

REFERENCE DATA & FILTERS

The reference data used to form the peer property population relies on Survey of Energy Consumption of Multi-Unit Residential Buildings (SECMURBs), which was commissioned by Natural Resources Canada (NRCan) and represents the energy use for the year 2018. The raw collected data file for this survey is not publicly available, but a report providing summary results is available on Natural Resources Canada’s website at: <https://oee.nrcan.gc.ca/corporate/statistics/neud/dpa/menus/murb/2018/tables.cfm>

Four types of filters are applied to analyse the building energy and operating characteristics in the survey. They are set to define the peer population for comparison and to overcome any technical limitations. Those filters are: Building Type Filters, Program Filters, Data Limitation Filters, and Analytical Filters.

A complete description of each category is given in the Technical Reference for the ENERGY STAR Score, at www.energystar.gov/ENERGYSTARScore. **Figure 1** summarizes each filter used to develop the ENERGY STAR score for the multifamily housing model and the rationale that supports the filter. After all filters are applied, the remaining data set has 219 observations. Due to the confidentiality of the survey data, NRCan is not able to identify the number of cases after each filter.

Figure 1 – Summary of Filters for the ENERGY STAR Score for Multifamily Housing

Condition for Including an Observation in the Analysis	Rationale
Buildings must be more than 50% multifamily housing and less than 50% of another building type	Building Type Filter – To be considered part of the multifamily housing peer population, the building must have a minimum multifamily housing space.
The building must be either four storeys above ground or have a horizontal footprint of 600 m ²	Program Filter – Requirement to be considered a commercial multifamily housing building.
Must have complete data for whole property energy use and operating characteristics	Program Filter – Complete data is necessary for analysis
Must be greater than 50% occupied by tenants	Program Filter – Must be majority occupied to be considered part of the multifamily housing peer population.
Must have electricity consumption data	Program Filter – Multifamily housing that do not use electricity are rare or non-existent and may indicate an omission in energy data. Electricity can be grid-purchased or produced on site.
Consumption data must be for energy year 2018	Data Limitation Filter – Multifamily housing with consumption data whose start and end date do not align or extend significantly outside of consumption year 2018 are not comparable to the peer population
Must not use any “other” fuels for which the consumption is not reported	Data Limitation Filter – Energy types defined as “other” could not be compared directly. In such cases, these observations were removed from the analysis.

Condition for Including an Observation in the Analysis	Rationale
The percent of the building that is heated must be greater than 50%	Program Filter – Multifamily housing must be at least 50% heated to be considered multifamily housing in Canada.
The size of the indoor or partially enclosed parking structures must be less than 50% of the gross floor area including indoor and partially enclosed parking structures	Program Filter – If the combined square foot of parking structures exceeds the size of the multifamily housing, then the overall structure is classified as parking, not multifamily housing. This is a standard policy in Portfolio Manager.
The count of bedrooms in the multifamily housing must be greater than or equal to the count of residential units	Program Filter – Requirement to be considered multifamily housing. Less bedrooms than units is likely due to missing data.
The multifamily housing must contain at least 2 residential units	Program Filter – Requirement to be considered multifamily housing.
The average gross floor per apartment must not exceed 4,000 ft ² (371.6 m ²)	Analytical Filter – Values determined to be statistical outliers.
Must have a source EUI less than or equal to 2.5 GJ/m ² and greater than or equal to 0.3 GJ/m ²	Analytical Filter – Values determined to be statistical outliers.
Must have a residential unit density less than or equal to 3 units per 100 m ²	Analytical Filter – Values determined to be statistical outliers.
Must have a bedroom density less than or equal to 4 bedrooms per 100 m ²	Analytical Filter – Values determined to be statistical outliers.

Of the filters applied to the reference data, some result in constraints on calculating a score in Portfolio Manager, and others do not. Building Type and Program Filters are used to limit the reference data to include only properties that are intended to receive a score in Portfolio Manager, and are therefore related to eligibility requirements. In contrast, Data Limitation Filters account for limitations in the data available during the analysis, but do not apply in Portfolio Manager. Analytical Filters are used to eliminate outlier data points or different subsets of data, and may or may not affect eligibility. In some cases, a subset of the data has a different behaviour from the rest of the properties in which case an Analytical Filter is used to determine eligibility in Portfolio Manager. In other cases, Analytical Filters exclude a small number of outliers with extreme values that skew the analysis, but do not affect eligibility requirements. A full description of the criteria you must meet to obtain a score in Portfolio Manager is available at <https://www.nrcan.gc.ca/energy-efficiency/energy-star-canada/benchmarking-frequently-asked-questions/3787#es17>

