Multivariate exploration of the questionnaire and typology of the surveyed people

The results are provided by the
EnQuireR package

July 28, 2010

Contents

List of Figures

1 Quick overview of the questionnaire

The analysis was performed on 120 individuals described by 13 variables:

- frequency.buy.chocolate (never , often , rarely , sometimes , very often)
- frequency.eat.chocolate (many times a day , many times a week , never , once a month , once a week)
- handle.lack.of.chocolate (absolutely , pretty much , sometimes , almost never , never)
- kind.chocolate (almond chocolate , black chocolate , cooking chocolate , lined chocolate , milk chocolate , white chocolate)
- kind.brands (discount , retailer , upmarket , well-known)
- buy.fair.chocolate (never , rarely , sometimes , often , always)
- Where (bed , dining room , others , outside , TV , work)
- When (after meal , after work , before sleep , snack , working)
- side.drink (alcohol , coffee , juices , nothing , tea , water , with milk)
- savoured.alone (disagree , slightly disagree , neither agree nor disagree , slightly agree , agree)
- savoured.with.people (disagree , slightly disagree , neither agree nor disagree , slightly agree , agree)
- keep.chocolate (cupboard , fridge)
- eaten.without.other.food (disagree , slightly disagree , neither agree nor disagree , slightly agree , agree)

Moreover, the dataset contained 0% of missing values.

2 Multivariate exploration of the questionnaire

2.1 Graphical representations of the questionnaire

The following results are obtained by performing a Multiple Correspondence Analysis (MCA) on the previous 13 variables. This method provides two important graphical displays, a representation of the individuals (surveyed people) and a representation of the categories (answers given by the surveyed people). The first two main axes of variability explain 11.36% of the information contained in the dataset (6.06% for the first factorial axis and 5.3% for the second one). In some cases the analyst may want to introduce supplementary quantitative variables.

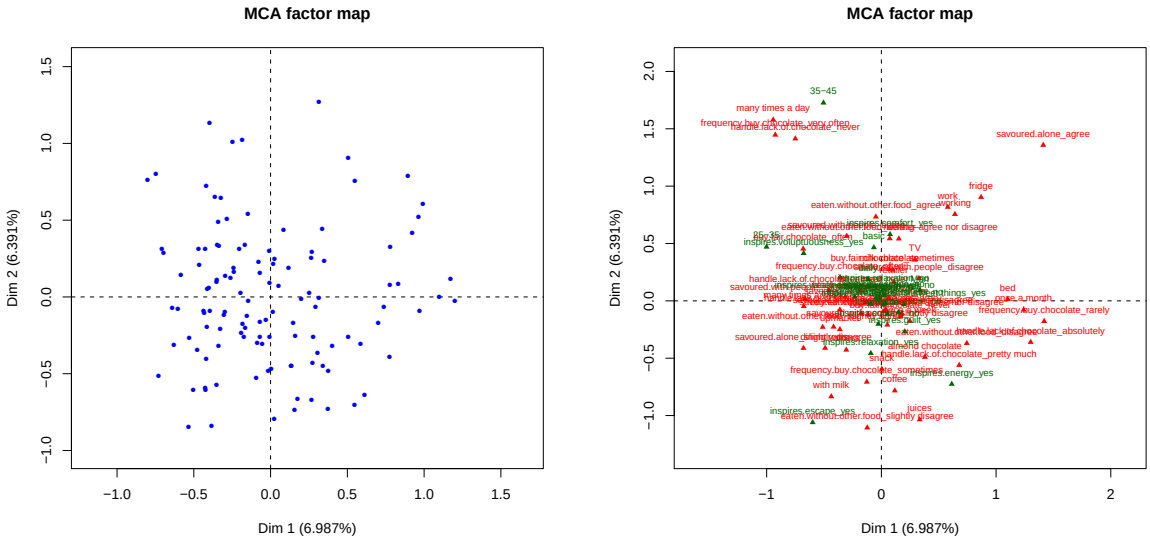

Figure 1: Representations of the individuals and of the categories on axes 1 and 2

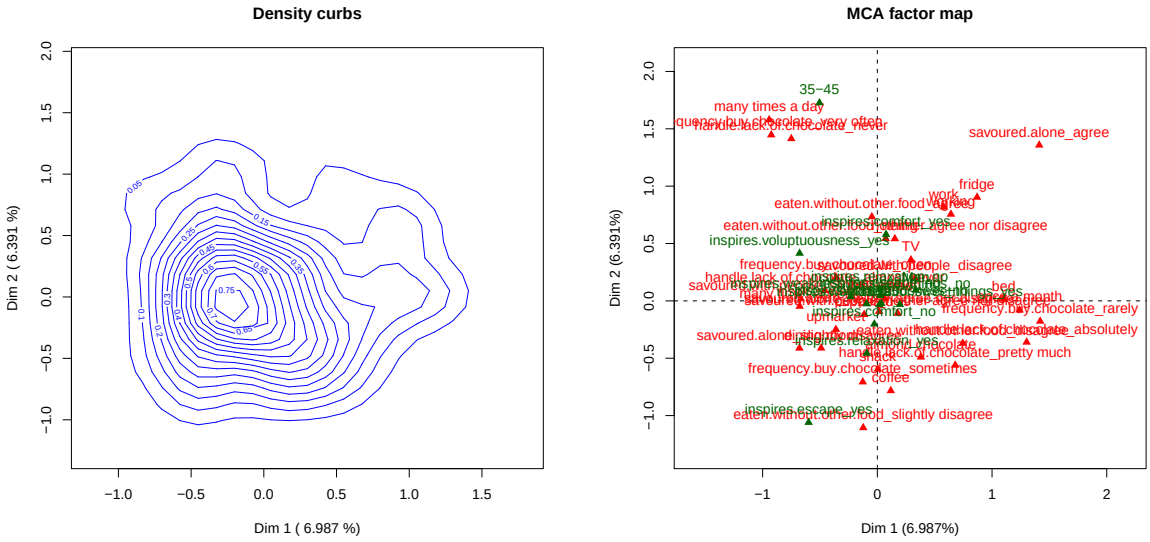

Figure 2: Representation of the individuals using density curbs and enhanced representation of the categories

