

1-1-2003

Winter tourist segments in Austria - Identifying stable vacation styles using bagged clustering techniques

Sara Dolnicar

University of Wollongong, s.dolnicar@uq.edu.au

Friedrich Leisch

Vienna University of Technology, Austria

Follow this and additional works at: <https://ro.uow.edu.au/commpapers>



Part of the [Business Commons](#), and the [Social and Behavioral Sciences Commons](#)

Recommended Citation

Dolnicar, Sara and Leisch, Friedrich: Winter tourist segments in Austria - Identifying stable vacation styles using bagged clustering techniques 2003.
<https://ro.uow.edu.au/commpapers/44>

Winter tourist segments in Austria - Identifying stable vacation styles using bagged clustering techniques

Abstract

Market segmentation is a very popular and broadly accepted way of increasing profitability. The number of reports published on a posteriori market segmentation studies has rapidly increased since Russel Haley's milestone publication on benefit segmentation in 1968. Nevertheless, it is common practice in market segmentation to use a single segmentation base only, thus choosing the main dimensions of interest a priori, and to run a single calculation of a single algorithm, which dramatically increases the chance of building an entire marketing plan on a random solution of the algorithm chosen. The application presented constructs winter vacation styles on the basis of Austrian Guest Survey data, avoiding both weaknesses mentioned before. Through the replicative framework provided by bagged clustering, potentially suboptimal random solutions are avoided. Independent partitioning of vacation activities and travel motives leads to more holistic market segments. By looking for over- and under-representation of all combinations of the behavioral and psychographic segmentation, vacation styles are identified and studied in detail.

Keywords

market segmentation, vacation styles, bagged clustering

Disciplines

Business | Social and Behavioral Sciences

Publication Details

Dolnicar, S & Leisch, F, Winter tourist segments in Austria - Identifying stable vacation styles using bagged clustering techniques, *Journal of Travel Research*, 2003, 41(3), 281-292. Copyright Sage Publications Ltd 2003. This article won the Charles R. Goeldner Article of Excellence Award for the best paper in the *Journal of Travel Research* 2003.

Winter Tourist Segments in Austria - Identifying Stable Vacation Styles for Target Marketing Action

This article won the Charles R. Goeldner Article of Excellence Award for the best paper in the Journal of Travel Research in 2003.

MMAG. DR. SARA DOLNICAR

Assistant Professor at the Institute for Tourism and Leisure Studies, Vienna University of Economics and Business Administration

Augasse 2-6
A-1090 Vienna, Austria
Telephone: ++43 (1) 313 36 / 4476
Fax: ++43 (1) 317 12 05
sara.dolnicar@wu-wien.ac.at

DI DR. FRIEDRICH LEISCH

Assistant Professor at the Department of Statistics, Probability Theory and Actuarial Mathematics, Vienna University of Technology

Wiedner Hauptstrasse 8-10/1071
A-1040 Vienna, Austria
friedrich.leisch@ci.tuwien.ac.at

ACKNOWLEDGMENTS

This research project was supported by the Austrian Science Foundation (FWF) under Grant SFB#010 ('Adaptive Information Systems and Modeling in Economics and Management Science').

ABSTRACT

Market segmentation is a very popular and broadly accepted way of increasing profitability. The number of reports published on *a posteriori* market segmentation studies has rapidly increased since Russel Haley's milestone publication on benefit segmentation in 1968. Nevertheless, it is common practice in market segmentation to use a single segmentation base only, thus choosing the main dimensions of interest *a priori*, and to run a single calculation of a single algorithm, which dramatically increases the chance of building an entire marketing plan on a random solution of the algorithm chosen. The application presented constructs winter vacation styles on the basis of Austrian Guest Survey data, avoiding both weaknesses mentioned before. Through the replicative framework provided by bagged clustering, potentially suboptimal random solutions are avoided. Independent partitioning of vacation activities and travel motives leads to more holistic market segments. By looking for over- and under-representation of all combinations of the behavioral and psychographic segmentation, vacation styles are identified and studied in detail.

Keywords: market segmentation, vacation styles, bagged clustering

INTRODUCTION

Austrian tourism officials like to speak of Austria as a “world champion” in tourism. Ignoring the existence of small islands such as the Maldives, Dominican Republic, etc., this actually is true in terms of tourism receipts per resident (about 1400 euros in 1999). However, even if the size of Austria is not taken into account, it ranks 10th in terms of arrivals worldwide and 8th in terms of tourism income according to WTO statistics. This illustrates that the tourism industry is highly developed in and extremely important for Austria, which of course is a purely static statement. From the dynamic point of view of structural changes in tourism, one dramatic development that has taken place in the past decades is a steady decrease in the summer season’s dominance. As illustrated in Figure 1, Austria used to be a typical summer destination. In the late 70s, the overnight stays in summer were about twice as frequent as overnight stays during the winter season. But the winter season has become more and more important every year, so that by the turn of the millennium winter tourism had almost caught up completely. The importance of winter tourism is not only increasing in terms of overnight stays, it is also known from the Austrian National Guest Survey that tourists visiting Austria during the winter season tend to spend 40 percent more money per day at their destinations than the visitors during the main summer season.

FIGURE 1

In sum, winter tourism plays a major role not only within the Austrian tourism industry but for the Austrian economy as a whole. And as competition is increasing among typical winter destinations, especially among destinations located in the Alps region, it is necessary to investigate which sub-markets might be of special interest to national tourism organizations (NTOs), regional tourism organizations (RTOs) or individual hotels, travel agencies, etc. Although market segmentation has long been recognized by the Austrian tourism industry, marketing action is mostly based on *a priori* segmentation approaches, especially by promoting so-called “holiday specialists”, including special products such as sports holidays with different focus, family vacation packages or culture trips. An excellent example can be found on the Austrian NTO’s web site (<http://www.austria-tourism.at/>).

When the horizon is extended beyond the borders of Austria, the density of segmentation studies focusing on winter tourism does not turn out to be very high either. Only a few publications in academic journals address the issue: Perdue (1996) investigates high potential market segments for the Colorado downhill skiing industry on the basis of market characteristics, access and competition, Ahmed, Barber & d’Astous (1998) segment winter travelers in Canada according to benefits sought, and Mills, Couturier and Snepenger (1986) describe the group of Texans that spend their leisure time skiing.

