

Memorandum

To: Jessica Aiona and Ben Mabee, Bonneville Power Administration

From: Kate Donaldson, Cadeo

Date: July 1, 2020

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, as well as the resulting data strengths and uncertainties. The content is organized in the following sections:

- Data Summary
- Outreach and Data Collection Process
- Data Structure
- Sources of Uncertainty
- Data Representativeness

Data Summary

The 2020 non-residential lighting data collection effort gathered distributor data from 2016-2019. Analysis included 2013 through 2015 data collected in previous studies. A total of 23 distributors submitted data to the research team in 2020. 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. There are three types of distributor business models: Full Line, MRO & Online, and Lighting Consulting. Full Line distributors sell a variety of electrical products, including lighting and maintenance, repair, and operations (MRO) products. MRO & Online distributors may sell only lighting or a mix of products with a focus on maintenance sales. Lighting Consulting distributors focus on energy efficiency products and project-based work. Project-based work includes new construction, renovation, and retrofit sales, which tend to be more efficient than maintenance sales.

Population data comes from the regional distributor database Bonneville Power Administration (BPA) compiled in 2016. In an effort to keep population data up to date, the research team updated firmographic data for distributors participating in the 2018, 2019, and 2020 studies. These updates were not comprehensive, as many data fields were populated from past in-depth distributor interviews and other research tasks. But, these changes added to the overall number of distributors represented in the

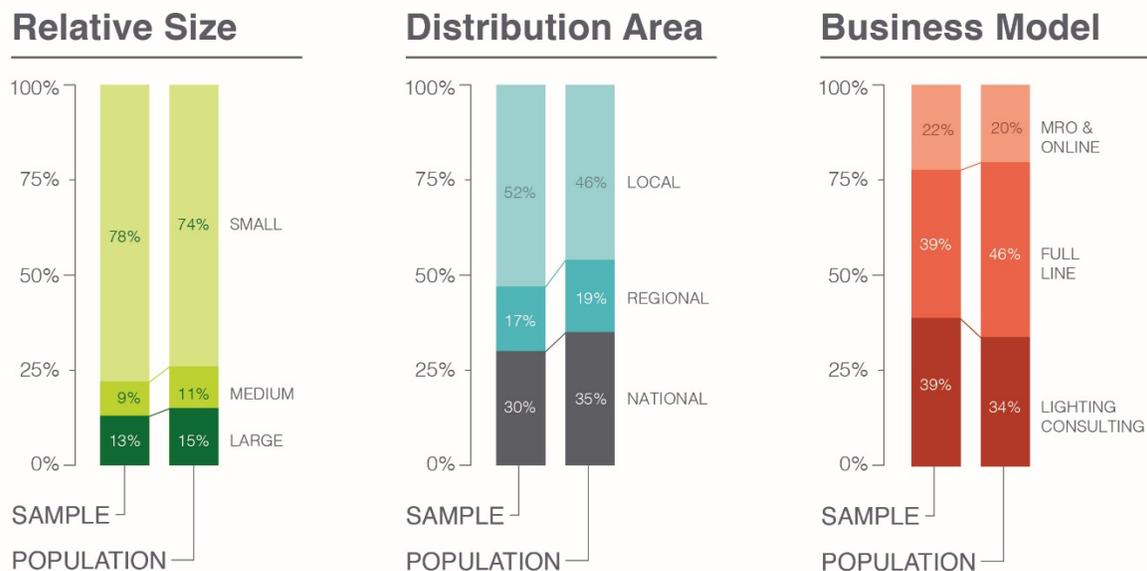
database and provided updated population distributions in Figure 1 and shares of branch locations by state in Figure 2 below.

Distributors participating in 2020 represented the following, shown in Figure 1:

- A representative mix of distributor sizes, relative to the population
- A slightly larger portion of local distributors and a slightly smaller portion of national distributors, relative to the population
- A slightly smaller portion of Full Line distributors and a slightly larger portion of Lighting Consulting distributors, relative to the population

Many of the new distributors participating in the last two studies were lighting consultants that only sell LEDs, which increased the representation of this business model type as well. This trend could indicate that lighting consulting businesses are becoming more common in the region. The research team will monitor this trend in future studies.

