

МЕТОДИКА СБОРА И ОБРАБОТКИ ДАННЫХ ДЛЯ ОЦЕНКИ СТРУКТУРЫ ПОТРЕБИТЕЛЕЙ УСЛУГ ТУРИСТСКОГО КОМПЛЕКСА РЕГИОНА

Мартышенко Н.С.

Владивостокский государственный университет экономики и сервиса

Для разработки стратегий развития туризма региона необходимо иметь специальный инструментарий оценки структуры потребления туристских продуктов и анализа структурных сдвигов в процессах потребления. Такие оценки могут быть получены при обработке и анализе информации, отражающей основные показатели туристской деятельности. Укрупненные показатели можно найти в отчетах и статистических сборниках Федеральной и региональной служб государственной статистики. В этих материалах содержится важная информация по демографическому составу населения, социально-экономическим показателям по отраслям народного хозяйства, въездным и выездным потокам и многое другое. Эти данные не могут быть получены самим исследователем. Однако в них не содержится данных, характеризующих процесс потребления продуктов туристской индустрии – почему и как формируется спрос на туристский продукт, как он и изменяется во времени и почему. Эти данные могут быть получены только самим исследователем в процессе опроса потребителей и производителей туристского продукта, а также опроса экспертов [1, 2].

Для анализа структуры потребителей туристских продуктов в Приморском крае кроме официальной статистики использовались данные, полученные в процессе проведения мониторинга потребительского рынка туристских услуг, производимого в течение более семи лет на кафедре маркетинга и коммерции Владивостокского государственного университета экономики и сервиса. О масштабе исследований можно судить по данным таблицы 1. В таблице представлен объем анкет, введенных в компьютерную базу данных. В опросах использовались несколько модификаций 6-ти видов анкет. Процесс сбора, ввода в компьютер и обработки является непрерывным и осуществляется круглогодично.

Мониторинг туристских услуг, предоставляемых на внутреннем рынке, осуществлялся путем анкетных опросов населения края, по большей части жителей г. Владивостока (около 77% опрошенных).

Для повышения информативности данных различные анкеты содержат пересекающиеся модули вопросов или отдельные общие вопросы. Рассмотрим структуру анкет, использованных в исследовании туристской сферы Приморского края.

Рассмотрим данные, собранных по трем видам анкет:

- анкете по изучению пляжно-купального отдыха (приложение 1);
- анкете по изучению времяпрепровождения отпусков и каникул (приложение 2);
- анкете по исследованию туристского потенциала Приморского края и перспектив

его развития.

Первая анкета - “Исследование пляжно-купального отдыха”. По этой анкете было собрано самое большое количество данных. Анкета содержит пять модулей вопросов (всего 36 вопросов). Количество вопросов - величина весьма условная. Это количество, которое воспринимает респондент. На самом деле, некоторые вопросы могут иметь достаточно сложную форму, например, когда респонденту предлагается заполнить несколько позиций и даже таблицу.

Первый модуль включает социально-демографические характеристики респондентов - 13 вопросов;

Второй модуль отражает сезонно-временные характеристики использования отпуска за последние три года - 1 сложный вопрос;

Третий модуль содержит пространственные характеристики используемых для отдыха пляжных зон Приморского края (без ночевки) – 6 вопросов;

Четвертый модуль содержит пространственные характеристики используемых для отдыха пляжных зон Приморского края (с ночевкой) - 7 вопросов;

Пятый модуль содержит характеристики процесса потребления туристских услуг, сопутствующих пляжно-купальному отдыху – 9 вопросов.

Таблица 1 - Масштаб исследований

	Название анкеты	Количество анкет							
		2003	2004	2005	2006	2007	2008	2009	Всего
1	“Исследование пляжно-купального отдыха”	---	2345	1217	1423	1344	505	353	7187
2	“Исследование времяпрепровождения отпусков и каникул”	2200	585	---	---	---	1350	1205	5340
3	“Исследование туристского потенциала Приморского края и перспектив его развития”	---				1004	997	911	2912
4	“Исследование времяпрепровождения свободного времени студентов ”	---	494	184	280	148	40	95	1241
5	“Исследование предпочтений потребителей г. Владивостока”	---	222	728	---	---	---		950
6	Будущее Приморья - Родители о детях (учеба, здоровье, отдых, досуг подростка 14-16 лет)	---	400	1208					1608
	ИТОГО	2200	4046	3337	1703	2496	2892	2564	19238

Специфику анкеты отражают третий и четвертый модули вопросов. В анкетном опросе исследовались две разновидности пляжно-купального отдыха, практикуемого жителями края, - без ночевки и с ночевками. Изучались пространственно-временные характеристики отдыха, транспортная составляющая, средства размещения и другие показатели системы потребления туристских услуг на фоне социально-демографического портрета респондентов.

Вторая анкета – “Анкета по изучению времяпрепровождения отпусков и каникул”. Анкета содержит семь модулей вопросов (всего 29 вопросов). Первый и второй модуль повторяют вопросы анкеты “Исследование пляжно-купального отдыха”.

Третий модуль включает один сложный вопрос, оформленный в форме трех таблиц, в которых указываются основные характеристики использования отпускного времени (“ГДЕ?”, “КАК?”, “С КЕМ?”). Этот модуль является основным для этой анкеты.

Четвертый модуль состоит из пространственных характеристик, используемых туристских продуктов (2 вопроса). Вопросы этого модуля взяты из анкеты “Исследование пляжно-купального отдыха”.

Пятый модуль вопросов характеризует затраты потребителями денежных средств в отпускной период за последние два года. Здесь исследуются преимущественные направления поездок жителей региона.

Шестой модуль характеризует туристскую мобильность потребителя за последние годы.

Седьмой модуль отражает впечатления по последнему проведенному отпуску.

Третья анкета - анкета “Исследование туристского потенциала Приморского края и перспектив его развития” является относительно новой. По ней производились опросы в течение 2007 – 2009 гг. (собрано и обработано около 3 тысяч анкет). Эта анкета содержит первый, второй, пятый, шестой и седьмой модули вопросов анкеты “Исследование времяпрепровождения отпусков и каникул” (всего 34 вопроса).

Четвертый модуль вопросов содержит характеристики отпускного времени. По смыслу он пересекается с вопросами третьего модуля анкеты “Исследование времяпрепровождения отпусков и каникул”, но отличается по форме представления вопросов.

Остальные модули вопросов – третий, восьмой и девятый в других анкетах не встречаются.

Третий модуль отражает материальное состояние респондентов.

Восьмой модуль представлен вопросом по качеству жизни.

