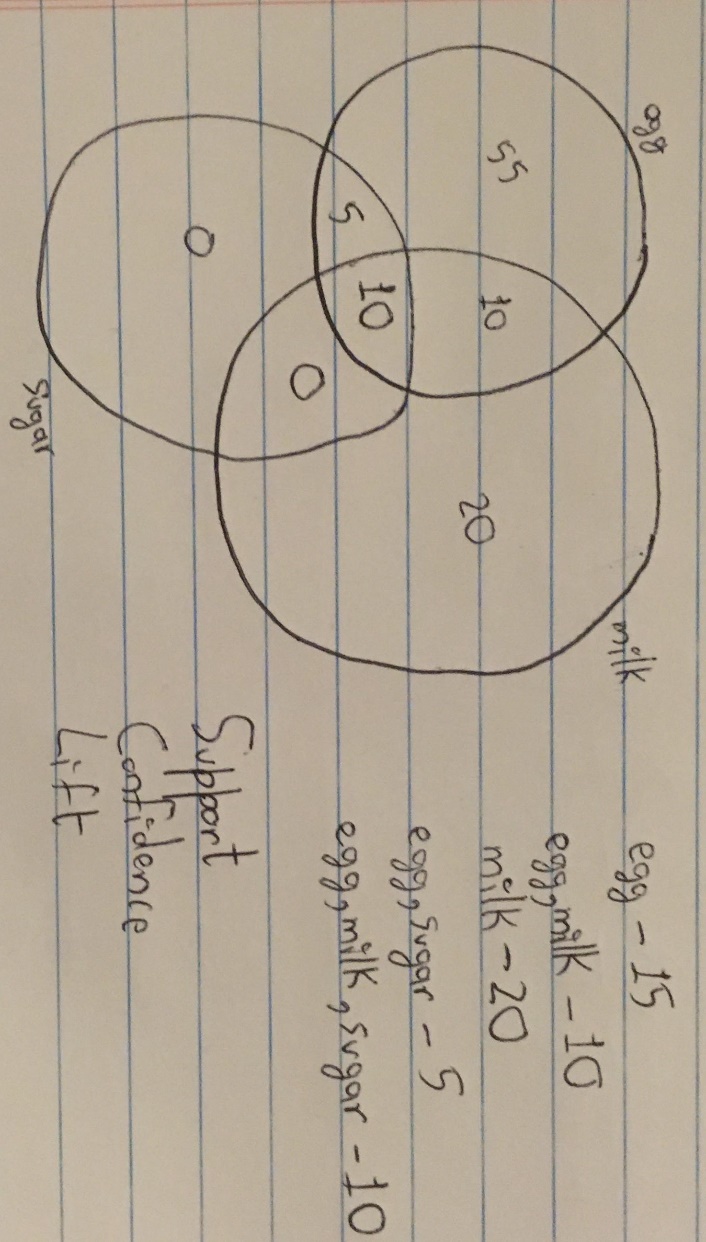
Assignment 4

273 Business Intelligence for Analytical Decisions

This assignment must be completed individually. Submit Word file to online drop box. Write your name in the Word file.

**Question 1.** George, the new manager of supermarket Ralph is trying to understand association rule with your help. He provides the following information. In April 2005, 80 transactions contain eggs; 40 transactions contain milk and 15 transactions contain sugar. Moreover, 10 transactions contain all three items (egg, milk and sugar) altogether; 20 transactions contain egg and milk together; 15 transactions contain egg and sugar together and 10 transactions contain milk and sugar together. (Note: these transactions may overlap, e.g. the 15 transactions with both egg and sugar are among the 80 transactions with egg.) George got a rule as follows: egg 🡪 milk with support 20%. Now he asks you to show him the support, confidence and lift for the rule: {egg, milk} 🡪 sugar. (Note: You may not need all the numbers provided here. This question is used to test your understanding about the concept of support, confidence and lift. You should be able to get the answer with a few simple calculations).

**Answer 1.**



Support for the rule is 0.1

Confidence for the rule is 0.5

Lift for the rule is 3.3333

**Question 2.**

Take the bank-data-final.arff file posted online, and perform association rule analysis using WEKA to answer the following questions. Change options for Apriori to obtain a longer list of rules as follows: change minMetric to 0.8 and change numRules to 100. Run Apriori. Copy resulting rules to MS Word so you can search for rules about pep and interpret those rules.

1). What types of customers have a higher chance of buying a personal equity plan (pep=YES)? What types of customers have a lower chance of buying a personal equity plan?

