Miroslav Foret

Marketing Research

Contents

FOREWORD

This textbook should provide at least a basic understanding of *methodological requirements and aspects* of marketing research. People often wrongly think that marketing research means a mere formulation of questions and a random distribution of questionnaires. In this textbook, however, we try to inform readers about possibilities of practical application of individual research techniques and methods of marketing research, above all with regard to the use of obtained results under conditions of concrete marketing activities. This means that we try *to acquaint students with problems of marketing research*. To make this textbook more illustrative and inspiring, we present twenty one *examples* of concrete, practical problems. These examples are based on results of research performed in the International Institute of Marketing, Communication and Entrepreneurship (IIMCE) in Brno, Czech Republic. To test how and how much the readers understood the explanation, all individual chapters are finished with more than eight tens of *discussion questions.*

Brno, January 2013

M. F.

1 MARKETING RESEARCH – ITS HISTORY, IMPORTANCE AND TYPES

Chapter objectives: Above all to emphasise importance of marketing research in connection with practical solutions of marketing problems. Six most important problems (themes) of marketing research are presented just at the beginning of this chapter. Thereafter the author explains basic types of marketing research and presents types of marketing data. Because the following text deals mostly with quantitative marketing research, the final part of this chapter informs students also about techniques of qualitative research and presents two concrete examples of qualitative marketing research.

Key words: marketing research, kinds and types of marketing research, quantitative marketing research, qualitative marketing research.

There is no doubt that efforts to learn more about customers and their needs stretch back to the past and have a very long history. It is very probable that already at the very beginning of the existence of human society the first manufacturers and vendors had to be interested in needs and requirements of their customers. Above all, they *observed* how the customers responded to their offer and how they selected the offered goods. At the same time, they also *listened to their customers and registered* what they needed and/or required.

In the course of the last century, the methods and procedures of cognition and research of customers and markets became more and more *elaborated and sophisticated*. Today, we cannot imagine that we would perform our business activities without using these important marketing tools.

From the thematic point of view, it is possible to define the following six basic domains of study:

1. *Studies on the socio-economic profile of customers,* i.e. efforts to learn more about customers and to describe their socio-demographic characteristics (sex, age, education, marital status, number of children, domicile, economic activities etc.);
2. *Studies on living conditions of customers*, i.e. surveys of their life standard resulting from their incomes and expenditures, ownership of movable and immovable assets and of households facilities and equipment; this research concerns above all consumer durables (appliances, electrotechnical products, automobiles, etc.);
3. *Studies on the lifestyle of customers*, which concern their working and out-of-service activities (e.g. recreation, sports, culture, education, tourism, household care, family care etc.);
4. *Studies on the value orientation of customers*, i.e. to learn about their beliefs and preferences, *weltanschauung*, life attitudes and orientation, political preferences etc.;
5. *Studies on the buying behaviour and decision-making,* i.e. efforts to determine which are the factors influencing decisions of customers and/or consumers in various life situations (above all when deciding which goods and services they are interested in and how much important

for them are such factors as quality, price, distribution or promotion);

6. *Studies on the perception and effects of marketing communication*, i.e. how they watch mass media and how they are influenced by marketing and business communication.

In all these domains the efforts of researchers should be focussed above all on *detection and unfolding of new problems* (i.e. needs, expectations, requirements and degree of satisfaction). Although the customers need not be fully aware of such problems, they can become a key challenge, possibility and direction of further business orientation.

1. 1 The essence and definition of marketing research

The efforts to learn something about customers should be always based on the application of *objectivised and systematic methods* created and developed by the current marketing research. For that reason this chapter briefly deals with the subject of *marketing research and its essence.* To explain it, the following two definitions may be cited:

- "Marketing research provides information helping marketing managers to distinguish and respond to marketing opportunities and threats ". /Tull & Hawkins; 1990, p. 5/
- „Marketing research involves planning, collection, and analysis of data which are relevant for marketing decision-making and presentation of results of this analysis to managers." /McDaniel & Gates; 2002, p. 6/

Marketing research provides *empiric data* about the situation in the market, above all about our customers. As already mentioned, these data are obtained on the base of objectivised and systematic *methodological procedures,* which respect specific features of the complex socio-economic reality.

So, for example, the methodological objectivity and the systematic approach of marketing research are projected into its understanding as a process (see the following Chapter 2), as well as into the methods of recognition (measuring) of socio-economical phenomena, which differ from natural, technical and other factors.

Second, however, we must be aware of the fact that even when using these objectivised and systematic methodological procedures it is possible to explain this complex socio-economic reality only *partially* and fragmentary. This partiality of results of marketing research stems above all from a *reduction of the complex socio-economic reality*, as it is performed when defining the subject of marketing research (see Chapter 3) and trying to operationalise it. In this context it is necessary to say that from this complex socio-economic reality we select only some, for us the most important, problems (phenomena). When analysing these complex phenomena, we study only some of their aspects (properties) and when doing this analysis, we use only some selected tools (i.e. concrete variables and/or variables).

The definition of its object is another part of the methodological reduction of empiric marketing research. In this case, again, it is necessary to select only a certain part (or segment) of the aforementioned very complex reality (see Chapter 5) and, moreover, we must divide it into individual units that will by finally investigated. However, for us the obtained results are interesting as a whole object, not as individual units.

Third, the partiality of results of marketing research results also from the very method of field data collection, which is performed in certain time intervals, sometimes even very long (e.g. once in ten years as in case of population census or houses and flats census). Data collected in this

way show a character of discrete (non-continuous) variables while that of socio-economic phenomena is continuous so that they are continually changing in real situations.

1. 2 Historical roots of marketing research

The history of marketing research is relatively long and stretches back to the 19th century. The first empiric survey of voters' behaviour and their decision-making took place in the course of presidential elections in the USA already in 1824. One hundred years later the generation of young researcher led by G. Gallup and E. Roper introduced into the practice statistically processed methods of selection of representative samples of respondents. In 1940s, P. F. Lazarsfeld and B. R. Berelson described the first explanation models of behaviour in their monographs *Voting* and *The People´s Choice*. These studies dealt above all with the manner and intensity of effects of opinion leaders on the decision-making of voters. Very soon, this knowledge of bandwagon and boomerang behaviour of voters was transferred into the practice of marketing research (above all when modelling the behaviour and decision masking of customers). Due to this fact these studies are rightly considered as the beginning of the marketing research. (McDaniel & Gates 2002, pp. 17 – 19).

This briefly outlined history also indicates that the marketing research follows up traditions of sociological research and studies of public opinion. Namely, the methodological starting points of this research of socio-economic phenomena and the applied statistical methods of processing of collected data are identical in all three aforementioned cases of empiric socio-economic studies. The differences among are determined above all by their *subject*. As already mentioned, marketing research deals above all with markets (customers, suppliers, and purchasers), sociological research is focused on social problems (life style, unemployment, family), and studies of public opinion investigate current social and, above all, political issues (electoral preferences etc.).

In contradistinction to *market research*, which represents a short-term activity enabling to analyse and evaluate a momentary situation in the market by means of selected research methods, the *marketing research* represents a long-term activity, which combines several research techniques, uses more complex statistic methods, compares and evaluates results obtained from several different sources, and reveals deeper relationships and pieces of knowledge (e.g. elaboration of already mentioned models of behaviour. Similarly as in case of market research, *the research of customers and their need will be understood as a simplified variant of marketing research*, i.e. as a „survey". The results of such surveys only *describe* phenomena occurring in a given domain. Such surveys deal with problems, which are interesting for marketers in a given situation and investigate *how* it looks like *now*. Market research does not look for deeper causal relationships and/or reveal developmental trends and tendencies.

1. 3 Some basic kinds and types of marketing research

When speaking about this subject, it is necessary to distinguish between *primary* and *secondary* marketing research. The *primary data* are collected or observed directly from respondents. These data are collected directly in the field and the research can be performed either by marketers themselves of by hired agencies. This means that the primary marketing research is a process, and this will be described and discussed in following chapters.

On the other hand, the *secondary marketing research* means usually a supplementary, subsequent use of data that were already collected by someone else in the course of primary marketing research, above all in the form of their modified statistical processing and new

interpretation; these data were collected for other purposes and other clients and usually are not fully relevant. However, in case of secondary research there is a substantial difference between situations when there are available either *non-aggregated data* (i.e. original ones concerning directly each research unit) or *aggregated data* (i.e. those that were already summarised and/or processed to provide statistical values (percents, means, coefficients of variability etc.). As it will be discussed below aggregated data can be used only for comparisons of different objects within identical time intervals or for the evaluation of the development of the same object in the course of different time periods or, possibly, for a combined application of both.

The main advantage of non-aggregated secondary data consists in the fact that they concern directly individual research units (i.e. respondents) so that they can be repeatedly processed by means of statistical methods corresponding with our needs and requirements. However, to obtain such non-aggregated data is more costly and also more lengthy (e.g. from archives of social research studies or by purchasing them directly of research institutions) than to acquire aggregated data, which have been already published (e.g. by the Czech Statistical Office, in final reports of marketing research, in professional magazines and journals etc.).

Example 1 Further processing of non-aggregated secondary data

In the final report of the survey of Znojmo population, there is the following table of distribution of relative frequencies of answers to the question concerning availability and numbers of playground for children:

Variant	Classification of the first degree					Total
	Definitely YES	More likely YES	More likely NO	Definitely NO	I don't know, I can't say	
Total	11 %	27 %	33 %	15 %	14 %	100 %

Table 1 – 1 Relative distribution of frequencies of the first degree classification

As one can see in this table, the in the opinion of respondents the numbers of playground are more or less low and they do not think that they are sufficient. The secondary processing of data aggregated in this way enables to obtain only broader categorisation (see Section 6. 1 in the Chapter Categorisation and, especially, Table 6 – 2). In this case we could learn that 38 % of respondents think that the number of playground in Znojmo is sufficient while 48 % of them are of the different opinion. The remaining 14 % could not answer this question. These aggregated data could be also presented as various (column or pie) graphs but this would be all what can we do with them from the statistical point of view.

If, however, there are available non-aggregated data (as it is fortunately in this case), they can be statistically processed in a greater detail. In such a case we can perform the categorisation of the second degree in which we will try to ascertain if there are not sex-dependent differences in answers of individual respondents. The resulting contingency table looks as follows:

Sex	Variants					Total
	Definitely YES	More likely YES	More likely NO	Definitely NO	I don't know, I can't say	
Females	13 %	23 %	35 %	16 %	13 %	100 %
Males	10 %	32 %	31 %	14 %	13 %	100 %
Total	11 %	27 %	33 %	15 %	14 %	100 %

Table 1 - 2 A contingency table of relative row frequencies

This table indicates that females (mothers and grandmothers) view the situation in a more critical manner (in the opinion of 51 % of them the numbers of playgrounds are not sufficient) than males (45 % of them answered that their numbers are not adequate). The fact that this statistical relationship is only weak was corroborated by means of subsequently calculated Pearson's contingency coefficient (its value was only 0.1).

In any case, however, also aggregated data are needed and useful in secondary marketing research; so, for example, they enable to *compare* different wholes (e.g. countries, regions, towns or organisations).

Example 2 Use of aggregated secondary data when comparing individual objects

An example of an international comparison of different objects (in this case of Czech towns and English Ipswich) and also of secondary research based on the use of aggregate data is presented in Tab. 1 – 3 which was performed within the framework of our international project Communicating Town (see www.komunikujici-město.cz):

Satisfaction with municipality activities	**Ipswich**	**Czech towns**
Satisfied	68 %	40 %
Unsatisfied	4 %	34 %
Neither satisfied, nor unsatisfied, I don't know	28 %	26 %
Total	**100 %**	**100 %**

Table 1 - 3 Comparison of the degree of satisfaction with activities of municipal offices

As one can see, people living in Ipswich are more satisfied with activities of their municipal office than inhabitants of similar Czech towns because approximately one third of them unsatisfied. Results of this international comparison thus indicate that Czech municipal offices can and should improve their work.

When presented as time series, aggregated data can *express the development*, dynamic changes and trends of social phenomena under study.

Example 3 Use of aggregated secondary data presented in the form of a time series for the expression of development

When performing regular annual surveys among Moravian Karst visitors, the following results were obtained (Tab. 1 – 4):

Year	Brno	Prague	Bohemia	Moravia	Outland
1991	26 %	10 %	23 %	23 %	18 %
1992	20 %	9 %	26 %	27 %	19 %
1993	23 %	9 %	26 %	30 %	12 %
1994	15 %	7 %	28 %	39 %	11 %
1995	28 %	9 %	21 %	31 %	11 %
1996	28 %	6 %	22 %	36 %	8 %
1997	21 %	9 %	32 %	31 %	7 %
1998	11 %	8 %	41 %	31 %	9 %
1999	34 %	8 %	16 %	36 %	6 %

Table 1- 4 The domicile of Moravian Karst visitors

As one can see, in 1990s the numbers of foreigners and Prague inhabitants gradually decreased while those of Moravians and visitors from Brno increased. This example also indicates the necessity and usefulness of regularly performed surveys because they enable to register developmental changes.

The importance of secondary aggregated data for surveys of customers and of their opinions concerns above all the following three domains:

1. First, statistical data can provide *introductory (preliminary) information about the problem under study* already in the initial stage of marketing research *;*
2. Second, the subsequent *selection of study sample* and its components (units) cannot be done without this information. This concerns above all quota sample (including the final *evaluation of representativeness* - see the Chapter 5));
3. Third, in the final stage of interpretation of marketing research results these aggregated statistical data enable for example to calculate (estimate, extrapolate) values concerning investigated problems as collected within *the whole population (the whole set)* under study.

In any case, however, we should be fully aware of a substantial difference between the recognition of technical and natural on the one hand and situations existing in the domain of socio-economic events on the other. The real features of abiotic nature, animals and plants can be studied more easily that the world of man and of human social systems (households, organisations etc.). Technical and natural phenomena can be measured and evaluated more exactly and they have been already studied for millennia. We know well such natural laws and regularities as for example exact physical and mathematical laws of celestial mechanics. Using them, we are able to predict movements of celestial bodies for many incoming decades. In case of socio-economic phenomena, however, the situation is much more complicated. In this domain it is not possible to describe processes of buying behaviour of people in such a degree that we will be able to predict what the customers will buy tomorrow. Moreover, it is necessary to take into account also effects of such sudden and unexpected changes in the market conditions as for instance devaluation of currency.

The primary marketing research can be divided into two categories: quantitative studies and qualitative studies. *Quantitative studies* deal with extensive sets of hundreds or even thousands of respondents and try to analyse sufficiently large and representative samples. These studies try to record opinions (cognition) of people in an as-standardized-way as possible. The obtained results are processed by means of statistic methods and thereafter they are generalized for the whole basic set of people (i.e. the whole population). Techniques of quantitative marketing research involve not only interviews, observations, experiments and questionnaires but also content analysis of texts and other symbolic statements. Quantitative studies are lengthier and also more expensive but usually they provide results in a well arranged numeric form which better describes and characterizes the extensive sample of respondents or even the whole population.

On the other hand, *qualitative studies* enable to learn more about the motives of people and about their behaviour. They reveal the essence and coherence of their beliefs, opinions, preferences, and attitudes and try to trace up their causes. The basic techniques of qualitative marketing research involve:

• *Individual in-depth interviews*, by means of which the inquirers try to ascertain and describe deeper causes of certain beliefs and/or certain behaviour of people. The interviewer uses well formulated questions and tries to stimulate the respondents in such a way that they would present their own opinion; the interviewers attentively listen to respondents and record their spontaneous narration. Finally, the whole interview is processed, analysed and summarised. When performing individual in-depth interviews, the inquirers use also so-called *projective techniques*, which should evocate associations and stimulate respondent's imagination. These projective techniques involve tests of word association, finishing of sentences, interpretation of pictures etc.

• In *focus groups* the discussion of 10 to 20 people is moderated by a leader (moderator); the sample of respondents is selected with regard to the structure of target population. The main purpose of this technique is to follow up the process of group behaviour and its effects on formation and confrontation of individual opinions of group members.

As compared with quantitative research, qualitative studies are quicker, less expensive and, above all, it is easier to perform them. However, it has also some significant limitations because the sample of respondents is small and non-representative. The obtained results cannot be generalised and related to the whole population. This research method is very suitable for the first contact description of the problem and possibly also for a subsequent deepening of results of quantitative research.

The qualitative marketing research helps to reveal the *uniqueness and individuality* of respondents. Its results then reflect more intensively the differences among them. On the other hand, the quantitative research reveals (thanks to its *standardization and unification*) more likely their similarities.

When considering advantages and disadvantages of these two types of marketing research (i.e. quantitative and qualitative studies) none of them should be preferred because they provide different views of the problem under study. In practice, they are often combined; so, for example, at first we use the qualitative techniques, which enable to formulate the problem more precisely and/or to improve the formulation of questions using a small group of respondents. The quantitative approach is then used in the next step and this enables to obtain more representative data and/or results. Another possibility is to start with a quantitative survey and thereafter to continue qualitative techniques, which help to concretize and elaborate results of performed extensive marketing research.

Textbooks of marketing research (including this one) are based above all on the application of quantitative studies. However, the following two examples illustrate advantages and benefits of qualitative marketing research. This techniques are substantially easier and does not require a laborious collection of field data (sufficient is already a small sample of respondents). The statistical processing and evaluation of results is also easier (often a mere summarisation of results is quite sufficient). Nevertheless, this approach can provide beneficial and stimulating conclusions. This is illustrated in the following example.

Example 4 Two qualitative marketing researches

The objective of the first example of practical application of a qualitative method of marketing research (or, more precisely, the technique of *focus groups*) was to obtain more detailed information about opinions of normal consumers about four samples of wine. Evaluated were the following wines: Malverina from the wine villages Sádek and Perná and Müller-Thurgau and from Sádek and Perná. The selected respondents tasted these four samples without any information that might lead to conscious or unconscious bias on their part (in the so-called blinded experiment). Bottles were labelled only with letters A, B, C and D. At first the respondents evaluated the following three properties of each sample:

1. Taste
2. Colour
3. Smell.

Thereafter they performed the paired comparison of wines originating from different regions (Perná and Sádek) and finally they evaluated all of them together and selected the best one.
In the two focus groups involved altogether 32 respondents (20 males and 12 females) from Bohemia, Moravia and Slovakia; their age and education ranged from 20 to nearly 70 years and from the primary school to university, respectively.

The processing and generalisation of their answers was performed according to steps mentioned above: At first they evaluated taste, colour and smell of each sample. Thereafter they compared the same varieties originating from different regions (Perná, Sádek) and finally they selected the best one of all.

Malverina from Sádek was evaluated as the best wine by 19 respondents, Müller-Thurgau from Sádek by 5, Müller-Thurgau from Perná by 4 and finally Malverina from Perná by 3. One respondent answered that the taste of all four samples was uniform.

As far as the colour of wine was concerned, Malverina from Sádek was again evaluated as the best one (17 respondents). Colour of Malverina and Müller-Thurgau from Perná was appreciated by 7 and 6 respondents, respectively, and Müller-Thurgau from Sádek by only 1. One respondent again answered that the colour of all four samples was the same.

The smell was evaluated as follows: Malverina from Sádek was evaluated as the best by 18 respondents, Müller-Thurgau from Perná by 8, Malverina from Perná by 3 and Müller-Thurgau from Sádek by 2. One respondent considered the smell of all four samples for identical.

When comparing the aforementioned characteristics of all four samples together, the respondents preferred Malverina (especially from Sádek).

The client (Wine and Cultural Centre Sádek) obtained in this way valuable marketing information and *empiric data about the attitude of customers to evaluated wine samples as well as about their opinions concerning evaluation of wine*. In spite of the fact that these results were surely different from the evaluation of producers (i.e. wine growers) and/or of experts evaluating wine at various competitions, it can be concluded that the knowledge of these opinions may stimulate the wine growers to modify their products according to the taste and requirements of their customers and to use the tools of marketing communication when informing people about (until now not fully appreciated) qualities and virtues of their products.

Ve druhém příkladu jednoduchého kvalitativního výzkumu byly naopak použity v prvé řadě techniky *non-standardized observation and individual in-depth interviews* s představiteli radnic navštívených měst a obcí jižní Moravy, stejně jako s místními producenty vína a majiteli vinných sklípků. Nedílnou součástí byla rovněž analýza nejrůznějších propagačních materiálů a produktů zmíněného projektu.

The main tourist route (the so-called backbone or “flag” trail) connects the towns of Znojmo and Uherské Hradiště. Its total length is 280 km and is symbolically called the *Moravian Wine Trail*.

Map of Wine Trails and Wine Growing Regions

Moravian Wine Trail - 280 km (Uherské Hradiště - Znojmo, red)
Brno Wine Trail – being prepared (medium blue)
Znojmo Wine Trail - 163 km (deep green)
Mikulov Wine Trail - 82 km (yellow-green)
Velké Pavlovice Wine Trail – 109 km (light blue)

Podluží Wine Trail - 113 km (orange)
Mutěnice Wine Trail - 60 km (rose)
Kyjov Wine Trail - 84 km (light green)
Bzenec Wine Trail - 24 km (silver)
Strážnice Wine Trail - 101 km (deep blue)
Uherské Hradiště Wine Trail - 74 km (violet)
Total 1090 km

The backbone trail is connected with another ten shorter trails that bear the names of wine growing regions through which they stretch, i.e.: *Brno Wine Trail, Znojmo Wine Trail, Mikulov Wine Trail, Velké Pavlovice Wine Trail, Podluží Wine Trail, Mutěnice Wine Trail, Kyjov Wine Trail, Bzenec Wine Trail, Strážnice Wine Trail*, and *Uherské Hradiště Wine Trail*. Their length is mostly about 100 km and only the *Bzenec Wine Trail* is a little shorter (24 km).

The logo of all Moravian Wine Trails is a picture of a small wine cellar and it is presented on all signposts, orientation tables and maps. Each trail has its own colour and this colour is used in all signs, legends, graphic symbols and other printed materials. For example, the colour of the Moravian Wine Trail is deep red, that of Brno Wine Trail blue, that of Znojmo deep green etc.

