

Mirosława Lasek

Katedra Informatyki Gospodarczej i Analiz Ekonomicznych, Uniwersytet Warszawski
00-241 Warszawa, ul. Długa 44/50, mlasek@wne.uw.edu.pl;

Ewa Nowak

Zakład Geografii Regionalnej, Instytut Geografii, Akademia Świętokrzyska
25-410 Kielce, ul. Świętokrzyska 15, ewa.nowak@pu.kielce.pl

Marek Pęczkowski

Katedra Informatyki Gospodarczej i Analiz Ekonomicznych, Uniwersytet Warszawski
00-241 Warszawa, ul. Długa 44/50, mpechkowski@wne.uw.edu.pl

**ZASTOSOWANIE REGUŁ ASOCJACJI
I SEKWENCJI ZDARZEŃ
DO ANALIZY DZIAŁALNOŚCI
INWESTYCYJNEJ GOSPODARSTW
AGROTURYSTYCZNYCH**

**ASSOCIATION AND SEQUENCE
RULES OF EVENTS
IN AN INVESTMENT
ANALYSIS OF
AGROTOURISM FARMS**

Zarys treści: W artykule przedstawiono próbę przewidywania kierunków działań rozwojowych gospodarstwa agroturystycznego. Zastosowano technikę znaną w marketingu jako analiza koszykowa. Analiza dotyczy współzależności przedsięwzięć podejmowanych przez poszczególne gospodarstwa agroturystyczne. Dzięki odkryciu współzależności między cechami, możemy tworzyć reguły postaci: jeżeli cecha A towarzyszy określonymu zdarzeniu, to możemy się spodziewać, że cecha B będzie także towarzyszyć temu zdarzeniu. W naszym przypadku będziemy rozważać, jakie inwestycje gospodarstwa agroturystyczne podejmują łącznie oraz jaka jest kolejność podejmowanych inwestycji – posłuży nam do tego analiza asocjacji oraz analiza sekwencji, przeprowadzane przy wykorzystaniu programu Enterprise Miner firmy SAS (2003). Analiza asocjacji pozwoli sformułować wnioski co do tego, jakie przedsięwzięcia poszczególne gospodarstwa realizują jednocześnie, a analiza sekwencji odnaleźć przedsięwzięcia realizowane przez gospodarstwa kolejno w określonych odstępach czasu. Analizę przeprowadzono na przykładzie gospodarstw agroturystycznych w regionie świętokrzyskim.

Słowa kluczowe: agroturystyka, gospodarstwo agroturystyczne, analiza koszykowa, analiza asocjacji, analiza sekwencji.

Abstract: The article presents an attempt to predict the course of development activity on agrotourism farms. A well-known marketing technique, *market basket analysis*, has been implemented which concerns the interdependencies of investments undertaken by individual agrotourism farms. Due to the discovery of interdependencies between different features, rules can be created: if feature A accompanies a particular event it may be expected that feature B will accompany this event as well. In this case, investments made jointly by agrotourism farms will be considered in the order in which they are made. Association analysis and sequence analysis were conducted with the SAS Enterprise Miner programme (2003). Association analysis will help find out what investments have been implemented at individual farms simultaneously, while sequence analysis will allow identification of investments consecutively implemented by farms measured at specific intervals. The analysis has been conducted using the example of the Świętokrzyskie Województwo.

Key words: agrotourism, agrotourism farm, *market basket analysis*, association analysis, sequence analysis.

1. WSTĘP

Agroturystyka, traktowana jako działalność gospodarcza człowieka, opisywana i badana jest przez takie dziedziny nauki, jak: geografia, ekonomia, ekologia, prawo.

Rozwój przedsiębiorczości agroturystycznej może być wspomagany badaniami w aspekcie eko-

1. INTRODUCTION

Agrotourism, treated as an economic activity, has been described and examined by disciplines such as geography, economics, ecology and law. The development of agrotourism entrepreneurship can be supported by research in aspects of economics,

nomiki, organizacji i zarządzania oraz marketingu. Działalność agroturystyczna, jak każda działalność gospodarcza, wymaga inwestycji w różne przedsięwzięcia. Działalność agroturystyczną można analizować metodami ekonomicznymi, takimi samymi jak te, które stosuje się do analizy działalności przedsiębiorstw.

Badania marketingowe to zaplanowane systematyczne zbieranie i analiza danych mające pomóc podmiotom działającym na rynku podjąć decyzje i monitorować rezultaty tych decyzji. Badania marketingowe mają na celu pomóc zrozumieć naturę rynku. Badania mogą być opisowe, tzn. pomagać odnaleźć faktyczne dane o tym, co się dzieje na rynku, lub też mogą być analityczne, usiłujące wyjaśnić relacje współzależności między zmiennymi, np. zrozumieć, dlaczego zachodzą dane zjawiska. Oba rodzaje badań mają swoje miejsce w procesie podejmowania decyzji przez właścicieli gospodarstw. Gospodarstwa domowe (agroturystyczne) mają całkowitą swobodę podejmowania decyzji, w szczególności podejmowania decyzji inwestycyjnych.

2. AGROTURYSTYKA – SPECYFIKA DZIAŁALNOŚCI

Agroturystyka ma w Europie długoletnie tradycje. Działalność z pogranicza rolnictwa i turystyki istniała już bardzo dawno. Ważnym elementem tego rodzaju aktywności jest pierwiastek „naturalny”, którym mogą być pola, lasy czy woda. Bez wątpienia głównym źródłem atrakcji jest samo przebywanie w wiejskim otoczeniu. Mamy tu do czynienia z pewnego rodzaju elementem „nieprodukcyjnego” użytkowania zasobów środowiska do spędzania wolnego czasu.

Turystyka i działalność usługowa jej towarzysząca postrzegana jest jako droga do zróżnicowania gospodarczego, umożliwiająca transfer dobrobytu z uprzemysłowionych obszarów miejskich i tworzenie nowych miejsc pracy na wsi. Występuje więc sprzyjający splot okoliczności, kiedy nowym formom podaży odpowiadają podobne oczekiwania rynku (popytu).

Agroturystyka definiowana jest jako forma rekreacji odbywająca się na obszarach wiejskich o charakterze rolniczym, która oparta jest na bazie noclegowej i aktywności rekreacyjnej związanej z gospodarstwem rolnym i jego otoczeniem. Obejmuje ona wiele rodzajów aktywności związanej z turystyką pobytową, krajoznawczą, aktywną specjalistyczną, kulturową (MŁYNARCZYK 2002). Opie-

organization and management, as well as marketing. As for any economic activity, agrotourism requires investments in a variety of investments and can be examined by means of the economic methods used in economic activity analysis.

Marketing research consists in planned, systematic data collection and analysis in order to help firms make decisions and monitor the outcome of these decisions. Market research is expected to clarify the nature of a market, it can be descriptive, helping to find out factual data about what is happening, or it can be analytical, trying to explain the relations between variables, for example why given phenomena occur. Both types have a place in the decision-making process. Household businesses (agrotourism farms) have full freedom in making decisions, especially those regarding investments.

2. AGROTOURISM – ITS SPECIFIC CHARACTER

Agrotourism in Europe has a long tradition and has existed for a long time, developing on the border of agriculture and tourism. An important part is the 'natural' element such as fields, forests or water. Beyond doubt, the main source of attraction is the stay itself in the countryside, a kind of 'unproductive' usage of natural resources.

Tourism and its accompanying services are perceived as a route to economic diversification, facilitating a transfer of goods from developed urban areas and creating new workplaces in the countryside. Therefore new forms of supply respond to such market expectations (demand).

Agrotourism is defined as a form of recreation that takes place in rural, farming areas and is based on the accommodation and recreational activity associated with a farm and its surroundings, including many types related to residential, active and cultural tourism (MŁYNARCZYK 2002). It is based on local initiatives in the creation of small firms, and it provides employment, particularly for those from the farming sector. Moreover, agrotourism gives a chance to local communities to gain alternative sources of income as such farms offer tourism services while maintaining their agricultural character, i.e. continuing in agricultural

ra się ona na lokalnych inicjatywach w zakresie tworzenia małych przedsiębiorstw, daje zatrudnienie przede wszystkim ludziom związanym z rolnictwem i jest szansą pomocy lokalnym społecznościom w uzyskiwaniu alternatywnych źródeł dochodów. Gospodarstwa agroturystyczne świadczą usługi turystyczne zachowując rolniczy charakter, czyli jest w nich prowadzona produkcja rolna. Ta produkcja jest m.in. wykorzystywana na potrzeby gości przybywających do gospodarstwa agroturystycznego. Gospodarstwa agroturystyczne świadczą usługi związane nie tylko z noclegiem, ale również oferują wyżywienie opierając się na własnych produktach.

Według Ministerstwa Rolnictwa i Rozwoju Wsi w 1993 r. funkcjonowało w Polsce ok. 1000 gospodarstw domowych na wsi świadczących usługi turystyczne, natomiast w 2000 r. było ich 11 260, a w roku 2005 już 18 500. Około połowy z nich stanowiły gospodarstwa agroturystyczne, reszta zaś to gospodarstwa, w których nie prowadzi się bezpośredniej produkcji rolnej.