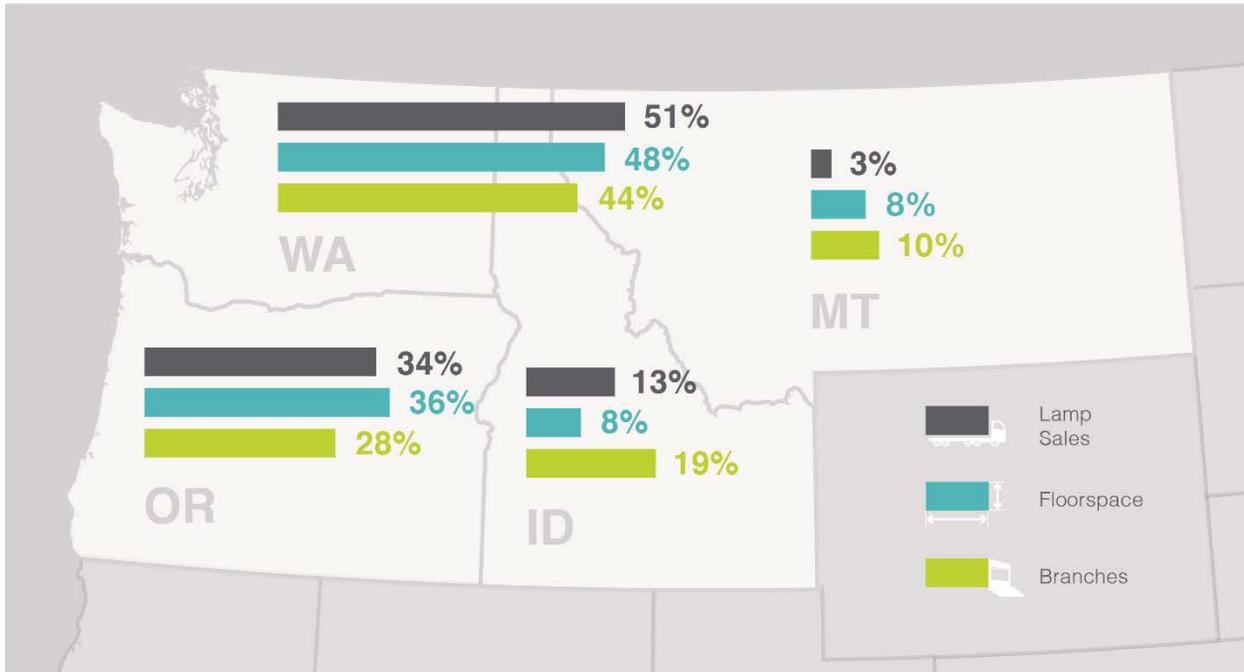
Figure 1: Mix of Distributors Submitting Data Compared to the Pacific Northwest Distributor Population



Source: Analysis of distributor interviews and online research as compiled in 2016 distributor database with updates to participating distributors from the 2018, 2019, and 2020 studies

Figure 2 shows lamp shipments by state, with state shares of total commercial floor space, and known distributor branch locations by state in the 2020 distributor sample for context. Relative to commercial floor space, sales quantity appears to be largely representative of the region, with most shipments going to Washington (51%) and Oregon (34%). With the updates to the distributor database, the research team did identify an increase in branch locations in Idaho, mirrored by an increase in sales data collected from Idaho branches beginning in 2018 and continuing in 2019.

Figure 2: Distributor Lamp Shipments, Floor Space, and Branches by State: 2019



Source: Distributor sales data analysis, 2019 Commercial Building Stock Assessment floor space by state, and 2016 distributor database with updates to participating distributors from the 2018, 2019, and 2020 studies.

Summary of Data Gaps

The research team identified the following data gaps based on the 2020 sales data participation and analysis:

- **Rapid LED tube growth between 2015 and 2018 increased uncertainty in extrapolating missing data.** The percentage of all LED sales extrapolated grew in sales year 2016, in part because the team assumed that non-participating distributors also saw significant sales growth in LED tubes, which was the fastest-growing product category between 2015 and 2018. Given the consistent growth across participating distributors and growth projections from past participants' prior submissions, the team believes this is a reasonable assumption. However, given this is an area of great change, there is more uncertainty for this product type than others. Additionally, the percentage of extrapolated LED tubes in 2019 was similar to 2018. In 2019, LED tube sales decreased by 9%, a shift from the rapid growth between 2015 and 2018. This could be due in part to a reduction in utility incentives targeting this product.
- **The 2018 and 2019 data requests improved the certainty in reporting on share of distributor sales to residential and non-residential customers, but uncertainty remains for some individual product types and for years prior to 2018.** This year, the research team asked participating distributors to provide the breakout of residential and non-residential sales for each technology and some more granular splits (e.g., LED downlights, LED tubes). This helped address uncertainty within each technology type. However, these updates only apply for the 2018 and 2019 participants and do not address historical uncertainties. In past years, distributors provided sector-split estimates of their sales at the technology level (e.g., CFL, incandescent, and LED). In

