An International Comparison of Fruit Consumption Patterns: A Cluster Analysis

Yan Heng[a] and Lisa A. House[b]

[a]Post-doctoral Associate, Food and Resource Economics Department, PO Box 110240, University of Florida, Gainesville, Florida, 32611, yheng@ufl.edu

[b]Professor, Food and Resource Economics Department, PO Box 110240, University of Florida, Gainesville, Florida, 32611, lahouse@ufl.edu

Selected Presentation at the Agricultural & Applied Economics Association's 2017 AAEA Annual Meeting, Chicago, IL, July 30-August 1, 2017

Abstract

This study clusters consumers based on consumption frequency of 15 fresh fruits and results in three segments: low-frequency consumer, common-fruit consumer, and high-frequency consumer. These consumer segments are heterogeneous in socioeconomics and shopping behavior across regions. Overall, the high-frequency consumer cluster had more individuals who were older, married, not single/never married, self-reported healthy and physically active, and living in North America and Europe. The low-frequency consumer group had a larger number of individuals who were younger, living alone, single/never married, self-reported unhealthy and not active, and living in Asian countries. Moreover, the high-frequency fruit consumers tended to focus on many fruit attributes, such as freshness, nutritional value, origin, and in the season, but not focus on the price. On the other hand, the low-frequency consumers tended to focus only on price and were less likely to pay attention to other attributes.

Introduction

As a major source of nutrition, fruit consumption is essential for human health. Studies have indicated that high consumption of fruits and vegetables may contribute to lower body mass index (BMI) and prevent several kinds of cardiovascular disease and cancer (Vecchia, 2004). To promote fruit consumption, diet recommendations and guidelines have been released by government agencies across countries. For example, the recently released U.S. Dietary Guidelines 2015-2020 recommend consumption of 2 cups of fruits daily based on a 2000-calorie intake level (ODPHP, 2015); dietary guidelines in UK and France recommend to eat at least five portions of fruits and vegetables every day (FAO, 2016); and the Japanese dietary guideline recommends to consume 2 servings of fruits per day (FAO, 2016a). However, the actual intakes of these foods across the world are still consistently below recommended levels (Konopacka et al, 2010; Dehghan et al., 2011). Therefore, identifying factors influencing fruit consumption and consumer groups with the low intake is critical to developing effective strategies to boost consumption.

Studies have shown that food consumption behavior is determined by various factors, including physiological factors, socioeconomics, behavior and lifestyle, and knowledge and attitudes. For example, higher intake of fruits and vegetables are found to be mostly associated with higher household income and higher educational attainment; and married people and females are found to have higher intake than their counterparts (Estaguio et al, 2008; Irala-Estevez et al. 2000; Tamers et al. 2009; Kamphuis et al. 2006). Availability, cost, and convenience are also important factors and most likely barriers for consumption (Baker and Wardle, 2003). Culture is another factor playing an important role in food consumption. For example, the Mediterranean diet contains a higher level of fruits than a conventional Western

diet (Hu et al, 2000) Although many studies have examined dietary patterns within a country or region (e.g. in Europe), few have attempted to compare across regions to understand the common and distinct factors that influence fruit intake among different cultures. Tamers et al. (2009) compared the fruit and vegetable consumption patterns in the U.S. and France and examined the relationship between fruit and vegetable consumption and demographic variables. They reported that French tended to eat fruits and vegetables more often than Americans and fruit and vegetable consumption was strongly associated with age, BMI, and educational attainment. Pingali (2007) reported the trend of westernization of Asian diets that increased consumption of temperate fruit, such as apples. Given the heterogenous diets in different regions, one interest of this study is to compare current fruit consumption cross regions.

The objective of this study is to compare and understand the fruit consumption patterns across regions. Previous research has found that fruit intake increases faster than vegetable intake, and this might be because fruits are perceived as more convenient and tasty than vegetables (Billson et al. 1999; Dinnella et al, 2016). Our study uses a unique dataset covering eight countries across North America, Europe, and Asia. Moreover, we consider not only socioeconomic and lifestyle variables but also factors influencing fruit choices in consumers' shopping trips. Our study will contribute to the literature by providing a comprehensive comparison of international fruit consumption patterns and identifying the important factors driving fruit purchase decisions. Such results will help to develop more effective strategies to enhance fruit purchase and consumption in different regions.

To achieve the goal of this study, we use a dataset from an international survey on fresh fruit consumption frequency on 15 types of fruits. The data was collected via an online survey that also collected information on lifestyles, health status, demographics, and shopping behavior.