2). Please examine the results and identify two interesting rules and explain why you think they are interesting.

Below is the description of the attributes in the data.

|  |  |
| --- | --- |
| **age** | age of customer in years |
| **sex** | MALE / FEMALE |
| **region** | inner\_city/rural/suburban/town |
| **income** | income of customer |
| **married** | is the customer married (YES/NO) |
| **children** | number of children |
| **car** | does the customer own a car (YES/NO) |
| **save\_acct** | does the customer have a saving account (YES/NO) |
| **current\_acct** | does the customer have a current account (YES/NO) |
| **mortgage** | does the customer have a mortgage (YES/NO) |
| **pep** | did the customer buy a PEP (Personal Equity Plan) after the last mailing (YES/NO) |

**Answer 2.**

**Rules found:**

1. income=43759\_max 80 ==> save\_act=YES 80 <conf:(1)> lift:(1.45) lev:(0.04) [24] conv:(24.8)

2. age=52\_max income=43759\_max 76 ==> save\_act=YES 76 <conf:(1)> lift:(1.45) lev:(0.04) [23] conv:(23.56)

3. income=43759\_max current\_act=YES 63 ==> save\_act=YES 63 <conf:(1)> lift:(1.45) lev:(0.03) [19] conv:(19.53)

4. age=52\_max income=43759\_max current\_act=YES 61 ==> save\_act=YES 61 <conf:(1)> lift:(1.45) lev:(0.03) [18] conv:(18.91)

5. children=0 save\_act=YES mortgage=NO pep=NO 74 ==> married=YES 73 <conf:(0.99)> lift:(1.49) lev:(0.04) [24] conv:(12.58)

6. sex=FEMALE children=0 mortgage=NO pep=NO 64 ==> married=YES 63 <conf:(0.98)> lift:(1.49) lev:(0.03) [20] conv:(10.88)

7. children=0 current\_act=YES mortgage=NO pep=NO 82 ==> married=YES 80 <conf:(0.98)> lift:(1.48) lev:(0.04) [25] conv:(9.29)

8. children=0 mortgage=NO pep=NO 107 ==> married=YES 104 <conf:(0.97)> lift:(1.47) lev:(0.06) [33] conv:(9.1)

9. income=43759\_max current\_act=YES 63 ==> age=52\_max 61 <conf:(0.97)> lift:(3.04) lev:(0.07) [40] conv:(14.31)

10. income=43759\_max save\_act=YES current\_act=YES 63 ==> age=52\_max 61 <conf:(0.97)> lift:(3.04) lev:(0.07) [40] conv:(14.31)

11. income=43759\_max current\_act=YES 63 ==> age=52\_max save\_act=YES 61 <conf:(0.97)> lift:(3.85) lev:(0.08) [45] conv:(15.72)

12. children=0 car=NO mortgage=NO pep=NO 62 ==> married=YES 60 <conf:(0.97)> lift:(1.47) lev:(0.03) [19] conv:(7.03)

13. age=0\_34 married=YES car=NO 69 ==> income=0\_24386 66 <conf:(0.96)> lift:(2.01) lev:(0.06) [33] conv:(9.06)

14. income=43759\_max 80 ==> age=52\_max 76 <conf:(0.95)> lift:(2.98) lev:(0.08) [50] conv:(10.91)

15. income=43759\_max save\_act=YES 80 ==> age=52\_max 76 <conf:(0.95)> lift:(2.98) lev:(0.08) [50] conv:(10.91)

16. income=43759\_max 80 ==> age=52\_max save\_act=YES 76 <conf:(0.95)> lift:(3.77) lev:(0.09) [55] conv:(11.97)

17. age=0\_34 car=NO mortgage=NO 68 ==> income=0\_24386 64 <conf:(0.94)> lift:(1.98) lev:(0.05) [31] conv:(7.14)

18. age=0\_34 region=INNER\_CITY married=YES 64 ==> income=0\_24386 60 <conf:(0.94)> lift:(1.97) lev:(0.05) [29] conv:(6.72)

19. age=0\_34 car=NO save\_act=YES 64 ==> income=0\_24386 60 <conf:(0.94)> lift:(1.97) lev:(0.05) [29] conv:(6.72)

20. age=0\_34 save\_act=YES current\_act=YES pep=NO 64 ==> income=0\_24386 60 <conf:(0.94)> lift:(1.97) lev:(0.05) [29] conv:(6.72)

