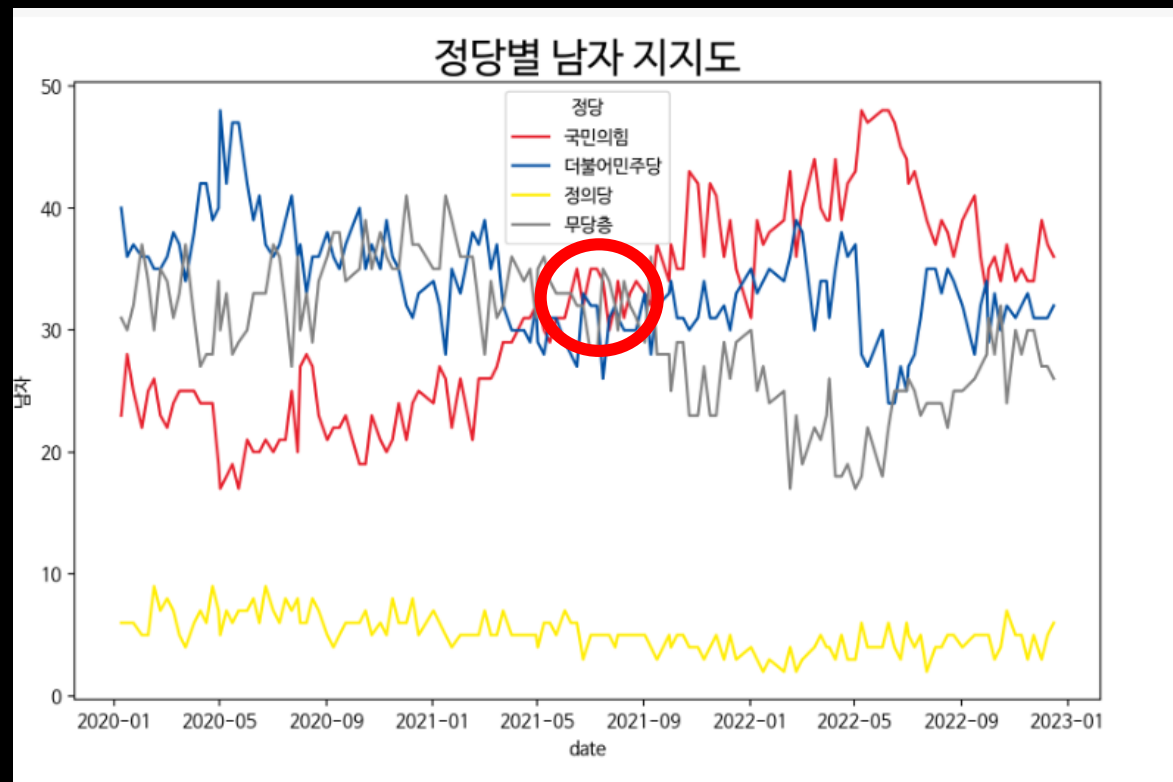
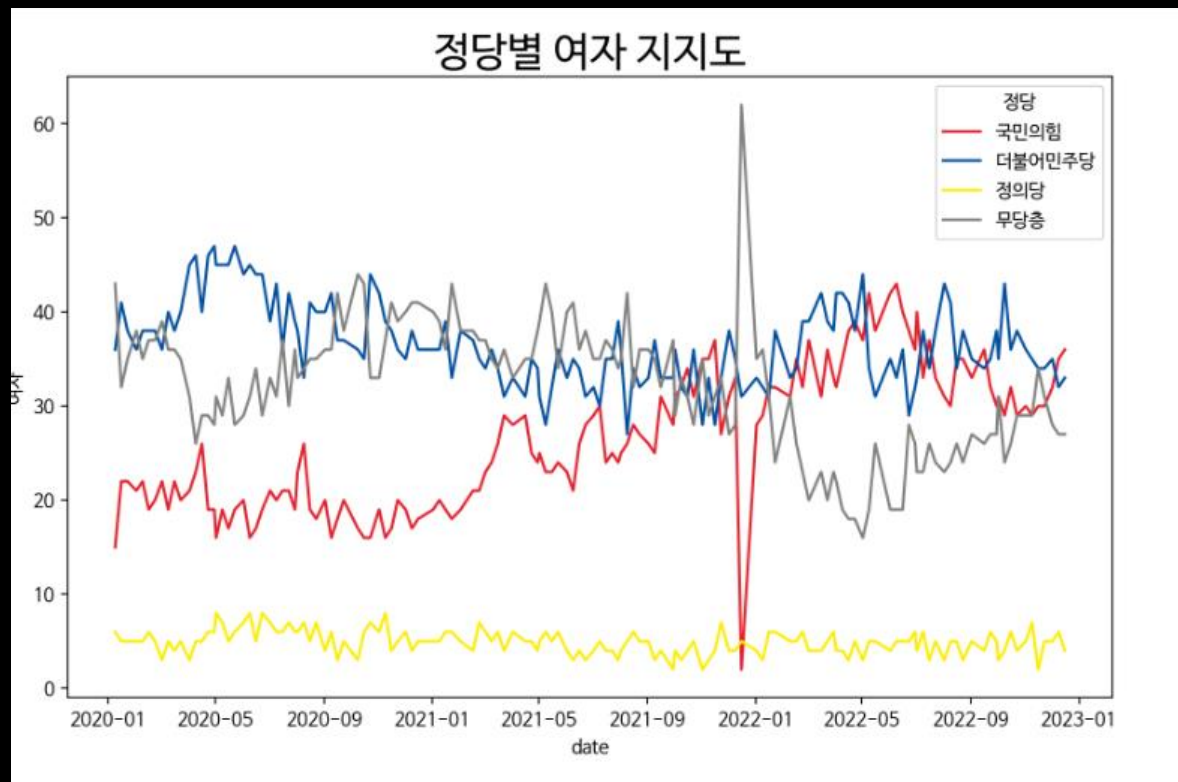
20대, 30대 더불어 민주당의 지지율 약세가 두드러짐

