Analysis of Smart Meter Data With Machine Learning for Implications Targeted Towards Residents

Ali Emrouznejad, Surrey Business Shcool, UK

(iD) https://orcid.org/0000-0001-8094-4244

Vishal Panchmatia, Aston University, UK

Roya Gholami, University of Illinois System, France*

Carolee Rigsbee, University of Illinois, Springfield, USA

(iD) https://orcid.org/0000-0001-5948-2401

Hasan B. Kartal, University of Illinois, Springfield, USA

ABSTRACT

Previous studies examining the electricity consumption behavior using traditional research methods, before the smart-meter era, mostly worked on fewer variables, and the practical implications of the findings were predominantly tailored towards suppliers and businesses rather than residents. This study first provides an overview of prior research findings on electric energy use patterns and their predictors in the pre and post smart-meter era, honing in on machine learning techniques for the latter. It then addresses identified gaps in the literature by: 1) analyzing a highly detailed dataset containing a variety of variables on the physical, demographic, and socioeconomic characteristics of households using unsupervised machine learning algorithms, including feature selection and cluster analysis; and 2) examining the environmental attitude of high consumption and low consumption clusters to generate practical implications for residents.

KEYWORDS

attitude, Cluster analysis, consumption classification, demographic, households, physical, resident, smart meter data, socioeconomic

INTRODUCTION

Planning for energy availability is critical in this era of rapid global economic and technological growth and climate change. Energy use is predicted to increase by approximately 15% and 70% by the year 2050 for OECD and non-OECD countries respectively (U.S. EIA, 2019). Researchers estimate the cost to construct infrastructure to support peak load increases of 18% in the U.S. would

*Corresponding Author

be approximately 180 billion dollars, assuming current infrastructure pricing (Auffhammer, Baylis, & Hausman, 2017). Nevertheless, the main issue is not even whether we will have enough fossil fuel, but whether we can afford to produce and burn the fossil fuel we have (Alexander, 2015). Experts warn the World will pass the crucial 2'C global warming limit and the carbon pledges from 147 nations to the Paris Climate Summit are not enough to stop such a temperature rise. As such, finding ways to not only plan for but also reduce future energy consumption are key.

While the industrial and transportation sectors will continue to consume the most energy globally, residential consumption also drives increased energy demand (U.S. EIA, 2019). Attitudinal and behavioral changes that reduce energy consumption could create needed 'quick wins' while longer-term systematic infrastructure upgrades are designed and implemented to address anticipated increased energy loads.

At the same time, data collection tools available for understanding current energy utilization have advanced. The use of smart meters enables the real-time data collection and more accurate billing, and facilitates unique insight on energy and water consumption behavior (Bhattarai et al., 2022; Ratner et al. 2022; Bhardwaj et al. 2022; Gohil et al. 2021; Dhanwani et al. 2021; Monks et al. 2021; Kumar et al. 2021; Ahuja and Khosla, 2019; Cango et al., 2018; Chaudhari and Mulay, 2021; Khosrojerdi et al., 2021; Kuhe and Bisu, 2020; Sharma and Singh 2015; Siepermann et al., 2021; Rathod and Garg, 2017; Agwu 2021; Baccarne & De Marez, 2021; Croci & Molteni 2021).

While Canada and the U.S. led the way in smart meter installation (30-40% current usage across North America), Europe is estimated to have around 72% installation in 2020 (IoT Analytics, 2020). This has led to the availability of large volumes of smart-meter generated data. For this reason, there is value in complementing traditional theory-driven research in energy consumption with research that uses machine learning techniques, or more broadly by artificial intelligence (AI)-driven approaches. Prior to the smart-meter era, studies tended to examine electricity consumption behavior using traditional theory-driven research methods, incorporating a limited number variables. The practical implications of the findings were predominantly tailored towards suppliers and businesses rather than residents.

To address these gaps in literature, the current study: 1) provides an overview of the findings from the studies on electric energy use patterns and their predictors in the pre and post smart-meter era; 2) combines unsupervised machine learning algorithms, feature selection, and cluster analysis for uncovering new, unexpected, or more nuanced insight into the determinants of energy usage, and patterns in its use, that might not otherwise have been predicted using conventional predictive research methods; and 3) analyzses the implications of smart meter data adoption for residential consumers by focusing on socio-economic characteristics, demographic characteristics, and pro-environmental behavior.

LITERATURE REVIEW

Pre-Smart Meter Era: Determinants of Energy Use

Prior to the advent of smart meters, household energy consumption research largely relied on data obtained through household member interviews and surveys. and secondary consumption details such as whether certain home appliances were in use or not in use (Klemenjak et al., 2019), examining the effects of "socioeconomic", "demographic", and "physical" determinants (Petersen and Gram-Hanssen, 2005). Some studies also examined attitudinal factors.

Socioeconomic Characteristics

Jones, Fuertes, and Lomas (2015) examined the relationship between *household income* and *electricity consumption* and found that higher household income is correlated with higher energy consumption. Cayla, Maizi, and Marchand (2011) examined the impact of income on residential energy consumption and found households with lower income were more restricted in their energy consumption.

Demographic Characteristics

Gram-Hanssen et al. (2004) identified the number of occupants in a household as the most significant predictor of electricity consumption, also finding 'number of occupants' along with 'household income' and 'size of the apartment' collectively explained most of the variance in electricity consumption related to appliances and lighting use. Zhou and Teng's (2013) research supported the importance of household occupancy as a predictor beyond Eurpean and US households, finding household electricity consumption in China increases by 8% for each additional occupant. Tiwari (2000) reported similar findings for households in India.

Physical Characteristics

As for physical characteristics, Leahy and Lyons (2010) used logistic and linear regression analysis to examine the relation between the size of the building (measured by the number of rooms) and energy consumption in Irish households. They found that as the number of rooms increases so does the consumption of electricity and the likelihood of owning appliances such as a washing machine, a tumble dryer, and a dishwasher. Gram-Hanssen et al. (2004) found total floor space ranked third in the overall explanatory power of energy consumption. Many regression-based studies of floor space and energy consumption imply this relation is due to larger houses requiring more energy to regulate temperature (Zhou and Teng, 2013; Parker, 2003).

Attitude

The Relationship Between Environmental Attitude and Energy Consumption Behavior

Previous research has also examined the relationship between consumer attitude towards the environment and energy consumption behavior, which is justified by increased global attention to environmental conservation. Martinsson, Lundqvist, and Sundstrom (2011) found environmental attitude impacts the behavior of energy consumption, particularly in high-income households.. These findings are supported by Hansla, Gamble, Juliusson, and Gärling (2008), who examined factors that affect consumer willingness to pay for green electricity. Regression analysis showed a high correlation between pro-environment attitude and willingness to adopt green electricity practices.

Factors Impacting Environmental Attitude

Oltra et al. (2013), using data collected in focus groups, explored the interaction between occupants' attitude and feedback on electricity consumption. They found resident motivation and pre-existing attitudes towards conserving energy are pivotal in saving energy. Previous research also suggests that offering rewards can change residents' attitudes towards energy consumption. Although these studies provide evidence of a strong correlation between psychological motivations and energy savings, they do not clarify whether an actual change in attitude causes energy conservation or whether energy conservation is prompted by the incentive to save money. This is supported by Mills and Schleisch (2012), who found that households with children were more likely to adopt energy-conservation practices.

Reframing Determinants: "Behavioral", "Lifestyle", and "Practice-based" Approaches.

Gram-Hanssen's (2013) in-depth review of the literature led to suggesting the study of energy consumption behavior could be more broadly viewed from three different approaches. A *behavioral approach* would explore demographic and psychological factors. A *lifestyle approach* would explore socioeconomic factors guided by consumer interests in material objects. A *practice-based approach* would place emphasis on understanding the the actual routines carried out by a group of occupants. As an example of the latter, beyond finding energy usage is related to the characteristics of the building, Fabi, Andersen, Corgnati, and Olesen (2012) observed that occupants would open and close windows to save energy.

Determinants of Energy Use Using Smart Meter Data: Cluster Analysis

Since the introduction of smart meters, researchers have attempted to use various data mining and unsupervised machine learning techniques to analyze the data generated to exploit the potential that comes from not assuming what to look for beforehand. Findings and implications from studies using smart meter data and which involve analysis using unsupervised machine learning techniques are discussed in this section. Included is a review of literature on consumption patterns and the determinants of energy consumption behavior (such as the demographic and socioeconomic factors) and their corresponding behavioral characteristics.

