



Exploring interdependencies in students' vacation portfolios using association rules

Anna B. Grigolon ^{1*}, Astrid D. A. M. Kemperman ² and Harry J. P. Timmermans ³

Received: 16/02/2012 Accepted: 20/04/2012

¹ Department of the Built Environment - Urban Planning Group, Eindhoven University of Technology, P.O. Box 513 – 5600 MB Eindhoven, The Netherlands, tel: +31 40 247 2861, fax: +31 40 243 8488, e-mail: a.b.grigolon@tue.nl

² Department of the Built Environment - Urban Planning Group, Eindhoven University of Technology, P.O. Box 513 – 5600 MB Eindhoven, The Netherlands, tel: +31 40 247 3291, fax: +31 40 243 8488, e-mail: a.d.a.m.kemperman@tue.nl

³ Department of the Built Environment - Urban Planning Group, Eindhoven University of Technology, P.O. Box 513 – 5600 MB Eindhoven, The Netherlands, tel: +31 40 247 2274, fax: +31 40 243 8488, e-mail: h.j.p.timmermans@tue.nl

* Corresponding author

Abstract

The transition into adulthood is usually associated with changes and events, accompanied by developments in young people's social status and consequently, some changes in travel and transport use. This makes the students' segment highly relevant for both marketing companies and policy-makers. The goal of this paper is to explore the overall vacation portfolio of decisions of students. Data about vacation history in terms of the long holidays of a sample of 211 students was used. To avoid any limiting a priori statistical assumptions in the analysis, the existence of any interdependencies in the vacation portfolios and their covariates are explored in this study using Association Rules Mining. Results indicate that, in addition to expected associations between socio-demographics characteristics, student's vacation choices are very heterogeneous.

© 2012 International University College. All rights reserved

Keywords: vacation portfolio, travel behaviour, students, association rules

Citation: Grigolon, A., A. Kemperman, H. Timmermans (2012) Exploring interdependencies in students' vacation portfolios using association rules. *European Journal of Tourism Research* 5(2), pp. 93-105

Introduction

The students segment represents an interesting and highly relevant subject of study in vacation marketing and tourism research. Over a relative short period of time, adolescents and young adults may experience

rapid changes in the contextual variables influencing their vacation decisions. They start as a family member and will likely try to influence their parents in their vacation choice. Then, at some age, they become increasingly reluctant to travel with their parents and prefer

to experience a vacation on their own or with their friends. At this stage, many will still have a limited budget, but may have a lot of time. Then, gradually, vacation partners will shift to spouses and increasing budgets will allow this segment of the population to choose from a wider set of travel options.

The study of the determinants of student's vacation choices is of both marketing and policy interest in the sense that the transition into adulthood is usually associated with changes and events, accompanied by developments in young people's social status and consequently, some changes in travel and transport use. Besides, semester breaks and other holidays gives them relatively large time blocks free from school commitments and/or work (Babin & Kim, 2001), making them quite time-flexible. Moreover, they may be more experienced tourists at an earlier age compared to the students of previous generations (Gibson & Yiannakis, 2002). Despite the relevance of the student's segment, relatively little is known about their vacation behaviour. The existent literature is focused on specific facets that make part of the vacation portfolio of decisions such as destination choice (Ross, 1993; Sirakaya, Sonmz and Choi, 2000; Michael, Armstrong and King, 2004), travel motivations (Josiam, Smeaton and Clements, 1999; Kim and Jogaratnam, 2003), accommodations (Murphy and Pearce, 1995). Other studies explore the vacation planning process of students (e.g Sung and Hsu, 1996) while others were interested in determining differences in behaviour given gender (Carr, 1998, 1999, 2001, 2002a, 2002b, 2002c, 2005) or nationality (Shoham, Schrage and van Eeden, 2004).

The contribution of our study is on the better understanding of the overall vacation portfolio of decisions, controlling for socio-demographic characteristics. The vacation portfolio in this context include the combined choice of destination, travel party, transport mode, length of stay, season and accommodation. To that effect, the vacation history of 211 students living in the Eindhoven region (The Netherlands) was explored. To avoid any limiting a priori statistical assumptions in the

analysis, the existence of any interdependencies between socio-demographic factors and the vacation portfolios are explored in this study using Association Rules Mining. This technique is derived from computer science and was first developed to solve various aspects of the database mining problem, such as discovering association rules and enhancing the database capability with classification queries over large sequences (Agrawal, Imielinski and Swami, 1993).

This paper is structured as follows. First the Association Rules Mining technique will be outlined. Next, the data collection and sample are described. This is followed by a description of the results of the three analyses that were performed. Finally we discuss the main conclusions of this study.

Association Rules Mining

Association rules represent a useful technique to find associations between variables in (large) databases. This method was first introduced by Agrawal, Imielinski and Swami (1993), who defined association rules for binary choices based on single-level variables.

In the context of the present paper however the interest is in finding associations between the multi-category attributes (for example destination and transport mode) that are part of the vacation profile of students. This implies the choice of multi-category association rule mining (Han and Fu, 1999). In the current study the data provided has variables with multiple attribute levels.

Assume given a database V consisting of a set of vacation profiles $\langle V_i, \{L_p, \dots, L_q\} \rangle$, where

V_i is a vacation profile identifier. Each vacation profile consists of a description of each attribute A with L_A levels in the form of $\langle L_i, description \rangle$, where $A_i \in I$ (for $i = p, \dots, q$). A rule has the form $X \Rightarrow Y$ (if X , then Y), where $X, Y \subseteq A$ and $X \cap Y = \emptyset$.