2.2 Highlights on the two principal axes of variability

2.2.1 Characterization of the first factorial axis

The most meaningful variables characterizing the first factorial axis are:

- `frequency.eat.chocolate`
- `handle.lack.of.chocolate`
- `frequency.buy.chocolate`
- `side.drink`
- `When`
- `kind.chocolate`
- `keep.chocolate`
- `eaten.without.other.food`
- `buy.fair.chocolate`
- `savoured.alone`
- `Where`

The most meaningful categories characterizing the positive side of the first axis are:

- `handle.lack.of.chocolate_absolutely`
  - Contribution: 13.91
  - V-Test: 7.84
  - Frequency in the population: 15 %
- `frequency.eat.chocolate_never`
  - Contribution: 9.31
  - V-Test: 6.01
  - Frequency in the population: 3.33 %
- `frequency.buy.chocolate_rarely`
  - Contribution: 10.92
  - V-Test: 7.16
  - Frequency in the population: 20 %
- `alcohol`
  - Contribution: 6.49
  - V-Test: 4.96
  - Frequency in the population: 0.83 %
- `frequency.buy.chocolate_never`
  - Contribution: 2.65
  - V-Test: 3.19
  - Frequency in the population: 2.5 %
- `white chocolate`
  - Contribution: 3.17

  - V-Test: 3.48
  - Frequency in the population: 1.67 %
- `buy.fair.chocolate_always`
  - Contribution: 2.94
  - V-Test: 3.35
  - Frequency in the population: 1.67 %
- `working`
  - Contribution: 2.77
  - V-Test: 3.5
  - Frequency in the population: 15 %
- `savoured.alone_agree`
  - Contribution: 2.71
  - V-Test: 3.35
  - Frequency in the population: 9.17 %
- `fridge`
  - Contribution: 2.51
  - V-Test: 3.27
  - Frequency in the population: 11.67 %

The most meaningful categories characterizing the negative side of the first axis are:

- `frequency.eat.chocolate_many times a week`
  - Contribution: 5.42
  - V-Test: -5.82
  - Frequency in the population: 40 %
- `frequency.buy.chocolate_often`
  - Contribution: 2.53
  - V-Test: -3.87
  - Frequency in the population: 36.67 %
- `frequency.eat.chocolate_many times a day`
  - Contribution: 0.92
  - V-Test: -1.92
  - Frequency in the population: 5.83 %
- `frequency.buy.chocolate_very often`
  - Contribution: 1.96
  - V-Test: -2.89
  - Frequency in the population: 11.67 %
- `tea`
  - Contribution: 1.29
  - V-Test: -2.45
  - Frequency in the population: 19.17 %

- `handle.lack.of.chocolate_almost never`
  - Contribution: 2.1
  - V-Test: -3.41
  - Frequency in the population: 32.5 %
- `handle.lack.of.chocolate_never`
  - Contribution: 1.54
  - V-Test: -2.53
  - Frequency in the population: 10 %
- `nothing`
  - Contribution: 0.02
  - V-Test: 0.37
  - Frequency in the population: 51.67 %
- `handle.lack.of.chocolate_sometimes`
  - Contribution: 1.19
  - V-Test: -2.44
  - Frequency in the population: 25 %
- `frequency.buy.chocolate_sometimes`
  - Contribution: 0.3
  - V-Test: -1.26
  - Frequency in the population: 29.17 %

2.2.2 Characterization on the second factorial axis

The most meaningful variables characterizing the second factorial axis are:

- `handle.lack.of.chocolate`
- `frequency.buy.chocolate`
- `side.drink`
- `kind.chocolate`
- `When`
- `eaten.without.other.food`
- `keep.chocolate`
- `frequency.eat.chocolate`
- `savoured.with.people`
- `savoured.alone`
- `kind.brands`
- `Where`
- `buy.fair.chocolate`

The most meaningful categories characterizing the positive side of the second axis are:

- `handle.lack.of.chocolate_never`
  - Contribution: 12.99
  - V-Test: 6.88
  - Frequency in the population: 10 %
- `frequency.buy.chocolate_very often`
  - Contribution: 10.82
  - V-Test: 6.34
  - Frequency in the population: 11.67 %
- `white chocolate`
  - Contribution: 9.41
  - V-Test: 5.6
  - Frequency in the population: 1.67 %
- `alcohol`
  - Contribution: 6.05
  - V-Test: 4.47
  - Frequency in the population: 0.83 %
- `eaten.without.other.food_agree`
  - Contribution: 3.77
  - V-Test: 4.02
  - Frequency in the population: 23.33 %
- `fridge`
  - Contribution: 3.54
  - V-Test: 3.62
  - Frequency in the population: 11.67 %
- `savoured.with.people_agree`
  - Contribution: 3.85
  - V-Test: 3.75
  - Frequency in the population: 10 %
- `TV`
  - Contribution: 2.62
  - V-Test: 3.64
  - Frequency in the population: 35 %
- `before sleep`
  - Contribution: 2.67
  - V-Test: 3.01
  - Frequency in the population: 3.33 %
- `frequency.eat.chocolate_many times a day`
  - Contribution: 3.03
  - V-Test: 3.25
  - Frequency in the population: 5.83 %

The most meaningful categories characterizing the negative side of the second axis are:

