



OTTAWA CANADA

13 | 15 | 20 | 22

SEPTEMBER

2022

Presentation of results from the Voorburg Group Alternative Data Survey

37th Voorburg Group Meeting
September 13-22, 2022

Virtual Meeting hosted by Canada

Dragos Ifrim (Canada)



Statistics
Canada

Statistique
Canada

Rohan Draper (Denmark)



Alternative Data Sources in Official Statistics

Alternative Data Survey - xlsx

Voorburg Group - Alternative Data Sources

2021

Purpose

- To map alternative data use per index for Voorburg members
- To enable Voorburg members to identify potential development partners
- To highlight trends in contemporary price index data source practises
- Improved decision making regarding SPPI development activities

Instructions

- This questionnaire is focused on alternative data sources for the compilation of price indexes.
- For each industry or product classification select the appropriate data collection method by clicking on the respective cell and chose from the list of available options.
- For definition support see the tab entitled 'definitions'.
- If you do not collect data for a given industry code no action is required (i.e. leave the cell as "<select>")
- "Supplementary information" cells are included if you would like to provide a short description of the alternative data (source, collection, use etc.) and/or include a link to a relevant document.

Completion Time

Estimated time for completion is circa 15min.

Administrative Questions

Which country do you represent?	Canada
Which department and/or team do you represent?	Statistics Canada
Are your PPIs product or industry based (or both)?	Both
What is your current applied classification	NAICS and NAPCS

Tips

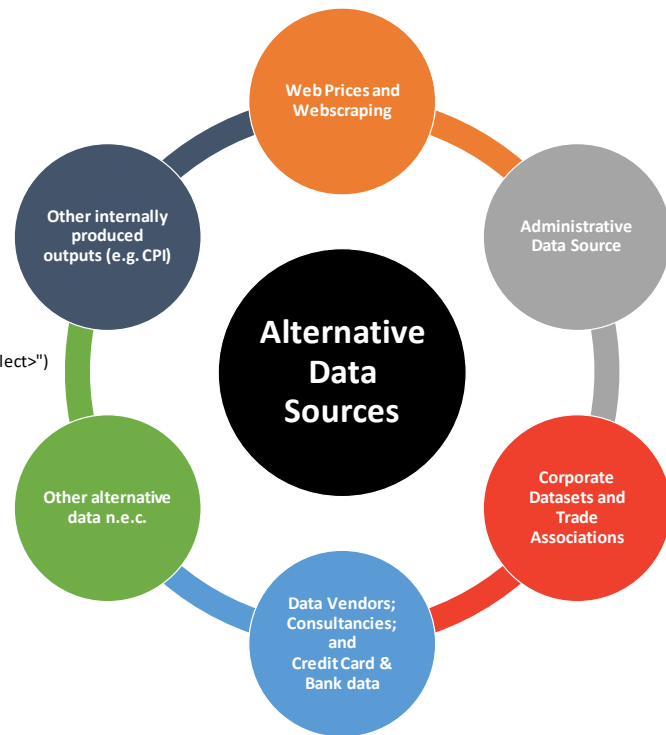
If your statistics office does not use the classification in the questionnaire a link is provided below where you can find a correspondence table via UN Stats and Eurostat

<https://unstats.un.org/unsd/classifications/Econ#corresp-isis-c>

https://ec.europa.eu/eurostat/ramon/rerelations/index.cfm?TargetUrl=LST_REL

▶ [click here](#) to go to *definitions* of alternative data

▶ [click here](#) to complete the survey for your country



Contents

- Purpose
- What's new?
- CPI Usage
- Multisource Usage
- Visualisation
- Survey and Results Guidelines
- What's next?

