"session 11 assignment2 " Ashish Saxena

output:

word\_document: default html\_document: df\_print: paged

--- ```{r}

#Variable Description

#age: age of client

#job : type of job

#marital : marital status

#education: highest educational achievement

#default: has credit in default?

#housing: has housing loan?

#loan: has personal loan?

#contact: contact communication type

#month: last contact month of year

#day\_of\_week: last contact day of the week

#duration: last contact duration, in seconds

#campaign: number of contacts performed during this campaign and for this client

#pdays: number of days that passed by after the client was last contacted from a previous campaign

(999 means client was not previously contacted)

#previous: number of contacts performed before this campaign and for this client

#poutcome: outcome of the previous marketing campaign

#emp.var.rate: employment variation rate - quarterly indicator

#cons.price.idx: consumer price index - monthly indicator

#cons.conf.idx: consumer confidence index - monthly indicator

#euribor3m: euribor 3 month rate - daily indicator #nr.employed: number of employees - quarterly indicator

#y - has the client subscribed a term deposit?

## The data set can be obtained from http://archive.ics.uci.edu/ml/datasets/Bank+Marketing

## DATASET UNDERSTANDING

library(readr) bank\_full <- read\_delim("C:/Users/Seshan/Desktop/Bank/bank-full.csv",

";", escape\_double = FALSE, trim\_ws = TRUE)

#Lets look at dataset and generate initial understanding about the column types str(bank\_full)

#A deep check for NA in a particular column let say age if(length(which(is.na(bank\_full$age)==TRUE)>0)){ print("Missing Value found in the specified column")

} else print("All okay: No Missing Value found in the specified column")

# Check another example say

if(length(which(is.na(bank\_full$campaign)==TRUE)>0)){print("Missing Value found in the specified column")} else print("All okay: No Missing Value found in the specified column")

head(bank\_full) ## Displays first 6 rows for each variable str(bank\_full) ## Describes each variables summary(bank\_full) ## Provides basic statistical information of each variable

## DATA EXPLORATION - Check for Missing Data

## Option 1 is.na(bank\_full) ## Displays True for a missing value

## Since it is a large dataset, graphical display of missing values will prove to be easier

##Option 2 require(Amelia) missmap(bank\_full,main="Missing Data - Bank ", col=c("red","grey"),legend=FALSE)

## No red colour stripes are visible. hence no missing values.

summary(bank\_full) ## displays missing values if any under every variable

#The Pearson’s chi-squared test of independence is one of the most basic and common hypothesis tests in the statistical analysis of categorical data. It is a significance test. Given two categorical random variables, X and Y, the chi-squared test of independence determines whether or not there exists a statistical dependence between them. Formally, it is a hypothesis test. The chi-squared test assumes a null hypothesis and an alternate hypothesis. The general practice is, if the p-value that comes out in the result is less than a pre-determined significance level, which is 0.05 usually, then we reject the null hypothesis.

#H0: The The two variables are independent

#H1: The The two variables are dependent

#The null hypothesis of the chi-squared test is that the two variables are independent and the alternate hypothesis is that they are related.

#To establish that two categorical variables (or predictors) are dependent, the chi-squared statistic must have a certain cutoff. This cutoff increases as the number of classes within the variable (or predictor) increases.

#i. Pearson’s chi-squared test of independence (significance test)

Is there any association between Job and default?

with(bank\_full, chisq.test( job, default)) with(bank\_full, table( job, default) )

# OR with(bank\_full, prop.table(table( job,default)))

#Pearson's Chi-squared test

|  |
| --- |
| data: job and default  X-squared = 60.343, df = 11, p-value = 8.008e-09  default job no yes admin. 5097 74 blue-collar 9531 201 entrepreneur 1432 55 housemaid 1218 22 management 9294 164 retired 2238 26 self-employed 1546 33 services 4079 75 student 935 3 technician 7467 130 unemployed 1273 30 unknown 286 2 default  job no yes admin. 1.127381e-01 1.636770e-03 blue-collar 2.108115e-01 4.445821e-03 entrepreneur 3.167371e-02 1.216518e-03 housemaid 2.694035e-02 4.866072e-04 management 2.055694e-01 3.627436e-03 retired 4.950123e-02 5.750813e-04 self-employed 3.419522e-02 7.299109e-04 services 9.022141e-02 1.658888e-03 student 2.068081e-02 6.635553e-05 technician 1.651589e-01 2.875406e-03 |

p-value = 8.008e-09

#Pearson's Chi-squared test

#since the p-value is < 2.2e-16 is less than the cu$t-off value of 0.05, we can reject the null hypothesis in favor of alternative hypothesis and conclude, that the variables,( job & default- p-value = 8.008e-09) are dependent to each other.

b. Is there any significant difference in duration of last call between people having housing loan or not?

with(bank\_additional\_full, chisq.test(duration,housing)) with(bank\_additional\_full, table( duration,housing) )

# OR with(bank\_additional\_full, prop.table(table(duration, housing)))

#data: duration and housing

#X-squared = 3162.3, df = 3086, p-value = 0.1657

#P value is above 0.05#

|  |
| --- |
| Chi-squared approximation may be incorrect  Pearson's Chi-squared test    data: duration and housing  X-squared = 3162.3, df = 3086, p-value = 0.1657    housing duration no unknown yes 0 1 0 3   1. 2 0 1 2. 1 0 0 3. 2 0 1 4. 2 0 10 5. 16 0 14 6. 13 0 24 7. 22 1 31 8. 27 3 39 9. 33 2 42 10. 36 1 35 11. 34 2 45 12. 24 2 39 13. 44 0 33 14. 25 2 43 15. 34 1 33 16. 35 1 44 17. 34 1 41 |

1. 43 1 40
2. 24 3 34
3. 29 0 32
4. 30 1 42
5. 35 2 39
6. 21 5 39 24 30 2 32 25 25 2 38
7. 23 1 37
8. 31 1 33
9. 25 0 24
10. 31 0 36
11. 17 2 35
12. 28 3 32
13. 20 1 21
14. 19 0 27
15. 31 3 34
16. 34 3 30
17. 42 1 39
18. 32 2 36
19. 26 2 32
20. 32 0 39
21. 26 2 38
22. 33 2 55
23. 35 1 45
24. 31 2 46
25. 33 3 49
26. 28 3 39
27. 37 3 37
28. 25 3 45
29. 43 3 46
30. 49 3 44
31. 41 3 52
32. 41 1 59
33. 48 4 50
34. 49 0 44
35. 46 2 58
36. 48 4 69
37. 51 2 57
38. 41 3 65
39. 44 0 66
40. 53 6 72
41. 47 1 57
42. 49 6 68
43. 49 0 59
44. 55 1 71
45. 63 2 74
46. 57 2 64
47. 57 3 48
48. 65 6 76
49. 68 5 67
50. 63 3 64
51. 59 3 67
52. 76 9 67
53. 63 7 91
54. 74 8 85
55. 64 3 69
56. 65 0 73 76 67 0 86
57. 56 3 87
58. 56 3 86
59. 71 0 67
60. 69 3 76
61. 69 6 72
62. 57 3 94
63. 59 6 84
64. 66 6 59
65. 85 1 84
66. 55 3 70
67. 70 3 89
68. 64 5 81
69. 68 3 82
70. 73 3 94
71. 75 2 70
72. 72 7 77
73. 76 6 69
74. 66 2 68
75. 63 5 80
76. 86 2 67
77. 79 3 76
78. 70 2 72
79. 55 3 71
80. 61 2 82
81. 62 3 80
82. 64 4 73
83. 68 1 78
84. 70 10 81
85. 62 4 79
86. 67 4 88
87. 60 6 84
88. 62 2 73
89. 66 1 91
90. 71 3 68
91. 72 2 86
92. 61 3 80
93. 62 1 79
94. 74 6 76
95. 58 1 76
96. 62 5 59
97. 64 2 68
98. 49 1 81
99. 72 4 79
100. 52 2 68
101. 61 4 76
102. 69 4 84
103. 69 6 74
104. 65 5 94
105. 67 5 80
106. 62 2 88
107. 68 3 75
108. 70 4 73
109. 64 2 70
110. 62 4 76 131 69 5 60

132 56 3 53 133 57 4 82

1. 74 1 58
2. 78 9 69
3. 77 8 83
4. 61 6 56
5. 51 4 64
6. 87 5 63
7. 64 2 79
8. 52 4 62
9. 57 1 58
10. 63 5 65
11. 64 2 71
12. 59 3 68
13. 54 3 45
14. 56 7 68
15. 51 6 66
16. 51 0 68
17. 49 2 82
18. 56 3 70
19. 49 2 64
20. 47 6 67
21. 62 8 66
22. 63 5 68
23. 47 5 76
24. 71 4 70
25. 43 4 82
26. 65 4 71
27. 48 4 85
28. 57 3 69
29. 46 3 57
30. 46 3 64
31. 55 4 71
32. 66 3 67
33. 73 1 68
34. 44 4 57
35. 59 1 71
36. 50 3 53
37. 37 3 60
38. 59 3 53
39. 53 2 68
40. 51 3 71
41. 57 2 63
42. 48 2 65
43. 39 4 58
44. 47 3 56
45. 59 5 63
46. 46 7 55
47. 58 3 56
48. 49 1 63
49. 39 5 57
50. 39 1 50 184 56 3 54 185 56 3 55
51. 37 1 48
52. 54 2 52
53. 58 1 55
54. 63 2 49 190 54 2 40
55. 54 2 53
56. 49 4 51
57. 45 3 58
58. 40 0 54 195 31 3 54
59. 40 0 48
60. 50 0 57
61. 56 2 49
62. 49 2 44
63. 60 1 47
64. 41 2 67
65. 45 4 47
66. 50 3 40
67. 47 2 42
68. 41 0 46
69. 50 1 41
70. 54 2 58
71. 46 3 41
72. 45 2 47
73. 42 2 56
74. 50 1 57
75. 44 3 46
76. 47 3 40
77. 39 2 50
78. 38 4 49
79. 33 3 44
80. 41 2 49
81. 36 1 39
82. 40 0 45
83. 37 1 48
84. 45 1 43
85. 52 2 39
86. 38 0 31
87. 30 1 42
88. 42 3 45
89. 29 3 61
90. 39 4 39
91. 28 1 44
92. 26 3 40
93. 36 6 40
94. 34 0 45
95. 38 1 36
96. 26 1 34
97. 37 4 37
98. 29 3 30
99. 27 1 44
100. 31 1 34
101. 32 1 43
102. 35 2 55
103. 31 4 52
104. 36 1 37
105. 23 3 32
106. 34 0 45
107. 34 0 41
108. 49 2 40
109. 39 3 37 247 31 2 52 248 24 3 43 249 33 2 41
110. 40 1 39
111. 34 1 48
112. 36 1 43
113. 31 1 33
114. 27 4 41
115. 38 1 33
116. 25 1 39
117. 36 3 32
118. 29 1 38
119. 33 2 36
120. 23 0 38
121. 29 1 32
122. 35 3 34
123. 37 1 36
124. 31 0 30
125. 23 0 43
126. 26 0 32
127. 33 1 38
128. 35 4 32
129. 23 0 31
130. 38 1 18
131. 27 5 25
132. 26 5 36
133. 23 0 37
134. 23 1 32
135. 21 1 34
136. 23 2 24
137. 33 1 29
138. 27 1 25
139. 19 2 22
140. 27 0 30
141. 30 0 31
142. 21 4 31
143. 31 1 30
144. 27 0 26
145. 29 1 30
146. 19 0 29
147. 25 0 21
148. 24 2 31
149. 23 2 28
150. 29 1 26
151. 29 2 28
152. 18 0 35
153. 28 0 34
154. 24 1 35
155. 35 4 28
156. 33 1 23
157. 28 3 32
158. 29 1 25
159. 22 1 24
160. 18 0 28
161. 35 2 25
162. 20 1 25
163. 16 0 23
164. 25 0 29
165. 26 1 35
166. 23 0 35
167. 18 0 19
168. 16 4 21
169. 16 1 27
170. 10 0 23
171. 24 5 16
172. 25 1 23

|  |
| --- |
| 1. 20 3 27 2. 27 1 27 3. 18 2 19 4. 23 2 20 5. 21 3 33 6. 18 1 37 7. 28 2 16 8. 16 1 30 9. 14 0 21 10. 31 1 25 11. 20 1 27 12. 20 0 33 13. 17 0 25 14. 19 0 31 15. 24 1 16 16. 19 1 29 17. 25 1 34 18. 10 0 18 19. 18 1 16 20. 27 2 24   [ reached getOption("max.print") -- omitted 1211 rows ] housing  duration no unknown yes 0 2.427892e-05 0.000000e+00 7.283675e-05   1. 4.855783e-05 0.000000e+00 2.427892e-05 2. 2.427892e-05 0.000000e+00 0.000000e+00 3. 4.855783e-05 0.000000e+00 2.427892e-05 4. 4.855783e-05 0.000000e+00 2.427892e-04 5. 3.884627e-04 0.000000e+00 3.399048e-04 6. 3.156259e-04 0.000000e+00 5.826940e-04 7. 5.341362e-04 2.427892e-05 7.526464e-04 8. 6.555307e-04 7.283675e-05 9.468777e-04 9. 8.012042e-04 4.855783e-05 1.019714e-03 10. 8.740410e-04 2.427892e-05 8.497621e-04 11. 8.254832e-04 4.855783e-05 1.092551e-03 12. 5.826940e-04 4.855783e-05 9.468777e-04 13. 1.068272e-03 0.000000e+00 8.012042e-04 14. 6.069729e-04 4.855783e-05 1.043993e-03 15. 8.254832e-04 2.427892e-05 8.012042e-04 |

