Graduates' Profile Survey Methodological Guidelines

561656 Voyage Project

Graduates' Profile Survey: Methodological Guidelines

The present guidelines are mainly based on the experience of the 561656-project coordinator AlmaLaurea in developing graduates' surveys in Italy. The guidelines want to provide some insights on the methodology to run the graduates' profile survey, from the selection of the cohort to the presentation of the results, as a guide for developing similar surveys in third countries.

Contents

I. What is the Graduates' Profile? The Graduates' Profile is a survey which looks at the features of Italian graduates in a certain year. The survey Report contains an analysis of the features and performance of graduates in that specific year, focusing on a number of variables. These include: age at graduation; examination grades; final degree grade; degree completion time; parents' educational qualifications; social background; secondary school qualification type and grade; class attendance; study experiences abroad; use of laboratories; traineeships or internships; time devoted to final dissertation; assessment of the university experience; knowledge of foreign languages; IT skills; work experiences during studies; intention to pursue further studies; preferred professional field; desired job characteristics.

Since 1998, when the Graduates' Profile was first launched, every year AlmaLaurea has been issuing a Report on graduates who achieved their title in the Consortium's universities. The aim is that of:

- outlining the main features of Italian graduates;
- providing a point of reference for those who have an interest in Italian universities;
- contributing to disseminate the culture of assessments within education systems;
- providing a guidance tool for the choice of university courses.

II. Who can benefit from the survey? The Graduates' Profile is addressed to all those who are interested in the Italian tertiary education system, including young people and their families, people in charge with guidance activities, companies (public/private, Italian/foreign), the Ministry of Education, University and Research as well as individual universities (in particular, the governing bodies of universities, faculties and assessment units).

III. Which cohort does the survey examine? The survey looks at the cohort of graduates who achieved their title in the Consortium's universities in a certain year. In 1998, when the Graduates' Profile was first launched, it involved a limited number of universities (13). It has then significantly expanded and, in 2014, it included 64 universities.

Sources. The content of the database is gathered from different sources and at different times. It includes:

- *information from administrative sources* – this is provided by universities and has to do with faculties, degree courses, degree grades, length of degree courses, final thesis titles and supervising professors, etc.;
- *information provided by the students just before their graduation* – this includes self-certifications on study experiences abroad, work experiences during studies, etc., as well as self-assessments on the knowledge of foreign languages and IT skills, the willingness to move for work reasons, etc.;

The combination of administrative information and the students' questionnaires results in a large database on Italian graduates, upon which the Graduates' Profile survey is based.

IV. How to access the data.

IV. How to access the data. Over the years, the survey data has become increasingly visible and easy to access. In addition to a printed volume, which is presented every year in May, at the annual AlmaLaurea Graduates' Profile Congress, the survey data is always available online (www.almalaurea.it/universita/profilo).

While in 1998 online data could only be broken down by university and faculty, it is now possible to refine a search by a number of variables, including degree course type, university, faculty, degree course, as well as by comparison with previous years.

Clearly, these variables are referred to the Italian university system and reflect the organization of tertiary education in Italy – also taking into account the University Reform implemented as of academic year 2000/2001. However, to be able to export AlmaLaurea's experience abroad, such variables need to be adjusted to the specific context.

Some useful advice. To compile a correct Graduates' Profile, a comprehensive and reliable database is needed, based on accurate information. To this purpose, the following actions are required:

- collaborating with university administrative offices, constantly following their work and motivating their staff, explaining the aims and benefits of the Graduates' Profile;
- ensuring adequate data matching (between administrative data and the questionnaires) to create a comprehensive database;
- constantly monitoring the share of students who fill in the questionnaires, trying to make sure that most of them do.

V. Exclusion criteria

While defining the cohort which will be analysed in the Graduates' Profile, AlmaLaurea applies several exclusion criteria, based on the Italian experience (when testing this kind of survey in another country, these criteria need to be adjusted to the specific university system features). These have been elaborated on the basis of the experience acquired in Italy by AlmaLaurea researchers over many years – and are therefore the result of great efforts spent on this project. The criteria used by AlmaLaurea in Italy are listed below, to serve as a guideline. However, for international partners who wish to test the Graduates' Profile survey for the first time, it is recommended that they use simpler criteria and adapt them to their country's university context. Exclusion criteria are applied in the following order: first, to graduates; second, to questionnaires; third, to universities.