Consumption Patterns, Usage Behavior, and Implications

Cluster analysis of meter data has provided useful insight into peak usage periods, which in turn is useful for planning electricity generation and distribution, and rate adjustments. Such analysis has revealed seasonal and daily/weekend use patterns in various geographical regions.

Flath, Nicolay, Conte, and van Dinther (2012), using the K-means and neural network algorithms for cluster analysis, identified relative load profiles for 9 clusters by examining consumption by day and by week for each season. They found residential electricity consumption is most prevalent throughout winter. The analysis revealed that more clusters are generated for weekdays during this season, indicating less variation in consumption patterns during the weekend.

Amri et al. (2017), using a K-means clustering algorithm, examined 370 households across four seasons and found households in all clusters consume more electricity in the summer months and the majority consume the least in winter which contrasted with the findings of Flath et al. (2012). The study specifically aimed at understanding seasonal patterns for appropriate adjustments to be made by electricity suppliers to ensure the necessary amount of electricity is supplied.

Implications of Socioeconomic, Demographic, and Physical
Determinants of Energy Consumption Behavior

Beckel et al. (2014), using data from 4232 Irish households, examined whether it is possible to gain insight on socioeconomic and demographic information from smart meter data to improve efficiency for utilities. They developed a model using unsupervised machine learning techniques. Gouveia et al. (2015), using additional energy consumption data of 230 households collected over three years, extended Becket et al.'s (2014) research. They found the most significant determinants of electricity consumption were the physical characteristics of the building and socioeconomic and demographic features of the household.

Kavousian, Rajagopal, and Fischer (2013) investigated the determinants of residential electricity consumption using data collected from 1628 households. They identified a detailed set of explanatory variables and split their analysis into maximum and minimum daily consumptions to identify which variables are significant for different levels of energy consumption (high and low peaks). They found geographic (weather and location) and physical characteristics (house size) were the most significant determinants of electricity consumption in general. The implications of these studies were also targeted at amending policies for improving energy efficiency in high-usage appliances.

Overall, the studies discussed above had implications mostly targeted towards energy providers who could use the cluster analysis results to inform policy revisions and improvements or increase energy distribution efficiency.

RESEARCH METHODS

In this section, we describe the dataset and how the data was cleaned and prepared for the cluster analysis. We also explain the modeling techniques used to determine the most appropriate variables to incorporate into the cluster analysis. We then provide descriptive insights into the final variables chosen to represent physical, demographic, socioeconomic characteristics, and environmental attitude.

Data

The dataset used in this paper was provided by the Commission for Energy Regulation (CER) (2012), collected in 2010 during smart metering trials for over 5000 Irish homes and businesses. Pre-trial survey data contains the occupants' attitudes towards energy consumption and conservation.

Our study focuses on residential electricity consumption data only, collated from 2506 homes, and investigates consumer energy consumption behaviors using detailed individual household physical, demographic, and socio-economic characteristics and unsupervised machine learning techniques. We use both electricity consumption data and pre-trial survey data in our analysis, and follow the CRISP-DM (Cross Industry Standard Process for Data Mining) framework during data preparation and analysis. We describe the data cleaning and attribute selection processes below.

Data Preparation and Attribute Selection

The dataset, stored in Excel and containsing 124 variables, is stored in Excel and was cleaned, and visualized with SPSS Modeler and Tableau. Variables with a high degree of missing values were removed. Also, semantically similar attributes or variables high correlated with each other were dropped or converged.For example, the variables 'property build year' and 'property age' are similar. Because both variables contain similar information, the latter has a smaller number of missing values, we removed 'property age,' and only kept 'property build year'. In the next step, we recoded the categorical variables of two variables, for example, converting 'yes' and 'no' into the binary values, 0 and 1. Then, we created dummy variables for the other categorical variables with nominal values.

Modeling: Determining Best Variables Before Cluster Analysis

At this stage, we used linear regression with numerical variables and feature selection with categorical variables to determine which of the 70 variables best predicted total consumption.

Part 1 – Linear Regression of Numerical Variables

We ran a linear regression node in SPSS Modeler for all the numerical variables from the remaining 70 variables. The predictor importance of this linear regression suggests 'No. of desktop computers' was the most predictive numerical variable of total consumption (Figure 1). Detailed regression output provided insight regarding the significance of each variable (Table 1).

The next step was to remove the statistically non-significant variables and run the linear regression again. These results suggested that 'No. of washing machine loads/day' was now the most important predictor (*Figure 2*). We removed the insignificant variables and ran the regression again. We ran the linear regression for a third time, this time removing the two insignificant variables.
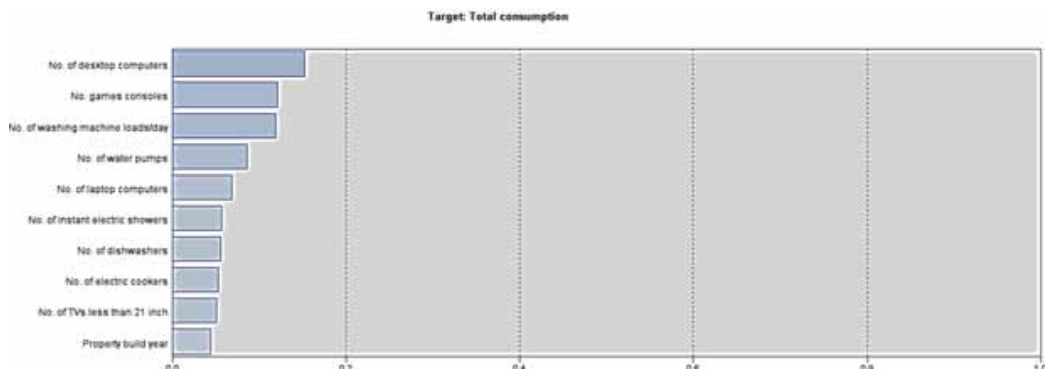
Figure 1.
Top 10 variables via predictor importance

Table 1.
Co-efficient output displaying the significance of each variable

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	328,894	253,861		1,296	0,195
Property build year	-0,130	0,127	-0,032	-1,019	0,309
Property size square feet.	0,001	0,002	0,019	0,590	0,555
No. of dishwashers	74,131	14,622	0,163	5,103	0,000
No. of instant electric showers	-10,229	11,952	-0,028	-0,856	0,392
No. of pumped electric showers	-6,848	10,965	-0,020	-0,625	0,532
No. of electric cookers	26,643	12,857	0,065	2,072	0,039
No. of electric heaters	-1,567	11,008	-0,005	-0,142	0,887
No. of standalone freezers	38,894	10,230	0,120	3,802	0,000
No. of water pumps	14,552	11,515	0,040	1,264	0,207
No. of Tv's less than 21 inches	17,909	10,733	0,054	1,669	0,096
No. of Tv's greater than 21 inches	23,935	19,217	0,041	1,245	0,213
No. of desktop computers	62,486	10,682	0,192	5,850	0,000
No. of laptop computers	24,815	10,941	0,074	2,214	0,024
No. games consoles	39,900	11,271	0,123	3,540	0,000
No. of washing machine loads/day	39,410	8,271	0,183	4,765	0,000
No. of tumble dryer loads/day	5,794	11,596	0,018	0,500	0,132

Figure 2.
Predictor importance output from second linear regression
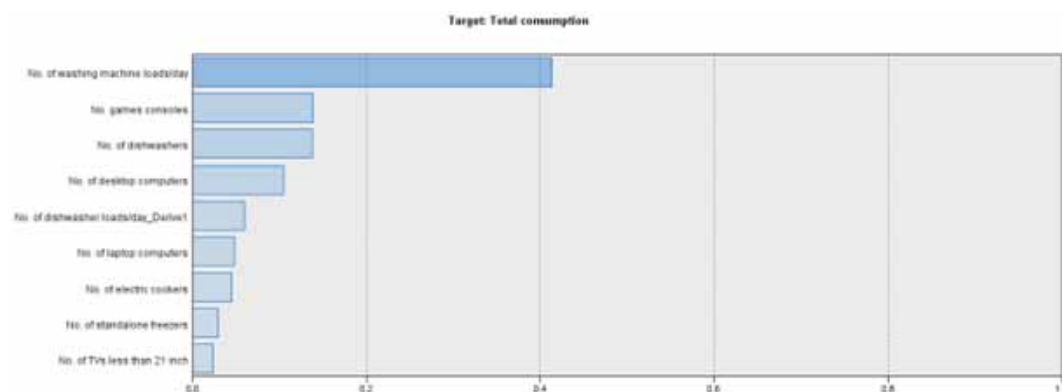
It might be first interesting to see that property size is not significant in the model. However, becasue many variable directly representing the number and frequency of usage for various electrical home appliances in households are included in the datasel and in our model, 'property size' may no longer be a significant indicator of the consumption because the most of the differences in electricity consumption can be captured by the other variables. For example, a household with a smaller house

but having more number of electrical home appliances and using them more frequently can be expected to consume more electricity than a household with a larger house but having fewer electrical appliances and using them less.