prior years, distributors provided a single estimate for each technology that the team applied to all sales. However, individual products within these technologies may be more or less likely to sell into a particular sector. For example, pin-base CFLs and some LED fixtures are most likely non-residential, whereas residential customers are more likely to purchase screw-in A-Type and reflector lamps. These mixes could also change over time. While some distributors reported that they used customer data to inform their sector split estimates, most provided approximate splits based on their judgment, creating additional uncertainty in these data.

- **The research team did not collect sales data from any specialized online distributors, but some online sales may already be reflected in the collected data.** One long-time online distributor stopped participating in the study in 2018, and since then the team has been unable to collect data from this segment. This year, the team sought to determine the size of this channel and, specifically, to determine Amazon's role in the commercial lighting space. In a webinar for traditional distributors put on by Oracle NetSuite, the team confirmed distributors are concerned about Amazon taking market share from traditional distribution. To protect the traditional distribution channel, distributors were encouraged both to seek out partnership opportunities with Amazon to use their online platform to facilitate online sales and to build out their own branded online strategies so they were not reliant on Amazon for these sales. In talking to one distributor that sells products on Amazon's platform, the team believes that the nonresidential lighting data collected may already contain online sales from traditional distributors. For example, one distributor has said anecdotally that online sales from its website represent around 50% of their annual sales. In future years of the study, the team will work to further define the size and importance of this sales channel as well as the percentage of online sales already submitted in the data collected in this study.
- **The research team did not collect sales data from any specialized national account distributors or manufacturers.** The team has made multiple attempts in this and previous study years to connect with national account distributors but has yet to secure participation from any of them. While 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 & 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 the impact on representation of national account sales.
- **Three longtime participants chose not to participate this year.** The team recruited three new distributors this year while losing three relatively small participants, which resulted in a small increase in the overall percent of sales the team had to extrapolate (see Table 1 below). One larger distributor has not participated in the last two years, but that distributor submits full-year, full-category data to NEEA. Because the team already has that distributor's annual sales data, additional extrapolation work is minimal.

Table 1: Percentage of Data Extrapolated by Year and Technology

Lighting Technology Type	2013	2014	2015	2016	2017	2018	2019	All Years
CFL	40%	26%	2%	17%	15%	17%	18%	19%
Halogen	24%	13%	1%	11%	15%	23%	24%	15%
High Intensity Discharge (HID)	7%	4%	1%	8%	11%	12%	13%	7%
Incandescent	30%	20%	10%	16%	16%	20%	21%	19%
LED	22%	13%	13%	22%	26%	27%	30%	25%
Linear Fluorescent	25%	5%	10%	5%	12%	12%	15%	12%
All Technologies	26%	11%	8%	12%	18%	20%	23%	17%

Source: Distributor sales data analysis

Considerations for Data Applicability

The research team offers the following considerations for BPA’s use of this sales data in future market analyses:

- Data completeness—and therefore reliability—generally increases until 2016, after which fewer distributors participated.** Inconsistent distributor participation contributes to uncertainty. The research team has developed an extrapolation methodology to estimate sales for distributors who contributed data in the past but did not contribute 2017, 2018, or 2019 data. The extrapolation process allowed the research team to estimate missing distributors’ sales based on historic market share and other participants’ sales. This allows the team to leverage data from past and present participants but does not fully resolve the uncertainty due to limited data from some distributors. Participation was relatively consistent in 2018 and 2019.
- Linear fluorescent sales remain the most reliable subset of the dataset to support quantitative analyses.** This is due to consistent availability of linear fluorescent data—largely due to the NEEA Reduced Wattage Lamp Replacement (RWLR) initiative—and the fact that the research team has consistently collected information on these two technology types since the first data collection effort in 2013.
- Linear fluorescent sales may be overrepresented relative to other technologies.** Analyses using this data should acknowledge the fact that linear fluorescent sales may be slightly overrepresented relative to other technology types based on the number of RWLR participants that only provide linear fluorescent data.
- This dataset best represents the technology mix of the distributor sales channel, which may not apply to other sales channels.** As discussed in previous memos on sales data gaps, the distributor sales channel is the largest single flow of lighting products in the non-residential market. However, if sales mixes are significantly different in the manufacturer direct or retail channels, and these channels become a significant portion of non-residential sales, this data could misrepresent some technologies.