A hierarchical clustering method is used to identify distinct consumer clusters and understand consumer heterogeneity based on the frequency of fruit intakes. The characteristics of each consumer segment are examined using both socioeconomic variables and factors influencing shopping behavior.

Methods and Data

Survey Instrument

A survey to collect information on fruit consumption, demographics, and lifestyle factors was developed and translated into English, French, Dutch, Korean, and Japanese. Respondents from the U.S., Canada, Japan, South Korea, the U.K., France, Netherlands, and Belgium were recruited from online panels (Toluna). After pre-testing, the survey was launched in April 2016. The survey consisted of screening questions, general questions, demographic questions, and a validation question. The screening questions narrowed the respondents to female primary grocery shoppers, at least 20 years of age. The general questions collected information on shopping behavior and consumption patterns on fresh fruits as well as their health status and physical activity level. Demographic information included country of residence, age, education, household status, employment status, and marital status. A validation question was included to ensure the sample quality (Jones et al, 2016, Heng et al. 2017). In this study, the validation question asked: "What month of the year it is?" If respondents failed to select "April" for the question, they were terminated from the survey (the survey was conducted in April).

Respondents were asked to rate their consumption frequency for 15 types of fruits in the past six months. The fruits include apples, oranges, bananas, berries (strawberries, blueberries, or raspberries), mangoes/papaya, grapes, pineapples, peaches/plums, grapefruit, melons (cantaloupes/honeydew), watermelons, lemons or limes, mandarin

oranges/tangerines/clementine, pear, and kiwi. Respondents chose their consumption frequency from “About every day”, “About every week”, “About every month”, “At least one time in the past six months”, and “Not consumed in the past six months”. Although the intake of fruits might vary throughout the year, these effects were equal for all respondents because they participated during the same period (Beck et al. 2015).

Socioeconomic data collected in this study included age, household status (live alone, live with a significant other but do not have children in household, live with a significant other and have children in household, live only with children), education (high school or lower, technical or associate or equivalent, university graduate or higher), marital status (single, married, other), and employment status (unemployed, employed full-time, employed part-time, and other). Self-reported health status and physical activity level were also collected. To capture cultural heterogeneity, country of residence was recorded. To simply the analysis, the country of residence was categorized into Asia (Japan and South Korea), North America (the U.S. and Canada), and Europe (France, Netherlands, the U.K., and Belgium).

In addition, respondents were asked to rate the importance of factors influencing their choices of fresh fruits using a 5-point Likert scale. The considered factors include price, freshness, sweetness, juiciness, nutritional value, grown in preferred origin, locally grown, quality appearance, ease of peeling, when fruit is in the season, trendy/try something new, available in primary shopping locations, and recipe diversity. Importance rating for influential factors on fresh fruit purchases was coded into binary variables (important, neutral or not important). Sample characteristics and associated frequenciey are presented in table 1.

Statistical Analysis

Cluster analysis can be used to determine distinct consumer segments based on their characteristics without any assumptions regarding the dataset (Visschers et al., 2013; Beck et al., 2015; Pohjolainen et al., 2016). In this study, we applied a hierarchical clustering method to identify consumer segments based on consumption frequency of 15 types of fruits. Since fruit consumption frequency was assessed on a 5-point scales, similar to Beck et al (2015) we recoded the options to an annual base using following steps: daily was coded as 365 times a year, weekly was coded as 104 times a year assuming an average twice a week, monthly was coded as 12 times a year, at least once in the past 6 months as coded as 3 times a year, and not consumed in the past 6 months were coded as zero times a year. Ward's method algorithm was used to perform the cluster analysis using SAS 9.4. This hierarchical clustering approach generates a tree-like gram representing observations, called a dendrogram. Euclidean distance is used to determine the similarity between two observations. Similar observations are assigned to the same cluster, whereas dissimilar observations are assigned into different clusters. While there are no strict rules to define the number of clusters, the final clusters from hierarchical clustering are determined by cutting the dendrogram based on the differences between clusters and the interpretability. A Generalized Linear Model (GLM) was applied to test whether the differences among clusters are statistically significant. To identify the association between cluster membership and cluster characteristics (socioeconomics and other factors), a chi-square test is applied for categorical variables through frequencies across clusters.

Results

Characteristics of Respondents

A total sample of 7,793 completes was collected from the U.S. (995), Canada (935), Japan (966), South Korea (951), France (1043), the U.K. (994), the Netherlands (954), and Belgium (955).