- `snack`
  - Contribution: 3.79
  - V-Test: -4.14
  - Frequency in the population: 27.5 %
- `almond chocolate`
  - Contribution: 1.07
  - V-Test: -2.14
  - Frequency in the population: 23.33 %
- `handle.lack.of.chocolate_pretty much`
  - Contribution: 2.08
  - V-Test: -2.87
  - Frequency in the population: 17.5 %
- `milk chocolate`
  - Contribution: 0.2
  - V-Test: -1.01
  - Frequency in the population: 36.67 %
- `juices`
  - Contribution: 2.36
  - V-Test: -2.86
  - Frequency in the population: 5.83 %
- `savoured.alone_neither agree nor disagree`
  - Contribution: 1.94
  - V-Test: -3.42
  - Frequency in the population: 45.83 %
- `cupboard`
  - Contribution: 0.47
  - V-Test: -3.62
  - Frequency in the population: 88.33 %
- `savoured.with.people_neither agree nor disagree`
  - Contribution: 1.74
  - V-Test: -3.17
  - Frequency in the population: 43.33 %
- `coffee`
  - Contribution: 1.87
  - V-Test: -2.66
  - Frequency in the population: 13.33 %
- `frequency.eat.chocolate_once a week`
  - Contribution: 1.48
  - V-Test: -2.65
  - Frequency in the population: 30.83 %

3 Typology on the individuals

3.1 Choice of the number of clusters

The ascendant hierarchical clustering (AHC) lead to a partition made of 3 clusters. Those clusters are displayed in the following representations: a graphical representation of the individuals according to the cluster they belong to, a representation of the center of gravity of each group enhanced by a confidence ellipse, a representation of the individuals according to the cluster they belong to by the use of density curbs.

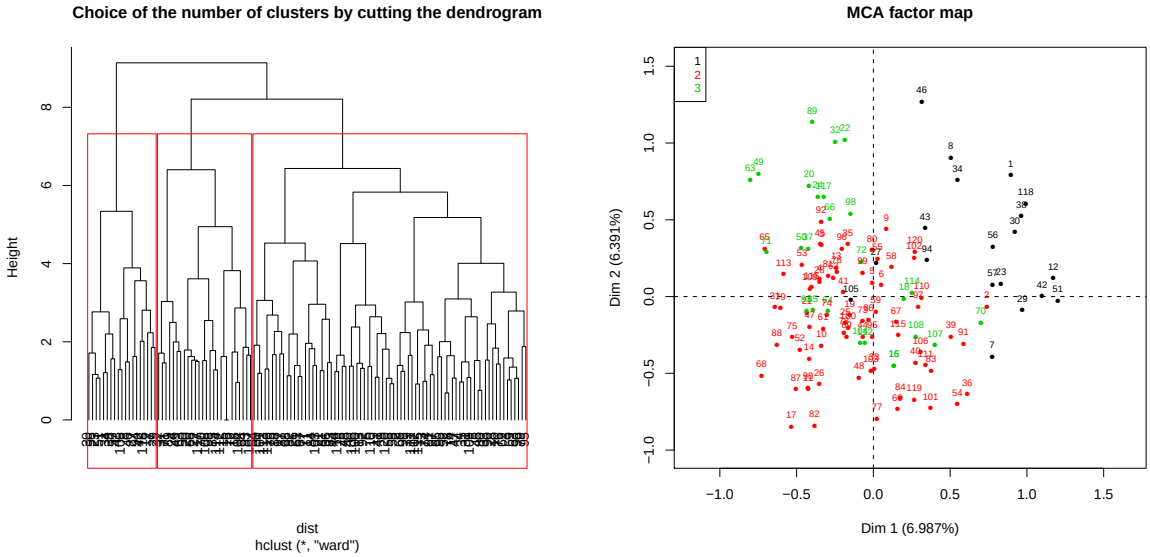

Figure 3: Number of clusters chosen by the analyst; representation of the individuals according to their cluster

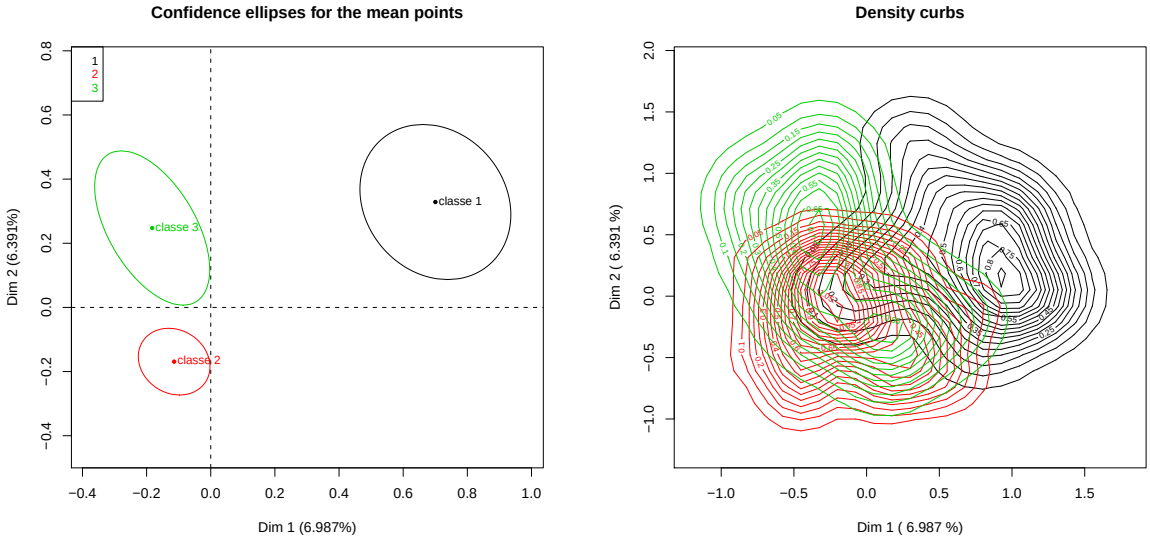

Figure 4: Centers of gravity with confidence ellipses; representation of the individuals according to their cluster with density curbs

