Big Data Analytics: LAB04
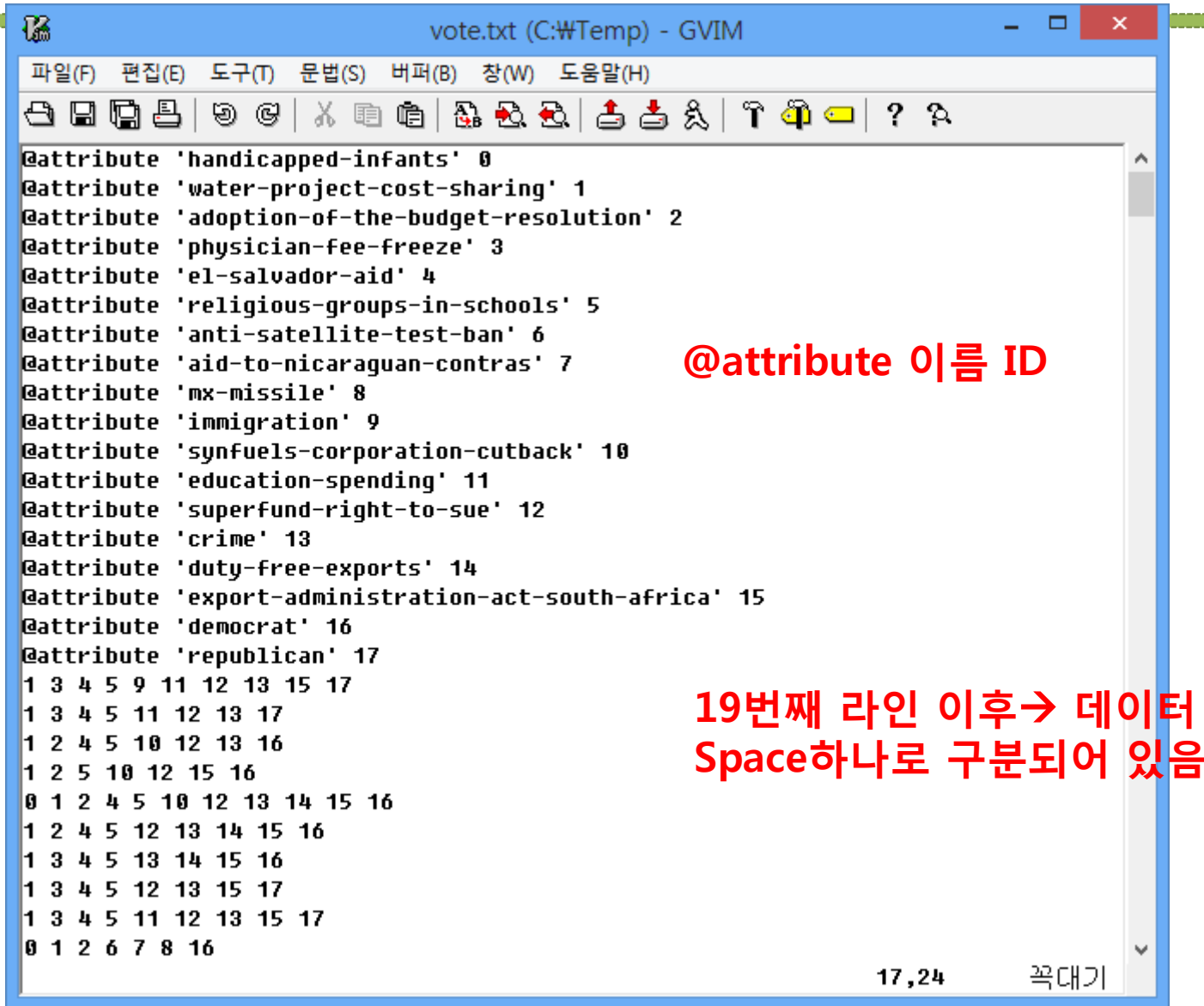
Younghoon Kim

(nongaussian@hanyang.ac.kr)

Problem 1

- Given
  - A transaction table available at
    - HY-IN
  - A minimum support (50 percentage ratio)
- Find
  - Every frequent item sets using Apriori Algorithm
  - (standalone program with Java)

Data format

vote.txt (C:\Temp) - GVIM

파일(F) 편집(E) 도구(T) 문법(S) 버퍼(B) 창(W) 도움말(H)

```
@attribute 'handicapped-infants' 0
@attribute 'water-project-cost-sharing' 1
@attribute 'adoption-of-the-budget-resolution' 2
@attribute 'physician-fee-freeze' 3
@attribute 'el-salvador-aid' 4
@attribute 'religious-groups-in-schools' 5
@attribute 'anti-satellite-test-ban' 6
@attribute 'aid-to-nicaraguan-contras' 7
@attribute 'mx-missile' 8
@attribute 'immigration' 9
@attribute 'synfuels-corporation-cutback' 10
@attribute 'education-spending' 11
@attribute 'superfund-right-to-sue' 12
@attribute 'crime' 13
@attribute 'duty-free-exports' 14
@attribute 'export-administration-act-south-africa' 15
@attribute 'democrat' 16
@attribute 'republican' 17
1 3 4 5 9 11 12 13 15 17
1 3 4 5 11 12 13 17
1 2 4 5 10 12 13 16
1 2 5 10 12 15 16
0 1 2 4 5 10 12 13 14 15 16
1 2 4 5 12 13 14 15 16
1 3 4 5 13 14 15 16
1 3 4 5 12 13 15 17
1 3 4 5 11 12 13 15 17
0 1 2 6 7 8 16
```

@attribute 이름 ID

19번째 라인 이후→ 데이터
Space하나로 구분되어 있음

17,24   꼭대기