



4월 6일

최종 파이썬 노트북 제작: 4/9 다같이 모여서 코딩하기 늦더라도 중간에 들어와서 같이하기! (10AM~~) 중간중간 진행사항 보고하기

오늘 정할 일

- 카테고리 정리하기
 - 외교/안보/국방 → 대외정책
 - IT분야 명칭 변경 → 과학기술
 - 금융정책 → 주거 정책 (재개발, 청년주택, ~~주택, 부동산세) & 금융정책 (주식관련, 금리, 대출)
 - 개헌/개혁 → 정치/사법/행정
-

TASK (~ 4/8)

- 목표/재원조달방안 삭제 → 이행방법에 있는 내용만! & 카테고리 수정(~4/7) @안지은 @hyemin song @김가인
- 주제분류별 엑셀 파일 명사 토큰화 엑셀파일 제작 (~4/7 밤까지 전달 or 4/8 오전 중으로)@안지은
- 지금까지 만든 크롤링, wordcloud, 네트워크 정리 파이썬 노트북 생성(~4/8) @안지은
- 주제 분류 엑셀 바탕으로 네트워크 시각화(~4/8) @안지은
- 주제분류 엑셀 바탕으로 각 주제분야별로 최다빈도 단어 시각화(바그래프) @hyemin song @김가인

- 후보별로 '청년','어린이','아동','노인','어르신','청년' 등 인구특성 관련된 키워드 카운트 분석 결과가 인구데이터 특성별 개표결과와 연관성있는지 파악해보기
 - **STEP0** : 후보별 인구특성 키워드 토큰화 및 카운트 DF 생성 @안지은
 - **STEP1**: 인구데이터 EDA (고령인구 비율 높은 지역/아동비율 높은 지역/청년층 높은지역/중장년층) → 이미 분석 한거로 각각 비율이 가장 높은 지역 5~10 개씩 뽑아서 코드 정리 해놓기 +바그래프 시각화 @김가인
 - **STEP2**: → 각 인구 특성이 뚜렷한 지역에서 얼마나 일관되게 개표(ex : 유소년 인구비율이 높은 a,b,c,d 지역에서 공약에 '어린이' 키워드가 가장 많이 나왔던 A후보의 득표율이 모두 전국 득표율에 비해 약간 높게 나타났다. / 혹은 a,b,c,d 모두 천차만별이다! 의미없다!) 결과나왔는지 파악. **4/9 최종 정리하면서 분석**

하면 좋고, 안 해도 괜찮고...

- JSON 파일 만들기 @hyemin song @김가인
 - 참고 : <https://park9eon.com/how-to-convert-to-korea-shp-geojson/>)

최종적으로 나올 그림?

- ☒ 워드클라우드
 - ☒ 전체 워클
 - ☒ 주제분야별 워드클라우드 (새로 주제 분류 한걸로 다시)
 - ☒ 후보별 전체 워클
- ☐ 네트워크
 - ☐ 주제분야별 네트워크 - 주제분야에서 최다 빈도 단어로 네트워크
 - ☐ 전체 최다빈도 네트워크 (강화, 확대, 추진)
- ☐ 바그래프
 - ☒ 주제 분야별 최다빈도 단어 그래프
 - ☒ 후보별 인구특성 키워드 카운트 그래프
 - ☐ 특징적 인구특성 보이는 지역 인구 분포 그래프(바그래프) (키워드 : '노인' → 1번 :37회 2번 :44회, 3번 33회 → 바 그래프로 후보별로 비교해볼 수 있도록)

+●해석



NOTE

- 정의당 네트워크 컬러맵 색상 재조정 필요
-