Therefore, sophisticated market segmentation approaches emphasizing the winter segment and taking a variety of information about the tourists into account would appear to be a promising way to gain competitive advantage and should be of great interest to a number of players in the Austrian tourism industry.

In the broader field of tourist market segmentation, attention has already been drawn to *a*

posteriori segmentation, and it seems to have been recognized that - as there is rarely something resembling a natural clustering structure in the data - competitive advantage can be achieved by generating interesting market segments by means of advanced methodology. A recent survey conducted by Baumann (2000) illustrates the common practice of segmentation studies within the field of economics and business administration and demonstrates how broad the field for competitive advantage still is. This meta-analysis is based on 250 studies, 47 of which are from the field of tourism. The results indicate that interest in the application of market segmentation is growing. 24% of the publications investigated are from the 1980s, 76% from the 1990s. Also, it is clear that the purpose of segmentation is almost exclusively to identify or construct groups of customers, as only two out of 47 studies state a different aim.

The following two insights motivate the application reported in this article:

- 77% state that they use psychographic variables and the remaining studies make use of behavioral information, with not one single study using two different kinds of information. The literature tends to ignore the fact that in the best case segments should reflect a number of integrated personal and behavioral characteristics and thus represent a vacation style (Wells and Tigert, 1971; Wells, 1974, Mazanec and Zins, 1993) rather than a one-dimensional activity or benefit type.
- Only a minority of the segmentation studies (21%) raise the issue of stability of results. Therefore, the results obtained by applying hierarchical (49%) and partition-based methods (49%) could end up being atypical and random.

In this article, we present an *a posteriori* segmentation study of winter tourists in Austria that differs from most of the studies in the field of tourist market segmentation so far: First, it is built upon two different and independent segmentation bases (psychographic and behavioral). Second, the bagged clustering framework is applied, which makes use of both partitioning and hierarchical methods and leads to stable results by aggregating replicated results, as demonstrated in a Monte Carlo simulation with empirical data by Dolnicar & Leisch (2000a). As a result, vacation styles are identified and described in detail.

DATA

Data from the Austrian National Guest Survey¹ was used as the basis for the segmentation task. The data set includes 2961 respondents questioned during their stays in Austria in the winter season of 1997/1998. Please note that tourists staying in the capital cities of Austria are not included in this sample, as a different questionnaire was used for this group. Hence, the number of cases used in our analysis amounts to 2961 out of the original 3599.

Two questions were of central interest for the segmentation task: travel motives and vacation activities.

The guest survey measures travel motives with an instrument developed by the *Europäisches Tourismus Institut GmbH* at the University of Trier. This instrument has been used in a number of consecutive guest surveys in Austria over the last decades. Thus the questions could not be influenced by the authors of this paper. The motives are operationalized by means of four-point scale ratings for personal statements on travel-related desires and consumption goals. The respondents' reactions are prompted by the phrase: "What is important to you during your present holiday? Please tell me for every one of the following statements, if it applies to you totally, mostly, a bit, or not at all." The 'totally' category is coded as agreement in binary format, the remaining possible answers are aggregated into an 'unimportant' category². The 26 motive statements available in the data are given in Table 1. The abbreviations used later in the exhibits are shown in Column 3.

TABLE 1

Although some of these statements are very precise, they are considered to be indicative of abstract travel motives here. It is also assumed that travel motives can be different for various travelers, even if they spend their vacation at the same destination engaging in the same leisure activities.

The leisure activities are also questioned in the survey and included as a second vacation style dimension in the segmentation study. The respondents were asked to check off activities that they did often, sometimes or never during their winter holiday in Austria. The following activities were included in the list (the percentage in brackets gives the proportion of all respondents who answered either 'often' or 'sometimes'): alpine skiing (76%), cross country skiing and touring (10%), snowboarding (10%), carving (5%), ski touring (2%), ice skating (5%), tobogganing/sleigh-riding (19%), tennis (5%), horseback riding (1%), going to a spa (11%), using health facilities (10%), hiking (29%), going for walks (63%), organized excursions (7%), excursions (30%), relaxing (72%), going out in the evening (59%), going to discos/bars (35%), shopping (53%), sightseeing (23%), museums, exhibitions (9%), theater, musicals, opera (2%), visiting a "Heurigen" (5%), visiting festivals, concerts (2%), visiting "Tyrolean evenings", amateur theaters (6%), visiting other local and regional events (16%), going to an indoor pool/sauna (46,2%). For the segmentation exercise, the categories 'sometimes' and 'often' were joined, resulting in a binary data set with 1

¹ The Austrian National Guest Survey is conducted every three years. The total winter sample size, including the city tourists, was 3599. The weighted data set is used for analysis, thus correcting country of origin and destination effects.

² Bagged clustering currently works only for binary and metric variables, because cluster means can be interpreted using usual Euclidean geometry for these types only. Extensions to general categorical variable types are under investigation.

denoting that an activity was done and 0 that it was not tried at all.

The rationale behind merging ordinal categories is different for the motive questions and the activities. In the case of motives, the original data format allows four possible answers. Only strong motivation is accepted as driving force, therefore only the strongest agreement level is coded as 1. In the case of activities, the main emphasis is to determine whether the tourists engage in a certain leisure activity (and pay for it) or not. Thus both the ‘often’ and the ‘sometimes’ category are joined in the same group, as opposed to the respondents who indicated that they had never spent time on these activities (the original data format allows three answer categories).

In addition to the variables used as a segmentation base, a number of demographic, socioeconomic, behavioral and psycho-graphic background variables are available in the extensive guest survey data set. A description of variables selected is provided in Table 2. Especially the monetary variables, expenditures and income, have a strongly skewed distribution with large positive outliers (as expected), which prompted us to give the robust median instead of mean values. For all categorical variables, observations are indicated as percentages in each category.

