The Bartlett School of Environment, Energy and Resources

MSc ESDA Title Page

UCL Candidate Code: NQMJ9

Module Code: BENV0092

Module Title: ENERGY DATA ANALYTICS IN THE BUILT ENVIRONMENT 24/25

Coursework Title: Exploring Time-of-use tariffs and sociodemographic factors as drivers of energy flexibility in Great Britain

Module Leader: Dr. Despina Manouselli

Date: 05/05/2025

Word Count: 2274

By submitting this document, you are agreeing to the Statement of Authorship:

I/We certify that the attached coursework exercise has been completed by me/us and that each and every quotation, diagram or other piece of exposition which is copied from or based upon the work of other has its source clearly acknowledged in the text at the place where it appears.

I/We certify that all field work and/or laboratory work has been carried out by me/us with no more assistance from the members of the department than has been specified.

I/We certify that all additional assistance which I/we have received is indicated and referenced in the report.

Please note that penalties will be applied to coursework which is submitted late, or which exceeds the maximum word count. Information about penalties can be found in your Course Handbook which is available on Moodle:
https://moodle.ucl.ac.uk/course/view.php?id=29435&section=0

- ***Penalties for late submission***
- ***Penalties for going over the word count***

Introduction

One of the many demand-side strategies utilities employ to shift demand on the grid in response to peak demand is time-of-use (ToU) tariffs. These tariffs incentivize consumers to use energy at cheaper prices during off-peak periods. Different variants of ToU tariffs exist, with some having price variations between weekdays and weekends, or high seasons and low seasons.

Consumers may respond to ToU tariffs by shifting their demand in an attempt to make savings or may be rigid in their energy use, with their energy flexibility continuing to be driven by their sociodemographic characteristics e.g., employment status or household size

This study analyzes data from the Energy Demand Research Project (EDRP) carried out in Great Britain between 2007 and 2010 with the aim to:

1) Determine what natural clusters of energy consumer profiles (energy flexibility) exist when ToU tariffs are employed across different timescales e.g., high season vs low season, weekend vs weekday.
2) Determine whether energy demand profiles are driven by ToU tariffs or consumer sociodemographic characteristics.

Literature Review

The EDRP was a major project in Great Britain whose aim was to deduce energy consumers' responses to additional information on their consumption. With over 60,000 participating households and 18,000 smart meters deployed, four energy supply companies, EDF, E.ON, Scottish Power, and SSE, oversaw the implementation of multiple interventions primarily directed at reducing domestic energy consumption, with a minority of interventions targeting the shifting of demand from off-peak periods. The interventions included energy efficiency advice, Customer engagement, real-time displays showing current electricity and gas use, and time-of-use tariffs (ToU). (AECOM, 2011.)

Time of Use tariffs have been extensively studied in recent years as a mechanism for demand-side management. Toretti et al (2021) examine how Time of Use (ToU) tariffs affect different residential electricity consumers based on both financial position and time availability and found regional differences in the distributional effects, such as positive effects for high-income groups in London, and similarities in household composition, like positive effects for households with children not in high-income groups.

Further, this study leverages socio-economic data as provided by ACORN (A Classification of Residential Neighbourhoods) to predict energy demand profiles. Developed by CACI Limited, ACORN provides a geodemographic segmentation of the UK's population and is used across the utility sector to predict consumption, shape communications strategies, and define product propositions (CACI Limited, 2009). For the Low Carbon London project, which included ToU tariff trials, ACORN was used to classify households according to geo-demographic groups (OASYS Group, 2023).

Several papers have combined clustering methodologies with sociodemographic factors to understand energy consumption patterns. Druckman, A. and Jackson, T. (2008) explore patterns of UK household energy use and associated carbon emissions at high levels of socio-economic and geographical disaggregation, examining specific neighborhoods with contrasting levels of deprivation, supporting this study's hypothesis that different segments have widely differing consumption patterns.

A comprehensive review by Frederiks et al. (2015) examines individual-level predictors of household energy usage, including socio-demographic factors such as income, employment status, dwelling type/size, and psychological factors such as beliefs, attitudes, and motives.

