

**ANALYZING ON-LINE CONSUMER ENQUIRY PATTERNS IN
THAI-TOURISM INDUSTRY: INSIGHTS FROM A K-MEANS CLUSTERING
SOLUTION**

Pongsak Hoontrakul
Sasin Graduate Institute of Business Administration,
Sasa Patasala Building,
Soi Chula 12,
Phyathai Road,
Bangkok 10330,
Thailand
Tel: 0-2218-3853
e-mail: pongsak@hoontrakul.com

And

Sunil Sahadev¹
School of Management
University of Sheffield
9 Mappin Street, Sheffield
S14DT
e- Mail: s.sahadev@shef.ac.uk
Phone: + 01142223434
UK

¹ Corresponding author

ANALYZING ON-LINE CONSUMER ENQUIRY PATTERNS IN THAI-TOURISM INDUSTRY: INSIGHTS FROM A K-MEANS CLUSTERING SOLUTION

ABSTRACT

As tourism booking is increasingly becoming an on-line activity, the patterns of on-line booking have become an important area of focus for on-line travel agents and destination marketing organizations. The study attempts to segment on-line tourist enquiries received for four popular destinations in Thailand. While segmentation research is a well established stream in the domain of tourism studies, the paper extends the existing domain by applying clustering algorithms to enquiries.

Key words: Thailand, Tourism enquiries, Segmentation, Clustering.

ANALYZING ON-LINE CONSUMER ENQUIRY PATTERNS IN THAI-TOURISM INDUSTRY: INSIGHTS FROM A K-MEANS CLUSTERING SOLUTION

Introduction:

The on-line travel intermediary sector has witnessed a surge in patronage over the past three years. It is estimated that about 55% of all American Internet users booked their travel on line at least once since 2001². The migration towards on-line travel intermediation continues unabated and large online intermediaries like expedia.com, Travelocity.com etc. are expanding fast to cater to the emerging demand. However, being a highly competitive industry, there is a mushrooming of small intermediaries concentrating on niche segments and offering products at a cheaper rate. The study attempts to explore the underlying patterns in tourist enquiries with regard to Thailand. Thailand is fast emerging as an important tourism destination. In 2005, Thailand attracted about 12 million tourists as per the estimates of the Tourism Authority of Thailand. Tourism is also a very important industry for Thailand. It is estimated that Tourism generated about \$11 billion direct revenue in 2005 in Thailand. The main objective of the study is to segment the enquiries to explore meaningful typologies of the tourists intending to visit Thailand. Data on enquiries across a twenty months period from November 2003 to September 2005 were used for the analysis. A total of 1385 enquiries were processed to derive the clustering solution. Data mining techniques like clustering, decision trees etc. have been widely used for successfully segmenting and targeting customers across various industries. Data mining provides an effective approach to discover and understand patterns in customer behavior thereby helping the decision maker to better group customers.

Segmentation studies in On-line marketing and Tourism literature:

The appropriate literature for the research domain is scattered across two main research streams: (i) research on segmentation of tourists (ii) research on segmentation based on on-line customer behavior. Studies on segmentation of tourists have been carried out

² quoted in <http://www.etourismnewsletter.com/ecommerce.htm> accessed on 12/07/'06

extensively in the past [For a general review refer Dolnicar(2004)]. These studies mostly seek to identify segments of tourists that have similar preferences and expectations. Several of these studies use psychographic bases (eg. Kastenholz, Davis and Paul, 1999; Sarigolu and Huang, 2005; Keng and Cheng, 1999) while few other studies deploy demographic characteristics and related attributes (eg. Laesser and Crouch, 2006). It is difficult to find a study that use enquiry based attributes for segmentation. In this study we concentrate on on-line enquiries to find out segments that could group tourists into homogenous groups based on their background profile and need characteristics.

The on-line marketing strategy literature has off late witnessed a number of studies which consider segmentation related issues (eg. Jainszeewski, 1998; Chen and Cooper 2001; Moe and Fader, 2003; Moe and Fader, 2004 etc.). Most of these studies attempt or propose to segment customers in terms of their search behavior. Further, most of these studies look at the search behavior in the context of B2C sites that sell consumer goods like Grocery. Unlike these low value items, at the time of purchase, Tourism is just an information good since the customer is only provided with information about the product and also, tourism is typically a high value item thus rendering click stream based segmentation approaches quite inadequate in describing customer behavior.