Turystyka wiejska oznacza spędzanie czasu wolnego w środowisku wiejskim. Cechą charakterystyczną tego środowiska jest odpowiedni charakter jego zabudowy i krajobrazu. Infrastrukturę noclegową tej turystyki tworzą hoteliki, pensjonaty i schroniska (DRZEWIECKI 1995). Turystyka wiejska jest pojęciem szerszym niż agroturystyka. Rozwój turystyki wiejskiej opiera się na zasobach naturalnych, historycznych i kulturowych danego obszaru, angażowaniu różnych podmiotów spośród miejscowej społeczności oraz zintegrowaniu turystyki z innymi aspektami rozwoju społeczności i gospodarki lokalnej.

Produktem turystyki wiejskiej nazywamy wszelkie dobra i usługi, które można uzyskać odpłatnie lub bezpłatnie na wsi, a które umożliwiają i uprzyjemniają pobyt i wypoczynek gości. Będą to przykładowo obiekty mieszkalne, zakłady przyrodolecznicze, środki komunikacji, placówki kulturalne, służby porządkowe dbające o bezpieczeństwo mieszkańców i turystów, ale także otoczenie naturalne oraz ekologiczne czy też empatia mieszkańców.

3. ANALIZA DZIAŁALNOŚCI AGROTURYSTYCZNEJ W REGIONIE ŚWIĘTOKRZYSKIM

Analiza podaży. Gospodarstwa w regionie świętokrzyskim oferują usługi agroturystyczne od 1993 r. Początkowo było ich zaledwie siedem, w 1997 r. zorganizowanych i zinventaryzowanych gospo-

production. The latter is also used to satisfy visitor needs as not only do the farms offer accommodation but full board based on their own products as well.

According to the Ministry of Agriculture and Rural Development, in 1993 there were about 1000 rural household businesses providing services for tourists, in 2000 – 11 260 and in 2005 – 18 500. Around half were agrotourism farms while the remaining were household businesses with no direct agricultural production.

Rural tourism means spending free time in the rural environment, with its characteristic buildings and landscape while its accommodation infrastructure includes small hotels, B&Bs, guest houses and hostels (DRZEWIECKI 1995). Rural tourism is a wider concept than agrotourism. Rural tourism development is based on the natural, historical and cultural resources of a given area, the involvement of the local community and the integration between tourism and other aspects of the community and local economic development.

The product of rural tourism includes all the goods and services which can be obtained in the countryside either for money or free of charge, enabling tourists to stay and relax e.g. residential buildings, healing institutions, transport infrastructure, cultural institutions, safety and security services, as well as the natural environment and 'empathy' with local residents.

3. AGROTOURISM ACTIVITY ANALYSIS IN THE ŚWIĘTOKRZYSKI REGION

Supply analysis. Household businesses in the Świętokrzyski region¹ have been offering agrotourism services since 1993. In the beginning there were only seven but by 1997 the number of organized and catalogued agrotourism farms had risen to 76, in 1998 to 106, 2002 to 240, in 2003 – 279 and the number is still growing (PAŁKA 2006).

The aim of the survey (carried out by geography students from the Geography Institute of Świętokrzyski Academy as a part of field studies) was to recognise the characteristic features of agrotourism farms and their investment preferences. It provided data on the size of the farms (the area owned), the age and education of the owners, the character of the buildings

darstw agroturystycznych było 76, w 1998 r. – 106, w 2002 – 240, w 2003 – 279 i ciągle powstają nowe gospodarstwa (PAŁKA 2006).

Celem badań ankietowych prowadzonych w ramach ćwiczeń terenowych przez studentów geografii z Instytutu Geografii Akademii Świętokrzyskiej było rozpoznanie charakterystycznych cech gospodarstw agroturystycznych oraz ich preferencji inwestycyjnych. Dostarczyły one informacji na temat wielkości gospodarstw (pod względem posiadanego arealu), wieku i poziomu wykształcenia ich właścicieli, charakteru obiektów przeznaczonych na działalność agroturystyczną i bazę noclegową, sezonowości, sposobów promocji gospodarstw, znaczenia i rodzaju współpracy z lokalnymi władzami samorządowymi (pomoc w promocji gospodarstwa, wspólne przedsięwzięcia, np. w budowie infrastruktury). Ze względu na sytuację ekonomiczną na wsi ważna jest również znajomość struktury i źródeł przychodów właścicieli gospodarstw czy ocena opłacalności przedsięwzięć agroturystycznych, sposoby pozyskiwania środków na rozwój gospodarstw i zamierzenia na przyszłość, planowanie i prognozowanie rozwoju działalności agroturystycznej.

Jako technikę badań ilościowych wybrano kwestionariusz ankiety, skierowanej do właścicieli gospodarstw agroturystycznych. Wyniki badań ankietowych prowadzonych w 2002 r. na próbie 248 gospodarstw agroturystycznych zlokalizowanych w województwie świętokrzyskim pokazały, że najczęściej, bo 33% badanych gospodarstw, mieści się w zakresie od 2 do 5 ha, 27% gospodarstw ma powierzchnię do 2 ha, 22% to gospodarstwa liczące powyżej 15 ha, 15% to gospodarstwa agroturystyczne o powierzchni 5–10 ha. Najniższy procent, bo zaledwie 3%, to gospodarstwa w przedziale 10–15 ha. Rozkład wielkości gospodarstw agroturystycznych jest zbliżony do rozkładu wielkości wszystkich gospodarstw w województwie świętokrzyskim, a więc działalność agroturystyczną podejmują gospodarstwa typowe dla tego regionu. Według wyników Powszechnego Spisu Rolnego 2002 rolnictwo województwa świętokrzyskiego charakteryzuje się ogromnym rozdrobnieniem, przeszło 60% gospodarstw posiada powierzchnię poniżej 5 ha.

47% badanych właścicieli gospodarstw agroturystycznych była w wieku 40–50 lat, powyżej 50 lat było 24% badanych, w grupie wiekowej 30–40 lat – 27% badanych, a w wieku poniżej 30 lat – 2% badanych. Przeważająca liczba właścicieli gospodarstw agroturystycznych posiadała wykształcenie średnie – 62%. Wykształcenie wyższe miało 22% właścicieli.

intended for agrotourism activity and accommodation, seasonality, methods of promotion, the importance and form of cooperation with local government (help in promotion, joint ventures, for instance installing infrastructure). Considering the economic situation in the countryside, it is important to know the kinds and sources of the owners' income, to estimate the cost-effectiveness of agrotourism investments, recognize ways of resourcing development, future goals, and to plan the development of agrotourism activity. A questionnaire directed at the owners of agrotourism farms was selected as a quantitative research tool. The results of the survey, carried out in 2002 among 248 agrotourism farms in Świętokrzyskie Województwo, showed that 33% (the largest group) of farms are between two and five hectares, 27% are less than two hectares, 22% are above 15 hectares, and 15% are between 5–10 hectares. The smallest number, just 3%, cover an area of 10–15 hectares. According to the Agricultural Census of 2002, agriculture in Świętokrzyskie Województwo is characterised by considerable fragmentation with 60% of the household businesses cover less than 5 hectares. Therefore it can be concluded that farm size distribution is close to the size distribution of all the household businesses in Świętokrzyskie Województwo and that agrotourism activity is undertaken by household businesses typical of this region.

As regards the age of farm owners, 47% were 40–50, 24% were over 50, 27% were between 30 and 40, and 2% below 30. The majority of agrotourism farm owners (66%) had had only secondary education, while 22% had had university education.

Accommodation is the most important part of rural tourism. The tourist whose visit includes an overnight stay is a valuable source of income for rural areas. Research shows that this tourist usually spends more (up to 300%) than a one-day visitor. The analysis also indicates that 70% of expenditure on accommodation in rural areas remains in the local economy and exceeds other expenditure made by tourists (MAJEWSKI 2000).

The hosts offered their guests a stay in a variety of buildings, with the majority of guest rooms being in their own houses (80% of households). In 20% the accommodation was offered in holiday 'cottages', i.e. separate buildings, fully reserved for tourist use, located mostly at the agrotourism farm

Baza noclegowa jest najważniejszą częścią składową turystyki wiejskiej. Turysta, którego pobyt obejmuje nocleg jest cenionym źródłem dochodów dla terenów wiejskich. Badania wykazują, że taki turysta wydaje z reguły więcej (nawet do 300%) niż jednodniowy odwiedzający. Analizy wykazują również, że do 70% wydatków związanych z zakwaterowaniem na obszarach wiejskich pozostaje w lokalnej gospodarce i jest to więcej niż w przypadku innych wydatków turystów (KRYŃSKI 2003).

Gospodarze oferowali swoim gościom noclegi (pobyt) w różnego typu obiektach, przy czym przeważały pokoje gościnne w domach gospodarzy (80% gospodarstw). W 20% przypadków oferowano noclegi w domach wakacyjnych, czyli oddzielne budynki, w całości przeznaczone dla turystów, najczęściej zlokalizowane w gospodarstwie agroturystycznym lub w znacznej od niego odległości. Obiekty tego typu to drewniane domy letniskowe lub niezamieszkały dom po rodzicach gospodarzy, specjalnie wyposażony dla gości. Gospodarstwa agroturystyczne oferowały też własne produkty żywnościowe w postaci posiłków lub na sprzedaż. Badania wykazały, że działalność agroturystyczna w regionie świętokrzyskim ma częściowo charakter sezonowy. Działalność całoroczną prowadziło 68% badanych gospodarstw, zaś 32% przyjmowało gości jedynie w sezonie letnim. Ta cecha ma szczególne odzwierciedlenie w infrastrukturze gospodarstw i ich lokalizacji.