The Sources of Uncertainty and Data Representativeness sections provide additional discussion of specific data strengths and weaknesses.

Outreach and Data Collection Process

Similar to the data collection strategy from the last three years, the research team worked in collaboration with two key partners, NEEA's RWLR Initiative and Evergreen Consulting (Evergreen), to facilitate distributor outreach. This group, led by BPA, made up this year's outreach team, which conducted outreach and recruited distributors to participate in the data collection effort.

Applying lessons learned from previous years, Evergreen maintained a prominent role in outreach and was responsible for initial engagement of most distributors. The two exceptions were RWLR participants, who were engaged by NEEA, and select longtime participants, who were engaged directly by Cadeo. The outreach team's outreach strategy followed these general steps:

1. One of the three outreach organizations – Evergreen, NEEA, or Cadeo – contacted distributors and asked them to participate, giving them a FAQ/information packet that BPA developed. Participating distributors were offered a monetary incentive of \$500-\$1,000, depending on the promptness of their data submission.
2. If the distributor agreed to participate, the research team followed up with the distributor with the sales data collection form, instructions for submitting data via ShareFile, and an optional non-disclosure agreement.
3. The research team logged all communication in the distributor data tracker and provided outreach updates in the 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. As in previous years, the outreach team coordinated communication with distributors to maximize data submissions. The team made the following improvements in this year's outreach process based on lessons learned from last year:

- **Continued early communication with corporate contacts.** Although it is efficient to engage larger corporations to collect data from multiple branches through a single point of contact, it can require multiple conversations to secure their willingness to participate. Beginning outreach to these contacts earlier in the process helped ensure participation during the study's timeframe and resulted in a corporate data submission covering multiple branches.
- **Established direct communication with long term participants.** In the 2019 study and again in the 2020 study, the research team at Cadeo reached out directly to select long term participants who proved highly responsive in past studies. The research team aimed to organize outreach assignments in order to reduce the number of touchpoints for distributors and streamline the total number of outreach steps where applicable. However, the research team highly values the relationships and connections that NEEA and Evergreen have and will only directly reach out to a limited number of distributors where applicable. NEEA, Evergreen, and Cadeo discuss and agree upon all outreach assignments at the beginning of each study.

The team also continued a change to the type of data collected started in 2019. In years prior to 2019, the only data the research team received from the NEEA RWLR initiative database was linear fluorescent and TLED data. However, in recent years NEEA has worked with distributors that participate 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 (with the exception of controls) for participating distributors, not just LFL And TLED data. This change is part of NEEA’s ongoing development of a distributor platform that supports distributor partnerships on multiple initiatives.

For the 2020 study, three long time non-res lighting data collection participants agreed to use their 2019 data dumps previously submitted to NEEA towards the data collection effort. For these three distributors, the research team analyzed and mapped the 2019 data sets to the standardized survey form categories to identify any gaps in the 2019 data. The team then sent these three distributors a modified survey form to fill in data gaps and confirm trends in their sales data. Despite the modified data collection approach, the distributors received the same \$1,000 incentive amount as other participating distributors. The benefits of this approach are twofold. First, this approach minimizes touchpoints with distributors, so they are not asked for the same data twice. Second, if these distributors decline to participate or become unresponsive to requests, the research team already has their annual data which reduces the amount of extrapolation in analysis.

Despite these benefits, there are drawbacks to this approach. In the 2019 and 2020 studies, mapping these three new format data sets was very time intensive. In both years, mapping these three data dumps took two of the three months in the outreach phase of the study, leaving only one month for these distributors to respond to data requests. These distributors work with NEEA on a variety of initiatives and often require more than a month’s notice to complete additional data requests in a timely manner. It is possible that this format will be less time intensive to work with in the future as the team is able to re-use mapping methodologies year over year. However, if distributors modify their product offerings by adding or removing product types from their catalogs, the team will need to map all new products and modify the mapping process for any removed products, increasing the level of effort needed for these year-over-year submissions. Furthermore, the addition of any new data dump submissions will add time and effort for that initial mapping.

