IS71069B Assignment 1: The Fast Food Paradox

John Downing

Introduction

This report relates to a Multiple Correspondence Analysis (MCA) performed on the *Fast Food Paradox* dataset from the EnQuireR website[1]. Per the notes accompanying this dataset, the paradox in question relates to the increasing popularity, in France, of fast food restaurants – despite their (presumed) negative image in the public perception. It is hoped that MCA will help identify the underlying factors which lead people to override these negative perceptions and eat fast food. This information could then be fed into public health policy, with a view to combating obesity by better understanding which issues people do or do not care about regards food quality and eating behaviour. The data could also be used by the fast food restaurants themselves, to address the perceived negative aspects of their products/services which may otherwise limit their appeal.

Method

166 students responded[2] to a survey of 48 questions concerning their perception of fast food restaurants and their pattern of consumption whilst at these restaurants. Some additional demographic information, such as gender, was collected – however this was not extensive.

The questionnaire results were downloaded in Excel format from the EnQuireR website[3], and exported in csv format for analysis with R. From the structure and ordering of the spreadsheet, it seems likely that questions 1 to 21 would have been presented to participants as relating to their perception of fast food restaurants, with questions 22 to 45 on their patterns of consumption and 46 to 48 as demographics. However, question 4 – pertaining to whether a participant regarded themselves as a good or bad customer of fast food restaurants – could be taken as a proxy for frequency of visit, and thus an indicator of behaviour rather than image.

In a departure from Berthelot, Brecheteau & Toupin (2010)[4], and following the *tea* example in Husson, Lê & Pagès (2011), the questions relating to behaviour were treated as the active variables – with those concerning image (perception) and demographics as supplementary. Question 4 was left as an indicator of perception rather than behaviour, since (a) not enough information was available from the original questionnaire to make a judgement on this question's intent, and (b) this

---

[1] http://enquirer.free.fr/case-studies/Fast-Food%20Paradox
[2] No information is available on how this survey was carried out, or on the actual text of the survey questions.
[3] http://enquirer.free.fr/case-studies/Fast-Food%20Paradox/Fast-Food%20Paradox.xls
[4] The authors of the report as quoted on the EnQuireR website; no survey data is provided, however the Excel file was originally created in June 2010.

question had four response categories compared with Yes|No for all other active variables – so may have contributed (dis)proportionally more to the variance. As such, the study comprised a total of 24 active variables and 24 qualitative supplementary. No variables were treated as quantitative supplementary, nor were there any supplementary individuals.

The raw data for the first 20 questions comprised numeric, likert-scale, responses. These responses were converted into factors, with the textual descriptions for each category replacing the numeric values[5]. Responses for one of the (ordinal) active questions – relating to level of expenditure during a visit – were collapsed from four to two categories, since two of the original categories had very low frequencies of response. Response frequencies were also low for some of the Yes|No variables; by their nature, these could not be directly re-assigned – and so a ventilation level of 0.05 was used in the MCA. A summary of the active variables is provided in the Annex.

Results/Discussion

MCA discovered 24 dimensions, with percentage of variance explained by each shown in Figure 1 below. Based on this distribution, it was decided to study only the first three dimensions – which accounted for a combined 25.1% of total inertia between them.

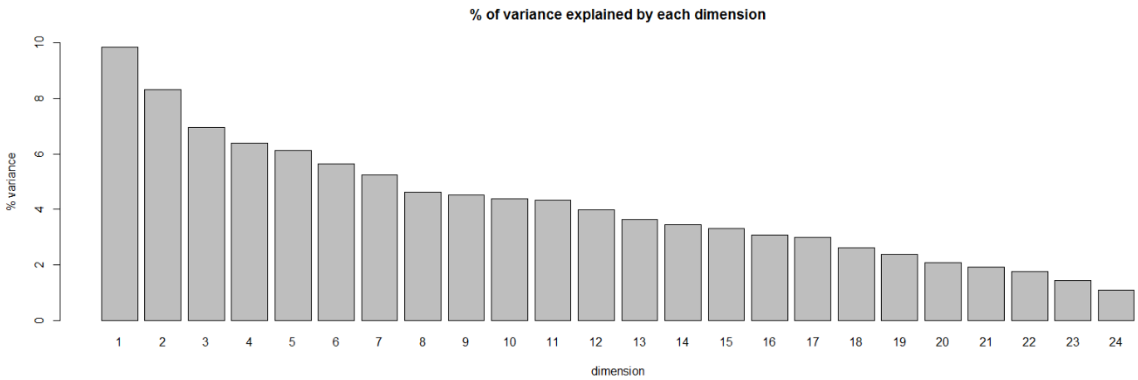

Figure 1: Percentage of variance explained by each of the 24 MCA dimensions.

[5] The textual responses were picked up from the .R file included on the EnQuireR website - http://enquirer.free.fr/case-studies/Fast-Food%20Paradox/Fast-Food%20Paradox.R.

Clouds of categories

Figure 2 shows the cloud of categories projected onto dimensions 1 and 2. From this can be seen some indications that dimension 1 potentially splits along a number of lines: higher vs lower spend; health-conscious vs not; families vs singles/couples; week vs weekend dining; fast food franchise (McDonald’s vs Quick). Dimension 2 conversely seems to oppose week vs weekend dining, take-away vs eating in, times of eating (night and lunchtimes vs other), and combinations of food consumed (burgers vs nuggets and ice cream).

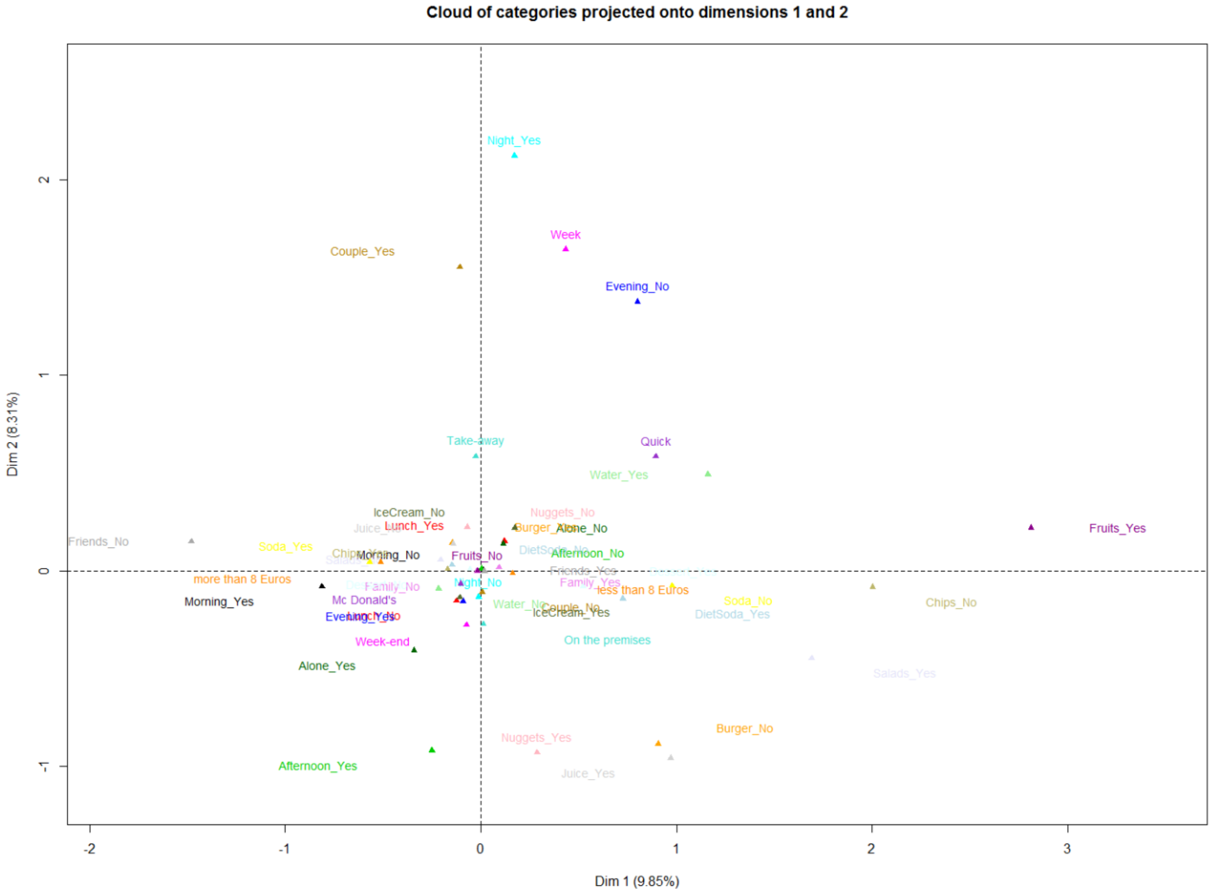

Figure 2: Cloud of categories projected onto dimensions 1 and 2

Figure 3 shows the cloud of categories projected onto dimensions 2 and 3. The distinctions along dimension 3 are less clear, however there are again indications of opposition between social aspects (family/not-family), times of eating (lunchtime vs other), and combinations of food (nuggets vs burger).

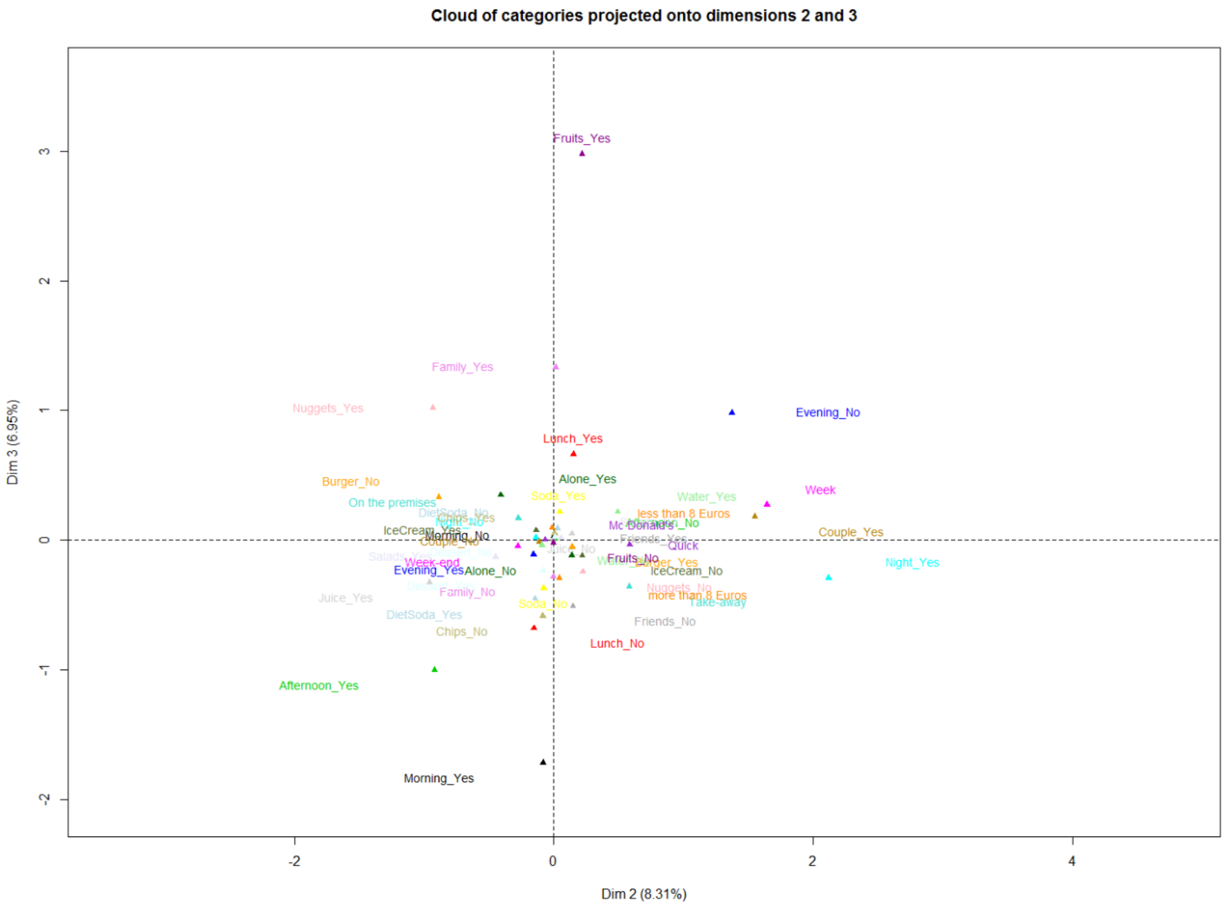

Figure 3: Cloud of categories projected onto dimensions 2 and 3

Figure 4 shows the cloud of categories in dimensions 1 and 2, with gender as a supplementary variable. There is a clear indication that dimensions 1 splits along gender lines, and similarly dimension 2 – albeit to a lesser extent.

