**6 - Correspondence Analysis**

Correspondence analysis (CA) is a dimension reduction technique for contingency tables/cross-tabulations of nominal or ordinal variables. It is similar to principal components analysis for continuous variables. The data matrix for simple correspondence analysis is typically a two-way contingency table, but it could be any other table of non-negative ratio-scale data where relative values are of primary interest. In other situations, for example multiple correspondence analysis, the data matrix may consist of individual level data with indicator variables denoting the levels these individuals have for set of qualitative variables. In either the case, singular value decomposition (SVD) of a properly scaled matrix is used to achieve the dimension reduction to typically k = 2 or k = 3 dimensions. Multiple correspondence analysis (MCA) allows for larger dimensional situations, thus additional dimensions beyond 2 or 3 may be considered. We will first look at how SVD can be used to create the lower dimensional representation of the data presented/summarized in a two-way contingency table.

**6.1 – First Look at Correspondence Analysis in R**

Example 6.1: Suicides in the West Germany  
The data for this example are taken from a study of suicides in the former West Germany in the years 1974 to 1977, reported by Van der Heijden and de Leeuw (1985). Nine methods of suicide were tabulated by sex and age category. The primary interest here is in the variation of suicide method patterns by age and by sex. The data can be regarded as a two-way contingency table, with the 34 age/sex categories forming the rows and the nine methods of suicide forming the columns. The contingency table in R format is displayed below.

> Suicide

pois cookgas toxgas hang drown gun knife jump other

m1015 4 0 0 247 1 17 1 6 9

m1520 348 7 67 578 22 179 11 74 175

m2025 808 32 229 699 44 316 35 109 289

m2530 789 26 243 648 52 268 38 109 226

m3035 916 17 257 825 74 291 52 123 281

m3540 1118 27 313 1278 89 299 53 78 198

m4045 926 13 250 1273 89 299 53 78 198

m4550 855 9 203 1381 71 347 68 103 190

m5055 684 14 136 1282 87 229 62 63 146

m5560 502 6 77 972 49 151 46 66 77

m6065 516 5 74 1249 83 162 62 92 122

m6570 513 8 31 1360 75 164 56 115 95

m7075 425 5 21 1268 90 121 44 119 82

m7580 266 4 9 866 63 78 30 79 34

m8085 159 2 2 479 39 18 18 46 19

m8590 70 1 0 259 16 10 9 18 10

m90p 18 0 1 76 4 2 4 6 2

pois cookgas toxgas hang drown gun knife jump other

w1015 28 0 3 20 0 1 0 10 6

w1520 353 2 11 81 6 15 2 43 47

w2025 540 4 20 111 24 9 9 78 67

w2530 454 6 27 125 33 26 7 86 75

w3035 530 2 29 178 42 14 20 92 78

w3540 688 5 44 272 64 24 14 98 110

w4045 566 4 24 343 76 18 22 103 86

w4550 716 6 24 447 94 13 21 95 88

w5055 942 7 26 691 184 21 37 129 131

w5560 723 3 14 527 163 14 30 92 92

w6065 820 8 8 702 245 11 35 140 114

w6570 740 8 4 785 271 4 38 156 90

w7075 624 6 4 610 244 1 27 129 46

w7580 495 8 1 420 161 2 29 129 35

w8085 292 3 2 223 78 0 10 84 23

w8590 113 4 0 83 14 0 6 34 2

w90p 24 1 0 19 4 0 2 7 0

The numbers in the table are frequencies, for example the 942 highlighted above indicates that there were 942 women between the ages of 50-55 that used poisoning to end their life.

> dim(Suicide)

[1] 34 9

We first compute the overall proportion of individuals in each cell defined by a specific gender-age and method of suicide employed.

> suicide.prop = Suicide/sum(Suicide)  
> options(digits=5)

