THESIS OF THE

DOCTORAL DISSERTATION

DOCTORAL (PhD) DISSERTATION

UWE RADTKE

HUNGARIAN UNIVERSITY OF

AGRICULTURE AND LIFE SCIENCES, KAPOSVÁR CAMPUS

2023

Hungarian University of Agriculture and Life Sciences
(MATE), KAPOSVÁR CAMPUS

Doctoral School of Economic and Regional Sciences

Head of the Doctoral (PhD) School

Prof. Dr. ZOLTÁN BUJDOSÓ, PhD

Supervisor:
Dr. Gyöngyi Bánkuti, PhD
Associate Professor

URBAN ENERGY SYSTEMS - ANALYSIS OF SMART METERING DISTRICT HEATING SYSTEM OF KAPOSVÁR

DOI: 10.54598/002700

Created
UWE RADTKE

KAPOSVÁR

2023

1 Research aim and objectives

The research will consist of several parts contributing to several questions regarding the reason and the use of smart meters in a district heating system. The aim is to focus on local needs and take a scientific, methodological approach to local problem-solving. The district heating system of Kaposvár will be explored and used as example to perform an economic and (customer) service research. With the analyses and the results, the heating plant has access to more information and could easier explore strategic suggestions. This can indirectly contribute to the efficiency of district heating, the reduction of environmental load and higher consumer satisfaction.

District heating (also known as heat networks or teleheating) is a system for distributing heat generated in a centralized location through a system of insulated pipes for residential and commercial heating requirements such as space heating and water heating. The heat is often obtained from a cogeneration plant burning fossil fuels or biomass, but heat-only boiler stations, geothermal heating, heat pumps and central solar heating are also used. Heat waste from factories and nuclear power electricity generation is also used and common. District heating plants can provide higher efficiencies and better pollution control than localized boilers. According to some research, district heating with combined heat and power (CHP) is the cheapest method of cutting carbon emissions, and has one of the lowest carbon footprints of all fossil generation plants. (Andrews 2009)

Consumer behavior is the study of how people make buying decisions. It attempts to understand how buyers choose and use products and services. By understanding how buyers think, feel and decide, businesses can determine how best to market their products and services (Southeastern

Oklahoma State University 2022). Understanding buyers can help marketers connect with them and influence their behavior. In regard to sustainability – personal behavior is one of the key success factors. Understanding energy users' consumption patterns benefits both utility companies and consumers, as it can support improving energy management and usage strategies. The heat usage of customers is crucial for effective district heating (DH) operations and management.

The deployment of smart meters offers a unique opportunity for researchers and district heating utilities to analyze large-scale data and discover both typical, as well as atypical, patterns in the network. Within the research a data-driven approach shall be used to partition district heating users into separate clusters such that users in the same cluster possess similar consumption and behavior pattern. Because of the unavailability of high-resolution, hourly or sub-hourly meter data before the installation of smart meters, the literature on analytics in district heating is still in its infancy. There are not many studies focusing on the analysis of heat load patterns in district heating systems. So far only a few reference papers could be identified.

The proposed methods for this research will include the use of the K-means algorithm to segment the different groups based on demand intensity and representative patterns according to measured values. Understanding of consumers' opinion about district heating with Q-method-based opinion categorization method is planned to be a major part of the research. By quantifying subjective data, the research will get information about consumers opinions and mentalities – e.g. regarding the tariff system and expected (sustainable) behavior.

The objective of the research project is to address the questions: RQ1: Does fully anonymous (but identified with a unique ID) data measured by smart meters provide new insight? Along with that it shall also be explored if personalized data will be needed for direct influence and strategies. RQ2: Can sampled data effectively capture the cluster information of large datasets in clustering analysis, potentially reducing the hardware requirements for processing and analysis? RQ3: Which strategies could the district heating company apply to leverage the available measured data the most in regard to the general sustainability goals? RQ4: Can the strategies be related to the personal behavioral style? RQ5: Would consumers follow the strategy and support it?

2 Methods and Materials

2.1 Literature review

According to (Radtke 2022) it’s important to use structured methods for unbiasing science. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) scheme was applied for the systematic review (Fink 2020; Guba 2008). PRISMA is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses. PRISMA primarily focuses on the reporting of reviews evaluating the effects of interventions, but can also be used as a basis for reporting systematic reviews with objectives other than evaluating interventions (Moher et al. 2009).