Methodology

As mentioned in the literature review, the four energy supply companies that participated in EDRP implemented a variety of interventions directed at energy consumers. Of the four companies, E.ON did not include Time-of-use tariffs as an intervention, and Scottish Power conducted the trials in phases, with different interventions trialed at different phases. (AECOM, 2011.)

Of the remaining two companies (EDF and SSE), it is beneficial to narrow the analysis down to energy consumers supplied by one company to control for differences in interventions between the two suppliers and differences in energy prices, thereby ensuring the analysis of energy flexibility is driven solely by ToU tariffs and sociodemographic characteristics.

The original dataset consolidated data from all four suppliers into one database as part of the anonymization process to ensure consumers could not be individually identified by directly linking them to their supplier (UK Data Archive, 2014). In this study, the Grid supply point (GSP) group code available in the geographical metadata was used as a proxy to infer consumer suppliers. SSE customers were chosen over EDF as the sample case because it was easier to infer consumers' energy suppliers from GSP codes, as per the GSP groupings in the literature findings given in Elexon ADR guidance note (Elexon, 2024) and National Grid's distribution code (National Grid, n.d.), because EDF is not explicitly mentioned therein. GSP codes "*P" and "*H" were used to filter for SSE-supplied consumers.

To analyze average daily load profiles, the data was converted from long to wide format, producing a dataframe grouped by household IDs and showing 24 columns for the average hourly consumption of the houses for the first trial year, 2009. Because the goal of energy flexibility analysis has to do with the shape of load profiles and when energy is used as opposed to the magnitude of consumption, and to avoid the impact of large consumers dominating the analysis of load profiles, row-wise normalization was carried out to produce the normalized hourly electricity consumption for each household.

With the data prepared, clustering algorithms were employed to determine which natural clusters of similar consumption profiles exist. K-means and Gaussian Mixture models (GMM) were considered. For the K-means algorithm, the elbow method and silhouette score were used to determine the optimal number of clusters.

For the GMM model, a function was implemented that generated several models and chose the one with the lowest Bayesian information criterion (BIC), a metric that measures the relative prediction errors of different models.

To further interpret the assigned clusterings, multinomial logistic regression was carried out. The independent predictor variables were the ACORN sociodemographic characteristics - ACORN Group Category, and the Time-of-use tariff indicator/flag, while the dependent target variable was the assigned clusters. The original dataset had the ACORN category and group variables alphanumerically encoded - this was decoded to textual descriptions of the sociodemographic variables as per ACORN's sociodemographic framework at the time of EDRP (CACI Limited, 2009).

Given that multinomial logistic regression requires a reference categorical value for result interpretation, the categorical values at the higher "end" of the sociodemographic scale were used as reference, i.e., "Wealthy achievers" and "Wealthy executives" were set as the reference category values for ACORN category and group values, respectively. This is because wealthy/affluent consumers may be presumed as the most energy flexible and hence can be used as a reference.

After conducting a baseline analysis with the entire dataset of SSE consumers for the year 2009, separate analyses were conducted to analyze load profiles on weekdays vs weekends and high season (November to February) vs low season (March to October).

Analysis and Results

i) Baseline analysis: K-means optimal number of clusters

Fig.1 shows the result of the elbow method analysis - it is observed that beyond **three** clusters, there is no significant improvement in the WSS, hence three clusters are chosen. The same is confirmed by the peak in silhouette score at three clusters seen in Fig. 2

