Full Length Research Paper

Post hoc segmentation combining conjoint analysis and a two-stage cluster analysis

Sérgio Dominique Ferreira Lopes

Management School, Polytechnic Institute of Cavado and Ave, Campus of IPCA, 4750 - 810 Barcelos, Portugal.
E-mail: sdominique@ipca.pt, sergio.dominique@gmail.com. Tel: +351 912 406 755.

Accepted 17 October, 2011

Tourism contributes to GDP and employment; therefore, several countries manage tourism resources using an R&D + I strategy (research and development of innovative methodologies). Thus, market segmentation adjusts tourism offer to consumers’ needs/preferences. Post hoc segmentation strategy has important advantages over a priori segmentation, allowing better adjustment of supply to tourists’ preferences. This paper illustrates the advantages of combined use of the conjoint analysis and cluster analysis when segmenting the tourist market - post hoc - according to consumers’ preferences. From the results, we identify six different clusters and the most important attributes are weather and price.

Key words: Market research, post hoc segmentation, conjoint analysis, cluster analysis, tourist preferences.

INTRODUCTION

Tourism has become one of the most important sectors of the global economy (Dominique-Ferreira et al., 2009). According to the World Tourism Organization (UNWTO, 2009), in 2008, the number of international tourist arrivals was 924 million, an increase of 2% (16 million) compared to 2007. The revenues increased to € 642 billion in 2008, that is, an increase of 1.7% compared to 2007. However, the results of 2009 show a decline because of an international financial crisis (Rial et al., 2010). Thus, every effort should be made to achieve a sustainable growth of the tourism sector.

In the concrete case of Portugal, according to the “Plano Estratégico Nacional do Turismo” that is, National Strategic Plan for Tourism (PENT, 2007), the Portuguese tourism industry accounts for about 13% of the national PIB and contributes to 12% of the employment in the country. In this context, special attention has to be paid in the moment of studying tourists’ preferences in order to configure products and services according to the consumers’ preferences.

THEORETICAL FRAMEWORK

As is currently known, a single product configuration is unlikely to meet everyone’s needs; thus, segmenting markets is necessary (Picón et al., 2004). Hence, business organizations need to cluster consumers into segments, not only to satisfy their needs, but also increase their level of satisfaction. Since the pioneering work of Smith (1956), the concept of market segmentation has evolved considerably.

Santesmases (1999) defines market segmentation as the process of dividing the market into homogeneous groups with the aim of implementing a marketing strategy for each group, in order to satisfy the needs of the tourists more effectively and achieve the commercial objectives of the company. Santesmases (1999) also said that segmenting markets presents several advantages, such as identifying markets that are not saturated, allowing to set priorities, facilitating the analysis of competition and thus obtaining information about the direct competitors and allowing companies to offer products/services that best suit the specific needs of each segment. Traditionally, there have been two ways of segmenting markets (Green et al., 1977; Wind, 1978):

1. *A priori* segmentation: In this kind of segmentation, the number of segments (groups) as well as their description is established before the study is carried out. The professionals choose the basis for the study, for example, the use of a product, needs, loyalty to a brand, at the...
beginning and then, the professionals assign consumers to the groups.

2. Post hoc segmentation: When the consumers' characteristics or their reactions to a new product are not known, it is better to realize a post hoc segmentation. This way, the number of groups, the number of subjects in each group and the description of the groups are known after the analysis is made. The resultant groups in post hoc segmentation are intra-groups that constitute consumers with more homogenous preferences than in a priori segmentation.

Therefore, the better the segmentation the better will the products be adjusted to the consumers' preferences. Nowadays, market researchers analyze products and services as a set of attributes, and therefore, products tend to be configured according to the needs and preferences of the consumers (Dominique-Ferreira et al., 2009). Hence, it is important to know the precise assessment of these attributes and how each attribute and the level of attribute contribute towards the purchase of a certain product or service.

In this context, conjoint analysis allows the analysis and modelling of consumer preferences; that is, it helps determine the importance given by consumers to each attribute in the set of a product. Conjoint analysis has its origins in the field of psychology and marketing. The work of Luce and Tuckey (1964) introduced multi-attribute models for analyzing and understanding consumer preferences and Wilkie and Pessemier (1973) were the first to discuss it. The conjoint algorithm can be represented as follows:

\[ U = f(U_{i1}, ..., U_{ik}) \]

Where \( U \): Total utility \( U_i \): Part-worth of the attribute \( i \); \( X_i \): Perceived level of the attribute \( i \)

Utility is the value that the subject assigns to a product/service through a combination of factors (attributes); thus, it holds maximum value for the consumer's choice in the set of choices (Ben-Akiva and Lerman, 1985). As a result, conjoint analysis is very useful for analyzing consumers' preferences (Wittink and Cattin, 1989; Wittink et al., 1994).