As mentioned above, the visitors can use a set of 11 maps, which provide information about all historical, cultural and natural monuments situated along individual trails. It is a pity that these maps are printed in different scales, obviously due to differences in and in dependence on the total length of individual trails so that their appearance is different, as well as their informative value. The scale of the majority of these maps is 1:50,000 but some are printed in the scale 1: 30, 000 while several others have the scale 1: 40,000, 1: 65,000 and even 1: 80,000. Unfortunately, there are also some gramatic/spelling errors in them, which should be removed and corrected in their possible future editions. The yellow colour of signposts and orientation tables, which is used for marking of all trails in the countryside, is problematic due to its impaired legibility and the graphical form of the logo of Moravian Wine Trails also should to be

a little modified. On the other hand, however, it is necessary to appreciate that their localisaton along individual trails is much better than that of other tourist signpost used until now in the Czech Republic. There is also a danger of the occurrence of a partly chaotic situation because all these cyclotrails are marked up in two different ways: either as trails with common signs for cyclists (i.e. yellow tables with black inscriptions and a black pictogram of a bicycle) or with special signs for cyclotourists (i.e. yellow-black signs with additional red arrows).

In addition to these specialised and local trails, the South Moravian wine growing region is crossed by other two important cyclotrails, namely the *Moravian – Silesian Trail* (Cracovia-Brno) and the *Amber Trail* (Brno – Vienna). In the spring of 2004, the Amber Trail was supplemented with new orientation tables with the inscription "Greenways Cracovia – Moravia – Vienna". Another problem exists in the neighbourhood of the territory of the Battle of Three Emperors 1805 near the town of Austerlitz (Slavkov u Brna) where the visitors can find another type of yellow tables with blue squares indicating the localisation of fighting French, Austrian and Russian troops.

This situation definitely requires a correction and it is necessary to try to coordinate activities of all participating institutions, to reduce the number of signposts (which are often doubled), and to try to unify the overall system of tourist marking in the whole territory of South Moravia. As an model it is possible to use the cycloroute Brno – Vienna, which mingles with Moravian Wine Trails because it has the following great advantage: there are several resting places along it with maps, orientation tables, photographs, marking of distances and litter bins that altogether increase the comfort of visitors. Similar accessories and the system of marking should be used also along other wine trails in South Moravia.

However, the crucial problem of Moravian Wine Trails lies in the fact that, for the time being, this project is *de facto* non-functioning because until now it has not been possible to establish a feeling of partnership among local people, enterpreneurs (wine growers) and local authorities. This unsatisfying situation can stem from the fact that the *mission* of this project has not been clearly formulated. It is necessary to emphasise that it is not possible to propagate in the local polulation an idea that the passing cyclists are the same customers and consumers of wine and sauser (young, half-fermented wine) in the same extent as participants of bus trips from Bohemia and/or North Moravia. It is true that these trails route very often through fields, vineyards, forests and/or meadows (i.e. outside of normal traffic) but also here the tumbles can be often of a very unpleasant, nasty and even dangerous nature so that it is not necessary to increase this risk through the consumption of wine.

In our opinion the importance of Moravian Wine Trails for cyclists lies not in the consumption of wine and/or sauser but (and above all) in:

- discovering, cognition and learning about an interesting, historically valuable and really beautiful South Moravian region;
- learning about the viticulture and life and work of wine growers and rural populaton in general;
- moderate tasting of wine and, first of all, purchasing local wines as presents for relatives and friends or for the consumption after the homecoming.

This means that the wine growers should inform the passing cyclotourists about secrets of wine growing and production, show them results of their work, explain them methods of vinification, try to attract them to become their customers etc. They can them also offer and sell their own agricultural products, refreshment, local specialities, and accomodation. Especially accomodation seems to be important because the cyclotourism is – in contradistinction to traveling in cars – much more dependent on weather and its unexpected changes.

In general, however, it can be said that in settlements situated along Moravian Wine Trails, the offer of services for bicyclists is much more limited than, for example, in Central and Northern Moravia, especially as far as repairs of bikes, refreshment and/or accomodation in private houses are concerned.

In our opinion the expectation of authors of the project of Moravian Wine Trails that this could be "…a long-term project of the development of viticulture.." is not correct. This project can and also must contribute to the creation of new jobs and to the development of infrastructure. The development of South Moravian viticulture should be based on other projects although it is quite clear that the rural tourism in this territory cannot exist without it and without local folklore and cultural traditions.

As far as the acceptation of this programme by the local population is concerned, it is necessary to say that the interest is generally not too high and that the wine growers underestimate it. In contradistinction to the wine growers in the neighbouring Lower Austria wine cellars in many villages are opened only when the people work there and the communication with visitors takes place practically only at local wine exhibitions and presentations that are organised from the end of January till the end of May. Too complicated buroeacratic procedures and legislative norms (Act on wine growing and production and Act on traffic on public roads) are another negative factors, which complicate selling activities of small wine producers.

Until now, the programme of Moravian Wine Trails has been considerably influenced by the fact that it was elaborated "at the table". The interest of local people is low; this concerns not only growers but also municipalities: in many villages the members of local adminsitration are not aware of the fact that there is a cyclotrial in their locality with signposts, orientation tables, maps etc.

Another, a closely related problem is the question how to integrate Moravian Wine Trails into the network of so-called microregions and their activities. According to statements of mayors of several settlements, Moravian Wine Trails are not a contribution to the regional development but more or less a source of troubles. Preparation and printing of maps, leaflets, brochures and other materials is costly and complicated and the final effect is very low (if not zero).

However, the management of microregions seems to be the most important factor of future improvement. There are sympathetic efforts to inform the public about individual localities, microregions and terroirs. People try to revive and dust off old traditions, organise local festivals and promote their villages. These efforts, however, must be supported not only by private organisations of the type of Environmental Partnership but also by governmental and regional bodies (South Moravian Regional Office, Viticultural Fund, universities etc.) because only in this way it will be possible to continue in the development of this important form of rural tourism and of the whole region.

It results from data discussed above that the Moravian Wine Trails require a *clear mission* and a *concrete and implementable vision*. Great attention should be also paid to the infrastructure enabling to carry out this type of activities (infromation centers, cheap accomodation, resting places, repair workshops etc.). Another problem is the development of partnership among all participating subjects (wine growers, local citizens, media and municipalities), as well as among microregions and their parts, regardless to the fact if they are localised near the cycloroutes of Moravian Wine Trails or not. To attract tourist is a common task of all. Important seems to be also the cooperation with neighbouring countries, especially with wine growers in Austria and Slovakia. The Euroregion South Moravia-Austrian Weinviertel-Western Slovakia exisists and for that reason it is necessary to continue in its development. The joint efforts and co-operation on

the international level can improve the competitiveness of local communities and to increase the standard of living in this beautiful part of Central Europe.

Discussion questions:

1. Which is the methodological essence of marketing research?
2. Which are the major thematic fields of learning about customers?
3. Which are the main causes of partiality of marketing research data?
4. When the first marketing research studies were performed and which was their content?
5. Which are the common denominators of marketing research, studies of public opinion, and sociological research? Can you describe their differences?
6. Which are advantages and disadvantages of primary research for researchers and ordering parties (clients)?
7. Which are advantages and disadvantages of secondary research for researchers and ordering parties (clients)?
8. Which are advantages and disadvantages of quantitative research?
9. Which are advantages and disadvantages of qualitative research?
10. How would you describe the difference between marketing research and market research?

2 MARKETING RESEARCH AS A PROCESS

Chapter objectives: Introduce marketing research as a process in which the preparatory phase with the most accurate definition of the marketing problem itself, including the goals of its entire research, is of fundamental importance. Subsequently, the basic five steps of marketing research are explained. Finally, the differences between the plan and the marketing research project are described.

Key words: process of marketing research, research objectives, marketing data, plan of marketing research, project of marketing research.

Each concrete case of marketing research has specific characteristics that result from the unique nature of investigated problems. In general, we can describe it as a process consisting of the following five steps:

1. Definition of a marketing problem and research objectives;
2. Creation of the research plan;
3. Collection of the data,
4. Statistical processing of collected data and their analysis;
5. Presentation of results (including practical recommendations),

.

In the course of the research itself we can distinguish two major phases - *stage of research preparation* and *stage of research implementation*, which includes also processing and analysis of collected data. Within the framework of these two major stages, there are several subsequent steps that are complementary and mutually interlinked. This means that possible shortcomings, which may occur during initial steps of the research activities, may cause serious weakening or even devaluation of obtained results. The preparatory stage of research involves steps leading to the formulation of prerequisites for starting of implementation and it can be said that that it is in many aspects of decisive importance for the methodological quality of research efforts.

2. 1 Definition of a marketing problem and research objectives

A good formulation of a *marketing problem* is one of the most important parts of the research. From the thematic point of view, it specifies and demarcates the domain in which the subsequent research steps should be performed. If the problem is not exactly formulated, the research costs may by higher than the value of results because there is a great risk that can be biased and will not meet the requirements of the client so that the whole process may become to be completely worthless from the practical point of view. Due to this fact it is necessary to elaborate at least a brief *theoretical definition* of the problem, i.e. to explain as it is perceived by experts, how it is described in professional literature etc. The definition of the most important and also of related *ideas/concepts* is an indispensable part of this step.

The need of investigation of a certain problem may result from initial, often indefinite signals (appearing for example in mass communication media, in written complaints of citizens, in claims of customers, resulting from the behaviour of employees etc.) or on the base of situational analysis (provided that there are stimuli of a more obvious and better pronounced character). It is also necessary to avoid and eliminate too limited and/or too broad definitions of individual marketing problems. A precise *problem formulation* enables to define further steps and techniques, which will result in collection of necessary data.

The performed definition of a problem is followed by the formulation of *a target (or even targets)* that should precisely define the required results of planned marketing research.

Further it is necessary to estimate *the value of information/data* obtained by means of marketing research. The final result should be a belief of conviction that the marketing research will bring new, more objective and more beneficiary data than those which are available at present and that their value will be higher than the research costs.

2. 2 Orientation analysis of situation

When preparing the marketing research, an *introductory orientation analysis of existing situation* should definitely not be missing. In this stage, the researcher tries to learn more about the environment and essence of a problem, studies already existing data, discusses with experts and looks for those data that could provide a deeper insight into the problem and indicate its possible solution. This analysis is very often based on the application of methods of secondary analysis, secondary research or, possibly, qualitative research.

2. 3 Specific features of marketing data

Information collected and studied within the framework of marketing research process can be classified according to many different criteria. The most usual categories of these data are the following:

- *Primary and secondary data*, i.e. those that were collected by means of primary or secondary research (see the preceding chapter);
- *Harddata* (depicting results of individual activities, occurrence individual phenomena and/or behaviour) and *softdata* (illustrating the perception and beliefs of respondents); and finally
- *Quantitative* (numeric, “measured”) and *qualitative* (verbal, descriptive) data. As it will be discussed in Chapter 6, data obtained by means of cardinal and ordinal variables can be qualified as quantitative information while qualitative information is obtained by means of nominal variables.

From the methodological point of view, it is generally required that data collected within the framework processes of marketing research should be as follows:

- *Relevant,* i.e. concerning the given marketing problem;
- *Valid,* i.e. expressing the content of things and facts of interest and enabling their measuring,
- *Reliable,* i.e. indicating the reliability of applied research tools and assuring repeatability of obtained results;
- *Effective,* i.e. showing that the results will be obtained within a reasonable time interval and that the costs will be also acceptable.

A thorough consideration of these methodological issues will undoubtedly contribute to a feedback and specification of the problem and targets of intended marketing research.

The selection of the method of collection of data applicable in individual concrete cases is influenced by the following two factors:

- *Purpose and targets of planned research*, which also determine the quantity and quality of required information, degree of accuracy of investigation, required generalisation of conclusions etc. and
- C*haracter of problems under study*, which above all determine the availability of these data.

Basin on important practical experiences it is also necessary to look for a most reasonable solution and very often even for a compromise.

2. 4 Plan of marketing research

This is the subsequent step, which follows the stage of problem definition and formulation of research targets. More exactly, it specifies the needed information, the method of data collection, and following steps of their processing, evaluation and interpretation. It also enables to check up the course of marketing research.

The research plan should involve:

- Formulation of the problem under study (i.e. the subject of research), its theoretical conception (i.e. theoretical starting point), and possible basic hypotheses concerning solution and determination of research objectives as well as the corresponding reasoning/justification;
- Preliminary concepts of expected results (as compared with the present knowledge).
- Definition of information needs, including their structure and sources, especially with regard to the initial secondary analysis.
- Definition of the basic set (i.e. object of research) and proposal of the sampling unit, justification of its size and structure, and proposal of the place and time of research implementation.
- Selection of the research technique and corresponding tools. Operationalisation of the subject and formulation of the final form of measurement tools.
- Definition of the technique of contacting the respondents (i.e. collection of field data).
- Preliminary research, which enables to specify the reliability of all aforementioned points on the base of testing the technique of field data collection in practice using a small set of respondents. As a rule, this is usually performed in a small ample of several tens of respondents (20 to 40).
- Field data collection.
- Statistical processing of collected data.
- Presentation and interpretation of results (including practical recommendations).
- Time schedule of research activities (including specification of responsible researchers).
- Research budget (calculation of expenses).

This means that it is necessary to assure and collect the maximum possible amount of essential and general data concerning the *essence of the marketing problem under study*. This can be based e.g. on data contained in final reports of earlier surveys, professional publications, interviews and discussion with experts, mass media, complaints of customers, statistical data etc.

If the technique of field data collection is already elaborated (including the corresponding tools – e.g. questionnaires or record sheets), it is necessary to test it in practice using a small set of sampling units. This stage is called "preliminary research", or briefly *pretest*. It tests the understandability and explicity of statements, easiness of manipulation with the questionnaire (distribution and collection), reaction of respondents when asked for an interview, course of interview, partial hypotheses, technical aspects of processing of collected data etc.

As already mentioned, the pretest is most frequently performed using a small sample (20 to 40 respondents) and is oriented to a practical verification of the overall conception of field data collection, i.e. selection and first contacts with respondents, testing of the quality of questionnaires, evaluation of its acceptance by respondents (are they interested or bored when answering it?), assessment of legibility of all questions or offered variants of possible answers, completeness of collected data, and a successful finalisation of survey.

2. 5 Project of marketing research

A more detailed description of individual steps of the aforementioned plan of marketing research results in the formulation of its *project*. The extent of this project may be approximately 20 pages and a proposal of the questionnaire (or the record sheet) is its integral part.

A properly elaborated project of marketing research then serves as the most important material used in negotiations with the client, in selective procedures and public competitions etc. Its definitive approval by both parties (especially if it becomes a supplement of signed business contract) should eliminate possible later protests of clients in the final stage of presentation of results and or after the handing over of written final report containing the collected data as well as client's objections that the presented results do not correspond with the order and problem formulation (see 2.1 Definition of a marketing problem and research objectives), that the research should present something else etc.

Discussion questions:

1. How many and which are the individual stages of the marketing research process?
2. Can you define primary and secondary data?
3. Can you define harddata and softdata?
4. Can you define quantitative and qualitative data?
5. Explain the meaning of terms relevance, validity, reliability and efficiency of marketing information?
6. Which is the purpose of marketing research plan and how does it look like?
7. What it is a pretest and why it is used?
8. What is a project of marketing research, why is it so important, and why is it used?
9. When asking exhibitors at the 41st International Engineering Fair (MSV) in Brno it was found out that already 87% of exhibiting firms intended to participate in also it in the next year. Another 9% did not know or were not able to answer. How would you test the

reliability of answers to the question “Do you intend to participate in this fair also in the next year”?

3 SUBJECT OF MARKETING RESEARCH – ITS DEFINITION AND OPERATIONALISATION

Chapter objectives: First of all, the reader will learn that the subject of marketing research (ie the researched marketing problem) must be decomposed and concretized into measurable features (variables, questions) for the needs of data collection in the field. Such operationalization represents a systematic process proceeding from a more general definition to specific observable items - variables. The creation of a research tool (questionnaire, record sheet), as shown in a specific example, should definitely not be understood as a random and spontaneous invention of questions and features. The previous chapter 2 already indicated that their final selection should be at the same time as invalid and as effective as possible with regard to the goals of marketing research.

Key words: subject of marketing research, marketing problem and its operationalization, questions, variables.

As we already know, a subject of the marketing research is usually a certain marketing problem. Its initial formulation may for instance as follows: *Satisfaction of our customers*. It is obvious already at the first sight that, from the viewpoint of an empiric study, this problem formulation is too broad and too general. For that reason it necessitates a further concretisation. Above all, it is necessary to explain such terms as "customers" and, above all, "their satisfaction". At the level of the subject of marketing research this requires *a definition and a subsequent operationalisation of these terms*. This problem is so important from the methodological point of view that it must be dealt with in a short but extraordinarily important chapter.

The term "Operationalisation" means a gradual *structuralisation, decomposition, and desaggregation of the defined problem* (i.e. subject of marketing research) with regard to objectives and possible practical use of obtained results. The problem under study should be gradually decomposed into more and more concrete variables, which are of essential importance for its practical solution. At the end, we should obtain empirically measurable values (conditions) of properties of the study object. These values can be obtained directly in the field (i.e. "on the spot") using its variables. This means that *the objective of operationalisation is to construct such variables (traits), which enable to express (measure) values (conditions) of those properties that of interest within a certain set of investigated units.* The whole process of operationalisation can be illustrated by means of the following simple example.

Both detailness and, on the contrary, broadliness of operationalisation of the problem under study should be based above all on the importance of research targets and possibilities of its practical use. If we do not intend to use the obtained results under practical conditions (e.g. for an improvement of our offer of products), it is enough to perform the operationalisation using a single question: „How are you satisfied with our outlet? ". The resulting distribution of

frequencies of individual answers then can either please or sudden us but in no case can we learn what should be improved. However, if we want to know if it is necessary to change opening hours, behaviour of our shop assistants when dealing with customers, interior equipment of our outlets etc., we must to operationalise the problem under study in a greater detail.

A similar situation may exist also when defining the object of our research – see Chapter 5. In the above example it seems to be quite clear but when thinking about this problem more thoroughly, we can realize that it is necessary to define exactly who if „our customer ": does this term indicate all people who show the interest in our products (e.g. those who visit our shop/outlet) or only those who buy our products?

Besides, such an operationalisation also influences a subsequent specification and concretisation of the object of marketing research. This will be illustrated in the following example.

Example 5 Operationalisation of the research subject and construction of a questionnaire

The following example concretely illustrated a gradual process of operationalisation of the marketing research subject to the final form of the questionnaire.

The subject of this research was "*Image of the Private College of Economic Studies in Znojmo"*. The operationalisation of this research subject was based on the *mission of this school*, which was formulated at the moment of its establishment in 2005 as follows: *The Private College of Economic Studies in Znojmo (hereafter mentioned only as PCES) wants to be a school of top quality and high study requirements. The tuition will take place in a pleasant environment and in an atmosphere of helpfulness and friendliness".*

In the first step, the image of PCES was structured into four levels, which resulted from the aforementioned definition of its mission:

1. Quality;
2. Demandingness;
3. Amenity,
4. Atmosphere.

In the subsequent second step, the following questions were formulated with regard of individual levels of research:

QUALITY:

1. How are you satisfied with the professionalism of teachers?
2. How are you satisfied with the quality of tuition at the PCES?
3. How are you satisfied with the approach of teachers?
4. Do you think that the content of study materials offered by the PCES is sufficient?

DEMANDINGNESS:

1. How would you evaluate the demandingness of studies at the PCES?
2. How would you evaluate the temporal demandingness of studies?
3. Do you think that the requirements of examining teachers are too high?

4. Are you satisfied with financial demands of the PCES?

ENVIRONMENT:

1. Are you satisfied with the accessibility of the PCES?
2. How can you orientate yourself within the school building?
3. Do you think that all classrooms are agreeable?
4. Are you satisfied with the technical equipment at the PCES?
5. Are you satisfied with the offer in the college snack bar?
6. Do you take advantage of beverage automatic machine?
7. Do you go to the college club Harvart?

ATMOSPHERE:

1. How are you satisfied with the approach of ladies in the study department?
2. How are you satisfied with the office hours in the study department?
3. Do you think that the college atmosphere is pleasant?

In the third step, these questions were further subdivided and the altogether 18 definite variables (traits) were constructed (respondents were offered several variants of answers). If not mentioned otherwise, we used the following scale: 1 = Very satisfied; 2 = More or less satisfied; 3 = More or less unsatisfied; 4 = Very unsatisfied and 5 = I don't know. This resulted in the following list of variables that helped us to measure (estimate) the image of PCES:

1. How are you satisfied with the professionalism of teachers?

1 2 3 4 5

2. How are you satisfied with the quality of tuition at the PCES?

1 2 3 4 5

3. How are you satisfied with the approach of teachers?

1 2 3 4 5

4. Do you think that the content of study materials offered by the PCES is sufficient?

1 2 3 4 5

1. How would you evaluate the demandingness of studies at the PCES?

(1-Very demanding; 2-More or less demanding; 3-Not so much demanding, 4-Not demanding; 5- I don't know)

1 2 3 4 5

2. How would you evaluate the time-consuming demandingness of studies?

(1-Very demanding; 2-More or less demanding; 3-Not so much demanding, 4-Not demanding; 5- I don't know)

1 2 3 4 5

3. Do you think that the requirement of examining teachers is too high?

(1-Very much; 2-More or less yes; 3-Not so much, 4-Not at all; 5- I don't know)

1 2 3 4 5

8. Are you satisfied with financial demands of the PCES?

1 2 3 4 5

9. Are you satisfied with the accessibility of the PCES?

(1-Very complicated, 2-More or less complicated, 3-Not too complicated, 4-Uncomplicated, 5-I don't know)

1 2 3 4 5

10. How can you orientate yourself within the school building?

(1-Very well; 2-Relatively well; 3-Not so much, 4-Not at all; 5-I don't know)

1 2 3 4 5

11. Do you think that all the classrooms are agreeable?

1 2 3 4 5

12. Are you satisfied with the offer in the college snack bar?

1 2 3 4 5

13. Do you take advantage of beverage automatic machine?

(1-Often, 2-From time to time, 3-Not too much, 4-Not at all, 5-I don't know)

1 2 3 4 5

14. Do you go to the college club Harvart?

(1-Often, 2-From time to time, 3-Not too much, 4-Not at all, 5-I don't know)

1 2 3 4 5

15. How are you satisfied with the approach of ladies in the study department?

1 2 3 4 5

16. How are you satisfied with the approach of ladies in the study department?

1 2 3 4 5

17. How are you satisfied with the office hours in the study department?

1 2 3 4 5

18. Do you think that the college atmosphere is pleasant?

1 2 3 4 5

In the fourth step the following four identification variables (traits) of respondents were created, which were used above all as independent valuables used in the statistical analysis and for the calculation of correlation coefficients:

1. Sex:

Male 1 Female 2

2. Form of studies:

Regular 1 Combined 2

3. Year of studies:

First 1 Second2 Third 3

4. Study programme:

Marketing and management	1
Accounting and financial management	2
Payroll accounting, taxes and human resources management	3
Economics of public administration and social services	4

In the final fifth step, a definitive electronic questionnaire was created on the base of a fusion of both groups of variables (i.e. of those under study and those used for identification of respondents) and elaboration of instructions for respondents (presented above all in its introductory part). The questionnaire was distributed and collected via e-mail and it is shown in Fig. 3 – 1.

