

King Sejong's National Referendum on Tax Reform : Data

20173204 광명빈

2020-06-08

Data

Reading Data

원본 자료는 세종실록 인터넷 판에 기록으로 있는데(세종12년 8월10일, 호조에서 공법에 대한 여러 의논을 갖추어 아뢰다) 3-4명이 팀을 이루어 요약할 때 1-2시간 정도 걸린다.

```
knitr::include_graphics("../pics/sejong_silok_12-08-10.png")
```

세종실록 49권, 세종 12년 8월 10일 무인 5번째기사 1430년 명 선덕(宣德) 5년

호조에서 공법에 대한 여러 의논을 갖추어 아뢰다

국역 원문 원본 보기

호조에서 중외(中外)의 **공법**(貢法)에 대한 가부(可否)의 의논을 갖추어 아뢰기를,

"서울의 지돈녕부사(知敦寧府事) **안수산**(安壽山)·총제 **이천**(李薦)·동지총제(同知摠制) **박규**(朴葵)·전 총제 **이순몽**(李順蒙)·전 동지총제 **이희귀**(李希貴)·전 도관찰사 **이정간**(李貞幹)·전 판목사(判牧事) **김사청**(金士淸)·전 중추원 부사(中樞院副使) **남실**(南實)·전 동지총제 **최견**(崔鑑)과, 3품 이하 현재 재직 중에 있는 2백 59명과, 전함(前銜) 4백 43명 등은 가하다고 하고, 전 병조 판서 **조말생**(趙末生)·전 판목사(判牧事) **황자후**(黃子厚) 등은 아뢰기를, '신 등이 옛 사람들의 나

아래 <표 1>은 오기수 교수님의 요약본을 옮긴 것이다. 전형적인 테이블 구조이나 아직 깔끔한(tidy) 데이터로 정리하지 않은 상태이다. 찬반 집계를 value 로 간주할 때, 각 value 들은 어느 지역, 어떤 계급의 사람들이 찬성이나 반대를 한 결과를 합산한 것이다. 지역, 계급, 찬반, 집계의 네 가지 변수가 등장하는 데 이를 집계에 대하여 정리하면 깔끔한 데이터가 된다.

```
knitr::include_graphics("../pics/sejong_ref_data.png")
```

<표 1> 공법관련 여론조사의 찬성과 반대 분석 (단위 : 명, %)

		대신·관찰사·도사 등				수령				품관·촌민				합계			
		찬성		반대		찬성		반대		찬성		반대		찬성		반대	
		인수	%	인수	%	인수	%	인수	%	인수	%	인수	%	인수	%	인수	%
대신 등		21 ^{*1}	9.8	194 ^{*2}	90.2									21	9.8	194	90.2
3품 이하	현직	259	39.7	393	60.3									259	39.7	393	60.3
	전직	443	79.1	117	20.9									443	79.1	117	20.9
유후사 ²³⁾										1,123	94.1	71	5.9	1,123	94.1	71	5.9
경기도						29	85.3	5	14.7	17,076	98.6	236	1.4	17,105	98.6	241	1.4
평안도				1		6	14.6	35	85.4	1,326	4.4	28,474	95.6	1,332	4.5	28,510	95.5
황해도						17	50.0	17	50.0	4,454	22.2	15,601	77.8	4,471	22.3	15,618	77.7
충청도				2		35	57.4	26	42.6	6,982	33.3	14,013	66.7	7,017	33.3	14,041	66.7
강원도						5	33.3	10	66.7	939	12.0	6,888	88.0	944	12.0	6,898	88.0
함길도				1		3	17.6	14	82.4	75	1.0	7,387	99.0	78	1.0	7,402	99.0
경상도						55	77.5	16	22.5	36,262	99.0	377	1.0	36,317	98.9	393	1.1
전라도				2		42	77.8	12	22.2	29,505	99.1	257	0.9	29,547	99.1	271	0.9
		723	50.5	710	49.5	192	58.7	135	41.3	97,742	57.1	73,304	42.9	98,657	57.1	74,149	42.9

