**- FOMC Meeting statements 분석 -**

**코퍼스 분석 대상 소개 및 선정 이유**

Federal Open Market Committee(FOMC – 연방공개시장위원회 : 연방 준비 제도의 단기 통화 정책 결정 기구)는 미국의 통화 및 금리 정책을 결정하는 기구로 매년 8회씩 회의를 통해 기준금리를 인상, 인하 또는 동결할지 결정한다. 한 나라의 경제에서 국민들도 피부에 직접적으로 와 닿도록 체감할 정도의 영향력을 끼치며 여러 가지 경제 요소에 막대한 영향을 미치는 것이 기준금리인데, 이러한 기준금리를 결정하는 회의록을 분석하여 금리 결정에 영향을 미치는 요인이 어떤 것인지 등을 분석하고자 한다.

**코퍼스 구축**

이에 따라 Federal Reserve(미국 연방준비 위원회) 홈페이지에 들어가 공개된 모든 회의록인 2011년부터 2016년 까지 모든 회의록을 전수조사 하였다.

코퍼스를 구축하여 도출해내려는 것은 금리 결정에 영향을 미치는 요인을 분석해내는 것이다. 기준금리를 결정하는데 있어서 계절적 요인이 중요하게 작용할지 안 할지는 분석하기 전까진 잘 모르겠지만, 경제에서 기본적으로 seasonal한 요인이 중요하기 때문에 모든 회의록을 담은 파일 외에 봄, 여름, 가을, 겨울로 나누어 파일을 생성하였고, 11년부터 13년까지 전기, 14년부터 16년까지 후기 파일을 만들었다. 회의는 일년에 총 8번 개최되는데 3월 6월 9월 12월에 고정적으로 4번 열리며 1월 또는 2월에 열리는 회의는 1월로, 4월 또는 5월에 열리는 회의는 4월로, 7월 또는 8월에 열리는 회의는 7월로, 10월 또는 11월에 열리는 회의는 11월로 통일하여 처리하였다. 따라서 각 파일의 제목은 년-월.txt 형식으로 작성했다. (예 : 13년 4월 회의록은 13-04.txt)

**전처리, 통계기법 사용 및 해석**

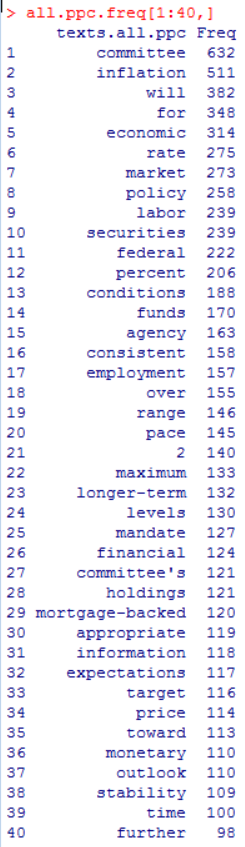
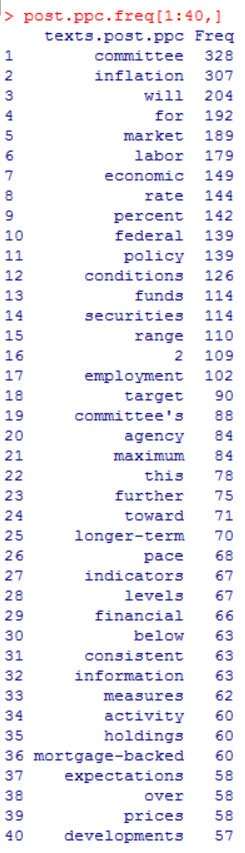
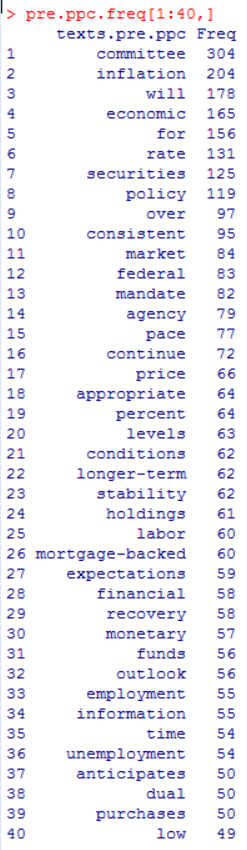
1. 대소문자가 구분되어 빈도표에 도출되는 것을 방지하기 위해 소문자 변환을 하였다.