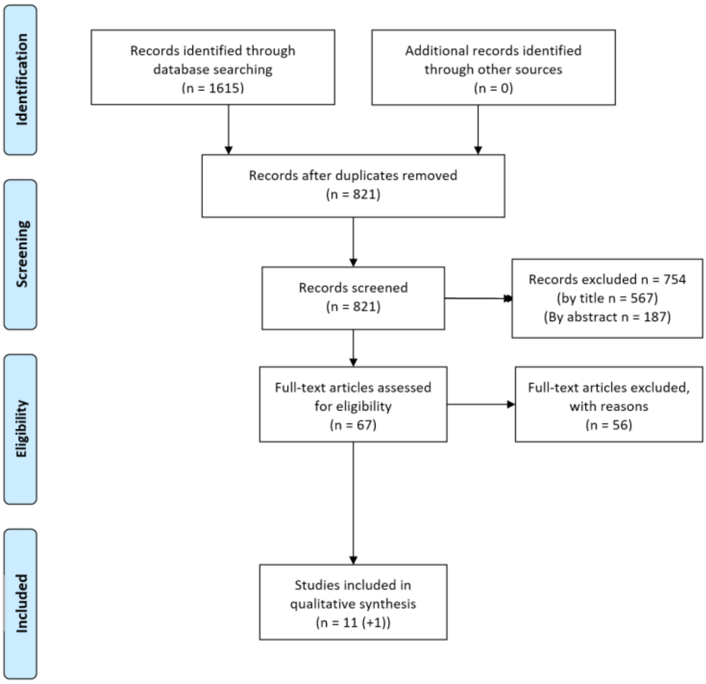

Figure 1: Prisma flow chart

Initially only the customer perspective or the consumers opinion on district heating (all aspects) was supposed to be included, but there were not enough published results. Almost all documents on the subject of consumer opinion were in relation to the pricing or sustainability. In order to broaden the basis, the different pricing models were included into the review scope. For pricing mechanisms in district heating and the customer opinion towards DH, similar methods were applied as already before. The summary of those methods can be described as follows:

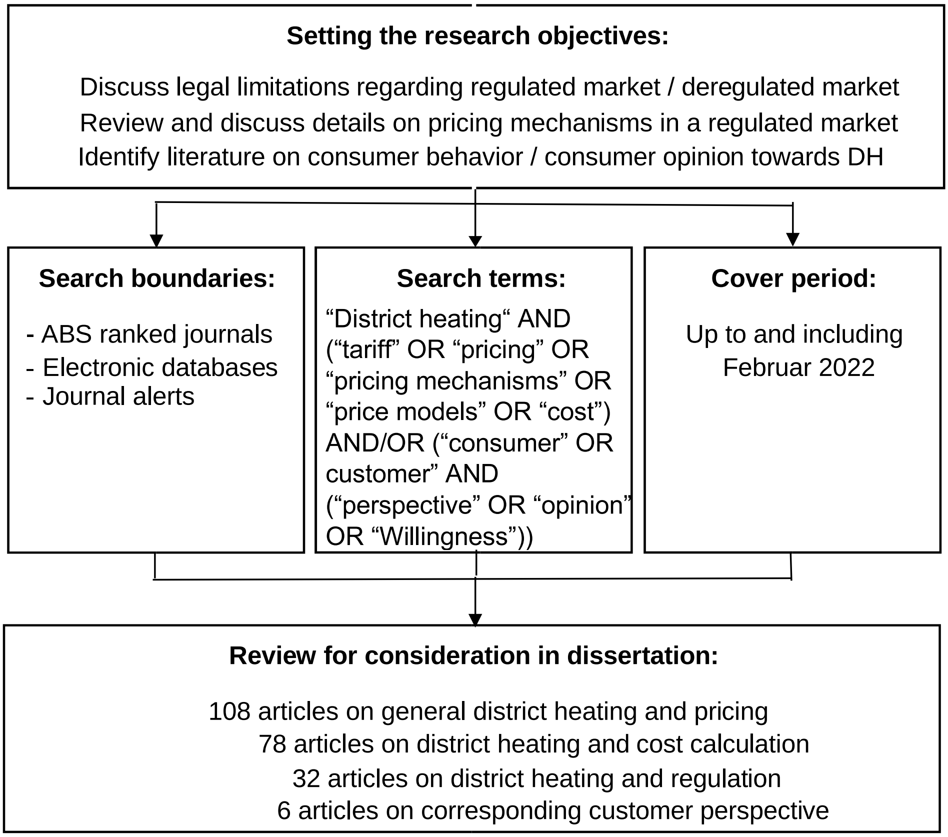

Figure 2: Summary on literature review

Given the fact that especially Hungary works with regulated prices throughout the energy sector in total, the publications having no relation to regulated price procedures were totally excluded.

The most prominent work of (Lajos Kerekes 2022) had to be excluded as it’s available only in Hungarian language. It provides an overview of the available biomass and geothermal resources, the technical characteristics and cost elements of the different technologies and estimates the costs of increasing the share of heat generation from biomass and/or geothermal and CHP in the district heating systems studied

2.1 K-means Cluster analysis of hourly measured power demand.

As found and described within the literature review, the clustering approach shall be done via K-Means. K-Means clustering is a method of vector quantization, originally from signal processing, that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean (cluster centers or cluster centroid), serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells. k-means clustering minimizes within-cluster variances (squared Euclidean distances), but not regular Euclidean distances, which would be the more difficult Weber problem: the mean optimizes squared errors, whereas only the geometric median minimizes Euclidean distances. For instance, better Euclidean solutions can be found using k-medians and k-medoids (MacQueen 1967). It is a well-known approach for clustering data (Jain 2010).