TABLE 2

METHODOLOGY

The bagged clustering framework was applied to the data from the Austrian National Guest Survey. The central idea behind bagged clustering is to stabilize partitioning methods like K -means (e.g., Anderberg, 1973) or learning vector quantization (LVQ, e.g. Ripley, 1996) by repeatedly running the cluster algorithm and combining the results. K -means is an unstable method in the sense that in many runs one will not find the global optimum of the error function but only a local optimum. Both initializations and small changes in the training set can have a strong influence on the actual local minimum where the algorithm converges.

By training repeatedly on new data sets, one obtains various solutions, the average of which should be independent of training set influence and random initializations. A collection of training sets can be obtained by taking samples from the empirical distribution of the original data, i.e., by bootstrapping. Then any partitioning cluster algorithm - called the *base cluster method* below - can be run on each of these training sets.

Bagged clustering simultaneously explores the independent solutions from several runs of the base method in an exploratory way using hierarchical clustering. The results of the base method are combined into a new data set, which is then used as input for a hierarchical method. This allows the researcher to identify structurally stable (regions of) centers which are found repeatedly.

The algorithm works as follows:

1. Construct B bootstrap training samples X_N^1, \dots, X_N^B by drawing with replacement from the original sample X_N .
2. Run the base cluster method (K -means, LVQ, ...) on each set, resulting in $B \times K$ centers $c_{11}, c_{12}, \dots, c_{1K}, c_{21}, \dots, c_{BK}$ where K is the number of centers used in the base method and c_{ij} is the j -th center found using X_N^i .

3. Combine all centers into a new data set $C^B = C^B(K) = \{c_{11}, \dots, c_{BK}\}$.
4. Run a hierarchical cluster algorithm on C^B , resulting in the usual dendrogram.
5. Let $c(x)$ in C^B denote the center closest to x . A partition of the original data can now be obtained by cutting the dendrogram at a certain level, resulting in a partition C_1^B, \dots, C_m^B of set C^B . Each point x in X_N is now assigned to the cluster containing $c(x)$.

The algorithm has been shown to compare favorably to several standard clustering methods on binary and metric benchmark data sets (Leisch, 1998, 1999). Dolnicar & Leisch (2000a, 2000b) demonstrate its usage in the field of tourist segmentation on the basis of summer activities and introduce visualization techniques for clusters of binary data using boxplots (Figures 3 and 4).

One of the original motivations of bagged clustering was the standard textbook approach of determining the number of clusters for k-means by running a hierarchical clustering algorithm on one or more sub-samples of the data. Bagged clustering uses several quantizations of the data (the outcome of k-means clustering on the bootstrap samples) instead of sub-samples. This reduces the effect of outliers by smoothing the data (quantization) and simultaneously explores data set variability (bootstrapping). See Leisch (1999) for a detailed discussion of advantages and limitations.

RESULTS

The bagged clustering procedure was calculated independently for the travel motives and vacation activities data. As in Dolnicar & Mazanec (in press), vacation styles are defined in the next step by cross-tabulating the behavioral and the psychographic segmentation results.

For both data sets (motives and activities), K-means was used with $K=10$ centers as the base method. The exact number of centers K has been shown not to be a very sensitive parameter within the bagged clustering framework (Leisch, 1999). The only critical issue is not to choose a number lower than the number of segments expected to exist in the data, as superfluous centers (possibly slitting a segment) are merged in the hierarchical step of the procedure. The base method was applied on $B=50$ bootstrap samples, resulting in a total of 500 centers, which were then hierarchically clustered using Euclidean distance and average linkage method (e.g., Anderberg 1973). These parameters were chosen because they performed best in empirical studies (Leisch, 1998) on simulated artificial data with characteristics similar to those of the present data set (Dolnicar et al., 1998).

All computations and graphics were done using the R software package for statistical computing (see <http://www.R-project.org>). R functions for bagged clustering can be obtained from the authors upon request.

Figure 2 depicts the dendrogram resulting from the hierarchical step of the bagged clustering procedure for the travel motives data. The dendrograms for both data sets look well-structured, and especially in the case of the motive data, there is an unusually clear recommendation in favor of the seven-cluster solution. In this case there is a strong indication that some kind of seven-cluster structure exists. In the behavioral segmentation, there is no such strong recommendation for the data structure, as three numbers of clusters are recommended by the dendrogram distances. Five clusters were chosen, because three clusters render too rough a picture of the tourists and eight

clusters seem too detailed from an interpretational standpoint³.

FIGURE 2

In the next section, both the behavioral and the psychographic types are outlined on the basis of the segmentation variables only. The vacation styles result from cross-tabulating the two partitions; the background variables are used to evaluate the attractiveness of the vacation styles even further.

Behavioral types of winter tourists

The following conclusions can be drawn about the behavioral tourist segments among the visitors to Austria during the winter season of 1997 on the basis of the boxplots for this particular bagged clustering solution:

Type 1 (62%) Skiers: The group with the fairly non-informative name 'skiers' is very large and therefore rather heterogeneous, as the profile chart in Figure 3 illustrates. Although there is a fairly high level of agreement regarding the high level of winter sports activity and low interest in cultural hobbies of any kind, this type differs greatly in regard to going for walks, relaxing, going out in the evening, going to discos and bars, shopping and going to indoor pools. This group consequently represents the majority of winter travelers in Austria very well.⁴

FIGURE 3

Type 2 (17%) Health tourists: This group represents the complete opposite to the typical winter tourist in Austria: no alpine skiing but a lot of time in spas and health facilities. As far as sports are concerned, this group prefers calm and relaxing activities, such as cross-country skiing, going for walks or hiking.

Type 3 (11%) Superactives: These tourists either say they do everything that is offered or they really are extremely active, as all activities are agreed with at an above-average level. The group is also very homogeneous in answering the questions. Unfortunately, it cannot be determined ex post whether this really is a behavioral type or an answer tendency.

³ Of course, bagged clustering is generally less concerned with the number-of-clusters problem, as the same bagged clustering solution can be explored arbitrarily for different numbers of tourist segments. Exploring the bagged clustering solutions allows exploration in the form of stepwise splitting and thus enhances insight into the unobserved heterogeneity structure inherent to the survey data.