QUESTIONNAIRE

„IMAGE OF THE PCES ZNOJMO“

Dear Colleagues,

Within the framework of our assignment we ask you for help and fulfil the enclosed questionnaire.

For the college management your opinions are very important. We will be pleased and appreciate, if you will have time to answer questions presented in this short questionnaire. Your answers will be used only for our internal use so that they should help to the improvement of your satisfaction.

Mark, please, you're personal in the enclosed scales your answers to individual questions and return the filled-up **questionnaire till 7 March 2008** at the latest**.**

If not mentioned otherwise, answer please according to the following scale:

1 = Very satisfied; 2 = More or less satisfied; 3 = More or less unsatisfied; 4 = Very unsatisfied, and 5 = Don't know

Highlight, please, your variant of answer in red.

Thank you very much for your answer and cooperation.

Students of the 1st year of study programme Marketing - Management

QUALITY:

1. How are you satisfied with the professionalism of teachers?

1 2 3 4 5

2. How are you satisfied with the quality of tuition at the PCES?

1 2 3 4 5

3. How are you satisfied with the approach of teachers?

1 2 3 4 5

4. Do you think that the content of study materials offered by the PCES is sufficient?

1 2 3 4 5

DEMANDINGNESS:

1. How would you evaluate the demandingness of studies at the PCES?

(1-Very demanding; 2-More or less demanding; 3-Not so much demanding, 4-Not demanding; 5- Don't know)

1 2 3 4 5

2. How would you evaluate the time-consuming demandingness of studies?

(1-Very demanding; 2-More or less demanding; 3-Not so much demanding, 4-Not demanding; 5- Don't know)

1 2 3 4 5

3. Do you think that the requirements of examining teachers are too high?

(1-Very much; 2-More or less yes; 3-Not so much, 4-Not at all; 5- Don't know)

1 2 3 4 5

4. Are you satisfied with financial demands of the PCES?

1 2 3 4 5

ENVIRONMENT:

1. Are you satisfied with the accessibility of the PCES building?

(1-Very well, 2-Relatively well, 3-Not too much, 4-Not at a, 5-Don't know)

1 2 3 4 5

2. How can you orientate yourself within the school building?

(1-Very well; 2-Relatively well; 3-Not so much, 4-Not at all; 5-Don't know)

1 2 3 4 5

3. Do you think that all the classrooms are agreeable?

1 2 3 4 5

4. Are you satisfied with the technical equipment at the PCES?

1 2 3 4 5

5. Are you satisfied with the offer in the college snack bar?

1 2 3 4 5

6. Do you take advantage of beverage slot machine?

(1-Often, 2-From time to time, 3-Not too much, 4-Not at all, 5-Don't know)

1 2 3 4 5

7. Do you go to the college club Harvart?

(1-Often, 2-From time to time, 3-Not too much, 4-Not at all, 5-Don't know)

1 2 3 4 5

ATMOSPHERE:

1. How are you satisfied with the approach of ladies in the study department?

1 2 3 4 5

2. How are you satisfied with the office hours of the study department?

1 2 3 4 5

3. Do you think that the college atmosphere is pleasant?

1 2 3 4 5

IDENTIFICATION DATA:

1. Sex:

Male 1 Female 2

2. Form of studies:

Regular 1 Combined 2

3. Year of studies:

First 1 Second 2 Third 3

4. Study programme:

Marketing and management	1
Accounting and financial management	2
Payroll accounting, taxes and human resources management	3
Economics of public administration and social services	4

We thank you once again for answering the above questions.

Students of the 1st year of study programme Marketing - Management

Fig. 3 – 1 The final questionnaire form

The whole process of operationalisation can be illustrated in the following block diagram (Fig. 3 – 2).

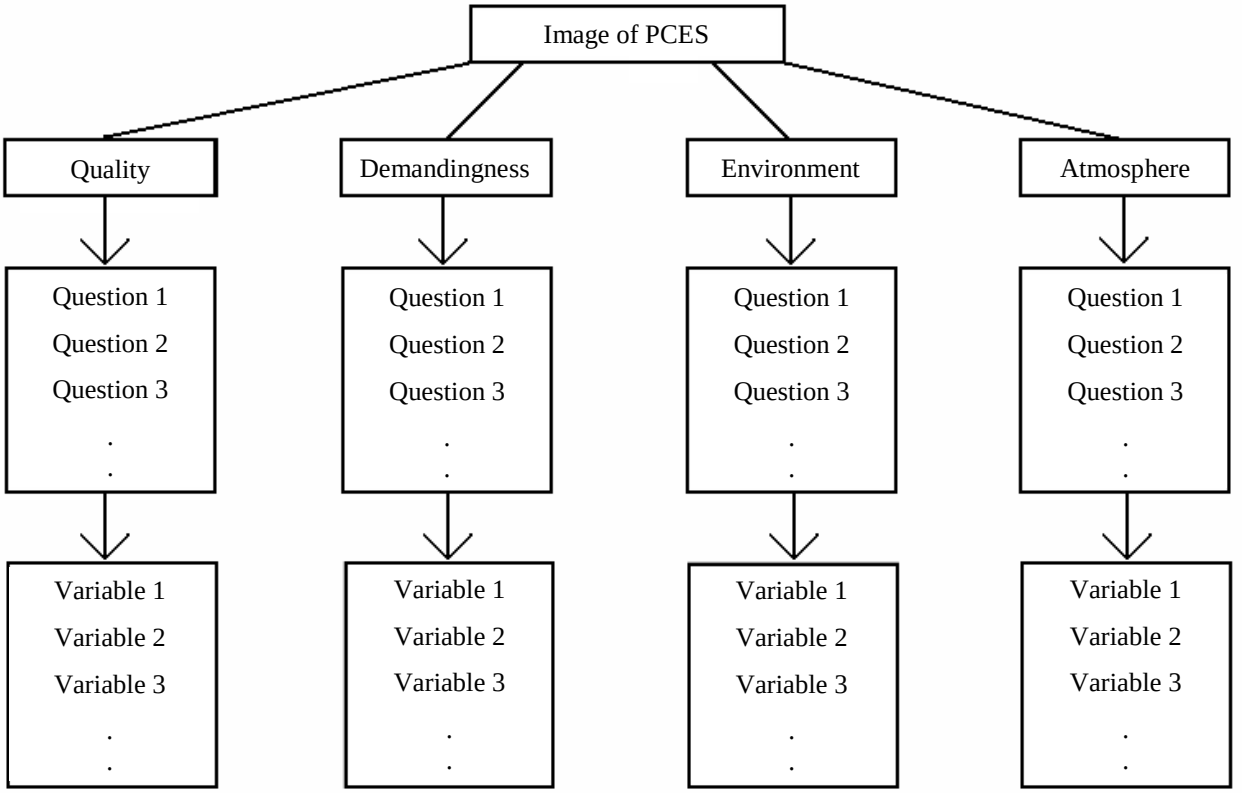

Figure 3 – 2 Process of operationalisation

In any case, the process of operationalisation should be finished by the elaboration of variables under study and by a proposal of the resulting research tool (questionnaire, record sheet). This will be discussed in the following chapter.

Discussion questions:

1. How would you measure the research subject "*Satisfaction of customers with our products*?"
2. How would you use the obtained results?
3. What would you investigate, who would be the object of your research and which would be the investigated units?
4. How would you operationalise this problem down to the level of concrete properties and of corresponding items under study?
5. How could you secondarily use the obtained statistical data (harddata)?
6. Where and how would you collect them?
7. Which are their advantages and disadvantages?

4 MARKETING RESEARCH TECHNIQUES

Chapter objectives: To draw attention to the advantages and disadvantages of using three techniques of marketing research - questioning, observation and experiment to obtain empirical data in the field. Point out the differences between direct and indirect inquiries, which are ultimately reflected in the obvious differences between the two tools - ie. questionnaire and record sheet. In this context, the advantages and disadvantages of various types of issues are also discussed, especially the basic three - open, closed and semi - open.

Key words: methods and techniques, personal interview, questionnaire, typology of questions, public inquiry, record sheet, telephone survey, focus group interview, observation, experiment.

In contradistinction to many Czech (but above all foreign) marketing textbooks we will make a more accurate discrimination between terms "method" and "technique ". *Methods* represent a generalized approach to the socio-economic problems and/or objects under study. In essence, each branch of science uses such tools as for example historical methods (following a problem and its development within a certain time period), comparative methods (studying changes of a problem in different objects), typological methods (investigating uniqueness of a problem with regard to specific objects) etc.

On the other hand, in the field of marketing research the term *technique* is used to specify a very *concrete way* of primary data collection, which enables not only to record and register the occurrence of various phenomena and/or to monitor behaviour of people but also to investigate their causes, opinions, attitudes and motives. In marketing research, we use three principal techniques, viz. *interviews (surveys), observation, and experiments*. All three will be discussed in greater detail in the following text.

4. 1 Interview

In marketing research, interviewing is the most frequent technique. Some people even believe that marketing research is a mere simple, unsophisticated questioning of customers and that it is nothing more than a distribution and filling of question forms. This mistake is nearly the same as that of a simple confusion of marketing with advertising.

Surveys are performed by interviewers who use tools called questionnaires and/or record sheets and select a suitable form of communication with respondents (addressed customers). This contact may either *direct* (as illustrated in Fig. 4 – 1) or *mediated* (Fig. 4 – 2).

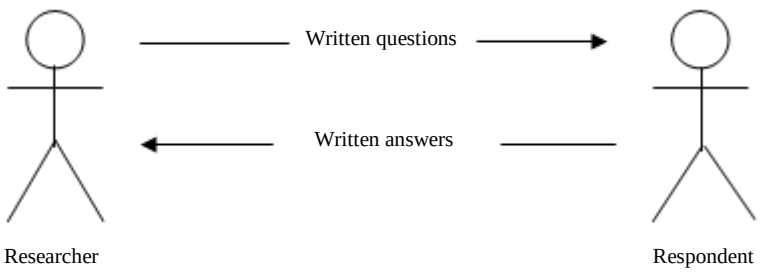

Figure 4 – 1 Direct (written) questioning

The situation presented in Fig. 4 – 2 is quite different. In this case the interrogation is *mediated by an interviewer* who enters between the researcher and the respondent. For that reason this research technique is called *personal interview.*

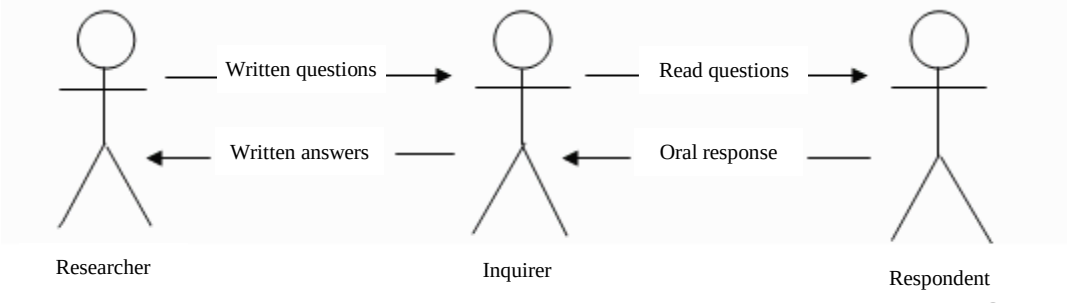

Figure 4 – 2 Personal interview

A comparison of both figures (Figs 4 - 1 and 4 – 2) indicates *significant differences* between direct questioning and personal interview. In the first case, the mediating person (inquirer) is missing and for that reason the *questionnaire* should be constructed in such a way that the interviewer would replaced by it as much as possible. It content should be explained to the respondents in great detail so that they would learn all essential facts (see paragraph 4. 2. 1 in this chapter), especially those contained in the introductory text of this questionnaire.

On the other hand, in case of personal interview the contact between researcher and respondent is mediated by a trained and well-instructed interviewer who does not need a detailed questionnaire but only a substantially shorter *record sheet*. Compare examples 4 – 1 (questionnaire) and 4 – 5 (record sheet).

4. 2 Written inquiry

The written communication is mediated by questionnaires and/or inquiry. Let's try to describe these techniques and to show the substantial differences between them.

4. 2. 1 Questionnaire

When constructing a questionnaire it is necessary to remember that it should be appropriate. A wrong questionnaire can negatively influence the obtained data and the final results need not corresponds with targets and requirements of research.

A good questionnaire should meet the following three major requirements:

1) *Purpose-technical needs*, i.e. the questions should be formulated and constructed in such a way that the respondents would answer the questions of interest with a maximum possible accuracy;
2) *Psychological needs*, i.e. it is necessary to create such conditions and atmosphere of research that the respondents would perceive the survey as an easy, pleasant and desired task. It is necessary that the respondent would answer briefly, frankly, and truthfully;
3) *Need of understandability*, i.e. the respondents should clearly understand all questions, they should also know what are they asked to do, how they should proceed and how they should fill up the questionnaire (this all concerns above all filter questions). The questionnaire must be able "to speak" and to explain (describe) to respondents intentions of the interviewer.

Above all, the following four aspects of a questionnaire contribute to the fulfilment of the requirements mentioned above:

a) Its overall design (appearance);
b) Formulation of individual questions;
c) Types of questions;
d) Manipulation with the questionnaire.

These aspects will be discussed more deeply in the following paragraphs.

The overall appearance of a questionnaire

The questionnaire must be attractive and its graphic layout should draw attention of the respondent *at the first sight*. This concerns its design and overall appearance, viz. its format, appearance of the first page, colours and paper quality. The respondents should be really asked and attracted to start with the filling up of the presented questionnaire; they should be willing to give us their valuable time, to take pains and to occupy themselves with our questions.

A too big format of the questionnaire evokes a feeling of a difficult handling (manipulation) with it and also a feeling that the amount of required data is too great. On the other hand, too small format may force respondents to concentrate themselves too much when reading it and answering individual questions. The optimum size of a questionnaire is the traditional format A4 (i.e. 210 x 297 mm).

The *first page* (especially its graphical layout) and the content of the introductory text are very important. *The introductory text* should above all:

1. *to interest the respondents*, to explain them the purpose of the performed research, and to emphasize the importance of their answers for the solution of practical problems under study as well as the potential personal benefits resulting from the solution of these problems (i.e. which improvements they may expect);
2. *to appeal respondents and ask them for cooperation*, to emphasize the purpose and significance of provided information, to persuade them that the filling up of the questionnaire as well as answering of possibly all questions are very important;
3. *to define the manner how to fill up the questionnaire* i.e. to explain if the they should encircle code numbers, cross blank squares or underline correct answers;
4. *to set up the date and method of returning* of fulfilled questionnaires; and
5. *to assure respondents that their anonymity will be respected*, e.g. that the researcher (submitter) will not receive individual questionnaires but only summarised, statistically processed results.

In some special cases, i.e. when questionnaires are distributed and collected by taskmasters themselves (for example among students in classrooms), the introductory text as well some other instructions for respondents may be omitted, because they can personally explain and describe the purpose of performed marketing research

It also makes a good impression, if the research is arched over and professionally warranted by a *respected scientific or public institution* (universities, Academy of Science etc.) or if it is supported by a third, not-interested and independent party (marketing research agency).

Example 6 Questionnaire

A concrete example of a questionnaire is presented in Fig. 4 – 3.

International Institute for Marketing, Communication and Entrepreneurship

Příkop 4, 602 00 Brno, Phone/Fax: 05/4517 6578

E-mail: info@iimce.cz http://www.iimce.cz

Dear citizens of the town of Dačice,

Students of the Private College of Economic Studies in Znojmo perform as a seminar assignment a marketing research study of your town. They would like to know your opinion about your mother town, how do you perceive it and you evaluate it. We ask you therefore for spent several minutes of your valuable time and to answer the following 24 questions. Processed results will be presented to the representatives of your town and they will use them when solving discovered problems and trying to improve the living conditions existing in Dačice.

This questionnaire is completely anonymous. We would be very pleased and appreciate if your answers will be maximally frank. The Municipal Office will receive only a final report containing results of statistical processing and summarisation of all recovered questionnaires.

These questionnaires can be filled up easily and very quickly. Mostly it is only necessary to select and encircle that numeric code of one of offered variants which corresponds at best with opinion and/or situation. In case of open-ended (i.e. those without offered variants of answers) be so kind and express briefly your opinion.

Basing on a mutual agreement with you, the filled-up questionnaires will be collected by students themselves.

Thank you very much in advance for your cooperation and wish you a maximally quick solution of ascertained shortcomings and problems.

International Institute of Marketing, Communication and Entrepreneurship Brno

1. How do you perceive the attractiveness of the town of Dačice with regard to:

	High	Average	Above average	None	Below average	Don't know
Natural beauties	**1**	**2**	**3**	**4**	**5**	**6**
Prospects of economic development	**1**	**2**	**3**	**4**	**5**	**6**
Hospitality	**1**	**2**	**3**	**4**	**5**	**6**
Employment opportunities	**1**	**2**	**3**	**4**	**5**	**6**
Business activities	**1**	**2**	**3**	**4**	**5**	**6**
Road network and transportation	**1**	**2**	**3**	**4**	**5**	**6**
Living conditions	**1**	**2**	**3**	**4**	**5**	**6**
Shopping opportunities and services	**1**	**2**	**3**	**4**	**5**	**6**

Offer of cultural events	**1**	**2**	**3**	**4**	**5**	**6**

2. Think over please and write down those priorities that should be solved by the Municipal Council of the town of Dačice: *(Encircle, please, **only three** most important domains)*

Possibilities of self realisation in sports	**1**
Possibilities of self realisation in culture	**2**
Care about community property	**3**
Problems of health care	**4**
Problems of social care	**5**
Environment	**6**
Educational problems	**7**
Housing	**8**
Town cleanness and amenities	**9**
Safety of citizens	**10**
Business opportunities	**11**
Approach of municipality officers to problems of citizens	**12**

3. Representatives of the town of Dačice elaborate plans and programmes of a further development in domains enumerated in Question 2. Do you want to:

Learn something about these plans	**1**
Co-operate and participate in the preparation of these plans	**2**
Express your opinion concerning these plans	**3**
Do nothing (no interest)	**4**

4. Do you agree with the decision of the Dačice Municipal Board to remain in the South Moravian Region?

Yes	1	Don’t know	2	No	3

5. Which shops/outlets do you miss in Dačice?

.....

6. Which services do you miss in Dačice?

.....

7. What else do you miss in Dačice?

.....

8. Is the offer of children and youth leisure time activities in Dačice sufficient?

Yes	1	I don't know	2	No	3

9. If not, what is missing in this domain according to your opinion?

.....

10. Is there an interesting place/object in Dačice that should deserve a greater attention?

.....

11. Which are in your opinion the most important qualities of the town of Dačice?

.....

12. Which are in your opinion the major shortcomings of the town of Dačice?

.....

13. Enumerate please three most important enterprises/businesses existing in Dačice:

.....

14. Do you think that you are sufficiently informed about:

	Yes	No
Activities of the Municipal Office?	1	2
Acts and Public Notices that concern you?	1	2
Decision-making of local authorities?	1	2
Plans of town development?	1	2
Management of community property?	1	2

15. How do you acquire information about events in the town and in the region:

	Yes	No
Directly at the Municipal Office	1	2
At meetings of the Municipal Council	1	2
Directly from town councillors	1	2
From municipality announcement board	1	2
From municipality bulletin	1	2
From local public service broadcasting	1	2
From public radio stations	1	2
From internet	1	2
From friends and acquaintances	1	2
From other sources (specify)		

.....

16. Be so kind please and specify which sources of data should be used to inform you better about activities of the Municipal Office and about events taking place in town. (You can extend and/or supplement variants mentioned in Question 15).

.....

.....

17. Which newspapers and journals do you regularly read?

.....

18. How long do you live in Dačice?

Less than 5 years	**1**
5 to 10 years	**2**
More than 10 years	**3**

19. Did you sometimes think about the possibility to move out of Dačice?

Definitely yes	1	Possibly no	3
Possibly yes	2	Definitely no	4

20. If yes, why?

.....

.....

21. Sex:

Male	**1**
Female	**2**

22. Your highest degree of finished education

Primary, skilled	**1**
Secondary	**2**
University	**3**

23. Age

Up to 24 years	**1**
25 to 44 years	**2**
45 to 64 years	**3**
65 years and more	**4**

24. Your current economic activity

Employee	**1**	**Student**	**5**
Worker	**2**	**Pensioner**	**6**
Private entrepreneur	**3**	**Other non-economic activities**	**7**
Other economic activities	**4**		

Thank you very much for your answers.