2.2 Data description and preliminary steps

After the initial analysis, the data was converted from the given SQL Server format into a readable STATA format. Using STATA for transforming the date and time as well as the measured demand (P1) into computerized forms was the first step. But the later steps were performed using R respectively R Script as the library and documentation was easier to access and the R Script did allow a computerized output. The used variables are:

- Cons_ID - Consumption place identifier [unit: --]
- Dat_Ro - Date of Readout [unit: --] -> transformed to YYYY-MM DD HH:MM:SS
- Op_hrs - Operation hours counter [unit: hour]
- P1 – Power/Demand (instantaneous value) [unit: kW]

Data description:

```
obs:      3,332,901
vars:             6                          24 Nov 2021 21:31

               storage   display    value
variable name   type     format     label      variable label
cons_id         long     %12.0g                Cons_id
Dat_Ro          str19    %19s                  Dat_Ro
Op_hrs          int      %8.0g                 Op_hrs
p1              str15    %15s                  P1
P1_numeric      float    %9.0g                 P1
datetime        double   %tc
```

Figure 3: description of the unprocessed dataset

2.3 Consumer's opinion on district heating using Q method

Q Method was developed by William Stephenson to examine individuals' psychological attitudes. Q method is a factor analysis, which analyses the people themselves not their characteristics (William Stephenson 1955). The method focuses on differences within individuals. The mathematical basis of the method is the same as the mathematical basis for factor analysis. By using Q method relatively large number of statements can be assessed with relatively small number of individuals involved in the research. Correlation coefficients that are calculated by the method show correlation between people. The Q method can be considered as an inverse factor analysis. The typical benefits of Q methodology include insight into the perceptions of individuals at a level where broad social forces are enacted within individual awareness. According to (Sue Ramlo 2016): "…methodological aspects of Q offer the ability to scientifically study subjectivity." The qualitative methods of Q-methodology allow participants to express their (subjective) opinions and the quantitative methods of Q-methodology use factor analytic data-reduction and induction to provide insights into opinion formation as well as to generate testable hypotheses (Valenta and Wigger 1997). Q-methodology research emphasizes the qualitative how and why people think the way they do; the methodology does not count how many people think a certain way. For that purpose, Q provides a more systematic approach and higher methodological transparency than purely qualitative methods (Brown 1996).

2.4 Defining the concourse, Developing Q-Set and Selection of P-set

Various sources and methods for constructing a concourse have been defined by Q-methodologists, for example (scientific) literature, (expert) interviews, focus groups, (social) media, websites. The material represents existing opinions and arguments, everything that representative organizations, professionals and other experts have to say about the topic. Prior to the development of the concourse, a research question must be identified to guide the entire research process. Therefore, the following questions were applied to define the concourse:

- What influences the opinion on district heating?
- What can be identified as everyday behavior of consumers?
- Which digital experience is expected and can contribute?
- Is sustainable behavior known to consumer?

To determine the statements, several sources were used, including the already reviewed literature, lectures and discussions with professors, expert interviews and also statements from websites. At the end of the concourse development, 61 statement items were identified with four themes: General district heating, prices, requirements/expectations for software/apps and personal behavior. The Q method often raises the question of the reliability of the method. In order to objectivize the Q-set definition, a Delphi-like technique was applied.

3. Results

3.1 Empirical Results on demand-based measurements

P1 was heavily skewed and is therefore log-transformed in the next steps:

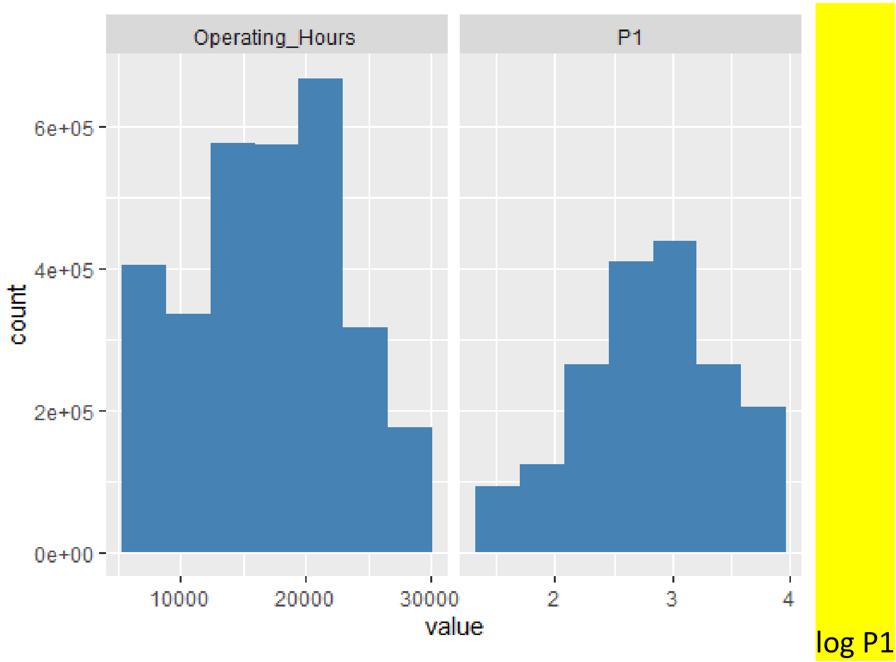

Figure 4: distribution of operating hours and P1 after transformation (log P1)

By this logarithmic transformation P1 was normalized, thus can be used for modelling.

3.2 Results of the K-means analysis

Sampling needed using seed (by setting the seed also at subsampling you force the same output at multiple runs) – a sample of 10 000 values was used. Sampling was needed in order to enable reproducible results. The initial dataset could not be handled on a PowerBook workstation with Xeon CPU with 6 cores and 12 Threads at 2.7 GHz and 64 GB RAM. The clustering was performed once using a virtual server cluster with 8 CPUs and 1024 GB RAM – but that system was not available after an initial run. Therefore, the sub-sampling approach was used. The results between the full dataset and sampled dataset differ only marginally. Initially for testing and setting up the algorithm a random number of k = 5 were used. The detailed data distribution is removed of the below script, only the sum of square was kept for later comparison.