16 8.497621e-04 2.427892e-05 1.068272e-03 17 8.254832e-04 2.427892e-05 9.954356e-04

1. 1.043993e-03 2.427892e-05 9.711566e-04
2. 5.826940e-04 7.283675e-05 8.254832e-04
3. 7.040886e 04 0.000000e+00 7.769253e-04
4. 7.283675e 04 2.427892e-05 1.019714e-03
5. 8.497621e 04 4.855783e-05 9.468777e-04
6. 5.098572e 04 1.213946e-04 9.468777e-04
7. 7.283675e 04 4.855783e-05 7.769253e-04 25 6.069729e 04 4.855783e 05 9.225988e-04
8. 5.584151e 04 2.427892e 05 8.983199e-04
9. 7.526464e 04 2.427892e 05 8.012042e-04
10. 6.069729e 04 0.000000e+00 5.826940e-04
11. 7.526464e 04 0.000000e+00 8.740410e-04 30 4.127416e 04 4.855783e-05 8.497621e-04
12. 6.798097e-04 7.283675e-05 7.769253e-04
13. 4.855783e-04 2.427892e-05 5.098572e-04
14. 4.612994e-04 0.000000e+00 6.555307e-04
15. 7.526464e-04 7.283675e-05 8.254832e-04
16. 8.254832e-04 7.283675e-05 7.283675e-04
17. 1.019714e-03 2.427892e-05 9.468777e-04
18. 7.769253e-04 4.855783e-05 8.740410e-04
19. 6.312518e-04 4.855783e-05 7.769253e-04
20. 7.769253e-04 0.000000e+00 9.468777e-04
21. 6.312518e-04 4.855783e-05 9.225988e-04
22. 8.012042e-04 4.855783e-05 1.335340e-03
23. 8.497621e-04 2.427892e-05 1.092551e-03
24. 7.526464e-04 4.855783e-05 1.116830e-03
25. 8.012042e-04 7.283675e-05 1.189667e-03
26. 6.798097e-04 7.283675e-05 9.468777e-04
27. 8.983199e-04 7.283675e-05 8.983199e-04
28. 6.069729e-04 7.283675e-05 1.092551e-03
29. 1.043993e-03 7.283675e-05 1.116830e-03
30. 1.189667e-03 7.283675e-05 1.068272e-03
31. 9.954356e-04 7.283675e-05 1.262504e-03
32. 9.954356e-04 2.427892e-05 1.432456e-03
33. 1.165388e-03 9.711566e-05 1.213946e-03
34. 1.189667e-03 0.000000e+00 1.068272e-03
35. 1.116830e-03 4.855783e-05 1.408177e-03
36. 1.165388e-03 9.711566e-05 1.675245e-03
37. 1.238225e-03 4.855783e-05 1.383898e-03
38. 9.954356e-04 7.283675e-05 1.578130e-03
39. 1.068272e-03 0.000000e+00 1.602408e-03
40. 1.286783e-03 1.456735e-04 1.748082e-03
41. 1.141109e-03 2.427892e-05 1.383898e-03
42. 1.189667e-03 1.456735e-04 1.650966e-03
43. 1.189667e-03 0.000000e+00 1.432456e-03
44. 1.335340e-03 2.427892e-05 1.723803e-03
45. 1.529572e-03 4.855783e-05 1.796640e-03
46. 1.383898e-03 4.855783e-05 1.553851e-03
47. 1.383898e-03 7.283675e-05 1.165388e-03
48. 1.578130e-03 1.456735e-04 1.845198e-03
49. 1.650966e-03 1.213946e-04 1.626687e-03
50. 1.529572e-03 7.283675e-05 1.553851e-03
51. 1.432456e-03 7.283675e-05 1.626687e-03
52. 1.845198e-03 2.185102e-04 1.626687e-03
53. 1.529572e-03 1.699524e-04 2.209381e-03
54. 1.796640e-03 1.942313e-04 2.063708e-03
55. 1.553851e-03 7.283675e-05 1.675245e-03
56. 1.578130e-03 0.000000e+00 1.772361e-03
57. 1.626687e-03 0.000000e+00 2.087987e-03
58. 1.359619e 03 7.283675e-05 2.112266e
59. 1.359619e 03 7.283675e-05 2.087987e 79 1.723803e 03 0.000000e+00 1.626687e 80 1.675245e 03 7.283675e-05 1.845198e

81 1.675245e 03 1.456735e 04 1.748082e 82 1.383898e 03 7.283675e 05 2.282218e

1. 1.432456e-03 1.456735e 04 2.039429e
2. 1.602408e-03 1.456735e 04 1.432456e
3. 2.063708e-03 2.427892e-05 2.039429e
4. 1.335340e-03 7.283675e-05 1.699524e
5. 1.699524e-03 7.283675e-05 2.160824e 88 1.553851e-03 1.213946e-04 1.966592e
6. 1.650966e-03 7.283675e-05 1.990871e-03
7. 1.772361e-03 7.283675e-05 2.282218e-03
8. 1.820919e-03 4.855783e-05 1.699524e-03
9. 1.748082e-03 1.699524e-04 1.869477e-03
10. 1.845198e-03 1.456735e-04 1.675245e-03
11. 1.602408e-03 4.855783e-05 1.650966e-03
12. 1.529572e-03 1.213946e-04 1.942313e-03
13. 2.087987e-03 4.855783e-05 1.626687e-03
14. 1.918034e-03 7.283675e-05 1.845198e-03
15. 1.699524e-03 4.855783e-05 1.748082e-03
16. 1.335340e-03 7.283675e-05 1.723803e-03
17. 1.481014e-03 4.855783e-05 1.990871e-03
18. 1.505293e-03 7.283675e-05 1.942313e-03
19. 1.553851e-03 9.711566e-05 1.772361e-03
20. 1.650966e-03 2.427892e-05 1.893755e-03
21. 1.699524e-03 2.427892e-04 1.966592e-03
22. 1.505293e-03 9.711566e-05 1.918034e-03
23. 1.626687e-03 9.711566e-05 2.136545e-03
24. 1.456735e-03 1.456735e-04 2.039429e-03
25. 1.505293e-03 4.855783e-05 1.772361e-03
26. 1.602408e-03 2.427892e-05 2.209381e-03
27. 1.723803e-03 7.283675e-05 1.650966e-03
28. 1.748082e-03 4.855783e-05 2.087987e-03
29. 1.481014e-03 7.283675e-05 1.942313e-03
30. 1.505293e-03 2.427892e-05 1.918034e-03
31. 1.796640e-03 1.456735e-04 1.845198e-03
32. 1.408177e-03 2.427892e-05 1.845198e-03
33. 1.505293e-03 1.213946e-04 1.432456e-03
34. 1.553851e-03 4.855783e-05 1.650966e-03
35. 1.189667e-03 2.427892e-05 1.966592e-03
36. 1.748082e-03 9.711566e-05 1.918034e-03
37. 1.262504e-03 4.855783e-05 1.650966e-03
38. 1.481014e-03 9.711566e-05 1.845198e-03
39. 1.675245e-03 9.711566e-05 2.039429e-03
40. 1.675245e-03 1.456735e-04 1.796640e-03
41. 1.578130e-03 1.213946e-04 2.282218e-03
42. 1.626687e-03 1.213946e-04 1.942313e-03
43. 1.505293e-03 4.855783e-05 2.136545e-03
44. 1.650966e-03 7.283675e-05 1.820919e-03
45. 1.699524e-03 9.711566e-05 1.772361e-03
46. 1.553851e-03 4.855783e-05 1.699524e-03
47. 1.505293e-03 9.711566e-05 1.845198e-03 131 1.675245e-03 1.213946e-04 1.456735e-03
48. 1.359619e-03 7.283675e-05 1.286783e-03
49. 1.383898e-03 9.711566e-05 1.990871e-03
50. 1.796640e 03 2.427892e-05 1.408177e
51. 1.893755e 03 2.185102e-04 1.675245e 136 1.869477e 03 1.942313e-04 2.015150e 137 1.481014e 03 1.456735e-04 1.359619e
52. 1.238225e 03 9.711566e 05 1.553851e
53. 2.112266e 03 1.213946e 04 1.529572e
54. 1.553851e 03 4.855783e 05 1.918034e
55. 1.262504e 03 9.711566e 05 1.505293e 142 1.383898e 03 2.427892e-05 1.408177e
56. 1.529572e 03 1.213946e-04 1.578130e
57. 1.553851e 03 4.855783e-05 1.723803e
58. 1.432456e-03 7.283675e-05 1.650966e
59. 1.311061e-03 7.283675e-05 1.092551e-03
60. 1.359619e-03 1.699524e-04 1.650966e-03
61. 1.238225e-03 1.456735e-04 1.602408e-03
62. 1.238225e-03 0.000000e+00 1.650966e-03
63. 1.189667e-03 4.855783e-05 1.990871e-03
64. 1.359619e-03 7.283675e-05 1.699524e-03
65. 1.189667e-03 4.855783e-05 1.553851e-03
66. 1.141109e-03 1.456735e-04 1.626687e-03
67. 1.505293e-03 1.942313e-04 1.602408e-03
68. 1.529572e-03 1.213946e-04 1.650966e-03
69. 1.141109e-03 1.213946e-04 1.845198e-03
70. 1.723803e-03 9.711566e-05 1.699524e-03
71. 1.043993e-03 9.711566e-05 1.990871e-03
72. 1.578130e-03 9.711566e-05 1.723803e-03
73. 1.165388e-03 9.711566e-05 2.063708e-03
74. 1.383898e-03 7.283675e-05 1.675245e-03
75. 1.116830e-03 7.283675e-05 1.383898e-03
76. 1.116830e-03 7.283675e-05 1.553851e-03
77. 1.335340e-03 9.711566e-05 1.723803e-03
78. 1.602408e-03 7.283675e-05 1.626687e-03
79. 1.772361e-03 2.427892e-05 1.650966e-03
80. 1.068272e-03 9.711566e-05 1.383898e-03
81. 1.432456e-03 2.427892e-05 1.723803e-03
82. 1.213946e-03 7.283675e-05 1.286783e-03
83. 8.983199e-04 7.283675e-05 1.456735e-03
84. 1.432456e-03 7.283675e-05 1.286783e-03
85. 1.286783e-03 4.855783e-05 1.650966e-03
86. 1.238225e-03 7.283675e-05 1.723803e-03
87. 1.383898e-03 4.855783e-05 1.529572e-03
88. 1.165388e-03 4.855783e-05 1.578130e-03
89. 9.468777e-04 9.711566e-05 1.408177e-03
90. 1.141109e-03 7.283675e-05 1.359619e-03
91. 1.432456e-03 1.213946e-04 1.529572e-03
92. 1.116830e-03 1.699524e-04 1.335340e-03
93. 1.408177e-03 7.283675e-05 1.359619e-03
94. 1.189667e-03 2.427892e-05 1.529572e-03
95. 9.468777e-04 1.213946e-04 1.383898e-03
96. 9.468777e-04 2.427892e-05 1.213946e-03
97. 1.359619e-03 7.283675e-05 1.311061e-03
98. 1.359619e-03 7.283675e-05 1.335340e-03
99. 8.983199e-04 2.427892e-05 1.165388e-03
100. 1.311061e-03 4.855783e-05 1.262504e-03
101. 1.408177e-03 2.427892e-05 1.335340e-03
102. 1.529572e-03 4.855783e-05 1.189667e-03
103. 1.311061e-03 4.855783e-05 9.711566e-04
104. 1.311061e-03 4.855783e-05 1.286783e
105. 1.189667e-03 9.711566e-05 1.238225e
106. 1.092551e-03 7.283675e-05 1.408177e 194 9.711566e-04 0.000000e+00 1.311061e 195 7.526464e-04 7.283675e 05 1.311061e 196 9.711566e 04 0.000000e+00 1.165388e
107. 1.213946e 03 0.000000e+00 1.383898e
108. 1.359619e 03 4.855783e-05 1.189667e 199 1.189667e 03 4.855783e-05 1.068272e