Among these participants, around 45% lived with children, over half held a bachelor or higher educational attainment, about 65% were married/partnered, and around 46% held a full-time job. Over 60% of participants reported that they were somewhat or very healthy, and nearly 40% reported that they were physically active and exercised for at least 30 minutes per day for at least 3 days per week (table 1).

Cluster Analysis

Three distinct clusters were identified through Ward's clustering method, and average consumption frequency annually for each fruit are reported in table 2. The differences in consumption frequency of fresh fruits are statistically significant at the 1% level across clusters. The first cluster represents approximate half of the sample. Participants in this cluster reported overall low consumption frequencies for all considered fruits compared to other clusters. Therefore, this cluster is referred to as the "low-frequency consumer". The second cluster consists of nearly one-quarter of the sample. Participants in this cluster reported overall higher consumption frequencies of fruits and particularly high consumption of apples, oranges, and bananas. Moreover, this cluster has the highest consumption frequency of apples and bananas across groups (consumption frequencies of apples and bananas were about 246 times and 222 times per year, respectively). We refer to this cluster as the "common-fruit consumer". The third cluster represents about 30% of the sample, and consumers in this cluster reported overall high consumption frequencies for all considered fruits. On average, this group consumed oranges, berries, grapes, mandarin, and kiwi more than 100 times per year besides their high consumption of apples and bananas. This cluster is referred to as the "high-frequency consumer". The results show that a larger proportion of consumers fell in the low-frequency consumer segment, which is

consistent with previous findings that fruit intake is generally low in consumers across countries (Konopacka et al, 2010; Dehghan et al., 2011).

Socioeconomics across Clusters

Examining demographics and lifestyle variables across clusters, the hypothesis that considered variables are independent of cluster membership was rejected at the 1% level, with the exception of educational level. As a result, we can say that these factors are drivers of fruit consumption across countries. The frequencies and adjusted standardized residuals for each variable are reported in table 3. If an adjusted standard residual is positive (negative), the frequency in the cell is higher (lower) than expected. If an adjusted standard residual is greater than 2 or smaller than -2, the difference between the cell frequency and expected value is statistically significant at 5% level (Beck et al 2015).

Results show that Asian respondents are more likely to be in the low-frequency cluster, while respondents from North America and Europe were more likely to be associated with the high-frequency consumer group. According to the Food Balance Sheets reported by the FAO (2016b), the average fruit intake per capita per day was 62 kcal in Japan and South Korea, 240 kcal in Canada and the U.S., and 115 kcal in the U.K., France, Netherlands, and Belgium. Our results were consistent with the report that fruit consumption in Asia was lower than in North America and Europe.

Younger consumers (20-39) were more likely to be associated with the low-frequency consumer group, while older consumers (60 or above) were more likely to be associated with the high-frequency fruit consumer group. Such results can be supported by previous findings that fruit intake usually increases with age and elderly consume more fruits and vegetables (Estaguio et al., 2008; Tamers et al., 2009; Dubuisson et al., 2010). Similarly, we found that being married

and having children are drivers for fruit intake, consistent with previous studies (Billson et al., 1999; Kamphuis et al., 2006; Estaguio et al., 2008). Our results show that the low-frequency consumer cluster had more consumers who lived alone and had fewer consumers who lived with children, whereas the high-frequency consumer cluster had significantly fewer consumers who lived alone. Also, the low-frequency consumer group had a higher number of single/never married individuals, whereas the high-frequency group had a significantly lower number of consumers who were single/never married. Both common-fruit consumer and high-frequency consumer groups had more individuals who are married/partnered. The common-fruit consumer group consisted of more individuals in middle-aged (40-59), married/partnered, and lived with children.

Though educational level has been reported to be associated with fruit and vegetable intakes in previous studies (Irala-Estevez et al., 2000; Perrin et al., 2002; Estaguio et al., 2008; Tamers et al., 2009), educational attainment is not a major driver of fruit intake in our sample. This suggests that there is no significant difference in educational levels across clusters. Also, our results show that the proportion of consumers holding a full-time job is significantly higher in the low-frequency consumer cluster but lower in the high-frequency consumer cluster.

The low-frequency consumer group had a significantly higher proportion of individuals who self-reported not being healthy or physically active, whereas the high-frequency consumer groups consisted of a significantly higher proportion of individuals who self-reported being healthy and physically active. The healthy status was not a driving factor in the common-fruit consumer group, while this group had a higher number of individuals who reported being physically active. This association between fruit intake frequency and physically activity is consistent with previous studies (Estaguio et al., 2008; Keller et al., 2014).