GRADUATES – Exclusion Criteria

Particularly between 2001 and 2005, Italian universities included a few peculiar cohorts of students. Some of them had their professional experience recognized for the purpose of their degree course completion and often did not attend their faculty, or barely knew it at all. Therefore, we had to ponder which graduates should be included in the Graduates' Profile. It was decided that for Graduates to be included in the survey, at least half of the credits required for their degree completion had to have been earned through educational activities organized by their university. Indeed, if this was not the case, they would not be able to provide a reliable and meaningful assessment of the university experience. Sometimes, graduates belonging to peculiar cohorts cannot be immediately recognized on the basis of the available administrative data. Significant efforts were therefore made to analyse and establish which operating parameters should be used to effectively recognize them.

Thus, the Italian survey leaves out some 'peculiar' cohorts of students, that is students whose work experience was recognized by the university as their main learning activity and earned them at least half of their university credits. This is the case of special conventions applied to health operators, members of the police or the armed forces, public officials and other professionals. Present in different universities, these graduates rarely fill in the AlmaLaurea survey questionnaire.

Once the graduates who are eligible for the survey have been selected, further exclusion criteria are applied to questionnaires.

QUESTIONNAIRS – Exclusion Criteria

Within the Graduates' Profile survey, questionnaires can be excluded on the basis of three main criteria. One has to do with *completeness*, another with the *plausibility of answers* and the third concerns the *time for completion*.

As far as the *completeness of answers* is concerned, two parameters are applied.

1) If a graduate has provided information for at least 3 out of the 6 aspects of the questionnaire which are regarded as essential for the purpose of the Graduates' Profile (namely, *parents' educational qualifications*, *class attendance*, *study experiences abroad*, *work experiences during studies*, *overall assessment of the university experience*, *intention to pursue further studies*), then the first completeness parameter is satisfied.

2) For each section of the questionnaire, excluding the personal data section, the ratio between the number of answers provided and the number of answers required is calculated. Then, the average of these ratios is determined. If the average is ≥ 25%, the questionnaire fulfils the second completeness parameter.

If all these parameters are satisfied, then the plausibility of a questionnaire should be *Plausibility criteria are applied to check whether a questionnaire is 'dubious' or contains inconsistent answers.* The Italian questionnaire contains 8 sets of questions (*knowledge of foreign languages*, *IT skills*, *relationship with the academic staff*, *student services and living conditions in university towns*, *preferred business areas*, *relevant aspects for the desired job*, *preferred geographic areas where to work*, *preferred types of work contracts*). A questionnaire is considered dubious when there is little variability among answers (the same answer has been provided for the whole set of questions), as the graduate could have answered without even reading the questions (see Fig. 1).

Fig. 1. Example of a suspicious set of answers

Are you interested in working in the following business areas? (please answer all questions)

Example of "same modality of answer"

Administration, accounting	Yes, definitely	More yes than no	More no than yes	No, definitely
Technical assistance	X			
Management	X			
Finance	X			
Legal	X			
Marketing, communication, public relations	X			
Organization, planning	X			
Human resources, recruitment, training	X			
Production	X			
Research and Development	X			
Information systems, EDP	X			
Purchases	X			
Logistics, distribution	X			
Secretarial, general affairs	X			
Marketing, sales	X			

If less than five sets of answers show poor variability, the questionnaire fulfils the first plausibility criterion.

Another plausibility parameter concerns possible inconsistencies between the answers to similar questions which may be located in different sections of the questionnaire. In the Italian questionnaire, 4 different cases have been established:

- inconsistency concerning the study experience;

- inconsistency concerning work;
- inconsistency concerning the knowledge of foreign languages;
- inconsistency concerning IT skills.

A consistency check is a simple comparison between the answers provided by a graduate to two similar questions. If these answers are inconsistent, it means there has been a misunderstanding or a lack of attention while filling in the questionnaire. If less than 2 inconsistencies are found in a questionnaire, this fulfils the second plausibility parameter. On the contrary, if many inconsistencies are found, the reliability of the answers is questioned.

