BACHELOR THESIS

Visualizing electricity consumption at a country scale

Rainer Pits
July 13, 2021

Supervisors:
dr. A. Kamilaris
dr. C.M. Epa Ranasinghe

Creative Technology
Faculty of Electrical Engineering,
Mathematics and Computer Science
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

UNIVERSITY OF TWENTE.

Abstract

Electricity use has come to play a crucial role in modern life, but it is also a major contributor of CO2 emissions. Considering its importance, many communities around the world also lack proper access to electricity and are therefore denied numerous opportunities. Although policy makers hold the power to address those issues on a national scale, domestic electricity consumption is a hugely complex topic to tackle where numerous parameters such as climate and demographic characteristics need to be considered. Data visualization methods can provide an effective method for intuitive analysis, but they face the issue of handling very large datasets. Moreover, methods for visualizing electricity consumption at a country scale are not well researched. The work of this project attempts to determine some of these methods by considering the electricity consumption in the Republic of Cyprus, using data collected by the Social Electricity project along side data for housing density, temperature, and expenses on utilities in the country. The results are four data stories that depict domestic electricity consumption in Cyprus with a variety of visualization methods that were determined through research and experimentation. The focus of these stories is on clearly delivering useful insights to the policy makers and people of Cyprus.

Table of Contents

Chapter 1: Introduction

1.1 Overview

Electricity is a fundamental commodity in the modern world. It is so ubiquitous that its presence in our lives is often completely taken for granted. Nearly all home appliances run on electricity and a most work could not be done without it. Yet for such an important commodity, the electricity sector faces severe challenges. First, electricity consumption is growing at a rapid rate. Global electricity production has risen from roughly 20000 TWh to 25000 TWh in the last 10 years [1]. In the wake of the climate crisis, this is a troubling development given that over half of the world's electricity is still produced with fossil fuels [1] and energy production accounts for roughly three-quarters of total greenhouse gas emissions [20]. The link between greenhouse gases and global temperatures is well established, with CO2 playing the largest role [21]. To become carbon neutral and effectively fight the climate crisis, nations worldwide must transition towards renewable energy sources and reduce electricity demand wherever possible. The latter of these solutions is increasingly relevant considering that green energy comes with its own environmental and social footprints. Constructing wind turbines and solar panels requires metals such as copper, which are carbon intensive to extract [2]. Moreover, many of the required resources are mined in the Global South, where local workforces are exploited for cheap labour [3]. Another challenge is that many people across the world still lack access to electricity. In 2016, 13% of the world population did not have access to electricity [4]. In a European context, where this project takes place, access to electricity is relatively guaranteed, but energy poverty is still prevalent [5]. This revelation is made even more troubling by the observation that many low-income households may lack adequate insulation, causing them to spend more electricity and money on heating and cooling [5]. These challenges correlate with the UN Sustainable Development Goals [6] and solving them is imperative for improving the wellbeing of human societies and the health of the planet. Nonetheless, approaching them is a challenge in its own right. Much of the weight falls on policy makers, who hold the power to shape the national electricity sector through public and economic policies. Yet for them to make effective decisions, they must know which issues are most problematic and urgent. When it comes addressing issues of electricity supply, getting a detailed understanding of a nation's electricity consumption is the first step.

Electricity consumption is by no means a simple topic to convey to policy makers (or most target groups). Simply knowing where and how electricity is consumed in a country does not give a full picture, especially when decision making is needed. There are numerous factors interacting with one another that make up a complete image of consumption. Geographic regions, economic sectors, demographic characteristics and weather conditions are just a few factors that dictate unique electricity consumption patterns. Due to the complexity of the topic, lengthy reports may not always be the best solution to present this information clearly. Although reports are still necessary, including visualization of the information can be very effective at generating understanding and encouraging exploration [7]. Electricity consumption at a country scale lends itself very well to geo-visualizations. These can be made interactive, allowing users to choose what they see by moving the map or selecting parameters and filters. Although visualizations have been proven to be effective in a variety of contexts, they have yet to become widely adopted [7]. For all their good qualities, it is also important to recognize the role of good design in making visualizations. Badly designed visualizations risk being confusing or at worst even misleading for viewers. As such, understanding the design conventions of good visualizations is imperative.

1.2 Electricity consumption in Cyprus

This paper addresses the electricity consumption in the Republic of Cyprus. Analyzing electricity consumption with the help of visualizations may be helpful for the authorities, scientists, and the public to help confront the problems Cyprus faces in the electricity sector. For one, Cyprus is still heavily reliant on the use of fossil fuels as their primary source of electricity despite the potential for solar and wind energy [8]. Nonetheless, the government of Cyprus has set goals for transitioning towards low carbon energy sources. A study analysing the costs of different energy scenarios for Cyprus found that the nation's power grid is highly susceptible to price volatility of fossil fuels, as those are largely imported [9]. The study identified renewable energy sources as clearly viable for future reductions in the price of electricity generation, with solar photovoltaics as the most competitive option. Although this can already contribute greatly to the nation's energy independence and help it meet EU goals, there is also the possibility of reducing electricity consumption in general, which can reduce the demand for imported fuel and carbon emissions. Decreasing electricity consumption is possible in a variety of sectors. An analysis of electricity consumption in Cyprus found that the commercial sector was the largest consumer of electricity, comprising over a third of the nation's total in 2004 [10]. A close second was the residential sector at approximately 30%. The remainder was made up of industry and agriculture. Research to increase the electricity efficiency in the commercial and residential sectors has priority, as they have the greatest potential for savings. In these sectors, weather has been identified as a critical factor when it comes to electricity consumption, which is tied to other factors such as income, quality of insulation and urban density [5]. However, how these factors affect electricity consumption in specific regions across the country has yet to be explored in detail.

Another problem that has received growing concern from the EU is energy poverty. Energy poverty is still prevalent throughout the low-income population of Europe and has been found to severely impact quality of life [5]. The authorities of Cyprus have defined energy poverty as the condition of consumers who may be in a difficult position because of low income, as evidenced by tax declarations in conjunction with their professional status, marital status and special health conditions, and therefore they are unable to meet the costs of the reasonable need of electricity supply, as these costs represent a significant proportion of their income [11]. A study of this in Cyprus found some notable contradictions [12]. On one hand, the government of Cyprus found that energy poor individuals made up only a very small proportion of the population. On the other hand, the study observed that Cyprus was performing poorly based on statistical indicators of energy poverty used across the EU. They conclude that government of Cyprus has taken appropriate action in consumer protection and energy poverty-mitigating measures but miss a coordinated approach to specify energy vulnerable groups and tackle the problem on a national level. Overall, to effectively respond to the problems of energy poverty and fossil fuel reliance, it is crucial to provide the authorities, as well as scientists and the public, a comprehensive overview of how factors such as demographic characteristics or weather may affect the electricity consumption of certain regions. This project seeks to provide that overview using geo-visualizations.

1.3 Project description

The aim of this project is to understand the domestic electricity consumption of Cyprus and how it changes through space and time. In doing so, various parameters such as weather conditions, seasons and demographic characteristics of the local population are considered. The goal is to develop explorative 2D or 3D geo-visualizations, aiming to identify patterns of electricity consumption and correlations of consumption. The data used in this project comes from the Social Electricity project that took place in Cyprus between 2011-2017. The data includes electricity consumption figures at the street level. As such, this paper addresses the questions:

- How to visualize electricity consumption at a country scale such that useful insights can be drawn from it?
- How to correlate electricity consumption at a country scale with different parameters such as weather conditions, seasons, and demographic characteristics of the local population?

Before addressing these questions with the project, an examination of the state of the art is needed.

Chapter 2: Ideation

In this chapter the ideation process of the project is summarized. As the starting point in the Creative Technology design process [35], the ideation phase relates to idea generation for the project, which can involve a variety of sub-processes such as examining stakeholder interests, examining related work, and tinkering. The results of this process are the product idea, interaction idea, and experience idea.

2.1 Core concepts

The first step in establishing a solid foundation for visualizing on a country scale is to consider the general principals of geo-visualizations established by cartographers. A paper on the challenges and opportunities for big data geo-visualizations by Robinson, et al. (2017) was used to establish the core concepts and broader research challenges which this project should address [33]. It should first be noted that it remains somewhat unclear whether the Social Electricity data can be classified as big data. The definition provided in the aforementioned paper also states that the definition is nebulous, but that most commonly big data is characterized by large volumes, high velocity, and a high degree of variety. As no tangible metrics are provided, and since the Social Electricity data is extensive in its coverage of Cyprus, it is reasonable to characterize it also as big data. These core concepts are instrumental in understanding the expectations of modern geo-visualizations and answering the research questions of this project.

The first core concept is the combination of place, space, and time. In this context space is an abstract concept that is given meaning by place and time. Geo-spatial big data allows for complex modelling and mapping of places, which can lead to complex representations of place and space. Combined with the temporal aspect, it can help cartographers tackle important issues through visualization. This complexity is addressed by the project through the inclusion of supplementary datasets, as well as the numerous parameters already included in the Social Electricity data. The next step from there is representing the data such that viewers can identify patterns and outliers, which forms the second core concept. Meeting this concept is less clear and will require experimentation with the data to determine visualizations work best.

Interaction is the third core concept and it determines how effectively geo-visualizations can convey their information. In the context of big data and the possibilities with modern technologies, interaction is necessary to allow users to navigate, search, filter, and compare using the data sources. The interaction needs for this project will primarily need to be met by the chosen software used for visualization. On a higher level, interaction becomes part of the user needs as it relates to usability. Users and usability form the fourth core concept, which strives to support a high degree of usability of the tools and techniques for mapping big data. As such, meeting this concept requires the design process to consider the needs of the intended users.

The fifth core concept is scale. In the context of geo-spatial big data, scale relates to challenges of showing detail on maps. Some degrees of abstraction will likely be necessary in this project, whether it means generalizing and simplifying or abstracting to smaller scales. In either case, the possible limitations need to be carefully considered and presented to the viewer.

Context is the sixth and last core concept. Contextual factors are a necessary consideration in the sense that mapping big data is inextricable from the relevance of real-world problems. As established before, electricity consumption is heavily tied to issues such as global warming and energy poverty. As such some important contextual factors for this project are the UN Sustainable Development Goals [6] and the Action for Climate Empowerment from the UNCCC [34]. These will need to be carefully considered throughout the project, such that it can provide the best possible aid in solving these global issues.

2.2 Research challenges

In their paper, Robinson, et al. (2017) also detailed several research challenges of big data geo-visualizations [33]. Many of these relate to software or analytical techniques which lie outside the scope of this project; however, some are also relevant for this project as they relate to the research questions. These research challenges go beyond the goals of this project, since they relate to current challenges faced by cartographers in visualizing geo-spatial big data. As such, the outcomes of this project can hope to provide a small contribution to tackling some of the challenges identified above.

The first research challenge is identifying effective methods for reducing complexity and creating overviews of geospatial big data. Traditional map-based overviews are often limited in their ability to accurately display the complexities of big data. The challenge is to create new approaches for generating overviews and determining which work best. This project tackles specifically the aspect of creating new approaches, as geo-visualizations of electricity consumption in relation to other parameters has received little attention from researchers. Good overviews are necessary for this project to provide quick understanding of the data and identify areas that should be examined closer.

The second research challenge is developing techniques for understanding change over time in geospatial big data. The challenge is in detecting change over time and determining patterns or subsequent changes. Algorithms already exist for performing such analysis, but visual interfaces are necessary to understand their results and guide future actions. The temporal aspect is an important consideration for this project as the Social Electricity data was collected over several years. Although the project does not utilize any algorithms for pattern analysis, showing changes in time clearly can help decision makers asses the impacts of past events and adjust their actions for the future.

The last research challenge covered in this project is developing spatiotemporal visualization methods for geo-spatial big data that support a variety of uses and users. This challenge relates mostly to creating meaningful visualizations that assist interpretation and build understanding. Moreover, these design choices need to be made with the needs of the target audience in mind. Although the primary target audience are decision makers, the visualizations are also intended for the public of Cyprus. Both groups have individuals with varying levels of expertise in a variety of domains. The challenge for this project, as well as for cartographers, is developing visualization solutions that can applied broadly for users with various backgrounds to understand.

