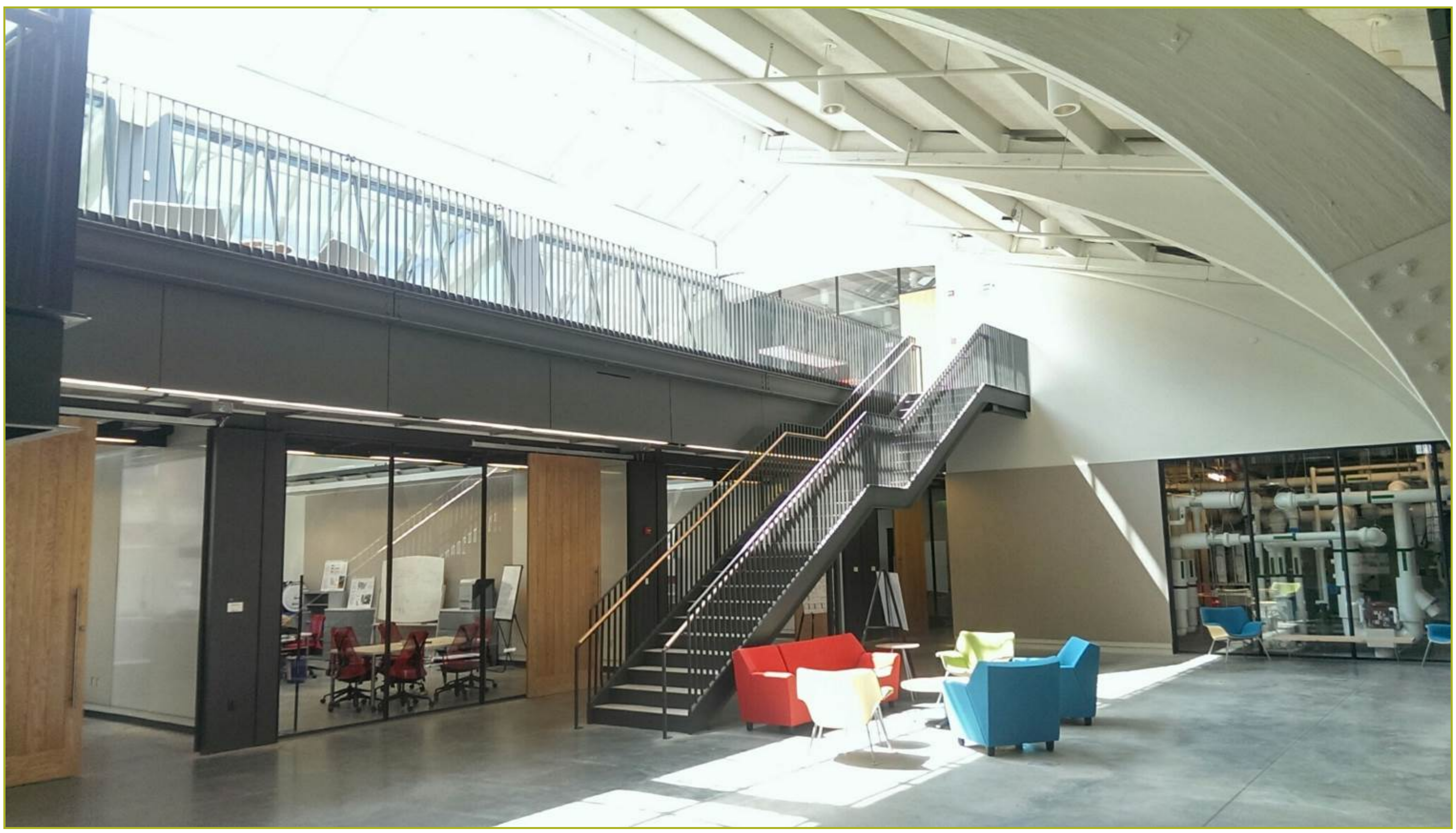
CONSORTIUM for
BUILDING ENERGY
INNOVATION

REPORT

Title: Introductory Level – A Guide to Community-Wide Benchmarking Analysis

Report Date: April 29, 2016

Report Author: Scott Wagner



Funded by U.S. DOE

Report Abstract

Project objectives.

Intended for property owners, portfolio managers and government officials, the *Introductory Level – A Guide to Community-Wide Benchmarking Analysis* guide offers strategies and guidance for data cleansing, parsing and basic analysis of energy performance and associated energy costs. This guide also allows benchmarking program managers the ability to develop an analytical foundation to understand what data has been collected via a benchmarking program and how this data can support efforts to improve the energy efficiency of various types of building stock.

Contact Information for lead researcher

Name: Scott Wagner
Institution: The Pennsylvania State University
Email address: swagner@engr.pse.edu
Phone number: 267-574-4239

Contributors

Ben Cohen – Penn State University

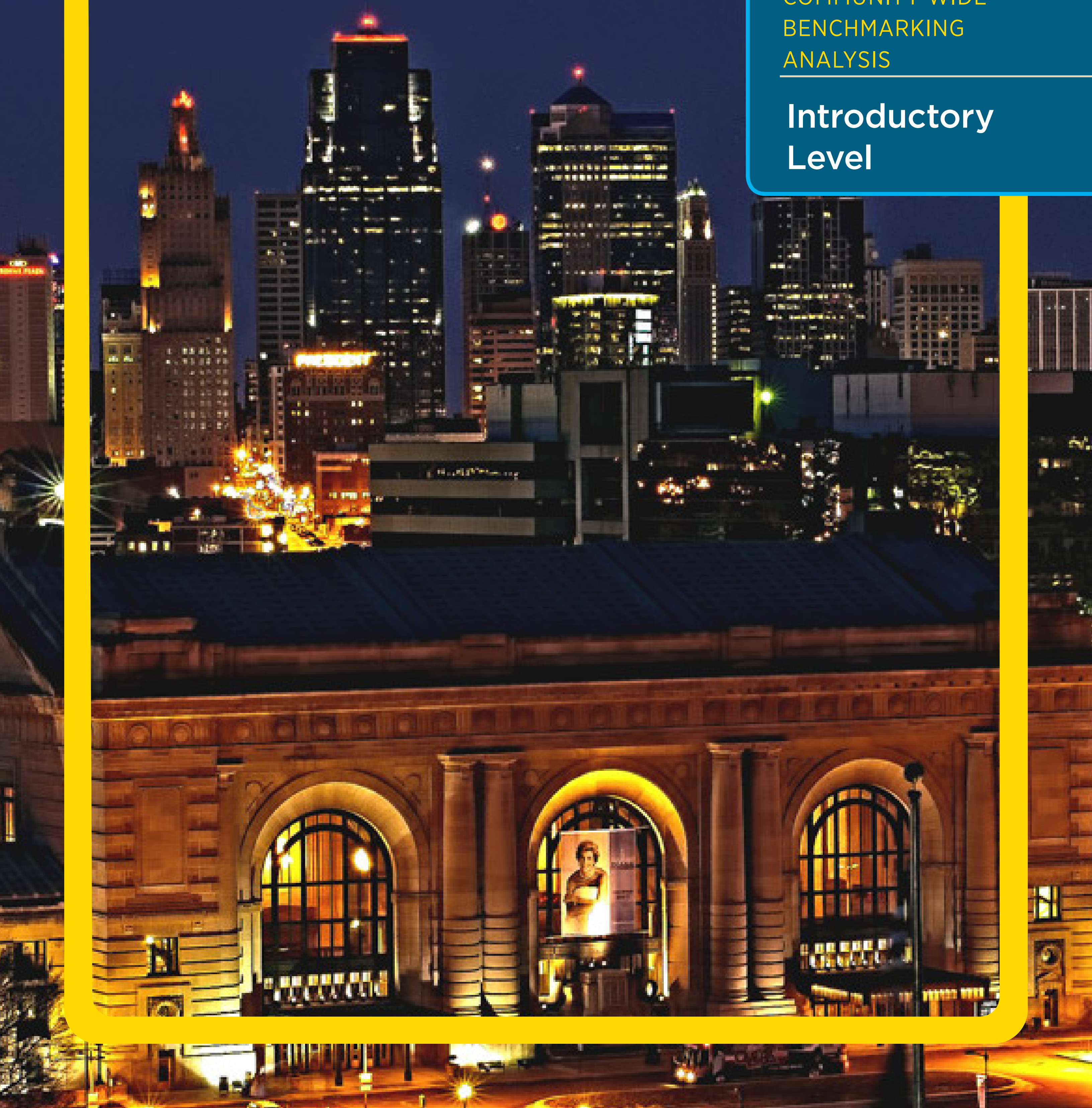
Mona Kahlil - DOE

CONSORTIUM for BUILDING ENERGY INNOVATION

U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy

A GUIDE TO COMMUNITY-WIDE BENCHMARKING ANALYSIS

Introductory Level



EXECUTIVE SUMMARY

Commercial buildings in the United States consume nearly half of all building-related energy use and roughly 20 percent of total energy consumption and greenhouse gas emissions. Experts have long believed that making the energy used by buildings more transparent is an essential step in helping to curb this energy use. Benchmarking building energy usage makes this information available to the owner and the market.

Property and portfolio managers can use benchmarking to measure how efficiently a building uses energy, to compare their buildings to other similar buildings, and to help identify building performance trends and opportunities for investment and energy savings. A large number of states and local governments regularly benchmark their building stock to track their energy performance and measure progress toward energy reduction goals. In addition, several states and a host of local governments have adopted community-wide benchmarking policies, programs, and initiatives that include the comparative measurement of commercial building energy performance across a portfolio of buildings across their jurisdictions.

This publication is the first of three documents that together make up a three-part series entitled "*Guide to Commercial Building Energy Benchmarking Analysis*," This guide was developed to help benchmarking program administrators analyze building energy benchmarking data and thereby develop data-driven strategies for prioritizing energy efficiency investments. The introductory level guide provides strategies and guidance for data cleansing, parsing and basic analysis of energy performance and associated energy costs. The intermediate level guide provides further approaches for understanding of how ESPM scores relate to fuel use and identifying and selecting a subset of inefficient properties from overall benchmarking data. The advanced level guide

Specifically, the introductory Level guide presents the essentials for analyzing benchmarking data, which includes the following focus areas:

a. **Cleansing data.** *Identifying suspect data and resolving data errors.* For benchmarking to be effective, the data that is collected must be accurate. This publication describes steps an analyst should take to ensure data quality before performing data analysis to better understand trends within the data.

b. **Data parsing.** *Breaking down data by building physical characteristics and use types.* The simplest form of data analytics, parsing benchmarking data gives fast, high-level insight into the entire building portfolio. The introductory level guide gives examples of the most common data parsing analyses.

c. **Evaluating basic energy performance.** *Assessing and understanding performance of benchmarked building stock, evaluating fuel mixes and energy costs, and analyzing energy use intensity alongside ENERGY STAR Scores*

Front Cover Photo Credit: Flickr User photoguyinmo, "Union Station in Kansas City Missouri at night"

Contents

Chapter 1: Introduction

Building energy benchmarking is a growing practice in the U.S. As of 2014, ten cities, one county, and two states have adopted benchmarking policies, programs, and initiatives that include the comparative measurement of commercial building energy performance across a portfolio of buildings.[1] Other organizations, such as real estate investment trusts and governmental agencies, have also initiated benchmarking programs for building portfolios under their purview.

As jurisdictions adopt these policies and programs, the body of building performance data continues to grow. With this growth comes the need to analyze the data meaningfully and accurately, as well as the need to identify clear "next steps" to realize energy and cost savings.

BACKGROUND

Commercial buildings comprise nearly half of building energy use and roughly 20% of total energy consumption and greenhouse gas (GHG) emissions in the United States. These energy expenditures amount to more than two dollars per square foot, providing significant opportunity for savings.[2,3]

To understand and manage building energy costs, data must be collected, evaluated, and analyzed in a way that facilitates decision making. Cost-effective energy reductions are often hindered by a lack of building energy use information. Benchmarking makes energy performance visible.

BOX 1. WHAT IS BENCHMARKING

Benchmarking is the process of measuring and comparing a building's energy performance to its energy baseline, i.e. to the energy performance of similar types of buildings (based on use, such as comparing the energy performance of an office building to that of other office buildings). Energy use is typically measured on a per-square-foot basis and normalized for a range of factors, such as building size, operational characteristics, tenancy, and climate. Benchmarking can be used to compare performance over time, within and between peer groups, or to document savings from installed energy conservation measures.

When combined with disclosure of energy performance data, benchmarking can be a powerful market-based policy tool for increasing building energy performance. By building awareness of building energy use among stakeholders, benchmarking and disclosure can catalyze demand for energy efficiency improvements.

For resources on benchmarking, benchmarking and disclosure policy, strategic energy management, energy use data access, and other related topics, visit DOE's State and Local Solution Center: energy.gov/eere/slsc/state-and-local-solution-center.

For example, in 2007, officials in Arlington County, Virginia, implemented a building benchmarking and energy efficiency improvement program for publicly owned buildings as part of the Arlington Initiative to Reduce Emissions. This initiative reduced the energy intensity of the county's building stock by 15% from 2007 to 2012, resulting in $850,000 of avoided energy costs each year.[4]

GUIDE OBJECTIVES

This guide is intended to help building managers, planners, and other officials understand, prioritize, and conduct the analyses of public and/or private building energy benchmarking data. These analyses may be required as part of a benchmarking and disclosure law or policy, as part of a voluntary or mandatory energy efficiency program, or to better understand the energy profile of a portfolio of buildings.

The analyses assume the use of off-the-shelf office computer software; the accompanying descriptions are presented with easy-to-understand, non-technical language; however, a basic understanding of statistics is helpful. Specifically, this guide:

- Introduces benchmarking data analysis concepts and provides a background to benchmarking
- Describes data collection field requirements
- Discusses considerations for evaluating data integrity and quality to remove suspect entries
- Outlines analysis techniques (supported by examples where appropriate)
- Provides considerations for how to interpret analysis results and identify or analyze trend

[1] For a detailed list of enacted state and local benchmarking policies, visit "BuildingRating.org." 2013. www.buildingrating.org.

[2] U.S. Department of Energy (DOE). Buildings Energy Data Book, Chapter 3. March 2011. http://buildingsdatabook.eren.doe.gov/ChapterIntro3.aspx.

[3] U.S. Environmental Protection Agency, Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2009, Table ES-8. April 2011. www.epa.gov/climatechange/emissions/usinventoryreport.html.

[4] Arlington County, Virginia. Memorandum: Achieving the AIRE 2012 Goal. (May 23, 2013). http://freshaireva.us/wp-content/uploads/2013/05/052113-CMO-AIRE-2012-goal-report-Final.pdf.

INTENDED AUDIENCE

This guide is intended for private entities, cities, communities, or states that have implemented or are considering implementing a benchmarking and/or disclosure program or policy or, more generally, for those interested in:

- **Evaluating building energy data quality and gaps**
- **Better understanding system-wide building inventory and energy use profiles**
- **Evaluating energy savings potential**
- **Determining compliance with benchmarking requirements**

BOX 2. ABOUT THE ANALYSES IN THIS GUIDE

Each analysis described in this guide contains, in detail, the following key attributes that help the reader to understand how to replicate the analysis methodology, as well as understand the value and utility of the analysis:

- Insights gained. Why perform this analysis—what could be learned from it? Is this analysis critical, or more of a "nice to know" analysis?
- Required data and dependencies. What data fields are required? What other analyses are required to be performed first.
- Analysis how-to. A detailed step-by-step description of how to perform the analysis using standard off-the-shelf office computer software.
- Example. A real-world example of each analysis, including sample charts and graphs, utilizing real community energy use benchmarking data where possible.

USING ENERGY STAR PORTFOLIO MANAGER AS AN ENERGY BENCHMARKING TOOL:

Although a number of different building energy benchmarking tools are available (e.g., ENERGY STAR Portfolio Manager, EnergyIQ, and Lab21 Energy Benchmarking Tool), this guide assumes use of ENERGY STAR Portfolio Manager. ESPM allows users to track and benchmark the energy use for a variety of building types, which makes it a good choice for municipalities, housing authorities, and large real estate owners. Additionally, ESPM provides many types of benchmarking data and metrics, which allows for deeper levels of analyses than do other tools.

