

# - 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 8 회씩 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 2011 년부터 2016 년 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 seasonal 한 요인이 중요하기 때문에 모든 회의록을 담은 파일 외에 봄, 여름, 가을, 겨울로 나누어 파일을 생성하였고, 11 년부터 13 년까지 전기, 14 년부터 16 년까지 후기 파일을 만들었다. 회의는 일년에 총 8 번 개최되는데 3 월 6 월 9 월 12 월에 고정적으로 4 번 열리며 1 월 또는 2 월에 열리는 회의는 1 월로, 4 월 또는 5 월에 열리는 회의는 4 월로, 7 월 또는 8 월에 열리는 회의는 7 월로, 10 월 또는 11 월에 열리는 회의는 11 월로 통일하여 처리하였다. 따라서 각 파일의 제목은 년-월.txt 형식으로 작성했다. (예 : 13 년 4 월 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. 3 번에서 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 .ppc 를 사용해 전처리를 한 변수를 생성한다.
5. 전처리 한 빈도 데이터 프레임을 생성한다.

> all.ppc.freq[1:40,]			> pre.ppc.freq[1:40,]			> post.ppc.freq[1:40,]		
	texts.all.ppc	Freq		texts.pre.ppc	Freq		texts.post.ppc	Freq
1	committee	632	1	committee	304	1	committee	328
2	inflation	511	2	inflation	204	2	inflation	307
3	will	382	3	will	178	3	will	204
4	for	348	4	economic	165	4	for	192
5	economic	314	5	for	156	5	market	189
6	rate	275	6	rate	131	6	labor	179
7	market	273	7	securities	125	7	economic	149
8	policy	258	8	policy	119	8	rate	144
9	labor	239	9	over	97	9	percent	142
10	securities	239	10	consistent	95	10	federal	139
11	federal	222	11	market	84	11	policy	139
12	percent	206	12	federal	83	12	conditions	126
13	conditions	188	13	mandate	82	13	funds	114
14	funds	170	14	agency	79	14	securities	114
15	agency	163	15	pace	77	15	range	110
16	consistent	158	16	continue	72	16	2	109
17	employment	157	17	price	66	17	employment	102
18	over	155	18	appropriate	64	18	target	90
19	range	146	19	percent	64	19	committee's	88
20	pace	145	20	levels	63	20	agency	84
21	2	140	21	conditions	62	21	maximum	84
22	maximum	133	22	longer-term	62	22	this	78
23	longer-term	132	23	stability	62	23	further	75
24	levels	130	24	holdings	61	24	toward	71
25	mandate	127	25	labor	60	25	longer-term	70
26	financial	124	26	mortgage-backed	60	26	pace	68
27	committee's	121	27	expectations	59	27	indicators	67
28	holdings	121	28	financial	58	28	levels	67
29	mortgage-backed	120	29	recovery	58	29	financial	66
30	appropriate	119	30	monetary	57	30	below	63
31	information	118	31	funds	56	31	consistent	63
32	expectations	117	32	outlook	56	32	information	63
33	target	116	33	employment	55	33	measures	62
34	price	114	34	information	55	34	activity	60
35	toward	113	35	time	54	35	holdings	60
36	monetary	110	36	unemployment	54	36	mortgage-backed	60
37	outlook	110	37	anticipates	50	37	expectations	58
38	stability	109	38	dual	50	38	over	58
39	time	100	39	purchases	50	39	prices	58
40	further	98	40	low	49	40	developments	57

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

공통적으로 will, outlook, expectations 뿐아니라 위 스크린샷에는 없지만 seeks ,likely 와 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, inflation 과 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 unemployment 가 빈도표에 등장하는 반면 후기에는 등장하지 않는다. 실제로 미국 고용상황이 2011 년 초반보다 후반에 개선되었는데 이를 반영한 증거가되는 지표라고 판단할 수 있다. Price 또한 전기에 더 많이 등장하고 recovery 도 마찬가지로 반면에 후기에는 전기에 등장하지 않는 단어인 progress 가 등장하는데 2008 년 미국발 금융위기에 여파가 전기에 더 많이 남아있어 그 경제상황을 반영했다고 볼 수 있겠다.