⁴ The differences between these segments concerning all metric background variables under investigation are significant. Except for the daily expenditures (p-value = 0.014), the error probability is below 1%. The ordinal variables also mainly differ in a highly significant manner, except for the following variables: getting information from travel agencies, media ads, friends and relatives, local tourism bureaus.

Type 4 (5%) Culture tourists: Sightseeing, visiting museums and exhibitions, going to the theater and the opera, shopping and going for walks best characterize the vacation of the culture tourist.

Type 5 (5%) Snowboarding and nightlife: This very small group seems to have very precise plans for a winter holiday: snowboarding during the day (sometimes also carving), and going to clubs and bars at night. This clear attitude is very well represented in the activity profile of this extremely homogeneous group.

Psychographic types of winter tourists

Type 1 (10%) In search of the destination of their dreams: This small group is characterized by a very interesting motivation in their vacations. Only three variables are important, and these are very strongly agreed upon within the group: sun, enough snow and the attractiveness of the skiing destination. At the same time, sports do not play a major role in their minds.⁵

Type 2 (22%) Pure relaxation: These travelers do not care very much about sports. They want to relax in beautiful surroundings and unspoiled nature. They enjoy the change of pace and want to take it easy during their stay. Although the group is not very large, the motives are not extremely homogeneous, but the main tendency to seek relaxation is common to all members of the group.

Type 3 (4%) No motives: The respondents within this group gave below-average answers to all travel motives listed in the questionnaire, in a very homogeneous manner. Of course, this group could represent an answer pattern. The inability to identify answer patterns *ex post* is fortunately not crucial in this case, as the pattern does not signal an interesting segment.

Type 4 (38%) Active fun: There is a fairly high level of agreement in this group that sports play a major role during their winter holidays. Therefore, sufficient snow in the region and a highly attractive skiing area are desirable. The group cares about sun, unspoiled surroundings, meeting people and having fun.

Type 5 (10%) Sportsfans: "Sports and sun" could be the motto of this group of tourists. Although only these two variables characterize this group, the agreement on the sports issue is not as high as could be expected.

Type 6 (3%) Budget travelers: If there is one motivation this group agrees upon, it is the concept of not exceeding their planned budget. Also, the holiday destination should be near home. Sun, snow and nature are important destination features for this group.

Type 7 (3%) Culture fans: This small group seems to be driven by one single but extremely dominating motive: culture. This factor almost seems to represent a knock-out criterion for any destination choice, and the culture fans strongly agree on this issue. Another interesting factor is the need for safety.

FIGURE 4

⁵ The differences between the psychographic segments in terms of the metric background variables under investigation are highly significant ($p < 0.001$), except for the insignificant difference in length of stay. Among the ordinal variables, the p-values for the following variables are higher than 0.001: sex (0.017), getting travel information from brochures (0.123, n.s.), from catalogs (0.009), friends and relatives (0.016) and traveling in a group (0.204, n.s.).

Vacation styles

It is reasonable to assume that travel motives and vacation behavior are not independent of one another. As a single cluster analysis of all variables would fall victim to the curse of dimensionality, vacation styles are constructed by cross-tabulating the results of the behavioral and the psychographic taxonomy. The mosaic chart (Friendly, 1994) in Figure 5 illustrates the cross-tabulation of the respondents.

FIGURE 5

Full lines around the squares indicate table entries that are larger than expected (given row and column totals and assuming independence) and hence can be interpreted as vacation styles. Anti-vacation styles (values smaller than expected) are represented by dashed lines. The darker the squares, the stronger the deviation of the theoretically expected number of respondents at the intersections among behavioral and psychographic groups.

Based on the data set investigated, we can thus identify 5 vacation styles which are worthy of thorough investigation. The background variable values for all vacation styles are given in Table 3, including their significance levels. The descriptions that follow include the variables used to group the individuals as well as the background variables. The vacation types constructed not only meet the criteria of mutual exclusiveness and exhaustiveness (Kotler, 1988) (a logical consequence of the algorithm used), most of them also turn out to be measurable (at the destination itself, the activity information might even be sufficient, but numerous significant descriptors can also be identified) and accessible (nearly all information channels show significantly different usage patterns by the vacation types). Differences in the responses to marketing strategy unfortunately can not be investigated, as no such questions were included in the questionnaire. Finally, the substantiality criterion cannot be judged independently of the unit actually choosing the optimum segment.

Vacation Style 1 (Activity Type 1 and Motive Type 4) / Fun & snow (31%): This vacation style could be called a mass market. Except for this one style, no other combination of behavioral and psychographic types reaches a size of more than 10%. The background variables do not reveal any valuable additional information as can be seen in Table 3.

Vacation Style 2 (Activity Type 2 and Motive Type 2) / Relaxation and Health (10%): The main aim of the vacation is to relax and enjoy healthy activities and body care. This group is about ten years older than the overall winter tourist population, and the proportion of men is lower. Their monthly disposable income is also very low, although this does not seem to affect daily per capita expenditures. Not only these factors but also the average length of stay and high level of loyalty to Austria as compared to the other vacation styles make this segment an interesting target group.

Vacation Style 3 (Activity Type 4 and Motive Type 2) / Moderate Culture Tourist (3%): Looking at the activity pattern of this group, it becomes obvious that there is only one reason for these tourists to visit Austria: culture. The only activities besides sightseeing and going to museums, exhibitions, the theater and operas are going for walks and spending some time shopping. Typical of culture tourists, the length of their stay is slightly shorter and the proportion of segment members spending their first vacation in Austria is extremely high. In addition, these tourists prefer hotels and tend to travel with their partners more often than others do. Media ads would appear to

be an excellent means of marketing communication for this group. All in all, this is a very profiled group which can be targeted in a highly effective manner. The rather short length of stay and low daily expenditures per person are the drawbacks of targeting this segment.

Vacation Style 4 (Activity Type 4 and Motive Type 7) / Pure Culture Tourist (1%): Culture is their main motive, and their activities reflect this attitude perfectly. This segment has an extremely high disposable monthly income, stays in Austria for a very short period of time only, is dominated by female travelers (only 42% are male), shows a strong intention to return to Austria at some point and usually travels either with partners or families. The extremely high proportion of 90% stays in hotels; travel agents seem to be the best way of reaching this group. Again, this vacation style is the perfect segment for all sorts of cultural attractions or destinations, although it definitely represents a niche market.