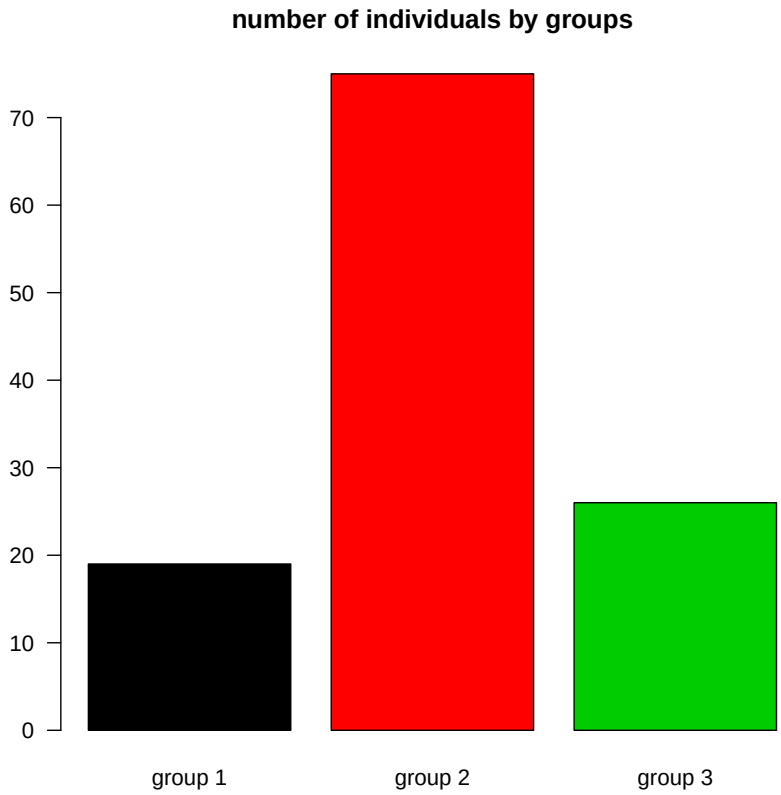

Figure 5: Number of individuals per cluster

3.2 Simultaneous comparison of the clusters with respect with the most relevant variables

3.2.1 Number of individuals by cluster for the variable savoured.alone

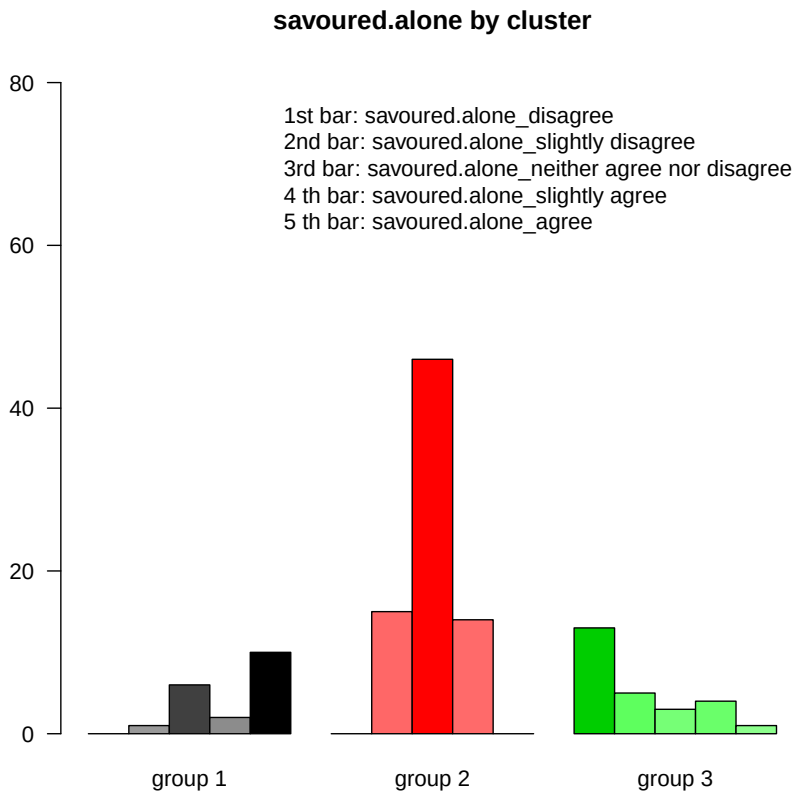

Figure 6: Variable savoured.alone

3.2.2 Number of individuals by cluster for the variable frequency.eat.chocolate

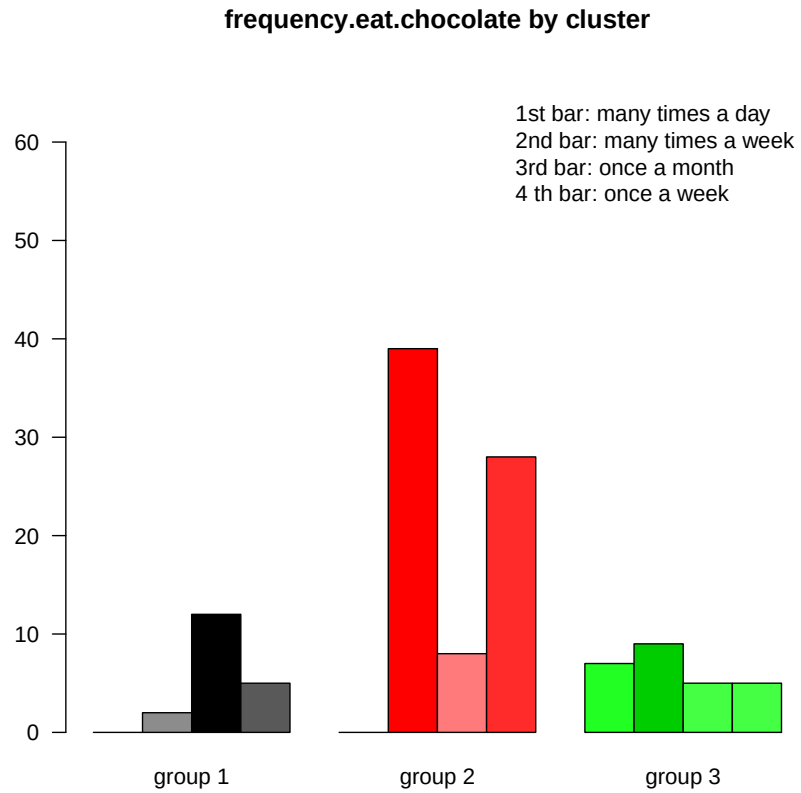

Figure 7: Variable frequency.eat.chocolate