30대의 지지율 하락이 20대보다 훨씬 크다는 걸 확인



전반적으로는 국민의힘 대약진, 더불어민주당 약세

<남자의 지지율 변동이 2021년 1-3월 기준 매우 컸음>

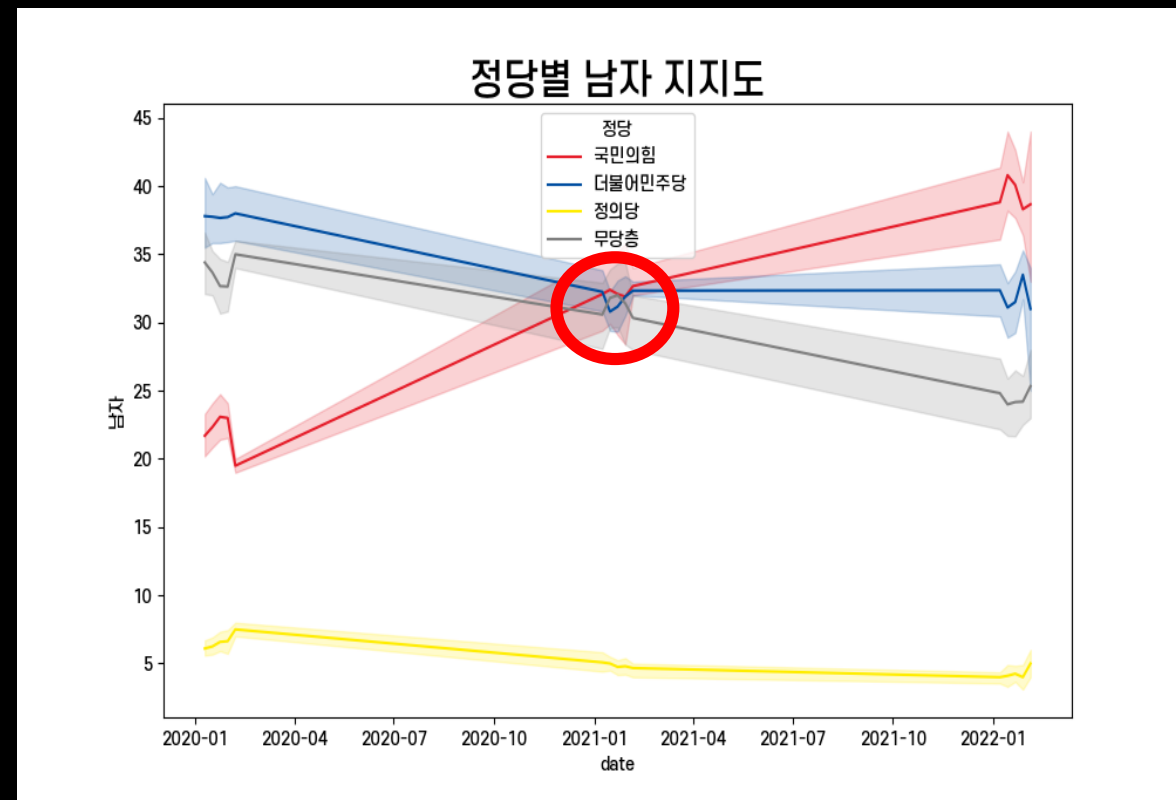
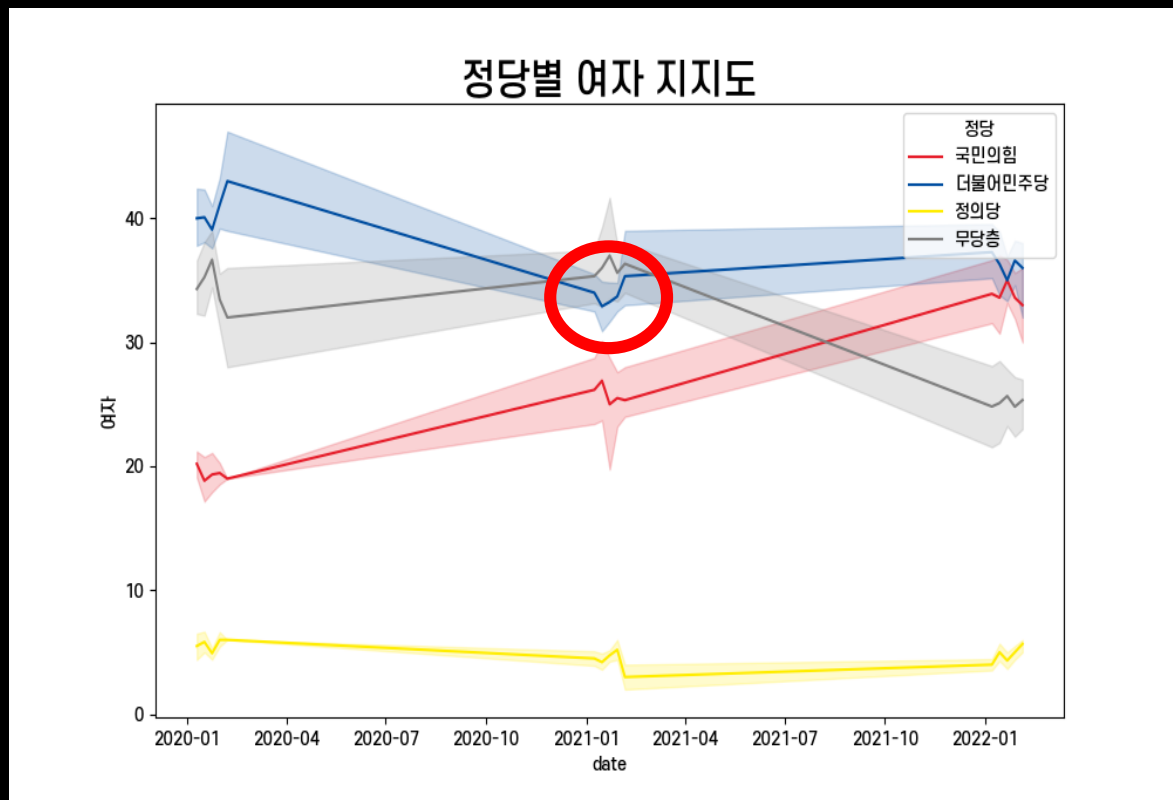


① 2020년 말 코로나 대 유행

② 2021년 3월 2일 LH 직원의 투기

③ 2021년 4월 7일 재보궐선거

남자가 여자보다 정부 정책에 더 빠르고 민감하게 반응 전반적으로는 모두 더불어민주당의 지지를 철회하는 시기



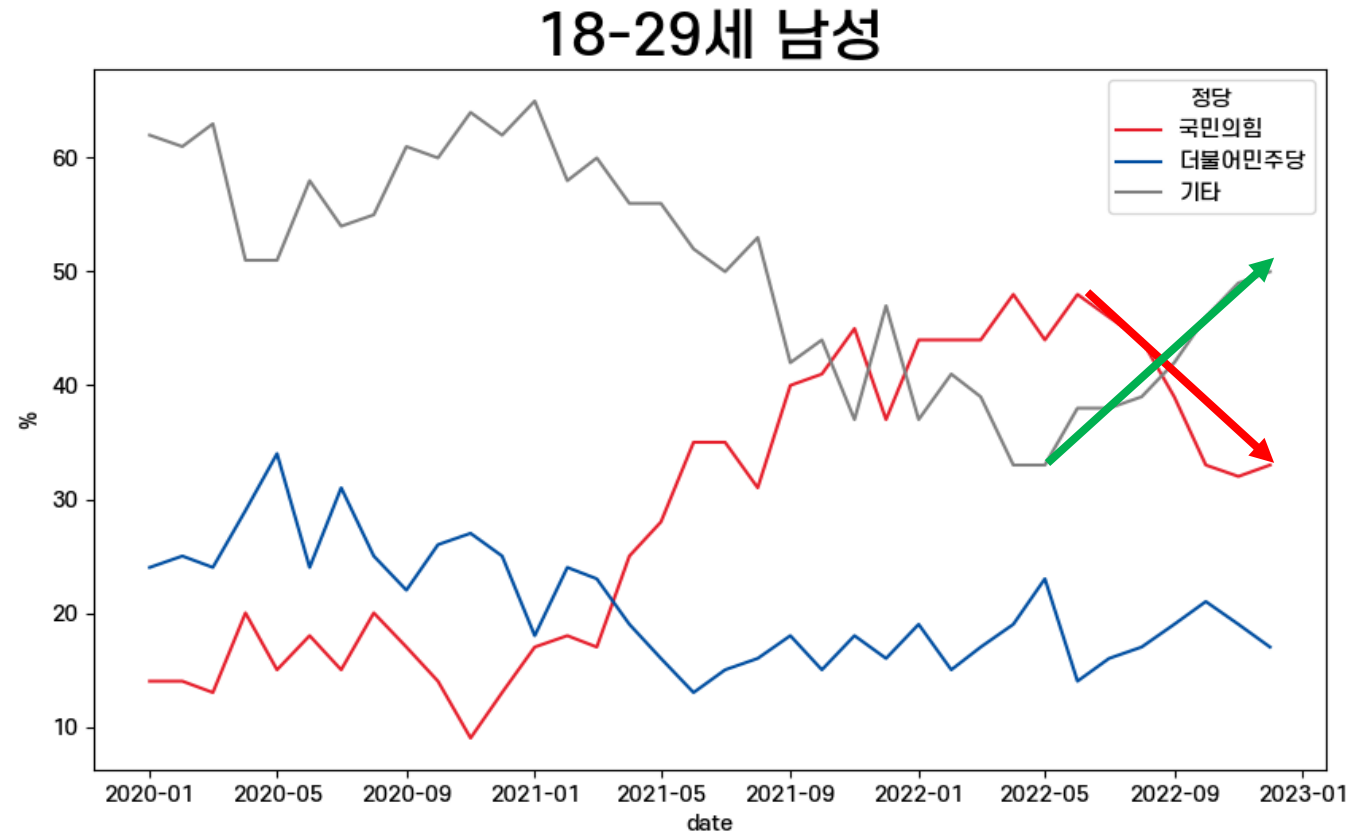
여성은 더불어민주당을 지지하고,
남성은 국민의힘을 지지한다?