Vacation Style 5 (Activity Type 5 and Motive Type 4) / Fun, Snow, Snowboards and Discos (3%): This vacation style is especially interesting because there is only a random probability of identifying it with single runs of a certain algorithm (e.g., only 9 out of the 50 K-means replications identified activity type 5). Nevertheless, it seems to be a very clearly profiled segment which would enable destinations or tourism businesses to offer a customized product. Action and fun being the motto of this niche segment, their days are dedicated to snowboarding and carving, their nights are reserved for clubs and bars. Their goal is entertainment throughout the entire vacation. The median age is 23; 82% of the members are male and travel with friends. In consequence, their main sources of information are friends and relatives as well. With only 33% staying in hotels, their daily expenditures are kept below average.

TABLE 3

STABILITY ANALYSIS

Cluster analysis is an "unsupervised learning technique" and thus we can not assess directly whether the outcome corresponds to real density clusters or if the partition is a mere artefact of the algorithm. One standard way of assessing the stability of a partitioning solution, i.e., if it is replicable over several runs of the algorithm, is to partition the data several times and compare the results for mutual agreement.

We have clustered 50 bootstrap samples (Efron & Tibshirani, 1993) drawn with replacement from the original data set, both using standard k-means and bagged clustering and looking for 2 to 10 clusters. All pairs of partitions for the same number of clusters were compared using the corrected Rand Index (Hubert & Arabie, 1985). This index is 0 if the agreement is due to pure chance and 1 if partitions agree completely.

FIGURE 6

FIGURE 7

Figure 6 and **Figure 7** show the mean corrected Rand Index over all pairwise comparisons for the motive and activity data, respectively. In both cases, *K*-means is only competitive in terms of stability for solutions with 2 or 3 clusters, which are not interesting when searching for niche segments. For higher numbers of clusters, bagging improves the stability of the solution (in terms of average agreement between partitions) considerably.

CONCLUSIONS

There is no doubt that market segmentation is one of the most powerful strategic tools in the tourism industry, be it on an organizational or a destination level. Using the bagged clustering approach, stable behavioral and psychographic winter tourist types in Austria are constructed from guest survey data. By cross-tabulating these results and looking for over- or under-representation of respondents at the intersections, vacation styles can be identified, and then used by Austrian NTOs, RTOs or the tourism industry as target segments for focused marketing action. All potential market segments (behavioral, psychographic or combined) seem valid and useful, as the background variables significantly differ between groups. Besides a very general vacation style representing one third of all tourists, four small segments are derived, offering interesting prospects for niche marketing action, with the young, fun and action-oriented group representing a very special niche. The members of this group are only half as old as the average winter tourist in Austria and they have a penchant for sun, fun and action that keeps them going day and night. Their daily expenses are kept low at the cost of the accommodation category, but members of this segment will come again ... and they will bring their friends.

Due to their sizes, the vacation styles constructed have little chance of being the result of single hierarchical or *K*-means analysis runs. Bagged clustering enables the identification of these niches on a stable basis and thus prevents us from choosing random results as a basis for marketing action on the one hand and not encountering the segment on the other. Besides the methodological advantage of the partitioning outcome, the combination of activities and travel motives renders a more holistic perspective of the tourist without running cluster analysis on too many variables. Hence, the concept of holistic vacation styles which integrate more than one dimension of tourist characteristics is reflected more accurately.

Limitations to the approach obviously include the fact that the empirical data is examined in a purely exploratory manner, and that there is no “proof” that the extracted segments are not artefacts of the partitioning procedure (as no significance tests are available for the outcome of cluster analysis). These limitations are not specific to bagged clustering, they are common to the entire family of partitioning algorithms typically used in the field of market segmentation and thus represent an obvious target for further research. Another interesting topic could be even more refined segmentation techniques, as we can still observe generic segments without distinct profiles, which might mirror answer tendencies more than actual motive or activity patterns.

BIBLIOGRAPHY

- Ahmed, S. A., M. Barber and A. d'Astous (1998). 'Segmentation of the Nordic Winter Sun Seekers Market', *Journal of Travel and Tourism Marketing*, vol. 7, pp. 39-53.
- Anderberg, M.R. (1973) *Cluster analysis for applications*. New York: Academic Press.
- Baumann, R. (2000) Marktsegmentierung in den Sozial- und Wirtschaftswissenschaften: eine Metaanalyse der Zielsetzungen und Zugänge. Master's thesis at Vienna University of Economics and Business Administration. Vienna, Austria.
- Dolnicar, S. & Leisch, F.(2000a). Getting More Out of Binary Data: Segmenting Markets by Bagged Clustering, Working Paper # 71, August 2000 SFB "Adaptive Information Systems and Modeling in Economics and Management Science", <http://www.wu-wien.ac.at/am>.
- Dolnicar, S. & Leisch, F.(2000b) Behavioral Market Segmentation Using the Bagged Clustering Approach Based on Binary Guest Survey Data: Exploring and Visualizing Unobserved Heterogeneity. *Tourism Analysis*, Vol. 5, No. 2-4, pp. 163-170.
- Dolnicar, S., Leisch, F. & Weingessel, A.(1998) Artificial Binary Data Scenarios. Working Paper # 20, SFB 'Adaptive Information Systems and Modeling in Economics and Management Science', Vienna.
- Dolnicar, S. & Mazanec, J. (in press) Vacation Styles and Tourist Types: Emerging New Concepts and Methodology, in: Gartner, W. and Lime, D.W. (eds.), *Trends in Outdoor Recreation, Leisure and Tourism*. CAB International, Wallingford, UK.
- Efron, B. & Tibshirani, R. J. (1993) *An introduction to the bootstrap*. Chapman & Hall, Monographs on Statistics and Applied Probability, New York, USA.
- Friendly, M. (1994), 'Mosaic displays for multi-way contingency tables.', *Journal of the American Statistical Association*, vol. 89, pp. 190-200.
- Haley, R. J. (1968), 'Benefit Segmentation: A Decision-Oriented Research Tool', *Journal of Marketing*, vol. 32, pp. 30-35.
- Hubert, L. & Arabie, P. (1995) 'Comparing partitions'. *Journal of Classification*, 2, 193-218, 1985.
- Kotler, P (1988). *Marketing Management*. Englewood Cliffs : Prentice Hall
- Leisch, F. (1998). *Ensemble methods for neural clustering and classification*. doctoral thesis, Institut für Statistik, Wahrscheinlichkeitstheorie und Versicherungsmathematik, Technische Universität Wien, Austria.
- Leisch, F. (1999). *Bagged Clustering*. Working Paper # 51, SFB "Adaptive Information Systems and Modeling in Economics and Management Science", <http://www.wu-wien.ac.at/am>.
- Mazanec, J.A. and A. Zins (1993), 'Tourist Behavior and the New European Lifestyle Typology: Exploring the Managerial Relevance for Tourism Marketing', in: Theobald, W. (ed.), *Global Tourism: The Next Decade*, pp. 199-216, Oxford: Butterworth-Heinemann.