Специфику анкеты “Исследование туристского потенциала Приморского края и перспектив его развития” определяет девятый модуль вопросов. Это ключевые вопросы анкеты. Именно в девятом модуле исследуются перспективы развития туристской отрасли в Приморском крае. Первая модификация этой анкеты, по которой было опрошено более 1,5 тысячи респондентов, включала больше вопросов по перспективам развития туристской индустрии. Она включала модуль вопросов, который начинается словами: “Сформулируйте в нескольких предложениях свое мнение по развитию туристского потенциала туристского комплекса Приморского края”? Далее следовали вопросы:

1. Каким условиям должен отвечать городской пляж, при которых вы стали бы посещать пляж чаще, чем в настоящее время?
2. Сформулируйте свои предложения по совершенствованию организации транспортных услуг по доставке отдыхающих на городские пляжи
3. Какие мероприятия и административные решения, на ваш взгляд, могут повысить туристскую привлекательность Приморского края?
4. Какие первостепенные экологические проекты необходимо реализовать в Приморском крае для обеспечения экологической безопасности?
5. Какие объекты развлечения и досуга необходимо ввести в г. Владивостоке?

Структура рассмотренных анкет представлена на рис. 1. Сплошными линиями связаны повторяющиеся модули анкет, пунктирными линиями связаны частично совпадающие модули.

Рассмотрим отличия использованного подхода к разработке анкет от традиционного.

Первое отличие. При традиционном подходе основными вопросами являются вопросы с определенным списком ответов. В рассмотренных анкетах ключевыми вопросами являются открытые вопросы.

Как показали многочисленные опыты, человек более точно (и с меньшими затруднениями) отвечает на вопросы качественного, например, сравнительного, характера, чем количественного. Наиболее приемлемой формой для респондента является форма выражения мнения в виде произвольных высказываний.

Произвольную форму ответа допускают открытые вопросы. Поскольку исследователь заинтересован в получении информации, ему и приходится подстраиваться под респондента, предоставляя вопросы в такой форме, в какой респондент сможет или пожелает ответить.

В отличие от закрытых, открытые вопросы не содержат подсказок, не “навязывают” тот или иной вариант ответа и рассчитаны на получение неформализованного мнения.

Кроме открытых вопросов мы используем и полузакрытые вопросы, который кроме определенного числа вариантов ответа, содержат позицию “другое – укажите какое (что, где,

как)”. В теории социологических и маркетинговых исследований известны и иные формы открытых вопросов: “завершение предложения”, ”подбор ассоциации” и другие.

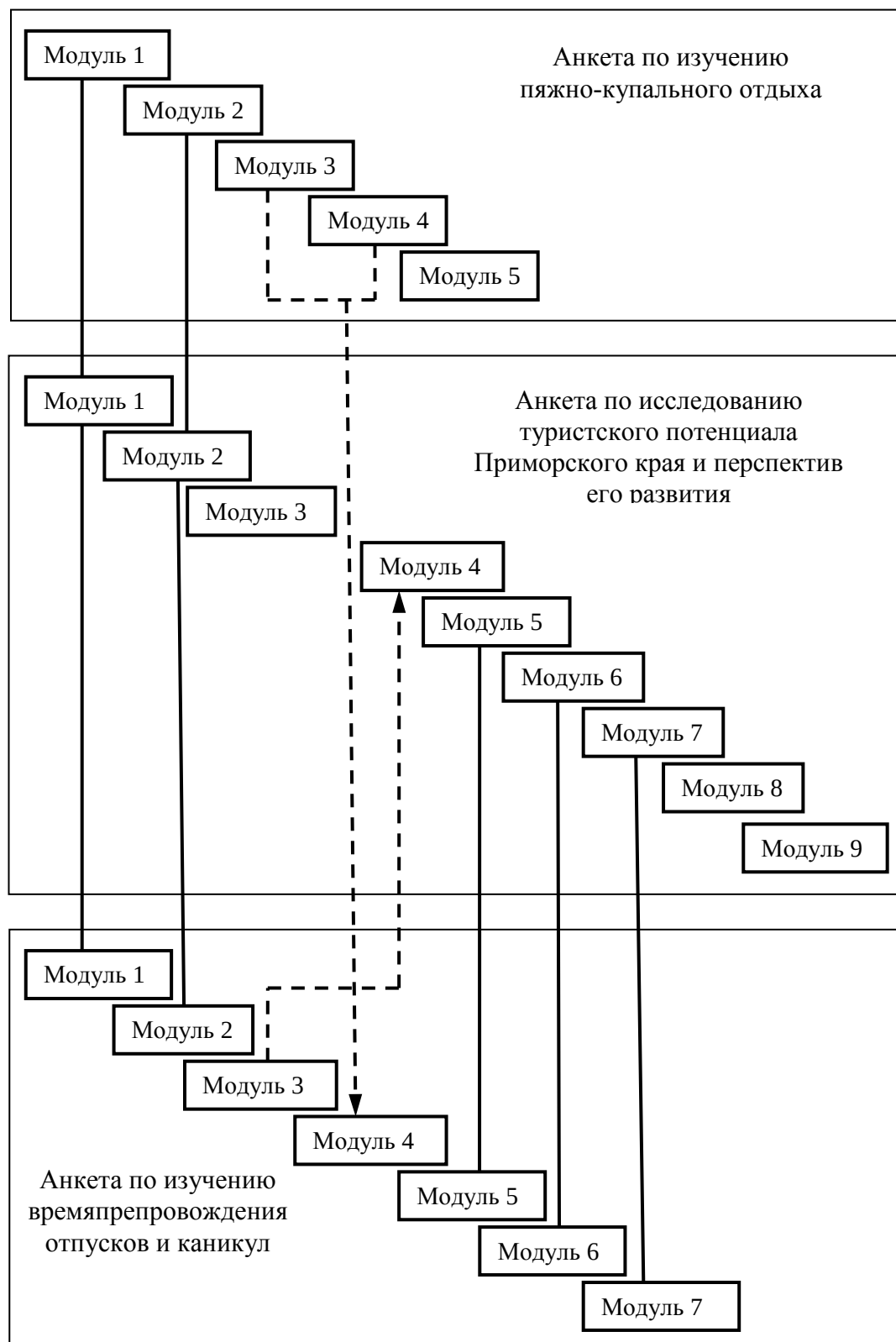


Рис. 1. Структура анкет

Второе отличие. Отличие состоит в большом разнообразии использованных измерительных шкал. Например, несколько ответов на один вопрос. Наличие такого

количества шкал вызвано стремлением получить от респондентов более достоверную информацию. Большинство исследователей не применяют компьютерную обработку данных открытых вопросов, а используют их в поисковых целях для получения информации для будущих исследований. Между тем, ответы на эти вопросы оказываются очень информативными.

Основная причина того, что исследователи мало используют открытые и сложные формы вопросов состоит в том, что они не обладают информационными технологиями обработки данных таких вопросов.

информацию. Для измерения одних и тех же свойств используются различные типы вопросов и шкал. Это позволяет производить сравнение различных подходов.

Третье отличие. В рассмотренных анкетах включены вопросы, предназначенные для исследования динамики процессов потребления туристского продукта. Примерами таких вопросов являются вопросы по продолжительности и сезону отпуска, вопросы по затратам денежных средств и т.п.

