Memorandum

To: Juan Carlos Blacker, Bonneville Power Administration

From: Kate Donaldson and Lucas Judson, Cadeo

Date: July 29, 2022

Subject: Non-Residential Lighting Distributor Sales Data Gaps

---

This memorandum describes the process the Cadeo team (the research team) used to collect and analyze distributor sales data for the non-residential lighting market in the Pacific Northwest. It also discusses the resulting data's strengths and uncertainties. The research team has organized the content of this memorandum into the following sections:

- Data Summary
- Data Representativeness
- Outreach and Data Collection
- Data Structure
- Summary of Data Gaps and Sources of Uncertainty
- Extrapolation Methodology
- Recommendations for Addressing Data Gaps

Data Summary

During the 2022 non-residential lighting data collection effort, the research team gathered distributor data from 2017 to 2021. A total of 24 distributors submitted data to the research team in 2022 for 2021. The research team merged these data with data collected in previous studies spanning the years 2013 to 2016, resulting in a data set spanning 2013 to 2021.

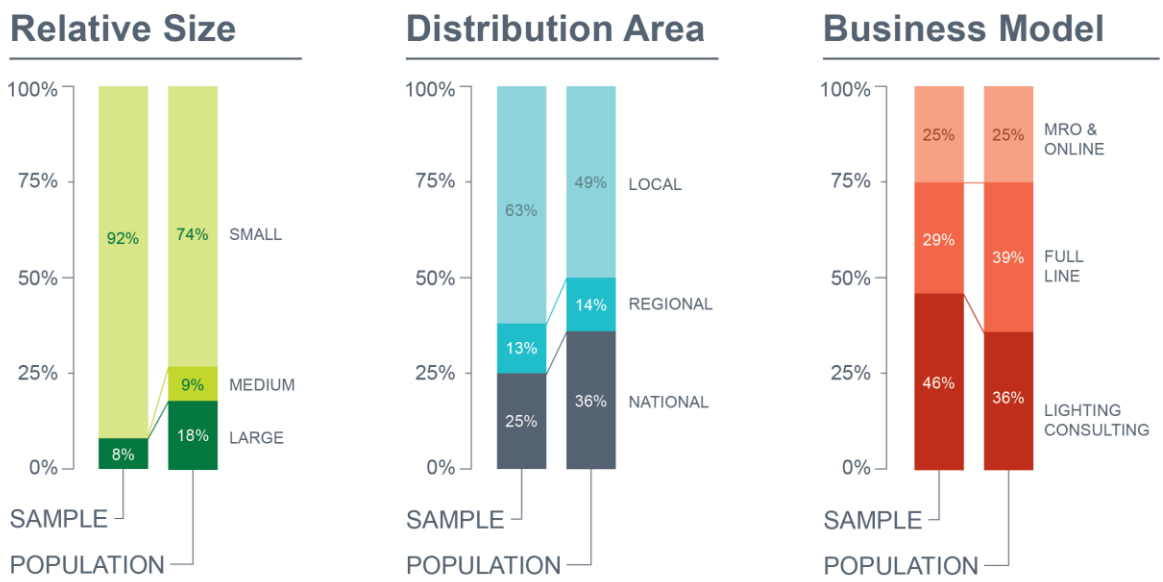
Data Representativeness

The research team reviewed this year's participation data and the cumulative data collected since 2013 to determine the extent that the data set accurately represents the non-residential lighting market. This review had two components: representation within specific segments (e.g., individual business models) and representation across channels (e.g., sales beyond traditional distributors). The following section uses two figures to demonstrate data representativeness and outlines possible data collection and analysis improvements that could lead to more accurate representation in the future.

Representativeness Findings

Figure 1 shows the mix of participating distributors by relative size, distribution area, and business model compared to the distributor population mix in the Pacific Northwest.[1]

Figure 1: Characteristics of Distributors Submitting Data Compared to the Pacific Northwest Distributor Population



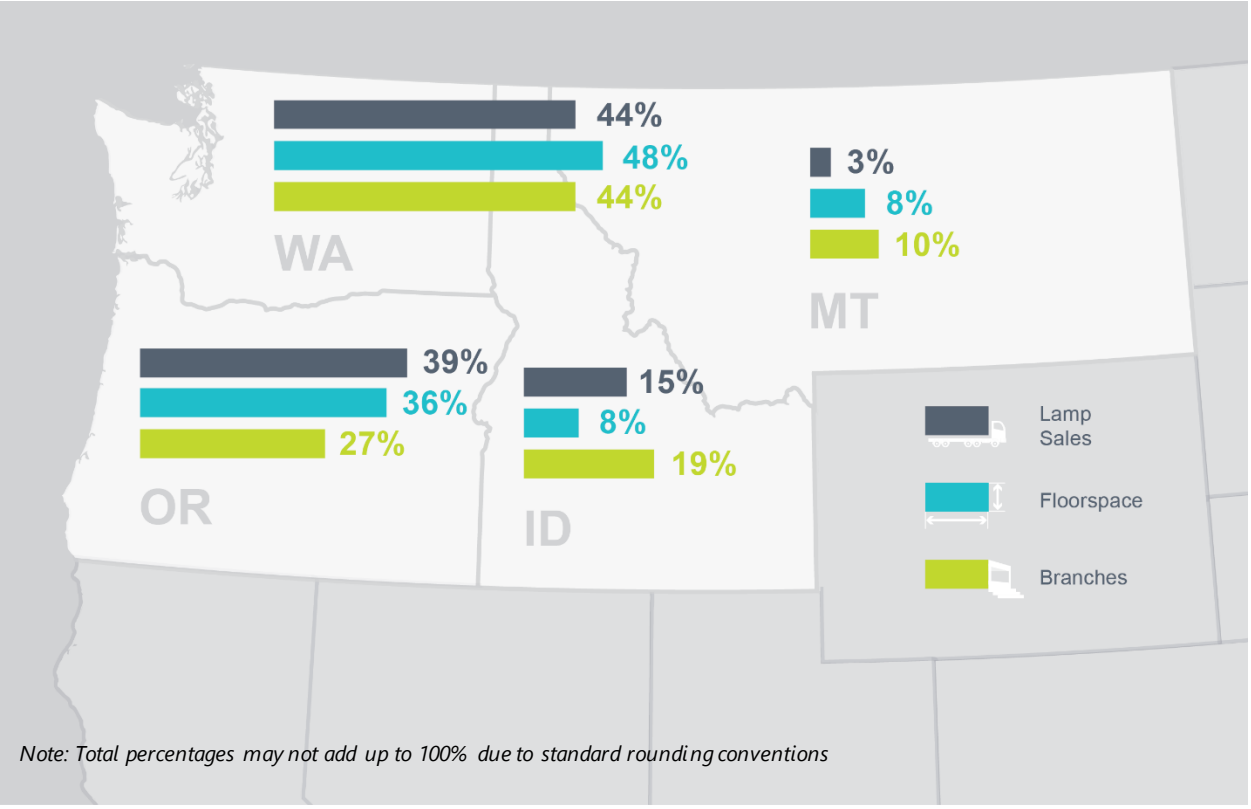
Source: BPA distributor database (population)

Population data comes from the regional distributor database Bonneville Power Administration (BPA) compiled in 2016. To keep population data up to date, the research team updated firmographic data for distributors participating in the 2018–2022 data collection efforts. The research team made these updates by conducting web searches for participating firms and populating the relevant firmographic details, like number and locations of branches, in the database. These updates were not comprehensive, as the research team collected most firmographic information through in-depth distributor interviews and other research tasks prior to 2016. However, these changes added to the overall number of distributors represented in the database and provided updated population distributions in Figure 1 and shares of branch locations by state in Figure 2 below.