Influential Factors on Fruit Purchases across Clusters and Regions

Food purchase behavior would also be expected to directly influence consumption patterns, therefore, factors influencing shopping may also impact cluster membership. We asked respondents to rate the importance for the following factors when making shopping decisions: price, freshness, sweetness, juiciness, nutritional value, grown in preferred origin, locally grown, quality appearance, ease of peeling, when fruit is in the season, trendy/try something new, available in primary shopping locations, and recipe diversity. Their answers were coded into "important" and "not important". The percentages of "important" and "not important" for each factor and corresponding adjusted standardized residuals are reported in table 4. The hypothesis of independence between these factors and the cluster membership was rejected at 5% level except for the sweetness.

A significantly larger number of respondents in the low-frequency consumer cluster considered price as an important factor influencing their fruit purchases, while this group has significantly less consumers who focus on other attributes. In the high-frequency consumer group, a smaller proportion of consumers prioritized the price, while a larger number of consumers in this group focus on freshness, juiciness, nutritional value, origin, quality appearance, in season, trendy, availability, and recipe diversity. Such shopping behavior might indicate that these high-frequency consumers have better awareness and knowledge about fruit attributes. The common-fruit consumer group focused on freshness, nutritional value, and season, and this group placed less attention on the ease of peeling, which could because that the fruits they mostly frequently consumed were apples and bananas. This group of consumers seems to understand the nutritional value of fresh fruits while more care about the basic attributes of fruits.

Since culture could influence fruit purchase and consumption, we compare the percentage of clusters across regions (table 5). Freshness and price were the most cared about factors in all regions. Generally, Europe had a larger proportion of consumers who prioritized freshness and availability in their primary shopping locations compared to consumers in North America and Asia. A larger proportion of consumers from North America placed higher importance on nutritional value, locally grown, quality appearance, and recipe diversity in all clusters compared to other regions. More Asian consumers focused on sweetness, preferred origin, ease of peeling, in season, and trendy than North American and European consumers in all clusters.

We further investigate the impacts of factors on cluster membership in each region. In North America, a significantly larger number of consumers in the high-frequency group value factors except price and appearance. The common-fruit group focused on nutritional value and trendy, and the low-frequency group paid less attention to these factors. In Europe, the high-frequency fruit consumers focused on all factors except price and ease of peeling. A significantly smaller number of consumers in the low-frequency group paid attention to freshness, juiciness, nutritional value, origins, in season, and availability. The common-fruit consumer cluster had fewer consumers who cared sweetness and origin, and this might indicate that the quality of common fruits, like apples and bananas, were consistent in term of sweetness and origin in Europe. In Asia, the low-frequency group paid less attention to freshness, nutritional value, origin, in season, trendy, and recipe diversity. The common-fruit consumers paid more attention to freshness, nutritional value, locally grown, and in season. The high-frequency group was less focused on price but more focused on nutritional value, origin, appearance, trendy, and recipe diversity.

Conclusion

Healthy food intakes are essential for human wellbeing, and fruits are considered a part of the healthy diet worldwide. However, consistent lower fruit intake than recommended levels across countries call for effective strategies to boost the consumption of healthy foods. Our study collected novel data covering eight countries in North America, Europe, and Asia to investigate and compare current fruit consumption patterns internationally. A cluster analysis was conducted to identify distinct consumer segments based on their consumption frequency of 15 types of fruit. The characteristics of each cluster were examined and compared across regions.

Our study shows that consumers can be classified into three distinct segments based on fruit intake frequency: “high-frequency consumer”, “common-fruit consumer”, and “low-frequency consumer”. Nearly half of the participants belong to the low-frequency consumer group, emphasizing the need for more efforts to boost fruit consumption. Comparing the three consumer groups, we found that consumers living in North American and European countries are more likely to be associated the high-frequency consumer group over consumers in Asian countries. Fruit intake frequency is positively associated with age, healthy status, and physical activity level. Also, participants who lived alone and who were single/never married are more likely to be associated with the low-frequency consumer group. Educational level is not a major factor influencing fruit consumption patterns in this sample, while female’s employment status is found to be negatively associated with fruit consumption frequency.