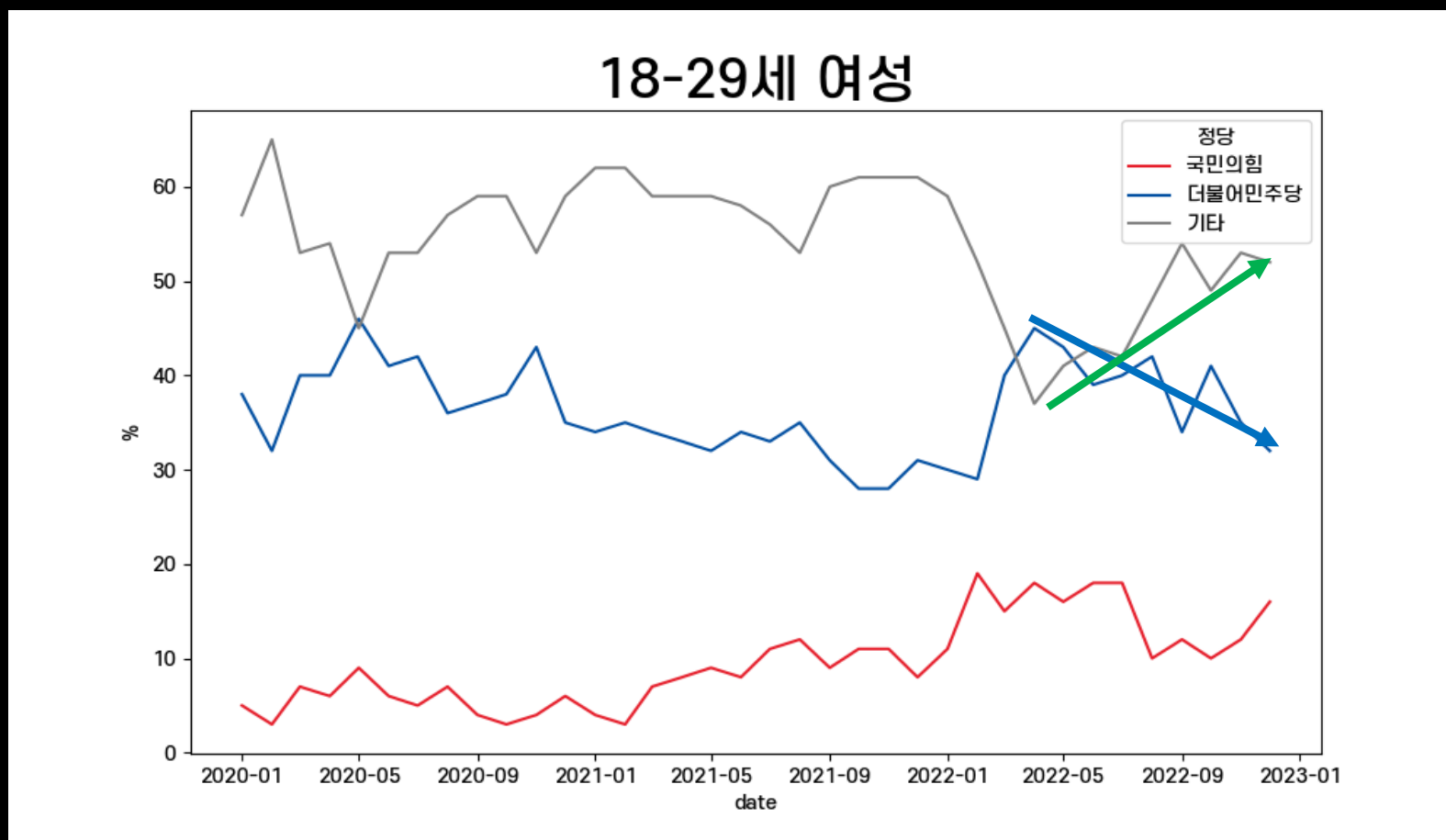
NO

꾸준하게 당시 여당에 대한 **불만**이 누적되고 있었음

18-29세 남성의 국민의힘 지지도는 2022.6.1 지방선거 이후, 지지를 철회하고 기타(무당층)로 바뀔

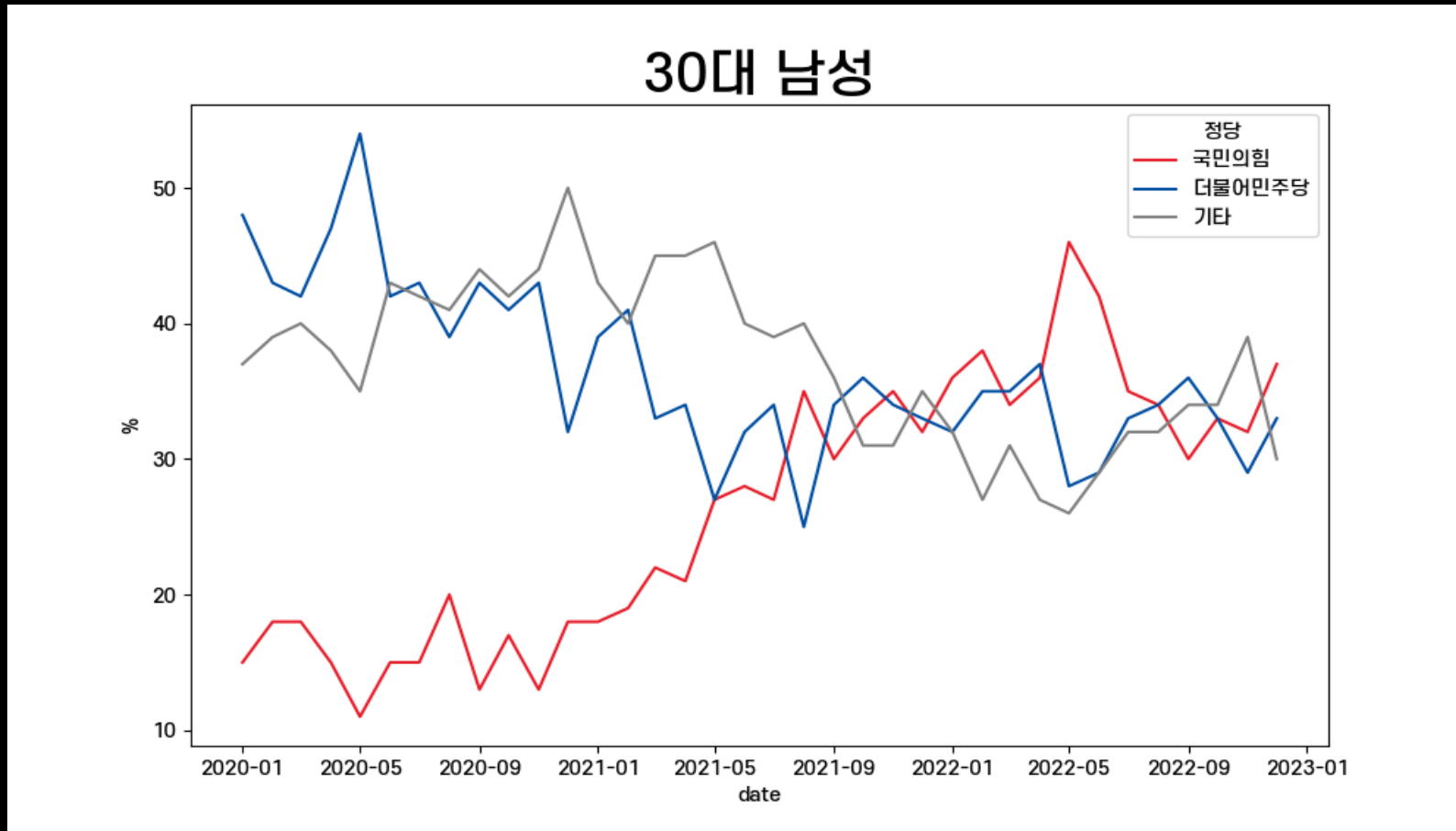


18-29세 여성의 경우도 무당층이 증가하는 중

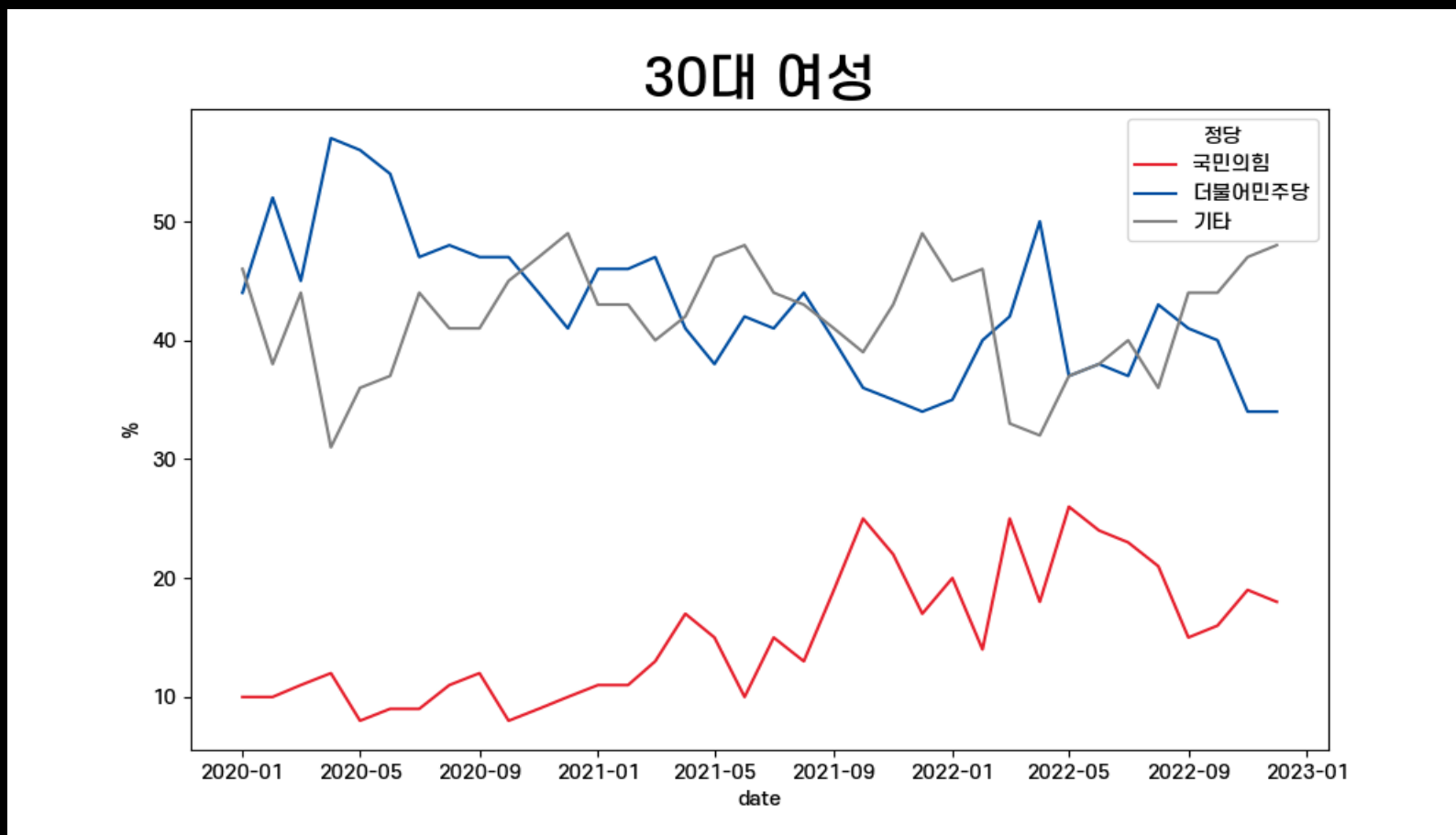


민심에 가장 민감하게 반응하는 세대로 확인

- 2020-2021, 더불어민주당 지지