2.3: State of the art

With a theoretical basis for making geo-visualizations of big data, it was necessary to also explore practical examples from recent work in the field. This section presents the state of the art in visualizing data on a country scale. The research papers and articles were found from Science Direct, Google Scholar, and Google Search. The Opower project was suggested by the project supervisor. Table 1 presents visualizations from the projects that are examined in this section.

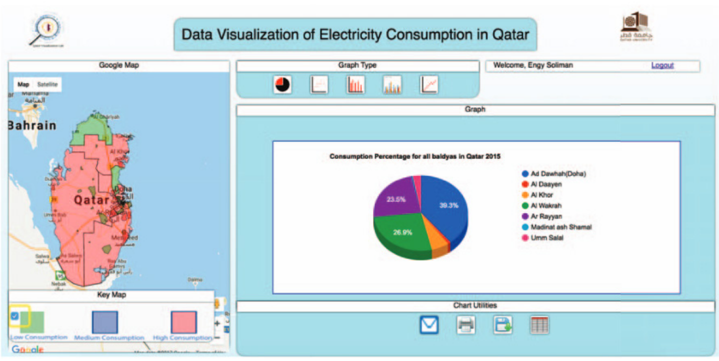
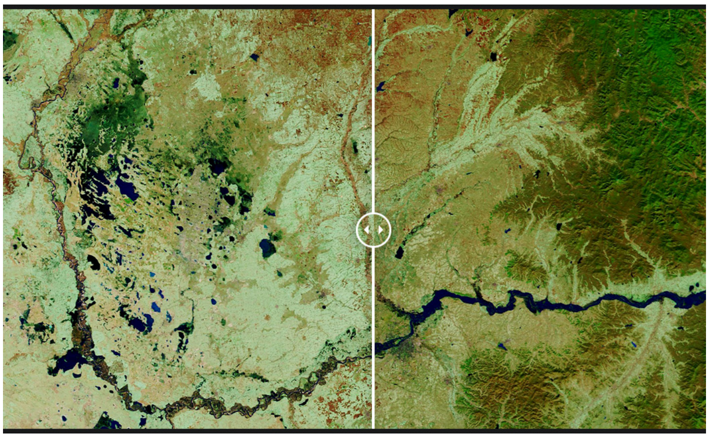
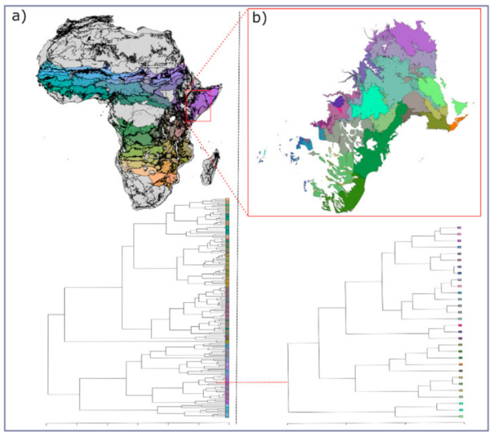
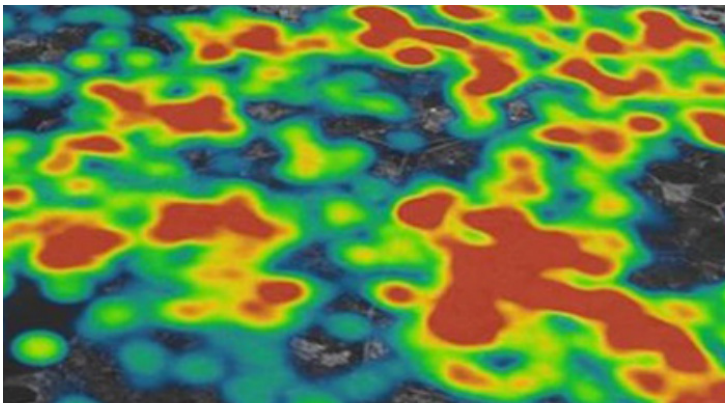
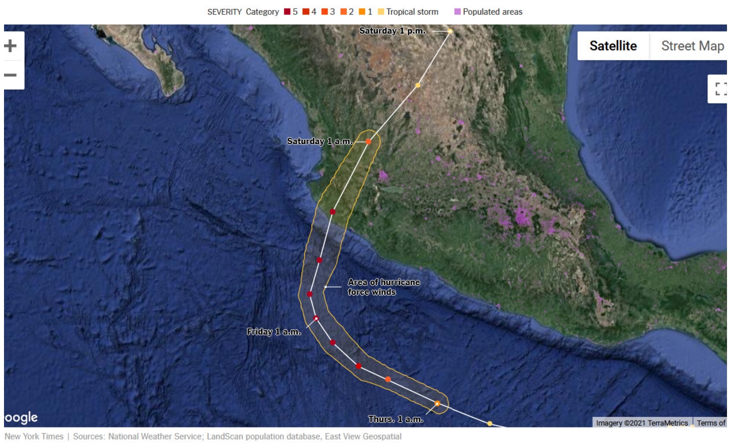
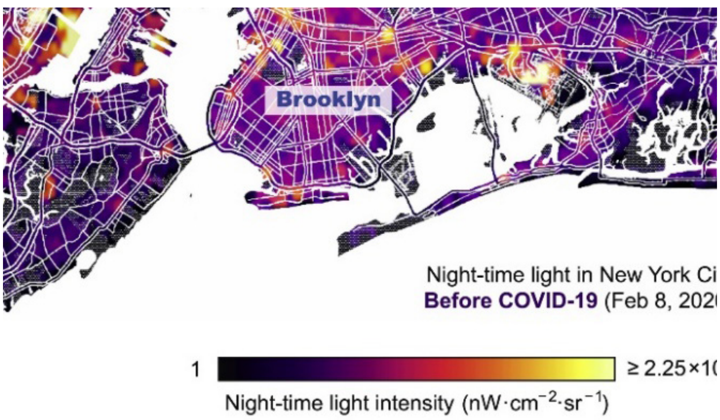
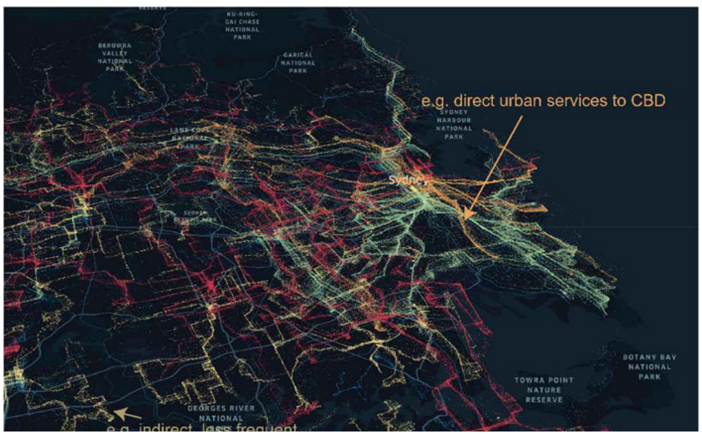
[14]	[15]
	
[23]	[24]
	

Table 1: Geo-visualizations from the state of the art

<table>
<tr><td>[16]</td><td>[22]</td></tr>
<tr><td></td><td>
</td></tr>
<tr><td>[25]</td><td>[17]</td></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Based on the literature examined it is clear that geo-visualizations of electricity consumption at a country scale are uncommon. Electricity consumption can be visualized through a wide variety of techniques, depending on the intentions and data [13]. Even though geo-visualization lends itself well to geographical data, many studies on this topic use techniques such as bar or line graphs, if any graphs at all. Although these are easier to develop and serve their purpose, they may hide details and correlations from viewers. Nonetheless, similar projects have been made in the past, where tools were developed with the purpose of visualizing the electricity consumption of a nation or territory. One project from Qatar was made with the Google Maps API [14]. The web-based tool was made by the researchers and allowed users to select the different municipalities and compare aspects of electricity consumption using various graphs. Another tool from the United Kingdom was made for the purpose of visualizing the relationships between multiple variables in the contexts of geography and scale [15]. As an example, the researchers used the country's energy consumption data. Perhaps the most relevant project is Opower, which is dedicated to promoting energy saving. Opower has used Power Map for Excel for visualizing energy consumption in specific areas for utilities companies [16]. Those visualizations showed how factors such as real-time communications could drive energy savings at peak usage times.

The projects highlighted above relate specifically to electricity consumption, however, geo-visualizations are used in a variety of scientific fields. Very prominent in this regard are visualizations for climate sciences, which often use satellite imagery in conjunction with sensor data to display weather patterns and behavior. An article on visualizing the size and strength of hurricane Patricia by the New York Times compiled several geo-visualizations to show the development of the hurricane [22]. Besides a video of satellite images showing the hurricane from space, other visualizations depict its path, increasing strength, wind speeds, and size. Furthermore, some visualizations are animated, making them more informative and engaging. In other examples, such as satellite imagery from NASA to visualize climate change, interaction is used to facilitate engagement [23]. Users can interact with sliders to view before and after images of certain regions which have changed over time due to climate change, urbanization, or other events. Besides studying the weather, geo-visualizations have also been used to study biodiversity [24], the impacts of COVID-19 on the electricity sector [25], and for designing intelligent transport systems in Sydney [17]. The latter two articles depicted particularly appealing geo-visualizations in terms of visual style.

For all the communicative and explorative benefits of visualizations, they hinge upon the abilities of their designer. Similarly, there are conventions and techniques that can unlock the full expressive potential of geo-visualizations, but it is also necessary to be mindful of what is being shown. A study of making attractive and unbiased visualizations noted that geo-visualizations can convey considerable amounts of context through attractive and familiar geographical patterns, but they can also distract from the actual data. Moreover, they can potentially mislead the viewers, for example, if data from a few localized points is extrapolated over far larger regions [7]. Aesthetically appealing visualizations have greater power to engage viewers [18], which is why it necessary to carefully consider how they present the data.

Geo-visualizations can also benefit from being interactive. Web-based tools or software can help audiences from non-scientific backgrounds to engage with information on their own terms and facilitate discovery [7]. Nonetheless, it is important to be mindful of when certain approaches may be more effective than others. For example, a study of Swiss electricity supply scenarios reported that the interactive web-tool had less user engagement and understanding from the public [19]. These results indicate that communicating certain scenarios may work better with static images of visualizations, rather than letting users discover them on their own. In the very least, it can provide new users with a foothold to understanding the visualization tool.

2.4 Tools

Having explored numerous examples from the state of the art, the next step was determining the tools which will be used for creating visualizations and managing the data. Understanding the possibilities and limitations of these software was necessary before making further planning for the visualizations. Each subsection will present a reasoning for why a software was chosen, as well as some observations from experimentation and tinkering.

2.4.1 Kepler.gl

Numerous software were identified in Chapter 2, which could be used for making geo-visualizations. These were Power Map, Cesium, and Kepler.gl. Between the three, the choice was made to use Kepler.gl as one of the visualization software for this project due to its speed, its good user interface and aesthetic visual style. For one, the speed at which Kepler.gl was able to visualize the large dataset of Social Electricity and cycle through different points in time made it great for usability. The user interface is relatively simple and has most necessary features included. This made it easy to learn and also more usable, although the list of features is quite limited. Lastly, Kepler.gl's visual style was attractive, which can play an important role when presenting the final visualizations.

Exploring the software was done through dummy data, which was smaller and available in CSV format, which is necessary for Kepler.gl. Additional exploration was research was done using the example projects provided by the software and later with the Social Electricity dataset. The software provides different styles for visualizations depending on the type of data. In the case of this project, with static coordinates, the point, hexbin, grid, cluster, and heatmap styles were suitable. In the point style, based on the selected data, the points can change color, radius, or outline weight. With the hexbin and grid styles, radius can be adjusted manually, and larger sizes will form averages of the postcodes covered. Besides that, both color and height be changed with the data, creating a 3D visualization. The heatmap style can be used to display areas of greater intensity using color, and radius can be changed manually. The cluster style forms aggregates of nearby points, where color and radius will change based on the average value.