200 1.456735e 03 2.427892e-05 1.141109e 201 9.954356e 04 4.855783e-05 1.626687e

1. 1.092551e-03 9.711566e-05 1.141109e
2. 1.213946e-03 7.283675e-05 9.711566e-04
3. 1.141109e-03 4.855783e-05 1.019714e-03
4. 9.954356e-04 0.000000e+00 1.116830e-03
5. 1.213946e-03 2.427892e-05 9.954356e-04
6. 1.311061e-03 4.855783e-05 1.408177e-03
7. 1.116830e-03 7.283675e-05 9.954356e-04
8. 1.092551e-03 4.855783e-05 1.141109e-03
9. 1.019714e-03 4.855783e-05 1.359619e-03
10. 1.213946e-03 2.427892e-05 1.383898e-03
11. 1.068272e-03 7.283675e-05 1.116830e-03
12. 1.141109e-03 7.283675e-05 9.711566e-04
13. 9.468777e-04 4.855783e-05 1.213946e-03
14. 9.225988e-04 9.711566e-05 1.189667e-03
15. 8.012042e-04 7.283675e-05 1.068272e-03
16. 9.954356e-04 4.855783e-05 1.189667e-03
17. 8.740410e-04 2.427892e-05 9.468777e-04
18. 9.711566e-04 0.000000e+00 1.092551e-03
19. 8.983199e-04 2.427892e-05 1.165388e-03
20. 1.092551e-03 2.427892e-05 1.043993e-03
21. 1.262504e-03 4.855783e-05 9.468777e-04
22. 9.225988e-04 0.000000e+00 7.526464e-04
23. 7.283675e-04 2.427892e-05 1.019714e-03
24. 1.019714e-03 7.283675e-05 1.092551e-03
25. 7.040886e-04 7.283675e-05 1.481014e-03
26. 9.468777e-04 9.711566e-05 9.468777e-04
27. 6.798097e-04 2.427892e-05 1.068272e-03
28. 6.312518e-04 7.283675e-05 9.711566e-04
29. 8.740410e-04 1.456735e-04 9.711566e-04
30. 8.254832e-04 0.000000e+00 1.092551e-03
31. 9.225988e-04 2.427892e-05 8.740410e-04
32. 6.312518e-04 2.427892e-05 8.254832e-04
33. 8.983199e-04 9.711566e-05 8.983199e-04
34. 7.040886e-04 7.283675e-05 7.283675e-04
35. 6.555307e-04 2.427892e-05 1.068272e-03
36. 7.526464e-04 2.427892e-05 8.254832e-04
37. 7.769253e-04 2.427892e-05 1.043993e-03
38. 8.497621e-04 4.855783e-05 1.335340e-03
39. 7.526464e-04 9.711566e-05 1.262504e-03
40. 8.740410e-04 2.427892e-05 8.983199e-04
41. 5.584151e-04 7.283675e-05 7.769253e-04
42. 8.254832e-04 0.000000e+00 1.092551e-03
43. 8.254832e-04 0.000000e+00 9.954356e-04
44. 1.189667e-03 4.855783e-05 9.711566e-04
45. 9.468777e-04 7.283675e-05 8.983199e-04
46. 7.526464e-04 4.855783e-05 1.262504e-03
47. 5.826940e-04 7.283675e-05 1.043993e-03
48. 8.012042e-04 4.855783e-05 9.954356e-04
49. 9.711566e-04 2.427892e-05 9.468777e-04
50. 8.254832e-04 2.427892e-05 1.165388e
51. 8.740410e-04 2.427892e 05 1.043993e 253 7.526464e 04 2.427892e-05 8.012042e-04 254 6.555307e 04 9.711566e-05 9.954356e-04
52. 9.225988e 04 2.427892e-05 8.012042e-04
53. 6.069729e 04 2.427892e-05 9.468777e-04
54. 8.740410e 04 7.283675e-05 7.769253e-04
55. 7.040886e 04 2.427892e-05 9.225988e-04
56. 8.012042e-04 4.855783e-05 8.740410e-04
57. 5.584151e-04 0.000000e+00 9.225988e-04
58. 7.040886e-04 2.427892e-05 7.769253e-04
59. 8.497621e-04 7.283675e-05 8.254832e-04
60. 8.983199e-04 2.427892e-05 8.740410e-04
61. 7.526464e-04 0.000000e+00 7.283675e-04
62. 5.584151e-04 0.000000e+00 1.043993e-03
63. 6.312518e-04 0.000000e+00 7.769253e-04
64. 8.012042e-04 2.427892e-05 9.225988e-04
65. 8.497621e-04 9.711566e-05 7.769253e-04
66. 5.584151e-04 0.000000e+00 7.526464e-04
67. 9.225988e-04 2.427892e-05 4.370205e-04
68. 6.555307e-04 1.213946e-04 6.069729e-04
69. 6.312518e-04 1.213946e-04 8.740410e-04
70. 5.584151e-04 0.000000e+00 8.983199e-04
71. 5.584151e-04 2.427892e-05 7.769253e-04
72. 5.098572e-04 2.427892e-05 8.254832e-04
73. 5.584151e-04 4.855783e-05 5.826940e-04
74. 8.012042e-04 2.427892e-05 7.040886e-04
75. 6.555307e-04 2.427892e-05 6.069729e-04
76. 4.612994e-04 4.855783e-05 5.341362e-04
77. 6.555307e-04 0.000000e+00 7.283675e-04
78. 7.283675e-04 0.000000e+00 7.526464e-04
79. 5.098572e-04 9.711566e-05 7.526464e-04
80. 7.526464e-04 2.427892e-05 7.283675e-04
81. 6.555307e-04 0.000000e+00 6.312518e-04
82. 7.040886e-04 2.427892e-05 7.283675e-04
83. 4.612994e-04 0.000000e+00 7.040886e-04
84. 6.069729e-04 0.000000e+00 5.098572e-04
85. 5.826940e-04 4.855783e-05 7.526464e-04
86. 5.584151e-04 4.855783e-05 6.798097e-04
87. 7.040886e-04 2.427892e-05 6.312518e-04
88. 7.040886e-04 4.855783e-05 6.798097e-04
89. 4.370205e-04 0.000000e+00 8.497621e-04
90. 6.798097e-04 0.000000e+00 8.254832e-04
91. 5.826940e-04 2.427892e-05 8.497621e-04
92. 8.497621e-04 9.711566e-05 6.798097e-04
93. 8.012042e-04 2.427892e-05 5.584151e-04
94. 6.798097e-04 7.283675e-05 7.769253e-04
95. 7.040886e-04 2.427892e-05 6.069729e-04
96. 5.341362e-04 2.427892e-05 5.826940e-04
97. 4.370205e-04 0.000000e+00 6.798097e-04
98. 8.497621e-04 4.855783e-05 6.069729e-04
99. 4.855783e-04 2.427892e-05 6.069729e-04
100. 3.884627e-04 0.000000e+00 5.584151e-04
101. 6.069729e-04 0.000000e+00 7.040886e-04
102. 6.312518e-04 2.427892e-05 8.497621e-04
103. 5.584151e-04 0.000000e+00 8.497621e-04
104. 4.370205e-04 0.000000e+00 4.612994e-04
105. 3.884627e-04 9.711566e-05 5.098572e-04
106. 3.884627e-04 2.427892e 05 6.555307e-04
107. 2.427892e 04 0.000000e+00 5.584151e-04 311 5.826940e 04 1.213946e-04 3.884627e-04
108. 6.069729e 04 2.427892e-05 5.584151e-04
109. 4.855783e 04 7.283675e-05 6.555307e-04
110. 6.555307e 04 2.427892e-05 6.555307e-04
111. 4.370205e 04 4.855783e-05 4.612994e-04
112. 5.584151e-04 4.855783e-05 4.855783e-04
113. 5.098572e-04 7.283675e-05 8.012042e-04
114. 4.370205e-04 2.427892e-05 8.983199e-04
115. 6.798097e-04 4.855783e-05 3.884627e-04
116. 3.884627e-04 2.427892e-05 7.283675e-04
117. 3.399048e-04 0.000000e+00 5.098572e-04
118. 7.526464e-04 2.427892e-05 6.069729e-04
119. 4.855783e-04 2.427892e-05 6.555307e-04
120. 4.855783e-04 0.000000e+00 8.012042e-04
121. 4.127416e-04 0.000000e+00 6.069729e-04
122. 4.612994e-04 0.000000e+00 7.526464e-04
123. 5.826940e-04 2.427892e-05 3.884627e-04
124. 4.612994e-04 2.427892e-05 7.040886e-04
125. 6.069729e-04 2.427892e-05 8.254832e-04
126. 2.427892e-04 0.000000e+00 4.370205e-04
127. 4.370205e-04 2.427892e-05 3.884627e-04
128. 6.555307e-04 4.855783e-05 5.826940e-04

[ reached getOption("max.print") -- omitted 1211 rows ]

Is there any association between consumer price index and consumer?

#Is there any association between consumer price index and consumer? with(bank\_additional\_full, chisq.test(cons.price.idx,cons.conf.idx)) with(bank\_additional\_full, table(cons.price.idx,cons.conf.idx))

# OR

with(bank\_additional\_full, prop.table(table(cons.price.idx,cons.conf.idx)))

#p-value < 2.2e-16 and it is very much less than 0.05.we can reject the null hypothesis in favor of alternative hypothesis and conclude, that the variables, (job & Marital-p-value < 2.2e16),(con.price.idx , consumer- are dependent to each other.

Chi-squared approximation may be incorrect

Pearson's Chi-squared test

data: cons.price.idx and cons.conf.idx X-squared = 1029700, df = 625, p-value < 2.2e-16

cons.conf.idx

cons.price.idx -50.8 -50 -49.5 -47.1 -46.2 -45.9 -42.7 -42 -41.8 -40.8 40.4

92.201 0 0 0 0 0 0 0 0 0 0

0

92.379 0 0 0 0 0 0 0 0 0 0 0

92.431 0 0 0 0 0 0 0 0 0 0

0

92.469 0 0 0 0 0 0 0 0 0 0

0

92.649 0 0 0 0 0 0 0 0 0 0

0

92.713 0 0 0 0 0 0 0 0 0 0

0

92.756 0 0 0 0 0 10 0 0 0 0

0

92.843 0 282 0 0 0 0 0 0 0 0 0

92.893 0 0 0 0 5794 0 0 0 0 0 0

92.963 0 0 0 0 0 0 0 0 0 715

0

93.075 0 0 0 2458 0 0 0 0 0 0 0

93.2 0 0 0 0 0 0 0 3616 0 0 0

93.369 0 0 0 0 0 0 0 0 0 0

0

93.444 0 0 0 0 0 0 0 0 0 0 0

93.749 0 0 0 0 0 0 0 0 0 0 0

93.798 0 0 0 0 0 0 0 0 0 0 67

93.876 0 0 0 0 0 0 0 0 0 0

0

93.918 0 0 0 0 0 0 6685 0 0 0

0

93.994 0 0 0 0 0 0 0 0 0 0

0

94.027 0 0 0 0 0 0 0 0 0 0

0

94.055 0 0 0 0 0 0 0 0 0 0 0

94.199 0 0 0 0 0 0 0 0 0 0 0

94.215 0 0 0 0 0 0 0 0 0 0

0

94.465 0 0 0 0 0 0 0 0 4374 0 0

94.601 0 0 204 0 0 0 0 0 0 0 0

94.767 128 0 0 0 0 0 0 0 0 0 0

cons.conf.idx

cons.price.idx -40.3 -40 -39.8 -38.3 -37.5 -36.4 -36.1 -34.8 -34.6 -33.6 33

92.201 0 0 0 0 0 0 0 0 0 0 0

92.379 0 0 0 0 0 0 0 0 0 0

0

92.431 0 0 0 0 0 0 0 0 0 0 0

92.469 0 0 0 0 0 0 0 0 0 178 0

92.649 0 0 0 0 0 0 0 0 0 0

|  |
| --- |
| 0  92.713 0 0 0 0 0 0 0 0 0 0 172  92.756 0 0 0 0 0 0 0 0 0 0 0  92.843 0 0 0 0 0 0 0 0 0 0  0  92.893 0 0 0 0 0 0 0 0 0 0  0  92.963 0 0 0 0 0 0 0 0 0 0 0  93.075 0 0 0 0 0 0 0 0 0 0 0  93.2 0 0 0 0 0 0 0 0 0 0  0  93.369 0 0 0 0 0 0 0 264 0 0 0  93.444 0 0 0 0 0 0 5175 0 0 0 0  93.749 0 0 0 0 0 0 0 0 174 0  0  93.798 0 0 0 0 0 0 0 0 0 0 0  93.876 0 212 0 0 0 0 0 0 0 0 0  93.918 0 0 0 0 0 0 0 0 0 0  0  93.994 0 0 0 0 0 7763 0 0 0 0 0  94.027 0 0 0 233 0 0 0 0 0 0 0  94.055 0 0 229 0 0 0 0 0 0 0  0  94.199 0 0 0 0 303 0 0 0 0 0  0  94.215 311 0 0 0 0 0 0 0 0 0  0  94.465 0 0 0 0 0 0 0 0 0 0  0  94.601 0 0 0 0 0 0 0 0 0 0  0  94.767 0 0 0 0 0 0 0 0 0 0 0  cons.conf.idx  cons.price.idx -31.4 -30.1 -29.8 -26.9 92.201 770 0 0 0  92.379 0 0 267 0  92.431 0 0 0 447  92.469 0 0 0 0  92.649 0 357 0 0  92.713 0 0 0 0  92.756 0 0 0 0  92.843 0 0 0 0  92.893 0 0 0 0  92.963 0 0 0 0  93.075 0 0 0 0  93.2 0 0 0 0  93.369 0 0 0 0  93.444 0 0 0 0  93.749 0 0 0 0  93.798 0 0 0 0  93.876 0 0 0 0  93.918 0 0 0 0 |