Alternative Data Sources in Official Statistics

Alternative Data Survey - xlsx

Voorburg Group - Alternative Data Sources

2021

Purpose

- To map alternative data use per index for Voorburg members
- To enable Voorburg members to identify potential development partners
- To highlight trends in contemporary price index data source practises
- Improved decision making regarding SPPI development activities

Instructions

- This questionnaire is focused on alternative data sources for the compilation of price indexes.
- For each industry or product classification select the appropriate data collection method by clicking on the respective cell and chose from the list of available options.
- For definition support see the tab entitled 'definitions'.
- If you do not collect data for a given industry code no action is required (i.e. leave the cell as "<select>")
- "Supplementary information" cells are included if you would like to provide a short description of the alternative data (source, collection, use etc.) and/or include a link to a relevant document.

Completion Time

Estimated time for completion is circa 15min.

Administrative Questions

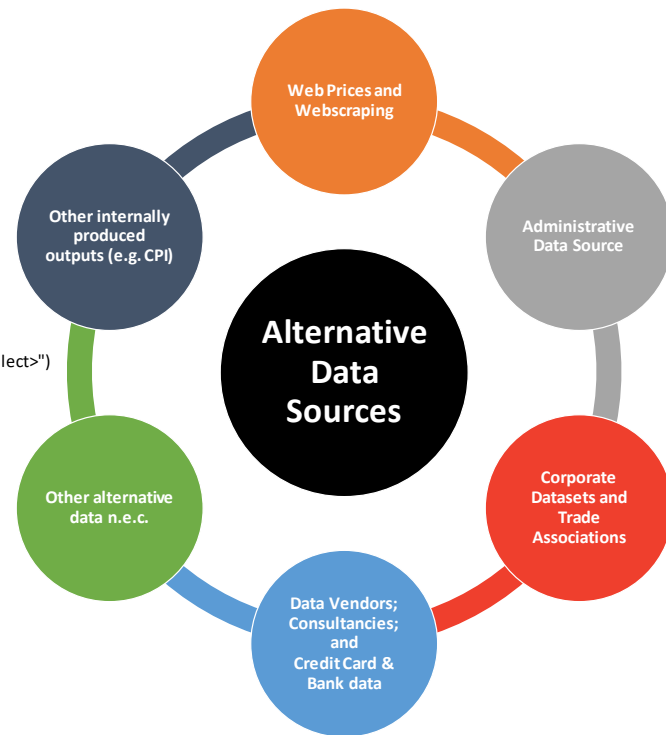
Which country do you represent?	Canada
Which department and/or team do you represent?	Statistics Canada
Are your PPIs product or industry based (or both)?	Both
What is your current applied classification	NAICS and NAPCS

Tips

If your statistics office does not use the classification in the questionnaire a link is provided below where you can find a correspondence table via UN Stats and Eurostat
<https://unstats.un.org/unsd/classifications/Econ#corresp-isis-un>
https://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_REL

► [click here](#) to go to *definitions* of alternative data

► [click here](#) to complete the survey for your country



Purpose

- o Map alternative data use per index
- o Enable Voorburg members to identify potential development partners
- o Highlight trends in contemporary price index data source practises
- o Enhance decision making regarding SPPI development activities



OTTAWA CANADA
13 | 15 | 20 | 22
SEPTEMBER
2022

What's new?

- o If members have requests and/or ideas on other ways the results can be presented please contact the presenters.
- o Attaining increased coverage remains a top priority.

Alternative Data Sources in Official Statistics

Presentation of Results.

The excel document now includes visualisations of the results plus an analysis of the prevalence of individual statistical agencies using multiple sources to produce a particular index.

Value Proposition.

Data from the survey has already provided value regarding investigations into CPI usage by PPI statistical programmes and provided insightful input into alternative data investigations.

Coverage.

Thank you to the countries that responded last year and those that have sent in responses and increased the coverage of the survey this year. We now have 16 countries contributing to the results.



Alternative Data Sources in Official Statistics

Category Recap

o Categories remain the same since the original survey.