Четвертое отличие. Рассмотренные анкеты предполагают широкое использование для их обработки методов многомерного анализа.

Пятое отличие. Это отличие заключается в системном характере исследований. Программа исследований организована в форме мониторинга системы потребления туристских продуктов. Согласованность различных анкет обеспечивается включением в них типовых модулей. Такой подход позволяет, с одной стороны, повысить достоверность результатов обработки данных, с другой стороны, контролировать достоверность данных.

Шестое отличие. Это отличие является следствием предыдущего. В процессе исследований по мере накопления данных и результатов их обработки производится анализ качества самих анкет. Это позволяет производить корректировку форм представления вопросов, с тем чтобы они были наиболее понятны для респондентов.

Седьмое отличие. Особое внимание уделяется информации учетного характера: такой, как номер анкеты, дата проведения опроса, фамилии интервьюеров. Эта информация используется для расчета динамических характеристик, а также используется для контроля качества информации.

Перечисленные особенности приводят к тому, что в конечном итоге мы имеем дело с огромными массивами информации, большая часть которой носит нечисловой характер и неструктурированную форму представления. Большинство распространенных пакетов прикладных программ по обработке статистической информации, напротив, нацелено на обработку числовой информации. Поэтому для обработки таких данных потребовалась разработка специальных методов обработки данных и программных средств [3]. Совокупность постановок задач анализа структуры потребления туристских продуктов, методов компьютерного представления данных, методов обработки данных и специализированных программных средств составляют новый инструментарий оценки структуры потребления туристских продуктов. Разработка собственных методов анализа ни в коей мере не говорит о том, что мы не используем традиционные методы анализа данных. Наш подход дополняет и расширяет возможности распространенных методов анализа.

Новые методы анализа основываются на новой форме представления данных. Прежде, чем сформулировать отличие использованной формы представления данных, рассмотрим классическую форму представления. Современная классическая форма представления данных предполагает использование для обработки данных методов многомерного анализа.

Многомерный статистический анализ определяется как совокупность глубоко формализованных статистических методов, базирующихся на представлении исходной информации в многомерном геометрическом пространстве и позволяющих определять неявные (латентные), но объективно существующие закономерности в организационной структуре и тенденциях развития изучаемых социально-экономических явлений и процессов.

В настоящее время теория многомерного анализа числовых признаков достаточно проработана. Но методы обработки нечисловых данных находятся в стадии развития [4].

Таким образом, данные анкетных опросов можно рассматривать, как наблюдения многомерной случайной величины (табл. 2).

Таблица 2. - Табличная форма представления данных анкетного опроса

Номер	Признак 1	Признак 2		Признак j		Признак m
1	x_{11}	x_{12}		x_{1j}		x_{1m}
2	x_{21}	x_{21}		x_{2j}		x_{2m}
....						
i	x_{i1}	x_{i2}		x_{ij}		x_{im}
....						
n	x_{n1}	x_{n2}		x_{nj}		x_{nm}

Ответы на вопросы могут быть представлены в виде некоторой таблицы данных, в которой строки представляют собой объекты (анкеты), а столбцы - значения признаков (ответы на вопросы). Такая таблица называется таблицей "объект-свойство"[4].

Можно сделать вывод, что сама структура данных опросов содержит предпосылки применения многомерных статистических методов. Но применение того или иного метода требует соблюдения ряда условий или требований к данным. И в первую очередь согласованность метода с типом данных (числовые, ранговые). Одни методы предназначены для работы с одним типом данных, другие - с другим. Как правило, в анкетах содержатся вопросы, порождающие множество типов данных. В зависимости от наличия групп однотипных данных исследователь выбирает и методы их обработки. Более того, даже при составлении анкет исследователь должен ориентироваться на те методы обработки, которыми он владеет.

Отличие нашего подхода к форме представления данных состоит в том, что мы, кроме обычных признаков, допускающих только одно значение исследуемого свойства, допускаем наличие в таблице "объект-свойство" составных признаков. Составной признак возникает тогда, когда при ответе на вопрос респондент может указать сразу несколько ответов. Например, при ответе на вопрос анкеты: "Какие крупные города вы посетили за последние три года", респондент может указать сразу несколько городов. Таким образом, составной признак состоит, как бы, из нескольких простых ответов. Для идентификации составного признака в компьютерном представлении мы вводим какой-либо единый знак разделителя. Простой ответ может состоять из нескольких слов или даже может быть сформулирован в форме целого предложения.

Таким образом, ответы на открытые вопросы в общем виде можно представить в виде составных признаков. Такая информация является неструктурированной, и для того, чтобы ее можно было использовать для содержательного анализа ее нужно подвергнуть предварительной обработке, то есть осуществить переход от неструктурированного представления информации к структурированному, допускающему количественную обработку. Такой переход осуществляется с помощью операции типизации.

Операция типизации - это замена исходного простого высказывания (в форме текста) на близкое или сходное по значению, или обобщающее высказывание (в форме текста). Схема компьютерной технологии обработки неструктурированных данных приведена на рис. 2. При выполнении операции типизации формируется вспомогательная таблица – “список значений признака”. Один из столбцов такой таблицы включает все уникальные значения исходного признака.

Если типизации подвергается составной признак, то при расчете таблицы – “список значений признака” учитываются все простые высказывания из которых состоит сложное или составное высказывание. Таблица – “список значений признака” содержит столбец, в котором рассчитаны частоты встретившихся значений. Таблица снабжается автофильтром.