- 2021-2022, 국민의힘 지지
- 2022-현재, 혼돈



30대 여성의 경우는 무당층이 계속 증가하는 중



여성은 더불어민주당을 지지하고,
남성은 국민의힘을 지지한다?

NO

모두 무당층이 늘어가고 있음 = 현 정치에 **불만** 많음
30대 남성이 가장 빠르게 정치 이슈에 반응함

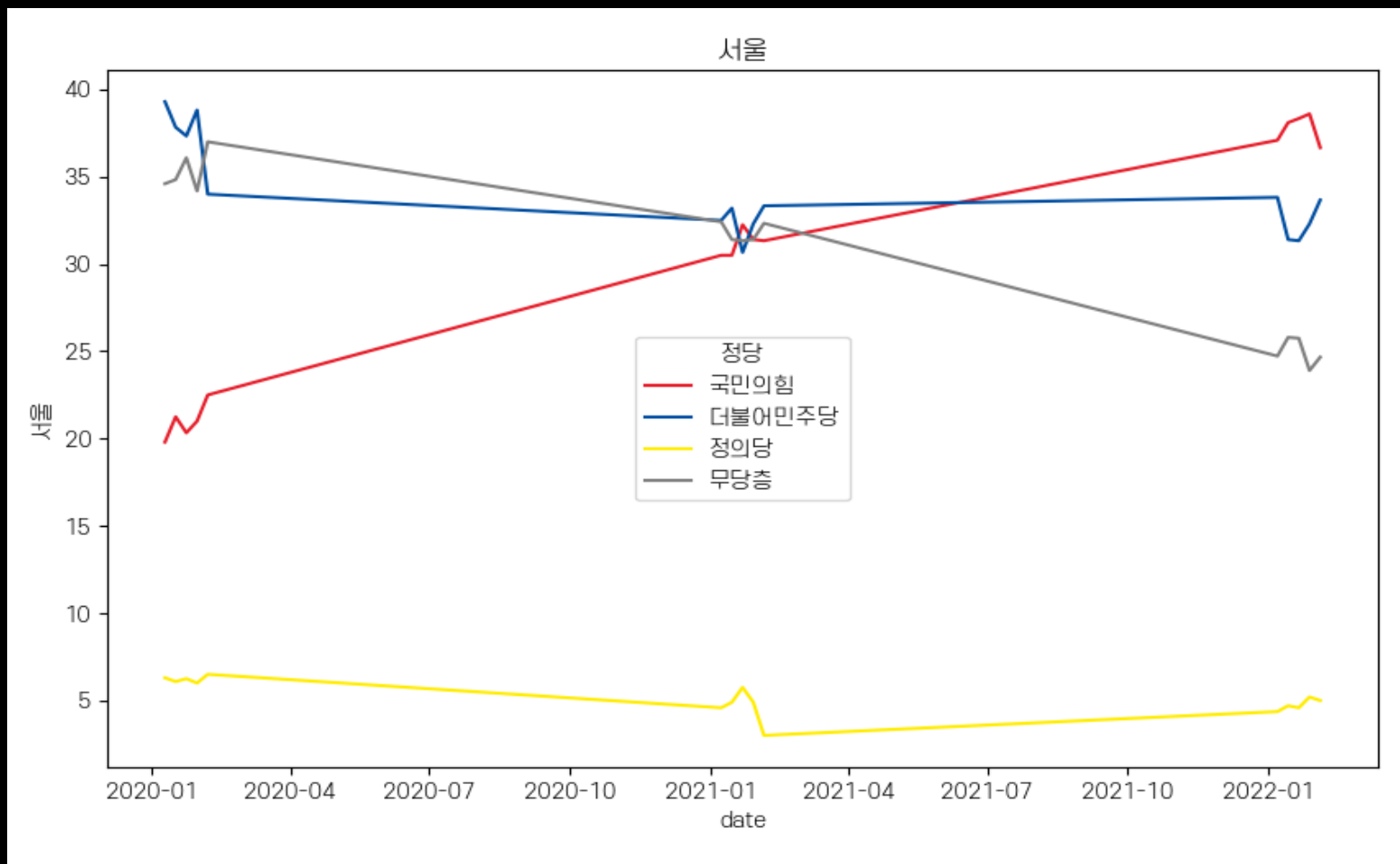
10대 여성인 제가 볼 때는

30대 남자가 핵심 타겟 아닐까요?

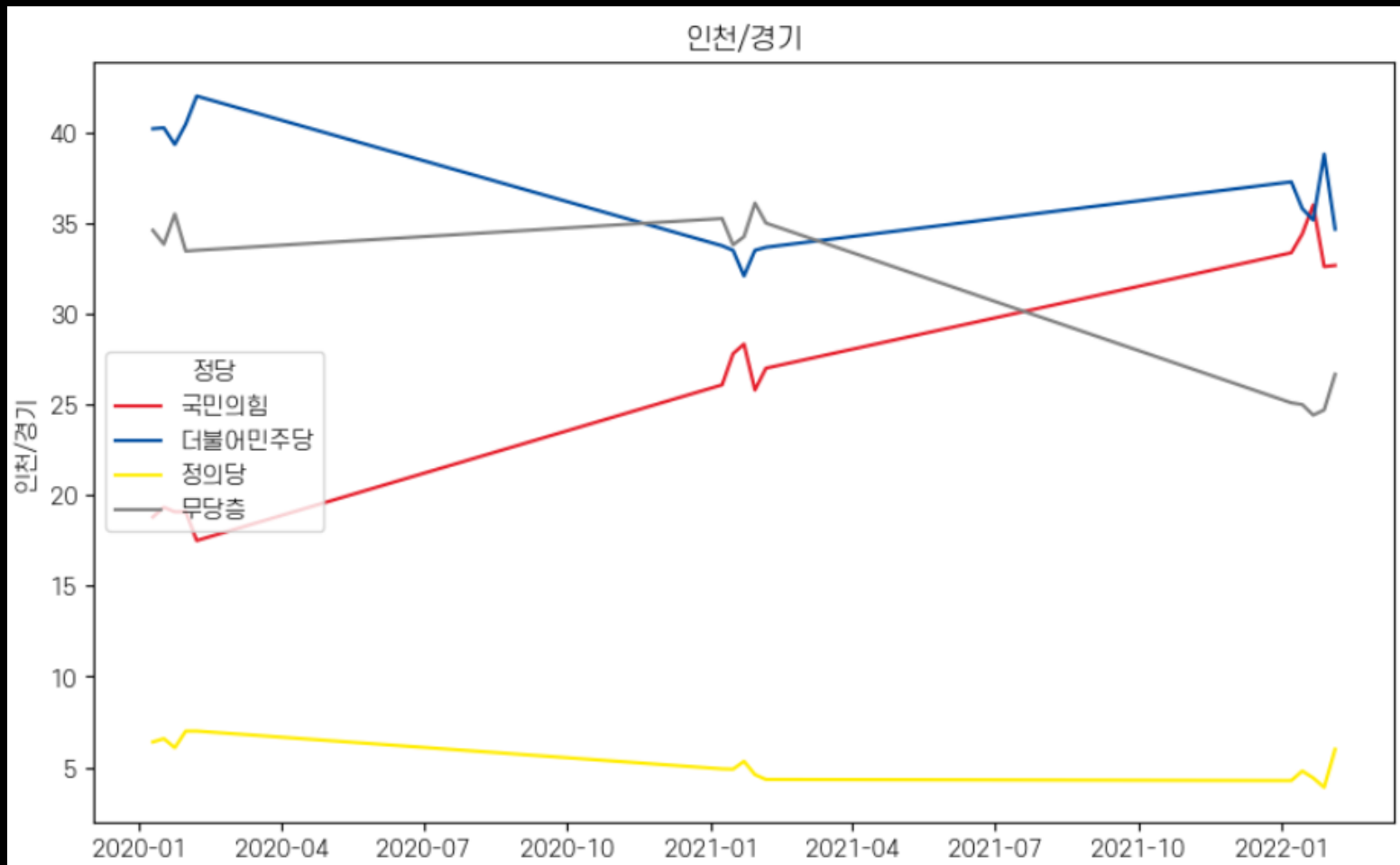
(변동성이 크고, 이슈에 민감하게 반응)