> suicide.prop

pois cookgas toxgas hang drown gun

m1015 0.00007532 0.00000000 0.00000000 0.00465099 0.00001883 0.00032011

m1520 0.00655281 0.00013181 0.00126160 0.01088369 0.00041426 0.00337055

m2025 0.01521457 0.00060256 0.00431205 0.01316211 0.00082852 0.00595025

m2530 0.01485680 0.00048958 0.00457567 0.01220178 0.00097916 0.00504642

m3035 0.01724820 0.00032011 0.00483929 0.01553468 0.00139341 0.00547950

m3540 0.02105184 0.00050841 0.00589376 0.02406462 0.00167586 0.00563014

m4045 0.01743650 0.00024479 0.00470748 0.02397047 0.00167586 0.00563014

m4550 0.01609957 0.00016947 0.00382247 0.02600410 0.00133692 0.00653398

m5055 0.01287966 0.00026362 0.00256087 0.02413994 0.00163820 0.00431205

m5560 0.00945261 0.00011298 0.00144990 0.01830267 0.00092267 0.00284332

m6065 0.00971623 0.00009415 0.00139341 0.02351856 0.00156288 0.00305045

m6570 0.00965974 0.00015064 0.00058373 0.02560868 0.00141224 0.00308811

m7075 0.00800271 0.00009415 0.00039543 0.02387633 0.00169469 0.00227842

m7580 0.00500876 0.00007532 0.00016947 0.01630670 0.00118628 0.00146873

m8085 0.00299396 0.00003766 0.00003766 0.00901953 0.00073437 0.00033894

m8590 0.00131809 0.00001883 0.00000000 0.00487695 0.00030128 0.00018830

m90p 0.00033894 0.00000000 0.00001883 0.00143107 0.00007532 0.00003766

w1015 0.00052724 0.00000000 0.00005649 0.00037660 0.00000000 0.00001883

w1520 0.00664696 0.00003766 0.00020713 0.00152522 0.00011298 0.00028245

w2025 0.01016815 0.00007532 0.00037660 0.00209012 0.00045192 0.00016947

w2530 0.00854878 0.00011298 0.00050841 0.00235374 0.00062139 0.00048958

w3035 0.00997985 0.00003766 0.00054607 0.00335172 0.00079086 0.00026362

w3540 0.01295498 0.00009415 0.00082852 0.00512174 0.00120511 0.00045192

w4045 0.01065773 0.00007532 0.00045192 0.00645866 0.00143107 0.00033894

w4550 0.01348222 0.00011298 0.00045192 0.00841697 0.00177001 0.00024479

w5055 0.01773777 0.00013181 0.00048958 0.01301147 0.00346470 0.00039543

w5560 0.01361402 0.00005649 0.00026362 0.00992336 0.00306928 0.00026362

w6065 0.01544053 0.00015064 0.00015064 0.01321860 0.00461333 0.00020713

w6570 0.01393413 0.00015064 0.00007532 0.01478148 0.00510291 0.00007532

w7075 0.01174986 0.00011298 0.00007532 0.01148624 0.00459450 0.00001883

w7580 0.00932081 0.00015064 0.00001883 0.00790856 0.00303162 0.00003766

w8085 0.00549833 0.00005649 0.00003766 0.00419907 0.00146873 0.00000000

w8590 0.00212778 0.00007532 0.00000000 0.00156288 0.00026362 0.00000000

w90p 0.00045192 0.00001883 0.00000000 0.00035777 0.00007532 0.00000000

knife jump other

m1015 0.00001883 0.00011298 0.00016947

m1520 0.00020713 0.00139341 0.00329523

m2025 0.00065905 0.00205246 0.00544184

m2530 0.00071554 0.00205246 0.00425556

m3035 0.00097916 0.00231608 0.00529120

m3540 0.00099799 0.00146873 0.00372832

m4045 0.00099799 0.00146873 0.00372832

m4550 0.00128043 0.00193948 0.00357768

m5055 0.00116745 0.00118628 0.00274917

m5560 0.00086618 0.00124277 0.00144990

m6065 0.00116745 0.00173235 0.00229725

m6570 0.00105447 0.00216544 0.00178884

m7075 0.00082852 0.00224076 0.00154405

m7580 0.00056490 0.00148756 0.00064022

m8085 0.00033894 0.00086618 0.00035777

m8590 0.00016947 0.00033894 0.00018830

m90p 0.00007532 0.00011298 0.00003766

w1015 0.00000000 0.00018830 0.00011298

w1520 0.00003766 0.00080969 0.00088501

w2025 0.00016947 0.00146873 0.00126160

w2530 0.00013181 0.00161937 0.00141224

w3035 0.00037660 0.00173235 0.00146873

w3540 0.00026362 0.00184533 0.00207129

w4045 0.00041426 0.00193948 0.00161937

w4550 0.00039543 0.00178884 0.00165703

w5055 0.00069671 0.00242906 0.00246672

w5560 0.00056490 0.00173235 0.00173235

w6065 0.00065905 0.00263619 0.00214661

w6570 0.00071554 0.00293747 0.00169469

w7075 0.00050841 0.00242906 0.00086618

w7580 0.00054607 0.00242906 0.00065905

w8085 0.00018830 0.00158171 0.00043309

w8590 0.00011298 0.00064022 0.00003766

w90p 0.00003766 0.00013181 0.00000000

Next we compute the proportion of individuals in each gender-age category.

> suicide.rowprop = apply(suicide.prop,1,sum)

> suicide.rowprop

m1015 m1520 m2025 m2530 m3035 m3540 m4045 m4550 m5055 m5560 m6065 m6570 m7075 m7580 m8085 m8590

0.0054 0.0275 0.0482 0.0452 0.0534 0.0650 0.0599 0.0608 0.0509 0.0366 0.0445 0.0455 0.0410 0.0269 0.0147 0.0074

m90p w1015 w1520 w2025 w2530 w3035 w3540 w4045 w4550 w5055 w5560 w6065 w6570 w7075 w7580 w8085

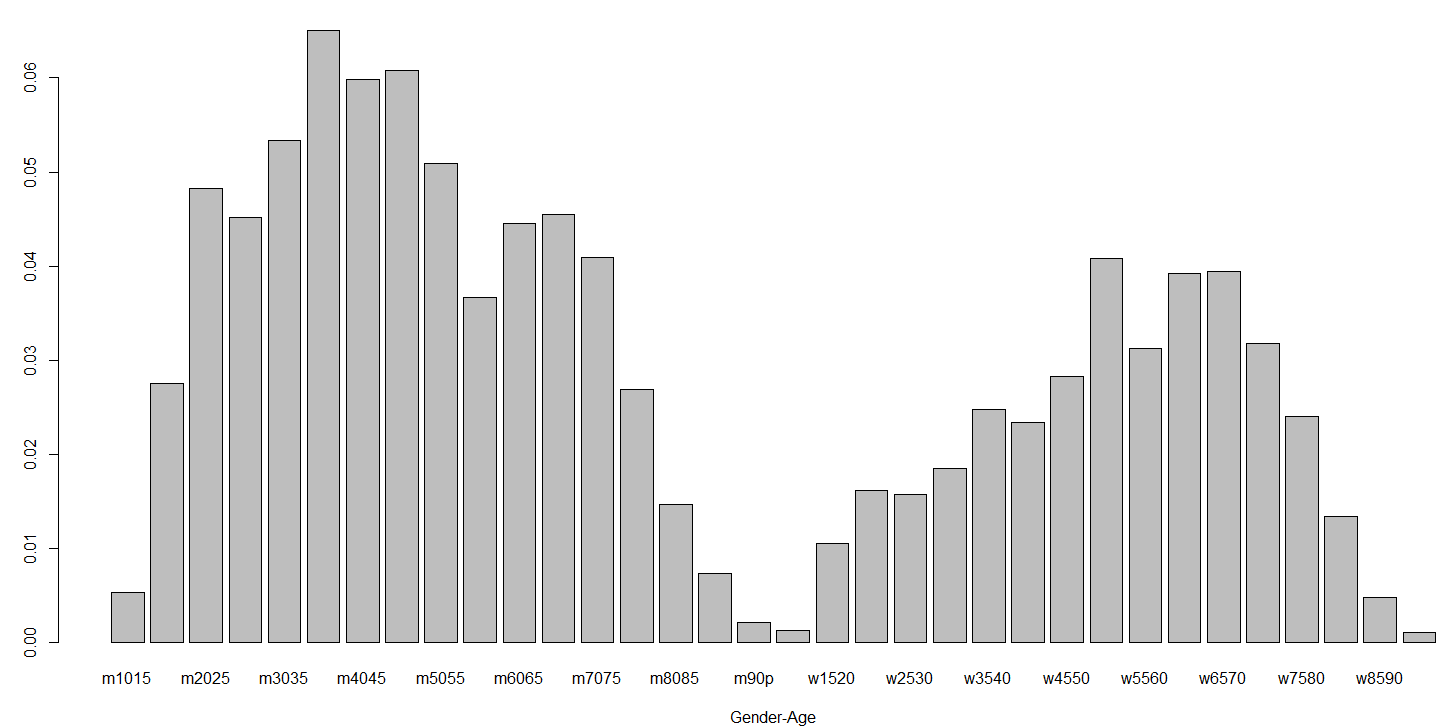
0.0021 0.0013 0.0105 0.0162 0.0158 0.0185 0.0248 0.0234 0.0283 0.0408 0.0312 0.0392 0.0395 0.0318 0.0241 0.0135

w8590 w90p

0.0048 0.0011

We can use these proportions to examine the gender-age distribution of suicides in W. Germany during the study period.

