TECHNICAL DOCUMENTATION
HOUSEHOLDSPEND 2022 AND FOODSPEND 2022
OCTOBER 2022

TECHNICAL DOCUMENTATION - 2022

HouseholdSpend FoodSpend

WHAT THEY ARE

HouseholdSpend provides current estimates of annual expenditures for 463 variables spanning 18 categories of goods and services used by Canadian households—everything from fashion apparel and household furnishings to cell phones and charitable donations. While HouseholdSpend includes a general food category, FoodSpend provides a more detailed breakdown of food-related purchases made by Canadian households, with current estimates of annual expenditures for 276 variables of food related purchases. With FoodSpend, users can gain valuable insights into what Canadian households are buying from grocery stores and restaurants—everything from specific vegetables, meat and fish products to meal types like lunches and dinners. Both databases also include basic variables describing number of households and household income from DemoStats, which are based on Statistics Canada's Survey of Household Spending (SHS).

Both wide-ranging databases let users analyze potential expenditures by average dollars spent per household and total dollars spent for any geographical level—from all of Canada to small custom-defined trade areas—regardless of where the consumer made the purchase.

Table 1 shows counts of variables by category for HouseholdSpend and FoodSpend.

Table 1: HouseholdSpend 2022 and FoodSpend 2022 categories by count of variables

Category	Count	Product
Code	1	HouseholdSpend and FoodSpend
Basics	4	HouseholdSpend and FoodSpend
Food Purchases (Category Summary)	12	FoodSpend
Bakery products	15	FoodSpend
Cereal grains and cereal products	17	FoodSpend
Dairy products and eggs	30	FoodSpend
Fish and Seafood	20	FoodSpend
Food purchased from restaurants	6	FoodSpend
Fruit, fruit preparations and nuts	36	FoodSpend
Meat	20	FoodSpend
Non-alcoholic beverages and other food products	77	FoodSpend
Vegetables and vegetable preparations	39	FoodSpend
Household Expenditures (Category Summary)	26	HouseholdSpend
Clothing	41	HouseholdSpend
Education	10	HouseholdSpend
Food	8	HouseholdSpend
Games of chance	5	HouseholdSpend
Gifts of money and contributions	8	HouseholdSpend
Goods and services by purchase method*	3	HouseholdSpend
Health care	22	HouseholdSpend
Household furnishings and equipment	32	HouseholdSpend
Household operation	36	HouseholdSpend
Miscellaneous expenditures	17	HouseholdSpend
Personal care	17	HouseholdSpend

Personal insurance and pensions	9	HouseholdSpend
Reading materials and other printed matter	6	HouseholdSpend
Real estate	33	HouseholdSpend
Recreation	61	HouseholdSpend
Shelter	78	HouseholdSpend
Tobacco and alcohol	7	HouseholdSpend
Transportation	38	HouseholdSpend

*Goods and services by purchase method reflects a different hierarchy from other categories (such as Clothing, Education, etc.) listed here. It offers purchase methods such as sales over the Internet or other types of direct sales, but does not include detailed categories.

HOW THEY ARE BUILT - KEY DATA SOURCES

The primary data source for HouseholdSpend 2022 and FoodSpend 2022 is the Survey of HouseholdSpend (SHS). Since 2010, the SHS has also tracked food purchased from stores in greater detail than before, effectively measuring what had been captured in the now-discontinued Food Expenditure Survey (FES). Multiple vintages of SHS are used to boost sample size and capture the temporal effect of household and food spending. A series of models is calibrated on the SHS, and these models are used to score a multi-dimensional cross-distribution of household demographics derived from Environics Analytics' DemoStats, PRIZM® and custom census cross-tabulations from Statistics Canada.

Data from statistics Canada's National Economic Accounts (NEA) and Provincial and Territorial Economic Accounts (PTEA) is augmented with administrative data from the Canadian Revenue Agency (CRA) to develop national, provincial and census tract level control totals. All modelled small area estimates are optimized to sum to these higher-level control totals.

HOW THEY ARE BUILT – MODELLING FRAMEWORK

The construction of HouseholdSpend 2022 and FoodSpend 2022 involved three distinct phases: the creation of the initial small-area behavioural estimates; the collection and projection of the control totals; and a mathematical reconciliation process that ensures everything "adds up."

