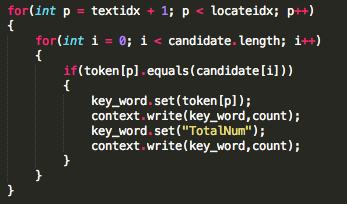
프로젝트 요약문서

# PROJECT’S subject

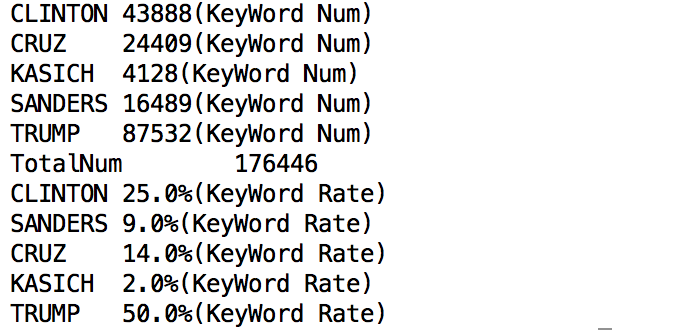
2016년 미국 대선 후보 5명(힐러리 클린턴, 버니 샌더스, 테드 크루즈, 존 케이식, 도널드 트럼프)의 인지도 분석하기

# Firtst Output

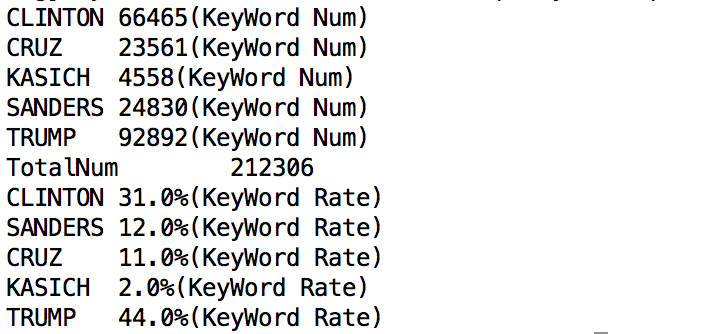
뉴욕과 펜실베니아에서 어떤 후보들이 트위터에서 가장 많이 언급되는지에 대해 분석하고, 통계를 내 보았다.



1. 뉴욕 결과



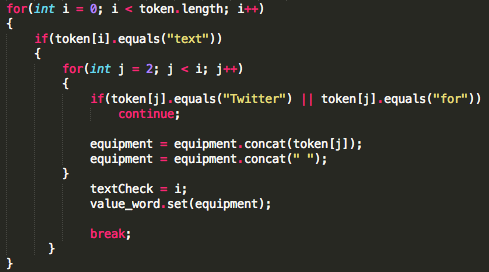
2. 펜실베니아 결과

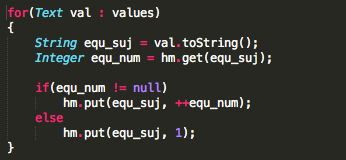


분석한 결과 뉴욕과 펜실베니아에서 모두 트럼프가 우세한 것을 알 수 있었다. 비율로 보면 뉴욕에서는 트럼프가 압도적으로 우세했지만, 펜실베니아에서는 두번째로 높은 클린턴과 10% 차이밖에 나지 않는 것을 확인할 수 있었다.

# second Output

트위터를 하는 사람들이 어떤 기기를 가장 많이 사용하는 지를 알고자, 뉴욕과 펜실베니아의 데이터를 합쳐서 하나의 데이터 셋으로 지역과 기기 사용 횟수를 나타내 보았다.





dataset을 하나로 합쳐서 지역별로 기기 종류에 대한 정보를 찾아서 카운트를 해 주었다.

1. 뉴욕



분석한 결과 뉴욕에서는 트위터를 하는 사람들 중에 아이폰을 사용하는 사람이 제일 많았다는 것을 확인할 수 있었다.