This way, we can carry out a post hoc segmentation that combines conjoint analysis and cluster analysis and obtain important benefits (Picón and Varela, 2000; Dominique-Ferreira et al., 2009). Specifically, the output of the conjoint analysis - attributes' importance and part-worths, that is, the structure of consumers' preferences - can be used as input for the cluster analysis to group consumers into segments on the basis of the similarity in their preferences (Picón, 2004; Picón and Varela, 2004; Varela et al., 2004; Dominique-Ferreira et al., 2009; Rial et al., 2010; Dominique-Ferreira, 2011; Dominique-Ferreira and Silva, 2011).

Thus, we can proceed to an integral segmentation, that is, a segmentation based on information on consumers' preferences and their socio-demographic profile. Each resultant group will be designated a different kind of destination (preference).

The conjoint analysis has been applied in various sectors of the economy because it helps obtain in depth knowledge of consumer preferences. The main fields of application are the tourism industry (Dellaert et al., 1995; Picón and Varela, 2000; Mazanec, 2002; Varela et al., 2004; Suh and McAvoy, 2005; Rial et al., 2010; Thyne et al., 2006), readers' preferences of journals (Braña et al., 2001), environmental preferences (Reig-Garcia and Coenders, 2002), waiting list for surgery (Rivera et al., 2004), human resources (Guerrero et al., 2003), teaching (Ramírez et al., 2005), and the system franchises (Ramírez, 2007).

So, the present work has two main goals. The first is to illustrate the importance and benefits of the combined application of the conjoint analysis and a two-stage cluster analysis. The second is to show how the conjoint algorithm aids in simulation analysis. This way, professionals and researchers can assign specific strategies to each cluster obtained.

METHODOLOGY

A total of 309 subjects were interviewed. They were aged between 18 and 35 years (mean = 22.96; SD = 5.637) and residents of Northern Portugal. Of the 309 tourists, 147 were men and 162 were women (mean = 43.29, SD = 14.01). The sample error is ±5.58% (p = q = 50) with a level of confidence of 95% (k = 2 sigma). Finally, all the subjects were asked to perform the task with the utmost concentration and seriousness.

**Selection of attributes and levels of attributes**

This work is part of a larger study, which had other goals, including analysing the satisfaction of Portuguese tourists and measuring Portugal's image and positioning as a tourist destination. However, the attributes and levels of attributes were selected on the basis of the existent bibliography (Goodrich, 1978; Muller, 1995; Baloğlu and McCleary, 1999; Gallarza et al., 2002; Picón and Varela, 2000; Varela et al., 2004; Rodríguez and Molina, 2007; Rial et al., 2008; Dominique-Ferreira et al., 2009; Dominique-Ferreira et al., 2010).

Finally, the selected attributes were weather, cultural offer, kind of destination, nightlife, price and time of permanency.

**Stimuli**

To perform the conjoint analysis, we selected six attributes of a tourism destination, with different levels for each attribute \( (3 \times 4 \times 2 \times 2 \times 4 \times 3) \). From the 576 possible combinations, we selected 16 using an orthogonal fractional factorial design; these 16 were eventually used in the data collection (with an orthoplan procedure of the SPSS software). 16 cards were created (Figure 1), each one representing one of the 16 combinations of the levels of attributes. For the cluster analysis, a two-stage clustering was used. We started with a hierarchical method (Euclidean distance) using Ward's (1963) linkage method and proceeded with the use of iterative k-means clustering, which is considered more reliable than the conventional single-stage procedures.
DESTINATION 4
SUN
HIGH CULTURAL OFFER
BEACH
LOW NIGHT FUN OFFER
MORE THAN € 1000
1 WEEK

Figure 1. Example of the destinations used.

