ENERGY STAR® PortfolioManager® Technical Reference

ENERGY STAR Score for Data Centers in the United States

OVERVIEW

The ENERGY STAR Score for Data Centers applies to spaces specifically designed and equipped to meet the needs of high density computing equipment such as server racks, used for data storage and processing. The objective of the ENERGY STAR score is to provide a fair assessment of the energy performance of a property relative to its peers, taking into account the climate, weather, and business activities at the property. To identify the aspects of building activity that are significant drivers of energy use and then normalize for those factors, a statistical analysis of the peer building population is performed. The result of this analysis is an equation that will predict the energy use of a property, based on its experienced business activities. The energy use prediction for a building is compared to its actual energy use to yield a 1 to 100 percentile ranking of performance, relative to the national population.

- **Property Types.** The ENERGY STAR score for data centers applies to spaces specifically designed and equipped to meet the needs of high density computing equipment such as server racks, used for data storage and processing. The score applies to individual buildings only and is not available for campuses.
- **Reference Data.** The analysis for data centers is based on survey data collected by EPA, in coordination with major industry associations, including Uptime Institute, Green Grid, 7x24 Exchange, and AFCOM,
- **Adjustments for Business Activity.** The analysis includes adjustments for Annual IT Energy.
- **Release Date.** The ENERGY STAR score for data centers was released in June 2010.

This document presents details on the development of the 1 - 100 ENERGY STAR score for data center properties. More information on the overall approach to develop ENERGY STAR scores is covered in our Technical Reference for the ENERGY STAR Score, available at www.energystar.gov/ENERGYSTARScore. The subsequent sections of this document offer specific details on the development of the ENERGY STAR score for data centers:

REFERENCE DATA & FILTERS

For the ENERGY STAR score for data center properties, the reference data used to establish the peer building population in the United States is based on survey data collected by EPA. EPA relies on publicly available external data sets to develop ENERGY STAR scores where feasible, but a sufficiently robust set of data center energy consumption information was not available. In its effort to collect survey data, EPA coordinated with major industry associations, including Uptime Institute, Green Grid, 7x24 Exchange, and AFCOM, to inform their members and encourage participation.

EPA collected data from stand alone data center facilities as well as those enclosed within larger buildings. In addition to collecting energy consumption data, EPA consulted with industry associations and partners to determine the specific operating parameters that were likely to influence energy consumption, and developed a list of operating characteristics that were requested from survey participants.

To analyze the building energy and operating characteristics in this survey data, four types of filters are applied to define the peer group for comparison and to overcome any technical limitations in the data: Building Type Filters, Program Filters, Data Limitation Filters, and Analytical Filters. A complete description of each of these categories is provided in our Technical Reference for the ENERGY STAR Score, at www.energystar.gov/ENERGYSTARScore. ***Figure 1*** presents a summary of each filter applied in the development of the ENERGY STAR score for data centers, the rationale behind the filter, and the resulting number of properties in the data set after the filter is applied. After all filters are applied, the remaining data set has 61 properties.

Figure 1 – Summary of Filters for the ENERGY STAR Score for Data Centers

Condition for Including an Observation in the Analysis	Rationale	Number Remaining
Must have complete data for energy use and operating characteristics	Program Filter - Complete data is necessary for analysis.	120
Must provide IT Energy measured at the output of the UPS Meter	Program Filter – In order to develop an equitable comparison between facilities, all IT Energy consumption must be measured at the same location.	108
Must be a Stand Alone data center	Analytical Filter – Data for Stand Alone data centers was more robust and resulted in higher significance for regression equations. Using stand alone facilities is also more consistent with the process used by EPA for other space types.	61

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 eligible 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 availability, 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 will have different behavior from the rest of the properties (e.g., office buildings smaller than 5,000 ft^2 do not behave the same way as larger buildings), in which case an Analytical Filter will be 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 get a score in Portfolio Manager is available at www.energystar.gov/EligibilityCriteria.

Related to the filters and eligibility criteria described above, another consideration is how Portfolio Manager treats properties that are 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 a campus of buildings. The applicability of the ENERGY STAR score depends on the type of property. For data center properties, the score is based on individual buildings, because the primary function of the data center is contained within a single building and because the properties included in the reference data are single buildings. In cases where multiple data centers are situated together, each individual building can receive its own ENERGY STAR score, but the campus cannot earn a score.