2. 쉼표나 구두점 등에 의해 동일한 단어가 구분되어 빈도표에 도출되는 것을 방지하기 위해 첫머리나 끝부분에 나오는 문장부호를 제거하였다.

3. 빈도 데이터 프레임을 생성한다.

4. 3번에서 생성한 데이터 프레임을 관찰한 후 별 의미 없다고 판단되는 단어들을(불용어 등) 제거한 전처리 텍스트 변수를 생성한다. 단, 나중에 연어분석 등에서 전처리를 하지 않은 텍스트가 필요할 수도 있으므로(물론 필요 없을 수도 있다.) 기존 텍스트에 덮어쓰는 것이 아니라 .ppc를 사용해 전처리를 한 변수를 생성한다.

5. 전처리 한 빈도 데이터 프레임을 생성한다.

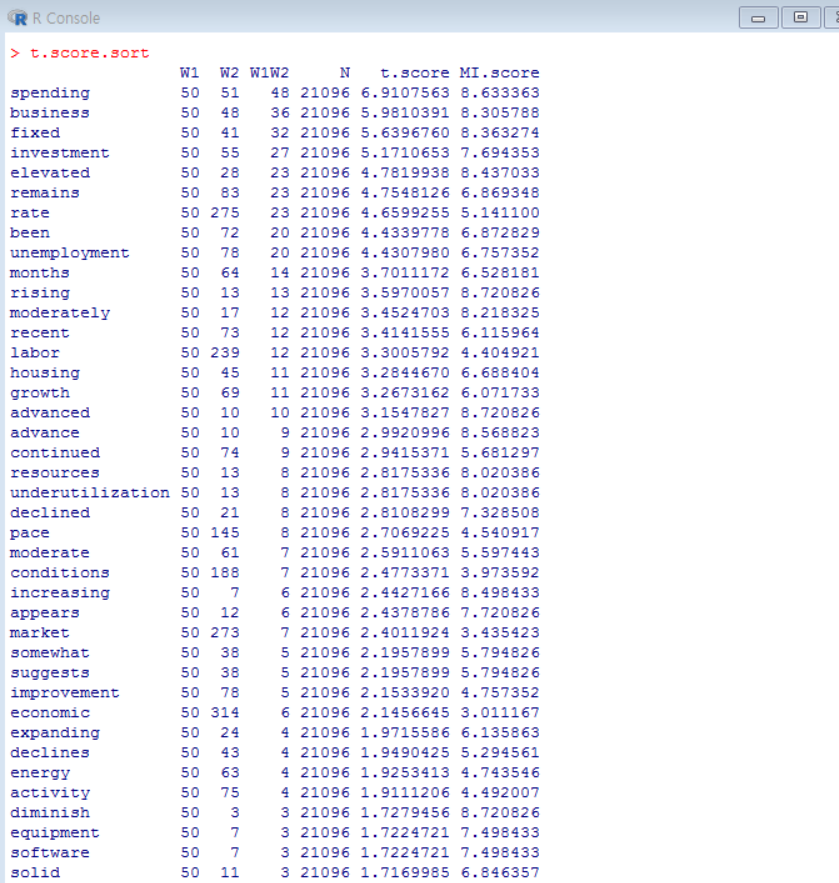
계절별로도 분석을 해봤으나 아쉽게도 그렇게 큰 차이점은 없었다. 따라서 모든 데이터, 11~13년도 전기, 14~16년도 후기 데이터의 빈도를 분석하기로 했다.

공통적으로 will, outlook, expectations 뿐아니라 위 스크린샷에는 없지만 seeks ,likely와 같은 단어가 많이 등장한다. 이는 금리를 결정할 때 과거의 사안에 집중하기 보다는 미래에 초점을 두고 앞으로 원하는 방향을 설정하여 추구하는 금리 정책을 펼친다고 볼 수 있겠다.