Fig. 4 – 3 Specimen of a questionnaire

The optimum length of a questionnaire ranges from 40 to 50 questions. The optimum duration of its filling-up should be approximately 20 minutes. The questionnaire should contain only really essential questions. Too long questionnaires discourage respondents at the first sight. There is no doubt that there is a substantial difference between answering only 15 or as much as 150 questions (or even more than one thousand as it is usual in some detailed surveys concerning for instance style of life). The higher the number of items (questions) in a questionnaire, the more impressively it seems that their answering and filling-up will be lengthy and time-consuming.

The sequence of questions contained in a questionnaire is also important. Each questionnaire should begin with interesting questions, which would be followed by the so-called meritory questions, i.e. questions that are directly related to the problem under study and require concentration; less important questions are usually posed in the final part of the questionnaire. The most suitable method of filling-up of a questionnaire is to encircle or cross selected variants of answers; it is also possible to use another ways of indication, especially when using electronic forms (via internet). Questions and answers contained in electronic forms may be involved directly into the process of computer processing.

The aforementioned general recommendations concerning construction of various questionnaires can be summarized in the following four rules:

1. The questionnaire should be *economic and rational;*
2. The questionnaire should be well-arranged and its filling-up should be *easy;*
3. The conception of the questionnaire should be *coherent*,
4. The conception must be *interesting and attractive* at the first sight.

There is no doubt that the formulation of individual questions need not be sufficient for the construction of a good, top quality questionnaire. Besides, also in this case the years of practical experience may pay off.

Formulation of questions

Explicitness and comprehensibility are the main requirements when formulating individual questions. Formulations forcing respondents to recall laboriously the facts, to calculate them exactly and/or only to estimate them, should be used very rarely and even exceptionally. This concerns for example such questions as:"When lastly..."; "How much/many in average..."; "How often ..."; "How many times recently..." etc. However, it is also more and more obvious that the more *concrete* questions are posed, the *clearer and more concrete* are the answers.

Questions should be formulated in such a way that they would be as *valid* as possible so that we could ask and receive answers really expressing what we want and need to ascertain. The validity of a certain question may be changed for example due to different temporal, social, and cultural conditions of the survey. Let's say, for example, that thirty years ago the ratings of TV stations involved also the frequency of TV watching, i.e. how many times and how often the respondents sat in front of their TV sets and watched movies. In that time it was quite clear that this concerned only movies involved into the programme of telecasting. Today, it would be necessary to eliminate also replays of video recordings and/or DVDs.

This is closely related to the problem of *reliability*. This factor should express the degree of stability (accuracy) of repeatedly ascertained data. Concretely, this means that we check up if each asked question is answered in the same way (i.e. similarly as in case of weighing when we check up if the weight of a thing is the same, identical after all weightings).

Questions starting with "Why..." are also problematic. It is not very probable that respondents will be able to explain a certain problem in general because they cannot describe its all aspects and things we are interested in, they cannot perform its objective analysis and very often they are also not aware of their motives and/or causes of some phenomena. However, if we want to ascertain and evaluate the respondent's capability to analyse various events then it is better to pose "Why" questions within the framework of qualitative research.

First of all, it is suitable to eliminate all factors that could show a negative effect on respondents (e.g. length of questionnaire, too long and too complicated questions etc.). It is also recommended to use *euphemisms*, i.e. such periphrastic expressions that attenuate such negative evaluations of behaviour as for example this one: "Today's hectic times are psychically very demanding. Especially after a hard working day, we can feel completely exhausted and need to drink something stimulating. Which alcoholic beverages do you usually drink?"

Unsuitable is also the use of *suggestive questions (or leading questions)*, i.e. those that already before being asked are formulated in such a way that they suggest the expected suitable answer. As an example it is possible to mention a slightly modified aforementioned question: "You definitely drink alcoholic beverages. How often do you drink them? "

Besides the suggestiveness of formulation there is also a suggestive sequence of questions, which is usually called "*halo effect"*. This effect occurs in those cases when several mutually related questions are posed in a sequence and when the first of them are answered either positively or, on the contrary, negatively. This stance is thereafter automatically transferred also to related question so that we can receive only negative (or positive) answers. This can be prevented in such a way that we interleave them with questions of a neutral character or that we pose offered variants of answers in a reversed sequence (i.e. we pose at first positive questions instead of negative ones an *vice versa*) and/or if we pose such questions that concern completely different themes.

It is also possible to use the so-called *projective questions* in a questionnaire, i.e. those that are based on an expectation that the respondent has no idea about our intentions. Projection questions are given in such a way that it seems that we are not interested in the personal stance of the respondent and that we ask him/her to evaluate the opinion of somebody else. In fact, however, we expect that the respondent will subconsciously identify himself/herself with the other person and/or situation. These questions should be used above all in those cases when we are afraid that the respondent could answer this question elusively or even untruly, e.g. "Provided that someone in an tenement house does not respect the period of night quietness, can the neighbours ring up his/her door bell and rebuke him/her?"

The so-called "*bubble tests*" represent an analogous form of projective questions. Similarly as in comics also in this case we show different characters in different situations with empty bubbles above them and the respondents should fill up them with sentences that would be used by these characters according to the respondent's opinion. In the above example this would concern the person embarrassed by violation of night quietness, scolding and ringing at the neighbour's door.

Typology of questions

From the formal point of view we can distinguish the following three basic types of questions:

1. *Open-ended questions* (free, non-standardized);
2. *Closed questions* (standardized, controlled),

3. *Combined questions (semi-opened questions),* which represent a compromise between both above types.

In the following text we will discuss and illustrate them using concrete examples showing possibilities of their use (including their subsequent statistical processing). In the end of this chapter we also explain what is called *filter questions*, and why

Open-ended questions

Open-ended questions do not contain any variant answers. Respondents can formulate quite freely their opinion(s) using words of their own. There are several subcategories of open-ended questions:

- Free – respondents are absolutely free when formulating their opinion;
- Associative – respondents are asked to quote the first work that comes to their mind when reacting to the word presented in the questionnaire;
- Free ending of a sentence – the respondent can freely finish the presented sentence;
- Ending of a short story – the respondent should finish the presented open-end story;
- Filling-up of a picture – the respondent should react to the statement of one of two people in the picture and fill up the response of the other person;
- Finishing of a theme – the respondent should invent a story about situation presented in the picture (i.e. express what happened and/or what can happen).

Note:

The last two examples are de facto illustrating situations described above as *projective qu*estions and *bubble test*. These are usually used in psychological research. Anyhow, they make the questionnaire more diversified and attractive.

On of the advantages of open-ended question consists in the fact that respondents are stimulated to think about the presented situation and to recollect facts that were the most interesting for them and that are the most important. A great advantage of these questions is also the fact that respondents can answer these questions quite freely, that they are not limited by any barriers (variants), and that they can think about all aspects of the problem. This means that they can tell us something new, something what we do not know, something that we did not think about before. As a rule, these answers are more important because these were recalled from the memory of respondents.

The "freedom" of respondents is the major disadvantage of these questions because it is the source of troubles during the subsequent statistical processing. Answers to open-ended questions must be at first read through and thereafter categorised, classified (i.e. "coded" – see the paragraph Coding in Chapter 6, Statistical processing of primary data); only after performing this analysis it is possible to feed the collected data into a computer.

Just the open-ended questions are typical for the qualitative research – see Example 1. 4 of qualitative marketing research.

Example 7 Formulation and evaluation of open-ended questions

When surveying visitors leaving caves of the Moravian Karst, they were required to answer also the following question:

Which other Czech and/or foreign caves have you already visited?

.....

In the Final Report, the answers to this question were summarised as follows:

This open-ended question was answered by approximately one fifth (19 %) of visitors.

Their answers were as follows:

- Český kras Caves – 66 times,
- Javořičské jeskyně Caves – 28 times ,
- Mladečská jeskyně Cave – 27 times,
- Chýnovská jeskyně Cave – 17 times,
- Jeskyně Na Pomezí Caves – 11 times,
- Jeskyně Na Špičáku Caves – 11 times ,
- Vysoké Tatry Mountains – Twice,
- Nízké Tatry Mountains – Once,
- Romania – Once

As one can see, without a cue, without presented variant answer the respondents sometimes cannot remember exact names of caves they visited. Instead, they only mentioned the locality – mountain range or even only the country.

Closed questions

Closed questions contain several possible variants of answers and the respondent must select one or several of them. Their advantage consists above all in a quick and easy answering of posed question and also in the fact that we can steer the respondent to those thing that interest us. Their certain disadvantage consists in the necessity to respond to those variants that need not be perceived by the respondent as important, exact and apt to describe the situation.

The presented variants of answers may also show a suggestive effect on the respondent and thus to result in conforming answers. They also enable that the respondents will answer the questions randomly so that their unawareness of or unconcern in the problem can be concealed. For this reason the answers to closed questions are sometimes considered to be less important as compared with open-ended questions. In English the answers to closed questions are defined as a mere "recognition", while those to opened ones as a "recall". For that reason the closed questions are used as a typical tool of quantitative research.

Closed questions may be divided to:

(i) Dichotomic (binary, alternative, dual), which admit only two possibilities (Yes – No, Male – Female). Their major advantage consists in the fact that they can be easily processed. However, they often force respondents to express their extreme points of view,

(ii) Selective (polytomic, with a possibility to select only one variant). These questions can be easily processed. Their major problem consists in the fact that they eliminate a possibility to select several variants;

(iii) Enumerative (polytomic, with a possibility to select several variants). They provide a wider selection and this corresponds better with real situations. Their major disadvantage consists in the fact that their statistical processing is rather difficult. This type of questions is usually considered as a battery of questions and individual variants are perceived as a dichotomic trait (variable) that either occurred or not;

(iv) Polytomic (with an order of variants). They show a positive effect on respondents because they provide a possibility to differentiate between variants and to define their ranking. However, their statistical processing is even more difficult than in the preceding case.

Scaling is used to measure different opinions and, especially, attitudes (stances). In this case we not only try to establish if the phenomenon under study occurred or not (as in case of dichotomic questions with variants "YES - NO") but also the degree of respondent's perception of this phenomenon. When using this scale, which has form of a closed question, the respondent is asked to express his/her opinion about a certain phenomenon (object property) only by indicating a certain position on the scale. However, different researchers have different opinions as far as the number of categories offered to respondents is concerned. An even number of possibilities forces respondents to take a certain attitude while the odd number enables to be also indifferent and answer "I don't know" or "I cannot say".

Also the concrete forms of individual categories may be quite different. This is above all determined by the *scale nicety (i.e. detailedness).* As we will see the following text, scales with five grades are the most common in practice. However, there are also more detailed scales, e.g. with ten or even one hundred grades.

The *form of a scale* represents another problem. This concerns above all the method of coding of individual variants of answers (grades) – see paragraph "Coding" in Chapter 6 Statistical processing of primary data. As shown in the following table, there are three possible variants:

Very unsatisfied	Slightly unsatisfied	I don't know	Slightly satisfied	Very satisfied
-2	-1	0	+1	+2
1	2	3	4	5
1	2	5	3	4

Table 4 – 1 Scale types

The methodologically best form is presented in the first row – here the variant "I don't know" represents the natural centre and individual grades are situated specularly to the right and to the left and coded either as negative or positive digits. The second row contains an ascending sequence of values from the most negative to the most positive value and the variant "I don't know" is again in the middle. The third row is similar to the second one but the variant "I don't know" is encoded as the last one. Just this scale form is used most frequently in practice. In the course of statistical processing, it enables an easier pre-categorisation (see again paragraph Coding in Chapter 6 Statistical processing of primary data).

To the end of the preceding Chapter 3 we have mentioned the problem of operationalization of a rather complicated term "Satisfaction" of visitors and exhibitors at the International Engineering Fair Brno. Finally two simple but (as already mentioned) *batteries of questions were obtained*. Both of them investigate the same phenomenon (i.e. satisfaction) but each of them from a different point of view, i.e. with a different content. The advantage of question batteries consists above all in a possibility of simpler, more economical form (the introductory cliché "How are you satisfied... ", is not repeated). We also ask directly about the property which in the

scope of our interest ("...with exhibited, offered products"). Besides, when processing statistically the results, a greater number of data (40 – 50 items/rows) can be advantageously processed by methods of multidimensional analysis, above all factor analysis.

The same five-grade scales were used in both batteries below. The visitors answered the following questions:

How much were you satisfied?

	Positive opinion		**Negative opinion**		**Neutral opinion**
	Very satisfied	More/less satisfied	More/less unsatisfied	Very unsatisfied	Don't know
With exposed, offered products	**1**	**2**	**3**	**4**	**5**
With acquired information and knowledge	**1**	**2**	**3**	**4**	**5**
With established and/or deepened business contacts	**1**	**2**	**3**	**4**	**5**
With the professional level of the trade fair	**1**	**2**	**3**	**4**	**5**

As far as the exhibitors were concerned, the situation was even simpler:

How much was your firm satisfied at the trade fair?

	Positive opinion		**Negative opinion**		**Neutral opinion**
	Very satisfied	More/less satisfied	More/less unsatisfied	Very unsatisfied	Don't know
With business negotiations	**1**	**2**	**3**	**4**	**5**
With the interest in your products	**1**	**2**	**3**	**4**	**5**
With the professional level of the trade fair	**1**	**2**	**3**	**4**	**5**

Example 8 Formulation and evaluation of a closed question

Within the framework of an inquiry performed among visitors leaving the International Engineering Fair in Brno the respondents were asked to answer among others also the following question:

When dealing with and discussion about future investments of your organisations, your vote is:

Decisive	**1**	I don't participate in such discussions	**4**
Co-decisive	**2**	I am not employed in the organisation	**5**
Advisory	**3**		

The resulting distribution of relative frequencies of answers, i.e. the classification of the first degree (see Chapter 6) was presented and interpreted in the final report as follows:

When discussing further investments of the company, your vote is:

Decisive	16 %
Co-decisive	18 %
Advisory	30 %

As one can see, nearly two thirds of respondents (more exactly 64 %) of respondents participated in dealings concerning further investments. The remaining one third (36 %) either did not participate in such discussions (18 %) or were not employed (18 %).

A more illustrative presentation of distribution of relative frequencies resulting from the first degree classification of their competency is presented in Fig. 4 - 1.

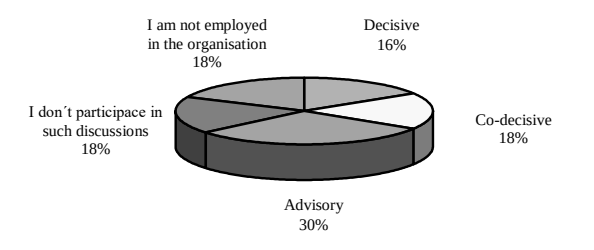

Figure 4 - 4: Competency of visitors to decide about company investments

To avoid the aforementioned disadvantages of closed questions we can use also the variant "Other", which *de facto* represents a free statement (similarly as an open-ended question). In this case the respondents can also give answers differing from presented variants. This corresponds with *semi-opened (semi-closed) questions*, which unify advantages of closed questions (i.e. presented answer variants) with those of an opened question (offering respondents to answer freely). Similarly it is also suitable to allow respondents to express their ignorance or indecisiveness by the selection of variants "I don't know", "I can't say", "I am not sure ", or "Neither - nor ".

Example 9 Formulation of semi-opened questions

Which is your favourite make automobile?

Škoda **1**

VW **2**

Ford **3**

BMW **4**

Mazda **5**

Honda **6**

Other (please specify) ……..

The open-ended variants of the answer "Other" must then by processed in a similar way as open-ended answers in the above Example 7.

Filter questions

Filter questions enable to classify respondents and to direct them by means of subsequent questions into different subgroups corresponding with their concrete situation (i.e. on the base of their answers to these filtering questions). When constructing the questionnaire, it is necessary to direct respondents clearly (and also highlight graphically) to subsequent questions they should answer with regard to their answers to the filter question and/or which questions should be omitted.

Example 10 Filter questions

So for instance, in a survey concerning automobilism the filter question could be that asking if the respondents own a car (motor vehicle). If yes, they are asked to answer also other questions (e.g. which make, how long do they own it, how many kilometres per year do they drive, how often and for which purpose the car is used etc.) while those who answer negatively are asked to answer quite different questions (e.g. why they do not have a car, if they intend to buy a car, if they already owned a car etc.)..

This means that car owners do not answer questions given to non-automobilists while those who do not own a car skip answers given to automobilists. After this filtering question (similarly as before it) all other questions given to respondents are undoubtedly identical (e.g. how and where the spent their holiday, how do they travel there, what are they doing during their holiday etc.).

Manipulation with the questionnaire

This part concerns mainly distribution and recollection of questionnaires. The most frequent method of distribution is either to send them by post (including e-mail and internet) or by personal contacts. Although the distribution by post, and above all electronic (e-mail), is cheaper and quicker but personal contacts may highlight and emphasize the importance of the current survey. A certain disadvantage of electronic distribution of questionnaires consists in the fact that it need not be delivered to the addressee, especially when it is evaluated as a SPAM by the software filter.

Filled-up questionnaires may be returned also by post (including e-mail), posted into special boxes or collected personally. In the first two cases the respondents usually feel more anonymous but in the third one the *response* is much higher. The response represents an important parameter of good construction and fruitfulness of a questionnaire because it indicates how the questionnaire was accepted by respondents and how many of them filled up received questionnaires and returned them back. The *response* (rate of return, recoverability) is calculated as the percentage of all distributed questionnaires. The response is often dependent on the attractiveness and interestingness of the questionnaire. This concerns above all its content (theme). So, for example, if the respondents are directly interested (or influenced) by a marketing problem (e.g. by construction of a new hypermarket near their place of residence), the response will be surely higher than in case of a general, trivial problem (e.g. malignance of smoking). The response can be also increased by tempting financial premiums and/or other awards given to cooperating customers (participants).

It is important to test the method of distribution already within the framework of preliminary research. The distribution and recollection of questionnaires by hired staff (e.g. by commissars participating in a census) is more expensive and demanding but the recoverability is usually nearly 100 %. Another advantage consists in the fact that the staff can explain the purpose of the survey and answer various questions put by respondents. On the other hand, the distribution of questionnaires by post is associated with a risk that the rate of return will be only 10 - 20 %. However, the recoverability may be in some cases as high as 90 %; this rate can be reached e.g. in cases of personal distribution and recollection of questionnaires or when the investigated problem directly influences the life of respondents.

The *checking up of completeness and correctness of filled and returned* questionnaires represents another important step of the marketing research. Incomplete and/or wrongly filled up

questionnaires should be eliminated from a further statistic processing because they bias and falsify the results. However, these shortcomings and deficiencies as well as a too high percentage of questions “I don’t know”, “I don’t understand”, “I can’t answer” should be identified and eliminated already in the stage of preliminary survey.

As it will be discussed and illustrated in Chapter 5, a low proportion of correctly filled up questionnaires seriously endangers the results and reduces the size of the sampling unit under study.

4. 2. 2 Public inquiry

The pubic inquiry can be used when introducing the company and addressing the public. It usually consists of only one or a few questions concerning a specific problem. The inquiry may be published in newspapers, broadcasted on radio or TV or distributed at various other occasions, e.g. among people doing shopping. Its authors (media, advertisement agencies etc.) use this tool to ask the maximum possible number of people and to ascertain their opinion about a certain problem. However, the public inquiry may be used also as means of public voting when organising various contests and competitions. Similarly as in case of written communication performed by means of questionnaires, also within the framework of a public inquiry the rate of return and participation of respondents may be increased by some promised attractive rewards, distributed, for example, after drawing of returned leaflets. The structure of the sampling unit need not be representative even in situations when a great number of answers are recollected. As known, filling up of leaflets, phoning or sending SMS attract above all only certain groups of respondents, namely those who have more time (pensioners, women on maternal leave, pupils, young students etc.) while employees, businessmen, managers etc. do not have time to participate and answer only very rarely. A *random self-selection of* participants is therefore one of the basic shortcomings of public inquiries.

In spite of the aforementioned problems is the use of public inquires in marketing research and communication with customers quite certainly justified. There is no doubt that is can interest and attract people, address them and establish firm relations with public but it is also necessary to emphasise that it should not be used as a base for concrete marketing decisions.

4. 3 Personal interview

As shown in Fig. 4 – 2, personal interviews are done as a dialog between the inquirer and only one respondent. The inquirer reads questions (possibly also with variant answers) formulated by the researcher and records respondent’s reactions (answers). As compared with questionnaires, the personal interviews are more demanding not only from the financial but also temporal and organisational points of view. First of all, it is necessary to recruit a sufficient number of capable (skilled) people – inquirers. They should be distributed (e.g. with regard to the population density) within the whole study territory (e.g. district, region, country etc.) and create a corresponding *network of inquirers*. Individual inquirers must be at first properly trained. Before each inquiry, it is necessary to instruct them about methods of selection of respondents, course of interview, and above supervision of their activities and field cooperation with respondents. Although they receive exact instructions how to proceed, they differ from each other, behave differently and sometimes even formulate questions quite differently. From this point of view we can imagine an ideal inquirer as pre-programmed robots who communicate with all respondents in the same manner (i.e. read identically all questions and record all the answers). Unfortunately, this is not the case: quite on the contrary, the inquirer’s ego undoubtedly influences the respondent both positively and negatively. For that reason it is

necessary to control their work adequately, and to upgrade and modulate regularly the network of inquirers. Another disadvantage o personal interviews consist in the fact that, as compared with questionnaires, it provides a lesser feeling of anonymity, especially if the inquiry takes place for instance in the respondent's flat.

On the other side, however, we know exactly who answered our questions. The inquirers themselves can also *accommodate to the situation*, establish contacts with respondents, explain problems, show various things and objects of presentation (e.g. labels or packages, photographs etc.), show cards with pre-scribed codes of variant answers etc.

Example 11 Cards with pre-set variants of answers

If we want to find out how many cigarettes the respondent smokes per day, we can expect that this is an undoubtedly sensitive, and for many smokers unpleasant questions. For that reason the inquirer usually offers to respondents a card with the following variants and codes of answers:

Up to 10	**1**
11 – 20	**2**
21 – 30	**3**
31 – 40	**4**
Over 40	**5**

For respondents this type of answers is undoubtedly more acceptable and the duration of the interview is quicker because they can only specify code numbers (e.g. **4**).

