Microgrid Customer Segmentation

Stuart King
In collaboration with HOMER Energy

Background

HOMER Energy is a world leader in accurately modeling microgrid optimization. The HOMER (Hybrid Optimization of Multiple Energy Resources) software allows users to understand how to build cost effective and reliable microgrids that combine traditionally generated and renewable power, storage, and load management.

The objective of this study was to extract structure and meaning from HOMER Energy's collection of user software simulations. By being able to segment its customer base, HOMER can enhance the market access branch of its business model by providing its vendor partners with more reliable information related to the microgrid consumer market.

Method

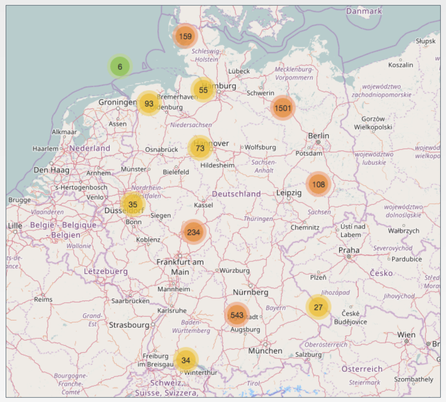
Because HOMER does not track which of its software users have become microgrid implementers, **k-modes clustering** was used to model the underlying structure of the data. K-modes is an extension of the popular k-means algorithm, however, instead of calculating distance, it quantifies the total number of mismatched categories between two objects: the smaller the number, the more similar the two objects. In addition, k-modes uses modes instead of means, in which the mode is a vector of elements that minimizes the dissimilarities between the vector and an individual data point.

The number of clusters was determined based on an evaluation of silhouette scores, resulting in k=4.

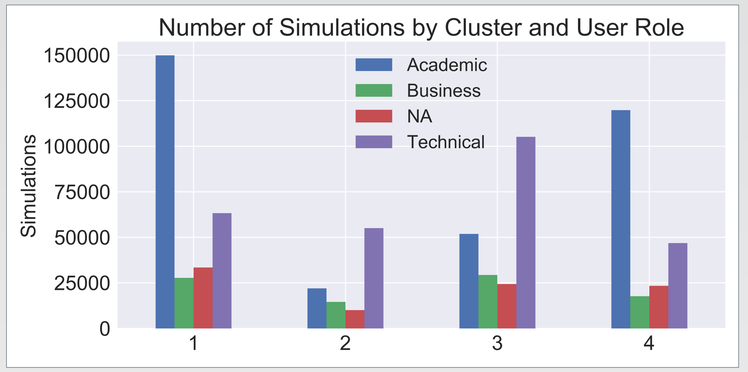
Results

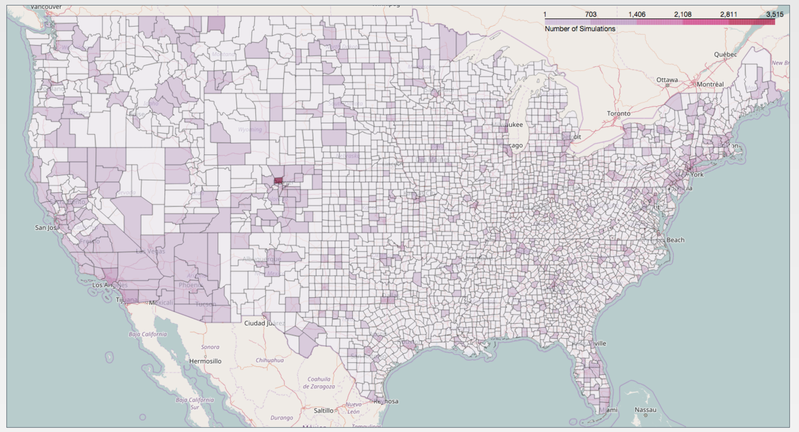
After clustering, the number of simulations are well distributed, with each cluster having distinct characteristics, including a majority user role (right) and propensity to model specific energy hardware.

Additional analysis demonstrated a propensity for users in Cluster 1 to perform simulations over the weekend, and Cluster 4 simulations excluded generators, likely indicating that these users were looking for supplemental energy sources.



A marker cluster map of simulations in Germany; map functionality includes clickable zoom to pinpoint simulation runs.





A county-by-county heat map of the United States showing Cluster 3 (engaged professionals) simulations.

To visualize the geographic concentration of simulations, cluster-specific **U.S. county heat maps** and country-by-country **marker cluster maps** were deployed to provide HOMER with additional insights.

Discussion

While the information gathered can be immediately helpful in targeting the right users, suggestions to expand upon the findings of this project include:

- Review **feature selection** and explore graph theory to model pairwise relationships between observations.
- Incorporate **time series** analysis to better understand the relevancy of simulations.
- Survey software users to allow for more advanced **predictive modeling**.
- Maps highlight the frequency of simulations near **airports** – a potential market for suppliers.

Cluster Profiles

1. Students and weekend warriors

2. Less active technicians with interest in wind

3. Engaged technical professionals simulating multiple system configurations

4. Academics and other professionals looking to supplement

Access the project's **web application** by following the QR code on the right to map simulations in any country in the world!

References

Huang, Zhexue. "Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values." *Data Mining and Knowledge Discovery* Volume 2. Issue 3 (1998): 283-304.

de Vos, Nico (2016). kmodes (Version 0.7) [software]. License: MIT. Retrieved from https://github.com/nicodv/kmodes

Contact Information:
sdking49@gmail.com
202-740-3685
github.com/stuart-d-king
linkedin.com/in/stuart-d-king