We also found that consumer segments are heterogeneous in their shopping behavior. Price is the major factor driving fruit consumption in the low-frequency consumer segment, whereas high-frequency consumers pay more attention to factors other than price. Common-fruit consumers understand the nutritional value of fruits but care more about basic attributes like freshness and in season. Comparing different regions in general, we observe that European

consumers are more likely to value freshness and availability, North American consumers focus more on nutritional value, locally grown, appearance, and recipe diversity, and Asian consumers tend to pay attention to many factors, including price, sweetness, origin, ease of peeling, in season, and trendy.

This study characterizes consumer segments by not only the socioeconomic status but also their shopping behavior factors. These results can be used to develop more effective region-specific strategies to target interested consumer groups and enhance fruit consumption.

References:

Baker A. H. and J. Wardle. 2003. Sex Differences in Fruit and Vegetable Intake in Older Adults. *Appetite*. 40:269-275.

Beck, T.K., S. Jensen, S.H. Simmelsgaard, C. Kjeldsen, and U. Kidmose. Consumer Cluster in Denmark based on Coarse Vegetable Intake Frequency, Explained by Hedonics, Socio-demographic, Health and Food Lifestyle Factors. A Cross-Sectional National Survey. *Appetite*, 91: 366-374.

Billson, H., JA. Pryer, and R. Nichols.1999. Variation in Fruit and Vegetable Consumption Among Adults in Britain. An Analysis from the Dietary and Nutritional Survey of British Adults. *European Journal of Clinical Nutrition*, 53: 946-952.

Dehghan, M., N. Akhtar-Danesh, and A.T.Merchant. 2011. Factors Associated with Fruit and Vegetable Consumption among Adults. *Journal of Human Nutrition and Dietetics*, 24(2):128-134.

Dinnella, C., D. Morizet, C. Masi, D. Cliceri, L. Depezay, K.M. Appleton, A. Giboreau, F.J. Perez-Cueto, H. Hartwell, and E. Monteleone. 2016. Sensory Determinants of Stated Liking for Vegetable Names and Actual Liking for Canned Vegetables: A Cross-Country Study among European Adolescents. *Appetite*, 107:339-347.

Estaguio C., N. Druesne-Pecollo, P. Latino-Martel, L. Dauchet, S. Hercberg, and S. Bertrais. 2008. Socioeconomic Differences in Fruit and Vegetable Consumption among Middle-Aged French Adults: Adherence to the 5 A Day Recommendation. *Journal of the American Dietetic Association*, 108(12): 2021-2030.

Food and Agriculture Organization of the United Nations (FAO). 2016a. Food-Based Dietary Guidelines. Available at http://www.fao.org/nutrition/education/food-dietary-guidelines/regions/en/.

Food and Agriculture Organization of the United Nations (FAO). 2016b. Food Balance Sheets. Available at http://faostat3.fao.org/download/FB/FBS/E.

Heng, Y., H. Kim, and L. House. 2017. An Overview of the Grapefruit Market in South Korea. Available at: https://edis.ifas.ufl.edu/pdffiles/FE/FE100300.pdf.

Hu, F.B., E.B. Rimm, M.J. Stampfer, A. Ascherio, D. Spiegelman, and W.C. Willett. 2000. Prospective Study of Major Dietary Patterns and Risk of Coronary Heart Disease in Men. *The American Journal of Clinical Nutrition*, 72(4): 912-921.

Irala-Estevez, J., M. Groth, L. Johansson, U. Oltersdorf, R. Prattala, and MA. Martinez-Gonzalez. 2000. A Systematic Review of Socio-Economic Differences in Food Habits in Europe: Consumption of Fruit and Vegetables. European Journal of Clinical Nutrition, 54:706-714.

Kamphuis, C.B.M., K. Giskes, G. Bruijn, W.Wendel-Vos, J. Brug, and F. Lenthe. 2006. Environmental Determinants of Fruit and Vegetable Consumption Among Adults: A Systematic Review. *British Journal of Nutrition*, 96(4):620-635.

Keller, K., S.R.. Lopes, M.M.C. Moreno, and P.A. Cantero. 2014. Associations between Food Consumption Habits with Meal Intake Behaviour in Spanish Adults. *Appetite*, 83:63-68.

Konopacka, D., K. Jesionkowska, D. Kruczyńska, R. Stehr, F. Schoor., A. Buehler, S. Egger, S. Codarin, C. Hilaire, I. Höller, and W. Guerra. 2010. Apple and Peach Consumption Habits Across European Countries. *Appetite*, 55(3): 478-483.