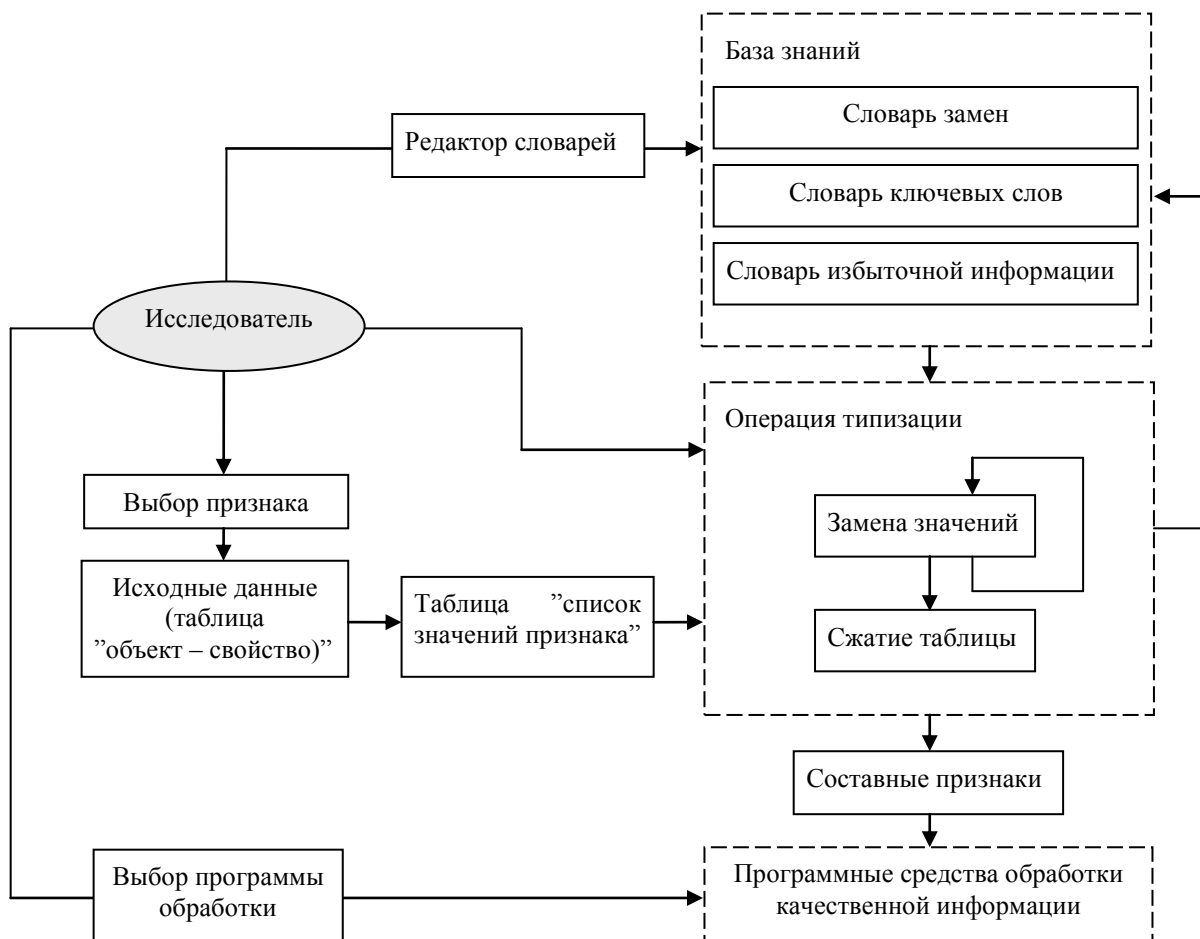


Рис. 2. Схема компьютерной технологии обработки неструктурированных данных

Операция типизации применяется не к исходным данным таблицы “объект-свойство”, а к данным таблицы “список значений признака”. В начале обрабатываются простые ситуации. Например, различное написание одного слова или различный порядок слов. Среди сходных высказываний выбирается наиболее удачная (или грамотная) форма написания высказывания, затем такое высказывание копируется в ячейки таблицы “список значений признака” со сходными высказываниями. Выполняя замену какого-то уникального высказывания на уже существующее из списка значений, мы, тем самым, сокращаем количество строк таблицы “список значений признака”. После выполнения серии замен целесообразно выполнять операцию “сжатия”, которая заключается в пересчете таблицы “список значений признака”. Постепенно таблица “список значений признака” сокращается и становится более наглядной.

После того, как простые ситуации обработаны, приступают к обработке более сложных случаев. В таблице “список значений признака” отыскивается группа редко встречающихся, но касающихся одной темы высказываний. Для этой группы простых высказываний исследователь подбирает в таблице некоторое обобщающее высказывание, и если такого не находит, то сам формулирует новое обобщающее высказывание, отражающее общий смысл или тему группы простых высказываний.

Например, отвечая на вопрос “Чем еще любите заниматься во время отдыха на море, кроме солнечных ванн и купания?” наряду с другими ответами различные респонденты давали такие ответы: “воспитание внуков”, “воспитание ребенка”, “играть с внуками”, “учить плавать детей”. Но эти высказывания встречались достаточно редко (менее 0,1%), поэтому мы заменили их на обобщающее высказывание - “заниматься с детьми”, которое нашли в таблице “список значений признака”. В принципе, смысл высказываний сохранился.

И все-таки, чтобы не потерять информацию, особенно при повторном проведении опросов, мы заменяем сходные высказывания на обобщающие, но с уточнением. Уточнение или нюанс указывается в скобках. Например, в рассмотренном выше случае, мы заменили оригинальные значения на:

- “заниматься с детьми (воспитание внуков)”;
- “заниматься с детьми (воспитание ребенка)”;
- “заниматься с детьми (играть с внуками)”;
- “заниматься с детьми (учить плавать детей)”.

Для нас важнее характер ответа, который определяет тип личности респондента (потребителя), а не конкретное содержание ответа. Если исходная таблица “список значений признака” может содержать до нескольких тысяч значений, то после обработки (типизации) данных такая таблица обычно содержит до трехсот значений с учетом значений с уточнениями. Созданием такой таблицы заканчивается **первый этап типизации** (первый уровень). Даже при автоматизации процесса работа требует достаточно много времени, опыта и большой внимательности от исследователя. И уж конечно, эта работа не может быть выполнена за один сеанс работы. Поэтому при завершении сеанса, результаты сохраняются, и в следующем сеансе работа продолжается с того места, где она была остановлена.

Полученный новый признак содержит все еще слишком много значений, чтобы его можно было анализировать с помощью числовых методов. Поэтому этот признак подвергается дополнительной обработке (**второй уровень типизации**). На этом этапе просто исключаются уточнения, содержащиеся в скобках, и формируется еще один столбец таблицы “список значений признака”, который мы называем подкласс, количество уникальных высказываний в котором будет уже от 30 до 50.

Наличие 30-50 вариантов значений - тоже большое количество для анализа измерений в номинальной шкале. Поэтому исследователь после формирования приемлемого списка действительно различных вариантов ответов, должен сгруппировать эти ответы, рассматривая их как некоторые характеристики непересекающихся классов, типов или тем, в зависимости от содержательного смысла признака и постановки задачи, для которой производится типизация. В нашем примере больше подходит определение типа личности потребителя туристских продуктов. Объединение простых высказываний в классы является **третьим уровнем типизации**. Для каждого класса исследователь сам формулирует название в зависимости от характера объединяемых высказываний.

На практике результаты группировки у разных исследователей получаются очень похожими. Различия могут возникать из-за того, что некоторые высказывания действительно могут занимать промежуточное состояние и могут быть отнесены сразу к нескольким классам. А вот названия классов каждый исследователь может дать совершенно разные.

Таким образом, в результате обработки данных открытого вопроса мы будем иметь (на выходе):

- три новых представления признака (свойства), которые включаются в исходную таблицу данных и могут быть подвергнуты дальнейшей обработке для получения содержательных выводов;

- таблицу “список значений признака”, которая может быть использована для обработки данных при повторении данного анкетного опроса или для выявления типизаций данных других анкет, которые предназначены для исследования данного процесса.

- базу знаний в форме трех словарей: “Словарь замен”, “Словарь ключевых слов”, “Словарь избыточной информации”.

Отметим, что в результате типизации составных признаков будут сформированы также составные признаки. Для их анализа разработаны специальные методы обработки [5, 6].

