Catalogue no. 37200001
ISBN 978-0-660-28785-0

Technical Reference Guides for the Education and Labour Market Longitudinal Platform (ELMLP)

Labour market outcomes for college and university graduates, 2010 to 2014

Release date: December 4, 2018

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at STATCAN.infostats-infostats.STATCAN@canada.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Table of contents

1. Introduction

Statistics Canada, in collaboration with the provinces and territories, Employment and Social Development Canada (ESDC), and other stakeholders, has developed the Education and Labour Market Longitudinal Platform (ELMLP).

The ELMLP allows longitudinal integration of administrative data related to education with other data sources to provide customized datasets for analytical purposes.

The ELMLP Program fills data gaps and enables a greater understanding of student and apprenticeship pathways, transitions to the labour market and outcomes over time.

Data from the ELMLP can help address a wide range of policy questions pertaining to student and apprenticeship persistence, completion, mobility and pathways as well as their labour market outcomes.

These data will allow policy makers to understand the different types of trajectories that students can take through their postsecondary education or apprenticeship training as well as student characteristics that may be related to these trajectories.

The target audience for the ELMLP includes provincial ministries of education, apprenticeship authorities, postsecondary institutions, federal government departments, students and parents interested in graduate outcomes and other stakeholder groups involved in education and the labour market.

2. The Education and Labour Market Longitudinal Platform (ELMLP)

2.1 The key features of the ELMLP

1. **Platform** – The Platform allows researchers to unlock information about past cohorts of college/university students and registered apprentices, to better understand their pathways and how their education and training affected their career prospects.
2. **Securely integrated datasets** – These integrated datasets allow us to know more than what a single dataset or survey can provide. Integrated datasets means that all of the datasets in the ELMLP may be linked with each other using an anonymous linkage identifier located on each file. After identifying which ELMLP datasets are needed to answer a specific research or policy question, researchers can then use the anonymous linkage identifier located on each file to bring these datasets together. The integration of datasets is carried out within the Statistics Canada Social Data Linkage Environment (SDLE), which maintains the highest privacy and data security standards.
3. **Longitudinal data** – The data available within the Platform are linked longitudinally, allowing researchers to better understand the behaviours and outcomes of students and apprentices over time.
4. **Accessible data** – All datasets prepared for the Platform will be made available to researchers through the Research Data Centres network across Canada.

2.2 Accessibility, confidentiality and privacy

The integrated datasets in the ELMLP are deemed sensitive statistical information and subject to the confidentiality requirements of the *Statistics Act*. Statistics Canada employees who build the integrated datasets for research purposes have access to the data only after it has been stripped of personal identifiers. Furthermore, only Statistics Canada employees and deemed employees who have an approved need to access the data for their analytical work are allowed access to the linked analytical files.

These data will be treated with the same level of confidentiality as surveys administered by Statistics Canada.

Findings from the ELMLP will be released through Statistics Canada's website.

The ELMLP data will also be available in Statistics Canada's Research Data Centres (RDC) to researchers with approved projects only. These researchers will be provided with access in a secure setting at the RDCs, which are staffed by Statistics Canada employees. The RDCs are operated under the provisions of the *Statistics Act* in

accordance with all confidentiality rules, and are accessible to researchers once they have been sworn in under the *Statistics Act* as "deemed employees."

2.3 Core and supplementary datasets

The ELMLP consists of two types of datasets: **core** and **supplementary**.

Core datasets will be updated in the ELMLP on an annual basis and include

1. **The Postsecondary Student Information System (PSIS)** - a data holding of all public college and university enrolments and graduates by type of program and credential, and field of study for each school year. The ELMLP consists of PSIS data from 2009 onwards for all provinces and territories, as well as from 2005 onwards for the Maritime Provinces.
2. **The Registered Apprenticeship Information System (RAIS**) - an administrative dataset of pan-Canadian (provincial and territorial) annual data on registered apprentices and trade qualifiers. The ELMLP consists of RAIS data from 2008 onward.
3. **Income tax** - select information from income tax data from 2004 onwards is available for all PSIS and RAIS records that were linked to the income tax data.

Supplementary datasets are additional datasets that will be integrated into the platform in order to add new indicators for research purposes. Over time, the number of supplementary datasets that are brought into the ELMLP will grow and could include survey data already collected by Statistics Canada, administrative data already obtained by Statistics Canada, and administrative data not yet available at Statistics Canada.