Methodology

This methodology broadly followed the approach advocated by Peltier and Schribrowsky (1997), in which a segmentation approach based on both the needs underlying a purchase and descriptive buyer data is implemented. This would help to better understand the customer in terms of certain important attributes that were considered to be critical in influencing the enquiry-to-conversion decision.

The present study attempts to find patterns of tourist enquiries based on the background profile as well as their spelt out requirements. The scope of the study is limited to the tourists enquiries for hotel rooms in various locations in Thailand. It was of course evident at the outset that the enquiry patterns varied from one destination to another. Thailand, offers tourists a variety of choices in terms of tourist attractions. The maximum

number of tourists however visits a few destinations like Bangkok, Phuket Pattaya etc. It is also expected that the nature and characteristics of the tourists who visit each of these destinations also would vary. Therefore in order to understand the background of tourists who visit Thailand, it was necessary to conduct a destination specific analysis of the tourist enquiries. Such an analysis would help us to understand the various categories of tourists in terms of their background characteristics. The study thus analysed the tourist enquiries to four major destinations in Thailand viz. (i) Bangkok, (ii) Phuket (iii) Krabi and (iv) Samui. Of the four destinations, Bangkok and Phuket was already quite well known to the outside world. Also, these two destinations offered more choice since they had more hotels and these hotels were also well represented in travel websites. Krabi and Samui were destinations which were relatively underdeveloped in terms of tourist arrivals.

The segmentation process progressed through three phases. The first phase was a qualitative phase where, five experts who had good knowledge of the tourism market and the general pattern of tourist enquiries were interviewed extensively. Insights from these interviews guided us in selecting the variables to be used for the cluster model. These variables were selected and categorized so that they meaningfully reflected the factors that according to the expert opinion differentiated the customer enquiries.

The transaction processing system that captured enquiries recorded data about the enquiries across several fields. A large number of variables could have been considered. However several of these variables were summarily rejected since they evidently had no association with enquiry-to-conversion process. For instance, variables like the currency used, the date of enquiry, the date of check in etc. were considered to be least influential in determining the enquiry-to-sale conversion behavior. After considerable discussion and analysis with experts familiar with the tourism scene in Thailand, six variables were selected for clustering the enquiries. Further, since the data was captured through different scales for each of these variables it was also decided to use a standard categorization for each of these variables for further analysis. The split of variables into its constituent categories was essential to make the clustering solution meaningful. The variables and the constituent categories are provided in table.1

||Take in Table.1 Variables included in the analysis||

In the second phase, the data were used to extract clusters of customer segments for each of the destinations. In each of the clusters that emerged, it was decided to look at the total number of enquiries as well as the total volume of enquiries. In order to obtain the initial solution, K-means clustering was carried out using Clementine for data from the four destinations separately. The results for each of the destinations are explained in the following sections. The K-means method is a clustering method used to group records based on similarity of values for a set of input fields. The basic idea is to try to discover k clusters, such that the records within each cluster are similar to each other and distinct from records in other clusters. K-means cluster algorithm involves an iterative process wherein each record (here an enquiry) is assigned to the cluster whose center is the closest. Closeness is measured by the squared Euclidian distance

$$D_{ij} = \| X_i - C_j \|^2 = \sum_{q=1}^Q (x_{qi} - c_{qj})^2$$

Where X_i is the vector of encoded input fields for record i , C_j is the cluster center vector for cluster j , Q is the number of encoded input fields, x_{qi} is the value of the q th encoded input field for the i th record, and c_{qj} is the value of the q th encoded input field for the j th record. The process involves assigning records to clusters with the nearest centers and then when all the records are grouped, the cluster centers are updated. This process is repeated till we reach the maximum distance from cluster centers indicated by the proximity distance. Both categorical as well as 'metric' variables were included in the analysis. The number of clusters were not specified initially and the program generated the best cluster solutions that maximized the inter centroid distances. The final solution considered for subsequent analysis was therefore the optimal partitioning generated as a consequence of iterations.