> barplot(suicide.rowprop,xlab="Gender-Age")



Next compute the proportion of individuals in each suicide method category and display them in a bar graph.

> suicide.colprop = apply(suicide.prop,2,sum)

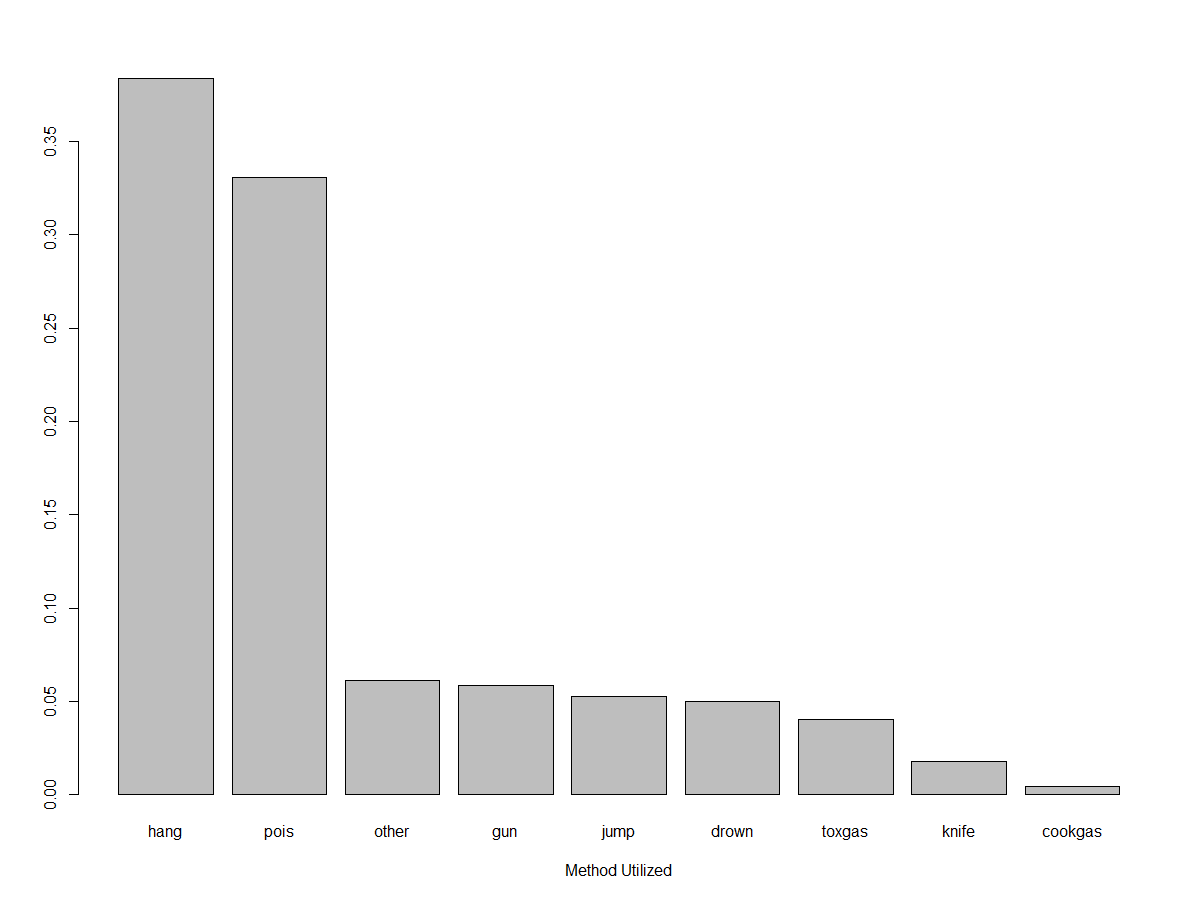
> suicide.colprop

pois cookgas toxgas hang drown gun knife jump other

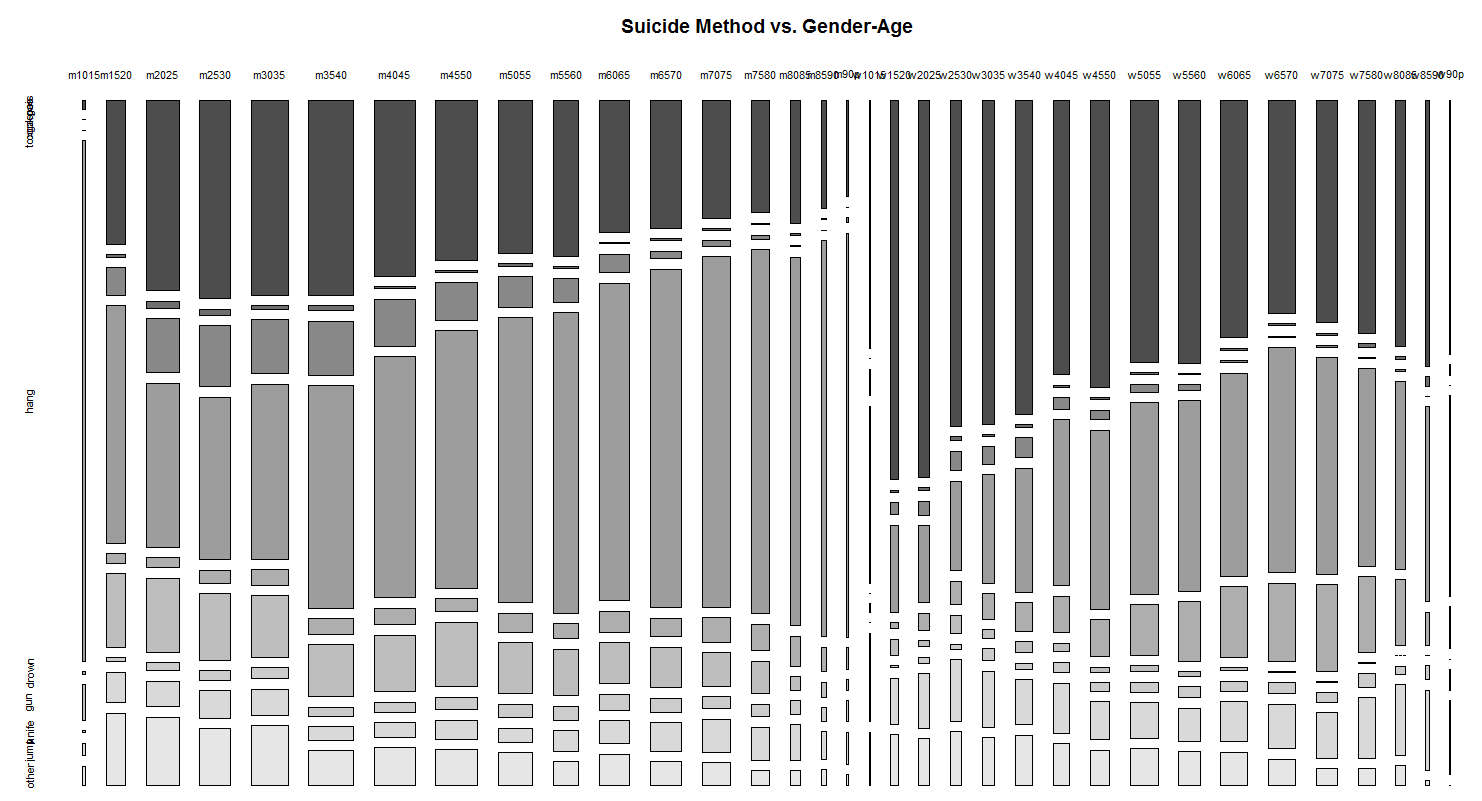
0.33075 0.00476 0.04056 0.38370 0.04992 0.05882 0.01791 0.05252 0.06107