The attribute sets X and Y are called *Body* (left-hand side or LHS) and *Head* (right-hand side or RHS) of the rule, respectively.

Table 1. A vacation profile table

Vacation profile	Attribute set
1	{prof1_att1, prof1_att2, prof1_att3, ...}
2	{prof2_att1, prof2_att2, prof2_att3, ...}
...	{..., ..., ...}

Table 2. A vacation profile example

Vacation profile	Transport mode	Accommodation	Destination	Travel party
1	Plane	Hotel	London	Family
2	Car	Camping	Marseille	Friends
3	Plane	Hotel	London	Alone
4	Plane	Hotel	Paris	Partner

To illustrate the concept, consider the next small trip-related example. The problem is to find multiple-level strong associations in the database of Table 1 for respondents who provided information about their vacation trips.

The trip {plane, hotel} ⇒ {London} could be an example of a rule. The best-known constraints on various measures of significance and interest that express the degree of uncertainty about the rule are called support and confidence. The support $sup(X)$ of an item set

X is defined as the proportion of profiles in the data set which contain the itemset. In Table 2, {plane, hotel, London} has a support of $2/4 = 0.5$ since it occurs in 50% of all profiles (2 out of 4 profiles). The confidence of a rule is defined

$conf(X \Rightarrow Y) = sup(X \cup Y) / sup(X)$. In the example, the rule {plane, hotel} ⇒ {London} has a confidence of $0.5/0.75$ in the database, which means that for 67% of the trips containing plane and hotel the rule is correct. Note that this example is extremely small and in practical applications a rule needs a support of several profiles before it can be considered statistically significant.

In the process of finding association rules, one must define the minimum support and confidence levels for a given database: if the minimum support and confidence thresholds are too low, it may lead to the generation of many uninteresting associations, whereas if the support and confidence levels are set too high, there are not enough associations to be

interpreted. Consequently, the settings of these minimum values are meant to find those items whose occurrences exceed the predefined minimum support and confidence levels.

Kotsiantis and Kanellopoulos (2006) provided a recent overview of the major theoretical issues with the concept of association rule mining in the field of computer sciences and engineering. Similarly, Law, Mok and Goh (2007) provided an overview of data mining techniques applied to tourism demand forecasting, but they did not include association rules. Koh and Gervais (2006) applied association rules to investigate the best tour options in China. Zhou *et al.* (2008) developed a semantic association rule mining algorithm based on word categories that appear in tourism emergency reports. Danubianu, Socaciu and Adina (2009) used data mining methods on data regarding developments of the economy aimed for by the tourism industry. Liao, Chen and Deng (2010) used association rules for mining customer knowledge of new tourism product development and customer relationship management in Taiwan. This overview indicates that the scant application of association rules in tourism has been targeted at topics other than their vacation behavior in general and vacation portfolios in particular.

Data and Sample

In order to explore interdependencies between socio-demographics variables, vacation portfolio data on vacation profiles were collected for a sample of 211 young adults, mostly students. Invitations were sent to

students from a sample of schools and universities in the province of Brabant, The Netherlands. In addition, invitation cards were personally given to young adults at train stations and city centers. They were briefly explained about the aims of the project and if interested guided to a web-site for the survey and an experiment. The goal of the experiment was to collect data on the choice of vacation portfolio, i.e. the combination of destination, transport mode, season choice, accommodation, and travel party. Five 100 euro's travel vouchers were raffled among participants.

Table 3. *Socio-demographic variables*

Variables	Levels	%
Gender	female	50.2
	male	49.8
Age	14 ÷ 18	10.1
	19 ÷ 24	55.0
	25 ÷ 30	30.3
	> 31	4.6
Education level	low	0.7
	medium	7.8
	high	91.5
Income	< €1000	75.4
	€1001 ÷ €2000	21.3
	> €2001	3.3
Number of people in household	1	23.3
	2	28.2
	3 or more	48.5
Drivers license possession	yes	77.0
	no	23.0
Car possession	yes	21.0
	no	79.0
Number of cars in the household	0	40.7
	1	30.8
	more than 2	28.5
Hours spending studying per week	Not studying	16.7
	1 ÷ 10	8.3
	11 ÷ 25	7.2
	26 ÷ 40	52.7
	> 41	15.2
Hours spending working per week	Not working	22.5
	1 ÷ 10	38.4
	11 ÷ 25	13.1
	26 ÷ 40	21.1
	> 41	4.9
Nationality (by continent)	Dutch	71.6
	Europe	18.6
	Asia	3.7
	Middle East	3.4
	South America	2.7

N=211 respondents

The composition of the sample is presented in Table 3. It shows an almost equal mix of women and men. Slightly more than fifty percent is in the 19-24 age group, while another thirty two percent is between 25-30 years of age. Respondents belong to households mostly consisting of three or more people. The largest group of students has a high education level and more than seventy five percent has less than €1000 per month to spend on their activities including leisure. Most respondents have a driver's license, but almost eighty percent does not possess a car, whereas more than fifty five percent has 1 or more cars available in the household. Most respondents have Dutch or European nationalities. The majority spends from 26 to 40 hours per week on studying, does not work, or works up to 10 hours per week.