In case of personal interview, however, problems with returnability as well as correctness and completeness of answers required in a questionnaire fall off. The interviewers only adhere to instructions and fulfil researcher's requirements. Besides, they are remunerated on the base and quality of performed interviews.

In personal interviews, we usually use *record sheets*. As we will see in the following example, they are much simpler than questionnaires. These sheets are only ancillary tools for the inquirers who are adequately instructed how to select respondents, how to address them, how to ask questions, how to record their answers etc. This means that in case of personal interviews are the introductory texts of questionnaires quite unnecessary, similarly as instructions concerning filter questions.

Example 12 Example of record sheet

The personal interview may be also advantageously associated with observation (see Figure 4−5). Customers leaving the outlets were addressed only in case when the inquirer could see that they bought beer. Within the framework of these interviews, the inquirers also tried to learn about prices of competitive brands Starobrno and Radegast directly on the spot.

International Institute for Marketing, Communication and Entrepreneurship
Příkop 4, 602 00 Brno, Phone/Fax: 05/4517 6578
E-mail: info@iimce.cz http://www.iimce.cz

Interviews concerning the Starobrno beer brand

!!!Addressed should be only those customers leaving the shop who purchased beer!!!

1. Which beer brand did you purchase?

…….……

2. Why did you purchase just this beer brand?

…….……

3. Why did you not purchase the beer brand Starobrno? *(This question concerns only those people who did not buy the Starobrno brand – see Q 1)*

…….……

4. Did you earlier purchase Starobrno beer?

Yes **1** No **2**

5. Why do you not buy it now? *This question concerns only those people who did not buy the Starobrno brand – see Q 1)*

…….……

Identification questions:

Sex:	Male	**1**	Female	**2**
Age:	18 – 24 years	**1**	25 – 39 years	**2**
	40 – 54 years	**3**	55 and more	**4**

Place of data collection:

	Hypermarket	**1**	Supermarket	**2**
	Brněnka	**3**	V.S.B.L.	**4**

Price of Starobrno as compared with Radegast:

Higher	**1**	Lower	**2**	Same **3**

Figure 4 – 5: Example of the record sheet used in personal interviews

In concrete situations, the interviews may be performed in two ways, namely as:

• *Standardized interviews*, which are prepared in advance and which pre-determine what and how will be investigated; the sequence of questions and, above all, their formulation are quite rigorous (including possible variant answers;

• *Non-standardized (non-structured, free) interviews*, which do not have pre-determined formulation and sequence of questions. This does not mean that the interview would not be prepared and pre-mediated in advance but it is purposefully hold in such a way that it should be similar to a free conversation as much as possible.

The non-standardized interview and the focus group interview (see below) represent typical technique used above all in the domain of qualitative marketing research while the standardized ones are used above all in the field of quantitative marketing research.

4. 3. 1 Telephone surveys

The telephone surveys represent a modified and very operative method of personal interviews. Its main advantages lay in its speed and price. Respondents are hidden in a certain anonymity and can answer all questions in a more honest and opened manner. On the other hand, however, it is clear that a telephone survey must be shorter than a personal face-to-face interview. Telephonic surveys are used above all in the so-called *express inquiries*, which are usually ordered by mass media. The subject of inquiry is usually a topical ("hot") social theme (phenomenon) and the public opinion is published within a period of several days (but, as usual, even within mere 24 hours).

Advantages of telephone surveys involve also comfort and safety of inquirers while their main disadvantage consists in the fact that it is not possible to use visual tools; so, for example, in contradistinction to personal interviews, it is not possible to display those products that are in the scope of our interest or to show cards with codes and variant answers.

4. 4 Observation

Another technique enabling to collect primary data – observation – is performed by trained staff - observers. The situation illustrated in Fig. 4 – 2 remains top unchanged and the difference consists in the fact that the observer only observes and registers reactions, behaviour and properties of members of participating people (sample).

In this case it is expected that both the observer and the subject of research are objective and independent so that they do not influence each other. Provided that these requirements are fulfilled, the situation of observation is considered to be a normal objective form of collecting data about sample units of a research object (e.g. customers of a hypermarket).

The principle of the observation technique thus consists in *recording (registration)* of properties and behaviour of sampling units under study (customers). In each concrete, practical case of the aforementioned hypermarket it would by therefore possible to use long-term records e.g. of numbers of customers registered in certain time intervals (hours, days, weeks, months, years), lists of their characteristics (males *vs.* females, young people *vs.* seniors), records of purchased goods or frequency of shopping etc.

Mechanical observation devices, e.g. electronic cash registers with barcode readers, cameras or electromagnetic credit card readers may become a suitable tool for observation of customers. They can record for example exact time and method of payment, frequency of purchases

(including a detailed list of individual purchased items) and/or movement of customers inside the shop. However, it is necessary to check up the validity of these data. Cash registers cannot inform us about numbers of customers but only about paid purchases – because not all customers buy something when visiting a shop and many of buy and pay goods together with their partners, friends and other relatives.

Similarly as in case of interviews we can distinguish between *standardized* and *non-standardized* observations. When using the method of non-standardized observation, we only follow a certain objective of observation and the observer can freely decide about its course and parameters. This freedom, however, rules off a possibility of comparison of results recorded by various observers. The non-standardized observation may be held for a method of qualitative research, which enables us to obtain only a preliminary, orientation idea (information) about the problem under study. It can be said that it is preparatory stage of the subsequent standardized observation.

In case of standardized observation all phenomena, which should be observed are exactly defined and the observer classifies the obtained results into individual categories; the manner of observation (time and place), results and behaviour of observed persons are also defined. This means that the standardized observation may be included among methods of quantitative research. As mentioned above, the technique of standardized observation is used when collecting harddata.

Especially in case of standardized observation, the observer must classify and summarise a really huge amount of observed phenomena. It is obvious that also here we can use the *record sheet*. This sheet involves individual variables and categories that should be recorded in the course of observation. Also in this case the record sheet should result from the operationalisation of problem under study; however, as compared with surveys, it shows certain differences because it contains variables that are not formulated as questions but only as simple items (e.g. “the customer purchased/did not purchase a certain product”; “number of purchased pieces”; “total spending”; “time of visit” etc.). Similarly, it is also possible to record some obvious identification parameters (e.g. “sex of visitors”; “age”; “did they arrived by car?”; “were the customer alone or accompanied by someone?” etc.

Observations may be *direct* or *indirect*: 1) observed persons are awarded of the fact that they are observed or 2) the observer may be hidden and the customers do not know that they are supervised or controlled. The direct observation is performed quite obviously, the observer is present amongst the observed participants and uses some mechanical devices (camera, dictaphone, notebook). On the other hand, the method of indirect (hidden) observation is used in situations when the obvious presence of the observed could disturb natural (spontaneous) behaviour of observed persons. The observer is hidden and observes customers by means of a camera system, one-way glass (mirror) etc.

When using the method of observation, the researchers should select representative samples of population – e.g. all visitors of a hypermarket. In this case we want to respect the basic principle of quantitative research, which says that the obtained results must be generalisable and that they must characterise the whole population. Recently, many modern technical devices (turnstiles, barcode readers, people meters, hidden camera systems, cash register systems, and magnetic card readers) enable to register and record the behaviour of people in a very exact and objective manner. Moreover, data are recorded in a digitalised form so that they can be statistically processed and analysed immediately (on line).

4. 5 Experiment

In the field of marketing research, the term experiment is in fact used to label *each change in the offer*, i.e. in individual components of the marketing mix. Recorded and registered *reactions of customers* are counted as dependent variables while the offer itself represents an independent variable. The behaviour of respondents is thereafter derived from these changes. So, for example, when innovating products, changing their prices, using different distribution channels or implementing new advertisement campaigns, we should also observe and register the response of customers to these changes, i.e. to record changes in the volume and/or way of purchasing. In case that the sales increase, the performed change may be evaluated positively while when the situation remains unchanged (or even gets worse) it is necessary to try something else, something better. As already mentioned above, the technique of observation enables to obtain harddata. We can also investigate the response of customers (their satisfaction with and acceptation of performed changes) by means of personal interviews; this, however, would result in collection of softdata.

In marketing research, the problem of experiment consists above all in the fact that the behaviour of customers may be influenced also by other changes taking place above all in the macroenvironment (inflation and overall economic situation) but also by changed offer of competitors. For that reason the practical application of experiments as a technique of marketing research is significantly less frequent than in the field of natural and/or technical sciences where experiments represent a base of all research activities. In marketing, where we study individuals, groups or even more complex socio-economic structures, interviews and observation techniques are used much more frequently.

Experiments may be divided into two major groups: Until now we have discussed *field experiments*, which are established under normal conditions, e.g. in a concrete shop or supermarket. However, similarly as in the domain of natural or technical sciences we can establish and perform *laboratory experiments.* These are implemented in a specially organised, artificial environment, under laboratory conditions. Experiments enable to measure *physiological reactions* of people (skin humidity, eye movements or, recently, even changes in the magnetic resonance of brain) to various marketing stimuli (e.g. new ads or changes in prices, package etc.). Main problems of laboratory experiments consist above all in costs (high prices of measuring and recording devices) and in limited numbers of respondents willing to participate in such methods of marketing research. Nevertheless, the accuracy of results (exactly measured harddata) represents undoubtedly their greatest advantage.

From the viewpoint of *organisation and establishment of experiments*, we can distinguish two major groups:

1) Experiments, in which we *follow only effects of the independent variable*. Measurements are performed in two groups, viz. experimental and control (controls are not influenced by the independent variable while the experimental group is). In case that the difference in data measured in both groups differs from zero, it is possible to conclude that this is caused just by the independent variable. To increase the probability of this conclusion, we try to establish such experimental conditions that eliminate effects of external factors (*ceteris paribus conditions*);

2) In the second case there is only one experimental group and this is measured *before* and *after the effect* of the independent variable. In case that the difference between these two measurements is again different from zero, we can conclude that this is also caused by the independent variable.

The application of experiments shows certain problems. As already mentioned, we can hardly be sure that no other effects as that of independent variable are influencing experimental conditions. Another problem results from the heterogeneity of compared two or more groups. This can be solved in such a way that the homogeneity of experimental group is evaluated only on the base of one or a few traits (age, education) and the effects of other factors are eliminated.

When generalising obtained experimental data it is necessary to be cautious and remember that they were collected under artificial, non-natural conditions and, moreover, that their extent is rather limited (*n* is relatively small).

In the end of this chapter we can repeat some major principles of marketing research (see also Chapter 1) and raise the following questions:

1. *Who are our customers* (description of their basic socio-demographic characteristics)?
2. *Which are their living conditions* (household facilities, ownership of immovable and movable assets)?
3. *Which is their life style?*
4. *Which is their value orientation?*
5. *How do they behave and make decisions?*
6. *How do they communicate and with whom?*

When looking at them from the viewpoint of the aforementioned techniques of marketing research it is quite obvious that in all six cases we can use the technique of interviewing (surveys). On the other hand, experiments and observations cannot be used at all (or to a very limited extent – when studying behaviour, decision-making and communication). This means that the technique of public surveys (interviews) is used most frequently when studying customers and their habits. It can be even said that this is the most frequent tool of marketing research and for that reason it was discussed so broadly.

Selection of a suitable technique is dependent above all on research objectives, nature and necessary extent of collected data, type of respondents, time consumption, and costs (financial possibilities). Each of discussed techniques has some advantages and some disadvantages and it is necessary to consider them very carefully. It is not possible to say that one of them is the most suitable tool of marketing research. In practice, individual marketing research techniques are usually combined; this was clearly shown in record sheets concerning purchases of the Starobrno beer brand (see Fig. 4 – 5).

Discussion questions:

1. Which is the difference between marketing research methods and techniques?
2. Can you enumerate the most important techniques of marketing research?
3. Which methods of public inquiry/interviewing do you know?
4. Which are the advantages and disadvantages of open-ended questions?
5. Which are the advantages and disadvantages of semi-opened questions?
6. Which are the advantages and disadvantages of closed questions?
7. How would you construct dichotomic questions concerning age and profession?
8. Which requirements should a questionnaire meet?
9. Give an example of suggestive questions.
10. What it is a "halo effect "? Give an example.
11. What it is a "filter question"? Give an example, including the subsequent content-targeted inquiry.
12. Fabricate some examples of open-ended, semi-opened and closed questions that could be used within the framework of a survey on the theme "*Satisfaction of our customers*". Which are their advantages and disadvantages as far as the respondent and the researcher are concerned?
13. What does mean the term "scaling"?
14. What do you understand when speaking about the detaileness of a scale? Are there some possibilities how to solve this problem?
15. What do you understand when speaking about the form of a scale? Are there some possibilities how to solve this problem?
16. What does mean the term battery of questions?
17. Which are the advantages and disadvantages of personal, telephonic and written questioning as far as the respondent and the researcher are concerned?
18. How would you use the observation technique within the framework of a survey on the theme "*Satisfaction of our customers*". Which are its advantages and disadvantages?
19. How would you use the experiment technique within the framework of a survey on the theme "*Satisfaction of our customers*". Which are its advantages and disadvantages?
20. In a survey of operational staff of the nuclear power station Dukovany the data recovery rate was 83 %. How would you evaluate this result and which were its most probable causes?
21. When you study the record sheet of a survey concerning the beer brand "Starobrno", which items could be investigated using the observation technique (see Fig. 4 – 5)?
22. Why is the final questionnaire of the survey *Image of the PCES Znojmo* (Fig. 3 – 1) less common, especially when comparing it with Figs 4 – 3 and/or 4 - 5?

5 OBJECT OF MARKETING RESEARCH – SAMPLE SPECIFICATION AND SELECTION OF UNITS

Chapter objectives: The object of marketing research can be understood as a set of units which, with regard to their number in the field, can be examined either exhaustively or made from them sample. For each sample survey, the representativeness of the final sample is a crucial issue for everyone. A specific example shows how the representativeness of the selected file can be checked. There are also very practical instructions for using random or quota selection.

Key words: object of marketing research, exhaustive research, sample, sample representativeness, probability (random) sample, quota sample.

At the beginning of Chapter 3, the problem (subject) of marketing research was gradually unfolded, concretized, and disaggregated with regard to the practical application of the obtained results. Now we will proceed in a similar way in the determination of the object and its units, which will be investigated. The *object of research* may be any socio-economic formation with duly determined basic temporal and spatial coordinates. In concrete cases, these may be for example Czech citizens of the age of eighteen and more, all childless families in the Southern-Moravian Region, or non-profit organizations in the Znojmo district, which showed any activity in 2007 etc.

A research object as an aggregate (set) consists of *units*. Units of the set may be not only individuals (persons), but also, as it results from the previous sentence, also groups (families, working collectives, organizational bodies, school classes, organizations, municipalities etc.). The unit is a carrier of properties, which are subject to the examination. If this unit is a person, we are usually interested in such personal qualities as basic socio-demographic characteristics (sex, age, the highest achieved level of education, domicile, economic activity etc.) and possibly also their income but above all their satisfaction with the product, the willingness to buy it, and many other marketing data. In case that the sampling unit is a group (e.g. a household, an organization, a municipality), these will be the properties of the whole group (e.g. its size, ownership of certain products, its needs etc.).

It is evident that in cases when the sampling unit is a group, the situation is getting complicated. Again (with respect to the goals, the practical implementation of marketing research and the application of the obtained results), it is necessary to decide if every single member of the group should be examined (for example each family member, each company employee, each municipality inhabitant), or possibly only some of them (e.g. only those over eighteen years), or just one of them will become a respondent to answer the question on behalf of the group as a whole (head of the household, chief of the department, municipality mayor).

In contradistinction to psychology, in the field of marketing research we are hardly ever interested in the respondent *per se*; for us, respondents are only members of a certain social group – object of research. The set of all sampling units, identically delimited from the viewpoint of our research, is called the *parent population*. We learn about it through an exhaustive or non-exhaustive (selective) survey of its units.

The exhaustive survey examines all units of the parent population, which may be very demanding from both financial and temporal points of view, especially in case of large parent

populations (population of the whole country). Submitting each unit to the examination need not necessarily mean that the evidence value of these results will automatically be higher than in the case of the non-exhaustive (selective) survey. The data collection and their statistic processing in large sets, such as in case of the population of the whole country, may last for years. The result then is the fact that the detected results are exact and characterising all units, but alas they are also outdated. A typical example of the exhaustive survey is the census, election results, or certain forms of statistical documents.

On the contrary, a big advantage of the exhaustive survey consists in the fact *that the problem of set representativeness and of calculation of the random sampling error falls off.*

The non-exhaustive (selective) sampling is used at most. It presumes the examination of only a certain part (sample) of the set of sampling units. This means that the whole research is then generally cheaper and less temporally demanding. Its disadvantage consists is the fact that obtained data as well as the calculated statistic characteristics relate only to a limited number of sampling units and that it is necessary to perform additionally their generalisation for the whole sampling set (population).

As examples of the non-exhaustive survey we can mention not only various types of customer marketing research, sociological research and/or public opinion research (especially at the macro-level) but also many surveys of the Czech Statistical Office (e.g. micro-censuses). For the whole of Czech Republic, the size of the examined sample is most frequently approximately 1,500 respondents but sometimes it may be even ten times bigger. In the following text we will demonstrate which selective procedures can be used and which advantages and disadvantages must be considered when using them.

A generalisation of findings discovered in a surveyed sample and their application in the whole population usually takes place by means of *statistic induction* (i.e. generalisation). However, the basic question in this case is whether the selected set of sampling units can be considered as *representative* – i.e. if and to what extent the selected sample represents as an exact miniature the basic parameters (variables) of the parent population. Or, in other words, to what extent it can be considered as a representative of the whole set of sampling units. If it is really so, then there is no obstacle for a generalization of the obtained results and their application within the framework of the whole population. For example, due to the representativeness of a 1,500 household panel in people-metre survey of TV ratings, the results observed *count* for all three million households of the Czech Republic. But if the sample is not representative enough, such a generalisation is not possible, or it will be erroneous and misleading. The obtained results are valid only for the examined sample.

Sample representativeness

In case of permanent populations, such as registered university students, employees of an organization or municipality inhabitants, there are available values of the basic identification (socio-demographic) characteristics of the whole object. The numeric data from these characteristics are then compared with values of equal traits of the examined sample – see Example 5 – 1 (Implementation and evaluation of the quota selection). *Representativeness of the surveyed sample therefore means the conformity (similarity) of the values of its selected basic identification characters (sex, age, marital status, education, domicile etc.) with the values of identical traits of the whole parent population (object).* The less the values differ, the more representative the sample is. This principle is then used in the procedure of the so-called purposeful, quota selection, as we will show in the paragraph 5. 2.

However, if we do not know these data (usually in the so-called *unknown, unsteady, "fluid" populations* such as visitors of various events, tourists etc.), *the representativeness is secured by the very process of the unit selection.* We rely on the mathematic *theory* of *probability*, on a genuine *incidentalness* (not only randomness) of the selection process, which in practice means that all units of the parent population have the same probability (chance) of being selected into the sample.

While considering the manner of examination (regardless if it is exhaustive or not), it is also necessary to consider if we select a *permanent respondent population* (the so-called *panel*), or if we perform a significantly more common, more often used *single (ad hoc) selection* for the given event. This decision also depends on the specific situation: if we examine the population more often or only on a one-off basis, if we have the means to continue with it etc. *The panel population* therefore provides a representative permanent set of respondents who are being monitored repeatedly for a certain longer period using a selected technique of interviewing or observation. It should offer information about changes of the observed phenomenon in time. Depending on the purpose of survey, the panel is formed either for a temporary period of time, or it may be of a relatively steady type. The formation and maintaining of a good panel is a very demanding task. The danger is that after some time the units stop "behaving normally", they realise their exceptionality (that they are objects of the examination). Therefore it is suitable to gradually modulate them, changing one part after another, for example one third after the first year, another third after the second year etc. A practical example of the panel is the aforementioned population of 1,500 households in the people-meter survey of TV rating in the Czech Republic.

5. 1 Probability (random) sampling

The probability, or random, selection begins with a clear *definition of the sampling unit* (customer, household, company etc.) of the population, and subsequently with a selection of a suitable *sampling technique*. For example, the most often used ones are ballot or random number generator. They require a list (record) of all units, which is sometimes difficult to realise. Other techniques are based on an *uncorrelated character*. For example, we observe living standards of households and we choose a random letter, with which the surname of the families addressed will begin, because the dependency between the surname's initial and the family's living standard is not presumed. Another possibility is represented by a *random walk*. The instructions itemized in detail of a standardized "wandering" will take the inquirer to a completely random (blind) selected domicile and person of the respondent.

The simplest form of the probability sampling is called the *simple random sampling*. It is a direct selection with the same probabilities for all units.

Example 13 Procedure of a simple random sampling in practice

If we have a record of all the units of the whole population, it is possible to apply the simple random sampling very easily, for example in the survey of municipality inhabitants or company customers. Let us imagine that the whole population contains 5,000 individuals, and we will suffice for the research with only 500. Therefore we will select randomly every tenth member of the sampling set.

For more clearness, we may imagine that all 5,000 names and addresses are on cards that will all be inserted in a large lottery wheel. Such as in every regular ballot we will mix them thoroughly and gradually select the required 500 from among them. It is evident that if we follow appropriate conditions, (the cards are well mixed, they are drawn blindfold, really at random), all those in the wheel-barrel have the same probability of being picked in every draw. The mathematic principle of the probability selection procedure therefore warrants that the selected 500 individuals are the representative sample of the whole population of 5,000 individuals.

In practice, however, we would realize such a procedure in a more simple way. We would take a list of the 5,000 names and select each tenth. To increase the probability, we would not start with the first one, but as it is, for instance, 13 April, we would pick the 13^{th} name from the list, and then 23^{rd}, 33^{rd} etc.