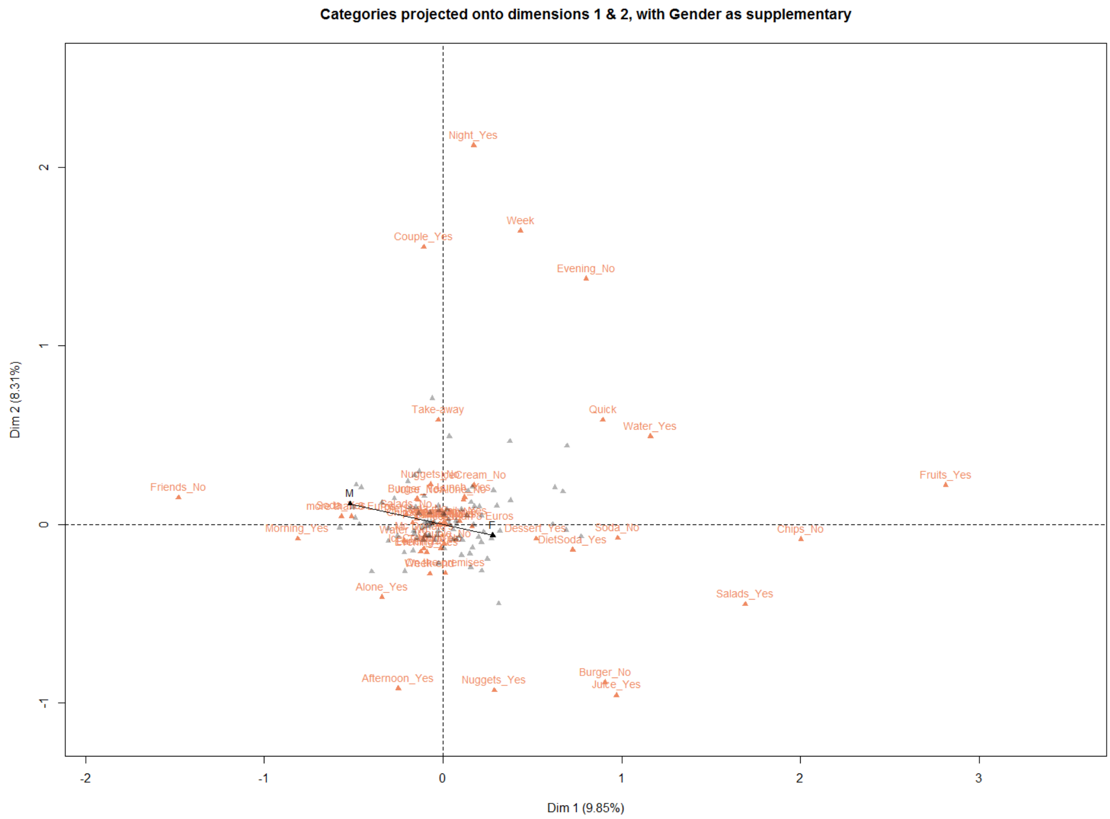

Figure 4: Cloud of categories in dimensions 1 and 2, with Gender as supplementary.

Figure 5 shows the cloud of categories in dimensions 1 and 2, with "image" (i.e. perception of fast food's reputation) as a supplementary variable. This suggests that dimension1 opposes students who have a very bad image of fast food from those who have a good image of fast food. Figure 6 suggests that dimension 3 splits along similar lines.

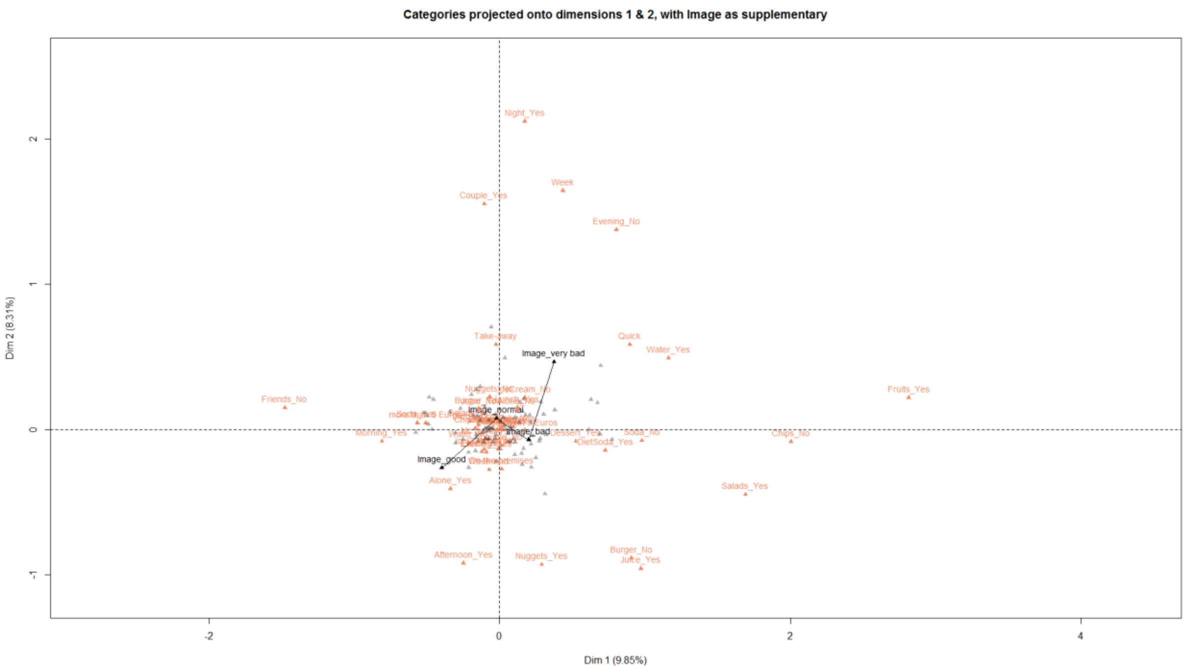

Figure 5: Cloud of categories in dimensions 1 and 2, with perception of fast food restaurants as supplementary.

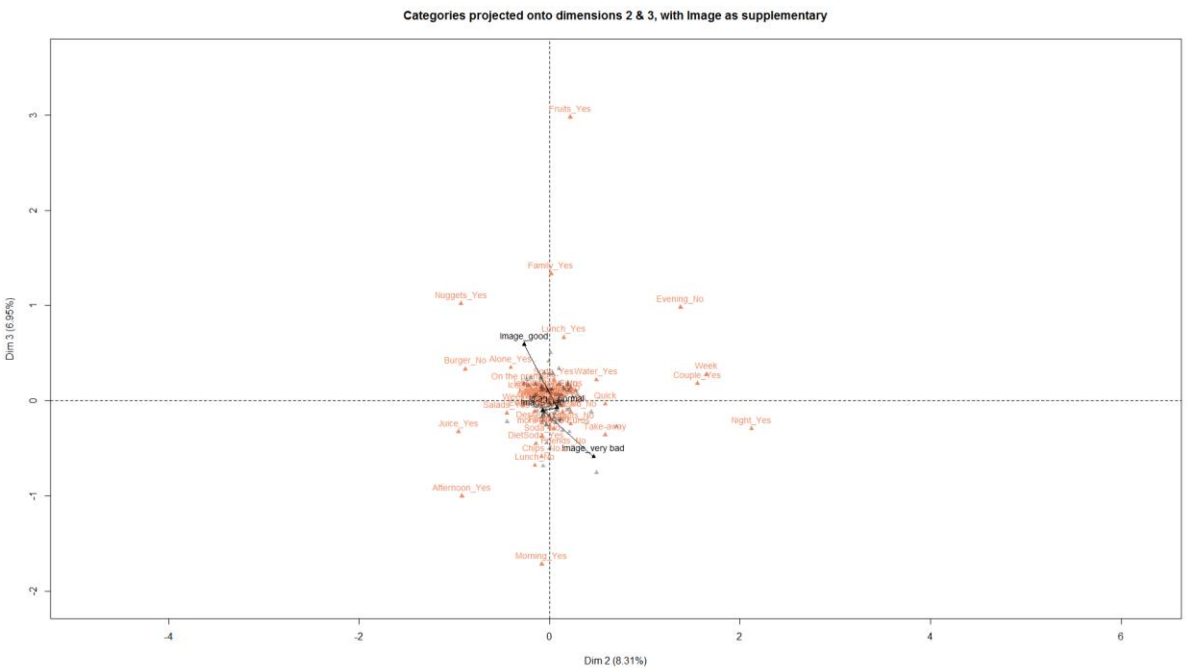

Figure 6: Cloud of categories in dimensions 2 and 3, with perception of fast food restaurants as supplementary.

Figure 7 shows the cloud of categories in dimensions 1 and 2, with “appreciation” (of fast food) as a supplementary variable. This suggests that dimension1 opposes students who enjoy fast food from those who do not.

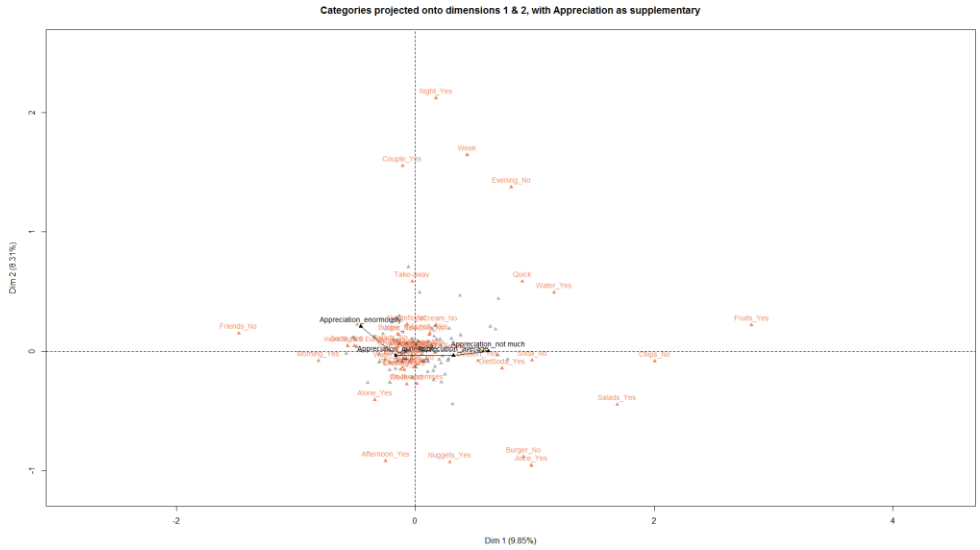

Figure 7: Cloud of categories in dimensions 1 and 2, with appreciation of fast food as supplementary.