Mills, A.S., H. Couturier and D.J. Snepenger (1986). 'Segmenting Texas Snow Skiers', *Journal of Travel Research*, vol. 25, pp. 19-23.

Perdue, R. (1996). 'Target Market Selection and Marketing Strategy: The Colorado Downhill Skiing Industry', *Journal of Travel Research*, Spring 1996, pp. 39-46.

Ripley, B.D. (1996). *Pattern recognition and neural networks*. Cambridge, UK: Cambridge University Press.

Wells, W. D. (1974), *Life Style and Psychographics*, Chicago: AMA.

Wells, W. D. and D. J. Tigert (1971), 'Activities, Interests and Opinions', *Journal of Advertising Research*, vol. 11, pp. 27-35.

Figure 1: Development of overnight stays in Austria, 1980 to 2000

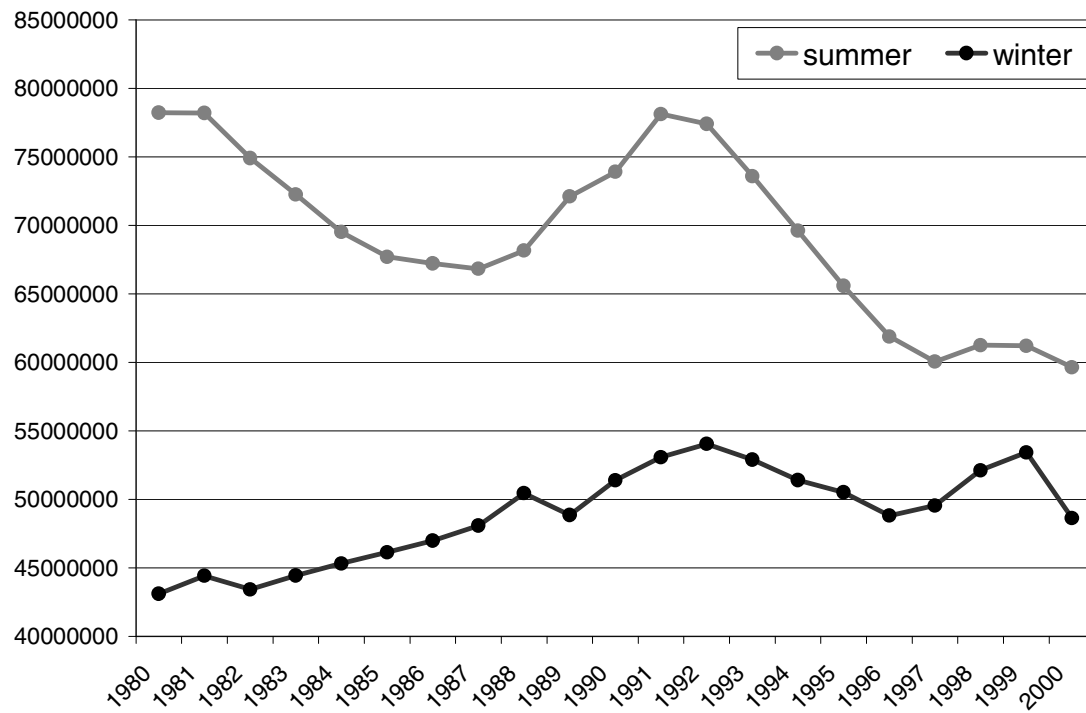


Table 1: Motive statements

Motives	Percent agreement	Abbreviation
'When I am on holiday I want to rest and relax.'	24%	relax
'On holiday I am looking for comfort and want to be spoilt.'	16%	comfort
'On holiday I want to exert myself physically and I play sports.'	42%	sports
'This holiday means excitement, a challenge and special experiences for me.'	18%	excite
'I do not want to exceed my planned vacation budget.'	22%	budget
'On holiday I want to realize my creativity.'	8%	creative
'On holiday I am looking for a variety of fun and entertainment.'	21%	fun
'On this holiday good company and getting to know people is important to me.'	23%	people
'I use my holiday for the health and beauty of my body.'	18%	body
'On this holiday I put much emphasis on free-and-easy-going.'	48%	easy
'I like to spend my holidays where there are many entertainment facilities.'	7%	entertain
'I do not want to care about prices and money.'	15%	prices
'On holiday I am interested in the lifestyle of the local people.'	21%	locals
'The special thing about my holiday is an intense experience of mountains, snow and sun.'	61%	sun
'On holiday I look for coziness and a familiar atmosphere.'	43%	cozy
'On holiday the efforts to maintain unspoiled surroundings play a major role for me.'	46%	surround
'On holiday it is important that everything is organized and I do not have to care about anything.'	16%	organized
'When I chose this holiday-resort an unspoiled nature and natural landscape played a major role for me.'	30%	nature
'Cultural offers and sights were a crucial factor.'	5%	culture
'I went on this holiday for a change to my usual surroundings.'	36%	change
'When choosing a holiday destination I put much emphasis on a romantic and nostalgic atmosphere.'	12%	romantic
'I want to stay near to my home.'	15%	near home
'When choosing a holiday destination it is important to me that there are offers and care for the children.'	5%	children
'When choosing a holiday destination it is important to me that I can feel safe.'	37%	safe
'I want to be sure that there is enough snow there.'	46%	enough snow
'I want an attractive skiing destination.'	41%	skiing dest.