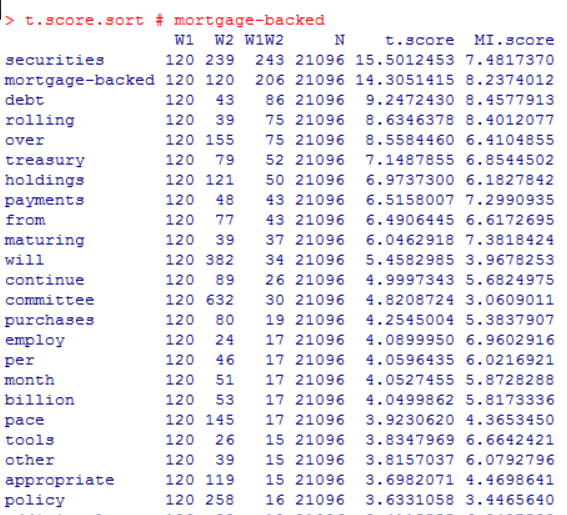
그리고 price, prices, employment, unemployment, inflation과 같은 단어가 많이 등장하는데 원글 을 읽어보면 미연준의 통화정책의 두가지 초점이 물가안정과 고용 극대화인 만큼, 각 시점에서 물가와 고용 상황에 따라 금리정책을 펼친다는 것을 알 수 있다. 따라서 이는 금리결정에 영향을 미치는 요인이라고 추측할 수 있다.

전기와 후기를 비교하여 키워드 분석을 했을 때 전기에는 unemployment가 빈도표에 등장하는 반면 후기에는 등장하지 않는다. 실제로 미국 고용상황이 2011년 초반보다 후반에 개선되었는데 이를 반영한 증거가되는 지표라고 판단할 수 있다. Price 또한 전기에 더 많이 등장하고 recovery도 마찬가지다 반면에 후기에는 전기에 등장하지 않는 단어인 progress가 등장하는데 2008년 미국발 금융위기에 여파가 전기에 더 많이 남아있어 그 경제상황을 반영했다고 볼 수 있겠다.

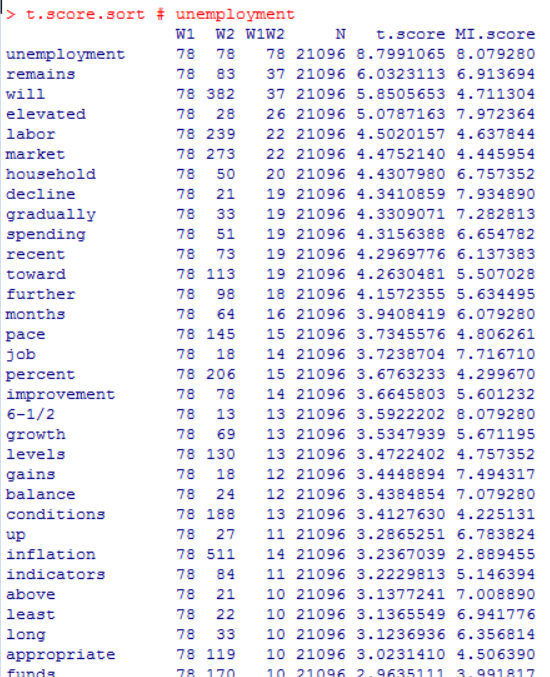
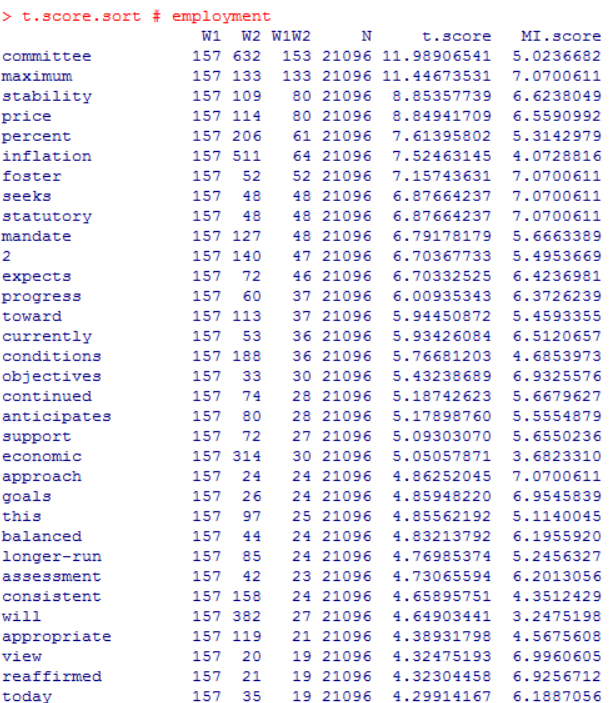
Mortgage-backed라는 단어 또한 자주 등장하는 것을 볼 수 있는데 이는 securities와 함께 쓰여 주택저당증권이라는 뜻으로, 가계부채와 연관될 수 있는 단어인데 가계부채는 금리와 상당히 밀접한 관계를 갖고 있기 때문에 중요한 경제 지표이므로 눈여겨 봐야겠다.