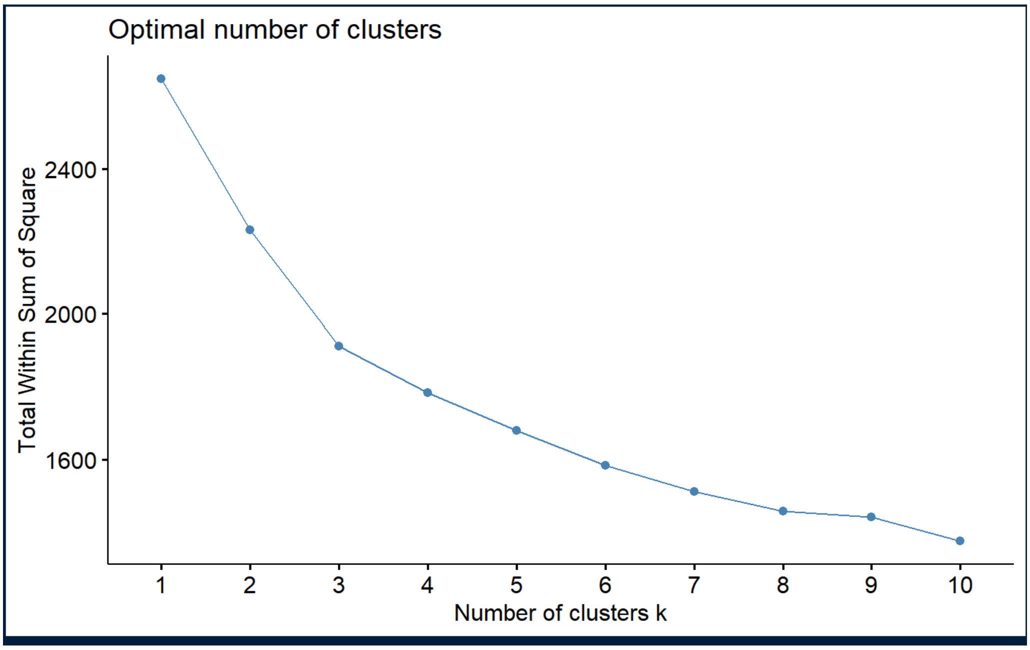

Fig 1: *Elbow method to determine optimal clusters*

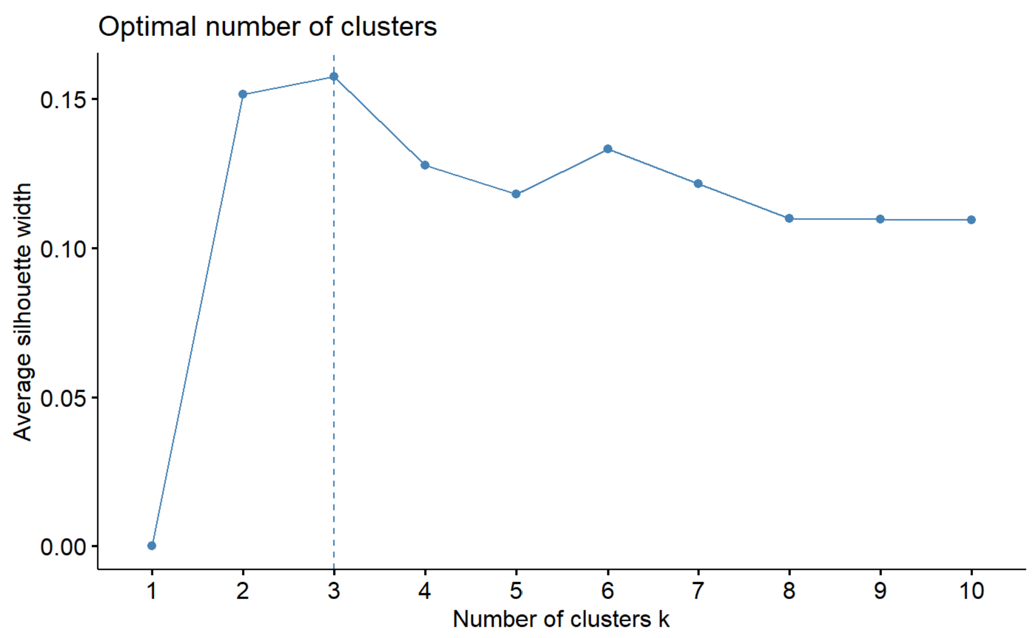

Fig 2: *Silhouette Score of k-means clusters*

ii) Load profile clusters

Fig. 3 shows the clusters as determined by a Gaussian mixture model (GMM). GMMs offer a more detailed breakdown of behavioral subgroups that may be valuable for targeted interventions or detailed analysis.

For interpretability sake, K-means was elected as the groupings produced by three clusters were expected to be more coherent and explainable by sociodemographic and ToU-tariff characteristics.