Результаты обработки анкетных данных по выявлению типологий, складывающихся в процессе потребления туристского продукта региона приведены в работах автора [7, 8]

Повышение эффективности работы компьютерной технологии обработки качественных данных достигается за счет создания и использования базы знаний. Компьютерные технологии, позволяющие использовать базы знаний, относятся к классу экспертных систем. Главная отличительная особенность экспертной системы заключается в умении делать правильные предсказания [4]. Производя всевозможные подсказки пользователю во время его работы, специальные программные средства позволяют существенно сократить время работы пользователя.

Подсказки пользователю осуществляются с помощью специальных словарей. Формирование словарей происходит в процессе работы пользователя над задачей типизации. Словари хранят опыт пользователя, приобретенный им в процессе решения задач типизации качественных признаков. Рассмотрим структуру и функции используемых словарей.

“Словарь замен” формируется автоматически при выполнении пользователем замен одних простых высказываний на другие. Словарь пополняется при работе пользователя с программой типизации. Словарь хранит все замены, выполненные пользователем. Все словари хранятся в единой базе данных Access. При сборе новых анкетных данных приходится опять производить операцию типизации. То есть пользователь должен подбирать аналоги для новых данных. Оказывается, что при повторном сборе информации, ситуации, обработанные пользователем на предыдущих этапах, в подавляющем числе случаев повторяются. Тогда, подключив “Словарь замен», пользователь получит подсказки по заменам. И тогда пользователю останется только обработать ситуации, которые ранее не встречались. Сопровождение словаря практически не требует дополнительных затрат времени. Время от времени целесообразно просматривать и редактировать словарь. Со временем в словаре начинают накапливаться неактуальные варианты замен. Такие записи необходимо удалять из словаря, потому что при очень больших объемах словаря, скорость работы программы типизации снижается. Пример “Словаря замен” представлен на рис. 3.

Рассмотрим функции “Словаря ключевых слов”. Этот словарь оказывается очень полезен в случае, когда фразы содержат много слов. При длинных фразах поиск подходящих синонимов в таблице “список значений признака” очень затруднен, потому что в начале работы список значений очень велик. Полное совпадение фраз встречается крайне редко. Поэтому часто приходится прибегать к фильтрам EXCEL, чтобы выделять более короткие списки, содержащие определенные сочетания отрывков фраз (рис. 4).

Такие фразы мы называем ключевыми словами, хотя отрывки фраз ключевыми словами являются условно. В таблице на рис. 4 выделены фразы по ключевому слову “экс”. В таком сокращенном списке такой фразе как “организ Экскурсии в заповедники” можно легко обнаружить подходящую замену: “разработать экскурсионные маршруты (по заповедным местам)”.

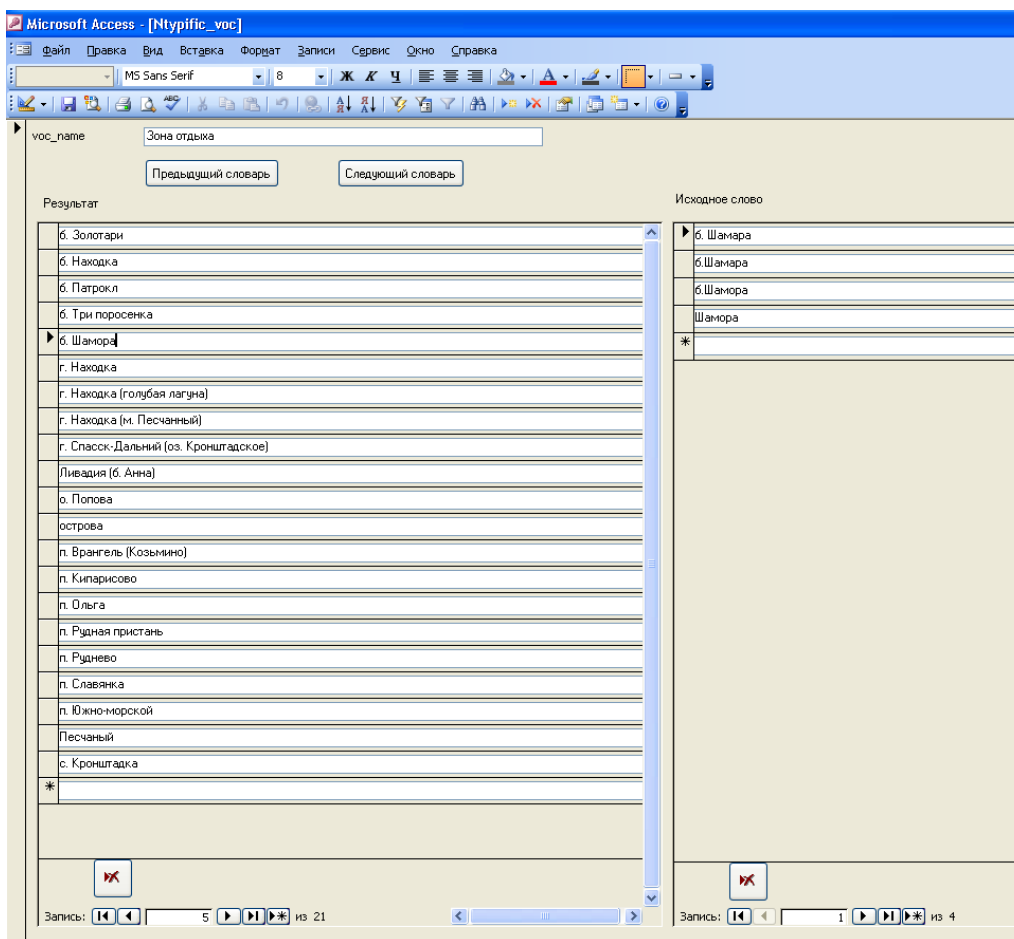


Рис. 3. – Пример “Словаря замен”

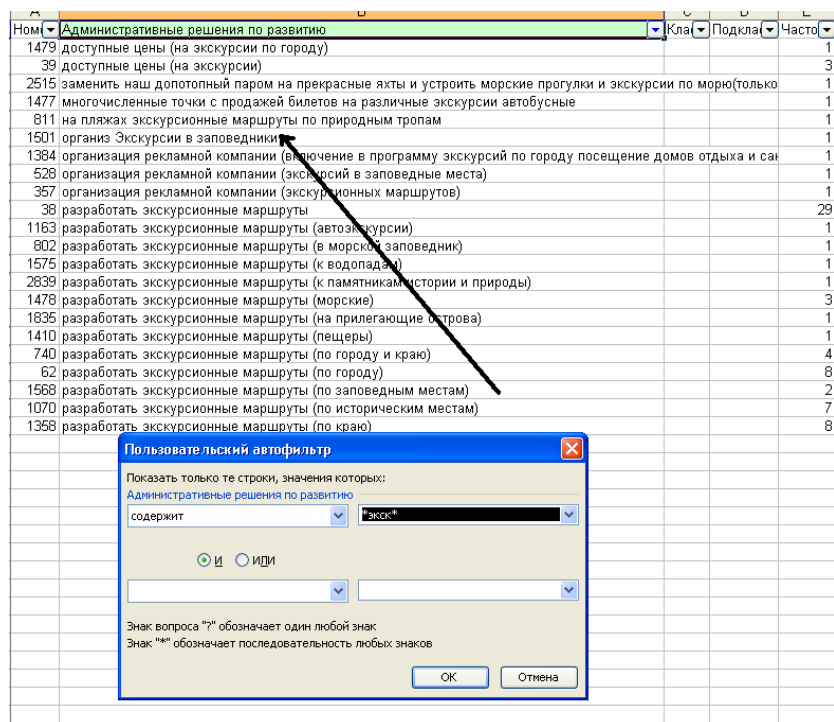


Рис. 4. Пример работы с пользовательским фильтром EXCEL при выполнении операции типизации