Figure 2 shows the distribution of lamp shipments and known branches for participating distributors by state. Figure 2 also shows the share of total commercial building floor space by state, as quantified by the regional 2019 Commercial Building Stock Assessment (CBSA). The team observed these metrics within and between each state to determine if sales data volume reasonably correlated with expected market size across the four-state region.

1 Appendix A: Summary of Distributor Business Model Types describes business model types in detail.

Figure 2: Distributor Lamp Shipments, Floor Space, and Branches by State, 2021



Source: Distributor sales data analysis, 2019 CBSA floor space by state, and BPA distributor database

Representativeness Analysis

The research team found some variations in the distribution of characteristics in the sample compared to the region's population. The research team concluded that these variations are within an acceptable range and therefore did not make any weighted adjustments to the data. With the goal of continually narrowing the gap between the data set and the actual regional population, this section outlines possible data collection and analysis improvements within the following topics: distributor characteristic mix, geographic coverage, participation, and specific segments.

Distributor Characteristic Mix

Figure 1 (above) shows the mix of participating distributors by relative size, distribution area, and business model compared to the distributor population mix in the Pacific Northwest. The research team uses these comparisons to evaluate whether the participating distributor pool is reasonably representative of the total population of distributors in the region. The research team found that the collected data include a mix of sizes, geographic scopes, and business models that are reasonably proportional to the population with the following variations:

- A higher portion of small distributors and a smaller portion of medium and large distributors submitting sales data, relative to the population

- A larger portion of local distributors and a smaller portion of national distributors submitting sales data, relative to the population
- A smaller portion of full-line distributors and a larger portion of lighting consulting distributors submitting sales data, relative to the population

As shown in Figure 1, the mix of business models in the population differed from the mix of business models that submitted data. In 2016, the research team reviewed the sales data in these categories and found that some sales trends (e.g., the portion of sales that are LED and the portion of sales that are lamps versus fixtures) may correlate with the business model types. However, the research team and BPA determined that weighting data submissions to reflect the relative presence of these three business models in the market may not improve the accuracy of the sales data because there is too much uncertainty around non-participating distributors' sales trends and business models. The research team did not find sufficient new data during 2022 data collection to change this conclusion.

Many of the new distributors participating since 2019 have been lighting consultants selling only LEDs, which increased the representation of this business model type. This trend could either indicate that lighting consulting businesses are becoming more common in the region or that the participating distributor pool overrepresents this business model. If this business model is overrepresented, this could lead to a bias toward efficient technologies in the data. The research team will monitor this trend in future studies to determine whether any adjustments to representativeness are needed.[2] The research team will also consider changes to the outreach strategy to bring the sampled full-line distributor portion more in-line with the population portion.

Geographic Coverage

As seen in Figure 2, lamp sales collected from 2021 are approximately proportional to each state's share of regional commercial floor space and distributor branches. The research team used floor space and branches as comparisons because these two metrics should nearly correlate with sales.

Figure 2 illustrates that the proportionality of sales versus floor space and branches is a close but imperfect match. With the updates to the distributor database, the research team identified an increase in branch locations in Idaho, mirrored by an increase in sales data collected from Idaho branches, beginning in 2018 and continuing through 2021. The result is a higher percentage of reported sales and branch locations in Idaho, proportionally, than commercial floorspace in Idaho. In Montana, reported lamp sales over the last four years of the study were proportionally lower compared to the state's branches and floorspace. This indicates a small deficit in distributor participation in this state. Despite this slight divergence, other states' metrics are approximately proportional, confirming that the research team has reasonable representation of distributors and their data.

[2] The research team developed the distributor population from an analysis of distributor interviews and online research compiled into a distributor database in 2016. This population assessment has not been updated comprehensively since then, but the research team does update participating distributor firmographic information each year. Further investment in comprehensively updating the distributor population would be necessary before making any adjustments to improve representativeness.

Participation

Participation in 2022 (24 distributors) was slightly lower than participation in 2021 (27 distributors). In 2022, five participants from the 2021 data collection effort did not participate because of time constraints and staffing changes. Non-participating distributors' sales may be different from participating distributors' sales. This would create non-response bias of an unknown direction and magnitude. However, the research team attempted to minimize this bias by performing repeated outreach. The research team attempts to connect with distributors at least three times before ceasing outreach and uses an incentive to encourage participation.

Specific Segments

The research team assumes submitted data from the full line; maintenance, repair, and operations (MRO) and online; and lighting consulting segment study participants to be representative of total sales within these distributor types in the total market. To validate this assumption, the research team reviewed this year's participating distributors' sales mixes by business model to determine whether participating distributors within a business model type show similar sales patterns. The research team calculated the normalized standard deviation of the mean within each business model type for participating distributors' sales mixes within that business model type and confirmed that participating distributor's sales mixes within each segment are generally similar.[3] This supports the research team's assumption that the submitted data are representative within each business model category.

The research team finds the representativeness of this year's data set acceptable and comparable with past years. The research team will continue to work to address representation gaps through the outreach recommendations covered in the next section. These recommendations will allow the research team to strategically target distributors in underrepresented or hard-to-reach segments using lessons learned over the previous study years.

Outreach and Data Collection

The research team, contracted by the Northwest Energy Efficiency Alliance (NEEA), worked in collaboration with the NEEA Distributor Platform team and Evergreen Consulting (Evergreen) to facilitate distributor outreach. The NEEA Distributor Platform team, including NEEA staff and D+R International staff, manages NEEA's relationships with distributors across various initiatives. The research team conducted outreach and recruited distributors to participate in the data collection effort.

The research team's strategy followed these general steps:

1. One of the three outreach organizations—NEEA, Evergreen, or Cadeo—contacted distributors and asked them to participate, giving them an information packet that BPA developed. The outreach organization offered a monetary incentive to participating distributors.

[3] The normalized standard deviation of the mean by business model type for participating full line, lighting consulting, and MRO and online distributors was 16%, 14%, and 23%, respectively.

2. If the distributor agreed to participate, the research team sent the distributor the sales data collection form, instructions for submitting data securely, and an optional non-disclosure agreement.
3. The research team logged all communication in a distributor data tracker and provided outreach updates in a weekly outreach team meeting.

NEEA's relationships with regional distributors and Evergreen's outreach to existing contacts were essential to the outreach team's success. In the last four years of the study, the research team at Cadeo reached out directly to select long-term participants who proved responsive in past studies. The research team aimed to organize outreach assignments to reduce the number of touchpoints for distributors and streamline the total number of outreach steps. NEEA, Evergreen, and Cadeo discussed and agreed upon all outreach assignments at the beginning of each study.