3.2.3 Number of individuals by cluster for the variable When

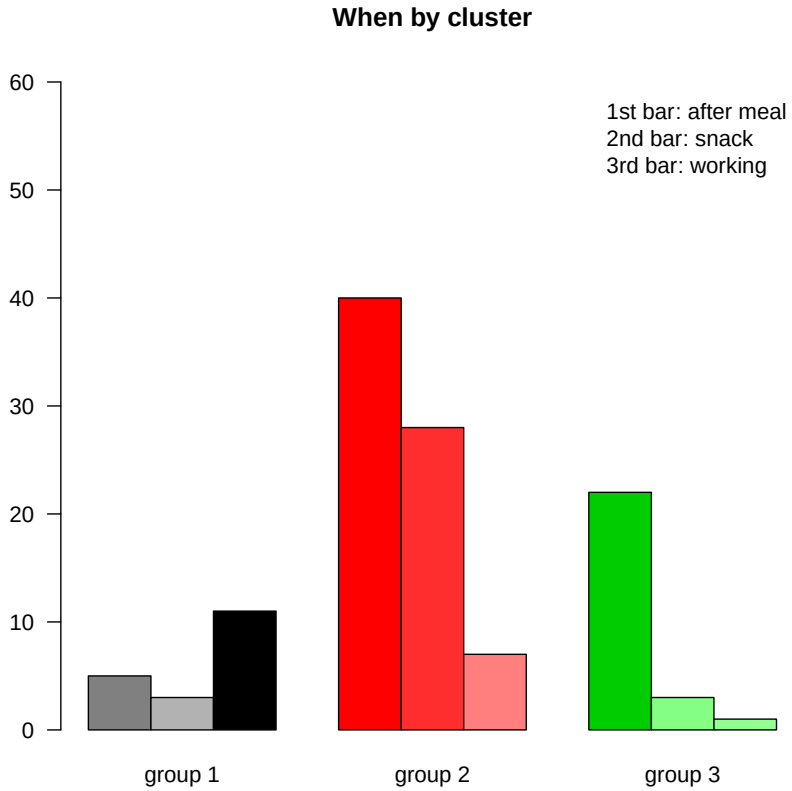

Figure 8: Variable When

3.2.4 Number of individuals by cluster for the variable frequency.buy.chocolate

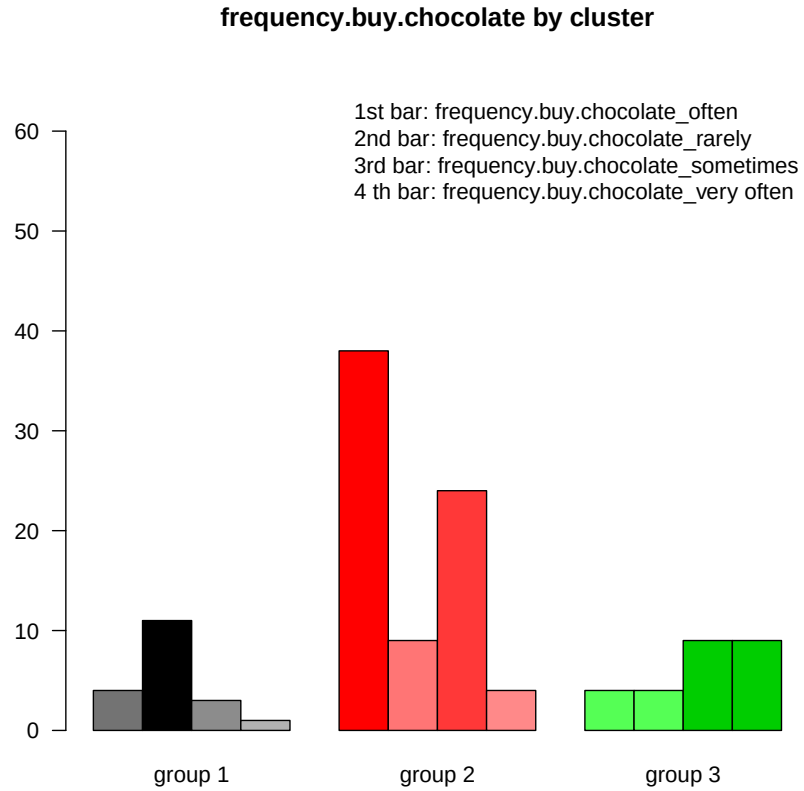

Figure 9: Variable frequency.buy.chocolate

3.2.5 Number of individuals by cluster for the variable keep.chocolate

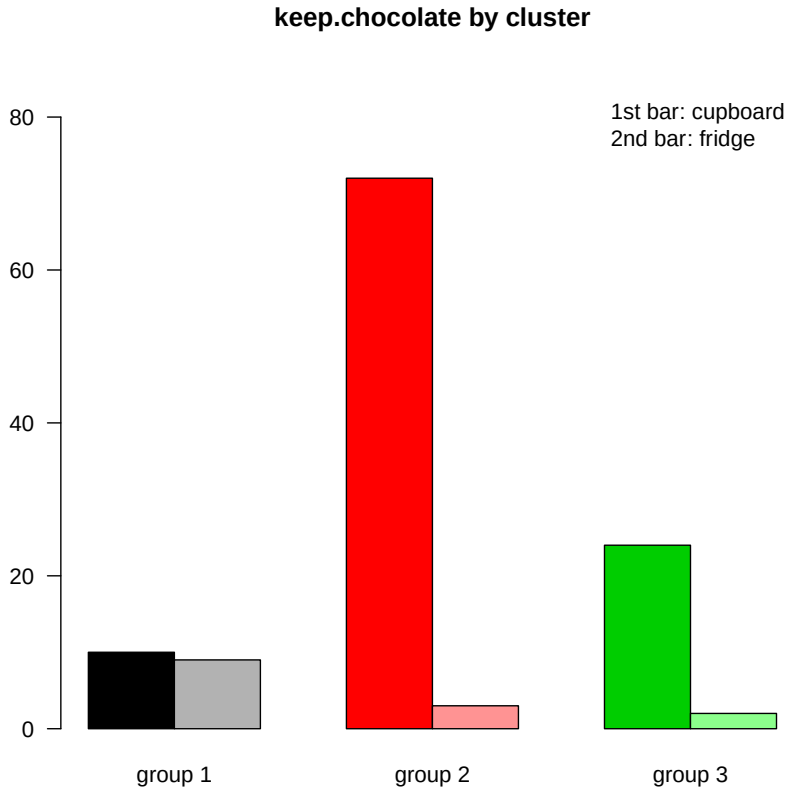

Figure 10: Variable keep.chocolate

3.2.6 Number of individuals by cluster for the variable Where