우리나라 선거의 핵심은 **수도권**

줄어든 무당층 ➡ 국민의힘 지지



줄어든 무당층 ➡ 국민의힘 지지



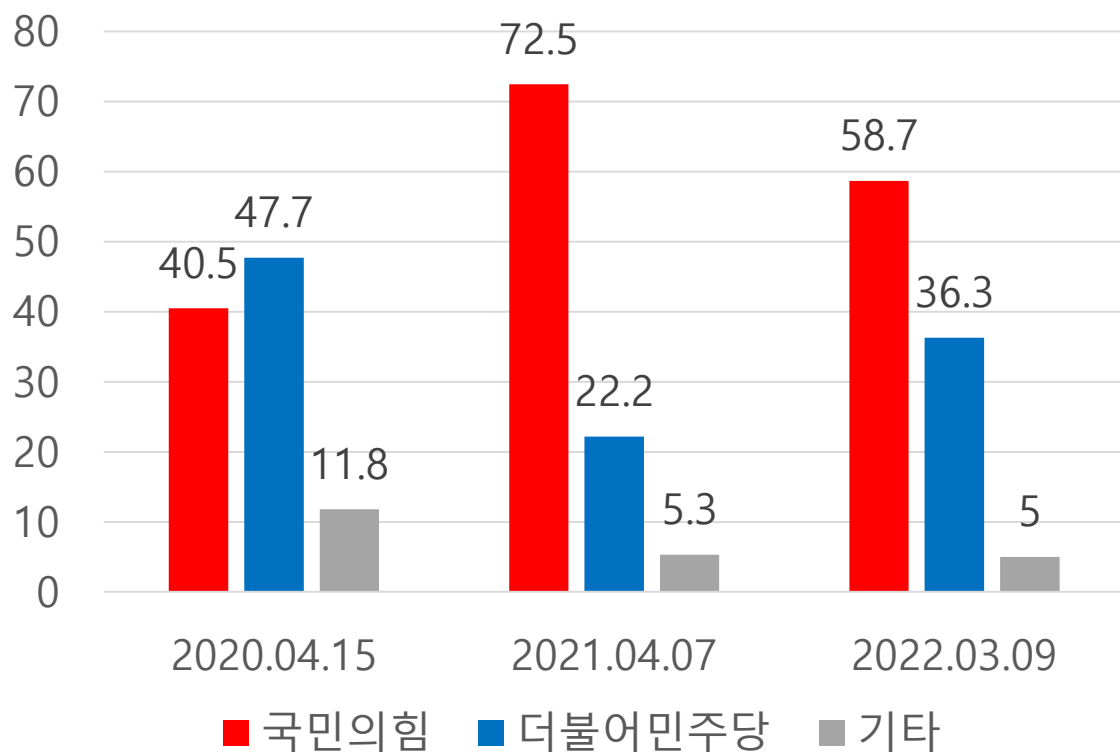
- ✓ 더불어민주당 지지도는 크게 줄지 않음
- ✓ 무당층이 국민의힘으로 대거 이동함

각 정당에 대한 지지도는 투표로

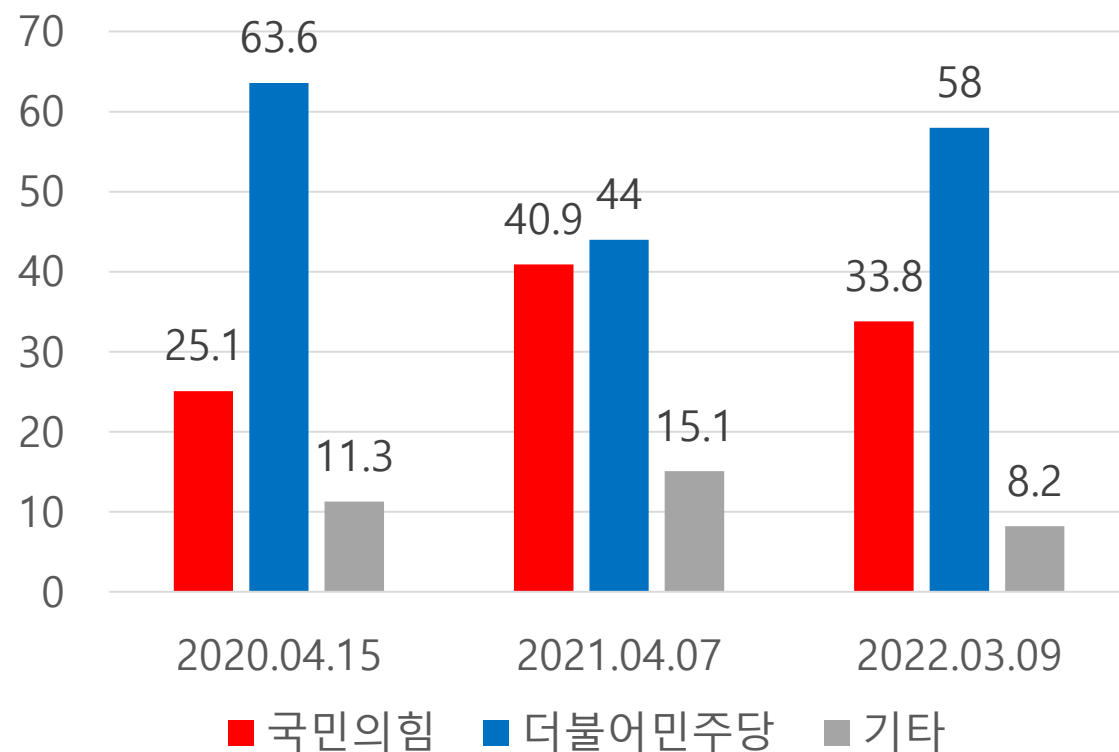
- ✓ 2020.4월 총선 (전국)
- ✓ 2021.4월 재보궐 선거(서울)
- ✓ 2022.3월 대통령선거 (전국)