> suicide.colprop = sort(suicide.colprop,decreasing=T)

> barplot(suicide.colprop,xlab=”Method Utilized”)



Finally, we use a mosaic plot to display the relationship between the gender-age categories and the methods of suicide employed. In order to do this we must convert the Suicide data frame to a matrix.

> suicide.mat = as.matrix(Suicide)  
> mosaicplot(suicide.mat,color=T,main=”Suicide Method vs. Gender-Age”)  


The mosaic plot above gives the breakdown of method chosen within each age/sex category. As some of the age/sex categories have very few people in them, the results are a bit hard to read. However, we can see some general trends with age and differences in methods chosen across gender. We will now see how correspondence analysis can be used to visualize these same relationships between age/sex and method used.

**6.2 – Mathematical Details of Simple Correspondence Analysis**

The mathematics behind this method of summarizing a contingency table involves performing a SVD of the “residuals” from chi-square test of independence. Pearson’s Chi-Square Test of Independence is based upon the -statistic which compares the observed frequencies in our contingency to the frequencies we would expect to see if the two nominal/ordinal variables being examined were truly independent. The statistic has a chi-square distribution with

.

where,

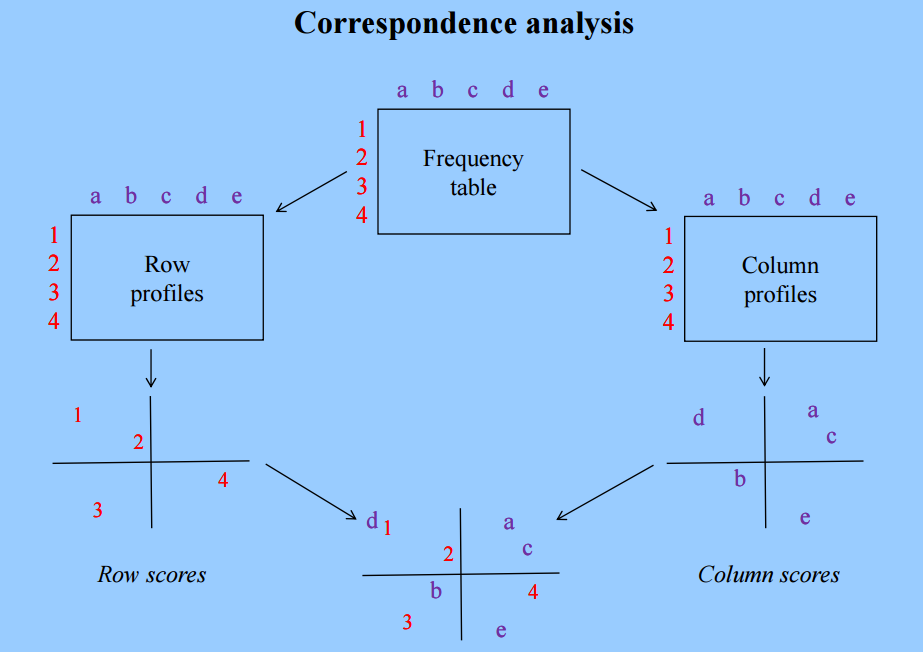
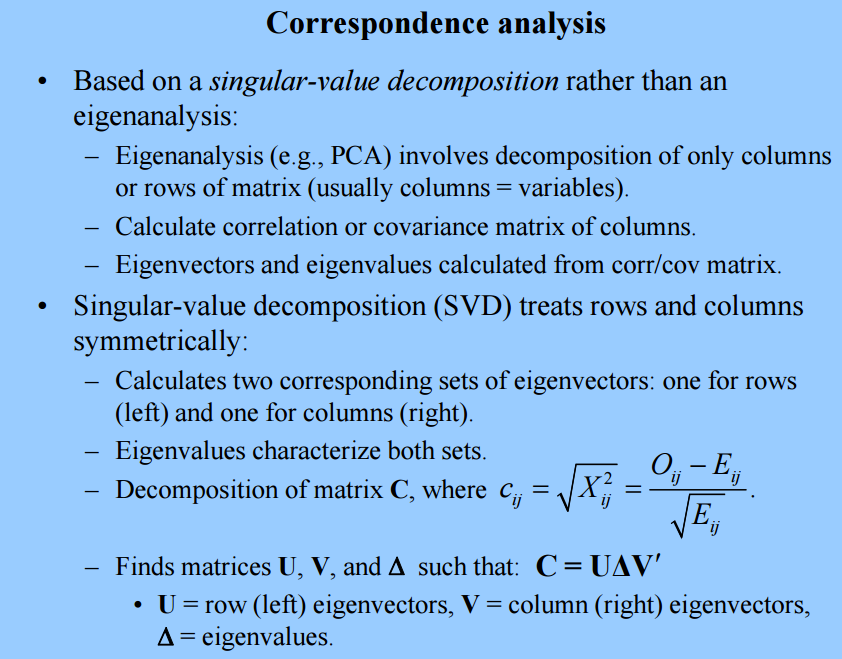
> chisq.test(Suicide)

