

## **T2-LEAP Data dictionary, 2000 to 2014**

The Longitudinal Employment Analysis Program (LEAP) is an administrative databank maintained by the Economic Analysis Division (EAD) at Statistics Canada. It was designed to provide longitudinal data on the behaviour of employment levels of Canadian businesses (Baldwin, Dupuy and Penner, 1992). The database, which covers the period 2000 to 2015, makes use of administrative tax records, data from the Business Register and from the Survey of Employment, Payrolls and Hours (SEPH) to derive the employment profile of businesses over time.

In order to track firms along with their performance, the LEAP file was linked to the Corporate Income Tax File (T2) to create T2-LEAP. The T2 file includes all incorporated firms that file a T2 tax return with the Canada Revenue Agency (CRA). The T2 file provides data on, among other things, sales, gross profits, equity and assets for all incorporated firms in Canada.

The universe of T2-LEAP covers all incorporated employers in Canada in the private sector. Public sector enterprise, classified in 2-digit NAICS 61, 62, and 91, are excluded. The dataset is constructed at the enterprise level.

This dictionary identifies and defines the T2-LEAP variables. General Index of Financial Information (GIFI) fields specific to farming corporation are not included. The variable definitions alphabetically lists each variable. The following information is provided:

- Definition of the variable and codes, if applicable;
- Source of the variable such as the line number on the tax form;
- Variable name.

### **T2-LEAP variable definitions**

#### **Assets**

Definition: The sum of current assets, capital assets, intangible assets, total accumulated amortization of intangible assets, and long-term assets.

Derived from: GIFI-Balance Sheet, Field 2599

Name: total\_assets

#### **Expenses**

Definition: Total expenses. The sum of operating expenses and cost of sales.

Derived from: GIFI-Income Statement, Field 9368  
Name: total\_expenses

### **First year**

Definition: The year the enterprise first appears in LEAP. It is equal to 1984 or the first year in which the enterprise hired employees.

Derived from: LEAP  
Name: leap\_first\_year

### **Gross profits**

Definition: Total sales of goods and services minus total cost of sales.

Derived from: GIFI-Income Statement, Field 8519  
Name: gross\_profits

### **Industry**

Definition: The 4-digit North American Industrial Classification System (NAICS) assigned to the enterprise. One of the characteristics of the LEAP is that the most recent industry code of the enterprise is pushed back, so that an enterprise has the same industry code each year. The industry code derived from LEAP represents the first most dominant longitudinal industry for the firm in LEAP.

Derived from: LEAP  
Name: naics

### **Labour**

Definition: Average labour units at the national level.

Derived from: LEAP  
Name: nalus

### **Last Year**

Definition: The last year that the enterprise has employees and is present in LEAP.

Derived from: LEAP  
Name: leap\_last\_year

### **Longitudinal Business Register Identifier (LBRID)**

Definition: Unique identifier for enterprises in T2-LEAP.

There may be certain lbrids with more than one observation per year in T2-LEAP. This is because some lbrids can be linked to more than one tax record. The information from LEAP for enterprises with more than one observation per year is the same across all observations for that year. Therefore, to obtain revenue for an enterprise in a particular year it may be necessary to add up revenue across all the relevant observations, but to obtain ALUS one needs only take the value from one of the observations.

Derived from: LEAP  
Name: lbrid

### **Net income/loss before taxes**

Definition: Net income/loss before taxes and extraordinary items. The sum of net non-farming income and net farming income.

Derived from: GIFI-Income Statement, Field 9970  
Name: netincome\_or\_loss\_beforetaxes

### **Payroll**

Definition: Sum of payroll from all T4 slips issued by enterprise.

Derived from: LEAP  
Name: leappayroll

### **Revenue**

Definition: Total revenue is the sum of sales of goods and services and other revenues, such as interest income.

Derived from: GIFI-Income Statement, Field 8299  
Name: revenue

### **Sales of goods and services**

Definition: Total sales of goods and services.

Derived from: GIFI-Income Statement, Field 8089  
Name: sales\_of\_goods\_and\_services

### **Shareholder's equity**

Definition: The sum of common shares, preferred shares, contributed and other surpluses, head office account and retained earnings/loss.

Derived from: GIFI-Balance Sheet, Field 3620  
Name: total\_shareholder\_equity

### **Size of firm**

Definition: National size code of firm based on national average labour units (nalus).

The codes are:

- '0' - 0 nalus
- '1' -  $0 < \text{nalus} < 5$
- '2' -  $5 \leq \text{nalus} < 9.99$
- '3' -  $10 \leq \text{nalus} < 19.99$
- '4' -  $20 \leq \text{nalus} < 49.99$
- '5' -  $50 \leq \text{nalus} < 99.99$
- '6' -  $100 \leq \text{nalus} < 499.99$
- '7' -  $\text{nalus} \geq 500$

Derived from: LEAP  
Name: nsize

### **Tangible assets**

Definition: Assets with a physical form, such as buildings, land, and machinery and equipment. They are sometimes called fixed, capital assets, or property, plants and equipment. Not included are financial and intangible assets.

Derived from: GIFI-Balance Sheet, Field 2008  
Name: total\_tangible\_assets

### **Tax year**

Definition: The reference of the filing corporation.

Derived from: T2 records  
Name: tax\_yr

### **Total liabilities**

Definition: Sum of current, long-term liabilities, subordinated debt and amounts held in trust.

Derived from: GIFI-Balance Sheet, Field 3499  
Name: total\_liabilities

### **Total liabilities and shareholder equity**

Definition: The sum of total liabilities and total shareholder equity.

Derived from: GIFI-Balance Sheet, Field 3640

Name: totalliabilities\_shrholderequity

### **Working capital**

Definition: Current assets minus current liabilities.

Derived from: GIFI-Balance Sheet, Field 1599 minus Field 3139

Name: working\_capital

Additional information: The sum of this variable across firms over time is very volatile.