Creating Indoor Water Use Benchmarks for Commercial, Industrial, and Institutional (CII) Customers: Methodology Explainer

Background and Goals

Thornton, CO is home to approximately 148,000 people living in Colorado's Front Range, with a drinking water service area population of 165,000. Currently, the City's utility billing database does not distinguish between business types within the commercial customer class, making it challenging to determine water use trends for different business categories and to project future water demand.

Thornton worked with WaterNow through its Project Accelerator program, with a goal of determining appropriate water use benchmarks for CII customers by business type as a foundation for establishing new CII water conservation goals and developing a more equitable tap fee structure.

The goal of this document is to explain WaterNow's methodology and learnings in simple terms, to support communities who may be interested in pursuing their own process of creating water use efficiency benchmarks for their commercial, institutional, and industrial users.

Setting Goals and Purpose

A first step in the process of water use benchmark creation is to identify the potential uses of the benchmarks. The structure and necessary statistical strength of the analysis will depend on the eventual use of the benchmarks created, as benchmarks which will be used primarily for, e.g., outreach and education to customers, may not require as much confidence as metrics used for setting water budgets or creating tap fees.

WaterNow's Accelerator project with the City of Thornton intended to use the new water efficiency benchmarks for outreach to commercial customers in the short term, and as a way of creating informed goals for individual users' efficiency. Ultimately, the benchmarks created in this process will be used to inform the development of customer category-specific tap fees for new development. For more information about utility readiness for taking on the benchmarking process, and how a local approach may vary depending on goals and purpose, see the Water Research Foundation's Water Use Analysis Guide for Commercial and Institutional Efficiency.

Selecting Customer Categories

Once goals for the benchmarking process have been established, a next step is identifying the categories of CII customer for which benchmarks will be created. Each category will be benchmarked separately. In selecting categories, consider:

- Any regulatory guidance. If your community is required to develop CII benchmarks by law, the legislation will likely provide guidance about which categories you should use. In particular, California communities should reference guidance aimed atcompliance with the Long Term Framework (SB 606/AB1668) Performance Measures. See for example this CalWEP Customer Classification Guide.
- The literature. The 2019 Water Research Foundation paper Developing Water Use Metrics for the Commercial and Institutional Sectors contains suggestions for 10 primary categories for CII water use benchmarking which may be helpful in selecting locally relevant categories.
- The local context. If there are user types which stand out as important priorities in local water use patterns in your CII landscape, creating categories focused on those users may be a way of capturing additional information about their usage.

CII Customer Categories for Thornton, CO

- Animal Clinic
- Auto Services
- Bank
- Car Wash
- City Building
- Eating/Drinking Place
- Gas Station
- Healthcare Facility
- Hotel/Motel
- Light Industrial
- Mini Storage
- Mixed Use
- Multifamily Residential
- Office
- Pre-K/Daycare
- Public Recreation
- Public Safety
- Religious Institution
- Retail Outlet
- School/College
- Senior Living
- Service-oriented Shopping

Keeping in mind the size of the community and the amount of available data may also be helpful in determining the categories selected. Ten to twenty categories is a reasonable range, but opting for fewer categories may be preferable in smaller communities or instances where less data is available, to maximize the data available and the robustness of each category's benchmark, or more in situations where a wide range of particular uses need to be captured. See the box above for the categories selected as part of the Thornton Accelerator project.

Data Selection and Collection

The next step in developing benchmarks for CII customers is to determine the amount and type of data available, and to collect that data.

There are three key types of data which are needed for producing CII water use benchmarks: (1) water use data, (2) category data, and (3) normalizing data.

Water use data records the amount of water used by each CII customer over time. We recommend capturing at least three years of historical data, but two years or even one year can be enough to create initial benchmarks. This data type is typically sourced from utility billing records. Some things to be aware of:

- Most utilities bill their CII customers monthly. Monthly water use data is a good fit for benchmarking.
- Many utilities will collect identifying information for each CII customer, including a unique ID, the name of the business or institution, and the address of the property being served. These can be helpful in confirming categories and cleaning the data – more on that later.
- In western states, summer irrigation produces a pattern of high use in the summer, with baseline levels of use in the winter. The goals of WaterNow's work in Thornton meant that we were most interested in indoor use, so we isolated winter months where no irrigation was taking place. In the dataset we used, these were the billing records from December through March. However, this may vary depending on whether a community's data is organized by billing date or use date, as well as local irrigation timing. We recommend assessing the average monthly use across the entire dataset to identify the seasonal pattern, and selecting only those months which do not show significant increases over the winter baseline.