Respondents were asked to list up to 20 different destinations they visited for a long holiday (at least four consecutive nights spent away from home) over the past 10 years. A total of 2011 trips were recorded, representing an average of 9.53 trips per respondent (almost 1 trip per year). Respondents were asked for each trip to specify the name of the visited destination and other travel-related facets (travel mode, accommodation, length of stay, travel party and season) by means of drop-down lists. The destinations were classified according to continent, type of trip and geographical location of the destination related to the nationality of the respondent. Table 4 shows the distribution of choices between the vacation portfolios and their levels.

Theoretical framework

Using the database of 2011 trips, the following association rules were explored:

- Analysis 1: Finding association rules between the socio-demographic variables
- Analysis 2: Finding association rules between the vacation portfolios variables
- Analysis 3: Finding association rules between both socio-demographic variables and vacation portfolios variables.

Table 4. *Vacation portfolios*

Variables	Levels	%
Travel mode	Car	37.6
	Regular airlines	31.5
	Discount airlines	13.3
	Bus	11.6
	Train	6.0
Accommodation	Hotel	35.7
	Camping	21.3
	Rented apartment	19.1
	Friend's/relative's house	12.7
	Hostel	9.3
	Boat	1.8
Length of stay	More 4 nights less 1 week	35.2
	Between 1 and 2 weeks	37.6
	Between 2 and 3 weeks	16.6
	Between 3 and 4 weeks	4.6
	More than 4 weeks	6.0
Travel Party	Family	46.4
	Friend(s)	27.6
	Partner	16.7
	Alone	9.2
Season	Summer	61.9
	Winter	16.5
	Spring	13.1
	Autumn	8.6
Destination - Continent	Europe	74.3
	Netherlands	6.3
	Middle East	5.5
	Asia	4.6
	North America	3.5
	South or Central America	2.8
	Africa	2.2
	Oceania	0.8
Destination - Type	Culture	63.3
	Beach	15.6
	Culture/Beach	9.6
	Domestic	6.0
	Sports	1.9
	Entertainment	1.9
	Nature	1.7
Destination - Geography x Nationality	Same continent	44.8
	Adjacent country	22.1
	Different country and continent	20.6
	Same country	12.5

N=2011 trips

The split in three different analyses is meant to find the association rules that are able to identify, first, the strongest associations between the variables making up the profiles of the respondents, second, the vacation portfolio variables that are most likely to be chosen and,

third, a combination of these, which is the exploration of the portfolio vacation variables that are most likely to be chosen depending on the profile of the respondents.

Analysis and Results

The results of the analyses are shown in terms of tables and 2D and 3D graphics. The tables express the support, confidence and correlation in percentages. The 2D graphical representation of the support values for the Body and Head portions of each association rule are indicated by the sizes and colours/shades of the circles. The thickness of each line indicates the confidence value; the sizes and colours/shades of the circles in the center, above the Implies label, indicate the joint support (for the co-occurrences) of the respective Body and Head. In the 3D representation, the confidence value is added and plotted in the vertical axis of the graphic.

Analysis 1: Association rules between the socio-demographic variables

In the first analysis the following socio-demographic variables were input to the algorithm finding association rules: gender, household size, possession of driver's license, car possession, number of cars in the household, age, age when the trip occurred (calculated with the informed year of the trip), income, hours per week expended studying, hours per week expended working, education and nationality. Satisfactory and interpretable association rules were found for 60% minimum support, confidence and correlation. Figure 1 shows the 2D graphical representation of the association rules found. It shows that the strongest support value is "education=high", associated with "nationality by category=Dutch", "car=no", "driver license=yes" and "income <=1000 euros". Thus, these association rules pick up the typical profile of the sample. The thicker line expresses the highest confidence value, found for the association between "driver license=yes" and "education=high".

It is also interesting to analyse the results of the 3D graphical representation, illustrated in Figure 2. It shows the same associations as Figure 1, but the axis z contains in addition the confidence levels of the association rules. It is

Exploring interdependencies in students' vacation portfolios using association rules.

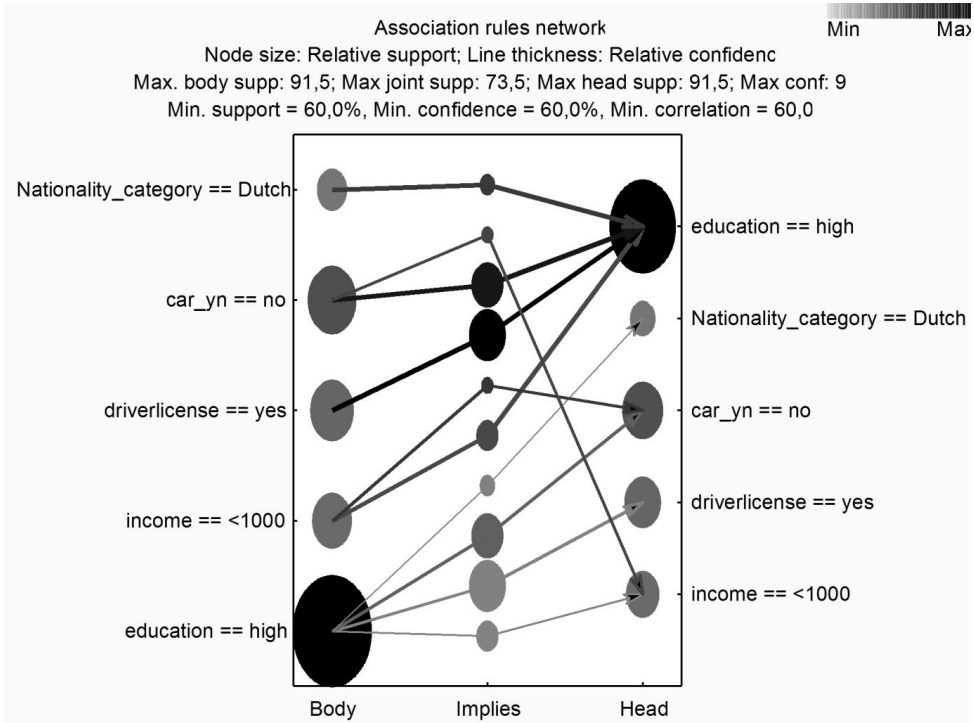


Figure 1. Association rules 2D graphical representation – socio-demographic variables only

