# Commodity Flow Survey—Expanded Hazardous Materials Supplement Experimental Data Product Methodology

# OVERVIEW

The Expanded Hazardous Materials Supplement (EHM) is a survey sponsored by the Pipeline and Hazardous Materials Safety Administration and is conducted by the U.S. Census Bureau. The primary goal of the EHM is to produce national and subnational estimates of the total number of establishments that ship hazardous materials (HAZMAT), the types of HAZMAT shipped, and the types of packaging used. These data are collected as a supplement to the first and fourth quarters of the 2022 Commodity Flow Survey (CFS) and pertain to CFS respondents who report shipping HAZMAT at any time during 2021 or 2022. For a given establishment, the respondent is asked to provide the top ten most frequent types of HAZMAT shipped. For the top three most frequently shipped HAZMAT, additional detailed questions are asked including the packaging, quantity, and mode of transportation used.

By definition, a HAZMAT is a substance or material that the Secretary of Transportation has determined is capable of posing an unreasonable risk to health, safety, and property when transported in commerce and has designated it as hazardous under section 5103 of Federal Hazardous Materials Transportation Law (49 U.S.C. 5103).

# INDUSTRY COVERAGE

The CFS (and subsequently the EHM) covers establishments in selected mining, manufacturing, wholesale trade, selected retail industries, selected transportation and warehousing industries, publishers, and selected auxiliary establishments (e.g., warehouses) of in-scope multiunit companies. Industries not covered by CFS include construction, most retail and services industries, farms, fisheries, foreign establishments, and most government-owned establishments. These industries are denoted by the North American Industry Classification System (NAICS). Refer to the tables at <www.census. gov/data/experimental-data-products/cfs-expandedhazardous-materials-estimates.html> for the specific industries included and notes on establishments not included in the scope.

# **GEOGRAPHY COVERAGE**

The data shown are for establishments at the national and Petroleum Administration for Defense District (PADD) levels. PADDs are geographic aggregations of the 50 states and the District of Columbia into five distinct districts. For the purposes of this EHM publication, PADDs 4 and 5 (Rocky Mountain and West Coast) are consolidated. The following values for geography, including descriptions and applicable states, appear in Columns A and B on all tables:

.A: United States (all PADD combined/all states).

PADD\_I: East Coast (Connecticut, Delaware, District of Columbia, Florida, Georgia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, North Carolina, Pennsylvania, Rhode Island, South Carolina, Vermont, Virginia, and West Virginia).

PADD\_II: Midwest (Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, Oklahoma, South Dakota, Tennessee, and Wisconsin).

PADD\_III: Gulf Coast (Alabama, Arkansas, Louisiana, Mississippi, New Mexico, and Texas).

PADD\_IV\_V: Rocky Mountain and West Coast combined (Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, Oregon, Utah, Washington, and Wyoming).

# SAMPLE DESIGN

The sample for the EHM is selected as part of the first stage of the 2022 CFS sample. There are three stages in the CFS sample; the first stage, which selects establishments, is the only sampling stage affecting the EHM. Under the CFS design, the sampling frame is stratified by geography (CFS area or state), NAICS code, and a measure of size variable. CFS areas represent selected combined statistical areas, metropolitan statistical areas, or remainder of state areas in a single state. There are 134 CFS areas; a list of them is available on request.

There are four components of the CFS sample. The first is a list of certainty establishments across all industries. These establishments are selected for shipment characteristics that are important to the CFS. The second

The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data (Project No. 7504831, Disclosure Review Board [DRB] approval number: CBDRB-FY23-100).

component comes from specific six-digit NAICS codes correlated with frequent HAZMAT shipments from prior CFS data. There are 35 of these NAICS codes and each of those is further stratified by state, though most of the industries are selected with certainty. The third component is the auxiliary industries. The establishments in these industries were selected with certainty since CFS did not perform its typical advance survey to identify shippers and nonshippers. The fourth (and largest) component comes from the remainder of the NAICS codes, further stratified by CFS area. The industry levels in this component of sampling match the publication industry levels, except for NAICS 212 being split up into four-digit levels for sampling.

For all noncertainty strata, there is a further stratification by a measure of size variable. This ensures that a larger proportion of larger establishments is selected compared to smaller establishments.

### Estimation

Estimated data items are produced as the sum of weighted establishment data. Proportions of estimated data items are derived using the appropriate estimated totals. Variances are estimated using the method of random groups and are used to determine if measured changes are statistically significant.

### Weighting and Unit Nonresponse

The weight for a given establishment is the product of two factors: the inverse of its probability of selection into the CFS (EHM) sample and a nonresponse adjustment factor.