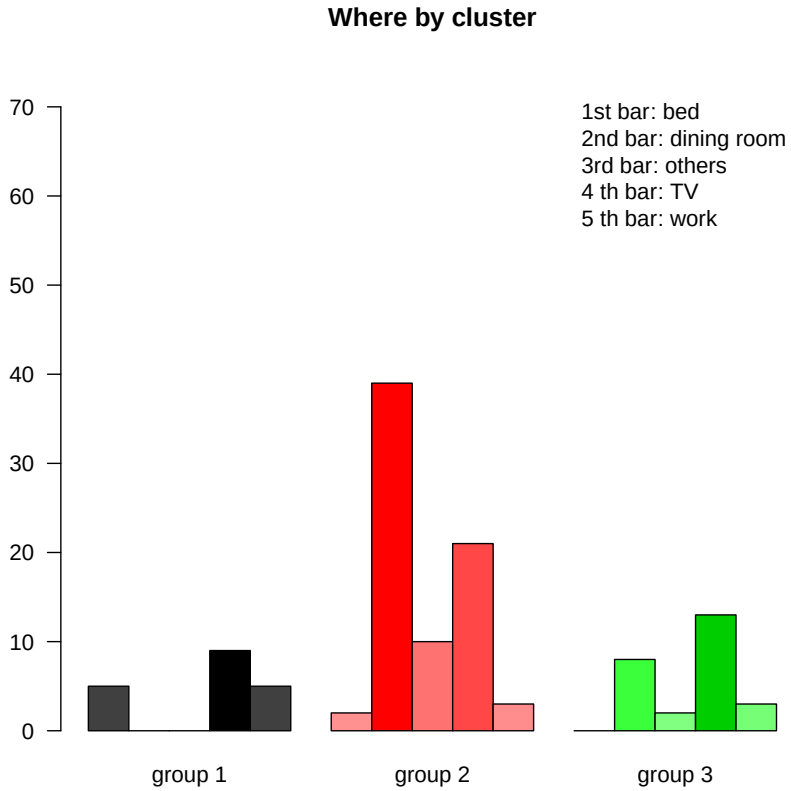

Figure 11: Variable Where

3.2.7 Number of individuals by cluster for the variable handle.lack.of.chocolate

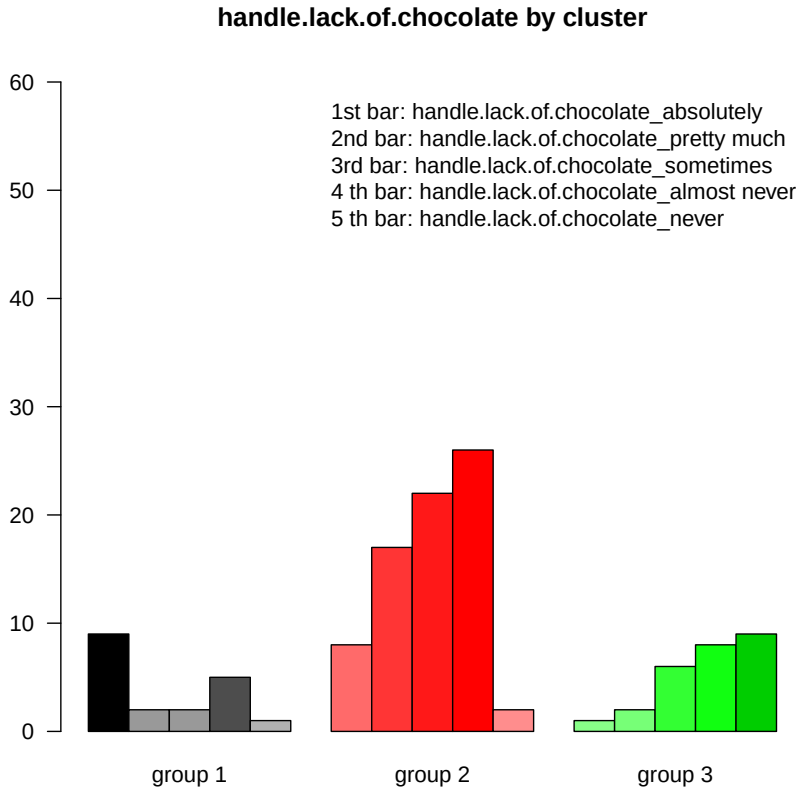

Figure 12: Variable handle.lack.of.chocolate

3.2.8 Number of individuals by cluster for the variable savoured.with.people

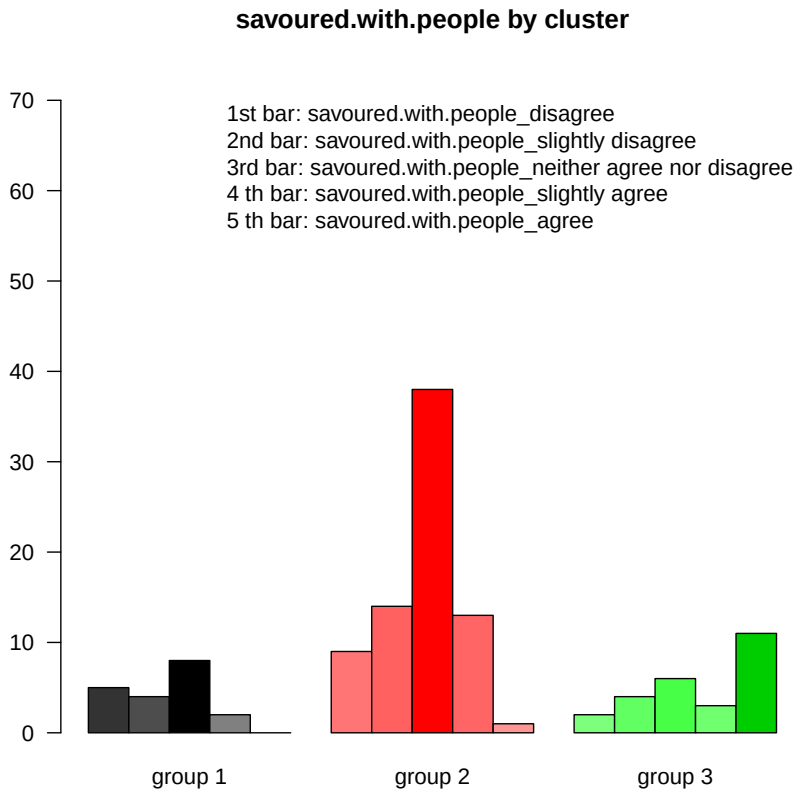

Figure 13: Variable savoured.with.people

3.2.9 Number of individuals by cluster for the variable sex