Map layers underneath the data points can also be changed. There are several map styles included by default, but custom styles can also be imported via Mapbox. For visualizing change over time, the timestamps included in the dataset need to be applied as a filter. The timestamps need to be presented in a specific format in order to be readable by Kepler.gl. There are three possible formats – the one used in this project is mm:dd:yyyy hh:mm. The date in each timestamp was set to the end of the month since the Social Electricity data was collected at that time. The time was set to 00:01 for each timestamp. Further points are the save function, which operates through Dropbox and the export function, which has an option to export an image of the visualization, an HTML file, or the dataset.

2.4.2 Tableau

The option of using Tableau in this project was not considered from the start, as it is mostly used for other types of data visualization. Nonetheless, it is among the most widely used data visualization software available and has an extensive variety of features. Tableau also includes a geo-visualization feature, although it is much more limited than Kepler.gl. The most crucial limitations are the inability to easily cycle through time and make 3D visualizations. Tableau was chosen for this project later in the project, in the specification phase, when it became apparent that some geo-visualizations would benefit from supplementary graphs to convey information. Tinkering with the software, however, revealed many new sides to the data which were inaccessible with geo-visualizations. As such, the decision was made to use supplementary Tableau graphs in all the geo-visualizations made in Kepler.gl. Further information on the use of Tableau is provided in Chapter 4.

2.4.3 phpMyAdmin

The Social Electricity data was available as an SQL database, meaning that an SQL database manager was needed to compile the correct dataset in a table and export it in a usable format to Kepler.gl and Tableau. The options considered for use were PostgreSQL and phpMyAdmin. After tinkering with both, the choice was made to use phpMyAdmin, as it had a clearer user interface which made it easier to use and interact with the database. phpMyAdmin was run using XAMPP. Tinkering with phpMyAdmin was exploring the contents of the database and experimenting with SQL programming. In the process it was discovered that the database holds a number of houses parameter associated with each postcode. This was later adopted into the visualizations to measure housing density. phpMyAdmin was used several times throughout the specification phase as new data was found and added to the database.

2.4.4 Microsoft Excel

After exporting the dataset from phpMyAdmin as a CSV file, some additional adjustments were necessary in formatting the data. The choice for this was Microsoft Excel. The software was necessary for correctly formatting the time stamps, removing some outliers from the data, and adding new parameters, such as the urban and rural tags for districts.

2.4.5 Sublime Text 3

Sublime Text 3 was used for the HTML and CSS coding of the final website presenting the data stories. The website was hosted by the University of Twente.

2.5 Stakeholder requirements

The most prominent stakeholders impacted by the project are the decision makers of Cyprus and the people of Cyprus. The decision makers of Cyprus can be considered the intended audience of the project. They include the government of Cyprus and the Electricity Authority of Cyprus (EAC), the nation's sole electricity provider. Both stakeholders can use the visualizations to make decisions about the supply of electricity. Government priorities may vary, but will generally be focused on providing fair access to electricity across the country such that all households can meet a baseline standard of living. After that, they may be most interested in understanding patterns of electricity consumption to innovate the electricity grid. This can further improve standard of living, stimulate the economy, or benefit the environment. All three of these are additionally motivated by obligations of being a European Union member state. Due to the mix of backgrounds and domains present in government it is certain that both experts and non-experts would interact with the visualizations.

The EAC is likely to have different priorities in this regard. As a private company that holds a monopoly over the nation's electricity, they would seek to maximize returns on their investments. Without competition from other companies, their drive to innovate will also be low, so sustainable development may receive less attention. It is likely that mostly experts will interact with the visualizations from this stakeholder group.

A general approximation for the interests of the citizens of Cyprus may be that everyone has fair electricity access. Looking closer at individuals, however, will likely reveal a multitude of different interests. For example, individuals may be most concerned about their personal access to electricity, which will be common for other households in their area. If they experience problems, they will want their area to receive attention from decision makers, even though other areas may be worse off. Having said that, many citizens can also have opinions on issues of social justice or climate change and be willing to act on them. Furthermore, the people of Cyprus are the stakeholders who may be most affected by the project. Based on this analysis it is likely that many citizens would be interested in viewing and interacting with the visualizations. With this stakeholder group there will also be experts and non-experts as users, however there will likely be more non-experts.

Considering all user groups, the visualizations will need to accommodate for users from various backgrounds, many of whom will not be experts on the electricity situation in Cyprus. This is also in line with meeting the third research challenge of supporting a variety of users and uses. Meeting this challenge will mean that the visualizations should provide easily digestible overviews as well as allow for experts to draw deeper information. Considering the tools available, interaction is best suited for this problem. Interaction is best enabled by Kepler.gl, which allows for maps to be exported in which the users can interact with the data on the same level as designers. On a lesser level, all visualizations display numerical information about datapoints when hovering over them, which is also the case for Tableau. Besides interaction, the format in which the visualizations are presented can also play a crucial role here. The stories made with the visualizations should include explanations for variable names, concepts, and instructions for how to interact with the visualizations. Importantly, these may all also need to be translated to Greek, as many viewers may have difficulties reading English.

2.6 Design notes from previous work

Having explored the available tools and the stakeholder requirements, the state of the art from Chapter 2 can be examined closer in the ideation phase to identify useful design practices which can be used in the project. The most relevant design choices from the projects are provided below.

• In their study of Swiss electricity supply scenarios, Xexakis G. and Trutnevyte, E. [19] found that static depictions of scenarios may at times be more effective at facilitating learning than interactive web-based tools. It provides interesting considerations for this project, as interaction is a crucial component. One potential design choice then for this project may be the inclusion of a few presets which display specific scenarios, which the user can then begin to interact with and modify. Another design choice may be to add two geo-visualizations side by side to depict change over time.

• In the design of a web-based electricity visualization tool in Qatar, Soliman, E., et al. [14] combined maps and other graph types by placing them side by side. The user could select different municipalities from the map and compare them using the graphs on the side. This form of interaction would likely engage the user and encourage them to explore the data on their own. Unfortunately, the chosen software for this project does not allow for quite the same type of interaction. Similar functions are available in Tableau, such as placing a geo-visualization next to a bar chart, where areas can be selected and highlighted.

• The geo-visualization tool developed by Goodwin, S. et al. [15] in the UK is very relevant for this project because it focused on depicting relationships between multiple variables. As this project will also use multiple parameters to study electricity consumption, finding solutions for clearly visualizing relationships between two or more parameters will require effort. The tool developed by Goodwin, S. et al. utilizes symbols and mosaics as a solution. Although their solution is elegant considering the amount of information that is fit onto a single map, it is also quite hard to read. Since simplicity and readability are important factors to this project, an alternative solution will need to be found. Kepler.gl is mostly limited to a maximum of two parameters, as adding a third parameter make visualizations too difficult to read. Tableau is better suited for three parameters, but the results vary based on the chosen parameters. Seeing the limitations of the software and maintaining a focus on clarity, the choice was made to avoid combining more than two parameters unless the results are reasonably clear.

• The study by Ruan, G. et al. [25] on analyzing the impacts of the COVID-19 pandemic on the US electricity sector was inspiring for the project in the layout and design of their geo-visualizations. The researchers combined two visualizations of New York City's night-time light intensity from before and during COVID-19. Combined with their chosen color scheme (dark blue to bright yellow), they were able to depict a sharp change in electricity use. Both the color scheme and layout design can be used in this project.

• The New York Times article on hurricane Patricia [22] served as the primary inspiration for the design of the final presentation of the visualizations and arrangement into stories. The article combined visualizations of the hurricane from different perspectives interspersed with text explanations and analysis. The design inspired the final presentation of visualization stories to be made on a website and utilize both text and graphics.

• Additional visual design inspiration for use in Kepler.gl was taken from the research on design of intelligent transport systems for Sydney by Lock, O. et al. [17] and the examples provided on the website of Kepler.gl [36].

2.7 Ethical considerations

Before developing the finalizing the ideation process with the product, interaction, and experience ideas, several ethical dilemmas also require addressing. Some of these issues have been lightly touched upon in previous sections, but not examined in detail. Resolving these issues is necessary for the project to minimize any potential harm. The solutions of this section provide further guidelines for the design of the visualizations and data stories. The information provided here is summarized from a more extensive ethical report done on the project. An email can be sent to the researcher of this project for access to it.

2.7.1 Ethical dilemmas

The first ethical dilemma in this project relates to a question of fair representation and accessibility. There is much more information held in the data that could feasibly be covered in this project, so ultimately some parts will receive focus while others will not. This decision will be based mostly on research, with some input coming also from the supervisor and critical observer of the project. Nevertheless, there will be aspects that some viewers find important that are not covered, potentially leading them to poor conclusions. A worse case may be that the reasoning or information presented in the stories is incomplete or incorrect. It is important for the viewer to be aware of the biases and limitations in the visualizations and for them to have means of engaging with the visualizations to draw deeper meaning. Awareness is important, because it helps stop the spread of potential misinformation that can wrongly influence the viewer. Engagement with the visualizations is also necessary as it promotes wider understanding of the data and allows more information to be extracted from it. These aspects of discovery and awareness of limitations need to be enabled by the design. Nevertheless, they should also not reduce the credibility of the work. Loss of credibility could happen for example if the efforts to communicate give the impression that the researchers are ignorant or that not enough research has been done. In the worst case, this could render efforts made in the project ineffective.

A similar ethical dilemma is understanding the visualizations and stories. The viewer needs to understand what exactly is meant by the names of certain parameters or filters to prevent misunderstanding. Moreover, they need to know the basics of navigating Kepler.gl and Tableau so they can explore the data themselves. The visualizations and stories may be very beneficial for policy makers who would use it as a source to base their decisions on. If they do not understand what is being shown, there is a risk of inaccurate decision making when it comes to policies that could affect many people in Cyprus.

Lastly, there is a general moral concern that the outcomes of the project should serve as a driver of change. The analysis which is performed through the visualizations will likely result in several interesting observations, but it should strive to highlight problematic areas of environmental or social harm. The cost for pursuing this is that other interesting goals will not receive attention.

2.7.2 Ethical guidelines

Based on the ethical dilemmas that have been defined, three ethical guidelines were developed to direct the design process of the visualizations and stories. These are: design as a driver of change, ensure accessibility, and enriching the viewer's knowledge. The decision of designing as a driver of change was the easiest to arrive at, as fighting climate change and benefitting society are hard to argue. Furthermore, these goals are in line with several of the UN sustainable development goals [6]. Some complications may arise in evaluating the potential benefits of different topics and making the choice of which to analyze further. The choices will likely be difficult due to limited information. Furthermore, choices between social needs and environmental needs will be difficult. They will need to be addressed with the information that is revealed in the specification phase. Overall, however, this guideline stands to ensure that social and environmental issues will receive priority in analysis.

Ensuring accessibility is the second ethical guideline for this project. Accessibility is defined here as the viewer's ability to have access to, as well as understand and explore the data. For one, this guideline relates to the agenda of the UNCCC in providing adequate data for decision making regarding climate policy. In 2016 the UNCCC set out the target to provide information about climate relevant data to the public in the Action for Climate Empowerment [34]. Ensuring accessibility would then be in line with that target. Furthermore, this guideline is very relevant for this project because the data covered in the visualizations concern all of Cyprus. It specifically considers domestic electricity consumption, so it is tied directly to the people of Cyprus. In that sense the data belongs to them and they have a right to see any analyses made from it. This guideline also implies that most people should be able to understand these analyses. As such, it will be necessary to add explanations and tutorials and ensure that versions in Greek are also available. Additionally, color gradients have to be chosen which are also visible for individuals with common types of colorblindness.