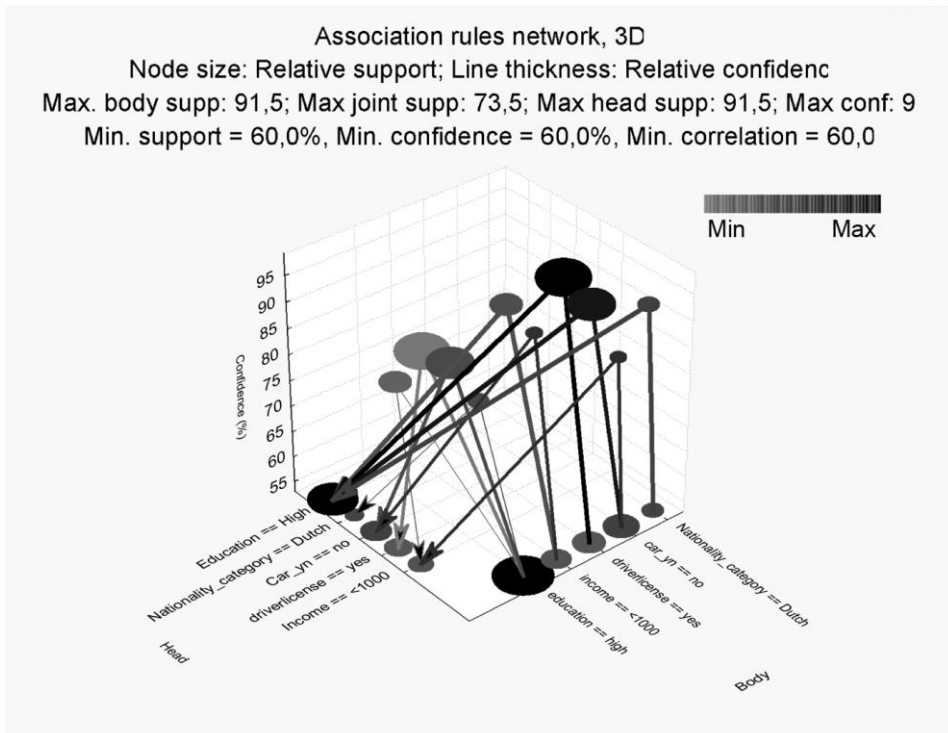


Figure 2. Association rules 3D graphical representation – socio-demographic variables only

Table 5. Association rules tabular representation – socio-demographic variables only

Summary of association rules - Analysis 1
 Min. support = 60.0%, Min. confidence = 60.0%, Min. correlation = 60.0% Max. size of body = 10, Max. size of head = 10

Body	Head	Support (%)	Confidence (%)	Correlation (%)
driver license = yes	education = high	73.55	95.48	87.61
education = high	driver license = yes	73.55	80.38	87.61
car_yn = no	education = high	71.71	90.75	84.33
education = high	car_yn = no	71.71	78.37	84.33
income = <1000	education = high	67.88	89.98	81.70
education = high	income = <1000	67.88	74.18	81.70
nationality = Dutch	education = high	65.29	91.24	80.69
education = high	nationality = Dutch	65.29	71.36	80.69
income = <1000	car_yn = no	64.05	84.90	82.96
car_yn = no	income = <1000	64.05	81.06	82.96

possible to see that the rule “driver license=yes” and “education=high” has the highest confidence level, as in the 2D representation, which can be also checked in the tabular representation showed in Table 5 (95.48%).

To summarise, the results of the association between the socio-demographic variables only shows that high education, Dutch nationality, income less than 1000 euros per month, no car and possession of a driver’s license are the variables with the strongest association rules in the sample. The most reliable rule, expressed by the highest confidence level, is the association between having a driver’s license and a high education.

Analysis 2: Association rules between the vacation portfolios variables

The following vacation portfolios variables were explored in the second analysis in order to find association rules: travel mode, accommodation, length of stay, travel party, season, destination by continent, destination by type and destination by combined geographical location and nationality. The minimum support, confidence and correlation values were set at 32%.

Figure 3 shows the 2D graphical representation of the association rules of analysis 2. The strongest support value was found for “destination by continent=Europe”, which was associated with “destination by geographical location and nationality=same continent”, “season=summer”, “travel party=family”, “travel

mode=car” and “destination by type=culture”. The highest confidence level is the association between “destination by geographical location and nationality=same continent” and “destination by continent=Europe”, which can also be seen in the 3D representation (Figure 4) and the table representation (97.44%, on Table 6).