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

R Console							
> t.score.sort							
	W1	W2	W1W2	N	t.score	MI.score	
spending	50	51	48	21096	6.9107563	8.633363	
business	50	48	36	21096	5.9810391	8.305788	
fixed	50	41	32	21096	5.6396760	8.363274	
investment	50	55	27	21096	5.1710653	7.694353	
elevated	50	28	23	21096	4.7819938	8.437033	
remains	50	83	23	21096	4.7548126	6.869348	
rate	50	275	23	21096	4.6599255	5.141100	
been	50	72	20	21096	4.4339778	6.872829	
unemployment	50	78	20	21096	4.4307980	6.757352	
months	50	64	14	21096	3.7011172	6.528181	
rising	50	13	13	21096	3.5970057	8.720826	
moderately	50	17	12	21096	3.4524703	8.218325	
recent	50	73	12	21096	3.4141555	6.115964	
labor	50	239	12	21096	3.3005792	4.404921	
housing	50	45	11	21096	3.2844670	6.688404	
growth	50	69	11	21096	3.2673162	6.071733	
advanced	50	10	10	21096	3.1547827	8.720826	
advance	50	10	9	21096	2.9920996	8.568823	
continued	50	74	9	21096	2.9415371	5.681297	
resources	50	13	8	21096	2.8175336	8.020386	
underutilization	50	13	8	21096	2.8175336	8.020386	
declined	50	21	8	21096	2.8108299	7.328508	
pace	50	145	8	21096	2.7069225	4.540917	
moderate	50	61	7	21096	2.5911063	5.597443	
conditions	50	188	7	21096	2.4773371	3.973592	
increasing	50	7	6	21096	2.4427166	8.498433	
appears	50	12	6	21096	2.4378786	7.720826	
market	50	273	7	21096	2.4011924	3.435423	
somewhat	50	38	5	21096	2.1957899	5.794826	
suggests	50	38	5	21096	2.1957899	5.794826	
improvement	50	78	5	21096	2.1533920	4.757352	
economic	50	314	6	21096	2.1456645	3.011167	
expanding	50	24	4	21096	1.9715586	6.135863	
declines	50	43	4	21096	1.9490425	5.294561	
energy	50	63	4	21096	1.9253413	4.743546	
activity	50	75	4	21096	1.9111206	4.492007	
diminish	50	3	3	21096	1.7279456	8.720826	
equipment	50	7	3	21096	1.7224721	7.498433	
software	50	7	3	21096	1.7224721	7.498433	
solid	50	11	3	21096	1.7169985	6.846357	

빈도분석에서 중요하다가 생각했던 것들에 대해 조사하기 위해 우선 빈도표에서 103 번에 등장하고 가계부채와도 연관이 있을 것 같은 household 를 키워드로 설정하여 연어분석을 해보았다. 아쉽게도 나오기를 기대하였던 mortgage-backed securities 는 등장하지 않지만, spending, business, investment 등과 같은 단어들이 보이는데 이는 기본적으로 매우 중요한 경제지표인 household spending(가계소비), business investment(기업투자)등 이 기준금리

결정할 때 역시 고려된다는 점을 알 수 있었다.

```
|> t.score.sort # mortgage-backed
      W1  W2 W1W2      N    t.score  MI.score
securities    120 239  243 21096 15.5012453  7.4817370
mortgage-backed 120 120  206 21096 14.3051415  8.2374012
debt          120  43   86 21096  9.2472430  8.4577913
rolling       120  39   75 21096  8.6346378  8.4012077
over          120 155   75 21096  8.5584460  6.4104855
treasury      120  79   52 21096  7.1487855  6.8544502
holdings      120 121   50 21096  6.9737300  6.1827842
payments      120  48   43 21096  6.5158007  7.2990935
from          120  77   43 21096  6.4906445  6.6172695
maturing      120  39   37 21096  6.0462918  7.3818424
will          120 382   34 21096  5.4582985  3.9678253
continue      120  89   26 21096  4.9997343  5.6824975
committee     120 632   30 21096  4.8208724  3.0609011
purchases     120  80   19 21096  4.2545004  5.3837907
employ        120  24   17 21096  4.0899950  6.9602916
per           120  46   17 21096  4.0596435  6.0216921
month         120  51   17 21096  4.0527455  5.8728288
billion       120  53   17 21096  4.0499862  5.8173336
pace          120 145   17 21096  3.9230620  4.3653450
tools         120  26   15 21096  3.8347969  6.6642421
other         120  39   15 21096  3.8157037  6.0792796
appropriate   120 119   15 21096  3.6982071  4.4698641
policy        120 258   16 21096  3.6331058  3.4465640
```