사람들의 마음은 언제든지 바뀐다

10-20대 남성 정당 지지도 출구조사

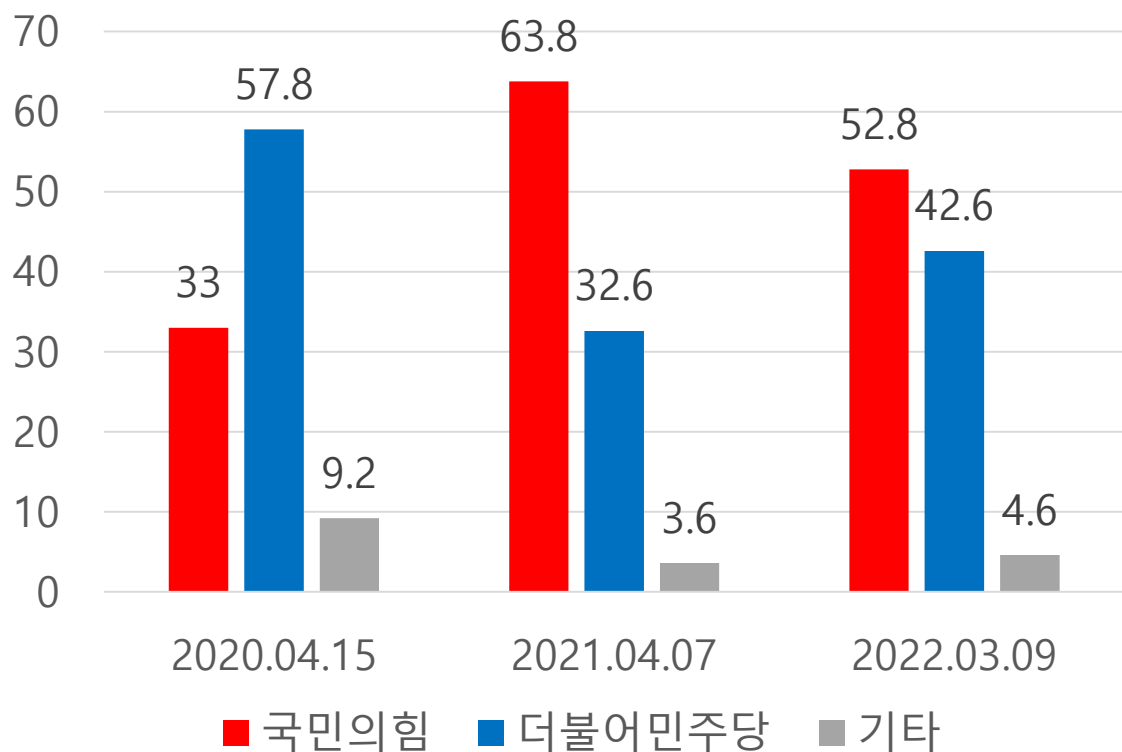


10-20대 여성 정당 지지도 출구조사

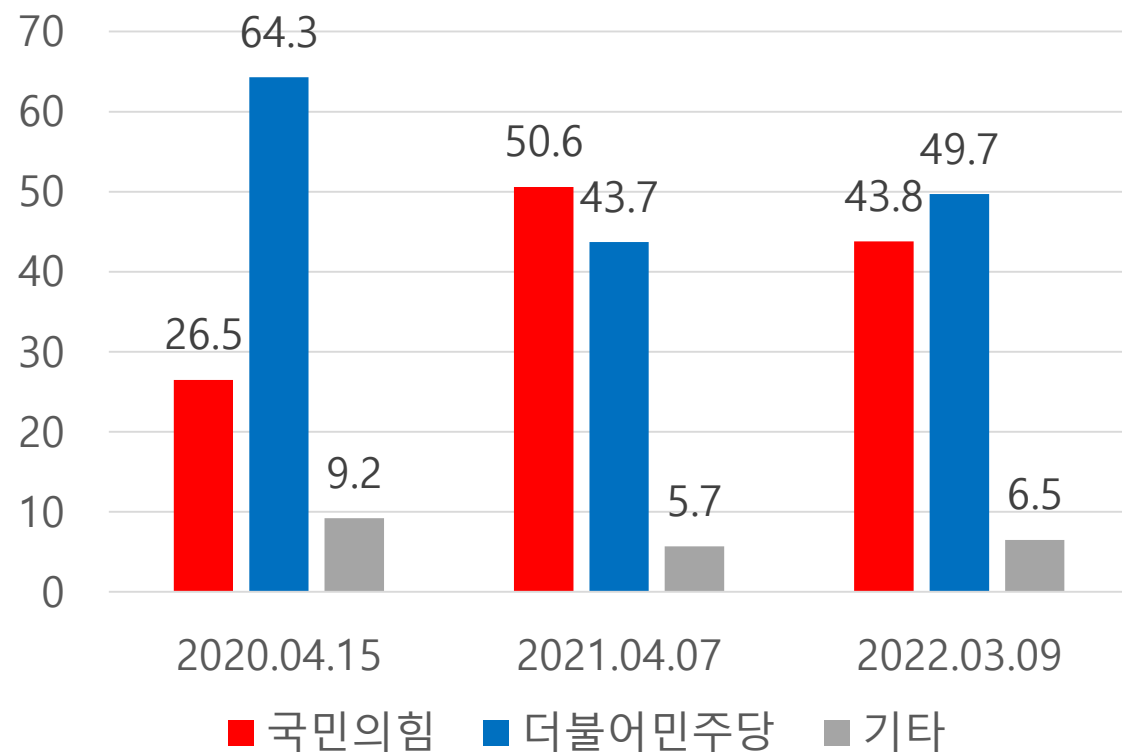


사람들의 마음은 언제든지 바뀐다

30대 남성 정당 지지도 출구조사



30대 여성 정당 지지도 출구조사



현재는?

MZ는 고정적이지 않다!

항상 움직인다!

2024년 총선 **MZ**는 어디로?

IV. 서비스 배포

- 대시보드 소개
- 측면 구성 요소
 - 홈 : 전반적인 프로젝트 소개
 - 갤럽 여론 조사 : 데이터 / 정당 지지도 추이(연령별/기간별)
 - 20대 대통령선거 : 지도 관련 데이터 확인
대선 득표율 데이터
후보자의 시도별 득표율

🖥️ 대시보드 메뉴

🏠 홈

📊 갤럽여론조사

📈 20대 대통령선거

홈 Workflow

대통령 선거 데이터 분석

10 - 30대는 어떻게 움직였는가?



Member	Skills	GitHub & Blog
김미나	분석 & 기획	https://github.com/180pp
권용석	분석 & 대시보드	https://github.com/MaestroYongseok
주강희	대시보드 & PPT	https://blog.naver.com/wnrkdgml10



갤럽여론조사

데이터 확인

	세대	연도	구분	비율
0	19~29세	2020	보수	5.7
1	19~29세	2020	중도	57.8
2	19~29세	2020	진보	36.5
3	19~29세	2021	보수	9.6
4	19~29세	2021	중도	58.9
5	19~29세	2021	진보	31.5
6	19~29세	2022	보수	9.4
7	19~29세	2022	중도	58.4
8	19~29세	2022	진보	32.2
9	30~39세	2020	보수	8.2

세대 : 연령

연도 : 2020~2022

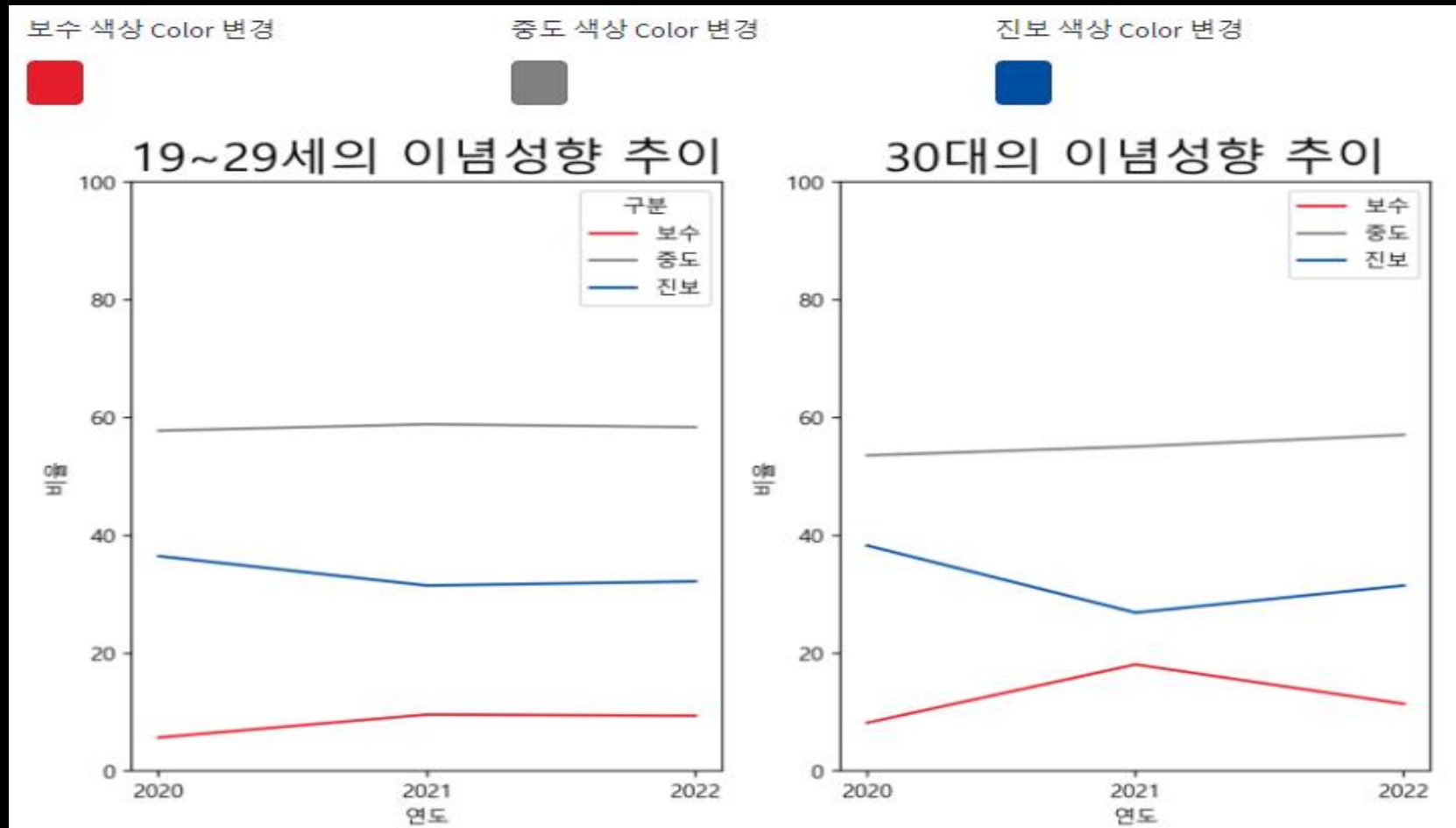
구분 : 보수(국민의힘)