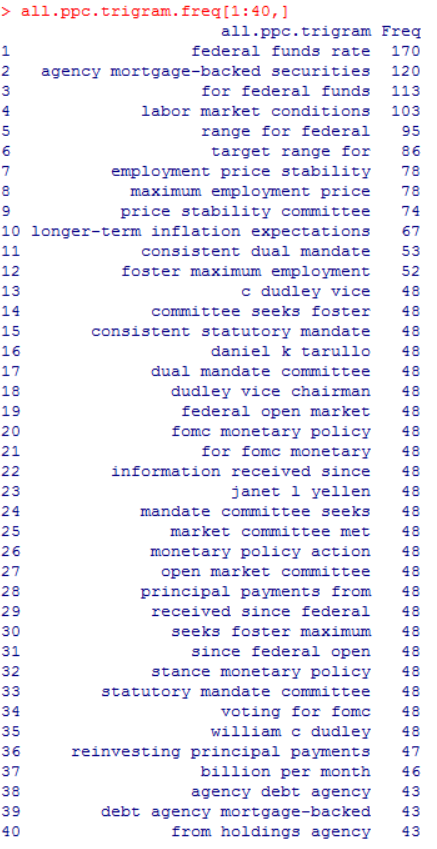
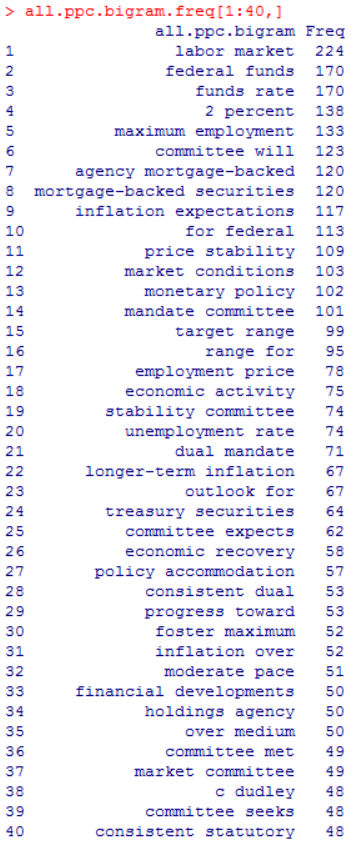
빈도분석에서 중요하다가 생각했던 것들에 대해 조사하기 위해 우선 빈도표에서 103번에 등장하고 가계부채와도 연관이 있을 것 같은 household를 키워드로 설정하여 연어분석을 해보았다. 아쉽게도 나오기를 기대하였던 mortgage-backed securities는 등장하지 않지만, spending, business, investment 등과 같은 단어들이 보이는데 이는 기본적으로 매우 중요한 경제지표인 household spending(가계소비), businees investment(기업투자)등 이 기준금리 결정할 때 역시 고려된다는 점을 알 수 있었다.



빈도분석에서 금리 결정에 영향을 미치는 요인으로 추측하였던 Mortgage-backed를 중심어로 연어분석을 한 결과이다. Securities(주택저당증권), debt(빚), holdings(자산)와 같은 단어들이 함께 많이 등장하는데 가계 부채가 금리결정에 영향을 미치는 요인임을 검증할 수 있었다.



