You will need to read the attached article before starting this week's graded discussion.

Tsanas, A., & Xifara, A. (2012). Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings*, *49(C)*, 560-567.

[M1\_CSE575\_AccurateQuantEstEnergyPerfUsingSML.pdf](https://asu.instructure.com/courses/45793/files/11859574/download?wrap=1)

[Actions](https://asu.instructure.com/courses/45793/discussion_topics/951427?module_item_id=2710145)

Meet the requirements of this graded assignment by completing two steps:

* Write and post a 100- to 150-word response to the following question: **Based on the methods used, the results achieved, and the conclusions drawn in this use case, what question or questions does this journal article answer or raise for you? Why?**
* Read through other students' responses and reply to at least two posts to explore the factors your peers considered and the reasoning they employed. Reply posts must be at least 50 words.

(133)

The primary purpose of this article was to determine the optimal values of variables affecting energy consumption (relative compactness, surface area, wall area, roof area, overall height, orientation, glazing area, and glazing area distribution) with the use of several data-driven statistical & machine learning techniques such iteratively reweighted last squares and classification of random forests in order to empirically measure heating and cooling load.

The results of the paper demonstrate that these methodologies are effective but display correlation rather than causation for energy efficiency. Relative Compactness, wall area and roof area appear as the most correlated with heating & cooling loads, with surface area being inversely proportional to relative compactness. The authors also strongly caution against using methods that rely on the normality of the data, as the data in this study doesn’t fit the normal distribution.

Venkat, it is important to note that the results in the study showed correlation between the 8 variables to heating load and cooling load (some more than others, particularly relative Compactness, wall area and roof area), but not causation.

The authors of the paper also had a disclaimer towards the end towards researchers to be wary of rushing into popular statistical methods and rigorously evaluate techniques based on the inherent properties of their own data, as their own data did not fit a normal distribution.

The authors did state, however, that this was a good general framework to include input variables that yield results averse to normal distribution and to obtain immediate, respectably accurate estimates of HL & CL

Parthav, it is important to note that the researchers were measuring the inherent structural properties of building design to determine how, through the usage of statistical & machine learning methodologies, they would affect heating load and cooling load. It doesn’t make sense to take temperature as an input variable in their study, as they are measuring static factors inherent to the design of buildings rather than dynamic factors such as temperature, which is contingent on the environment

Temperature is more fitting, categorically speaking, as a result of analyzing the relation presented by the 8 input variables, although that is not accurate, as Heat Load is measured by kilowatts or British thermal units