The association rules found for the inclusion of vacation portfolio variables shows that European nationality, destination within Europe, summer holidays, with family, by car and performing cultural activities during the holiday are the strongest associated variables in the sample. Based on these association rules, profiles of most common vacation profiles are identified. The most reliable rule found, expressed by the highest confidence level, was for people with European nationality spending their holiday trips within Europe.

Analysis 3: Association rules between socio-demographic and vacation portfolios variables

The third analysis includes the socio-demographic and vacation portfolio variables together. The minimum support, confidence and correlation were defined at 59%, leading to a satisfactory number of interpretable association rules. Note that these association rules pick up the strongest associations between particular facets of vacation choices and socio-demographics. It will be evident that the results are important for vacation marketing purposes as it allows marketers and planners to identify the most responsive segments for particular vacation facets.

Exploring interdependencies in students' vacation portfolios using association rules.

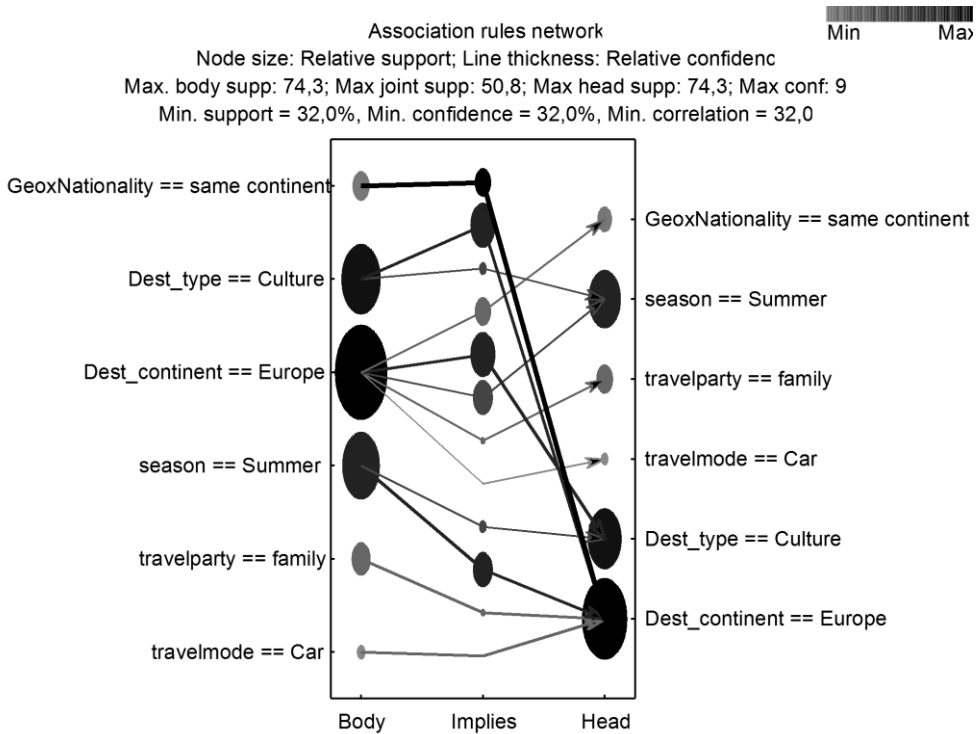


Figure 3. Association rules 2D graphical representation – vacation portfolio variables only

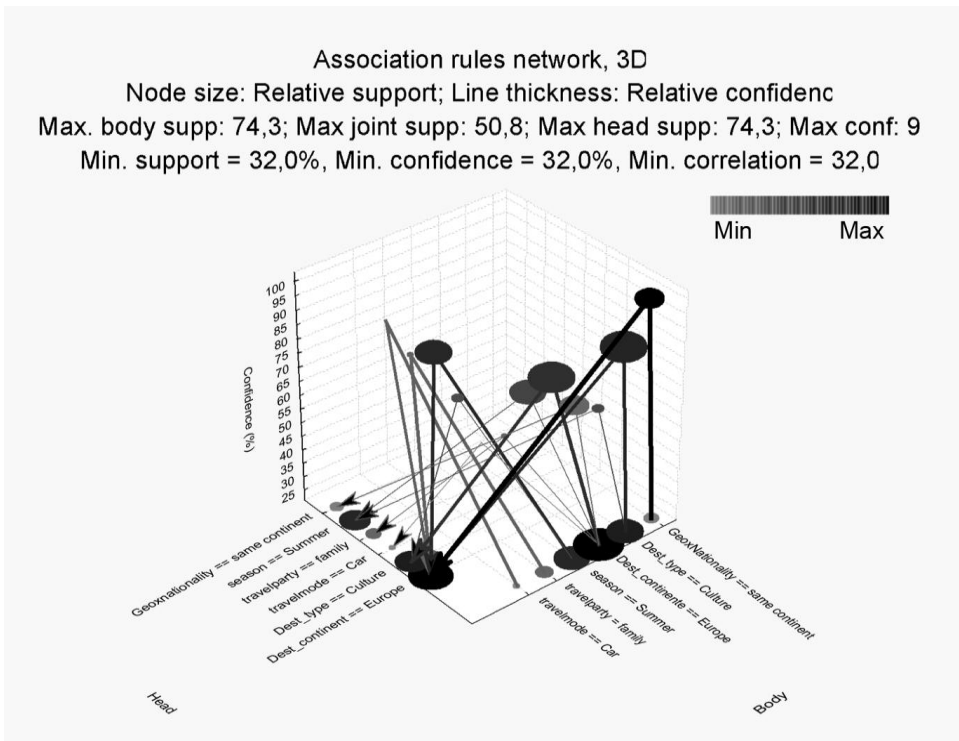


Figure 4. Association rules 3D graphical representation – vacation portfolio variables only