Related to the filters and eligibility criteria described above, another consideration is how Portfolio Manager treats properties situated on a campus. The main unit for benchmarking in Portfolio Manager is the property, which may be used to describe either a single building or campus of buildings. The applicability of the ENERGY STAR score depends on the type of property. For multifamily housing, the score applies to the entire property, whether it is a single building or a campus of buildings. Multifamily housing may have multiple buildings that are all integral to the primary activity. One building may be leasing office and laundry, another, the actual residential units. In these cases, the campus can get an ENERGY STAR score as long as the energy for all the buildings is metred and reported. For cases where all the activities are contained within one building, that multifamily housing building can get an ENERGY STAR score.

SURVEY WEIGHTS

A comprehensive national listing of multifamily buildings was not available to create a nationally representative survey, so an area probability sample was used in eight census metropolitan areas (CMAs). The eight CMAs sampled are: Vancouver, Calgary, Winnipeg, Toronto, Ottawa-Gatineau, Hamilton, Montreal, and Halifax. As such, the sample weights represent the estimated number of multifamily properties located within these CMAs.

VARIABLES ANALYSED

To normalize for differences in business activity, NRCan performed a statistical analysis to understand what aspects of building activity are significant with respect to energy use. The filtered reference data set, described in the previous section, was analysed using a weighted ordinary least squares regression, which evaluated energy use relative to business activity (e.g. number of workers, operating hours per week, floor area, and climate). This linear regression gives an equation used to compute energy use (also called the dependent variable) based on a series of characteristics that describe the business activities (also called independent variables). This section details the variables used in the statistical analysis for multifamily housing in Canada.

Dependent Variables

The dependent variable is what NRCan tries to predict with the regression equation. For the multifamily housing analysis, the dependent variable is energy use, expressed in source energy use intensity (source EUI). This is equal to the total source energy use of the property divided by the gross floor area. The regression analyses the key drivers of source EUI—those factors that explain the variation in source energy use per square metre in multifamily housing. The units for source EUI in the Canadian model are annual gigajoules per square metre (GJ/m²).

Independent Variables

The reference survey contains numerous property operation questions that NRCan identified as likely to be important for multifamily housing. Based on a review of the variables found in the reference data, and following the criteria for inclusion in Portfolio Manager,¹ NRCan initially analysed the following variables in the regression analysis:

- Gross floor area (m²)
- Cooling degree days (CDD)
- Heating degree days (HDD)
- Percentage of floor space that is cooled
- Percentage of floor space that is heated
- Number of residential units
- Number of bedrooms
- Residence population type
- Residential unit ownership type
- Percent common area
- Number of dishwasher hookups

¹ For a complete explanation of these criteria, refer to the Technical Reference for the ENERGY STAR Score, at www.energystar.gov/ENERGYSTARScore.

- Presence of laundry
- Location of laundry services (In unit versus common area)
- Number of laundry hookups
- Presence of fitness room
- Presence of party room
- Presence of charging stations for electric vehicles
- Presence of hot tubs, whirlpools and Jacuzzi tubs

NRCan, with the advice of the Environmental Protection Agency (EPA) and its contractor, performed an extensive review on all of these operational characteristics individually and in combination with each other (e.g. heating degree-days times percent heated). As part of the analysis, some variables were reformatted to reflect the physical relationships of building components. For example, the number of residential units can be evaluated in a density format: units per 100m². The unit density (as opposed to the gross number of units) is more closely related to the energy use intensity. In addition, using analytical results and residual plots, variables were assessed using different transformations (such as the natural logarithm, abbreviated as Ln). Overall, the analysis consists of multiple regression formulations, structured to find the combination of statistically significant operating characteristics that explained the greatest amount of variance in the dependent variable: source EUI.

The final regression equation includes the following variables:

- Number of bedrooms per 100m², with a ceiling at 2.3 (Bedroom Density)
- Number of residential units per 100m² (Unit Density)
- The percentage of residential units in a mid-rise or high-rise setting
- The percentage of the building that is cooled times the number of Cooling Degree Days (Percent Cooled x CDD)
- The percentage of the building that is heated times the number of Heating Degree Days (Percent Heated x HDD)

These variables are used together to compute the predicted source EUI for multifamily housing. The predicted source EUI is the mean EUI for a hypothetical population of buildings that share the same values for each of these characteristics. It is the mean energy for buildings that operate like your building. There were several variables for which NRCan did not have adequate data to make a determination regarding inclusion in the model.