Final Outreach Disposition and Results

Table 2 provides details on the final disposition for the distributors included in the team’s outreach.

Table 2: Summary of Distributor Outreach: Final Disposition

Category	Number of Distributors
Total distributors included in outreach*	68
Distributors submitting data	23
• Repeat participants	20
• New participants	3
Distributors declining to participate	9
• Lack of time and/or interest	5
• Data reporting limitations	0
• Prohibited by company policy	0
• Extenuating circumstances (Covid-19 or other)	4
Distributors unresponsive to outreach	36

*Includes individual distributor branches that operate independently

Source: *Distributor outreach tracking*

This year, 23 distributors submitted data, slightly higher than 2019's 22 participants. This year also saw a very small number of distributors outright decline to participate. About half of the time, these distributors declined after initially agreeing to participate either by ceasing to respond to outreach efforts or citing a lack of time or interest. The other half of declining distributors referenced extenuating circumstances. For example, one of these distributors filed for bankruptcy in early 2020 and did not have the time or staff to complete the request. Two more distributors declined due to extenuating circumstances related to Covid-19, whether that be due to lack of time due to reduced staff or due to temporarily closing lighting divisions.

While the number of declines this year was lower than in prior years, a higher number of distributors were unresponsive to outreach efforts entirely this year in relation to last year. The team assumes that a portion of this increase is due to Covid-19 and other extenuating circumstances. Four distributors that provided data in prior years did not provide data in 2020. However, one of these distributors submitted a 2019 data dump to NEEA prior to the study, so the analysis team was able to use their 2019 data leaving only three additional distributors that required extrapolation for 2019.

Outreach Lessons Learned/Recommendations

- Follow up with long-time participants who do not participate in the current year's study to clarify sales magnitude and aid with extrapolation.** In some cases, long-time participants respond to outreach to let the team know they cannot participate in the current study (this year, for example, one distributor did not participate due to staffing limitations during Covid-19). In these cases, the team should follow up with the non-participating distributor to gather information about the magnitude of their sales and product offerings to aid the extrapolation effort. For example, the team could reach out to ask if the distributor saw any substantial changes to their sales in key technology categories or if they stopped selling any products from the prior

year to the current year. With their past data years and the answers to these follow up questions, the team can more accurately extrapolate that distributor's data.

- **Proactively gather NEEA data sets for mapping ahead of outreach.** The research team should work with NEEA to receive these data sets as soon as possible to begin the mapping process. If possible in contracting, NEEA and the research team should plan to collect these data dumps and coordinate with NEEA and D+R to collect these datasets as soon as final 2020 data is submitted. The research team can then start mapping as soon as possible at the beginning of 2021. That way, these steps can occur ahead of outreach to both extend the mapping timeline for the Cadeo team and to give these distributors ample time to complete the data request. This will become increasingly important if other distributors that work with NEEA may submit data dumps to the research team as part of future studies.
- **Continue to seek out corporate or division level contacts for data collection.** In the past three years, the research team received data from one distributor for over 40 branches in the Northwest. That approach consolidated outreach under one contact while resulting in a significant amount of sales data from a large organization. The team recommends trying to replicate this approach with distributors where applicable (i.e., high number of regional branches within a corporate territory).
- **Develop and implement a targeted strategy for defining the online sales channel and collecting online sales data.** The team made an effort to better understand how traditional distribution is being changed by online market actors like Amazon and how large of a channel online sales represent. The team needs to continue this work by understanding if and how much of the sales data collected in this study includes online sales from participating distributors. The team also needs to continue to build knowledge of the changing online landscape. Potential strategies for addressing these questions include:
 - Modify the data collection instrument to ask distributors what percentage of their sales come from online sales, if online sales are included in their reported data, and if those sales are included in the sales reported through the study
 - Compile a list of follow up questions devoted to online sales to send to participating distributors that indicate online sales in their data collection instrument
 - Determine specific outreach plans for reaching online distributors (e.g. Amazon, LED King, 1000bulbs.com)
- **Continue engaging national account distributors, and consider a relationship building strategy.** The team has made multiple attempts to connect with national account distributors but has yet to secure participation from any of them. NEEA and BPA should consider investing additional effort during future research or data collection to attract and incent national account distributors to share their sales data. Based on past qualitative research, the team estimates that this channel could represent anywhere from 5% to 15% of non-residential lighting sales. Potential strategies for attracting participants in this market segment include:
 - Collaborating with NEEA to develop a coordinated outreach strategy, which may include value for the distributor in the form of opportunities to participate in NEEA's initiatives.
 - Developing a streamlined data request that would significantly reduce the level of effort required for national account distributors to participate.