An EHM unit (establishment) is considered a "response" if they indicate whether they ship HAZMAT, and if they do ship HAZMAT, if they provide packaging details for at least one of their most frequently shipped HAZMATs. Unit nonresponse occurs when this response criterion is not met and is handled in the estimation procedure by a nonresponse adjustment. The sample is divided into mutually exclusive adjustment cells defined by NAICS code and PADD. A nonresponse adjustment factor is then computed for each adjustment cell and is equal to the ratio of establishments selected into the sample to the number of responses received within each cell.

# Item Nonresponse

Item nonresponse occurs either when a particular question is unanswered or the response to the question fails computer or analyst edits. Nonresponse to the "Total Number of HAZMATs (UN/NAs) Shipped" question is corrected by imputation and uses information provided for the top ten most frequent types of HAZMAT shipped (referred to as the respondent's "HAZMAT listing").

If the respondent provides between one and nine United Nations or North American (UN/NA) hazardous material

codes in their HAZMAT listing, the total number of HAZMATs (UN/NAs) shipped is imputed to equal this count. If the respondent provides exactly ten UN/NA numbers in their HAZMAT listing, the imputed value is instead calculated with an imputation cell method. Here, the sample is divided into mutually exclusive cells defined by publication NAICS code; for each cell, the average total number of HAZMATs (UN/NAs) shipped is computed among respondents that reported exactly ten UN/NA numbers in their HAZMAT listing. Missing values are then replaced with the appropriate average values.

For all other data items, no imputation is performed. Instead, separate estimates are published in a "Did Not Respond" or equivalent category. For example, a respondent who did not indicate the type of outer packaging used to ship a particular HAZMAT would be included in the estimate for the "No Outer Packaging Type Reported" category.

# Sampling Error and Suppression

The sampling error of an estimate based on a sample survey is the difference between the estimate and the result that would be obtained from a complete census conducted under the same survey conditions. This error occurs because characteristics differ among sampling units in the population and only a subset of the population is measured in a sample survey. The particular sample used in this survey is one of a large number of samples of the same size that could have been selected using the same sample design. Because each unit in the sampling frame had a known probability of being selected into the sample, it was possible to estimate the sampling variability of the survey estimates.

Common measures of the variability among these estimates are the sampling variance, the standard error, and the coefficient of variation (CV), which is also referred to as the relative standard error. The sampling variance is defined as the squared difference, averaged over all possible samples of the same size and design, between the estimator and its average value. The standard error is the square root of the sampling variance. The CV expresses the standard error as a percentage of the estimate to which it refers. For example, an estimate of 200 units that has an estimated standard error of 10 units has an estimated CV of 5 percent. The sampling variance, standard error, and CV of an estimate can be estimated from the selected sample because the sample was selected using probability sampling. Note that measures of sampling variability, such as the standard error and CV, are estimated from the sample and are also subject to sampling variability. It is also important to note that the standard error and CV only measure sampling variability. They do not measure any systematic biases in the estimates.

The Census Bureau recommends that individuals using these estimates incorporate sampling error information into their analyses, as this could affect the conclusions drawn from the estimates.

Variances for published statistics are estimated using the method of random groups, with ten random groups. To implement this method of variance estimation, a random group number between one and ten is assigned to each sampling unit not selected with certainty. Certainty units are not considered part of any of the ten groups. Variance estimates are then computed by calculating ten replicate totals, which estimate a given statistic using only members of one of the random groups, and the certainties and calculating the variability in the ten replicate totals.

An estimate from a particular sample and the standard error associated with the estimate can be used to construct a confidence interval. A confidence interval is a range about a given estimator that has a specified probability of containing the result of a complete enumeration of the sampling frame conducted under the same survey conditions. Associated with each interval is a percentage of confidence, which is interpreted as follows. If, for each possible sample, an estimate of a population parameter and its appropriate standard error were obtained, then:

- For approximately 90 percent of the possible samples, the interval from 1.833 standard errors below the estimate to 1.833 standard errors above the estimate would include the result as obtained from a complete enumeration of the sampling frame conducted under the same survey conditions.
- For approximately 95 percent of the possible samples, the interval from 2.262 standard errors below the estimate to 2.262 standard errors above the estimate would include the result as obtained from a complete enumeration of the sampling frame conducted under the same survey conditions.

The 1.833 and 2.262 values, used to compute the 90 percent and 95 percent confidence intervals, are taken from the t-distribution with nine degrees of freedom (one less than the number of random groups used to produce the CV estimates). The t-distribution takes into account the uncertainty in the estimation of the CVs.