Figure 8 shows the cloud of categories in dimensions 1 and 2, with consumer type as a supplementary variable. This suggests that dimension1 opposes the students who are regular customers of fast food restaurants, from those who are not. Taken together, this (unsurprisingly) suggests a narrative where students who do not enjoy fast food have a negative perception of it (or vice versa), and consequently do not often eat it. With the converse also being true.

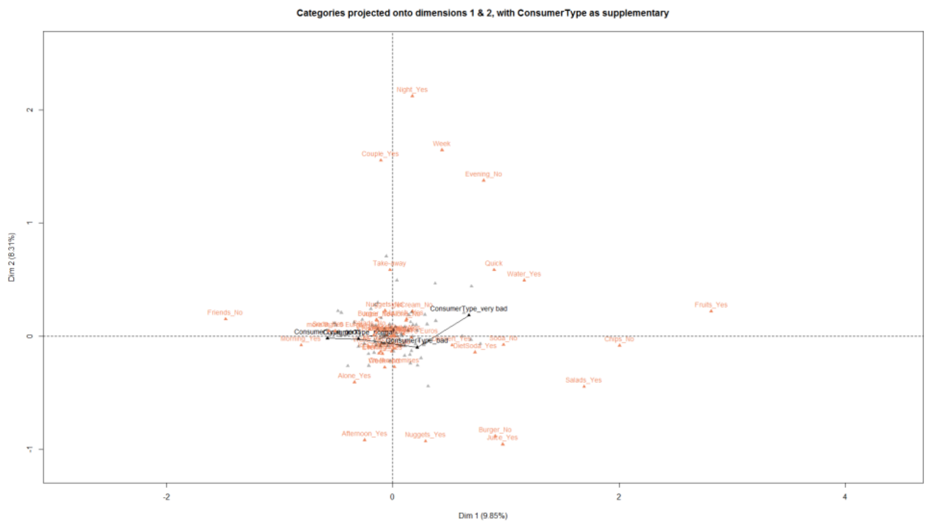

Figure 8: Cloud of categories in dimensions 1 and 2, with consumer type as supplementary.

Cloud of individuals

Figures 9 and 10 show the cloud of individuals projected onto dimensions 1 & 2 and 2 & 3, respectively. The clouds appear to be reasonably homogenous, in that there are no obvious discontinuities or isolated groups. Landmark individuals are highlighted in both cases, chosen to represent opposing sides on each dimension.

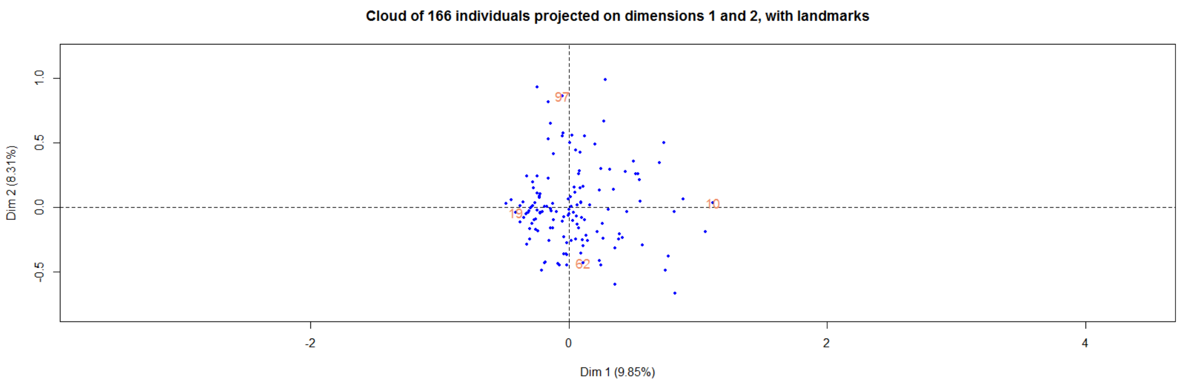

Figure 9: Cloud of individuals projected onto dimensions 1 and 2.

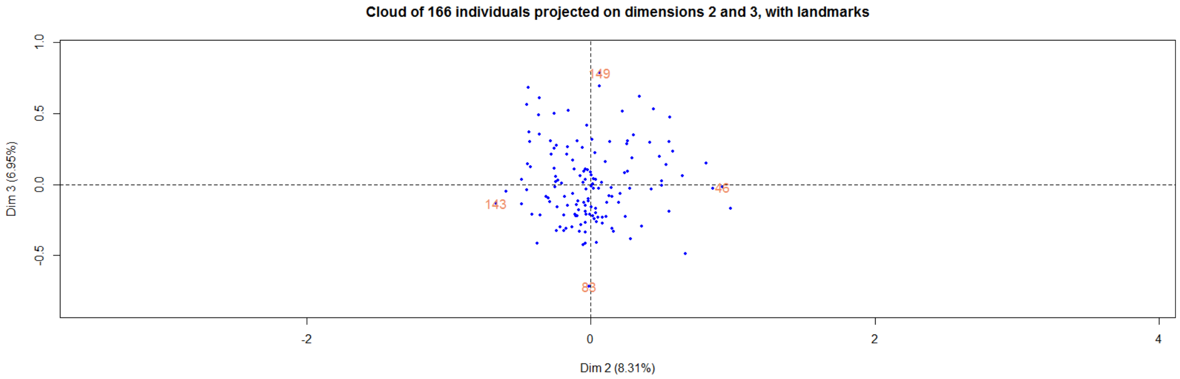

Figure 10: Cloud of individuals projected onto dimensions 2 and 3.

Table 1 highlights the differences in category responses between landmark individuals 10 and 19, opposed on dimension 1. This is consistent with the idea that individuals who score positively on this dimension do not much consume/appreciate fast food; individual 10 (hypothetically) pops in on their lunch break, and grabs a salad and diet coke. Whereas individual 19 goes to the restaurant in the evening for a burger and chips.

	Lunch	Evening	Alone	EatIn	Burger	Salads	Soda	Water	DietSoda	Chips	IceCream	EUR
10	Yes	No	No	EatIn	No	Yes	No	Yes	Yes	No	No	< 8
19	No	Yes	Yes	Takeout	Yes	No	Yes	No	No	Yes	Yes	> 8

Table 1: Landmark individuals 10 and 19.

Table 2 highlights the differences in category responses between landmark individuals 62 and 97, opposed on dimension 2. This is consistent with themes of week vs weekend and eat-in vs take-away.

	Evening	Night	Week/end	Alone	EatIn	Salads	Burger	Nuggets	DietSoda
62	Yes	No	Weekend	Yes	Eat-In	Yes	Yes	Yes	Yes
97	No	Yes	Week	No	Take-away	No	Yes	No	No

Table 2: Landmark individuals 62 and 97.

Figure 11 shows subclouds of the "consumer type" variable, projected onto dimensions 1 and 2. Good consumers are those individuals who responded with the categories "good" or "very good"; bad consumers responded "bad" or "very bad". The remainder responded "average". Generally speaking, this supports the notion that dimension 1 splits along the lines of appreciation and enjoyment of fast food – and hence how regularly an individual consumes it. This being the case, there is also the hint that some people who rate themselves as bad or very bad consumers nevertheless group with the good consumers – although it is not clear that this can be interpreted as them eating fast food despite having a negative image of it, in line with the proposed "paradox"[6].

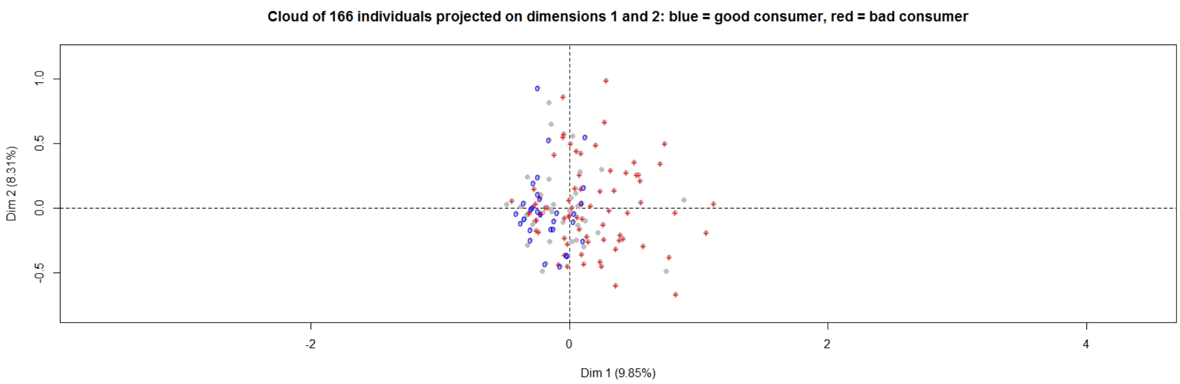

Figure 11: Subclouds of good and bad consumer types, projected onto dimensions 1 and 2.

Figure 12 shows subclouds of individuals who eat (or not) fast food at lunch, projected onto dimensions 2 and 3 – supporting the idea that this variable is important to dimension 3.

[6] It could equally be that these individuals have a neutral or positive image of fast food, but have interpreted this question as "how often do you eat it", as discussed in the introduction.

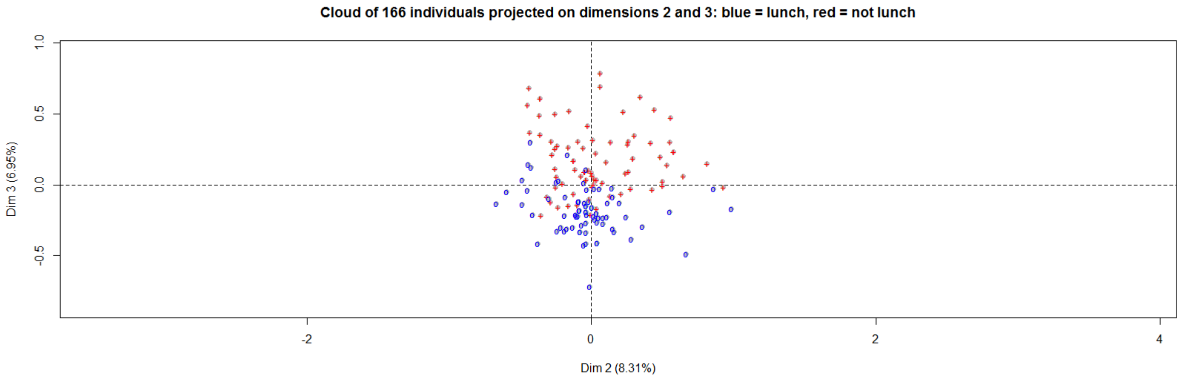

Figure 12: Subclouds of lunchtime and non-lunchtime individuals, projected onto dimensions 2 and 3

Category Contributions

Figures 13 and 14 show the top 15 categories contributing to dimensions 1 & 2, and 2 & 3, respectively.