Source	Type	Origin	Code	Description
TRADITIONAL	STRUCTURED	Questionnaire (paper, phone and/or electronic)	QNR	The traditional way of collection price information by asking firms for the information via telephone, paper or electronic questionnaires. This is the default and not an alternative data source.
	UNSTRUCTURED	Web Prices (manual)	WPR	The price collection data is derived from websites manually. The data source itself is considered an unstructured data source not designed for statistical purposes.
Webscraping (automated)		WSC	The price collection data is derived from websites automatic systems. The data source itself is considered an unstructured data source not designed for statistical purposes.	
EXTERNAL	STRUCTURED	Administrative Data Source	ADM	Data which are derived from the operation of administrative systems by public agencies (e.g. data collected by government agencies for the purposes of registration, transaction, regulation and record keeping). Data is often structured for administrative purposes and is highly transferable for statistical purposes.
		Corporate Datasets	COR	Survey respondent provided datasets obtained directly from corporate headquarters in lieu of data collectors collecting data in respondent stores or on their websites. Data pertains to the particular company that is providing said data is often structured for organisational purposes and is highly transferable for statistical purposes.
		Trade Associations	TAD	Industry based surveys that the target industry is producing for themselves.
		Data Vendors (commercially available structured data)	DVS	Data acquired from companies that actively collect and sell data as a business activity. Often such companies provide data on a contractual basis with defined terms and conditions.
		Consultancies (mandated specific task) (transformed data)	CON	Consulting company and/or specialist company is contracted to collect and/or compile data for a specific purpose (mandated or otherwise). Often such companies are utilised on a contractual basis with defined terms and conditions.
		Credit card and bank data	CCD	Financial information collected at the moment of a transfer of funds between a card holder's account and a business account. Data is graded based on the level of metadata available about the transaction. This source is considered a structured data source.
	BOTH	Other alternative data sources n.e.c.	OTH	Other types of alternative data sources not elsewhere classified. For example, transaction-level data from email receipts (like UBER email receipt data). Other special data delivery from third party data collectors not elsewhere classified.
INTERNAL	STRUCTURED	Consumer Price Index	CPI	Data is sourced directly from the Consumer Price Index
		Producer Price Index	PPI	Data is sourced directly from the Producer Price Index
		Structural Business Statistics	SBS	Surveys utilised for benchmarking purposes
		National Accounts	NA	Price indices derived from volume and value data (implicit price indices)

CPI Usage

- o CPI remains the predominant alternative data source for PPI.
- o Guidelines for using CPI as an alternative datasource can be found on the Voorburg website.

Alternative Data Sources in Official Statistics

CPI frequency as an alternative data source for a particular 4-digit classification.		
Class	Name	Frequency
H	Transportation and storage	
4911	Passenger rail transport, interurban	8
4912	Freight rail transport	1
4921	Urban and suburban passenger land transport	9
4922	Other passenger land transport	7
4923	Freight transport by road	2
5011	Sea and coastal passenger water transport	6
5021	Inland passenger water transport	1
5110	Passenger air transport	6
5221	Service activities incidental to land transportation	4
5310	Postal activities	3
5320	Courier activities	2
I	Accommodation and food service activities	
5510	Short term accommodation activities	7
5520	Camping grounds, recreational vehicle parks and trailer parks	4
5590	Other accommodation	2
5610	Restaurants and mobile food service activities	11
5621	Event catering	1
5629	Other food service activities	4
5630	Beverage serving activities	5
J	Information and communication	
5811	Book publishing	4
5813	Publishing of newspapers, journals and periodicals	5
5819	Other publishing activities	1
5820	Software publishing	1
5911	Motion picture, video and television programme production activities	2
5913	Motion picture, video and television programme distribution activities	1
5914	Motion picture projection activities	4
5920	Sound recording and music publishing activities	1
6020	Television programming and broadcasting activities	2
6110	Wired telecommunications activities	5
6120	Wireless telecommunications activities	3
6312	Web portals	1
6399	Other information service activities n.e.c.	1

Alternative Data Sources in Official Statistics

CPI Usage

- o CPI remains the predominant alternative data source for PPI.
- o Guidelines for using CPI as an alternative datasource can be found on the Voorburg website.