93.994 0 0 0 0

|  |
| --- |
| 94.027 0 0 0 0  94.055 0 0 0 0  94.199 0 0 0 0  94.215 0 0 0 0  94.465 0 0 0 0  94.601 0 0 0 0  94.767 0 0 0 0 cons.conf.idx  cons.price.idx -50.8 -50 -49.5 -47.1 46.2  92.201 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.379 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.431 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.469 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.649 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.713 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.756 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.843 0.0000000000 0.0068466544 0.0000000000 0.0000000000  0.0000000000  92.893 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.1406720404  92.963 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.075 0.0000000000 0.0000000000 0.0000000000 0.0596775760  0.0000000000  93.2 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.369 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.444 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.749 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.798 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.876 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.918 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.994 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.027 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.055 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.199 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.215 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.465 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.601 0.0000000000 0.0000000000 0.0049528989 0.0000000000  0.0000000000  94.767 0.0031077013 0.0000000000 0.0000000000 0.0000000000 0.0000000000  cons.conf.idx |

cons.price.idx -45.9 -42.7 -42 -41.8 -

|  |
| --- |
| 40.8  92.201 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.379 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.431 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.469 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.649 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.713 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.756 0.0002427892 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.843 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.893 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.963 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0173594251  93.075 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.2 0.0000000000 0.0000000000 0.0877925609 0.0000000000  0.0000000000  93.369 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.444 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.749 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.798 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.876 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.918 0.0000000000 0.1623045547 0.0000000000 0.0000000000  0.0000000000  93.994 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.027 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.055 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.199 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.215 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.465 0.0000000000 0.0000000000 0.0000000000 0.1061959794  0.0000000000  94.601 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.767 0.0000000000 0.0000000000 0.0000000000 0.0000000000 0.0000000000  cons.conf.idx  cons.price.idx -40.4 -40.3 -40 -39.8 38.3  92.201 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.379 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.431 0.0000000000 0.0000000000 0.0000000000 0.0000000000 |

92.469 0.0000000000 0.0000000000 0.0000000000 0.0000000000

|  |
| --- |
| 0.0000000000  92.649 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.713 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.756 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.843 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.893 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.963 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.075 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.2 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.369 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.444 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.749 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.798 0.0016266874 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.876 0.0000000000 0.0000000000 0.0051471302 0.0000000000  0.0000000000  93.918 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.994 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.027 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0056569875  94.055 0.0000000000 0.0000000000 0.0000000000 0.0055598718  0.0000000000  94.199 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.215 0.0000000000 0.0075507429 0.0000000000 0.0000000000  0.0000000000  94.465 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.601 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.767 0.0000000000 0.0000000000 0.0000000000 0.0000000000 0.0000000000  cons.conf.idx  cons.price.idx -37.5 -36.4 -36.1 -34.8 34.6  92.201 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.379 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.431 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.469 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.649 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.713 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.756 0.0000000000 0.0000000000 0.0000000000 0.0000000000 |

92.843 0.0000000000 0.0000000000 0.0000000000 0.0000000000

|  |
| --- |
| 0.0000000000  92.893 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.963 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.075 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.2 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.369 0.0000000000 0.0000000000 0.0000000000 0.0064096339  0.0000000000  93.444 0.0000000000 0.0000000000 0.1256433913 0.0000000000  0.0000000000  93.749 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0042245314  93.798 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.876 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.918 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.994 0.0000000000 0.1884772264 0.0000000000 0.0000000000  0.0000000000  94.027 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.055 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.199 0.0073565116 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.215 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.465 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.601 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.767 0.0000000000 0.0000000000 0.0000000000 0.0000000000 0.0000000000  cons.conf.idx  cons.price.idx -33.6 -33 -31.4 -30.1 29.8  92.201 0.0000000000 0.0000000000 0.0186947655 0.0000000000  0.0000000000  92.379 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0064824706  92.431 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.469 0.0043216471 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.649 0.0000000000 0.0000000000 0.0000000000 0.0086675731  0.0000000000  92.713 0.0000000000 0.0041759736 0.0000000000 0.0000000000  0.0000000000  92.756 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.843 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.893 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  92.963 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.075 0.0000000000 0.0000000000 0.0000000000 0.0000000000 |

93.2 0.0000000000 0.0000000000 0.0000000000 0.0000000000

|  |
| --- |
| 0.0000000000  93.369 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.444 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.749 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.798 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.876 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.918 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  93.994 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.027 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.055 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.199 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.215 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.465 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.601 0.0000000000 0.0000000000 0.0000000000 0.0000000000  0.0000000000  94.767 0.0000000000 0.0000000000 0.0000000000 0.0000000000 0.0000000000  cons.conf.idx cons.price.idx -26.9 92.201 0.0000000000  92.379 0.0000000000  92.431 0.0108526755  92.469 0.0000000000  92.649 0.0000000000  92.713 0.0000000000  92.756 0.0000000000  92.843 0.0000000000  92.893 0.0000000000  92.963 0.0000000000  93.075 0.0000000000  93.2 0.0000000000  93.369 0.0000000000  93.444 0.0000000000  93.749 0.0000000000  93.798 0.0000000000  93.876 0.0000000000  93.918 0.0000000000  93.994 0.0000000000  94.027 0.0000000000  94.055 0.0000000000  94.199 0.0000000000  94.215 0.0000000000  94.465 0.0000000000  94.601 0.0000000000  94.767 0.0000000000 |

Is the employment variation rate consistent across job types?

# with(bank\_additional\_full, chisq.test( job,emp.var.rate))

with(bank\_additional\_full, table( job,emp.var.rate) )

# OR

with(bank\_additional\_full, prop.table(table( job,emp.var.rate)))

#p-value < 2.2e-16 is very much less than 0.05

|  |
| --- |
| Pearson's Chi-squared test    data: job and emp.var.rate  X-squared = 4676.8, df = 99, p-value < 2.2e-16    emp.var.rate  job -3.4 -3 -2.9 -1.8 -1.7 -1.1 -0.2 -0.1 1.1 1.4 admin. 321 47 562 2231 246 187 3 940 1601 4284 blue-collar 64 9 99 2519 58 33 3 575 2295 3599 entrepreneur 24 1 38 306 14 7 0 265 289 512 housemaid 32 9 41 120 18 16 1 70 229 524 management 98 12 121 593 47 38 0 522 553 940 retired 193 33 181 338 96 83 0 72 215 509 self-employed 40 6 60 287 24 12 0 187 253 552 services 32 2 88 1040 47 40 0 311 932 1477 student 62 20 144 311 72 73 0 21 66 106 technician 145 22 234 1243 110 115 2 575 1060 3237 unemployed 44 9 76 164 31 28 1 141 171 349 unknown 16 2 19 32 10 3 0 4 99 145 emp.var.rate  job -3.4 -3 -2.9 -1.8 -  1.7  admin. 7.793532e-03 1.141109e-03 1.364475e-02 5.416626e-02  5.972613e-03  blue-collar 1.553851e-03 2.185102e-04 2.403613e-03 6.115859e-02  1.408177e-03  entrepreneur 5.826940e-04 2.427892e-05 9.225988e-04 7.429348e-03  3.399048e-04  housemaid 7.769253e-04 2.185102e-04 9.954356e-04 2.913470e-03  4.370205e-04  management 2.379334e-03 2.913470e-04 2.937749e-03 1.439740e-02  1.141109e-03  retired 4.685831e-03 8.012042e-04 4.394484e-03 8.206274e-03  2.330776e-03  self-employed 9.711566e-04 1.456735e-04 1.456735e-03 6.968049e-03  5.826940e-04  services 7.769253e-04 4.855783e-05 2.136545e-03 2.525007e-02  1.141109e-03  student 1.505293e-03 4.855783e-04 3.496164e-03 7.550743e-03  1.748082e-03  technician 3.520443e-03 5.341362e-04 5.681266e-03 3.017869e-02  2.670681e-03  unemployed 1.068272e-03 2.185102e-04 1.845198e-03 3.981742e-03  7.526464e-04 |

unknown 3.884627e-04 4.855783e-05 4.612994e-04 7.769253e-04

2.427892e-04

emp.var.rate

job -1.1 -0.2 -0.1 1.1

1.4

admin. 4.540157e-03 7.283675e-05 2.282218e-02 3.887054e-02

1.040109e-01

blue-collar 8.012042e-04 7.283675e-05 1.396038e-02 5.572011e-02

8.737982e-02

entrepreneur 1.699524e-04 0.000000e+00 6.433913e-03 7.016607e-03

1.243081e-02

housemaid 3.884627e-04 2.427892e-05 1.699524e-03 5.559872e-03

1.272215e-02

management 9.225988e-04 0.000000e+00 1.267359e-02 1.342624e-02

2.282218e-02

retired 2.015150e-03 0.000000e+00 1.748082e-03 5.219967e-03

1.235797e-02

self-employed 2.913470e-04 0.000000e+00 4.540157e-03 6.142566e-03

1.340196e-02

services 9.711566e-04 0.000000e+00 7.550743e-03 2.262795e-02

3.585996e-02

student 1.772361e-03 0.000000e+00 5.098572e-04 1.602408e-03

2.573565e-03

technician 2.792075e-03 4.855783e-05 1.396038e-02 2.573565e-02

7.859085e-02

unemployed 6.798097e-04 2.427892e-05 3.423327e-03 4.151695e-03

8.473342e-03

unknown 7.283675e-05 0.000000e+00 9.711566e-05 2.403613e-03

3.520443e-03

Is the employment variation rate same across education? Which group is more confident?

with(bank\_additional\_full, chisq.test( education,emp.var.rate)) with(bank\_additional\_full, table( education, emp.var.rate) )

# OR

with(bank\_additional\_full, prop.table(table( education,emp.var.rate)))

Pearson's Chi-squared test

data: education and emp.var.rate

X-squared = 1451.6, df = 63, p-value < 2.2e-16

emp.var.rate

education -3.4 -3 -2.9 -1.8 -1.7 -1.1 -0.2 -0.1 1.1 1.4 basic.4y 141 17 106 843 75 59 3 238 993 1701 basic.6y 36 0 35 584 18 9 0 154 592 864 basic.9y 69 16 110 1628 53 27 0 504 1428 2210 high.school 216 36 358 2366 183 143 4 809 1857 3543 illiterate 0 0 3 3 0 0 0 3 2 7 professional.course 131 19 196 1041 93 113 3 470 887 2290 university.degree 411 70 758 2403 301 242 0 1414 1627 4942

|  |
| --- |
| unknown 67 14 97 316 50 42 0 91 377 677 emp.var.rate  education -3.4 -3 -2.9 -1.8 basic.4y 3.423327e-03 4.127416e-04 2.573565e-03 2.046713e-02 basic.6y 8.740410e-04 0.000000e+00 8.497621e-04 1.417889e-02 basic.9y 1.675245e-03 3.884627e-04 2.670681e-03 3.952608e-02 high.school 5.244246e-03 8.740410e-04 8.691852e-03 5.744392e-02 illiterate 0.000000e+00 0.000000e+00 7.283675e-05 7.283675e-05 professional.course 3.180538e-03 4.612994e-04 4.758668e-03 2.527435e-02 university.degree 9.978635e-03 1.699524e-03 1.840342e-02 5.834224e-02 unknown 1.626687e-03 3.399048e-04 2.355055e-03 7.672138e-03 emp.var.rate  education -1.7 -1.1 -0.2 -0.1 basic.4y 1.820919e-03 1.432456e-03 7.283675e-05 5.778382e-03 basic.6y 4.370205e-04 2.185102e-04 0.000000e+00 3.738953e-03 basic.9y 1.286783e-03 6.555307e-04 0.000000e+00 1.223657e-02 high.school 4.443042e-03 3.471885e-03 9.711566e-05 1.964164e-02 illiterate 0.000000e+00 0.000000e+00 0.000000e+00 7.283675e-05 professional.course 2.257939e-03 2.743518e-03 7.283675e-05 1.141109e-02 university.degree 7.307954e-03 5.875498e-03 0.000000e+00 3.433039e-02 unknown 1.213946e-03 1.019714e-03 0.000000e+00 2.209381e-03 emp.var.rate  education 1.1 1.4 basic.4y 2.410896e-02 4.129844e-02 basic.6y 1.437312e-02 2.097698e-02 basic.9y 3.467029e-02 5.365640e-02 high.school 4.508595e-02 8.602020e-02 illiterate 4.855783e-05 1.699524e-04 professional.course 2.153540e-02 5.559872e-02 university.degree 3.950180e-02 1.199864e-01 unknown 9.153151e-03 1.643683e-02 |

bank\_marketing\_data <- read\_delim("C:/Users/Seshan/Desktop/bank\_marketing\_data.csv",

";", escape\_double = FALSE, trim\_ws = TRUE) head(bank\_marketing\_data)

