Data Mining Assignment 1

Vedic Sharma (2014B4A7424H)

Mudit Pandey (2014A7PS017H)

**Language used**: C++

**Pre-processing:**

We process the data as follows.

1. Treat each attribute answer (yes / no) as an item
2. Treat each candidate class (republican / democrat) as an item
3. For each candidate, consider a transaction
   1. If the candidate answers a question yes, then an item corresponding to that question answered yes is inserted in the transaction.
   2. Similarly, if the candidate answers no to that question.
   3. If the candidate doesn’t answer, then no item is inserted in the transaction

Here’s the key for each of the possible answers / class of each candidate

1. handicapped-infants=n
2. handicapped-infants=y
3. water-project-cost-sharing=n
4. water-project-cost-sharing=y
5. adoption-of-the-budget-resolution=n
6. adoption-of-the-budget-resolution=y
7. physician-fee-freeze=n
8. physician-fee-freeze=y
9. el-salvador-aid=n
10. el-salvador-aid=y
11. religious-groups-in-schools=n
12. religious-groups-in-schools=y
13. anti-satellite-test-ban=n
14. anti-satellite-test-ban=y
15. aid-to-nicaraguan-contras=n
16. aid-to-nicaraguan-contras=y
17. mx-missile=n
18. mx-missile=y
19. immigration=n
20. immigration=y
21. synfuels-corporation-cutback=n
22. synfuels-corporation-cutback=y
23. education-spending=n
24. education-spending=y
25. superfund-right-to-sue=n
26. superfund-right-to-sue=y
27. crime=n
28. crime=y
29. duty-free-exports=n
30. duty-free-exports=y
31. export-administration-act-south-africa=n
32. export-administration-act-south-africa=y
33. Class=republican
34. Class=democrat

A republican candidate who only answered to spending for education will have a transaction basket like

From here we can apply Apriori algorithm to find interesting rules

**Compilation Steps:**