In some cases, long-time participants respond to outreach to let the research team know that they cannot participate in the current study (this year, for example, five distributors did not participate due to time constraints). In 2021, the research team piloted an effort to follow up with non-participating distributors to gather basic information about the magnitude of their sales and product offerings. The research team continued that strategy this year but did not receive any sales trend information from these distributors.

Final Outreach Disposition and Results

Table 1 provides details on the final disposition for the distributors included in the research team's outreach.

Table 1: Summary of Distributor Outreach, Final Disposition

Category	Number of Distributors
Total distributors included in outreach*	**69**
Distributors submitting data	24
• Repeat participants	23
• New participants	1
Distributors declining to participate	12
• Lack of time and/or interest	10
• Data reporting limitations	2
• Prohibited by company policy	0
Distributors unresponsive to outreach	33

**Includes individual distributor branches that operate independently*
Source: Distributor outreach tracking

This year, 24 distributors submitted data, which was three fewer than in 2021. This year also saw two more distributors decline to participate compared to the previous year. These distributors declined either by ceasing to respond to outreach efforts or citing a lack of time or interest. Two of the declining distributors indicated their company's data storage and reporting capabilities prevented them from reporting data at

the level of granularity requested. Of the 27 participants from the previous year, five did not participate this year due to lack of time or lack of interest. However, the research team successfully re-recruited one participant that had historically participated but did not participate last year.

Data Structure

Distributors submitted 2021 sales data in two different formats:

- The standardized survey tool provided. Of the 24 participants, 23 submitted data using the standard data collection tool.
- A custom protocol developed through discussions with the distributor to understand and work around data reporting limitations. One distributor participated using a custom protocol in 2022.

Between 2019 and 2021 (data years 2018 to 2020) NEEA worked with distributors that participated in various initiatives to streamline requests by requesting raw data extracts of all product sales from some of their partner distributors. These "data dumps" represented full-year, full-category data (except for controls) for participating distributors, not just linear fluorescent and tubular LED (TLED) data. This change was part of NEEA's ongoing development of a distributor platform that supports distributor partnerships on multiple initiatives. Starting in 2021, NEEA did not collect data dumps from distributors, so the research team discontinued the data dump option for participating distributors; all distributors participating in 2022 submitted data using the standardized survey tool or the custom protocol.

The research team merged the data collected from both data submission formats with historic data submissions from 2013 to 2020 in a single SQL server database. Collectively, these data span sales from 2013 to 2021. The research team used five fields to organize the sales data by lamp and luminaire characteristics, which Table 2 summarizes.

Table 2: Lighting Product Description Fields Used in Database

Field Name	Description
Lighting_Technology_Type	The technology category includes either the lighting technology type (e.g., LED, linear fluorescent, etc.) or controls.
General_Category	This field lists the lamp shape (e.g., T8, T5, A-Type, Reflector, etc.), fixture type, or type of high intensity discharge (HID) lamp (e.g., high-pressure sodium, metal halide, or mercury vapor).
Dimension	Where applicable, this field provides the length or dimensions of the lamp or fixture. This field primarily applies to linear lamps and fixtures (e.g., 4-foot and 8-foot lamps).
Subcategory	This field provides additional detail on lamp and fixture characteristics. Details may include a specific wattage, wattage range, lumen output, or more specific lamp shape (e.g., MR16 within the "Reflectors" general category).
Base_Type	This field specifies whether the product is a fixture; for lamps it indicates whether the base type is screw-in, mogul-base screw-in, or pin.

Appendix A: Recommendations for Addressing Data Gaps

The research team will continue refining the outreach process to maximize distributor participation and fill existing data gaps caused by participation. The research team recommends the following refinements to the outreach process:

- **Increase efforts to reengage past participants, focusing on large full-line distributors.** Over the last five years, several large, regional, and national full-line distributors who previously participated in the study have dropped out. These distributors declined to participate due to timing constraints. In the next year of the study, the research team should focus outreach efforts in the first month of outreach on reengaging these large regional and national full-line distributors in the study. The research team recommends discussions, including in-person meetings where possible, among Cadeo, NEEA, and the distributor to underscore the distributor's importance as a partner. In these discussions, the research team can offer options to ease the burden of participation, including extending data submission deadlines, developing custom data collection strategies, or increasing incentives. If past participants are unresponsive in the first month, the research team should engage other reasonable contacts at a given distributor to rebuild rapport with the unresponsive contact. Many unresponsive cases this year were impacted by staffing changes, so this process may result in an introduction to the new best point of contact.
- **Collect updated firmographic information about all distributors in the next round of data collection.** The research team collected firmographic information (through a combination of interviews and website/online review) in previous rounds of data collection and populated a distributor database with these details. For example, interviews with distributors confirmed their supply chain relationships to prevent double counting of collected sales data, and online research allowed the research team to record how many branch locations each distributor has in the Northwest and whether each distributor promotes energy efficiency on their websites. These

firmographic details help the research team check for potential biases in the data set. In the 2018–2022 study years, the research team updated what firmographic information was found easily on distributor websites for those participants in these outreach efforts. The information for other distributors in the database has not been updated since 2016. NEEA and BPA should collect updated firmographics in a future round of data collection. These updates should include confirming supply chain relationships for national and local distributors.

- **Determine whether a targeted recruitment strategy to increase lamp sales representation in Montana is necessary.** Based on the updated distributor database recommended above, the research team recommends assessing distributor participation for Montana (reported lamp sales and branch locations in the state) to determine whether the research team should make an effort to increase participation from Montana-based distributors.

In addition to refining the existing outreach process, BPA and NEEA should consider strategies to expand data collection to include non-residential lighting sales channels that have previously not been included in this study. The research team proposes BPA pursue the following research activities to understand the size and sales mix of channels not previously included in data collection (retailers, online distributors, national accounts distributors, etc.):

- **Review collected distributor data to investigate the sales mix.** The research team recommends reviewing collected data submissions from participating distributors in non-distribution sales channels – namely one participating retailer and one previously participating online distributor – to develop an initial hypothesis about the sales mix in each of these sales channels.
- **Interview participating distributors and (other market actors) to understand the sales channel.** The research team recommends conducting market actor interviews to investigate changes to non-residential lighting sales channels in terms of volume or the specific flow of products. In particular, BPA should seek to understand changes in the volume of sales flowing through retail and online channels to non-residential lighting projects. The research team recommends interviewing a few distributors that participate year over year to understand any shift in sales channels from the distribution side. The research team also recommends interviews or a survey with lighting, electrical, and/or general contractors to understand their lighting purchase habits and if or when they use different sales channels for projects. BPA can leverage the research team's existing relationships with distributors to identify interview targets.

Based on this initial research, the research team will recommend expanding outreach to one or multiple additional non-residential lighting sales channels. Findings from initial research will also inform tailored outreach strategies for each sales channel.