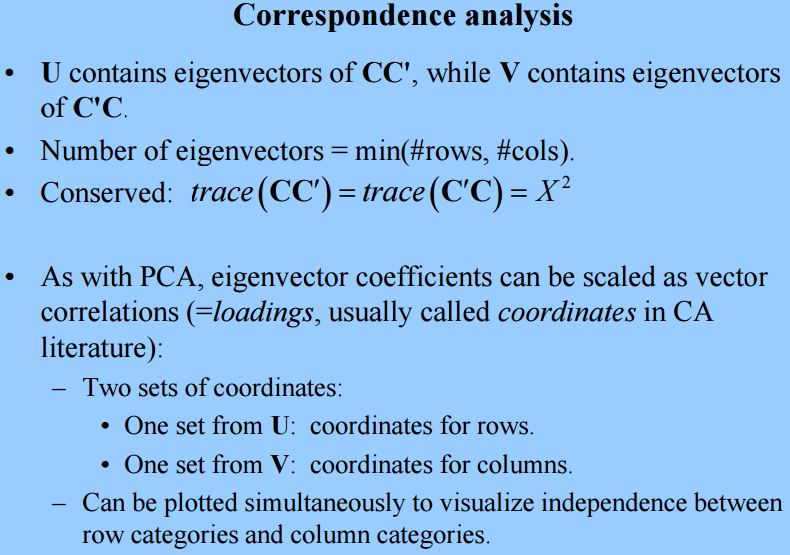
Pearson's Chi-squared test

data: Suicide

X-squared = 10061, df = 264, p-value < 2.2e-16

Here is general idea:



In terms of proportions the matrix C has elements

The dimension of *C* for the suicide data is 34 rows and 9 columns.   
  
In order to form the elements of *C* we need to form the following matrices in R:

> suicide.prop = Suicide/sum(Suicide)

> suicide.rowprop <- apply(suicide.prop,1,sum)  
> suicide.colprop <- apply(suicide.prop,2,sum)

> rdiag = diag(1/sqrt(suicide.rowprop))

> cdiag = diag(1/sqrt(suicide.colprop))

> suicide.prop = as.matrix(suicide.prop)  
> suicide.rowprop = as.matrix(suicide.rowprop)

> suicide.colprop = as.matrix(suicide.colprop)

> C = rdiag%\*%(as.matrix(suicide.prop) -suicide.rowprop%\*%t(suicide.colprop))%\*%cdiag

> C

[,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]

[1,] -0.04034 -5.1e-03 -0.01475 0.0571 -1.5e-02 0.00025 -0.00788 -0.01006 -0.00874

[2,] -0.02669 6.6e-05 0.00436 0.0032 -2.6e-02 0.04356 -0.01286 -0.00135 0.03941

[3,] -0.00582 2.5e-02 0.05327 -0.0393 -3.2e-02 0.05846 -0.00696 -0.00954 0.04602

[4,] -0.00069 1.9e-02 0.06409 -0.0390 -2.7e-02 0.04635 -0.00328 -0.00657 0.02850

[5,] -0.00312 4.1e-03 0.05744 -0.0346 -2.5e-02 0.04172 0.00074 -0.00922 0.03555

[6,] -0.00309 1.1e-02 0.06342 -0.0056 -2.8e-02 0.02919 -0.00487 -0.03330 -0.00384

[7,] -0.01679 -2.4e-03 0.04626 0.0066 -2.4e-02 0.03554 -0.00226 -0.02987 0.00121

[8,] -0.02820 -7.1e-03 0.02735 0.0176 -3.1e-02 0.04950 0.00583 -0.02216 -0.00218

[9,] -0.03048 1.4e-03 0.01093 0.0330 -1.8e-02 0.02409 0.00848 -0.02876 -0.00644

[10,] -0.02423 -4.7e-03 -0.00094 0.0358 -2.1e-02 0.01481 0.00820 -0.01554 -0.01665

[11,] -0.04130 -8.1e-03 -0.00971 0.0492 -1.4e-02 0.00842 0.01310 -0.01254 -0.00810

[12,] -0.04396 -4.5e-03 -0.02938 0.0616 -1.8e-02 0.00794 0.00839 -0.00460 -0.01879

[13,] -0.04763 -7.2e-03 -0.03105 0.0651 -7.7e-03 -0.00266 0.00351 0.00194 -0.01913

[14,] -0.04124 -4.7e-03 -0.02791 0.0589 -4.3e-03 -0.00287 0.00378 0.00198 -0.02474

[15,] -0.02689 -3.9e-03 -0.02290 0.0448 -2.5e-05 -0.01791 0.00463 0.00334 -0.01806

[16,] -0.02283 -2.8e-03 -0.01732 0.0382 -3.5e-03 -0.01184 0.00321 -0.00252 -0.01240

[17,] -0.01375 -3.2e-03 -0.00726 0.0215 -3.0e-03 -0.00782 0.00603 0.00012 -0.00810

[18,] 0.00504 -2.5e-03 0.00063 -0.0052 -8.0e-03 -0.00651 -0.00479 0.01476 0.00393

[19,] 0.05350 -1.8e-03 -0.01067 -0.0396 -1.8e-02 -0.01356 -0.01100 0.01087 0.00950

[20,] 0.06551 -2.3e-04 -0.01098 -0.0524 -1.3e-02 -0.02542 -0.00711 0.02111 0.00859

[21,] 0.04598 4.3e-03 -0.00523 -0.0476 -6.0e-03 -0.01443 -0.00898 0.02742 0.01441

[22,] 0.04910 -5.4e-03 -0.00752 -0.0446 -4.4e-03 -0.02505 0.00244 0.02430 0.00999

[23,] 0.05230 -2.2e-03 -0.00563 -0.0452 -9.9e-04 -0.02640 -0.00859 0.01498 0.01424

[24,] 0.03323 -3.4e-03 -0.01613 -0.0265 7.7e-03 -0.02795 -0.00022 0.02030 0.00506

[25,] 0.04252 -1.9e-03 -0.02056 -0.0235 9.5e-03 -0.03482 -0.00496 0.00782 -0.00174

[26,] 0.03645 -4.5e-03 -0.02866 -0.0212 3.2e-02 -0.04093 -0.00127 0.00616 -0.00052

[27,] 0.03236 -7.6e-03 -0.02818 -0.0188 3.8e-02 -0.03670 0.00025 0.00229 -0.00399

[28,] 0.02167 -2.6e-03 -0.03611 -0.0149 6.0e-02 -0.04372 -0.00163 0.01270 -0.00508

[29,] 0.00771 -2.7e-03 -0.03813 -0.0029 7.1e-02 -0.04662 0.00033 0.01899 -0.01457

[30,] 0.01187 -3.1e-03 -0.03384 -0.0066 7.5e-02 -0.04284 -0.00259 0.01851 -0.02445

[31,] 0.01511 3.3e-03 -0.03066 -0.0139 5.3e-02 -0.03665 0.00551 0.03270 -0.02119

[32,] 0.01567 -9.6e-04 -0.02176 -0.0135 3.1e-02 -0.02814 -0.00340 0.03289 -0.01357