Figure 3 highlights the predictor importance ('No. of washing machine loads/day') and the regression output highlighting the significance of each variable. As indicated in *Table 2*, we achieved the most parsimonious model and therefore used the remaining numerical variables in our cluster analysis.

Part 2 – Feature Selection of Categorical Variables

This section describes how a "feature selection node" was used to determine which categorical variables most predicted average total consumption. The feature selection algorithm does not employ the same method as a regression/correlation but rather identifies the variables most important to the dependent variable for any future modeling. It first removes inputs/records with too many missing values or with very small variation to be deemed useful. It then ranks the remaining inputs based on importance using a likelihood ratio of 0-1 (with 1 being the most important). With the default threshold to define importance set at '0.9 +', the results suggest the variables 'Highest Day' and 'Lowest Day' were unimportant. However, 'Highest Day'

Figure 3.
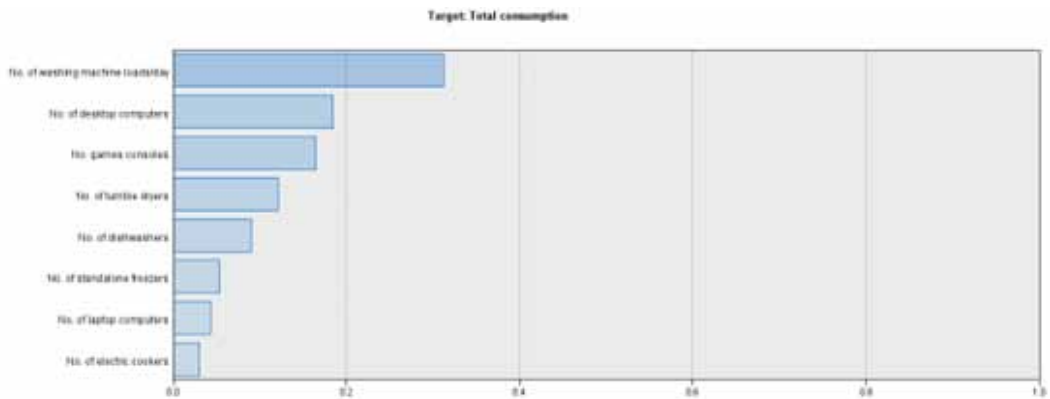**Predictor importance from the parsimonious linear regression model**

Table 2.
The most parsimonious linear regression predictor output

Model	Standardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	72,623	8,723		8,325	0,000
No. of tumble dryers	42,473	6,163	0,124	6,892	0,000
No. of dishwashers	55,466	6,143	0,165	9,029	0,000
No. of electric cookers	28,075	6,285	0,076	4,467	0,000
No. of Standalone freezers	35,377	5,351	0,113	6,612	0,000
No. of desktop computers	49,810	5,540	0,160	8,990	0,000
No. of laptop computers	15,361	5,574	0,049	2,756	0,006
No. of games consoles	43,555	6,250	0,132	6,969	0,000
No. of washing machine loads/day	47,034	3,182	0,225	12,339	0,000

had an importance value of 0.837, whereas 'Lowest Day' had an importance value of 0.668. Therefore, we decided to keep only 'Highest Day' in the variable-set. Additionally, both variables regarding 'Heating Space' were automatically discarded due to insignificant variation within the data (*Figure 4*).

The final variable set is presented in Table 3. It includes a total of 26 demographic, socioeconomic, physical characteristics, and consumption related variables (as well as the variables capturing environmental attitude, which will be cleaned accordingly, later). The variables representing usage patterns were kept despite being deemed unimportant by the feature selection. This is because the feature selection decides how important it is for modeling and not for descriptive analysis.

Figure 5 shows the histogram of the dependent variable, which is the average total consumption. 275 kWh is the most common average for electricity consumption among residents and the variable follows a normal distribution.
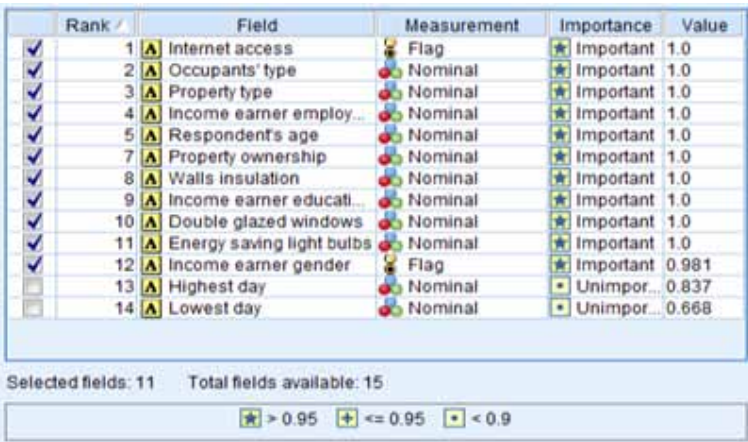
Cluster Analysis

Partitioning-based clustering and neural networks in general work better for large datasets and a large number of variables in comparison to hierarchical clustering techniques which are more appropriate for applications such as document retrieval (Äyrämö and Kärkkäinen, 2006). We used the k-means algorithm, one of the most applied partitioning-based clustering algorithms. Because we have variables with three distinct characteristics (physical, demographic, and socioeconomic), we created different cluster analysis models to better interpret the results for each characteristic (all include the 'total consumption' variable).

To determine the right number of clusters in each model, we ran each model several times with a varying number of clusters and choose the models with the highest silhouette values for high-quality clusters. The clustering models were run for first physical characteristics, then, demographic characteristics, and finally, socioeconomic characteristics. For all three-clustering analysis, the models with five clusters have the highest silhouette values greater than 0.5.

Variables Measuring Environmental Attitude

After the cluster analysis with physical, demographic, and socioeconomic characteristics, we investigated the highest and lowest electricity consumption across all clustering models and the environmental attitude of the clusters. *Table 4* displays the variables where environmental attitude was measured by 5-point Likert-scale (1: Strongly Agree, 2, 3, 4, 5: Strongly Disagree). The survey data from before the smart meter trial capture pre-trial attitude towards energy, electricity consumption, and electricity bill as well as pre-existing willingness to reduce energy consumption.

Figure 4.
A screenshot of the feature-selection output

Table 3.
Final cleaned demographic, socioeconomic, and physical characteristics of residents

Variable Name	Data Type	No. of Missing	Characteristic
Total Consumption	Numerical	0	Target Field
Highest Consumption	Numerical	0	Target Field
Lowest Consumption	Numerical	0	Target Field
Highest Day	Categorical	44	Target Field
Lowest Day	Categorical	44	Target Field
Property Type	Categorical	0	Physical
Property Ownership	Categorical	0	Physical
Energy Saving Light Bulbs	Categorical	0	Physical
Double Glazing Windows	Categorical	0	Physical
Wall insulation	Categorical	0	Physical
Income Earner Gender	Flag / Binary	0	Socioeconomic
Respondent Age	Categorical	0	Demographical
Income Earner Employment Type	Categorical	0	Socioeconomic
Occupants' type	Categorical	1	Demographical
15 years old and over	Categorical	9	Demographical
Under 15 years old	Categorical	28	Demographical
Income Earner Education Level	Categorical	0	Socioeconomic
Internet Access	Flag / Binary	0	Physical
No. of Tumble Dryers	Numerical	5	Physical
No. of Dishwashers	Numerical	5	Physical
No. of Electric Cookers	Numerical	5	Physical
No. of Standalone Freezers	Numerical	5	Physical
No. of Desktop Computers	Numerical	12	Physical
No. of Laptop Computers	Numerical	12	Physical
No. of Game Consoles	Numerical	12	Physical
No. of Washing Machine Loads/day	Numerical	49	Physical

RESULTS

In this section, we provide insight into consumption patterns by focusing on the 'highest' and 'lowest' electricity consumption and the characteristics in each cluster. Then we present the findings of our regression analysis for determining the effect of 'environmental attitude' on 'total consumption' in each cluster.