Each year of data collection brings new insights and challenges as distributors submit additional data. This year, as in past years, the team found distributor participation to reasonably represent the Northwest region, as reflected in the mix of distributor sizes, distribution areas of sales, and business model types. As NEEA and BPA continue to study non-residential lighting sales, the team recommends expanding the study's focus to capture non-distributor sales channels. This expansion will improve the validity and usability of the non-residential lighting data collected in this study in future years.

Summary of provides a complete list of possible entries for these fields, and Appendix B: Summary of Unique Database Category Entries describes the analysis process.

Under the 2013 data collection effort, distributors provided data back to 2010. But, for the 2022 analysis, the research team focused only on sales in years 2013–2021. Data quality and number of participants increased significantly in 2013 due to expanded outreach and revisions to the data collection instrument.

Summary of Data Gaps and Sources of Uncertainty

The research team identified the following data gaps and sources of uncertainty based on the 2022 study participation and analysis:

- **Inconsistent distributor participation contributes to uncertainty.** Changes in distributor participation year over year create gaps in reported sales data. The research team can fill these gaps if the distributor participates again in a future year (e.g., if a distributor participated in 2020, did not participate in 2021, and then participates in 2022, the research team will ask that distributor to supply the missing year of data). Additionally, the research team has developed an extrapolation approach, discussed below, to estimate sales for distributors who contributed data in the past but did not contribute data in every year. The extrapolation process allows the research team to estimate missing distributors' sales based on historic market shares and other participants' sales. This allows the research team to leverage data from past and present participants but does not fully resolve the uncertainty due to limited data from some distributors.
- **Collected sales data only represent the distribution sales channel.** Market intelligence gathered by the research team indicates that the distribution channel is the largest sales channel for non-residential lighting. However, the sales data collected does not include sales from the following channels:
    - **National Account Distributors:** The research team did not collect sales data from any specialized national account distributors. Despite multiple attempts over previous study years to connect with national account distributors, the research team has yet to secure participation from any of them. Although some participating distributors serve national account customers among other types of customers, previous research on national accounts suggests that many national accounts get their lighting products from specialized distributors that serve national accounts exclusively. It is possible that some lighting consultants serve more national accounts than traditional full line or MRO and online, but the research team did not explore this possibility specifically as part of this study. If participation from lighting consulting distributors remains high in future studies, the research team will explore if having increased lighting consulting representation impacts the representation of national account sales.
    - **Retailers:** The research team collected sales data from only one low-volume sales retailer, giving a small, non-representative look into the retail channel. Through market intelligence gathered over previous study years and modeling efforts, the research team believes that the retail channel could represent a meaningful volume of non-residential lighting sales, particularly in the linear product technology categories. The research team

has not had success in its efforts to collect sales data from other retrial channels like Home Depot. In future years of the study, the research team will work to further define the size and importance of this sales channel as well as the sales mix of technologies sold through the retail channel.

- **Online Sales Channels:** The research team did not collect sales data from any specialized online distributors or retailers, but some online sales are reflected in the collected data. One long-time online distributor stopped participating in the study in 2018, and since then the research team has been unable to collect data from this segment. In previous years, the research team determined through webinars and informal interviews with traditional distributors that most traditional distributors are attempting to reach the online sales channel both through their own online sales platforms and through partnerships with third-party vendors like Amazon. In 2021, the research team added questions to the standard data collection tool to ask what percentage of sales were from online channels and if these sales were from their own websites or through third-party vendors. In 2022, eight of the 24 participants reported online sales. All except one of the distributors with reported online sales indicated that they included these sales in the sales reported through the standard data collection tool, indicating that the research team is already collecting online sales from participating distributors. In future years of the study, the research team will work to further define the size and importance of this sales channel as well as the percentage of online sales within data already collected in this study.

- **The research team adjusted some reported sales data, with distributor verification, to account for residential sales.**[4] For 2021 sales data, all distributors indicated what percentage of their sales go to residential versus non-residential customers in the data collection tool. The research team confirmed with each distributor the split of sales and whether their reported sales included any residential sales. Distributors with both residential and non-residential sales confirmed the split between these sales and provided additional insights into which product categories contained residential sales. For these distributors, the research team normalized reported sales to only include non-residential sales.

Extrapolation Methodology

To maintain consistency in the distributor database over time, the research team applied an extrapolation approach to leverage information submitted by non-participants in previous study years.[5] Extrapolation provides a mechanism to use the best available information in any given year by projecting sales for previous study participants based on overall market trends and each participant's past market activity.

[4] Over the course of the study, not all participating distributors have provided percentages indicating what portion of their sales go to residential versus non-residential customers. Because the target market for wholesale distributors is the commercial market, the team has assumed 100% of reported sales go to non-residential customers in instances where no other percentages were reported. This assumption is consistent with prior years' analysis and with the research team's qualitative understanding of those distributors' sales.

[5] New distributors are not back-cast to years prior to the first year included in their submissions.

In this year's study, the research team applied the same methodology to extrapolate missing sales data for each combination of distributor and technology as in previous years. The research team leveraged distributor-specific sales data to inform the extrapolation process. One new participating distributor provided sales data from 2017 through 2021. The team updated previous extrapolation results to incorporate this additional data submission.

The extrapolation methodology relies on two core assumptions:

1. **Distributors that did not participate retained a similar market share within lighting technologies and categories (e.g., LED downlights) as they did in the previous year.** In other words, distributors that did not participate experienced a similar trend in sales as the rest of the market within each technology and category. For example, if participants that submitted data for this study year (2021) and the previous year (2020) observed a 25% decrease in full-line CFL A-Type lamp sales, then the research team assumed each non-participating distributor's full-line CFL A-Type lamp sales decreased by 25% in 2021.
2. **The mix of subcategories[6] (lamp shape, wattage, and base type) for a given distributor remains consistent over time.** For example, if a participating distributor's sales show that in 2020 5% of full-line 4-foot T8 lamp sales are 28W lamps, the analysis assumes that 5% of the missing distributors' 4-foot T8 lamp sales were 28W lamps in 2021.

At a high level, the extrapolation methodology is broken into three steps:

1. Identify distributors that participated in both the extrapolation year (i.e., 2021) and the previous year (i.e., 2020).
2. Calculate the year-over-year percentage change in sales within each technology category (e.g., LED a-type lamps) for the distributors that participated in both years (identified in step 1).
3. Apply the percentage change calculated in step 2 to each non-participating distributor's previous year sales within each technology and category.

After extrapolating each non-participating distributor's technology category sales, the research team applied each distributor's subcategory breakdown to the extrapolated numbers using the distributor's subcategory proportions from the previous year.[7]

Extrapolation Findings

Table 3 shows the percentage of extrapolated data by year and technology from 2013 through 2021. The overall percentage of extrapolated data increases year over year. With each additional distributor that drops out of the study, their data requires extrapolation. Therefore, the cumulative amount of

[6] Subcategories for each technology and general category are provided in Appendix B: Summary of Unique Database Category Entries.