Finally, when all criteria for completeness and plausibility are fulfilled, the most complex of criteria is applied. This concerns *the average time spent on each answer*.

Thanks to its many years of experience, AlmaLaurea has conducted a careful analysis and empirically proved that the questionnaire cannot be correctly compiled if the average time spent on each question is less than 4 seconds.

To conclude, if a questionnaire fulfils the three sets of parameters (completeness, plausibility and time spent on the questionnaire), this can be used for the purpose of the Graduates' Profile survey and is included in the analysis.

UNIVERSITIES – Exclusion Criteria

Provided that it has successfully gathered administrative data for all its graduates, a university is included in the Graduates' Profile survey if its rate of filled-in questionnaires is high. This rate is not affected by graduates who are not compatible with the survey, on the basis of the exclusion criteria. The rate of filled-in questionnaires is the ratio between the number of questionnaires which can be used for the survey and the number of graduates. If the rate is 50% or more, the university is included in the survey. Otherwise, it will not be considered. This last criterion is a significant one in the case of the Italian survey, which involves almost all Italian universities with very high rates of filled-in questionnaires (above 90%). However, it is not very meaningful in pilot surveys which involve few universities.

VI. Cleaning the database and assessing representativeness

Have all or almost all questionnaires been collected? Among Italian universities which are part of AlmaLaurea, for example, not all the faculties have always managed to collect questionnaires from all their graduates – or at least a good percentage of them. Sometimes, whole universities or single faculties have not succeeded to do so. To tackle this issue, a specific division of AlmaLaurea constantly monitors the rate

of filled-in questionnaires for each university and interacts with administrative offices to spot – and possibly solve – problems which can be triggered by different causes.

In general, participation in the survey increases when, within a faculty, students are aware that the questionnaire is a tool to assess their university experience, but it also works as a CV which can help them find a job. In particular, graduates who already have a job or do not intend to immediately look for one are also required to fill in the questionnaire, in order to provide their feedback on the degree course they have just completed.

Three times in a year – at the end of each graduation session – AlmaLaurea gets in contact with all the graduates who have not yet filled in the questionnaire, inviting them to do so (except for those graduates who have explicitly stated that they do not wish to be part of the database). In some universities, this has led to a significant increase in the number of collected questionnaires.

Has the questionnaire been entirely filled in? In addition to the issue of graduates not filling in the questionnaire, some of them may also do so only partially or in an unreliable manner. Sometimes, some sections are not completed, with no specified reason. All these situations should be monitored. In general, these questionnaires are not considered in the survey, but they are still taken into account in the proportion of results.

Checking that the questionnaire is clear and consistent. Checking that the questionnaire is clear and consistent is the most difficult task when assessing a questionnaire. Sometimes, interviewees do not follow the rules. It may happen, for example, that they select two answers when they are required to only choose one (an error in the questionnaire programming), or that they answer a question when they should not – because they disregard the relevant filter question – or that they select the 'Other' option without adding any information in the relevant space.

Has the interviewee taken the questionnaire seriously? Explaining how to assess whether a questionnaire has been filled out in a serious manner is not easy. For example, when questions are to be answered with a score between 1 and 4, it can be assumed that the lack of an obviously schematic set of answers (the same answer for the whole batch of questions) suggests seriousness. In addition, it may happen that some graduates deliberately provide nonsensical answers. It is therefore recommended to have a general idea of the features which may suggest that the questionnaire was not taken seriously. In any case, only during the analysis of data is it possible to systematically assess this.

Statistical procedures. Having checked the questionnaires and sorted the cases of <u>missing answers</u> (for example, when an interviewee does not reply to specific questions) as well as the <u>missing cases</u> (for

example, when data is unavailable) and having decided whether to include them in the survey or not, the analysis of data can begin. At this stage, a statistical software should be employed (SPSS, SAS, Stata or others). To begin with, simple statistical operations should be carried out to perform a descriptive analysis of the main outcomes (frequencies and charts with percentages, based on the examined variables).

Correlation tables are useful to spot a possible relationship between two variables – an independent variable and a dependent variable.