The third ethical guideline in the development of this project is enriching the viewer's knowledge. This means that the visualizations and stories will strive to give the viewers as much new and relevant information as possible, as well as strive to facilitate learning wherever possible. This is related to the guideline of accessibility, but the focus is more so on learning and individual exploration. Clear naming, explanations, and tutorials are all aspects that help the viewer understand the visualizations and stories better. Furthermore, tutorials for Kepler.gl and Tableau can encourage the viewers to explore the data on their own, allowing them to get a much deeper understanding than they will from the stories alone.

2.8 Guidelines and ideas

This section presents the outcomes of the ideation phase. Having considered the challenges in modern big data geo-visualization, the state of the art, the stakeholder requirements, the available software, and the ethical dilemmas a number of guidelines could be developed which would direct the project towards the best possible results. Although many guidelines were initially established, overlap between them allowed for four core guidelines to be derived. These are:

- Accommodate viewers from different backgrounds.
- Design as a driver of change.
- Enrich the viewer's knowledge.
- Develop techniques for understanding change over time

With the guidelines in mind, a combination of brainstorming and tinkering resulted in the product, interaction, and experience ideas. These ideas will be explored further in the specification phase.

2.9.1 Product idea

The product of this project will be a website that presents the visualizations as a data story to the viewer. The data story will highlight the relationship between domestic electricity consumption in Cyprus and other parameters, such as the climate conditions. The story will be a mix of visualizations depicting relationships and text blocks that convey analysis or elaboration. A separate page will be available for instructions on how to use Kepler.gl and Tableau. There will also be an option to switch between Greek and English.

2.9.2 Interaction idea

The product aims to enable interaction wherever possible. Most of the interaction will be done on the visualizations. Geo-visualizations in Kepler.gl will allow users to navigate and zoom to certain areas, get numerical information for data points, adjust the settings of the visualizations, and export images, HTML files, or the data used in the visualizations. The visualizations in Tableau allow for data points to be highlighted and filtered, as well as shared as links or images. Both visualization types can be opened on full screen. The interaction on the website will be navigating between the pages.

2.9.3 Experience idea

The experience of using the site should be defined by exploration and learning. The viewer should understand well the tools made available to them and be able to use them with ease. The site can be used to become acquainted with electricity consumption in Cyprus or as a tool to look up or verify information. The analysis of visualizations will be limited and not make bold suggestions. They are more so meant to guide the viewer towards relationships in the data that they can explore further.

Chapter 3: Specification

As the second step in the Creative Technology design process [35], the specification phase is defined by the construction of several prototypes to test the ideas from the ideation phase. Several rounds of prototypes were created, evaluated, and improved upon. The design workflow was fluid and several times the failed results from prototypes required a step back to the ideation phase for adjustments. As extensive user testing was not a requirement in the project, evaluations were mostly made by the designer. Informal feedback was occasionally taken from acquaintances of the designer. Formal feedback on prototypes was given by the project supervisor during meetings.

Due to the nature of the project, certain deviations occurred from the Creative Technology design process. Most importantly, the specification was less focused on detailed planning than it was on discovering the information held in the data, as it was not immediately clear what information the data holds and how best to visualize it. As such, several prototypes were constantly being adjusted based on the information revealed by the data or feedback received. Several iterations of the dataset also occurred throughout the process, as formatting was adjusted, the data filtered, and parameters were added or removed. Because of this, the distinction between specification and the realization phases is also somewhat blurred.

The specification phase began with data collection. This was followed by the creation of a series of functional and experience prototypes, which were often combined as single visualizations or stories. The specification phase resulted in a detailed list of visualizations and stories, which were created and combined in the realization phase.

3.1 Data collection

The specification phase began with data collection. This process was a combination of research and exploration. Due to limited resources, free data sources had to be found for the supplementary parameters. This meant that certain data was not available or was much more limited compared to the Social Electricity data. The final parameters used in the dataset were: district, postcode, houses in postcode, kWh, temperature, Latitude, Longitude, utilities expenditures by area, district area and DateTime. Besides these, six additional parameters were left in the dataset, which were not found to be relevant enough for the stories, but can be explored by the viewer.

3.1.1 Social Electricity and database management

The electricity consumption data used in this project comes from the Social Electricity project [37], collected between January 2011 and August 2016 from domestic electricity consumers. The highest granularity of the data is on the street level, however, coordinates for the streets are not provided. Coordinates are provided for postcodes, so the electricity consumption of postcodes is used. To clarify, a postcode in Cyprus is used for all buildings in a given area and it is comprised of four numbers. In total, the Social Electricity project retrieved the electricity consumption figures of 1127 postcodes out of the 8999 throughout the country. Not all the postcodes are included in every month, but their number stays consistent throughout the time period to provide stable readings. Moreover, the examined postcodes are distributed throughout the country, with greater densities in larger cities. Every postcode also has a mark for the district it belongs to. In total there are five districts, which correspond to the names of the largest cities in the country. The electricity consumption was recorded bimonthly and divided in two to get readings for individual months. The unit of measurement was kilowatt hours (kWh). The kWh readings in the database are normalized, meaning that the electricity consumption of all households in a postcode was added up and divided by the number of households.

The number of households data also deserves mention in this section. It was not used until late in the specification phase, when it was found in the database and included for additional context. The results were quite insightful, so the data was eventually made into a parameter on its own, termed housing density. This parameter is especially valuable because it is as detailed as the electricity consumption data. Since this parameter was added late, research about it was also lacking. Nonetheless, the concept did occur in the

research for the climate parameter as the phenomenon of urban heat bubbles, which is the overheating of densely packed urban areas [5]. Moreover, a study of Danish households found that the highest electricity consuming households live in large, single family detached houses [29]. These houses can be expected to lie outside of densely packed urban areas and therefore belong to postcodes with lower housing density.
The electricity consumption data was inside an SQL database. The data was well organized when it was received and included a large variety of different variables, most of which were unnecessary for this project. Through simple SELECT and JOIN commands the necessary data was compiled in a table, which was then exported as a CSV file. The columns selected for the table were: district, postcode, houses in postcode, year, month, kWh, Latitude, and Longitude. Before uploading the data to Kepler.gl, some additional formatting needed to be done in Microsoft Excel for the dates and the names of parameters.

3.1.2 Geographical and climate data

Knowledge of the geographic and climate context in the Republic of Cyprus is necessary for understanding the physical conditions that shape domestic electricity consumption. Climate may often be linked to geographical features, however both factors can also influence electricity consumption independently. The geographical features relevant to this project are coastal and mountainous areas. Certain coastal areas in Cyprus are famous for their beaches and are therefore the largest tourist attractions in the country. In the peak tourism seasons, these areas may consume more electricity than other areas. On the other hand, mountainous areas are more sparsely populated, have more mild weather conditions, and accommodate less tourists than other areas. As a result, residents in mountainous areas are likely to consume far less electricity. Mountainous areas could be identified on the satellite map provided in Kepler.gl, depicted clearly in dark green. Hotspots of tourism were identified from simple online searches. The three largest hotspots are the coastal cities of Paphos and Limassol, as well as the Cape Greco peninsula. Although electricity consumption from hotels or resorts is not included in the dataset, these areas provide a lot of alternative lodging for tourists, such as homestays with Airbnb. The electricity consumption habits of tourists in homestays may be higher than regular households. Since the electricity consumption of houses or apartments accommodating tourists are included, the impact of tourism on domestic electricity consumption could be observed.

Like geography, the effects of climate conditions on domestic electricity consumption may be very significant. Climate does vary across different parts of Cyprus, but it can also vary significantly due to larger climatic patterns. First there are different seasons. In the summer, temperatures in Cyprus are typically around 30°C in the capital city of Nicosia [26] but can go much higher up to 46°C [27]. Warm summers mean that more electricity is spent on air conditioners. On the other hand, winter months are on average 13°C in Nicosia [26], so heating is required. The effects of seasons are generally predictable, however there are also possibilities for abnormal weather phenomena, such as heat spikes. In these cases, electricity consumption can be expected to be much higher also. There are many aspects of weather that may influence electricity consumption and analyzing how all of them influence electricity consumption is outside the scope of this project. Considering the different aspects of weather and the data most easily available, temperature was chosen. In terms of electricity consumption, temperature relates primarily to heating and cooling, which can both constitute a large portion of domestic electricity consumption.

Historical temperature data on the level of postcodes was not available, however it could be found for the major cities of Cyprus: Nicosia, Limassol, Paphos, Larnaca, and Famagusta. As each of these cities belong to separate districts, the temperatures recorded in the cities could be extrapolated to all the postcodes in the district. The small size of Cyprus means that significant variations in temperatures inside regions should not occur, however mountainous areas can be expected to be colder than coastal or flat inland areas. The historical temperature data from 2011 to 2016 was gathered from provided by the Cyprus Department of Meteorology [28] and arranged into a separate table in the SQL database, where it could be joined with electricity consumption data.

3.1.3 Demographic characteristics

There are numerous demographic characteristics that could be analyzed together with electricity consumption, however the most important one may be household income. Numerous studies conducted throughout Europe have shown links between income and high electricity consumption. The findings can, however, contradict themselves. A study of Danish households found that high-income households generally consume more electricity [29], which may also be the case in Cyprus. High-income households may have larger houses or electricity intensive features such as pools. On the other hand, low-income households can also consume excessive amounts of electricity due to reasons such as poor insulation or inefficient appliances Insulation can play a significant role in electricity consumption in the winter or summer months, as it can help maintain stable indoor temperatures without resorting to heating or cooling appliances. In impoverished areas throughout Europe, poor insulation has been shown to significantly increase electricity consumption and decrease standard of living [5][30]. Similarly, switching to electricity efficient cold appliances has been shown in Swedish households to significantly decrease electricity consumption, however they also come with higher price points [38]. A study in Denmark found that low-income households had significantly less flexibility to adopt more efficient appliances [29]. Households can have many reasons for not adopting efficient appliances, but high costs generally pose the largest barrier.

Detailed data for income on the postcode level is available for Cyprus, but was inaccessible due to financial restrictions. Instead, free data from the Statistical Service of Cyprus on household spending could be used [31]. This dataset provides the mean annual consumption expenditure of households per region, for both urban and rural areas. This distinction was not included earlier, but was later also added to the dataset for additional context. It also provides more detailed expenditure data, such as expenditure on utilities, which include gas, water, and electricity. Since the utilities expenditure directly relates to electricity consumption, it was chosen as the data for the demographic characteristics parameter. It can be used to investigate tp what extent spending on utilities correlates with electricity consumption. It is important to note also that electricity for domestic use is charged at the same rate across Cyprus [40]. The data for separating urban and rural postcodes was available on the website of Just About Cyprus [32]. Both datasets were added to the SQL database. After joining these with the table of electricity consumption per postcode, some additional formatting was necessary for both the expenditure data and timestamps before the CSV file could be imported to Kepler.gl and Tableau.

3.2 Prototyping

In the case of this project, the specification phase largely consisted of extensive prototyping. Since the data was not known before hand, establishing the functional specifications meant that it had to be explored for information until the best visualization methods could be found. The functionalities of the website were less relevant, as it was intended to be quite basic. The functional specifications were also much more related to answering the research questions and adhering to the second and fourth guidelines. The experience specifications were largely already established by the interaction and experience ideas. Nonetheless, in the course of prototyping, the experience specifications were also established in greater detail and required following the first and third guidelines. Having completed the final dataset, determining the functional specifications meant exploring the data and discovering what information it included, as well as how best to extract it. Since the data collection process was done over the course of three weeks, the exploration of some datasets began earlier than others. This process of exploration resulted in three rounds of prototypes, which mostly focused on functionality. These were presented to the project supervisor in biweekly meetings. Feedback from him, as well as the continuous exploration of new methods and combinations was used to improve upon the prototypes.