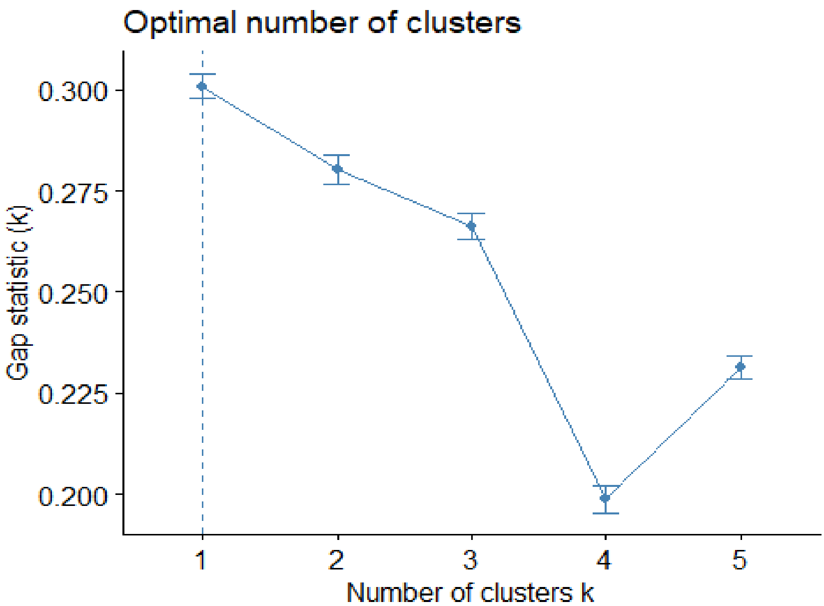

Figure 5: Optimal number of k using the gap statistics

The optimal k is calculated and visible with 4. Within the next steps the clustering will be performed using k = 4. This indicates there are 4 clusters – the same result was achieved with the full set of data. For the number of clusters there is no difference between the sampled data and the full set of data.

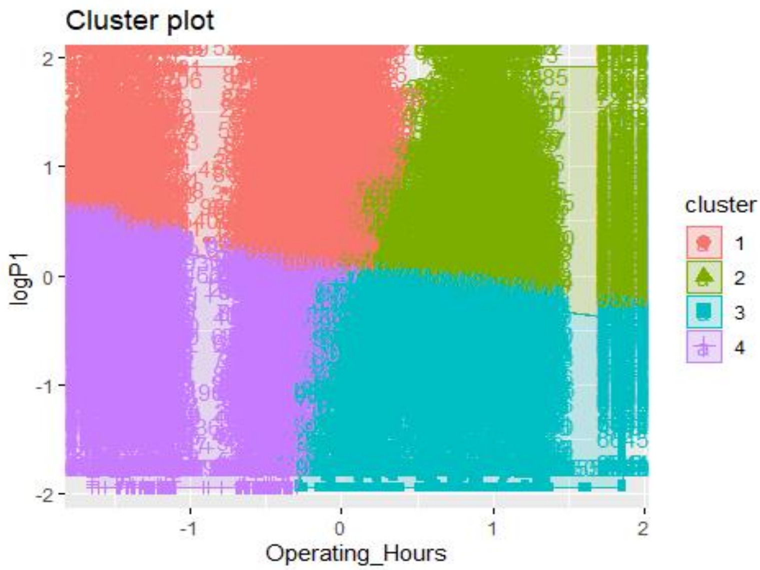

Figure 6: cluster plot sampled data

Retransforming of the cluster centers is needed in order to evaluate the result. Just the log-transformed demand value will not enable to evaluate the found clusters. The cluster centers can be easily determined using the corresponding command. But for the retransformed centers the earlier calculated values for standard deviation and the mean is needed.

The values for standard deviation and mean were already calculated within the metric statistics and if the formula is now applied to. For the P1 values which was transformed to a logarithmic value, the calculation has to be reversed using the exponential function.

Table 1: Retransformed cluster values

cluster	Operating_Hours	P1
1	13678.97	37.45434
2	23021.50	27.16575
3	21419.25	11.63607
4	10439.83	13.10257

3.3 Q-Method Results

Investigation of factor loadings, eigenvalues, explained variance, factor correlations and composite reliability scores suggests the four-factor solution, accounting for 57.4 % of the variance (see table 2).

Table 2: Characteristics of factor loading

	F1	F2	F3	F4
Average reliability coefficient	0.80	0.80	0.80	0.80
Number of loading Q-sorts	9	4	7	1
Eigenvalues	4.1	3.3	3.2	1.5
Percentage of explained variance	19.5	15.6	15.3	7.1
Composite reliability	0.97	0.94	0.97	0.80
Standard error of factor scores	0.16	0.24	0.19	0.45

The same total amount of loading Q-sorts was achieved by using 5 factors. The loadings per factor were different, but in total all 21 sorts could be directly allocated to a factor. As the loading of four factors did provide already more than 55% of explanation, this number was taken for further evaluation. The composite reliability as above the recommended value of 0.7 (Hair, Black, William & Anderson 2010) for all four extracted factors.

