INFORMATION BENEFIT OF GRADUATES FOLLOW-UP SURVEYS

Zuzana Pacáková, Andrea Jindrová
Czech University of Life Sciences Prague
pacakova@pef.czu.cz

ARTICLE INFO

Article type
Full research paper

Article history
Received: August 22, 2011
Received in revised form: September 18, 2011
Accepted: September 29, 2011
Available on-line: September 30, 2011

Abstract

The article aims to evaluate the benefit of follow-up surveys which are a useful tool to get feedback from university graduates. Questionnaire surveys are of irreplaceable sources of information from students and graduates, but several rules have to be taken into account when conducting such survey. The questionnaire is then one of the most often used tools for data collection in such types of surveys. The article presents not only the essential methods and principles of data collection and analysis, but it also evaluates the surveys efficiency and their impact on the quality of the educational process.

Various methodological principles are presented within the context of large questionnaire survey of CULS graduates that was held in 2008. In connection with the survey, the process is presented from the very beginning (list of questions, data matrix) through the introduction of the most often used statistical methods up to the analysis of the survey profit.

Key Words

Graduates, follow-up survey, information efficiency, data analysis, statistical methods

Introduction

With growing amount of graduates, the universities have realized the necessity of relationship between the university and its graduates. There are several studies (Lambert et al., 2006; Brew, 2008; Liu et al., 2008; Rymešová, Kolman, 2010) presenting results of surveys among graduates, usually carried out in order to reflect the quality of studies and career development of students, generally to get a feedback from the graduates. A timely collection of feedback on the quality of studies from graduates is very important. However, it is difficult to collect the feedback, when the graduates are no more in contact with the university.

Students who have left the university are called alumni of the university. „Alumni are counted as pillars of universities’ greatness. They play many critical roles for the university – returning to teach and to learn, counseling graduating students, serving on advisory boards, providing financial resources for programs to execute the university’s mission“ (Střížová, 2009).

An alumni system is a social system connecting (not only) graduates having the same field of interest according to the universities’ programme structure. Alumni system includes both forward (to the alumni) and backward (to the university) communication, mostly supported by IT/ICT. To get the students’ feedback, various alumni systems, supported either by the university or published as a separate websites, were published.

Questionnaire surveys (follow-up survey) are an essential tool for getting graduates’ feedback. Follow-up surveys in education are those carried out after students have left the university by the administration of questionnaires to graduates and/or their employers (according to Liu et al., 2008). In practice, follow-up surveys are widely used. For example, Queensland University of Technology explored the employment outcomes of its students (Powell, Partridge, 2010). The survey aimed to find out where the graduates live, work and what type of position they hold. Efendiev and Balabanova (2010) investigated careers of graduates of the management faculty at the Higher School of Economics State University.

They wanted to determine the extent to which the graduates are in demand on the labour market and the degree to which the content and the level of their education are in keeping with labour market needs. They also outlined some problems often occurring with follow-up surveys. Firstly they mentioned the problem that the graduates are often not connected to the institution they have graduated from. Too small samples and not representative „qualitative“ research methods are then other often occurring limitations.

A large study of medical graduates in the United Kingdom was performed in 2002 (Lambert et al., 2006). It aimed to report the specialty choices of UK medical graduates and to compare the results with those from previous surveys. This is an example of a large study (2778 respondents) focused not on graduates from a particular university but on graduates with the same specialization.

The scheme 1 summarizes the main topics of interest of follow-up surveys and their information efficiency.
The questionnaire is a type of research technique that provides various types of information in accordance to its subsequent manipulation (data analysis). The data can either be in form of a text, qualitative or quantitative variables, based on questions formulation. For all that, the most common data are type of qualitative variables, resulting from questions with predefined or partly predefined answers.

Questionnaire surveys are the most common survey tools to be used for the data collection in various researches activities.

Questions formulation, their type and the variety of answers are the basic problems to be solved when constructing a questionnaire. An incorrectly presented questionnaire can discourage respondents from completing it, an unprofessional data analysis can lead to misinterpreted results and the aim of the study does not have to be fulfilled.