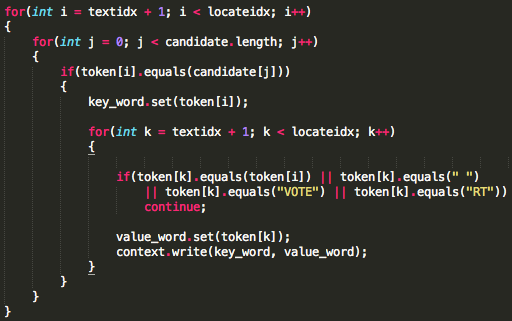
2. 펜실베니아



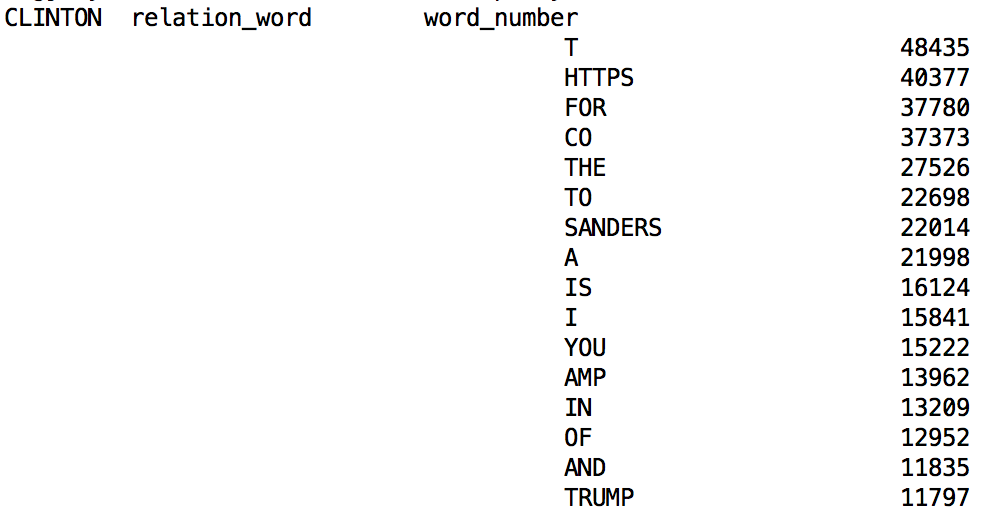
분석한 결과 펜실베니아에서도 트위터를 하는 사람들 중에 아이폰을 사용하는 사람이 제일 많았다는 것을 확인할 수 있었다.

# third output

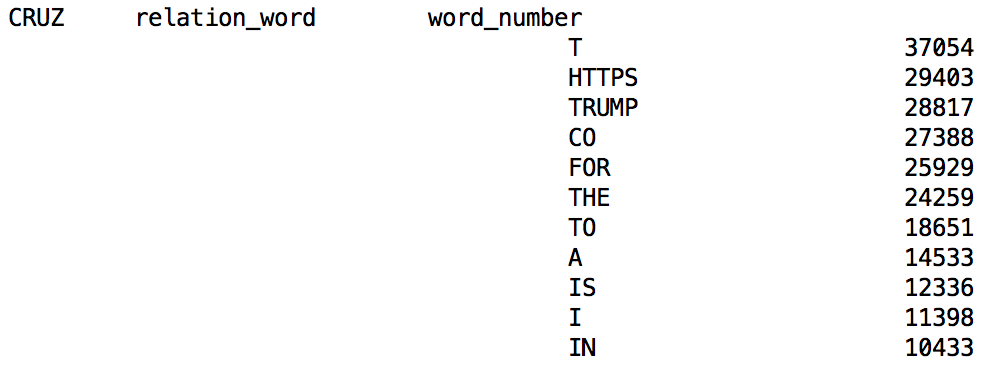
트위터에 언급된 대선 후보들에게 어떤 단어들이 연관되어 나오는지를 조사하기 위해 단어 상관 관계를 나타내보았다.