The first prototyping round was done with the electricity consumption, temperature, and income data. The electricity consumption data was initially visualized using dots of uniform radius, which change color based on their value. A base map to highlight the Troodos mountains in the middle of the country was created with Mapbox and imported to Kepler.gl as the “elevation map.” The elevation map was used as a base in all of the visualizations in the first prototyping round, which were all also geo-visualizations. The temperature data was harder to work with, since visualizing it together with electricity consumption resulted in geo-visualizations that were difficult to read accurately. The choice was made to use the hexbin map layer for the visualizations as it could show one parameter as the color of a column and another parameter as the height of the column. This map layer had a crucial flaw however, as it calculated the range of values differently, which reduced the functionality of the visualization. Nonetheless, since a better alternative could not be found, the hexbin visualization shown in Figure 1 was kept for feedback. Lastly, the income data was still being searched for at that point. The data that had been found came in the form of an interactive map in Arcgis. Since the data could not be imported on time, an attempt was made to make a static visualization by overlaying an image of the visualization from Kepler.gl on top of the Arcgis map [reference]. This is shown in Figure 2. In the feedback round, the income visualization was suggested to be made interactive and fully in Kepler.gl and the temperature visualization received several suggestions for possible fixes. The electricity consumption visualization was satisfactory.

Figure 1: Hexbin map of temperature and kWh

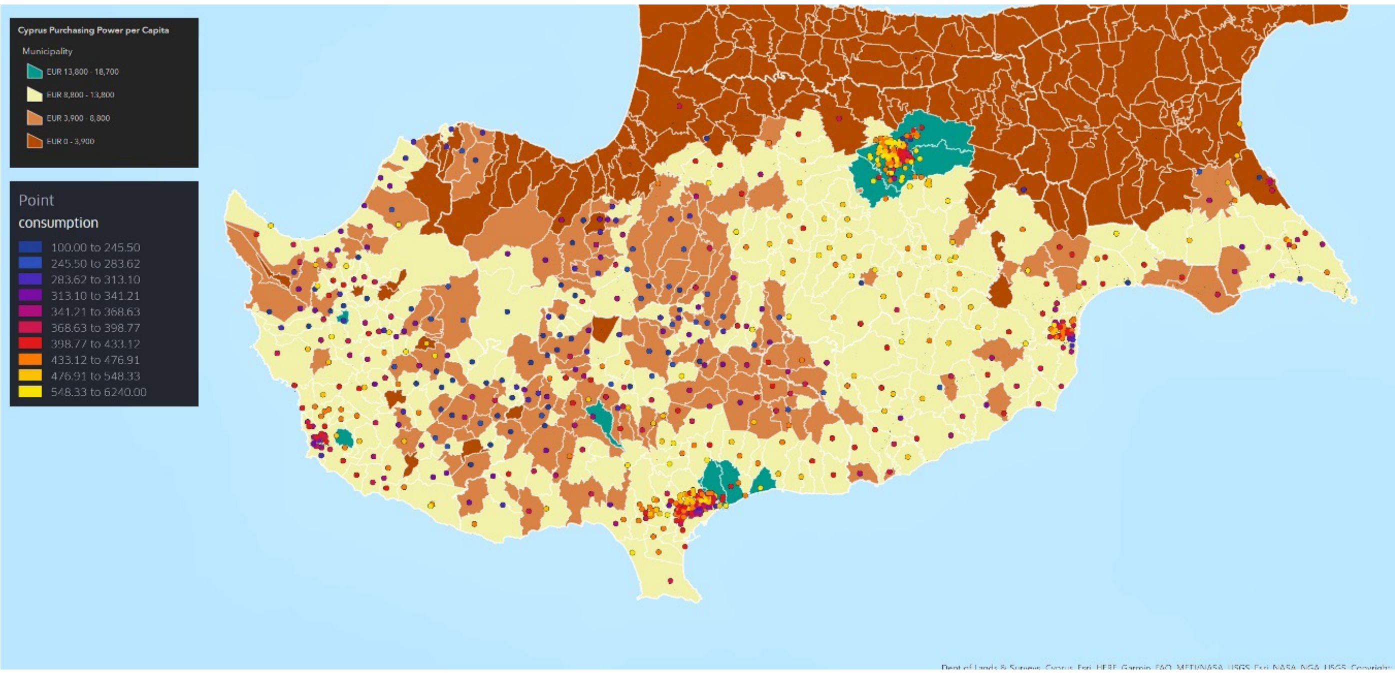


Figure 2: Electricity point map on Arcgis map of purchasing power

For the second round of prototypes, the most important factor was to find a clear way to depict the temperature and electricity data together. The software issue with calculating the range meant that only the point layer could be used. Although both the point color and radius can be connected to data, the results were difficult to read. A subsequent attempt was to create an index value by dividing the electricity consumption with temperature to get a single column of values. This was made into a prototype that can be seen on Figure 3. The Arcgis income data was abandoned due to its high price point. Instead, the expenditure data was found and integrated into the dataset. It was believed that this data would also be unusable with the hexbin map, so a visualization of index values was also created. Over the course of some experimentation, the housing density data was also found from the Social Electricity database. Visualizing that together with a heatmap of electricity consumption proved very clear and insightful. Similar methods could not be used for the temperature and expenditures data, however. In the feedback session, it became clear that the index values could not be used because they did not present the data accurately. The housing density visualization received positive feedback. A suggestion for all parameters was made to utilize a few scatter plots with Tableau alongside the geo-visualizations to display the data with greater accuracy.

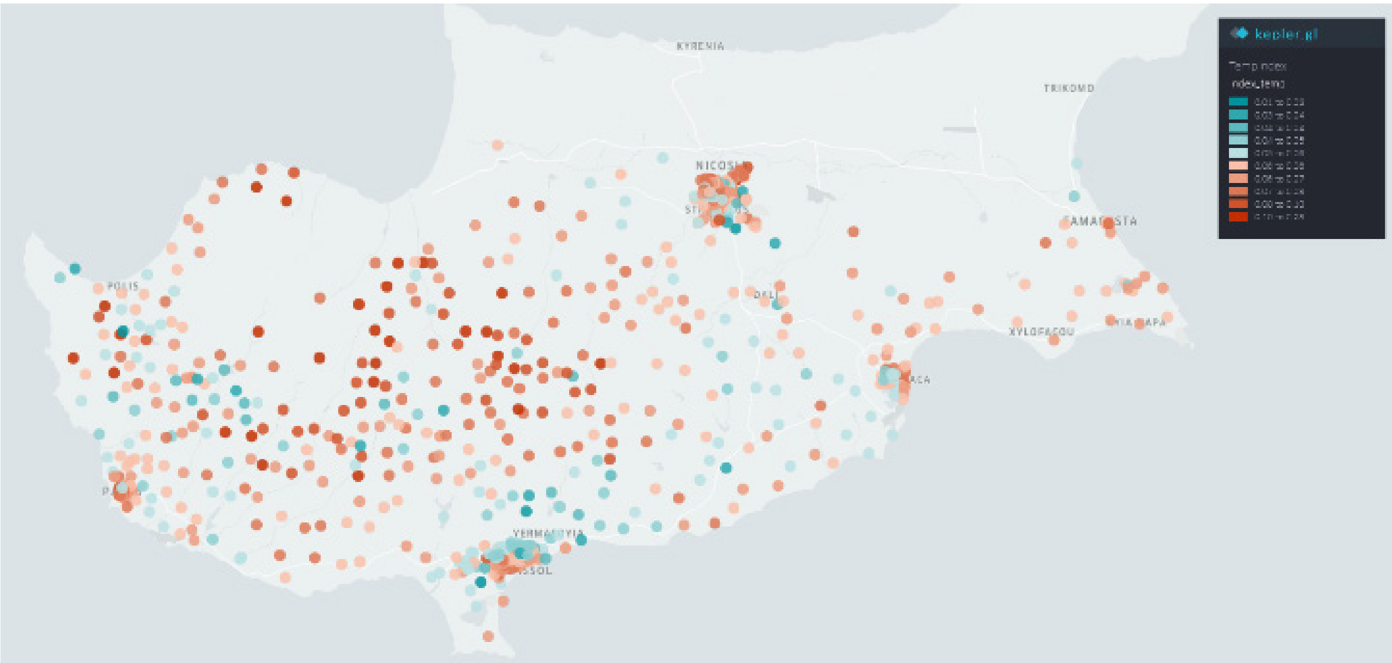


Figure 3: Index graph showing the quotient of kWh and temperature in Celsius

The third prototyping round was the most productive, which was largely due to the inclusion of Tableau in the project. So far, all the prototypes had been geo-visualizations and alternatives had not been explored. The value of utilizing scatter plots, bar graphs, and line graphs became clear immediately, as they revealed many insights that could not be found on the geo-visualizations. Numerous visualizations were made with Tableau and a minimum of two were chosen for each parameter. A prototype of the website was also made, and the greater number of visualizations also meant that a single data story become too lengthy. An alteration was made from the initial product idea and the website was divided into different pages, such that each parameter had a smaller, separate story on its own page. Besides this, a good method was found for making the geo-visualization with the temperature data that involved having two separate visualizations side-by-side. This was inspired by an example that was found in the sample visualizations on the Kepler.gl website [36]. Further experimentation with the utilities expenditure data revealed that the hexbin map layer was suitable for it. The dots electricity consumption geo-visualization also received some minor adjustments and another geo-visualization was created with the hexbin layer to depict electricity consumption in the rural and urban areas of districts. The feedback for this prototyping round was generally positive since a lot of progress was made. Comments were made for improving certain aspects in all of the data stories. The knowledge gained from the rounds of prototyping and the feedback sessions became the basis for establishing the functional and experience specifications.

3.3 Functional specifications

Generally, the functional specifications are a list of functions that a system must perform. In the case of this project, the functional specifications were a list of visualizations for each story and a description for their intentions. These are listed below.

The first data story includes a total of six visualizations. It focuses on the domestic electricity consumption of Cyprus, how it relates to the country's geography, and how it has changed between January 2011 and August 2016.

• The first and second visualizations are geo-visualization that are intended for identifying areas of electricity consumption and geo-spatial consumption patterns over time. They will be a point and a heat map made in Kepler.gl. The elevation base map will be used on both.

• A third geo-visualization is also needed to depict where the urban and rural areas of districts are located on a map. This can be done with a hexbin layer on a light or dark base map in Kepler.gl. Functionally, the hexbin layer is chosen because its different range calculating method can reveal additional insights.

• The fourth visualization will made in Tableau. It will show six smaller geo-visualizations of the average electricity consumption of postcodes per year.

• The fifth and sixth visualizations will be line or bar graphs made in Tableau. These will show the average changes in electricity consumption by year and month in the districts of Cyprus.

The second data story focuses on the relationship between domestic electricity consumption and the housing density in Cyprus. Geography and changes over time are also presented. This data story includes three visualizations.

• The first is a geo-visualization that combines electricity consumption on a heat map and housing density on a point map, where the points are black and the radius changes. This serves the function of correlating the electricity consumption of specific areas or individual postcodes with the housing density.

• The second and third visualization are made in Tableau and are intended to work together. One is a scatter plot that has kWh on one axis and housing density on the other. Each postcode has one dot on the plot and average values are used. The other is a geo-visualization that is intended as way to interact with the scatter plot. Its functional intention is to highlight individual or groups of postcodes and show them in greater detail on the scatter plot.

The third data story focuses on the relationship between temperature and domestic electricity consumption in Cyprus. Geography and changes over time are also presented. In total, this story includes six visualizations.

• The first is a Kepler.gl geo-visualization that is split in two. One side has a point map of temperature, where color changes. The other side is a point map of electricity consumption, similar to that in the first story. The function of this visualization is to provide insight to how electricity consumption in certain areas changes based on the temperature. By default it is on a dark base map.

• The second and third visualizations are made in Tableau. They are a similar combination of geo-visualization and scatter plot that was in the second data story. The scatter plot presents the data, while the map is used to highlight individual or groups of postcodes. The scatter plot shows how the electricity consumption of each postcode changes relative to temperature.

• The third is a Tableau visualization that is composed of several smaller geo-visualizations. Each geo-visualization shows the average electricity consumption of postcodes in bimonthly or seasonal periods. This visualization does not show temperature and thus works in conjunction with the fourth and fifth visualization. Its function is to show change over time from a different perspective compared to the first geo-visualization in the story.