As for the part-worths of the levels of attributes, the results show that the preferred level of weather is sun with a part-worth of 1.61; contrarily, the levels that show negative part-worths are tepid or humid (−0.65) and cold or snow (−0.97). For the attribute cultural offer, the preferred level is high offer with a part-worth of 0.62 and the opposite (low offer) has a part-worth of −0.62. For the attribute kind of destination, the preferred level is beach, with a part-worth of 0.57. The other two levels considered, namely, nature-mountain and city show negative part-worths, that is, −0.18 and −0.39 respectively. Concerning the attribute nightlife, the preferred level is high offer, which has a part-worth of 0.41 and the opposite level (low offer) has a negative value of −0.41. With regard to the attribute price, the favourite levels are up to € 300, which has a part-worth of 0.95 and between € 300 and 600 with a part-worth of 0.48. The levels between € 600 and 1000 and more than € 1000 have negative part-worths, that is, −0.36 and −1.07 respectively. Finally, regarding the attribute time of permanency, there are two preferred levels, namely, 2 weeks (0.57) and 1 week (0.31); however, the level 2/3 days or weekend has a negative part-worth of −0.88. These results can be observed in Figure 3.

Ideal destination

However, the conjoint analysis also allows us to identify the ideal product/service, which is estimated as the sum of the maximum part-worth of each attribute and the value of the constant. Therefore, in the present case, the ideal destination would be constituted as in Figure 4. Thus, if we sum the highest part-worths of each attribute and the value of the constant, we get:

\[ 1.61 + 0.62 + 0.57 + 0.41 + 0.95 + 0.57 + 8.17 = 12.9 \]

Thus, the global importance of the ideal destination would be 12.9.

Market segmentation

As mentioned earlier, the second goal of this work is to illustrate the advantages of combining the conjoint analysis and cluster analysis. On the basis of the part-worths of all the subjects of the sample (output of the conjoint analysis), a two-stage cluster method was applied in order to segment the markets according to the similarity in the structure of consumers (tourists). Subsequently, the results show the existence of six well-defined clusters (Figure 5):

Group 1: Seeks cheap destinations (6.8%)

This group of tourists gives considerable importance to the price for travel. The importance they have given to

Procedure

The subjects were asked to analyze every card in the first minute and then sort these cards according to their preferences. This procedure is called full profile and uses simulated stimuli and the sorted cards (sequence). Finally, it should be noted that the method of estimation of partial utilities was ordinary least squares (OLS), which is the most frequently used method (Wittink and Cattin, 1989).

RESULTS

Modelling consumers’ preferences

Attributes’ importance and part-worths values

The model fit is very high, so we can conclude that the validity of the results is high (Pearson’s r = 1.000; Kendall’s tau = 0.983). By applying the conjoint analysis to the entire sample, we find that the most important attribute is weather, with an importance of 28.414%, followed by price with 22.214%. In the third place, we find time of permanency with an importance of 15.986%, and in the fourth place, we find cultural offer with an importance of 13.626%. Next, we find kind of destination with an importance of 10.645%, and finally, nightlife with an importance of 9.116% (Figure 2).

Figure 2. Importance of attributes.

https://www.example.com/figure2.png

Figure 3. A chart showing the importance of various attributes.

https://www.example.com/figure3.png

Figure 4. The ideal destination.

https://www.example.com/figure4.png

Figure 5. Six well-defined clusters.

https://www.example.com/figure5.png

As mentioned earlier, the second goal of this work is to illustrate the advantages of combining the conjoint analysis and cluster analysis. On the basis of the part-worths of all the subjects of the sample (output of the conjoint analysis), a two-stage cluster method was applied in order to segment the markets according to the similarity in the structure of consumers (tourists). Subsequently, the results show the existence of six well-defined clusters (Figure 5):

Group 1: Seeks cheap destinations (6.8%)

This group of tourists gives considerable importance to the price for travel. The importance they have given to
Figure 3. Part-worths of all the attributes’ levels.

IDEAL DESTINATION
SUN (1.61)
HIGH CULTURAL OFFER (0.62)
BEACH (0.57)
PASSIVE NIGHTLIFE (0.41)
MORE THAN €1000 (0.95)
1 WEEK (0.57)

Figure 4. Ideal destination.

the attribute price is 37.68%.

Group 2: Seekers of long-time vacations (13.7%)
This group gives importance to the time they spend at a destination (importance = 28.28) and, understandably, the price they pay (maybe per day) is also important (importance = 20.53), being the preferred level up to €300 (u = .800). These tourists also give particular importance to the weather that characterizes the destination (importance = 21.31).