3. Data sources

3.1 Data sources and record linkages

Different administrative data files are linked through a record linkage process to create the ELMLP analytical dataset. The data used for calculation of the graduate outcome indicators are derived primarily from the Postsecondary Student Information System (PSIS) and selected tax variables from the T1 Family File (T1FF).

Postsecondary Student Information System (PSIS) data

The PSIS is a national annual administrative database that enables Statistics Canada to provide detailed information on enrolments and graduates of Canadian public postsecondary institutions in order to meet policy and planning needs in the field of postsecondary education. PSIS collects information pertaining to the programs and courses offered at an institution, as well as information regarding the students themselves and the program(s) and course(s) in which they were registered, or from which they have graduated.

Reference period: The start date for the PSIS report is the day after the end of the institution's previous winter term, which is usually a date in April, May or June. The reference period is one year from this start date.

PSIS data is currently available in the ELMLP-integrated files in the Research Data Centres (RDCs) as follows:

- All provinces, reference years 2009/2010 to 2015/2016.
- Territories, partial data for 2010/2011, complete data for 2011/2012 to 2015/2016
- Three Maritime Provinces, additional data for universities for 2005/2006 to 2008/2009.

PSIS data for the reference years 2009/2010 to 2014/2015 and T1FF data for 2011 to 2016 have been used for the December 4, 2018 release of the indicator tables.

Please note: Information about the students must be available from all sources to be used in the graduate outcome indicators calculations. PSIS data that have been imputed[1] or graduate records for which no tax data were available were not used.

Tax Family File (T1FF)

The T1FF is a database that combines variables from the T1 and T4 tax files and the Canada Child Tax Benefit into a family composition file. It includes income, demographic and geographic variables for each tax filer and their spouse, family and children.

Reference period: Calendar year (January to December).

Selected variables from the T1FF database are currently available in the RDCs with keys for matching to other ELMLP datasets for the tax years 2004 to 2015. Newer years of data will be added when they become available.

4. Methodology to derive the graduate outcome indicators

4.1 Deriving the target population from the Postsecondary Student Information System (PSIS) data: Tables: 37-10-0122-01, 37-10-0114-01 and 37-10-0115-01

Using both PSIS and tax data, five graduating cohorts from the 2010 to 2014 calendar years were derived for the indicators released on December 4, 2018.

In PSIS, graduate counts refer to the total number of degrees, diplomas and certificates awarded to graduates. For example, if an individual graduates with two credentials in a calendar year, this individual appears twice in the PSIS graduate file for that year.

In contrast to the PSIS published counts, the approach used in the Graduate Outcome Indicators focuses on outcomes for unique persons, rather than for each reported credential earned. This method is more appropriate for matching PSIS graduates to tax data and for simplifying the interpretation of the outcomes over time.

The derivation of the target population thus requires some adjustments to the source population data and is described below.

Creation of the cohort of graduates

Students graduate at different times over the year. PSIS information on graduates can be assessed according to either the institution's PSIS reporting year or a calendar year. As tax data are compiled based on a calendar year, the graduating cohorts were constructed based on the unique individuals who graduated in each calendar year to facilitate the analysis of the linkage results. Complete cohorts were constructed for calendar years 2010 to 2014. Calendar year graduates are collected from two reporting years e.g. 2010 graduates come from the second part of the 2009/**2010** reporting year data and the first part of **2010**/2011.

Exclusion of some PSIS programs

A number of program types, credential types and fields of study are excluded from the main graduate population definition. Some of these are considered 'out-of-scope' for looking at postsecondary graduate outcomes as they are either related to the in-class components of apprenticeship training, are non-postsecondary in nature, do not result in a credential or they specifically prepare students to enter another postsecondary program rather than the labour market. They are:

1. Apprenticeship programs;
2. High school/secondary diploma and certificate programs;
3. Pre-technology education/pre-industrial art programs;
4. Basic education programs;

1. Records may be imputed for PSIS using other information if a particular institution does not respond to PSIS in a given year. These records cannot be matched to tax data and therefore are removed from the counts.

5. Undergraduate or graduate qualifying programs;
6. Micro-programs (related to co-operative education terms in Quebec); and
7. Non programs (PSIS records for students taking courses or involved in educational activities that are not officially part of a credential program).

Grouping by educational qualification

Graduates are grouped by 'educational qualification' using the 'Classification of programs and credentials' (a combination of the PSIS program type and credential type variables) according to the credentials they receive.