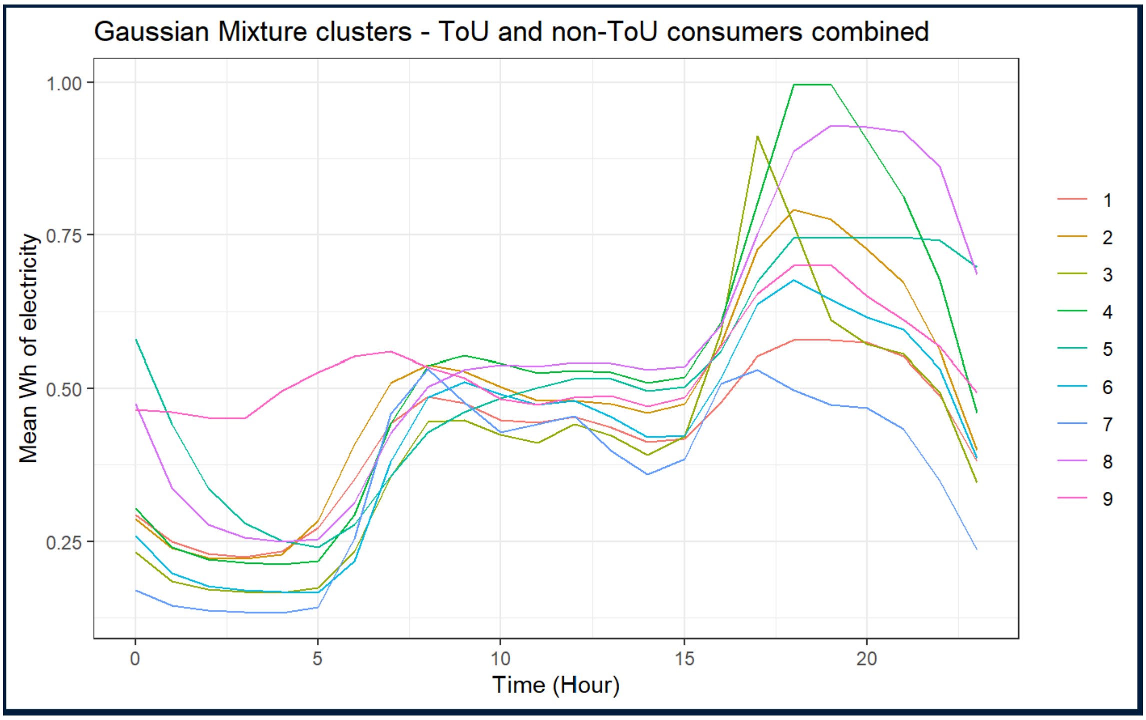

Fig 3: *Gaussian mixture models of energy consumption*

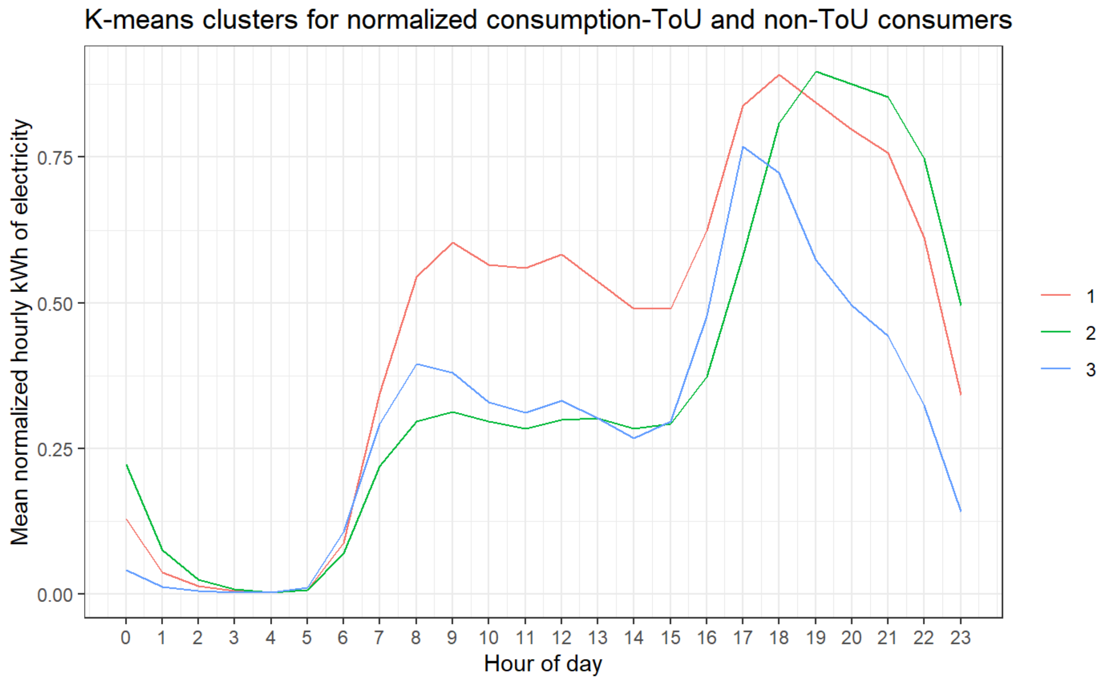

Fig 4: *Daily Load profile clusters derived by k-means*

Fig. 4 shows the three clusters of daily load profiles. Cluster 1 in red shows the highest daytime consumption with an earlier peak that starts rising around 3-4 PM Cluster 2 in green shows lower daytime consumption but the highest evening peak (especially around 8-9 PM), while Cluster 3 in blue shows a moderate consumption pattern with a less pronounced evening peak.