The third phase comprised of interpretation of the clusters. In this phase the services of the experts were again utilized. Based on the segment profiles, the clusters were labeled and interpreted. The interpretation of clusters that were extracted for the four destinations are explained below

Cluster Profiles for Bangkok Enquiries

Bangkok being the capital of Thailand, attracted the maximum number of tourists. Bangkok also attracted business travelers on as well as transit passengers on a large scale. A total of 593 enquiries for Bangkok were considered for the analysis. Initially the entire group was divided into 5 clusters. Based on the analysis of the proximity values the five cluster solution was found to be stable. The details of the cluster solution for Bangkok is provided in table.2

|| Take in Table.2 Cluster output for Bangkok||

Based on subsequent discussions with experts and booking staff, it was decided to label the clusters into: (i) New on-transit customers (ii) New Holiday makers (iv) Loyal business customers (iv) Economy seeking holiday makers and (v) Loyal on-transit customers. These nomenclatures have been based on the characteristics as defined by the clustering solution as well as the observations by the experts. A fact which is very evident in terms of Bangkok enquiries is the relative preference for low rates across all types of customers, short duration stay and the short advance booking days. This may be because a majority of the customers who visit Bangkok are either in-transit customers or business travelers. Across all the segments, more than 60% of the enquiries are for rooms which are less than 50\$ a day. Almost 90% of all enquiries are for duration of stay of less than 7 days and further, across all clusters, about 30% of all enquiries have an advance booking period of less than 7 days.

New on-transit customers arrive during the low season and are probably on their way to some other country. They are not necessarily tourists since they didn't arrive during the season. They are from different continents and are economy seeking. The only feature that distinguishes them from the members on the last cluster (loyal in-transit customers)

is that they are approaching this on-line intermediary for the first time. Hence they are either visiting Thailand for the first time or are utilizing the services of this intermediary for the first time. About 43% of all enquiries for Bangkok belong to these two categories. While these two categories together account for about half of all the enquiries, the value of these enquiries together amounts to about 32%.

Another prominent category of customers are holiday makers. They are represented by clusters 2 and 3. Together they account for another 40 % of all the enquiries for Bangkok. These customers visit Bangkok during the peak season. They fall under two clusters as per this cluster solution. They typically book in advance with about 60% from this category booking with more than a fortnight in advance although another 30% have a very short advance booking period. A majority of them are new customers and only about 35% are loyal customers. This category is further subdivided into two equal groups. One category comprise of moderate spenders in which almost 40% prefer costly hotels and the other category comprise of high economy seekers in which just about 20% prefer costly rooms. However, the 'holiday seekers' together accounted for about 46% of the total value of enquiries.

About 15% of all enquiries are from loyal business-customers. These types of enquiries occur throughout the year spread evenly across high season as well as low season. For this category, the number of days of advance booking is particularly low with about 60% booking with less than a fortnight in advance. All these enquiries are from loyal customers in that these customers have already booked through this particular on-line intermediary before. They originate evenly from all continents and are not particularly economy seeking. Only less than 8% of these customer demand a room with less than 25\$ a night. They are lucrative to serve since they represent about 20% of the value of enquiries.

To summarize, the enquiries for Bangkok are mostly differentiated in terms of the season, loyalty and to a certain extent the cost of rooms asked for. Other attributes like number of days of advance booking, nationality, length of stay etc. doesn't have much impact in

segmenting enquiries. Most of the customer segments show remarkable similarities across these variables.

Cluster Profiles for Phuket Enquiries

Phuket is a very popular tourist destination in south Thailand on the Andaman sea. According to the Tourism Authority of Thailand's estimates (TAT 2003), Phuket, attracted about 3.5 million tourists in 2003. The numbers have seen a marginal increase since then. In the case of Phuket about 416 enquiries were processed to extract the segmentation model. The analysis of the clusters shows a more clear segmentation. Table.3 gives details about the cluster profiles for Phuket

||Take in Table.3 Cluster output for Phuket ||

The five segments extracted from the K-means cluster analyses were labeled based on the advice of the experts. The following nomenclature was adopted: (i) New economy seeking weekenders (ii) High season rich long stayers (iii) Low season long stayers (iv) High season short stay holiday makers and (v) high season loyal holiday makers. A prominent set of customers are those who visit Phuket for weekends or just short duration visits during the low season. About 24% of the total enquiries are from this segment. They predominantly (about 90%) stay only for less than a week and are new customers. However in terms of number of days of advance booking, they seem to be evenly divided across the four possibilities. Most of these enquiries are from new customers. In terms of price preferences, they are more prone to look for cheap rooms with more than 65% looking for rooms with less than 50\$ per day. While this category represents about 24% of all the enquiries, they hardly generate 20% of the value of enquiries.