5. 2 Quota sampling

In comparison with the simple random sampling, the intentional, quota sampling hast the advantage of not needing a record (list) of all the units of the parent population, but there is a sufficient and perquisite knowledge of its basic statistical characteristics. Based on these values we will then establish the specific procedure of the unit selection. For example, from the Czech statistic office's data we learn about the socio-demographic characteristics of the inhabitants of a certain territory, and according to these data (their percentage rate) we will then draw quotas for the final selection so that also in the selected population the appropriate proportionality of the individual socio-demographic groups remains to be maintained. In the very selection, we usually proceed by establishing the so-called *control* characters, significant for the population characteristics (for example age, sex, education level, profession, domicile). According to their values in the parent population, we establish the numeric values of selective control characters (quotas), which must be kept (fulfilled) by the inquirer.

Example 14 Definition of the quota sample

Tab. 5 – 1 shows that from the 2001 census we learnt that out of the Brno inhabitants of over 18 years of age, 21 % live directly in the city centre and 79 % in other town quarters. Therefore, as far as the control character of domicile is concerned, each inquirer would receive a direction that among his ten respondents, there should be 2 from the city centre and 8 from elsewhere. And it is exactly these specifications (2 and 8) that are called quotas, from which the whole selection procedure takes its name. Quotas for other control characters in Tab. 5 – 1 are established in a similar way.

A necessary condition for percentually expressed shares of control characters of a selective population is their knowledge per the parent population. From the statistical point of view, it is possible to say that the random selection is more preferable because it secures better the statistical representativeness, and it is possible to calculate the random sampling error in it, but its implementation in practice is not always easy. The quota selection therefore represents a sort of fall-back increases when we are not able to realize the probability selection, but on the other

hand we have the necessary statistical characteristics of the whole parent population. The quota selection principle rests in trying to create a perfect simulation, a miniature of the parent population.

Sometimes, in the selected populations there occurs a so-called *auto-selection*. We encounter it especially, as we have said in the previous chapter, in an opinion poll, or possibly also in a questionnaire. During visits of homes we do not catch all selected respondents at home, or they refuse to reply. If the percentage of the non-communicating respondents or of refusals starts to appear within the range of 15 - 20 %, there is a threat, similar to the opinion poll, of a non-representative sample. Therefore it is necessary to consider what respondents did not respond to us and why, and most significantly, how (by tracking back or other adjustment of the already examined sample) we will create a representative set.

In some cases it is suitable to realize a random selection with a quota back-tracking. The point is that after the first random selection we will check the values of the check characters, and on the basis of the discovered differences against the parent population, we will carry out the back-tracking with an intentional quota selection. But also here it is necessary to accept the premise that the differences from the first step are not dramatically significant.

Example 15 Implementation and evaluation of quota sampling

For the research of inhabitants of the city of Brno for the Brno Municipal Office performed in the period of November – December 2001, the method of quota selection was used. The following table shows the comparison of the individual quota characteristics of the whole population and the set of sampling units.

Socio-professional position

	*Whole population **	*Quota sample*
Owner of a firm, businessman	8 %	*10 %*
Employee, administrative worker	33 %	*38 %*
Other professions	8 %	*1 %*
Manual worker, operational staff	16 %	*16 %*
Student, apprentice	9 %	*6 %*
No economic activity	26 %	*29 %*

The highest finished education

	*Whole population **	*Quota sample*
Primary, unfinished	18 %	*22 %*
Incomplete secondary, apprentice	36 %	*36 %*
Leaving examination, A level	35 %	*31 %*
University	11 %	*11 %*

Age

	*Whole population **	*Quota sample*
18 – 24 years	15 %	*13 %*
25 – 39 years	28 %	*26 %*
40 – 59 years	37 %	*36 %*
60 and more years	20 %	*25 %*

Sex

	*Whole population **	*Quota sample*
Male	45 %	*47 %*
Female	55 %	*53 %*

Domicile

	*Whole population **	*Quota sample*
City centre	21 %	*19 %*
Other urban districts	79 %	*81 %*

* *Source: Data published by the Czech Statistical Office (preliminary results of 2001 Census and of a survey performed in the 2nd Q 2001)*

Table 5-1 Comparison of basic sampling units and quota samples

A comparison of values in both columns (i.e. the whole population and the set of sampling units) shows that the differences are quite insignificant. Therefore the sample of almost 500 respondents can be considered a *representative* selective population of the inhabitants of Brno older than18 years of age.

Example 16 Selection of sampling units in practice

In examining the satisfaction of visitors and exhibitors at international trade fairs, we were choosing the units in two ways. It was because the fact that in both cases they responded to the different subjects as well as objects of both researches.

The interviewing of visitors was taking place during the whole time of the fair, i.e. from the first day until the last. Equally, in the course of each day, the leaving visitors were counted and consequently selected at random. Duly trained and controlled inquirers carried out prepared standardized interviews with the individuals selected that way. Out of the total number of over 100 thousand visitors, about a thousand interviews took place.

On the contrary, the interviewing of exhibitors took place only during the last three days, i.e. at the moment when the exhibiting companies could balance and evaluate their participation at the fair on the basis of results achieved and experience. In total, at least 10% of the exhibiting companies were selected at random from the catalogue. In addition, the task of our interrogators was to carry out a standardized interview with the possibly highest-ranked and therefore duly competent representative of the exhibiting firm. That is testified by the data about the position of the interviewed persons at the 41st International Engineering Fair:

- in 12 % it was either the owner or co-owner of the company,
- in 33 % it was a director or manager of the company,
- in 8 % it was a direct guarantor of the exposition,
- in 23 % it was other managing staff,
- in 24 % they were other staff.

The manner of selection of the visitors as well as exhibitors, as well as the chosen technique of a personal interview led in practice to trying to minimise the number of inquired characters (number of questions asked)

5. 3 Size of a sampling unit

While selecting the population, also its size is substantial. It has its significance especially for the determination of the random sampling error, furthermore also for the possibility of higher grades sorting and the significance level of the observed statistic relations. On the other side, it is evident that the larger the population is, the more exigent the survey will be from the temporal and financial point of view. Sometimes it is erroneously considered that the larger the sampling population, the more representative it is. The *size of the population and its representativeness are two different domains*. Even the practical results of pre-election surveys in the USA at the beginning of the 20th century convinced the researchers that a large population counting several thousand, but originated by auto-selection in various public enquiries will hardly be representative, whereas a few hundred appropriately (probability or quota-based) selected units will meet the requirement of representativeness much better. If you have large (tens of thousands of respondents), but unrepresentative sets (originated by auto-selection in public enquiries), the

generalisation of these results onto the whole population will bring absolutely erroneous conclusions.

As was stated at the very beginning of this chapter, the common size of an examined sample for the whole of our republic oscillates around 1,500 respondents. In some cases, however, it tends to be essentially larger. It may come close to 20,000. The reason for such a large sample is, beside an effort for a bigger precision of the results observed, most of all the fact that for further statistic processing (classification) it is necessary that also the little numerous phenomena should be represented in a sufficiently large numbers. For example, in a media research, the buying public of certain local media (radios, newspapers) form only a negligible percentage in the national sample. But to be able to describe their socio-demographic profile, we need to have their absolute percent occurrence minimally in the order of tens. And this is why the whole sample will grow several times.

Discussion questions:

1. Try to characterise the difference between the population (universe) and a sampling unit. Describe their advangtages and disadvantages.
2. Which are the typical features of a panel?
3. Which requirement should a sampling set meet?
4. Which are the most frequently used sampling methods? Which are their advantages and disadvantages?
5. Within the framework of elaboration of the strategy of a further development of the microregion Podluží, a questionnaire survey was performed in altogether seven villages (Hrušky, Kostice, Lanžhot, Moravská Nová Ves, Moravský Žižkov, Tvrdonice and Týnec). The total number of processed, correctly filled up questionnaires was 2,778. Why was this number so high? Which was a sampling unit in this marketing research and were its advantages and disavantages?
6. How many and which quota variables were used in the example of quota sampling performed among Brno inhabitants? How was in this case exactly defined the object of this research?
7. How would you characterise random samplings described in the example of survey performed among visitors of and exhibitors at the Internatonal Engineering Fair in Brno? How would you perform a random sampling of exhibitin firms by means of the fair catalogue (describe this process step by step)?
8. Why and when is the size of the sampling set so important?
9. Which is the relaltionship between the size of the sampling set and its representativeness?
10. How can we test the representativeness of the selected sample?
11. Thera are 600 persons over 18 in a village. The municipal office wants to learn about their satisfaction with the local situation. Which are three possible methods of implementation of this survey?
12. What do you unerstand under the term *marketing probe*?

6 STATISTICAL PROCESSING OF PRIMARY DATA

Chapter objectives: The basis of the whole process of statistical processing of primary data from marketing research is the distinction of individual features into nominal, ordinal and cardinal. Subsequently, the appropriate statistical procedures are selected. Another necessity is the coding and categorization of input data. On the other hand, the level of sorting is first selected for the resulting outputs, when especially the sorting of the second stage already allows to search for and present statistical connections between two features. Depending on the characteristics, suitable coefficients can be selected to measure the statistical context.

Key words: variables, categorisation, coding, classification, contingency tables, statistical relationships among variables, absolute and relative frequencies, statistical coefficients, interpretation.

In contradistinction to chapters 3 and 5, in which both the subject and object of marketing research were gradually decomposed, the principle of statistical processing is a *synthesis, i.e. aggregation of partial data into an organic whole*. This process runs on both levels, i.e. at the level of marketing research subject and also of object. At the level of subject we try to obtain an answer to the maximum possible proportion of the problem under study. Similarly, also at the level of object we do our best to learn as much as possible about individual surveyed units so that we can get a better knowledge of the whole investigated or (better to say) basic set of units (population).

Statistic methods used for processing of data collected in the course of marketing research differ not only as far as the survey objectives are concerned but also according to the their character. As we already know, the values of traits under study are estimated on the base of *variables.* A variable serves as a tool enabling to measure properties of individual units making a sampling set (population).

When using the technique of *interviewing*, a question with pre-defined variants of answers represents a variable. In case of opened questions, these variants are created after a supplementary categorisation of obtained answers.

So, for example, the following variable (i.e. a closed question) originates from a record sheet used in interviews about purchasing the Starobrno beer brand (see Fig. 4 – 4):

Did you earlier purchase Starobrno beer?

Yes **1** No **2**

In the following example, however, the investigated variable was represented by an open-ended question (it was used in the same record sheet):

Which beer brand did you purchase?

…….....…….....……

When using the technique of *observation* the variable used in the same survey was formulated as follow:

Did the respondent buy beer?

Yes **1** No **2**

Variables may be considered from different points of view. Above all, they can be classified according to their content as *identification marks* and observed *variables*.

Identification variables inform researchers about basic (essential) socio-demographic parameters of a sampling unit. Within the framework of statistical processing (especially in case of grouping of the second and/or higher degrees) they are usually classified as independent variables. They are used when calculating statistical coefficients of correlation (dependence), which indicate the effect of socio-demographic variables (i.e. differences) on reactions of respondents (their answers).

Observed variables are used when we try to identify the investigated problem (subject of research), e.g. when ascertaining which transportation means do the respondents use to get to their workplace, how are they satisfied with the attitude of their supervisors etc. Both groups of variables were already illustrated in Example 3 – 2 (operationalisation of research subject) and construction of a questionnaire (Fig. 3 – 1).

Properties of sampling units may be of either *qualitative* (verbal) or *quantitative* (numeric) character. Variables used when estimating the value of a concrete sampling unit are then constructed in dependence on this fact. Variables are classified either as *nominal* (qualitative, verbal) or as *numeric* (quantitative).

The adjective *nominal* (qualitative, verbal) is used in case that the collected data are expressed in words (e.g. male - female, healthy – sick, the most favourable colour, preferred products, name of preferred communication media etc.). As we already know, answers may be *dichotomic* with only two alternative answers (e.g. male – female) or *polytomic* with several variants (e.g. nationality, citizenship etc.). All variants are equal and it is not possible to say that one is better or worse. So, for example, the question: "Where do you preferably go on your summer holiday?" is of nominal character. If it would be open-ended, the respondent should mention his/her preferred destination. Anyway, these answers must be at first categorised and coded; individual variants are numbered (but these code numbers do not express any numeric values as in case of quantitative variables). Nominal variables express only different qualities values of a phenomenon and digital marks are used only for an easier computer processing. When answering closed questions, the respondent can select directly one of pre-set and coded variant.

Numeric variables express quantitative character of collected data. They can be classified either as ordering or cardinal (measurable). In this case, questions and answers may be evaluated either as dichotomy or as polyatomic – e.g. pupils with the average study result up to 1.5 belong to one group while others, with the average grade higher than 1.5 (dichotomic) to the other. When grouping pupils according to their individual grades, we use a polytomic classification.

Ordinal (ranking) *variables* express different, graded, and changing levels of a variable and its occurrence. Their typical examples are classification of students in a class, wage categories, accomplished education etc. In this case we can say which variant is higher and/or lower, which means more and which less, and which evaluation is better and which worse. After all, however, s compared with measurable variables, the informative content of ordinal variables is lower. As

an example of ordinal variable it is possible to mention the age differentiation of visitors of the theatre performance for children into three groups: “Children”, “Parents” and “Grandparents “.

Cardinal (measurable*) variables* are collected by means of common methods of numeric classification (i.e. measuring, weighing, and counting). As typical examples not only time, temperature, height, or weight may be mentioned but also some economic values (volume of production, costs, income, profit, credit etc.) expressed either in natural or in monetary units. In this case we can express not only difference in quantity but also how much is the difference (e.g. temperature) or how many times is one value greater than the other (age). As an example it is possible to mention the following (closed) question:”How old are you?” with pre-set variant answers 15 – 19 years; 20 – 34 years; 35 – 49 years; and above 50 years.

There is no doubt that this rank of variables characterises also the *accuracy* of measuring of a given property. Categorisation of age into categories “Young”, “Middle-aged” and “Old” is definitely less exact than the aforementioned variants.

However, this classification is important above all from the viewpoint of statistical processing and analysis. As shown below, we can (or cannot) use certain statistic characteristics in dependence on the nature of variables under study. So, for example, there is no point in calculating such statistical parameters as mean or median. In this case we must be satisfied only with modus, i.e. with the most frequent variant occurring in a set of sampling units. Similarly, there is no point to use such characteristics as dispersion and/or standard error. Coefficients of correlation of two different variables will be mentioned in the end of this chapter.

For the sake of completeness it is also necessary to add that the measurable variables are further classified as *continuous* and *discrete variables***.** Within the framework of statistical analysis of socio-economic phenomena we usually use only discrete characteristics.

When summarising the *methodological essence* of marketing research it should be obvious that in the investigate *object* (or in its units) interesting are for us only *properties* (e.g. “Age”). Selection of these properties should be based on the *operationalisation* of a problem (subject of research). As already mentioned, c*ondition* (value) *of a property*, is estimated (investigated, measured) by means of *variables*. Depending on the research technique, each trait/variable under study may be defined as an item (in case of observation) or as a question (in case of interviewing). So, for example, the aforementioned property of “Age “ could be defined either as “Young”, “Middle-aged”, and “Old” (in case of observation) or as variants 15 – 19 years, 20 – 34 years, 35 – 49 years and 50 and more (in case of interviewing)..

As one can see, not only the *nature of a property* (sex, preferred brand etc.) but also the *construction of a trait/variable* and *technique of its ascertaining* determine its type, i.e. if we will operate with a nominal (ordinal) or even with a cardinal (measurable) variable. It is quite clear that in the former case (observation) we can receive *harddata* while in the latter (interviewing) *softdata*.

The method of construction of a variable also decides about *accuracy of obtained values.* The age of respondents can be ascertained not only by means of the aforementioned variants (15 – 19 years, 20 – 34 years, 35 – 49 years, and 50 and more years) but also more accurately by using the year of their birth or even quite exactly on the base of their date of birth.

6. 1 Categorisation

The term “Categorisation” means an exact specification of individual variants of answers (estimated values) of variables, which will be further processed. Sometimes the categorisation results directly from the operationalisation of a property (such as in case of “Sex” with variants

"Male" or "Female"). More often it is selected with regard to the question what should individual answers contribute to the analysis, which operations will be performed with individual categories, and which techniques will be used for their identification.

Thanks to different categorisation we can proceed with a different rate of accuracy. As already mentioned, the age of respondents can be specified either by means of variants 15 – 19 years, 20 – 34 years, 35 – 49 years, and 50 and more years or, in a more detailed manner, according to individual years of age (15 years, 16 years, 17 years etc.).

The categorisation is important above all when proceeding open-ended (or semi-open ended) questions when we must at first record all answers of respondents. Thereafter the most frequent of them are selected and the remaining one are involved into the category "Others ". In this way we could link to the example *Formulation and evaluation of opened questions* presented in Chapter 4 – Techniques of marketing research, where the question "Which other Czech and/or foreign caves have you already visited?" was answered as follows:

• Český kras Caves – 66 times • Javořičské jeskyně Caves – 28 times • Mladečská jeskyně Cave – 27 times • Chýnovská jeskyně Cave – 17 times Jeskyně Na Pomezí Caves – 11 times	• Jeskyně Na Špičáku Caves – 11 times, • Vysoké Tatry Mountains – Twice, • Nízké Tatry Mountains – Once, • Romania – Once

Basing on obtained absolute frequencies, the answers could be categorised for example into the following four categories:

Český kras Caves	**1**
Mladečská jeskyně Cave	**2**
Chýnovská jeskyně Cave	**3**
Others	**4**

The first three of them represent the most frequent caves and the fourth one would contain all others.

A similar categorisation could be used also in case of semi-opened questions. Its initial categorisation of variant answers looked out as follows:

Brno	**1**		Central Bohemian Region	**6**
Prague	**2**		West Bohemian Region	**7**
South Moravian Region	**3**		North Bohemian Region	**8**
North Moravian Region	**4**		South Bohemian Region	**9**
East Bohemian Region	**5**		Foreign countries.....	**10**

Regarding results of processing of open-ended variant 10 and obtained frequencies of individual mentioned countries (see again the Chapter 4 Techniques of marketing research) we could re-categorise them in such a way that the Variant 10 would be Slovakia and Variant 11 would involve all other countries.

Example 17 Re-categorisation of contingency table

Sometimes, when the frequencies of some variants are too low, we can perform *re-categorisation* as late as during the process of classification into contingency tables. This is simply performed through unifying low-frequency categories into more extensive units. Low frequencies occurring in a higher number of boxes may for example bias the test criterion χ^2 (chi-squared), i.e. if the value of more than one fifth of all expected frequencies occurring in the contingency table is lower than 5 or if in one box the expected frequency is lower than 1. In such a case it is better to re-categorise variables presented in the contingency table (Tab. 6 – 1).

Satisfaction / Education	Definitely satisfied	More or less satisfied	More or less unsatisfied	Definitely unsatisfied	Total
Primary	7	30	22	1	60
Apprentice	19	15	8	2	44
Secondary	18	22	9	7	56
University	3	11	10	16	40
Total	47	62	49	42	200

Table 6 – 1 The original contingency table

As one can see in this table, some absolute frequencies presented in both extreme columns ("Definitely satisfied" and "Definitely unsatisfied") are even lower than 5. For that reason they were re-categorised into a dichotomic trait (see Tab. 6 – 2).

Satisfaction / Education	Satisfied	Unsatisfied	Total
Primary	37	23	60
Apprentice	34	10	44
Secondary	40	16	56
University	14	26	40
Total	109	91	200

Table 6 - 2 A contingency table with re-categorised variables

The categorisation is thereafter followed by coding.

6. 2 Coding

The term "coding" means that each categorised variable variant receives a certain (most often numeric) index (code). The major rule of coding is that the coding must be explicit (clear-cut). This means that coding contributes enables a quicker and well-arranged processing and storage of primary data. This concerns above all nominal (verbal) variables.

There are several methods of coding. Because of practical reasons (computer processing), the aforementioned numeric codes of each variant of answers seem to be the most suitable (see, for example, codes used in two model categorisations, i.e. (i) which other trade fairs the respondents already visited and (ii) from where they arrived) presented in above Section 6. 1. Respondents and/or inquirer should only circle the corresponding number. Recently, blank boxes attached to each variant became very popular and it is only necessary to check the selected variant of answer. However, this is not simpler than the aforementioned circling and besides it is also necessary to attach a numeric code to each box.

6. 3 Classification

The result of a field marketing survey is a great volume of collected, but unclassified (non-aggregated) primary data. Imagine that the ascertained values of individual variables represent in case of each respondent a vector and that the size of this vector results from the total number of variables. This means that if in a set of 200 respondents altogether 30 variables are ascertained, the matrix is 200 x 30 and the total number of items is 6,000. For that reason, our first task is to create a meaningful aggregation of collected data so that it would be possible to describe the whole set and to draw reliable conclusions.

The purpose of *classification* is to express how many respondents from the set under study selected one (or several) variants of a variable. When classifying the answers, we use the term *degree of classification*. *Classification of the 1st degree (one way frequency tables) means a mere listing of absolute (or relative) frequencies expressed as variants of one variable.* The relative frequency is usually expressed as percentages (or graphically as shown in Fig. 4 – 1 about competencies of visitors when deciding about investments). Of other statistical characteristics it is possible to calculate average values (mean, median and modus) as well as distribution parameters (e.g. variance, standard error etc.). The table presenting data collected among visitors of Znojmo is

an example of classification of the 1st degree. Among others, the respondents were also asked why did they arrive to Znojmo and what did they do at the Horní náměstí square. Absolute and relative frequencies are presented in Tab. 6 – 3.

Frequency	Doing shopping	Only passing	Going to the office	Going to work	Going to school	Going to doctor's	Tourism, sightseeing	Going home	Other	Total
Absolute	140	99	85	85	29	16	14	6	3	477
Relative	29 %	21 %	18 %	18 %	6 %	3 %	3 %	1 %	1 %	100 %

Table 6 - 3 Classification of the first degree

For a further statistical processing it would be suitable to re-categorise this variable, i.e. link together at least two (or possibly even four) last categories into only one category "Others".