При такой замене смысл фразы совершенно не искажается. В “Словарь ключевых слов” заносятся фразы из таблицы ”список значений признака”, имеющие высокую встречаемость (частоту). Для этих фраз вносится список ключевых слов для поиска сходных фраз. Эти операции выполняются пользователем с помощью команд специального диалогового окна “показать словари ключевых слов” (рис. 5).

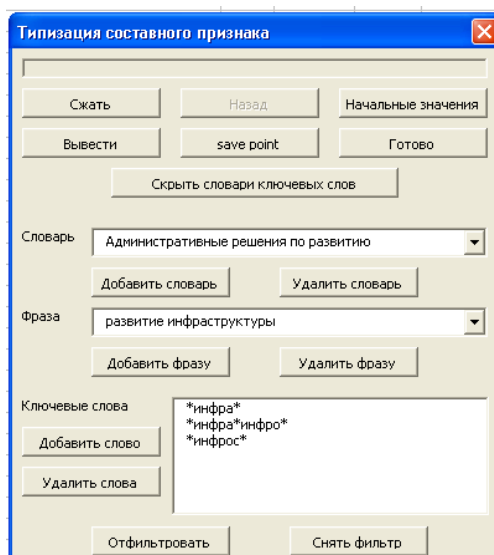


Рис. 5. Набор инструментальных средств для работы со “Словарем ключевых слов”

В диалоговом окне, представленном на рис. 5, представлены все функции по созданию, пополнению и редактированию “Словарей ключевых слов”. На рис. 6 представлены диалоговые окна, предназначенные для пополнения “Словарей ключевых слов”. Работа со словарем организована по принципу работы расширенного фильтра.

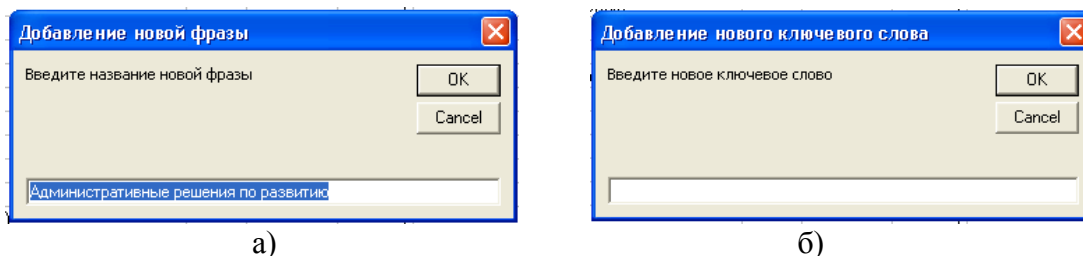


Рис. 6. Ввод новых фраз (а) и ключевых слов (б) в “Словарь ключевых слов”

Отличие работы словаря от работы расширенного фильтра состоит в том, что используемые ранее для поиска подходящих синонимов ключевые слова сохраняются в словаре. Этот словарь создается и пополняется только самим пользователем. Словарь создается для облегчения работы в будущем. Когда словарь будет содержать достаточное количество данных, его можно будет использовать, как базу знаний. При подключении этого словаря перед началом работы программы типизации в таблице ”список значений признака” формируется дополнительный столбец – “замены из словаря ключевых слов”. В этом столбце предлагаются замены фраз из “Словаря ключевых слов”. Предлагаемая фраза выбирается по наибольшему совпадению ключевых слов исходной фразы и фразы из “Словаря ключевых слов”. После просмотра, предлагаемых замен пользователь может принять только те замены, которые он считает подходящими. При этом неподходящие замены отвергаются.

Третьим словарем из базы знаний является ”Словарь избыточной информации”. В отличие от двух предыдущих словарей, которые создаются для каждого качественного признака, этот словарь работает сразу со всеми качественными признаками и даже с

различными анкетами. Словарь используется на первом этапе обработки качественной текстовой информации. С помощью этого словаря удаляются или корректируются высказывания, содержащие различную избыточную и несодержательную информацию.

Например, с помощью словаря могут быть исключены такие словосочетания, как: "Я думаю, что", "По моему мнению", "Это, в свою очередь" и т.п. В этот словарь включаются и слова с типовыми обобщениями. Например, сокращения "о" или "о." программа будет предлагать заменять на слово "остров". "Словарь избыточной информации" оказывается очень полезным при обработке длинных фраз и предложений. С помощью этого словаря удастся существенно сократить таблицу "список значений признака" на первом этапе работы, тем самым снизить трудоемкость обработки данных.

Все словари хранятся в одном файле Access. Компьютерная программа, обеспечивающая работу исследователя по выполнению операции типизации, обеспечивает связь таблиц EXCEL с базой данных Access.

После выполнения операции типизации все неструктурированные данные приводятся к структурированному виду и могут быть обработаны, как признаки, измеренные в номинальной шкале. Отличие в представлении данных состоит в том, что признаки остаются составными и соответственно для их обработки не применимы стандартные средства. Поэтому были разработаны несколько специальных компьютерных программ для анализа структуры потребления туристских продуктов.

Эти программы предназначены для решения двух основных задач:

- задачи выявления и анализа структуры потребления туристских продуктов;
- задачи анализа структурных изменений в потреблении туристского продукта.

Формирование базы знаний – это достаточно длительный и трудоемкий процесс. Но сформированные базы знаний не только существенно сокращают трудоемкость работы пользователя при повторных исследованиях, но и могут быть переданы другим исследователям, занимающимся в данной области, в виде файлов. То есть, можно наладить процесс обмена знаниями.

Методы выявления и анализа структуры потребления туристского продукта являются важнейшим инструментом при разработке стратегий развития туризма региона [9-11].