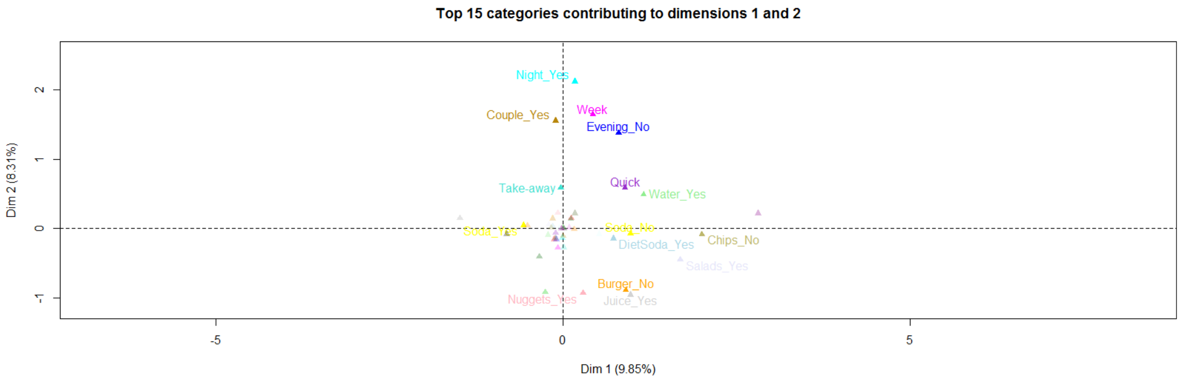

Figure 13: top 15 categories contributing to dimensions 1 and 2

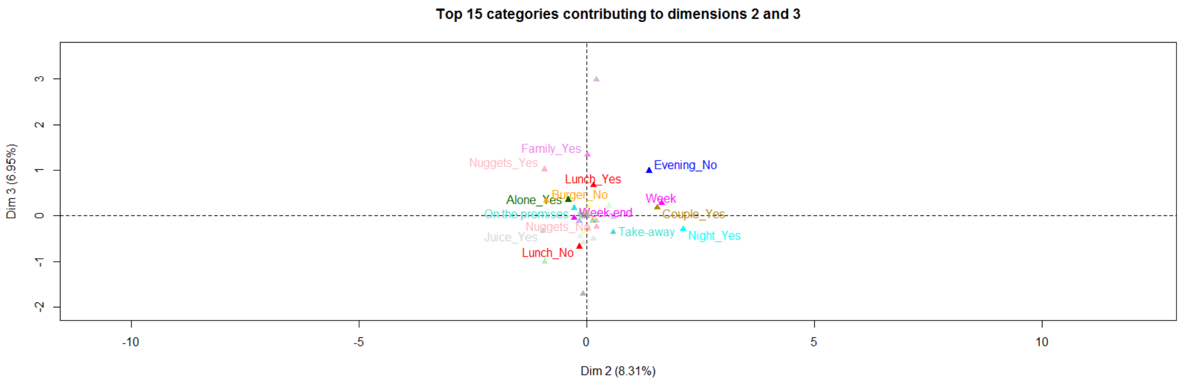

Figure 14: top 15 categories contributing to dimensions 2 and 3

Figures 15, 16 and 17 show the categories which make the highest contributions to each of dimensions 1, 2 and 3, respectively. Following Le Roux and Rouanet (2010), those categories which made more than the average contribution (100/48 = 2.08%) were selected. Suggested labels have been added to the extremities of each dimension, based on these and earlier findings.

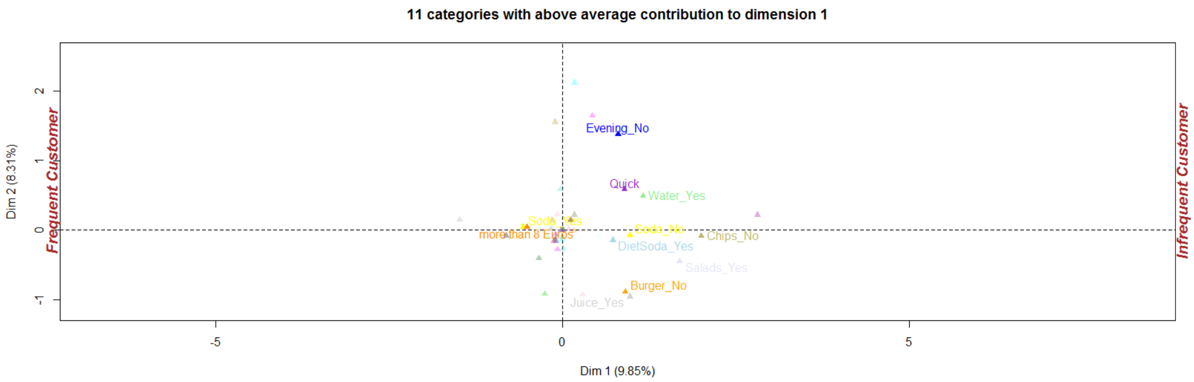

Figure 15: top categories contributing to dimension 1

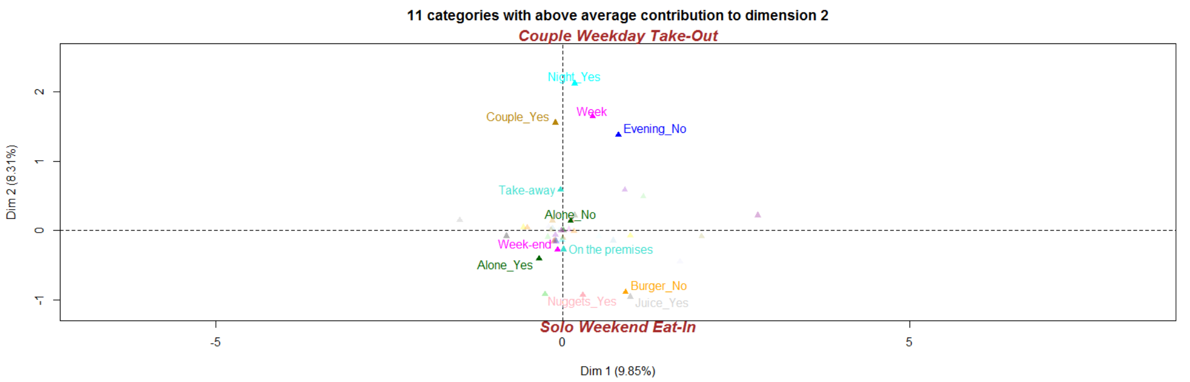

Figure 16: top categories contributing to dimension 2

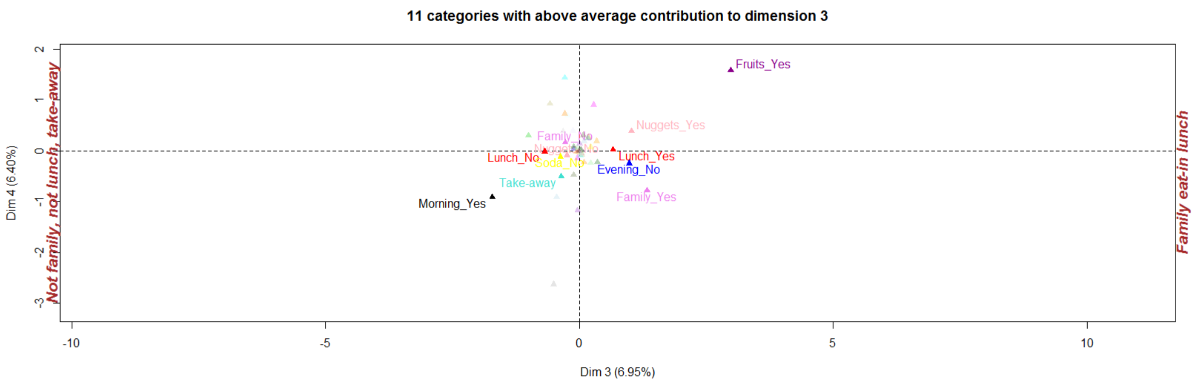

Figure 17: top categories contributing to dimension 3

Confidence Ellipses

Figures 18, 19, 20 and 21 show confidence ellipses for the variables ConsumerType, Lunch, WeekOrWeekend and Family. There is considerable overlap between the categories for ConsumerType, making it hard to support a significant distinction between these groups. Categories of the other three variables show less overlap, although none are entirely distinct.