CPI frequency as an alternative data source for a particular 4-digit classification.

Class	Name	Frequency
N	Administrative and support service activities	
7710	Renting and leasing of motor vehicles	1
7722	Renting of video tapes and disks	1
7729	Renting and leasing of other personal and household goods	2
7740	Leasing of intellectual property and similar products, except copyrighted works	1
7911	Travel agency activities	2
7912	Tour operator activities	3
7990	Other reservation service and related activities	1
8121	General cleaning of buildings	1
8129	Other building and industrial cleaning activities	1
P	Education	
8510	Pre-primary and primary education	2
8521	General secondary education	2
8522	Technical and vocational secondary education	3
8530	Higher education	1
8541	Sports and recreation education	1
8542	Cultural education	1
8550	Educational support activities	1
R	Arts, entertainment and recreation	
9311	Operation of sports facilities	4
9312	Activities of sports clubs	2
9321	Activities of amusement parks and theme parks	1
9329	Other amusement and recreation activities n.e.c.	1
S	Other service activities	
9511	Repair of computers and peripheral equipment	1
9512	Repair of communication equipment	1
9521	Repair of consumer electronics	1
9522	Repair of household appliances and home and garden equipment	1
9523	Repair of footwear and leather goods	1
9601	Washing and (dry-) cleaning of textile and fur products	3
9602	Hairdressing and other beauty treatment	5
9603	Funeral and related activities	4
9609	Other personal service activities n.e.c.	2



Multisource Usage

o This table shows the frequency of multisource statistics per industry and class. For example, a result/count of six means that six unique statistical institutions have reported using multiple sources of data to produce a particular index.