Analysis of High, Mid, and Low-Rise Buildings

Analysis of building heights in multifamily housing showed a trend towards shorter buildings consuming less energy than taller buildings. To further analyse this trend, buildings were placed into three height bins: low-rise (1-4 storeys), mid-rise (5-9 storeys), and high-rise (10 or more storeys). It was noted that between low-rise and mid or high-rise, there was a large difference in energy use intensity (EUI) with the low-rise bin consuming less energy per square metre than taller buildings. Among other operational differences, this is attributed to low-rise buildings having less common space (hallways, amenities, etc.), which has a higher energy intensity than simple residential units. This aligns with trends present in Portfolio Manager data, as well as trends observed by the EPA in their multifamily datasets.

In our analysis, each property was 100% low-rise, mid-rise or high-rise. However, in Portfolio Manager, it is possible for a community to have buildings of different heights. Therefore, the percent mid-high rise variable is determined as the percentage of the property's units that are in a mid or high-rise building. For instance, if a property consists of two buildings—one mid-rise building with 60 units and one low-rise building with 40 units—the mid-high rise variable value will be 0.6 (mid-high rise = $60/(60+40) = 0.6$).

Laundry

The presence of laundry was analysed to see if it contributed to increased energy use in multifamily buildings. Unfortunately, a conclusion could not be made, as there was little variability in the survey data. 95% of properties had either laundry facilities in a common area, or had the necessary hookups for tenants to install laundry in their own unit. Because so few buildings without laundry responded, it was not possible to determine a statistically significant difference between the population with laundry, and the population without laundry.

Climate Variables

Climate is one characteristic that was examined closely. NRCan analysed the relationship between EUI and both Cooling Degree Days (CDD) and Heating Degree Days (HDD). While HDD was found to be consistently significant in the models, analysis showed that CDD also contributed to increased energy use, although to a lesser extent compared to HDD. Due to the limited variance of CDD within a single sample year, and the complex relationship between HDD and CDD in different climate regions in Canada, NRCan had concerns that the impact of cooling was not being fully represented in the model. To adjust, NRCan used a combination of analysis techniques, including linear regression of EUI and Percent Cooled x CDD in the survey data, review of engineering models, and comparisons to Portfolio Manager data, to determine an appropriate factor to account for cooling energy. From the analysis, the rate of increase of source energy to Percent Cooled x CDD was determined. Applying this adjustment resulted in a more comprehensive model with both cooling and heating terms that is better adapted to changing climate trends.

Testing

NRCan further analysed the regression equation using actual data entered in Portfolio Manager. In addition to the survey data, this analysis provided another set of buildings to examine the ENERGY STAR scores and distributions to assess the impacts and adjustments. It also confirmed that there are minimal biases when it comes to fundamental operational characteristics, such as Unit Density or Percent Heated x HDD, and that there was no regional bias or bias for the type of energy used for heating.

It is important to reiterate that the final regression equation is based on the nationally representative reference data from the survey, not on data previously entered into Portfolio Manager.

REGRESSION EQUATION RESULTS

The final regression is a weighted ordinary least squares regression across the filtered data set of 219 observations. The dependent variable is source EUI. Each independent variable is centred relative to the weighted mean value, presented in **Figure 2**. The final equation is presented in **Figure 3**. All variables in the regression equation are considered significant at a 85% confidence level or better, as shown by their respective significance levels.

The regression equation has a coefficient of determination (R^2) value of 0.1545, indicating that this equation explains 15.45% of the variance in source EUI for multifamily housing. Because the final equation is structured with energy per unit area as the dependent variable, the explanatory power of the area is not included in the R^2 value, and thus this value appears artificially low. Recomputing the R^2 value in units of source energy² demonstrates that the equation actually explains 87.75% of the variation in total source energy of multifamily housing. It is an excellent result for a statistically based energy model.

For detailed information on the ordinary least squares regression approach, go to the Technical Reference for the ENERGY STAR Score, at www.energystar.gov/ENERGYSTARscore.