ЛИТЕРАТУРА

1. Артемова, Е.Н. Современные технологии стимулирования развития туризма в России / Е.Н. Артемова, В.В. Анпилогова // Российское предпринимательство. - 2007. - № 9. - С. 28-32.
2. Артемова Е. Н. Сравнительная характеристика самопозиционирования турфирм и их позиционирования на основе потребительских оценок / Е.Н. Артемова, В.А. Козлова, Л.И. Шмаркова / Территория науки: мультидисциплинарный научно-практический журнал. – Воронеж, 2008 г. - № 7 (8). – С. 110-121.
3. Мартышенко Н.С., Мартышенко С.Н., Кустов Д.А. Совершенствование математического и программного обеспечения обработки первичных данных в экономических и социологических исследований // Вестник Тихоокеанского государственного экономического университета. — 2006. — №2. С. 91–103.
4. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. – Новосибирск: Изд-во Ин-та математики, 1999. – 270 с.
5. Мартышенко Н.С., Мартышенко С.Н. Практические вопросы обработки анкетных данных // Практический маркетинг. — 2007. — №4. С. 2–8.
6. Мартышенко Н.С., Мартышенко С.Н. Методы обработки нечисловых данных в социально-экономических исследованиях // Вестник Тихоокеанского государственного экономического университета. — 2006. — №4. С. 48–57.

7. Мартышенко Н.С. Анализ транспортной составляющей в процессе формирования туристско-го кластера приморского края// Вестник ИНЖЭКОНа. — 2009. — №5. С. 107 – 117.
8. Мартышенко Н.С., Старков А.С. Методические основы оценки структуры туристско-рекреационного потенциала (на примере Приморского края) // Вестник национальной академии туризма — 2009. — №3 (11). С.21-27.
9. Мартышенко Н.С. Управление структурой туристского кластера на основе маркетингового подхода // Региональная экономика: теория и практика. — 2008. — №34 (91). С. 74 – 80.
10. Мартышенко Н.С. Формирование туристского кластера и управление его развитием на территории Приморского края // Регион: системы, экономика, управление. — 2008. — №1. С. 122–132.
11. Мартышенко Н.С. Принципы формирования туристского кластера в Приморском крае // Экономика региона. — 2009. — №1. С. 204 – 208.

Приложение 1

АНКЕТА по изучению пляжнооздоровительного отдыха

Место жительства _____ (населенный пункт)

Анкетирование провел _____

Дата _____ Номер _____

1. Пол: (подчеркнуть) [муж] [жен]

2. Возраст (подчеркнуть):

[до 20 лет] [20 ÷ 25 лет] [26÷30 лет] [31÷ 35 лет] [36÷40 лет] [41÷45 лет] [46÷50 лет] [51÷55 лет] [56 ÷60 лет] [более 61 лет]

3. Семейное положение (в браке): [да] [нет]

4. Наличие детей:

количество детей до 7 лет _____ количество детей 7-10 лет _____ количество детей 11-17 лет _____

количество детей 18-22 лет _____ количество детей 23-25 лет _____ количество детей более 25 лет _____

5. Тип жилья (подчеркнуть): [общежитие] [гостинка] [с подселением]

[многоквартирный дом] [собственный дом] [коттедж]

6. Количество комнат (подчеркнуть): [1ком.] [2ком.] [3ком.] [4ком.] [более 4 ком.]

7. Отношение к жилью (подчеркнуть): [муниципальное] [собственность предприятия] [свое] [съемное]

8. На одной жилплощади всего проживают: _____ человек;

9. Ваше образование (подчеркнуть): [8 классов] [среднее] [среднее специальное] [высшее] [ученая степень]

10. Специальность по последнему образованию: (для студентов специальность обучения) _____

11. Учусь (подчеркнуть): [нет] [высшее учебное заведение] [среднее учебное заведение] [другое] (что) _____

12. Форма обучения (подчеркнуть если обучаетесь): [очная] [заочная] [экстернат] [дистанционная]

13. Занятость (подчеркнуть): [работаю] [не работаю] [на пенсии] [на пенсии и работаю]

14. В каком сезоне у вас был отпуск (каникулы) последние три года (если отпуска не было в позиции дней указать "нет"):

(2009г. в месяце _____ дней _____) (2008 г. в месяце _____ дней _____) (2007 г. в месяце _____ дней _____)

ИСПОЛЬЗОВАНИЕ ПЛЯЖНОГО ОТДЫХА ПО ВЫХОДНЫМ ДНЯМ И В СВОБОДНОЕ ВРЕМЯ (БЕЗ НОЧЕВОК)

1. Чаще всего посещаю пляж: _____ (название бухты, острова или ближайшего населенного пункта)

2. Всего в сезоне 2009 года посещал пляжи (без ночевки): июнь _____ раз; июль _____ раз; август _____ раз; сентябрь _____ раз;

3. Обычно на пляже нахожусь: _____ часов.

4. Обычно добираться до пляжа (подчеркнуть): [пешком] [автобусом] [троллейбусом] [личным автотранспортом] [служебным автотранспортом] [электричкой] [рейсовым плавсредством] [личным плавсредством] [трамваем]

5. На дорогу до пляжа обычно затрачиваю: _____ часов.

6. Сколько человек обычно составляет ваша компания при посещении пляжа: _____ чел.

ИСПОЛЬЗОВАНИЕ ДЛЯ ОТДЫХА ПЛЯЖНЫХ ЗОН КРАЯ (С НОЧЕВОЙ)

1. Чаще всего посещаю зону отдыха: _____ (название ближайшего населенного пункта, бухты или острова)

2. Сколько дней обычно отдыхаете за одну поездку в пляжной зоне отдыха: _____ дней

3. Для ночевки использовал: [палатка] [комнаты турбаз] [гостиничный комплекс] [автомобиль] [катер/яхта] [жилплощадь в частном секторе] [жилье родственников и знакомых] [собственное жилье] [дом отдыха/санаторий] [другое] (указать что) _____

4. Обычно добираться до места (подчеркнуть): [пешком] [автобусом] [личным автотранспортом] [электричкой] [служебным автотранспортом] [рейсовым плавсредством] [личным плавсредством] [поездом] [другое] _____

5. На дорогу до места обычно затрачиваю: _____ часов.