Figure 3: Profile chart for Behavioral Type 1

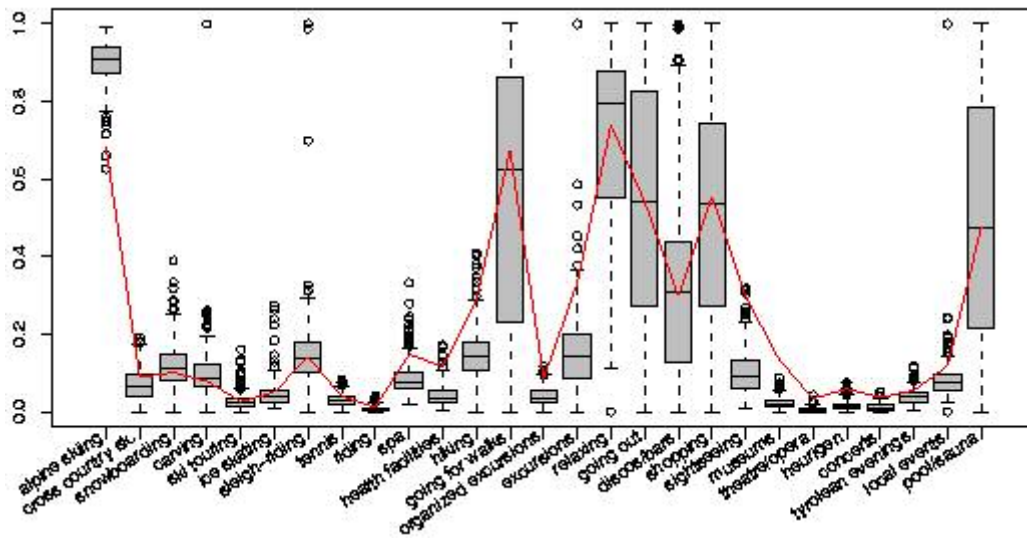


Figure 4: Profile chart for the culture fans

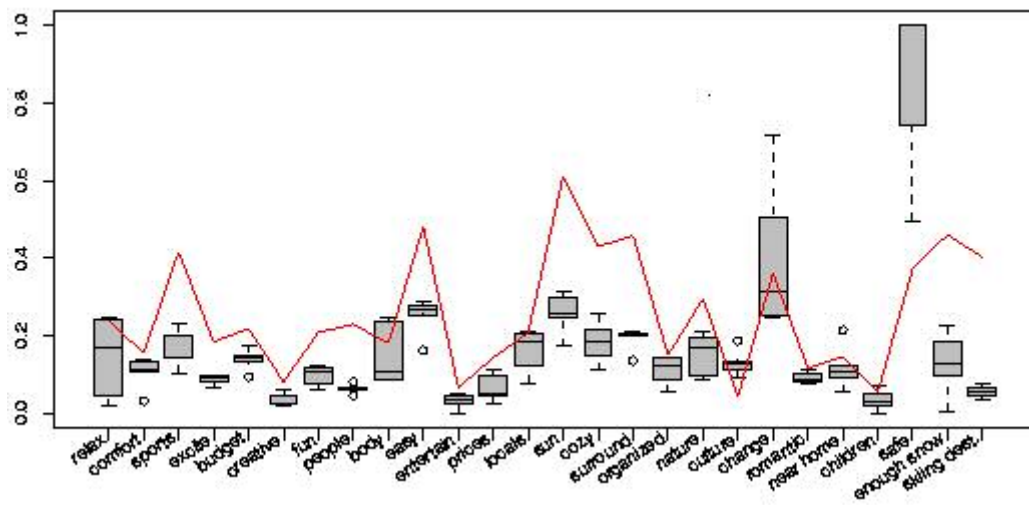


Figure 5: Mosaic chart of cross-tabulation

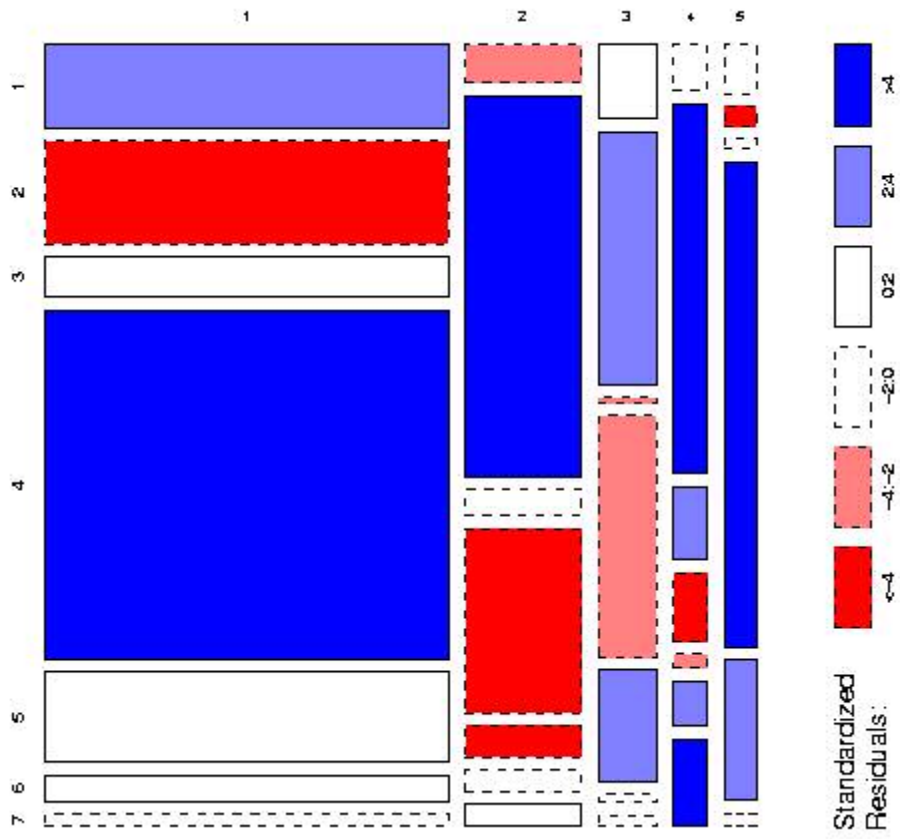


Table 2: Description of background variables⁶

Variable (Type)	Units / Levels	% per level or Median
Age (metric)	years	40
Daily expenditures per person (metric)	EUR	74.56
Monthly disposable income (metric)	EUR	2258.08
Length of stay (metric)	days	7
Sex (dichotomous)	male	59
Prior vacations in Austria (3 ordered categories)	Never	2
	Once	5
	Twice or more	93
Intention to revisit Austria (4 ordered categories)	Certainly	36
	Probably	27
	Probably not	19
	Certainly not	18
Accommodation (5 nominal categories)	Hotel	57
	Bed and Breakfast	13
	Apartment	17
	Private room	10
	Private room on a farm	3
Travelling in company of ... (4 variables)	Partner	69
	Family	44
	Friends	36
	Travel group	5
Sources of information (8 variables)	Brochures	8
	Media ads	2
	Tour operator catalog	8
	Friends and relatives	30
	Travel agent	9
	Local/regional tourism bureau	10
	Internet	4
	No information needed	36

⁶ For metric variables we list the median, for categorical variables the percentage of people per category.

Table 3: Vacation style background variables

		Vacation style 1	Vacation style 2	Vacation style 3	Vacation style 4	Vacation style 5	p- value ⁷