ESPM is a no-cost online benchmarking tool and is the most widely used software across the public and private commercial buildings sectors. It generates a statement of performance detailing a building's Energy Use Intensity (EUI). Energy Use Intensity is the amount of energy used by a building per square foot each year, often expressed in kBtu/sf/yr. The EUI of a building can be used to compare it to other peer buildings, allowing a better understanding of relative overall building energy efficiency. For many common building types, Portfolio Manager also scores each uploaded building on a scale from 1-100, enabling building owners to compare their property to similar buildings nationwide. Those buildings achieving an ESPM score of 75 or higher are eligible for ESPM certification which demonstrates that a building has been verified by a licensed Professional Engineer or Registered Architect as performing among the top 25 percent of similar buildings nationwide. Studies show that ENERGY STAR certification provides value to building owners, as they achieve higher occupancy rates, rental prices, and sale prices per square foot than non-certified buildings. As of December 31, 2014, more than 400,000 buildings, or over 40 percent of the total U.S. commercial building market, have been tracked using Portfolio Manager.

In light of this, CBEI and the DOE recommend using ESPM for most benchmarking efforts.

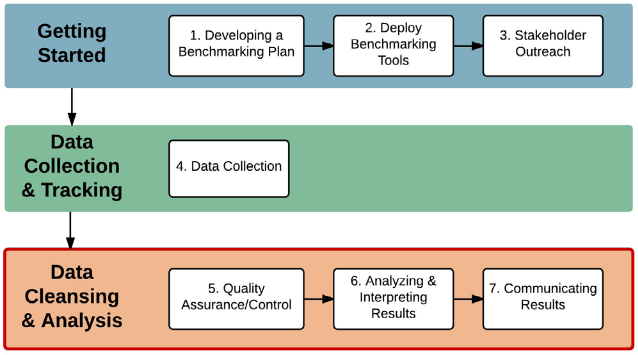


Figure 1. The full life-cycle of building benchmarking, from planning to communicating results. This report focuses on the Data Cleansing & Analysis phase.

[8] United States Environmental Protection Agency's ENERGYSTAR.gov website. What is Energy Use Intensity (EUI)? Accessed: 4/28/15. Available at: http://www.energystar.gov/buildings/facility-owners-and-managers/existing-buildings/use-portfolio-manager/understand-metrics/what-energy

[9] Miller, Norm (et. al). Does Green Pay Off? (July 2008). Accessed: 4/28/15. Available at: http://www.usgbc.org/Docs/Archive/General/Docs5537.pdf

[7] United States Environmental Protection Agency. Energy Use Benchmarking: Data Trends. Accessed: 4/28/15. Available at: http://www.energystar.gov/buildings/about-us/research-and-reports/portfolio-manager-datatrends

Chapter 2: How to Use This Guide

This guide is intended to help analysts, energy planners, or community officials understand, prioritize, and conduct energy performance benchmarking analyses. This guide assumes the reader has an available raw dataset in a standard database or spreadsheet format (e.g., through a dataset extracted from Portfolio Manager[8]) and therefore does not describe the data collection process.[9] Additionally, while tools exist that can help analysts cleanse and analyze data, this guide emphasizes the value of the analytic process itself. This guide focuses on a core set of analyses based on typical data collected through voluntary and mandatory benchmarking programs—jurisdictions may opt to perform additional analyses in addition to those covered by this guide.

COMPOSITION OF THE GUIDE:

To better address an audience with varying levels of knowledge regarding building energy use and benchmarking, this guide series is composed of three stand-alone sections, each at a different level of complexity: introductory, intermediate, and advanced.

Introductory Level Section: Intended for an audience either new to energy benchmarking or interested in gaining a general understanding of benchmarked building stock and its associated energy use. The guide provides strategies and guidance for data cleansing, parsing and basic analysis of energy performance and associated energy costs. *The guide shows how to analyze and develop enough benchmarking information that it may be used as the foundation of program reporting (such as for a municipality's annual benchmarking report).*

KEY STEPS

While jurisdictions' programs, policies, objectives, data, and resources may vary, the basic steps data analysts must take are the same. These are:

STEP 1: DEFINE THE ANALYSIS OBJECTIVES

Identify what you are trying to gain from the analysis—is it policy driven, program driven, or curiosity driven? For example, gaining a simple understanding of your building portfolio stock (e.g., comparing your benchmarking data building portfolio to a property tax database) would require far fewer analyses–and fewer resources—than identifying sector-specific energy savings potential and opportunities. The analysis objectives should drive what analyses to perform.

STEP 2: COLLECT AND CONFIRM DATASETS

Identify which datasets are available for analysis. These may include benchmarking data downloads from ESPM, property tax records, consolidated reports from utilities, among others. Evaluate available fields within each set, and confirm compatibility across datasets where appropriate (e.g., unique building identifiers). Confirm that time intervals are consistent across sets (e.g., monthly, annually) to ensure an "apples to apples" comparison.

STEP 3: PERFORM DATA QUALITY ANALYSIS

Following the guidance provided in Chapter 3, carefully and systematically review the datasets and remove erroneous data based on pre-determined parameters and cleansing criteria. Determine which buildings may require follow-up with building owners/operators regarding missing data.

STEP 4. DETERMINE ANALYSES TO PERFORM

Based on the objectives defined in Step 1, and available data fields identified in Step 2, refer to Chapter 4 to determine which analyses best suit your needs. This introductory guide covers analyses

[5] For an overview of EPA ENERGY STAR Portfolio Manager, visit: www.energystar.gov/portfoliomanager.

[6] For a list of resources benchmarking, benchmarking and disclosure policy, strategic energy management, energy use data access, and other related topics, visit DOE's State and Local Solution Center: www.eere.energy.gov/wip/solutioncenter/data_management.html.

to parse the data based on specific attributes (e.g. building types or square footage) and to assess basic energy performance.

STEP 5. CONDUCT ANALYSES AND COMMUNICATE RESULTS

Following the methods described in the introductory, intermediate, and advanced guides, perform the analyses using the tools of choice. Perform quality checks on the analyses to validate accuracy of results, and compare results with other publicly available data to "truth test" the data. Finally, communicate results based on the analysis objectives defined in Step 1.

TOOLS AND RESOURCES

The analyses in the guide were purposefully selected to be performed using only standard, off-the-shelf office computer software and available datasets.[10] Nonetheless, there are additional tools and resources available that are recommended to help perform the analyses and gain a better understanding of the results and underlying background, assumptions, and methodologies. These include the following, which are referenced throughout this guide:

- **Commercial Building Energy Consumption Survey (CBECS).** CBECS is a robust dataset that is a frequent source for commercial sector-wide EUI information. It is a compendium of national survey data collected on the stock of U.S. commercial buildings, energy-related building characteristics, and energy consumption and expenditures. The CBECS website contains maps, terminology, definitions, and additional information. www.eia.gov/consumption/commercial.
- **EPA ESPM.** Portfolio Manager is an online tool used to measure and track energy and water consumption, as well as GHG emissions. It benchmarks the performance of one building or a whole portfolio of buildings. Datasets from Portfolio Manager can be extracted and downloaded, allowing for additional custom analysis. The website contains a plethora of technical resources. www.energystar.gov/portfoliomanager.
- **Buildings Performance Database.** DOE's Buildings Performance Database (BPD) provides access to data on the actual energy performance and physical and operational characteristics of commercial and residential buildings. www.eere.energy.gov/buildings/commercial/bpd.html.
- **State and Local Solution Center.** DOE's State and Local Solution Center and associated Technical Assistance Program (TAP) provides resources, events, and one-on-one technical assistance to state, local, tribal, and K-12 school district leaders advancing high-impact clean energy policies, programs, and projects. The Solution Center contains many resources on benchmarking, benchmarking and disclosure policy, energy use data access, and other related topics. wip.energy.gov/solutioncenter.

[10] All example analyses performed in this guide utilized Microsoft Excel, though any standard spreadsheet and graphics software will suffice.

Chapter 3: Data Quality Analysis

In general, data quality analysis of benchmarking data can be broken down into two different areas: 1) Understanding why critical benchmarking data is missing, and 2) Data cleansing.

UNDERSTANDING WHY CRITICAL BENCHMARKING DATA IS MISSING PRIOR TO CLEANSING

To be most effective, a benchmarking program must obtain two critically important energy benchmarking metrics: the Energy Use Intensity (EUI) and, when applicable, the benchmark score. In reality, for most benchmarking programs, a significant subset of properties do not report one or both of these metrics. This section helps program managers and others identify why critical benchmarking data is missing or incorrectly supplied, making it possible to follow up with building owners/operators in order to improve data quality in the short- and long- term.

Understanding why critical data is missing from a benchmarking dataset usually requires a close examination of the data along with using some type of building "forensics." ESPM generates two alerts to improve data quality – "Energy Alerts" and "Property Use Detail Alerts" – that can typically explain why a metric was not calculated. Information contained in these alerts can help understand why critical benchmarking data was not generated.

A summary of the data quality issues that can be identified using a property's Energy Alerts and Property Use Detail Alerts fields when using ESPM can be found below. These issues fall into two categories: 1) When a property is eligible to receive an ESPM score, but no score or EUI was generated; and 2) When a property was not eligible to receive an ESPM score, but no EUI was generated. In most cases, resolution of these issues requires following up with building owners/operators (or third parties) who submitted the data to alert them to the missing data. Without follow-up, the person submitting the data may continue to submit incomplete data and the issue will reoccur each year.

DATA QUALITY ISSUE 1: PROPERTY IS ELIGIBLE TO RECEIVE AN ESPM SCORE, BUT NO SCORE OR EUI WAS GENERATED BY ESPM:

Check for the following issues:

1. **25% or more of gross floor area is associated with a non-eligible space use type;**
2. **Building/space details were not defined for the whole calendar year, e.g., the number of occupants was not defined throughout the year;**
3. **Certain input values were not within range, e.g., weekly operating hours were too low;**
4. **Energy use meters did not have monthly data for the whole year;**
5. **Energy use meters did not account for all energy usage of property;**
6. **Source EUI was determined to be out of range;**

DATA QUALITY ISSUE 2: PROPERTY IS *NOT* ELIGIBLE TO RECEIVE AN ESPM SCORE, BUT NO EUI WAS GENERATED:

Check for the following issues:

1. **Energy use meters did not have monthly data for the whole year;**
2. **Energy use meters do not account for all energy usage of property;**
3. **Source EUI was determined to be out of range;**
4. **Property type was reclassified to Not Available.**

A systematic approach to understanding the general quality of benchmarking data can prove to be very useful for a benchmarking program. In the Data Cleansing subsection, Table 2 can be used both as a way of separating good benchmarking data from suspect data and as a guide for offering feedback communication to building owners/operators that have submitted questionable data.

DATA CLEANSING

To identify suspect data, look for:

- **Very high or very low ESPM scores and EUIs**
- **Properties that should have received PM scores or just an EUI but did not**
- **Properties that used default input data**

Before analyzing benchmarking data, it is critically important to "cleanse" the data for accuracy and consistency. While this guide does not provide a rigorous "how-to" in regards to data cleansing, this section discusses the importance of accurate

data, provides a high-level walkthrough of the data cleansing process, and offers recommendations for data cleansing.

THE IMPORTANCE OF WORKING WITH CLEAN DATA

Data cleansing is the process of carefully and systematically reviewing benchmarking data and removing suspected erroneous data based on pre-determined parameters and cleansing criteria. The use of correct and thoroughly vetted data for benchmarking analysis is paramount to the integrity and interpretation of results. Incomplete, inaccurate, or misinterpreted data can skew analysis results and contribute to misinformed conclusions. Data analyzed without first being cleansed can lead to high variances and uncertainties in data, making it difficult or impossible to compare results of a building portfolio against its peers, perform year-to-year trending, or gain a firm understanding of which buildings are truly performing poorly (versus merely appearing to be perform poorly because of inaccurate data). These consequences will frustrate not only the community energy planner, but also building owners and managers, public officials, and other stakeholders. Further, such problems potentially undermine the credibility of the underlying energy benchmarking program or policy.

Having a documented, standardized, and replicable data cleansing process in place, supplemented by scientifically sound parameters and data thresholds, will go a long way toward minimizing these headaches. It will help yield credible and defensible analytic results and conclusions. Table 2 is a list of the most common benchmarking data quality issues typically found when data quality analytics are applied to a benchmarking dataset. The table offers information about what properties (records) should be excluded from further analysis along with explanations as to why the data should be cleansed and when it should be used for feedback purposes.

Table 1. the 21 ESPM property types (out of a total of 84 property types) that can receive a benchmarking score.

Bank Branch	Residence Hall/Dormitory
Barracks	Office
Financial Office	Courthouse
K-12 School	Wastewater Treatment Plant
Supermarket/Grocery Store	Worship Facility
Wholesale Club/Supercenter	Retail Store
Hospital (General Medical & Surgical)	Data Center
Medical office	Distribution Center
Senior Care Center	Non-Refrigerated Warehouse
Hotel	Refrigerated Warehouse
Multi-Family	

PROCESS OVERVIEW

The cleansing process for a specific dataset may vary depending on its size and format, the number of data fields it contains, and the quality of the raw data. The raw data typically include general building characteristics, namely floor area and measured or billed energy consumption data. A general process should 1) identify and fix incorrect data types, 2) identify and fix missing or erroneous data values, 3) identify and fix other data inconsistencies, and 4) ensure internal consistency.[11]

BOX 3. WHAT ARE PROPERTY TYPES?

In this guide, the term "property type" refers to the primary end-use activity or function carried on within a building or a group of buildings, in alignment with EPA ESPM definitions of property types.[12] Typical property types include office, education, food service, healthcare, and warehouse/storage.

Energy consumption per area, or energy use intensity (EUI), can vary widely between property types. For example, a restaurant with its heating and refrigeration appliances is expected to be more "energy intensive" (on a per square foot basis) than a simple warehouse. Therefore, it is both important and informative when examining a portfolio of properties to parse by individual property type.

A property's type is also a critical piece of information in explaining its consumption patterns. Municipalities have a wide variety of property types. If property types are well characterized, this variety can improve the quality of information about the overall building stock. If poorly characterized, such variety appears to reveal confusing trends and scattered consumption patterns.

Learn more by visiting Portfolio Manager's guidance on identifying property types for benchmarking: www.energystar.gov/buildings/facility-owners-and-managers/existing-buildings/use-portfolio-manager/identify-your-property-type.

[11] These steps were developed in collaboration with Shankar Earni of the Lawrence Berkeley National Laboratory. A tutorial on benchmarking data cleansing can be found on DOE's State and Local Solution Center at energy.gov/eere/slsc/downloads/benchmarking-data-cleansing-rite-passage-along-benchmarking-journey

[12] Identify Your Property Type." U.S. Environmental Protection Agency ENERGY STAR Portfolio Manager. (2013). www.energystar.gov/buildings/facility-owners-and-managers/existing-buildings/use-portfolio-manager/identify-your-property-type

Step 1: Identify and fix incorrect data types

[Note: If ESPM is being used as the benchmarking tool, then incorrect data types are not a problem. However, if a benchmarking effort involves other types of data collection and development of associated metrics, then close attention should be paid to incorrect data types.]

A common challenge when working with any data collected by external sources is to ensure that all of your data remains consistent in terms of nomenclature, data class (e.g. text vs number), and formatting. The first step, and generally the simplest, is to go through your dataset to identify and resolve any inconsistencies in data types. Good practice for data management is to adopt a standard data dictionary that defines proper nomenclature, class, and formatting for each field to maintain consistency. Data providers may not include reference tables to identify fields and their characteristics. In this case, it is up to the data manager to devise her own reference table based on interviews with the data providers, and other inputs in the dataset. Some common errors associated with data type are highlighted below with corrective actions.

- **Convert data values to standard nomenclature**

  A typographical error may result in the appearance of multiple values where only one should exist. For example, the building type Hospital (General Medical & Surgical) may be inputted as Hospital (General Medical & Surgical). The correction is to simply correct the misspelled entry to condense the values into a single value, as shown in Figure 2.