The next chapter aims to present techniques of statistical analysis, especially the analysis of qualitative variables. Valid results cannot be obtained on the basis of incorrect data, that’s why not only the data analysis techniques, but also the whole research process from planning to presentation of results, is introduced in the chapter.

**Material and Methods**

Properly prepared data are the basic for good-quality and correct results of the data analysis.

During the data preparation phase it is necessary to focus on: cleaning the data, quality assessment, missing values analysis or weighting the data.

The scheme 2 illustrates the follow-up survey process.
Planning

Planning is the first step of survey during which following problems have to be solved: aim of the survey, time schedule, funding and the target group. In case of using data stored in a database, the sample size and variables to be analyzed have to be stated. But if the data are to be collected through survey (e.g. questionnaire) the planning process has to encompass also a selection of proper data collection technique.

Data collection

The next step which is closely connected with planning is the data collection phase. The questions have to be formulated in accordance with the previously stated hypothesis and the aim of the study. The structure of the sample should be in accordance with the structure of population the sample is selected from, so the representativeness of the sample is another important aspect. Data collection can be realized using different principles and methods, e.g. personal interview, telephone survey, internet based questionnaire and others.

Data matrix preparation

After finishing the data collection phase, the next step is the data matrix preparation. In the following text basic procedures connected with the data matrix preparation are presented. The data collected within a questionnaire survey has to be transformed into a data format suitable for analysis by the statistical software. Some questions in the questionnaire are represented by one variable in the data matrix only, but some questions are represented by several variables. Let’s imagine a question “educational attainment” with possible answers: primary school – high school – university degree. While one answer only is to be selected, the question can be transformed into one variable with three possible values. But imagine the question “Which information sources do you use for your study?” with offer of answers: presentations - textbooks – internet – scientific books – other materials. In this case several answers can be selected. For this question more variables, in the concrete five, have to be created in the data matrix. Each variable can only reaches one of two values, e.g. 0 – not selected, 1 – selected. The use of simple codes (usually numbers) as it is shown in the last example (0; 1) is called coding. This step makes the data analysis using SW much easier. If the data are collected from different sources, the aggregation of the data has to be accomplished. Common statistical software offers both possibility to aggregate variables and units (cases).

The process of the quality adjustment and the data cleaning focuses on identification of extreme values and outliers. The one-dimensional approach to outliers identification consists in graphical analysis mainly. In case of multidimensional data, values that are extremely different in the viewpoint of one variable and also values that do not reflect the overall structure of the data have to be investigated. Johnson and Wichern (2007) propose following steps for outliers detection: graphical analysis of separate variables, graphical analysis of paired variables, examination of standardized values and squared distances.

Missing values analysis aims to identify and substitute missing values by the use of suitable method. The main sources of missing values are: non-response, unreliable or uncertain data. There are two basic approaches to handling missing data: elimination of the data (either object or variable) or substitution of missing values using imputation methods. For missing values analysis of categorical data, so called mining methods can be used (see e.g. Batista, Monard, 2003).
To balance the structure of the sample to make sure it reflects the structure of the population, weights should be estimated. While stratification weights are projected together with the design of the study to handle the unequal probabilities of selecting units, post-stratification weights adjust unbalances detected after the data are collected.

To meet the presumptions necessary for the use of selected statistical methods, transformation of variables is sometimes needed. Most often the transformation is made to ensure the normality of data. While the presumption of normal distribution refers to quantitative variables which are not as often collected through questionnaire surveys, detailed information is not presented in the article. For details it is possible to see e.g. Hebák et al. (2005) and Johnson, Wichern (2007).

Data analysis

The fourth phase of the follow-up survey process is data analysis. The procedures of statistical data analysis are mainly determined by the type of the data. Quantitative variables are numerical information often expressed on a scale (percentage, thousands crowns). Qualitative variables take verbal values (e.g. gender: female-male). There are two types of qualitative variables: values of ordinal variable can be sorted (educational attainment: primary school – high school – university degree), while the values of nominal variable cannot. By the use of questionnaire mostly qualitative data are collected. The basic principles to deal with qualitative data are based on frequency tables: one-dimensional, two-dimensional or multidimensional.

