## Explanation of the United States Census Bureau's Supplemental Quarterly ECommerce Data Product ${ }^{1}$

## 1 Introduction

A dynamic retail economy garners great attention every day as it evolves to meet consumers’ changing needs and to adapt to emerging technologies. The evolution of e-commerce is a particularly noteworthy topic to retail industry experts, analysts, policymakers, and consumers. As a result, there is increased interest in the Census Bureau's retail data products that measure the e-commerce component of the retail economy. In order to meet the needs of its data users, the United States Census Bureau has made it a priority to:

1. Evaluate the current definition of e-commerce.
2. Produce additional e-commerce data products where quality standards and resources allow.

The first priority is a long-term goal. A group of Census Bureau retail and classification subject matter experts are working with retail industry experts to better understand how different retail industry stakeholders-be it retailers, real estate developers, financial institutions, or retail trade groups, among others-define and track e-commerce.

The second priority is a short-term goal. A Census Bureau E-Commerce Data Analysis Team was tasked with researching the feasibility of using already collected data to produce one or more supplemental e-commerce tables. These tables would be published in conjunction with the Quarterly E-Commerce Report. The Census Bureau currently produces the following ecommerce data products on a monthly or quarterly basis:

- On a monthly basis, the Monthly Retail Trade Report includes estimates for companies classified by the North American Industry Classification System (NAICS) as Nonstore Retailers (NAICS 454) and a subset of those companies in NAICS 4541 called Electronic Shopping and Mail-Order Houses. However, other retailers that may not be conducting ecommerce, including electronic auctions and mail-order houses, are also included in these estimates.
- On a quarterly basis, the Quarterly E-Commerce Report is published. This report provides a single quarterly estimate measuring all retail e-commerce sales captured by the Monthly Retail Trade Survey (MRTS).

The purpose of this document is to share the findings of the work done on the second priority. Using existing MRTS data, a quarterly supplemental e-commerce table has been developed to provide more granular e-commerce sales estimates.

[^0]
## 2 Overview of Census Retail Programs

The retail trade program currently covers retail companies as defined by NAICS and represents all retail companies (NAICS Sector 44-45) with and without paid employees. These retail businesses may be large retailers with many store locations, single-unit retailers with only one location, or retailers operating solely as e-commerce businesses.

The Census Bureau measures the retail economy every five years in the Economic Census and on a more frequent basis in monthly and annual surveys. In years ending in " 2 " and " 7 ", the Economic Census asks for detailed sales and product-level information as well as employment and payroll and business characteristics for each physical store location that a retailer operates. Response to the Economic Census is mandatory and the data collected are used to update the Census Bureau's Business Register from which the sampling frames for many Economic Directorate surveys - including the annual and monthly retail trade surveys - are created. Each year, the Annual Retail Trade Survey (ARTS) collects data at the company or retailer level nationally; no store location data are collected. The ARTS collects annual sales, e-commerce sales, end-of-year inventories, and expenses data as well as some retailer characteristics; the annual data are released approximately 15 months after the data year ends.

The MRTS is a voluntary survey conducted at the retailer or company level and collects sales/receipts as well as end-of-month inventories and e-commerce sales from all retail industries. Estimates from this survey are released approximately six weeks after month's end.

The MRTS is a subsample of the ARTS and a selection of the MRTS sample occurs approximately every five years to ensure the sample remains representative and to redistribute the burden for small and mid-size businesses. A stratified simple random sampling method is used to select approximately 10,800 retail firms whose sales are then weighted and benchmarked to represent the complete universe of retail firms. The MRTS sample is probability based and represents all employer firms engaged in retail activities as defined by NAICS. Coverage includes all retailers whether or not they are engaged in e-commerce. Online travel services, financial brokers and dealers, and ticket sales agencies are not classified as retail and are not included in either the total retail or retail e-commerce sales estimates. Nonemployers are represented in the estimates through benchmarking to prior annual survey estimates that include nonemployer sales based on administrative records and imputation and e-commerce sales through imputation. E-commerce sales are included in the total monthly sales estimates.

The MRTS sample is updated on an ongoing basis to account for new retail employer businesses (including those selling via the Internet), business deaths, and other changes to the retail business universe. Firms are asked each month to report e-commerce sales separately. For each month of the quarter, data for nonresponding companies are imputed from responding companies falling within the same kind of business and sales size category. Responding firms account for approximately 67 percent of the e-commerce sales estimate and about 72 percent of the estimate of U.S. retail sales for any quarter.