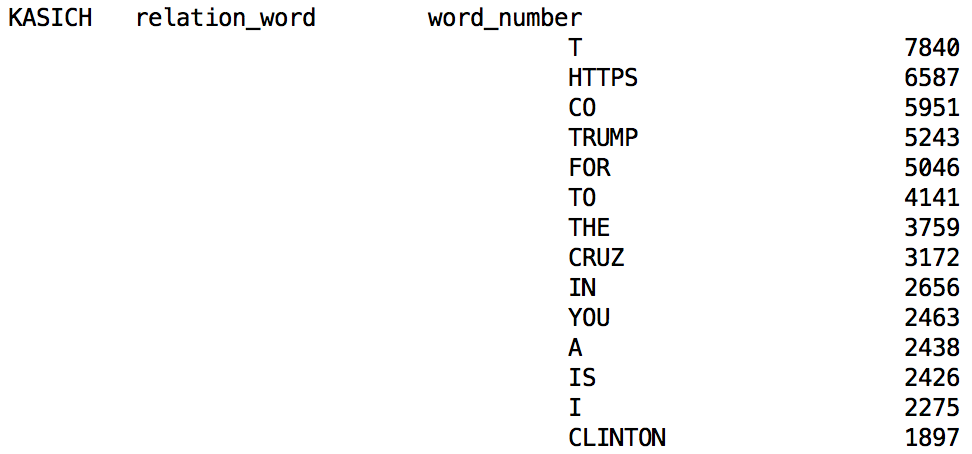
1. 힐러리 클린턴



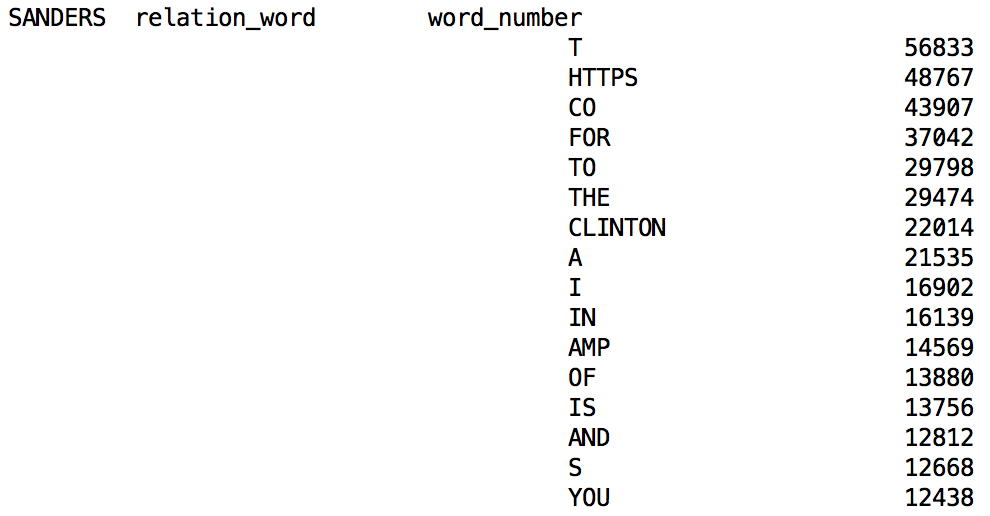
2. 테드 크루즈



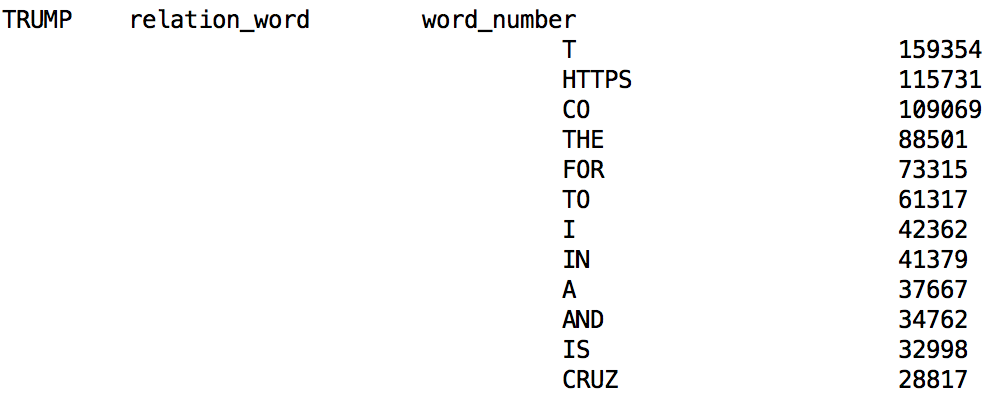
3. 존 케이식



4. 버니 샌더스



5. 도널드 트럼프



# conclusion

이번 프로젝트를 진행한 결과, 뉴욕과 펜실베니아에서의 대선 후보들의 인지도와 해당 후보들과 관련하여 어떤 단어가 많이 연관되었는지, 그리고 어떤 기기를 많이 사용하여 트위터를 하는지를 알 수 있었다. 이를 바탕으로 많이 이용하는 기기들을 이용하여 후보들이 자신의 선거 전략 등을 홍보하면 더 많은 유권자들을 모을 수 있을 것이라고 예상한다.