National Health Service (NHS). 2016. The Eatwell Guide. Available at http://www.nhs.uk/Livewell/Goodfood/Pages/the-eatwell-guide.aspx

Office of Disease Prevention and Health Promotion (ODPHP). 2015. 2015-2020 Dietary Guidelines for Americans. Available at https://health.gov/dietaryguidelines/2015/

Perrin, A.E., C. Simon, G. Hedelin, D. Arveiler, P. Schaffer, and J.L. Schlienger. 2002. Ten-Year Trends of Dietary Intake in a Middle-Aged French Population: Relationship with Educational Level. *European Journal of Clinical Nutrition*, 56(5): 393-401

Pohjolainen, P., P. Tapio, M. Vinnari, P. Jokinen, and P. Rasanen. 2016. Consumer Consciousness on Meat and the Environment-Exploring Differences. *Appetite*, 101: 37-45.

Tamers, SL., T. Agurs-Collins, KW. Dodd, and L. Nebeling. 2009. US and France Adult Fruit and Vegetable Consumption Patterns: An International Comparison. *European Journal of Clinical Nutrition*, 63: 11-17.

Vecchia, CL. 2004. Mediterranean diet and cancer. *Public Health Nutrition*, 7(7):965-968.

Visschers, V.H.M., C. Hartmann, R. Leins-Hess, S. Dohle, and M. Siegrist. 2013. A Consumer Segmentation of Nutrition Information Use and Its Relation to Food Consumption Behavior. *Food Policy*, 42:71-80.

Table 1. Descriptions and frequency of variables

Variable	Description	Frequency
		N= 7793
Age	20-29	18.05%
	30-39	21.57%
	40-49	20.80%
	50-59	21.47%
	60 and over	18.11%
Household status	Live alone	17.62%
	Live with others	37.34%
	Live with children	45.04%
Education	Lower education (high school or lower)	17.14%
	Intermediate education (technical or associate or equivalent)	27.52%
	Higher education (university graduate or higher)	55.33%
Marital status	Single	20.86%
	Married	67.98%
	Other	11.15%
Employment	Unemployed	5.06%
	Employed full-time	46.41%
	Employed part-time	15.95%
	Other	32.58%
Health status	Not healthy	13.78%
	Neither healthy nor unhealthy	19.67%
	Healthy	66.55%
Activity level	Not active	35.94%
	Neither active nor not active	25.78%
	Active	38.28%
Location	Asia	24.77%
	North America	24.60%
	Europe	50.64%

Table 2. Ward's cluster analysis on consumption frequency (average consumption times per year)

Variable	Three clusters			F-test p-value
	Low frequency consumer	Common fruit consumer	High frequency consumer	
Apples	48.07	245.57	204.26	<0.001
Oranges	29.68	117.92	136.89	<0.001
Bananas	52.08	221.20	202.87	<0.001
Berries	27.42	38.88	142.68	<0.001
Mangoes	5.22	8.60	28.52	<0.001
Grapes	19.19	29.75	108.26	<0.001
Pineapples	7.68	11.09	43.54	<0.001
Peaches	4.68	8.10	36.69	<0.001
Grapefruit	6.38	8.32	42.55	<0.001
Melons	4.98	7.10	32.26	<0.001
Watermelons	3.60	4.68	27.31	<0.001
Lemons or limes	13.48	17.65	96.58	<0.001
Mandarin	21.39	31.70	149.37	<0.001
Pear	15.25	24.22	97.76	<0.001
Kiwi	16.32	23.14	103.86	<0.001
Percentage	48.64%	23.05%	28.31%	

Table 3. Statistics of factors by cluster (adjusted standardized residual)