Univariate data analysis is based on description of separate variables using descriptive characteristics, graphic techniques or frequency tables. These characteristics provide basic information about the data. Measures of central tendency are measures that represent with a proper value the tendency of most data to gather around this value. In case of nominal variable a proper measure of central tendency is the mode $x^\sim$. For ordinal variable the tendency is often described using the median $x^\sim$. Median is the fiftieth quantile, it is the middle value of sorted data, it cuts the sorted data into two halves. For quantitative variables arithmetic mean $x$ is often a suitable characteristic to express the central tendency of values. When using the arithmetic mean one should consider the presence of extreme values, because the arithmetic mean is a measure which is sensitive to extremes.

Based on two qualitative variables the data can be sorted into a contingency table. Contingency table is then a basic setting for hypothesis testing and computation of measures of association. The basic statistical test for evaluation of relation (symmetric) between two qualitative variables is the chi-square test. The chi-square test can only be used if the presumptions about expected frequencies (frequencies expected under the null hypothesis about independency) are fulfilled. The expected frequencies have to take values 5 or more in at least 80% of cells and in other cells the values have to be equal at least 1. In case of not fulfilling the presumptions for the use of chi-square test, exact tests are to be used. Fisher’s exact test is used in case of 2x2 table. There are various measures of strength of association. Classical measures of relationship between two nominal variables, which are based on the chi-square statistic, are Pearson’s and Cramer’s coefficients. The coefficients take values between 0 and 1, or between -1 and 1, while zero indicates no relation between the variables.

In addition to the symmetric measures mentioned above, it is possible to construct asymmetric measures for evaluation of
one-sided dependency. Coefficient Sommers’s $d$ which takes values from interval $<-1,1>$ is an example of these asymmetric measures.

For selected tables of 2x2 dimension it is possible to describe the relationship between two factors by the use of odds ratio (OR) and relative risk (RR). The odds ratio ranges from 0 to infinity. When OR equals 1, there is no association between the row and the column variable. The odds ratio is a useful measure of association regardless of how the data are collected. The relative risk is the risk of developing a particular condition for one group compared to another group. (Stokes et. alter, 2000)

Example of the use and interpretation of OR and RR is given further in the article. For details about the computation see e. g. Stokes et. alter (2000) or Řezanková (2007).

For detailed analysis of differences of frequencies in the contingency table, sign scheme is a very helpful graphical tool. The scheme is based on the differences between observed and expected counts. If there is no significant difference between the observed and expected count, the appropriate cell in the table is labeled “0”, in case of statistically significant difference the cell is labeled “+” or “−” depending on the difference (the observed value is above or below the expected). For $\alpha=0.05$ one sign is used, for $\alpha=0.01$ two and in case of statistically significant difference at $\alpha=0.001$ three signs.

Another graphical method for representing associations in a table of frequencies or counts is correspondence analysis (see e. g. Johnson, Wichern, 2007 or Manly, 2005).

When considering one dependent variable only or when analysing the relationship between more than two variables advanced multivariate modelling techniques are to be used. Logistic regression is a suitable method when analysing the impact of both qualitative and quantitative variables on one qualitative variable - both alternative and multinominal (see Field, 2005).

**Presentation of the results**

The last step of the process is report and presentation of the results. The results presented have to be understandable, presented using suitable graphical outputs and the outputs should be exported so that they are readable by commonly used software tools.

**Results**

Follow-up survey held in 2008 at Czech University of Life Sciences Prague The data were collected through an internet-based questionnaire. The information about the survey was distributed by email, while there were two sources of the email addresses used. The first source included CULS graduates registered at public servers such as www.absolventi.cz, www.spoluzaci.cz or www.facebook.com. Second source of email addresses was the database of graduates of the Faculty of Economics and Management (http://www.pef.czu.cz/dotazniky). Based on the second source of addresses the rate of return was estimated, it was about 30%.