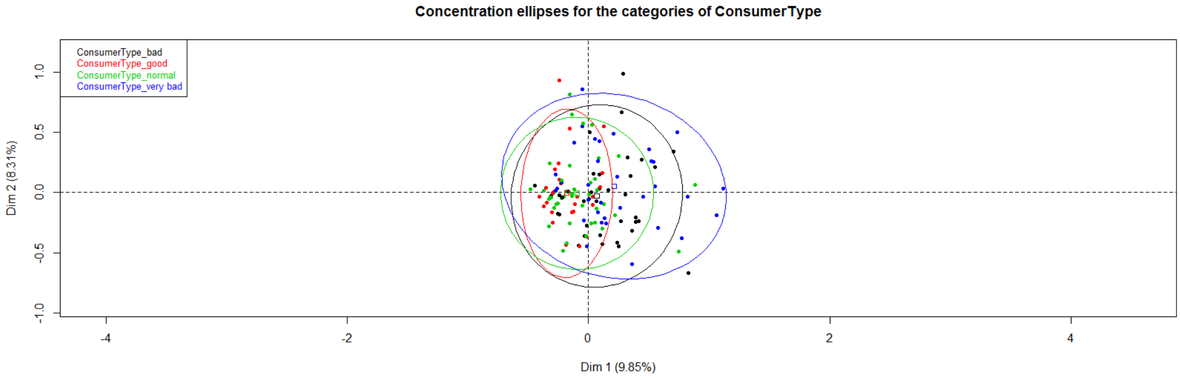

Figure 18: Confidence ellipses for the categories of variable ConsumerType

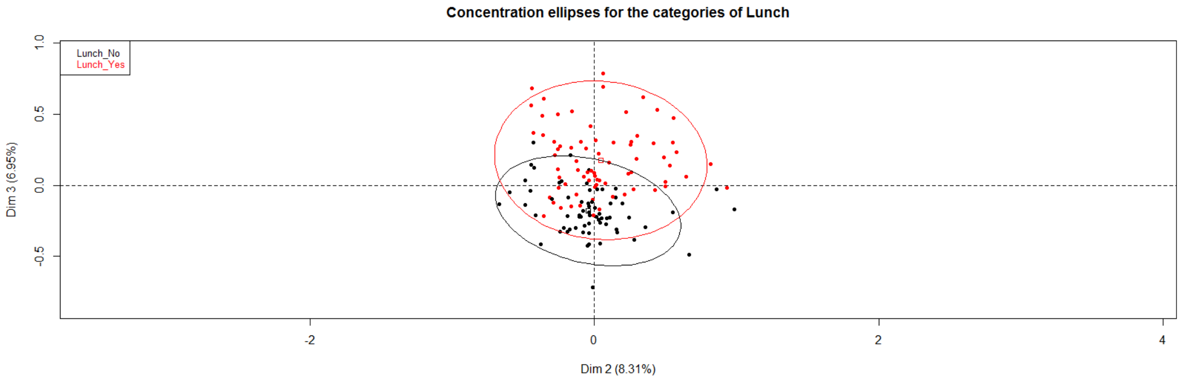

Figure 19: Confidence ellipses for the categories of variable Lunch

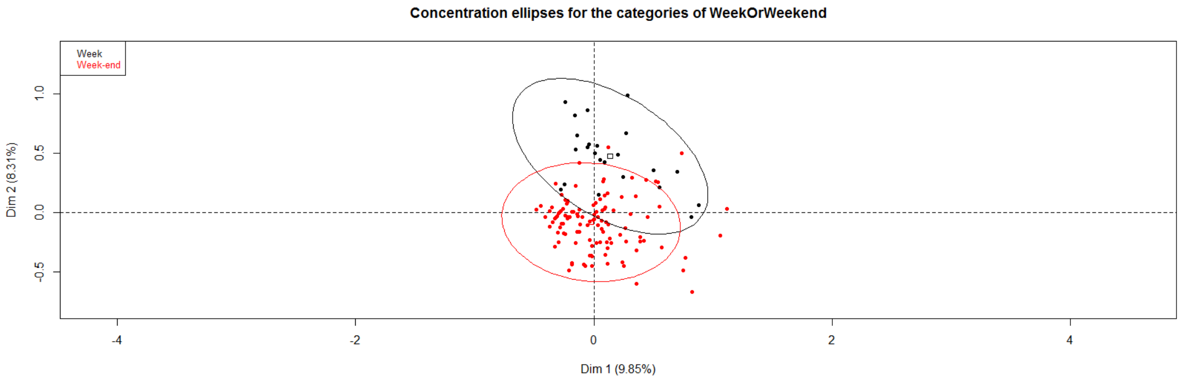

Figure 20: Confidence ellipses for the categories of variable WeekOrWeekend

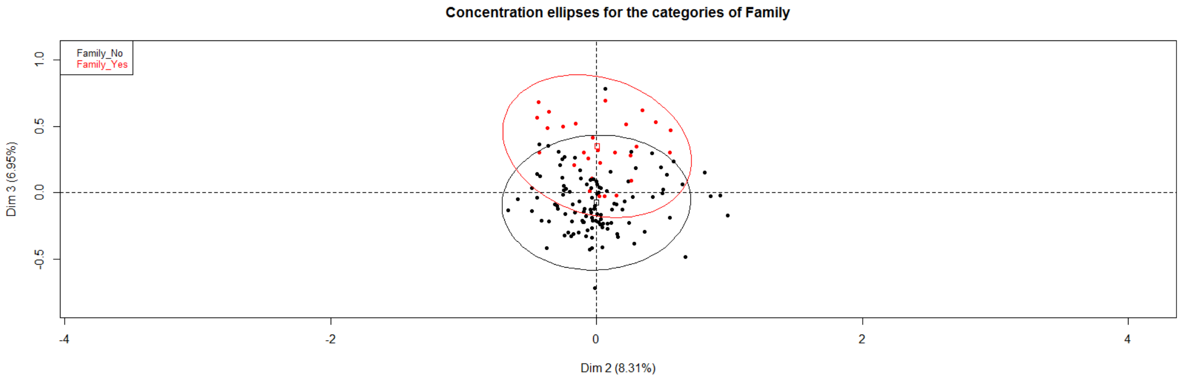

Figure 21: Confidence ellipses for the categories of variable Family

Clustering

Figures 21 and 22 show the results of Hierarchical Clustering on Principal Components (HCPC) applied to results of the MCA. Cutting the dendrogram at 5 factors results in a reasonable separation of groups along dimensions 1 and 2. If we were to try and label these factors then we might, tentatively, start with something like the following[7]:

- Cluster 1: People who like fast food, and get take-outs during the week.
- Cluster 2: People who like fast food, and eat-in at the weekend.
- Cluster 3: People who neither like nor dislike fast food, and get take-outs during the week; possibly out of convenience or if they are with other people who want to go.
- Cluster 4: People who neither like nor dislike fast food, with no particular pattern of eating; possibly they go because they are with other people who want to go.
- Cluster 5: People who do not like fast food restaurants, and only eat healthy options if they do go.

[7] Pending a deeper dive into the data behind each proposed group of individuals

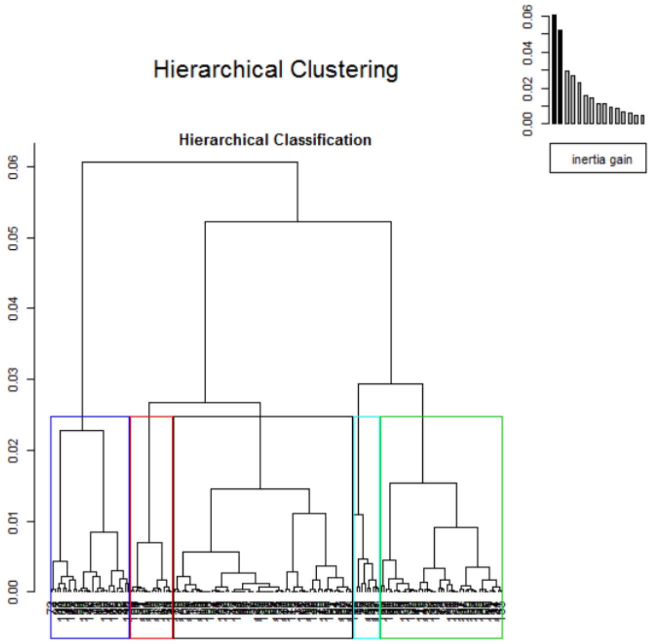

Figure 22: HCPC dendrogram

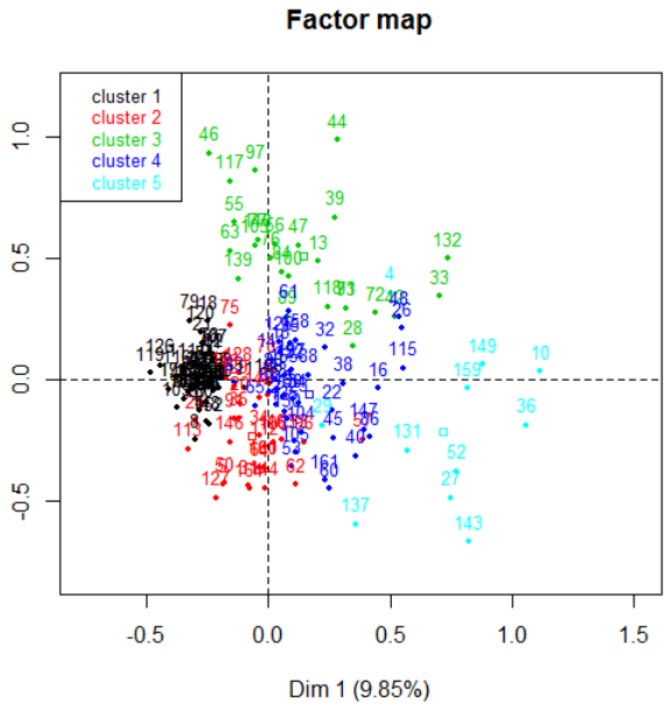

Figure 23: HCPC Factor Map

Conclusion

MCA did not reveal any obvious evidence of the proposed paradox. It seems, from the data, as though people who go to fast food restaurants – and eat typical fast food (burgers, nuggets, chips, ice cream, etc.) – mostly enjoy it, and do not have a particularly negative image of it. There is some indication of different usage patterns, in terms of weekday vs weekend, eat-in vs take-away, families vs others – as well as the types of fast food (burgers vs nuggets, ice cream yes vs no) interacting with these categories. To the extent that there is a paradox, in that people who actively dislike fast food still sometimes visit fast food restaurants, then the success of these restaurants seems mostly down to their healthy options such as salads and juices[8]. There is no indication that these people eat the more typical fast food, in spite of their dislike for it.

References

Berthelot, S., Brecheteau, J., Toupin, L. (2010). *Multivariate exploration of the questionnaire and typology of the surveyed people*. Retrieved from http://enquirer.free.fr/case-studies/Fast-Food%20Paradox/Multivariate_report.pdf. Accessed March 1st 2017.

Husson, F., Lê, S., & Pagès, J. (2011). *Exploratory multivariate analysis by example using R. Boca Raton*: CRC Press.

Le, R. B., & Rouanet, H. (2010). *Multiple correspondence analysis*. Los Angeles [Calif.: SAGE.

[8] The chain "Quick" seems to have a better perception in this regard than McDonalds

Annex

Active Variables / Categories Summary

```
Morning    Lunch     Afternoon Evening    Night      WeekOrWeekend  Friends    Couple     Alone      Family     EatInOrTakeaway        Burger
No :163    No :82    No :164   No : 17    No :156    Week     : 24  No :  2    No :155    No :124    No :137    On the premises:113    No : 23
Yes:  3    Yes:84    Yes:  2   Yes:149    Yes: 10    Week-end:142   Yes:164    Yes: 11    Yes: 42    Yes: 29    Take-away      : 53    Yes:143

Nuggets    Salads     Soda      Water      Juice      DietSoda   Chips      Fruits     IceCream  Dessert    Expenditure
No :134    No :148    No : 61   No :140    No :145    No :138    No : 13    No :165    No : 64   No :150    less than 8 Euros:126
Yes: 32    Yes: 18    Yes:105   Yes: 26    Yes: 21    Yes: 28    Yes:153    Yes:  1    Yes:102   Yes: 16    more than 8 Euros: 40

Brand
Mc Donald's:149
Quick      : 17
```