• The fourth and fifth are Tableau visualizations that use a mix of bar and line graphs to show both electricity consumption and temperature changing over time. One shows the change over the whole period of 2011 to 2016 in monthly readings in different districts, while the other shows the average monthly changes.

The fourth data story depicts the relationship between electricity consumption and the expenditure on utilities in Cyprus and how it changes over time. This story includes three visualizations.

3.4 Experience specifications

The experience specifications were largely already defined in the experience and interaction ideas. Nonetheless, the process of prototyping revealed some additional details that improved the experience of using the website. Most important was determining the layout of the website and adding instructions or explanations for the visualizations.

The division of the data stories into separate pages was done in the interests of giving the information to the viewer in smaller, more digestible chunks. The intention is that the viewer spends more time with each data story and they understand it better. The visualizations are arranged in each story such that they build information on top of each other. This means that the visualizations will not jump between different styles constantly. The analysis section for each story is included at the bottom of the page to avoid clutter. The analysis is a few short paragraphs that can be read in a few minutes. It highlights only the most important insights in each story. Each visualization is accompanied by a title and a short description of what is being shown. Navigating between the pages is done with a navigation bar which is always present. A Home page will be added which introduces the project and provides a short tutorial for how to interact with the visualizations. Another button on the navigation bar can direct the viewer to a version of the page in Greek or Turkish.

The visualizations will encourage interaction. The geo-visualizations made in Kepler.gl will be fully interactable and the viewer can save or share their edited visualizations. Refreshing the page will reset the changes. The Tableau visualizations have less possibilities for interaction, but encourage it in other ways, such as combining visualizations together into dashboards. For example, the combination of geo-visualization and scatter plots is designed with interaction in mind. The Tableau visualizations can also be shared and all visualizations can be opened in full screen. Instructions for interactions for each visualization will be written next to them.

Chapter 4: Realization

The realization phase is the third and last step in the Creative Technology design process [35]. The specification phase ended with detailed feedback from the project supervisor about the last prototypes, after which a set of detailed functional and experience specifications were established. Knowing these, the final versions of the visualizations and stories could be made. Where possible, they include tips for what interactions are possible and how to do them. Furthermore, they include legends and colorblindness friendly color palettes. The stories were then integrated to the website, which was also improved upon based on guidelines from the ideation phase. With the finished stories, some analysis could also be made, which was added to the respective pages. This section presents the separate visualizations as well as the finished web design.

4.1 Electricity consumption

The first data story focuses on how electricity consumption varies over time and location. Location is shown using the first two visualizations with the use of geographic maps and the areas of districts. These maps can also be used to examine changes over time in greater detail, using the time slider. The Tableau visualizations present broader trends in changes over time.

Figure 4 depicts the electricity consumption of postcodes across Cyprus and how they relate to geography. The map on the left shows a heatmap of areas with more intense electricity consumption. Bright areas can occur either from higher electricity consumption or from postcodes being close. The map on the left shows the consumption of individual postcodes. A tip was added alongside this visualization to switch between the geographic (written as Elevation in the visualization) and dark background map.

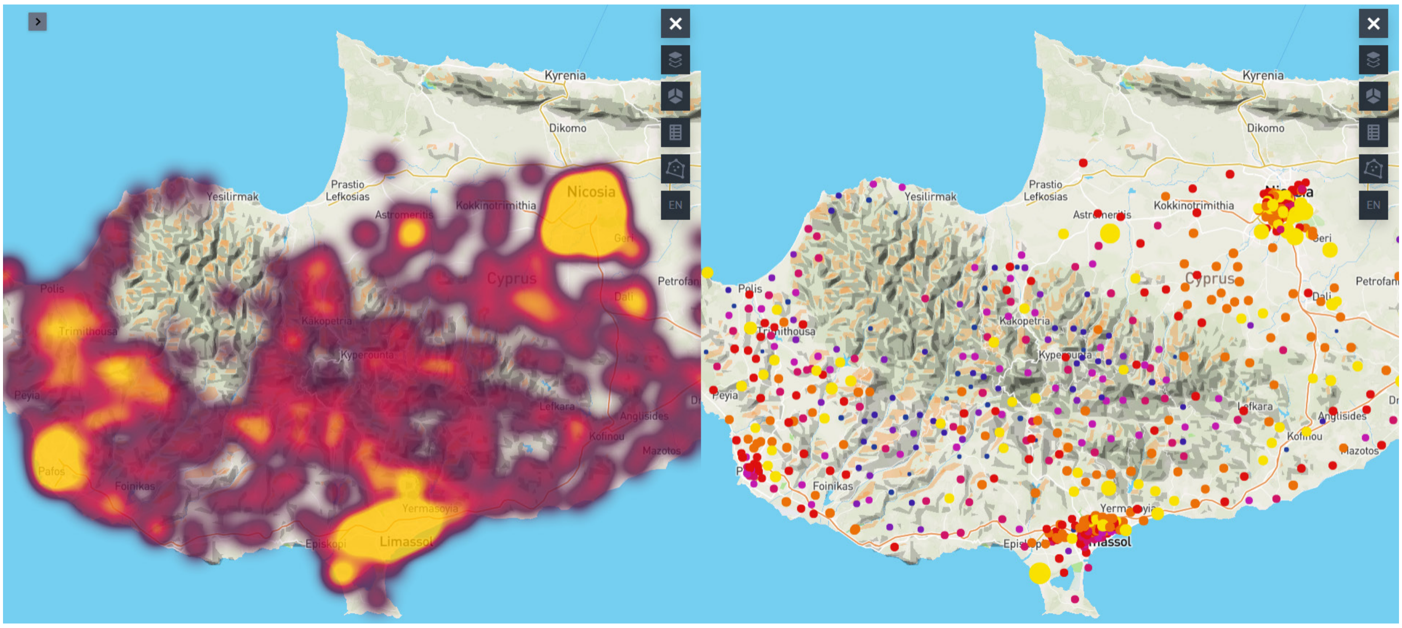


Figure 4: Electricity consumption across Cyprus

Figure 5 depicts the electricity consumption of postcodes, with color varying on the district area. Due to the hexbin map layer, some postcodes are combined in pillars and their average kWh readings are displayed. It is also important to note that the hexbin layer makes a new range each time the visualization refreshes. This means that the height of pillars corresponds to the range of kWh values that are being shown for the dates that are currently being viewed.

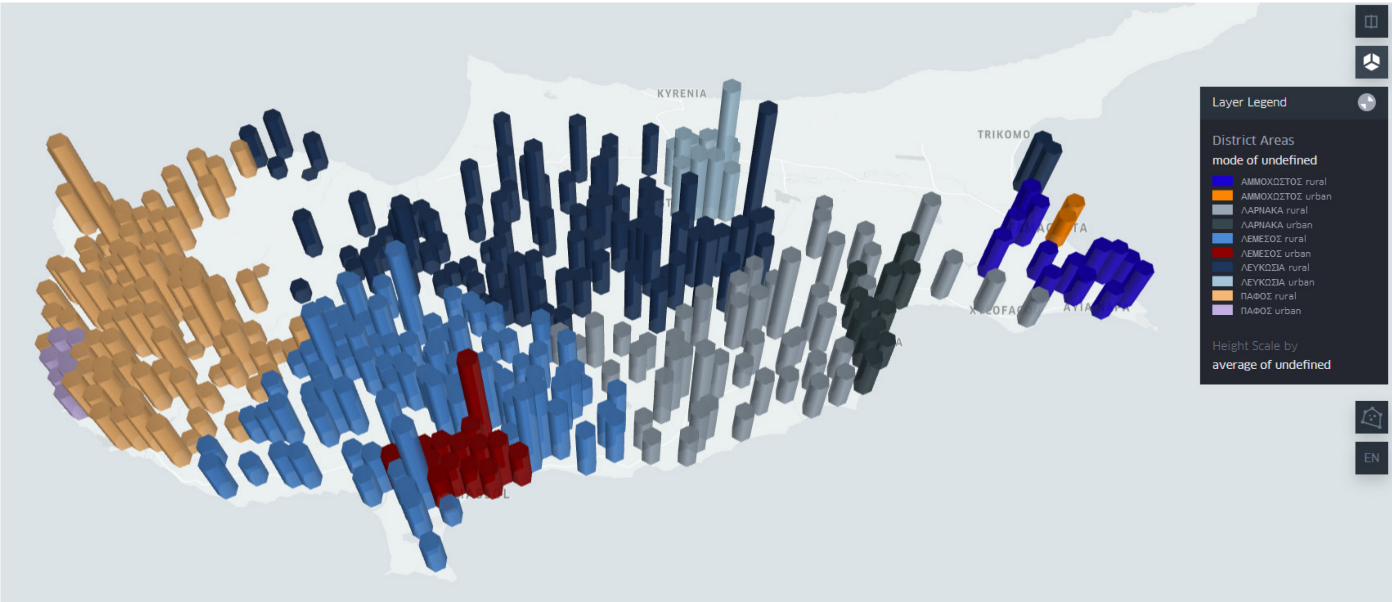


Figure 5: Differences between rural and urban areas in districts

Figure 6 presents the yearly changes in the average electricity consumed across Cyprus on a postcode level. On the website, Figure 6, 7 and 8 and displayed together under the title of Changes over time.

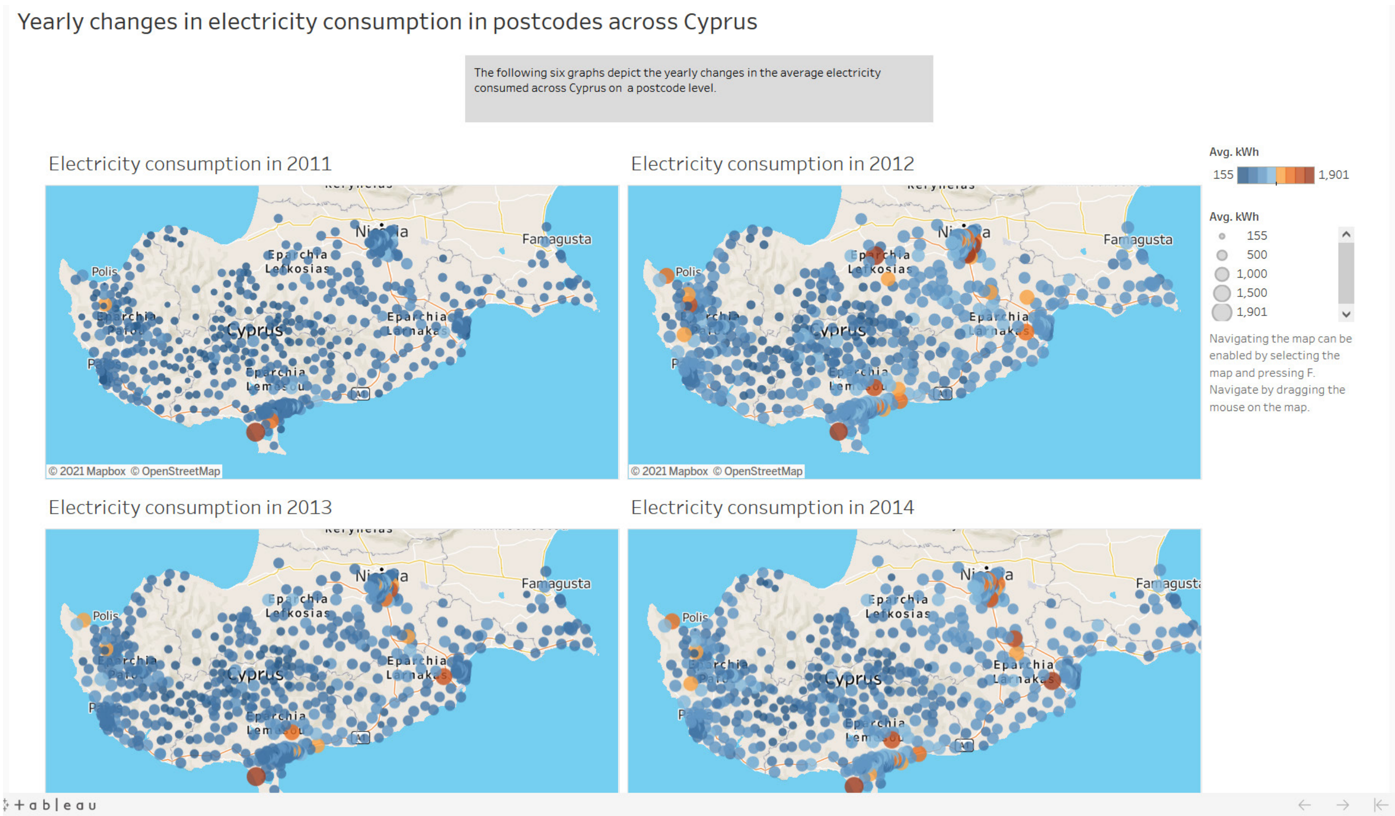


Figure 6: Yearly changes in electricity consumption in postcodes across Cyprus

Figures 7 and 8 are combined in a dashboard and present several line graphs depicting the average electricity consumption in district areas in monthly or yearly intervals. Interaction is possible by selecting and highlighting lines across both visualizations.