21. age=0\_34 car=NO 107 ==> income=0\_24386 100 <conf:(0.93)> lift:(1.97) lev:(0.08) [49] conv:(7.02)

22. age=0\_34 car=NO current\_act=YES 86 ==> income=0\_24386 80 <conf:(0.93)> lift:(1.96) lev:(0.07) [39] conv:(6.45)

23. age=0\_34 current\_act=YES mortgage=NO 95 ==> income=0\_24386 88 <conf:(0.93)> lift:(1.95) lev:(0.07) [42] conv:(6.23)

24. age=0\_34 married=YES current\_act=YES mortgage=NO 67 ==> income=0\_24386 62 <conf:(0.93)> lift:(1.95) lev:(0.05) [30] conv:(5.86)

25. age=0\_34 car=NO pep=NO 66 ==> income=0\_24386 61 <conf:(0.92)> lift:(1.95) lev:(0.05) [29] conv:(5.77)

26. age=0\_34 married=YES current\_act=YES pep=NO 66 ==> income=0\_24386 61 <conf:(0.92)> lift:(1.95) lev:(0.05) [29] conv:(5.77)

27. married=YES children=0 save\_act=YES current\_act=YES 87 ==> pep=NO 80 <conf:(0.92)> lift:(1.69) lev:(0.05) [32] conv:(4.97)

28. age=0\_34 current\_act=YES pep=NO 95 ==> income=0\_24386 87 <conf:(0.92)> lift:(1.93) lev:(0.07) [41] conv:(5.54)

29. age=0\_34 save\_act=YES pep=NO 82 ==> income=0\_24386 75 <conf:(0.91)> lift:(1.93) lev:(0.06) [36] conv:(5.38)

30. age=0\_34 children=0 81 ==> income=0\_24386 74 <conf:(0.91)> lift:(1.92) lev:(0.06) [35] conv:(5.32)

31. married=YES children=0 save\_act=YES mortgage=NO 80 ==> pep=NO 73 <conf:(0.91)> lift:(1.68) lev:(0.05) [29] conv:(4.57)

32. married=YES children=0 current\_act=YES mortgage=NO 88 ==> pep=NO 80 <conf:(0.91)> lift:(1.67) lev:(0.05) [32] conv:(4.47)

33. age=0\_34 married=YES current\_act=YES 98 ==> income=0\_24386 89 <conf:(0.91)> lift:(1.91) lev:(0.07) [42] conv:(5.15)

34. age=0\_34 sex=FEMALE current\_act=YES 75 ==> income=0\_24386 68 <conf:(0.91)> lift:(1.91) lev:(0.05) [32] conv:(4.92)

35. age=0\_34 save\_act=YES current\_act=YES 95 ==> income=0\_24386 86 <conf:(0.91)> lift:(1.91) lev:(0.07) [40] conv:(4.99)

36. age=0\_34 mortgage=NO 125 ==> income=0\_24386 113 <conf:(0.9)> lift:(1.9) lev:(0.09) [53] conv:(5.05)

37. age=0\_34 current\_act=YES 153 ==> income=0\_24386 138 <conf:(0.9)> lift:(1.9) lev:(0.11) [65] conv:(5.02)

38. age=0\_34 sex=MALE 102 ==> income=0\_24386 92 <conf:(0.9)> lift:(1.9) lev:(0.07) [43] conv:(4.87)

39. age=0\_34 save\_act=YES mortgage=NO 80 ==> income=0\_24386 72 <conf:(0.9)> lift:(1.89) lev:(0.06) [34] conv:(4.67)

40. age=0\_34 mortgage=NO pep=NO 80 ==> income=0\_24386 72 <conf:(0.9)> lift:(1.89) lev:(0.06) [34] conv:(4.67)

41. sex=FEMALE married=YES children=0 mortgage=NO 70 ==> pep=NO 63 <conf:(0.9)> lift:(1.66) lev:(0.04) [24] conv:(4)

42. married=YES children=0 save\_act=YES 119 ==> pep=NO 107 <conf:(0.9)> lift:(1.65) lev:(0.07) [42] conv:(4.18)

43. age=0\_34 married=YES mortgage=NO 89 ==> income=0\_24386 80 <conf:(0.9)> lift:(1.89) lev:(0.06) [37] conv:(4.67)

44. age=0\_34 married=YES pep=NO 88 ==> income=0\_24386 79 <conf:(0.9)> lift:(1.89) lev:(0.06) [37] conv:(4.62)