Overall Consumption Patterns

The electricity consumption dataset contains the following variables that represent usage patterns: 'Highest Day' (of electricity consumption), 'Lowest Day' (of electricity consumption), 'Highest Consumption' (in KW/h), and 'Lowest Consumption' (in KW/h). The histogram in *Figure 6* shows the frequency distribution of 'Highest Consumption' for households. The average level of 'highest consumption' across households was 54.077 kW/h, and the highest reading recorded was 250.704 kW/h.

Figure 5.
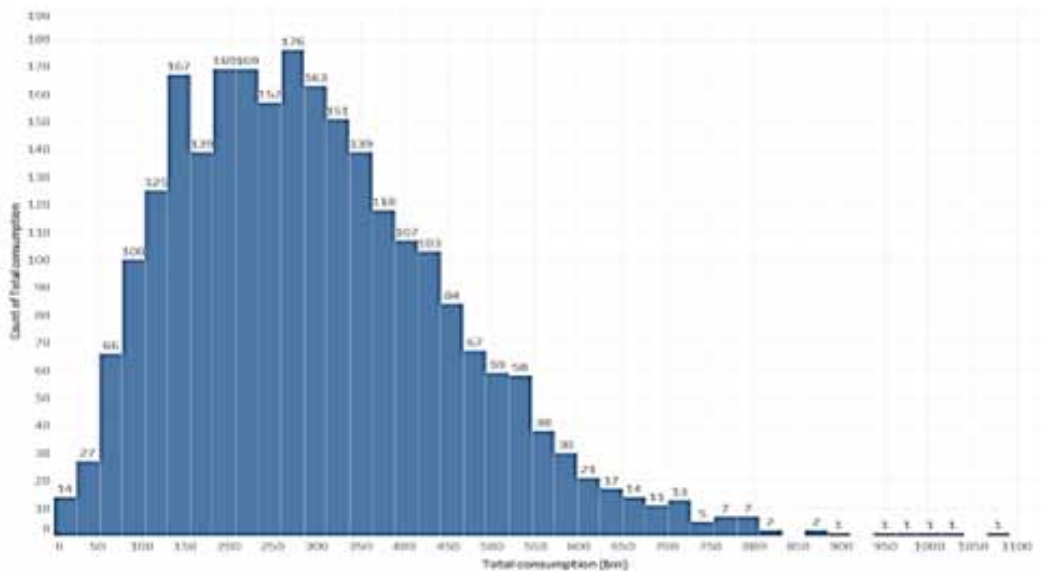**Histogram displaying the frequency distribution of average total consumption**

Table 4.
Variables representing environmental attitude used in the analysis

Variable Name	Original Statement from the Questionnaire
Change energy usage if it reduces bills	I/we am/are interested in changing the way I/we use electricity if it reduces the bill
Change energy usage if it helps the environment	I/we am/are interested in changing the way I/we use electricity if it helps the environment
Can change energy usage if it reduces the bills	I/we can reduce my electricity bill by changing the way the people I/we live with use electricity
Want to reduce energy usage	I/we would like to do more to reduce electricity usage
Know how to reduce energy usage	I/we know what I/we need to do to reduce electricity usage
Other occupants do not want to reduce energy usage	I am not able to get the people I live with to reduce their electricity usage
Not enough time to reduce energy	I do not have enough time to reduce my electricity usage
Not want to be instructed about energy usage	I do not want to be told how much electricity I can use

Cluster Analysis with Physical Characteristics

The cluster analysis in *Figure 7* is done based on the physical characteristics of the household. The most important predictors in this cluster analysis were 'No. of game consoles', 'No. of dishwashers', and 'No. of standalone freezers'. Cluster 2 (in red) accounts for 18.1% of the dataset and represents a high consumption cluster with the highest electricity consumption (mean average of 395.17 KW/h), and Cluster 1 (in light blue) accounts for 21% of the dataset represents a low consumption cluster with the lowest electricity consumption (mean average 191.14).

We applied linear regression to the high consumption Cluster-2 to determine which environmental attitude variables were significantly correlated with the 'total consumption' variable. Table 5 shows the variable 'Not enough time to reduce energy' is the most significant, followed by 'Change energy usage if it helps the environment.' However, the latter was significant at the 10% significance level.

Figure 6.
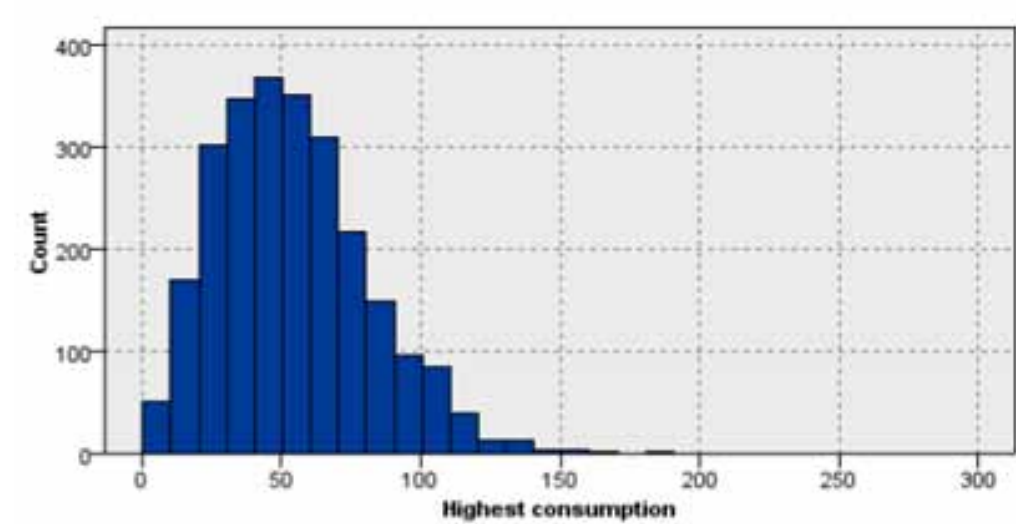**Histogram of the 'Highest Consumption'**

Figure 7.
K-means cluster analysis representing physical characteristics of the household and electricity consumption
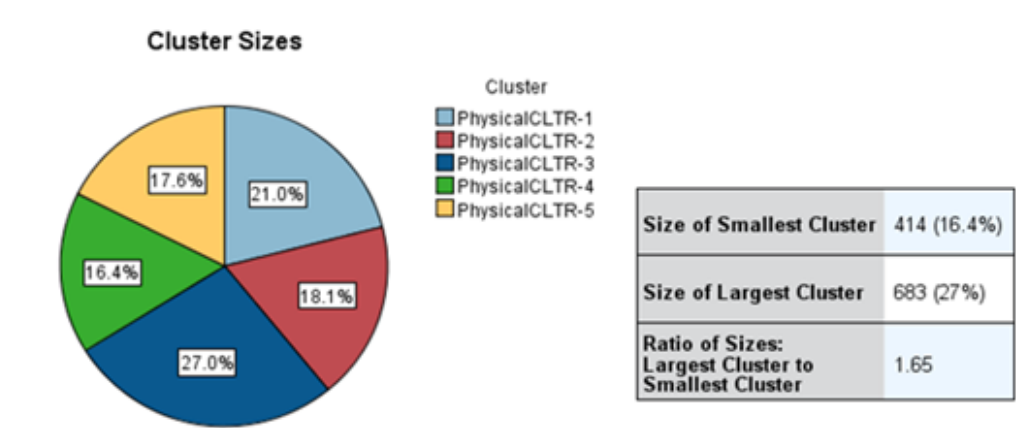

Table 6 displays the results of the linear regression for the low consumption cluster-1 suggesting the variable 'Other occupants do not want to reduce energy usage' narrowly misses the 5% significance level. However, since it is significant at the 10% significance level, we can interpret the results with caution. *Table 7* summarizes the physical characteristics of the Low and High Consumption Clusters along with significant environmental attitudes determined in the regression analyses in *Tables 5* and *6*.