To determine whether the average normalized demand profiles of consumers on ToU tariffs were clustered together, the percentage membership of ToU consumers in the different clusters was analyzed.

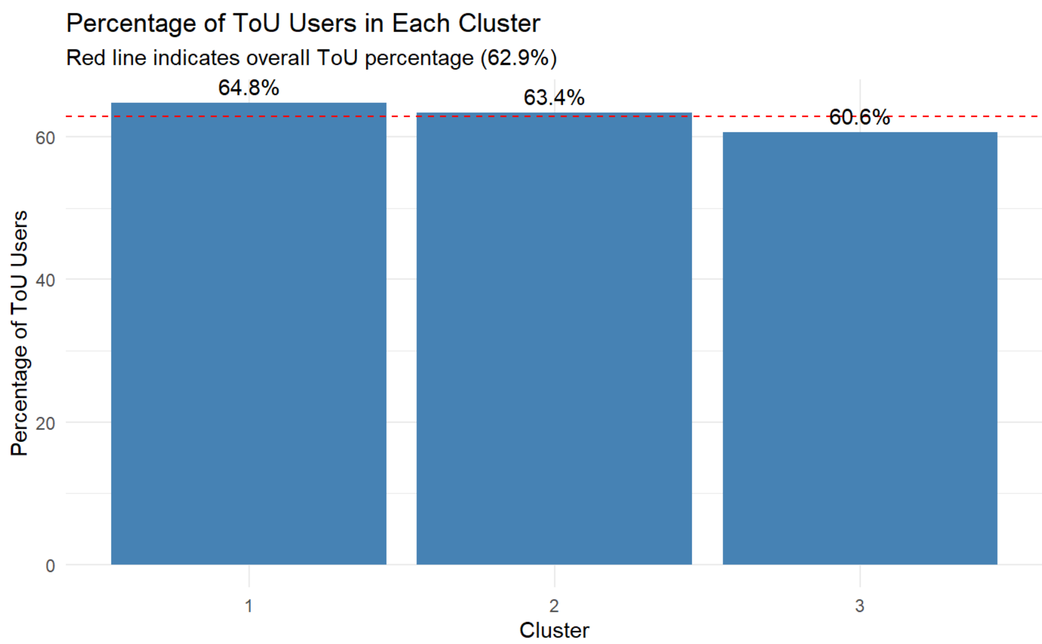

Fig 5: *Percentage of ToU consumers across the three clusters*

As seen in Fig. 5, the ToU consumers are evenly distributed across the three clusters and the percentage within each cluster is close to the percentage of ToU consumers across the entire population. This suggests that ToU consumers are indifferent to the tariffs and do not assume any particular load profile, i.e., their daily load profiles are evenly distributed across the whole population.