Z badań ankietowych wynika, że 59% badanych to rolnicy i ich jedynym źródłem przychodów jest praca w gospodarstwie. Natomiast 41% ankietowanych oświadczyło, że ich głównym źródłem dochodów jest praca poza rolnictwem (handel, usługi, szkolnictwo). Wydawałoby się, że osoby, które prowadzą gospodarstwo agroturystyczne i pracują poza gospodarstwem powinny zatrudniać do pracy pracowników najemnych. Tymczasem tylko 7% gospodarstw agroturystycznych zatrudnia pracowników najemnych i to nie do obsługi gości czy przy pracach porządkowych, ale przy pracach polowych. Badani właściciele gospodarstw w 93% deklarowali, że w ich gospodarstwach agroturystycznych pomagają im członkowie rodziny, a jedynie w wyjątkowych sytuacjach mogą liczyć na nieodpłatną pomoc sąsiadów.

Inwestycje turystyczne to ponoszenie wydatków, które mają na celu przyciągnięcie turystów i zaspokojenie ich potrzeb. Jest to zatem pojęcie używane w bardzo szerokim znaczeniu, obejmujące zarówno inwestowanie w nowe obiekty lub modernizację istniejących, jak i w zasoby ludzkie (DAVIDSON 1993).

or within a close distance. The buildings are most often purpose built holiday cottages or houses, now uninhabited, that used to be lived in by the owner's parents, equipped specially for tourists. The research showed that agrotourism activity in the Świętokrzyskie Województwo has a partly seasonal character with 68% of those studied running all year long activity, while 32% took tourists only in the summer season. This feature is a particular reflection of the infrastructure of the business and its location.

Agrotourism gives an opportunity to create other sources of income apart from arable and pastoral farming on which to base a household business existence. According to the survey, 59% were farmers whose only income was from their work on the farm, however 41% stated that their main source of revenue was from outside work (business, services, education). It might seem that those who run an agrotourism farm, and work outside it as well, should be looking for employees. However, only 7% do and not for serving the visitors or carrying out cleaning duties, but to help with work in the fields. 93% declared that those helping at their farms are members of the family and only in exceptional circumstances could they count on voluntary support of their neighbours.

Tourism investments mean higher expenses in order to attract tourists and meet their demands. Therefore, the term is used in a wide range of meanings including investment in new buildings, the modernisation of the existing ones, and investment in human resources as well (DAVIDSON 1993).

Running an agrotourism activity is connected with the investment and funds necessary for its achievement. Investment activity may assume different forms and respondents quoted those made over the past three years such as adapting rooms for guests, purchase of beds and room furnishings. Some of them had rebuilt the sanitary system, adapted rooms for bathrooms, built extra rooms or extended the house. Not many farmers built new buildings especially for tourists. In the majority of cases funds for investment came from personal savings. Only 1.5% of the respondents took a bank loan in the first year of their activity and only 8% of the farmers received a grant from the EU both of which required drawing up a business plan. 69% of respondents had development plans for their agrotourism

Prowadzenie działalności agroturystycznej związane jest z inwestycjami i środkami finansowymi na ich realizację. Działalność inwestycyjna może przybierać tutaj różne rodzaje i rozmiary. Ankietowani podawali swoje działania inwestycyjne prowadzone w ciągu trzech lat. Były to działania typu: dostosowanie pokoi dla gości, zakup łóżek, wyposażenie pokoi. Niektórzy przebudowywali węzeł sanitarny w domu, dokonywali adaptacji pomieszczeń na łazienki, dobudowywali pomieszczenia, rozbudowywali dom. Niewielu gospodarzy budowało nowe budynki, specjalnie dla turystów. Środki na inwestycje pochodziły w większości z własnych oszczędności. Jedynie 1,5% ankietowanych w pierwszym roku działalności zaciągnęło kredyt bankowy, zaledwie 8% rolników uzyskało dotacje z Unii Europejskiej. Zaciągnięcie kredytu lub pozyskanie dofinansowania ze środków unijnych wymagało opracowania planu biznesowego. Plany rozwoju działalności agroturystycznej posiadało 69% ankietowanych rolników, natomiast 31% z nich określiło, że prowadzą swoją działalność spontanicznie, bez planowania wydatków i dochodów.

4. PRZYDATNOŚĆ REGUŁ ASOCJACJI ORAZ REGUŁ SEKWENCJI W ANALIZACH PRZEDSIĘWZIĘĆ INWESTYCYJNYCH GOSPODARSTW AGROTURYSTYCZNYCH

Celem naszej analizy jest znalezienie związku (asocjacji) oraz kolejności (sekwencji) pomiędzy działaniami inwestycyjnymi gospodarstwa, czyli wspólnego ich występowania w „koszyku” badanego gospodarstwa agroturystycznego. Analiza dotyczy współzależności inwestycji podejmowanych przez poszczególne gospodarstwa agroturystyczne. Gospodarstwo agroturystyczne opisało swoje działania inwestycyjne za okres trzech kolejnych lat. Usystematyzowanie wiedzy o powiązaniach między prowadzonymi inwestycjami przez gospodarstwa agroturystyczne pozwoliłoby doradcom zaproponować dodatkowe zakupy, wspierać działania w formie dotacji czy kredytów, rozpoznawać rynkowe potrzeby. Taka wiedza może być również pomocna w polityce zagospodarowania przestrzennego badanych gmin.

Reguły asocjacyjne i sekwencyjne umożliwiają odkrywanie współwystępowania i sekwencji zdarzeń. Odkrywanie asocjacji polega na znalezieniu przedsięwzięć (w szczególności inwestycji), które są dokonywane razem. Zagadnienie to ilustruje ogólna reguła: „jeżeli zdarzenie A wystąpiło w danej sytu-

activity, while 31% stated that it was run without planning expenses or incomes.

4. USEFULNESS OF ASSOCIATION AND SEQUENCE RULES IN THE ANALYSIS OF INVESTMENTS ON AGROTOURISM FARMS

The aim of the analysis is to find the connection (association) and the order (sequence) of investment activities at a farm, i.e. their co-occurrence in the 'basket' of a given agrotourism farm. The analysis applies to the interdependencies between investments made at individual agrotourism farms over a period of three consecutive years. Systematised knowledge about the links between investments would help advisors suggest additional purchases, support the agrotourism activity in the form of grants or loans, and to identify market demands. This knowledge could be helpful in the spatial management policy of the communities studied.

Association and sequence rules enable exploration of event co-occurrence and sequence. Association discovery consists in finding investments that are undertaken jointly. The issue is illustrated by a general rule: 'if an event A occurred in the given situation a certain number of times, then event B occurred in the same situation a certain number of times'.

Association rules should not be interpreted as rules directly illustrating causality (if A, then B) but as interdependencies (relations) between two events or a set of events. Association analysis does not consist in creating rules concerning cause-and-effect relationships: if event A has occurred, then event B will occur with a certain probability. Such research is not about whether the potential buyer will purchase one or more items of 'good B' but whether the presence of 'good B' in the basket is related to the purchase of 'good A'. The basket analysis can be used to establish rules such as: if A then probably B as well. For instance, if investment A is a purchase of a car then investment B will mean starting to provide catering services.

In this case the investments agrotourism farms make jointly were considered. For example if the farm provides catering services then it buys a dishwasher, and land purchases co-occur with the development

acji pewną liczbę razy, to zdarzenie B także wystąpiło w tej samej sytuacji określoną liczbę razy”.

Reguły asocjacyjne nie powinny być interpretowane jako reguły ujmujące bezpośrednio przyczynowość (jeżeli A, to B), lecz jako współzależności (związki) między dwoma zdarzeniami lub zespołem zdarzeń. Analiza asocjacji nie polega na tworzeniu reguł o związkach przyczynowo-skutkowych postaci: jeżeli wystąpiło zdarzenie A, to wystąpi zdarzenie B z określonym prawdopodobieństwem. W badaniach tych nie chodzi o to, czy potencjalny nabywca kupi jedną czy więcej sztuk towaru B, ale czy obecność towaru B w „koszyku” jest związana z zakupem towaru A. Analizę koszykową możemy użyć do znalezienia reguł typu: jeżeli A, to prawdopodobnie również B. Przykładowo, jeżeli inwestycja A będzie zakupem samochodu, to inwestycja B będzie oznaczać podjęcie świadczenia usług gastronomicznych.

W naszym przypadku rozważaliśmy, jakie przedsięwzięcia gospodarstwa agroturystyczne realizują łącznie, np. czy jeżeli gospodarstwo realizuje usługi gastronomiczne, to zakupuje zmywarkę do naczyń, a zakupom ziemi towarzyszy rozbudowa bazy do rekreacji i zakup zwierząt do celów terapeutycznych i rekreacyjnych.