The educational qualification category definitions make two further adjustments compared to PSIS to permit greater consistency and homogeneity of grouping when studying graduate outcomes:

1. Undergraduate or post-baccalaureate non-graduate degrees in the six fields of study of Dentistry (DDS, DMD), Law (LLB, JD, BCL), Medicine (MD), Optometry (OD), Pharmacy (PharmD, BS, BSc, BPharm) or Veterinary medicine (DVM) are moved to their own category named "professional degree".
2. Post-baccalaureate non-graduate degrees in "Education" or in "Social work" are regrouped with the undergraduate degrees group for more consistency across jurisdictions. A small number of post-baccalaureate non-graduate degrees in a handful of other selected fields of study are also regrouped with undergraduate degrees due to the small remaining category size. Thus, very few graduates remain in the post-baccalaureate non-graduate degree category and the category no longer represents these graduates.

Please note: These changes made to the groupings of programs and credentials do not impact the final total target population count, only the classification within it.

Removing multiple records in the same cohort for an individual

To simplify the analysis of graduate outcomes by educational qualification groups, only one record was retained for individuals obtaining more than one credential in the same calendar year and thus appearing more than once in the year's PSIS data.

To retain only ONE credential record per graduate, a set of sequential rules are applied in the following order to choose the most relevant record for the outcome analysis:

1. Keep the record with the highest PSIS program type, e.g. graduate program level is retained over undergraduate level;
2. If program types are the same, then keep the record with the highest PSIS credential type, e.g. degree is retained over diploma;
3. If credential types are also the same, but one credential is in the field of study of 'Education' then it is retained over the other fields of study;
4. If credential types are also the same, but neither credential is in the field of study of 'Education' then one of the fields of study is chosen at random.

Missing information and out-of-scope individuals

Finally, graduate records with missing key demographic information (sex or age) are flagged and removed since not enough information was present to classify them. Graduates less than age 15 or 65 years old and over were also removed and deemed out-of-scope as the study focused on outcomes in the labour market. Age is defined as reported on December 31 of the calendar year of graduation. It is assumed that graduates over age 64 are more likely to complete programs for personal achievement and development, rather than as an asset for the labour market. However, these numbers are generally quite small.

Fields of study for non-postsecondary, non-credit and/or personal improvement programs are not included.

Final target population

The final target population of graduate data contains one record per graduate aged 15 to 64 with an in-scope postsecondary credential (no multiple records), and excludes individuals missing selected key (sex or age) demographic variables.

4.2 Criteria used to derive the analysis subpopulations for the median employment income indicator calculations

For a graduate to be included in the analysis subpopulation for calculating the median employment income indicators, the following criteria must be met. The criteria were applied slightly differently for the cross-sectional analysis versus the longitudinal cohort analysis (see below).

a) Criteria based on tax record availability

Information on the graduates after graduation, including their earnings and student status for those returning to school, comes from the tax records. These graduate outcome indicator calculations only include graduates that could be linked to a tax record for each year following graduation that was included in the analysis.

b) Criteria based on school attendance

Graduates who returned and attended school full-time during the year(s)/period of analysis after the PSIS graduation are excluded from the median employment earnings analysis subpopulation because their employment and earnings profiles differ from graduates who are working and attending school part-time and from graduates who are working and not attending school.

The full-time and part-time 'returners' to school are identified through tax variables indicating eligibility for full-time or part-time education-related credits. Tuition credits and longitudinal PSIS data are not used to identify returning students for the December 2018 indicator release.

The group of graduates who returned to school part-time only (and were eligible for tax credits) is not substantially different from the non-returners in terms of average earnings or earnings distribution one year after graduation. For the outcomes analysis, it was decided that part-time returners would be grouped with non-returners.

In contrast, a large difference exists in average employment earnings (mean, median and distribution) for the full-time returners compared to the part-time returners and non-returners. These results confirm that full-time returners should not be included in the same median earnings analysis with the other graduates.

Full-time returners are excluded from the graduate cohort subpopulation for earnings analysis in the year(s) of analysis.

Two different types of analysis are then used to develop the graduate outcome indicators:

The **cross-sectional** analysis label means that the exclusion criteria above used to define the analysis subpopulation of graduates are only applied in the year of the outcome measurement e.g. in the second or fifth year after graduation for these indicators.

Cross-sectional data for two and five years after graduation should not be compared to each other, because the underlying populations are different.

The **longitudinal** analysis label means that graduates who met the exclusion criteria in any of the tax years after the graduation year up to the last year of analysis (e.g. from the first year after up to year five) were excluded from the analysis for all years of the longitudinal period. Doing this ensures that the analysis subpopulation is the same in every year of the analysis period and the results can be compared over time.