921 graduates participated in the study, most of the respondents were FEM graduates ($n=696$), then FAFNR (Faculty of Agrobiology, Food and Natural Resources) graduates ($n=98$), Faculty of Engineering ($n=73$). The rest ($n=36$) were students of other faculties and institutes of CULS. There were also twenty students who studied more than one faculty at CULS. In the study 54.7% of women and 45.3% of men participated. The mean age of respondents was 36 years (median 32 years).
While not all questionnaires were totally complete, all the analysis are based on the real number of answers. That is the reason why number of respondents for different questions can differ.

The data were analysed using the statistical software IBM SPSS, v. 17. Data were automatically saved into a database. The first step then was to identify the invalid questionnaires. We set down number of criteria to identify invalid questionnaires:

- IP address,
- more than twenty percent of the items incomplete,
- questionnaire without key information,
- misleading data.

First criterion that was checked was the IP address. In case of similar IP addresses, other aspects were further investigated (agreement in sex, date of birth, etc.).

Questionnaires with more than twenty percent of missing values were set aside. Also some basic logical structure among the answers was evaluated, e.g.:

- year of graduation >year of birth+22,
- length of study ≥3 (bachelor degree),
- etc.

Following these criterion 25 questionnaires were taken out.

To optimize the process of data preparation the data collection process was made in order to automate some procedures. The data matrix was schemed out along with the on-line questionnaire so that as many steps as possible were realized together. Following steps of data matrix preparation were realized along with the questionnaire preparation:

- projection of the questions into the variables,
- variables coding,
- variable types of measure,
- etc.

Though, there were still many steps of data preparation to be done. At first, invalid questionnaires were identified and taken out following the criteria mentioned above. The data were provided with value labels. Basic graphic and descriptive techniques were than used to clear the data. Each qualitative variable was firstly described separately using simple frequency table or bar graph which allows detection of unusual or senseless values.

It has already been mentioned that most of the follow-up surveys among the graduates aims at either:

- employment outcomes of graduates,
- correspondence of graduates knowledge and labour market needs,
- evaluation of study programmes.

The first and the second points are those probably studied most often. One of the indicators that are monitored in order to assess the quality of universities, is the ability to find a suitable position on the labour market. The unemployment among university graduates is generally not as high as it is among other employees. In the Czech Republic the unemployment rate among the graduates has decreased since 2004 (7.7%) to 2008 (2.5%). In 2009 (3.1%) there was a little increase (Úlovcová, H. et al., 2010).

Although the unemployment rate is not high, it is necessary to investigate the structure of graduates unemployed. Follow-up surveys bring information on position and specialization of graduates. Based on these information, the management of
the university can create new study programmes in accordance with the labour market demand.

The employability of graduates or let say the ability of university sector to provide graduates with such skills that employers need, was since 1990’s studied by many authors (e. g. Tomlinson, 2007; Mason et. alter, 2009).

In our survey a set of questions to this topic was oriented. The respondents were asked several questions on their present, but also previous work: specialization (work in field) - employee/employer/sole trader - management position - satisfaction with work and others.

In table 1, results on consistence between current specialization of graduates and their study programme at university are presented.

<table>
<thead>
<tr>
<th>Work in field</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>330</td>
<td>35.8</td>
</tr>
<tr>
<td>partly</td>
<td>253</td>
<td>27.5</td>
</tr>
<tr>
<td>other specialization, use knowledge</td>
<td>209</td>
<td>22.7</td>
</tr>
<tr>
<td>other specialization, do not use knowledge</td>
<td>129</td>
<td>14.0</td>
</tr>
<tr>
<td>total</td>
<td>921</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Table 1: Current specialization of graduates according to their study programme

Our analysis showed that almost 36% of graduates work in field they have studied. But 129 graduates (14%) do not use the knowledge they have got during university studies. This answer is more often by women (60.5%) than men (39.5%). We can also indicate that working without using the university knowledge is much more often by young graduates; more than 50% of them have maximum the second work after graduation. What is also interesting is the fact that almost 50% of those 129 graduates mentioned that the graduation does not have an impact on their salary.

Next results are connected with the graduate’s management position. The results are shown in table 2.