Category data helps to identify which customer category (e.g., Retail Outlet, Restaurant, Hotel) a given meter falls into. This data type can be sourced from a variety of places, including county tax assessor data, city sales tax NAICS codes, county or city GIS departments, (WaterNow's selection for the Thornton Accelerator project) or existing internal utility records. Some things to keep in mind:

- Data sourced from county tax assessors is generally accurate and helpful, but may not align perfectly with the categories selected for benchmarks. See the section on Categorization and Data Cleaning below for more information on selecting categories and aligning your data with selections.
- Information created by the utility to categorize CII customers may be more likely to align with the categories the utility is interested in benchmarking, but should be carefully checked for accuracy. Data that has not been updated frequently may contain misidentified businesses or business categories. Comparing these data sources against tax assessor data or other external sources may be a helpful step in the benchmarking process.
- If using external data, it may be necessary to cross-reference water users against parcels in tax assessor data in order to match users with their identifying category. This can be done through geospatial analysis, e.g. matching the geographic location of a meter to the geographic location of a parcel in the parcel-level data using Geographical Information System (GIS) software. In WaterNow's analysis of Thornton's CII users, Thornton's GIS staff was able to perform this cross-reference by spatially joining the locations of Thornton's meters with the parcels included in the City's data. In some cases this involved combining water usage for multiple meters on the same parcel, where, e.g. separate tenants in a shopping center or strip mall each have their own meter. Situations like this, with various users on one parcel, can present difficulties in benchmark creation – see the section on categorization below.

Normalization data allows water use to be compared fairly across different-sized users within a category. For example, a very large school may use more water than a small school even if the large school is highly efficient and the small school is not. Scaling users' water use by their size, or "normalization," allows comparisons of efficiency across a user category. Sources for this data may include county tax assessor data, fire marshal data, or in some cases data gathered directly from customers or by the project team through additional research.

- The most commonly used parameter for normalizing users is building square footage (*n.b.: total square footage, rather than building footprint, is the best measure of a building's size*). In general, customers within a category use more water in proportion with the size of their buildings. Building square footage for all parcels is often available from a county's tax assessor or GIS department, making this an easily obtainable and straightforward normalization factor. In rural counties or other areas where building square footage may not be available for free, data vendors exist which may be able to supply this information for a fee.

- Studies have shown that some customer types, including restaurants, hotels, and schools, can be normalized more accurately by alternate factors targeting the number of individuals served. The total number of seats in a restaurant, the total number of rooms in a hotel, and the total number of students at a school can be more effective normalizing factors for these users. Alternative normalization factors like these can be sourced through local research or from, e.g., local fire marshal's offices, which may collect it already.
- In collecting the data, decisions may need to be made about format and data handling tools. In general, non-spatial data is available in Microsoft Excel's data format (.xlsx files) and the more generic Comma Separated Variable (.csv) file type. Spatial data is available in a variety of file types, most often as ESRI Shapefiles or GeoJSON files. WaterNow used the open source research tool and programming language R to handle, filter, and manipulate utility billing water usage and customer category data, and the open source GIS tool QGIS to work with the spatial components of the parcel data in the data cleaning and verification process. A practitioner implementing a CII benchmarking process may opt for different tools based on their own experience and expertise; manipulating data in Microsoft Excel or using a more complete programming language like Python may be preferable if the user is more comfortable with one of these tools.

Categorization and Data Review

Once data has been collected and the categories selected, the next step is to sort users by category. This can be done using the category data sourced in the previous step. In the Thornton analysis, WaterNow categorized users based on the "parcel activity" column included in the city development parcel database, combining some activities into single categories, and in some cases, separating a single activity into several categories based on business name and further research. WaterNow's analysis in Thornton also excluded mixed-use parcels, where several businesses share a single parcel ID and use category in the City's GIS data. These types of parcels show a wide range of use rates, varying with the relative sizes and use rates of the individual businesses which make them up, and so are not a good candidate for the standard benchmarking process. Future iterations on benchmark creation may include attempts to categorize mixed-use parcels by their dominant water user type, separate them into individual businesses, or other approaches that may allow them to be included in the benchmarking analysis.

In order to verify that as many users as possible are correctly categorized, some level of data review and cleaning will be necessary. In the most extreme case, each entry in the dataset can be confirmed by comparing the user's name to its assigned category; however, this may be prohibitively time-intensive depending on the number of users in the dataset. In order to clean the data while minimizing the investment of time and effort, methods for targeting likely miscategorized users can be helpful. In WaterNow's analysis, preliminary rate-of-use statistics were calculated for each water user by dividing each user's winter usage by its square footage. Users in each category were then screened for outliers using the interquartile method, a statistical approach that sets up a "fence" scaled to the size of the interquartile range, or the difference between the 25th and 75th percentile users. All outliers detected using this method were screened for incorrect categorization by review of and recategorized, or excluded, based on the findings. Following this step, the preliminary benchmarks were re-created and outliers evaluated a second time to catch any additional miscategorized or anomalous entries.