Group 3: Seekers of sun and culture (26.4%)
These tourists give special importance to weather (importance = 25.49), that is, destinations characterized by sunny days and the culture it offers (importance = 19.788), specifically, a high cultural offer (u = 0.908)

Group 4: Seekers of sun and cheap destinations (20.2%)
These tourists give greater importance to weather
Figure 5. Attributes' importance by clusters.

(Importance = 38.495), specifically, sun (u = 2.139). They simultaneously look for tourism products with affordable prices, specifically, a part-worth of 1.130 (up to €300) and 0.514 (between €300 and 600).

Group 5: Seekers of sun, culture and cheap destinations (21.6%)

Although group may seem similar to Group 4,
because it is a group that gives high importance to weather (importance = 28.457) and price (importance = 26.178; \( u = \) up to €300 = 1.056), Group 5 gives special importance to high cultural offer (importance = 13.110) compared to Group 4 (importance = 7.264, which is almost double the importance of Group 5).

**Group 6: Seekers of sun, beach and cheap destinations (11.3%)**

Finally, Group 6 may also be similar to Groups 4 and 5 because they give considerable importance to weather (importance = 28.457) and price (importance = 26.178), but they give more importance to the attribute kind of destination (importance = 16.249), the preferred level being the beach (\( u = 0.780 \)).

The subjects have been grouped into clusters, which show a significant difference in the choice of part-worh. Through the F-values (Table 1), we can see that the level sun distinguishes the clusters the most (\( f = 141.782 \)), followed by more than €1000 (\( f = 121.056 \)) and high cultural offer (\( f = 81.717 \)).

These results validate the advantages of the current application of the cluster analysis because what differentiates the subjects of the other five clusters is their preferences and not their socio-demographic characteristics. This is one of the reasons why it is advisable to carry out a post hoc segmentation according to the consumers’ preferences (Picón and Varela, 2000; Dominique-Ferreira et al., 2009).

Therefore, the authors cross-tabulated the results present in the questionnaire on the basis of groups, and used the analysis of residuals (Haberman, 1973) to identify responses that differed significantly from those expected to have no effect of the group (Tables 2 and 3). This information is of interest but analysis of residuals allows us to take the analysis one step further and consider the relative strength of this preference (Haberman, 1973).

From Tables 2 and 3, we can see that the clusters present a socio-demographic with a very similar composition. Therefore, the preferences of young tourists of Northern Portugal cannot be explained on the basis of only socio-demographic characteristics because there is no concrete socio-demographic profile. Subjects have been grouped into clusters not because they share socio-demographic characteristics, but because they share tourism preferences.

**Conclusions**

Nowadays, tourism is an important sector for growth in many countries (for example, Portugal, Spain, France, the USA). The share of tourism in Portugal's Gross Domestic Product (about 13%) and employment development are some of the examples indicating the importance of tourism.

In the current competitive environment, managing the tourism resources of a country under a marketing approach has become a key strategy to success. This way, many researchers refer to the importance of investing on the basis of R + D + i strategy (research, development

### Table 1. ANOVA.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cluster</th>
<th>Error</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quadratic mean</td>
<td>df</td>
<td>Quadratic mean</td>
<td>df</td>
</tr>
<tr>
<td>Sun</td>
<td>200.620</td>
<td>5</td>
<td>1.415</td>
<td>303</td>
</tr>
<tr>
<td>Cold-snow</td>
<td>114.493</td>
<td>5</td>
<td>2.449</td>
<td>303</td>
</tr>
<tr>
<td>Tepid</td>
<td>45.000</td>
<td>5</td>
<td>2.586</td>
<td>303</td>
</tr>
<tr>
<td>High cultural offer</td>
<td>57.841</td>
<td>5</td>
<td>0.708</td>
<td>303</td>
</tr>
<tr>
<td>Low cultural offer</td>
<td>57.841</td>
<td>5</td>
<td>0.708</td>
<td>303</td>
</tr>
<tr>
<td>Beach</td>
<td>84.496</td>
<td>5</td>
<td>1.602</td>
<td>303</td>
</tr>
<tr>
<td>City</td>
<td>32.774</td>
<td>5</td>
<td>1.552</td>
<td>303</td>
</tr>
<tr>
<td>Nature or mountain</td>
<td>18.260</td>
<td>5</td>
<td>1.815</td>
<td>303</td>
</tr>
<tr>
<td>High nightlife offer</td>
<td>3.282</td>
<td>5</td>
<td>1.057</td>
<td>303</td>
</tr>
<tr>
<td>Low nightlife offer</td>
<td>3.282</td>
<td>5</td>
<td>1.057</td>
<td>303</td>
</tr>
<tr>
<td>Up to €300</td>
<td>193.224</td>
<td>5</td>
<td>2.481</td>
<td>303</td>
</tr>
<tr>
<td>€300 - €600</td>
<td>52.032</td>
<td>5</td>
<td>2.131</td>
<td>303</td>
</tr>
<tr>
<td>€600 - €1000</td>
<td>23.936</td>
<td>5</td>
<td>2.099</td>
<td>303</td>
</tr>
<tr>
<td>More than €1000</td>
<td>264.830</td>
<td>5</td>
<td>2.188</td>
<td>303</td>
</tr>
<tr>
<td>2 to 3 days (Weekend)</td>
<td>59.057</td>
<td>5</td>
<td>1.567</td>
<td>303</td>
</tr>
<tr>
<td>1 Week</td>
<td>6.023</td>
<td>5</td>
<td>1.571</td>
<td>303</td>
</tr>
<tr>
<td>2 Weeks</td>
<td>46.201</td>
<td>5</td>
<td>1.549</td>
<td>303</td>
</tr>
</tbody>
</table>
and innovation), that is, market research and development of innovative methodologies, which have become added value.