Clearly, different types of analysis can be performed depending on the kind of variables under consideration. As for quantitative variables, summary measures can be used, such as mean, median and mode – which are very intuitive and facilitate the interpreting of data. Similarly, distance measurements – such as standard deviation or variance – can be used to analyse some variables. As for qualitative variables, charts can certainly be a first useful tool to sum up the main features; then, contingency tables – or cross tabulations – can be employed to display the frequencies related to two categorical variables at once (absolute values).

For a more in-depth study, following a first descriptive analysis of the main features, a multivariate statistical analysis may be performed. It could be used, for example, to determine the influence of some variables on others and the correlations between different variables. Multivariate analyses need to be interpreted carefully and with the help of specialists.

VII. Graphical display of the results

Representing and disseminating the results of statistical analyses in an effective manner is another crucial step in the process. Indeed, statistical accuracy and a clean database are not sufficient to guarantee a good analysis of data. The main results must be summed up in charts and tables which can make them clear for the general public, and not just for specialists.

An incomplete or unclear table or a poor communication of the results can compromise the whole survey.

It is therefore important to spend a few minutes looking at the features of good graphical representations. Here, the attention will be focused on the presentation of the results of descriptive analyses (univariate, bivariate or multivariate analyses).

Results can be described using tables or charts. What is important is that the table/chart and its title are self-explicative. In other words, they need to include all the elements required to interpret the data, in order for the reader to understand without having to rely on the accompanying text. Both for charts and tables, the title is crucial and should communicate:

- the subject of the analysis;
- the examined variables;

- the type of statistical analysis performed.

Displaying results in a table. A well-designed table should be:

- *efficient* (it should cost as little as possible)

- *comprehensive* (it should contain all the information required to correctly understand the data)

- *appropriate* (it should be consistent with the aims of the survey)

In other words, each table should fulfil the so-called *principle of parsimony* – only information which is necessary to understand the table itself should be displayed, as an excess of figures may confuse the reader. Compared to charts, tables allow for a larger quantity of information to be displayed.

Displaying results in a chart. Graphical representations are very efficient tools to help readers understand the data. Usually, charts contain little, condensed information and do not add any details, compared to tables. However, charts certainly have a strong communicative impact and, again, they are easy to understand for the general public, even with little statistical skills.

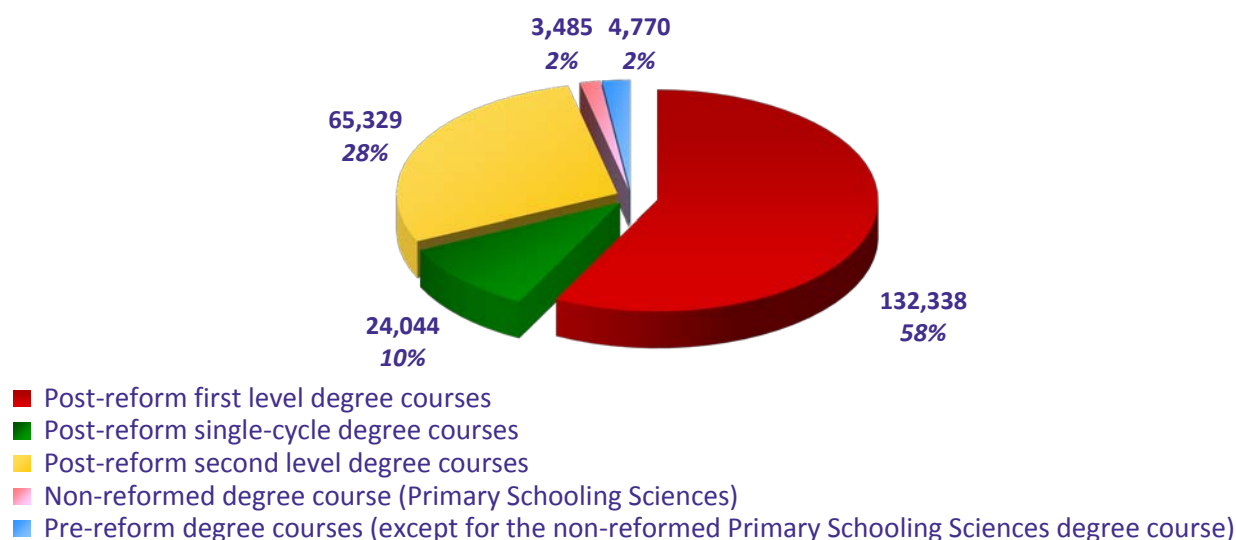To draw a chart correctly, the following aspects should be taken into account.