45. age=0\_34 sex=MALE current\_act=YES 78 ==> income=0\_24386 70 <conf:(0.9)> lift:(1.89) lev:(0.05) [32] conv:(4.55)

46. age=0\_34 region=INNER\_CITY 97 ==> income=0\_24386 87 <conf:(0.9)> lift:(1.89) lev:(0.07) [40] conv:(4.63)

47. married=YES children=0 mortgage=NO 116 ==> pep=NO 104 <conf:(0.9)> lift:(1.65) lev:(0.07) [40] conv:(4.07)

48. age=0\_34 married=NO 67 ==> income=0\_24386 60 <conf:(0.9)> lift:(1.89) lev:(0.05) [28] conv:(4.4)

49. married=YES children=0 car=NO mortgage=NO 67 ==> pep=NO 60 <conf:(0.9)> lift:(1.65) lev:(0.04) [23] conv:(3.82)

50. age=0\_34 pep=NO 124 ==> income=0\_24386 111 <conf:(0.9)> lift:(1.88) lev:(0.09) [52] conv:(4.65)

51. age=0\_34 save\_act=NO 76 ==> income=0\_24386 68 <conf:(0.89)> lift:(1.88) lev:(0.05) [31] conv:(4.43)

52. age=0\_34 region=INNER\_CITY current\_act=YES 76 ==> income=0\_24386 68 <conf:(0.89)> lift:(1.88) lev:(0.05) [31] conv:(4.43)

53. age=0\_34 195 ==> income=0\_24386 174 <conf:(0.89)> lift:(1.88) lev:(0.14) [81] conv:(4.65)

54. age=0\_34 save\_act=YES 119 ==> income=0\_24386 106 <conf:(0.89)> lift:(1.88) lev:(0.08) [49] conv:(4.46)

55. age=0\_34 married=YES 128 ==> income=0\_24386 114 <conf:(0.89)> lift:(1.88) lev:(0.09) [53] conv:(4.48)

56. age=0\_34 pep=YES 71 ==> income=0\_24386 63 <conf:(0.89)> lift:(1.87) lev:(0.05) [29] conv:(4.14)

57. age=0\_34 married=YES save\_act=YES 77 ==> income=0\_24386 68 <conf:(0.88)> lift:(1.86) lev:(0.05) [31] conv:(4.04)

58. age=52\_max married=YES car=YES 68 ==> save\_act=YES 60 <conf:(0.88)> lift:(1.28) lev:(0.02) [13] conv:(2.34)

59. age=0\_34 sex=FEMALE 93 ==> income=0\_24386 82 <conf:(0.88)> lift:(1.86) lev:(0.06) [37] conv:(4.07)

60. children=0 car=NO pep=NO 91 ==> married=YES 80 <conf:(0.88)> lift:(1.33) lev:(0.03) [19] conv:(2.58)

61. region=INNER\_CITY children=0 pep=NO 73 ==> married=YES 64 <conf:(0.88)> lift:(1.33) lev:(0.03) [15] conv:(2.48)

62. age=0\_34 mortgage=YES 70 ==> income=0\_24386 61 <conf:(0.87)> lift:(1.83) lev:(0.05) [27] conv:(3.67)

63. children=0 car=NO current\_act=YES pep=NO 69 ==> married=YES 60 <conf:(0.87)> lift:(1.32) lev:(0.02) [14] conv:(2.35)

64. sex=FEMALE children=0 pep=NO 90 ==> married=YES 78 <conf:(0.87)> lift:(1.31) lev:(0.03) [18] conv:(2.35)

65. car=NO mortgage=NO pep=YES 89 ==> current\_act=YES 77 <conf:(0.87)> lift:(1.14) lev:(0.02) [9] conv:(1.65)

66. children=1 save\_act=YES current\_act=YES 73 ==> pep=YES 63 <conf:(0.86)> lift:(1.89) lev:(0.05) [29] conv:(3.61)

67. sex=FEMALE children=0 current\_act=YES pep=NO 70 ==> married=YES 60 <conf:(0.86)> lift:(1.3) lev:(0.02) [13] conv:(2.16)

68. car=YES save\_act=YES mortgage=NO pep=NO 74 ==> married=YES 63 <conf:(0.85)> lift:(1.29) lev:(0.02) [14] conv:(2.1)

69. region=INNER\_CITY current\_act=YES mortgage=NO pep=NO 78 ==> married=YES 66 <conf:(0.85)> lift:(1.28) lev:(0.02) [14] conv:(2.04)