To illustrate the computation of a confidence interval for an estimate of the total number of establishments with HAZMAT shipments, assume that an estimate of the total is 10,750 and the coefficient of variation for this estimate is 1.8 percent, or 0.018. First obtain the standard error of the estimate by multiplying the value of the total estimate by its CV. For this example, multiply 10,750 by 0.018. This yields a standard error of 193.5. The upper and lower bounds of the 90 percent confidence interval are computed as 10,750 plus or minus 1.833 times 193.5, or 354.7. Consequently, the 90 percent confident interval is 10,395 to 11,105. If corresponding confidence intervals were constructed for all possible samples of the same size and design, approximately 9 out of 10 (90 percent) of these intervals would contain the result obtains from a complete enumeration.

In this publication, estimates that have a high CV (greater than 50 percent) are suppressed. Some of these suppressed estimates can be derived directly from the EHM tables by subtracting published estimates from their respective totals. However, the suppressed estimates obtained by such subtraction would be subject to poor response, high sampling variability, or other factors that may make them potentially misleading. Estimates derived in this manner should not be attributed to the Census Bureau.

For quality purposes, estimates about HAZMAT estimated to be shipped by fewer than 500 establishments nationally are not published on any table.

# Nonsampling Error

Nonsampling error encompasses all factors other than sampling error that contribute to the total error associated with an estimate. This error may also be present in censuses and other nonsurvey programs. Nonsampling error arises from many sources: inability to obtain information on all units in the sample; response errors; differences in the interpretation of the questions; mismatches between sampling units and reporting units, requested data and data available or accessible in respondents' records, or with regard to reference periods; mistakes in coding or keying the data obtained; and other errors of collection, response, coverage, and processing.

Although no direct measurement of nonsampling error was obtained, precautionary steps were taken in all phases of the collection, processing, and tabulation of the data in an effort to minimize its influence. Precise estimation of the magnitude of nonsampling errors would require special experiments or access to independent data and, consequently, the magnitudes are often unavailable.

The Census Bureau recommends that individuals using these estimates factor in this information when assessing their analyses of these data, as nonsampling error could affect the conclusions drawn from the estimates.

The Unit Response Rate (URR) is defined as the ratio (expressed as a percentage) of the total unweighted number of establishments that provided usable data to the total number of establishments that were eligible (or potentially eligible) for data collection. URRs are indicators of the performance of the data collection process in obtaining usable responses. The overall URR is 47.7 percent for data collection year 2021 with a similar rate for data collection year 2022. The URR by PADD and by NAICS can be found in Table 1 and Table 2, respectively.

#### Table 1. Response Rates by PADD<sup>1</sup>

(In percent)

PADD	Unit response rate
PADD_I	45.8
PADD_II	53.1
PADD_III	42.5
PADD_IV_V	46.7

<sup>1</sup> Petroleum Administration for Defense District.

Source: 2022 Commodity Flow Survey, Expanded Hazardous Materials Supplement.

#### Table 2.

#### **Response Rates by NAICS<sup>1</sup>**

(In percent)

NAICS	Unit response rate
212	44.5
311	49.9
312	50.8
313	56.7
314	53.2
315	45.7
316	54.4
321	54.3
322	55.8
323	54.0
324	46.5
325	50.7
326	55.2
327	49.0
331	60.6
332	58.2
333	56.4
334	52.8
335	56.9
336	53.1
337	53.6
339	52.8
4231	42.0
4232	47.9
4233	41.6
4234	46.2
4235	56.5
4236	42.2
4237	46.3
4238	48.1
4239	46.5
4241	50.2
4242	38.3
4243	48.3
4244	44.4
4245	49.8
4246	41.3
4247	43.6
4248	59.9
4249	42.3
4541	49.3
45431	60.0
4931	30.5
5111	49.0
551114	41.8

<sup>1</sup> North American Industry Classification System.

Source: 2022 Commodity Flow Survey, Expanded Hazardous Materials Supplement.

#### **Disclosure Avoidance**

Disclosure is the release of data that reveals information or permits deduction of information about a particular survey unit through the release of either tables or microdata. Disclosure avoidance is the process used to protect each survey unit's identity and data from disclosure. Using disclosure avoidance procedures, the Census Bureau modifies or removes the characteristics that put information at risk of disclosure. Although it may appear that a table shows information about a specific survey unit, the Census Bureau has taken steps to disguise or suppress a unit's data that may be "at risk" of disclosure while making sure the results are still useful.

Cell suppression is a disclosure avoidance technique that protects the confidentiality of individual survey units by withholding cell values from release. The cells that must be protected are called primary suppressions. To make sure the cell values of the primary suppressions cannot be closely estimated by using other published cell values, additional cells may also be suppressed. These additional suppressed cells are called complementary suppressions. For this EHM publication, disclosure-related suppressions are not presented in tables.

The process of suppression does not usually change the higher-level totals. Values for cells that are not suppressed remain unchanged. Before the Census Bureau releases data, computer programs and analysts ensure primary and complementary suppressions have been correctly applied.

