Individualizing Energy Benchmarking

Ian Xiao, Ben Steers, Te Du

Agenda

- Project Motivation
- Approach and Result Overview
- Demo
- Discussion and Next Steps

Why did we focus on this topic?

Our team chose building energy benchmarking is because …

- It is **challenging**: data quality issue due to from self reporting and user readiness[1]
- Its impact on specific **policy action remains unclear**: which buildings should put more effort in reducing energy consumption? And by how much so that it is fair?
- It is **controversial**: how do we define peer groups?

“The selection of peer groups is crucial to insightful benchmarking.”

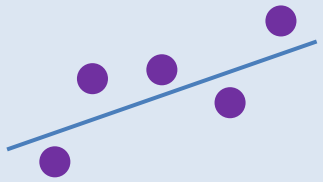
- Raul Valdes-Perez, Harvard Business Review, 2015

Source:

1. NYC’s Energy and Water Use 2014 and 2015 Report; published on Oct, 2017

What benchmark methods did we use?

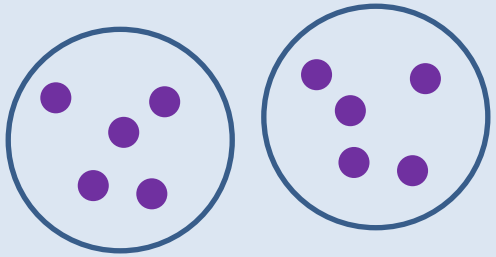
1 Linear Regression



Approach

- Fit a line using key building parameters as repressors
- The fitted value is the expected energy usage; then, compare actual and expected energy use
- This is used as a baseline method

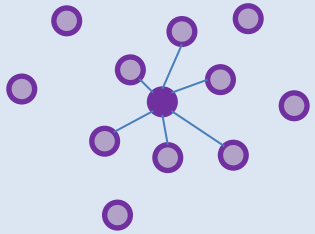
2 K-Means



Approach

- Develop N-clusters using key building parameter; having 5 clusters is optimal based on Silhouette Scores
- Compute average, minimum, and maximum energy use in each cluster
- Actual building energy is then measured against cluster level stats

3 Collaborative Filtering

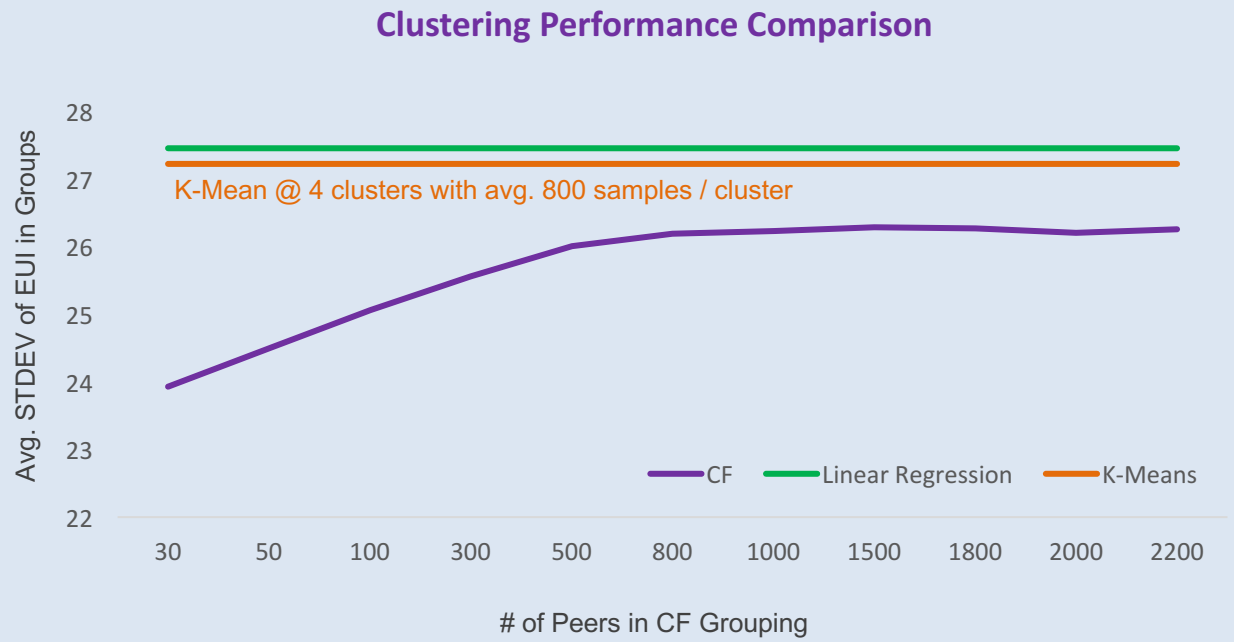


Approach

- Compute pair-wise Euclidean Distance based on key building parameters
- Find the closest N Peers; measuring using 700 peers is optimal
- Calculate average, minimum, and maximum energy use based on each building's peer group
- Measure each building's actual energy with its own peer groups

How do we measure peer selection?