  Another issue that results from a lack of a standard data dictionary is having multiple specifications for essentially the same building type. To correct this issue, simply standardize these building type specifications in line with data types from the dictionary.

- **Convert data classes to standard data classes**

  The source of this error is that data has been classified inconsistently with the data dictionary. For example, data may be classified as string/text instead of a number. Thus data from this record

Table 2. The most common benchmarking data quality issues typically found when data quality analytics are applied to a benchmarking dataset

TYPE OF DATA QUALITY ISSUE	CLEANSE FOR ANALYSIS	USE FOR FEEDBACK PURPOSES	CRITERIA	IMPACT
Duplicate Property Entries or "dummy data" entries:	X	-		
Too Small Building Square Footage:	X	X	Property's square footage is below minimum program requirement. Use for feedback.	Building square footage may be incorrect
No Property Type	X	X	Property type was reclassified to "Not Available" as defined in ESPM's Primary Property Type – EPA Calculated field. Use for feedback.	Property does not have building gross floor area defined for complete timeframe
No EUI	X	X	Property did not report an EUI. Use for feedback.	See Understanding Why Critical Benchmarking Data Is Missing Prior to Cleansing
No ESPM Benchmark Score for Property Type That Should Have Received a Score	-	X	Use for feedback.	See Understanding Why Critical Benchmarking Data Is Missing Prior to Cleansing
Extremely High or Low ESPM Benchmark Score	X	X	Remove properties with score of 100, 99, 2, or 1. Use for feedback.	Total energy use or building square footage may be too high or too low for property
Extremely High or Low EUI	X	X	In general, Properties with site EUIs less than 2 kBtu/sf/yr or greater than 800 kBtu/sf/yr, except for Industrial/Manufacturing or Waste Water Treatment properties. Use for feedback.	Total energy use or building square footage may be too high or too low for property
Zero Electric Use	X	X	Virtually all buildings in the U.S. use some amount of electrical energy; total energy use is incorrect. Use for feedback.	Total energy use of property was not accounted for properly
Default Data Use	-	X	Use for feedback.	Benchmark score may not have been calculated correctly

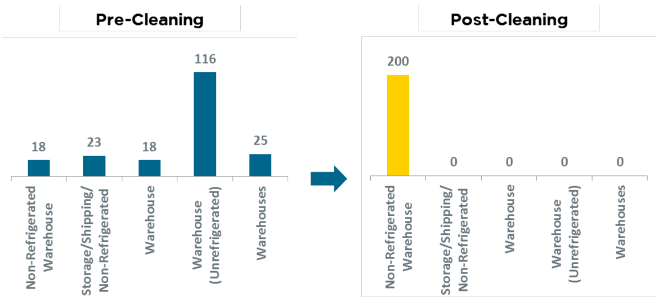


Figure 2. An example of converting data to standard nomenclature.

cannot be sorted or analyzed with other records that are classified as number. Correct data classifications to make them consistent with data classifications from the data dictionary. Tools such as Openrefine can help identify and correct these issues.

- **Standardize data formats**

Data may be incorrectly formatted. This is a common issue with zip codes. Variations in format may include _1234, 01234, and 01234-5678. Zip codes should be classified as text or string with the standardized format ABCDE.

*Step 2: Identify and fix **missing or erroneous data values***

Nearly all datasets will contain data with errors. These can be due to data entry or inputting default values instead of actual values, or simply incomplete data. Some of the missing and erroneous values are easy to identify through a cursory analysis in Excel: sorting the values and looking for zero, empty, or peculiar values (such as N/A), or too many records with the same value.

Flag all records that may be imputed (inferred) with additional research. For example, some fields (e.g. Gross Square Footage (GSF) or Zip Code) may be missing for the current year, but be available from a previous year's record.

Refer to Table 2 when performing a data quality review to identify and potentially fix missing or erroneous data values. Cleanse (remove) records with missing key pieces of data that cannot be imputed (e.g. GSF, Energy Use Intensity (EUI), Property Type).

BOX 4. WHAT IS ENERGY USE INTENSITY (EUI)?

When you benchmark your building in Portfolio Manager, one of the key metrics you'll see is energy use intensity, or EUI. Essentially, the EUI expresses a building's energy use as a function of its size or other characteristics.

For most property types in Portfolio Manager, the EUI is expressed as energy per square foot per year. It's calculated by dividing the total energy consumed by the building in one year (measured in kBtu or GJ) by the total gross floor area of the building. And don't worry — Portfolio Manager automatically does the conversion to kBtu or GJ for you, so you can just enter your energy use information as you get it on your utility bills.

Both site and source EUI are available in Portfolio Manager, though EPA relies on source EUI as the basis for the ENERGY STAR score. Learn the difference between source and site energy.

Note: This definition was taken from the EPA Energy Star Portfolio Manager website, http://www.energystar.gov/buildings/facility-owners-and-managers/existing-buildings/use-portfolio-manager/understand-metrics/what-energy

*Step 3: Identify and fix **other data inconsistencies***

This step involves establishing the rules for permissible values based on the data type and realistic values for each of the fields that are collected. For example, the maximum and minimum gross square footage (GSF) numbers, and Building Vintage (year built) can be obtained from other sources like the tax assessors dataset. The reasonable ranges for EUI for different building types can be obtained from established databases like CBECS.

Once the limits are established, the next step is to perform a distribution analysis on key variables like GSF and EUI to look for data points that fall outside the established limits. These outliers are to be investigated for their plausible cause and possibly exclusion from further analysis.

Figure 3 shows the distribution of a set of buildings reviewed for inclusion in DOE's Building Performance Database. This graph shows the histogram of GSF shown (in 1000s SqFt) on X-axis in buckets of 500,000 SqFt. The Y-axis is the frequency of buildings that fall in each bucket on a log scale. As you can see the GSF for buildings in the dataset vary quite a bit.

Four entries that have a GSF greater than 7 million SqFt. Buildings with a GSF of more than 7 million square feet are relatively uncommon; hence, further investigation needs to be conducted as to why these buildings have reported a GSF of this magnitude.

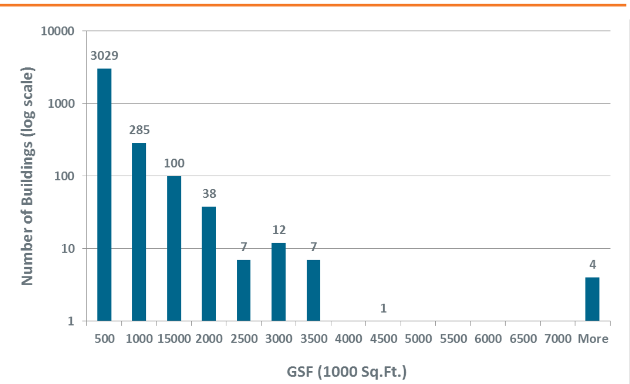


Figure 3. An example histogram of building gross square footage.

Figure 4 was obtained from Seattle's benchmarking analysis report[13]. The plot shows each building's site EUI represented by a dot and the full range of EUI performance for each building type. Although most building EUIs fall within a reasonable range, as indicated by the distribution of 10th (blue) to 90th (gold) percentiles, extremely high and low EUIs also occur in the dataset, as shown by the dots that fall outside the blue and gold bars. While some of these more extreme EUIs may represent a legitimate intense energy use, such as an office with a very large data center, other high EUIs may indicate extremely poor building performance or a possible data error, which would be worthy of investigation by the building owner or manager. In general, this type of analysis should be performed with the help of someone with considerable knowledge of statistics prior to cleansing suspicious records.

Step 4: Ensure internal consistency

The next step in the process is to check and make sure the data is consistent internally especially if historical data exists. For example, take a dataset with multiple years of available benchmarking data for a group of properties. In this dataset identify records with big variations in the reported EUI. Variations of higher than 50% are considered suspicious and should be investigated to understand their root cause. Based on the Seattle report referenced above, improvement or degradation of EUI of 50% or greater is considered an outlier. The high variation in Site EUI can be caused by high variation in either GSF or Site Energy Use.

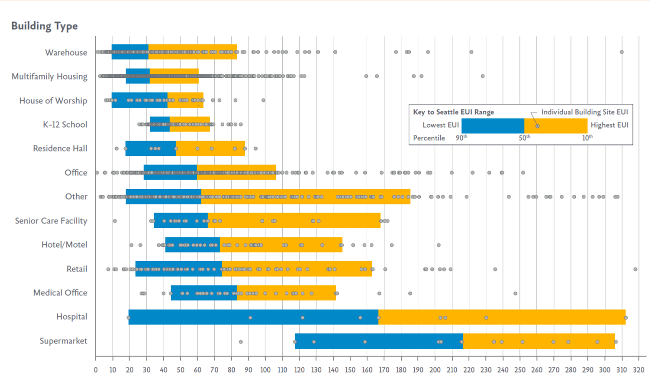


Figure 4. 2012 EUI Performance Range and Distributions by Building Type

RECOMMENDATIONS

Based on experience gained from performing multiple data cleansing analyses, several important tips, tricks, and lessons learned have been developed that can help analysts.[14] These include the following:

- Data are dirty until proven clean. The Building Performance Database throws out roughly 20% of buildings provided by data contributors due to identified data quality issues such as missing key fields or estimated energy use values. There is a trade-off between data quality and quantity: while including more buildings in your analysis will reduce uncertainty, bad data will lessen confidence in your results.
- Avoid making assumptions about building asset data if there is any doubt about its accuracy.
- Include a well-defined glossary of terms, such as BEDES or ESPM, to avoid assumptions and ambiguity.
- If individual data fields seem suspect, it may not be necessary to throw out the whole building record; some data fields may still be valid for analysis.
- Become familiar with standard building characteristics and usage trends by major building types to help spot errors.
- Learn data manipulation techniques to spot outliers: sorting values, identifying missing data, plotting the distribution by GSF or EUI.
- Perform statistical analysis to further characterize portfolio and identify additional data issues.

TIPS

- Very similar energy uses for different building types in different locations can be an indication of default values, instead of actual values, used in either energy use or gross floor area (see Identifying Estimated Values below).
- Data with inconsistent time intervals require extra effort to aggregate, as do buildings with sub-meters or meters out of service.
- For multifamily buildings, watch for inconsistent whole-building energy use (e.g., where the heating fuel has reported consumption for the whole building, but electricity only for common areas).
- Inconsistencies in building identification codes, floor area, etc., can make mapping the individual characteristic of a single building into a common

[13] Seattle Office of Sustainability & Environment. 2011/2012 Seattle Building Energy Benchmarking Analysis Report. January 2014.

[14] These lessons learned, as well as general input into this section, were provided by staff of the Lawrence Berkeley National Laboratory, based on their experience in analyzing energy benchmarking datasets in support of DOE's Buildings Performance Database. For more information, visit: www1.eere.energy.gov/buildings/commercial/bpd.html.

data format more time-consuming.

- Bounding ranges for building characteristics—such as single buildings with floor areas greater than 100,000 sqft or non-zero electricity data—need to be reasonable yet non-exclusive.
- Additional data fields might be required to understand peculiar building data. Multi-building facilities can have large gross floor areas spread out over multiple floors. Apartment properties in the same portfolio can have similar floor areas and energy uses. Other times, monthly electricity data can have zero or negative values because of on-site generation. These are still valid data. In some cases, data cleansing may require someone with knowledge of building science to make judgments on what values are reasonable.
- Data about facilities with on-site power generation tends to be more difficult to cleanse. For instance, a facility with a central generating plant that serves buildings outside of the reported square footage should be excluded from further analysis.

By using both ESPM as a benchmarking tool and the data cleaning guide listed in Table 2, raw benchmarking datasets can be cleansed to an acceptable level.

The following analysis examines the potential data quality impacts of properties with ESPM scores using default data

BOX 5. EXAMPLE: IDENTIFYING ESTIMATED VALUES

As part of the process of integrating datasets into the Buildings Performance Database (BPD), Lawrence Berkeley National Laboratory (LBNL) performs rigorous data cleansing to identify and remove suspect or erroneous values from the raw data provided by sources. While the process described here and threshold recommendations used for BPD provide general guidelines for identifying common issues with data quality, each dataset offers unique data cleansing challenges. We describe here an example in which a number of estimated values were identified in an initially cleansed dataset. This instance of bad data resulted from a difference in data quality standards between the BPD and one of its data contributors, which became apparent only through further analysis of the dataset.

The dataset consisted of over 11,000 residential buildings in the Northeastern United States. About 8% of these buildings were removed during the initial cleansing due to failure to meet the cleansing thresholds. A histogram of the dataset using standard cleansing rules, shown in Figure 5(a), revealed that nearly 8% of the buildings reported energy use intensities (EUI) equal to exactly 32 kBtu/ft2/year. Although data are not expected to follow smooth distributions, this highly unlikely distribution prompted further investigation. The data source contributor was able to confirm that energy use readings were estimated using a default EUI for buildings that were unable to provide one complete year of data.

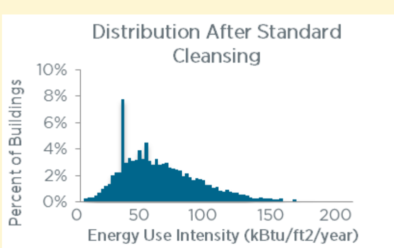
A new cleansing rule was devised to address this specific issue. To guarantee the quality of data imported into the BPD database from a dataset identified to contain erroneous data, only buildings with unique EUIs were accepted. While this rule may be considered excessively strict for most datasets, the BPD prioritizes data quality over number of buildings and therefore implemented a rule that would guarantee high data quality. With this new cleansing rule, 16% of the buildings in original dataset were eliminated.

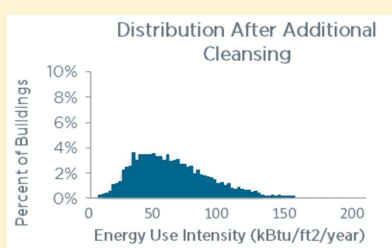
Analysis of the fully cleansed data reveals some minor differences in the net characteristics of the dataset (given in Table 3). Visual comparison of the two figures below (Figure 5), confirms the absence of the spike in the histogram of EUI, which initially caused suspicion about the data quality. The histogram follows an approximately log-normal distribution, which is consistent with expectations based on previous studies.

Table 3. Dataset characteristics, before and after standard and supplemental cleansing.

DATA CHARACTERISTIC	BEFORE CLEANSING	AFTER CLEANSING
Number of Buildings	11,485	8,758
Aggregate Gross Floor Area	22 million ft2	17 million ft2
Total Annual Energy Consumption	1.4 billion kBtu	0.97 billion kBtu
Annual Energy Use Intensity (EUI)	65.0 kBtu/ft2	57.1 kBtu/ft2
Median Annual EUI	52.5 kBtu/ft2	56.7 kBtu/ft2

Figure 5. Histograms showing the underlying distribution of a residential dataset in the Northeastern United States by energy use intensity (a) after standard cleansing [left] and (b) after further cleansing designed to eliminate an identified data quality issue [right].





DISTRIBUTION OF PROPERTIES USING DEFAULT DATA

To generate an ESPM score, ESPM allows for the use of default values as input variables if a building owner/operator does not know the actual values (input variables such as number of computers, number of workers, or hours of operation can be defaulted to predetermined values). This use of default data allows ESPM to calculate a score, but the default data may not correctly reflect energy usage patterns at a property, and the resulting score may not accurately reflect the actual energy efficiency score of the property.

To minimize uncertainty about the energy efficiency of the building portfolio, the use of default data should be investigated to determine its potential impact on ESPM scores. This can be done by calculating the percentage of properties that received a score that used default data. ESPM's "Default Data Flags" fields track whether or not default data was used for a property.