Reguły asocjacyjne są regułami budowanymi w postaci reguł logicznych, złożonych z poprzednika (lewa strona reguły) i następnika (prawa strona reguły). Zarówno poprzednik, jak i następnik mogą być złożone (zawierać koniunkcję wielu pozycji). Przykładem reguły złożonej może być reguła postaci: Jeżeli A i B to C, gdzie zdarzenia A, B, C dotyczą zakupu trzech określonych towarów.

Współczynnik ufności (*confidence factor*) jest to udział procentowy pojawienia się następnika, jeżeli pojawi się poprzednik. Na przykład, jeżeli 10% gospodarstw zakupiło towary A i B, a 8% zakupiło towary A, B, i C, to współczynnik ufności wynosi $8\% : 10\% = 0,8$. Jeżeli 80% gospodarstw zakupiło towary A i B, a 40% zakupiło towary A, B i C, to współczynnik ufności wynosi $40\% : 80\% = 0,5$. Widzimy, że w pierwszym przypadku frakcja gospodarstw, które jednocześnie zakupiły towary A, B, i C jest mała (8%), jednak współczynnik ufności reguły wynosi 0,8 i jest większy niż dla drugiego przypadku (0,5), chociaż wtedy frakcja gospodarstw, które dokonały jednoczesnego zakupu towarów A, B i C wynosi 40%. Im większa wartość współczynnika ufności, tym „lepsza” reguła.

Współczynnik wsparcia (*level of support*) jest to częstość współwystępowania (jednoczesnego występowania) poprzednika i następnika reguły.

of recreational infrastructure and the purchase of animals for therapeutic and recreational purposes.

Association rules have the form of logical rules consisting of a *predecessor* (left side of the rule) and a *consequent* (right side of the rule). Both the *predecessor* and the *consequent* can be complex (including the conjunction of several items). An example of a complex rule can be the following: if A and B then C, where the events A, B, C concern the purchase of three different goods.

The *confidence factor* is the percentage of the *consequent*, providing the *predecessor* appears. For instance, if 10% of household businesses have purchased goods A and B, and 8% have purchased goods A, B and C, then the confidence factor amounts to $8\% : 10\% = 0.8$. If 80% of household businesses have bought goods A and B, and 40% have bought goods A, B and C, then the confidence factor amounts to $40\% : 80\% = 0.5$. In the first example it can be seen that those household businesses that purchased goods A, B and C at the same time is low (8%), however the confidence factor comes to 0.8 and is higher than in the second example (0.5), even though those household businesses which bought goods A, B and C simultaneously was 40%. The higher the value of the confidence factor, the 'better' the rule is.

The *level of support* is the co-occurrence (simultaneous occurrence) frequency of the *predecessor* and *consequent* in the rule. In the given examples the level of support comes to 8% and 40% respectively.

Let us assume that 50% of household businesses purchased commodity C (regardless of buying A and B). Then we can say that the probability of purchasing commodity C is 0.5. If goods A and B (*predecessor*) and C (*consequent*) were independent, then probability $P(A \text{ and } B \text{ and } C) = P(A \text{ and } B) \cdot P(C)$ would result in the first case $0.1 \cdot 0.5 = 0.05$, and in second case $0.8 \cdot 0.5 = 0.4$. In reality $P(A, B, C) = 0.08$ (first case) and $P(A, B, C) = 0.5$ (second case). The information that the household business purchased goods A and B allows us to conclude that the household business is more likely to purchase commodity C. A relative increase in this probability comes to $0.08 : 0.05 = 1.6$ in the first case and $0.5 : 0.4 = 1.25$ in the second case. This quotient is called the *lift factor* and it demonstrates the increase in probability of the *consequent* if the

W podanych przykładach współczynnik wsparcia wynosi odpowiednio 8% i 40%.

Przyjmijmy, że 50% gospodarstw zakupiło towar C (niezależnie od tego, czy zakupiło towary A i B). Wówczas możemy powiedzieć, że prawdopodobieństwo zakupu towaru C wynosi 0,5. Gdyby zakupy A i B (poprzednik reguły) i C (następnik reguły) były niezależne, to prawdopodobieństwo $P(A \text{ i } B \text{ i } C) = P(A \text{ i } B) \cdot P(C)$ wynosiłoby w pierwszym przypadku $0,1 \cdot 0,5 = 0,05$, a w drugim przypadku $0,8 \cdot 0,5 = 0,4$. W rzeczywistości mamy $P(A \text{ i } B \text{ i } C)$ równe 0,08 (pierwszy przypadek) oraz $P(A \text{ i } B \text{ i } C)$ równe 0,5 (drugi przypadek). Informacja o tym, że gospodarstwo zakupiło towary A i B pozwala nam wywnioskować, że gospodarstwo z większym prawdopodobieństwem zakupi towar C. Względny wzrost tego prawdopodobieństwa wynosi: $0,08 : 0,05 = 1,6$ – w pierwszym przypadku i $0,5 : 0,4 = 1,25$ – w drugim przypadku. Iloraz ten nazywamy współczynnikiem wzrostu (*lift*). Współczynnik wzrostu wyraża wzrost prawdopodobieństwa następnika, jeżeli wystąpił poprzednik w stosunku do prawdopodobieństwa następnika w całym zbiorze. „Dobre” reguły mają wartość współczynnika wzrostu większą niż 1¹. W naszym przykładzie reguła dla pierwszego przypadku ma większy współczynnik wzrostu.

Liczba transakcji (*transaction count*) oznacza liczbę różnych par: gospodarstwo – przedsięwzięcie (zdarzenie) w zbiorze danych.

Reguły sekwencji (*sequence discovery*) będą informować o kolejności działań. Jeżeli zamierzamy je budować, to w zbiorze danych musi istnieć zmienna określająca czas. Wówczas reguła: jeżeli A, to B wskazuje, że B pojawiło się po zajściu A. Zmienna określająca czas jest zmienną numeryczną, której wartość jest miernikiem czasu dla każdej obserwacji. Czas może być podany w godzinach, dniach, tygodniach, latach. Zmienna może przyjmować wartości całkowite, co oznacza dyskretny pomiar czasu.

5. ANALIZA ASOCJACJI – PRZYPADEK GOSPODARSTW AGROTURYSTYCZNYCH WOJEWÓDZTWA ŚWIĘTOKRZYSKIEGO

Analiza dotyczy współzależności przedsięwzięć podejmowanych przez gospodarstwa agroturystyczne województwa świętokrzyskiego. Do przedsięwzięć zaliczono działania podawane w ankiecie, takie jak zaciągnięcie kredytu, rozbudowa

predecessor appeared in relation to the probability of the *consequent* in the whole set. 'Good' rules have a growth ratio value higher than 1¹. In our example the rule for the first case has a higher growth ratio.

The *transaction count* represents a number of different pairs: household businesses – investments (events) in a data set.

Sequence discovery will predict the order of events and if it is going used there has to be one variable in the data set that defines time. Therefore the rule 'if A then B' indicates that B appeared after A. The variable which defines time is a numerical variable whose value is a measure of time for any observation. Time can be quoted in hours, days, weeks or years. It can be a whole number which indicates a discrete measurement of time.

5. ASSOCIATION ANALYSIS – A CASE STUDY OF AGROTOURISM FARMS IN THE ŚWIĘTOKRZYSKIE WOJEWÓDZTWO

The analysis concerns the interdependencies between the investments made by agrotourism farms in the Świętokrzyskie Województwo. The investments include those mentioned in the questionnaire such as taking a loan, house extension, investments in the internet and computer equipment, preparing relaxation areas, purchase of sports equipment and animals for therapeutic and recreational purposes. A list of all investments quoted in the questionnaire can be found in Table 1.

The research covered a three year period, therefore the time variable is measured in years and takes the values 1, 2, and 3. In each year activity to adapt the household business for agrotourism activity was analysed.

Investments were looked for that had been made simultaneously by agrotourism farms. In order to find such relations the data set had to include two variables: identification of the subject (here the name of the agrotourism farm) and a variable containing data about the type of investment undertaken by the farm. Table 2 gives some data.

T a b e l a 1. Lista inwestycji w gospodarstwie agroturystycznym

Zmienne jakościowe wielokrotnych odpowiedzi Zakres czasowy t = 1, 2, 3 (dla trzech lat)			
1	Zaciągnięcie kredytu	20	Sprzęt do rekreacji w domu i ogrodzie
2	Pozyskanie środków unijnych	21	Sprzęt do rekreacji w ogrodzie
3	Rozbudowa domu	22	Sprzęt sportowy
4	Nadbudowa domu	23	Rowery górskie
5	Przebudowa domu	24	Usługi gastronomiczne
6	Remont domu	25	Zakup chłodziarki
7	Adaptacja poddaszy	26	Zmywarka do naczyń
8	Zwiększenie liczby łazienek	27	Zakup maszyn ogrodniczych
9	Sieć kanalizacyjna	28	Zakup ziemi
10	Zakup samochodu	29	Minizoo
11	Zagospodarowanie terenu	30	Zakup zwierząt do celów rekreacyjnych
12	Budowa parkingu	31	Zakup zwierząt do celów terapeutycznych
13	Wyposażenie pokoi	32	Rozbudowa bazy noclegowej
14	Artykuły do urządzenia mieszkań	33	Ekologiczne ogrzewanie
15	Artykuły do wyposażenia łazienki	34	Budowa stawu
16	Wyposażenie pomieszczeń dla gości	35	Konie
17	Urządzenie miejsc do wypoczynku	36	Zakup lasu
18	Sprzęt do rekreacji w domu	37	Budowa kortu tenisowego
19	Internet i sprzęt komputerowy	38	Budowa basenu

Ź r ó d ł o: Opracowanie na podstawie ankiet (pozycje inwestycji podawane przez przedstawicieli gospodarstw).