#### **Table Notes**

The estimates presented are based on data from the EHM as of October 18, 2023. Estimates may not be additive due to rounding. To maintain confidentiality, the Census Bureau suppresses data to protect the identity of any business or individual. Estimates are based on a sample of establishments and are subject to both sampling and nonsampling error. Data users who create their own estimates using data from this file should cite the Census Bureau as the source of the original data only.

# TABLES

This experimental data product includes Excel files with EHM estimates for 2021 and 2022. Each file includes 11 tables, with estimates at the U.S. and PADD geographic levels. The following list describes the tables available with this data product. The tables are located on the EDP webpage at <www.census.gov/data/experimental-dataproducts/cfs-expanded-hazardous-materials-estimates. html>.

# Table 1: Number of Establishments That Ship AnyHazardous Materials by Industry and Geography

Columns A and B: Geography and Description of Geography (PADD). PADD detail is provided only for industry total estimates.

Columns C and D: 2017 NAICS Code and Meaning of NAICS Code. NAICS is provided at selected two-, three-, four-, five-, and six-digit levels.

Column E: Number of HAZMATs (UN/NAs) Shipped From Establishment. These metrics are presented as categorical whole numbers in increments of five or more.

Column F: Year. The reference year for the data.

Column G: Number of Establishments With HAZMAT Shipments. The number of establishments that reported shipping at least one HAZMAT shipment (with corresponding UN/NA) for the reference year.

Column H: Proportion of Establishments Shipping HAZMAT. The calculated proportion of the number of establishments with HAZMAT shipments by the number of total establishments for a geography or industry. This estimate is provided only for the total number of HAZMATs shipped and is presented as a range.

Column I: Average Number of HAZMATs (UN/NAs) Shipped Among HAZMAT Shippers. This estimate represents an average taken of the reported total number of HAZMATs shipped for each tabulation level. This estimate is provided only for the total number of HAZMATs shipped.

Column J: Coefficient of Variation for Number of Establishments With HAZMAT Shipments. CV is a measure of sampling error that represents the ratio of the standard deviation to the estimate. Data users should exercise caution when using estimates with high CVs. Estimates with a CV greater than 50 percent are suppressed due to high sampling variability and are represented by a "S."

Column K: Standard Error for Proportion of Establishments Shipping HAZMAT. Data users should exercise caution when using estimates with high standard errors. These data are being provided because aggregates of the tabulated estimates can be useful.

Column L: Coefficient of Variation for Average Number of HAZMATs (UN/NAs) Shipped Among HAZMAT Shippers. CV is a measure of sampling error that represents the ratio of the standard deviation to the estimate. Data users should exercise caution when using estimates with high CVs. Estimates with a CV greater than 50 percent are suppressed due to high sampling variability and are represented by a "S."

# Table 2: Number of Establishments That Ship Selected Hazardous Materials by Industry and Geography

Columns A and B: Geography and Description of Geography (PADD). PADD detail is provided only for industry total estimates.

Columns C and D: 2017 NAICS Code and Meaning of NAICS Code. NAICS is provided at selected two-, three-, four-, five-, and six-digit levels.

Columns E and F: HAZMAT Code and HAZMAT Description. HAZMAT code is designated by UN/NA codes and class/ division categories. HAZMAT codes starting with a "D" denotes classes or divisions and "NA" or "UN" denotes individual UN/NA codes. The reference for the HAZMAT description is the hazardous materials table found in Title 49 of the Code of Federal Regulations, Section § 172.101 at <www.ecfr.gov/current/title-49/subtitle-B/chapter-I/ subchapter-C/part-172/subpart-B/section-172.101>. Descriptions were consolidated when multiple instances of a UN/NA code were present.

Column G: Year. The reference year for the data.

Column H: Number of Establishments. This is an estimate of the number of establishments that reported shipping a type of hazardous material (HAZMAT Code) for the reference year.

Column I: Proportion of HAZMAT Shipping Establishments Shipping This Material. The calculated proportion of the number of establishments shipping a specified HAZMAT code divided by the number of establishments shipping any HAZMAT code for a geography or industry. This column is presented as a percentage.

Column J: Coefficient of Variation for Number of Establishments Shipping This Material. CV is a measure of sampling error that represents the ratio of the standard deviation to the estimate. Data users should exercise caution when using estimates with high CVs. Estimates with a CV greater than 50 percent are suppressed due to high sampling variability and are represented by a "S."

Column K: Standard Error for Proportion of HAZMAT Shipping Establishments Shipping This Material. Data users should exercise caution when using estimates with high standard errors.

# Table 3: Outer Packaging Most Frequently Used byIndustry, Geography, and Hazardous Material

Columns A and B: Geography and Description of Geography (PADD). PADD detail is provided only for industry total estimates.