Alternative Data Sources in Official Statistics

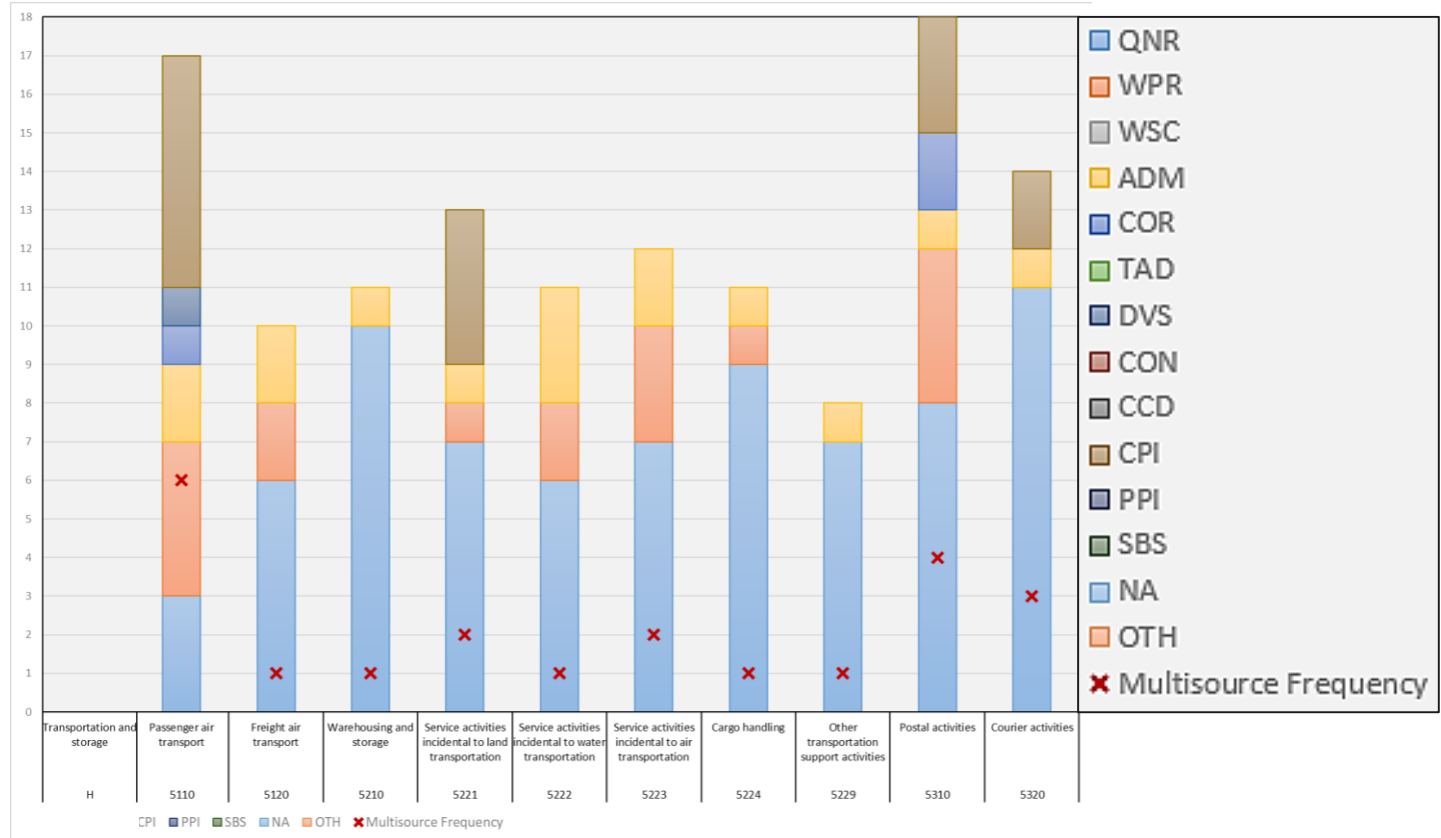
Occurance of institutions using multiple data sources for a particular 4-digit classification.			
Industr	Class	Name	Count
H	5110	Passenger air transport	6
H	5310	Postal activities	4
J	6110	Wired telecommunications activities	4
L	6820	Real estate activities on a fee or contract basis	4
M	6910	Legal activities	4
H	4922	Other passenger land transport	3
H	4923	Freight transport by road	3
H	5320	Courier activities	3
J	6120	Wireless telecommunications activities	3
K	6419	Other monetary intermediation	3
H	4912	Freight rail transport	2
H	5011	Sea and coastal passenger water transport	2
H	5012	Sea and coastal freight water transport	2
H	5221	Service activities incidental to land transportation	2
H	5223	Service activities incidental to air transportation	2
I	5510	Short term accommodation activities	2
M	7120	Technical testing and analysis	2
N	7710	Renting and leasing of motor vehicles	2
N	7740	Leasing of intellectual property and similar products, except copyrighted works	2
N	7810	Activities of employment placement agencies	2
N	8129	Other building and industrial cleaning activities	2

H – Transportation and Storage (example)

Visualisation

○ This visualization shows the frequency that each price collection method is being utilised on a per class basis (4-digit level)

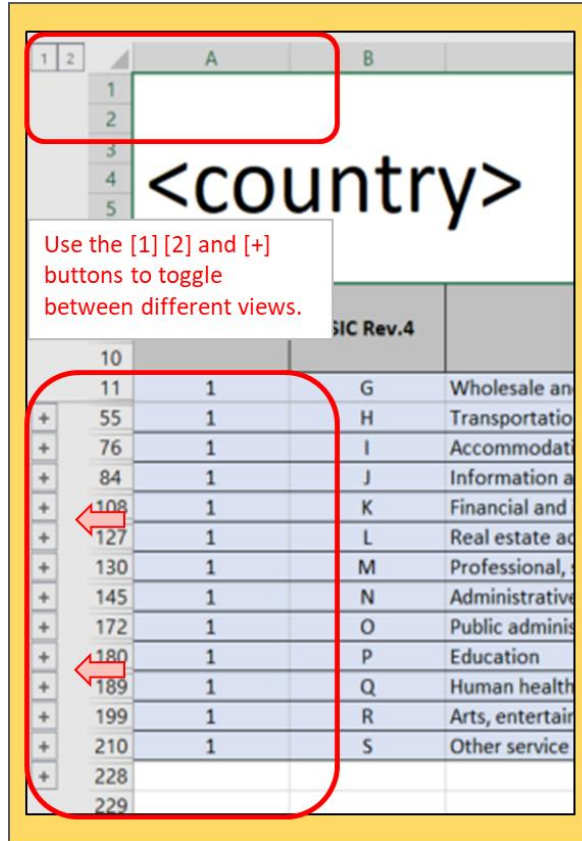
○ The **x** shows the frequency that unique NSIs are using multiple sources to produce a particular 4-digit index