[33,] 0.01336 1.1e-02 -0.01398 -0.0067 1.5e-03 -0.01684 0.00287 0.02433 -0.01496

[34,] 0.00514 6.1e-03 -0.00660 -0.0027 3.0e-03 -0.00795 0.00421 0.01005 -0.00810

> suicide.sva = svd(C)

> suicide.sva

$d

[1] 3.1e-01 2.7e-01 1.0e-01 7.1e-02 5.1e-02 3.1e-02 2.6e-02 2.4e-02 4.9e-17

$u

[,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]

[1,] -0.138 -0.214 -0.1394 0.0413 -0.1164 -0.43758 0.1539 0.2091 0.290

[2,] -0.163 0.079 -0.0757 0.5724 -0.3123 0.00765 0.1566 0.2196 -0.182

[3,] -0.179 0.329 0.1278 0.3399 0.0012 -0.08003 0.2709 -0.3765 0.051

[4,] -0.151 0.313 0.1747 0.0864 0.2368 -0.13647 -0.0063 -0.1096 0.035

[5,] -0.144 0.283 0.1479 0.1202 0.0014 -0.06274 -0.4603 0.0439 0.064

[6,] -0.182 0.173 0.1457 -0.4865 0.1409 -0.31954 0.2057 0.0637 -0.247

[7,] -0.201 0.098 0.1167 -0.2722 -0.0406 -0.02373 -0.0313 0.3168 -0.016

[8,] -0.235 0.034 0.0105 -0.0738 0.0104 0.52413 -0.0240 0.2999 0.066

[9,] -0.185 -0.072 0.0240 -0.1595 -0.1209 0.14021 0.0813 -0.2833 -0.134

[10,] -0.141 -0.111 -0.1105 -0.1643 0.0372 0.23640 0.0291 -0.0318 0.357

[11,] -0.152 -0.187 -0.0756 0.0139 -0.0703 0.06487 -0.3393 -0.1901 -0.237

[12,] -0.144 -0.262 -0.2032 0.0911 0.0372 0.17603 0.1047 -0.1379 -0.124

[13,] -0.114 -0.298 -0.1514 0.1382 0.0593 -0.09030 -0.0064 0.0797 -0.282

[14,] -0.095 -0.278 -0.1155 0.0606 0.1386 -0.00925 0.1068 0.0786 0.081

[15,] -0.032 -0.226 -0.1065 -0.0092 0.0921 -0.25406 -0.1270 -0.0768 0.036

[16,] -0.046 -0.178 -0.0979 -0.0266 0.0011 -0.22181 -0.0604 -0.1129 0.031

[17,] -0.028 -0.100 -0.0556 -0.0350 0.0547 -0.09727 -0.2136 -0.0957 -0.042

[18,] 0.026 0.026 -0.1106 0.0907 0.1380 -0.20771 -0.1640 0.2407 -0.040

[19,] 0.147 0.169 -0.3294 -0.0714 -0.1034 0.10104 0.2820 0.1321 -0.019

[20,] 0.216 0.197 -0.3326 -0.0854 0.0294 0.07245 0.1003 0.0221 -0.180

[21,] 0.169 0.181 -0.1998 0.1374 0.1330 -0.02952 0.0893 0.0823 -0.026

[22,] 0.190 0.151 -0.2253 -0.0026 0.0818 0.09283 -0.3897 -0.0101 -0.061

[23,] 0.192 0.166 -0.1858 -0.0445 -0.1028 -0.15594 -0.0633 0.0689 -0.114

[24,] 0.173 0.053 -0.1347 0.0404 0.0177 -0.04242 -0.2173 -0.0313 0.211

[25,] 0.193 0.038 -0.1544 -0.1552 -0.1358 -0.09728 0.1162 -0.0831 0.172

[26,] 0.227 -0.017 0.0018 -0.0915 -0.2722 -0.04405 -0.0122 -0.1679 -0.310

[27,] 0.217 -0.035 0.0771 -0.1144 -0.3073 0.10523 -0.0171 -0.0742 0.086

[28,] 0.254 -0.099 0.2271 0.0607 -0.2209 -0.04252 0.0378 -0.0956 0.410

[29,] 0.247 -0.176 0.3165 0.0995 -0.0545 -0.08752 -0.0331 0.0027 -0.180

[30,] 0.255 -0.166 0.3736 0.0043 0.0413 0.05583 0.1614 0.2438 -0.229

[31,] 0.235 -0.109 0.1958 0.1062 0.3374 0.16690 0.0060 -0.0443 -0.037

[32,] 0.182 -0.056 0.0486 0.1305 0.3165 -0.00084 0.0045 0.3027 0.108

[33,] 0.096 -0.023 -0.1170 0.0262 0.4368 0.01442 0.1944 -0.2209 -0.035

[34,] 0.045 -0.017 -0.0291 -0.0017 0.2091 0.05107 0.0649 -0.2030 0.078

$v

[,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]

[1,] 0.486 0.380 -0.290 -0.37 -0.115 0.154 0.180 0.023 0.575

[2,] -0.026 0.097 0.119 0.10 0.322 -0.275 0.568 -0.677 0.069

[3,] -0.346 0.458 0.383 -0.36 0.314 -0.376 -0.295 0.167 0.201

[4,] -0.359 -0.633 -0.153 -0.10 -0.054 -0.221 0.026 0.031 0.619

[5,] 0.416 -0.242 0.824 0.10 -0.131 0.076 0.042 0.073 0.223

[6,] -0.509 0.252 0.144 0.28 0.013 0.620 0.325 0.173 0.243

[7,] -0.030 -0.075 0.037 -0.11 0.157 0.470 -0.584 -0.613 0.134

[8,] 0.278 -0.016 -0.175 0.55 0.669 -0.043 -0.185 0.241 0.229

[9,] -0.060 0.329 -0.032 0.56 -0.541 -0.313 -0.273 -0.207 0.247

> delta = suicide.sva$d

> U = suicide.sva$u[,1:2]

> V = suicide.sva$v[,1:2]

> U[,1] = delta[1]\*U[,1]/sqrt(suicide.rowprop)

> U[,2] = delta[2]\*U[,2]/sqrt(suicide.rowprop)

> V[,1] = delta[1]\*V[,1]/sqrt(suicide.colprop)

> V[,2] = delta[2]\*V[,2]/sqrt(suicide.colprop)

> CA = rbind(U,V)

> inertia = sum(delta^2)

> inertia

[1] 0.19

> per1 = delta[1]^2/inertia

> per1

[1] 0.52 🡨 the first coordinate explains 52% of the variation

> per2 = delta[2]^2/inertia

> per2

[1] 0.38 🡨 the second coordinate explains 38% of the variation

> options(digits=5)