VARIABLES ANALYZED

To normalize for differences in business activity, we perform 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 is analyzed using a weighted ordinary least squares regression, which evaluates energy use relative to business activity (e.g., IT energy). This linear regression yields an equation that is 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 data centers.

Dependent Variable

The dependent variable is what we try to predict with the regression equation. For the data center analysis, the dependent variable is power usage effectiveness, or PUE. This variable is unique to the Data Center space type. Source energy use intensity (source EUI) is used for the majority of EPA's ENERGY STAR scores, but was not considered to be the best choice for data centers. Source EUI can vary widely for data center facilities, and does not take into consideration the varying densities of IT equipment that can be present in these facilities.

EPA consulted with industry associations and data center operators to identify an appropriate metric to evaluate energy use in data center facilities. The dependent variable of PUE is defined as:

$$\text{PUE} = \text{Total Energy} / \text{IT Energy}$$

where both Total Energy and IT Energy are expressed in Source kBtu.

Total Energy includes the annual energy consumption for all fuels at the data center. In many cases, the only energy consumption at data centers is electricity. However, it is important to capture any other fuel use (e.g., chilled water, natural gas), in order to evaluate the total energy performance of the facility. This practice is consistent with all ENERGY STAR scores.

IT Energy is defined as the total amount of energy required by the server racks, storage silos, and other IT equipment in the data center. For the purposes of ENERGY STAR, this should be measured at the output of the Uninterruptible Power Supply (UPS).

EPA considered alternate locations for measuring IT Energy consumption, and requested data from both the UPS and PDU meters from survey participants. Measurements at the PDU meter or closer to the racks can provide a more accurate representation of IT Energy. However, these measurements are still not commonplace in the industry, and were not provided by a large number of survey participants. EPA prefers a common metric that can be used by the majority of data center operators.

By setting PUE as the dependent variable, the regressions analyze the key drivers of PUE – those factors that explain the variation in Power Usage Effectiveness in data centers

Independent Variables

The reference survey collects numerous property operating characteristics that were identified as potentially important for data centers. Based on a review of the available variables in the data, in accordance with the criteria for inclusion in Portfolio Manager[1], the following variables were analyzed:

- Building Square Footage
- Data Center Square Footage
- Tier Level (four levels denoting increasing equipment redundant capacity)
- Number of racks
- UPS Utilization
- Annual IT Energy
- Building Type (Stand alone data center vs. Enclosed in another building)
- Data Center type (options included Hosting, Hybrid, Internet, Traditional, and Telecom)
- HDD
- CDD

We perform extensive review on all of these operational characteristics. In addition to reviewing each characteristic individually, characteristics were reviewed in combination with each other (e.g., IT Energy / Square Foot). Based on analytical results and residual plots, variables are examined using different transformations (such as the natural logarithm, abbreviated as Ln). The analysis consists of multiple regression formulations. These analyses are structured to find the combination of statistically significant operating characteristics that explain the greatest amount of variance in the dependent variable: Power Usage Effectiveness (PUE).

The final regression equation includes the following variable:

- Annual IT energy

This variable is used to compute the predicted PUE for data centers. The predicted PUE is the mean PUE for a hypothetical population of buildings that share the same values for each of these variables. That is, the mean PUE for a building that operates just like your building.

IT Energy Analysis

The Annual IT Energy variable warrants additional discussion, because it is unlike most operating characteristics included in ENERGY STAR scores. Most variables can be entered as a single value that remains relatively constant throughout the year (i.e., square foot, hours of operation). IT Energy, on the other hand, must be metered on a regular basis. Users can enter monthly entries in Portfolio Manager. The total use over a period of one year is calculated and used in the regression equation.

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

The regression analysis shows that facilities with higher IT Energy loads have lower PUE values on average. This phenomenon can be understood as an economy of scale: the larger, more intensive data centers can have more opportunities for efficiencies than their smaller counterparts. Similar behavior has been observed in other types of commercial buildings, where larger buildings may use less energy per square foot. The relationship between PUE and Annual IT Energy was only observed up to a certain Annual IT Energy value, after which the average PUE values level off (i.e. there are no longer economies of scale beyond a certain size). Therefore, the adjustment of IT Energy within the equation is applied over that range, and capped at a maximum adjustment at the value of 0.4 Source TBtu. That is, Data Centers with IT energy values of greater than 0.4 TBtu will receive the same regression adjustment as Data Centers with IT energy equal to 0.4 TBtu.