Please note: Results from the cross-sectional and longitudinal analysis should not be compared because the underlying populations are different. The results may in fact be quite different for some groups due to different definitions. Actual dollar amounts should be interpreted with this in mind.

Graduate outcome indicators released in Tables: 37-10-0122-01, 37-10-0114-01 and 37-10-0115-01

The counts of graduates in the target population who during the cross-sectional year/longitudinal period of analysis:

- had no tax information (i.e. no employment income);
- returned to school full-time;
- had employment income (wages, salaries, commissions, etc. or self-employment income);
- had only wages, salaries, or commissions, (but no self-employment income).

Earnings-related indicators:

- median employment income
- median wages, salaries and commissions.

These indicators are released for different graduating year cohorts (where data allows) and according to dimensions such as province of postsecondary institution of graduation, educational qualification (type of program and credential), field of study, sex, age group, and student status in Canada. Estimates are available for two and five years after graduation for both a cross-sectional analysis and a longitudinal analysis. Additional results for other years (example: year one, three, four, etc. after graduation) are available by request. The estimates are subject to annual revision.

Treatment of graduates with no wages, salaries or commissions income in the calculations:

Graduates with no employment income in a given year are excluded from the median employment income calculations for the cross-sectional analysis. The proportion is usually quite small and it is assumed that most graduates who had no employment income for the entire calendar year were probably out of the labour market. In this way, the earnings measure only includes those who actually worked in that year. Additionally, doctoral degree graduates with no employment income at all, but who are in a special situation because they have post-doctoral fellowships (not included with employment income), would not pull down the medians for those with regular employment income. Note that doctoral degree graduates with small amounts of employment income and post-doctoral fellowships are still included.

Graduates with no employment income in the year are included in the counts of the different characteristics of graduates for the cross-sectional analysis to significantly reduce the amount of suppression of data cells required for confidentiality reasons.

Graduates with no employment income are also included in the counts and the median employment income calculations for the longitudinal analysis. Since the aim of the longitudinal analysis is to measure the change in income characteristics by following the exact same group of graduates over the five year period, the proportion of graduates with no employment over the five year period is very small and is kept in the calculations.

5. Quality analysis

Due to data gaps for some Ontario colleges, graduate outcome indicators are available for Ontario university graduates only.

Quebec CEGEP graduates with a diploma in the stream leading to labour market entry are included in the outcomes analysis. The pre-university diploma graduates from CEGEPs are excluded from the outcomes analysis by educational qualification. University micro-programs and attestations are also excluded from the analysis at the undergraduate, master's and doctoral levels.

Due to data gaps for some New Brunswick colleges, graduate outcome indicators for the 2010 cohort are available for New Brunswick university graduates only.

Due to data gaps, graduate outcome indicators for the 2010 cohort are not available for Manitoba college graduates and for graduates who obtained Career, technical or professional training program certificates from Manitoba universities.

The Territories include Northwest Territories, Yukon and Nunavut. Due to data gaps, graduate outcome indicators for the Territories start with the 2012 cohort.

Postsecondary Student Information System (PSIS) – T1 Family File (T1FF) Linkage Quality

- Over 90% of PSIS graduates link to the T1FF at least once.
- Over 94% of Canadian graduates link at least once.
- Although they comprise approximately 8% of graduates, international students account for about 50% of never-linked graduates within a two year period after graduation.

5.1 Confidentiality and rounding

All data are subject to the confidentiality procedures of rounding and suppression.

Constant dollar conversion and rounding:

All median employment income figures are adjusted for inflation using the Consumer Price Index (CPI) and are presented in 2016 constant dollars to correspond with the most recent year of tax data available when the indicators were calculated. To protect the confidentiality of graduates, counts and amounts are rounded. Rounding may increase, decrease, or cause no change to counts and amounts.

When producing estimates on graduate incomes, the disclosure control rules as outlined on the webpage for the T1FF were used.

5.2 Limitations

Limitations for data linkages and the data interpretation should be acknowledged.