Conducting Benchmark Analysis

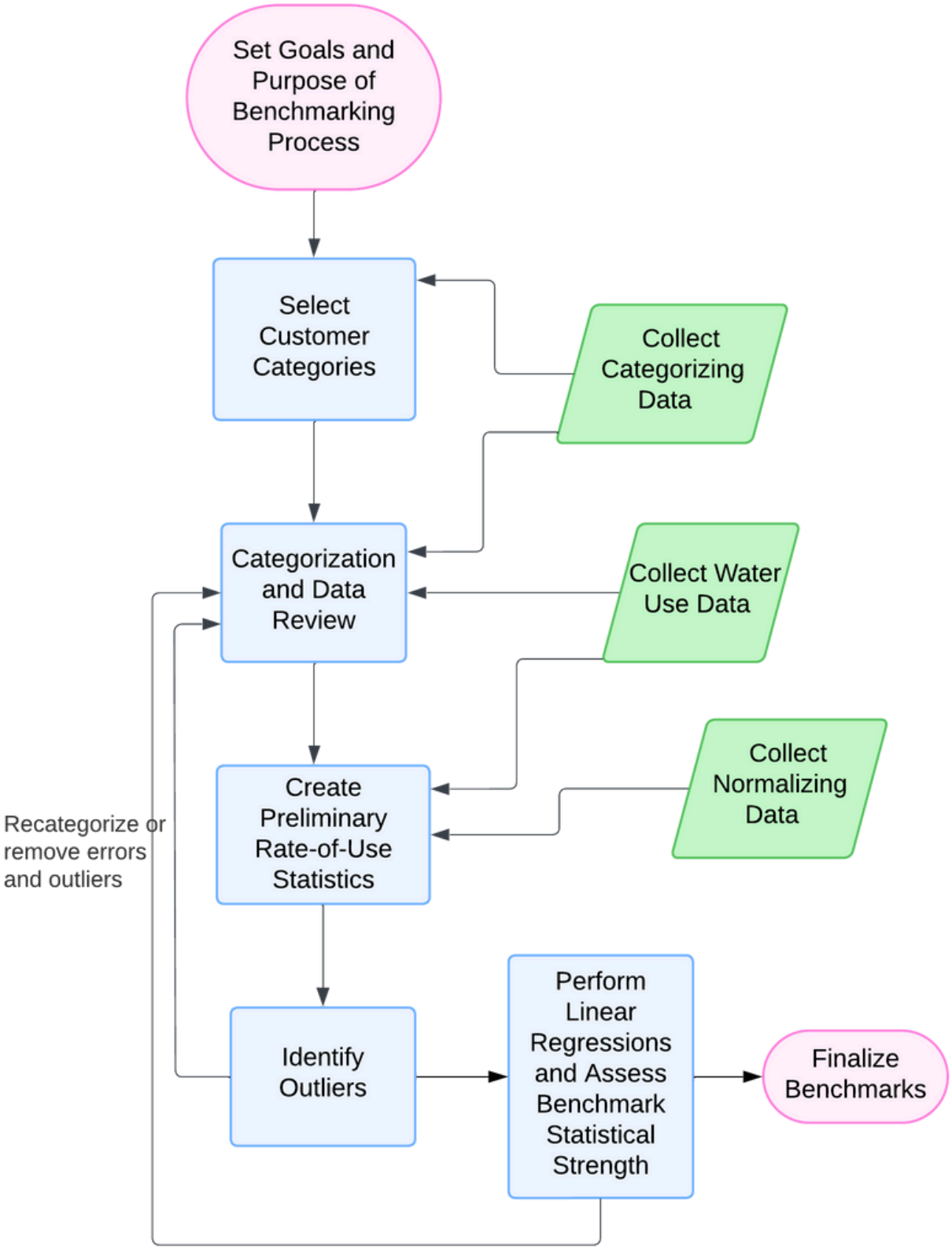
With the categories correctly evaluated, the next step is to create a benchmark for each category. Approaches to creating the benchmark can vary depending on the intended use of the finished benchmark: in cases where the benchmark is being used for planning purposes, such as water budgeting for future development, a slight overestimation may be helpful, to account for uncertainty and to avoid underestimating the amount of demand that will result from development. On the other hand, where benchmarks will be used primarily to set goals for existing users, aligning the benchmark to the median user and providing information about high-efficiency users along with it may be the most relevant approach.

Calculating a descriptive benchmark can be as simple as calculating the average gallons of indoor use per square foot for each user in a category. This can be accomplished by totaling the use during winter months, dividing by the number of months, and multiplying by twelve to establish average annual indoor use. This figure can then be divided by the normalizing factor – in most cases, the square footage of the building. If using alternate normalizing factors such as the number of students in a school or the number of seats in a restaurant, simply divide by the alternate factor instead. Once this rate-of-use metric has been calculated for each user in a category, the median rate-of-use for the category provides a simple benchmark for that category.