Quantitative Measurement



Qualitative Assessment

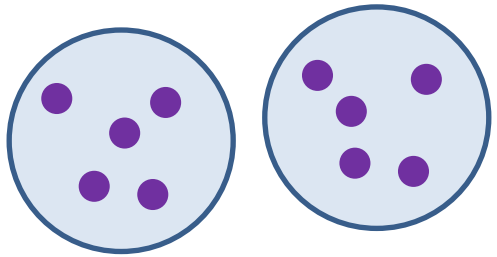
Collaborative Filtering offers:

- **Traceability** that provide clear explanation of why certain buildings are grouped over others
- **Dynamic, yet fair grouping** that does not require pre-defined boundary (except # of peers); it helps to avoid subjectivity
- **Scalable to higher dimension**, unlike K-Means, when more building parameters are included in the model
- **Deterministic**, instead probabilistic like many predictive modelling approach using decision trees

Note: *average standard deviation of Linear Regression is simply the standard deviation of EUI 2016 of total sample population; the one for K-Mean is the average STD of EUI 2016 of cluster = 4*

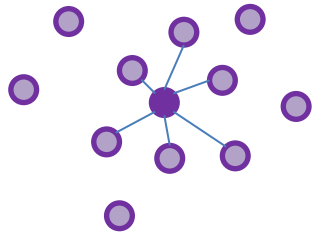
Is there a significant difference?

Summary of K-Mean Avg. Stdev of EUI 2016



	Clusters	avg_std_km
count	11.000000	11.000000
mean	7.000000	28.212471
std	3.316625	0.613608
min	2.000000	27.228395
25%	4.500000	27.959380
50%	7.000000	28.087397
75%	9.500000	28.404299
max	12.000000	29.403644

Summary of C.F. Avg. Stdev of EUI 2016



	Peers	avg_std_cf
count	11.000000	11.000000
mean	934.545455	25.684934
std	819.528689	0.829439
min	30.000000	23.930930
25%	200.000000	25.318520
50%	800.000000	26.198768
75%	1650.000000	26.246033
max	2200.000000	26.290484

Welch T-Test (sample size < 30 & Different Variance)

$$t = \frac{m_A - m_B}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}}$$

Null Hypothesis: There is no significant difference in the mean standard deviation in EUI based on K-Mean and C.F.

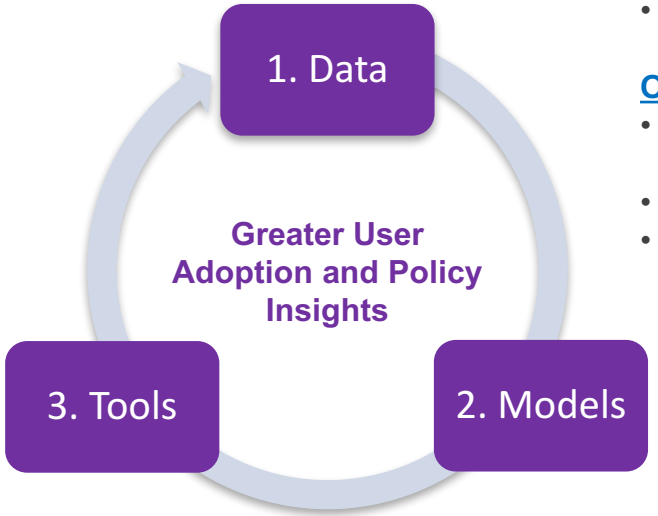
M_a = 28.21
S_a = 0.613^2 = 0.376
N_a = 11
M_b = 25.68
S_b = 0.829^2 = 0.687
N_b = 11

T = (28.21 – 25.68) / sqrt (0.376/11 + 0.687/11)
= **8.13 > 1.96**

Therefore, Null Hypothesis can be rejected @ alpha = 0.05. **C.F. peer grouping provides a statistically different standard deviation in EUI 2016 compared to K-Mean approach.**

Demo

Discussions & Next Steps



Limitation
- Less reliable self-reporting energy data
- Limited time horizon for meaningful trend analysis
- Year over year data is not standardized

Opportunities
- Introduce granular tenant data for analysis and validation for occupancy and energy usage
- Perform more rigorous duplication removal
- Develop automatic processing for new yearly data