Required Data and Dependencies

- **Property type: *ESPM data field – "Primary Property Type - EPA Calculated"***
- **ESPM data field – *"Default Data Flag"***
- **Subset of properties having an ESPM score: *ESPM data field – "ESPM Score"***

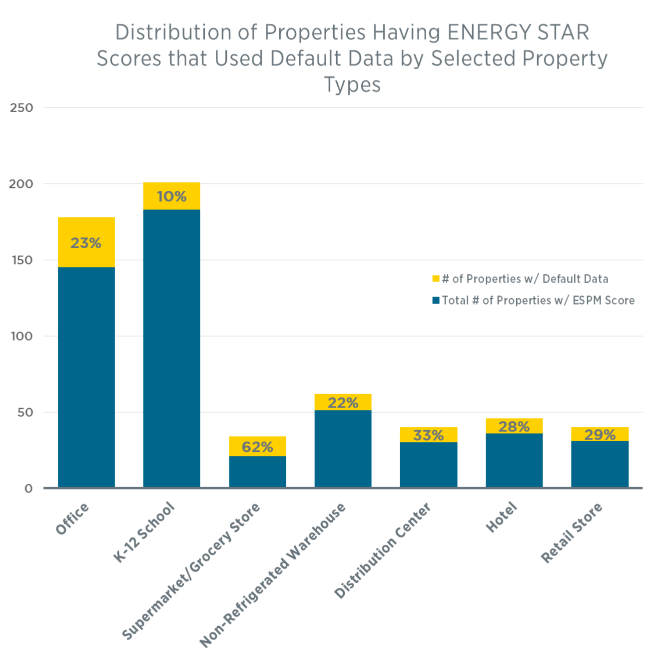


Figure 6. Distribution of properties having ESPM scores that used default data by Property Type for Philadelphia, 2013

Example

In Figure 6, properties from Philadelphia's 2013 cleansed benchmarking dataset having ESPM scores (an analysis performed in Chapter 4 below) are compared with those properties that reported using default data grouped by selected property types. In order to perform this analysis, properties with default values are first parsed out of the dataset Supermarkets/grocery stores have the highest incidence of using default data (62%), which brings into question the validity of many of the ESPM scores attributed to this property type. Distribution centers also have a relatively high incidence of default data use (33%), which may indicate data quality issues for this property type as well. About 23% of office properties reported using default data, which is high enough to raise some concern, since offices represent the majority of benchmarked gross floor area.

Although the impact that default data can have on the actual calculation of an ESPM score can vary from one property type to another, it is interesting to note both distribution centers and offices also had some of the highest (73 and 71, respectively) median scores for the property types shown in Figure 26. Further analysis outside the scope of this guide would have to be performed to determine if using default data in these cases significantly skewed the scoring of these properties.

CHAPTER SUMMARY

This guide opens with a strong focus on assessing the quality of your dataset. The methodology demonstrated in this chapter will allow the program manager and data analyst to not only clean up data so that future analyses are more accurate, but to ascertain any systematic causes of the poor quality data. Understanding the types of data quality issues within your dataset can also help you develop educational and training materials that will improve the quality of future data and lead to decreased time required to clean data for analysis.

Chapter 4: Data Parsing

FLOW CHART FOR PARSING AND ANALYSIS OF BENCHMARKING DATA:

Data parsing forms the foundation for energy benchmarking analysis. To evaluate a portfolio of buildings, you must begin by characterizing the building stock by parsing the building data. Characterizing a portfolio's energy use profile by property type, age, size, and other parameters helps stakeholders understand the physical nature of the building stock and identify sub-groups with high energy efficiency potential (these sub-groups may be buildings of a certain age, size, or use, for example), Additionally, parsed data can be cross-checked against other records (such as real property and/or tax data or other city records) for verification.

The flow chart in Figure 7 illustrates the steps involved in the analysis of a dataset from ESPM. Using this approach, the dataset is first cleansed, and suspect data are set aside. The next step parses the properties in the dataset into those that are eligible to receive an ESPM score, and those that are not eligible based on their property type, which of these actually received a score, which used default data, and, finally, analyzing select property types in detail.

Performing this analysis allows the analyst to uncover relationships among subsets of data (such as the relationship between fuel type and property type). By uncovering these relationships, the benchmarking analyst will gain insight into how their building stock uses energy. Such parsing allows more complex analyses based on score and allows the analyst to determine specific energy efficiency strategies that maximize the ESPM score.

Disclaimer:
Philadelphia's 2013 Benchmarking Dataset for commercial buildings was used with this benchmarking data analytics guide to show various types of data parsing and benchmarking analyses. This dataset is comprised of both public and private sector buildings with a minimum gross square footage of 50,000 sqft. The use of this data is for educational purposes only and any analytical results shown in this guide are not intended to represent any type of "official" results related to Philadelphia's benchmarking program.

SECTION 1: CHARACTERISTICS OF TOTAL BUILDING STOCK

This section offers different analyses at the "all-properties" level and is intended to give information about the total building stock assessed by a benchmarking program. The different types of analyses shown in this section are the following:

- Distribution of all properties by property type
- Distribution of total gross floor area by property type
- Distribution of calculated total energy consumption by property type
- Impact of property age
  - Distribution of properties by decade built
  - Distribution of calculated gross floor area by decade built for all properties
  - Distribution of median site and source EUI by decade built
- Distribution of property size for all properties
- Distribution of properties by fuel share
- Comparison of median site and source EUIs for property types with largest gross floor area
- Parsing all properties by ESPM benchmarking score eligibility

DISTRIBUTION OF PROPERTIES BY PROPERTY TYPE

One of the most basic analyses involves understanding how many total properties are within a property portfolio, particularly within a given property type. While the number of properties within a given property type does not typically correlate to energy consumption (stronger correlations exist

between energy consumption and floor area), it is still helpful for the energy planner to understand the distribution of properties among property types.

This data can serve as a basic initial check for compliance with energy reporting requirements. For example, if only two hotels have reported data, yet the planner knows of more than ten hotel buildings downtown that pass the size threshold, then there may be a gap in compliance.

Required Data and Dependencies

- **Property type: *ESPM data field – "Primary Property Type - EPA Calculated"***

Example

The example shown in Figure 8 indicates the reported benchmarked properties for Philadelphia's 2013 cleansed benchmarking dataset are dominated by K-12 school, office, college/university, non-refrigerated warehouse and distribution center property types. The figure shows two different graphical representations to convey this information, a bar chart and a pie chart. If an analysis showed these results, an energy planner might conclude that creating an energy efficiency program that addressed K-12 school properties could involve reaching out to the largest number of building operators. Based on

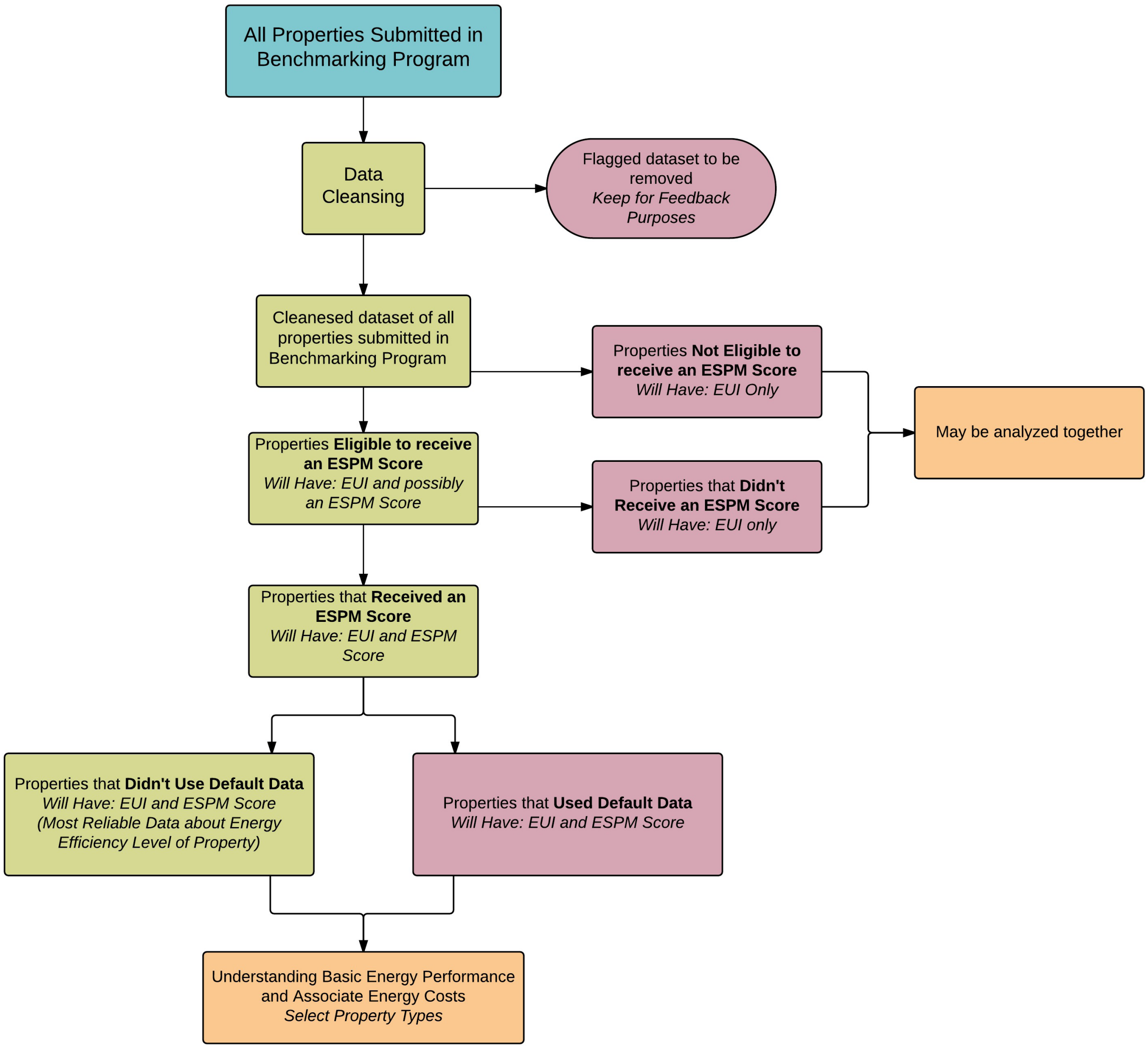


Figure 7. Flow Chart for Parsing Benchmarking Data for deeper/further Analysis

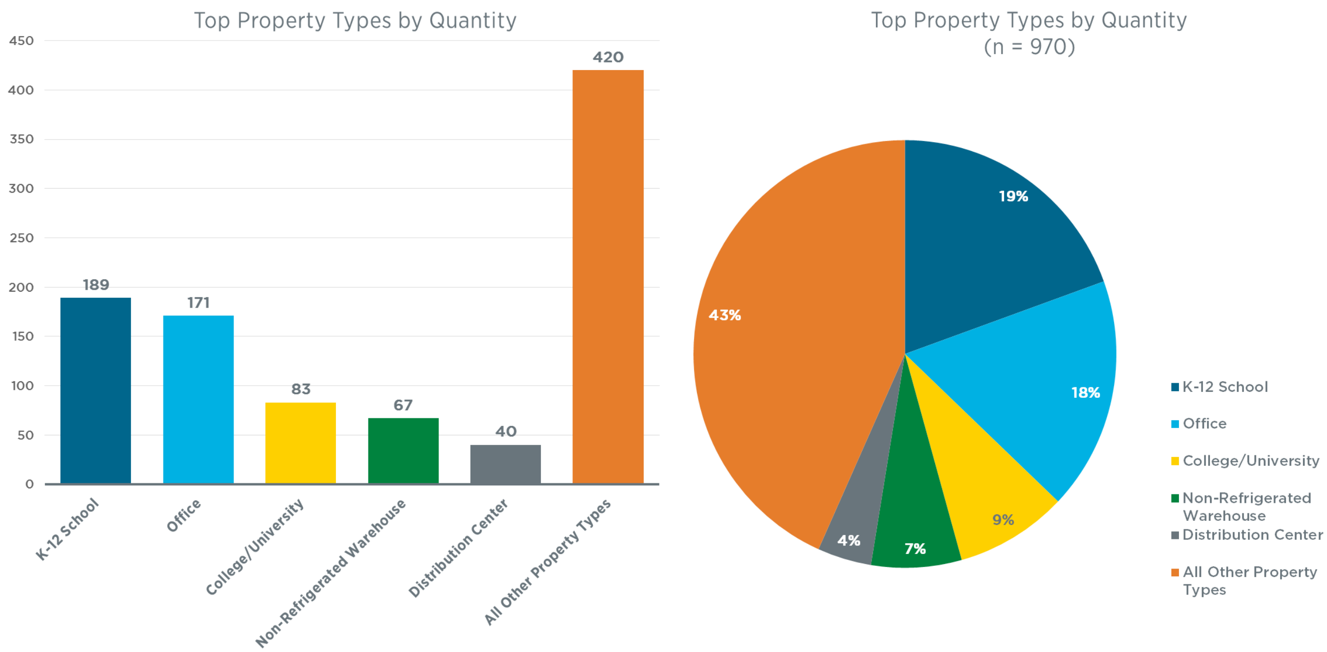


Figure 8. Distribution of properties by property type for Philadelphia, 2013

the distribution of property types, schools and office buildings are two property types that may represent the highest potential opportunity for targeted EE programs, however further analysis should be used to confirm this preliminary hypothesis. When planning a strategy for EE programs, this analyses points to the types of decision-makers, building owners and managers that would need to be engaged as the primary audience of a targeted EE program.

DISTRIBUTION OF TOTAL GROSS FLOOR AREA BY PROPERTY TYPE

This analysis examines the distribution of property types within a dataset by showing how much total gross floor area is represented by each designated property type.

Property energy consumption scales with gross floor area to varying degrees, depending on property type and other factors, such as operating hours and surface-to-volume ratio. Knowing how much square footage is represented by each property type can help jurisdictions understand which energy efficiency programs will likely have the greatest impact on their building stock or portfolios. (Property square footage data can also help managers of utility incentive programs better target energy efficiency incentives.)

For example, New York City's 2012 benchmarking report indicated that multifamily buildings comprised nearly two-thirds of the total benchmarked building area.[15] This indicates that multifamily buildings and their occupants and associated energy use play a very important role in the overall city energy profile.

Showing which sectors comprise the largest floor-area portions of the building stock or specific portfolios enables administrators to strategically target them in order to meet goals of implementing efficiency measures, as these sectors likely have significant energy impact. For example, if a portfolio shows ten dominant sectors by floor area, one with a 50% representation, a program administrator might prioritize program efforts by targeting this largest sector.

BOX 6. CAMPUS-TYPE ENERGY USE AND SUB-METERING

Since properties with more than one building often share energy systems and meters, understanding how much floor space is attributable to campus-type properties with more than one building is important. Shared energy systems and meters can make it difficult to analyze energy use at the building level. Not accounting for this complexity, however, can skew benchmarking results, Suffice it to say, collecting and analyzing this kind of benchmarking data requires special care. (ESPM offers a detailed methodology for properly collecting and reporting energy use for multi-building properties.)

[15] New York City. New York City Local Law 84 Benchmarking Report (September 2013). http://www.nyc.gov/html/planyc/downloads/pdf/publications/ll84_year_two_report.pdf

Required Data and Dependencies

- **Property floor area (typically reported in gross square feet (GSF)):** *ESPM data field – "Property Floor Area (Building(s)) (ft2)"*
- **Property type:** *ESPM data field – "Primary Property Type - EPA Calculated"*

Example

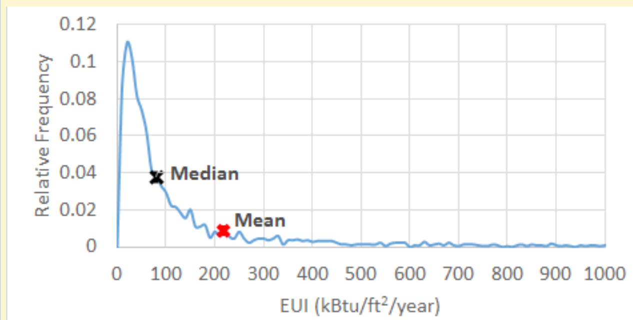
The example shown in Figure 9 indicates the breakdown of total gross building floor area for Philadelphia's 2013 benchmarking cleansed dataset is dominated by K-12 schools with office buildings, colleges/universities, hospitals and non-refrigerated warehouses also representing significant, but far smaller, portions. The figure shows two different graphical representations, a bar chart and a pie chart. To add context to the pie chart, a conversion of 1% to approximately 2,200,000 GSF is provided. When the analysis shifts from quantity of buildings to gross square footage of, the Distribution Center building type drops out from the graphs, replaced by the Hospital type as a potentially significant subsector. Both visualizations communicate that the office building type presents potentially great opportunity for EE savings because of the large proportion of floor area it represents.