Cluster Analysis with Demographic Characteristics

The cluster analysis in *Figure 8* represents the demographic characteristics of the households. Cluster 2 (in red) shows the highest consumption (mean average of 434.55 KW/h) and accounts for 10.8% of the dataset whereas Cluster 1 (in light blue) shows the lowest electricity consumption (mean average 171.0KW/h) and accounts for 19.3% of the dataset. The most important predictors in this cluster analysis were 'Occupant Type,' 'No. of Over 15s' and 'No. of Under 15s'.

Table 5.
Linear regression on environmental attitude variables and total consumption

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	481,897	37,411	-	12,881	0,000
Change energy usage if it reduces bills	-3,604	10,301	-0,019	-0,350	0,727
Change energy usage if it helps environment	17,366	9,796	0,093	1,773	0,077
Changing energy usage behavior reduces bills	-7,873	7,117	-0,057	-1,106	0,269
Want to reduce energy usage	-5,681	8,721	-0,034	-0,651	0,515
Know how to reduce energy usage	-2,209	6,709	-0,016	-0,329	0,142
Other occupants do not want to reduce energy usage	-5,483	5,064	-0,054	-1,083	0,280
Not enough time to reduce energy	-17,151	5,961	-0,143	-2,871	0,004
Not want to be instructed about energy usage	0,510	5,818	0,004	0,088	0,930

a. Dependent Variable: Total consumption

Table 6.
Linear regression indicating the most significant variables

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	242,961	22,259		10,915	0,000
Change energy usage if it reduces bills	-1,731	5,148	-0,018	-0,336	0,737
Change energy usage if it helps environment	-0,453	5,785	-0,004	-0,078	0,938
Changing energy usage behavior reduces bills	-3,335	3,884	-0,043	-0,858	0,391
Want to reduce energy usage	-1,945	4,442	-0,022	-0,438	0,662
Know how to reduce energy usage	1,904	3,894	0,022	0,489	0,625
Other occupants do not want to reduce energy usage	-5,741	3,147	-0,385	-1,824	0,069
Not enough time to reduce energy	-4,713	3,894	-0,056	-1,210	0,227
Not want to be instructed about energy usage	0,026	3,443	0,000	0,007	0,994

a. Dependent Variable Total consumption

Table 8 displays the results of the linear regression and suggests the variable 'Not enough time to reduce energy' very narrowly misses the 5% significance level by 0.4%. However, since it is significant at the 10% significance level, we can interpret the results with caution.

Table 9 displays the results of the linear regression for Cluster 1. It suggests that the variables 'Change energy usage if it helps the environment' and 'Not want to be instructed about energy usage' are significantly correlated with 'Total Consumption.'

Table 10 summarizes the demographic characteristics of the Low and High Consumption Clusters along with significant environmental attitudes determined in the regression analyses in *Table 8* and *Table 9*.

Cluster Analysis with Socio-Economic Characteristics

Figure 9 presents the results of cluster analysis with the socio-economic characteristics of the households. Cluster 3 (dark blue) represents the highest consumption (mean average of 341.78 KW/h)

Table 7.
Low and high consumption clusters of the analysis with physical characteristics

Attribute Types	Cluster 2	Cluster 1
	High Consumption	**Low Consumption**
Day of the week	(1) Tuesday, (2) Monday	(1) Sunday, (2) Saturday
Appliances	Owns one or more game consoles Owns standalone freezer Owns tumble dryer Owns desktop/laptop computers No electric cookers Owns dishwasher (avg. 2 loads/day)	No game consoles No standalone freezer No tumble dryer No desktop/laptop computers Owns electric cookers Owns dishwasher (avg. 1 loads/day)
Property	Detached house Ownership is with a mortgage Wall insulations Double-glazed windows Energy-saving bulbs (avg. 25% of bulbs)	Semi-detached house Ownership is outright No wall insulations No double-glazed windows No energy-saving bulbs
Environmental Attitudes	'*Not enough time to reduce energy*' (64% disagree or strongly disagree, only around 21% agree or strongly agree) '*Change energy usage if it helps the environment*' (89% agree or strongly agree, only around 6% disagree or strongly disagree)	'*Other occupants do not want to reduce energy usage*' (39% disagree or strongly disagree, only around 22% agree or strongly agree)

Figure 8.
K-Means cluster analysis of energy consumption with demographic characteristics
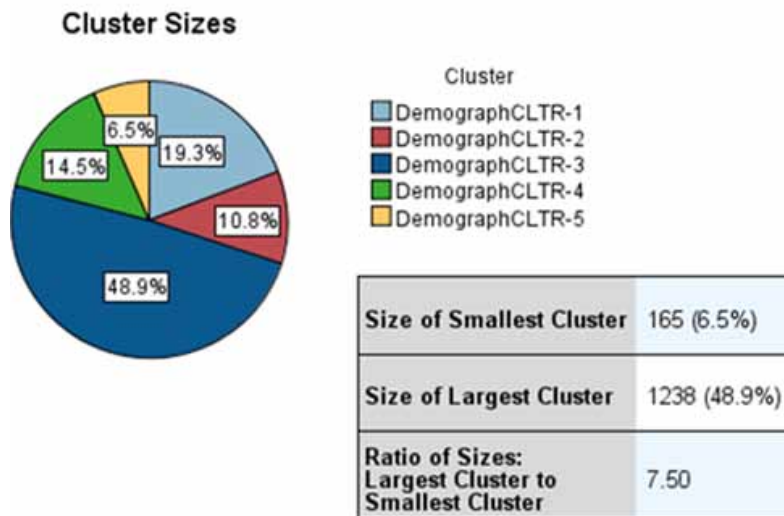
and accounts for 17.8% of the dataset while cluster 1 (in light blue) represents the lowest electricity consumption (mean average 214.48 KW/h) and accounts for 16.1% of the dataset. The most important predictors in this cluster analysis were 'Income Earner Education Level' and 'Internet Access.'

Table 11 displays the results of the linear regression for the high consumption Cluster-3 and suggests the variables 'Change energy usage behavior if it reduces bills', 'Other occupants do not want to reduce energy usage', and 'Not want to be instructed about energy usage' are significantly correlated with 'Total Consumption.'

Table 8.
Linear regression displaying the most significant variables (Cluster 2) (Dependent Variable: Total Consumption)

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	556,575	55,289		10,067	,000
Change energy usage if it reduces bills	-3,962	17,385	-,016	-,228	,820
Change energy usage if it helps environment	-4,683	12,590	-,025	-,372	,710
Changing energy usage behavior reduces bills	4,650	11,602	,026	,401	,689
Want to reduce energy usage	-13,110	12,279	-,073	-1,068	,287
Know how to reduce energy usage	5,775	9,334	,038	,619	,537
Other occupants do not want to reduce energy usage	-10,567	7,274	-,093	-1,453	,147
Not enough time to reduce energy	-16,747	8,653	-,127	-1,935	,054
Not want to be instructed about energy usage	3,340	8,483	-,026	-,394	,694

Table 9.
Linear regression displaying the most significant variables (Cluster 1) (Dependent Variable: Total Consumption)

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
1 (Constant)	195,421	22,115		8,836	0,000
Change energy usage if it reduces bills	-5,493	4,833	-0,065	-1,137	0,256
Change energy usage if it helps environment	12,706	5,715	0,126	2,223	0,027
Changing energy usage behavior reduces bills	-3,037	3,406	-0,046	-0,892	0,373
Want to reduce energy usage	-1,374	4,045	-0,018	-0,340	0,734
Know how to reduce energy usage	-0,395	3,631	-0,005	-0,109	0,913
Other occupants do not want to reduce energy usage	-0,126	2,953	-0,002	-0,043	0,966
Not enough time to reduce energy	2,273	3,542	0,030	0,642	0,521
Not want to be instructed about energy usage	-8,106	3,523	-0,109	-2,301	0,022

Table 10.
Low and high consumption clusters of the analysis with demographic characteristics

Attribute Types	Cluster 2	Cluster 1
	High Consumption	Low Consumption
Gender (of primary income earner)	Male	Female
Age	Respondent is in 46-55 range, all over 15	Over 65
Household occupants	Lives with other people	Lives alone
Environmental Attitudes	'Not enough time to reduce energy' (62% disagree or strongly disagree, only around 19% agree or strongly agree)	'Change energy usage if it helps the environment' (89% agree or strongly agree, around 6% disagree or strongly disagree) 'Not want to be instructed about energy usage' (74% disagree or strongly disagree, around 15% agree or strongly agree)

