



[DACON]제 20대 대선 후보 정책·공약 시각화 경진대회

🕒 Created	@2022년 3월 12일 오후 8:28
🏷️ Tags	

팀원

- 김가인
- 송혜민
- 안지은

규칙

- 회의 시간 늦지 않기
- 회의주기 : 일주일
- 매주 토요일 밤 9시
- 비정기회의 : 필요할 때마다

회의록

🔥 03.12(토) 회의

🔥 03.19(토) 회의

🌵 03.27(토)

👯 4월 3일

💥 4월 6일

👤 4월 9일

4월 10일

일정표

Aa 이름	📅 날짜	☰ 태그
회의	@2022년 3월 19일	
회의	@2022년 3월 26일	
🔥 <u>세미해커톤..(10am~~)</u>	@2022년 4월 9일	

할 일(~3/26)

1. 정책 크롤링
2. 자연어처리(빈도수 ~~~) 후 워드클라우드

- 정당 전체로 볼 때
- 공약에 써있는 [분야] 별로 세분화해서 워드클라우드

>> 분야별로 세분화에 있어서 정책목록 팝업창 소제목별로 워클 해오기

>> 스트레드시트에 불용어 추가하기 (전체 / 분야별 불용어 이런 것도 하면서 슬랙에서 상의)

>> 3/26까지는 무조건 완료!!!

>>> 다음주까지 무조건 데이터 정제는 완료해야 됨. 그래야 시각화에 집중 가능.

3. 개표 데이터
4. 인구 데이터 (성별, 연령)
5. 시각화 툴 / 정교화 툴 해오기

*생각해와야 되는 거 - 시각화 툴, 자연어처리 정교화 (툴??) : 대신 빨리 할 수 있게...

RNR

- 더불어민주당 : 안지은
- 국민의 힘 : 송혜민

- 정의 : 김가인
- 개표데이터 정제 : 송혜민 - json 관련
- 인구데이터 정제 : 김가인 -
- 시각화 툴 해오기 (네트워크 쪽으로 최대한!) : 안지은

*개표데이터, 인구데이터 지역 범위 : 서울시 00구, 경기도 —시, 강원도 —군

Task 업로드

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/ff864197-919f-44e3-b842-dd7cea9fa61c/정당별_공약_크롤링.ipynb

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/81089e0f-4834-4adc-8977-37e9a0351767/공약_전체_크롤링.ipynb

https://docs.google.com/spreadsheets/d/1Ty3Uyo7PdmjICnQ3B75e_eXuxbRux713kgIeKfgh-4/edit?usp=sharing → 불용어사전

▼ 인구데이터

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/468145ac-18a8-43ba-acc4-cda8e5513a90/전국_시군구_인구데이터.ipynb

[https://s3-us-west-2.amazonaws.com/secure.notion-static.com/7e0155af-4f0f-4585-8b54-ee065f955143/202202_202202_전국청년층\(18-39\)인구현황_월간.csv](https://s3-us-west-2.amazonaws.com/secure.notion-static.com/7e0155af-4f0f-4585-8b54-ee065f955143/202202_202202_전국청년층(18-39)인구현황_월간.csv)

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/348515ac-39f3-41eb-bd5f-a236cb14b5a9/202202_202202_전국연령별유소년층인구현황_월간.csv

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/0f63ee88-624c-4528-af3c-b708791c055f/202202_202202_전국연령별인구현황_10세단위_월간.csv

Idea

1. 연령별-성별별 공약 분류해보기! - 세대/성별별 관심사와 부합하는 정책 매칭하기