# We look at difference between mean and median in summary if it's more there might be outliers

boxplot(bank\_marketing\_data$age, main="Age Box plot",yaxt="n", xlab="Age", horizontal=TRUE, col=terrain.colors(2))

# By plotting histogram we can ensure if there are outliers or not

## DATA VISUALISATION

## Use Box plots (Only for continuous variables)- To Check Ouliers

boxplot(bank\_marketing\_data$age~bank\_marketing\_data$contact, main=" AGE",ylab="age of customers",xlab="contact")

boxplot(bank\_marketing\_data$age~bank\_marketing\_data$job, main=" AGE",ylab="age of customers",xlab="job")

boxplot(bank\_marketing\_data$age~bank\_marketing\_data$education, main=" AGE",ylab="age of customers",xlab="education")

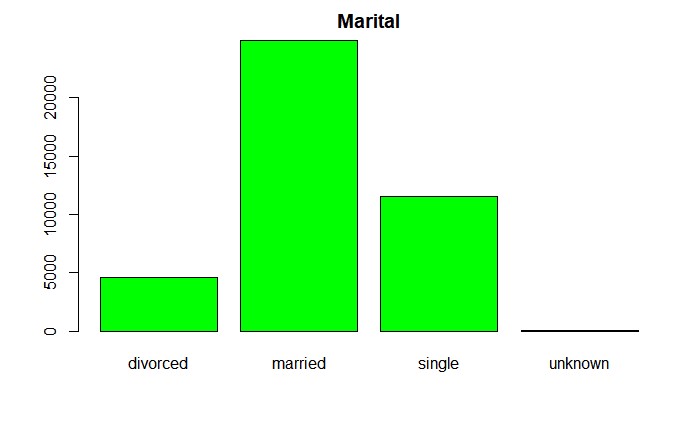
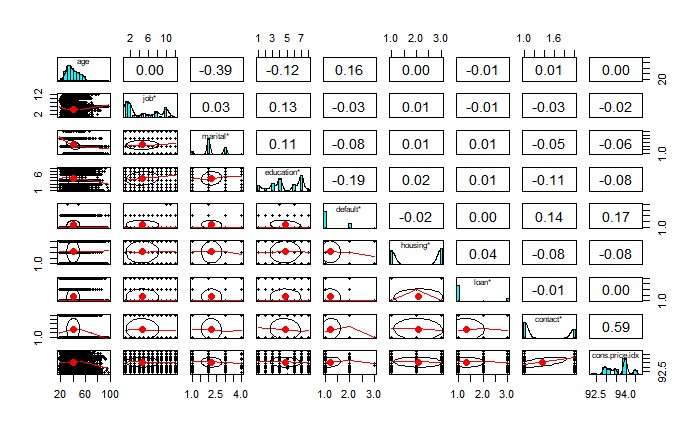
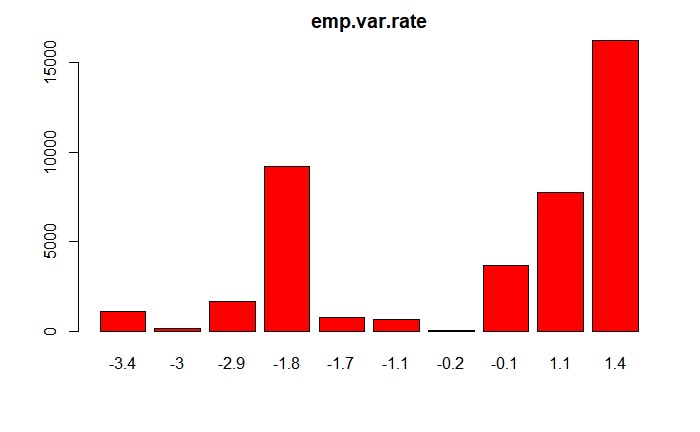
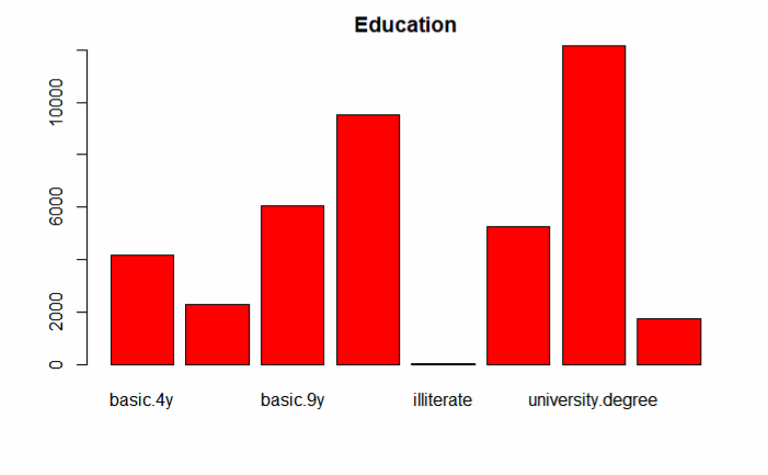
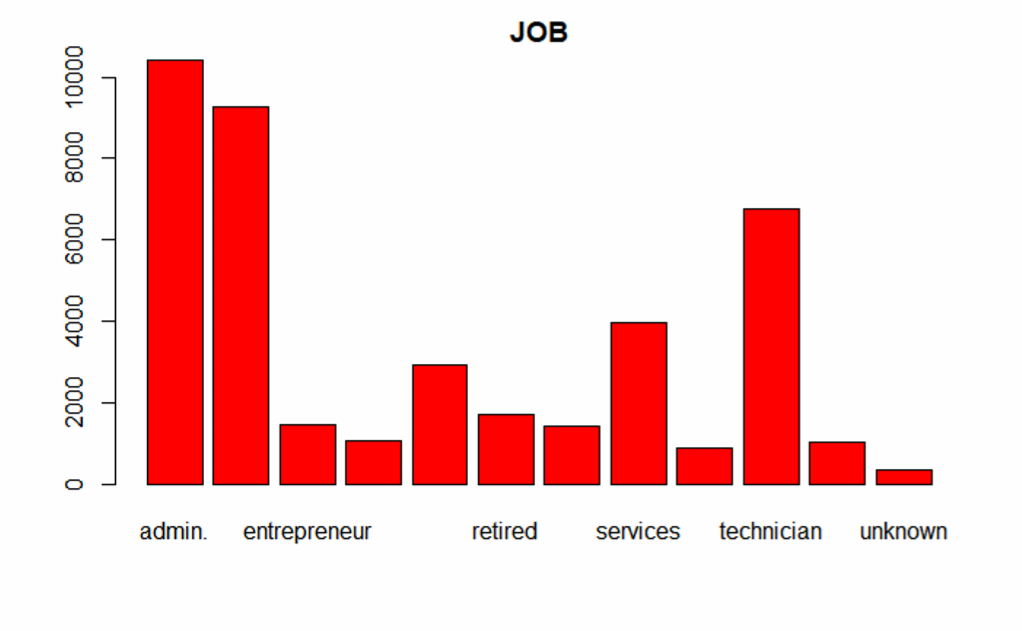
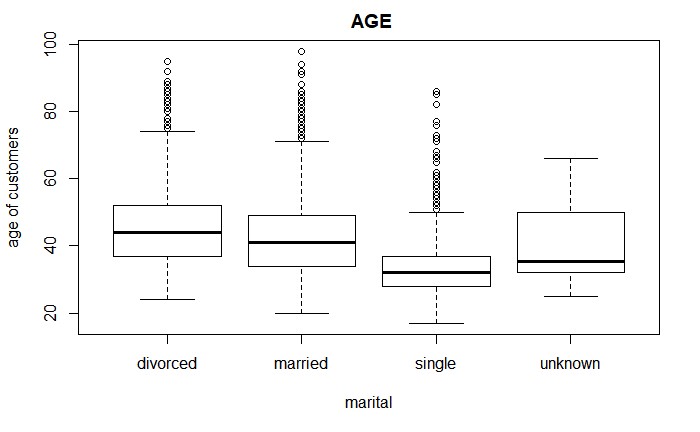
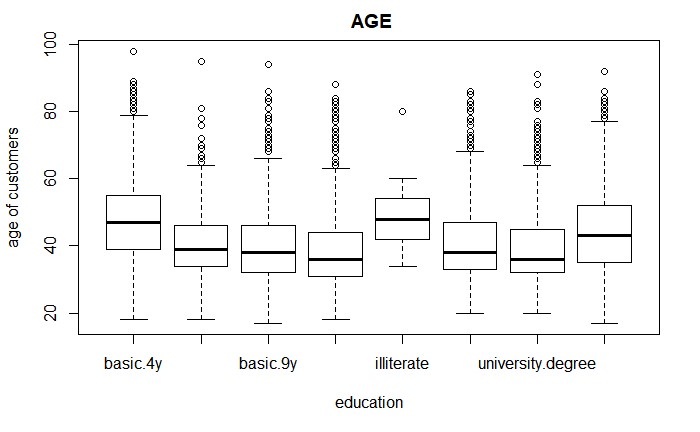
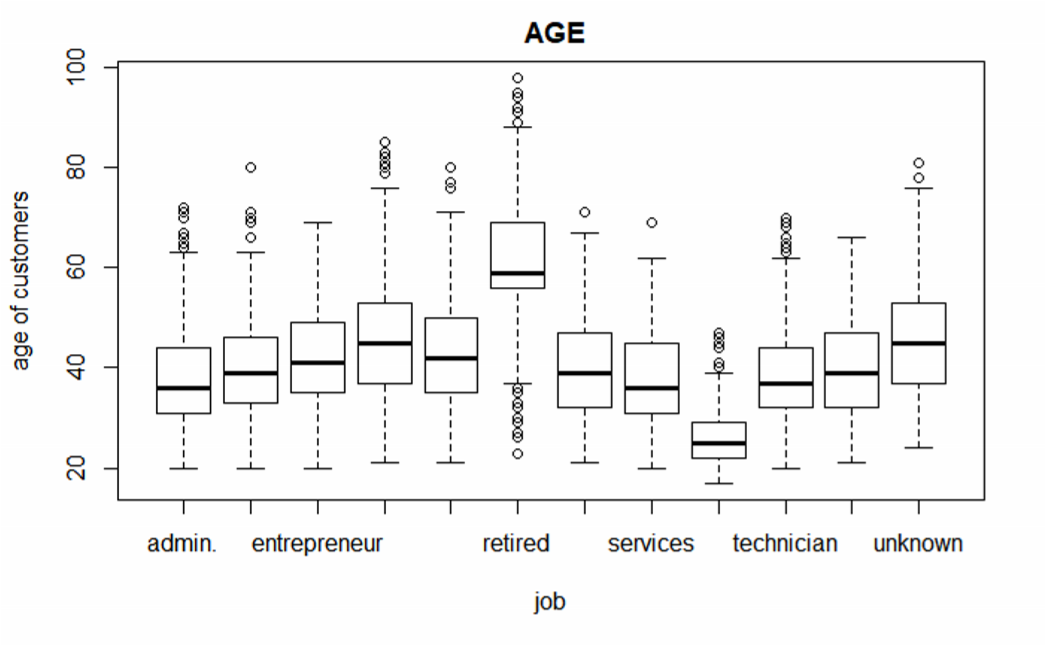
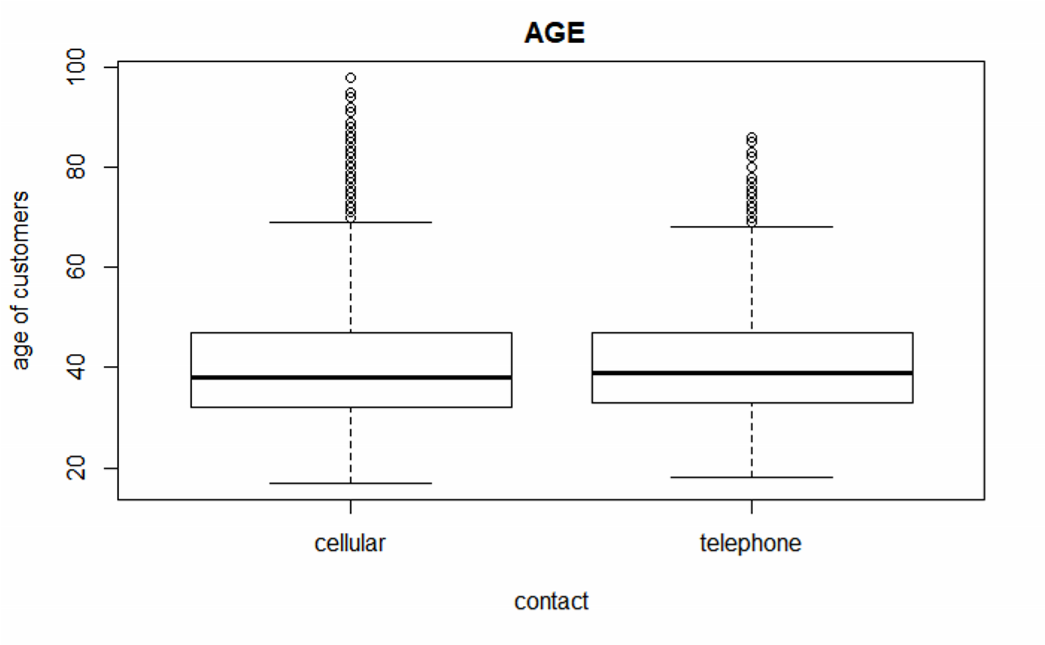
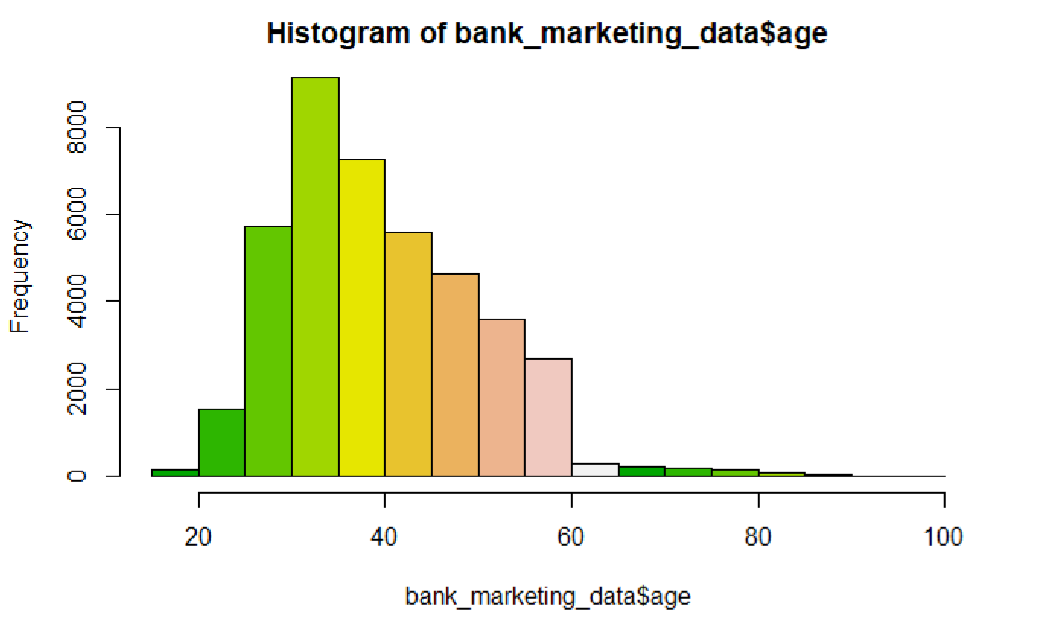
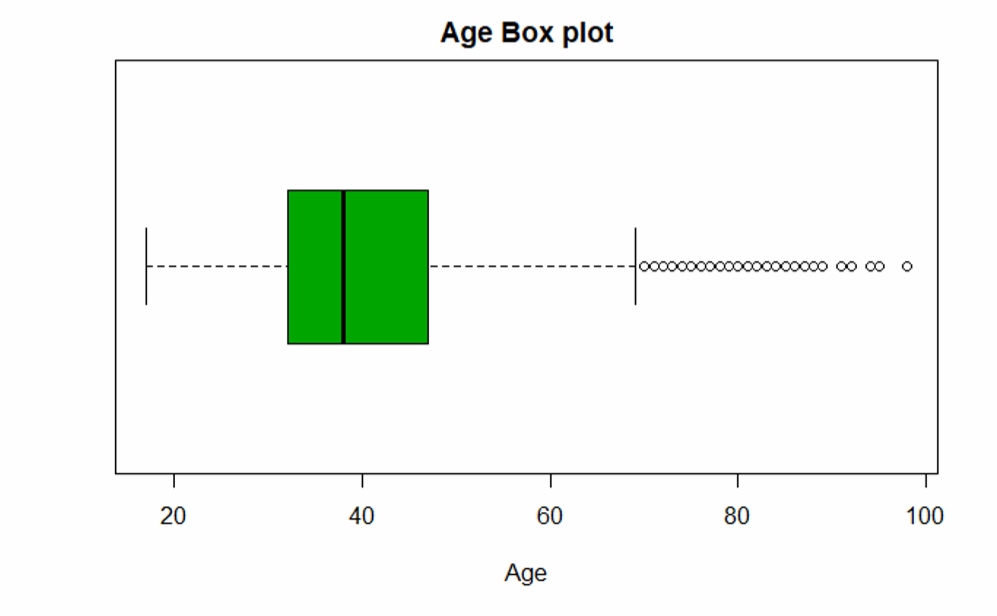
boxplot(bank\_marketing\_data$age~bank\_marketing\_data$marital, main=" AGE",ylab="age of customers",xlab="marital")