Retailers are sampled by primary kind-of-business as classified by NAICS code. However, if a retailer engages in more than one kind-of-business (e.g, food and beverage stores, general
merchandise stores, etc), it can have more than one tabulation part. This means that a company contributes to estimates of each of the different kinds-of-business it conducts. Note that in general, companies with a tabulation part in NAICS 4541 may have multiple kinds of brick-andmortar businesses, but all e-commerce sales are reported under one NAICS 4541 part.

Retail e-commerce sales are estimated from the same sample used for the MRTS to estimate preliminary and final U.S. retail sales.

## 3 Project Objective

The core objective of the E-Commerce Data Analysis Team is to analyze existing e-commerce data and determine if the data and data quality would allow for publishing more granular ecommerce sales. ${ }^{2}$

This effort mimics the effort that led to the publication of an annual supplemental e-commerce table released as part of the 2015 Annual Retail Trade Survey (ARTS). ARTS has long published a table that contains total sales and total e-commerce sales for each 3-digit NAICS code.

Figure 1 shows an example of this annual data presentation for 2016 and 2017.

| NAICS Code | Kind of Business | 2017 |  | $2016{ }^{\text {r }}$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Total | E-commerce | Total | E-commerce |
|  | Total Retail Trade | 5,046,894 | 461,034 | 4,851,774 | 397,307 |
| 441 | Motor vehicle and parts dealers | 1,174,417 | 34,273 | 1,142,261 | 31,913 |
| 442 | Furniture and home furnishings stores | 113,783 | 1,335 | 110,695 | 1,122 |
| 443 | Electronics and appliance stores | 99,401 | 2,078 | 99,297 | 1,954 |
| 444 | Building mat. and garden equip. and supplies dealers | 365,651 | 2,583 | 349,372 | 2,260 |
| 445 | Food and beverage stores | 725,915 | 3,802 | 699,362 | 2,435 |
| 446 | Health and personal care stores | 333,219 | S | 327,031 | D |
| 447 | Gasoline stations | 452,856 | S | 418,684 | D |
| 448 | Clothing and clothing access. stores | 258,472 | 9,898 | 259,840 | 8,854 |
| 451 | Sporting goods, hobby, musical instrument, and book stores | 84,264 | 2,423 | 86,562 | 2,288 |
| 452 | General merchandise stores | 683,854 | 453 | 675,374 | 362 |
| 453 | Miscellaneous store retailers | 125,500 | 4,053 | 121,740 | 3,766 |
| 454 | Nonstore retailers | 629,562 | 399,178 | 561,556 | 341,430 |
| 4541 | Electronic shopping and mail-order houses | 552,214 | 397,490 | 488,619 | 339,928 |

D - Denotes an estimate withheld to avoid disclosing data of individual companies; data are included in higher-level totals; S - Suppressed
Figure 1: Estimated Annual U.S. Retail Trade Sales - Total and E-commerce: 2016-2017 Estimates are shown in millions of dollars and are based on data from the Annual Retail Trade Survey.

## Source: https://www2.census.gov/programs-surveys/arts/tables/2017/ecommerce.xls

In recent years, data users have requested more detailed e-commerce sales information. In 2017, ARTS introduced a new supplemental e-commerce table that is now included as part of its annual publication. This data product uses existing data collected from nonstore retailers in

[^1]NAICS 4541 and further breaks out that data by primary kind-of-business for those nonstore retailers.

If a company operates in more than one industry, it is asked to report for each industry separately. This allows the company's data to be tabulated in the correct industries. ${ }^{3}$ When a company has a large e-commerce segment - typically with separate warehousing facilitiesARTS considers this a separate industry from the company's brick-and-mortar NAICS classifications. Appendix A provides examples of how these types of company structures are created and tabulated in retail and e-commerce sales estimates.

For companies with separate store and e-commerce components as described above, the supplemental e-commerce table reallocates the sales of the NAICS 4541 component to the primary business activity (3-digit NAICS code) of the company. 'Primary business activity' refers to the 3-digit NAICS code having the most sales for the company when the sample was selected. For example, consider two separate companies (Company A and Company B) that both have a brick-and-mortar component (NAICS code 448110, which is Men's Clothing Stores) and an ecommerce component (NAICS code 454111 , which is Electronic Shopping). When the sample was selected, the majority of Company A's sales were under NAICS code 448110. Consequently, the primary business activity of Company A is 448, and the sales of the NAICS 4541 component of Company A are under 448 in the supplemental table. Conversely, when the sample was selected, the majority of Company B's sales were under NAICS code 454111 . Thus, for Company B, the primary business activity is 454 , and the sales of the NAICS 4541 component of Company B are under 454 in the supplemental table. Companies without a brick-and-mortar component remain classified under NAICS 454. Figure 2 shows an example of this annual data presentation for 2016 and 2017.