70. children=1 mortgage=NO 84 ==> pep=YES 71 <conf:(0.85)> lift:(1.85) lev:(0.05) [32] conv:(3.26)

71. save\_act=YES mortgage=NO pep=NO 142 ==> married=YES 120 <conf:(0.85)> lift:(1.28) lev:(0.04) [26] conv:(2.1)

72. income=0\_24386 children=0 pep=NO 71 ==> married=YES 60 <conf:(0.85)> lift:(1.28) lev:(0.02) [13] conv:(2.01)

73. sex=FEMALE married=YES children=0 current\_act=YES 71 ==> pep=NO 60 <conf:(0.85)> lift:(1.56) lev:(0.04) [21] conv:(2.7)

74. children=0 pep=NO 167 ==> married=YES 141 <conf:(0.84)> lift:(1.28) lev:(0.05) [30] conv:(2.1)

75. age=52\_max pep=NO 83 ==> save\_act=YES 70 <conf:(0.84)> lift:(1.22) lev:(0.02) [12] conv:(1.84)

76. save\_act=YES current\_act=YES mortgage=NO pep=NO 108 ==> married=YES 91 <conf:(0.84)> lift:(1.28) lev:(0.03) [19] conv:(2.04)

77. children=1 save\_act=YES 95 ==> pep=YES 80 <conf:(0.84)> lift:(1.84) lev:(0.06) [36] conv:(3.23)

78. children=0 car=YES pep=NO 76 ==> save\_act=YES 64 <conf:(0.84)> lift:(1.22) lev:(0.02) [11] conv:(1.81)

79. age=0\_34 car=YES 88 ==> income=0\_24386 74 <conf:(0.84)> lift:(1.77) lev:(0.05) [32] conv:(3.08)

80. age=52\_max sex=MALE 80 ==> save\_act=YES 67 <conf:(0.84)> lift:(1.21) lev:(0.02) [11] conv:(1.77)

81. region=INNER\_CITY children=0 pep=NO 73 ==> current\_act=YES 61 <conf:(0.84)> lift:(1.1) lev:(0.01) [5] conv:(1.36)

82. region=INNER\_CITY mortgage=NO pep=NO 96 ==> married=YES 80 <conf:(0.83)> lift:(1.26) lev:(0.03) [16] conv:(1.92)

83. married=NO car=NO save\_act=YES 72 ==> current\_act=YES 60 <conf:(0.83)> lift:(1.1) lev:(0.01) [5] conv:(1.34)

84. children=1 current\_act=YES 101 ==> pep=YES 84 <conf:(0.83)> lift:(1.82) lev:(0.06) [37] conv:(3.05)

85. married=YES children=1 89 ==> pep=YES 74 <conf:(0.83)> lift:(1.82) lev:(0.06) [33] conv:(3.02)

86. sex=FEMALE region=INNER\_CITY pep=NO 77 ==> current\_act=YES 64 <conf:(0.83)> lift:(1.1) lev:(0.01) [5] conv:(1.33)

87. married=NO save\_act=YES pep=YES 77 ==> current\_act=YES 64 <conf:(0.83)> lift:(1.1) lev:(0.01) [5] conv:(1.33)

88. married=NO save\_act=YES pep=YES 77 ==> mortgage=NO 64 <conf:(0.83)> lift:(1.28) lev:(0.02) [13] conv:(1.92)

89. sex=FEMALE save\_act=YES mortgage=NO pep=NO 77 ==> married=YES 64 <conf:(0.83)> lift:(1.26) lev:(0.02) [13] conv:(1.87)

90. sex=FEMALE married=YES children=0 94 ==> pep=NO 78 <conf:(0.83)> lift:(1.53) lev:(0.04) [26] conv:(2.53)

91. sex=MALE mortgage=NO pep=NO 94 ==> married=YES 78 <conf:(0.83)> lift:(1.26) lev:(0.03) [15] conv:(1.88)

92. age=0\_34 sex=FEMALE income=0\_24386 82 ==> current\_act=YES 68 <conf:(0.83)> lift:(1.09) lev:(0.01) [5] conv:(1.32)

93. income=24387\_43758 mortgage=NO pep=NO 76 ==> married=YES 63 <conf:(0.83)> lift:(1.26) lev:(0.02) [12] conv:(1.85)

94. age=52\_max married=YES 128 ==> save\_act=YES 106 <conf:(0.83)> lift:(1.2) lev:(0.03) [17] conv:(1.73)