[7] The extrapolation is based on the prior year of data for each non-participating distributor. In the 2021 study, the team explored using a three-year average to extrapolate market share. Due to year-over-year changes in distributor participation, combined with regular market share changes, the three-year average produced less accurate extrapolated sales, when compared with distributor's actual sales, than the prior year when evaluated using participants that participated in consecutive years. For more information on this work, consult the Non-Residential Lighting Distributor Sales Data Gaps memo from 2021.

extrapolated data increases. Decreases in overall extrapolated data can occur if a distributor participates again after a period of non-participation.

A non-participant in 2020 underwent extrapolation and captured a large percentage of the sales mix for CFLs, halogens, HIDs, and incandescent lamps. In 2020, that distributor captured as much as 15% and no less than 9% of the market share within these specific technologies. The research team has not received sales data from this high-volume sales distributor since 2020, resulting in a substantial rise in extrapolation percentages for these technologies between 2019 and 2021. If this distributor participates in future years, these extrapolation percentages will likely decline. For CFLs, specifically, two distributors with high volumes of CFL sales who participated in 2021 did not participate in 2022, resulting in an additional jump in CFL extrapolation percentage for that technology type between 2020 and 2021.

The extrapolation percentage for linear fluorescents increased sharply between 2020 and 2021. NEEA's Reduced Wattage Lamp Replacement (RWLR) initiative collected sales data on linear fluorescent and TLED lamps for all its participating distributors through 2020. NEEA has ceased RWLR data collection, so 2020 was the last year of this study that included RWLR data. Several of these RWLR participants also participated in this study (i.e., submitted sales data for other technologies), but for some distributors the research team only received the RWLR linear product (linear fluorescent and TLED) data.[8] In 2021, eight distributors for which the research team had RWLR data did not participate in the data collection study. The research team therefore extrapolated these eight distributors' sales of linear products in 2021 for the first time, resulting in an increase in the extrapolated percentage in this technology category.[9]

Table 3: Percentage of Data Extrapolated by Year and Technology

Lighting Technology Type	2013	2014	2015	2016	2017	2018	2019	2020	2021	All Years
CFL	40%	26%	2%	17%	15%	17%	18%	37%	41%	21%
HID	7%	4%	1%	8%	10%	11%	13%	24%	27%	9%
Halogen	24%	13%	1%	11%	15%	23%	24%	34%	35%	17%
Incandescent	30%	20%	6%	16%	16%	20%	21%	36%	34%	20%
LED	22%	13%	13%	18%	22%	23%	26%	37%	37%	27%
Linear Fluorescent	25%	5%	7%	5%	11%	12%	14%	15%	45%	14%
All Technologies	26%	11%	6%	11%	16%	18%	21%	30%	39%	20%

Source: Distributor sales data analysis

The following section describes the strengths and weaknesses of the extrapolation methodology.

Strengths and Weaknesses of Extrapolation Methodology

The analysis team identified one main strength of the extrapolation process: **using historical data provides a more comprehensive and representative view of the market, regardless of participation.**

[8] See Appendix C: Detailed Data Cleaning Approach for more details on how the team included RWLR sales data in this study.

[9] See Appendix C: Detailed Data Cleaning Approach for more details on how the research team adjusted reported sales data for linear products to account for RWLR sales data from non-study participants.

In every year, the data set comprises not only collected data from participating distributors but also extrapolated data from previous participants. This provides the research team with a more comprehensive and representative data set than only including participant data.

The research team leveraged past information to identify what products certain companies sell and estimate how much they sell within certain market segments. Over the course of the study, the research team collected sales data from 44 distributors. Of those, 24 are consistent participants with no extrapolated data in the 2021 data set (55% of participants). For the other 45%, extrapolation leverages current market trends and distributors' historical data to predict sales for non-participants. This allows the research team to leverage the best available information to estimate market sales in each study year. The basic assumption is that the historical market share for non-participants will continue unless new evidence suggests otherwise.

Full-line distributors have decreased and lighting consulting distributor participation have increased as a percentage of participants, but the extrapolation methodology allows the research team to represent a larger percentage of full-line distributor sales in the final extrapolated data set than what could be represented without extrapolated data.

The analysis team identified one main weakness of the extrapolation process: **extrapolation introduces uncertainty.** The extrapolation process applies the market's sales trend for distributors who did not participate. As a result, the research team assumed that non-participating distributors possess the same sales trends and patterns as the rest of the market. The strategy is susceptible to inaccurate representations for specific distributors whose sales may differ from the overall market. **However, where possible, the research team mitigates this weakness by applying adjustments based on distributor-reported trends.**

To better quantify uncertainty resulting from the extrapolation process, the research team applied an uncertainty analysis to the extrapolated sales quantities in 2021.[10] The uncertainty analysis found that though the extrapolation methodology used does introduce some margin for error and uncertainty in the final estimates of annual lamp sales, the error introduced appears small. The overall margin of error for each technology ranges from 2% to 22% when using a 90% confidence interval. The margin for error due to extrapolation shown is small enough to reinforce the research team's understanding of directional trends (i.e., if a lighting technology is generally increasing or decreasing and at what magnitude) within most individual lighting technologies. Although uncertainty related to extrapolation will naturally increase over time as the percentage of extrapolated data increases, the results of the uncertainty analysis confirm a high degree of confidence in the extrapolated data.

Recommendations for Addressing Data Gaps

The research team will continue refining the outreach process to maximize distributor participation and fill existing data gaps caused by participation. The research team recommends the following refinements to the outreach process:

[10] For more information on this uncertainty analysis, consult the Uncertainty Analysis Memo.

- **Increase efforts to reengage past participants, focusing on large full-line distributors.** Over the last five years, several large, regional, and national full-line distributors who previously participated in the study have dropped out. These distributors declined to participate due to timing constraints. In the next year of the study, the research team should focus outreach efforts in the first month of outreach on reengaging these large regional and national full-line distributors in the study. The research team recommends discussions, including in-person meetings where possible, among Cadeo, NEEA, and the distributor to underscore the distributor's importance as a partner. In these discussions, the research team can offer options to ease the burden of participation, including extending data submission deadlines, developing custom data collection strategies, or increasing incentives. If past participants are unresponsive in the first month, the research team should engage other reasonable contacts at a given distributor to rebuild rapport with the unresponsive contact. Many unresponsive cases this year were impacted by staffing changes, so this process may result in an introduction to the new best point of contact.
- **Collect updated firmographic information about all distributors in the next round of data collection.** The research team collected firmographic information (through a combination of interviews and website/online review) in previous rounds of data collection and populated a distributor database with these details. For example, interviews with distributors confirmed their supply chain relationships to prevent double counting of collected sales data, and online research allowed the research team to record how many branch locations each distributor has in the Northwest and whether each distributor promotes energy efficiency on their websites. These firmographic details help the research team check for potential biases in the data set. In the 2018–2022 study years, the research team updated what firmographic information was found easily on distributor websites for those participants in these outreach efforts. The information for other distributors in the database has not been updated since 2016. NEEA and BPA should collect updated firmographics in a future round of data collection. These updates should include confirming supply chain relationships for national and local distributors.
- **Determine whether a targeted recruitment strategy to increase lamp sales representation in Montana is necessary.** Based on the updated distributor database recommended above, the research team recommends assessing distributor participation for Montana (reported lamp sales and branch locations in the state) to determine whether the research team should make an effort to increase participation from Montana-based distributors.