Figure 2 – Descriptive Statistics for Variables in Final Regression Equation

Variable	Minimum	Median	Maximum	Mean
Source energy per square metre (GJ/m ²)	0.3611	1.067	2.465	1.115
Bedroom Density*	0.5893	1.767	2.3	1.726
Unit Density	0.4071	1.140	2.298	1.228
Percent Units in Mid-Rise or High-Rise	0	1	1	0.7346
Percent Cooled x CDD	0	0	496	113.1
Percent Heated x HDD	2016	3828	5456	3804

*Bedroom Density has a ceiling at 2.3 beds per 100 m²

² The R^2 value in Source Energy is calculated as: $1 - (\text{Residual Variation of } Y) / (\text{Total Variation of } Y)$. The residual variation is sum of $[\text{Weight} * (\text{Actual Source Energy}_i - \text{Predicted Source Energy}_i)]^2$ across all observations. The total variation of Y is the sum of $[\text{Weight} * (\text{Actual Source Energy}_i - \text{Weighted Mean Source Energy})]^2$ across all observations.

Figure 3 – Final Regression Results

Summary				
Dependent variable	Source energy use intensity (GJ/m ²)			
Number of observations in analysis	219			
R ² value	0.1700			
Adjusted R ² value	0.1545			
F statistic	10.96			
Significance (p-level)	< 0.0001			
	Unstandardized Coefficients	Standard Error	T Value	Significance (p-level)
Constant	1.115	2.462E-02	45.29	<.0001
Bedroom Density*	0.1235	7.924E-02	1.56	0.1205
Unit Density	0.2363	7.576E-02	3.12	0.0021
Percent Units in Mid-Rise or High-Rise	0.2138	5.879E-02	3.64	0.0003
Percent Cooled x CDD**	0.0002	n/a	n/a	n/a
Percent Heated x HDD	5.950E-05	2.939E-05	2.02	0.0441

- Notes:
- *Bedroom Density has a ceiling at 2.3 beds per 100 m²
- **The Percent Cooled x CDD coefficient was restricted to the average GJ/m² for cooling degree-days x percent cooled identified through analysis performed by NRCAN. The analysis showed that, on average, source heating EUI increases by 0.0002 GJ/m² for every cooling degree days x percent cooled.
- The regression is a weighted ordinary least squares regression, weighted by the dataset variable "finalwt."
- All model variables are centered. The centered variable is equal to the difference between the actual value and the observed mean. The observed mean values are presented in Figure 2.
- Heating and cooling degree days are sourced from Canadian weather stations included in the U.S. National Climatic Data Center system.

ENERGY STAR SCORE LOOKUP TABLE

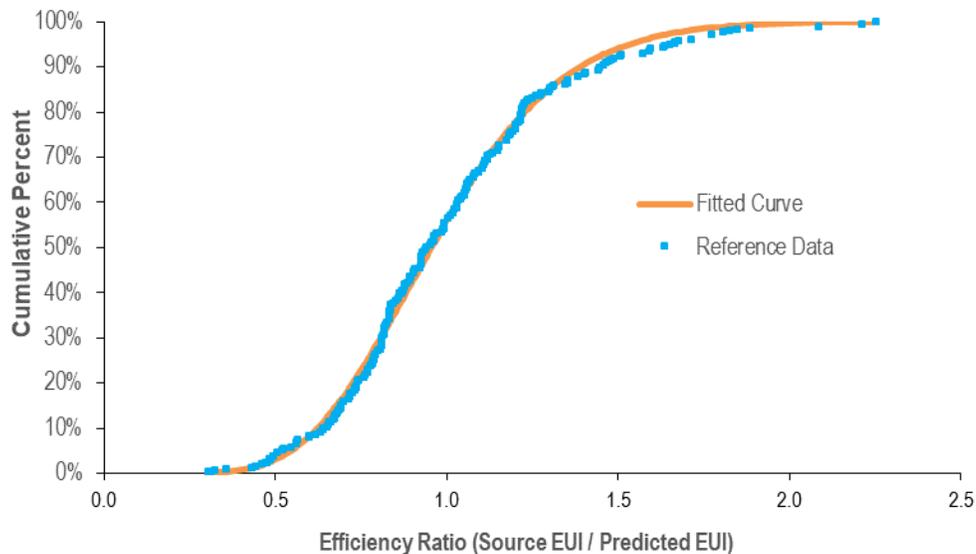
The final regression equation (presented in **Figure 3**) gives a prediction of source EUI based on a building's operating characteristics. Some buildings in the SCIEU data sample use more energy than predicted by the regression equation, while others use less. The *actual* source EUI of each reference data observation is divided by its *predicted* source EUI to calculate an energy efficiency ratio:

$$\text{Energy Efficiency Ratio} = \frac{\text{Actual Source Energy Intensity}}{\text{Predicted Source Energy Intensity}}$$

An efficiency ratio lower than one indicates that a building uses less energy than predicted and consequently is more efficient. A higher efficiency ratio indicates the opposite.