Another long season category comprises of the low season long stayers. They are of course just a minority constituting about 12% of the total enquiries. About 60% of them stay for more than one week. They typically book well in advance and a majority of them are loyal customers. They are not characterized by preference for any particular room prices. They also substantially add to the value of the enquiries-contributing almost 17% of the total value.

Phuket has another category of long stayers who visit Phuket during the high season. They are quite rich, with about 64% of them preferring rooms with a rent of more than 50\$ per day. This category constitutes about 20% of all enquiries and 26% of the total value of enquiries. They mostly hail from the rich countries of UK, Australia, New Zealand, Western Europe and North America.

About 40% of the enquiries are from high season short stayers. Within this larger group about 57% are new customers to this on-line intermediary and about 43% are loyal customers. They don't seem to have any significant characteristic in terms of number of advance days or price preferences. They together accounted for about 36% of the value of enquiries.

The enquiry profile for Phuket is therefore quite different from that of Bangkok. In terms of the Percentage of people who opt for long stay and the also the percentage of people who visit during the high and low season. Typically the average length of stay for Phuket is longer than that of Bangkok. In Bangkok for instance less than five percent of enquiries across all segments had opted for long stays while the case for Phuket is much different. Also, a significant percentage of people across all segments originated from South and East Asia as far as Bangkok is concerned, while in the case of Phuket it is not so.

Cluster Output for Samui Enquiries

Samui is an Island in the south west of Thailand. Samui is fast emerging as a very important tourist destination with its natural beauty and well developed infrastructure. According to the Tourism Authority of Thailand, Samui attracted about 1 million tourists in 2003. For this study, a total of 500 enquiries received for Samui were processed and five clusters were extracted. Based on the proximity values the five cluster solution was found to be stable. fig.5. It should be born in mind that the typical tourist season in Samui is different from the tourist season in other parts of Thailand. Thus, high season and low season has different meaning for Samui. Nevertheless, in our analysis we considered this change in season duration before keying in the data and therefore the high season as indicated in the clustering solution is in fact the high season in Samui and not that of entire Thailand. As is evident from the clustering solution, the enquiries for Koh Samui (or Samui for short) can be classified into five clusters. Just about 35% of all enquiries are received during the high season in Samui. This may be because of the fact that tourists are still not aware of the season difference with regard to Samui. The customer segments for Samui have been labeled as: (i) New low season short stayers (ii) High seasoners (iii) Low season loyal customers (iv) Low season economy seeking long stayers and (v) Low season affluent long stayers.

Of the high season cluster comprising of about 28% of all the enquiries and 31% of the total value, there is a distinct preference for short stays of less than one week. They typically book much in advance with more than 70% booking about a fortnight before the trip. There is also a distinct preference for good rooms with only less than 10% preferring rooms with a rate less than 25\$. About 40% of the enquirers in fact prefer rooms with more than 75\$ per day.

||Take in Table.4 Cluster output for Koh Samui||

The low season enquiries can be subdivided into about four categories. The first division is in terms of the duration of stay. There is a distinct division in terms of short stayers and long stayers. The first and third clusters consist of customers who are short

stayers. They hardly stay for more than a week. The together constitute about 45% of all enquiries as well as about 30% of the value of enquiries. The low season short stayers can be subdivided into two distinct categories in terms of affluence. The first cluster consists of less affluent customers with a large majority of the enquiries (78%) with preference for rooms below 75\$ and below. This category comprise of more than 28% of the total enquiries received by Koh Samui. The second category comprise of relatively affluent short stayers with more than 86% preferring rooms with a price tag of 50\$ or more per day with about 40% preferring rooms with more than 75\$ per day. The second segment of short stayers seems to be a loyal bunch having had past experience with this particular on-line travel intermediary.