## Barplots for Categorical Variables

barplot(table(bank\_marketing\_data$job),col="red",main="JOB") barplot(table(bank\_marketing\_data$marital),col="green",main="Marital") barplot(table(bank\_marketing\_data$education),col="red",main="Education") barplot(table(bank\_marketing\_data$emp.var.rate ),col="red",main="emp.var.rate") hist(bank\_marketing\_data$age,col=terrain.colors(10))

#Correlation Analysis What we saw in the box plot can be emphasized by correlation plot, It can tell if predictor is a good predictor or not a good predictor. This analysis can help us decide if we can drop some columns/predictors depending upon its correlation with the outcome variable. library(psych) pairs.panels(bank\_marketing\_data[, c(1:8,17)]) pairs.panels(bank\_marketing\_data[, c(9:17)])

pairs.panels(bank\_marketing\_data[, c(1:8,19)])



Subset Selection/ Feature

-

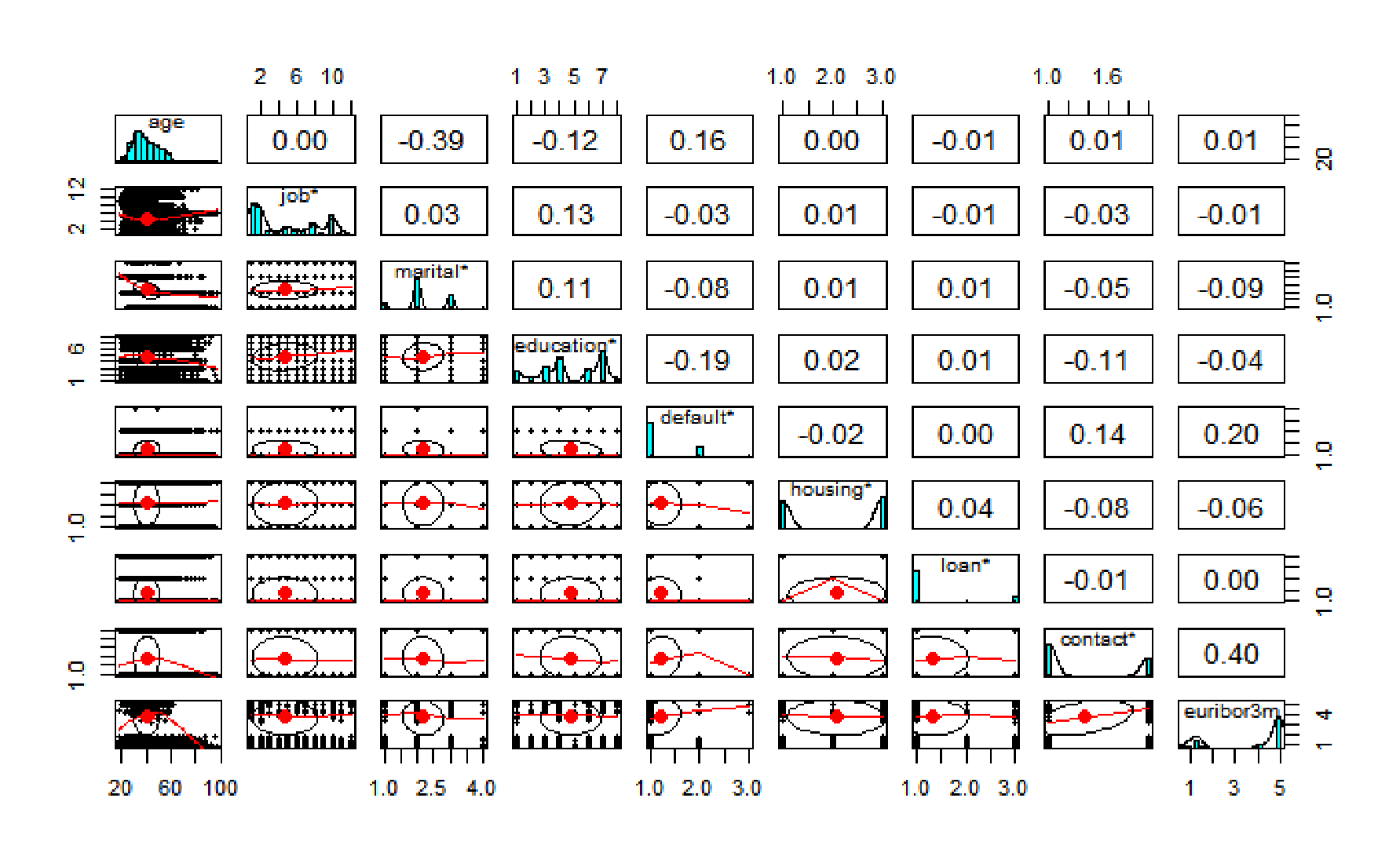
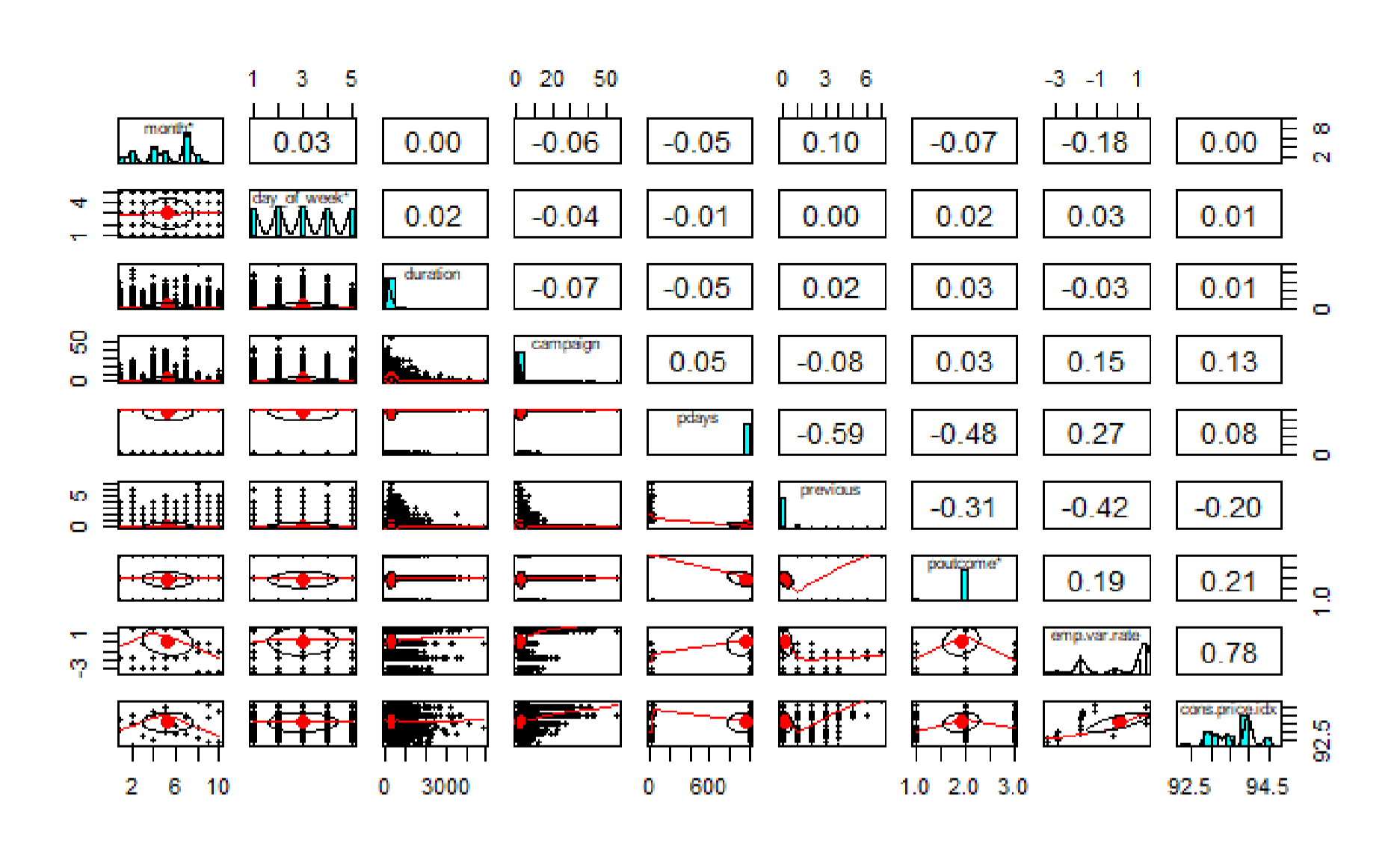
space reduction: Features

-

space can be reduced by selecting s

ubsets

based upon correlation values obtained



#################Subset Selection################# lib bank\_marketing\_data\_sub<-bank\_marketing\_data[, c(1:4,7:9,12,14,15,17)] str(bank\_marketing\_data\_sub) pairs.panels(bank\_marketing\_data\_sub)

#3.4. Data transformation and Binning We do data transformation and binning for better modeling. We convert categorical variable into numerical using binning.

