# Category characteristics survey for the IRI Marketing Science dataset

Data collected by Datta, Van Heerde, and Ailawadi (2017) on Amazon Mechanical Turk (May 2016) to measure category characteristics for the IRI Marketing Science data set (Bronnenberg, Kruger, and Mela 2008).

**Files**

grocery\_survey.sav Raw data in SPSS format

grocery\_survey.csv Raw data in CSV format

preclean.R Converts SPSS raw data to constructs, and validates them (e.g., computation of Cronbach alphas). Writes construct means and standard deviations for constructs to output file (survey.csv). Converts raw SPSS data to CSV file.

survey.txt Report file, generated by preclean.R

survey.csv Final data set with means and standard deviations per construct.

**Constructs**

| **Construct** | **Measures (variables in brackets)** | **Scale** | **Adapted from** |
| --- | --- | --- | --- |
| Hedonic nature of category *[cat\_hedonic]*  (Cronbach α = .81) | Please rate category X on how Not Fun / Fun it is *[fun]*.  Please rate category X on how Unenjoyable / Enjoyable it is *[enjoyable]*.  (Dummy variable *hedonism* set to 1 if responded evaluated this construct). | 7-point semantic differential scale:  Not Fun = 1; Fun = 7   Unenjoyable = 1; Enjoyable = 7 | Voss, Spangenberg, and Grohmann, (2003) |
| Functional / performance risk *[cat\_perfrisk]*  (Cronbach α = .60) | There is much to lose if you make the wrong choice in category X *[muchtolose]*.  In category X, there are large differences in quality between the various products *[largedifferences]*.  (Dummy variable *performance* set to 1 if responded evaluated this construct). | 5-point Likert scale:  1= strongly disagree; 2 = somewhat disagree;  3 = neither agree nor disagree;  4 = somewhat agree; 5 = strongly agree. | Steenkamp and Geyskens (2014)  Own development |
| Social value / social demonstrance *[cat\_socdemon]*  (Cronbach α = .85) | You can tell a lot about a person from the brand of category X he or she buys *[tellalot]*.  The brand of category X a person buys says something about who they are *[sayswhotheyare]*.  (Dummy variable *socialdemonstrance* set to 1 if responded evaluated this construct). | 5-point Likert scale:  1= strongly disagree; 2 = somewhat disagree; 3 = neither agree nor disagree;  4 = somewhat agree; 5 = strongly agree. | Steenkamp and Geyskens (2014) |
| Involvement *[cat\_involv]*  (Cronbach α = .73) | The X category is very important to me *[important]*.  The X category interests me a lot *[interest]*.  (Dummy variable *investment* set to 1 if responded evaluated this construct). | 5-point Likert scale:  1= strongly disagree; 2 = somewhat disagree; 3 = neither agree nor disagree;  4 = somewhat agree; 5 = strongly agree. | Steenkamp, van Heerde and Geyskens 2010 |
| Utilitarian nature of category *[cat\_utilit]*  (Cronbach α = .82) | Please rate category X on how Practical / Not practical it is [practical].  Please rate category X on how Necessary / Unnecessary it is [necessary].  (Dummy variable *utility* set to 1 if responded evaluated this construct). | 7-point semantic differential scale:  Not Fun = 1; Fun = 7   Unenjoyable = 1; Enjoyable = 7 | Voss, Spangenberg, and Grohmann, (2003) |

**Description**

We ask each respondent to rate all categories on two of the constructs listed below, easing the burden to answer too many questions. Each construct is measured using two items. Dummy variables in the raw data indicate which construct was evaluated by which respondent.

The order of categories is randomized across respondents, but fixed for each respondent.

We select US adults 21 years and older when admistring the survey. For each of the categories, we ask respondents whether they have at least purchased it once in the past six months. If the answer to this question is no, the category is not shown to respondents.

**Additional variables**

1. We ask respondents to indicate how much of their grocery shopping is done by the respondent *[amtofshopping]*:

Please tell us how much of the grocery shopping for your household is done by you:  
1 = None of it; 2 = Some of it; 3 = About half of it; 4 = Most of it; 5= All of it

1. What is your age *[age]*?

1 = 16-27 years

2 = 28-34 years

3 = 35-44 years

4 = 45-54 years

5 = 55-64 years

6 = 65-79 years

7 = 80 years or older

1. Please indicate your gender *[gender]*:

1 = Male

2 = Female

**Categories**

1. Beer
2. Carbonated Soft Drinks
3. Cigarettes
4. Coffee
5. Cold (Ready to Eat) Cereal
6. Deodorants
7. Disposable Diapers
8. Frozen Pizza and Dinners
9. Household Cleaners
10. Ketchup
11. Laundry Detergents
12. Margarine and Spreads
13. Mayonnaise
14. Milk
15. Mustard
16. Pasta Sauce
17. Peanut Butter
18. Razors and Blades
19. Salty Snacks
20. Shampoo
21. Soup
22. Sugar Substitutes
23. Toilet Tissue
24. Toothpaste
25. Yogurt

**References**

Bronnenberg, B. J., Kruger, M. W., & Mela, C. F. (2008). Database paper—The IRI marketing data set. *Marketing science*, *27*(4), 745-748.

Datta, H., Ailawadi, K. L., & van Heerde, H. J. (2017). How Well Does Consumer-Based Brand Equity Align with Sales-Based Brand Equity and Marketing-Mix Response?. *Journal of Marketing*, *81*(3), 1-20.

Steenkamp, J. B. E., & Geyskens, I. (2013). Manufacturer and retailer strategies to impact store brand share: Global integration, local adaptation, and worldwide learning. *Marketing Science*, *33*(1), 6-26.

Steenkamp, J. B. E., Van Heerde, H. J., & Geyskens, I. (2010). What makes consumers willing to pay a price premium for national brands over private labels?. *Journal of Marketing Research*, *47*(6), 1011-1024.

Voss, K. E., Spangenberg, E. R., & Grohmann, B. (2003). Measuring the hedonic and utilitarian dimensions of consumer attitude. *Journal of Marketing Research*, *40*(3), 310-320.