진보(더불어민주당)

중도(무당층)

비율 : 지지율

20대와 30대의 비율 차이



정당 지지도 추이

	연도	월	주	정당	전체	서울	인천/경기	대전/세종/충청	광주/전라	대구/경북	부산/울산/경남	남자	여자	10~20	30대	40대	50대	60대이상	date
0	2,020	1	2	국민의힘	23	27	20	22	4	32	32	23	15	25	15	16	29	37	2020-01-09 00:00:00
1	2,020	1	2	더불어민주당	40	36	42	46	64	27	33	40	36	41	49	50	38	33	2020-01-09 00:00:00
2	2,020	1	2	정의당	6	6	5	4	11	4	8	6	6	7	5	11	6	5	2020-01-09 00:00:00
3	2,020	1	2	무당층	31	31	33	28	21	37	27	31	43	27	31	23	27	25	2020-01-09 00:00:00
4	2,020	1	3	국민의힘	25	20	24	27	5	39	30	28	22	20	15	19	28	35	2020-01-16 00:00:00
5	2,020	1	3	더불어민주당	39	41	38	41	66	21	32	36	41	30	47	50	39	31	2020-01-16 00:00:00
6	2,020	1	3	정의당	5	6	7	2	7	4	5	6	5	5	4	8	8	3	2020-01-16 00:00:00
7	2,020	1	3	무당층	31	33	31	30	22	36	33	30	32	45	34	23	25	31	2020-01-16 00:00:00
8	2,020	1	4	국민의힘	24	23	21	23	5	35	33	25	22	16	17	15	26	37	2020-01-23 00:00:00
9	2,020	1	4	더불어민주당	38	38	39	41	61	21	30	37	38	31	41	45	39	24	2020-01-23 00:00:00
10	2,020	1	4	정의당	6	6	7	4	7	3	5	6	5	5	10	9	5	2	2020-01-23 00:00:00
11	2,020	1	4	무당층	32	33	33	32	27	41	32	32	35	48	32	31	30	37	2020-01-23 00:00:00
12	2,020	2	1	국민의힘	22	22	17	19	3	36	29	22	21	9	14	13	24	38	2020-02-02 00:00:00
13	2,020	2	1	더불어민주당	36	34	38	30	58	19	34	36	36	23	39	52	36	31	2020-02-02 00:00:00
14	2,020	2	1	정의당	5	6	6	4	8	3	4	5	5	7	4	5	7	4	2020-02-02 00:00:00
15	2,020	2	1	무당층	37	38	39	47	31	42	33	37	38	61	43	30	33	27	2020-02-02 00:00:00
16	2,020	2	2	국민의힘	24	21	20	26	4	42	32	25	22	17	18	15	24	39	2020-02-09 00:00:00
17	2,020	2	2	더불어민주당	37	34	41	37	66	18	26	36	38	29	46	51	35	28	2020-02-09 00:00:00
18	2,020	2	2	정의당	5	8	4	5	5	2	7	5	5	4	4	6	8	4	2020-02-09 00:00:00
19	2,020	2	2	무당층	34	37	35	32	25	38	35	34	35	50	32	28	33	29	2020-02-09 00:00:00

갤럽 여론조사

기간

2020/06/01

2023/06/21

시작
날짜

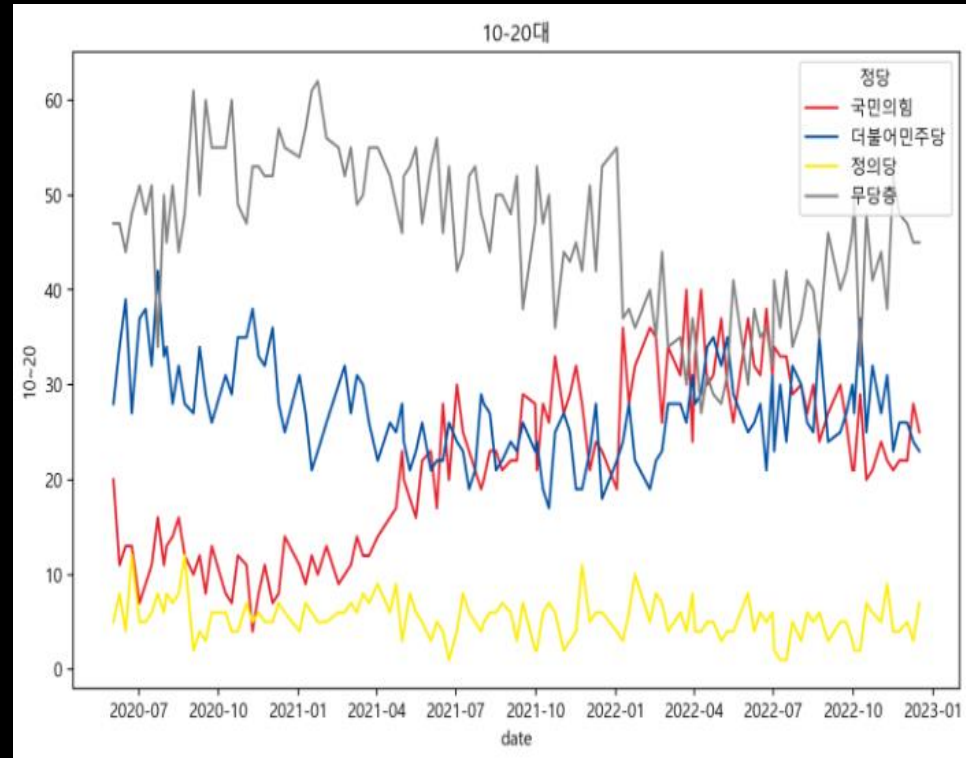
마지막
날짜

요소별

10~20대

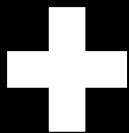
요소 선택

< June 2020 >						
Su	Mo	Tu	We	Th	Fr	Sa
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				



지도관련 데이터 확인

	CTPRVN_CD	CTP_ENG_NM	CTP_KOR_NM	geometry
0	11	Seoul	서울특별시	POLYGON ((966987.226 1941110.946, 961574.146 1...
1	26	Busan	부산광역시	MULTIPOLYGON (((1138209.365 1673609.179, 11381...
2	27	Daegu	대구광역시	POLYGON ((1107819.047 1781088.551, 1112278.678...
3	28	Incheon	인천광역시	MULTIPOLYGON (((909743.413 1946023.117, 909736...
4	29	Gwangju	광주광역시	POLYGON ((932712.687 1696168.692, 935873.826 1...