The used software package "qmethod" based on 'R' provide the following automated flagging for the factor loadings (table 3):

Table 3: Factor matrix provided by qmethod()

Participants	F1	F2	F3	F4
P01	0.11	0.21	0.52*	0.36
P02	0.03	0.26	0.43*	-0.13
P03	0.48*	0.17	0.33	0.04
P04	0.05	-0.07	-0.01	0.86*
P05	0.23	-0.16	0.68*	-0.25
P06	-0.06	0.04	0.91*	-0.18
P07	0.13	0.22	0.37*	-0.19
P08	-0.31	0.36	0.61*	0.02
P09	0.20	-0.10	0.64*	0.35
P10	0.49*	0.36	-0.01	0.20
P11	0.11	0.72*	0.22	-0.15
P12	0.10	0.66*	0.24	-0.06
P13	0.19	0.78*	-0.06	-0.11
P14	0.82*	-0.58	0.03	0.26
P15	0.81*	0.06	-0.23	-0.21
P16	0.75*	0.28	-0.06	-0.07
P17	0.70*	0.32	-0.17	0.06
P18	0.54*	-0.01	0.37	0.10
P19	0.63*	0.23	-0.13	0.03
P20	-0.05	0.75*	0.16	0.21
P21	0.54*	-0.30	0.21	-0.29

Note: Values calculated after factor rotation; marked values indicating a defining sort (a significant loading) automatically by “qmethod”.

3.4 Quantitative data analysis of Q-Methods

For statement 14 “District heating supports (large) investments in renewable energies.” all factors distinguish. But for statement 19: “An app should always have several divisions integrated (district heating, electricity, water, gas, ...)” consensus is achieved. The implications will be drawn later. To investigate how the perspectives (factors) differ within this thesis the analysis will focus on the distinguishing statements.

While it did not achieve a consensus, the statement 29 did achieve a positive attitude (agree, most important).

Factor 1: Space Savings and Convenience: The Appeal of District Heating

Factor 1 encompasses customers' perspectives on district heating, emphasizing the appeal of space savings, convenience, environmental considerations, and the role of government intervention and support.

Table 4; Table distinguishing statements for factor 1 ($p < 0.05$)

Statement	Q-SV	Z-Score
1. It is mainly the space savings that encourage many customers to buy district heating.	-4	-2.26
12. The state should intervene even more in the energy market and make and enforce regulations regarding the technologies used.	0	-0.11
14. District heating offers price stability.	2	0.04
21. District heating should be supported and promoted by the government.	1	-1.08
22. Customers should have the option to choose between different district heating providers.	2	-0.10
35. District heating is a sustainable and environmentally friendly heating solution.	1	0.93
38. District heating should be the primary heating solution for new residential buildings.	0	0.00

Factor 2: Advantages of District Heating

For factor 2, the highest-rated statements are "Security of supply is always the top priority in a district heating system" and "District heating supports (large) investments in renewable energies." These ratings indicate that customers highly value the security of energy supply provided by district heating systems and recognize the positive impact of district heating on promoting investments in renewable energy sources. The high ratings suggest that customers perceive these aspects as important and desirable qualities of district heating systems. The least important part – similar to factor 1 – is the space saving aspect

Table 5: Table distinguishing statements for factor 2 ($p < 0.05$)

Statement	Q-SV	Z-Score
1. It is mainly the space savings that encourage many customers to buy district heating.	-4	-1.66
8. District heating should also be available in sparsely populated areas. (Preferably from a local resource.)	3	1.43
14. District heating offers price stability.	4	1.52
39. Consumption and billing via app (if available) must be displayed graphically.	-1	-0.69

Factor 3: Energy Market Dilemma: Convenience in District Heating

Factor 3 The integration of smartphone apps that offer personalized rate suggestions based on consumption caters to the convenience and preferences of customers. Additionally, the willingness to embrace changes in habits and adopt energy-conscious behaviors reflects a positive attitude towards sustainability. However, differing perspectives may arise concerning the optimal level of state intervention and the enforcement of regulations in the energy market. Striking a balance between market dynamics and sustainable practices is vital for the successful implementation of district heating systems.

Table 6: Table distinguishing statements for factor 3 ($p < 0.05$)

Statement	Q-SV	Z-Score
5. Combined heat and power (CHP) make optimum use of fuels.	0	0.11
12. The state should intervene even more in the energy market and make and enforce regulations regarding the technologies used.	-4	-1.76
14. District heating supports (large) investments in renewable energies.	2	0.75
26. A smartphone app must offer/suggest the cheapest rate for me depending on consumption.	2	-0.97
30. I would be willing to change my habits (turn off the heating on vacation, ventilate now and then instead of constantly ventilate, …).	2	0.87
35. The state should let the market regulate prices (market economy).	0	0.15

Factor 4: Debunking Myths: District Heating's Value and Sustainability

Factor 4 examines the aspects of simplicity, security, and sustainability in district heating systems based on varying viewpoints. While opinions differ on the simplicity of district heating for housing temperature control, there is a strong rating on governmental influence in investments and the need for insulation. District heating is recognized as a catalyst for investments in renewable energies. Additionally, aligning energy taxation with climate protection goals is considered significant. Factor 4 explores the dimensions of simplicity, security, and sustainability in district heating systems.

