

Abstract #: 442

Author Name: Theodore Love

Author Company: Green Energy Economics Group, Inc.

Second Author's Name:

Abstract Title: Using Open Data to Predict Energy Usage: What Tax Lot Data can tell us about Energy Usage Intensity in New York City

Abstract Text:

New York City (NYC) is on the forefront of a movement towards levels of transparency never seen before through its efforts at providing open data. Through the city's Socrata platform, NYC has over 13,000 data sets that provide countless opportunities for using modern analysis techniques to gain new insights. A large portion of New York City's data is focused on buildings, and figuring out ways to use this building data is an important task step in understanding how NYC uses energy. NYC's Local Law 84 requires large buildings to report energy usage intensity (EUI) on an annual basis, along with information on property type, water usage, and greenhouse gas emissions. Data for 2011 and 2012 has been made public through NYC's open data initiative, and more than 14,000 buildings provided information for 2012. NYC also has extensive public data for each tax lot in the city stored in the Primary Land Use Tax Lot Output database (PLUTO), including data on zoning, building classification, assessment, and square footage estimates by usage type. This project uses supervised machine learning techniques to see how much of the variance in EUI can be predicted using the separate land use data from PLUTO. It also discusses the implications from what the model is unable to explain. Finally, it explores how the model could be used for policy and planning purposes, and how the rise in "open data" can provide new opportunities for addressing climate change.