[^2]| NAICS Code of Primary Business Activity ${ }^{3}$ | Kind of Business | 2017 |  | $2016{ }^{\text {r }}$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Total | E-commerce | Total | E-commerce |
|  | Total Electronic Shopping and Mail-Order Houses | 552,214 | 397,490 | 488,619 | 339,928 |
| 441 | Motor vehicle and parts dealers | 530 | 528 | 496 | 496 |
| 442 | Furniture and home furnishings stores | 8,214 | 7,756 | 7,275 | 6,878 |
| 443 | Electronics and appliance stores | 30,791 | 30,762 | 25,639 | 25,627 |
| 444 | Building material and garden equipment and supplies dealers | D | 10,363 | D | 7,542 |
| 445 | Food and beverage stores | 1,531 | 1,400 | 1,476 | 1,358 |
| 446 | Health and personal care stores | D | 3,722 | D | 3,296 |
| 447 | Gasoline stations | ZZ | ZZ | ZZ | ZZ |
| 448 | Clothing and clothing accessories stores | 25,206 | 24,778 | 23,635 | 23,176 |
| 451 | Sporting goods, hobby, musical instrument, and book stores | 5,949 | 5,548 | 5,959 | 5,634 |
| 452 | General merchandise stores | D | 31,744 | 27,477 | 27,167 |
| 453 | Miscellaneous store retailers | D | 11,518 | D | 10,154 |
| 454 | Nonstore retailers | 371,757 | 269,371 | 329,381 | 228,600 |

D - Denotes an estimate withheld to avoid disclosing data of individual companies; data are included in higher-level totals. ZZ - Estimate is less than \$500,000.
Figure 2: Supplemental Estimated Annual Sales for U.S. Electronic Shopping and Mail-Order Houses (NAICS 4541) - Total and E-commerce Sales by Primary Business Activity: 2016-2017 Estimates are shown in millions of dollars and are based on data from the Annual Retail Trade Survey.
Source: https://www2.census.gov/programs-surveys/arts/tables/2017/supecommerce4541.xls

## 4 Supplemental Quarterly E-Commerce Table

After the annual supplemental table was published, data users expressed the usefulness of this new table but requested that the estimates in the table be published on a more frequent basis: either monthly or quarterly. The E-Commerce Data Analysis Team initially replicated the tables displayed in Figures 1 and 2 at a quarterly frequency using existing monthly retail sales data collected by the Monthly Retail Trade Survey. ${ }^{4}$ The Team used the same methodology in creating these tabulations that was employed by ARTS. When these tables were created, numerous sales estimates required suppression due to disclosure protections and data quality concerns. For many survey data products, including the retail surveys, the current quality practice is to suppress as follows:

- If $50 \%$ or more of the estimate is imputed, the data should not be published.
- If the coefficient of variation equals or surpasses $30 \%$, the data should not be published.
- If the estimate does not pass disclosure review, the data cannot be published and is further protected so a reader cannot calculate the suppressed value.
- Data collected in the MRTS are subject to legal confidentiality protections (U.S. Code Title 13, Section $9[1,5])$.

[^3]- MRTS uses cell suppression for disclosure avoidance. Cell suppression is a disclosure avoidance technique that protects the confidentiality of individual survey units by withholding cell values from release and replacing the cell values with a symbol, usually a "D". If the suppressed cell value were known, it would allow one to estimate an individual survey unit's value too closely. 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. The process of suppression does not change the higher-level totals.

Balancing the requests and needs of data users with the value of publishing a table with limited information, the Team sought feedback from internal and external stakeholders. One idea proposed during the feedback gathering was to combine the data in Figure 1 and Figure 2 to create a whole picture of e-commerce sales by primary kind-of-business. For example, rather than looking at clothing store e-commerce sales for clothing store retailers (NAICS 448 in Figure 1) and clothing nonstore retailers (the 448 breakout of 4541 in Figure 2) individually, this table would combine the e-commerce sales for the two and create total clothing store and nonstore sales and e-commerce sales estimates.

To produce the MRTS sales estimates, the following methodology is used. For each month of the quarter, Horvitz-Thompson estimates are obtained by summing weighted e-commerce sales (either reported or imputed). Benchmarked monthly estimates are computed by multiplying each Horvitz-Thompson estimate by the carry-forward factor calculated during the most recent benchmarking. Estimates for the quarter are obtained by summing the monthly benchmarked estimates. For companies with separate store and e-commerce components, the e-commerce components are tabbed in MRTS under NAICS 4541.