Variable	Three Clusters			Chi-square test p-value
	Low frequency consumer %	Common fruit consumer %	High frequency consumer %	
Age				<0.001
20-39	44.08% (7.82)*	36.41 (-3.17)*	34.59 (-5.71)*	
40-59	42.02% (-0.43)	44.93 (2.61)*	40.53 (-1.96)	
60 and over	13.90% (-9.38)*	18.65 (0.69)	24.89 (9.77)*	
Household status				<0.001
Live alone	20.21 (5.84)*	14.81 (-3.56)*	15.46 (-3.15)*	
Live with others	36.30 (-1.86)	37.53 (0.19)	38.98 (1.88)	
Live with children	43.50 (-2.66)*	47.66 (2.55)*	45.56 (0.58)	
Education				<0.124
Lower education	16.57 (-1.17)	17.09 (0.02)	17.95 (1.28)	
Intermediate education	28.22 (1.35)	26.67 (-0.92)	27.02 (-0.63)	
Higher education	55.18 (-0.26)	56.24 (0.88)	54.85 (-0.54)	
Marital status				<0.001
Single	24.61 (7.92)*	18.65 (-2.63)*	16.23 (-6.33)*	
Married	65.10 (-5.31)*	70.77 (2.88)*	70.67 (3.19)*	
Other	10.29 (-2.36)*	10.58 (-0.88)	13.10 (3.44)*	
Employment				0.001
Unemployed	5.35 (1.17)	4.84 (-0.47)	4.71 (-0.86)	
Employed full-time	48.40 (3.43)*	44.88 (-1.49)	44.24 (-2.41)*	
Employed part-time	16.12 (0.39)	16.54 (0.77)	15.19 (-1.16)	
Other	30.12 (-4.50)*	33.74 (1.2)	35.86 (3.88)*	
Health status				<0.001
Not healthy	15.06 (3.19)*	13.36 (-0.59)	11.92 (-2.99)*	
Neither healthy nor unhealthy	23.11 (7.43)*	18.54 (-1.37)	14.69 (-6.96)*	
Healthy	61.83 (-8.59)*	68.10 (1.59)	73.39 (8.05)*	
Activity level				<0.001
Not active	43.79 (14.05)*	30.46 (-5.52)*	26.93 (-10.42)*	
Neither active nor not active	26.27 (0.97)	26.95 (1.29)	23.98 (-2.28)*	
Active	29.94 (-14.74)*	42.59 (4.29)*	49.09 (12.34)*	
Location				<0.001
North America	20.68 (-8.13)*	23.94 (-0.92)	32.46 (9.88)*	
Asia	32.47 (15.71)*	25.06 (0.51)	10.70 (-17.90)*	
Europe	46.85 (-6.51)*	51.00 (0.35)	56.84 (6.89)*	

Note: *indicates statistically significant at 5% level

Table 4. Statistics of shopping factors across clusters (adjusted standardized residual)

Variable	Three Clusters			Chi-square test p-value
	Low frequency consumer %	Common fruit consumer %	High frequency consumer %	
Price				0.011
Important	83.14 (2.09)*	82.80 (0.73)	80.15 (-3.00)*	
Not important	16.86 (-2.09)*	17.20 (-0.73)	19.85 (3.00)*	
Freshness				<0.001
Important	90.24 (-7.49)*	94.21 (3.08)*	95.10 (5.43)*	
Not important	9.76 (7.49)*	5.79 (-3.08)*	4.90 (-5.43)*	
Sweetness				0.230
Important	68.4 (-0.59)	67.71 (-1.05)	70.08 (1.63)	
Not important	31.6 (0.59)	32.29 (1.05)	29.92 (-1.63)	
Juiciness				<0.001
Important	67.53 (-5.42)*	70.82 (0.44)	75.02 (5.61)*	
Not important	32.47 (5.42)*	29.18 (-0.44)	24.98 (-5.61)*	
Nutritional value				<0.001
Important	61.80 (-13.67)*	72.72 (3.73)*	78.88 (11.68)*	
Not important	38.20 (13.67)*	27.28 (-3.73)*	21.12 (-11.68)*	
Preferred origin				<0.001
Important	33.74 (-6.07)*	35.80 (-1.35)	44.11 (7.99)*	
Not important	66.26 (6.07)*	64.20 (1.35)	55.89 (-7.99)*	
Locally grown				<0.001
Important	43.60 (-8.12)*	47.10 (-1.18)	57.43 (10.11)*	
Not important	56.40 (8.12)*	52.90 (1.18)	42.57 (-10.11)*	
Quality appearance				<0.001
Important	77.31 (-6.34)*	80.85 (0.72)	84.81 (6.36)*	
Not important	22.69 (6.34)*	19.15 (-0.72)	15.19 (-6.36)*	
Ease of peeling				0.035
Important	53.26 (0.25)	50.72 (-2.31)*	54.81 (1.88)	
Not important	46.74 (-0.25)	49.28 (2.31)*	45.19 (-1.88)	
In season				<0.001
Important	72.80 (-7.23)*	78.73 (2.68)*	80.60 (5.51)*	
Not important	27.20 (7.23)*	21.27 (-2.68)*	19.40 (-5.51)*	
Trendy				0.002
Important	17.49 (-2.19)*	17.48 (-1.24)	20.99 (3.59)*	
Not important	82.51 (2.19)*	82.52 (1.24)	79.01 (-3.59)*	
Available in primary shopping locations				<0.001
Important	78.32 (-6.04)*	81.74 (0.82)	85.27 (5.94)*	