Alternative Data Sources in Official Statistics

Survey and Results Guidelines

- o Please remember to expand the industries via the '+' icon or [1] / [2] buttons
- o This enables completion of information at the 4-digit level



Use the [1] [2] and [+] buttons to toggle between different views.

	10		SIC Rev.4	
	11	1	G	Wholesale and
+	55	1	H	Transportation
+	76	1	I	Accommodation and
+	84	1	J	Information and
+	108	1	K	Financial and
+	127	1	L	Real estate and
+	130	1	M	Professional, scientific
+	145	1	N	Administrative and
+	172	1	O	Public administration
+	180	1	P	Education
+	189	1	Q	Human health and
+	199	1	R	Arts, entertainment and
+	210	1	S	Other services
+	228			
	229			

Alternative Data Sources in Official Statistics

Survey and Results Guidelines

- o In the “Results_Per_Country” sheet, data can be sorted by 4-digit level classification via “Column2”
- o For those developing or reviewing an index this provides a good overview of coverage per country

Detailed results per country

Source.Name	Column1	Column2	Column3	Column4	Column5	Column6
Australia	Sortér fra A til Å		Freight transport by road QNR		PD	<select>
Canada	Sortér fra Å til A		Freight transport by road QNR		PU	<select>
China	Sortér fra Å til A		Freight transport by road QNR		PD	OTH
Denmark	Sortér efter farve		Freight transport by road QNR		PU	<select>
Ireland			Freight transport by road QNR		PU	<select>
Japan	Fjern filter fra "Column2"		Freight transport by road QNR		PU	TAD
Latvia	Filtrer efter farve		Freight transport by road QNR		PU	<select>
México	Tekstfiltre		Freight transport by road QNR		PD	ADM
Poland			Freight transport by road QNR		PU	<select>
Republic of Korea			Freight transport by road QNR		PU	<select>
Spain			Freight transport by road QNR		PU	<select>
Sweden			Freight transport by road QNR		PU	<select>
Switzerland			Freight transport by road CPI		PU	<select>
United Kingdom			Freight transport by road QNR		PU	<select>
United States			Freight transport by road QNR		PU	<select>

Note: A visualisation of this information has been suggested for a future version of the results



OTTAWA CANADA
13 | 15 | 20 | 22
SEPTEMBER
2022

Alternative Data Sources in Official Statistics

What's next?

- o If members have requests and/or ideas on other ways the results can be presented please contact the presenters.

- o Attaining increased coverage remains a top priority. Please contact the Voorburg Bureau via Ifrim and Rohan with contributions and/or queries.

Visualisation per index.

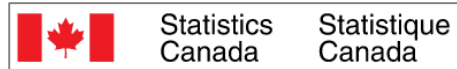
Country and sources shown.

Visualisation of weights information.

Other.

Feedback and/or ideas from participants can be sent to Ifrim and Rohan for consideration by the Voorburg Bureau. If time permits, additions to the excel summary document will be actioned by the 22nd September.

Dragos Ifrim (Canada)
dragos.ifrim@statcan.gc.ca



Rohan Draper (Denmark)
rjd@dst.dk