Figure 9.
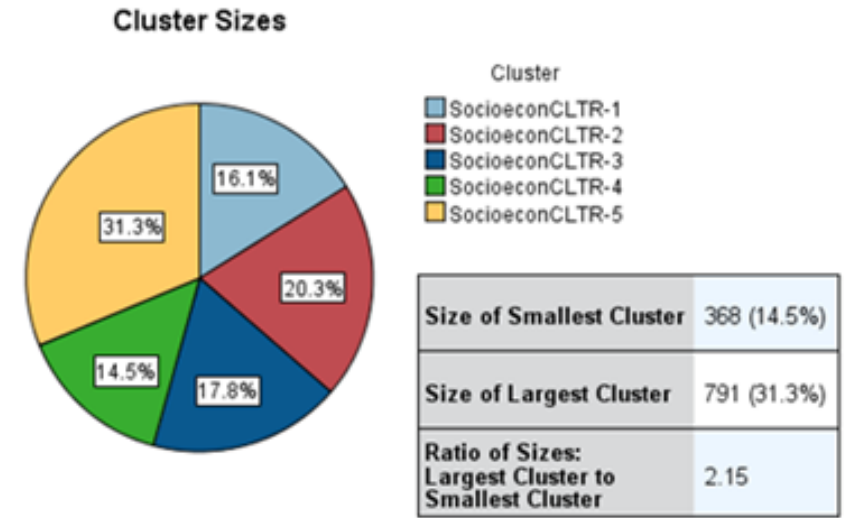**A 'K-means' cluster analysis representing the socio-economic characteristics**

Table 11.
Linear regression displaying all the significant variables

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	473,330	40,091		11,806	0,000
Change energy usage if it reduces bills	15,593	10,653	0,082	1,464	0,144
Change energy usage if it helps environment	11,231	11,130	0,056	1,009	0,314
Changing energy usage behavior reduces bills	-15,624	7,053	-0,116	-2,215	0,027
Want to reduce energy usage	-10,854	8,621	-0,067	-1,258	0,209
Know how to reduce energy usage	-0,648	7,089	-0,004	-0,091	0,927
Other occupants do not want to reduce energy usage	-11,949	5,067	-0,117	-2,358	0,019
Not enough time to reduce energy	-12,803	6,360	-0,102	-2,013	0,045
Not want to be instructed about energy usage	-7,479	5,545	-0,064	-1,349	0,178

Table 12 displays the results of the linear regression for the low consumption Cluster-3 and suggests the variables 'Change energy usage behavior if it reduces bills', 'Not enough time to reduce energy', and 'Not want to be instructed about energy usage' are significantly correlated with 'Total Consumption.'

Table 13 summarizes the demographic characteristics of the Low and High Consumption Clusters along with significant environmental attitudes determined in the regression analyses in *Table 12*.

Table 12.
Linear regression displaying the most significant attitude variables in Cluster 1

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	243,317	26,478		9,189	0,000
Change energy usage if it reduces bills	11,562	5,497	-0,127	-2,103	0,036
Change energy usage if it helps environment	9,888	6,216	0,094	1,591	0,112
Changing energy usage behavior reduces bills	-3,474	4,237	-0,046	-0,820	0,413
Want to reduce energy usage	-2,352	4,702	-0,028	-0,500	0,617
Know how to reduce energy usage	-2,025	4,471	-0,023	-0,453	0,651
Other occupants do not want to reduce energy usage	2,081	3,840	0,029	0,542	0,588
Not enough time to reduce energy	-12,027	4,678	-0,137	-2,571	0,010
Not want to be instructed about energy usage	8,626	4,103	0,109	2,102	0,036

Table 13.
Low and high consumption clusters of the analysis with socio-economic characteristics

Attribute Types	Cluster 3	Cluster 1
	High Consumption	Low Consumption
Employment	Currently working/employee	Retired
Education (of primary income earner)	Secondary to intermediate education	Primary education
Internet access (of home)	Yes	No
Environmental Attitudes	'Changing energy usage behavior reduces bills (77% agree or strongly agree, only around 11% disagree or strongly disagree) 'Not enough time to reduce energy' (68% disagree or strongly disagree, only around 18% agree or strongly agree) 'Other occupants do not want to reduce energy usage' (45% disagree or strongly disagree, only around 30% agree or strongly agree)	'Change energy usage if it helps the environment' (64% agree or strongly agree, only around 18% disagree or strongly disagree) 'Not enough time to reduce energy' (74% disagree or strongly disagree, only around 15% agree or strongly agree) 'Not want to be instructed about energy usage' (68% disagree or strongly disagree, only around 20% agree or strongly agree)

DISCUSSION OF FINDINGS AND IMPLICATIONS

Our study was designed to (1) investigate residential consumer behaviors with detailed physical, demographic, and socio-economic characteristics of each household using cluster analysis techniques that address prior methodological constraints, while (2) determining the implications of socioeconomic, demographic, and physical characteristics and corresponding attitudes towards energy conservation by residential users. The first component of the study's objective was achieved through examining the compiled data of both smart meter consumption and household survey though data cleaning, feature-selection, and linear regression methods to determine variables that are most significant to the 'total consumption'. The second was achieved by conducting the cluster analysis using the final 25 chosen variables. Unlike prior studies, we conducted three K-means cluster analyses splitting them by the three different characteristics. We also extended previous research by examining the

environmental attitude for the highest and the lowest electricity consumption within each cluster. As such, we expanded the cluster analysis of Flath et al (2012) by including more detailed data and investigating the physical, demographic, and socio-economic characteristics of each cluster.

Physical Characteristics and Environmental Attitude

Based on clustering around physical characteristics, we found respondents in the high-consumption cluster had more appliances and electronic devices such as computers, games consoles, as well as a tumble dryer and a standalone freezer. In contrast, the low consumption cluster did not have all these appliances and electronics but had an electric cooker and dishwasher - the latter of which was used less often than in the high consumption cluster. This finding, that ownership and use of more appliances lead to higher energy consumption, is logical and confirmed by previous research. The responses to the environmental attitude questions in the high and low consumption clusters did not suggest attitudinal barriers to reducing energy consumption – this suggests the savings offset was unintentional.

Demographic Characteristics and Environmental Attitude

When clustering by demographics, we found the high consumption cluster tended to be composed of males, 46-55 years of age, who lived with other adults (over 15). The low consumption cluster tended to be single females, over the age of 65 and who lived alone. Again, the responses to the environmental attitude questions in the high and low consumption clusters did not suggest the overall presence of attitudinal barriers to reducing energy consumption - in either cluster. However, the attitudes which significantly impacted energy use were different. In particular, the environmental attitude which found to be significant within the high consumption group was *disagreement with* the statement about not having enough time to reduce energy usage (62%). In contrast, the low consumption group showed a very positive overall attitude in learning about energy usage and about related practices that help the environment.

On the other hand, an earlier study, that used results from a 1993 to 2002 national "Environmental Attitudes, Values and Behavior in Ireland" survey, suggests that mid-range ages show higher concern (Kelly, Tovey, and Faughnan 2007). Our results might then be more consistent with prior research suggesting there is an interaction between gender and age as it relates to pro-environmental behavior (i.e., increase in pro-environmental behavior by women with age, relative to men (Steel, 1996)).

Socio-economic Characteristics and Environmental Attitude

When clustering by socio-economic characteristics, we found that in the high consumption cluster, the primary income earner was currently employed and had secondary to intermediate education, with access to the Internet. In contrast, the primary income earner in the low consumption group tended to be retired, with primary level education, and with no Internet. Both the high and low energy consumption groups were characterized by pro-environmental attitudes, especially as it relates to having time to reduce energy use. However, the high consumption cluster indicated an interest in changing energy use to reduce bills and disagreed that others in the home did not want to reduce energy use. In contrast, the low consumption cluster was characterized by an interest in reducing energy to help the environment and seemed to be more receptive to advice on how to reduce their energy use.

The findings seem generally straightforward in this case, except that the low consumption group, which tended to have a primary level of education, tended to want to reduce energy use to help the environment (more so than to reduce energy costs as with the generally higher educated, high consumption cluster). This result is interesting because national surveys in Ireland (Kelly et al, 2007) and more broadly in Britain (Brennan et al, 2015) show that concern over the environment tends to increase with education level, though the results of the survey in Britain shows that the GCSE group (closest to secondary level) showed a lower level of concern for the environment than those individuals with no attained education level.