DISTRIBUTION OF CALCULATED TOTAL ENERGY CONSUMPTION BY PROPERTY TYPE

Another basic analysis involves understanding how total energy use (normalized by kBtu for all fuel types combined) is distributed by property type. This kind of high-level analysis shows which property types consume the most overall energy out of all benchmarked properties and can be an indicator of where substantial energy savings are potentially available. For energy planners and utilities, knowing

BOX 7. MEDIAN VERSUS MEAN?

The mean and median describe the "central tendencies" of a dataset. Both statistics provide information about a set of values by identifying a type of "center." Which "center" is more appropriate depends on several factors. Advantages of using the mean (i.e., "average") include ease of calculation and the fact that its calculation includes all values in the dataset. Changing any one value in the dataset will influence the mean, but may not influence the median. However, the value of the mean is more susceptible to the influence of the values of outliers. If a dataset may have outliers that are significantly larger and smaller than the majority of the values, then characterizing it with the mean could be misleading. Furthermore, if a dataset is asymmetrical, or skewed, then the median is considered a more representative center because it is not affected by the skew like the mean. Assuming skewed distributions and the possibility of significant outliers is prudent when examining a dataset. Therefore, these analyses more often employ the median as the central tendency of choice. There may be other analyses, however, for which the mean would be more appropriate.



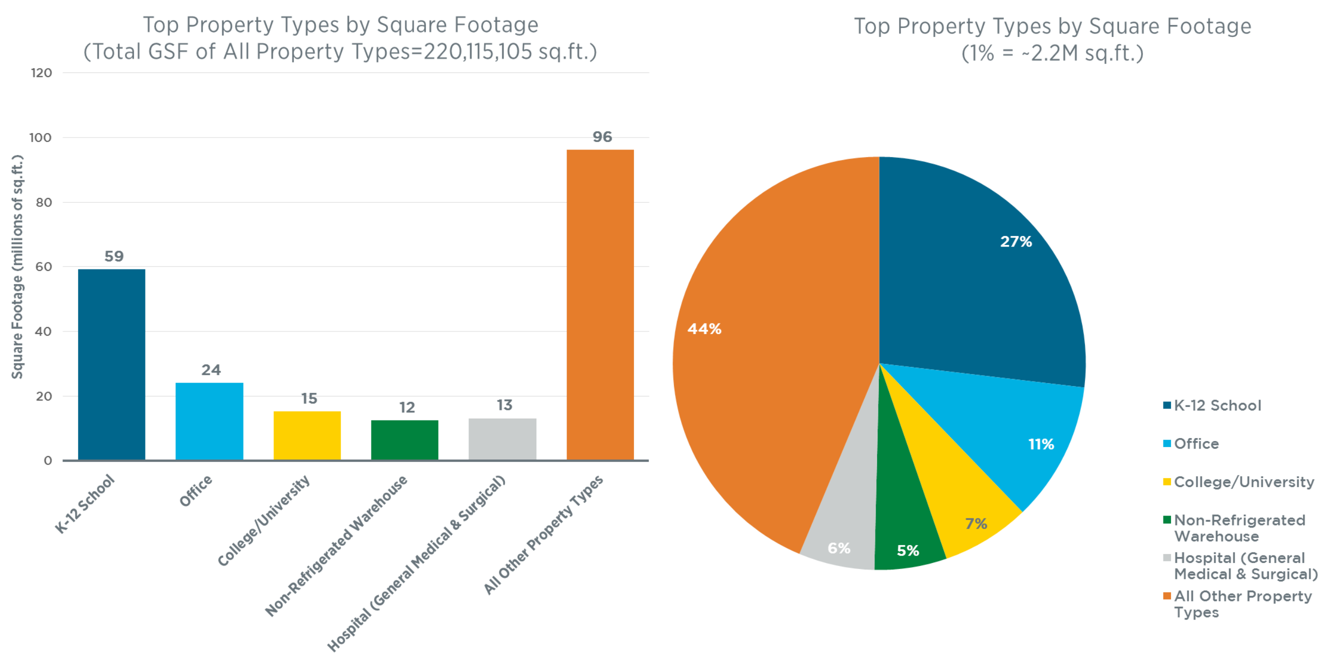


Figure 9. Distribution of calculated gross floor area by property type for all properties for Philadelphia, 2013

which property types consume the most energy can be useful in directing energy efficiency efforts and programs, such as potential energy reduction goals or utility incentive programs. This data combined with data of the commensurate property types that have the largest number of properties can help target energy efficiency programs for the greatest impact.

Required Data and Dependencies

- **Calculated total property energy use based on summing individual fuel shares across all fuel types (normalized to kBtu): typical** *ESPM data fields – "Electricity Use - Grid Purchase and Generated from Onsite Renewable Systems (kBtu)", "Natural Gas Use (kBtu)", "Fuel Oil #2 Use (kBtu)" and "District Steam Use (kBtu)"*

Example

The example shown in Figure 10 shows the total site energy (kBtu) of all reported benchmarked properties for Philadelphia's 2013 cleansed benchmarking dataset based on the summation of individual fuel types. Within this dataset, property types relating to offices, hospitals, colleges/universities, manufacturing/industrial plants and K-12 schools represent the largest energy consumers.

IMPACT OF DECADE WHEN PROPERTIES WERE CONSTRUCTED:

A property's energy consumption depends on a variety of factors, some significant and some more peripheral. One factor that may be potentially significant is a property's age. Looking at property age data in a portfolio of properties may illuminate important characteristics about the dataset and should be part of general analytical due diligence when parsing benchmarking data. This type of data analysis could provide important information for a municipality to understand the relationship between age and energy use across their building stock or portfolio. For example, if much of a portfolio's building stock was built during a certain decade, then this stock likely includes similar building materials and mechanical equipment. An energy efficiency program could then be designed to target specific retrofits to benefit these properties.

Three different, but related, analyses are shown below that look at the potential relationship between property age, combined gross floor area and energy use. The first deals with the distribution of properties by decade built, while the second shows combined property gross square footage by decade built. The third analysis looks at the energy use, represented by the median EUI, of those properties parsed by decade built.

BOX 8. A WORD ON SITE AND SOURCE ENERGY USE

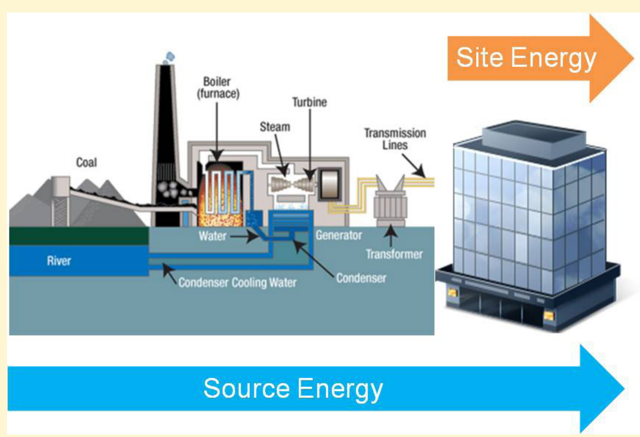
Site energy use intensity refers to the amount of energy consumed per square foot of gross building area, typically measured by meters at the site (i.e., through utility bills). Since site energy only involves energy consumed at the building site, it truly reflects building performance.

Source energy use intensity considers the fuels consumed in the generation, transmission, and distribution of electricity, as well as the losses from distributing and dispensing natural gas; this is done by site-to-source fuel factors to site energy use intensity. Thus, source energy is a more rigorous calculation, and better reflects the overall environmental impact of energy consumption.

ESPM rates building performance using source, not site, energy. Converting site EUIs to source EUIs is a general function of Portfolio Manager, using national source-site ratios. This guide also recommends performing analyses using source energy, where applicable.

For more information on site and source energy, visit: www.energystar.gov/buildings/facility-owners-and-managers/existing-buildings/use-portfolio-manager/understand-metrics/difference.

Fuel Type	Source-to-Site Ratio
Electricity	3.14
Natural Gas	1.05
Fuel Oil	1.01
Steam	1.20



Note: although increasing building age is often thought to correlate with decreasing efficiency, construction techniques and building materials have improved over time, so building envelopes (i.e., exterior surfaces like walls, windows, and roof) should theoretically leak less energy. The efficiencies of equipment used for lighting, HVAC, and miscellaneous loads have also increased. There are factors that complicate this general trend, however. For example, if building age outlasts equipment lifetime, then equipment could be replaced with newer, more efficient equipment. A 40 year old building, therefore, might have more efficient HVAC equipment in some instances than a 10 year old building. Additionally, newer buildings tend to be outfitted with more energy-intensive miscellaneous "bells and whistles" equipment (i.e., plug loads). Complicating factors notwithstanding, understanding the range of ages in a building stock provides valuable information about it.

> DISTRIBUTION OF PROPERTIES BY DECADE BUILT

Required Data and Dependencies

- **Property age. Municipal records, such as construction permits, should indicate when buildings went into service. Dates of initial construction, instead of renovation, are typically most useful for this type of analysis.** *ESPM data field – "Year Built"- aggregated into decades.*

Example

Figure 11 shows the distribution of the decades (bins) when properties were built for Philadelphia's 2013 cleansed benchmarking dataset. The size of the age bin can be varied if more "resolution" is desired as a way of gaining better understanding about the relationship between number of properties and age, but in most cases a timeframe of a decade is sufficient.

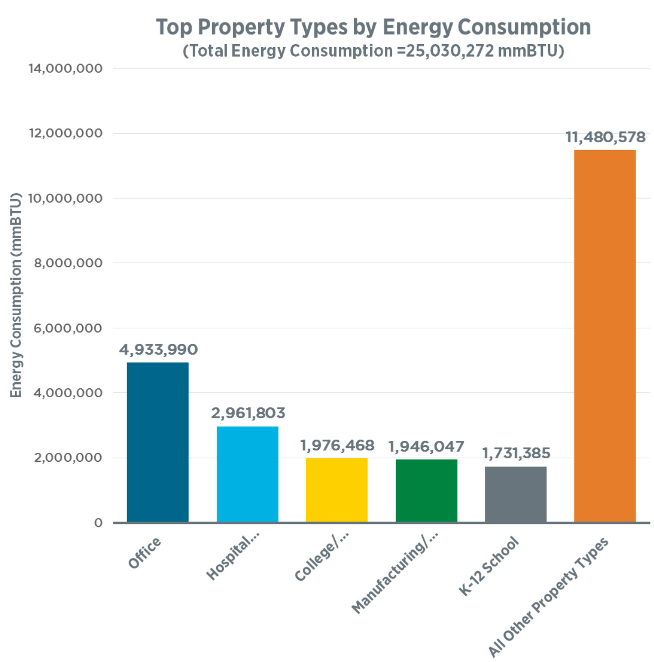
As would be expected, the decades in which benchmarked properties were built in Philadelphia run the gamut from the 1700's till present, with about 44% being built prior to the 1960's.

BOX 9. WHAT TYPE OF EUI DO I USE FOR ANALAYSIS - SITE, SOURCE, WEATHER-NORMALIZED?

When using ESPM for benchmarking, four Energy Use Intensities are typically generated for each property: site, source, weather normalized site, and weather normalized source. Which type of EUI you use for analysis depends on the intended outcome of your analysis. If the intent of the analysis is to reduce energy consumption through energy efficiency retrofits in commercial buildings, then using site EUI makes sense, as it more closely aligns with how building owners/operators manage energy use. For utility incentive programs, site EUIs are used in determining energy savings and associated incentives.

If the intent is to understand the best way to reduce greenhouse gas emissions from a municipality's building stock, for example, then source EUI would be more appropriate, since source EUI is better aligned with total energy consumed at the point of generation and subsequent transmission. If the intent of the analysis is to compare building performance from year to year without the influence of weather conditions, then either weather-normalized site or source EUI is appropriate. In many cases, performing the same analysis using all four types of EUI can yield important insights into overall energy use and energy savings.

Note: in some cases ESPM will generate both site and source EUIs for a property but will not generate commensurate weather normalized EUIs if, for example, utility data spans more than a one-month period.



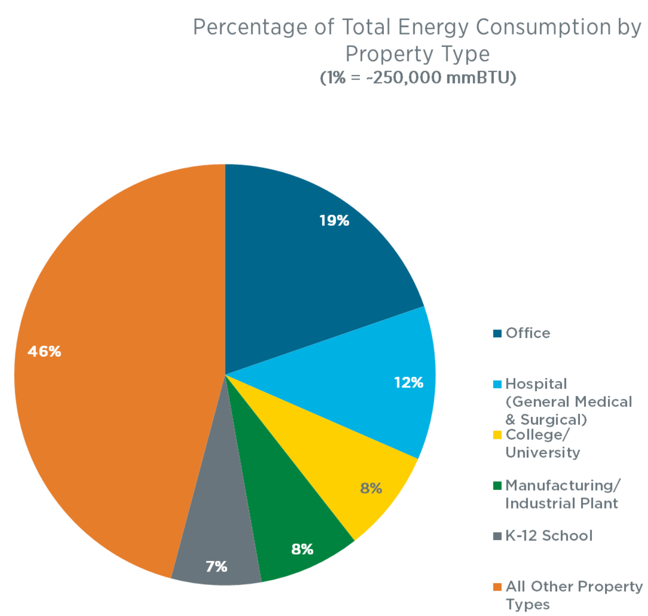


Figure 10. Total site energy of all reported benchmarking properties for Philadelphia

> DISTRIBUTION OF CALCULATED GROSS FLOOR AREA BY DECADE CONSTRUCTED

This analysis examines the distribution of all property gross floor square footage by decade built and will indicate which decades the largest amounts of property square footage were built.

Required Data and Dependencies

- **Property floor area (typically reported in gross square feet (GSF)):** *ESPM data field – "Property Floor Area (Building(s)) (ft²)"*
- **Building age:** *ESPM data field – "Year Built"- aggregated into decades.*

Example

In the example shown in Figure 12, property age is shown by decade along with the total property gross floor area for all properties built in that decade for Philadelphia's 2013 cleansed benchmarking dataset. The majority of floor area is represented by buildings constructed between 1950 and 2009. A smaller peak in construction also occurred in the 1920s. When compared to Figure 11, about 69% of benchamrked gross floor area is contained in buildings built from 1950 to present, which reperesent about 66% of the total number of buildings benchmarked.