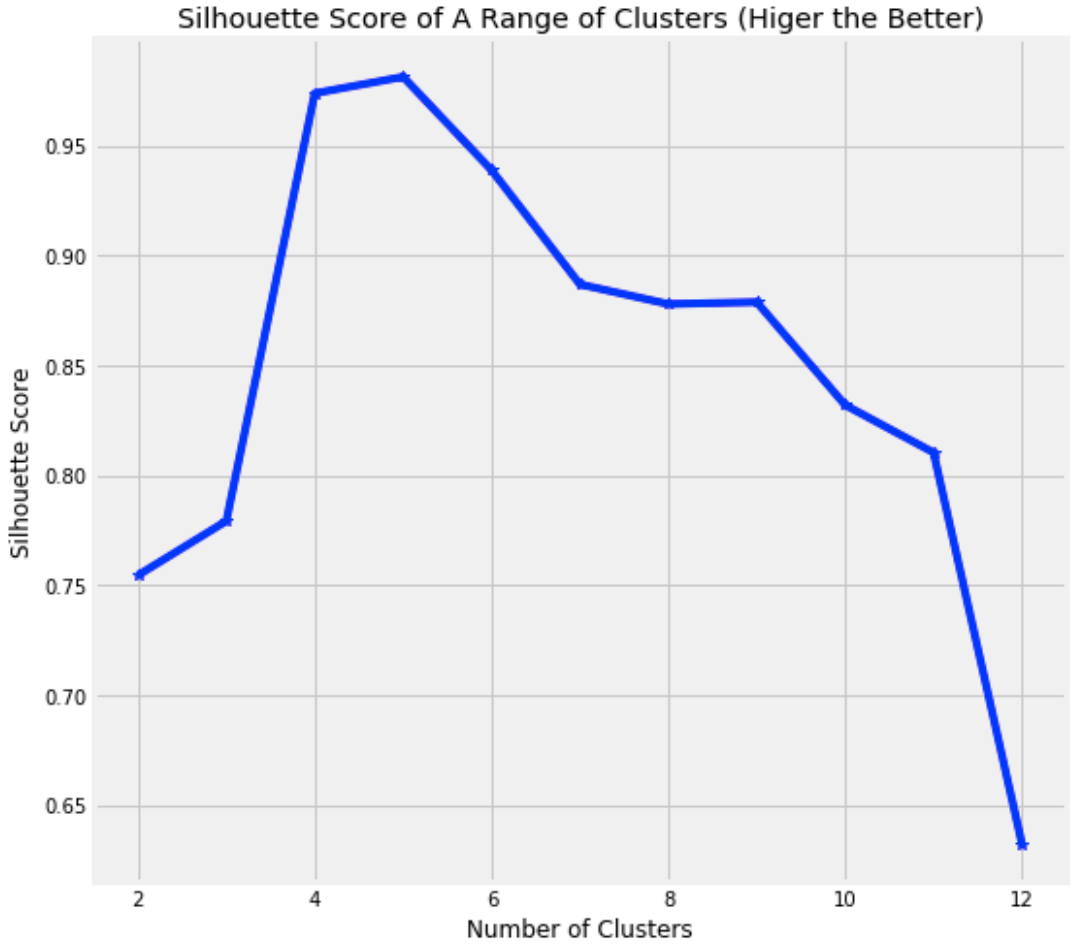
Limitation
- Run time based on dynamic users interaction
- More Tailored UX based on specific user need