The efficiency ratios are sorted from smallest to largest, and the cumulative percent of the population at each ratio is computed using the individual observation weights from the reference data set. **Figure 4** presents a plot of this cumulative distribution. A smooth curve (shown in orange) is fitted to the data using a two-parameter gamma distribution. The fit is performed in order to minimize the sum of squared differences between each building’s actual percent rank in the population and each building’s percent rank with the gamma solution. The final fit for the gamma curve gives a shape parameter (alpha) of 10.58 and a scale parameter (beta) of 0.0932. The sum of the squared error for this fit is 0.0499.

Figure 4 – Distribution for Multifamily Housing



The final gamma shape and scale parameters are used to calculate the efficiency ratio at each percentile (1 to 100) along the curve. For example, the ratio on the gamma curve at 1% corresponds to a score of 99; only 1% of the population registers such a small or even smaller ratio. The ratio on the gamma curve at the value of 25% corresponds to the ratio for a score of 75; only 25% of the population registers such a small or even smaller ratio. **Figure 5** shows the complete score lookup table.

Figure 5 – ENERGY STAR Score Lookup Table for Multifamily Residential

ENERGY STAR Score	Cumulative Percent	Energy Efficiency Ratio >=	Energy Efficiency Ratio <	ENERGY STAR Score	Cumulative Percent	Energy Efficiency Ratio >=	Energy Efficiency Ratio <
100	0%	0.0000	0.4193	50	50%	0.9549	0.9624
99	1%	0.4193	0.4670	49	51%	0.9624	0.9699
98	2%	0.4670	0.4992	48	52%	0.9699	0.9775
97	3%	0.4992	0.5244	47	53%	0.9775	0.9851
96	4%	0.5244	0.5455	46	54%	0.9851	0.9928
95	5%	0.5455	0.5640	45	55%	0.9928	1.0005
94	6%	0.5640	0.5806	44	56%	1.0005	1.0083
93	7%	0.5806	0.5957	43	57%	1.0083	1.0162
92	8%	0.5957	0.6097	42	58%	1.0162	1.0242
91	9%	0.6097	0.6227	41	59%	1.0242	1.0322
90	10%	0.6227	0.6351	40	60%	1.0322	1.0404
89	11%	0.6351	0.6468	39	61%	1.0404	1.0487
88	12%	0.6468	0.6579	38	62%	1.0487	1.0570
87	13%	0.6579	0.6686	37	63%	1.0570	1.0655
86	14%	0.6686	0.6789	36	64%	1.0655	1.0741
85	15%	0.6789	0.6889	35	65%	1.0741	1.0829
84	16%	0.6889	0.6985	34	66%	1.0829	1.0918
83	17%	0.6985	0.7079	33	67%	1.0918	1.1009
82	18%	0.7079	0.7170	32	68%	1.1009	1.1101
81	19%	0.7170	0.7260	31	69%	1.1101	1.1195
80	20%	0.7260	0.7347	30	70%	1.1195	1.1291
79	21%	0.7347	0.7433	29	71%	1.1291	1.1389
78	22%	0.7433	0.7517	28	72%	1.1389	1.1490
77	23%	0.7517	0.7599	27	73%	1.1490	1.1593
76	24%	0.7599	0.7681	26	74%	1.1593	1.1699
75	25%	0.7681	0.7761	25	75%	1.1699	1.1807
74	26%	0.7761	0.7840	24	76%	1.1807	1.1919
73	27%	0.7840	0.7919	23	77%	1.1919	1.2034
72	28%	0.7919	0.7996	22	78%	1.2034	1.2153
71	29%	0.7996	0.8073	21	79%	1.2153	1.2276
70	30%	0.8073	0.8149	20	80%	1.2276	1.2404
69	31%	0.8149	0.8225	19	81%	1.2404	1.2537
68	32%	0.8225	0.8300	18	82%	1.2537	1.2676
67	33%	0.8300	0.8374	17	83%	1.2676	1.2821
66	34%	0.8374	0.8448	16	84%	1.2821	1.2973
65	35%	0.8448	0.8522	15	85%	1.2973	1.3134
64	36%	0.8522	0.8596	14	86%	1.3134	1.3304
63	37%	0.8596	0.8669	13	87%	1.3304	1.3484
62	38%	0.8669	0.8742	12	88%	1.3484	1.3678
61	39%	0.8742	0.8815	11	89%	1.3678	1.3886
60	40%	0.8815	0.8888	10	90%	1.3886	1.4113
59	41%	0.8888	0.8961	9	91%	1.4113	1.4361
58	42%	0.8961	0.9034	8	92%	1.4361	1.4638
57	43%	0.9034	0.9107	7	93%	1.4638	1.4951
56	44%	0.9107	0.9180	6	94%	1.4951	1.5314
55	45%	0.9180	0.9253	5	95%	1.5314	1.5747
54	46%	0.9253	0.9327	4	96%	1.5747	1.6289
53	47%	0.9327	0.9400	3	97%	1.6289	1.7029
52	48%	0.9400	0.9474	2	98%	1.7029	1.8239
51	49%	0.9474	0.9549	1	99%	1.8239	>1.8239