Table 7: Table distinguishing statements for factor 4 ($p < 0.05$)

Statement	Q-SV	Z-Score
4. Simple operation: In terms of comfort, district heating is probably the simplest way of housing temperature control.	-2	-0.95
5. Combined heat and power (CHP) make optimum use of fuels.	0	-0.95
10. No more research and investment should be made in district heating.	1	1.89
13. Security of supply is always the top priority in a district heating system.	-3	-1.42
14. District heating supports (large) investments in renewable energies.	-3	-1.42
25. If an app suggests savings measures, e.g., home insulation, different thermostat, ... I would take the appropriate steps to implement the recommendations.	2	0.95
33. The taxation of energy products in Hungary / in the EU must be more closely aligned with climate protection aspects.	3	1.42

Finding common socio-demographic factors for each factor group was not possible. For example, factor group 3 included participant 5, 6, 7, 8 and 9 – all having different educational degree, different preferences on used devices and partnership. The age range also differs. Only the persons living in the household (HM2 or HM3) were close by. But with P01 also being part of F3 that did not match any more. Therefore, the factor groups represent different attitudes towards sustainability and district heating – but there is no unified socio-demographic characteristic identified.

4 Conclusion and Recommendations

The K-means analysis results show:

- K-means clustering on large data requires high performing hardware
- Data sampling reveals the same number of clusters
- Even data without private data like building type or addresses clustering can be performed
- Four clusters were identified in a two-dimensional space
- LOW operating hours with a HIGH demand is the least preferrable option under normal circumstances

Any additional steps like identifying meters and the behind consumers for cluster #1 from sampled data or performing the test for maybe an additional year should be part of a follow-up research. These have to be kept as gaps as the primary goal was to evaluate the possibility of clustering and the identification of the clusters. Follow-up research could also perform structuring / data analysis containing private data (house size, type of use (private, office, ...)).

The result did reveal a significant heat demand at low operating hours – given the starting date of the measurements, this is in the early phase of the heating period. The use of waste heat like the heat and methane gas produced by the sugar factory could be enhanced by additional sources as well. That would be a measure which could be performed without any further research on the details of each cluster.

Consumers have a positive attitude towards district heating as a key to sustainability. All participants except one use electronic devices for banking, and trust in smartphone apps is high. However, they don't trust apps to regulate their heating. Digitalization is seen as beneficial for environmentally friendly heat generation. A small number of participants want paper bills. Consumers support further research in district heating.

The integrated usage of several divisions is important for smart meter apps. The socio-demographic group is tech-savy and some prefer smartphones over PCs.

Actually, Hungary has state regulated prices – and many other European countries (e.g. Germany) are introducing state regulated prices as well. Nevertheless, the results of the previous chapters of the dissertation can be used to describe a rather new pricing model. The best pricing method for district heating varies depending on several factors (IEA 2022), including the characteristics of the system, the customer base, and the regulatory framework. Some of the most commonly accepted pricing methods include:

1. Flat rate pricing: This method involves charging a fixed amount per unit of energy consumed, regardless of the time of day or season.
2. Time-of-use pricing: This method involves charging different rates based on the time of day, with higher prices during peak demand periods and lower prices during off-peak periods.
3. Dynamic pricing: This method involves using advanced metering and billing systems to charge customers in real-time based on the prevailing market price for energy.
4. Conservation pricing: This method involves offering incentives to customers who reduce their energy consumption, such as lower rates for those who use energy during off-peak periods.

Ultimately, the best pricing method for district heating will depend on the specific needs and goals of the system and its customers. The first two are quite common and also widely accepted as found within the literature review. Conservation pricing might work for electricity, but heating periods depend on outside temperature – when it’s warm outside it rare to find consumers needing a lot of heat. Dynamic pricing is something which would have to be explored deeper based on the findings of this dissertation, but the idea development shall be started.

With this in mind, the dynamic pricing method could be designed as follows:

A. High demand for a long period of time (cluster 1)

During peak periods of high demand, prices could be adjusted to reflect the increased usage

For example, if the average price per unit of energy during off-peak periods is €0.10, during peak periods of high demand the price could increase to €0.15

Additionally, the price could be made transparent by showing users how it compares to the average consumption of a selected comparison group

This would allow users to adopt more sustainable habits and make more informed decisions about their energy usage

B. High demand for a short period of time (cluster 2)

During short periods of high demand, prices could be adjusted to reflect the increased usage

For example, if the average price per unit of energy during off-peak periods is €0.10, during a short period of high demand the price could increase to €0.12

This would encourage users to shift their energy usage to off-peak periods, which would help to balance the demand on the system

C. Low demand for a long period of time (cluster 3)

During low demand periods, prices could be adjusted to reflect the reduced usage

For example, if the average price per unit of energy during off-peak periods is €0.10, during low demand periods the price could decrease to €0.08

This would encourage users to continue their energy usage during these periods, which would help to balance the demand on the system

D. Low demand for a short period of time (cluster 4)

During short periods of low demand, prices could be adjusted to reflect the reduced usage

For example, if the average price per unit of energy during off-peak periods is €0.10, during a short period of low demand the price could decrease to €0.09

This would encourage users to shift their energy usage to these periods, which would help to balance the demand on the system.