대선 득표율 데이터 확인

	sido_id	시명	선거인수	투표수	이재명	윤석열	이재명_득표율	윤석열_득표율
0	11	서울	8,346,647	6,501,831	2,944,981	3,255,747	45.2946	50.0743
1	26	부산	2,921,510	2,200,224	831,896	1,270,072	37.8096	57.7247
2	27	대구	2,046,714	1,611,512	345,045	1,199,888	21.4113	74.4573
3	28	인천	2,519,225	1,883,504	913,320	878,560	48.4905	46.645
4	29	광주	1,209,206	985,492	830,058	124,511	84.2278	12.6344

20대 대통령선거

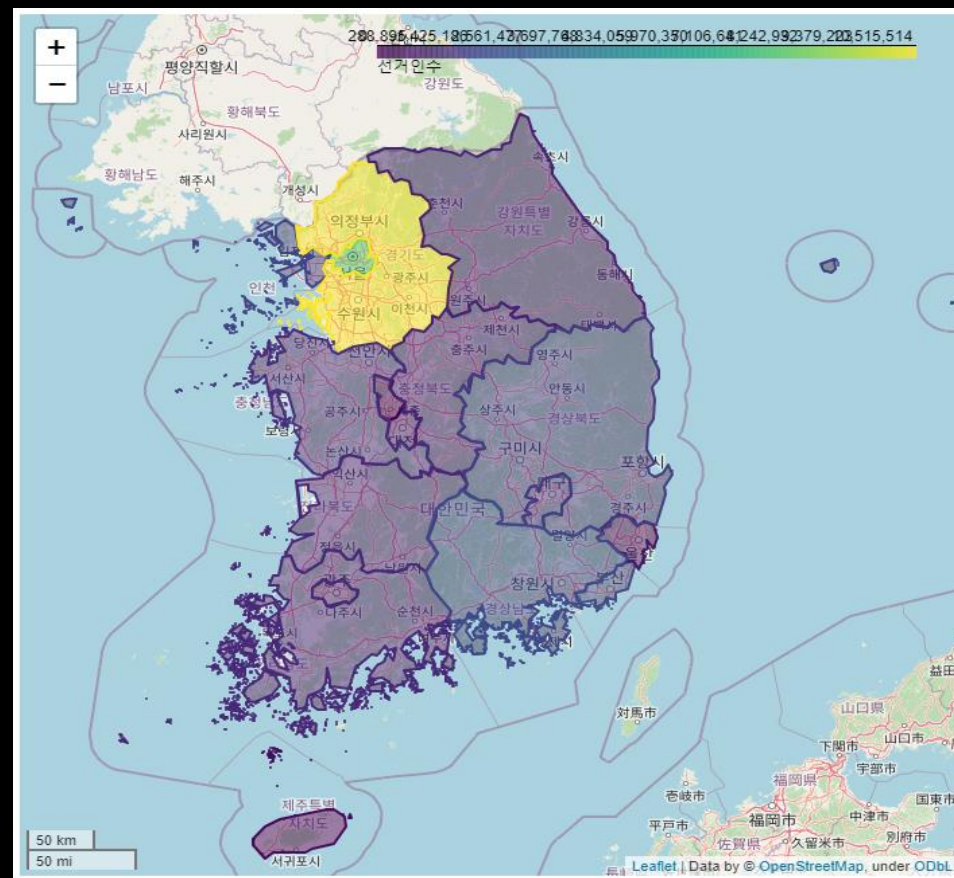
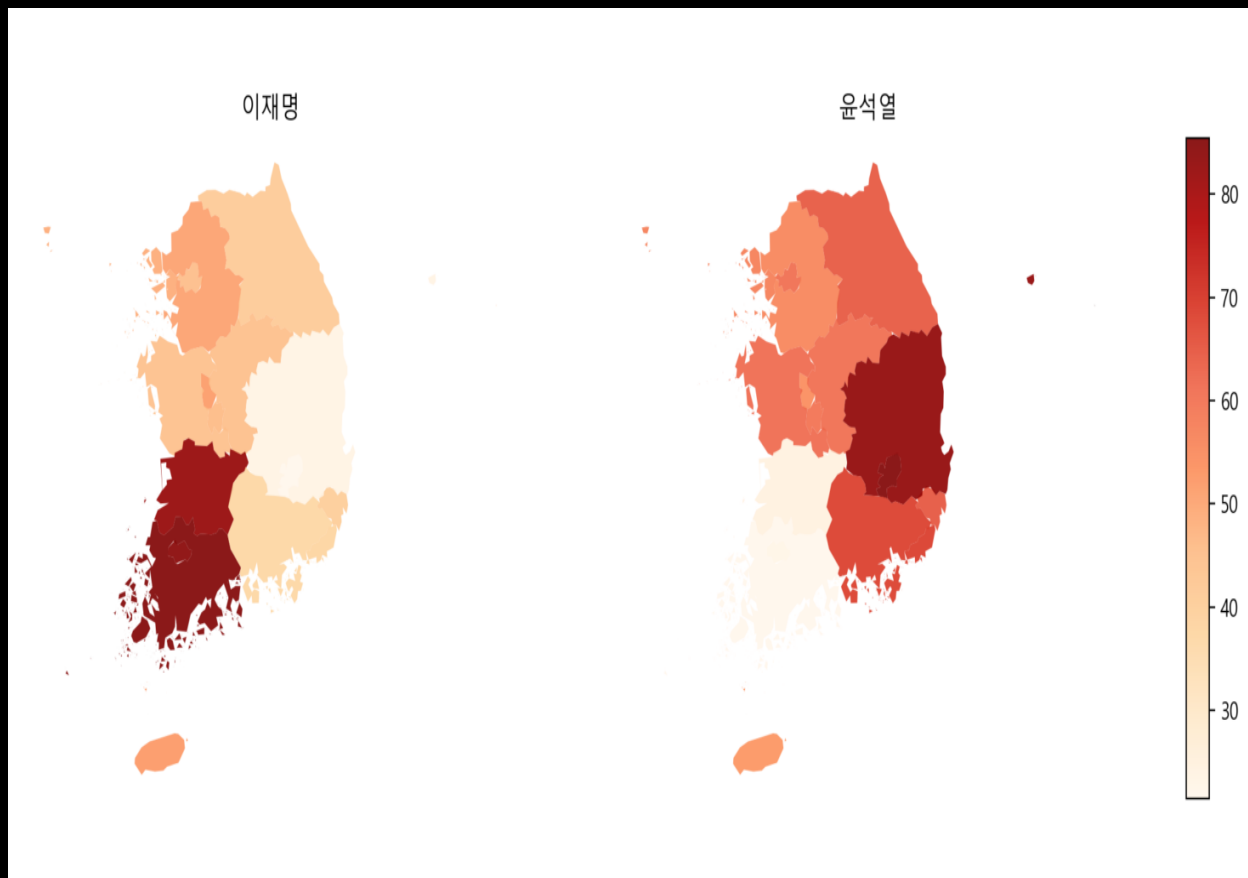
GeoDataFrame과 DataFrame 병합

	sidoid	geometry	시명	선거인수	투표수	이재명	윤석열	이재명_득표율	윤석열_득표율
0	11	POLYGON ((966987.226 1941110.946, 961574.146 1...	서울	8346647	6501831	2944981	3255747	45.294641	50.074310
1	26	MULTIPOLYGON (((1138209.365 1673609.179, 11381...	부산	2921510	2200224	831896	1270072	37.809605	57.724668
2	27	POLYGON ((1107819.047 1781088.551, 1112278.678...	대구	2046714	1611512	345045	1199888	21.411258	74.457280
3	28	MULTIPOLYGON (((909743.413 1946023.117, 909736...	인천	2519225	1883504	913320	878560	48.490473	46.644977
4	29	POLYGON ((932712.687 1696168.692, 935873.826 1...	광주	1209206	985492	830058	124511	84.227777	12.634400

20대 대통령선거

후보자의 시도별 득표율

선거인단



V. 프로젝트 한계

이번 프로젝트를 통해서 처음 정치를 접해 기본 지식이 아예 없어서 어려웠다

빨강색이 어디 정당이고
파란색이 어디 정당인데 . . ?



여당 / 야당 / 무당??
이게 다 어딘데 . .

- 기본 지식도 없다 보니 초반에 방향을 잡기 어려웠음
- 뒤늦게 지식을 익히는 탓에 더 심도 깊은 분석을 하지 못했음

VI. 참고 자료

논문

1) 사전 태도 및 투표 의향과 실제 투표행동간 상관성 연구
A Study on Correlation of Voting Behavior and Attitude and Vote Intention in the Poll Survey
조사연구(2011. 03)

2) 빅데이터로 선거 결과를 예측할 수 있을까?
조재근 경성대 수학과응용통계학부 교수
(2018년 05월호 칼럼)

3) 제20대 대통령선거 분석
< 국회입법조사처 >
2022. 8. 26.

기사

https://www.hani.co.kr/arti/politics/politics_general/1034593.html
20대 대선...절묘한 민심, 모두가 패자

<https://fleishmanhillard.co.kr/2022/03/20%EB%8C%80-%EB%8C%80%EC%84%A0%EC%97%90%EC%84%9C-%EC%9D%BD%EC%96%B4-%EB%B3%B8-%EA%B5%AD%EB%AF%BC%EC%9D%98-%EB%A7%88%EC%9D%8C/>
20대 대선에서 읽어 본 국민의 마음

https://www.hani.co.kr/arti/politics/politics_general/1019522.html
역대급 '비호감' 대선, 주목해야 할 관전 포인트 3가지