- Data coverage: PSIS data exhibits some institution-level non-response or non-tax-linking data gaps for some years with the newer years generally being more complete. These gaps are mostly concentrated by jurisdiction and by type of institution. These gaps are imputed for regular PSIS estimates, but imputed data cannot be used when linking to tax data. The gaps thus affect which graduating cohorts and educational qualifications are included in the tables for each jurisdiction. Improvements to response and linkage rates for these institutions are ongoing and time will yield more complete years of longitudinal data.
- Survey data and administrative data are not free of errors and inconsistencies. There may be differences in the way in which some institutions report different variables or program records for PSIS. For some variables, some institutions may provide best proxies rather than not respond at all. Recognized inconsistencies can sometimes be dealt with by adapting the indicator methodology.
- There may be a possible bias or measurement errors to graduate outcome indicators due to a small proportion of incorrect links between records and the presence of unlinked records caused by missing linkage information.
- Another possible bias to graduate outcome indicators may be due to T1 non-filers (or late filers). These individuals do not file their taxes for various reasons including some people not being required to file. Therefore they have no information in the T1FF annual database. As some classes of individuals (e.g. lower earners and those out of the labour force) might be more likely to be non-filers, this might affect the analytical results. This study cannot distinguish between non-filers and individuals that are unlinked due to missing data.

This analysis does not identify students with multiple credentials to assess if they have different outcomes. This kind of analysis requires a different and more complex methodology and will require more years of graduation data to get better results. Work has started on this analysis, however, results/findings are not available as of this data release.

Appendix A

Glossary of terms

Cross-sectional analysis: Analysis of results for cohorts of graduates with tax records for the tax year of analysis, two or five years after graduation and who did not return to school full-time in that year.

Educational qualification: This variable classifies the kind of credential a student obtained (e.g., undergraduate degree, Master's certificate) according to the 'Classification of programs and credentials' which essentially combines the two PSIS variables of program type (e.g., undergraduate program, master's program) and credential type (e.g., degree, diploma or certificate) and also creates a separate category for Professional degree. Some combinations were renamed for easier identification by data users (e.g., master's degree, and doctoral degree). The educational qualification obtained is expected to be an important contributor to determination of the new graduate's earnings.

Employment income: Employment income includes wages, salaries and commissions (wages and salaries, commissions from employment, training allowances, tips and gratuities, tax exempted Indian employment income) and net self-employment income (net income from business, profession, farming, fishing and commissions). Net self-employment income may be positive, negative or zero. In analysis subpopulations where the self-employed are excluded, wages, salaries or commissions will make up all of employment income.

Field of study and Classification of Instructional Programs (CIP) 2016: The CIP is used to classify the main field of study of postsecondary programs at four standardized levels of aggregation (classes, sub-series, series and 13 Primary Groupings). The 'Variant of CIP 2016 – Alternative primary groupings' is used in the release tables and is the most aggregated of these. CIP is most useful for analysis when combined with information on educational qualification.

Longitudinal cohort analysis: This refers to graduates of a given calendar year (cohort) who have tax records for all years of the analysis period, and who did not return to school full-time at any time during the period from one year after graduation to the final year of analysis (e.g. five years after graduation) as of the December 4th 2018 table release.

Median employment income: The median employment income of a specified group is the amount that divides the employment income distribution of that group into two halves, i.e., the incomes of half of the units in that group are below the median, while those of the other half are above the median.

Note that median income after graduation does not control for the effect of additional education that may have been obtained since the cohort graduation year for the cross-sectional cohorts.

Earnings estimates for doctoral graduates do not include post-doctoral fellowships since they cannot be separated from the 'Other income' category on the T1 form including other amounts from lump sum payments, retiring allowance, death benefits, and other kinds of income.

All dollar figures are rounded to the nearest hundred. All median employment income figures are adjusted for inflation using the Consumer Price Index (CPI) and are presented in 2016 constant dollars to correspond with the most recent year of tax data available when the indicators were calculated.

Full-time returners: Full-time returners are individuals identified as having returned to school after their initial graduation using the T1FF education credit eligibility variables or the PSIS database. In this study, those who return to school full-time are treated differently (and removed from the analysis subpopulation) than those who only return part-time.

Self-employed: Self-employed individuals are identified using the tax variable for any non-zero gross self-employment income or net income, since a negative or exactly zero net income can be reported. Gross self-employment income includes at least one of the following types of earnings: gross farming income, gross commission income, gross business income, gross fishing income or gross professional income.

Student Status in Canada: Student status in Canada is defined at the end of the winter term, during the year of graduation. 'Canadian students' include Canadian citizens and permanent residents. 'International students' include students studying in Canada on student visas, non-Canadian students in Canada on other types of visas, non-Canadians whose status is unknown, and students studying from outside Canada (e.g., by Internet).

T1 Family File (T1FF): The T1FF combines variables from the T1 and T4 tax files and the Canada Child Tax Benefit into a family composition file. It includes income, demographic and geographic variables for each tax filer and their spouse, family and children. Only selected tax variables are added to the PSIS postsecondary graduate records.