Columns C and D: 2017 NAICS Code and Meaning of NAICS Code. NAICS is provided at selected two-, three-, four-, five-, and six-digit levels. Column E and F: HAZMAT Code and HAZMAT Description. HAZMAT code is designated by UN/NA codes and class/ division categories. HAZMAT codes starting with a "D" denotes classes or divisions and "NA" or "UN" denotes individual UN/NA codes. The reference for the HAZMAT description is the hazardous materials table found in Title 49 of the Code of Federal Regulations, Section § 172.101 at <www.ecfr.gov/current/title-49/subtitle-B/chapter-I/ subchapter-C/part-172/subpart-B/section-172.101>. Descriptions were consolidated when multiple instances of a UN/NA code were present.

Column G: Outer Packaging. Refers to the reported outer packaging in which the establishments shipped a specified material or HAZMAT code.

Column H: Year. The reference year for the data.

Column I: Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently. These are among establishments reporting a single outer packaging most frequently used for a given hazardous material.

Column J: Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material. This column is the calculated proportion of the number of establishments using a specified outer packaging type shipping a specified HAZMAT code divided by the number of establishments in any outer packaging for a geography or industry.

Column K: Coefficient of Variation for Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently. CV is a measure of sampling error that represents the ratio of the standard deviation to the estimate. Data users should exercise caution when using estimates with high CVs. Estimates with a CV greater than 50 percent are suppressed due to high sampling variability and are represented by a "S."

Column L: Standard Error for Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material. Data users should exercise caution when using estimates with high standard errors.

## Table 4: Annual Shipment Quantity for Performance-Oriented Packaging by Industry, Geography, Hazardous Material, and Outer Packaging Most Frequently Used

Columns A and B: Geography and Description of Geography (PADD). PADD detail is provided only for industry total estimates.

Columns C and D: 2017 NAICS Code and Meaning of NAICS Code. NAICS is provided at selected two-, three-, four-, five-, and six-digit levels.

Column E and F: HAZMAT Code and HAZMAT Description. HAZMAT code is designated by UN/NA codes and class/ division categories. HAZMAT codes starting with a "D" denotes classes or divisions and "NA" or "UN" denotes individual UN/NA codes. The reference for the HAZMAT description is the hazardous materials table found in Title 49 of the Code of Federal Regulations, Section § 172.101 at <www.ecfr.gov/current/title-49/subtitle-B/chapter-I/ subchapter-C/part-172/subpart-B/section-172.101>. Descriptions were consolidated when multiple instances of a UN/NA code were present.

Column G: Outer Packaging. Refers to the reported outer packaging in which the establishments shipped a specified material or HAZMAT code. Performance-oriented packaging types only.

Column H: Annual Shipment Quantity. These metrics are presented as number of packages (annual) in categorical whole numbers of varying increments.

Column I: Year. The reference year for the data.

Column J: Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently by Annual Shipment Quantity. These are among establishments reporting a single outer packaging most frequently used for a given hazardous material and annual shipment quantity.

Column K: Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material by Annual Shipment Quantity. This column is the calculated proportion of the number of establishments using a specified outer packaging type, shipping a specified HAZMAT code, and shipping a specified annual shipment quantity divided by the number of establishments with any annual shipment quantity for an outer packaging and geography or industry.

Column L: Coefficient of Variation for Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently by Annual Shipment Quantity. CV is a measure of sampling error that represents the ratio of the standard deviation to the estimate. Data users should exercise caution when using estimates with high CVs. Estimates with a CV greater than 50 percent are suppressed due to high sampling variability and are represented by a "S".

Column M: Standard Error for Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material by Annual Shipment Quantity. Data users should exercise caution when using estimates with high standard errors.

#### Table 5: Average Net Weight for Performance-Oriented Packaging by Industry, Geography, Hazardous Material, and Outer Packaging Most Frequently Used

Columns A and B: Geography and Description of Geography (PADD). PADD detail is provided only for industry total estimates.

Columns C and D: 2017 NAICS Code and Meaning of NAICS Code. NAICS is provided at selected two-, three-, four-, five-, and six-digit levels.

Column E and F: HAZMAT Code and HAZMAT Description. HAZMAT code is designated by UN/NA codes and class/ division categories. HAZMAT codes starting with a "D" denotes classes or divisions and "NA" or "UN" denotes individual UN/NA codes. The reference for the HAZMAT description is the hazardous materials table found in Title 49 of the Code of Federal Regulations, Section § 172.101 at <www.ecfr.gov/current/title-49/subtitle-B/chapter-I/ subchapter-C/part-172/subpart-B/section-172.101>. Descriptions were consolidated when multiple instances of a UN/NA code were present.