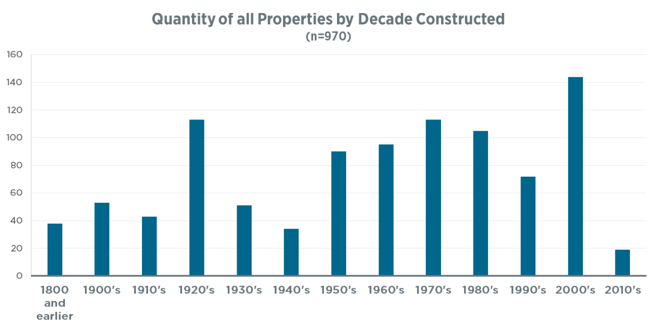


Figure 11. Quantity of Properties by Decade Constructed

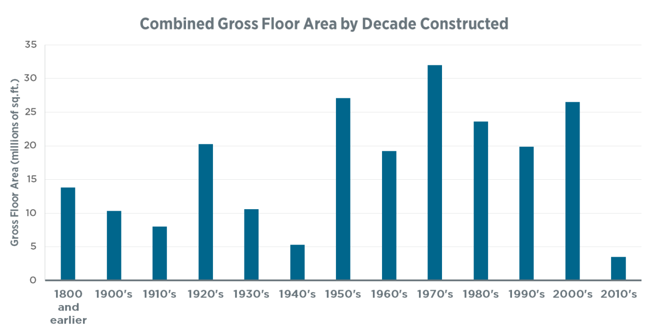


Figure 12. Gross Floor Area by Decade Constructed

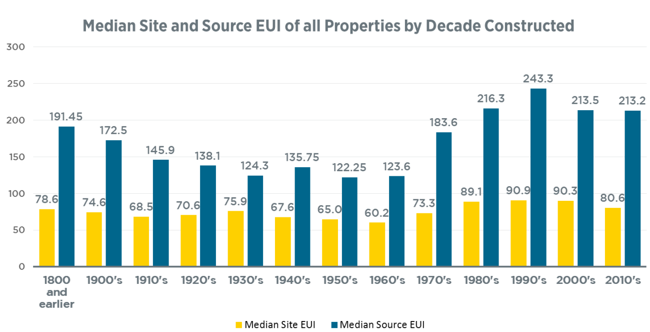


Figure 13. Median Site and Source EUI by Decade Constructed

> DISTRIBUTION OF MEDIAN SITE AND SOURCE EUI BY DECADE CONSTRUCTED

This analysis examines the distribution of the annual median site and source EUI for all properties grouped by decade built for Philadelphia's 2013 cleansed benchmarking dataset and will indicate which decades had the highest and lowest median EUI.

Required Data and Dependencies

- **Annual Site EUI (kBtu/sq/yr):** *ESPM data field – "Site EUI (kBtu/ft²)", "Source EUI (kBtu/ft²)"*
- **Building age:** *ESPM data field – "Year Built"- aggregated into decades.*

Example

Shown in Figure 13 is the relationship between annual median site and source EUI by decade built for all properties in Philadelphia's 2013 cleansed benchmarking dataset.

For the properties that were built prior to the 1960s (approximately 44% of all properties), annual median site EUIs ranged from a low of 65.0 kBtu/sf/yr to a high 78.6 kBtu/sf/yr, while for the remaining properties annual median site EUIs ranged from a low of 60.2 kBtu/sf/yr to a high of 90.9 kBtu/sf/yr . In particular, properties built in the 1970s to present show the highest set of annual median site EUIs, which is an indication they typically use more energy than properties built prior to the 1960s. When the properties' energy use is assessed in terms of median *source* EUI and the related generation of greenhouse gases, the differential between properties built prior to the 1960s and those built after the 1960s grows significantly. This may be an indication that the additional energy use attributed to properties built after the 1960s could be the result of increased use of electric internal loads.

This analysis can also be done using a comparison of weather normalized site and source EUIs; however, care must be taken with this approach. For Philadelphia's 2013 cleansed benchmarking dataset used for examples in this guide, ESPM generated site and source EUIs for all 970 properties, but it only generated weather normalized site and source EUIs for a subset of these 970 properties. Hence, the calculation of metrics like median site and source EUIs and weather normalized site and source EUIs will involve datasets that are not commensurate in terms of the number of data points used.

DISTRIBUTION OF PROPERTY SIZE FOR ALL PROPERTIES

This technical analysis examines the distribution of buildings by floor area within a dataset.

This could be particularly useful when considering next steps after benchmarking, for example, for improving energy efficiency. Not only are activities such as energy audits, retro-commissioning, or whole building retrofits approached differently based on building size (including the level of resources required), but larger buildings tend to have different mechanical system types (e.g., a centralized heat or chiller plant) than smaller buildings (e.g., packaged rooftop or window units). Additionally, when looking at a portfolio as a whole, larger buildings typically have a greater impact potential on energy consumption than smaller buildings, on a per building basis. Understanding these characteristics can guide building managers and energy planners when making decisions about appropriate strategies for energy efficiency within a portfolio of buildings.

Required Data and Dependencies

- **Property floor area (typically reported in gross square feet (GSF)):** *ESPM data field – "Property Floor Area (Building(s)) (ft²)" – aggregated into gross floor area bins*

Example

For this analysis, sorting property data by floor area into appropriately sized bins will allow you to create a histogram of the number of properties in your pre-defined size ranges. Grouping by 100,000 GSF is a good place to start for properties that range from 100,000 to 500,000 GSF. The smallest bin may be set by any thresholds that impacted what data were collected (e.g., if only buildings larger than 50,000 GSF were covered by a benchmarking ordinance, then the first bin would range from 50,000 GSF to 100,000 GSF). For bins covering the largest properties, two bins – one for 500,000 GSF to 1 million GSF and any property over 1 million GSF - may be more applicable.

In the example shown in Figure 14, properties contained in Philadelphia's 2013 Benchmarking cleansed dataset were distributed by a set of nonlinear GSF bins since the overall range of property GSF was quite large. Although a majority of properties (83%) were found to be 300,000 GSF or less, Philadelphia (as would be expected for a large US city) has benchmarking data on more than 120 larger properties. This analysis can also be performed for a subset of buildings of a specific building type, such as for office buildings. This can be useful to further understand.

DISTRIBUTION OF PROPERTIES BY FUEL SHARE

Benchmarking properties requires the collection of utility billing information for the individual properties assessed. For the vast majority of commercial buildings in the U.S., utility billing information offers energy use data related to electricity and either natural gas, fuel oil or district steam (collectively known as a property fuel shares or mix). By summing each type of fuel share and parsing by property type, valuable insight can be gained about how various property types consume energy. Fuel share information can be used by managers of utility incentive programs to target - based on the type of energy they sell - specific property types for energy efficiency rebates.

Required Data and Dependencies

- **Calculated sum of individual fuel shares for each property type (normalized to kBtu): typical** *ESPM data fields – "Electricity Use - Grid Purchase and Generated from Onsite Renewable Systems (kBtu)", "Natural Gas Use (kBtu)", "Fuel Oil #2 Use (kBtu)" and "District Steam Use (kBtu)"*

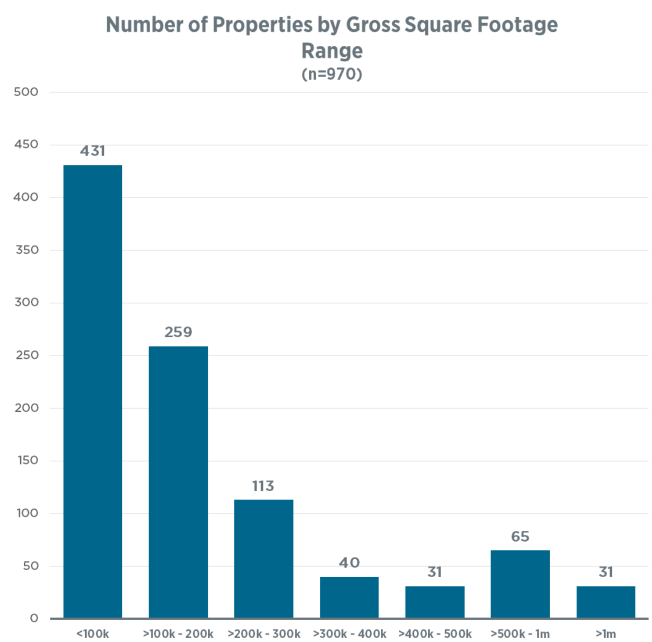


Figure 14. Distribution of Property Size for All Properties for Philadelphia, 2013

Example

Figure 15 shows the various fuel ratios of the different fuel shares for the top energy consuming property types found in Philadelphia's 2013 cleansed benchmarking dataset. The disbribution of fuel shares can vary significantly from one property type to another. Among the top energy consuming property types, offices have a very high ratio of electric energy use compared to the other property types which have, in contrast, higher ratios of natural gas and district steam use. This type of fuel share analysis can also be done using the average fuel share value (as opposed to the total value) for the top property types.

COMPARISON OF MEDIAN SITE AND SOURCE EUI FOR PROPERTY TYPES WITH LARGEST GROSS FLOOR AREA

Each type of EUI offers a "different" look at how a property uses energy (for more information on this refer to "Box 9. What type of EUI do I use for Analaysis - Site, Source, Weather-Normalized?" on page 20). For this analysis, a comparison is made of the site and source EUIs for the four property types having the largest amount of total gross floor area.

Required Data and Dependencies

- **Annual site and source EUI (kBtu/sq/yr): ESPM data field – "Site EUI (kBtu/ft²)", "Source EUI (kBtu/ft²)"**
- **Property type: ESPM data field – "Primary Property Type - EPA Calculated"**

Example

Figure 16 shows both the median site and source EUI for the four property types having the largest amount of gross floor area found in Philadelphia's 2013 cleansed benchmarking dataset. The difference between the magnitude of a property type's median and source EUI can be indication of the impact of electric energy use since ESPM utilizes a significantly larger site-to-source multiplier for electric energy use. In this example, the difference between median site and source EUIs for offices is much greater than the EUI difference found with K-12 schools, which makes sense since the electric fuel shares shown in the Fuel Share analysis indicates a much larger electric fuel share for offices as opposed to K-12 schools.

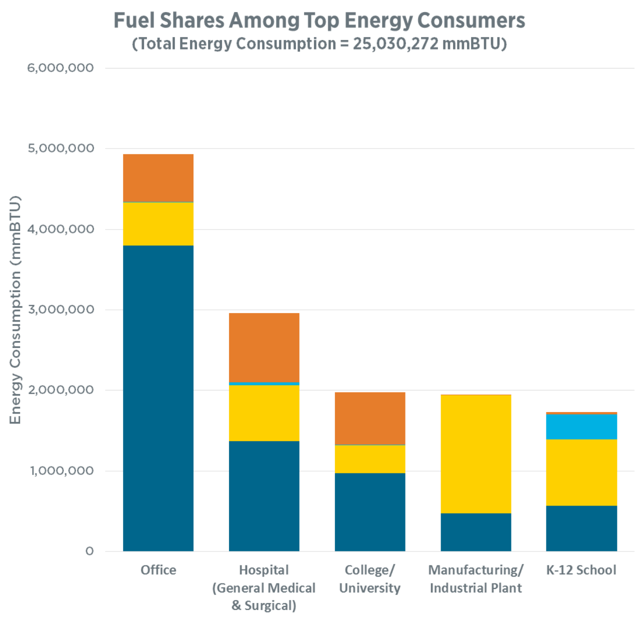


Figure 15. Distribution of Properties by Fuel Share (Mix) for Philadelphia, 2013

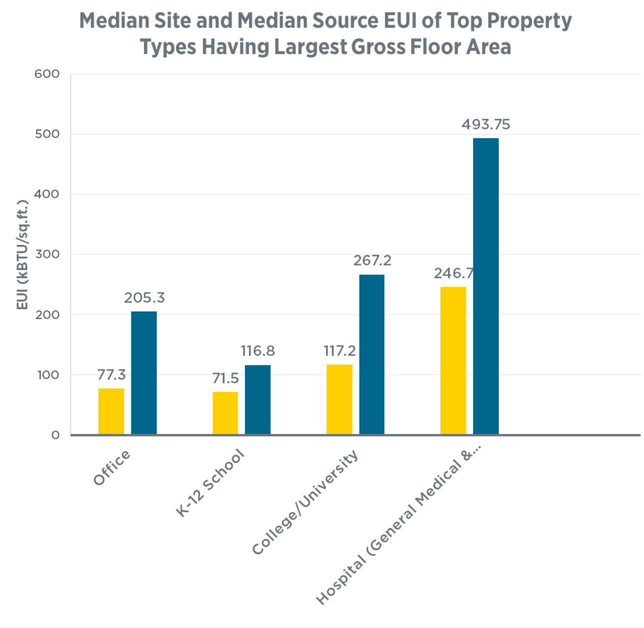


Figure 16. Comparison of Median Site and Source EUI for Philadelphia, 2013

SECTION 2: CHARACTERISTIC OF BUILDING STOCK WITH ESPM SCORES

This section describes analyses related to the subset of benchmarked properties with an ESPM score. Because the ESPM normalizes a property's predicted EUI based on actual building conditions, including "energy driving" variables (e.g. number of occupants) and weather to generate an ESPM score, this score allows a more relevant comparison across buildings and across time than using the EUI alone. The following analyses can be undertaken by first parsing the subset of eligible properties that received ESPM score from the larger set of properties *eligible* to receive a score (as illustrated above) with the subset of eligible properties not receiving scores set aside to determine why scores were not generated and how to remedy any data quality problems (see Chapter 3).

> The benchmarking score is considered a better metric for measuring the energy efficiency of a property than a property's EUI. Although the EUI can be used for benchmarking purposes (especially when buildings are not eligible to receive and ESPM score), this metric does not account for factors that drive energy use in a property and that contribute the considerable variance in EUI across a property type. ESPM normalizes the predicted EUI based on actual building conditions, including weather and "energy driving" variables (e.g. number of occupants, hours of operation, etc.) and compares this predicted EUI to the actual EUI. ESPM generates a score based on this this normalized comparison process, allowing comparison with other buildings and across time. In general, both the EUI and its associated benchmark score offer the best assessment of a property's energy efficiency and potential energy savings.

The following analyses are recommended to be performed for the subset of properties with ESPM scores and can be reviewed in Section 1:

- **Distribution of gross floor area and number of properties with ESPM score by property type**
- **Distribution of ESPM scores for all properties and by selected property types**
- **Distribution and comparison of median site EUI and median source EUI by selected property types**
- **Distribution of median ESPM scores, median site EUI and median source EUI by decade built**
- **Distribution of Property Gross Floor Area by ESPM score**

PARSING PROPERTIES BY ESPM BENCHMARKING SCORE ELIGIBILITY

Energy benchmarking programs are undertaken to understand the energy efficiency of the building stock. A property's EUI is one simple measure of energy efficiency, but it does not adequately account for the types of energy-using activities that occur in a building. For this reason, the EUI usually needs to be compared against some type of "predicted EUI" that has been normalized for energy-driving variables. ESPM, like some other benchmarking tools, can generate normalized predicted EUIs based on assessed energy use from properties across the United States, and compare them to individual property EUIs. Hence, the ESPM benchmark score can be considered an assessment of the energy efficiency of a property based on its source energy use and relative to a set of theoretical "peers," which ultimately allows for comparison of one property to the next. Unfortunately, although ESPM can collect and warehouse energy use and property data on 84 different property types, only 21 of those types can actually receive a benchmark score.

Most benchmarking efforts are based on broad categorizations of building stock (such as: all commercial buildings with a gross floor area above a minimum cut-off point). By casting such a wide net, benchmarking programs will almost certainly gather property and energy use data on a set of property types that far exceeds the 21 types eligible to receive a benchmark score. In light of this, it is important to parse overall benchmarking data into two general subsets of property types—those eligible and those not eligible to receive an ESPM score.

Table 4. Reported benchmarked properties for Philadelphia's 2013 cleansed dataset

	No. of Property Types	No. of Properties:	Gross Floor Area:
Properties Eligible to Receive ESPM Score:	18	651	145,804,625
Properties Not Eligible to Receive ESPM Score:	38	319	74,310,480
Total	56	970	220,115,105

Required Data and Dependencies

- **Property type:** *ESPM data field – "Primary Property Type - EPA Calculated"*
- **Property floor area (typically reported in gross square feet (GSF)):** *ESPM data field – "Property Floor Area (Building(s)) (ft²)"*

Example

Table 4, Figure 17, and Figure 18 show the parsing of the reported benchmarked properties for Philadelphia's 2013 cleansed benchmarking dataset into properties eligible and not eligible to receive an ESPM. About 67% of total reported properties are eligible to receive a score, which represents about 66% of all reported gross floor area, which are distributed across 18 different eligible property types. For property types not eligible to receive a score, the smaller portion of gross floor area (34%) is related to 38 non-eligible property types.

The figures contain a miscellaneous category of "other" properties. This category often represents a large portion of the total properties and thus deserves clarification. The graphical representations of the analyses in this guide are merely shown as guidance. Your specific graphs may have any number of categories and an "other" category of varying magnitude The graphs here only contain 6 categories, including "other" which inflates the magnitude of that category, whereas you may choose to show all categories and eliminate the other category entirely. If only a subset of categories are graphed, it is essential to capture the remaining data in an "other" category so that your audience has a complete understanding of your data.