Another segment of customers who enquire about (and probably visit) Samui are long stayers. These long stayers visit during the low season in Samui. Together they constitute about 27% of all enquiries. More than 85% of them stay for more than one week. Being long stayers they book sufficiently in advance with more than 50% of them booking almost one month in advance. Among the long stayers about a third are marginally affluent with about 30% from this segment preferring rooms that costs more than 75\$.

The segmentation profile of Samui presents a unique picture in terms of the split between high season and low season customers as well as a clear categorization between short stayers and long stayers.

Cluster Output for Krabi Enquiries

Krabi is a very attractive destination fast emerging as an important tourism attraction in Thailand. Krabi is estimated to have attracted about 1 million tourists in 2003 as per the estimates of TAT. For the present study, 358 enquiries that were received for Krabi were processed. Based on the proximity values, a four cluster solution was found to be stable. Upon careful analysis of the cluster profiles as well as the discussion with experts, the following cluster labels were developed: (i) New High Season Long Stayers (ii) High Season Economy seeking Short Stayers (iii) Low Seasoners and (iv) High Season Affluent Short Holiday Makers. Of the four clusters, three consist of customer enquiries that are either entirely or predominantly for the high season. One segment of customer enquiries entirely pertains to low seasoners.

The low season enquiries comprise of about 22% of the total enquiries and 14% of the value. The low seasoners are also short stayers with all of them staying for not more than a week. A significant portion of them are from south east Asia (about 30%) and hence can be presumed as week end holiday makers. Low seasoners also don't book too much in advance. About 70% of them book hardly one month in advance. They also prefer moderately priced rooms with about 60% of them preferring rooms in the 25 to 75\$ range.

The High seasoners can be subdivided into basically two broad categories – short stayers and long stayers. The long stayers who prefer to stay for about a fortnight comprise of about 14% of all enquiries and about 19% of the value of enquiries. The long stayers book much in advance with about 80% of them booking more than a

month in advance. The long stayers prefer rooms that are moderately priced with about 80% of them preferring rooms in the 25 to 100\$ per day.

The high season short stayers comprise of two categories, the economy seekers and the affluent. The economy seekers in the high season short stayers comprise of about 22% of all the enquiries and about 14% of the value. They book sufficiently in advance with about 65% of them booking more than 30 days in advance. Also true to their label, almost 60% of them prefer rooms with less than 75\$ per night. This is in fact the largest segment of enquiries to Krabi. In fact about 40% of the enquiries are in this segment. The high season, affluent short holiday makers also account for a significant proportion of the enquiries. Together they account for about 24% of all enquiries and 28% of the total value of enquiries. The affluent short stayers too book much in advance and are mostly new customers. Clusters 1, 2 and 4 pertain to enquiries for high season while cluster 3 comprises of enquiries for low season. Cluster.1 comprises of relatively long stay enquiries with the length of stay being more than one week while the clusters 2,3,and 4 comprises of enquiries with less than one week duration. Apart for cluster.4, the other three clusters have predominantly had room rate between 25 to 75\$ per night. Also predominantly all the clusters consist of enquiries from new customers. In terms of volume of enquiries, cluster.2 has the highest value and the firm has done well to convert the entire group of enquiries. However cluster.4 which constitutes the second highest value in terms of the total volume of enquiry is a matter of concern. The main distinction between enquiries in cluster.4 and other clusters is higher incidence of

enquiries for rooms with a rent greater than 75\$ per day. Further the incidence of new customers is more in cluster. 4 than in other clusters.

||Take in Table. 5 Cluster Output for Krabi||

Discussion

The analysis point towards certain unique patterns in terms of the enquiries received across the four destinations in Thailand. It is to be noted that the enquiries received in these destinations do not fall in the same pattern and the destinations exhibit differences if all the enquiries for a destination are pooled together and analysed. The paper puts forward a particular methodology for analyzing customer traffic to destinations. The analysis framework is useful in gaining practical insights about the segmentation profile of customer enquiries across destinations. The results also go to endorse the view that even within a country, there could be significant variations across destinations. The results could help practitioners to realize the variety in tourist enquiries and provides a useful format for segmenting their markets. The research also extends the existing knowledge in the tourism segmentation literature. Future studies could look at the possibility of validating these results as well as adding different variables to increase the usefulness of the segmentation scheme.