In addition to refining the existing outreach process, BPA and NEEA should consider strategies to expand data collection to include non-residential lighting sales channels that have previously not been included in this study. The research team proposes BPA pursue the following research activities to understand the size and sales mix of channels not previously included in data collection (retailers, online distributors, national accounts distributors, etc.):

- **Review collected distributor data to investigate the sales mix.** The research team recommends reviewing collected data submissions from participating distributors in non-distribution sales channels – namely one participating retailer and one previously participating online distributor – to develop an initial hypothesis about the sales mix in each of these sales channels.

- **Interview participating distributors and (other market actors) to understand the sales channel.** The research team recommends conducting market actor interviews to investigate changes to non-residential lighting sales channels in terms of volume or the specific flow of products. In particular, BPA should seek to understand changes in the volume of sales flowing through retail and online channels to non-residential lighting projects. The research team recommends interviewing a few distributors that participate year over year to understand any shift in sales channels from the distribution side. The research team also recommends interviews or a survey with lighting, electrical, and/or general contractors to understand their lighting purchase habits and if or when they use different sales channels for projects. BPA can leverage the research team's existing relationships with distributors to identify interview targets.

Based on this initial research, the research team will recommend expanding outreach to one or multiple additional non-residential lighting sales channels. Findings from initial research will also inform tailored outreach strategies for each sales channel.

Each year of data collection brings new insights and challenges as distributors submit additional data. This year, as in past years, the team found distributor participation to reasonably represent the Northwest region, as reflected in the mix of distributor sizes, distribution areas of sales, and business model types. As NEEA and BPA continue to study non-residential lighting sales, the team recommends expanding the study's focus to capture non-distributor sales channels. This expansion will improve the validity and usability of the non-residential lighting data collected in this study in future years.

Appendix A: Summary of Distributor Business Model Types

The research team defines each business model as follows:

- **Full line:**
    - Traditional electric distributors selling all general electric products, including (but not limited to) scheduled regular maintenance orders
    - Larger businesses that typically have in-house lighting and/or electrical staff
    - Lighting is usually a small portion of the overall business
- **Maintenance, repair, and operations and online:**
    - Primarily serve scheduled regular maintenance orders
    - Often receive orders online or via email
    - May sell a variety of products or just lighting
    - Tend to have a higher proportion of lamps to fixtures in sales
- **Lighting consultants:**
    - Small companies with a focus on energy efficiency projects
    - May only sell LED products and lighting controls
    - May sell only lighting or a variety of products

Appendix B: Summary of Unique Database Category Entries

Tables B-1 through B-7 summarize the possible entries for the General_Category, Dimension, Subcategory, and Base_Type fields for each lighting technology type.

Table B-1: Lighting Technology Type: Linear Fluorescent[11]

General Category	Dimension	Subcategory	Base Type
T12	4-foot	34W	Lamp
	4-foot	40W	Lamp
	4-foot	U-Shape	Lamp
	4-foot	Other	Lamp
T8 - High Performance 800 Series or Better	4-foot	25W	Lamp
	4-foot	28W	Lamp
	4-foot	32W	Lamp
	4-foot	Other	Lamp
T8 - Standard 700 Series	4-foot	32W	Lamp
	4-foot	U-Shape	Lamp
	4-foot	Other	Lamp
T5	4-foot	28W	Lamp
	4-foot	54W	Lamp
	4-foot	Other	Lamp
T12	8-foot	Slimline	Lamp
	8-foot	High Output	Lamp
	8-foot	Other	Lamp
T8	8-foot	Slimline	Lamp
	8-foot	High Output	Lamp
	8-foot	Other	Lamp

[11] The distributor data collection tool does not include non-standard linear products (e.g., 1-foot 32W T8s) because the team assumed these products were sold at low volumes. The research team verified this assumption when the data collection tool was built through review by subject matter experts. In 2021, the research team tested this assumption by asking distributors with high sales of other linear fluorescent products if non-standard linear products are still a low sales volume category. The three distributors consulted confirmed low volume of non-standard linear fluorescent products sales.

Table B-2: Lighting Technology Type: HID

General Category	Dimension	Subcategory	Base Type
Mercury Vapor		≤400W	Mogul-Base Lamp
		>400W	Mogul-Base Lamp
High Pressure Sodium		≤400W	Mogul-Base Lamp
		>400W	Mogul-Base Lamp
Metal Halide		≤400W	Mogul-Base Lamp
		>400W	Mogul-Base Lamp

Table B-3: Lighting Technology Type: LED

General Category	Dimension	Subcategory	Base Type
A-Type		100W Incandescent Equivalent	Screw-Base Lamp
		75W Incandescent Equivalent	Screw-Base Lamp
		60W Incandescent Equivalent	Screw-Base Lamp
		40W Incandescent Equivalent	Screw-Base Lamp
Reflectors		MR16	Pin-Base Lamp
		PAR	Screw-Base Lamp
		R/BR	Screw-Base Lamp
		Other Reflectors	Screw-Base Lamp
LED Downlights		PL Replacement	Pin-Base Lamp
		≤4-inch	Fixture
		>5-inch	Fixture
		≤4-inch	Retrofit Kit
		>5-inch	Retrofit Kit
LED Tubes	4-foot	UL Type A - Plug-and-Play/Direct Replacement	Lamp
	4-foot	UL Type B – Ballast Bypass	Lamp
	4-foot	UL Type A/B – Dual-Mode/Hybrid	Lamp
	4-foot	UL Type C – Remote Driver	Lamp
	4-foot	T5 Replacements	Lamp
	4-foot	Other	Lamp
Other LED Linear Fixtures		Linear Strip Fixture (Lightbar)	Fixture
		Panels < 3,500 Lumens	Fixture
		Panels ≥ 3,500 Lumens	Fixture
		Troffers and Wraps <3,500 Lumens	Fixture
		Troffers and Wraps ≥3,500 Lumens	Fixture

General Category	Dimension	Subcategory	Base Type
Decorative			Screw-Base Lamp
Flood Light		Screw Terminal Base Flood Lamps	Lamp
Flood Light Fixtures/Luminaires			Fixture
LED Decorative Post-Top and Bollard			Fixture
LED Other Outdoor Area and Site Fixtures		<15,000	Fixture
		≥15,000	Fixture
LED Track Head			Fixture
LED Garage Fixtures			Fixture
LED Canopy Fixtures (e.g., Gas Stations)			Fixture
LED Roadway (e.g., Cobra type)			Fixture
LED Other Form Factors			Fixture
LED Wall Packs			Fixture
Industrial Applications		High-bay ≥ 15,000	Fixture
		Low-bay 5000-15,000	Fixture
		High-bay ≥ 15,000	Mogul-Base Lamp
		Low-bay 5000–15,000	Mogul-Base Lamp
		Medium Base Corn Lamps	Lamp

