



## 03.27(토)

### ▼ 자연어 처리

- 불용어처리
  - 지프의 법칙 적용하기

### ▼ EDA

- 각 후보별로 무슨 단어가 가장 많이 나왔는지 순위 5순위까지(명사만)
- 각 후보별로 어떤 지역명이 가장 많이 나오는지 - 개표결과와 함께 보기
- 각 정당 명 구글 트렌드 입력하고 상위 트렌드 5개 확인
- (아동&어린이), (노인&어르신) 단어 등장 빈도 → 실제 지역별 연령층 비율과 함께 보기

+) 연관검색어(구글트렌드, 네이버 데이터랩)

---

## 할일

1. EDA : 후보상관없이 전체 \_ 🌟진짜진짜 완료해야함.
2. 시각화
  - 인구/개표 할 지도 시각화 → json파일도
  - 인터랙티브 시각화 툴 찾기
  - 자연어 처리 관련 시각화