|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Supplemental table 2. Goodness-of-fit (R2 and pseudo F-statistic) by number of cluster solutions for beverages and food store shopping patterns** | | | | | | | | |
| **Beverages** | | | | | | | | |
| Numer of cluster solutions | R2 | pseudo  F-statistic | Proportion of HH within cluster (%) | Proportion of purchases made by households in each cluster by store-type (%) | | | | |
| Convenience stores | Supermarkets | Wholesalers | Traditional stores | Others |
| 1 | 0.41 | 3844.97 | 57 | 5 | 13 | 1 | 76 | 5 |
| 2 | 43 | 11 | 33 | 5 | 23 | 28 |
|  |  |  |  |  |  |  |  |  |
| 1 | 0.63 | 4696.62 | 28 | 14 | 47 | 6 | 26 | 6 |
| 2 | 55 | 5 | 12 | 1 | 77 | 5 |
| 3 | 17 | 4 | 11 | 2 | 23 | 60 |
|  |  |  |  |  |  |  |  |  |
| 1 | 0.71 | 4579.95 | 48 | 2 | 12 | 1 | 80 | 5 |
| 2 | 14 | 4 | 65 | 4 | 20 | 6 |
| 3 | 21 | 24 | 23 | 7 | 38 | 7 |
| 4 | 17 | 4 | 11 | 1 | 23 | 61 |
|  |  |  |  |  |  |  |  |  |
| 1 | 0.77 | 4554.52 | 21 | 4 | 33 | 7 | 49 | 7 |
| 2 | 11 | 5 | 69 | 5 | 15 | 6 |
| 3 | 42 | 3 | 9 | 1 | 83 | 5 |
| 4 | 17 | 4 | 11 | 1 | 23 | 61 |
| 5 | 10 | 45 | 15 | 2 | 31 | 5 |
| **Foods** | | | | | | | | |
| Numer of cluster solutions | R2 | pseudo F-statistic | Proportion of HH within cluster (%) | Proportion of purchases made by households in each cluster by store-type (%) | | | | |
| Convenience stores | Supermarkets | Wholesalers | Traditional stores | Others |
| 1 | 0.63 | 9216.03 | 57 | 2 | 68 | 5 | 21 | 4 |
| 2 | 43 | 1 | 21 | 3 | 71 | 3 |
|  |  |  |  |  |  |  |  |  |
| 1 | 0.76 | 8504.96 | 35 | 2 | 79 | 3 | 13 | 3 |
| 2 | 36 | 2 | 45 | 8 | 41 | 5 |
| 3 | 29 | 1 | 14 | 2 | 80 | 3 |
|  |  |  |  |  |  |  |  |  |
| 1 | 0.82 | 8233.01 | 28 | 2 | 83 | 2 | 11 | 2 |
| 2 | 23 | 1 | 11 | 1 | 84 | 3 |
| 3 | 24 | 3 | 53 | 12 | 27 | 5 |
| 4 | 25 | 2 | 37 | 3 | 55 | 4 |
|  |  |  |  |  |  |  |  |  |
| 1 | 0.86 | 8138.36 | 26 | 2 | 84 | 2 | 9 | 3 |
| 2 | 23 | 2 | 36 | 2 | 56 | 4 |
| 3 | 22 | 1 | 10 | 1 | 85 | 2 |
| 4 | 6 | 3 | 33 | 40 | 18 | 6 |
| 5 | 23 | 2 | 60 | 2 | 31 | 4 |
| Source: Authors’ own analyses and calculations based on data from Nielsen through its Mexico Consumer Panel Service (CPS) for the food and beverage categories for January 2015 -December 2015 The Nielsen Company, 2016. Nielsen is not responsible for and had no role in preparing the results reported herein. Cluster solutions were derived from k-means cluster analyses using SAS FASTCLUST, SAS version 9.4, with an iterative process of 1,000 replications using randomly selected seeds. We used the pseudo F-statistic (Calinski and Harabasz) and the r-squared (R2) of each solution to identify the optimal number of clusters. A higher F- statistic value indicates a better intra-cluster homogeneity and inter-cluster heterogeneity while a higher R2 indicates a larger proportion of the variance of purchases volume explained by the cluster solution. | | | | | | | | |