LIMITATIONS AND CONCLUDING REMARKS

In our analysis, we created clusters based on physical, demographic, and socio-economic attributes. This provided us more stable clusters with meaningful and easier interpretation, and we gained insight into attitudes that aligned specifically with these clusters. Due to the use of a separate set of attributes in the analysis, possible interaction effects amongst the different sets of characteristics have not been included in our study.

Also, we used a *K*-means cluster analysis algorithm due to its simplistic nature with the justification that it is better for large datasets. A widely known limitation of a K-means is the pre-requisite of determining the optimal number of clusters. By following the common practice, we ran varying numbers of clusters with K-means for each characteristic and settled with the cluster analysis with the highest silhouette quality. Nevertheless, the ideal number of clusters is highly dependent on the judgment of the decision-maker and is subjective.

In addition, while the starting dataset contained 124 variables, 34 in total were used for analysis. The removal of most variables was due to a large number of missing values and high correlation or semantic similarity among the variables. The remaining variables are those retained by the feature selection process with multi-stage regression. Variables such as property age and size omitted by our feature selection processes, which have empirical advocacy in the literature, need further investigation in future studies.

Similarly, some of the other variables in the initial dataset which represent different environmental attitudes, such as "Inconvenient to reduce energy usage", "Energy usage reduction would not significantly reduce bills", "Cannot control own energy usage" and "Self-reported potential energy savings" may be worth further investigation in future studies to gain a wider perspective of the household's environmental views. Particularly, the first three variables can be helpful to investigate the value of developing a gamifying energy consumption using smart meter data may be futile for some households. Additionally, the environmental attitude variables used in the current study represent occupants' pre-trial attitude. Future studies are needed to investigate changes in pre-trial versus post-trial attitude.

Also, it is important to recognize that environmental attitude does not necessarily translate into expected behavior. In fact, research on ethical consumerism suggests there can be a very large gap between attitude and behavior (Carrington, Neville, and Whitwell 2010; Hassan, Shiu, and Shaw 2016). Moreover, this gap is not sufficiently explained by intention as would be suggested by the Theory of Reasoned Action (Fishbein and Ajzen, 1975). However, it may be more holistically explained by considering a mediating effect between intentions and behavior, of *having a plan to take action* (Carrington et 2010; Hassan et al, 2016). Constraints that exist around having the actual ability and control over taking an action, and from the situation in which the behavior occurs (e.g., consider moods/health, time of day, ease of access), may also moderate effects (Carrington et 2010). This may also explain why respondents in the high consumption cluster that focused on socio-economic characteristics tended to believe that 'changing energy usage behavior reduces bills' and disagreed with statements such as "not enough time to reduce energy' and 'other occupants do not want to reduce energy usage'.

Finally, the data used in this study was collected in the pre-pandemic, pre-climate catastrophic era. Looking ahead, longitudinal studies will be needed to understand the effects of pandemics such as Covid-19 on energy usage, and the degree to which energy use effects persist once the pandemic subsides. The World Economic Forum (2020) reports that the COVID-19 pandemic leads to an increase in multimedia activities with more time spent at home and is proliferating electronic gaming. How will high levels of unemployment, and changing activities such as working from home affect energy use patterns? Will changes in energy consumption behavior revert to the pre-pandemic state? Another anticipated impact will be the result of regional effects from climate change. Multiple climate change models predict energy demand will go up "by more than 25% in the tropics and southern regions of

the USA, Europe, and China" (van Ruijven et al, 2019). For all these reasons, and more, there is a continued and increasing need for more data-driven research into residential energy consumption that captures the evolving physical, demographic, and socio-economic factors and related environmental attitudes surrounding residential consumer energy consumption.

REFERENCES

Agwu, J. O. (2021). Emerging Themes, Concepts, and Solutions to Urban Challenges in Africa: Comparing Smart City and Sharing City. *International Journal of Urban Planning and Smart Cities*, *2*(1), 30–44. doi:10.4018/IJUPSC.2021010103

Ahuja, K., & Khosla, A. (2019). Data analytics criteria of IoT enabled smart energy meters (SEMs) in smart cities. *International Journal of Energy Sector Management*, *13*(2), 402–423. doi:10.1108/IJESM-11-2017-0006

Alexander, S. (2015). *Prosperous Descent: Crisis as Opportunity in an Age of Limits*. The University of Melbourne.

Amri, Y., & Fadhilah, A., Fatmawati, Setiani, N., & Rani, S. (2017). Analysis Clustering of Electricity Usage Profile Using K-Means Algorithm. *IOP Conference Series. Materials Science and Engineering*, *105*, 6.

Auffhammer, M., Baylis, P., & Hausman, C. H. (2017). Climate change is projected to have severe impacts on the frequency and intensity of peak electricity demand across the United States. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(8), 1886–1891. doi:10.1073/pnas.1613193114 PMID:28167756

Äyrämö, S., & Kärkkäinen, T. (2006). Introduction to partitioning-based clustering methods with a robust example. *Reports of the Department of Mathematical Information Technology. Series C, Software engineering and computational intelligence*, (1/2006).

Baccarne, B., & De Marez, L. (2021). Co-Shaping Smart Cities: Participation Inequalities in Civic Crowdsourcing. [IJUPSC]. *International Journal of Urban Planning and Smart Cities*, *2*(2), 34–47. doi:10.4018/IJUPSC.2021070103

Beckel, C., Sadamori, L., Staake, T., & Santini, S. (2014). Revealing household characteristics from smart meter data. *Energy*, *78*, 397–410. doi:10.1016/j.energy.2014.10.025

Bhardwaj, A., Dagar, V., Khan, M. O., Aggarwal, A., Alvarado, R., Kumar, M., Irfan, M., & Proshad, R. (2022). Smart IoT and Machine Learning-based Framework for Water Quality Assessment and Device Component Monitoring. *Environmental Science and Pollution Research International*, *29*(30), 46018–46036. doi:10.1007/s11356-022-19014-3 PMID:35165843

Bhattarai, T. N., Ghimire, S., Mainali, B., Gorjian, S., Treichel, H., & Paudel, S. R. (2022). Applications of smart grid technology in Nepal: Status, challenges, and opportunities. *Environmental Science and Pollution Research International*. doi:10.1007/s11356-022-19084-3 PMID:35138545

Brennan, J., Chanfreau, J., Finnegan, J., Griggs, J., Kiss, Z., & Park, A. (2015). *The Effect of Higher Education on graduates' attitudes: Secondary Analysis of the British Social Attitudes Survey.* U.K. Department for Business, Innovation, and Skills (BIS Research Paper No. 200).

Cagno, E., Micheli, G. J. L., & Di Foggia, G. (2018). Smart metering projects: An interpretive framework for successful implementation. *International Journal of Energy Sector Management*, *12*(2), 244–264. doi:10.1108/IJESM-08-2017-0009

Carrington, M. J., Neville, B. A., & Whitwell, G. J. (2010). Why ethical consumers don't walk their talk: Towards a framework for understanding the gap between the ethical purchase intentions and actual buying behaviour of ethically minded consumers. *Journal of Business Ethics*, *97*(1), 139–158. doi:10.1007/s10551-010-0501-6

Cayla, J., Maizi, N., & Marchand, C. (2011). The role of income in energy consumption behaviour: Evidence from French household data. *Energy Policy*, *39*(12), 7874–7883. doi:10.1016/j.enpol.2011.09.036

Chaudhari, A. Y., & Mulay, P. (2021). *Unleashing analytics to reduce electricity consumption using incremental clustering algorithm*. International Journal of Energy Sector Management., doi:10.1108/IJESM-11-2019-0016

Commission for Energy Regulation (CER). (2012). *CER Smart Metering Project - Electricity Customer Behaviour Trial, 2009-2010* [dataset]. 1st Edition. Irish Social Science Data Archive. SN: 0012-00. https://www.ucd.ie/issda/data/commissionforenergyregulationcer/

Croci, E., & Molteni, T.Croci & Molteni. (2021). Business Models for Smart City Solutions: An Overview of Main Archetypes. [IJUPSC]. *International Journal of Urban Planning and Smart Cities*, *2*(2), 94–109. doi:10.4018/IJUPSC.2021070106

Dhanwani, R., Prajapati, A., Dimri, A., Varmora, A., & Shah, M. (2021). Smart Earth Technologies: A pressing need for abating pollution for a better tomorrow. *Environmental Science and Pollution Research International*, *28*(27), 35406–35428. doi:10.1007/s11356-021-14481-6 PMID:34018104

Fabi, V., Andersen, R., Corgnati, S., & Olesen, B. (2012). Occupants' window opening behaviour: A literature review of factors influencing occupant behaviour and models. *Building and Environment*, *58*, 188–198. doi:10.1016/j.buildenv.2012.07.009

Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Addison-Wesley.