To calculate the sales and e-commerce estimates in the supplemental quarterly e-commerce table, first subtotal Horvitz-Thompson estimates are obtained by summing weighted sales and ecommerce sales within 4541 by the primary kind-of-business of the company. Note these companies are only tabbed under their primary kind-of-business even though they may operate in multiple kinds-of-business. To ensure the overall total retail estimates are consistent with MRTS, these kind-of-business subtotals are raked to the benchmarked monthly estimate for NAICS 4541 from MRTS. This is done by calculating the proportion of each kind of business to the HorvitzThompson estimate of NAICS 4541, then applying each of these proportions to the benchmarked monthly estimate for NAICS 4541 to get the corresponding benchmarked monthly estimate for that kind of business subtotal. Finally, these kind-of-business nonstore subtotals are added to the MRTS estimate for that kind of business to get a total store and nonstore estimate. ${ }^{5}$ Note that the Nonstore line of the table contains the remaining NAICS 454 companies, including NAICS 4541 companies whose primary kind-of-business is nonstore, after the other companies' NAICS 4541 parts are removed.

[^4]These combined estimates also presented a number of disclosure and data quality issues. To overcome these issues while still publishing as many e-commerce sales estimates as possible, the Team created groups of kinds-of-businesses. Total store and nonstore sales estimates are provided in the table at the grouping levels but not at the separate kinds-of-business. However, total e-commerce sales estimates are provided at the grouping level and at the separate kinds-ofbusiness level. For example, total sales for retailers whose kinds-of-business are in furniture and home furnishings; electronics and appliance; and building materials, garden equipment and supplies are combined to produce a furniture, building materials, and electronics total store and nonstore estimate. This grouping approach was necessary to produce a table with sufficient disclosure protections.

Based on this methodology, a new supplemental quarterly e-commerce table was originally released as an experimental data product on August 28 ${ }^{\text {th }}, 2019$. As of the 4th Quarter 2021 release on February 18, 2022, this is now a standard data product. Data in this table are available back to 2018Q2 when the current MRTS sample was introduced. Figure 3 shows an example of this table.

| Kind of Business | Retail Sales (millions of dollars) |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $2019 Q 2$ <br> (p) |  | 2019Q1 <br> (p) |  | 2018Q4 <br> (p) |  |
|  | Total Store and Nonstore Sales | Total <br> E- <br> Commerce | Total Store and Nonstore Sales | Total <br> E- <br> Commerce | Total Store and <br> Nonstore Sales | Total <br> E- <br> Commerce |
| Total Retail Trade | 1,377,719 | 139,671 | 1,241,540 | 129,015 | 1,407,934 | 160,894 |
| Motor vehicle and parts | 322,912 | 9,609 | 289,748 | 8,697 | 297,397 | 9,169 |
| Furniture, building materials, and electronics | 173,939 | S | 146,516 | S | 171,411 | 18,970 |
| Furniture and home furnishings | * | 2,505 | * | 2,344 | * | 2,934 |
| Electronics and appliance | * | S | * | S | * | S |
| Building materials, garden equipment and supplies | * | 3,995 | * | 3,018 | * | 3,449 |
| Clothing and general merchandise | 257,621 | 17,679 | 230,698 | 16,275 | 306,894 | 27,368 |
| Clothing and clothing accessories | * | 9,028 | * | 8,833 | * | 13,710 |
| General merchandise | * | 8,651 | * | 7,442 | * | 13,658 |
| All other | 489,958 | 9,131 | 445,976 | 8,962 | 487,336 | 10,735 |
| Food and beverage | * | 1,843 | * | 1,840 | * | 1,935 |
| Health and personal care | * | 1,219 | * | 1,065 | * | 1,609 |
| Sporting goods, hobby, musical instrument, and books | * | 1,347 | * | 1,399 | * | 2,072 |
| Miscellaneous including gasoline stations | * | 4,722 | * | 4,658 | * | 5,119 |
| Nonstore retailers | 133,289 | 85,716 | 128,602 | 79,337 | 144,896 | 94,652 |

S - Suppressed; * - not published
Figure 3: Supplemental Estimated Quarterly U.S. Retail Trade Sales - Total and E-commerce ${ }^{1}$ : 2019Q2. Estimates are not adjusted and based on data from the Monthly Retail Trade Survey and administrative records.

Source: https://www.census.gov/retail/mrts/www/data/excel/19q2supptables.xls

## 5 Conclusions

A new supplemental quarterly e-commerce table was created and originally published as an experimental data product in response to data user requests for more frequent and more detailed measures of e-commerce sales on August 28, 2019. As of the 4th Quarter 2021 release on February 18,2022 , this table is now a standard data product. The Team creating the table encountered data suppression issues in creating a quarterly table to mimic the annual tables produced by the Annual Retail Trade Survey. Rather than publishing a table comparable to the annual table but with limited data, the Team created a new table layout that presents a whole picture of e-commerce sales by retailers' kinds-of-business. Going forward, this table will be included in the Quarterly E-Commerce Report as a supplemental table.