Employment를 연어분석 했을 시 눈에 뛰는 단어는 역시 추측대로 price, maximum, 그리고 stability였다. 미연준의 두가지 목표인 고용 극대화(maximum employment)와 물가 안정(price stability)을 방향삼아 금리결정 실행함을 알 수 있었다. Price로 연어분석을 했을 때 또한 비슷한 결과가 나온다. 그리고 unemployment 또한 연어분석을 해봤는데 그 결과 employment를 검색했을 때 등장하지 않는 household, decline, spending과 같은 단어들이 등장한다. 실업과 같은 단어의 상황에선 가계 소비 감소와 같은 상황이 동반된다는 사실 또한 알 수 있다.



바이그램을 관찰해 보면 노동시장, 실업률, 최대 조장 등과 같이 역시 고용을 중요시하고 있음을 재차 확인할 수 있으며 longer-term inflation과 같은 단어가 보이는데 인플레이션 같은 경우는 단기적이 아닌 장기적 관점에서 다루고 있음을 알 수 있다. 또한 바이그램과 트라이그램을 관찰했을 때 트라이그램에서 위원들의 이름이 나오는 것을 제외하면 대부분 바이그램과 일치함을 알아냈다. 그래서 원문으로 돌아가 몇 개의 회의록을 관찰한 결과 매 회의록 톳시하나 틀리지 않고 그대로 사용하는 어구가 있음을 발견할 수 있었다.



전기와 후기의 바이그램을 비교 클라우드로 시각화하여 관찰해 보았다.

전기 같은 경우는 economic recovery나 unemploy rate와 같은 단어가 눈에 띄는 반면 후기에는 labor market, holdings longer- term, inflation this와 같은 단어가 눈에 띈다. 이는 전기의 좋지 않은 경제 상황과, 후기의 상대적으로 나은 경제상황을 반영하는 결과라고 볼 수 있는데 후기에 등장한 단어들은 일반적으로 경제를 안정화에 고려되는 키워드들이 등장했다고 볼 수 있다. 앞서 빈도분석과 연어분석시 하였던 키워드 분석과 일맥상통하는 내용임을 알 수 있다.

**과정 요약 및 결론**

차이점이 있을까 하는 궁금증을 품고 일단 구축한 계절별 분류간 차이는 거의 없는 것으로 판단.

전기 / 후기는 차이가 존재하여 비교

주로 모든 회의록이 담긴 all 파일을 통해서 금리결정에 영향을 미치는 요인을 확인함

빈도분석에서 추측한 사항을 연어분석을 통해 검증하는 하였다. (키워드 분석 수반)

all파일의 바이그램과 트라이그램을 분석하였다. (키워드 분석 수반)

전기와 후기의 BI-gram을 비교 워드클라우드화 하여 시각화 하고 분석해 보았다.

이번 프로젝를 통해서 얻고자 한 유의미한 결과는 금리 결정에 영향을 미치는 요인을 분석 검증하여 파악하는 것이었다. (기준금리는 두말할 것 없이 경제에서 중요한 변수일 뿐 아니라 국민 개개인에도 큰 영향을 미치는 요소이다. 그래서 이를 이미 예측할 수 있어서 그에 맞는 행동을 취할 수 있다면 매우 좋지 않을까 싶다.)

미 연준의 Fomc meeting statements를 분석한 결과 금리 인상, 인하, 동결을 결정할 때 고려하는 주요 요소로는 고용, 물가, 부채, 물가상승률, 가계소비, 기업투자 등이 있다. 앞으로 이번에 찾아낸 위와 같은 요소들을 금리에 선행하는 요소로 상정하고 관찰하여 한국은행이나 미연준의 금리정책을 발표전에 미리 예측하고 이득을 취하는 방향으로 경제적 행동을 할 수 있을 것이다.

주 목표에 더하여 부수인 사항으로는 인플레이션 같은 경우는 단기적이기 보다는 장기적 측면에서 대처하는 양상을 보였으며, 11, 12, 13년은 2008년 미국발 금융위기의 여파로 인해 14, 15, 16년에 비해 미국의 경제 상황이 좋지 않았음을 확인할 수 있었다.