T a b l e 1. Investment in an agrotourism farm

Qualitative changes found repeatedly Time span t = 1, 2, 3 (indicating years)			
1	Raising a loan	20	Home and garden recreation equipment
2	Receiving an EU grant	21	Garden recreation equipment
3	House extension	22	Sports equipment
4	Additional house floors	23	Mountain bikes
5	House reconstruction	24	Catering services
6	House renovation	25	Purchase of a freezer
7	Attic conversions	26	Purchase of a dishwasher
8	Increasing the number of bathrooms	27	Garden equipment purchase
9	Sewage network	28	Land purchase
10	Car purchase	29	Mini zoo
11	Land management	30	Purchase of animals for recreational purposes
12	Building of a car park	31	Purchase of animals for therapeutic purposes
13	Room furnishing	32	Accommodation extension
14	Furnishings	33	Environmentally-friendly heating
15	Bathroom items	34	Building of a pond
16	Guest rooms furnishings	35	Horses
17	Relaxation facilities	36	Purchase of forest
18	Home recreational equipment	37	Building of a tennis court
19	Internet and computer equipment	38	Building of a swimming pool

S o u r c e: Based on questionnaires (entries on the list provided by household business representatives).

domu, inwestycje w Internet i sprzęt komputerowy, urządzenie miejsc do wypoczynku, zakup sprzętu sportowego, pozyskanie zwierząt do celów terapeutycznych i rekreacyjnych. Zestawienie wszystkich przedsięwzięć podanych przez gospodarstwa agroturystyczne w ankiecie znajduje się w tab. 1.

Badaniami objęto okres trzech lat, zatem zmienna czasowa mierzona jest w latach i przyjmuje wartości 1, 2, 3. W każdym roku analizowano przedsięwzięcia podejmowane dla lepszego przystosowania gospodarstwa do prowadzenia działalności agroturystycznej. Poszukiwaliśmy, jakie przedsięwzięcia gospodarstwa agroturystyczne realizują łącznie, np. czy jeżeli gospodarstwo świadczy usługi gastronomiczne, to zakupuje zmywarkę do naczyń, czy zakupom ziemi towarzyszy rozbudowa bazy do uprawiania sportu i rekreacji. Aby znaleźć takie zależności zbiór danych musi zawierać dwie zmienne: identyfikator obiektu (u nas przyjęliśmy nazwę gospodarstwa agroturystycznego) i zmienną zawierającą dane o rodzaju przedsięwzięcia podejmowanego przez gospodarstwo. Tabela 2 pokazuje postać kilku obserwacji w zbiorze danych.

The analysis concerns transactions. Each transaction is a subject (here an agrotourism farm identified by name) and an event (here starting an investment, e.g. taking a loan, implementing ecological heating or purchase of a certain commodity: a car, some land, or a dishwasher). The first line in Table 2 indicates that household business Gnojno 1 invested in room furnishing, the second line shows that the same one invested in bathroom articles. The fifth line shows that household business Busko Zdrój 1 installed the internet and purchased computer equipment, while the sixth line – that ecological heating was installed.

In order to find out what the studied agrotourism farms mostly invested in, a graph based on the *Investment* variable was made (Table 2, Fig. 1).

T a b e l a 2. Zbiór danych przygotowany na potrzeby analizy asocjacji (fragment)

Lp.	Gospodarstwo agroturystyczne (identyfikator miejsce lokalizacji) Zmienna „gospodarstwo”	Przedsięwzięcie Zmienna „inwestycje” (target)
1	Gnojno 1	Wyposażenie pokoi
2	Gnojno 1	Artykuły do wyposażenia łazienki
3	Gnojno 2	Pozyskanie środków unijnych
4	Gnojno2	Usługi gastronomiczne
5	Busko-Zdrój 1	Internet i sprzęt komputerowy
6	Busko-Zdrój 1	Ekologiczne ogrzewanie
7	Busko-Zdrój 2	Rozbudowa bazy noclegowej
8	Busko-Zdrój 2	Zaciągnięcie kredytu
9	Busko-Zdrój 2	Zakup samochodu

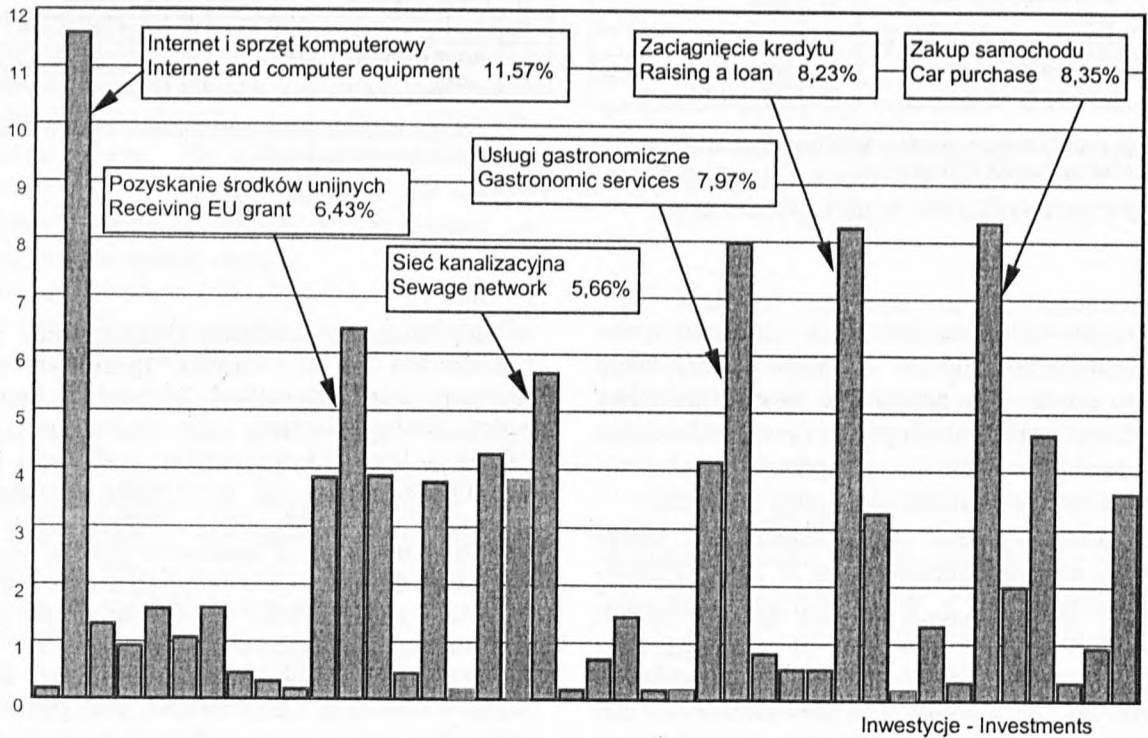
Ź r ó d ł o: Opracowanie własne przy wykorzystaniu programu SAS Enterprise Miner.

T a b e l e 2. Data prepared for association analysis purposes (partial)

	Agrotourism farm (identified by location)	Investment made
1	Gnojno 1	Room furnishings
2	Gnojno 1	Bathroom furnishings
3	Gnojno 2	Receiving an EU grant
4	Gnojno 2	Catering services
5	Busko-Zdrój 1	Internet and computer equipment
6	Busko-Zdrój 1	Environmentally-friendly heating
7	Busko-Zdrój 2	Accommodation extension
8	Busko-Zdrój 2	Raising a loan
9	Busko-Zdrój 2	Car purchase

S o u r c e: Author with the help of the SAS Enterprise Miner programme.

Procent
Percentage



Rys. 1. Wykres rozkładu zmiennej „inwestycje” (pierwszy rok działalności gospodarstw)
(ź r ó d ł o: opracowanie własne przy wykorzystaniu programu SAS Enterprise Miner)

Fig. 1 Investment distribution (first year of agrotourism activity)
(source: author with the help of the SAS Enterprise Miner programme)

Przedmiotem analizy są tzw. transakcje. Każda transakcja, to obiekt (tu: gospodarstwo agroturystyczne identyfikowane przez nazwę gospodarstwa agroturystycznego) i zdarzenie (tu podjęcie działalności, np. zaciągnięcie kredytu, wprowadzenie eko-

The graph presented in Fig. 1 illustrates the percentage of individual investments made by agrotourism farms. Due to the fact that each investment is treated in terms of a 'made/ not made' category, it can be concluded that farms, in their first year

logicznego ogrzewania, zakup określonego dobra: ziemi, samochodu, zmywarki do naczyń). W tabeli 2 pierwszy wiersz sygnalizuje transakcje: gospodarstwo Gnojno 1 zainwestowało w wyposażenie pokoi, drugi wiersz, że to samo gospodarstwo zainwestowało w artykuły do wyposażenia łazienki, a np. piąty wiersz, że gospodarstwo Busko Zdrój 1 zainstalowało Internet i zakupiło sprzęt komputerowy, a szósty wiersz, że zainstalowało ekologiczne ogrzewanie.