iii) Temporal/Seasonal flexibility

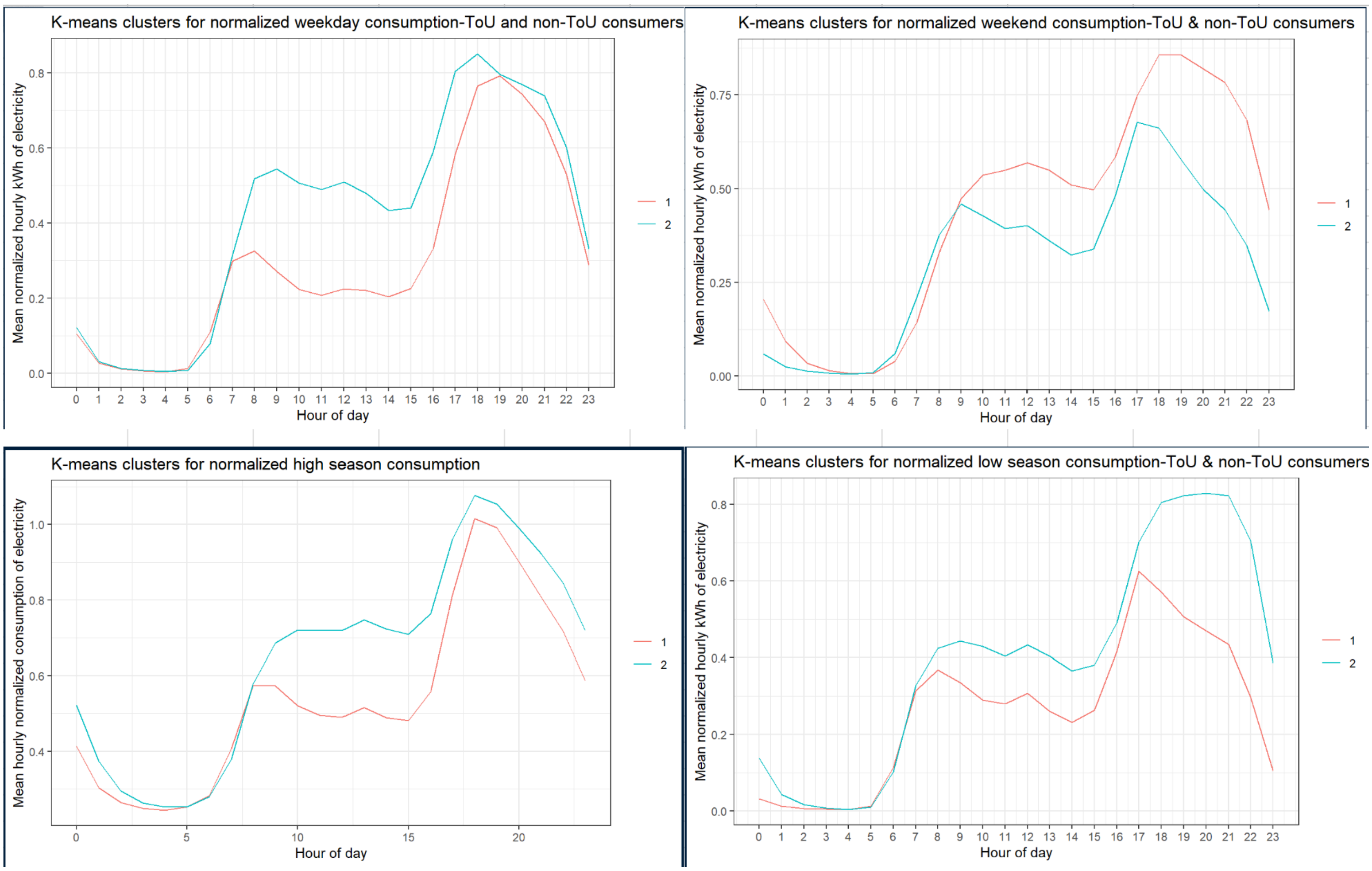

Fig 6: *Temporal energy flexibility*

Four separate k-means clustering models were run on the subset of normalized consumption data comprising weekdays, weekends, high season and low season, respectively. The weekday vs weekend analysis revealed two distinct clusters, with a wider gap in consumption during work hours (9AM-5PM) for weekdays than weekends, likely indicating that one group is comprised mostly of working professionals

The high-season vs low-season analysis revealed an overall consumption level higher in the high season than the low season for both clusters.

iv) Socio-demographic predictors

Table 1 shows the result of the multinomial logistic regression for the baseline analysis.