Not important	21.68 (6.04)*	18.26 (-0.82)	14.73 (-5.94)*	
	Receipt diversity			<0.001
Important	27.75 (-5.14)*	29.12 (-1.45)	36.36 (7.05)*	
Not important	72.25 (5.14)*	70.88 (1.45)	63.64 (-7.05)*	

Note: *indicates statistically significant at 5% level

Table 5. Influential purchase factors across regions in frequency (adjusted standardized residual)

Variable	North America			Europe			Asia		
	Low-frequency %	Common-fruit %	High-frequency %	Low-frequency %	Common-fruit %	High-frequency %	Low-frequency %	Common-fruit %	High-frequency %
Price	84.44	84.88	83.8	80.63	79.59	77.99	85.95	87.33	80.51*
	(0.14)	(0.38)	(-0.47)	(1.52)	(0.03)	(-1.66)	(0.57)	(1.2)	(-2.38)
Freshness	90.56*	94.88	94.69*	91.67*	94.32	95.77*	87.98*	93.33*	92.80
	(-3.57)	(1.69)	(2.17)	(-4.46)	(1.04)	(3.82)	(-3.59)	(2.81)	(1.61)
Sweetness	69.90	70.93	74.72*	63.80	61.14*	66.51*	74.09	78.00	75.00
	(-1.63)	(-0.52)	(2.10)	(-0.29)	(-2.09)	(2.20)	(-1.4)	(1.62)	(-0.04)
Juiciness	66.58*	72.33	74.02*	64.98*	70.74	75.28*	71.81	69.56	76.69
	(-3.22)	(0.88)	(2.52)	(-5.7)	(0.87)	(5.3)	(-0.09)	(-1.26)	(1.76)
Nutritional value	70.28*	83.26*	86.17*	56.19*	65.83	74.48*	64.50*	76.67*	80.08*
	(-7.85)	(2.42)	(5.89)	(-9.54)	(1.14)	(9.16)	(-6.07)	(3.89)	(3.84)
Preferred origin	36.48*	39.07	47.91*	29.73*	30.02*	40.35*	37.77*	44.44	52.54*
	(-3.55)	(-1.06)	(4.53)	(-4.16)	(-2.31)	(6.54)	(-4.03)	(1.62)	(3.80)
Locally grown	53.57*	55.81	65.22*	42.85*	42.58*	53.35*	38.34*	48.00*	55.51*
	(-3.56)	(-1.23)	(4.67)	(-3.73)	(-2.46)	(6.22)	(-5.19)	(2.59)	(4.24)
Quality appearance	84.06	87.91	86.87	81.48	81.77	85.41*	67.02*	72.22	75.42*
	(-1.99)	(1.32)	(0.89)	(-1.98)	(-0.94)	(2.97)	(-2.87)	(1.55)	(2.19)

Ease of peeling	51.02	50.00	56.42*	51.86	49.78	53.19	56.70	53.33	58.47
	(-1.29)	(-1.32)	(2.45)	(0.07)	(-1.39)	(1.19)	(0.68)	(-1.37)	(0.78)
In season	72.83*	80.23	81.28*	70.44*	74.89	78.95*	76.20*	85.11*	87.29*
	(-4.17)	(1.48)	(2.97)	(-4.85)	(0.56)	(4.67)	(-5.04)	(3.29)	(3.11)
Trendy	15.82*	14.88*	23.74*	13.80	13.54	16.19*	23.88*	28.00	38.14*
	(-2.55)	(-2.22)	(4.51)	(-1.13)	(-0.94)	(2.06)	(-3.61)	(0.77)	(4.28)
Available in primary shopping locations	78.44*	80.93	84.64*	83.56*	86.68	88.12*	70.67	72.44	72.03
	(-2.66)	(-0.22)	(2.89)	(-3.53)	(0.94)	(2.92)	(-0.76)	(0.64)	(0.28)
Receipt diversity	30.74*	33.49	42.74*	28.89	27.51	31.66*	24.21*	28.22	41.95*
	(-3.84)	(-1.14)	(4.88)	(-0.70)	(-1.47)	(2.08)	(-4.11)	(0.48)	(5.38)

Note: *indicates statistically significant at 5% level