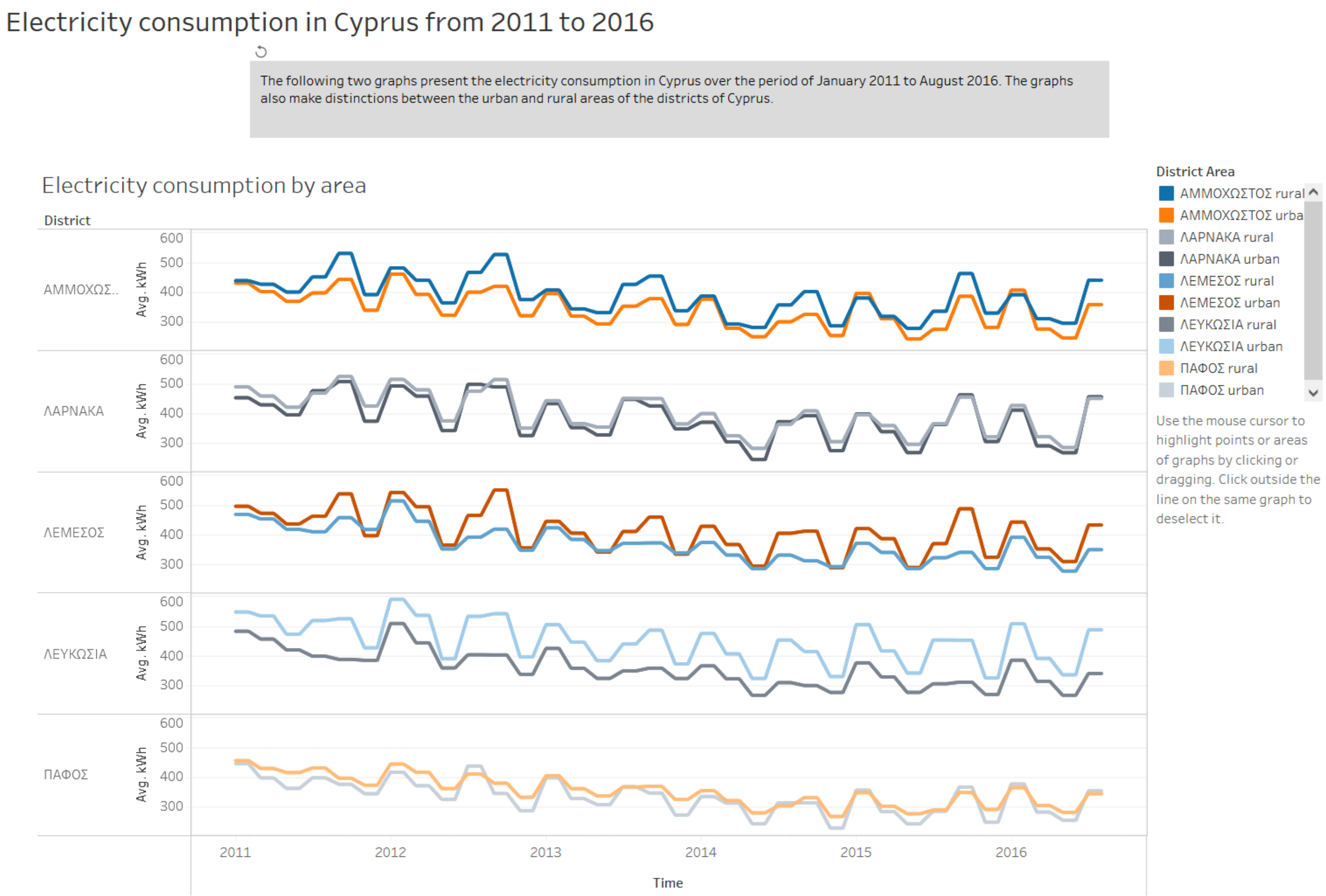


Figure 7: Electricity consumption in Cyprus from 2011 to 2016

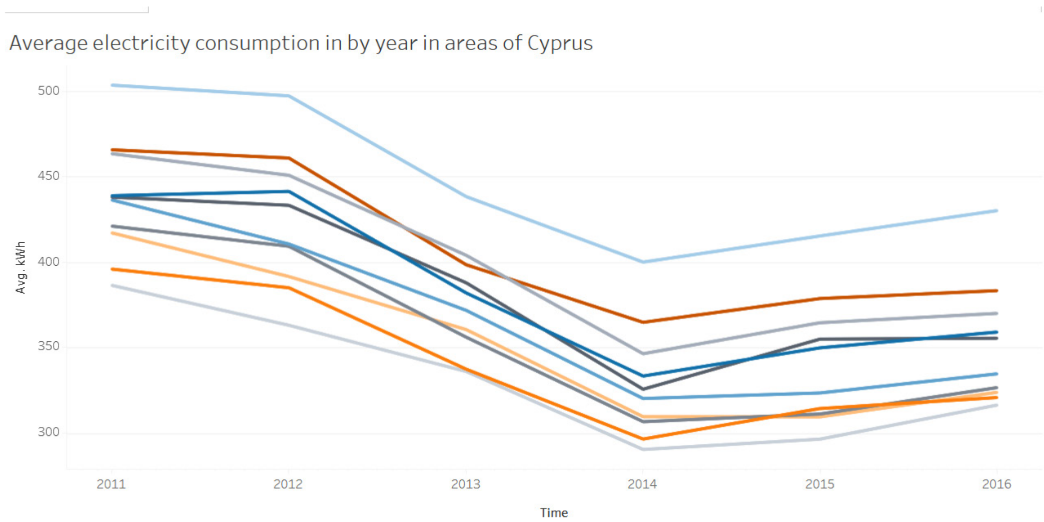


Figure 8: Average electricity consumption by year in the district areas of Cyprus

4.2 Housing Density

The second data story intends to depict the relationship between electricity consumption of postcodes and the number of households that were registered under it, which is used to describe the housing density. The two parameters are shown in a geo-visualization in Figure 9, which combines a heat map depicting kWh with points that vary in size based on the number of houses. This map can also be used to examine changes over time in greater detail, using the time slider. The Tableau visualizations present broader trends in changes over time of both parameters.

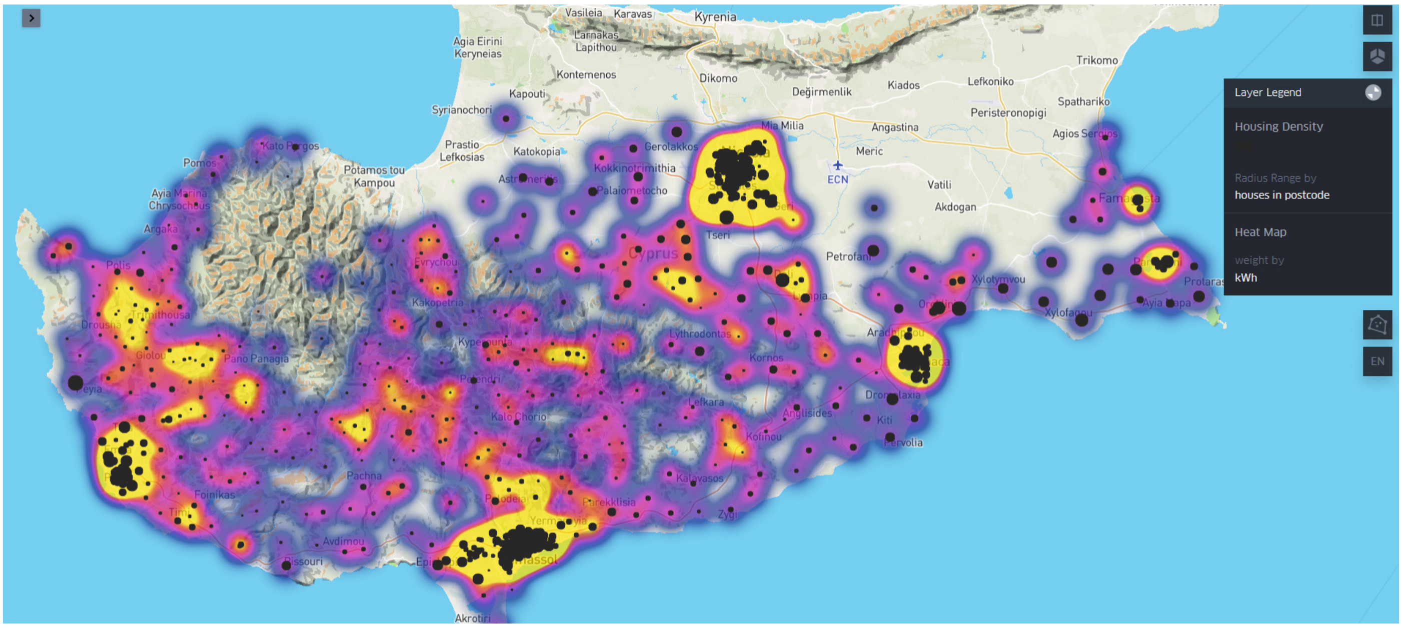


Figure 9: Electricity consumption and number of households at postcodes

Figure 10 presents a scatter plot of the average kWh and housing density of postcodes by year.

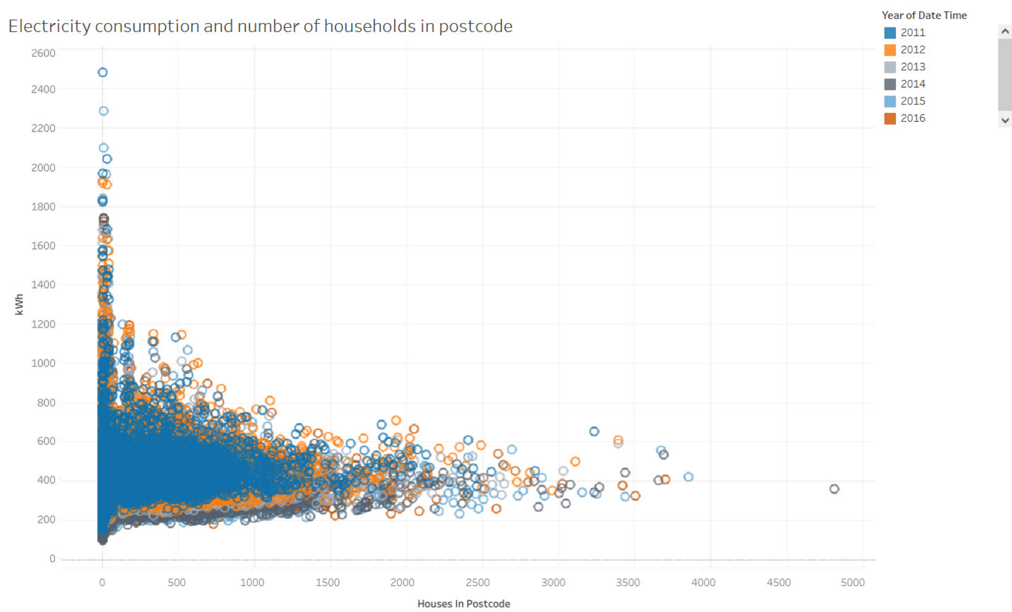


Figure 10: Electricity consumption and number of households in postcode, part 2 of Changes over time

4.3 Temperature

The third data story presents the relationship between domestic electricity consumption and temperature from a variety of perspectives. Temperature is shown in the geo-visualization on the left through color, whereas the electricity consumption is shown on the right with color and radius. These maps can also be used to examine changes over time in greater detail, using the time slider. The Tableau visualizations present broader patterns over time and also include interactive elements.