References

Albers, S and Manfred, K (2000) "Approaches for the Segmentation of Customers, How suitable are the traditional concepts", *Schmalenbach Business Review*, Vol.52 (3), pp. 312-313.

Chen, H.M and Cooper, M. D (2001), "Using Clustering Techniques to Detect Usage Patterns in a Web-Based Information System", *Journal of the American Society for Information Science and Technology*, Vol.52 (11), pp.888-904.

Dolnicar, S (2004) "Beyond Commonsense Segmentation: A Systematics of Segmentation Approaches in Tourism", *Journal of Travel Research*, Vol.42 (February), pp.244-250.

Jainszeewski, C (1998), "The Influence of Display Characteristics on Visual Exploratory Search Behavior", *Journal of Customer Research*, Vol.25 (3), pp.290-301.

Kastenholz, E., Davis, D and Paul, G (1999) "Segmenting Tourism in Rural Areas: The Case of North and Central Portugal", *Journal of Travel Research*, Vol.37 (May), pp.353-363.

Keng, K.A. and Cheng, J.L.L. (1999) "Determining Tourist Role Typologies: An exploratory Study of Singapore Vacationers", *Journal of Travel Research*, Vol.37, pp.382-390

Laesser, C and Crouch, G., I, "Segmenting Markets by Travel Expenditure Patterns: The case of International Visitors to Australia", *Journal of Travel Research*, Vol.44 (May), pp.397-406.

Moe, W and Fader, P. (2003), *Capturing Evolving Visit Behavior in Clickstream Data*, *Journal of Interactive Marketing*

Moe, W and Fader, P., S (2004), “Dynamic Conversion Behavior at E-Commerce Sites”, *Management Science*, Vol.50 (3), pp.326-335.

Palacio, V. and Mc Cool, S. F. (1997) “Identifying Ecotourists in Belize through Benefit Segmentation: A Preliminary Analysis, *Journal of Sustainable Tourism*, Vol.5(3), pp.234-243.

Peltier, J., and Schribrowsky, J (1997), “The use of need-based segmentation for developing segment-specific direct marketing strategies”, *Journal of Direct Marketing*, Vol.11 (4). Pp.53-62.

Sarigollu, E and Huang, R., (2005) “Benefit Segmentation of Visitors to Latin America”, *Journal of Travel Research*, Vol.43 (February), pp.277-293.

Smith, S.L. J and Smale, B (1980),“Classification of Visitors to Agreements for Recreation and Conservation Sites, National Parks and Related Sites. *Contact: Journal of Urban and Environmental Affairs*, Vol. 12 (1), pp.35-53

Taylor, G.D. (1986) “Multi- Dimensional Segmentation of the Canadian Pleasure Travel Market”, *Tourism Management*, 7 (3), pp,146-153

Table.1 Variables considered

| Variable name | Categories |
|--|---------------------------|
| The Price of the room desired | Less than 25\$ per day |
| | 25 to 50\$ per day |
| | 50 to 75\$ per day |
| | 70 to 100\$ per day |
| | Above 100\$ per day |
| The number of days in advance of the check in date that the enquiry was made | 0 to 7 days |
| | 8 to 14 days |
| | 15 to 30 days |
| | Above 30 days |
| The length of stay | 0 to 7 days |
| | 8 to 14 days |
| | 15 to 30 days |
| | Above 30 days |
| The repeat customer status of the enquirer | New customer |
| | Repeat customer |
| The nationality group of the enquirer | Northern Europe |
| | Western Europe |
| | East Asia |
| | North America |
| | Australia and New Zealand |
| | South Asia |
| The season when the enquiry was made | High season |
| | Low season |