The utilization of cross-company waste heat, where waste heat that cannot be utilized internally is redirected to third parties in commercial or residential buildings, presents a viable option. However, there are key challenges associated with this approach, such as the need for reliable data to match waste heat potential with demand, as they may not always align. Currently, the most economically feasible utilization of waste heat relies on the spatial proximity between the waste heat source and the demand. Heat recovery or heat displacement techniques serve as efficient and straightforward technological approaches to harness waste heat, thereby enhancing overall energy and cost efficiency. Heat exchangers are commonly employed in this process, facilitating the transfer of waste heat to a transport medium that can distribute it to other units. However, some losses occur during this heat transfer process. Furthermore, the effective utilization of waste heat by third parties necessitates additional transport infrastructure, including local and district heating pipes, buffer storage, and more. Local and district heating networks offer advantages in their flexibility to accommodate various heat sources, both centralized and decentralized. They can integrate diverse energy sources at different levels and points throughout the year, irrespective of seasonal variations. This enables the extraction and profitable integration of economically viable waste heat into the heating network. Consequently, companies can reduce cooling water costs, generate income from the sale of heat energy, and significantly contribute to CO_2 emissions reduction by utilizing waste heat that would otherwise require alternative energy generation methods.

4.1 Further research aspects

Further research on the topic of district heating and its impact on energy systems could include:

- Examining the potential for energy efficiency measures to be combined with electrification in order to meet heating demand where it is significant.
- Examining local challenges facing implementation of renewable energy infrastructure projects and investigating solutions to overcome them.
- Investigating innovative solutions for the urban energy systems, like waste heat recovery and combined heat and power.

- Examining the potential for demand-side management strategies, such as load shifting and peak shaving, to reduce the strain on urban energy systems during times of high demand.
- Investigating consumer preferences and behavior in relation to energy consumption and willingness to support further investments in district heating and urban energy systems.

Especially the last part would be crucial to expanding involvement of the European union. Country-specific distinctions would result in more specific programs dependent on local needs and possibilities. Countries such as Norway with their huge availability of waterpower generations are in need for totally different funding programs than especially East European countries. Here is still gap on research towards consumer behavior and expectation (in Eastern Europe).

The work has shown it's possible to get some results on anonymous measured values – but better and further recommendation could only be given on more detailed data. Building size, number of inhabitants or detailed definition of business hours could further contribute to recommendation such as insolation of the buildings itself or even indicate missing pipelines for easier consumption and low-loss transmission. The transmission itself could also be part of further research. The infrastructure is already provided and available in Kaposvár – but could there be secondary usages such as de-icing: Heat transmission pipelines can be used to melt snow and ice on roads and sidewalks to improve safety and mobility during winter conditions. Another option could be research in the area of soil remediation: heat transmission pipelines can be used to heat contaminated soil to temperatures high enough to destroy pollutants and make the soil safe for use again or enable and help farmers grow crops year-round, regardless of the outside temperature.

5 New scientific findings

The new scientific results are based on answering the research questions of the dissertation. The dissertation consists of mainly three parts: literature review, K-means analysis of the measured values and the consumer opinion.

The research gaps identified in the literature review could be closed to some extent. Mainly the lack of research in Eastern Europe based data and the lack of research in consumer opinion in district heating. According to the literature review, the most used and efficient method for clustering district heating data is K-means. Other methods have been discussed and evaluated, but no consensus on an alternative method was found among the reviewed literature. Each researcher who tried a different approach only compared their results to K-means. Main findings are:

1. As the coverage of Eastern European countries in this field is nonexistent, this research can be considered to be a pilot one at this field.

While performing the actual K-means analysis, the smart metering based (containing hourly measured data) district heating dataset is extremely large and requires high-performance hardware and careful data analysis. Cloud-based techniques are not sufficient for this type of investigation. It is possible to use K-means clustering on them, although it requires high-performance hardware. Main finding here is:

2. The dataset was sampled to reduce hardware demands, and the same number of clusters and values were identified.

3. Four clusters were identified in the two-dimensional space (of operating hours and heat demand variables) by using K-means algorithm. The cluster centers from the full and from the sample database have proved to be the same – was proven by t-test at significancy level of 0.05. The result also revealed a large heat demand at low operating hours as well as on a larger number of operating hours.

While analyzing the consumer opinion closes another gap within the research in general, the performed method reveals all consumers generally accept the necessity and effectiveness of district heating. To conclude the finding:

4. Q-method analysis reveals four distinct perceptual factors influencing attitudes towards district heating: environmental sustainability, convenience and reliability, cost-effectiveness, and concerns about individual autonomy.

Based on the identified K-means clusters and groups of the Q-method findings, a new idea came up:

5. In case of smart metering-based dynamic regulation available for the customers, a dynamic pricing model should be created as it encourages customers to reduce energy consumption during peak demands, that makes a step into sustainable direction.

At the peak demand times the heat comes mostly from less-sustainable resources thus, this financial technique makes a step into sustainable direction. Obviously, this will only be implementable if the official prices for district heating get liberalized. This method can encourage customers to reduce energy consumption during peak demand, increase revenue, and align supply and demand.

Based on the technical details(and/or additional parameters) it is desirable to use the waste heat for district heating from the glass factory that is under construction in Kaposvár. This solution is not only CO_2 emission reducing but also environmentally friendly by reducing the environmental heat load that the cooling technology would imply. It is also economical friendly, as it eliminates the need to install cooling technology, which would require a significant investment.