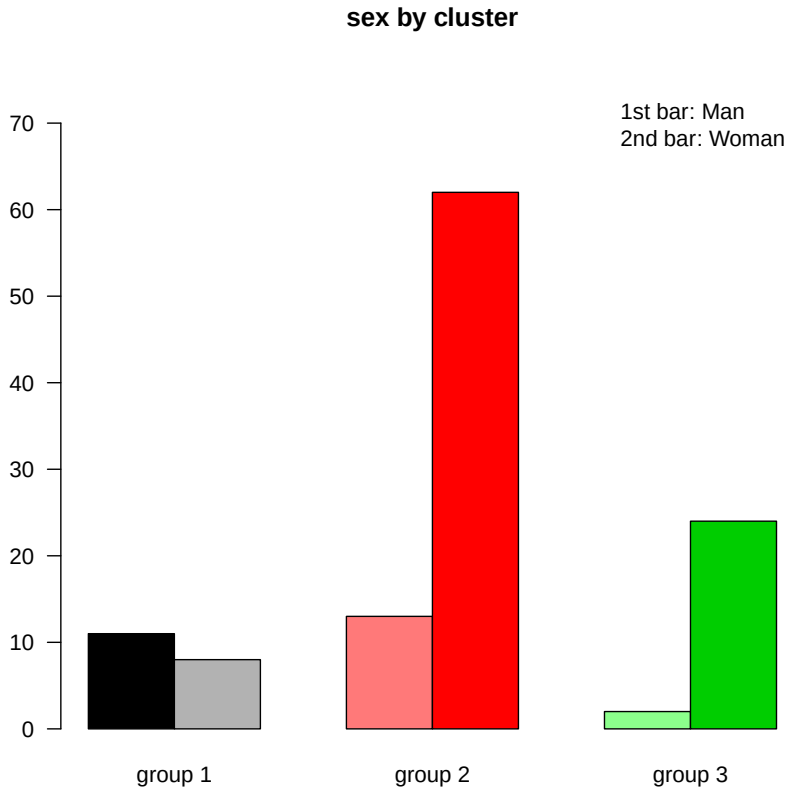

Figure 14: Variable sex

3.2.10 Number of individuals by cluster for the variable inspires.relaxation

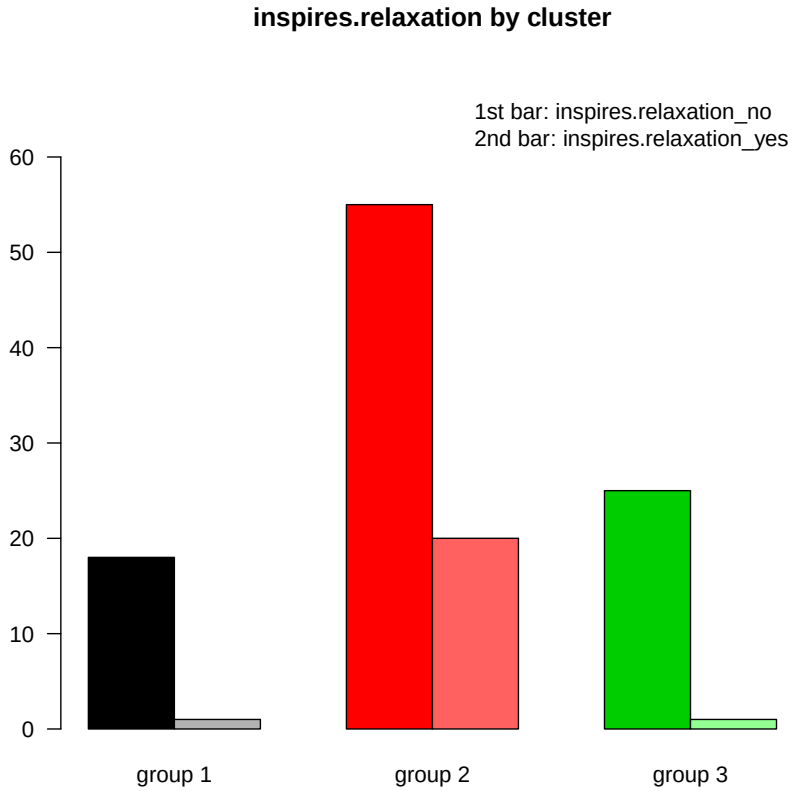

Figure 15: Variable inspires.relaxation

3.3 Automatic description of each cluster

The cluster 1 (19 individuals) includes the individuals possessing the following categories:

- `savoured.alone=savoured.alone_agree`

  9.17 % of the individuals possess this category in the global population versus 52.63 % in the cluster 1 .