Table.2 Clustering for Bangkok

| Cluster No. | | Cluster.1 | Cluster.2 | Cluster.3 | Cluster.4 | Cluster.5 |
|---|--------------------|---------------------------------|--------------------|---------------------------|--------------------------------|---|
| Number of members | | 175 | 109 | 93 | 135 | 81 |
| Cluster label | | New on-transit, economy-seeking | New Holiday makers | Repeat business customers | Economy seeking holiday makers | Repeat on transit economy seeking customers |
| Check in season | High | 0% | 100% | 53% | 100% | 0% |
| | Low | 100% | 0% | 47% | 0% | 100% |
| Length of stay | 1 to 7 nights | 95% | 96% | 90% | 94% | 95% |
| | 8 to 15 nights | 5% | 4% | 5% | 3% | |
| | Above 15 nights | 0% | - | 5% | 3% | 5% |
| Nationality status | South-Eastern Asia | 23% | 15% | 16% | 27% | 27% |
| | North America | 15% | 16% | 16% | 5% | 15% |
| | Western Europe | 17% | 11% | 18% | 13% | 16% |
| | United Kingdom | 13% | 15% | 18% | 27% | 27% |
| | Australia & NZ | - | 17% | 18% | 28% | 6% |
| | others | 32% | 26% | 14% | | 9% |
| No. of advance days | 0 to 7 days | 35% | 30% | 43% | 26% | 32% |
| | 8 to 14 days | 13% | 14% | 19% | 14% | 16% |
| | 15 to 30 days | 33% | 18% | 12% | 23% | 20% |
| | Above 30 days | 9% | 38% | 26% | 37% | 32% |
| Repeat customer status | New customers | 100% | 67% | 0 | 65% | 0% |
| | Repeat customers | 0% | 33% | 100% | 35% | 100% |
| Price range | Less than 25\$ | 30% | 11% | 8% | 18% | 30% |
| | 25 to 50\$ | 44% | 47% | 48% | 62% | 40% |
| | 50 to 75\$ | 14% | 24% | 14% | 6% | 18% |
| | 75 to 100\$ | 6% | 7% | 12% | 7% | 8% |
| | Above 100\$ | 6% | 11% | 18% | 7% | 4% |
| Percentage of total number of enquiries | | 29% | 18% | 15% | 22% | 13.6% |
| % of the total volume of enquiries | | 23% | 24% | 19.5% | 22% | 9% |

Table.3 Clustering for Phuket

| Cluster No. | | Cluster.1 | Cluster.2 | Cluster.3 | Cluster.4 | Cluster.5 |
|------------------------------------|--------------------|--------------------------------|--------------------------------|-------------------------|---------------------------------------|----------------------------------|
| Number of members | | 101 | 85 | 59 | 99 | 72 |
| Cluster label | | New economy seeking weekenders | High season, rich long stayers | Low season long stayers | High season short stay holiday makers | High season loyal holiday makers |
| Check in season | High | 0% | 100 | 34% | 100% | 84% |
| | Low | 100 | 0% | 66% | 0% | 16% |
| Length of stay | 1 to 7 nights | 88% | 37% | 37% | 93% | 90% |
| | 8 to 15 nights | 12% | 51% | 56% | 4% | 5% |
| | Above 15 nights | 0% | 12% | 7% | 3% | 5% |
| Nationality status | South-Eastern Asia | 16% | 6% | 0% | 17% | 15% |
| | North America | 10% | 10% | 25% | 14% | 15% |
| | Western Europe | 13% | 10% | 24% | 7% | 20% |
| | United Kingdom | 27% | 32% | 27% | 18% | 20% |
| | Australia & NZ | 8% | 4% | 24% | 8% | 12% |
| | others | 26% | 38% | | 36% | 18% |
| No. of advance days | 0 to 7 days | 30% | 4% | 17% | 28% | 20% |
| | 8 to 14 days | 23% | 6% | 12% | 17% | 25% |
| | 15 to 30 days | 27% | 15% | 15% | 27% | 29% |
| | Above 30 days | 20% | 75% | 56% | 28% | 26% |
| Repeat customer status | New customers | 67% | 88% | 40% | 78% | 10% |
| | Repeat customers | 33% | 12% | 60% | 22% | 90% |
| Price range | Less than 25\$ | 22% | 12% | 14% | 16% | 23% |
| | 25 to 50\$ | 45% | 22% | 39% | 23% | 34% |
| | 50 to 75\$ | 16% | 28% | 17% | 24% | 11% |
| | 75 to 100\$ | 11% | 18% | 9% | 12% | 10% |
| | Above 100\$ | 6% | 18% | 23% | 25% | 22% |
| Percentage of total enquiries | | 24% | 20% | 12% | 24% | 17% |
| % of the total volume of enquiries | | 19% | 26% | 17% | 22% | 14% |