MCA Summary

```
Call:
MCA(X = data, quali.sup = quali.sup, graph = FALSE, level.ventil = 0.05)



Eigenvalues
                       Dim.1   Dim.2   Dim.3   Dim.4   Dim.5   Dim.6   Dim.7   Dim.8   Dim.9  Dim.10  Dim.11  Dim.12  Dim.13  Dim.14
Variance               0.099   0.083   0.070   0.064   0.061   0.056   0.052   0.046   0.045   0.044   0.043   0.040   0.036   0.034
% of var.              9.851   8.314   6.951   6.396   6.121   5.635   5.240   4.616   4.519   4.384   4.328   3.970   3.640   3.447
Cumulative % of var.   9.851  18.165  25.116  31.512  37.633  43.267  48.507  53.124  57.643  62.026  66.355  70.325  73.966  77.412
                      Dim.15  Dim.16  Dim.17  Dim.18  Dim.19  Dim.20  Dim.21  Dim.22  Dim.23  Dim.24
Variance               0.033   0.031   0.030   0.026   0.024   0.021   0.019   0.017   0.014   0.011
% of var.              3.307   3.069   2.983   2.607   2.383   2.066   1.923   1.748   1.426   1.075
Cumulative % of var.  80.720  83.789  86.772  89.379  91.762  93.829  95.752  97.500  98.925 100.000

Individuals (the 10 first)
                                   Dim.1    ctr   cos2    Dim.2    ctr   cos2    Dim.3    ctr   cos2
1                               | -0.014  0.001  0.000 | -0.368  0.983  0.158 |  0.490  2.077  0.280 |
```

```
2                             |  0.032  0.006  0.001 | -0.104  0.078  0.015 | -0.219  0.417  0.068 |
3                             | -0.224  0.306  0.225 | -0.046  0.015  0.010 |  0.094  0.076  0.040 |
4                             |  0.501  1.533  0.109 |  0.359  0.932  0.056 | -0.294  0.750  0.038 |
5                             | -0.187  0.213  0.040 | -0.432  1.353  0.215 |  0.367  1.165  0.155 |
6                             | -0.346  0.731  0.349 | -0.082  0.049  0.020 | -0.180  0.280  0.094 |
7                             | -0.186  0.213  0.132 |  0.005  0.000  0.000 |  0.063  0.034  0.015 |
8                             | -0.298  0.545  0.176 | -0.249  0.448  0.122 |  0.020  0.003  0.001 |
9                             |  0.097  0.057  0.016 |  0.042  0.013  0.003 | -0.412  1.470  0.293 |
10                            |  1.116  7.620  0.577 |  0.034  0.008  0.001 |  0.035  0.011  0.001 |

Categories (the 10 first)
                                Dim.1    ctr   cos2 v.test    Dim.2    ctr   cos2 v.test    Dim.3    ctr   cos2 v.test
Morning_No                    |  0.015  0.009  0.012  1.416 |  0.002  0.000  0.000  0.142 |  0.032  0.059  0.054  2.990 |
Morning_Yes                   | -0.812  0.504  0.012 -1.416 | -0.082  0.006  0.000 -0.142 | -1.716  3.189  0.054 -2.990 |
Lunch_No                      | -0.124  0.323  0.015 -1.577 | -0.155  0.593  0.023 -1.965 | -0.680 13.700  0.452 -8.633 |
Lunch_Yes                     |  0.121  0.315  0.015  1.577 |  0.151  0.579  0.023  1.965 |  0.664 13.374  0.452  8.633 |
Afternoon_No                  |  0.003  0.000  0.001  0.356 |  0.011  0.006  0.010  1.303 |  0.012  0.009  0.012  1.421 |
Afternoon_Yes                 | -0.251  0.032  0.001 -0.356 | -0.919  0.510  0.010 -1.303 | -1.001  0.724  0.012 -1.421 |
Evening_No                    |  0.802  2.785  0.073  3.479 |  1.374  9.691  0.215  5.962 |  0.982  5.920  0.110  4.261 |
Evening_Yes                   | -0.091  0.318  0.073 -3.479 | -0.157  1.106  0.215 -5.962 | -0.112  0.675  0.110 -4.261 |
Night_No                      | -0.011  0.005  0.002 -0.558 | -0.136  0.870  0.288 -6.896 |  0.019  0.020  0.006  0.956 |
Night_Yes                     |  0.171  0.075  0.002  0.558 |  2.120 13.573  0.288  6.896 | -0.294  0.312  0.006 -0.956 |

Categorical variables (eta2)
                              Dim.1 Dim.2 Dim.3
Morning                       | 0.012 0.000 0.054 |
Lunch                         | 0.015 0.023 0.452 |
Afternoon                     | 0.001 0.010 0.012 |
Evening                       | 0.073 0.215 0.110 |
Night                         | 0.002 0.288 0.006 |
WeekOrWeekend                 | 0.032 0.457 0.013 |
Friends                       | 0.027 0.000 0.003 |
Couple                        | 0.001 0.171 0.002 |
```

```
Alone                            | 0.039 0.057 0.041 |
Family                           | 0.002 0.000 0.375 |

Supplementary categories (the 10 first)
                                  Dim.1   cos2 v.test    Dim.2   cos2 v.test    Dim.3   cos2 v.test
Image_bad                        |  0.178  0.013  1.457 | -0.077  0.002 -0.635 | -0.120  0.006 -0.985 |
Image_good                       | -0.413  0.036 -2.438 | -0.259  0.014 -1.532 |  0.582  0.072  3.440 |
Image_normal                     | -0.009  0.000 -0.108 |  0.060  0.003  0.728 | -0.081  0.006 -0.985 |
Image_very bad                   |  0.375  0.010  1.283 |  0.595  0.025  2.036 | -0.433  0.013 -1.481 |
HowExpensive_a little expensive |  0.086  0.003  0.725 | -0.070  0.002 -0.592 |  0.196  0.017  1.651 |
HowExpensive_average             | -0.075  0.003 -0.695 | -0.054  0.002 -0.506 |  0.078  0.003  0.725 |
HowExpensive_quite expensive     |  0.010  0.000  0.084 |  0.006  0.000  0.048 | -0.236  0.024 -1.991 |
HowExpensive_very expensive      | -0.059  0.000 -0.182 |  0.704  0.028  2.165 | -0.270  0.004 -0.830 |
ValueForMoney_bad                |  0.044  0.001  0.400 |  0.110  0.006  0.991 | -0.067  0.002 -0.604 |
ValueForMoney_good               | -0.101  0.003 -0.684 |  0.010  0.000  0.066 |  0.503  0.070  3.397 |

Supplementary categorical variables (eta2)
                                  Dim.1 Dim.2 Dim.3
Image                            | 0.048 0.039 0.079 |
HowExpensive                     | 0.004 0.029 0.034 |
ValueForMoney                    | 0.043 0.007 0.090 |
ConsumerType                     | 0.200 0.013 0.004 |
HowWellBalanced                  | 0.005 0.016 0.023 |
Appreciation                     | 0.113 0.009 0.044 |
AreSmallPortions                 | 0.049 0.025 0.006 |
IsPoorNutrition                  | 0.021 0.004 0.032 |
HowPleasurable                   | 0.060 0.015 0.037 |
PollutesEnvironment              | 0.013 0.001 0.072 |
```

MCA DIMDESC

```
$`Dim 1`
$`Dim 1`$quali
                     R2      p.value
Soda           0.55481514 1.258980e-30
Salads         0.34765780 6.389834e-17
Chips          0.34138360 1.413165e-16
Water          0.25054431 6.547525e-12
ConsumerType   0.19986834 6.689446e-08
Gender         0.14523722 4.073037e-07
Juice          0.13662967 9.528815e-07
Burger         0.13221280 1.470123e-06
DietSoda       0.10700664 1.694665e-05
Brand          0.09132352 7.588196e-05
Expenditure    0.08294253 1.681659e-04
Appreciation   0.11272065 2.223909e-04
Overall        0.09397418 3.213340e-04
Evening        0.07336691 4.160701e-04
Unsatisfying   0.09173661 1.346835e-03
Fruits         0.04800799 4.565170e-03
Alone          0.03926830 1.048887e-02
DietAfter      0.07621298 1.207639e-02
HowPleasurable 0.05988565 1.827305e-02
WeekOrWeekend  0.03184327 2.143418e-02
Dessert        0.02906474 2.808942e-02
Friends        0.02672880 3.531918e-02
SuitEverybody  0.06110830 3.696681e-02
AreSmallPortions 0.04856562 4.419135e-02
Image          0.04823648 4.532197e-02
```

```
$`Dim 1`$category
                                     Estimate      p.value
Soda_No                            0.24245300 1.258980e-30
Salads_Yes                         0.29759241 6.389834e-17
Chips_No                           0.34128477 1.413165e-16
Water_Yes                          0.21612478 6.547525e-12
F                                  0.12543712 4.073037e-07
Juice_Yes                          0.17449855 9.528815e-07
Burger_No                          0.16516491 1.470123e-06
ConsumerType_very bad              0.21195237 3.302873e-06
DietSoda_Yes                       0.13708810 1.694665e-05
Quick                              0.15641799 7.588196e-05
less than 8 Euros                  0.10567843 1.681659e-04
Evening_No                         0.14019919 4.160701e-04
Unsatisfying_slightly agree        0.21583395 1.672487e-03
Appreciation_not much              0.16676330 2.522224e-03
Fruits_Yes                         0.44435263 4.565170e-03
HowPleasurable_not much pleasure   0.17563413 6.720582e-03
Alone_No                           0.07153169 1.048887e-02
ValueForMoney_very bad             0.19419362 1.155185e-02
normal satisfying                  0.05000660 1.212690e-02
AreSmallPortions_slightly agree    0.16993448 1.242463e-02
SuitEverybody_slightly disagree    0.10286996 1.970431e-02
Week                               0.07962914 2.143418e-02
WouldBeMissed_not at all           0.10536241 2.588495e-02
Appreciation_average               0.07901504 2.677182e-02
Dessert_Yes                        0.09065475 2.808942e-02
often                              0.07547898 2.892724e-02
Friends_Yes                        0.23516128 3.531918e-02
City                               0.06172454 3.999526e-02
HowPleasurable_quite a lot pleasure -0.08727372 4.455457e-02
Friends_No                        -0.23516128 3.531918e-02
IsCheaper_slightly agree          -0.10253170 3.455917e-02
Unsatisfying_disagree             -0.16954363 3.437642e-02
Country                           -0.06716398 3.370422e-02
Dessert_No                        -0.09065475 2.808942e-02
Appreciation_quite a lot          -0.07686288 2.475666e-02
Week-end                          -0.07962914 2.143418e-02
Appreciation_enormously           -0.16891546 2.120467e-02
Image_good                        -0.13980072 1.430062e-02
```

```
$`Dim 2`
$`Dim 2`$quali
                       R2      p.value
WeekOrWeekend   0.45656067 1.751595e-23
Night           0.28820998 8.925182e-14
Evening         0.21543713 3.003387e-10
Nuggets         0.20692989 7.412812e-10
Couple          0.17101976 3.062522e-08
EatInOrTakeaway 0.15945351 9.857191e-08
Juice           0.13262123 1.412441e-06
Burger          0.12608239 2.676372e-06
Alone           0.05656755 2.031478e-03
Water           0.04488838 6.138470e-03
Brand           0.03897566 1.078668e-02
IceCream        0.03018816 2.517422e-02
Salads          0.02442465 4.435352e-02
Lunch           0.02339174 4.915556e-02
```

```
SuitEverybody_agree                  -0.14085775 1.080713e-02
Alone_Yes                            -0.07153169 1.048887e-02
ConsumerType_normal                  -0.09644211 6.042572e-03
never                                -0.12073846 5.315706e-03
Fruits_No                            -0.44435263 4.565170e-03
ConsumerType_good                    -0.18128541 4.672500e-04
Evening_Yes                          -0.14019919 4.160701e-04
more than 8 Euros                    -0.10567843 1.681659e-04
Mc Donald's                          -0.15641799 7.588196e-05
satisfying                           -0.13912632 6.989610e-05
DietSoda_No                          -0.13708810 1.694665e-05
Burger_Yes                           -0.16516491 1.470123e-06
Juice_No                             -0.17449855 9.528815e-07
M                                    -0.12543712 4.073037e-07
Water_No                             -0.21612478 6.547525e-12
Chips_Yes                            -0.34128477 1.413165e-16
Salads_No                            -0.29759241 6.389834e-17
Soda_Yes                             -0.24245300 1.258980e-30

$`Dim 2`$category
                                Estimate      p.value
Week                          0.27700645 1.751595e-23
Night_Yes                     0.32529893 8.925182e-14
Evening_No                    0.22071564 3.003387e-10
Nuggets_No                    0.16625498 7.412812e-10
Couple_Yes                    0.23969081 3.062522e-08
Take-away                     0.12348958 9.857191e-08
Juice_No                      0.15794407 1.412441e-06
Burger_Yes                    0.14817868 2.676372e-06
Alone_No                      0.07887486 2.031478e-03
Water_Yes                     0.08404400 6.138470e-03
Quick                         0.09387924 1.078668e-02
PleasentSide_average          0.09210473 2.108131e-02
IceCream_No                   0.05146579 2.517422e-02
HowExpensive_very expensive   0.16082312 2.995166e-02
Image_very bad                0.14857936 4.143280e-02
Salads_No                     0.07246653 4.435352e-02
Lunch_Yes                     0.04410371 4.915556e-02
Lunch_No                     -0.04410371 4.915556e-02
Salads_Yes                   -0.07246653 4.435352e-02
IceCream_Yes                 -0.05146579 2.517422e-02
Mc Donald's                  -0.09387924 1.078668e-02
Water_No                     -0.08404400 6.138470e-03
Alone_Yes                    -0.07887486 2.031478e-03
```

```
Burger_No                    -0.14817868 2.676372e-06
Juice_Yes                    -0.15794407 1.412441e-06
On the premises              -0.12348958 9.857191e-08
Couple_No                    -0.23969081 3.062522e-08
Nuggets_Yes                  -0.16625498 7.412812e-10
Evening_Yes                  -0.22071564 3.003387e-10
Night_No                     -0.32529893 8.925182e-14
Week-end                     -0.27700645 1.751595e-23
```

```
$`Dim 3`
$`Dim 3`$quali
                             R2      p.value
Lunch                0.45165201 3.680834e-23
Family               0.37539375 1.746147e-18
Nuggets              0.24831017 8.393411e-12
Evening              0.11001917 1.268413e-05
Soda                 0.08057315 2.104793e-04
EatInOrTakeaway      0.06010220 1.455108e-03
ValueForMoney        0.08983161 1.581194e-03
Morning              0.05418575 2.544096e-03
Fruits               0.05367584 2.669710e-03
Image                0.07891002 3.925388e-03
PollutesEnvironment 0.07208351 6.864657e-03
DietSoda             0.04167317 8.336533e-03
Alone                0.04071028 9.138718e-03
Overall              0.05173301 1.317792e-02
Chips                0.02927321 2.752331e-02
Expenditure          0.02742793 3.297324e-02
FeelBadAfter         0.05121882 3.601678e-02
```

```
$`Dim 3`$category
                                                Estimate      p.value
Lunch_Yes                                     0.17719543 3.680834e-23
Family_Yes                                    0.21270679 1.746147e-18
Nuggets_Yes                                   0.16652022 8.393411e-12
Evening_No                                    0.14421609 1.268413e-05
Soda_Yes                                      0.07761269 2.104793e-04
Image_good                                    0.15692465 4.852053e-04
ValueForMoney_good                            0.15660335 5.757055e-04
On the premises                               0.06932105 1.455108e-03
Morning_No                                    0.23034841 2.544096e-03
Fruits_Yes                                    0.39467952 2.669710e-03
Satisfying                                    0.08829386 4.295017e-03
DietSoda_No                                   0.07186325 8.336533e-03
Alone_Yes                                     0.06118059 9.138718e-03
PollutesEnvironment_neither agree nor disagree  0.07893684 1.282639e-02
Chips_Yes                                     0.08394888 2.752331e-02
less than 8 Euros                             0.05104796 3.297324e-02
IsPoorNutrition_neither agree nor disagree    0.06104444 4.698711e-02
HowExpensive_quite expensive                 -0.04695349 4.609890e-02
PollutesEnvironment_agree                    -0.08393632 3.859423e-02
more than 8 Euros                            -0.05104796 3.297324e-02
Chips_No                                     -0.08394888 2.752331e-02
ValueForMoney_very bad                       -0.15654952 2.538531e-02
Appreciation_not much                        -0.10839633 1.417474e-02
FeelBadAfter_not much                        -0.16211670 9.486141e-03
Alone_No                                     -0.06118059 9.138718e-03
DietSoda_Yes                                 -0.07186325 8.336533e-03
Fruits_No                                    -0.39467952 2.669710e-03
Morning_Yes                                  -0.23034841 2.544096e-03
Take-away                                    -0.06932105 1.455108e-03
Soda_No                                      -0.07761269 2.104793e-04
Evening_Yes                                  -0.14421609 1.268413e-05
Nuggets_No                                   -0.16652022 8.393411e-12
Family_No                                    -0.21270679 1.746147e-18
Lunch_No                                     -0.17719543 3.680834e-23
```