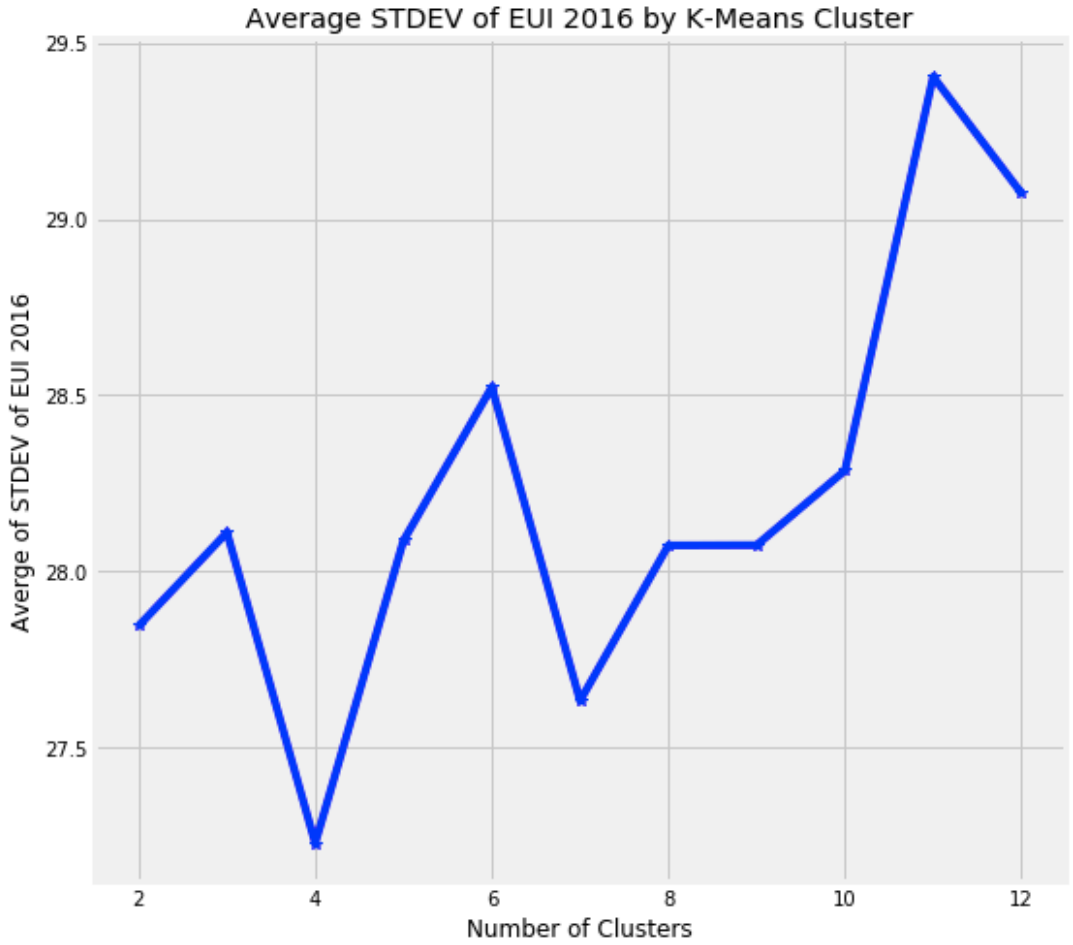
Opportunities
- Refine front and back end process and integration
- Conduct user discovery session to further define needs and the best UX elements

Limitation
- Run time when feature and sample sizes scale

Opportunities
- Develop weighted average based on distance for more robust benchmarking
- Ensemble models to integrate strength of various techniques
- Integrate more granular features from tenant data

Appendix A – K Mean Analysis





Appendix B – K Mean Cluster Details

	Clusters	avg_cluster_size	avg_std_km	max_cluster_size	min_cluster_size
0	2	1599.500000	27.847696	2938.0	261.0
1	3	1066.333333	28.819750	2658.0	82.0
2	4	799.750000	28.999299	2323.0	47.0
3	5	639.800000	29.380010	1924.0	37.0
4	6	533.166667	29.588011	1870.0	13.0
5	7	457.000000	29.888705	1672.0	13.0
6	8	399.875000	29.501516	1604.0	10.0
7	9	355.444444	29.789397	1609.0	2.0
8	10	319.900000	29.751046	1475.0	2.0
9	11	290.818182	29.571487	1155.0	2.0
10	12	266.583333	29.365383	1121.0	2.0

Appendix C – CF Peer Group Details (1)

	Avg_STDEV_EUI_30	Avg_STDEV_EUI_50	Avg_STDEV_EUI_100	Avg_STDEV_EUI_300	Avg_STDEV_EUI_500	Avg_STDEV_EUI_800
count	3193.000000	3193.000000	3193.000000	3193.000000	3193.000000	3193.000000
mean	23.930930	24.491923	25.069338	25.567701	26.010171	26.198768
std	7.119235	5.972372	4.351275	2.637836	2.656476	2.307777
min	11.054881	11.513088	13.374352	18.498774	19.961016	22.051839
25%	19.231347	20.721933	22.283085	23.915837	24.434498	24.756902
50%	22.960524	23.632545	24.661588	25.302923	25.270763	25.298941
75%	27.262962	27.275743	27.354332	26.812243	26.579906	26.423995
max	76.314235	62.327245	46.679634	36.575953	34.243297	31.549737

Appendix C – CF Peer Group Details (2)

Avg_STDEV_EUI_1000	Avg_STDEV_EUI_1500	Avg_STDEV_EUI_1800	Avg_STDEV_EUI_2000	Avg_STDEV_EUI_2200
3193.000000	3193.000000	3193.000000	3193.000000	3193.000000
26.236173	26.290484	26.268088	26.214801	26.255894
1.952052	1.359133	1.146079	0.988842	0.802864
22.731816	24.393274	24.706538	24.962888	25.093293
25.044064	25.371277	25.431934	25.557814	25.746654
25.460707	25.614276	25.734972	25.745743	25.904565
26.278735	26.926769	27.360180	27.229908	27.176872
30.589434	29.155651	28.579181	28.379640	28.000337