Testing and Variables not Correlated with PUE

The ENERGY STAR score for data centers includes fewer operating characteristics than most other ENERGY STAR scores. It was determined that the Annual IT Energy is the primary factor influencing PUE, and that other operating characteristics do not show any statistically significant correlation with PUE.

Climate is one characteristic that was examined closely. EPA found no statistically significant relationship between heating and cooling degree days and PUE, which was initially considered to be surprising. However, upon further review, it was determined that the energy required for cooling a data center is dominated by the high internal loads generated by the IT equipment, and that climate has a relatively low contribution to the building cooling load.

EPA also examined Tier level, a measure of redundancy of equipment capacity, and Data Center type, with options that included Hosting, Hybrid, Internet, Traditional, and Telecom. These characteristics were both excluded from the final data center ENERGY STAR score. The dependence of PUE on both variables was not shown to be statistically significant. Additionally, the variables were determined to be hard to define. Operators reported that there could be multiple Tiers or Data Center types within one facility. For Tier, it was also determined that data centers may have unnecessarily high Tier levels, and normalization for Tier could provide a disincentive for efficient design.

REGRESSION EQUATION RESULTS

The final regression is a weighted ordinary least squares regression across the filtered data set of 61 observations. The dependent variable is PUE. Each independent variable is centered relative to the mean value, presented in ***Figure 2***. The final equation is presented in ***Figure 3***. The variable in the regression equation is significant at the 95% confidence level or better, as shown by the significance level.

The regression equation has a coefficient of determination (R^2) value of 0.1138, indicating that this equation explains 11.38% of the variance in PUE for data center buildings. Because the final equation is structured with PUE as the dependent variable, the explanatory power of IT energy is not included in the R^2 value, thus this value appears artificially low. Re-computing the R^2 value in units of source energy[2], demonstrates that the equation actually explains 92.2% of the variation of source energy of data centers. This is an excellent result for a statistically-based energy model.

Detailed information on the ordinary least squares regression approach is available in our Technical Reference for the ENERGY STAR Score, at www.energystar.gov/ENERGYSTARScore.

[2] The R^2 value in Source Energy is calculated as: 1 – (Residual Variation of Y) / (Total Variation of Y). The residual variation is sum of (Actual Source Energy$_i$ – Predicted Source Energy$_i$)^2 across all observations. The Total variation of Y is the sum of (Actual Source Energy$_i$ – Mean Source Energy)^2 across all observations.

Figure 2 - Descriptive Statistics for Variables in Final Regression Equation

Variable	Mean	Minimum	Maximum
PUE	1.924	1.362	3.598
Annual IT Energy (in Source TBtu)	0.2091	0.0129	0.7204

Figure 3 - Final Regression Results

Summary				
Dependent Variable	Power Usage Effectiveness (PUE)			
Number of Observations in Analysis	61			
R^2 value	0.1138			
Adjusted R^2 value	0.0988			
F Statistic	7.579			
Significance (p-level)	0.0078			
	Unstandardized Coefficients	Standard Error	T value	Significance (p-level)
Constant	1.924	0.0523	36.81	0.0000
C_ Annual IT Energy	-0.9506	0.3453	-2.753	0.0078

Notes:
- *The prefix C_ on Annual IT Energy indicates that the value is centered. The centered variable is equal to the difference between the actual value and the observed mean. The observed mean Annual IT Energy is 0.2091 Source TBtu.*
- *Annual IT Energy is computed in Source Energy, and entered in Tera Btu (TBtu):*
  Annual IT Energy kWh(3.412kBtu/kWh)*(3.14 Source/Site Electric)*(1 TBtu/10^9kBtu)*
- *The Annual IT Energy adjustment is capped at 0.4 TBtu*