In this context, the destination managers should have in-depth knowledge of the preferences of the tourists, as well as their needs, habits and lifestyles. As such, focusing on methodologies that allow optimizing of the management of tourism resources has become important.

In this context, countries such as the USA, Spain and France are the most competitive countries in the world (in terms of tourism revenue and visits), and they have a competitive advantage when it comes to analyzing the tourists' preferences: The use of conjoint analysis to study the real structure of tourists' preferences (Varela et al., 2003; Dominique-Ferreira et al., 2009). Consequently, these countries can design new strategies that better suit the tourists' preferences.

Thus, from a methodological point of view, the present work illustrates the considerable potential of the combined use of the conjoint analysis and cluster analysis in tourism management. The results show that tourists' preferences can vary greatly between some clusters obtained, and, at the same time some, clusters that seem to be similar, in reality, differ from the preference in just one or two attributes. However, the real importance of those preferences can mean not to choose a destination over another.

We could thus identify up to six different segments (in a seemingly homogeneous population such as the young residents of Northern Portugal), with each segment having a specific structure of preferences.

The characterization of each group probably has important implications for destination managers in the case of the post hoc segmentation than in the case of a priori segmentation. Thus, on precisely knowing the preferences of tourists, segmenting markets post hoc allows professionals to identify heterogeneous groups (referring

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cluster</th>
<th>Age</th>
<th>Sex</th>
<th>Occupation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>18-24</td>
<td>5.975</td>
<td>5.334</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>More than 30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cluster</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td>-0.52</td>
<td>-1.50</td>
<td>-0.37</td>
<td>0.84</td>
<td>0.57</td>
<td>0.75</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td>1.22</td>
<td>1.33</td>
<td>0.46</td>
<td>-1.26</td>
<td>-0.69</td>
<td>-0.56</td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
<td>-0.58</td>
<td>0.59</td>
<td>0.01</td>
<td>0.20</td>
<td>-0.04</td>
<td>-0.40</td>
</tr>
<tr>
<td>Students</td>
<td></td>
<td>0.50</td>
<td>-1.32</td>
<td>-0.95</td>
<td>0.52</td>
<td>1.02</td>
<td>0.36</td>
</tr>
<tr>
<td>Workers</td>
<td></td>
<td>0.13</td>
<td>1.63</td>
<td>0.40</td>
<td>-0.89</td>
<td>-0.07</td>
<td>-1.22</td>
</tr>
</tbody>
</table>

Table 2. Pearson's chi-square values.

Table 3. Residuals (standardized and adjusted) of the cross-tabulation of the questionnaire results for each of the six groups.
to consumers’ preferences) and to develop specific products for each group (Picón et al., 2004).

LIMITATIONS AND FUTURE RESEARCH

The main limitation of the present work is its small scale, it has a sampling error of 5.58%. With 383 tourists (74 more), the sample error should be less than 5%, a desirable number for studies in the social sciences. This way, the results could have been interpreted with greater certainty (about the preferences of young tourists from Northern Portugal). However, the main goal of the present work was only to illustrate the potential of the conjoint analysis. Future research should include tourists aged up to 65 years and from other countries.

REFERENCES