95. children=0 current\_act=YES pep=NO 127 ==> married=YES 105 <conf:(0.83)> lift:(1.25) lev:(0.04) [21] conv:(1.88)

96. sex=FEMALE region=INNER\_CITY save\_act=YES 86 ==> current\_act=YES 71 <conf:(0.83)> lift:(1.09) lev:(0.01) [5] conv:(1.3)

97. current\_act=YES mortgage=YES pep=NO 86 ==> save\_act=YES 71 <conf:(0.83)> lift:(1.2) lev:(0.02) [11] conv:(1.67)

98. region=INNER\_CITY married=YES mortgage=NO pep=NO 80 ==> current\_act=YES 66 <conf:(0.82)> lift:(1.09) lev:(0.01) [5] conv:(1.29)

99. married=NO save\_act=YES 137 ==> current\_act=YES 113 <conf:(0.82)> lift:(1.09) lev:(0.02) [9] conv:(1.32)

100. region=INNER\_CITY children=0 save\_act=YES 74 ==> current\_act=YES 61 <conf:(0.82)> lift:(1.09) lev:(0.01) [4] conv:(1.28)

**Rules with pep=NO :-**

27. married=YES children=0 save\_act=YES current\_act=YES 87 ==> pep=NO 80 <conf:(0.92)> lift:(1.69) lev:(0.05) [32] conv:(4.97)

31. married=YES children=0 save\_act=YES mortgage=NO 80 ==> pep=NO 73 <conf:(0.91)> lift:(1.68) lev:(0.05) [29] conv:(4.57)

32. married=YES children=0 current\_act=YES mortgage=NO 88 ==> pep=NO 80 <conf:(0.91)> lift:(1.67) lev:(0.05) [32] conv:(4.47)

41. sex=FEMALE married=YES children=0 mortgage=NO 70 ==> pep=NO 63 <conf:(0.9)> lift:(1.66) lev:(0.04) [24] conv:(4)

42. married=YES children=0 save\_act=YES 119 ==> pep=NO 107 <conf:(0.9)> lift:(1.65) lev:(0.07) [42] conv:(4.18)

47. married=YES children=0 mortgage=NO 116 ==> pep=NO 104 <conf:(0.9)> lift:(1.65) lev:(0.07) [40] conv:(4.07)

49. married=YES children=0 car=NO mortgage=NO 67 ==> pep=NO 60 <conf:(0.9)> lift:(1.65) lev:(0.04) [23] conv:(3.82)

73. sex=FEMALE married=YES children=0 current\_act=YES 71 ==> pep=NO 60 <conf:(0.85)> lift:(1.56) lev:(0.04) [21] conv:(2.7)

90. sex=FEMALE married=YES children=0 94 ==> pep=NO 78 <conf:(0.83)> lift:(1.53) lev:(0.04) [26] conv:(2.53)

**It looks like if children=0 and are married=YES, very often pep is also equal to NO. Also it seems more likely that married women have pep=NO**

**Rules with pep= YES :-**

66. children=1 save\_act=YES current\_act=YES 73 ==> pep=YES 63 <conf:(0.86)> lift:(1.89) lev:(0.05) [29] conv:(3.61)

70. children=1 mortgage=NO 84 ==> pep=YES 71 <conf:(0.85)> lift:(1.85) lev:(0.05) [32] conv:(3.26)

77. children=1 save\_act=YES 95 ==> pep=YES 80 <conf:(0.84)> lift:(1.84) lev:(0.06) [36] conv:(3.23)

84. children=1 current\_act=YES 101 ==> pep=YES 84 <conf:(0.83)> lift:(1.82) lev:(0.06) [37] conv:(3.05)

85. married=YES children=1 89 ==> pep=YES 74 <conf:(0.83)> lift:(1.82) lev:(0.06) [33] conv:(3.02)

**From these rules it looks like children=1 is a good predictor of pep=YES. Also it looks like if people have a savings account or a current account and no mortgage there’s a higher probability of pep=YES.**

**2)** The 2 interesting rules are as follows:-

1. income=43759\_max 80 ==> save\_act=YES 80 <conf:(1)> lift:(1.45) lev:(0.04) [24] conv:(24.8)

This shows that people in the highest income group all seem to prefer savings accounts. This means that people who are rich all believe in the idea of saving money.

2. income=43759\_max 80 ==> age=52\_max 76 <conf:(0.95)> lift:(2.98) lev:(0.08) [50] conv:(10.91)

This shows that belonging to the highest income group is a good predictor of the individual belonging to the highest age group.