Column G: Outer Packaging. Refers to the reported outer packaging in which the establishments shipped a specified material or HAZMAT code. Performance-oriented packaging types only.

Column H: Average Net Weight. These metrics are presented as averages of weight (in pounds) per package in categorical whole numbers of varying increments.

Column I: Year. The reference year for the data.

Column J: Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently by Average Net Weight. These are among establishments reporting a single outer packaging most frequently used for a given hazardous material and average net weight.

Column K: Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material by Average Net Weight. This column is the calculated proportion of the number of establishments using a specified outer packaging type, shipping a specified HAZMAT code, and shipping a specified average net weight divided by the number of establishments with any average net weight for an outer packaging and geography or industry.

Column L: Coefficient of Variation for Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently by Average Net Weight. CV is a measure of sampling error that represents the ratio of the standard deviation to the estimate. Data users should exercise caution when using estimates with high CVs. Estimates with a CV greater than 50 percent are suppressed due to high sampling variability and are represented by a "S." Column M: Standard Error for Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material by Average Net Weight. Data users should exercise caution when using estimates with high standard errors.

# Table 6: Primary Mode of Transportation for PerformanceOriented Packaging by Industry, Geography, HazardousMaterial, and Outer Packaging Most Frequently Used

Columns A and B: Geography and Description of Geography (PADD). PADD detail is provided only for industry total estimates.

Columns C and D: 2017 NAICS Code and Meaning of NAICS Code. NAICS is provided at selected two-, three-, four-, five-, and six-digit levels.

Column E and F: HAZMAT Code and HAZMAT Description. HAZMAT code is designated by UN/NA codes and class/ division categories. HAZMAT codes starting with a "D" denotes classes or divisions and "NA" or "UN" denotes individual UN/NA codes. The reference for the HAZMAT description is the hazardous materials table found in Title 49 of the Code of Federal Regulations, Section § 172.101 at <ww.ecfr.gov/current/title-49/subtitle-B/chapter-I/ subchapter-C/part-172/subpart-B/section-172.101>. Descriptions were consolidated when multiple instances of a UN/NA code were present.

Column G: Outer Packaging. Refers to the reported outer packaging in which the establishments shipped a specified material or HAZMAT code. Performance oriented packaging types only.

Column H: Primary Mode of Transportation. These metrics are presented as categories of types of transportation used primarily to ship specified materials.

Column I: Year. The reference year for the data.

Column J: Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently by Primary Mode of Transportation. These are among establishments reporting a single outer packaging most frequently used for a given hazardous material and primarily using a specified mode of transportation.

Column K: Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material by Primary Mode of Transportation. This column is the calculated proportion of the number of establishments using a specified outer packaging type, shipping a specified HAZMAT code, and using a specified primary mode of transportation divided by the number of establishments using any primary mode of transportation for an outer packaging and geography or industry. Column L: Coefficient of Variation for Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently by Primary Mode of Transportation. CV is a measure of sampling error that represents the ratio of the standard deviation to the estimate. Data users should exercise caution when using estimates with high CVs. Estimates with a CV greater than 50 percent are suppressed due to high sampling variability and are represented by a "S."

Column M: Standard Error for Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material by Primary Mode of Transportation. Data users should exercise caution when using estimates with high standard errors.

# Table 7: Packaging Type Used for Performance Oriented Packaging by Industry, Geography, Hazardous Material, and Outer Packaging Most Frequently Used

Columns A and B: Geography and Description of Geography (PADD). PADD detail is provided only for industry total estimates.

Columns C and D: 2017 NAICS Code and Meaning of NAICS Code. NAICS is provided at selected two-, three-, four-, five-, and six-digit levels.

Column E and F: HAZMAT Code and HAZMAT Description. HAZMAT code is designated by UN/NA codes and class/ division categories. HAZMAT codes starting with a "D" denotes classes or divisions and "NA" or "UN" denotes individual UN/NA codes. The reference for the HAZMAT description is the hazardous materials table found in Title 49 of the Code of Federal Regulations, Section § 172.101 at <www.ecfr.gov/current/title-49/subtitle-B/chapter-I/ subchapter-C/part-172/subpart-B/section-172.101>. Descriptions were consolidated when multiple instances of a UN/NA code were present.

Column G: Outer Packaging. Refers to the reported outer packaging in which the establishments shipped a specified material or HAZMAT code. Performance oriented packaging types only.

Column H: Packaging Type. These metrics are presented as categories of types of packaging types and material used primarily to ship specified outer packaging materials.

Column I: Year. The reference year for the data.

Column J: Number of Establishments. These are among establishments reporting a single outer packaging most frequently used for a given hazardous material and primarily using a packaging type or material.