Aby poznać, w co przede wszystkim inwestowały badane gospodarstwa agroturystyczne możemy sporządzić wykres rozkładu zmiennej „inwestycje” (por. tab. 2 i rys. 1).

Wykres przedstawiony na rys. 1 informuje nas o rozkładzie procentowym poszczególnych przedsięwzięć podejmowanych przez gospodarstwa agroturystyczne. Ponieważ każde przedsięwzięcie jest traktowane w kategoriach „podejmowane/nie podejmowane” można wnioskować, że najczęściej gospodarstw w pierwszym roku działalności zainwestowało w Internet i sprzęt komputerowy (11,57%), zakup samochodu (8,35%) i usługi gastronomiczne (7,97%), sieć kanalizacyjną (5,66%). Gospodarstwa zaciągały kredyty (8,23%) i pozyskiwały środki unijne (6,43%).

of activity, invested in the internet and computer equipment (11.57%), car purchase (8.35%), catering services (7.97%) and sewage network (5.66%), while 8.23% of farms raised a loan and 6.43% obtained EU funds.

In order to make an association analysis with the help of *Enterprise Miner*, the procedural parameters have to be stated:

– *Minimum Transaction Frequency to Support Associations* – 5% – it must be remembered that the lower the value, the higher the number of analyses that have to be conducted and the number of possible dependencies will increase with the decrease of the value

– The maximum number of items considered in the association rule – 4;

Tables 3 and 4 present a part of the association analysis results regarding investments in 248 agrotourism farms in the Świętokrzyskie *Województwo*. Rules with two relations were placed in Table 3 and rules with three relations in Table 4.

T a b e l e 3. Rules in text form with their characteristics: number of relations – 2; confidence, support, lift factors, transaction count. An example of the rules from one of the three years under study

Relation	Confidence (%)	Support (%)	Lift factor	Transaction count	Association rules
2	57.14	11.72	2.17	28	Receiving EU grant == > Raising a loan
2	35.94	9.61	1.39	23	Car Purchase == > Catering Services
2	40.82	8.37	1.12	20	Receiving EU grant == > Internet and computer equipment
2	30.16	7.95	2.18	19	Raising a loan == > Accommodation extension
2	44.19	7.95	1.21	19	Sewage network == > Internet and computer equipment
2	46.65	5.84	1.28	14	House expansion == > Internet and computer equipment

S o u r c e: Author with the help of SAS Enterprise Miner programme.

T a b e l a 3. Reguły w postaci tekstowej wraz z charakterystykami: liczba relacji w regule drugiej, zaufanie, wsparcie, współczynnik wzrostu, liczba transakcji – przykład reguł z jednego z trzech kolejnych badanych lat

Liczba relacji	Zaufanie (%)	Wsparcie (%)	Współczynnik wzrostu	Liczba transakcji	Reguły asocjacji
2	57,14	11,72	2,17	28	Pozyskanie środków unijnych == > zaciągnięcie kredytu
2	35,94	9,61	1,39	23	Zakup samochodu == > usługi gastronomiczne
2	40,82	8,37	1,12	20	Pozyskanie środków unijnych == > Internet i sprzęt komputerowy
2	30,16	7,95	2,18	19	Zaciągnięcie kredytu == > rozbudowa bazy noclegowej
2	44,19	7,95	1,21	19	Sieć kanalizacyjna == > Internet i sprzęt komputerowy
2	46,65	5,84	1,28	14	Rozbudowa domu == > Internet i sprzęt komputerowy

Ź r ó d ł o: Opracowanie własne przy wykorzystaniu programu SAS Enterprise Miner.

Aby przeprowadzić analizę asocjacji za pomocą programu Enterprise Miner należało określić parametry procedury:

– minimalną częstość pojawienia się transakcji, która pozwala podejrzewać, że zachodzi asocjacja (*Minimum Transaction Frequency to Support Associations*) – założono wartość 5% – należy pamiętać, że im niższa wartość, tym musi być prowadzona większa liczba analiz i wraz ze zmniejszeniem wartości, liczba możliwych zależności, która będzie rozważana, rośnie;

– maksymalną liczbę pozycji, jaka będzie brana pod uwagę w przesłance oraz następniku reguł asocjacji – przyjęto 4;

Fragment wyników przeprowadzonej metodą asocjacji analizy współzależności przedsięwzięć inwestycyjnych dla 248 gospodarstw agroturystycznych z województwa świętokrzyskiego przedstawiono w tab. 3 oraz 4. W tabeli 3 zamieszczono reguły z dwiema pozycjami w regule, a w tab. 4 – reguły z trzema pozycjami.

Graphs illustrating individual association rules can be drawn. The *predecessor* is presented on the vertical and the *consequent* on the horizontal axis. The value of the *confidence* factor is marked with symbols (geometrical figures) and the value of the *support* factor with colours. The size of the symbol represents the value of the *lift* factor.

In Table 3 the rules are sorted out decreasingly according to the *support* size. Two of them: 'Car Purchase = = > Catering Services' and 'House expansion = = > Internet and computer equipment' have been placed in Fig 2. These are rules with fairly high *support* and *confidence* values. The precise values of *support* and *confidence* ratios can be read in Table 3 and the range they fit into can be seen in the key under the figure. For comparison, two rules with

T a b l e 4. Rules in text form with their characteristics: number of relations – 3; *confidence*, *support*, *lift* factors, *transaction count*. An example of the rules from one of the three years under study

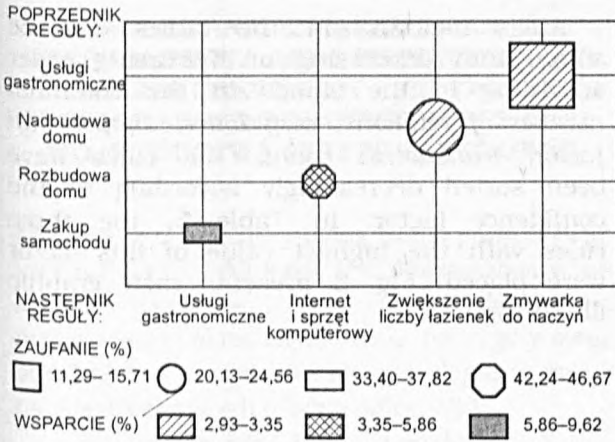
Relation	Confidence (%)	Support (%)	Lift factor	Transaction count	Association rule
3	22.73	6.28	3.02	15	Raising a loan == > Accommodation extension & Receiving EU grant
3	29.41	6.28	3.35	15	Receiving EU grant == > Raising a loan & Accommodation extension
3	53.57	6.28	3.46	15	Raising a loan & Receiving EU grant == > Accommodation extension
3	14.71	4.18	3.20	10	Purchase of room furnishing == > equipment for guest rooms & facilities for relaxation areas
3	35.71	4.18	1.64	10	Raising a loan & Receiving EU grant == > Catering services
3	66.67	4.18	2.41	10	Catering services & receiving EU grant == > Raising a loan

S o u r c e: Author with the help of the SAS Enterprise Miner programme.

T a b e l a 4. Reguły w postaci tekstowej wraz z charakterystykami: liczba relacji w regule trzeciej, zaufanie, wsparcie, współczynnik wzrostu, liczba transakcji – przykład reguł z jednego z trzech kolejnych badanych lat

Liczba relacji	Zaufanie (%)	Wsparcie (%)	Współczynnik wzrostu	Liczba transakcji	Reguły asocjacji
3	22,73	6,28	3,02	15	Zaciągnięcie kredytu == > rozbudowa bazy noclegowej & pozyskanie środków unijnych
3	29,41	6,28	3,35	15	Pozyskanie środków unijnych == > zaciągnięcie kredytu & rozbudowa bazy noclegowej
3	53,57	6,28	3,46	15	Zaciągnięcie kredytu & pozyskanie środków unijnych == > rozbudowa bazy noclegowej
3	14,71	4,18	3,20	10	Zakup wyposażenia pokoi == > wyposażenie pomieszczeń dla gości & urządzenie miejsc wypoczynku
3	35,71	4,18	1,64	10	Zaciągnięcie kredytu & pozyskanie środków unijnych == > usługi gastronomiczne
3	66,67	4,18	2,41	10	Usługi gastronomiczne & pozyskanie środków unijnych == > zaciągnięcie kredytu

Ź r ó d ł o: Opracowanie własne przy wykorzystaniu programu SAS Enterprise Miner.