Predictors	*Odds Ratios*	*CI*	*p*	*Response*
Elec Tout [non-ToU]	1.03	0.83 – 1.28	0.776	2
ACORN Category [Comfortably Off]	1.99	1.50 – 2.64	**<0.001**	2
ACORN Category [Hard-Pressed]	3.09	2.30 – 4.15	**<0.001**	2
ACORN Group [Flourishing Families]	1.96	1.33 – 2.89	**0.001**	2
ACORN Group [Aspiring Singles]	2.50	1.37 – 4.57	**0.003**	2
ACORN Group [Burdened Singles]	1.66	1.05 – 2.63	**0.030**	2
Elec Tout [non-ToU]	1.17	0.95 – 1.43	0.140	3
ACORN Category [Comfortably Off]	1.59	1.22 – 2.07	**0.001**	3
ACORN Category [Hard-Pressed]	2.17	1.63 – 2.88	**<0.001**	3
ACORN Group [Settled Suburbia]	1.96	1.31 – 2.94	**0.001**	3
ACORN Group [Post-Industrial Families]	1.55	1.08 – 2.21	**0.017**	3
ACORN Group [Struggling Families]	1.46	1.02 – 2.08	**0.040**	3

Table 1: *Odds ratios of belonging to clusters 2 and 3 vs cluster 1 compared to the wealthy*

The multinomial logistic regression model used the ACORN socio-demographic and ToU tariff variables to predict the **three** assigned clusters as derived by k-means in the baseline analysis. The prediction model had a modest accuracy of 0.4 and sensitivity parameters across the clusters as seen in Fig. 7

```
                Accuracy : 0.4007
                  95% CI : (0.3603, 0.4421)
     No Information Rate : 0.3589
     P-Value [Acc > NIR] : 0.02102

                   Kappa : 0.0828

  Mcnemar's Test P-Value : 1.013e-10

Statistics by Class:

                     Class: 1 Class: 2 Class: 3
Sensitivity            0.5248  0.20482   0.4369
Specificity            0.5753  0.89951   0.6060
Pos Pred Value         0.4015  0.45333   0.3830
Neg Pred Value         0.6903  0.73547   0.6578
Prevalence             0.3519  0.28920   0.3589
Detection Rate         0.1847  0.05923   0.1568
Detection Prevalence   0.4599  0.13066   0.4094
Balanced Accuracy      0.5500  0.55216   0.5214
```

Fig. 7*: Multinomial regression model performance*

Discussion and Conclusion

Table 1 shows, for each sociodemographic category, the odds ratios of belonging to clusters 2 and 3 vs cluster 1, compared to the wealthy. Analysis of the table revealed the following key findings:

Cluster 2 compared to Cluster 1:

- Hard-Pressed households have the highest odds ratio (3.09), indicating they are 3.09 times more likely to be in Cluster 2 than Cluster 1 compared to Wealthy Achievers.
- Aspiring Singles have 2.5 times higher odds of being in Cluster 2.
- Comfortably Off households are 1.99 times more likely to be in Cluster 2 than Cluster 1.

Cluster 3 compared to Cluster 1:

- Hard-pressed households are 2.17 times more likely to be in Cluster 3 than Cluster 1.
- Settled Suburbia and Flourishing Families both have odds ratios of 1.96.
- Struggling Families have the lowest significant odds ratio at 1.46.

These results point to Wealthy Achievers/Executives (reference) being most strongly associated with Cluster 1's pattern of lower daytime usage but pronounced evening peak as seen in Fig. 4. This could be explained by larger homes that are empty during workdays, higher-powered appliances used in evening hours, and/or less price sensitivity to peak evening rates.