Column K: Proportion of Establishments Using the Specified Packaging Type Among Establishments Shipping This Material in This Outer Packaging. This column is the calculated proportion of the number of establishments using a specified outer packaging type, shipping a specified HAZMAT code, and using a specified packaging type divided by the number of establishments using any packaging type for an outer packaging and geography or industry.

Column L: Coefficient of Variation for Number of Establishments. CV is a measure of sampling error that represents the ratio of the standard deviation to the estimate. Data users should exercise caution when using estimates with high CVs. Estimates with a CV greater than 50 percent are suppressed due to high sampling variability and are represented by a "S."

Column M: Standard Error for Proportion of Establishments Using the Specified Packaging Type Among Establishments Shipping This Material in This Outer Packaging. Data users should exercise caution when using estimates with high standard errors.

# Table 8: Specification Marking Groups for SpecificationPackaging by Industry, Geography, Hazardous Material,and Outer Packaging Most Frequently Used

Columns A and B: Geography and Description of Geography (PADD). PADD detail is provided only for industry total estimates.

Columns C and D: 2017 NAICS Code and Meaning of NAICS Code. NAICS is provided at selected two-, three-, four-, five-, and six-digit levels.

Column E and F: HAZMAT Code and HAZMAT Description. HAZMAT code is designated by UN/NA codes and class/ division categories. HAZMAT codes starting with a "D" denotes classes or divisions and "NA" or "UN" denotes individual UN/NA codes. The reference for the HAZMAT description is the hazardous materials table found in Title 49 of the Code of Federal Regulations, Section § 172.101 at <www.ecfr.gov/current/title-49/subtitle-B/chapter-I/ subchapter-C/part-172/subpart-B/section-172.101>. Descriptions were consolidated when multiple instances of a UN/NA code were present.

Column G: Outer Packaging. Refers to the reported outer packaging in which the establishments shipped a specified material or HAZMAT code. Specification Packaging types only.

Column H: Specification Marking Group. These metrics are presented as groups of specification markings used primarily to ship specified materials.

Column I: Year. The reference year for the data.

Column J: Number of Establishments. These are among establishments reporting a single outer packaging most frequently used for a given hazardous material and using a specified specification marking group. Column K: Proportion of Establishments Using the Specified Specification Marking Group Among Establishments Shipping This Material in This Outer Packaging. This column is the calculated proportion of the number of establishments using a specified outer packaging type, shipping a specified HAZMAT code, and using a specified specification marking group divided by the number of establishments using any specification marking group for an outer packaging and geography or industry.

Column L: Coefficient of Variation for Number of Establishments. CV is a measure of sampling error that represents the ratio of the standard deviation to the estimate. Data users should exercise caution when using estimates with high CVs. Estimates with a CV greater than 50 percent are suppressed due to high sampling variability and are represented by a "S."

Column M: Standard Error for Proportion of Establishments Using the Specified Specification Marking Group Among Establishments Shipping This Material in This Outer Packaging. Data users should exercise caution when using estimates with high standard errors.

## Table 9: Annual Shipment Quantity for Specification Packaging by Industry, Geography, Hazardous Material, and Outer Packaging Most Frequently Used

Columns A and B: Geography and Description of Geography (PADD). PADD detail is provided only for industry total estimates.

Columns C and D: 2017 NAICS Code and Meaning of NAICS Code. NAICS is provided at selected two-, three-, four-, five-, and six-digit levels.

Column E and F: HAZMAT Code and HAZMAT Description. HAZMAT code is designated by UN/NA codes and class/ division categories. HAZMAT codes starting with a "D" denotes classes or divisions and "NA" or "UN" denotes individual UN/NA codes. The reference for the HAZMAT description is the hazardous materials table found in Title 49 of the Code of Federal Regulations, Section § 172.101 at <www.ecfr.gov/current/title-49/subtitle-B/chapter-I/ subchapter-C/part-172/subpart-B/section-172.101>. Descriptions were consolidated when multiple instances of a UN/NA code were present.

Column G: Outer Packaging. Refers to the reported outer packaging in which the establishments shipped a specified material or HAZMAT code. Specification Packaging types only.

Column H: Annual Shipment Quantity. These metrics are presented as number of packages (annual) in categorical whole numbers of varying increments.

Column I: Year. The reference year for the data.

Column J: Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently by Annual Shipment Quantity. These are among establishments reporting a single outer packaging most frequently used for a given hazardous material and annual shipment quantity.

Column K: Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material by Annual Shipment Quantity. This column is the calculated proportion of the number of establishments using a specified outer packaging type, shipping a specified HAZMAT code, and shipping a specified annual shipment quantity divided by the number of establishments with any annual shipment quantity for an outer packaging and geography or industry.