Table 6. Association rules tabular representation – vacation portfolio variables only

Summary of association rules - Analysis 2				
Min. support = 32.0%, Min. confidence = 32.0%, Min. correlation = 32.0% Max. size of body = 10, Max. size of head = 10				
Body	Head	Support (%)	Confidence (%)	Correlation (%)
Dest_type = Culture	Dest_continent = Europe	50.77	80.27	74.04
Dest_continent = Europe	Dest_type = Culture	50.77	68.29	74.04
season = Summer	Dest_continent = Europe	46.54	75.18	68.61
Dest_continent = Europe	season = Summer	46.54	62.61	68.61
GeoxNationality = same continent	Dest_continent = Europe	43.61	97.44	75.61
Dest_continent = Europe	GeoxNationality = same continent	43.61	58.66	75.61
season = Summer	Dest_type = Culture	36.70	59.28	58.64
Dest_type = Culture	season = Summer	36.70	58.02	58.64
travelparty = family	Dest_continent = Europe	34.36	73.98	58.48
Dest_continent = Europe	travelparty = family	34.36	46.22	58.48
travelmode = Car	Dest_continent = Europe	32.17	85.47	60.82
Dest_continent = Europe	travelmode = Car	32.17	43.28	60.82

In the analysis of the association rules in Figure 5, the strongest support value was found for “education=high”, which was associated with “destination by type=culture”, “destination by continent=Europe”, “nationality by category=Dutch”, “car=no”, “driver license=yes” and “income=<1000”. The highest confidence level is the association between “driver license=yes” and “education=high”, which can also be seen in the 3D representation (Figure 6) and the table representation (Table 7), with the value of 95.48%. It evidences that because association rules examine all possible associations, these association rules between socio-demographics only are redundant if the goal is more specific to find associations between vacation choice facets and socio-demographics.

Summarising the findings of the third analysis, which included both socio-demographic and vacation portfolio variables for the responses of the present sample: having a high education was associated with Dutch nationality, no car, income less than 1000 euros, possession of a driver’s license, holidays within Europe and performing cultural activities. The most reliable association found was, like in the first analysis, the possession of a driver’s license and a higher education.

Conclusions and Discussion

The goal of this paper was to explore vacation portfolio decisions of a sample of students, as part of a larger study on vacation planning of this segment of the population. Revealed data

about vacation history in terms of the long holidays of a sample of 211 students was used. The vacation portfolio variables included joint combinations of destination, transport mode, accommodation type, length of stay, travel party and season. Association rules mining was used to find associations between (a) the socio-demographic variables only, (b) the various facets of the vacation portfolio only and (c) between the socio-demographic variables and the vacation portfolio choice facets.

Results of the first analysis found interdependencies between having a higher education and Dutch nationality, income less than 1000 euros, possession of a driver’s license and no car ownership. These patterns reflect the composition of the sample.

The second analysis included the vacation portfolio variables only, which were: destination (divided in three categories), transport mode, length of stay, travel party, season, and accommodation. Association rules were found between season and the three categories of destination. This is an indication that although the sample is quite homogeneous by means of its socio-demographics, it reflects a substantial amount of heterogeneity in taste preferences regarding their vacation decisions. The interdependencies found suggest that the holiday destination of a European student is associated with choosing a destination in Europe where cultural activities will be conducted, during the summer.

Exploring interdependencies in students' vacation portfolios using association rules.

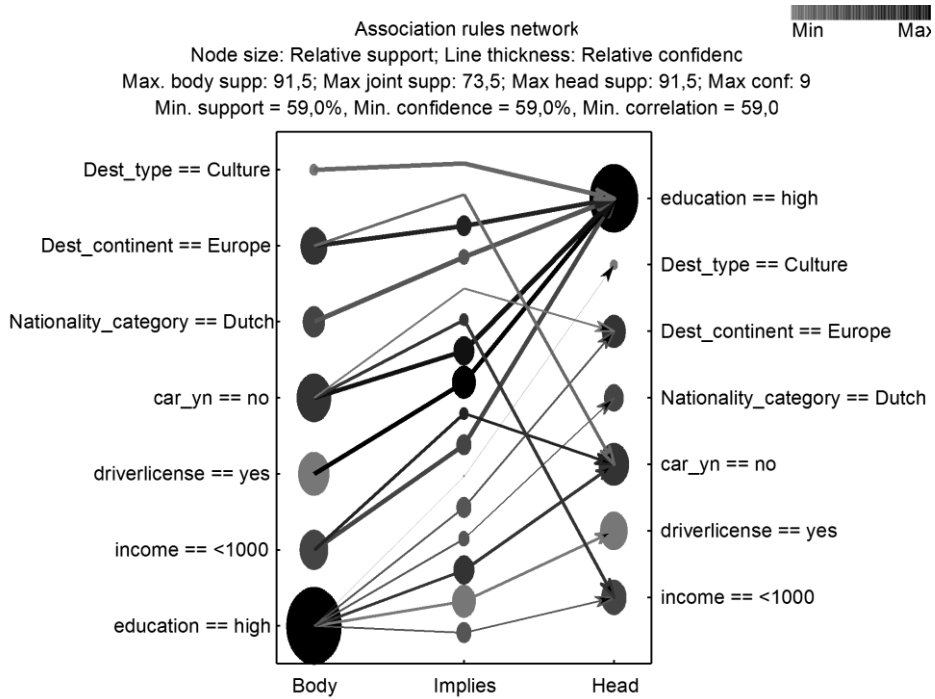


Figure 5. Association rules 2D graphical representation– socio-demographic variables and vacation portfolios variables