> plot(CA[,1],CA[,2],type="n",xlab = paste("coord 1% inertia =",format(per1\*100)),

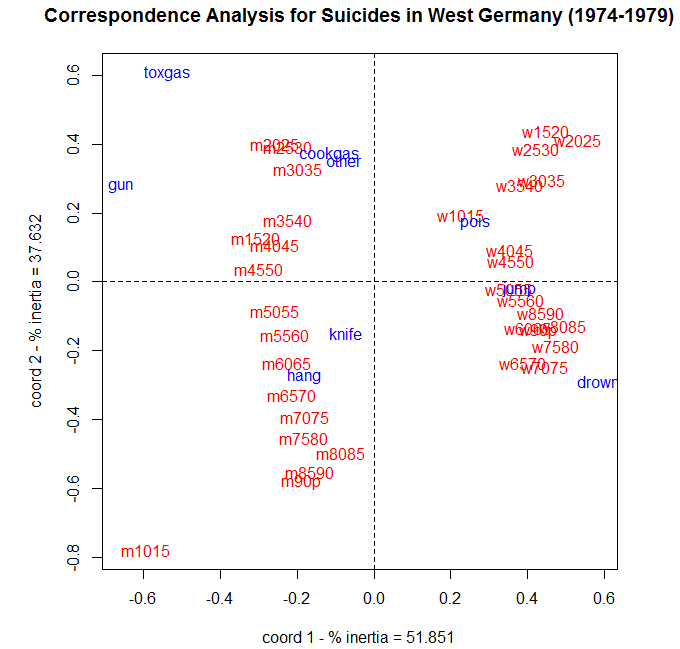
+ ylab = paste("coord 2% inertia =",format(per2\*100)))

> text(CA[,1],CA[,2],labels=c(dimnames(Suicide)[[1]],dimnames(Suicide)[[2]]))

> abline(h=0,lty=2)

> abline(v=0,lty=2)

> title(main="Correspondence Analysis for West German Suicides (1974 - 1977)")

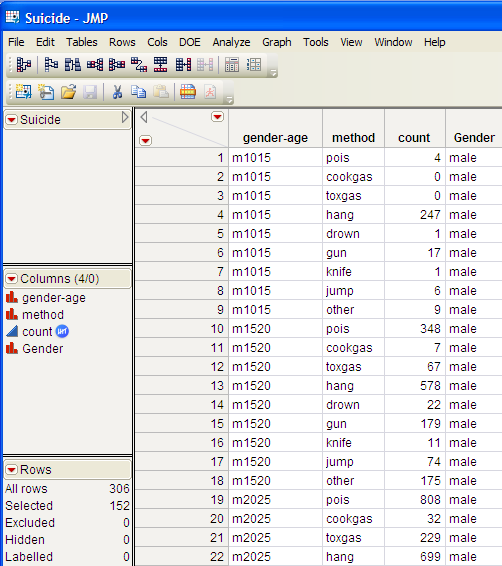


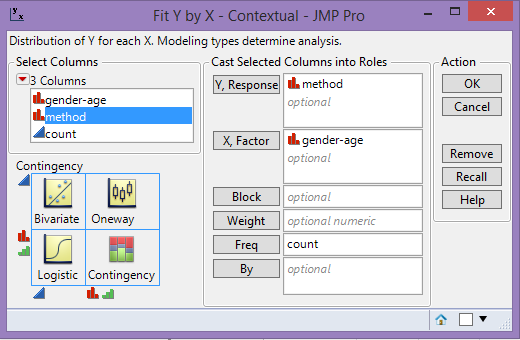
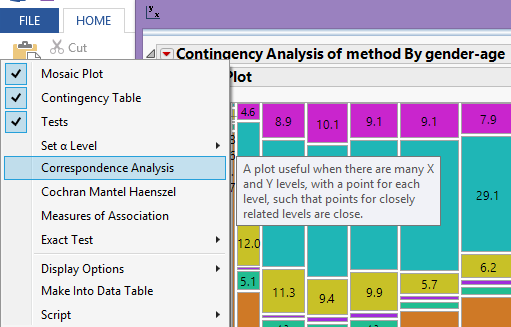
**General Function for Conducting Correspondence Analysis in R (corresp)**Here is function that performs all of the above operations given a two-way contingency table as input. You can change the title to whatever you would like to be. The only argument required is a data frame or matrix (x) representing the cells of a contingency table.

|  |
| --- |
| Custom R Function |
| corresp = function(x,title="Correspondence Analysis") {  x.prop <- as.matrix(x/sum(x))  x.row <- apply(x.prop,1,sum)  x.col <- apply(x.prop,2,sum)  rdiag <- diag(1/sqrt(x.row))  cdiag <- diag(1/sqrt(x.col))  x.row <- as.matrix(x.row)  x.col <- as.matrix(x.col)  E <- rdiag%\*%(x.prop - x.row%\*%t(x.col))%\*%cdiag  x.sva <- svd(E)  delta <- x.sva$d  U <- x.sva$u[,1:2]  V <- x.sva$v[,1:2]  U[,1] <- delta[1]\*U[,1]/sqrt(x.row)  U[,2] <- delta[2]\*U[,2]/sqrt(x.row)  V[,1] <- delta[1]\*V[,1]/sqrt(x.col)  V[,2] <- delta[2]\*V[,2]/sqrt(x.col)  U <- rbind(U,V)  inertia <- sum(delta[delta>0]\*delta[delta>0])  per1 <- (delta[1]\*delta[1]/inertia)\*100  per2 <- (delta[2]\*delta[2]/inertia)\*100  dim1 <- dim(x)[1]  ds <- as.integer(dim1+1)  dim2 <- dim(x)[2]  dt <- dim1 + dim2  plot(U[,1],U[,2],type="n",xlab=paste("coord 1 - % inertia =",format(per1)),  ylab=paste("coord 2 - % inertia =",format(per2)))  text(U[1:dim1,1],U[1:dim1,2],labels=dimnames(x)[[1]],col=2)  text(U[ds:dt,1],U[ds:dt,2],labels=dimnames(x)[[2]],col=4)  abline(h=0,v=0,lty=2)  title(title)  } |