With your data parsed into the buildings that are eligible and those ineligible for ESPM scores, you can proceed to dig into deeper parsing and analysis.

COMPARISON OF PROPERTIES WITH ESPM SCORES TO ELIGIBLE PROPERTIES THAT DID NOT RECEIVE ESPM SCORES

This analysis addresses the subset of properties eligible to receive an ESPM score but that, for one reason or another, did not. Chapter 3 lists the many possible reasons scores may not be generated; the lack of a score usually indicates data quality issues (often building owners/operators omitting critical building information). While the lack of ESPM scores limits the analyses that can be performed, in most cases, eligible properties lacking ESPM scores can be analyzed like the other subset of properties without ESPM scores, i.e. those property types ESPM deems ineligible to receive a score (such as colleges/ universities, libraries, and museums). The discussion that follows offers data analyses typically applied to *ineligible* property types.

For the subset of properties that are *eligible* to receive scores but that did not, gauging the data quality of these properties is important. One way to do this is to compare these properties with the ones that *did* receive a score. Those property types with a significant fraction of eligible properties without scores may highlight properties for which building

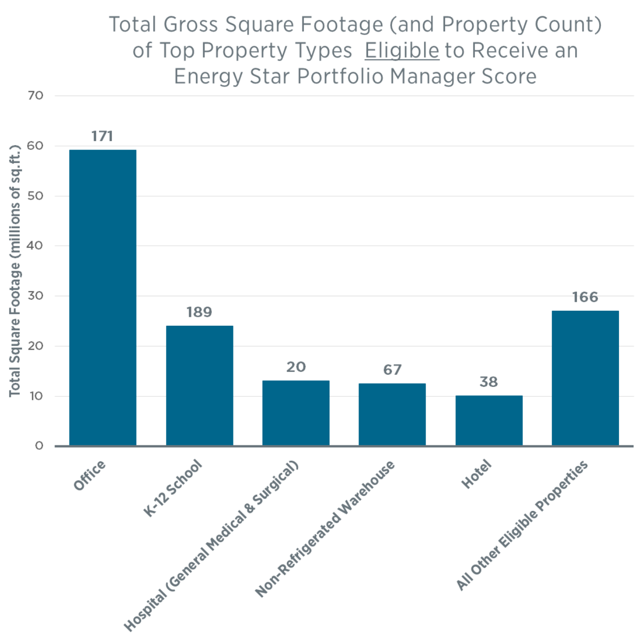


Figure 17. Properties eligible for ESPM Scores

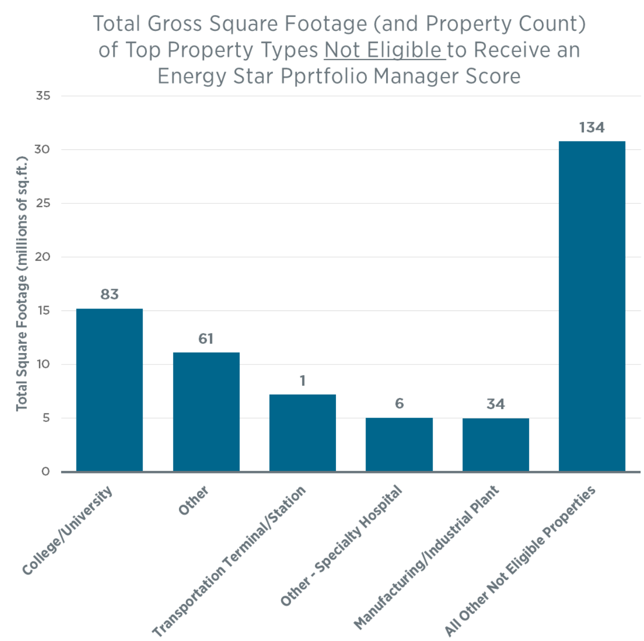


Figure 18. Properties ineligible for ESPM Scores

owners/operators are not correctly inputting data; this information can help benchmarking program administrators tailor advice to these challenging property types in future energy benchmarking training sessions.

Required Data and Dependencies

- **Property type: ESPM data field – "Primary Property Type - EPA Calculated"**
- **Subset of properties eligible to receive an ESPM score but didn't receive a score**
- **Subset of properties having an ESPM score: ESPM data field – "ESPMScore"**

Example

In Figure 19, the subset of eligible properties from Philadelphia's 2013 cleansed benchmarking dataset having ESPM scores (total of 567 properties) are compared to the subset of eligible properties not receiving scores (total of 84 properties) by property type. In particular, distribution centers had a relatively high percentage of eligible properties without ESPM scores (25% of all eligible distribution centers), with non-refrigerated warehouse also showing a significant percentage as well (24% of all eligible non-refrigerated warehouse). Although offices had a smaller percentage (15% of all eligible offices) of eligible properties without ESPM scores, the impact from these 26 office properties without scores is of concern since this property type represents the largest amount of benchmarked square footage.

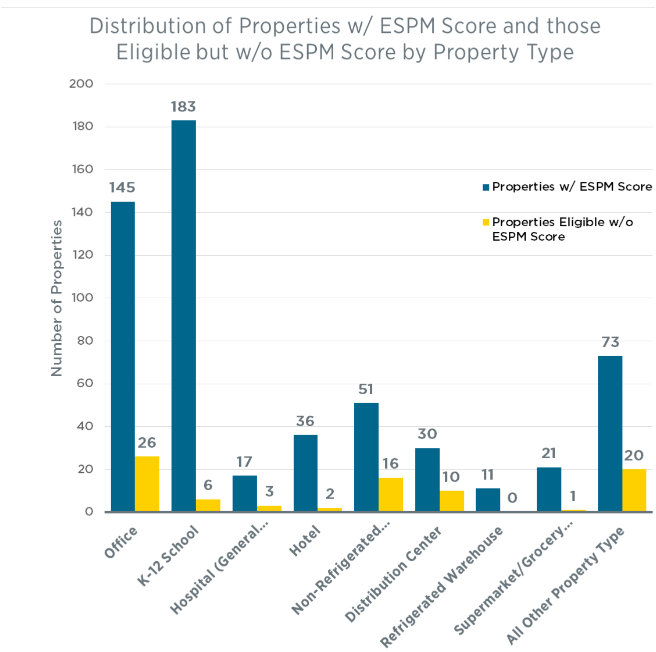


Figure 19. Distribution of properties with ESPM score and those eligible but without a score for Philadelphia, 2013

SECTION 3: PROPERTIES NOT ELIGIBLE TO RECEIVE AN ESPM SCORE

Of the 84 property types defined in ESPM, approximately 63 types are not eligible to receive a benchmark score. Typically, for a benchmarking program designed to gather data on general commercial building stock, the number of ineligible property types will far exceed the number of eligible property types. Generally, the reason so many buildings are ineligible to receive an ESPM score is that there is insufficient statistical correlation between easily defined variables and building energy use in these building types.

Unfortunately, the analyses that can be done on ineligible property types is somewhat limited since 1) they cannot receive an ESPM score, and 2) often, relatively few properties constitute many of the ineligible property types (typically 1 to 6 properties). Having such small numbers of data points in a dataset makes calculations of metrics like median site EUI much less meaningful; this is especially true if there is extremely wide variation in the data.

When the ESPM score is not available for interpretation, the only metric available to assess the energy efficiency of an ineligible property is its EUI; this is problematic since this metric hasn't undergone a normalized comparative process. Valuable information can be derived from an analysis of ineligible properties through the following analyses:

- **Distribution of number of ineligible properties and total gross floor area by property type**
- **Distribution of calculated total energy consumption and fuel share (mix) by select ineligible property type**
- **Distribution of median site EUI by select ineligible property type**

The following analyses should be performed on properties not eligible to receive an ESPM score as part of general analytical due diligence:

- **Distribution of Property Size for Ineligible Properties**
- **Distribution of Number of Properties, Total Gross Floor Area and Median Site and Median Source EUI for Ineligible Properties by Decade Built**

All of these analyses can be reviewed in Section 1 of this chapter.

Chapter 5: Basic Energy Performance and Associated Energy Costs

This chapter offers various analyses dealing with the energy performance and associated energy costs of individual benchmarked properties which can help lay the analytical foundation for identifying properties with a high energy savings potential. In addition to examining the relationships between EUI, ESPM score and energy cost per square foot, other relationships dealing with ESPM score, total annual energy use and total annual energy cost are also investigated in terms of potential energy savings. Moreover, comparisons of these various relationships are made between two selected property types, Office and K-12 School (highlighted because they represent a significant fraction of the Philadelphia building stock and both types are eligible to receive ESPM scores), which will allow the user of this guide to understand the logical consistency between some of the analyses described in Chapter 4 and those found in this chapter.

This chapter also offers an analysis and comparison of the relationship of property gross floor area and total annual energy cost for a subset of selected properties which, "historically," represent the hard to reach (in terms of improved energy efficiency) segment of small- to medium-sized properties 200,000 square feet and smaller.

The types of analyses shown in this Chapter are the following:

- **Comparison and screening of calculated EUI vs ESPM generated EUI**
- **Calculated site and source EUI by ESPM score for select property types**
- **Total energy cost per square foot by calculated site EUI for select property types**
- **Total energy cost per square foot by ESPM score for select property types**
- **Total annual site energy use by ESPM score for select property types**
- **Total annual energy cost by ESPM score for select property types**

COMPARISON AND SCREENING OF CALCULATED EUI VS ESPM GENERATED EUI

The calculated EUI values used for analysis in this chapter are values calculated by summing all annual energy use (normalized to kBtu) across all fuel types for each property divided by the property's gross floor area. These *calculated* EUIs can vary from the value generated by what ESPM *reports* for a property if, for instance, the gross floor area changed during the annual period the EUI was calculated by ESPM, in which case the reported EUI is considered a weighted average value. In light of this potential difference, reported EUIs and calculated EUIs must be compared to identify which properties have two differing values. The properties having a different reported EUI and calculated EUI should be removed prior to performing the analyses described in this chapter. (Note: if the reported EUI is different than the calculated EUI for a property, the associated ESPM score is based on the reported EUI not the calculated EUI.) For the purposes of consistency, properties having the same reported and calculated EUI implies the ESPM score has been generated appropriately for the calculated EUI.

Required Data and Dependencies

- **Calculated site EUI - sum of individual site fuel shares for each property type (normalized to kBtu): typical ESPM data fields –** *"Electricity Use - Grid Purchase and Generated from Onsite Renewable Systems (kBtu)", "Natural Gas Use (kBtu)", "Fuel Oil #2 Use (kBtu)" and "District Steam Use (kBtu)"divided by property floor area*
- **Calculated source EUI - sum of individual site fuel shares for each property type (normalized to kBtu): typical ESPM data fields –** *"Electricity Use - Grid Purchase and Generated from Onsite Renewable Systems (kBtu)", "Natural Gas Use (kBtu)", "Fuel Oil #2 Use (kBtu)" and "District Steam Use (kBtu)" multiplied by the appropriate Source-to-Site Ratio found in the following table then divided by property floor area*
- **Property type: ESPM data field –** *"Primary Property Type - EPA Calculated"*
- **Property floor area (typically reported in gross square feet (GSF)): ESPM data field –** *"Property Floor Area (Building(s)) (ft^2)"*
- **Annual site and source EUI (kBtu/sq/yr): ESPM data field –** *"Site EUI (kBtu/ft^2)", "Source EUI (kBtu/ft^2)"*
- **Subset of properties that received an ESPM score**

Example

For the analyses shown in this Chapter, the properties with ESPM scores contained in the Office and K-12 School property type from Philadelphia's cleansed 2013 benchmarking dataset were screened for differences in ESPM generated and calculated source and site EUIs. The results of the screening process are shown in Table 5.

SITE AND SOURCE EUI BY ESPM SCORE FOR SELECT PROPERTY TYPES

One of the most common ways to assess the overall energy efficiency of a set of properties classified under a specific property type is to plot the relationship of the properties' ESPM scores and associated site and source EUIs. By doing this, variations in EUIs for properties with same ESPM score and, conversely, differences in scores for properties with the same EUI can be seen. Understanding these two types of differences is critical to gaining an overall understanding of how ESPM measures the energy efficiency of benchmarked properties.

Required Data and Dependencies

- **Property type: ESPM data field –** *"Primary Property Type - EPA Calculated"*
- **Subset of properties that received an ESPM score and have screened calculated site and source EUIs (see first analysis of this chapter)**

Example

Figure 20 and Figure 21 show the distribution of calculated site and source EUI by associated ESPM score for office and K-12 School properties, respectively, from Philadelphia's 2013 cleansed dataset. For both figures, a general trend is evident where properties with high ESPM scores typically have lower site and source EUIs, and vice versa. In other words, this data shows that site or source EUI can be an indicator of a property's level of energy efficiency.

However, while this may be generally true when analyzing a large sample size, there are instances

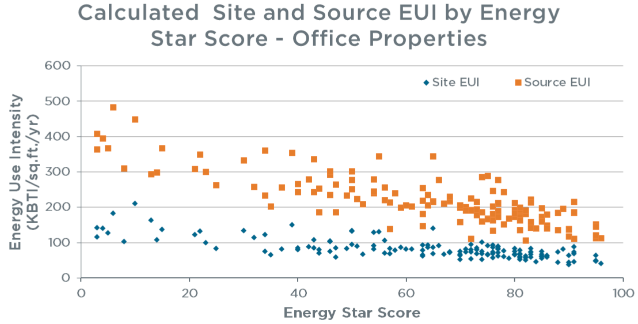


Figure 20. Distribution of calculated site and source EUI by associated ESPM score for Office properties for Phila., 2013

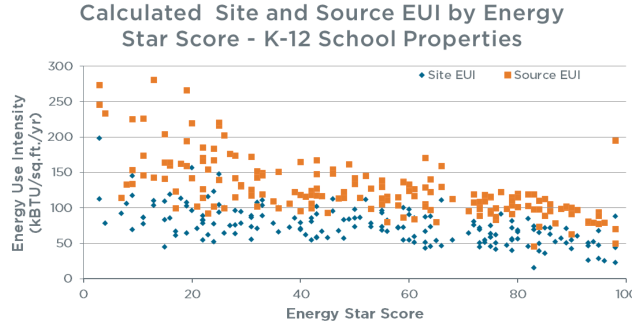


Figure 21. Dist. of calculated site and source EUI by associated ESPM score for K-12 School properties for Phila., 2013

when the EUI and energy score may not correlate well. As an example, Figure 23 shows that it is possible for a group of buildings to have similar EUIs, yet have significantly different energy star scores. This discrepancy is due to the way that the ENERGY STAR score is generated, which considers building operation as well as total energy use. Site or source EUI may not be adequate indicators of a property's level of energy efficiency when the energy loads and usage patterns in the building vary significantly from an average building, such as when an office building is operated 24 hours per day.

Figure 37 Office and K-12 school properties with the same source EUI but different ENERGY STAR score for Philadelphia, 2013

The opposite can also be seen. Figure 22 shows that it is possible for a group of buildings to have similar energy star scores, yet have significantly different EUIs, which is due to the same factor.

In general, a property's EUI reflects the total amount of energy consumed, while the property's ENERGY STAR score reflects how efficiently the energy was used.

When comparing the site and source EUI distribution by ENERGY STAR score, note also the relative difference in magnitude between the properties' site and source EUIs. In general, the greater the source-to-site ratio, the greater the electricity use fuel share, with "all-electric" properties having the largest source-to-site ratio of 3.14. Figure 24 is a scatter plot of the source-to-site EUI ratios for both office and K-12 school properties; it shows that the office properties have larger ratios than the K-12 School properties, a strong indication that these office properties tend to have larger electricity fuel shares. The grouping of K-12 school properties typically have smaller source-to-site EUI ratios when compared to office properties, which is commensurate with the office and K-12 school property total fuel share distribution shown in Figure 15 (Section 1, Chapter 4).