ENERGY STAR SCORE LOOKUP TABLE

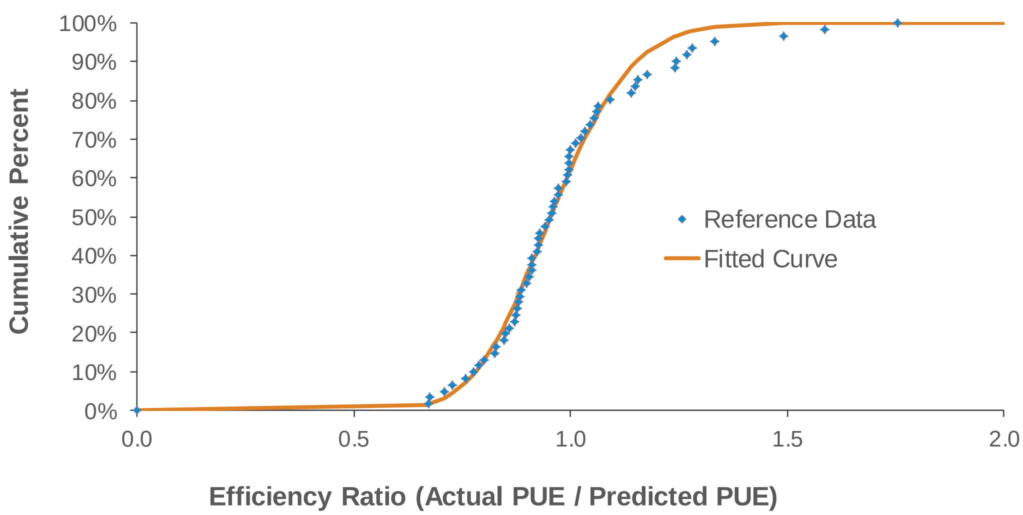
The final regression equation (presented in ***Figure 3***) yields a prediction of PUE based on a building's operating characteristics. Some buildings in the reference data sample use more energy than predicted by the regression equation, while others use less. The *actual* PUE of each reference data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

$$Energy\ Efficiency\ Ratio = \frac{Actual\ PUE}{Predicted\ PUE}$$

A lower efficiency ratio 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 yielded a shape parameter (alpha) of 43.93 and a scale parameter (beta) of 0.0219. For this fit, the sum of the squared error is 0.0555.

Figure 4 – Distribution for Data Center



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 has a ratio this small or smaller. The ratio on the gamma curve at the value of 25% will correspond to the ratio for a score of 75; only 25% of the population has ratios this small or smaller. The complete score lookup table is presented in ***Figure 5***.

Figure 5 – ENERGY STAR Score Lookup Table for Data Center

ENERGY STAR Score	Cumulative Percent	Energy Efficiency Ratio		ENERGY STAR Score	Cumulative Percent	Energy Efficiency Ratio	
		> =	<			>=	<
100	0%	0.0000	0.6569	50	50%	0.9548	0.9584
99	1%	0.6569	0.6879	49	51%	0.9584	0.9620
98	2%	0.6879	0.7082	48	52%	0.9620	0.9657
97	3%	0.7082	0.7236	47	53%	0.9657	0.9694
96	4%	0.7236	0.7364	46	54%	0.9694	0.9731
95	5%	0.7364	0.7474	45	55%	0.9731	0.9768
94	6%	0.7474	0.7571	44	56%	0.9768	0.9805
93	7%	0.7571	0.7659	43	57%	0.9805	0.9843
92	8%	0.7659	0.7739	42	58%	0.9843	0.9880
91	9%	0.7739	0.7814	41	59%	0.9880	0.9919
90	10%	0.7814	0.7883	40	60%	0.9919	0.9957
89	11%	0.7883	0.7949	39	61%	0.9957	0.9996
88	12%	0.7949	0.8011	38	62%	0.9996	1.0035
87	13%	0.8011	0.8071	37	63%	1.0035	1.0075
86	14%	0.8071	0.8127	36	64%	1.0075	1.0115
85	15%	0.8127	0.8182	35	65%	1.0115	1.0156
84	16%	0.8182	0.8235	34	66%	1.0156	1.0198
83	17%	0.8235	0.8285	33	67%	1.0198	1.0240
82	18%	0.8285	0.8335	32	68%	1.0240	1.0282
81	19%	0.8335	0.8383	31	69%	1.0282	1.0326
80	20%	0.8383	0.8429	30	70%	1.0326	1.0370
79	21%	0.8429	0.8475	29	71%	1.0370	1.0415
78	22%	0.8475	0.8520	28	72%	1.0415	1.0461
77	23%	0.8520	0.8563	27	73%	1.0461	1.0508
76	24%	0.8563	0.8606	26	74%	1.0508	1.0556
75	25%	0.8606	0.8648	25	75%	1.0556	1.0605
74	26%	0.8648	0.8689	24	76%	1.0605	1.0656
73	27%	0.8689	0.8730	23	77%	1.0656	1.0708
72	28%	0.8730	0.8770	22	78%	1.0708	1.0761
71	29%	0.8770	0.8810	21	79%	1.0761	1.0816
70	30%	0.8810	0.8849	20	80%	1.0816	1.0873
69	31%	0.8849	0.8888	19	81%	1.0873	1.0932
68	32%	0.8888	0.8926	18	82%	1.0932	1.0994
67	33%	0.8926	0.8964	17	83%	1.0994	1.1058
66	34%	0.8964	0.9002	16	84%	1.1058	1.1125
65	35%	0.9002	0.9039	15	85%	1.1125	1.1195
64	36%	0.9039	0.9076	14	86%	1.1195	1.1269
63	37%	0.9076	0.9113	13	87%	1.1269	1.1348
62	38%	0.9113	0.9150	12	88%	1.1348	1.1432
61	39%	0.9150	0.9186	11	89%	1.1432	1.1521
60	40%	0.9186	0.9223	10	90%	1.1521	1.1619
59	41%	0.9223	0.9259	9	91%	1.1619	1.1725
58	42%	0.9259	0.9295	8	92%	1.1725	1.1842
57	43%	0.9295	0.9331	7	93%	1.1842	1.1974
56	44%	0.9331	0.9367	6	94%	1.1974	1.2126
55	45%	0.9367	0.9403	5	95%	1.2126	1.2306
54	46%	0.9403	0.9439	4	96%	1.2306	1.2530
53	47%	0.9439	0.9475	3	97%	1.2530	1.2831
52	48%	0.9475	0.9512	2	98%	1.2831	1.3315
51	49%	0.9512	0.9548	1	99%	1.3315	>1.3315