출처 : 「세종실록」 12년(1430) 8월 10일

sejong_ref.txt 는 위의 표를 깔끔한 구조로 정리한 텍스트 파일이다.

```
sejong_ref <- read.table("../data/sejong_ref.txt",
                        header = TRUE,
                        stringsAsFactors = FALSE)

str(sejong_ref)
```

```
## 'data.frame':   44 obs. of  4 variables:
## $ Counts: int  21 194 259 393 443 117 1123 71 29 5 ...
## $ Vote : chr  "찬성" "반대" "찬성" "반대" ...
## $ Class : chr  "대신_등" "대신_등" "3품이하현직" "3품이하현직" ...
## $ Region: chr  "서울" "서울" "서울" "서울" ...
```

```
kable(sejong_ref[4:1])
```

Region	Class	Vote	Counts
서울	대신_등	찬성	21
서울	대신_등	반대	194
서울	3품이하현직	찬성	259
서울	3품이하현직	반대	393
서울	3품이하전직	찬성	443
서울	3품이하전직	반대	117
유후사	품관촌민	찬성	1123
유후사	품관촌민	반대	71
경기	수령	찬성	29
경기	수령	반대	5
경기	품관촌민	찬성	17076
경기	품관촌민	반대	236
평안	대신_등	반대	1
평안	수령	찬성	6
평안	수령	반대	35
평안	품관촌민	찬성	1326
평안	품관촌민	반대	28474
황해	수령	찬성	17
황해	수령	반대	17
황해	품관촌민	찬성	4454
황해	품관촌민	반대	15601
충청	대신_등	반대	2

Region	Class	Vote	Counts
충청	수령	찬성	35
충청	수령	반대	26
충청	품관촌민	찬성	6982
충청	품관촌민	반대	14013
강원	수령	찬성	5
강원	수령	반대	10
강원	품관촌민	찬성	939
강원	품관촌민	반대	6888
함길	대신_등	반대	1
함길	수령	찬성	3
함길	수령	반대	14
함길	품관촌민	찬성	75
함길	품관촌민	반대	7387
경상	수령	찬성	55
경상	수령	반대	16
경상	품관촌민	찬성	36262
경상	품관촌민	반대	377
전라	대신_등	반대	2
전라	수령	찬성	42
전라	수령	반대	12
전라	품관촌민	찬성	29505
전라	품관촌민	반대	257

Factor conversion

Vote, Class, Region 은 factor 로 전환하여야 한다. 이들을 character 로 놓아 두면, table() 로 요약할 때 알파벳 순, 가나다 순으로 하기 때문이다. 다음을 실행해 보면 쉽게 알 수 있다.

```
table(sejong_ref$Vote)
```

```
##
## 반대 찬성
##    24    20
```

```
table(sejong_ref$Class)
```

```
##
## 3품이하전직 3품이하현직 대신_등 수령 품관촌민
##          2          2          6          16          18
```

```
table(sejong_ref$Region)
```

```
##
## 강원 경기 경상 서울 유후사 전라 충청 평안 함길 황해
##      4      4      4      6      2      5      5      5      5      4
```

따라서 `factor()` 를 활용하여 변환한다. `Vote` 는 “찬성”, “반대”의 순으로 `level` 을 설정한다.

```
sejong_ref$Vote <- factor(sejong_ref$Vote,
                          levels = c("찬성", "반대"))
```

`Class` 는 중앙관료, 지방관료, 품관촌민의 순으로 `levels` 를 설정한다. 단, “대신_등”을 “대신 등”으로 표시하기 위하여 `labels` 도 설정한다.

```
Class_levels <- c("대신_등", "3품이하현직", "3품이하전직", "수령", "품관촌민")
Class_labels <- c("대신 등", "3품이하현직", "3품이하전직", "수령", "품관촌민")
sejong_ref$Class <- factor(sejong_ref$Class,
                           levels = Class_levels,
                           labels = Class_labels)
```