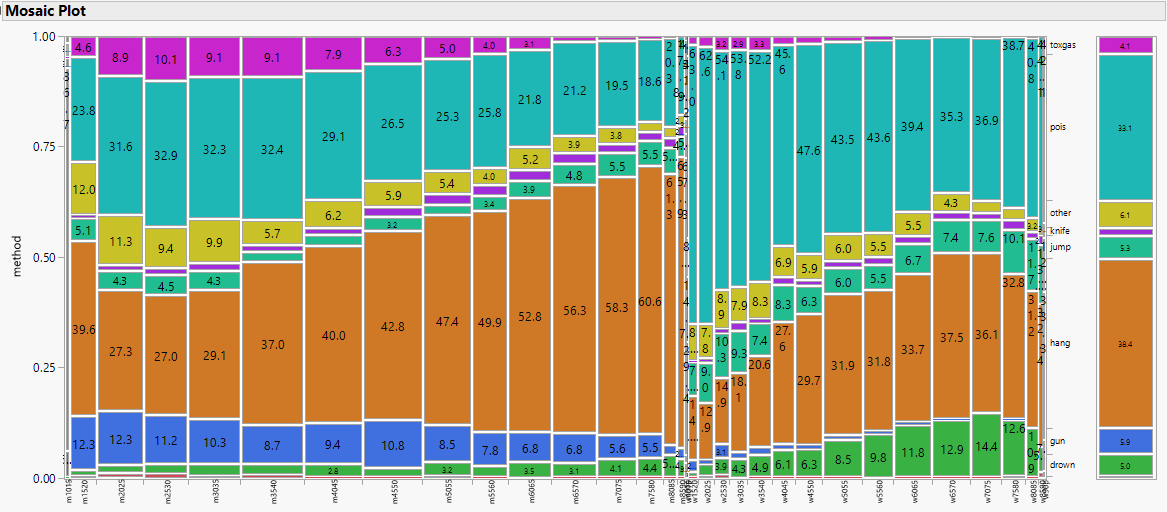
> corresp(Suicide,title=”Correspondence Analysis of Suicides in West Germany (1974-1979)”)

**6.3 - Correspondence Analysis in JMP**JMP has both simple and multiple correspondence analysis functionality.



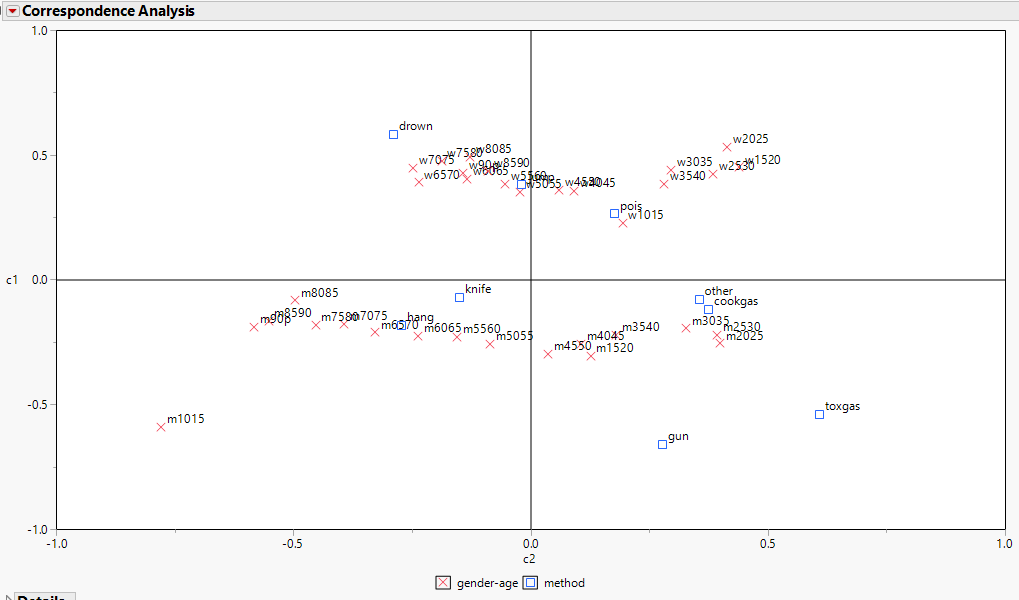
  


Here we have a much more attractive mosaic plot of the West Germany suicide data.

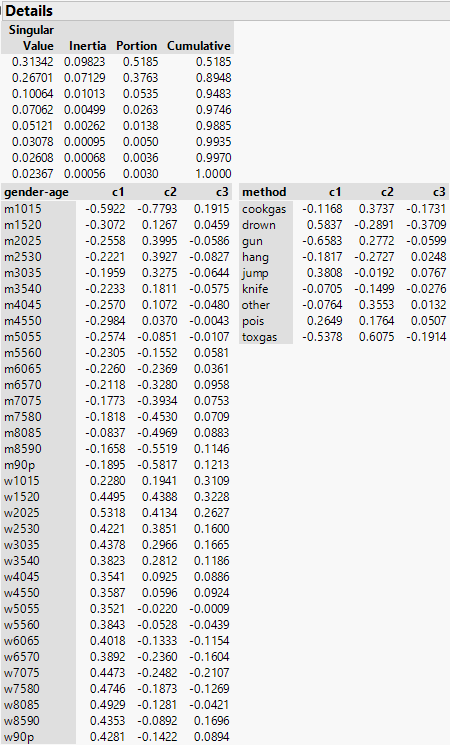
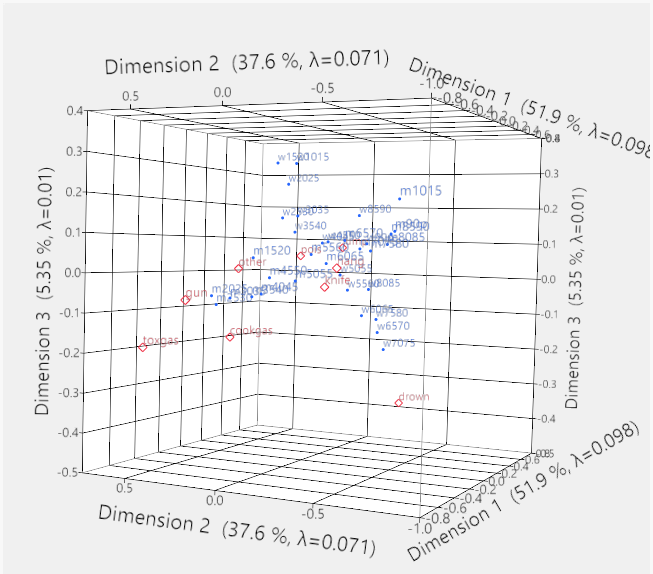


Right-click on the cells in the mosaic plot and select **Cell Labeling > Show Percents**.

The correspondence analysis plot is show below. Notice that JMP reverses the axes for the first two dimensions from our plot constructed in R above.



SVD details () 3-D Correspondence Analysis

We will examine multiple correspondence analysis in JMP later in this handout.