EXAMPLE CALCULATION

As detailed in our Technical Reference for the ENERGY STAR Score, at www.energystar.gov/ENERGYSTARScore, there are five steps to compute a score. The following is a specific example for the score for data center:

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 (IT energy)

Energy Data	Value
Electricity	15,000,000 kWh
Natural gas	20,000 therms

Property Use Details	Value
Annual IT Energy (kWh)	8,500,000

2 Portfolio Manager computes the actual power usage effectiveness

- 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

Computing Actual Source Energy

Fuel	Billing Units	Site kBtu Multiplier	Site kBtu	Source kBtu Multiplier	Source kBtu
Electricity	15,000,000 kWh	3.412	51,180,000	3.14	160,705,200
Natural gas	20,000 therms	100	2,000,000	1.05	2,100,000
			Total Source Energy (kBtu)		162,805,200

- Portfolio Manager must also convert Annual IT energy (kWh) into Site kBtu, and then from Site kBtu to Source kBtu.

Computing IT Source Energy

Fuel	Billing Units	Site kBtu Multiplier	Site kBtu	Source kBtu Multiplier	Source kBtu
Electricity	8,500,000 kWh	3.412	29,002,000	3.14	91,066,280

- Total source energy is then divided by Annual IT source energy to determine actual PUE
  - PUE = 162,805,200 kBtu / 91,066,280 kBtu = 1.788

3 Portfolio Manager computes the predicted power usage effectiveness

- Using the property use details from Step 1, Portfolio Manager computes the building variables needed for the regression equation
    - For data centers, it is necessary to convert the IT Energy to TBtu
    - 91,066,280 kBtu / (10^9 TBtu/kBtu) = 0.0911
- The centering values are subtracted to compute the centered variable for each operating parameter.
- The centered variables are multiplied by the coefficients from the regression equation to obtain a predicted PUE.

Computing Predicted PUE

Variable	Actual Building Value	Reference Centering Value	Building Centered Variable	Coefficient	Coefficient * Centered Variable
Constant	--	--	--	1.924	1.924
Annual IT Energy	0.0911	0.2091	-0.1180	-0.9506	0.1122
				Predicted PUE	**2.036**

4 Portfolio Manager computes the energy efficiency ratio

- The ratio equals the actual PUE (Step 2) divided by predicted PUE (Step 3)
- Ratio = 1.788 / 2.036 = 0.8781

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 0.8781 is greater than or equal to 0.8770 and less than 0.8810.
- ***The ENERGY STAR score is 71.***