6. 4 Contingency table

The classification of the 2nd degree enables to pick up those sampling units that show identical values of two variables – two ways frequency tables, crosstabulations. Especially in this case we look for relationships existing between frequencies of variants of variables of operationalised problem (subject of research) under study on the one hand and their identification parameters, e.g. domicile, size of the household (number of family members), their age, sex (male - female), education (the highest degree of finished school), socio-economic situation (income) etc. The result of classification of the first degree is the *distribution of frequencies* of individual variants of each variable (Tab. 6 – 3) while the classification of the second degree informs about frequencies of combinations of individual variants of two variables. The result of this classification is presented in the so-called contingency table (see Tabs. 6 – 1 and 6 – 2).

Independent (identification) variables (e.g. sex, age, education , domicile etc.) are usually presented in rows while variables under study, i.e. the holiday destination of respondents (see Tab. 6 – 10) are arranged in columns.

The total number of possible contingency tables of the second degree classification is relatively huge. It can be calculated as a half of the product $n \,.\, n\text{-}1$ where "n" is the number of variables under study. In practice, however, a substantially lower number of classified contingency tables can be sufficient. So for example it is not necessary to meddle with activities of the Czech Statistic Office and look for relationships existing between identification (usually demographic) variables, e.g. to investigate the relationship between the age and education.

Similarly, it is also not necessary to inquire into more general marketing problems within the framework of a simple market research performed among customers (e.g. which relationship exists between brand perception of a product and its quality), i.e. creation of contingency tables from investigated variables *per se*. This should be more likely done within the framework of marketing research performed at the macro level.

Example 18 Statistic relationships existing among variables under study

In a nation-wide research campaign performed at the turn of years 2005 – 2006, we used the technique of questionnaires and the total number of respondents (Czech households) was more than one thousand (exactly 1,070)

Among others, we also tried to find out which is the statistical relationship between and how is for people buying foodstuffs important *brand, package* and *advertisement.* It was found out that the relationship between the brand and the advertisement was weak (Kendall's tau coefficient was 0. 22) as well as between the brand and package (Kendall's tau = 0. 25). On the other hand, however, the relationship between the brand and the package was stronger (Kendall's tau = 0, 4). This means that for *households influenced pretty much by promotion and advertisements the package of purchased foodstuffs was also important*. On the other hand, however, households not influenced too much by promotion of foodstuffs the package of purchased products was not very important.

When performing simple surveys and investigations, it is not necessary to use too many contingency tables with pairs of identification and investigated variables. Their number may be relatively low and the possible maximum would be as product of multiplication *number of variables under study time's number of identification variables*. And even these need not be usually processed all. Namely, it is well-known the of identification variables the respondents mention more frequently age, education and sex than e.g. marital status

Even classifications of the 3rd and higher degrees are usually transformed into two-dimensional tables. So, for example, we could try to find out if the relationship existing between place of living and holiday is not influenced also by the sex, i.e. if the respondent is male or female (Tab. 6 – 4). In this case we would classify answers into two tables – one for men and the other for women. Thereafter we could use some of contingency coefficients and try to ascertain if and which difference exists between them. This method is also used to identify the so-called *spurious, apparent, false correlations*.

Example 19 A spurious statistical relationship

As an example of spurious, „false" statistic correlation, i.e. relationship existing between two variables (variables) we can mention the value of correlation coefficient expressing the relationship between the place of business (seat of other company) and the degree of satisfaction of customers. It is quite logical that there is no reason why the customers should be more satisfied with companies operating their business in the country than with those having place of business in big towns. A more probable reason of popularity would be more probably in their size and, above all, assortment of goods. We can also expect that small firms are in a closer contact with their customers and that they can respond to their (often very specific) wishes, needs, and requirements more flexibly than medium or big ones. There is no doubt that small family businesses are situated mostly in the country and small towns while medium and big (especially international and supranational) companies have their headquarters in big towns and urban areas. To reveal a spurious (false) relationship it would be therefore necessary to perform classification of the 3rd degree, i.e. to create two, formally identical tables containing pairs of variables: (i) place of business (independent variable, written in rows with variants "Town" *vs.* "Country ") and (ii) satisfaction of customers (dependent variable, written in columns with variants "Satisfied customers" *vs.* "Unsatisfied customers"). The first table would involve only the subset of small businesses while the other that of medium and big ones. A higher coefficient of correlation in the first table would therefore confirm the suspicious "false" statistical relationship between variables *place of business* and *satisfaction of customers*.

The classification thus enables to construct a contingency table. This table then help us to ascertain and measure statistic relationships between two variables.

Example 20 Example of a contingency table and method of calculation of contingency coefficients

The following exhibit contains a contingency table, in which the respondents are classified according to their domicile and holiday destinations.

Respondent's domicile	Holiday destination					n_i
	At home	CR	SR	Seashore	Other	
Znojmo – City	**33**	**13**	**26**	**13**	**3**	**88**
Znojmo – district	**87**	**34**	**24**	**11**	**4**	**160**
Other districts	**62**	**24**	**17**	**11**	**4**	**118**
n_j	**182**	**71**	**67**	**35**	**11**	**366**

Table 6 - 4 A contingency table of absolute frequencies

As one can see at the first sight, the respondents from Znojmo want (in contradistinction to those from Znojmo-district and/or other districts) to spent their holiday out of their homes, most frequently in Slovakia. This relationship is even more evident in following tables that contain relative frequencies (presented in rows or in columns).

Row relative frequencies indicate relative distribution (most frequently in percent) of frequencies of a certain variant (row) of one variable in all variants (columns) of the second classification variable of the contingency table (see Tab. 6 – 5). These row relative frequencies enable to compare distribution of frequencies in different variants (rows) of the variable under study and/or with regard to its overall distribution.

As compared with people from the Znojmo district and/or other districts, the Znojmo citizens want more frequently (60 %) spent their holiday out of their homes, most frequently in Slovakia (30 %); they also want to go to the seashore (15 %) or, at least, somewhere in the Czech Republic (Tab. 6 – 5). People living out of Znojmo answered that they intend to spent their holiday at home (more than half of them) or that they to visit some destinations in the Czech Republic (one fifth of them). Less than one quarter of these respondents plan to spent holiday abroad (in Slovakia or on the seashore).

Respondent's place of living	Where do you go on holiday?					n_i
	At home	CR	SR	Sea	Other	
Znojmo – City	**37 %**	**15 %**	**30 %**	**15 %**	**3 %**	**100 %**
Znojmo – District	**54 %**	**21 %**	**15 %**	**7 %**	**3 %**	**100 %**

Other districts	**53 %**	**20 %**	**14 %**	**9 %**	**4 %**	**100 %**
n_j	**50 %**	**19 %**	**18 %**	**10 %**	**3 %**	**100 %**

Table 6 - 5 A contingency table of relative row frequencies (percentages)

Analogically *column relative frequencies* express relative distribution (most frequently in percent) of frequencies of a certain variant (column) of one variable in all variants (rows) of the other classification trait of the contingency table (Tab. 6 – 6). These column relative frequencies enable to compare distribution of the variable under study of frequencies in different variants (columns) and/or with regard to its overall distribution.

Respondent's place of living	**Where do you go on holiday?**					**n_i**
	At home	**CR**	**SR**	**Sea**	**Other**	
Znojmo – City	**18 %**	**18 %**	**39 %**	**37 %**	**27 %**	**24 %**
Znojmo – District	**48 %**	**48 %**	**36 %**	**32 %**	**36 %**	**44 %**
Other districts	**34 %**	**34 %**	**25 %**	**31 %**	**37 %**	**32 %**
n_j	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**	**100 %**

Table 6 - 6 A contingency table of relative column frequencies (percentages)

Table 6 – 6 illustrates differences in intentions of respondents to spend their holiday in different destinations from a different point of view, i.e. with regard to their domicile. Stay at home or somewhere in the Czech Republic is preferred by people from the Znojmo district (and also from other districts while Znojmo citizens intend mostly to go abroad (to Slovakia and/or to the sea).

Total relative frequencies provide minimum information for an interpretation (Tab. 6 – 7). They express percentages of relative frequencies of a certain variant of one variable (column) simultaneously with a certain variant of the other (row) from the total number of all sampling units (i.e. within the whole sampling set). This means that they present the share of each individual box of the contingency table in the total sum of answers.

Respondent's place of living	**Where do you go on holiday?**					**n_i**
	At home	**CR**	**SR**	**Sea**	**Other**	
Znojmo – City	**9 %**	**3 %**	**7 %**	**4 %**	**1 %**	**24 %**
Znojmo – District	**24 %**	**9 %**	**7 %**	**3 %**	**1 %**	**44 %**
Other districts	**17 %**	**7 %**	**4 %**	**3 %**	**1 %**	**32 %**
n_j	**50 %**	**19 %**	**18 %**	**10 %**	**3 %**	**100 %**

Table 6 – 7 A contingency table of total frequencies (percentages)

So for example the first box informs us that of the total number of respondents decided to stay at home 9 % live in Znojmo but this information is not interesting for us. More likely we would like to know how want all Znojmo respondents to spent their holiday (see relative row frequencies in Tab. 6 – 5) or as differ domiciles of those who plan to stay at home see relative column frequencies in Tab. 6 – 6).

If we want express statistically the relationship between the domicile of respondents and their intentions to spent their holiday (Tabs 6 – 5 and 6 – 6) we must calculate some of coefficients of contingency because these variables are nominal. We shall use only two coefficients of them – Pearson's and Crammer's ones. In both cases we must know chi-squared (χ^2) contingency. When calculating it, the first step represents the value of theoretical (expected) frequencies, which should be calculated for each box of the contingency table using the following equation:

$$n'_{ij} = \frac{n_i \cdot n_j}{n}$$

where:

n´= theoretical (expected) frequency

Index i is row index i = 1 ……k (k = 3)

Index j is column index j = 1 ……l (l = 4)

Concretely, for the first box of the first box it is:

$$n'_{1,1} = \frac{88 \cdot 182}{366} = \underline{\underline{43{,}8}}$$

Similarly, for the third box of the second row it is:

$$n'_{2,3} = \frac{160 \cdot 67}{366} = \underline{\underline{29{,}3}}$$

etc.

To illustrate this in detail, an overall survey of theoretical (expected) frequencies is presented in Tab. 6 – 8.

Respondent's place of living	**Where do you go on holiday?**					**n_i**
	At home	**CR**	**SR**	**Sea**	**Other**	
Znojmo – City	**43,8**	**17,1**	**16,1**	**8,4**	**2,6**	**88**
Znojmo – District	**79,6**	**31,0**	**29,3**	**15,3**	**4,8**	**160**
Other districts	**58,7**	**22,9**	**21,6**	**11,3**	**3,5**	**118**
n_j	**182**	**71**	**67**	**35**	**11**	**366**

Table 6 - 8 A contingency table of calculated theoretical (expected) frequencies (percentages)

Similarly as the table of total relative frequencies (Tab. 6 – 7), also a contingency table containing theoretical (expected) frequencies is not very helpful for the reader. That's why it has no

sense to present it in the final report. It represents only an interstage when calculating the characteristic χ^2.

The more the real (empiric) frequencies are different from theoretical (expected) ones, the higher is the probability that the relationship between both variables will move from discontinuity to continuity. The aforementioned chi-squared (χ^2) contingency represents a characteristic, which informs us about these differences by means of only one number. We can calculate it using the following equation:

$$\chi^2 = \sum_{i=1}^{k}\sum_{j=1}^{l}\frac{\left(n_{ij} - n'_{ij}\right)^2}{n'_{ij}}$$

This formula means that we should calculate this fraction for each box of the contingency table and summarise then all calculated values (in this case all 15 boxes). The resulting value of χ^2 equals in this example 16.79.

As already mentioned in association with re-categorisation, the χ^2 test requires a sampling set of sufficient size. It cannot be used in cases when we expect frequencies lower than 5 in one fifth of boxes of contingency table or when we expect that in at least one box the expected frequency will be lower than 1. In such cases the calculated test criterion χ^2 is considered to be incorrect.

The squared χ^2 contingency enables to derive also other coefficients used for measuring of the strength of relationship existing between two nominal variables. The most popular are the following:

Pearson's coefficient of contingency

$$P = \sqrt{\frac{\chi^2}{\chi^2 + n}} = \sqrt{\frac{16{,}79}{382{,}79}} = \underline{\underline{0{,}21}}$$

and

Cramer's coefficient of contingency

$$C = \sqrt{\frac{\chi^2}{n \cdot \min(k-1; l-1)}} = \sqrt{\frac{16{,}79}{366 \cdot 2}} = \underline{\underline{0{,}15}}$$

Similarly as in case of Chuprov's coefficient, which is another contingency coefficients used for nominal variables derived from the characteristic χ^2, values of these two coefficients also range from 0 to 1.

Both aforementioned examples of calculation of contingency coefficients indicate that the relationship existing between the respondent's domicile and place of intended holiday is not too strong and that it can be more probably characterised as weak.

As far as the ordinal variables are concerned, these relationships are measured by means of *Spearman's rand correlation coefficient.* Its values range within the interval from –1 to + 1 and it is calculated as follows:

$$R = 1 - \frac{6 \cdot \sum d^2}{n(n^2 - 1)}$$

where

d = the difference between compared pairs of rank values and

n = number of compared pairs.

If the statistical relationship (dependence) is measured between a *pair of variables of different level* (e.g. one of them is nominal and the other ordinal or even cardinal) there is no other possibility than to select the coefficient with regard to the level of the lower one (i.e. of nominal). Analogically, in case that one of variables would be ordinal and the other cardinal, it is not possible to use the correlation coefficient but only one of coefficients used for ordinal variables.

Example 21 Calculation of Spearman's coefficient

Numbers of citizens visiting meetings of municipal boards and percentages of people satisfied with activities of municipal offices were surveyed in altogether ten municipalities (Tab. 6 – 9):

Municipality	Number of inhabitants	Ranking	Number of satisfied people	Ranking
A	850	6	40.9	6
B	924	4	46.7	3
C	1012	3	34.6	9
D	578	9	42.4	5
E	773	8	42.8	4
F	1149	1	40.2	7
G	851	5	47.0	2
H	794	7	36.6	8
CH	1078	2	49.2	1
I	400	10	31.6	10

Table 6 - 9 Frequencies and ranks of two ordinal variables

$$R = 1 - \frac{6 \cdot 116}{10 \cdot 99} = 1 - \frac{696}{990} = 1 - 0.703 = 0.297$$

As one can see, the statistical relationship (correlation) between numbers of citizens visiting the meetings of municipal board meetings and those satisfied with activities of the local municipal office is very low.

For ordinal variables we can use Kendall's tau (τ) coefficient or Somers' delta, the values of which also range within the interval from – 1 to + 1.

Similarly, also *correlation coefficients measure the dependence of two cardinal variables* within the range from – 1 to +1

This indicates that, as compared with coefficients of contingency, the measuring of relationships between two variables by means of Spearman's, Kendall's, Somers' or correlation coefficients is at a higher informative level, among others also due to the fact that these coefficients not only determine the strength but also the direction of the relationship, i.e. if the measured relationship is proportional directly or indirectly. As mentioned above, ordinal and above all cardinal variables measure values of investigated properties more exactly and, besides, it is also obvious that their coefficients enable to define statistical relationships and dependences at a higher level.

Example 22 The using of Kendall's coefficient

The nation-wide inquiry was performed on the turn of years 2005/2006 using a survey method of personal interviews. The questionnaire was distributed among and answered by altogether 1,070 households in the whole territory of the Czech Republic. When trying to evaluate the strength of relationships existing between individual pairs of variables we decided to use the Kendall coefficient tau because of it simplicity and practical and pedagogical advantages. The coefficient was used to measure the degree of correspondence between two rankings and assessing the significance of this correspondence. Factors influencing the buying behaviour and decision-making of Czech households were analysed by these statistical coefficients and all values were significant at the level of 95 % and more

The dependence of the effect of ten factors on the net annual income of households is presented in Tab. 6 -.10 . The rank of these factors was calculated using Kendall coefficient tau.

Factor	**Kendall coefficient tau**
Habit	- 0. 01
Product parameters	0. 1
Price	- 0. 15
Quality	0. 1
Brand	0. 11
Price-off offers	- 0. 12
Package	0. 06
Promotion	0. 05
Recommendation of other people	0. 02
Effort to try an innovation	0. 1

Table 6 - 10 Effect of ten factors on food purchase as dependent on the household income

As shown above, the *relationship between the net annual income and effects of the aforementioned ten factors on the decision-making and buying behaviour of households was statistically non-significant*. As far as the relatively strongest factors were concerned (such as price and price-off offer), their effects were naturally indirectly proportional. This means that, in general, for households with higher annual incomes the price and price-off offers are less important when buying foodstuffs than for those with lower annual incomes.

Benchmark recommendations concerning interpretation of calculated absolute values of individual coefficients are as follows: if they are lower than 0.3, the statistical relationship existing between two variables under study is weak (or even none); if they fluctuate within the interval of 0.3 to 0.6 the ascertained relationship is worth of interest and if it is higher than 0.6 we can be very happy that we have revealed a strong statistical relationship.

As we already know, these coefficients measure *statistic strength of the relationship between two variables,* and, possibly, also its direction (direct, indirect). However, it is also important to calculate the statistical *significance* of this relationship. In case of contingency coefficients, their significance may be tested by χ^2 while the significance of Spearman's coefficient by t-test. In any case, however, we must decide at first about the level of significance, which indicates the probability that the statistical relationship arose quite randomly, i.e. that there is no statistical relationship between both variables (phenomena). In case of socio-economic phenomena, the significance level is usually five percent ($p = 0.05$) or maximally one percent ($p = 0.01$). If the value of calculated empiric coefficient is higher or (at least) equal to the value of test criterion at the selected level of significance, then we accept the ascertained relationship as significant at this level. If not, it is evaluated as non-significant.

In our example $\chi^2 = 16.79$ at the statistically significant level of at least 5 % because the table value for a five-percent level of significance is lower than the calculated one (15.51 *vs.* 16.79). This means that the *zero hypothesis of independence of holiday destination and respondent's domicile* is rejected. However, low values of both coefficients indicate that this relationship is not too strong but more probably it is rather weak.

The aforementioned coefficients express the relationship between two variables so that we can speak about a two-dimensional analysis. However, in case of cardinal and possibly also ordinal variables we can use also *multidimensional analysis,* which measures simultaneously relationships existing among several variables. The most popular of them are factor analysis, cluster analysis, discrimination analysis, regression analysis and many others. Above a combination of factor and cluster analysis represents a very suitable means of market (customer) segmentation.

Discussion questions:

1. What it is a variable?
2. Which is the difference between identification variable and followed variable?
3. Which are the properties of nominal, ordinal and cardinal variables?
4. What determents is the variable will be nominal, ordinal, or even cardinal (measurable)?
5. What it is categorisation?
6. What it is and why we perform pre-categorisation?
7. What it is coding?
8. What it is classification?
9. What it is classification of the first and the second degree?
10. What is performed at first – categorisation or coding?
11. Which kinds of relative frequencies in the contingency table do you know and is their meaning?
12. Which are correlation coefficients for nominal and ordinal variables, which are their values and what they express?

13. How can we test statistical significance of a calculate coefficient?
14. Which is the principle of multidimensional analysis and which are its best known methods?
15. How would you formulate the question concerning the entrance fee into a cultural monument that it would be evaluated as a) nominal, b) ordinal, and c) cardinal variable?
16. Which is the *methodological principle* of marketing research?
17. How would you formulate a polytomic nominal (verbal) variable?
18. Which coefficient would you use to evaluate the relationship between:
    a) two verbal (nominal) variables,
    b) two ordinal variables,
    c) two cardinal (measurable) variables?
19. Which are the values of each of them?
20. Which is the relationship between the number of questions and the number of variables (variables)?
21. The are the following there contingency tables:

a)

Number of children / **Household income**	**Low**	**Median**	**High**	**Total**
High	100	100	100	300
Median	100	100	100	300
Low	100	100	100	300
Total	300	300	300	**900**

b)

Number of children / **Household income**	**Low**	**Median**	**High**	**Total**
High	300	0	0	300
Median	0	300	0	300
Low	0	0	300	300
Total	300	300	300	**900**

c)

Number of children / Household income	Low	Median	High	Total
High	0	0	300	300
Median	0	300	0	300
Low	300	0	0	300
Total	300	300	300	**900**

21.1. Define values of correlation, Spearman's, Pearson's, and Cramer's coefficients in all three tables.

21.2. How would you interpret distribution frequencies and values of coefficients presented in these three contingency tables (including practical recommendation for the Ministry of Labour and Social Affairs of the Czech Republic?

21.3. Evaluate suitability (correctness) of application of each of the aforementioned coefficients with regard to both variables and decide which of them would be in this case the most suitable from the statistical point of view.

7 INTERPRETATION AND THE FINAL REPORT

Chapter objectives: Using a simple example, the author describes the whole process of marketing research, i.e. from the specification and operationalisation of a marketing problem, through collection and processing of data to the interpretation of statistical results and their final presentation (including practical marketing measures).

Key words: marketing problem and its operationalisation, sample representativeness, classification, contingency tables, statistical relationships among variables, relative frequencies, statistical coefficients, interpretation, practical recommendations.

The aim of a simple marketing research was to *find out which business organisation and/or company was considered by Znojmo citizens as the most important from the viewpoint of the town development (including its positive and negative aspects).* We were also interested what task of Znojmo Municipal Council was the most important for town's citizens. Thereafter we also tried to learn if the respondents think that the municipality think is fulfilling this task. This question was answered by means of a five-grade scale. Identification questions concerned sex of respondents, their age and the highest finished type of education.

The same inquiry was performed already in the year 2010 so that it was possible to simplify considerably the process of operationalisation of the research subject. As already mentioned in Chapter 2, this process is usually the *most* principal, *the most demanding and the most important step within the framework of the introductory (preparatory) stage of the marketing research*. In this case, however, it was necessary to find out in which direction the opinions of Znojmo citizens did change at most and which of them remained unchanged. This means that we tried to compared answers obtained in 2010 with the newer ones obtained in 2011. For that reason we used a practically unchanged questionnaire. This questionnaire consisted of five predominantly opened questions and of three identification questions.

The final questionnaire form is presented in Fig. 7–1. It contained only five predominantly opened questions and three identification questions.