Code

```
# Developed in RStudio 1.0.136 on Windows 10.

# The graphics used in this presentation were exported to the clipboard by RStudio, and imported into Word directly.

> version
               _
platform       x86_64-w64-mingw32
arch           x86_64
os             mingw32
system         x86_64, mingw32
status
major          3
minor          3.2
year           2016
month          10
day            31
svn rev        71607
language       R
version.string R version 3.3.2 (2016-10-31)
nickname       Sincere Pumpkin Patch
```

```
# Initial set-up
library(FactoMineR)
setwd("C:/Users/john/dev/goldsmiths/gda")

# Read in the data table
data <- read.csv(
  file = "FastFoodParadox.csv",
  header = TRUE
)

# Columns 1-20 are raw numeric responses from (presumed) likert scale questions. To
# perform MCA, these need to be converted into factors. And following the lead from
# http://enquirer.free.fr/case-studies/Fast-Food%20Paradox/Fast-Food%20Paradox.R, we
# are also converting these to the equivalent textural responses, to aid interpretation.

# Creating variables for each type of likert scale, for re-use across questions.
likert1 <- c('very bad','bad','normal','good','very good')
likert2 <- c('not expensive','a little expensive','average','quite expensive','very expensive')
likert3 <- c('not balanced','badly balanced','average','quite well balanced','well balanced')
likert4 <- c('not at all','not much','average','quite a lot','enormously')
```

```
likert5 <- c('disagree','slightly disagree','neither agree nor disagree','slightly agree','agree')
likert6 <- c('no pleasure','not much pleasure','average','quite a lot pleasure','great pleasure')
likert7 <- c('not convivial','not much convivial','average','quite convivial','very convivial')
likert8 <- c('not practical','average','quite practical','very practical')
likert9 <- c('nothing pleasant','few pleasant things','average','some pleasant things','a lot of pleasant things')
likert10 <- c('not at all','a little','average','not much')
likert11 <- c('never','rarely','sometimes','often','always')

# Now convert each of columns 1-20 to textual factors.
data$Image <- factor(data$Image, labels = likert1)
data$Expensive <- factor(data$Expensive, labels = likert2)
data$Good.value.for.money <- factor(data$Good.value.for.money, labels = likert1)
data$Kind.of.consumer <- factor(data$Kind.of.consumer, labels = likert1)
data$Not.balanced <- factor(data$Not.balanced, labels =  likert3)
data$Products.assessment <- factor(data$Products.assessment, labels = likert4)
data$Don.t.eat.enough <- factor(data$Don.t.eat.enough, labels = likert5)
data$Bad.nutritionnal.quality <- factor(data$Bad.nutritionnal.quality, labels = likert5)
data$Pleasure <- factor(data$Pleasure, labels = likert6)
data$Agree.with.pollution <- factor(data$Agree.with.pollution, labels = likert5)
data$Convivial <- factor(data$Convivial, labels = likert7)
data$Practical <- factor(data$Practical, labels = likert8 )
data$Play.side <- factor(data$Play.side, labels = likert9)
data$Not.varied.enough <- factor(data$Not.varied.enough, labels = likert5)
data$Satisfy.everybody <- factor(data$Satisfy.everybody, labels = likert5)
data$A.lack.of.it <- factor(data$A.lack.of.it, labels = likert4)
data$Feel.bad <- factor(data$Feel.bad, labels = likert10)
data$Food.adjust <- factor(data$Food.adjust, labels = likert11)
data$Unstatisfying.products <- factor(data$Unstatisfying.products, labels = likert5)
data$Cheaper.meal <- factor(data$Cheaper.meal, labels = likert5)

# Also renaming the columns, as the translations from French are not always helpful.
colnames(data)[1]="Image"
colnames(data)[2]="HowExpensive"
colnames(data)[3]="ValueForMoney"
colnames(data)[4]="ConsumerType"
colnames(data)[5]="HowWellBalanced"
colnames(data)[6]="Appreciation"
colnames(data)[7]="AreSmallPortions"
colnames(data)[8]="IsPoorNutrition"
colnames(data)[9]="HowPleasurable"
colnames(data)[10]="PollutesEnvironment"
colnames(data)[11]="HowConvivial"
colnames(data)[12]="HowPractical"
colnames(data)[13]="PleasentSide"
colnames(data)[14]="NotVariedEnough"
colnames(data)[15]="SuitEverybody"
colnames(data)[16]="WouldBeMissed"
colnames(data)[17]="FeelBadAfter"
colnames(data)[18]="DietAfter"
```

```
colnames(data)[19]="Unsatisfying"
colnames(data)[20]="IsCheaper"
colnames(data)[21]="Overall"
colnames(data)[22]="Morning"              #Active
colnames(data)[23]="Lunch"                #Active
colnames(data)[24]="Afternoon"            #Active
colnames(data)[25]="Evening"              #Active
colnames(data)[26]="Night"                #Active
colnames(data)[27]="WeekOrWeekend"        #Active
colnames(data)[28]="Friends"              #Active
colnames(data)[29]="Couple"               #Active
colnames(data)[30]="Alone"                #Active
colnames(data)[31]="Family"               #Active
colnames(data)[32]="EatInOrTakeaway"      #Active
colnames(data)[33]="Burger"               #Active
colnames(data)[34]="Nuggets"              #Active
colnames(data)[35]="Salads"               #Active
colnames(data)[36]="Soda"                 #Active
colnames(data)[37]="Water"                #Active
colnames(data)[38]="Juice"                #Active
colnames(data)[39]="DietSoda"             #Active
colnames(data)[40]="Chips"                #Active
colnames(data)[41]="Fruits"               #Active
colnames(data)[42]="IceCream"             #Active
colnames(data)[43]="Dessert"              #Active
colnames(data)[44]="Expenditure"          #Active
colnames(data)[45]="Brand"                #Active
colnames(data)[46]="Gender"
colnames(data)[47]="RegularSports"
colnames(data)[48]="Location"

active <- c(22:45)
quali.sup <- c(1:21,46:48)

# Check summary stats (looking for e.g. low response frequencies)
print(summary(data[, active]))

# Collapse Expenditure categories into two
levels(data$Expenditure) <- c("less than 8 Euros", "more than 8 Euros", "less than 8 Euros", "more than 8 Euros")

# Re-check modified summary stats
print(summary(data[, active]))

# Perform the MCA
data.mca <- MCA(
  X = data,
  level.ventil = 0.05, # since we can't automatically combine the other Yes|No categories
  quali.sup = quali.sup,
  graph = FALSE
```

```
)

# Check summary and dimdesc
print(summary(data.mca))
print(dimdesc(data.mca))

# Extract info from the MCA response model into more user-friendly variables
num.ind <- dim(data)[1]
num.categories <- dim(data.mca$var$contrib)[1]
num.dimensions <- dim(data.mca$eig)[1]
max.variance <- ceiling(max(data.mca$eig[[2]]))
ind.coords <- data.mca$ind$coord
sup.coords <- data.mca$quali.sup$coord
cat.names <- rownames(data.mca$var$coord)
sup.cat.names <- rownames(sup.coords)

# Check the number of dimensions discovered / percentage of variance/inertia explained by each
print(data.mca$eig)

# Barplot of the above
barplot(
  data.mca$eig[[2]],
  main = "% of variance explained by each dimension",
  xlab = "dimension",
  names = as.character(1:num.dimensions),
  ylab = "% variance",
  ylim = c(0, 10)
)

# Invg cloud of categories.
plot(
  x = data.mca,
  invisible = c("ind", "quali.sup"),
  title = "Cloud of categories projected onto dimensions 1 and 2",
  habillage = "quali",
  autoLab = "yes",
  axes = c(1,2),
  cex = 0.9
)

plot(
  x = data.mca,
  invisible = c("ind", "quali.sup"),
  title = "Cloud of categories projected onto dimensions 2 and 3",
  habillage = "quali",
  autoLab = "yes",
  axes = c(2,3),
  cex = 0.9
)
```

```
# Helper function for investigations using supplementary categories
plot_with_sup <- function(
  dim1,
  dim2,
  sup.names,
  sup.label) {

  title = sprintf("Categories projected onto dimensions %d & %d, with %s as supplementary", dim1, dim2, sup.label)

  plot(
    x = data.mca,
    invisible = c("ind"),
    col.var = "#ef8a62",
    title = title,
    col.quali.sup = "black",
    cex = 0.9,
    axes = c(dim1, dim2),
    selectMod = c(sup.names, cat.names)
  )
  points(
    sup.coords[sup.names, dim1:dim2],