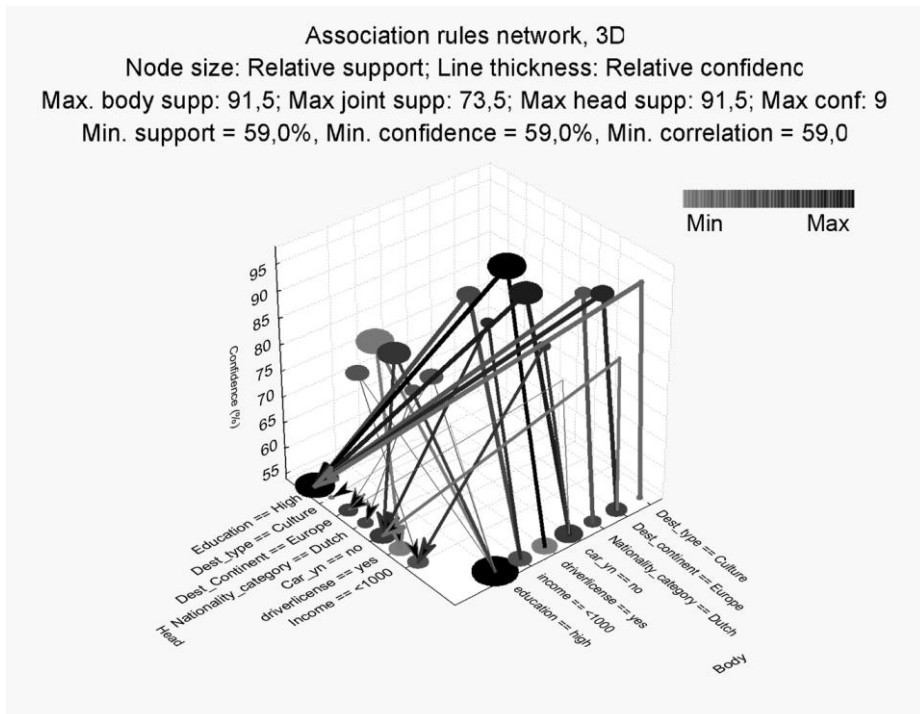


Figure 6. Association rules 3D graphical representation – socio-demographic variables and vacation portfolios variables

Table 7. Association rules tabular representation – socio-demographic variables and vacation portfolios variables

Summary of association rules - Analysis 3
 Min. support = 59.0%, Min. confidence = 59.0%, Min. correlation = 59.0% Max. size of body = 10,
 Max. size of head = 10

Body	Head	Support (%)	Confidence (%)	Correlation (%)
driver license = yes	education = high	73.55	95.48	87.61
education = high	driver license = yes	73.55	80.38	87.61
car_yn = no	education = high	71.71	90.75	84.33
education = high	car_yn = no	71.71	78.37	84.33
Dest_continent = Europe	education = high	67.88	91.30	82.30
income = <1000	education = high	67.88	89.98	81.70
education = high	Dest_continent = Europe	67.88	74.18	82.30
education = high	income = <1000	67.88	74.18	81.70
nationality = Dutch	education = high	65.29	91.24	80.69
education = high	nationality = Dutch	65.29	71.36	80.69
income = <1000	car_yn = no	64.05	84.90	82.96
car_yn = no	income = <1000	64.05	81.06	82.96
Dest_type = Culture	education = high	59.22	93.63	77.85
education = high	Dest_type = Culture	59.22	64.73	77.85
Dest_continent = Europe	car_yn = no	59.03	79.40	77.01
car_yn = no	Dest_continent = Europe	59.03	74.70	77.01

Vacation portfolios and socio-demographic variables were included in the third analysis. This analysis led to the conclusion that the strongest association rules were found between the same socio-demographic variables of the first analysis (higher education, Dutch nationality, income less than 1000 euros, possession of a driver's license and no possession of a car) and only two variables of the vacation portfolio (holiday destination within Europe and cultural activities mostly performed during the holiday). This finding confirms that students of this sample were very heterogeneous regarding their vacation preferences.

The main advantage of using Association Rules Mining in this context was the possibility to explore the associations between sample composition and vacation choices separately, and jointly. This allows the exploration of different patterns in the data. Therefore it is highly relevant for other applications, combining commercial interest with intriguing research questions (Agrawal, Imielinski and Swami, 1993). A limitation of the study may be related to the set-up values of support, confidence and correlation values. To get interpretable results the minimum values have to be set in order to provide visual and understandable outputs. This decreases the number of interdependencies that could be

found, in other words, smaller minimum set-up values would provide more association rules in the data.

It is important to highlight that although the usage of data mining techniques provided the analysis of patterns in the data, the analysis of the behaviour should still be allied to methods that are not vulnerable to collinearity because of unknown interrelations. To address this sort of issue, the analysis of the vacation behaviour of students was completed with a portfolio experimental design, addressed in another paper (Grigolon, Kemperman and Timmermans, 2012).