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

```

> t.score.sort # employment
      W1 W2 W1W2   N   t.score  MI.score
committee 157 632 153 21096 11.98906541 5.0236682
maximum    157 133 133 21096 11.44673531 7.0700611
stability   157 109 80 21096 8.85357739 6.6238049
price       157 114 80 21096 8.84941709 6.5590992
percent     157 206 61 21096 7.61395802 5.3142979
inflation   157 511 64 21096 7.52463145 4.0728816
foster      157 52 52 21096 7.15743631 7.0700611
seeks       157 48 48 21096 6.87664237 7.0700611
statutory   157 48 48 21096 6.87664237 7.0700611
mandate     157 127 48 21096 6.79178179 5.6663389
2           157 140 47 21096 6.70367733 5.4953669
expects     157 72 46 21096 6.70332525 6.4236981
progress    157 60 37 21096 6.00935343 6.3726239
toward      157 113 37 21096 5.94450872 5.4593355
currently   157 53 36 21096 5.93426084 6.5120657
conditions  157 188 36 21096 5.76681203 4.6853973
objectives  157 33 30 21096 5.43238689 6.9325576
continued   157 74 28 21096 5.18742623 5.6679627
anticipates 157 80 28 21096 5.17898760 5.5554879
support     157 72 27 21096 5.09303070 5.6550236
economic    157 314 30 21096 5.05057871 3.6823310
approach    157 24 24 21096 4.86252045 7.0700611
goals       157 26 24 21096 4.85948220 6.9545839
this        157 97 25 21096 4.85562192 5.1140045
balanced    157 44 24 21096 4.83213792 6.1955920
longer-run  157 85 24 21096 4.76985374 5.2456327
assessment  157 42 23 21096 4.73065594 6.2013056
consistent  157 158 24 21096 4.65895751 4.3512429
will        157 382 27 21096 4.64903441 3.2475198
appropriate 157 119 21 21096 4.38931798 4.5675608
view        157 20 19 21096 4.32475193 6.9960605
reaffirmed  157 21 19 21096 4.32304458 6.9256712
today       157 35 19 21096 4.29914167 6.1887056

> t.score.sort # unemployment
      W1 W2 W1W2   N   t.score  MI.score
unemployment 78 78 78 21096 8.7991065 8.079280
remains       78 83 37 21096 6.0323113 6.913694
will          78 382 37 21096 5.8505653 4.711304
elevated      78 28 26 21096 5.0787163 7.972364
labor         78 239 22 21096 4.5020157 4.637844
market        78 273 22 21096 4.4752140 4.445954
household     78 50 20 21096 4.4307980 6.757352
decline       78 21 19 21096 4.3410859 7.934890
gradually     78 33 19 21096 4.3309071 7.282813
spending      78 51 19 21096 4.3156388 6.654782
recent        78 73 19 21096 4.2969776 6.137383
toward        78 113 19 21096 4.2630481 5.507028
further       78 98 18 21096 4.1572355 5.634495
months        78 64 16 21096 3.9408419 6.079280
pace          78 145 15 21096 3.7345576 4.806261
job           78 18 14 21096 3.7238704 7.716710
percent       78 206 15 21096 3.6763233 4.299670
improvement   78 78 14 21096 3.6645803 5.601232
6-1/2        78 13 13 21096 3.5922202 8.079280
growth        78 69 13 21096 3.5347939 5.671195
levels        78 130 13 21096 3.4722402 4.757352
gains         78 18 12 21096 3.4448894 7.494317
balance       78 24 12 21096 3.4384854 7.079280
conditions    78 188 13 21096 3.4127630 4.225131
up            78 27 11 21096 3.2865251 6.783824
inflation     78 511 14 21096 3.2367039 2.889455
indicators    78 84 11 21096 3.2229813 5.146394
above         78 21 10 21096 3.1377241 7.008890
least         78 22 10 21096 3.1365549 6.941776
long          78 33 10 21096 3.1236936 6.356814
appropriate   78 119 10 21096 3.0231410 4.506390
funds         78 170 10 21096 2.9635111 3.991817