- **Consider collecting updated firmographic information about distributors in the next round of data collection.** The 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, the team recorded 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 team check for potential biases in the data set. Last year, the team updated what firmographic information was easily found on distributor websites for those participants in the 2018 and 2019 outreach efforts. This year, the team repeated the process exclusively for new participants in the 2020 study. The information for other distributors in the database has not been updated since 2016, and NEEA and BPA could consider collecting updated firmographics in a future round of data collection.

Data Structure

Distributors submitted 2019 sales data in three different formats in 2020:

- Using the standardized survey form provided
- Through NEEA’s lighting initiative database
- In raw data extracts of all sales by model number and/or product description

The research team merged the data from all three formats with historic data submissions from 2013, 2014, 2015, 2016, 2017, and 2018 in a single SQL Server database. Collectively, these data span sales from 2013-2019. The research team used five fields to organize the sales data by lamp and luminaire characteristics, which Table 3 summarizes.

Table 3: 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 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: Summary of Unique Database Category Entries provides a complete list of possible entries for these fields, and Appendix B: Detailed Data Cleaning Approach describes the analysis process.

Under the 2013 data collection effort, distributors provided data back to 2010. But, for this year's analysis, the research team focused only on sales years 2013-2019. Data quality and number of participants increased significantly in 2013 due to expanded outreach and revisions to the data collection instrument, and BPA has received feedback from distributors that more recent data trends are more valuable market intelligence than analysis of data back to 2010.

Sources of Uncertainty

The analysis team found two types of data gaps within individual distributors' data submissions:

- **Missing years and technologies.** Some distributors were not able to participate in all years or provide data for certain technologies. This increases uncertainty for the missing years and technologies. The three situations that led to data gaps included:
 - New participants only provided data back to 2016, resulting in missing 2013-2015 data. In each year of the study, distributors are only asked to provide back data for four years. This is a deliberate choice by the research team to minimize the distributor's effort.
 - Some participants stopped participating resulting in missing sales.
 - The team added new sales data categories in 2016, 2017, 2018, and 2019. Not all participants provided data for these new categories for years prior to 2015 and 2016. These categories are as follows:
 - In 2016, the research team added CFL A-Type lamps. Participants in years prior to 2015 have only provided CFL A-Type data since sales year 2015.
 - In 2017, the research team added flood lights, halogen MR16 lamps, and decorative lamps. There is little data for these categories in 2013 and distributors who participated in 2016 and prior did not all provide data for sales years 2014 and 2015 for these categories.
 - In 2018, the research team added 2-pin and 4-pin classifications to CFL pin-based lamps. Previously, those lamps were reported in aggregate regardless of pin base.
 - In 2019, the research team added breakouts for TLED Type C and Type A-B lamps. Previously, those lamps were reported in aggregate in the Other TLED category. The team also added Medium Base Corn Lamps to LED Industrial Applications.
- **No data on percent of sales to non-residential sector.** Some distributors did not indicate what percent of their sales go to residential versus non-residential customers. For these distributors, the research team assumed that all sales for all technologies went to non-residential customers. This assumption is consistent with prior years' analysis and with the research team's qualitative understanding of those distributors' sales. This question was a new addition to the 2016 data collection form.

The following sections describe how the team addressed the most significant data gap: missing years and technologies.

Missing Years and Technologies

The research team applied the same methodology to extrapolate missing sales data for each combination of distributor and technology as in previous years. The team leveraged distributor-specific sales data to inform the extrapolation process. Because many distributors provided sales data back to 2015, the team updated previous extrapolation results from 2015-2018, along with extrapolating 2019 for the first time. This methodology rests on two key assumptions:

- First, that **distributors' market share of each general product category remains relatively constant** over time relative to other distributors of their type. For example, if Distributor A represented an average of 25% of all full-line CFL A-Type lamp sales in 2018, then the team assumed they sold 25% of all full-line CFL A-Type lamps in 2019.
 - This assumption has the effect of scaling extrapolated sales to trends in other distributors' sales, ensuring that the analysis captures technology shifts seen at the distributor-type level. Continuing the example above, if reported CFL A-Type sales declined over time, maintaining a constant market share would force extrapolated sales to decline as well.
 - This does not account for the fact that distributors may gain or lose market share over time due to business reasons. For example, a full-line distributor may invest in an aggressive sales strategy that increases their market share among full-line distributors for a specific product category.
- Second, that **the mix of subcategories (lamp shape, wattage, and base type) for a given distributor remains relatively constant over time**. For example, if participating distributors' sales show that in 2018 5% of full-line 4-foot T8 lamp sales are 28W lamps, the analysis assumes that missing distributors' 4-foot T8 lamp sales were also 5% 28W lamps in 2019. Subcategories for each technology and general category are provided in Appendix A.

The team developed code in R to calculate consistent market shares for each distributor. This code ensured that market shares for each data gap by general technology were filled with the previous-year market share for that general technology for each distributor.¹ With the total sales in a technology category filled out for each distributor based on their respective previous-year market share, the research team applied the subcategory breakdown to the extrapolated numbers using that particular distributor's subcategory breakdown.

In the prior year's 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 can produce distortion in the extrapolated data results. Therefore, the team decided to continue using the previous-year method discussed above. The team verified that using the previous year's market share was the best approach by assessing how variable distributor market shares are within technologies. The team focused this assessment on LED and linear fluorescent technologies, since they are the fastest changing technologies and comprise the largest shares of total market sales. Overall market share for most distributors for these technologies shifted less than one percent from year to year. This means that in 95% of cases, the previous-year assumption would have been correct (within 1%) for

¹ New distributors are not back-cast to years prior to the first year included in their submissions.

distributors that provided data in all study years. There were two outlying distributors in the dataset that showed dramatic market shares changes in LEDs over the study period (both increasing and decreasing). However, expanding the one-year extrapolation methodology to a larger time period would increase the 'lag' in market share estimation, resulting in a worse extrapolation. Based on this assessment, the team determined that the previous-year method is a reasonable approach and preferable to the three-year method. While growth or decline over the study period can be large, year-over-year changes relative to the overall market are still quite small, which minimizes distortion introduced through extrapolation.

Data Representativeness

As shown in Figure 1, the mix of distributor types in the population differed from the mix of distributor types that submitted data. All sales data represent an unweighted mix of sales reported by all participating distributors.² In 2016, the 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 models defined below. However, the team and BPA determined that weighting data submissions to reflect the relative presence of these three business models in the market would not improve the accuracy of the sales data, because there is too much uncertainty around non-participating distributors' market shares. The research team did not collect sufficient new data for 2019 to change this conclusion.

The research team defines each distributor type 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
- **MRO & 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

² In 2016, the research team investigated the effect of adjusting the distributor type weights to better align with the overall market share analysis. The research team used the sales data, branch counts, and publicly available financial reports for those distributors that submitted data to estimate market shares for distributors that did not submit data as part of the 2016 analysis.

- May sell only lighting or a variety of products

The following sections outline possible data collection and analysis improvements that could lead to greater certainty around representativeness in future years.

Representativeness Assessment

The research team reviewed this year's participation data and the cumulative data collected since 2013 to determine the extent that the dataset accurately represents the non-residential lighting market. This review had two components: representation within specific segments (e.g., individual distributor types) and representation across segments channels (e.g., sales beyond traditional distributors). The team's findings include:

- Geographic Coverage
 - As seen in Figure 2 above, lamp sales collected in 2019 are approximately proportional to each state's share of regional commercial floor space and distributor branches. The team uses floor space and branches as comparisons because we believe these two metrics should approximately correlate with sales.
 - There is some variation, as the three bars representing lamp sales, floor space, and branches do not align perfectly in each state, but each metric is proportional, confirming that we have reasonable representation from each state.
- Distributor Characteristic Mix
 - 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. The team uses these comparisons to evaluate whether our participating distributor pool is reasonably representative of the total population of distributors in the region. The team found:
 - A representative mix of distributor sizes and business models
 - A slightly larger portion of local distributors and a slightly smaller portion of national distributors participate relative to the northwest population
 - A slightly smaller portion of Full Line distributors and a slightly larger portion of Lighting Consulting distributors, relative to the population
- Participation
 - Inconsistent participation from distributors does introduce concerns about representativeness. However, the team applies the extrapolation methodology to reduce these gaps. Participation in 2019 was consistent with previous years.
 - There is a possibility that nonparticipating distributors sales look different than participating distributors sales. Extrapolation can soften this risk through filling gaps with known market trends, but it is still a possible data gap.
- Specific segments