Column L: Coefficient of Variation for Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently by Annual Shipment Quantity. CV is a measure of sampling error that represents the ratio of the standard deviation to the estimate. Data users should exercise caution when using estimates with high CVs. Estimates with a CV greater than 50 percent are suppressed due to high sampling variability and are represented by a "S."

Column M: Standard Error for Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material by Annual Shipment Quantity. Data users should exercise caution when using estimates with high standard errors.

# Table 10: Average Net Weight for Specification Packaging by Industry, Geography, Hazardous Material, and Outer Packaging Most Frequently Used

Columns A and B: Geography and Description of Geography (PADD). PADD detail is provided only for industry total estimates.

Columns C and D: 2017 NAICS Code and Meaning of NAICS Code. NAICS is provided at selected two-, three-, four-, five-, and six-digit levels.

Column E and F: HAZMAT Code and HAZMAT Description. HAZMAT code is designated by UN/NA codes and class/ division categories. HAZMAT codes starting with a "D" denotes classes or divisions and "NA" or "UN" denotes individual UN/NA codes. The reference for the HAZMAT description is the hazardous materials table found in Title 49 of the Code of Federal Regulations, Section § 172.101 at <www.ecfr.gov/current/title-49/subtitle-B/chapter-I/ subchapter-C/part-172/subpart-B/section-172.101>. Descriptions were consolidated when multiple instances of a UN/NA code were present.

Column G: Outer Packaging. Refers to the reported outer packaging in which the establishments shipped

a specified material or HAZMAT code. Specification Packaging types only.

Column H: Average Net Weight. These metrics are presented as averages of weight (in pounds) per package in categorical whole numbers of varying increments.

Column I: Year. The reference year for the data.

Column J: Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently by Average Net Weight. These are among establishments reporting a single outer packaging most frequently used for a given hazardous material and average net weight.

Column K: Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material by Average Net Weight. This column is the calculated proportion of the number of establishments using a specified outer packaging type, shipping a specified HAZMAT code, and shipping a specified average net weight divided by the number of establishments with any average net weight for an outer packaging and geography or industry.

Column L: Coefficient of Variation for Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently by Average Net Weight. CV is a measure of sampling error that represents the ratio of the standard deviation to the estimate. Data users should exercise caution when using estimates with high CVs. Estimates with a CV greater than 50 percent are suppressed due to high sampling variability and are represented by a "S."

Column M: Standard Error for Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material by Average Net Weight. Data users should exercise caution when using estimates with high standard errors.

## Table 11: Primary Mode of Transportation for Specification Packaging by Industry, Geography, Hazardous Material, and Outer Packaging Most Frequently Used

Columns A and B: Geography and Description of Geography (PADD). PADD detail is provided only for industry total estimates.

Columns C and D: 2017 NAICS Code and Meaning of NAICS Code. NAICS is provided at selected two-, three-, four-, five-, and six-digit levels.

Column E and F: HAZMAT Code and HAZMAT Description. HAZMAT code is designated by UN/NA codes and class/ division categories. HAZMAT codes starting with a "D" denotes classes or divisions and "NA" or "UN" denotes individual UN/NA codes. The reference for the HAZMAT description is the hazardous materials table found in Title 49 of the Code of Federal Regulations, Section § 172.101 at <www.ecfr.gov/current/title-49/subtitle-B/chapter-I/ subchapter-C/part-172/subpart-B/section-172.101>. Descriptions were consolidated when multiple instances of a UN/NA code were present.

Column G: Outer Packaging. Refers to the reported outer packaging in which the establishments shipped a specified material or HAZMAT code. Specification Packaging types only.

Column H: Primary Mode of Transportation. These metrics are presented as categories of types of transportation used primarily to ship specified materials.

Column I: Year. The reference year for the data.

Column J: Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently by Primary Mode of Transportation. These are among establishments reporting a single outer packaging most frequently used for a given hazardous material and primarily using a specified mode of transportation.

Column K: Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material by Primary Mode of Transportation. This column is the calculated proportion of the number of establishments using a specified outer packaging type, shipping a specified HAZMAT code, and using a specified primary mode of transportation divided by the number of establishments using any primary mode of transportation for an outer packaging and geography or industry.

Column L: Coefficient of Variation for Number of Establishments Using the Outer Packaging to Ship This Material Most Frequently by Primary Mode of Transportation. CV is a measure of sampling error that represents the ratio of the standard deviation to the estimate. Data users should exercise caution when using estimates with high CVs. Estimates with a CV greater than 50 percent are suppressed due to high sampling variability and are represented by a "S."

Column M: Standard Error for Proportion of Establishments Using the Outer Packaging to Ship This Material Most Frequently Among Establishments Shipping This Material by Primary Mode of Transportation. Data users should exercise caution when using estimates with high standard errors.