#################Binning and Data Transformation#################

#bank\_marketing\_data\_sub$age <- cut(bank\_marketing\_data\_sub$age, c(1,20,40,60,100))

#bank\_marketing\_data\_sub$is\_divorced <- ifelse( bank\_marketing\_data\_sub$marital ==

"divorced", 1, 0)

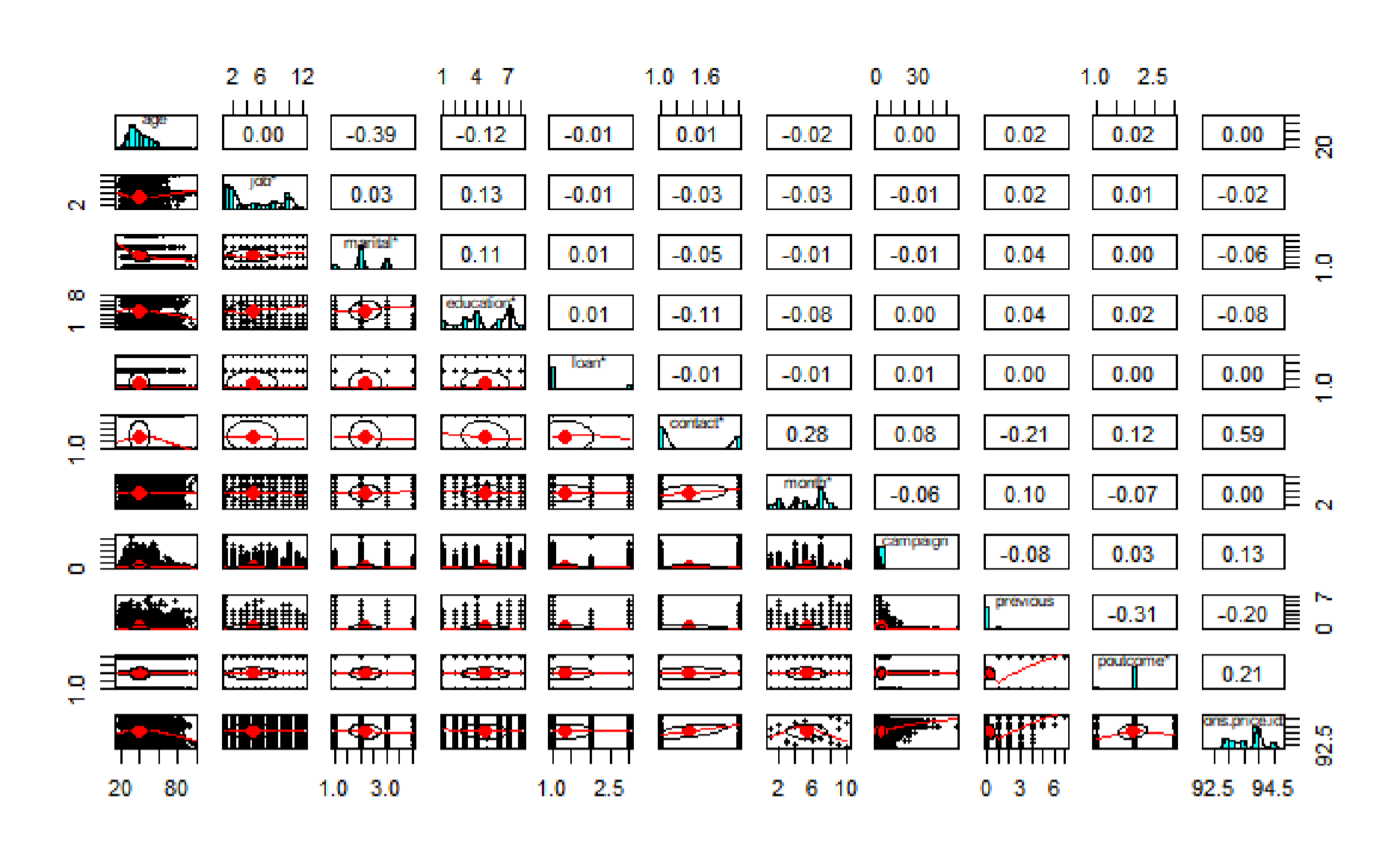
bank\_marketing\_data\_sub$is\_nr.employed <- ifelse( bank\_marketing\_data\_sub$education == "employed", 1, 0)

#bank\_marketing\_data\_sub$is\_single <- ifelse( bank\_marketing\_data\_sub$marital == "single",

1, 0)

bank\_marketing\_data\_sub$nr.employed <- NULL

str(bank\_marketing\_data\_sub)



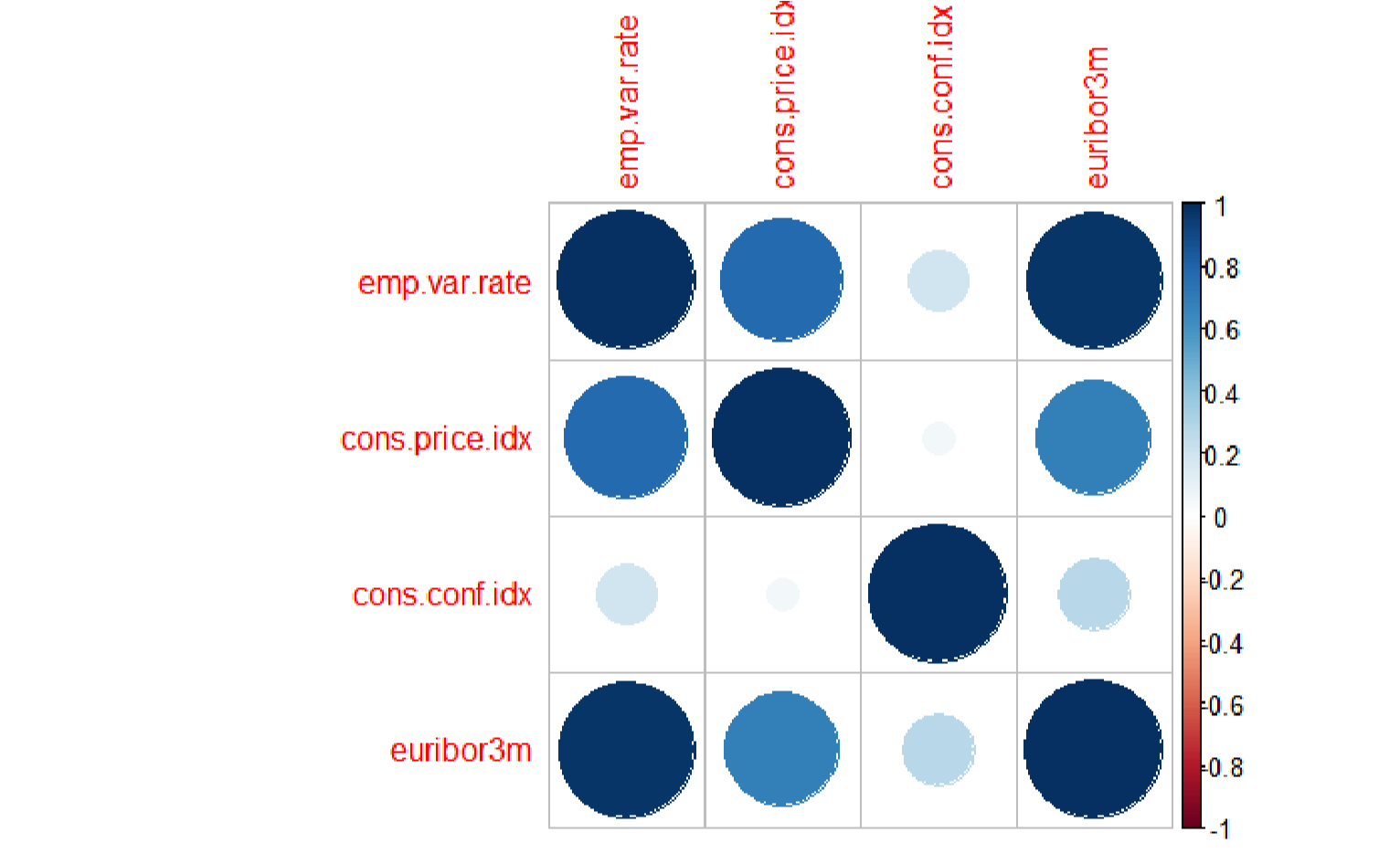
#scatter.smooth(x=bank\_marketing\_data$job, y=bank\_marketing\_data$emp.var.rate, main="emp.var.rate ~ job") # scatterplot

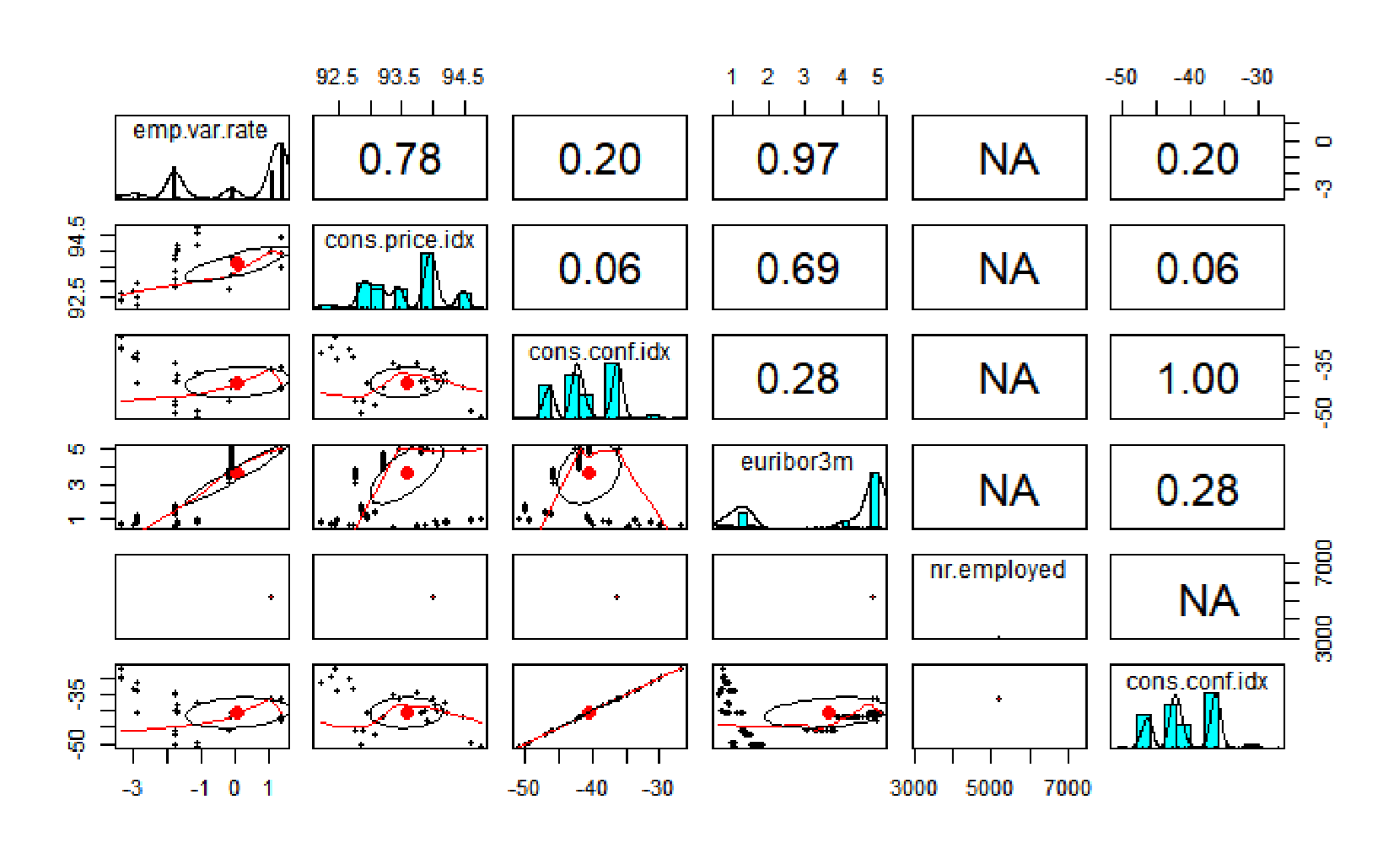
# load library library(corrplot) # load the data data<-bank\_marketing\_data data(bank\_marketing\_data\_sub) # calculate correlations

correlations <- cor(bank\_marketing\_data[,16:19])