EXAMPLE CALCULATION

According to the Technical Reference for the ENERGY STAR Score at www.energystar.gov/ENERGYSTARScore, there are five steps to compute a score for multifamily housing. The following is an example:

1 User enters building data into Portfolio Manager

- 12 months of energy use information for all energy types (annual values, entered in monthly meter entries)
- Physical building information (size, location, etc.) and use details describing building activity (hours, etc.)

Energy Data	Value
Electricity	300,000 kWh
Natural gas	218,141 m ³

Property Use Details	Value
Gross floor area (m ²)	9,000
Total Number of Units	180
Total Number of Bedrooms	250
Units – Low Rise	20
Units – Mid Rise	0
Units – High Rise	160
Percent That Can Be Heated	100%
Percent That Can Be Cooled	100%
HDD (provided by Portfolio Manager, based on postal code)	5751
CDD (provided by Portfolio Manager, based on postal code)	117

2 Portfolio Manager computes the actual source EUI

- Total energy consumption for each fuel is converted from billing units into site energy and source energy.
- Source energy values are added across all fuel types.
- Source energy is divided by gross floor area to determine actual source EUI.

Computing Actual Source EUI

Fuel	Billing Units	Site GJ Multiplier	Site GJ	Source Multiplier	Source GJ
Electricity	300,000 kWh	3.600E-03	1,080	1.960	2,117
Natural gas	218,141 m ³	3.843E-02	8,383	1.010	8,467
Total Source Energy (GJ)					10,584
Source EUI (GJ/m²)					1.176

3 Portfolio Manager computes the predicted source EUI

- Using the property use details from Step 1, Portfolio Manager computes each building variable value in the regression equation (determining the density as necessary).
- The centering values are subtracted to compute the centered variable for each operating parameter.
- The centered variables are multiplied by the coefficients from the multifamily housing regression equation to obtain a predicted source EUI.

Computing Predicted Source EUI

Variable	Actual Building Value	Reference Centering Value	Building Centered Variable	Coefficient	Coefficient x Centered Variable
Constant	-	-	-	1.115	1.115
Bedroom Density*	2.300	1.726	0.5740	0.1235	7.089E-02
Unit Density	2.000	1.228	0.7720	0.2363	0.1824
Percent Units in Mid-Rise or High-Rise	0.8900	0.7346	0.1554	0.2138	3.322E-02
Percent Cooled x CDD	117.0	113.1	3.900	0.0002	7.800E-04
Percent Heated x HDD	5751	3804	1947	5.950E-05	0.1158
				Predicted Source EUI (GJ/m²)	1.518

*Bedroom Density has a ceiling at 2.3 beds per 100 m²

4 Portfolio Manager computes the energy efficiency ratio

- The ratio equals the actual source EUI (Step 2) divided by the predicted source EUI (Step 3).
- Ratio = 1.176 / 1.518 = **0.7747**

5 Portfolio Manager uses the efficiency ratio to assign a score via a lookup table

- The ratio from Step 4 is used to identify the score from the lookup table.
- A ratio of 0.7747 is greater than 0.7681 and less than 0.7761.
- **The ENERGY STAR score is 75.**