Interview about the most important businesses in Znojmo

1. Which company is in your opinion the most important for a further development of Znojmo?

……

2. Which are the most important positive features of this company for Znojmo citizens?

……

……

3. And which are its major negatives for Znojmo citizens?

……

……

4. Which are the most important tasks of Znojmo Municipal Council as far as the further development of town is concerned?

……

……

5. And does the Znojmo Municipal Council fulfil this task?

Definitely yes	More or less yes	More or less no	Definitely no	Don't know
1	2	2	4	5

Sex

Male	1
Female	2

The highest acquired education

Basic, skilled worker, incomplete secondary school	1
Secondary school	2
University	3

Age

18 – 39	1
40 – 59	2
>60	3

Fig. 7 – 1: Final questionnaire

The data collection took place in Znojmo from 9th to 16th March 2011. Interviews were performed in different town districts and the total number of respondents was 578 persons. Respondents were selected on the base of quota sampling. The *representative sample* is characterised below.

Sex	Basic set	Quota sampling
Male	48 %	50 %
Female	52 %	50 %

The highest acquired education	Basic set	Quota sampling
Basic, skilled worker, incomplete secondary school	60 %	59 %
Secondary school	30 %	30 %
University	10 %	11 %

Age	Basic set	Quota sampling
18 – 39	42 %	41 %
40 – 59	36 %	39 %
>60	22 %	20 %

As mentioned in Chapter 5, conclusions of the aforementioned comparison of relative frequencies in both columns is one of the most important conclusions of the final interpretation of this research. Thanks to this it can be said that the obtained results are valid not only for the whole set of 578 selected respondents but that they can be generalised for the whole basic set of Znojmo citizens in the age of 18 years and more. The main results of the statistical processing of this survey are presented below.

Results

within the framework of this marketing research, Znojmo citizens were asked, *which company is, in their opinion, the most important for the further development of their town.*

As shown in Tab. 7–1, from the viewpoint of citizens, the most important was the company PEGAS NONWOVENS, s.r.o., which was mentioned 129-times.

Rank	**Company**	**Frequency**
1.	PEGAS NONWOVENS, s.r.o.	129
2.	ZNOVÍN ZNOJMO, a.s.	82
3.	FREEPORT LEISURE (CZECH REPUBLIC) s.r.o.	66
4.	Znojmo Public Transport Company – Psota, s.r.o.	33
5.	TOS ZNOJMO	26
6.	EGSTON SYSTEM ELECTRONIC, s. r.o.	21
7.	Private College of Economic Studies in Znojmo	15
8.	A.S.A. EKO Znojmo, s.r.o.	12
9.	ELEKTROKOV	11
10.	KAUFLAND	10
11.	Not mentioned	37
12.	Others	136
Total		**578**

Table 7 – 1: The most important companies in Znojmo

Comparison of results obtained in years 2010 and 2011

The company PEGAS NONWOVENS, s.r.o.has been prospering already for a long time and offers many working opportunities (jobs).

The firm Znovín Znojmo, a.s. occupied the second place both in 2010 and 2011. This company contributes to the development of tourism and excellently promotes the wine growing region of Znojmo.

The Znojmo Transport Company – Psota s.r.o. occupied the "virtual" third place because it supplied good-quality transport of passengers and offered a wide assortment of traffic channels and operated new, modern buses.

In 2011, the fourth and the fifth place were occupied by TOS Znojmo and Freeport Leisure (Czech Republic) s.r.o. companies. These two firms were mentioned above all due to the fact that they contributed to the regional development through a permanent offer of jobs. Besides, the company Freeport Leisure (Czech Republic) s.r.o. supported the development of local tourism not only because of an offer of services in the domain of shopping but also in spending leisure time, entertainment and body care; these services are used above all by our Austrian neighbours.

As compared with 2010, also Private College of Economic Studies in Znojmo, Elektorkov, Kaufland were mentioned among ten most important subject contributing to the development of Znojmo in 2011.

Statistical relationships

Using the following contingency tables that were the result of classification (grouping) of the 2nd degree of a triplet of identification questions we were able to find out *statistical relationships. This possibility represents the fundamental (or the most important) contribution of classification of the 2nd degree and/or of contingency tables.* However, the decisive first information contains already the classification of the 1st degree in the form of cumulated total relative frequencies of the most frequently mentioned Znojmo companies. It is quite obvious that Pegas was the most frequently mentioned company (22 % of answers), followed by Znovín (14 %) and Freeport (11 %). Psota (6 %), TOS (4 %) and other companies were mentioned less frequently.

Education	Variant answers						Total
	Pegas	Znovín	Freeport	Psota	TOS	Others	
Basic	18 %	16 %	11 %	5 %	5 %	44 %	59 %
Secondary	27 %	10 %	12 %	7 %	3 %	41 %	30 %
University	28 %	18 %	5 %	5 %	5 %	39 %	11 %
Total	22 %	14 %	11 %	6 %	4 %	43 %	100 %

Tab. 7 – 2: Differences in the evaluation of Znojmo companies by respondents with different finished education

This contingency table indicates that also differences in answers of respondents with different highest finished education are relatively small; this is corroborated also by a low value of Pearson's coefficient of contingency (0.15).

Similarly, the contingency Table 7 – 3 was created when analysing the correlation with the age of respondents.

Age	Variant answers						Total
	Pegas	Znovín	Freeport	Psota	TOS	Others	
18–39	22 %	14 %	16 %	4 %	3 %	41 %	41 %
40–59	25 %	16 %	10 %	7 %	3 %	39 %	39 %
>60	18 %	11 %	4 %	7 %	8 %	52 %	20 %
Total	22 %	14 %	11 %	6 %	4 %	43 %	100 %

Tab. 7 – 3: Differences in evaluation of Znojmo companies by respondents of different age

This contingency table indicates that differences in answers of respondents of different age were not too great. Young people mentioned most frequently Freeport, the middle-age category Pegas, and the oldest one TOS (or some other companies. these relatively small differences were corroborated also by a low value of Pearson's coefficient of contingency (0.20).

The preceding three tables indicate that, as far as the further development of the town was concerned, Pegas was evaluated as the most important company in Znojmo. This opinion expressed 22 % of the total number of respondents. Positive answers were recorded among males, people with secondary and university education, and in the middle-age category. Znovín occupied the second position (14 % of respondents) and was followed by Freeport (11 %), Psota (6 %), and TOS (4 %). Remaining 43 % of respondents mentioned some other or even none company.

Similar results were obtained also in case of the fourth question that concerned *the most important task of Municipality Council in the future development of the town.* A survey of received answers is presented in Tab. 7 – 4 (including differences in answers of both sexes). The most frequent answers were the following: "Elimination of town's debts" (21 %), "Reduction of unemployment" (9 %), "Construction of a bypass road" (7 %) and "Reduction of corruption" (6 %).

Sex	Variant answers					Total
	Eliminate town's debts	Reduction of unemployment	Reduction of corruption	Construction of a bypass road	Others	
Male	20 %	13 %	7 %	7 %	53 %	50 %
Female	22 %	5 %	5 %	7 %	61 %	50 %
Total	21 %	9 %	6 %	7 %	57 %	100 %

Tab. 7 – 4: Differences in opinions of respondents of both sexes concerning the most important tasks of Znojmo Municipal Council

The contingency table indicates that differences between answers of males and females were relatively small and the value of Pearson's coefficient of contingency was low as well (0.14).

The following contingency table presents differences in opinions of individual groups of respondents with different levels of finished education (Tab. 7 – 5).

Education	Variant answers					Total
	Eliminate town's debts	Decrease unemployment	Reduce corruption	Construct a bypass road	Others	
Basic	21 %	10 %	6 %	6 %	57 %	59 %
Secondary	25 %	7 %	6 %	8 %	55 %	30 %
University	13 %	10 %	8 %	8 %	61 %	11 %
Total	21 %	9 %	6 %	7 %	57 %	100 %

Tab. 7 –5: Differences in opinions of respondents with different finished education about the most important tasks of Znojmo Municipal Council

As one can see, differences in answers of individual categories of respondents with different levels of finished education were also very small. The value of Pearson's coefficient of contingency was really low (0.10).

Similar differences were observed also among answers of respondents of individual age categories (Tab.7 –6).

Age	Variant answers					Total
	Eliminate town's debts	Decrease unemployment	Reduce corruption	Construct a bypass road	Others	
18 – 39	22 %	10 %	8 %	6 %	54 %	41 %
40 – 59	23 %	8 %	4 %	7 %	58 %	39 %
>60	16 %	10 %	6 %	8 %	60 %	20 %
Total	21 %	9 %	6 %	7 %	57 %	100 %

Tab. 7 – 6: Differences in opinions of respondents of different age concerning the most important tasks of Znojmo Municipal Council

Also in this case it was found out that differences in answers of respondents from individual age groups were relatively small because the value of Pearson's coefficient was also very low (0.10).

Next three contingency tables clearly indicate that in the opinion of the major part of respondents (21 %) the most important task of Znojmo Municipal Council is the elimination of town's indebtedness; 9 % of citizens answered that the Municipal Council should reduce the unemployment rate, in the opinion of 7 % of them the construction of a bypass road was very important and 6 % answered that the Municipal Council should reduce corruption.

When answering the last question, respondents should express their opinion *how much the activities of Znojmo Municipal Council coincide with their expectations and opinions.* As one can see in the following contingency table, the respondents are not satisfied too much with activities of the Municipal Council because more than a half of them (55 %) answered that the Municipal Council does not fulfil this task, one third (32 %) is of the opposite opinion and remaining 13 % were not able to answer YES or NO (Tab. 7 – 7).

Sex	Variant answers					Total
	Definitely YES	More or less YES	More or less NO	Definitely NO	Don't know	
Male	6 %	31 %	36 %	18 %	9 %	50 %
Female	3 %	23 %	37 %	20 %	17 %	50 %
Total	4 %	28 %	36 %	19 %	13 %	100 %

Tab. 7 – 7: Differences in evaluation of respondents of both sexes concerning the fulfilment of tasks by Znojmo Municipal Council

Differences in answers of males and females were again very small and this was corroborated also by a low value of Pearson's coefficient of contingency (0.14). In this case, more critical were above all females.

The following contingency table 7 – 8 shows even smaller differences in opinions of people with different finished education.

Education	Variant answers					Total
	Definitely YES	More or less YES	More or less NO	Definitely NO	Don't know	
Basic	3 %	28 %	38 %	16 %	15 %	59 %
Secondary	5 %	26 %	35 %	22 %	12 %	30 %
University	7 %	28 %	31 %	25 %	9 %	11 %
Total	4 %	28 %	36 %	19 %	13 %	100 %

Tab. 7 – 8: Differences in evaluation of Znojmo Municipal Council by respondents with different finished education

Once more we can see that there were relatively small differences in answers of respondents with a defferent education. This was corroborated also by a low value of Spearman's coefficient (0.11).

The greatest differences in answers concerning activities of Znojmo Municipal Council were recorded in individual age categories of respondents (Tab.7 – 9).

Age	Variant answers					Total
	Definitely YES	More or less YES	More or less NO	Definitely NO	Don't know	
18 – 39	6 %	25 %	33 %	21 %	15 %	41 %
40 – 59	3 %	25 %	40 %	18 %	14 %	39 %
>60	4 %	36 %	37 %	16 %	7 %	20 %
Total	5 %	27 %	36 %	19 %	13 %	100 %

Tab. 7 – 9: Differences in evaluation of Znojmo Municipal Council by respondents of different age

This contingency table shows that differences in answers of respondent of different age were relatively small (the value of Spearman's coefficient was only 0.15). However, it is obvious, that above all respondents of the middle-age category criticised activities of Znojmo Municipal Council at most.

Positive and negative features of the most important companies in Znojmo

The following survey presents results of processing of answers to two opened questions considering positive and negative features of five most important companies in Znojmo.

Pegas Nonwovens, s.r.o.

Positive features	
Job opportunities	97 x
Prosperity and development	17 x
Foreign trade	9 x
Good wages	8 x
Representation of Znojmo	7 x
Sponsoring	4 x
Geographical situation	4 x
Presence in stock market	3 x
Payment of taxes to the municipality	3 x
Others	5 x

Nearly all respondents answere that in their opinion "Job opportunities" were the most important positive feature of the Pegas company. "Prosperity and development" was evaluated as another positive feature of this company.

Negative features	
Don't know, nothing	58 x
Low wages	13 x
Environmental pollution	12 x
Occupation of arable land	7 x
Out of town center (commutation)	7 x
Lack of job opportunities	7 x
Non-stop operation	6 x
Noise, trucks	4 x
Low contribution for sport and cultural organisations	4 x
A Luxembourg firm	3 x
Permanent expansion	2 x
Corruption	2 x
Others	5 x

As far as the negative features of the Pegas company were concerned, "Low wages" occupied the first place and were followed by answers concerning "Environmental pollution". The third and fourth place were occupied by answers concerning "Occupation of arable land" and "Non-stop operation".
Nevertheless, in this case, positive features markedly predominated over negative ones.

Znovín Znojmo, a.s.

Positive features	
Good representation of Znojmo	27 x
Tourism development	17 x

Quality wine	16 x
Job opportunities	8 x
Development of viticulture	8 x
Others	2 x

The highest number of respondents answered that in their opinion a very good representation of the town of Znojma was the most important positive feature of Znovín company. Very positively were evaluated also activities supporting the development of local tourism movement. Other positives involved production of quality wines, job opportunities and care about the development of viticulture.

Negative features	
None	41 x
Bad accessibility	10 x
Higher alcohol consumption	6 x
Decreasing quality of wine	5 x
Lack of job opportunities	5 x
Others	11 x

As far as negatives were concerned, the respondents most frequently answered that they did not know about any negative feature of the Znovín company. As negatives, "Bad accessibility", "Higher alcohol consumption", "Decreasing quality of wine" and "Lack of job opportunities" were mentioned by nearly one third of respondents.

Freeport Leisure (Czech Republic), s.r.o.

Positive features	
Job opportunities	24 x
Attractive for tourists	14 x
Branded clothing	11 x
Good accessibility	8 x
Shopping centre	6 x
Low prices	4 x
Wider assortment	4 x
Leisure	3 x
Modern industry	1 x
Discounts/sales	1 x
Culture	1 x
Bargain cheap purchasing	1 x

Altogether 24 respondents answered that the most important positive feature of this company is the offer of job opportunities. On the other hand, the distance between the seat of the company and the town centre was mentioned most frequently (21 answers) as a negative feature.

Negative features	
Distance from the town centre	21 x
Efflux of customers	17 x
None	7 x
Don't know	3 x
High prices	2 x
Influx of foreigners	2 x
Competition	2 x
High attendance	2 x
Small size	1 x

Bad transportation services	1 x
Low cultural self-realisation	1 x
Bad advertising	1 x
Vietnamese market – bad image	1 x
Outlet	1 x
Low numbers of customers	1 x

Znojemská dopravní společnost – Psota, s.r.o.

Positive features	
Public transport	16 x
Good connections	10 x
The largest transport company in Znojmo	2 x
New buses	2 x
Discount fares	2 x
Ecological transport	1 x
Reliability	1 x
Don't know	1 x

In this case, a good "Public transport" was mentioned as the most important positive feature while the "Reliability" was evaluated as the most negative one.

Negative features	
Don't know	9 x
Non-ecological transport	7 x
High fares	4 x
Intensive traffic	3 x
Changes in time tables	3 x
Inadequate traffic channels	2 x
Not enough traffic means	2 x
Bad working conditions for employees	2 x

TOS Znojmo, a.s.

Positive features	
Employment	19 x
Extended company	5 x
Ligh engineering firm	3 x
Locality	2 x
Qualification of employees	1 x

In this case, offer of jobs was evaluated predominantly as a positive feature.

Negative features	
Don't know	10 x
Low wages	7 x
Non-ecological	4 x
High competition	2 x
Lack of jobs	1 x
Dismissal	1 x

Summary and practical recommendations

This marketing survey revealed that according to Znojmo citizens that PEGAS NONWOVENS, s.r.o. was the most important company in town. This opinion expressed 22 % of respondents. The company ZNOVÍN ZNOJMO, a.s. occupied the second place with 14 % of positive answers, the third most important company was FREEPORT LEISURE (CZECH REPUBLIC) s.r.o. with 11 %, the fourth one was the transport company PSOTA, s.

r. o. with 6 % and the company TOS ZNOJMO was mentioned in the fifth place (4 %). Remaining 43 % of respondents mentioned either other companies or did not answer at all.

Women preferred the company FREEPORT LEISURE but in other cases the differences were not too obvious. It is interesting that the company PEGAS NONWOVENS was more appreciated by respondent with university education and also by people of middle-age category (40 - 59 years).

Basing on obtained answers it can be recommended that companies TOS ZNOJMO, a.s. a PEGAS NONWOVENS, s.r.o. should pay attention to low wages of their employees and that they also should pay greater attention to their negative environmental effects. On the other hand, ZNOVÍN companies ZNOJMO, a. s. and FREEPORT LEISURE, s.r.o. should try to improve their accessibility for Znojmo citizens. Last but not least, the company PSOTA, s.r.o. should buy more ecological transportation means and decrease fares. Other companies (that were not mentioned among the five most frequently mentioned ones) should pay an increased attention to advertisements and promotion so that they could penetrate into the awareness of Znojmo citizens.

It is also interesting that 57 % of respondents were not able to answer the question concerning the most important tasks of the municipal council. Remaining 43 % of respondents mentioned most frequently the eliminate the town's indebtedness. This measure was preferred by secondary school and university graduates and also by respondents of middle-age category (40 - 59 years).

The question "Which are the most important tasks of Znojmo Municipal Council as far as the further development of town is concerned?" was negatively answered by 55 % of respondents; 32 % of them were of the opinion that this task is being fulfilled and 13 % answered that they do not know. From viewpoints of age category and of finished education the opinions of respondents were in principle identical, more marked differences were found out in the category of sex because 37 % and 26 % males and females, respectively, answered this question positively. On the other hand, 57 % of women gave a negative answer to this question.

It can be therefore concluded that Znojmo Municipal Council should pay greater attention to debt elimination, for example through selling plots owned by the town to people wanting to construct their family houses or by means of leasing of both commercial and housing premises. This idea should be published on town websites, on billboards situated in Znojmo and its neighbourhood and in regional broadcasting.

Discussion questions:

1. Why it was possible to omit the process of operationalisation of the subject of marketing research in the above example (although it is otherwise of key importance)?
2. Basin on what is it possible to say that the selected set of 578 respondents is really representative as far as this example is concerned?
3. What fundamental results from this conclusion for the subsequent interpretation and above all for a practical application of obtained results and conclusions?
4. Where can one see the classification of the 1st degree in contingency tables?
5. Which information results from this classification?
6. How and from what is a contingency table constructed?
7. Which information the contingency table presents, what can we learn from it?

8. By means of which fundamental calculations can we obtain the first orientation idea?
9. By means of what can we exactly measure and express it?
10. In tables 7 – 2 to 7 – 8 the statistical relationships were expressed by means of Pearson's coefficient of contingency. Why it was replaced by Spearman's coefficient in tables 7 – 9 and 7 – 10?
11. On what is the selection of the most suitable coefficient dependent when measuring statistic relationships in contingency tables?

CONCLUSIONS

The foregoing six chapters represent a really brief introduction into problems of marketing research. Especially the last Chapters 6 and 7, which deal with statistical processing of marketing data, represents a mere outline of this complex problem. However, as already mentioned in Chapter 2 – Marketing research as a process, the statistic processing of obtained data does not mean that whole process of marketing research is ended. Quite on the contrary, even the best and the most careful statistical analysis provides only a base for final interpretation and presentation of obtained results. These results should provide answers to current problems and present suggestions of concrete, practical marketing solutions of issues creating the subject of marketing research. However, these tasks cannot be solved only on the base of current knowledge of and experiences in the field of marketing research. To be able to do it, we must have a good knowledge not only of marketing *per se* and of its possibilities but also of managerial and entrepreneurial aspects of problems requiring solution.

This textbook tries to point out some already tested and objectivised methods, procedures, and tools of marketing research. However, possibilities of their application (when studying customers and trying to learn more about them and their behaviour) are much wider. Very important are our experiences and our capability to formulate new questions and to solve new problems. Methods of marketing research should become an integral part of our day-to-day practice in the field of collecting and evaluating data concerning the market and our customers. It is necessary to point out that these activities must be *systematic* and *regular*. However, in practice unfortunately we can often see that top managers of both commercial firms and non-profit organisations became to be interested in marketing research issues as lately as it is too late – i.e. in situations when the customers lost the interest in their offer and prefer their competitors. If we want to avoid such situations as much as possible, we should apply marketing research methods and tool from the very beginning of our managerial activities, i.e. already in the stage of product development and before its introduction into the market. Moreover, a systematic and regular use of methods of marketing research would be positively reflected also in a better formulation of new tasks as well as when exploiting the obtained empiric data in processes of managerial and entrepreneurial decision-making, i.e. when trying to improve our working methods and to increase the fruitfulness of work of managers and businessmen.

REFERENCES

BLALOCK, H. M., Jr. *Causal Inferences in Nonexperimental Research*. The University of North Carolina Press, Chapel Hill, 2nd Printing, 1964.

BLALOCK, H. M., Jr. *Social Statistics*. McGraw – Hill Boook Company, New York, 2nd Edition, 1972.

McDANIEL,C. - GATES, R. *Marketing Research: The Impact of the Internet.* 5th edition, South – Western, 2002

1986

TULL, D.S., - HAWKINS, D. I. *Marketing Research.* 5th edition, Macmillan Publishing Company, New York, 1990

BURNS, A. C. - BUSH, R. F. *Marketing research.* 6. vyd. Upper Saddle River, N.J.: Prentice Hall, 2010

MALHOTRA, N. K. *Marketing research : an applied orientation.* 5. vyd. Upper Saddle River, New Jersey: Pearson Prentice Hall, 2007

BRADLEY, N. *Marketing research : tools & techniques.* Oxford: Oxford University Press, 2007

STEVENS, R. E. a kol. *The marketing research guide.* 2. vyd. New York: Best Business Books, 2006