A more statistically robust benchmark can be created using linear regression. A regression reveals the strength of the relationship between the size of a building and its water use for each category. The regression produces an R2 value, which represents the proportion of variation in the data that can be explained by the normalizing variable. A high R2 value (in our analysis, we set a threshold of 0.3, while other studies have used 0.5 as a threshold) indicates a strong relationship and a reliable benchmark, while a low R2 value suggests a weaker relationship between size and water use, suggesting that other variables may be responsible for differences between users in a category. A linear regression also provides a way to build a safety factor into the benchmark – the upper limit of the 95% confidence interval can be used in situations where a benchmark is desired to err on the side of overestimation of water use.

Putting It All Together

The flowchart on the following page captures the entire process of data collection, categorization, and benchmark creation. Green boxes indicate places where data enters the process, blue boxes represent data manipulation steps, and pink circles represent high-level decision-making steps at the beginning and end of the process.



Putting It All Together: *This flowchart summarizes the process of creating CII Water Use Efficiency Benchmarks.*

References and Resources

A wide range of resources are available to support communities in learning more about the CII benchmarking process, potential uses for benchmarks and approaches to creating them, as well as accompanying programs for outreach and education with CII water users focused on water conservation. A selection of these resources are included here:

Methodology Reports / Guides:

- Kiefer, Jack C, Lisa R Krentz, and Benedykt Dziegielewski. "Methodology for Evaluating Water Use in the Commercial, Institutional, and Industrial Sectors." Water Research Foundation, 2015. https://www.waterrf.org/research/projects/methodology-evaluating-water-use-commercial-institutional-and-industrial-sectors.
- Fedak, Rebecca, Derek Hannon, Zach Taylor, and Amy Volckens. "Developing Water Use Metrics for the Commercial and Institutional Sectors." The Water Research Foundation, 2019. https://www.waterrf.org/research/projects/developing-water-use-metrics-commercial-and-institutional-sectors.
- Volckens, Amy, Katie Kershman, Derek Hannon, Zach Taylor, Rebecca Fedak, and Russ Sands. "Water Use Analysis Guide for Commercial and Institutional Efficiency." The Water Research Foundation, 2019. https://www.waterrf.org/resource/water-use-analysis-guide-commercial-and-institutional-efficiency.

Examples of Regional CII Benchmark Development:

- Brendle Group. "Benchmarking Task Force Collaboration for Industrial, Commercial & Institutional (ICI) Water Conservation." Colorado WaterWise, June 2007. https://coloradowaterwise.org/Resources/Documents/ICI_toolkit/docs/Brendle Group and CWW ICI Benchmarking Study.pdf.
- Dziegielewski, Benedykt, Jack C Kiefer, Eva M Opitz, Gregory A Porter, Glen L Lantz, William B DeOreo, Peter W Mayer, and John Olaf Nelson. "Commercial and Institutional End Uses of Water." AWWA Research Foundation and the American Water Works Association, 2000. https://www.waterrf.org/research/projects/residential-commercial-and-institutional-end-uses-water.
- Morales, Miguel A., James P. Heaney, Kenneth R. Friedman, and Jacqueline M. Martin. "Estimating Commercial, Industrial, and Institutional Water Use on the Basis of Heated Building Area." Journal AWWA 103, no. 6 (2011): 84–96. https://doi.org/10.1002/j.1551-8833.2011.tb11475.x.

Alternative Methods

- A&N Technical Services, Inc. "Analysis of Water Use Efficiency Metrics and Benchmarking." Mesa Water District, October 2013. https://www.epa.gov/system/files/documents/2024-03/ws-commercial-bmp-watersenseatwork_section2.3_benchmarking.pdf.
- -Morales, Miguel, and James Heaney. "Benchmarking Nonresidential Water Use Efficiency Using Parcel-Level Data." Journal of Water Resources Planning and Management 142, no. 3 (March 1, 2016): 04015064. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000616.

Existing CII Efficiency Programs

- Dziegielewski, Benedykt. "National Survey of Commercial, Industrial and Institutional Water Efficiency Programs," 2016. https://www.awwa.org/wp-content/uploads/AWWAs-Utility-Survey-of-CII-Water-Efficiency-Programs-Report.pdf.
- US EPA. "Water Efficiency in the Commercial and Insitutional Sector: Considerations for a WaterSense Program." US EPA, August 20, 2009. https://www.epa.gov/sites/default/files/2017-03/documents/ws-commercial-ci-whitepaper.pdf.

CII Classification

- Cuellar, Lisa, Michelle Maddaus, Linda Higgins, and Chris Tull. "Update on CII Water Use Classification." March 2022. https://calwep.org/wp-content/uploads/2022/03/06-CalWEP-Winter-Plenary-2022_CII-Classifications_FINAL.pdf.