Table B-4: Lighting Technology Type: Incandescent

General Category	Dimension	Subcategory	Base Type
A-Type		100W Incandescent Equivalent	Screw-Base Lamp
		75W Incandescent Equivalent	Screw-Base Lamp
		60W Incandescent Equivalent	Screw-Base Lamp
		40W Incandescent Equivalent	Screw-Base Lamp
Reflectors		R/BR	Screw-Base Lamp
		PAR	Screw-Base Lamp
		Other Reflectors	Screw-Base Lamp
Decorative			Screw-Base Lamp
Flood Light		Screw Terminal Base Flood Lamps	Lamp

Table B-5: Lighting Technology Type: Halogen

General Category	Dimension	Subcategory	Base Type
A-Type		100W Incandescent Equivalent	Screw-Base Lamp
		75W Incandescent Equivalent	Screw-Base Lamp
		60W Incandescent Equivalent	Screw-Base Lamp
		40W Incandescent Equivalent	Screw-Base Lamp
Flood Light			Fixture
		Screw Terminal Base Flood Lamps	Lamp
Reflectors		R/BR	Screw-Base Lamp
		PAR	Screw-Base Lamp
		MR16	Pin-Base Lamp
		Other Reflectors	Screw-Base Lamp
Decorative			Screw-Base Lamp

Table B-6: Lighting Technology Type: CFL

General Category	Dimension	Subcategory	Base Type
A-Type		100W Incandescent Equivalent	Screw-Base Lamp
		75W Incandescent Equivalent	Screw-Base Lamp
		60W Incandescent Equivalent	Screw-Base Lamp
		40W Incandescent Equivalent	Screw-Base Lamp
Flood Light			Fixture
Spiral GU24 Base Type		<20W	Fixture
		≥20W	Fixture
Single, Double, Triple Tube		<20W	Pin-Base Lamp
		≥20W	Pin-Base Lamp
Reflectors		R/BR	Screw-Base Lamp
		PAR	Screw-Base Lamp
		Other Reflectors	Screw-Base Lamp
Decorative			Screw-Base Lamp

Table B-7: Controls

General Category	Dimension	Subcategory	Base Type
Daylight Dimmers			
Networked and Other Advanced Controls		Luminaire Level Networked	
		Other Networked/Advanced Controls	
Occupancy/ Vacancy Sensors		Ceiling Mounted	
		Wallbox	
Photocells		Indoor	
Scheduling Clock/Timers			
Wireless			
Wireless Relays			

Appendix C: Detailed Data Cleaning Approach

The research team performed two levels of quality control (QC) review on incoming data submissions. First, the research team reviewed submissions within 48 hours of receipt to identify any notable data gaps that required follow-up requests to the distributor. Second, the research team reviewed the final submitted data relative to previous submissions (where applicable) and to other distributors' sales trends. The research team aggregated all sales data into a common format in a SQL server database to perform this second review. The following sections describe these processes.

Initial Data QC Review

The research team used a standard QC checklist to review all data submissions. The checklist covers the major areas where errors in data input are likely to occur. It also ensures the appropriate information for aggregating the data for analysis is present. A summary of the initial data QC follows.

1. **Scope of review.** The research team ensured distributors did not report sales outside of the Pacific Northwest region (Idaho, Montana,[12] Oregon, and Washington), branch information was included for data sets spanning multiple locations, and sales totals were expressed in units and dollars. The research team also checked that the data did not violate any data validation rules.
2. **Data gaps.** The research team reviewed all data sets, flagged any sales field that was missing data (either a 0.00 or blank cell), and followed up with the distributor to confirm the gap. If the distributor confirmed zero products were sold, the research team filled in zeros. However, if a product was sold but the distributor could not report it, the research team left it blank. For example, some distributors could not extract sales for controls or fixtures due to reporting system limitations.
3. **Data magnitude.** The research team also reviewed each tab to ensure the magnitude of sales for each application was reasonable and flagged any cell that could have been an error (e.g., 0.25 or 250,000,000 sales for a particular application, or an unusual increase or decrease in sales year-over-year).
4. **Data reporting.** Lastly, the research team confirmed that distributors reported all sales data in terms of individual lamps and not packages of multiple lamps.

Merging Data from Multiple Sources

Mapping Sales Survey Forms to Standardized Fields

The research team created a data import process to bring data from the Excel survey forms into SQL. Using a data extraction template created by the research team, Cadeo extracted all key data fields from the survey forms into a comma separated values (CSV) file. The research team then imported these CSV

[12] The research team accepted data from all of Montana; however, the Pacific Northwest region only includes Western Montana. Thus, using data from the entire state for regional analysis assumes that the sales mixes of the eastern and western portions of the data are similar.

files directly into the SQL Server database. The research team repeated this process for each of the returned distributor surveys.

Mapping Historic Data (2013–2016) to Standardized Fields

Each year, the research team revisits the Excel-based sales survey form sent to distributors to capture any new product categories and to look for opportunities to collapse categories to ease the reporting burden. Thus, the research team had to ensure that data from previous collection efforts (originally stored in Excel) mapped to the proper lamp and luminaire characteristics in the new database. (See the 2017 memo for tables summarizing this mapping and for specific cases when the category names changed in the 2016 version of the sales survey.)

Mapping Reduced Wattage Lamp Replacement Data to Standardized Fields

The Northwest Energy Efficiency Alliance (NEEA) Reduced Wattage Lamp Replacement (RWLR) initiative collected sales data on linear fluorescent and tubular LED (TLED) lamps for all participating distributors through 2020. NEEA cleaned this data to identify the share of 4-foot T8 lamps that are reduced wattage (25W and 28W) and standard wattage (32W). Participating distributors submitted sales data for all 4-foot T8 lamps, including T12 and T5 lamps, on a monthly basis. As NEEA's goals and data structure differed from this project's goals and data structure, the research team had to map NEEA's data to the standardized fields in the database and, in some cases, add granularity to the data provided. The research team took the following steps in this process:

1. **Identify and eliminate products that are not linear fluorescent or TLED.** Because the research team does not have complete product data for these RWLR participants outside of linear fluorescent and TLED data, the additional sales in other categories are incomplete. The research team removes these sales because the research team cannot verify that reported sales in non-linear fluorescent and TLED categories represent all of that distributor's sales in those additional product categories. The research team also removed black light lamps, gold tubes, and germicidal lamps. Some distributors submitted sales of hardware and wiring, CFLs, and metal halide lamps; the research team also removed these products. The research team used the following fields to identify products to exclude, in the order listed:
    a. Technology type
    b. Bulb description
    c. Bulb type
2. **Standardize the naming conventions for incorporation in the SQL server database.** In some cases, multiple product descriptions mapped to the same category in the SQL server database. Table C-1 provides an example of 15 unique combinations of the NEEA fields "Shape" and "Category" that map to U-shape lamps. In this step, the research team expanded definitions to incorporate new categories: UL Type A/B TLEDs, and UL Type C TLEDs.