In developing the new Table, the Team found that data users wanted a better understanding of how e-commerce sales are tabulated. In response, the Team developed a guide to help with this that is included as Appendix A. We also welcome feedback on improving this guide.

## Appendix A

As e-commerce has evolved over the past few decades, retailer operations around e-commerce have evolved. As part of its outreach efforts regarding e-commerce, the Team realized that data users would like to better understand how individual retailers' total sales and total e-commerce sales are tabulated and where that data is included in Census Bureau's monthly, quarterly, and annual publications. This appendix first walks through how retailers' retail and e-commerce operations are classified in the NAICS structure. Next, this appendix walks through how total retail sales and e-commerce sales tabulation parts are defined for retailers. And last, the appendix shows how these individual tabulation parts are included in the retail sales estimates published on a monthly, quarterly, and annual basis.

To guide these explanations, three example retailers are used. These are simplified and hypothetical examples used to convey e-commerce sales survey reporting and tabulation scenarios. E-commerce operations and survey reporting arrangements can vary by retailer. These three example retailers represent the following three potential e-commerce operations scenarios:

- Single store location where e-commerce sales are fulfilled from within the store.
- Large retail chain with a separate e-commerce distribution center.
- Online retailer that has no brick and mortar presence and only operates online.

STORE


- Retailer A is a small women's clothing store.
- Like other mom and pop style shops, Retailer A has only one store location.
- Retailer A has a website that they conduct e-commerce on where all online orders are fulfilled from within the store.

- Retailer B is a women's clothing chain with many store locations across the country.
- Retailer B has a website for e-commerce. Online orders that are shipped are fulfilled from one distribution center.
- Retailer B offers the omnichannel e-commerce experience allowing customers to buy online and pick up in-store.

- Retailer C is a women's clothing store that operates solely online ("pure play").


On its Economic Census form, Retailer A reported:

- One physical store location operating in NAICS code 44812 (Women's Clothing Store).


On its Economic Census form, Retailer B reported:

- Multiple physical store locations operating in NAICS code 44812 (Women's Clothing Store).
- One distribution center for e-commerce which is classified in NAICS Code 45411 (Nonstore retailer).


On its Economic Census form, Retailer C reported:

- One distribution center for e-commerce which is classified in NAICS Code 45411 (Nonstore retailer).


## Retailer A receives one MRTS form for retail activity in NAICS 44812 and <br> - Reports Total Store Sales (equal to brick \& mortar sales plus e-commerce sales) <br> - Answers "yes" to the question "Did you have e-commerce sales this month?" <br> - Reports Total E-commerce Sales (e-commerce sales only) <br> - Total E-commerce Sales should be less than or equal to Total Store Sales <br> Retailer B receives two MRTS forms for retail activity in NAICS 44812 and 45411. On the 44812 form, Retailer B <br> - Reports Total Store Sales equal to brick \& mortar only <br> - Reports zero for Total E-commerce Sales <br> On the 45411 form, Retailer B <br> - Reports Total Store Sales equal to total e-commerce sales. <br> - Reports Total E-commerce Sales. <br> - Total Store Sales equals Total E-commerce Sales unless there are catalog or phone orders. <br> Retailer C receives one MRTS form for retail activity in NAICS 45411 and <br> - Reports Total Store Sales equal to total e-commerce sales <br> - Reports Total E-Commerce Sales <br> - Total Store Sales equals Total E-commerce Sales unless there are catalog or phone orders.







[^0]:    ${ }^{1}$ The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied. (Approval ID: CBDRB-FY19-EID-B00004z)

[^1]:    ${ }^{2}$ Information on the history of the existing E-Commerce definition and measurements can be found in Thomas L. Mesenbourg's Measuring the Digital Economy working paper. Working paper available here:
    https://www.census.gov/content/dam/Census/library/working-papers/2001/econ/umdigital.pdf

[^2]:    ${ }^{3}$ Note that this only refers to operating different types of stores, not to selling multiple types of products within a single type of store.

[^3]:    ${ }^{4}$ The scope of the e-commerce work currently only covers Retail Trade. Food Services (NAICS 722) are excluded from the work.

[^4]:    ${ }^{5}$ For more information on the current quarterly E-commerce methodology as well as the supplemental table, please visit https://www.census.gov/retail/ecommerce/how_surveys_are_collected.html