Flath, C., Nicolay, D., Conte, T., van Dinther, C., & Filipova-Neumann, L. (2012). Cluster Analysis of Smart Metering Data. *Business & Information Systems Engineering*, *4*(1), 31–39. doi:10.1007/s12599-011-0201-5

Gouveia, J., Gram-Hanssen, K. (2004). Domestic electricity consumption - Consumers and appliances. *The ecological economics of consumption*, 132-150.

Gram-Hanssen, K. (2013). New needs for better understanding of household's energy consumption-behaviour, lifestyle or practices? *Architectural Engineering and Design Management*, *10*(1-2), 91–107. doi:10.1080/17452007.2013.837251

Hansla, A., Gamble, A., Juliusson, A., & Gärling, T. (2008). Psychological determinants of attitude towards and willingness to pay for green electricity. *Energy Policy*, *36*(2), 768–774. doi:10.1016/j.enpol.2007.10.027

Hassan, L. M., Shiu, E., & Shaw, D. (2016). Who says there is an intention–behaviour gap? Assessing the empirical evidence of an intention–behaviour gap in ethical consumption. *Journal of Business Ethics*, *136*(2), 219–236. doi:10.1007/s10551-014-2440-0

IoT Analytics. (2020*). Smart Meter Market 2019: Global penetration reached 14% – North America, Europe ahead*. IOT. https://iot-analytics.com/smart-meter-market-2019-global-penetration-reached-14-percent/

Jones, R., Fuertes, A., & Lomas, K. (2015). The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings. *Renewable & Sustainable Energy Reviews*, *43*, 901–917. doi:10.1016/j.rser.2014.11.084

Kavousian, A., Rajagopal, R., & Fischer, M. (2013). Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behaviour. *Energy*, *55*, 184–194. doi:10.1016/j.energy.2013.03.086

Kelly, M., Tovey, H., & Faughnan, P. (2007). *Environmental Attitudes, Values and Behaviour in Ireland (2001-MS-SE1-M1) (Prepared for Environmental Protection Agency)*. EPA. https://www.epa.ie/pubs/reports/research/econ/ertdi%2065%20kelly%20report%20for%20web.pdf

Khosrojerdi, F., Akhigbe, O., Gagnon, S., Ramirez, A., & Richards, G. (2021). *Integrating artificial intelligence and analytics in smart grids: a systematic literature review*. International Journal of Energy Sector Management. doi: doi:10.1108/IJESM-06-2020-0011

Kuhe, A., & Bisu, D. Y. (2020). Influence of situational factors on household's energy consumption behaviour: Towards an effective energy policy. *International Journal of Energy Sector Management*, *14*(2), 389–407. doi:10.1108/IJESM-03-2019-0017

Leahy, E., & Lyons, S. (2010). Energy use and appliance ownership in Ireland. *Energy Policy*, *38*(8), 4265–4279. doi:10.1016/j.enpol.2010.03.056

Martinsson, J., Lundqvist, L., & Sundström, A. (2011). Energy saving in Swedish households. The (relative) importance of environmental attitudes. *Energy Policy*, *39*(9), 5182–5191. doi:10.1016/j.enpol.2011.05.046

Mills, B., & Schleich, J. (2012). Residential energy-efficient technology adoption, energy conservation, knowledge, and attitudes: An analysis of European countries. *Energy Policy*, *49*, 616–628. doi:10.1016/j.enpol.2012.07.008

Oltra, C., Boso, A., Espluga, J., & Prades, A. (2013). A qualitative study of users' engagement with real-time feedback from in-house energy consumption displays. *Energy Policy*, *61*, 788–792. doi:10.1016/j.enpol.2013.06.127

Parker, D. (2003). Research highlights from a large scale residential monitoring study in a hot climate. *Energy and Building*, *35*(9), 863–876. doi:10.1016/S0378-7788(02)00244-X

Rathod, R. R., & Garg, R. D. (2017). Design of electricity tariff plans using gap statistic for K-means clustering based on consumers monthly electricity consumption data. *International Journal of Energy Sector Management*, *11*(2), 295–310. doi:10.1108/IJESM-02-2016-0005

Ratner, S., Salnikov, A. A., Berezin, A., Sergi, B. S., & Sohag, K. (2022). Customer engagement in innovative smart grid deployment projects: Evidence from Russia. *Environmental Science and Pollution Research International*, *29*(4), 5902–5911. doi:10.1007/s11356-021-16095-4 PMID:34427884

Seixas, J., Luo, S., Bilo, N., & Valentim, A. (2015). *Understanding electricity consumption patterns in households through data fusion of smart meters and door to door surveys. pp.* ECEEE Summer Study Proceedings.

Sharma, D. D., & Singh, S. N. (2015). Aberration detection in electricity consumption using clustering technique. *International Journal of Energy Sector Management*, *9*(4), 451–470. doi:10.1108/IJESM-11-2014-0001

Siepermann, M., Rehtanz, C., Liebenau, V., Lackes, R., & Gebauer, M. (2021). The potential of shifting residential energy consumption for the energy transition. *International Journal of Energy Sector Management*, *15*(3), 628–646. doi:10.1108/IJESM-01-2020-0006

Steel, B. S. (1996). Thinking globally and acting locally?: Environmental attitudes, behavior and activism. *Journal of Environmental Management*, *47*(1), 27–36. doi:10.1006/jema.1996.0033

Tiwari, P. (2000). Architectural, Demographic, and Economic Causes of Electricity Consumption in Bombay. *Journal of Policy Modeling*, *22*(1), 81–98. doi:10.1016/S0161-8938(98)00003-9

U.S. Energy Information Administration (U.S. EIA). (2019). *International Energy Outlook 2019*. EIA. https://www.eia.gov/outlooks/ieo/

van Ruijven, B. J., De Cian, E., & Sue Wing, I. (2019). Amplification of future energy demand growth due to climate change. *Nature Communications*, *10*(1), 2762. doi:10.1038/s41467-019-10399-3 PMID:31235700

World Economic Forum. (2020). *COVID-19 is taking gaming and eSports to the next level.* World Economic Forum. https://www.weforum.org/agenda/2020/05/covid-19-taking-gaming-and-esports-next-level/

Zhou, S., & Teng, F. (2013). Estimation of urban residential electricity demand in China using household survey data. *Energy Policy*, *61*, 394–402. doi:10.1016/j.enpol.2013.06.092

Ali Emrouznejad (Professor and Chair in Business Analytics) is Director of the Centre for Business Analytics in Practice (CBAP) at Surrey Business School, UK. His areas of research interest include performance measurement and management, efficiency and productivity analysis as well as data mining and big data. He holds an MSc in applied mathematics and received his PhD in operational research and systems from Warwick Business School, UK. Dr Emrouznejad is editor / associate / member of editorial boards of several scientific journals. He has published over 200 articles in top-ranked journals and authored / edited several books.

Vishal Panchmatia was a MSc student in Business Analytics at Aston University.

Prior to joining NEOMA Business School in France as a Full Professor, Roya Gholami worked at University of Illinois (US), Carnegie Mellon University (US), Aston University (UK), and National University of Singapore. Gholami published and presented more than 70 papers in refereed international journals, conferences, and book chapters in areas of IT Value, Green IT/IS (Information Systems and Sustainability), IT Adoption and Diffusion, & IT for Development.

Carolee Rigsbee is Assistant Professor of Management at College of Business and Management, University of Illinois Springfield. She has extensive business and industry leadership experience in quality assurance and project/program management (software-based systems) and specializes in strategic management. Her research is focused on the design of innovative organization governance and value chain structures and processes that strengthen stakeholder voice and reduce obstacles to social goal achievement. Currently, Carolee is studying the potential of vertically integrated value chains (meta-organizations) enabled by blockchain technology and shared governance structures to deliver goods and services in support of the U.N. Sustainable Development Goals.

Hasan Kartal is an Associate Professor of MIS at the University of Illinois Springfield. His main research areas include Data Privacy, Information Economics, and Medical and Health Informatics.