<table>
<thead>
<tr>
<th>Management position</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>top management</td>
<td>113</td>
<td>12.4</td>
</tr>
<tr>
<td>middle management</td>
<td>204</td>
<td>22.3</td>
</tr>
<tr>
<td>lower management</td>
<td>120</td>
<td>13.1</td>
</tr>
<tr>
<td>specialist</td>
<td>363</td>
<td>39.8</td>
</tr>
<tr>
<td>ordinary employee</td>
<td>113</td>
<td>2.4</td>
</tr>
<tr>
<td>total</td>
<td>913</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Table 2: Management position of graduates

Most of the graduates stated that they work as an independent specialist (39.8%), the second most frequent position is in the middle management (22.3%).

More detailed analysis confirmed that there are large differences between men and women in the consistence of their current specialization and their study programme. For the next analysis the original variable was transformed into two categories only (consistent – not consistent). Current specialization is consistent with the study programme much often by men than by women (p=0.006; n=921).

In table 3, there are odds ratio and relative risk describing the relation between sex and work in field.
When studying the success of CULS graduates on the labour market, we were also interested in a fact whether the later work in field is affected by students’ interest of the university and its study programme. That is whether the student is interested in the field of study/work already before studies. The analysis is based on a set of questions on motives to study at the CULS Prague. One of the questions was interest in the field/study programme. For the purpose of this analysis the variable was recoded into two categories: yes, I was interested – no, I was not interested. There is significant relation between the interest in field and later employment in the field (p<0.0001; n=921).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>2.13</td>
<td>1.61</td>
</tr>
<tr>
<td>RR</td>
<td>1.35</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Table 5: OR a RR – work in field in relation to interest in the field before studies

As it is obvious from table 5, students who were strongly interested in the field already before starting the studies have 2.13 times higher chance of working in the field.

Czech University of Life Sciences Prague consists of five faculties and two institutes. Further analysis focuses on different chances of graduates from the faculties on the labour market. It was found that there is significant difference between the students from various faculties and their ability to find work in consistence with their study programme (p<0.0001; n=896). For example when comparing two largest faculties in the study, FEM (Faculty of Economics and Management) graduates are often employed in consistence with their study programme, FFFWW (Faculty of Forestry and Wood Sciences) graduates assert in different fields also.
The graduates were also asked how many jobs they passed after the graduation. When concerning only those graduates who are more than five but maximum ten years after graduation, most of the students (more than 40%) answered that the job they have is still the first after graduation. More than 80% had maximum two previous job experiences. It can be pointed out that young people maximum 10 years after they had graduated, have very small work fluctuation. Tomlinson (2007) realized a qualitative study upon 53 final-year undergraduates to examine the way students, making the transition from higher education into the labour market, construct, understand and begin to manage their employability. His outcomes are somewhat different. His work pointed out that students realized the necessity of changing jobs. "The evidence of this study suggests that fewer students are anticipating their careers to be played out within the secure confines of single jobs and organisations which would form the basis of their long-term career progression. Movement between jobs and organisations was seen by some students as an inevitable reality." (Tomlinson, 2007, pp 288).

The second part of our analysis is connected with the student’s evaluation of quality of study at the university – availability of study materials, possibility to study abroad and other aspects. While in our study also students who graduated before 1990 were participating, our results provided us to make a retrospective study.

Set of questions was connected with the evaluation of availability of study materials including availability of IT. Availability of selected study supports was evaluated on a one-to-five scale (1-excellent; 5-unsufficient). Although lot of Czech and foreign literature is available today, no significant difference was found in evaluation of study materials availability and quality in relation to the graduation period (p=0.266; resp. p=0.286). We can then assume that materials available for various periods were sufficient for students. Unfortunately, lot of present students are not very interested in other literature than common textbooks, so they do not see the availability of literature (especially foreign) helpful.

On the other hand there are wide differences in evaluation of availability of information technologies (p<0.0001).

The differences in evaluation in relation to the graduation period are to be obviously seen from the table 6. The results are presented using the sign scheme.