- The scale of the axes, as any variation of the scale would considerably change the reader's visual perception, minimizing or augmenting the differences of value reported in the chart.

- Ordering the bars on the basis of a certain answer – if differences between male and female interviewees are to be highlighted, for example, the chart bars can be ordered on the basis of one type of answer, as this helps visualize the main differences between the two cohorts.

- Choosing colours to represent different answers within each variable. A certain visual impact is important when displaying scale variables. For example, when portraying a feedback on a specific service or the overall assessment of the university experience, a scale of colours should be used, ranging from warm red shades to cold green shades (red shades for negative feedbacks, orange/yellow for intermediate results and green shades for positive feedbacks). This also applies to cases where the possible answers have a natural bottom-to-top order – for example, about parents' educational qualifications. Here, one colour may be employed in its different shades – for example, gradually fading from dark blue to light blue – so as to reflect the natural order of the possible answers.

- Choosing the chart title, which should include – in a few lines – the subject of the analysis, the examined variables and the type of statistical analysis performed.

An important feature of charts is that they allow to select and highlight the key (and also most important) information reported in a table. Since a chart cannot contain all the values reported in a table, the most important data should always be selected, in order to produce a simple and immediate graphical

representation. Univariate, bivariate or multivariate analyses have different graphical representations. Since they are very easy to read, charts are used to highlight the main features of the population analyzed. *Pie charts* are used to provide an overview of the situation (see fig. 2).

Fig. 2 – Graduates by course type

- ■ Post-reform first level degree courses
- ■ Post-reform single-cycle degree courses
- ■ Post-reform second level degree courses
- ■ Non-reformed degree course (Primary Schooling Sciences)
- ■ Pre-reform degree courses (except for the non-reformed Primary Schooling Sciences degree course)

Simple bar charts are used to represent frequency distributions in absolute values or percentages (fig. 3).

Fig. 3 – Percentage of graduates who gained study experiences abroad through EU programmes, by degree subject area

Grouped or two-sided bar charts are often used to compare how the same variable is distributed in different populations. Fig. 4 is an example of a *grouped bar chart* displaying the distribution of graduates

who benefited from student support services, divided on the basis of their university's geographical location.

Fig. 4 – Percentage of graduates who benefited from student support services, by university's geographical location
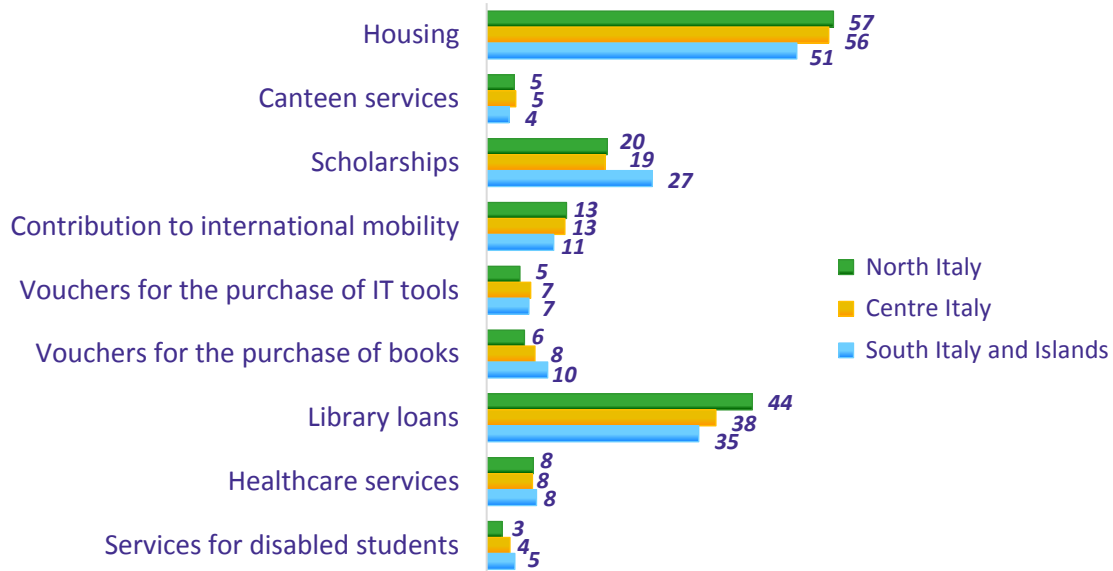