Hard-pressed households have the strongest association with Clusters 2 and 3, suggesting more daytime occupancy (potentially due to unemployment, part-time work, or retirement) as well as more distributed energy usage throughout the day and possibly higher price sensitivity affecting energy flexibility.

Moderate Means and Comfortably Off categories fall between these extremes, with patterns that align with typical middle-class lifestyles.

The analysis shows no statistically significant relationship between tariff type (ToU vs. non-ToU) and cluster membership (p=0.776 for Cluster 2 and p=0.140 for Cluster 3). While the odds ratios appear close to unity (1.03 and 1.17), suggesting minimal difference between ToU and non-ToU consumers in terms of cluster assignment, these findings are not statistically reliable. Therefore, we cannot conclusively determine whether load profiles are sensitive or indifferent to time-of-use tariffs based on these results. By contrast, sociodemographic factors such as ACORN categories and groups demonstrate statistically significant relationships with cluster membership, with Hard-Pressed households showing the strongest association

The strong statistical significance ($p<0.001$ for many variables) indicates these are robust findings that energy policy and utility companies could consider when designing tariffs and energy programs.

The report findings notwithstanding, the study was limited by several factors. The filtering of the dataset using Grid supply point (GSP) codes as a proxy for energy supplier may have introduced some error, as there is a possible overlap with other energy suppliers for any given GSP. The study did not take into account other interventions that were trialled in the EDRP project that may have influenced demand profiles. Other drivers of energy flexibility, such as the quality of building fabric and services, were not taken into account. Lastly, the study is limited by errors introduced in the modelling process, such as sampling errors, model selection and parameter tuning, e.g., setting the number of clusters in K-means introduces inherent bias.

Future work should better control for other interventions and geographical characteristics in the EDRP dataset to produce clusters that are solely influenced by sociodemographic variables. Use of the entire dataset should also be explored as well as the use of Gaussian mixture models to infer subcategories of consumers with unique load profiles and how sociodemographic characteristics may explain these. Finally, more temporal analysis, such as year-on-year analysis of load profiles, should be performed to determine demand-shifting trends.

References

AECOM, 2011. *Energy Demand Research Project: Final Analysis*. London: Ofgem.

CACI Limited. (2009). *ACORN user guide*. London: CACI. Available at: https://manualzilla.com/doc/6892872/acorn-user-guide (Accessed: 4 May 2025).

Druckman, A. and Jackson, T. (2008) 'Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model', Energy Policy, 36(8), pp. 3177-3192. https://doi.org/10.1016/j.enpol.2008.03.021

Elexon, 2024. *Draft Elexon ADR guidance note*. [online] Available at: https://bscdocs.elexon.co.uk/guidance-notes/draft-elexon-adr-guidance-note [Accessed 3 May 2025].

Frederiks, E. R., Stenner, K., & Hobman, E. V. (2015). The Socio-Demographic and Psychological Predictors of Residential Energy Consumption: A Comprehensive Review. Energies, 8(1), 573-594. https://doi.org/10.3390/en8010573

National Grid. (n.d.) *Distribution code*. Available at: https://www.nationalgrid.co.uk/our-network/distribution-code (Accessed: 3 May 2025).

OASYS Group (2023) *Data_Low_Carbon_London: Repository of energy consumption data in London households measured by smart meter systems* [Dataset]. GitHub. Available at: https://github.com/groupoasys/Data_Low_Carbon_London (Accessed: 4 May 2025).

UK Data Archive, 2014. *Smart Meter Energy Demand Research Project: Data Release UK Data Archive Study Number 7591 - Energy Demand Research Project: Early Smart Meter Trials, 2007–2010*. [data collection]. UK Data Service. SN: 7591. DOI: 10.5255/UKDA-SN-7591-1

Yunusov, T. and Torriti, J. (2021) 'Distributional effects of Time of Use tariffs based on electricity demand and time use', Energy Policy, 156, 112412. Available at: https://centaur.reading.ac.uk/98828/9/Distributional_effects_of_ToU_accepted.pdf (Accessed: 4 May 2025).