    type = "l"
  )
}

# Supplementary - Gender
sup.cat.names.gender = c("M","F")
plot_with_sup(1, 2, sup.cat.names.gender, "Gender")
plot_with_sup(2, 3, sup.cat.names.gender, "Gender")

# Supplementary - Image
sup.cat.names.image = c("Image_very bad","Image_bad","Image_normal","Image_good")
plot_with_sup(1, 2, sup.cat.names.image, "Image")
plot_with_sup(2, 3, sup.cat.names.image, "Image")

# Supplementary - Appreciation
sup.cat.names.appreciation = c(
  "Appreciation_not much",
  "Appreciation_average",
  "Appreciation_quite a lot",
  "Appreciation_enormously"
)
plot_with_sup(1, 2, sup.cat.names.appreciation, "Appreciation")

# Supplementary - ConsumerType
sup.cat.names.consumer = c(
  "ConsumerType_very bad",
  "ConsumerType_bad",
```

```
  "ConsumerType_normal",
  "ConsumerType_good"
)
plot_with_sup(1, 2, sup.cat.names.consumer, "ConsumerType")

# Cloud of individuals
plot(
  x = data.mca,
  choix = "ind",
  col.ind = "#ef8a62",
  invisible = c("var", "quali.sup"),
  title = "Cloud of 166 individuals projected onto dimensions 1 and 2",
  axes = c(1,2),
  label = "none"
)

# Helper function
plot_with_landmarks = function(
  dim1,
  dim2,
  landmarks
  ) {

  title = sprintf(
    "Cloud of %d individuals projected on dimensions %d and %d, with landmarks", num.ind, dim1, dim2)

  plot(
    x = data.mca,
    choix = "ind",
    invisible = c("var", "quali.sup"),
    title = title,
    label = "none",
    axes = c(dim1, dim2),
    cex = 0.5
  )
  text(
    x = data.mca$ind$coord[landmarks,dim1:dim2],
    y = as.character(landmarks),
    col = "#ef8a62",
    cex = 1.1
  )
}

# Dimensions 1 and 2.
landmarkind12 <- c(10, 19, 62, 97) # each ends of each axis
print(data[landmarkind12,active])
plot_with_landmarks(1, 2, landmarkind12)

# Dimensions 2 and 3.
```

```
landmarkind23 <- c(46, 83, 143, 149)
print(data[landmarkind23, active])
plot_with_landmarks(2, 3, landmarkind23)

# Subclouds

# helper function

plot_subclouds = function(
  dim1,
  dim2,
  group1,
  group2,
  group1text,
  group2text) {

  title = sprintf(
    "Cloud of %d individuals projected on dimensions %d and %d: blue = %s, red = %s",
      num.ind, dim1, dim2, group1text, group2text
  )

  plot(
    x = data.mca,
    choix = "ind",
    invisible = c("var", "quali.sup"),
    title = title,
    label = "none",
    axes = c(dim1,dim2),
    col.ind = "grey"
  )
  text(
    x = ind.coords[group1,dim1:dim2],
    "+",
    col = "red",
    cex = 0.8
  )
  text(
    x = ind.coords[group2,dim1:dim2],
    "o",
    col="blue",
    cex=0.8
  )
}

# Subclouds - ConsumerType.
indBadVeryBad <- rep(FALSE, num.ind)
indBadVeryBad[data[, "ConsumerType"] == "very bad"] <- TRUE
indBadVeryBad[data[, "ConsumerType"] == "bad"] <- TRUE
```

```
indGoodVeryGood <- rep(FALSE, num.ind)
indGoodVeryGood[data[, "ConsumerType"] == "good"] <- TRUE
indGoodVeryGood[data[, "ConsumerType"] == "very good"] <- TRUE
print(length(indBadVeryBad[indBadVeryBad]))
print(length(indGoodVeryGood[indGoodVeryGood]))

plot_subclouds(1, 2, indBadVeryBad, indGoodVeryGood, "good consumer", "bad consumer")

# Subclouds - WeekOrWeekend.
indWeek <- rep(FALSE, num.ind)
indWeek[data[, "WeekOrWeekend"] == "Week"] <- TRUE
indWeekend <- rep(FALSE, num.ind)
indWeekend[data[, "WeekOrWeekend"] == "Week-end"] <- TRUE
print(length(indWeek[indWeek]))
print(length(indWeekend[indWeekend]))

plot_subclouds(1, 2, indWeek, indWeekend, "weekday", "weekend")
plot_subclouds(2, 3, indWeek, indWeekend, "weekday", "weekend")

# Subclouds - EatInOrTakeaway.
indEatIn <- rep(FALSE, num.ind)
indEatIn[data[, "EatInOrTakeaway"] == "On the premises"] <- TRUE
indTakeAway <- rep(FALSE, num.ind)
indTakeAway[data[, "EatInOrTakeaway"] == "Take-away"] <- TRUE
print(length(indEatIn[indEatIn]))
print(length(indTakeAway[indTakeAway]))

plot_subclouds(1, 2, indEatIn, indTakeAway, "eat-in", "take-away")
plot_subclouds(2, 3, indEatIn, indTakeAway, "eat-in", "take-away")

# Subclouds - Lunch vs Not Lunch.
indLunch <- rep(FALSE, num.ind)
indLunch[data[, "Lunch"] == "Yes"] <- TRUE
indNotLunch <- rep(FALSE, num.ind)
indNotLunch[data[, "Lunch"] == "No"] <- TRUE
print(length(indLunch[indLunch]))
print(length(indNotLunch[indNotLunch]))

plot_subclouds(1, 2, indLunch, indNotLunch, "lunch", "not lunch")
plot_subclouds(2, 3, indLunch, indNotLunch, "lunch", "not lunch")

# Contributions of categories to dimensions.

# Top 15 dimension 1 & 2.
plot(
  x = data.mca,
  invisible = c("ind", "quali.sup"),
  title = "Top 15 categories contributing to dimensions 1 and 2",
  habillage = "quali",
```

```
  selectMod = "contrib 15"
)

# Top 15 dimension 2 & 3.
plot(
  x = data.mca,
  invisible = c("ind", "quali.sup"),
  title = "Top 15 categories contributing to dimensions 2 and 3",
  habillage = "quali",
  axes = c(2,3),
  selectMod = "contrib 15"
)

# Find categories which contribute more than the average - to a single dimension.

# Dimension 1
avg.contr <- 100 / num.categories
dim1.contr <- data.mca$var$contr[,1]
dim1.above.avg.contr <- dim1.contr[dim1.contr > avg.contr]
dim1.above.avg.contr.names = names(dim1.above.avg.contr)
print(round(dim1.above.avg.contr, digits = 1))

# 11 categories with total 81.37% contribution
print(
  sum(data.mca$var$contr[dim1.above.avg.contr.names,1])
)

plot(
  x = data.mca,
  invisible = c("ind", "quali.sup", "ind.sup"),
  selectMod = dim1.above.avg.contr.names,
  habillage = "quali",
  autoLab = "yes",
  title = "11 categories with above average contribution to dimension 1",
  cex = 1.0
)
mtext(
  side = 2,
  text = "Frequent Customer",
  col = "brown",
  font = 4,
  cex = 1.3
)
mtext(
  side = 4,
  text = "Infrequent Customer",
  col = "brown",
  font = 4,
  cex = 1.3
```

```
)

dim2.contr <- data.mca$var$contr[,2]
dim2.above.avg.contr <- dim2.contr[dim2.contr > avg.contr]
dim2.above.avg.contr.names = names(dim2.above.avg.contr)
print(round(dim2.above.avg.contr, digits = 1))

# 11 categories with total 83.88% contribution to dimension 2
print(
  sum(data.mca$var$contr[dim2.above.avg.contr.names,2])
)

plot(
  x = data.mca,
  invisible = c("ind", "quali.sup", "ind.sup"),
  selectMod = c(dim2.above.avg.contr.names, "Alone_No"),
  habillage = "quali",
  autoLab = "yes",
  title = "11 categories with above average contribution to dimension 2",
  cex = 1.0
)
mtext(
  side = 3,
  text = "Couple Weekday Take-Out",
  col = "brown",
  font = 4,
  cex = 1.3
)
mtext(
  side = 1,
  text = "Solo Weekend Eat-In",
  col = "brown",
  font = 4,
  cex = 1.3
)

# Dimension 3
dim3.contr <- data.mca$var$contr[,3]
dim3.above.avg.contr <- dim3.contr[dim3.contr > avg.contr]
dim3.above.avg.contr.names = names(dim3.above.avg.contr)
print(round(dim3.above.avg.contr, digits = 1))

# 11 categories with total 82.28% contribution to dimension 3
print(
  sum(data.mca$var$contr[dim3.above.avg.contr.names,3])
)

plot(
  x = data.mca,
```

```
  invisible = c("ind", "quali.sup", "ind.sup"),
  selectMod = dim3.above.avg.contr.names,
  habillage = "quali",
  axes = c(3,4),
  autoLab = "yes",
  title = "11 categories with above average contribution to dimension 3",
  cex = 1.0
)
mtext(
  side = 4,
  text = "Family eat-in lunch",
  col = "brown",
  font = 4,
  cex = 1.3
)
mtext(
  side = 2,
  text = "Not family, not lunch, take-away",
  col = "brown",
  font = 4,
  cex = 1.3
)


# Concentration Ellipses.
plotellipses(
  model = data.mca,
  keepvar = "ConsumerType",
  label = "none",
  means = "FALSE",
  axes = c(1,2)
)

plotellipses(
  model = data.mca,
  keepvar = "WeekOrWeekend",
  label = "none",
  means = "FALSE",
  axes = c(1,2)
)

plotellipses(
  model = data.mca,
  keepvar = "Family",
  label = "none",
  means = "FALSE",
  axes = c(2,3)
)
```

```
plotellipses(
  model = data.mca,
  keepvar = "Lunch",
  label = "none",
  means = "FALSE",
  axes = c(2,3)
)

# Hierarchical clustering.
res.hc.ind <- HCPC(data.mca,  nb.clust=5)
print(res.hc.ind)

res.hc.var <- HCPC(data.frame(data.mca$var$coord, nb.clust=-1))
print(res.hc.var)
```