Rys. 2. Ilustracja graficzna reguł asocjacji (źródło: opracowanie własne przy wykorzystaniu programu SAS Enterprise Miner)

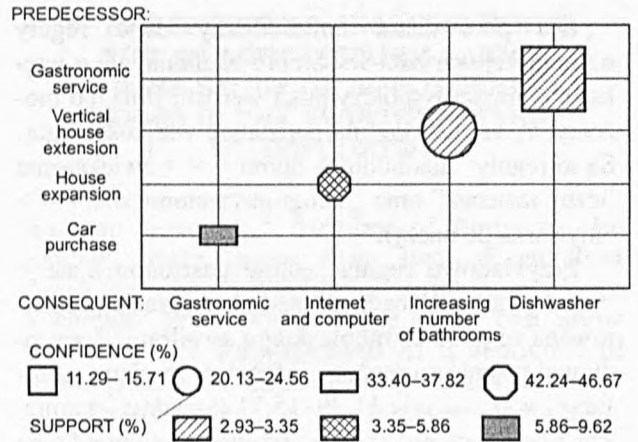


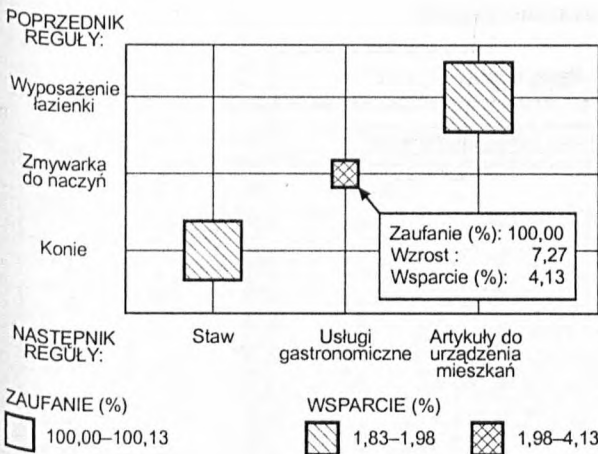
Fig. 2. Illustration of association rules (source: author with the help of the SAS Enterprise Miner programme)

Możemy sporządzić wykresy ilustrujące poszczególne reguły asocjacji. Poprzednik reguły jest przedstawiony na osi pionowej, a następnik na osi poziomej. Wielkości wskaźnika zaufania są oznaczane symbolami (kształtami figur), a wielkości wskaźnika wsparcia – szrafem. Rozmiar (wielkość) symbolu reprezentuje wartość współczynnika wzrostu (*lift*).

W tabeli 3 reguły były posortowane malejąco według wielkości wsparcia. Dwie z nich: „zakup samochodu ==> usługi gastronomiczne” oraz „rozbudowa domu ==> Internet i sprzęt komputerowy” zamieszczono na rys. 2. Są to reguły o dość dużych wartościach wsparcia i wartości zaufania. Dokładne wartości wskaźników wsparcia i zaufania możemy odczytać z tab. 3, a z legendy umieszczonej pod rysunkiem przedziały wartości, w jakich się mieszczą.

lower *support* and *confidence* values but with a higher *lift* factor value (which can be seen in the size of the symbols) are presented. They are 'House extension ==> Increased number of bathrooms' and 'Catering Services ==> Dishwasher'.

For instance, the rule *Catering services ==> Dishwasher* is presented on the graph as a blue square. It can be seen from the key under the figure that *confidence* is within the range 11.29–15.71 (the lowest symbol) and *support* within the range of 2.93–3.35 (the lowest colour). The size of the symbol represents the value (strength) of the *lift* factor. The picture shows that for the analysed rule the factor is high. Similarly, other rules can be analysed based on the figure.



Rys. 3. Ilustracja graficzna reguł asocjacji (źródło: opracowanie własne przy wykorzystaniu programu SAS Enterprise Miner)

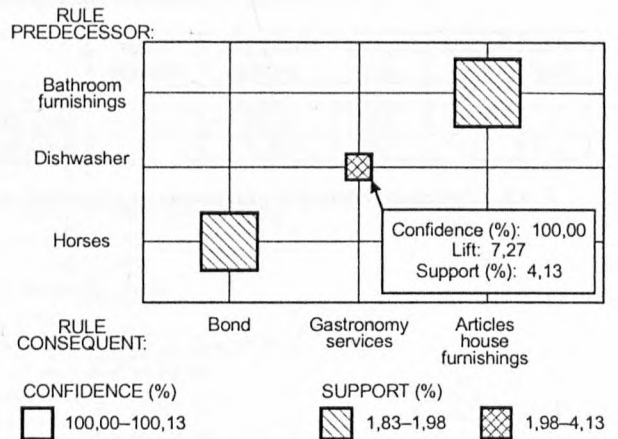


Fig. 3. Graphical illustration of association rules (source: author with the help of the SAS Enterprise Miner programme)

Dla porównania umieściliśmy dwie reguły o mniejszej wartości wsparcia i zaufania, ale o większej wartości współczynnika wzrostu (*lift*), co możemy wywnioskować na podstawie wielkości figur. Są to reguły: „nadbudowa domu == > zwiększenie liczby łazienek” oraz „usługi gastronomiczne == > zmywarka do naczyń”.

Przykładowo reguła „usługi gastronomiczne == > zmywarka do naczyń” jest na wykresie prezentowana za pomocą niebieskiego kwadratu. Z legendy pod rysunkiem możemy odczytać, że mamy „zaufanie” w przedziale 11,29–15,71 (kwadrat – najniższe z wymienionych) oraz „wsparcie” w przedziale 2,93–3,35 (kolor niebieski – najniższe z wymienionych). Wielkość symbolu reprezentuje wartość (siłę) wskaźnika *lift*. Rysunek pokazuje, że dla analizowanej reguły jest on wysoki. Podobnie możemy na podstawie rysunku analizować inne reguły.

Reguły zawarte w tabelach możemy sortować w porządku malejącym lub rosnącym według wartości kolumn: „liczba relacji”, „zaufanie”, „wsparcie”, „współczynnik wzrostu”, „liczba transakcji”. Posortowaliśmy reguły malejąco wg wskaźnika zaufania. W tabeli 5. umieściliśmy trzy reguły o najwyższej wielkości tego wskaźnika. Na rysunku 3. przedstawiono ich ilustrację graficzną.

Rules included in the tables can be sorted into decreasing or increasing order according to the values in the columns: *number of relations, confidence, support, lift factor, transaction count*. The rules have been sorted decreasingly according to the *confidence* factor. In Table 5, the three rules with the highest value of this factor were placed. Fig. 3 presents their graphic illustration.

T a b e l e 5. Rules according to the *confidence* factor (partial)

Relation	Confidence (%)	Support (%)	Lift factor	Transaction count	Association rule
2	100	1.83	19	4	Horses == > Pond
2	100	4.13	7.27	9	Dishwasher == > Catering Services
2	100	1.83	24.22	4	Bathroom furnishings == > House furnishings

S o u r c e: Author with the help of the SAS Enterprise Miner programme.

T a b e l a 5. Zestawienie z posortowanymi regułami według wskaźnika zaufanie (fragment)

Liczba relacji	Zaufanie (%)	Wsparcie (%)	Współczynnik wzrostu	Liczba transakcji	Reguły asocjacji
2	100	1,83	19	4	Konie == > staw
2	100	4,13	7,27	9	Zmywarka do naczyń == > usługi gastronomiczne
2	100	1,83	24,22	4	Wyposażenie łazienki == > artykuły do urządzania mieszkań

Ź r ó d ł o: Opracowanie własne przy wykorzystaniu programu SAS Enterprise Miner.

6. MODEL ANALIZY SEKWENCJI DLA GOSPODARSTWA AGROTURYSTYCZNEGO. ANALIZA SEKWENCJI – PRZYPADEK GOSPODARSTW AGROTURYSTYCZNYCH WOJEWÓDZTWA ŚWIĘTOKRZYSKIEGO

Możemy przyjąć maksymalną liczbę pozycji (więcej niż dwie), jaka może być umieszczona w sekwencji. Każda pozycja określa przedsięwzięcie realizowane w określonym czasie, np. reguły mogą przedstawiać, że realizowane są kolejno w trzech punktach czasowych (równo odległych):

– zaciągnięcie kredytu == > artykuły do wyposażenia łazienki == > sprzęt do rekreacji w domu i w ogrodzie,

– remont domu == > zakup wyposażenia pokoi == > urządzenie miejsc do wypoczynku,

– zaciągnięcie kredytu == > artykuły do wyposażenia łazienki i ekologiczne ogrzewanie == > sprzęt do rekreacji w domu i w ogrodzie.

Reguły zawierające dwie pozycje rozpatrywane w czasie możemy przedstawiać w postaci wykresu.

6. THE SEQUENCE ANALYSIS MODEL FOR AN AGROTOURISM FARM – A CASE STUDY OF AGROTOURISM FARMS IN THE ŚWIĘTOKRZYSKIE WOJEWÓDZTWO

We can assume a number of items in the longest chain (more than two). Each item describes an investment implemented at a certain time, for example rules can show that they are implemented in a sequence of three points in time (evenly distributed):

– Raising a loan == > Bathroom articles == > Home and garden recreation equipment,

– House renovation == > Purchase of room furnishings == > Relaxation areas facilities,

– Raising a loan == > Bathroom articles & Environmentally-friendly heating == > Home and garden recreation equipment.

Rules including two items considered from the time perspective can be presented in the form of a graph.