Table.4 Clustering output for Koh Samui

| Cluster No. | | Cluster.1 | Cluster.2 | Cluster.3 | Cluster.4 | Cluster.5 |
|---|--------------------|------------------------------|----------------|----------------------------|---|----------------------------------|
| Number of members | | 140 | 142 | 85 | 85 | 48 |
| Cluster label | | New low season short stayers | High seasoners | Low season loyal customers | Low season economy seeking long stayers | Low season affluent long stayers |
| Check in season | High | 0 | 100% | 34% | 13% | 27% |
| | Low | 100% | | 65% | 87% | 72% |
| Length of stay | 1 to 7 nights | 99% | 78% | 100% | 15% | |
| | 8 to 15 nights | 1% | 16% | 0 | 77% | 96% |
| | Above 15 nights | | 6% | 0 | 8% | 4% |
| Nationality status | South-Eastern Asia | 13% | 8% | 16% | - | 18% |
| | North America | 10% | 19% | 10% | 3% | 17% |
| | Western Europe | 15% | | 22% | 23% | 21% |
| | United Kingdom | 20% | 25% | 24% | 32% | 31% |
| | Australia & NZ | 9% | 12% | 8% | 6% | |
| | others | 33% | 36% | 20% | 36% | 13% |
| No. of advance days | 0 to 7 days | 18% | 17% | 21% | 12% | 5% |
| | 8 to 14 days | 15% | 12% | 16% | 6% | 18% |
| | 15 to 30 days | 35% | 15% | 29% | 13% | 21% |
| | Above 30 days | 32% | 56% | 33% | 69% | 56% |
| Repeat customer status | New customers | 73% | 65% | 10% | 82% | |
| | Repeat customers | 27% | 35% | 90% | 18% | |
| Price range | Less than 25\$ | 4% | 8% | 5% | 3% | 10% |
| | 25 to 50\$ | 40% | 24% | 10% | 26% | 9% |
| | 50 to 75\$ | 34% | 29% | 44% | 54% | 50% |
| | 75 to 100\$ | 6% | 17% | 13% | 9% | 10% |
| | Above 100\$ | 15% | 22% | 28% | 8% | 21% |
| Percentage of the total number of enquiries | | 28% | 28% | 17% | 17% | 10% |
| % of the total volume of enquiries | | 12% | 31% | 18% | 24% | 13% |

Table.5 Clustering output for Krabi

| Cluster No. | | Cluster.1 | Cluster.2 | Cluster.3 | Cluster.4 |
|---|--------------------|------------------------------|---|---------------|---|
| Number of members | | 49 | 146 | 78 | 85 |
| Cluster label | | New high season long stayers | High season economy seeking short holidaymakers | Low seasoners | High season affluent short holiday makers |
| Check in season | High | 85% | 100% | | 100% |
| | Low | 15% | 0% | 100% | |
| Length of stay | 1 to 7 nights | - | 100% | 100% | 100% |
| | 8 to 15 nights | 100% | | | |
| | Above 15 nights | - | - | - | - |
| Nationality status | South-Eastern Asia | - | 10% | 30% | 10% |
| | North America | 8% | 10% | 14% | 13% |
| | Western Europe | 4% | 8% | 4% | 12% |
| | United Kingdom | 35% | 26% | 14% | 22% |
| | Australia & NZ | 7% | 6% | 4% | 10% |
| | others | 46% | 40% | 34% | 33% |
| No. of advance days | 0 to 7 days | 4% | 6% | 21% | 13% |
| | 8 to 14 days | 10% | 10% | 17% | 13% |
| | 15 to 30 days | 4% | 20% | 33% | 22% |
| | Above 30 days | 82% | 64% | 29% | 52% |
| Repeat customer status | New customers | 65% | 60% | 60% | 76% |
| | Repeat customers | 35% | 40% | 40% | 24% |
| Price range | Less than 25\$ | 8% | 9% | 21% | 4% |
| | 25 to 50\$ | 24% | 32% | 47% | 20% |
| | 50 to 75\$ | 37% | 28% | 18% | 20% |
| | 75 to 100\$ | 20% | 12% | 14% | 26% |
| | Above 100\$ | 11% | 19% | - | 30% |
| Percentage of total enquiries | | 14% | 40% | 22% | 24% |
| Percentage of volume of enquiries in each cluster | | 18.5% | 39% | 14% | 28% |