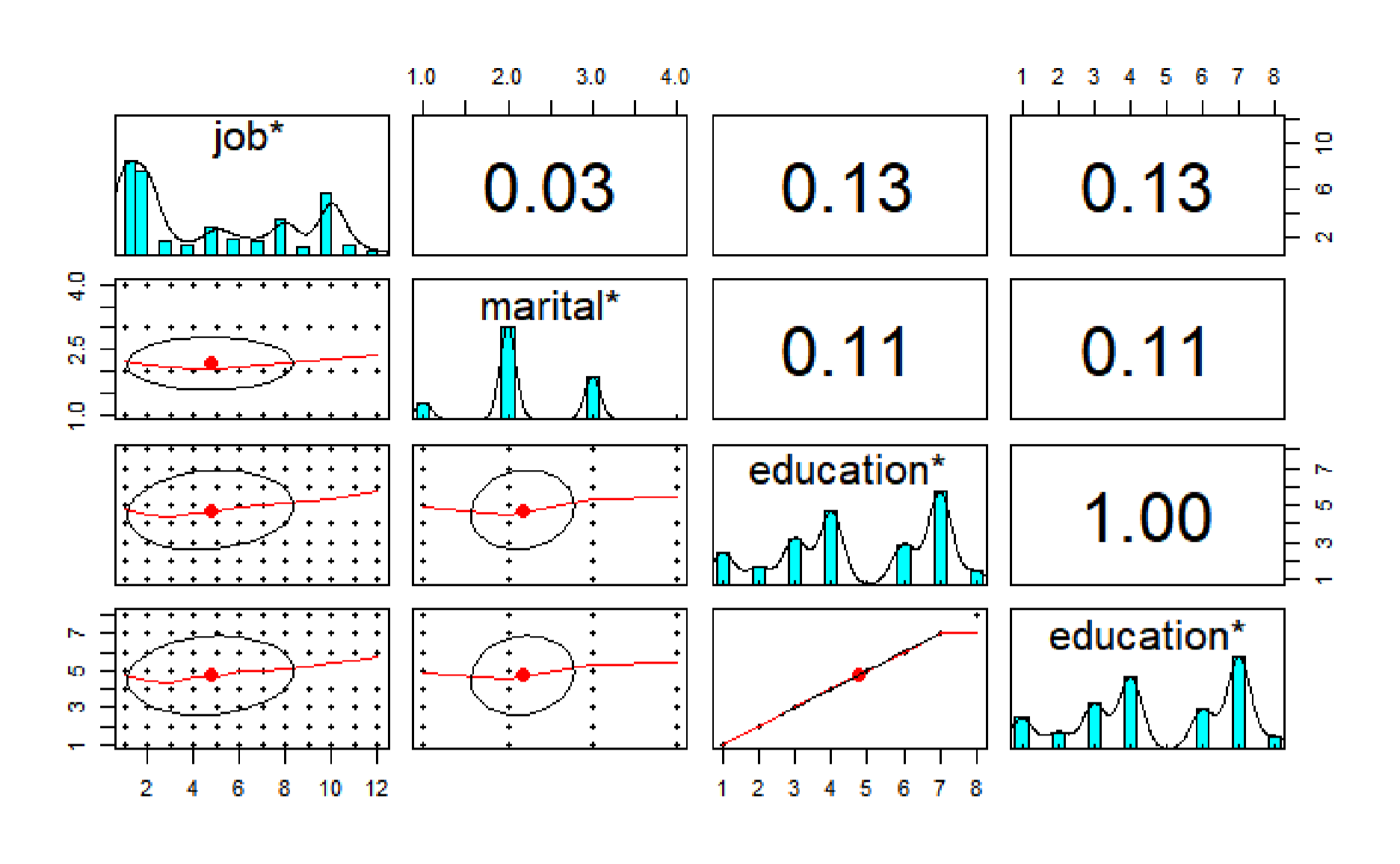
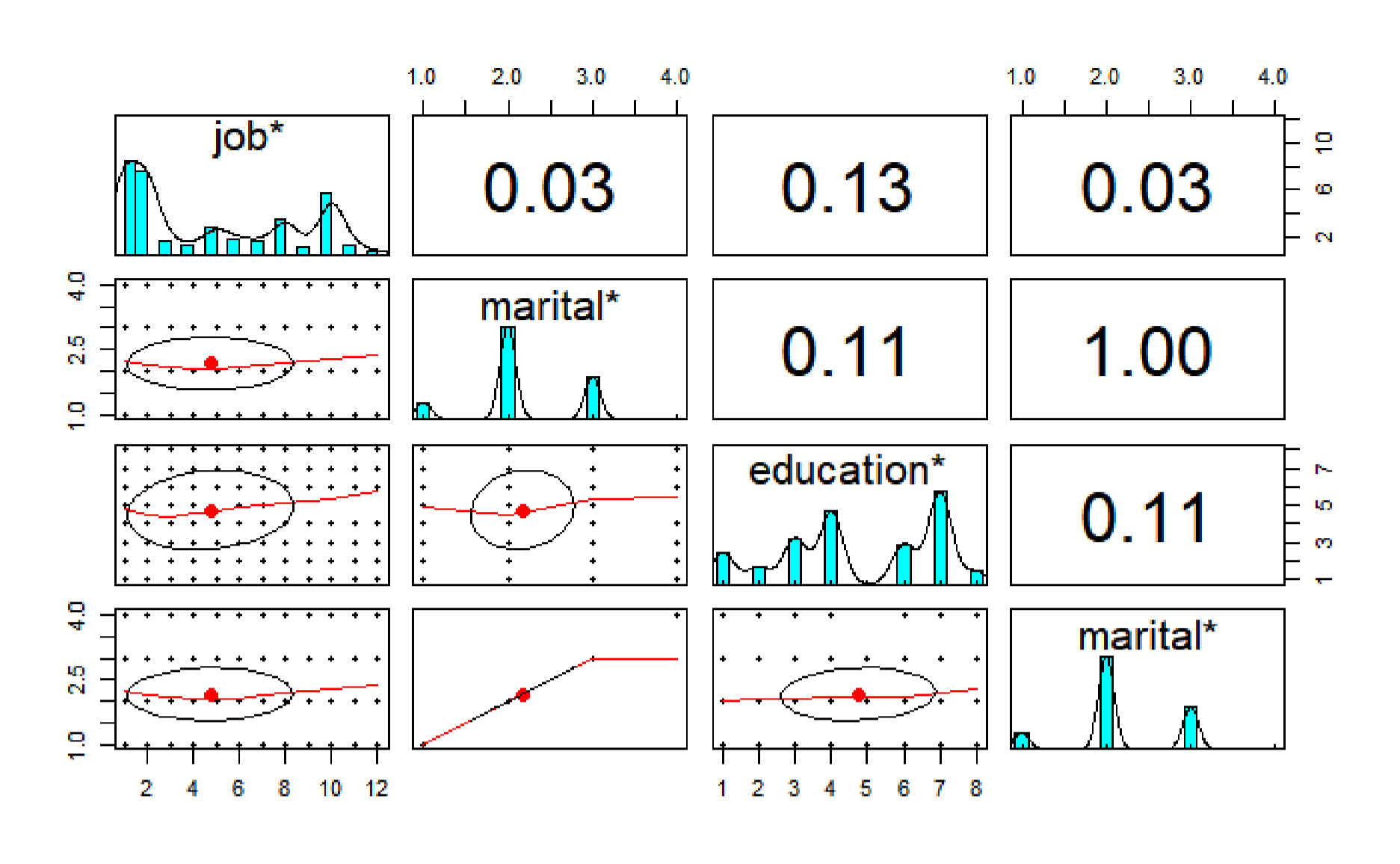
# create correlation plot

corrplot(correlations, method="circle")

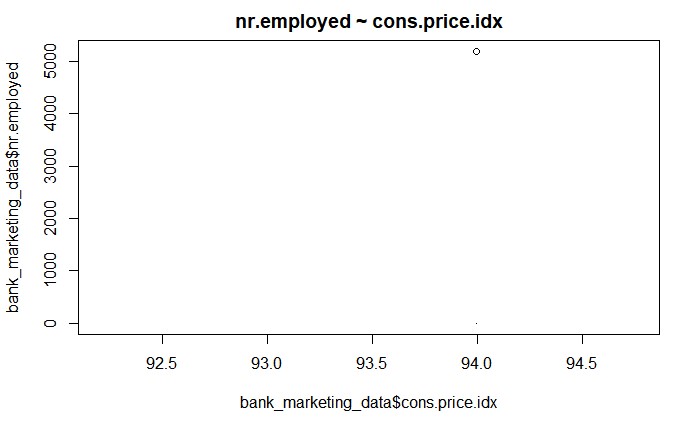
 airs.panels(bank\_marketing\_data[, c(16:20,18)]) pairs.panels(bank\_marketing\_data[, c(2:4,3)]) pairs.panels(bank\_marketing\_data[, c(2:4,4)])



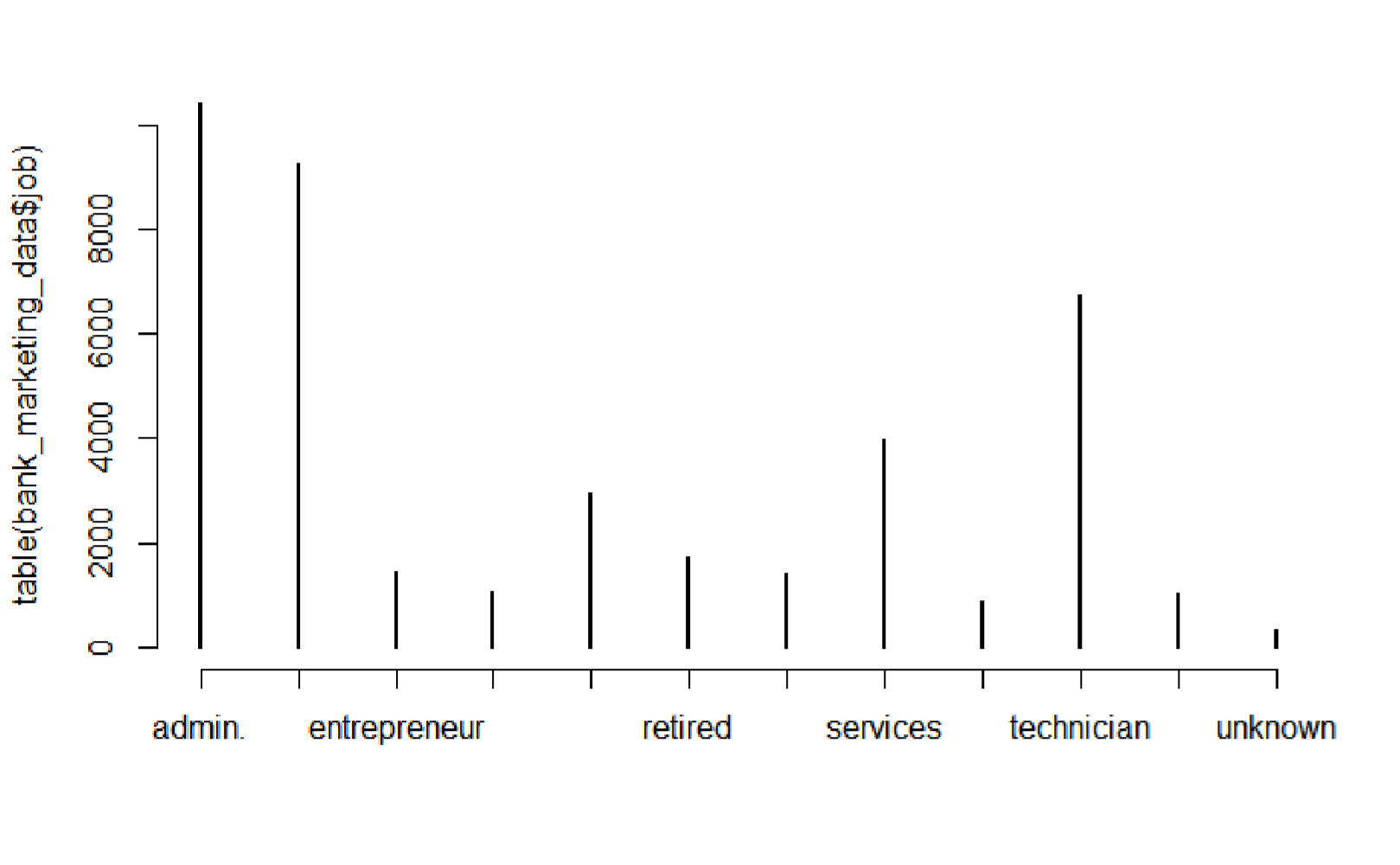
head(bank\_marketing\_data)



scatter.smooth(x=bank\_marketing\_data$cons.price.idx, y=bank\_marketing\_data$nr.employed, main="nr.employed ~ cons.price.idx")



#cor(bank\_marketing\_data$age, bank\_marketing\_data$emp.var.rate) head(bank\_marketing\_data) table(bank\_marketing\_data$job) table(bank\_marketing\_data$marital) plot(table(bank\_marketing\_data$job))

 library(psych)

pairs.panels(bank\_marketing\_data[,1:6])