Acknowledgment:

The study was conducted as part of the project “The value of recreation: Now, and in a completely different future”, which is part of the DBR (Duurzame Bereikbaarheid van de Randstad - Sustainable Accessibility of the Randstad) programme. It was financially supported by the Netherlands Organization for Scientific Research (NWO).

References

Agrawal, R., Imielinski, T. & Swami, A. N. (1993). *Mining association rules between sets of items in large databases*. Paper presented at ACM SIGMOD International Conference on

- Management of Data. Washington D.C., ACM Press, 207-216.
- Carr, N. (1998). The young tourist: a case of neglected research. *Progress in Tourism and Hospitality Research*, 4(4), 307-318.
- Carr, N. (1999). A study of gender differences: young tourist behaviour in a UK coastal resort. *Tourism Management*, 20(2), 223-228.
- Carr, N. (2001). An exploratory study of gendered differences in young tourists perception of danger within London. *Tourism Management*, 22(5), 565-570.
- Carr, N. (2002a). A comparative analysis of the behaviour of domestic and international young tourists. *Tourism Management*, 23(3), 321-325.
- Carr, N. (2002b). Defining young tourists visiting beach-oriented resorts: a behavioural analysis. *Anatolia: An International Journal of Tourism and Hospitality Research*, 13(1), 49-62.
- Carr, N. (2002c). Going with the flow: an assessment of the relationship between young people's leisure and holiday behaviour. *Tourism Geographies*, 4(2), 115-134.
- Carr, N. (2005). Poverty, debt, and conspicuous consumption: university students' tourism experiences. *Tourism Management*, 26(5), 797-806.
- Danubianu, M., Socaciu, T. & Adina, B. (2009). Toward a distributed data mining system for tourism industry. *Annals of Faculty of Economics*, 4(1), 922-926.
- Gallarza, M. G. & Saura, I. G. (2006). Value dimensions, perceived value, satisfaction and loyalty: an investigation of university students' travel behaviour. *Tourism Management*, 27(3), 437-452.
- Grigolon, A. B., Kemperman, A. D. A. M. & Timmermans, H. J. P. (2012). The influence of low-fare airlines on vacation choices of students: Results of a stated portfolio choice experiment. *Tourism Management*, 33(5), 1174-1184.
- Han, J. & Fu, Y. (1999). Mining multiple-level association rules in large databases. *IEEE Transactions on Knowledge and Data Engineering*, 11(5), 798-805.
- Josiam, B., Smeaton, G. & Clements, C. J. (1999). Involvement: travel motivation and destination selection. *Journal of Vacation Marketing*, 5(2), 167-175.
- Kim, K. Y. & Jogaratnam, G. (2003). Travel motivations: a comparative study of Asian and domestic American college students. *Journal of Travel & Tourism Marketing*, 13(4), 61-82.
- Kim, K. Y., Jogaratnam, G. & Noh, J. (2006). Travel decisions of students at a US university: Segmenting the international market. *Journal of Vacation Marketing*, 12(4), 345-357.
- Koh, H. C. & Gervais, G. (2006). *Using data mining for travel and tourism research*. Paper presented at the International Conference on Business and Information, Singapore.
- Kotsiantis, S. & Kanellopoulos, D. (2006). Association rules mining: a recent overview. *International Transactions on Computer Science and Engineering*, 32(1), 71-82.
- Law, R., Mok, H. & Goh, C. (2007). *Data mining in tourism demand analysis: a retrospective analysis*. In: Proceedings of the 3rd international conference on advanced data mining and applications, Harbin, China.
- Liao, S., Chen, Y. & Deng, M. (2010). Mining customer knowledge for tourism new product development and customer relationship management. *Expert Systems with Applications*, 37(6), 4212-4223.
- Michael, I., Armstrong, A. & King, B. (2004). The travel behaviour of international students: The relationship between studying abroad and their choice of tourist destinations. *Journal of Vacation Marketing*, 10(1), 57-66.
- Murphy, L. L. & Pearce, P. L. (1995). Young budget travellers: backpackers in Australia. *Annals of Tourism Research*, 22(4), 819-843.
- Pizam, A., Jeong, G. H., Reichel, A., Boemmel, H., Lusson, J. M., Steynberg, L., Costache, O. S., Volo, S., Kroesbacher, C., Kucerova, J. & Montmany, N. (2004). The relationship between risk-taking, sensation-seeking, and the tourist behavior of young

- adults: a cross-cultural study. *Journal of Travel Research*, 42(3), 251-260.
- Ross, G. F. (1993). Destination evaluation and vacation preferences. *Annals of Tourism Research*, 20(3), 477-489.
- Siracaya, E., Sonmez, S. F. & Choi, H. S. (2000). Do destination images really matter? Predicting destination choices of student travellers. *Journal of Vacation Marketing* 7 (2), 125-142.
- Shoham, A., Schrage, C. & van Eeden, S. (2004). Student travel behavior, *Journal of Travel & Tourism Marketing*, 17(4), 1-10.
- STATISTICA (2011). *Electronic Statistics Textbook*. StatSoft Inc., Tulsa, USA. <http://www.statsoft.com/textbook>.
- Sung, S. & Hsu, C. H. C. (1996) International students' travel characteristics: an exploratory study. *Journal of Travel and Tourism Marketing*, 5(3), 277-283.
- Zhou, Y., Du, J., Zeng, G. & Tu, X. (2008). *Association rule mining based on the semantic categories of tourism information*. Paper presented at the 5th international symposium on Neural Networks: Advances in Neural Networks, Part II, Beijing, China.
-