The utilization of cross-company waste heat, where waste heat that cannot be used internally is used by third parties in commercial or residential buildings, is a promising solution to increase energy efficiency and reduce CO_2 emissions. The most economically feasible utilization of waste heat requires spatial proximity of the waste heat source and demand, and heat recovery or heat displacement through the use of heat exchangers is the most efficient and simplest technological approach. Thus, not only the technology of energy creation as such should be considered, but also the context from which the technology originates. The last finding would be:

6. The Utilization of Waste Heat for District Heating from a Glass Factory in Kaposvár Promotes Energy Efficiency and Sustainability.

6. Topic related publications

Radtke, U. (2022). K-MEANS CLUSTER ANALYSIS OF HOURLY MEASURED POWER DEMAND IN THE DISTRICT HEATING SYSTEM IN KAPOSVÁR. Regional and Business Studies (2022) Vol 14 No 1, 57–72. https://doi.org/10.33568/rbs.3343

Radtke, U. (2022). A Brief Literature Review of Structuring District Heating Data based on Measured Values. The Ohio Journal of Science, 122(2), 75–84. https://doi.org/10.18061/ojs.v122i2.8845

Radtke, U. & Kaempf, D. (2021). RELATION BETWEEN DISTRICT HEATING SYSTEM OF KAPOSVÁR AND SUSTAINABILITY: in Challenges in the Carpathian Basin - Global Challenges - Local Answers. Interdependencies or Slobalisation?, 1197–1215. http://csik.sapientia.ro/en/news/global-challenges-local-answers-interdependencies-or-slobalisation

Uwe Radtke & Doreen Kaempf (2021). Relation between Urban energy systems - analysis of smart metering district heating system of Kaposvár and sustainability. IOSR Journal of Environmental Science, Toxicology and Food Technology, 15(3 Ser. II), 59–65. https://www.iosrjournals.org/iosr-jestft/papers/Vol15-Issue3/Series-2/G1503025965.pdf

References

Andrews, Dave. 2009. “"Carbon Footprints of Various Sources of Heat - Biomass Combustion and CHPDH Comes Out Lowest " - William Orchard.” *Claverton Group*, February 7, 2009. https://claverton-energy.com/carbon-footprints-of-various-sources-of-heat-chpdh-comes-out-lowest.html.

Brown, Steven R. 1996. “Q Methodology and Qualitative Research.” *Qual Health Res* 6 (4): 561–67. https://doi.org/10.1177/104973239600600408.

Fink, Arlene. 2020. *Conducting Research Literature Reviews: From the Internet to Paper.* Fifth edition. Los Angeles: SAGE.

Guba, Beate. 2008. “Systematische Literatursuche.” [Systematic literature search]. *Wiener medizinische Wochenschrift (1946)* 158 (1-2): 62–69. https://doi.org/10.1007/s10354-007-0500-0.

Hair, Black, William & Anderson. 2010. *Multivariate Data Analysis.* 7th ed. Upper Saddle River NJ: Prentice Hall.

IEA. 2022. “District Heating.” *IEA*. https://www.iea.org/reports/district-heating. Accessed December 20, 2022.

Jain, Anil K. 2010. “Data Clustering: 50 Years Beyond K-Means.” *Pattern Recognition Letters* 31 (8): 651–66. https://doi.org/10.1016/j.patrec.2009.09.011.

Lajos Kerekes. 2022. “Assessment of Price Regulation in Hungarian District Heating Systems - REKK.” Accessed May 30, 2022. https://rekk.hu/analysis-details/285/assessment-of-price-regulation-in-hungarian-district-heating-systems.

MacQueen, J. B. 1967. “Some Methods for Classification and Analysis of Multivariate Observations.” *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability* 1: 281–97. https://projecteuclid.org/Proceedings/berkeley-symposium-on-mathematical-statistics-and-probability/proceedings-of-the-fifth-berkeley-symposium-on-mathematical-statistics-and/toc/bsmsp/1200512974.

Moher, David, Alessandro Liberati, Jennifer Tetzlaff, and Douglas G. Altman. 2009. “Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement.” *PLoS medicine* 6 (7). https://doi.org/10.1371/journal.pmed.1000097.

Radtke, Uwe. 2022. “Unbiasing Science by Using Search Engines and Structured Methods – Analyzing or Measuring Scientific Output.” *IJARBM* 3 (1): 19–28. https://doi.org/10.51137/ijarbm.2022.3.1.3.

Southeastern Oklahoma State University. 2022. “Gain a Better Understanding of Consumer Behavior.” Accessed March 20, 2022. https://online.se.edu/articles/mba/gain-a-better-understanding-of-consumer-behavior.aspx.

Sue Ramlo. 2016. “Centroid and Theoretical Rotation: Justification for Their Use in Q Methodology Research.” *Mid-Western Educational Researcher* 28 (1). https://www.mwera.org/MWER/volumes/v28/issue1/v28n1-Ramlo-SPECIALIZED-RESEARCH-STATISTICAL-METHODS.pdf. Accessed December 17, 2022.

Valenta, A. L., and U. Wigger. 1997. “Q-Methodology: Definition and Application in Health Care Informatics.” *Journal of the American*

Medical Informatics Association : JAMIA 4 (6): 501–10. https://doi.org/10.1136/jamia.1997.0040501.

William Stephenson. 1955. “The Study of Behavior: Q-Technique and Its Methodology.” *American Journal of Sociology* 61 (2): 167–69. https://doi.org/10.1086/221704.