TOTAL ENERGY COST PER SQUARE FOOT BY CALCULATED SITE EUI FOR SELECT PROPERTY TYPE

This analysis examines total energy cost per square foot relative to site EUI.

Required Data and Dependencies

- **Subset of properties that received an ESPM score and have screened calculated site EUIs (see first analysis of this chapter)**
- **Fuel cost ratios ($ per kBtu): either local average $/kBtu (see Table 5) or ESPM fuel share cost data divided by appropriate fuel share kBtu**
- **Estimated total energy cost per square foot – Individual site fuel shares for each property type (normalized to kBtu): typical ESPM data fields –** *"Electricity Use - Grid Purchase and Generated from Onsite Renewable Systems (kBtu)", "Natural Gas Use (kBtu)", "Fuel Oil #2 Use (kBtu)" and "District Steam Use (kBtu)" multiplied by the appropriate fuel cost ratios (see Table 6then summed and divided by property floor area*
- **Property floor area (typically reported in gross square feet (GSF)): ESPM data field –** *"Property Floor Area (Building(s)) (ft²)"*

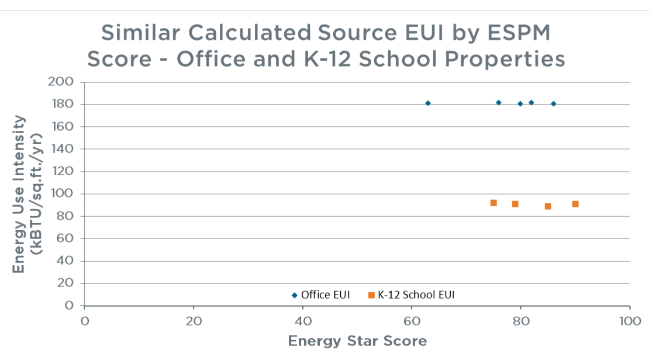


Figure 22. Office and K-12 school properties with the same source EUI but different ESPM score for Philadelphia, 2013

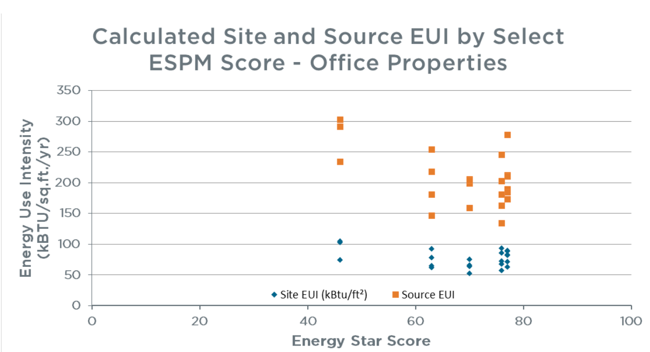


Figure 23. Office properties with the same ESPM score but different site and source EUIs for Philadelphia, 2013

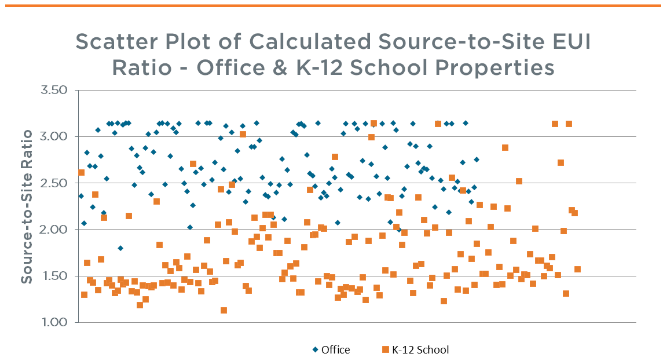


Figure 24. Calculated source-to-site EUI ratio for office and K-12 school properties for Philadelphia, 2013

Table 5. Office and K-12 School properties with ESPM scores

Dataset	Office	K-12 School
Total # of Properties w/ ESPM Score:	145	183
# of Properties w/ ESPM Score and Equivalent Reported and Calculated EUI:	142	177

Example

For this analysis, the cost per square foot of office and K-12 school properties (with calculated EUIs and ESPM scores) from Philadelphia's 2013 cleansed dataset are plotted in relationship to their associated EUI. In Figure 25, costs per square foot are typically greater for offices than for K-12 schools for the range of EUIs shown, which makes logical sense since offices typically have a larger electric fuel share when compared to K-12 Schools and electricity has one of the highest $/kBtu. (Note: the various $/kBtu shown in Table 6 are representative of typical fuel costs found in Philadelphia and are based on various analysis of utility billing data).

For the majority of offices, costs per square foot range from about $1.00/sf to $3.00/sf, while for the majority of K-12 schools the $/sf range is about $0.50 to $2.00. Even for offices and K-12 schools with the same EUI, offices will typically have higher costs per square foot due the larger percentage of electric use. Based on these $/sf distributions, offices properties may offer better energy efficiency investment opportunities when compared to K-12 schools since higher $/sf can lead to larger overall cost savings.

TOTAL ENERGY COST PER SQUARE FOOT BY ESPM SCORE FOR SELECT PROPERTY TYPES

This analysis deals with understanding the relationship between total energy cost per square foot and ESPM score for a property.

Required Data and Dependencies

- **Subset of properties that received an ESPM score and have screened calculated site EUIs (see first analysis of this chapter)**
- **Calculated total energy cost per square foot (see previous analysis for calculating $/sf)**

Example

In order to see the relationship of cost per square foot and ESPM score, properties from Philadelphia's 2013 cleansed data contained in the Office and K-12 School property types (with calculated EUIs and ESPM scores) are plotted in relationship to their associated

Table 6. Typical fuel costs found in Philadelphia (based on analysis of billing data, 2014)

FUEL SHARE:	$/KBTU
Electric:	$0.0293
Natural Gas:	$0.0126
Fuel Oil:	$0.0205
Steam:	$0.0340

ESPM score. Figure 26 shows that although costs per square foot are typically larger for properties having lower, less efficient ESPM scores, costs per square foot can still vary significantly for a subset of properties having similar scores. Moreover, properties with similar $/sf can have significantly different levels of efficiency (e.g., offices with a $/sf of about $2.50). In general, properties with a high $/sf and a low ESPM score can offer better opportunities for energy efficiency improvement when compared to more efficient properties having a much lower $/sf.

TOTAL ANNUAL SITE ENERGY USE BY ESPM SCORE

This analysis deals with understanding the relationship between total annual energy use and ESPM score for a property. As opposed to an EUI where energy use has been normalized by the size of a property (sf), an assessment of a property's total annual energy use will offer some indication of what limits exist for installing certain types of energy efficiency measures. Energy service providers, in many cases, will look at both the total annual amount of energy use and current level of energy efficiency of a property to gauge the types of energy efficiency measures that may be installed to meet energy reduction goals, such as a mandated 20% reduction in total energy use. In general, properties with high total annual energy use and low ESPM scores can represent relatively good candidates for realizing significant energy savings.

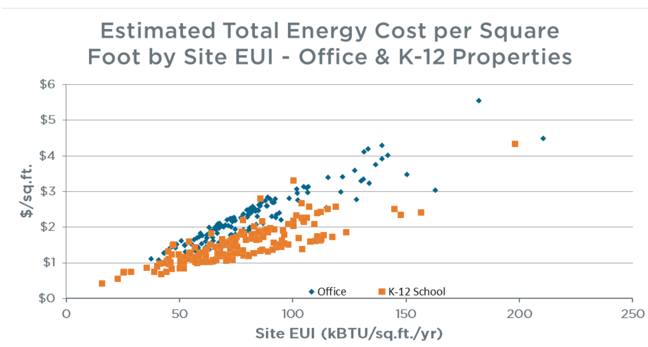


Figure 25. Estimated total energy cost per floor area by site EUI for office and K-12 school properties for Philadelphia, 2013

Required Data and Dependencies

- **Subset of properties that received an ESPM score and have screened calculated site EUIs (see the first analysis of this chapter)**
- **Calculated sum of individual fuel shares for each property type (normalized to kBtu): typical ESPM data fields –** *"Electricity Use - Grid Purchase and Generated from Onsite Renewable Systems (kBtu)", "Natural Gas Use (kBtu)", "Fuel Oil #2 Use (kBtu)" and "District Steam Use (kBtu)"converted to million Btu (mmBtu)*

Example

Figure 27 shows the total energy use of office and K-12 school properties (with calculated EUIs and ESPM scores) from Philadelphia's 2013 cleansed dataset plotted in relationship to their associated ESPM score. The majority of K-12 Schools use less than 20,000 mmBtu per year of energy and show, at the annual level, little correlation with ESPM score, indicating the energy efficiency of this building type is quite variable. Offices, on the other hand, show a grouping of larger energy consuming properties with higher ESPM scores, which is commensurate with the distribution of larger sized (sf) offices having higher scores (see Section 2, Chapter 4).

SUMMARY

Many municipalities across the United States see benchmarking programs as an initial step in laying a foundation for driving energy retrofits in their building stock. Based on all of the analyses found in this guide, it would be understandable to ask: "Which are the building types and characteristics that may be good candidates for retrofits?" While the answer to this question may be a bit complicated, the following suggestions should provide guidance in identifying properties that could benefit from energy efficiency retrofits.

1. **Identify building types that have received an energy star score.**
2. **Select a subset of properties in these building types having ESPM scores of 74 or less.**
3. **Identify the properties with site EUIs equal to or greater than the median EUI from the subset in #2.**
4. **Identify the properties with $/sf equal to or greater than the median value from the subset in #3.**

Once these four steps have been applied to the building types that have received an energy star score, combined the final subset (#4 above) of all the building types into one dataset of properties. This combined dataset should represent those properties that have a fairly high potential of benefiting from energy retrofits.

For further information on identifying properties for further investigation related to efficiency measures, please refer to the Intermediate Guide on Energy Benchmarking.

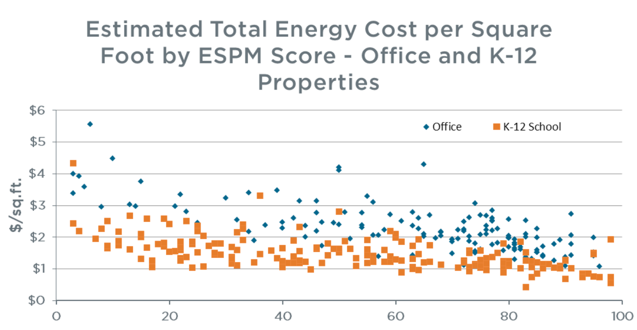


Figure 26. Total energy cost per floor area by ESPM score for office and K-12 school properties for Philadelphia, 2013

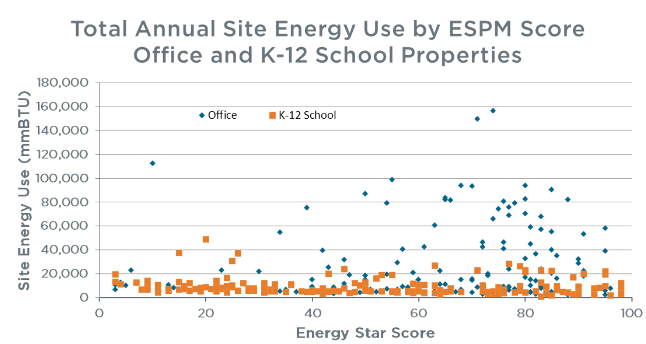
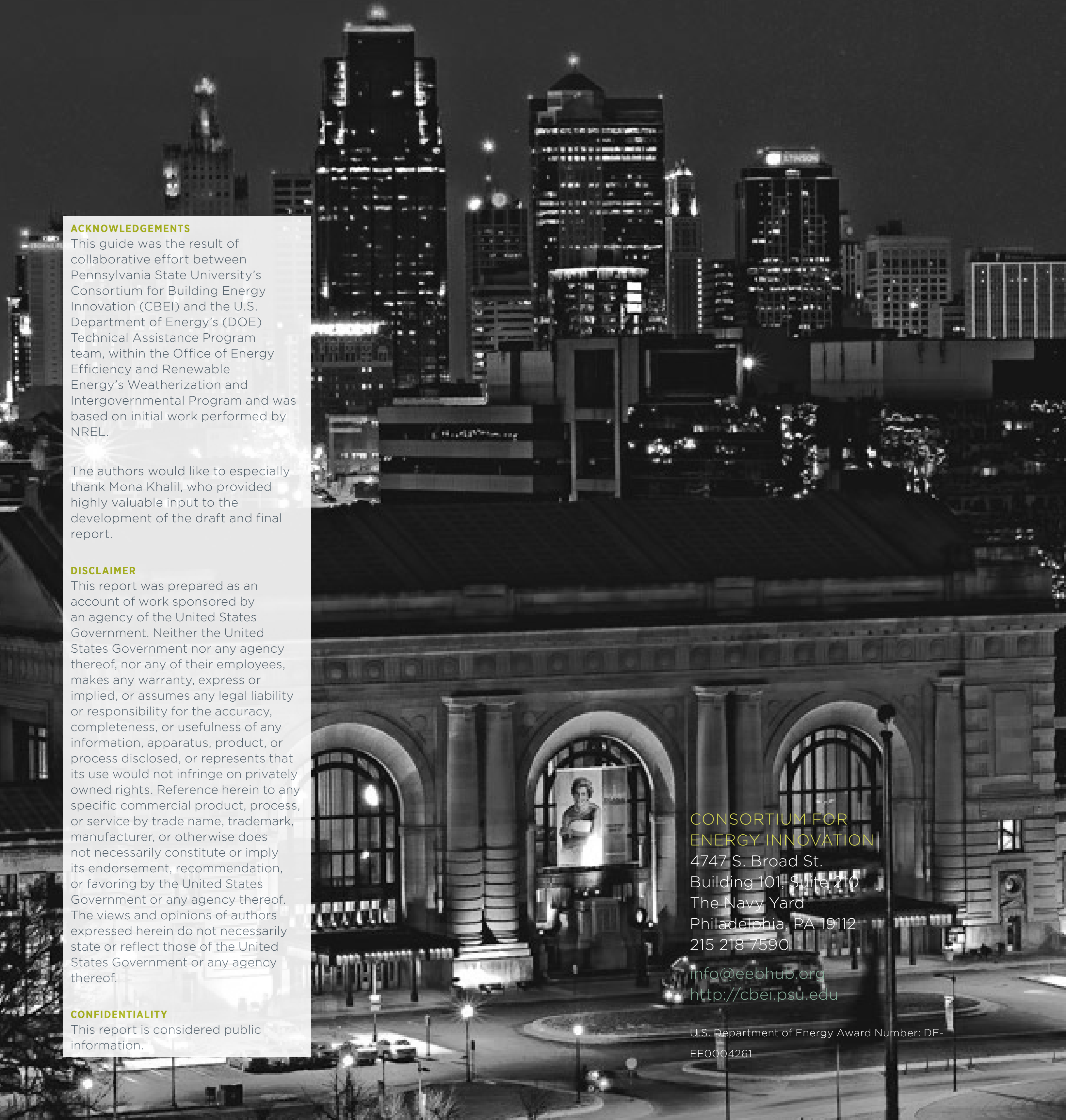


Figure 27. Total site energy use by ESPM score for office and K-12 school properties for Philadelphia, 2013

CONSORTIUM for
BUILDING ENERGY
INNOVATION



ACKNOWLEDGEMENTS

This guide was the result of collaborative effort between Pennsylvania State University's Consortium for Building Energy Innovation (CBEI) and the U.S. Department of Energy's (DOE) Technical Assistance Program team, within the Office of Energy Efficiency and Renewable Energy's Weatherization and Intergovernmental Program and was based on initial work performed by NREL.

The authors would like to especially thank Mona Khalil, who provided highly valuable input to the development of the draft and final report.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe on privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

CONFIDENTIALITY

This report is considered public information.

CONSORTIUM FOR
ENERGY INNOVATION
4747 S. Broad St.
Building 101, Suite 210
The Navy Yard
Philadelphia, PA 19112
215 218 7590

info@eebhub.org
http://cbei.psu.edu

U.S. Department of Energy Award Number: DE-EE0004261