<table>
<thead>
<tr>
<th>Graduation period</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>until 1990</td>
<td>---</td>
<td>---</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>1991-2000</td>
<td>---</td>
<td>0</td>
<td>++</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>after 2000</td>
<td>+++</td>
<td>++</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

Table 6: Sign scheme – evaluation of availability of IT in relation to the graduation period

The results that today’s graduates evaluation of IT availability is better, is not to very surprising (especially due to overall (un)availability of IT before 1990), but when studying the results more deeply, we can see that there are significant differences between the two groups of graduates after 1990 also.

Another result that seems to be interested at the present time, is connected with the evaluation of student team and its cohesion. No difference was found in relation to the graduation period (p=0.802). While most of present students work regularly during their studies, it was expected that they would evaluate the student team worse than their older colleagues. However, the evaluation of student collective is very positive across all
graduates. Especial campus of CULS Prague and intensive use of various modern communication tools such as email, Facebook, Twitter and so on could be the reasons.

Next set of questions was aimed on student’s evaluation of benefit of their study on CULS Prague. Various statements on this topic were evaluated on a scale: agree – rather agree – rather disagree – disagree. There is no difference in evaluation of graduates from different periods in case of development of communication skills (p=0.583) or ability to speak foreign language (p=0.065).

Very positive is the improving evaluation of team work and presentation skills. Today’s students mention that university studies have improved their ability to work in team much more often than it was by older graduates (p<0.0001). They also have better presentation skills (p=0.005).

The changes within the graduation period are depicted in figure 1.

![Figure 1: Level of agreement with the statement “During the studies at CULS Prague I have learned to present and argue for the results of my work in relation to the graduation period”](image)

Discussion

The follow-up survey data enable detailed analysis of graduate’s position on the labour market. It is possible to find out a proportion of graduates that work in the field of interest, what is their position in the management structure and it is also possible to identify factors that influence graduate’s ability to find an appropriate work. Our results turned out, that FEM students can easier find work in the field of interest than students of some
other more specialized faculties. The largest differences were than between FEM graduates and FFWS graduates. Rymešová and Kolman (2010) who studied the differences between the FEM graduates and FAFNR (Faculty of Agrobiology, Food and Natural Resources) graduates found that in both groups of graduates the percent of those who work in the field reaches almost 70%.

Our analysis presented a new trend also: younger graduates work in the field of interest they have studied more often than those who had graduated earlier (in 2000 and earlier).

The interest in specialization and study programme is one of the key factors of later employment of graduates. Students who were really interested in the field of study are often employed in the top management. Younger graduates can easily find work in the field than graduates a decade ago. This trend is in consistence with the active approach of today’s companies looking for later employments already at universities.

What seems to be surprising with students’ evaluation of different aspects of study is the fact that no difference was found in evaluation of availability and quality of study materials among younger and older students. We have expected younger students to appreciate possibilities to use foreign literature or online databases providing foreign studies and articles.

The authors have also expected that there are differences in evaluation of student team and its cohesion. While most of today’s are employed and do not visit the campus regularly, we expected them not to evaluate the student team much positive. However, both younger and older graduates evaluate the collective of student very positive. We suppose that today the face-to-face communication is widely substituted by various modern communication tools such as Facebook. Also another result that shows positive evaluation of team work and improvement of team work abilities during studies is in agreement with the finding of “good team”.

**Conclusion**

Modern technologies provide various useful tools to organize an extensive follow-up survey.

Manipulation and analysis of large datasets is easily realizable by the use of statistical programmes also. However, the quality of study and its results is primarily defined by the quality of the whole survey process starting with a properly stated aim of the study, followed by the quality of questionnaire and the data itself. No good and valid results can be obtained upon incorrectly prepared data. No valuable data can be obtained without proper aim and hypothesis stated. To get valid information on the basis of follow-up survey all phases of the process have to be implemented with attention.

**Acknowledgements**

The authors gratefully acknowledge the support from the Faculty of Economic and Management, Czech University of Life Sciences, via IGA grant, No. 11170/1312/3142 – „Questionnaire survey on employability of CULS graduates“.

**References**