```

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



```

> all.ppc.bigram.freq[1:40,]
      all.ppc.bigram Freq
1      labor market  224
2      federal funds 170
3      funds rate   170
4      2 percent    138
5      maximum employment 133
6      committee will 123
7      agency mortgage-backed 120
8      mortgage-backed securities 120
9      inflation expectations 117
10     for federal 113
11     price stability 109
12     market conditions 103
13     monetary policy 102
14     mandate committee 101
15     target range 99
16     range for 95
17     employment price 78
18     economic activity 75
19     stability committee 74
20     unemployment rate 74
21     dual mandate 71
22     longer-term inflation 67
23     outlook for 67
24     treasury securities 64
25     committee expects 62
26     economic recovery 58
27     policy accommodation 57
28     consistent dual 53
29     progress toward 53
30     foster maximum 52
31     inflation over 52
32     moderate pace 51
33     financial developments 50
34     holdings agency 50
35     over medium 50
36     committee met 49
37     market committee 49
38     c dudley 48
39     committee seeks 48
40     consistent statutory 48

> all.ppc.trigram.freq[1:40,]
      all.ppc.trigram Freq
1      federal funds rate 170
2      agency mortgage-backed securities 120
3      for federal funds 113
4      labor market conditions 103
5      range for federal 95
6      target range for 86
7      employment price stability 78
8      maximum employment price 78
9      price stability committee 74
10     longer-term inflation expectations 67
11     consistent dual mandate 53
12     foster maximum employment 52
13     c dudley vice 48
14     committee seeks foster 48
15     consistent statutory mandate 48
16     daniel k tarullo 48
17     dual mandate committee 48
18     dudley vice chairman 48
19     federal open market 48
20     fomc monetary policy 48
21     for fomc monetary 48
22     information received since 48
23     janet l yellen 48
24     mandate committee seeks 48
25     market committee met 48
26     monetary policy action 48
27     open market committee 48
28     principal payments from 48
29     received since federal 48
30     seeks foster maximum 48
31     since federal open 48
32     stance monetary policy 48
33     statutory mandate committee 48
34     voting for fomc 48
35     william c dudley 48
36     reinvesting principal payments 47
37     billion per month 46
38     agency debt agency 43
39     debt agency mortgage-backed 43
40     from holdings agency 43

```

바이그램을 관찰해 보면 노동시장, 실업률, 최대 조장 등과 같이 역시 고용을 중요시하고 있음을 재차 확인할 수 있으며 longer-term inflation 과 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 토시하나 틀리지 않고 그대로 사용하는 어구가 있음을 발견할 수 있었다.





전기와 후기의 바이그램을 비교 클라우드로 시각화하여 관찰해 보았다.

전기 같은 경우는 economic recovery 나 unemploy rate 와 같은 단어가 눈에 띄는 반면 후기에는 labor market, holdings longer- term, inflation this 와 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)

all 파일의 바이그램과 트라이그램을 분석하였다. (키워드 분석 수반)

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

이번 프로젝트를 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 statements 를 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, 13 년은 2008 년 미국발 금융위기의 여파로 인해 14, 15, 16 년에 비해 미국의 경제 상황이 좋지 않았음을 확인할 수 있었다.