Variable	Unites / Levels	% or median	% or median	% or media	% or median	% or median	
Age	Years	39	51	54	54	23	<0.001
Daily expend. p.p.	EUR	74.71	73.76	63.08	66.93	60.32	<0.001
Monthly income	EUR	2543.55	2197.7	2459.98	3606.53	2301.4	<0.001
Length of stay	days	7	7	6	5	7	<0.001
Sex	male	60	51	56	42	82	<0.001
Prior vacations	Never	2	6	23	10	2	<0.001
	Once	4	2	9	5	5	
	twice and more	94	92	68	85	93	
Intention to revisit	Certainly	29	34	26	25	34	0.005
	Probably	25	23	24	56	38	
	Probably not	15	11	25	13	14	
	Certainly not	32	32	25	6	14	
Accommodation	Hotel	58	62	83	90	33	<0.001
	Bed and Breakfast	10	12	10	5	27	
	Apartment	13	12	5	5	16	
	Private room	13	10	1	0	1	
	Private room (farm)	3	5	1	0	23	
Company	Partner	66	75	85	100	55	<0.001
	Family	42	22	26	42	36	<0.001
	Friends	48	20	7	5	82	<0.001
	Travel group	6	5	15	5	17	<0.001
Info sources	Brochures	11	7	11	16	2	
	media ads	1	1	12	0	1	<0.001
	tour operator	7	3	20	11	11	<0.001
	Friends & relatives	26	25	28	16	42	0.001
	travel agent	9	9	20	21	11	0.012
	tourism bureau	11	6	3	5	5	0.004
	Internet	3	4	0	0	1	0.066
	no info needed	41	41	23	47	10	<0.001

⁷ The Kruskal-Wallis Test was applied for the metric and ordinal data, the Chi-square test was computed for nominal variables.

Figure 6: Stability comparison (activity data)

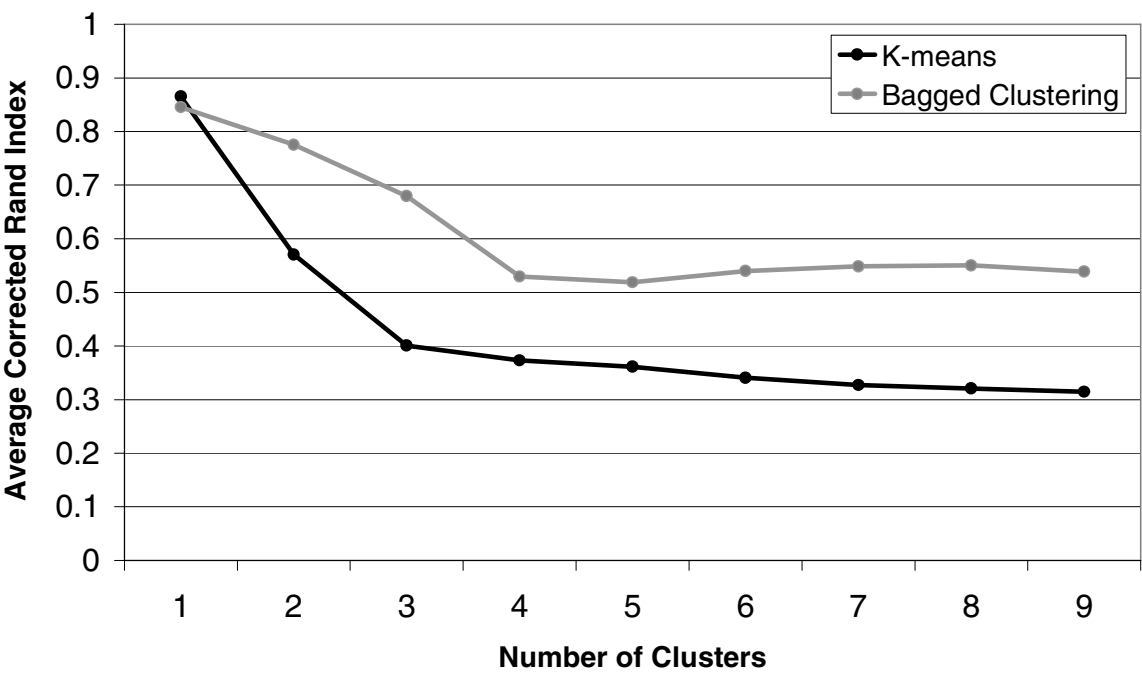


Figure 7: Stability comparison (motive data)