Two-sided bar charts (see fig. 5) can display, for example, the different distribution of the population with tertiary education among age groups, in different countries. Here, this kind of chart is effective in that it immediately displays the differences between two different cohorts. The bars can be set in an increasing/decreasing order based on one of the two aspects (in this case, data was displayed in decreasing order in relation with the 25-34 years of age group). This is an effective way to highlight the differences between the two cohorts under consideration.

Fig. 5 – Population with tertiary education, by age group (percentage values)
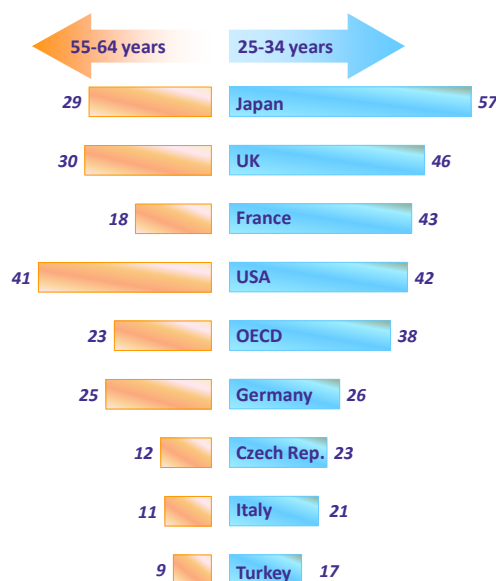
Broken line graphs are used to display the development of a phenomenon over time (see fig. 6).

Fig. 6 –Age to graduation – components (average values)
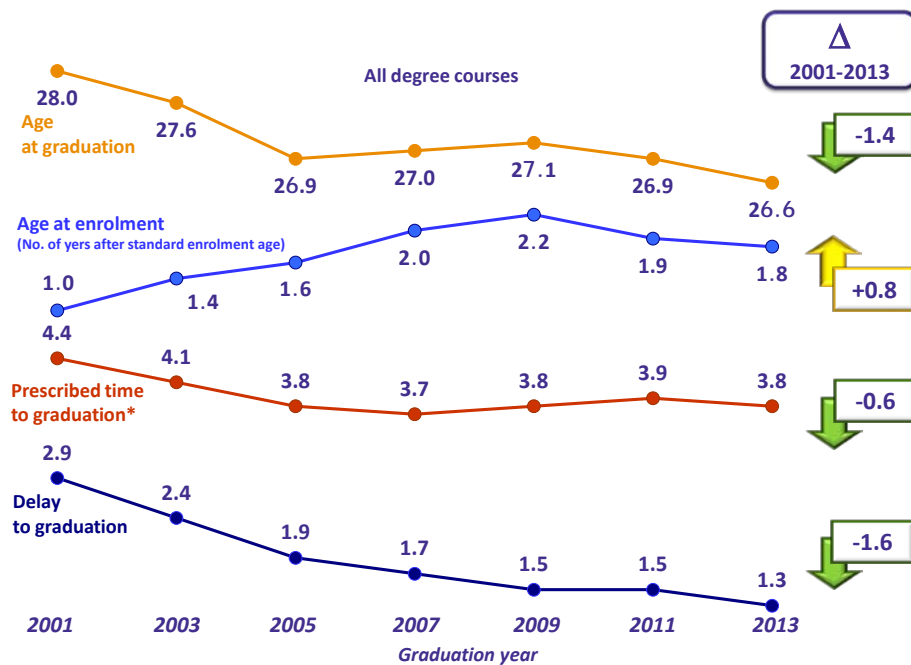
* 5 instead of 2 years for single-cycle degree courses.