Table 6. Sequence rules
(part of a table with two-item sequence rules sorted according to Support)

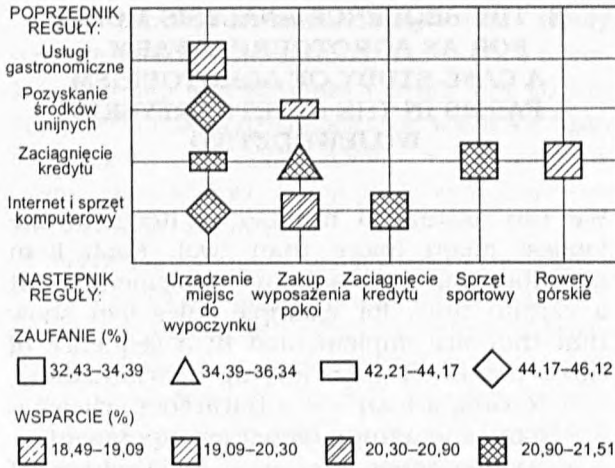
Chain length	Confidence (%)	Support (%)	Transaction count	Sequence rule
2	44.37	21.51	67	Internet and computer equipment == > Relaxation area facilities
2	42.57	21.49	63	Raising a loan == > Relaxation area facilities
2	35.81	21.13	53	Raising a loan == > Purchase of room furnishings
2	33.77	21.12	51	Internet and computer equipment == > Raising a loan
2	33.78	21.03	50	Raising a loan == > Sports equipment
2	48.08	21.01	50	Receiving EU grant == > Relaxation area facilities
2	32.45	20.59	49	Internet and computer equipment == > Purchase of room furnishings
2	33.80	20.28	48	Catering services == > Relaxation area facilities
2	32.43	20.17	48	Raising a loan == > Mountain bikes
2	42.31	18.49	44	Receiving EU grant == > Purchase of room furnishings

Source: Author with the help of SAS Enterprise Miner programme.

Table 6. Reguły sekwencyjne
(fragment tablicy z regułami sekwencyjnymi dwuelementowymi posortowanymi wg wielkości „wsparcia”)

Liczba pozycji w łańcuchu	Zaufanie (%)	Wsparcie (%)	Liczba transakcji	Reguły sekwencji
2	44,37	21,51	67	Internet i sprzęt komputerowy == > urządzenie miejsc do wypoczynku
2	42,57	21,49	63	Zaciągnięcie kredytu == > urządzenie miejsc do wypoczynku
2	35,81	21,13	53	Zaciągnięcie kredytu == > zakup wyposażenia pokoi
2	33,77	21,12	51	Internet i sprzęt komputerowy == > zaciągnięcie kredytu
2	33,78	21,03	50	Zaciągnięcie kredytu == > sprzęt sportowy
2	48,08	21,01	50	Pozyskanie środków unijnych == > urządzenie miejsc do wypoczynku
2	32,45	20,59	49	Internet i sprzęt komputerowy == > zakup wyposażenia pokoi
2	33,80	20,28	48	Usługi gastronomiczne == > urządzenie miejsc do wypoczynku
2	32,43	20,17	48	Zaciągnięcie kredytu == > rowery górskie
2	42,31	18,49	44	Pozyskanie środków unijnych == > zakup wyposażenia pokoi

Źródło: Opracowanie własne przy wykorzystaniu programu SAS Enterprise Miner.



Rys. 4. Ilustracja graficzna reguł sekwencji wymienionych w tab. 6
(źródło: opracowanie własne przy wykorzystaniu programu SAS Enterprise Miner)

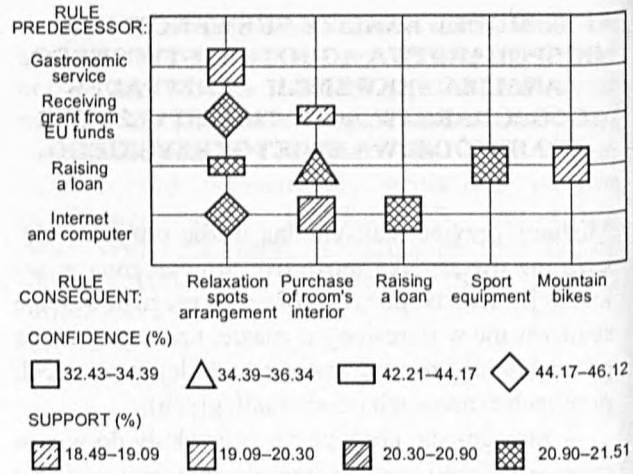


Fig. 4. Graphic illustration of the sequence rules included in Table 6
(źródło: opracowanie własne przy wykorzystaniu programu SAS Enterprise Miner)

7. PODSUMOWANIE

Z przeprowadzonej analizy możemy dowiedzieć się, w co przede wszystkim inwestowały gospodarstwa agroturystyczne. Rozkład inwestycji był następujący: najczęściej inwestowano w Internet i sprzęt komputerowy, wyposażenie pomieszczeń dla gości, sprzęt do rekreacji w domu i ogrodzie, zakup wyposażenia pokoi, rozbudowę bazy noclegowej, artykuły do wyposażenia łazienki, rozbudowę domu. Po zastosowaniu analizy asocjacji i sekwencji możemy stwierdzić, że zazwyczaj zakupowi samochodu towarzyszy rozpoczęcie usług gastronomicznych (oferowanie posiłków dla gości), zakup zmywarki do naczyń jest również zdarzeniem najczęściej towarzyszącym rozpoczęciu usług gastronomicznych. Z rozbudową bazy noclegowej najczęściej wiąże się pozyskanie kredytu i środków unijnych. Gospodarstwo, które kupiło konie, często rozpoczynało budowę stawu, gospodarstwo, które kupiło ziemię organizowało minizoo, zaś gospodarstwo, które inwestowało w Internet i sprzęt komputerowy pozyskiwało środki unijne.

PRZYPISY

¹ Współczynnik wzrostu (*lift*) wynosi 1, gdy szansa zdarzenia B nie zależy od tego, czy wystąpiło zdarzenie A. Gdy współczynnik wzrostu jest większy od 1, to szansa zdarzenia B zwiększa się, gdy otrzymujemy informację, że wystąpiło zdarzenie A. Stąd duża wartość współczynnika wzrostu oznacza, że gdy wystąpi zdarzenie A, to bardziej prawdopodobne jest, że wystąpi zdarzenie B, niż w przeciwnym przypadku. Ogólnie, zdarzenie B może występować rzadko, ale wystąpienie zdarzenia A zwiększa wtedy szansę wystąpienia zdarzenia B.

7. SUMMARY

From the analysis above the nature of investment on agrotourism farms can be discovered. The investment distribution was as follows: the majority invested in the internet and computer equipment, furnishings for guest rooms, equipment for home and garden recreation, room furnishings, accommodation extension, bathroom articles and house expansion. After implementing association analysis and sequence analysis, it can be stated that generally car purchase occurred together with an introduction of catering services (offering meals for guests). The purchase of a dishwasher was also an event usually co-occurring with catering services. Accommodation extension was most often connected with raising a loan or receiving a grant from the EU. Farms which bought horses often began to build a pond, those which bought land organised mini zoos, while those who invested in the internet and computer equipment received EU funds.

FOOTNOTES

¹ An area of ancient highlands in the centre and east of Poland.

² The lift factor equals 1 when a chance of event B does not depend on whether or not event A has occurred. When the factor is higher than 1, the chance of event B increases after we receive information that event A has occurred. Hence, the high value of the *lift* factor means that if event A occurs, it is more probable, that event B will occur. In general, event B may occur rarely, but the occurrence of event A increases the chance of event B taking place.

BIBLIOGRAFIA – BIBLIOGRAPHY

- BERRY, M.J.A., LINOFF, G.S., 2004, *Data Mining Techniques. For Marketing, Sales, and Customer Relationship Management*, Wiley Publishing, Inc, Indianapolis, Indiana.
- CLARKE, J., 1999, *Marketing Structures for Farm Tourism: Beyond the Original Provider of Rural Tourism*, Journal of Sustainable Tourism, vol. 7, Channel View Publications/Multilingual Matters Ltd.
- Data Mining Using SAS Enterprise Miner: A Case Study Approach*, 2003, SAS Institute Inc., Cary, NC, USA.
- DAVIDSON, R., 1993, *Turystyka*, Longman Group Limited.
- DRZEWIECKI, M., 1995, *Agroturystyka*, Instytut Wydawniczy „Świadectwo”.
- KRYŃSKI, Z., 2003, *Stan rozwoju bazy turystycznej w gospodarstwach agroturystycznych Regionu Karpackiego i Gór Świętokrzyskich*, *Ekonomika Rolnictwa*, 1, z. 9, s. 125.
- MLYNARCZYK, K., 2002, *Agroturystyka*, Wydawnictwo UWM, Olsztyn.
- PALKA, E., 2006, *Agroturystyka a ochrona środowiska w regionie Gór Świętokrzyskich*, [w:] *Funkcje turystyki i krajoznawstwa w strategii rozwoju regionów w Polsce*, red. B. Wójtowicz, Instytut Geografii Akademii Świętokrzyskiej w Kielcach, Kielce.