Table C-1. Inconsistent Naming Convention Example: U-Shape Lamps

Shape	Category
T8-6U	T8LEDU
T8-6U	Other
U-Bent	Other
T8-1-5/8 (U-Bend)	Other
T8-1-5/8 (U-Bend)	U-Bend-T8
T8-6U	U-Bend-T8
T8-6U	32W
U-Bend	Other
T8-U	T8
T8-U	32W
T8 U-Bend	Other
T8-6U	T8-6U
T8-6U	T8
T8-U	T8LEDU
T8 U-Bend	U-Bend-T8

Source: NEEA RWLR database

3. **Review for incorrectly categorized lamps.** Table C-2 provides examples of the classifications given to four model numbers that correspond with a single LED lamp type (with two different correlated color temperatures). In two cases, these lamps were incorrectly identified as 4-foot fluorescent lamps, and, in two cases, they were correctly classified as LED U-shape lamps. The research team reviewed online manufacturer catalogs to verify correct product classification.

Table C-2: Incorrect Product Categorization Example

Model Number	Description	Shape	Technology	Watts	Lumen	Length
PHIL 16.5T8/244000 IF6U		T8-6U	Fluorescent	16.5W	1900	48
PHIL 16.5T8/24-4000 IF-6U		T8-6U	Fluorescent	16.5W	1900	48
PHIL 16.5T8/245000 IF6U	Fluorescent - Tube - T8	T8	LED	16.5W	1950/2150/290*	24
PHIL 16.5T8/24-5000 IF-6U	Fluorescent - Tube - T8	T8	LED	16.5W	1950	24

**Lumen output with different ballast options.*
Source: NEEA RWLR database

4. **Map data to SQL data categories.** The research team used the following data to map to the SQL Server database linear fluorescent and TLED data categories:

a. Lighting technology type (LED or linear fluorescent)

b. Lamp shape (T8, T5, T12, or LED tubes)

c. Lamp wattage

d. Lamp length

e. T8 series (for 4-foot or 2-foot U-shape T8 lamps between 29W and 32W: 700 series or 800 series based on color rendering index)

f. UL Type for TLEDs. The two main UL Types are Type A which is ballast compatible and often referred to as Plug-and-Play and Type B which requires bypassing the ballast. This year, the research team also added Type A/B Dual Mode/Hybrid and Type C Remote Driver TLEDs.

For any lamp where one of these fields was missing, the research team used text strings from other product fields or an online web search to fill in the missing data. Table C-3 summarizes the linear fluorescent and LED categories in the SQL Server database.

Table C-3: Standardized Data Fields for Linear Lamps

Lighting Technology Type	General Category	Dimension	Subcategory
Linear Fluorescent	T12	4-foot	34W
Linear Fluorescent	T12	4-foot	40W
Linear Fluorescent	T12	4-foot	Other
Linear Fluorescent	T12	4-foot	U-Shape
Linear Fluorescent	T8 - High Performance 800 Series or Better	4-foot	25W
Linear Fluorescent	T8 - High Performance 800 Series or Better	4-foot	28W
Linear Fluorescent	T8 - High Performance 800 Series or Better	4-foot	32W
Linear Fluorescent	T8 - High Performance 800 Series or Better	4-foot	Other
Linear Fluorescent	T8 - Standard 700 Series	4-foot	32W
Linear Fluorescent	T8 - Standard 700 Series	4-foot	Other
Linear Fluorescent	T8 - Standard 700 Series	4-foot	U-Shape
Linear Fluorescent	T5	4-foot	28W
Linear Fluorescent	T5	4-foot	54W
Linear Fluorescent	T5	4-foot	Other
Linear Fluorescent	T12	8-foot	Slimline
Linear Fluorescent	T12	8-foot	High Output
Linear Fluorescent	T12	8-foot	Other
Linear Fluorescent	T8	8-foot	Slimline
Linear Fluorescent	T8	8-foot	High Output
Linear Fluorescent	T8	8-foot	Other
LED	LED Tubes	4-foot	UL Type A - Plug-and-Play/Direct Replacement
LED	LED Tubes	4-foot	UL Type B - Ballast Bypass
LED	LED Tubes	4-foot	UL Type A/B – Dual Mode/Hybrid
LED	LED Tubes	4-foot	UL Type C – Remote Driver
LED	LED Tubes	4-foot	T5 Replacements
LED	LED Tubes	4-foot	Other

Source: Distributor sales data structure

Adjusting for Linear Product Overrepresentation

Over the course of the study, linear fluorescent sales have been slightly over-represented relative to other technologies. The NEEA RWLR data the research team collected included linear lamp sales from non-participating distributors (i.e., distributors that did not submit full-category lighting data). Linear product sales may be overrepresented relative to other technology types because these distributors only submitted linear fluorescent and TLED (linear product) data and did not submit sales data for other categories. The research team created an adjustment to address the overrepresentation and ensure the sales mix reported in the adjusted data accurately reflects the best estimate of the overall product mix of the lighting market.

The research team identified four distributors that only provided linear product data through RWLR in the study period. The research team had two core priorities in developing the adjustment:

- Adjust the *overall quantities* of linear products to account for over-representation from these four distributors and ensure the *overall* product sales mix across technologies is accurate
- Maintain the *sales mix* within the linear product category to incorporate the information from these RWLR-only participants and ensure the data reflect the most accurate sales mix

The research team developed an annual percentage adjustment to linear product sales data to reduce the quantity of these lamps and better reflect the best estimate of the overall product mix. The linear adjustment is included in the associated Market Results Spreadsheet[13] as *Linear Adjustment* and is multiplied by the original *Sales Qty* column to derive an adjusted value (*Sales Qty with Linear Adjustment*). The adjustment is segmented by year, so it preserves the mix within the linear product category—the adjustment utilizes the strong RWLR data while providing an unbiased estimation of overall sales mix. Table C-1 shows the adjustments by year.

[13] Contact Juan Carlos Blacker (jcblacker@bpa.org) for more information on the Market Results Spreadsheet.

Table C-1: Linear Adjustments by Year

Year	Linear Adjustment
2013	93.1%
2014	91.8%
2015	92.7%
2016	94.8%
2017	95.2%
2018	95.6%
2019	95.3%
2020	95.5%
2021	96.0%

The largest impact of this adjustment is in 2014 where linear products are adjusted from 55% to 53% of the overall sales mix. The smallest impact is in 2021 where linear products are adjusted from 30% to 24% of the overall sales mix. The adjustments the research team developed are equal to the total linear product sales in each year excluding the four distributors, divided by the total linear product sales in each year including the four distributors (e.g., in 2021, 4% of total linear product sales were accounted for by these four distributors, resulting in an adjustment of 100% - 4% = 96%).

Processing Raw Sales Data (Data Dumps) and Mapping to Standardized Fields

In 2018, 2019, and 2020, three distributors did not use the Excel-based form to provide sales. The research team used a combination of R and SQL logic, web scraping, and manual classification to map these raw data to the SQL Server database categories. Lighting product descriptions and model numbers do not follow consistent formatting and often vary by manufacturer, making automated classification of products often just as time intensive as a strictly manual approach.

The research team used Python, R, and SQL to:

1. Identify products from previous years leveraging existing mappings to classify as many products as possible
2. Leverage an automated process to lookup model numbers on known manufacturer websites
3. Extract specific product characteristics from model number and product description fields

The research team manually mapped products using a combination of available data fields and online model number searches for products that the R and SQL code could not easily map.