SMALL-AREA ESTIMATES

A series of Heckman selection, log-linear and multinomial logit models are calibrated using respondent-level SHS micro-data. These models predicted consumption using a combination of demographic, location, PRIZM®, and seasonality data as independent parameters. The log-linear and Heckman selection models estimate the base-level dollar consumption totals like "total current consumption" or "purchases of primary real estate". These estimates are then broken down into the finer level spending categories by applying percentages of spend derived from the multinomial logit models. To do so, the model coefficients are scored against a multi-dimensional cross-distribution of household demographics derived from Environics Analytics' DemoStats, PRIZM and custom census cross-tabulations from Statistics Canada in a multi-level parent-child hierarchical fashion. This effectively partitions the consumption total into increasingly more detailed spending categories, while preserving all intrinsic variable hierarchies defined by the SHS. Altogether, the initial development of

HouseholdSpend and FoodSpend required over 800 estimates of consumption using over 200 models and 150,000 coefficients for over 1,000,000 small area geographic units in Canada.

PROVINCIAL, TERRITORIAL, AND SMALL-AREA CONTROL TOTALS

The spending control totals fall into three categories: NEA/PTEA-derived provincial control totals, the SHS-derived provincial control totals, and the administrative data-derived small-area (census tract level) control totals.

The NEA and PTEA's consumption estimates form a set of authoritative control numbers, which are originated primarily from Statistics Canada's Quarterly Survey of Financial Statements (QSFS) and augmented by the SHS. In practice, the QSFS program produces robust consumption estimates, with its data derived from the financial statements of household serving institutions rather than relying on the memory recall of household members. In 2019, the matching mechanism between NEA/PTEA and SHS was enhanced to more precisely reconcile the coverage gaps between the two sources. This enhancement allows us to derive more accurate controls from NEA/PTEA wherever applicable, and hence correct for response and reporting biases within our SHS-derived models. The most recent available year of PTEA and NEA data is 2020 and 2021, respectively. Provincial shares of expenditure from PTEA as well as national NEA dollar values are projected to the current year (2022) by leveraging historical data to fit linear and non-linear lines for each category. Provincial controls are then obtained by rolling out projected provincial shares onto NEA dollar values. The projection is also cross checked with data from trade association if necessary.

The shelter and miscellaneous expenditure categories from NEA/PTEA cannot be reconciled directly with the SHS due to inconsistent definitions between the two sources. The estimated consumption for these two categories is controlled directly to SHS. On the other hand, the CRA data are prepared at the census tract level of geography, and the data are used to guide our estimates for household income taxes paid, employment insurance premiums, Canada/Quebec Pension Plan payments, charitable contributions, etc. These financial-related outlays account for about 20 percent of total Canadian household expenditures. Furthermore, the process of census tract level CRA data also involves imputation of missing values, where we apply proprietary imputation techniques that factor in spatial-temporal patterns.

RECONCILIATION PROCESS

For each year, the control totals have to be reconciled with the initial small-area estimates at the postal code and dissemination area level. This reconciliation is achieved using a set of non-linear mathematical optimizations that adjust the initial small-area estimates to agree with higher geographic level control totals. The controlled category totals at the small-area level are then allocated using the initial estimate shares of the category components. This reconciliation process results in estimates that match the control totals at different levels of geography while deviating as little as possible from the estimates derived from the raw data.

SPECIAL NOTE ABOUT COMPARABILITY WITH WEALTHSCAPES

Starting with the 2016 release, WealthScapes (and its variations), HouseholdSpend and FoodSpend all use identical definitions for disposable income and discretionary income. In fact, WealthScapes' disposable and

discretionary income data are produced directly from the HouseholdSpend and FoodSpend databases. The only difference between the disposable and discretionary income figures for each database is the year of reference. All of the statistics in HouseholdSpend 2022 and FoodSpend 2022 are for the current year (2022), and are identical to those found in DemoStats 2022. But WealthScapes 2022's population and income statistics come from 2021 data, while WealthScapes 2022's historical year statistics refer to 2020.

REFERENCE DOCUMENTS

For HouseholdSpend Release Notes, Variables List, and other reference materials, please visit:
https://community.environicsanalytics.com/hc/en-us/articles/360034824732-HouseholdSpend

For FoodSpend Release Notes, Variables List, and other reference materials, please visit:
https://community.environicsanalytics.com/hc/en-us/articles/360035199611-FoodSpend