6. Сколько человек обычно составляет ваша компания при посещении зоны отдыха: _____ чел

7. Сколько раз примерно посещал за последние три года известные зоны пляжного отдыха края:

а) бухта Шамора и прилегающие бухты _____ раз;

б) остров Русский _____ раз

в) острова (Попова, Рейнике, Путятин) какие другие острова " _____ ", _____ раз

г) полуостров Песчаный _____ раз;

д) бухта Ливадия и прилегающие бухты _____ раз;

е) район поселка Славянка _____ раз;

ж) Район поселка Безверхово и Перевозной _____ раз;

з) Район бухты Троица (Витязь, с. Андреевка) _____ раз;

и) Посьетский залив _____ раз;

к) бухта Находка _____ раз;

л) бухта Врангеля _____ раз; м) ДРУГИЕ ИЗЛЮБЛЕННЫЕ МЕСТА ОТДЫХА _____, _____ раз

ВОПРОСЫ ПО ВСЕМ ВИДАМ ОТДЫХА НА МОРЕ

1. Какие плавсредства использовались во время отдыха на море:

[надувные лодки] [лодка] [катер] [водный мотоцикл] [скутер] [яхта] [другое] (что) _____

2. Чем еще любите заниматься во время отдыха на море, кроме солнечных ванн и купания:

3. Пользуетесь ли вы какими-либо услугами, предоставляемыми на пляже (подчеркнуть чем):

[киоски по продаже продуктов питания] [кафе] [аттракционы] [стоянка автотранспорта] [прокат] [другое] (что) _____

4. Оборудован ли пляж элементарными удобствами (туалет, кабинки для переодевания, урны для мусора): [да] [нет]

5. Что омрачало ваш отдых в пляжной зоне: _____

6. В какую сумму вы оцениваете свои расходы на день отдыха на море на человека (без учета транспортных расходов) _____

7. Количество людей в посещаемой вами зоне с годами: [увеличивается] [остается неизменным] [уменьшается]

8. За последние годы вы стали посещать зоны морского отдыха: [больше] [на прежнем уровне] [меньше]

9. С течением времени благоустройство посещаемых вами зон отдыха: [улучшается] [остается неизменным] [ухудшается]

Приложение 1

Анкета по изучению времяпрепровождения отпусков и каникул

Место жительства _____ (населенный пункт)

Анкетирование провел _____

Дата _____ Номер _____

1. Пол: (подчеркнуть) [муж] [жен]

2. Возраст (подчеркнуть):

[до 20 лет] [20 ÷ 25 лет] [26 ÷ 30 лет] [31 ÷ 35 лет] [36 ÷ 40 лет] [41 ÷ 45 лет] [46 ÷ 50 лет] [51 ÷ 55 лет] [56 ÷ 60 лет] [более 61 лет]

3. Семейное положение (в браке): [да] [нет]

4. Наличие детей:

количество детей до 7 лет _____ количество детей 7-10 лет _____ количество детей 11-17 лет _____

количество детей 18-22 лет _____ количество детей 23-25 лет _____ количество детей более 25 лет _____

5. Тип жилья (подчеркнуть): [общеежитие] [гостинка] [с подселением]

[многоквартирный дом] [собственный дом] [коттедж]

6. Количество комнат (подчеркнуть): [1ком.] [2ком.] [3ком.] [4ком.] [более 4 ком.]

7. Отношение к жилью (подчеркнуть): [муниципальное] [собственность предприятия] [свое] [съемное]

8. На одной жилплощади всего проживают: _____ человек;

9. Ваше образование (подчеркнуть): [8 классов] [среднее] [среднее специальное] [высшее] [ученая степень]

На вопросы 10 и 11 отвечают лица, которые проходят обучение в учебных заведениях (если нет перейти к вопросу 12).

10. Учусь (подчеркнуть если обучаетесь): [высшее учебное заведение] [среднее учебное заведение] [другое] (что) _____

11. Форма обучения (подчеркнуть если обучаетесь): [очная] [заочная] [экстернат] [дистанционная]

12. Занятость (подчеркнуть): [работаю] [не работаю] [на пенсии] [на пенсии и работаю]

13. Должность: (для студентов специальность обучения) _____

14. В каком сезоне у вас был отпуск (каникулы) последние два года (если отпуска не было, в позиции дней указать "нет", если отпуск использовался частями, то заполнить две строки):

Первая часть отпуска (2009г. в месяце _____ дней) (2008 г. в месяце _____ дней)

Вторая часть отпуска (2009г. в месяце _____ дней) (2008 г. в месяце _____ дней)

15. Где вы провели отпуск в 2008 - 2007 годах (в таблице ГДЕ запишите названия, в таблицах КАК и С КЕМ знак ×):

ГДЕ?	2009	2008
Страна (заполняется, если не Россия)		
Край или область (если не Прим. край)		
Населенный пункт (или вблизи него)		
Район (не заполняется, если указан город)		

КАК?	2009	2008
Дома		
На даче		
В санатории, доме отдыха		
На турбазе		
В турпоездке за рубежом		
На берегу моря		
Навещал родственников		
Другое (что)		

С КЕМ?	2009	2008
Один		
С друзьями		
С другом/подругой		
С мужем/женой		
С детьми, без мужа/жены		
Всей семьей		
Другое (что)		

16. Укажите излюбленное место (а) для отдыха на море БЕЗ НОЧЕВКИ (название бухты, острова или ближайшего населенного пункта): _____

17. Укажите излюбленное место (а) для отдыха на море С НОЧЕВКОЙ (название бухты, острова или ближайшего населенного пункта): _____

18. Сколько примерно денежных средств (в тысячах рублей) было истрачено вами в отпуске на человека (подчеркнуть):

в 2009 г. [до 2 т. р.] [2-4] [5-7] [8-10] [11-14] [15-20] [21-30] [31-40] [41-50] [51-60] [более 60]

в 2008 г. [до 2 т. р.] [2-4] [5-7] [8-10] [11-14] [15-20] [21-30] [31-40] [41-50] [51-60] [более 60]

19. Есть ли у вас загранпаспорт: да/нет

20. Сколько раз вы были в Китае: [не был] [1 раз] [2 раза] [3-5 раз] [более 5 раз]

21. В каких городах Китая вы были: [СуйфэньХе] [Далянь] [Харбин] [Муданьцзянь] [Шеньян] [Пекин] [Шанхай] другие _____

22. Укажите страны мира (кроме Китая), которые посещали: _____

23. Какие крупные города России и мира вы посетили за последние три года: _____

24. Расположите в порядке вашего предпочтения турпоездки, если бы вам представился выбор (в позициях [] проставляется порядковый номер предпочтения 1, 2, и т.д., можно добавить свое направление в ответ – другое)

Тайланд []; Токио []; Сингапур []; Вьетнам []; Шанхай []; Пекин []; Индия [] другое _____ []

25. Что определило (или определяло бы) ваш выбор турфирмы: [рекламные объявления в СМИ]

[мнения родственников и знакомых, ранее пользовавшихся услугами турфирмы] [услуги турфирмы, не предоставляемые другими] другое, что _____

26. Куда планируете совершить поездку в ближайший год _____

27. Ваше любимое занятие во время отпуска _____

28. Что произвело на вас самое яркое впечатление в отпуске _____

29. Что омрачило вам отдых в _____

Аннотация

В работе рассмотрена методика оценки структуры потребления регионального туристского продукта. Для решения задачи используются данные анкетных опросов, в которых основное место занимают открытые вопросы. Рассмотрена методика перехода от неструктурированного представления информации открытых вопросов к структурированному.

Ключевые слова: регион, туристский продукт, открытый вопрос, обработка неструктурированных данных

Abstract

The technique of an estimation of structure of consumption of a regional tourist product is considered. For the problem decision the data of questionnaires in which the basic place is occupied with open questions are used. The technique of transition from not structured representation of the information of open questions to the structured is considered.

Keywords: region, the tourist product, an open question, processing of not structured data