`Region` 은 다음 순서대로 `levels` 를 설정한다.

```
Region_levels <- c("서울", "유후사", "경기", "평안", "황해", "충청", "강원", "함길", "경상",
                  "전라")
sejong_ref$Region <- factor(sejong_ref$Region,
                           levels = Region_levels)
```

구조를 살피고, 'Region', 'Class', 'Vote', 'Counts' 순으로 출력한다. 맨 처음 줄을 해석하면 서울의 대신 등은 21명이 찬성하였고, 194명이 반대하였다고 읽을 수 있다. 이 숫자가 매우 중요하다. 분명히 세종대왕이 임명한 최고급관료 중 90%이상이 세종대왕의 개혁 정책에 반대하고 있는 것이다. 21이라는 값은 “서울”, “대신 등”, “찬성”이라는 key 값에 대응한다.

```
str(sejong_ref)
```

```
## 'data.frame':   44 obs. of  4 variables:
## $ Counts: int  21 194 259 393 443 117 1123 71 29 5 ...
## $ Vote : Factor w/ 2 levels "찬성","반대": 1 2 1 2 1 2 1 2 1 2 ...
## $ Class : Factor w/ 5 levels "대신 등","3품이하현직",...: 1 1 2 2 3 3 5 5 4 4 ...
## $ Region: Factor w/ 10 levels "서울","유후사",...: 1 1 1 1 1 1 2 2 3 3 ...
```

```
kable(sejong_ref[4:1])
```

Region	Class	Vote	Counts
서울	대신 등	찬성	21
서울	대신 등	반대	194

Region	Class	Vote	Counts
서울	3품이하현직	찬성	259
서울	3품이하현직	반대	393
서울	3품이하전직	찬성	443
서울	3품이하전직	반대	117
유후사	품관촌민	찬성	1123
유후사	품관촌민	반대	71
경기	수령	찬성	29
경기	수령	반대	5
경기	품관촌민	찬성	17076
경기	품관촌민	반대	236
평안	대신 등	반대	1
평안	수령	찬성	6
평안	수령	반대	35
평안	품관촌민	찬성	1326
평안	품관촌민	반대	28474
황해	수령	찬성	17
황해	수령	반대	17
황해	품관촌민	찬성	4454
황해	품관촌민	반대	15601
충청	대신 등	반대	2
충청	수령	찬성	35
충청	수령	반대	26
충청	품관촌민	찬성	6982
충청	품관촌민	반대	14013
강원	수령	찬성	5
강원	수령	반대	10
강원	품관촌민	찬성	939
강원	품관촌민	반대	6888
함길	대신 등	반대	1
함길	수령	찬성	3
함길	수령	반대	14
함길	품관촌민	찬성	75

Region	Class	Vote	Counts
함길	품관촌민	반대	7387
경상	수령	찬성	55
경상	수령	반대	16
경상	품관촌민	찬성	36262
경상	품관촌민	반대	377
전라	대신 등	반대	2
전라	수령	찬성	42
전라	수령	반대	12
전라	품관촌민	찬성	29505
전라	품관촌민	반대	257

- Save the working directory image.

```
save.image(file = "sejong_ref_data.RData")
```

Comments

기존 조세제도가 관료들이 직접 관측학양을 바탕으로 세금으로 거둘 양을 결정하였기 때문에 부정부패가 있을 수 있어 백성들에게 이로운 조세제도를 만들게 되었다. 이는 비옥도와 날씨에 따라 세금을 부과하겠다는 것이었다. 결과를 보게 되면 백성들의 투표결과가 관료들의 투표결과에 영향을 미친다. 때문에 관료들과 백성이 비슷한 극단적인 결과가 나타났다고 볼 수 있다. 시대적 배경에 따라 백성들이 제대로 된 글을 읽을수 없어 나타났다고 생각할 수 있다. 데이터 과학자는 데이터를 만들뿐만아니라 데이터에 함축되어있는 시대적배경등 상황을 이해 할 수 있어야 하는것을 알게되었다.