  Moreover, 90.91 % of the individuals possessing this category belong to the cluster 1 .

- `When=working`

  15.83 % of the individuals possess this category in the global population versus 57.89 % in the cluster 1 .

  Moreover, 57.89 % of the individuals possessing this category belong to the cluster 1 .

- `frequency.eat.chocolate=once a month`

  20.83 % of the individuals possess this category in the global population versus 63.16 % in the cluster 1 .

  Moreover, 48 % of the individuals possessing this category belong to the cluster 1 .

- `keep.chocolate=fridge`

  11.67 % of the individuals possess this category in the global population versus 47.37 % in the cluster 1 .

  Moreover, 64.29 % of the individuals possessing this category belong to the cluster 1 .

- `frequency.buy.chocolate=frequency.buy.chocolate_rarely`

  20 % of the individuals possess this category in the global population versus 57.89 % in the cluster 1 .

  Moreover, 45.83 % of the individuals possessing this category belong to the cluster 1 .

- `sex=Man`

  21.67 % of the individuals possess this category in the global population versus 57.89 % in the cluster 1 .

  Moreover, 42.31 % of the individuals possessing this category belong to the cluster 1 .

- `handle.lack.of.chocolate=handle.lack.of.chocolate_absolutely`

  15 % of the individuals possess this category in the global population versus 47.37 % in the cluster 1 .

  Moreover, 50 % of the individuals possessing this category belong to the cluster 1 .

- `Where=bed`

  5.83 % of the individuals possess this category in the global population versus 26.32 % in the cluster 1 .

  Moreover, 71.43 % of the individuals possessing this category belong to the cluster 1 .

- `side.drink=nothing`

  52.5 % of the individuals possess this category in the global population versus 78.95 % in the cluster 1 .

  Moreover, 23.81 % of the individuals possessing this category belong to the cluster 1 .

- `Where=work`

  9.17 % of the individuals possess this category in the global population versus 26.32 % in the cluster 1 .

  Moreover, 45.45 % of the individuals possessing this category belong to the cluster 1 .

The cluster 2 (75 individuals) includes the individuals possessing the following categories:

- `savoured.alone=savoured.alone_neither agree nor disagree`

  45.83 % of the individuals possess this category in the global population versus 61.33 % in the cluster 2 .

  Moreover, 83.64 % of the individuals possessing this category belong to the cluster 2 .

- `Where=dining room`

  39.17 % of the individuals possess this category in the global population versus 52 % in the cluster 2 .

  Moreover, 82.98 % of the individuals possessing this category belong to the cluster 2 .

- `frequency.buy.chocolate=frequency.buy.chocolate_often`

  38.33 % of the individuals possess this category in the global population versus 50.67 % in the cluster 2 .

  Moreover, 82.61 % of the individuals possessing this category belong to the cluster 2 .

- `keep.chocolate=cupboard`

  88.33 % of the individuals possess this category in the global population versus 96 % in the cluster 2 .

  Moreover, 67.92 % of the individuals possessing this category belong to the cluster 2 .

- `inspires.relaxation=inspires.relaxation_yes`

  18.33 % of the individuals possess this category in the global population versus 26.67 % in the cluster 2 .

  Moreover, 90.91 % of the individuals possessing this category belong to the cluster 2 .

- `frequency.eat.chocolate=many times a week`

  41.67 % of the individuals possess this category in the global population versus 52 % in the cluster 2 .

  Moreover, 78 % of the individuals possessing this category belong to the cluster 2 .

- `When=snack`

  28.33 % of the individuals possess this category in the global population versus 37.33 % in the cluster 2 .

  Moreover, 82.35 % of the individuals possessing this category belong to the cluster 2 .

The cluster 3 (26 individuals) includes the individuals possessing the following categories:

- `savoured.alone=savoured.alone_disagree`

  10.83 % of the individuals possess this category in the global population versus 50 % in the cluster 3 .

  Moreover, 100 % of the individuals possessing this category belong to the cluster 3 .

- `savoured.with.people=savoured.with.people_agree`

  10 % of the individuals possess this category in the global population versus 42.31 % in the cluster 3 .

  Moreover, 91.67 % of the individuals possessing this category belong to the cluster 3 .

- `frequency.eat.chocolate=many times a day`

  5.83 % of the individuals possess this category in the global population versus 26.92 % in the cluster 3 .

  Moreover, 100 % of the individuals possessing this category belong to the cluster 3 .

- `handle.lack.of.chocolate=handle.lack.of.chocolate_never`

  10 % of the individuals possess this category in the global population versus 34.62 % in the cluster 3 .

  Moreover, 75 % of the individuals possessing this category belong to the cluster 3 .

- `frequency.buy.chocolate=frequency.buy.chocolate_very often`

  11.67 % of the individuals possess this category in the global population versus 34.62 % in the cluster 3 .

  Moreover, 64.29 % of the individuals possessing this category belong to the cluster 3 .

- `When=after meal`

  55.83 % of the individuals possess this category in the global population versus 84.62 % in the cluster 3 .

  Moreover, 32.84 % of the individuals possessing this category belong to the cluster 3 .

- `inspires.relaxation=inspires.relaxation_no`

  81.67 % of the individuals possess this category in the global population versus 96.15 % in the cluster 3 .

  Moreover, 25.51 % of the individuals possessing this category belong to the cluster 3 .