- As in previous years' data collection, the team understands the submitted data for the Full Line, MRO & Online, and Lighting Consulting segments to be representative of total sales in these distributor types. While it is not possible to assess the sales of nonparticipating distributors to confirm this statement, the team's prior market research indicates that sales patterns within distributor type are generally similar. Therefore, the team believes it is reasonable to assume that submitted data are representative within each distributor type.
- As the team was not able to collect any sales data from national account distributors or other manufacturer direct sales, the sales mix of these channels remains unknown.
- Across segments and channels
 - The research team believes that traditional distribution through the three distributor segments defined in this memo represents most lighting sales in the Pacific Northwest, but this is difficult to confirm given the uncertainty around three aspects of the market:
 - The size of the Lighting Consulting segment
 - The mix and volume of products from the manufacturer direct sales channel
 - The mix and volume of products from the retail channel
 - The magnitude of the Lighting Consulting segment remains unknown, but the team still believes this segment is much smaller than the other two based on the number of distributor consultants identified in the market and the low level of sales from those who have submitted sales data. This is partially a product of the Lighting Consulting business model; these distributors sell retrofit projects at a smaller volume than the larger distributors selling to new construction and retrofit markets. Therefore, the research team believes this segment has a relatively small effect on the overall reliability of the sales data.
 - The relative size of the manufacturer direct channel remains unknown.
 - BPA has conducted additional qualitative research to better understand national account distributors. However, this research did not produce a quantitative assessment of the relative size of this sales channel in the non-residential lighting market.

Recommendations for Addressing Representation Gaps

The primary way the research team will address representation gaps will be through the outreach recommendations above. These recommendations will allow the team to strategically target distributors in underrepresented or hard to reach segments from lessons learned over the previous seven study years.

As part of the upcoming BPA nonresidential lighting market model update, the team will review the 2019 Commercial Building Stock Assessment (CBSA) to compare the commercial lighting stock in the CBSA data against the stock levels calculated in the BPA market model. This review process will provide the team with empirical data to verify sales assumptions made in the lighting market model and the trends observed in the annual data collection study. Any discrepancies between the CBSA findings and the

market model and data collection findings will highlight opportunities where additional data collection efforts may have a significant impact on data representativeness.

Appendix A: Summary of Unique Database Category Entries

The following tables summarize the possible entries for the General_Category, Dimension, Subcategory, and Base_Type fields for each lighting technology type.

Table 4: Lighting Technology Type: Linear Fluorescent

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

Table 5: 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 6: 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
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 7: 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

Table 8: 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 9: 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 10: Controls

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

Appendix B: 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 an 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,³ Oregon, and Washington), branch information was included for datasets spanning multiple locations, and sales totals were given for both units and dollars. The team also checked that the data did not violate any data validation rules.
- 2. Data gaps:** The research team reviewed all datasets, 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

³ 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.

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

Mapping Historic Data (2013-2015) 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 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 NEEA RWLR Data to Standardized Fields

NEEA's RWLR initiative collects sales data on linear fluorescent and tubular LED (TLED) lamps for all its participating distributors. NEEA cleans this data to identify the share of 4-foot T8 lamps that are reduced wattage (25W and 28W) and standard wattage (32W). Participating distributors must submit sales data for all 4-foot T8 lamps, including T12 and T5 lamps, on a monthly basis. As NEEA's goals and data structure differ 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.** For example, some distributors also submitted sales of CFLs and metal halide lamps. The research team also removed black light lamps, gold tubes, and germicidal lamps. The 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 11 provides an example of 15 unique combinations of the NEEA fields “Shape” and “Category” that map to U-shape lamps. In this step, the team expanded definitions to incorporate new categories –UL Type A/B TLEDs, and UL Type C TLEDs.

Table 11: Inconsistent Naming Convention Example: U-Shape Lamps

Shape	Category
T8-6U	T8LEDU
T8-6U	Other
U-Bend	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 12 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 12: 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 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 13 summarizes the linear fluorescent and LED categories in the SQL Server database.

Table 13: 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

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

Four 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.