Figure 11 compares the temperature in districts with the electricity consumption data on the postcode level. Both sides of the visualization are linked by the time stamp.

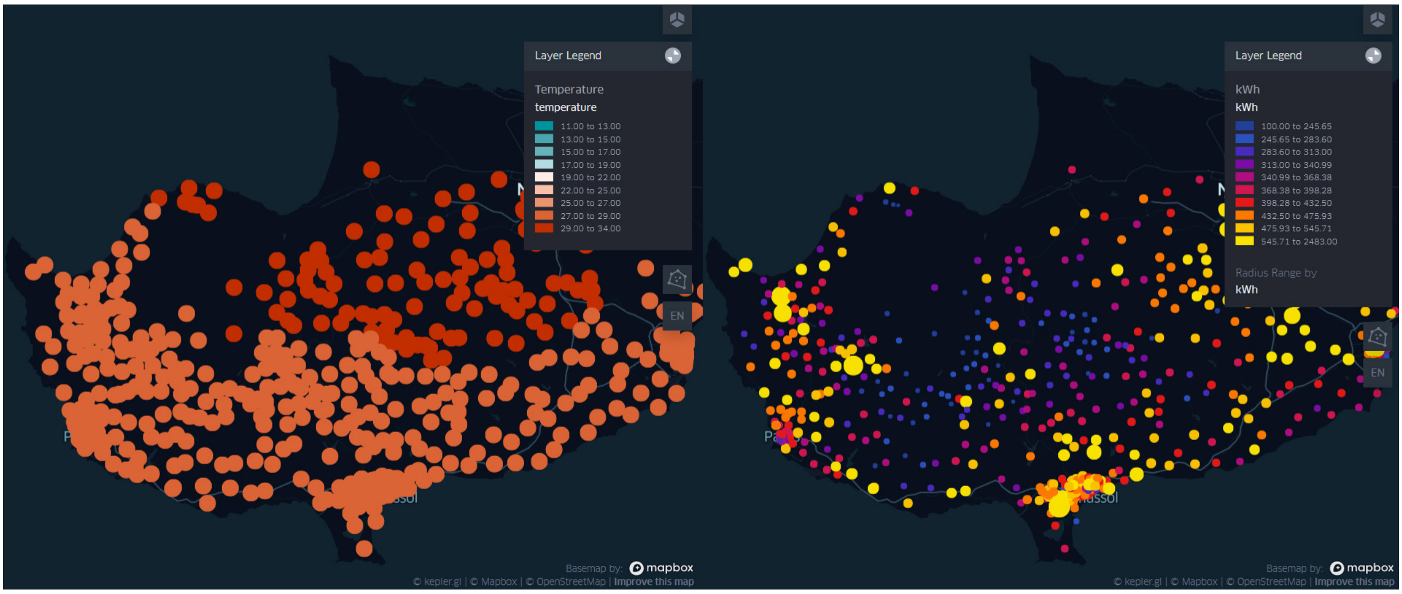


Figure 11: Electricity consumption and temperatures in districts

Figure 12 compares the average electricity consumption with the outside temperature at the postcode level. The first graph presents the average electricity consumption on a postcode level over the period 2011-2016. The geo-visualization is intended as a means of selecting postcodes from the map in order to highlight them in the scatter plot below.

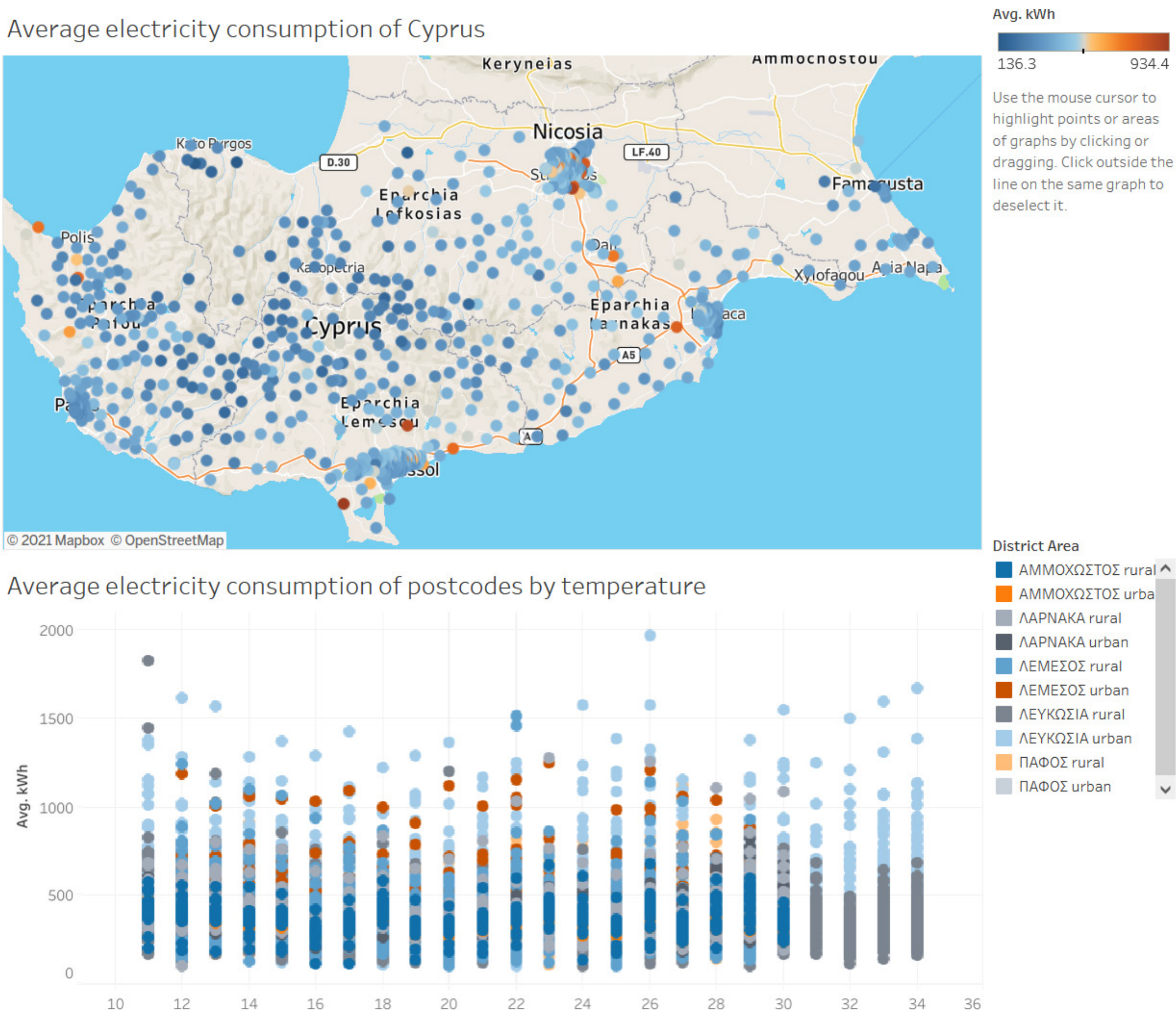


Figure 12: Electricity consumption and temperature in Cyprus

Figure 13 depicts electricity consumption in Cyprus on the postcode level. The graphs are separated on a bimonthly basis, corresponding with the methods used in the Social Electricity project. The electricity consumption per each two-month period is taken as an average over the period 2011-2016. The maps can be navigated, and selecting an area on one map will highlight it on all of them, also adjusting the value range to match. The figure is a cropped image of the full visualization.

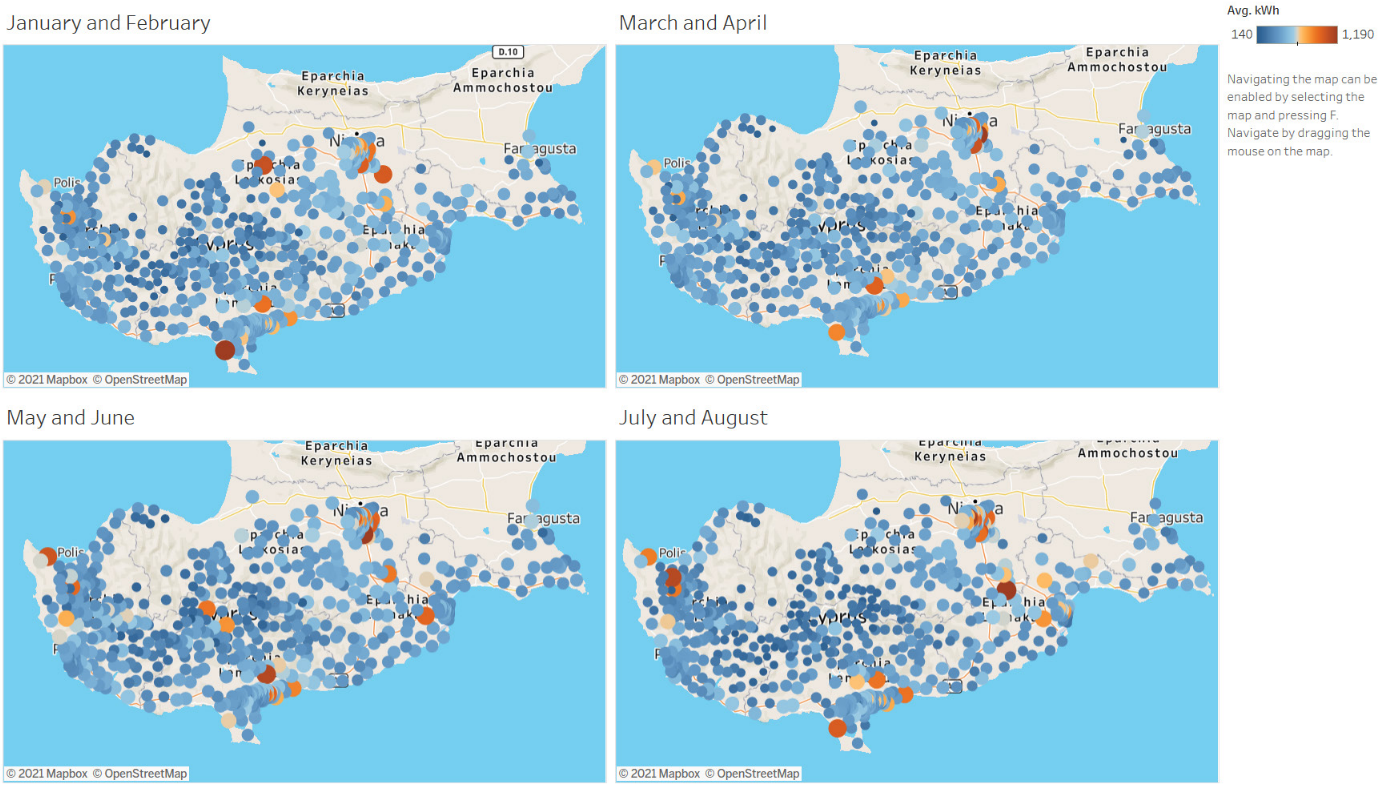


Figure 13: Bimonthly electricity consumption across Cyprus

Figures 14 and 15 depict the recorded temperatures and average electricity consumption in the districts of Cyprus over the period 2011-2016. The second graph shows the average electricity consumption per month in each district, aggregated over the same period.

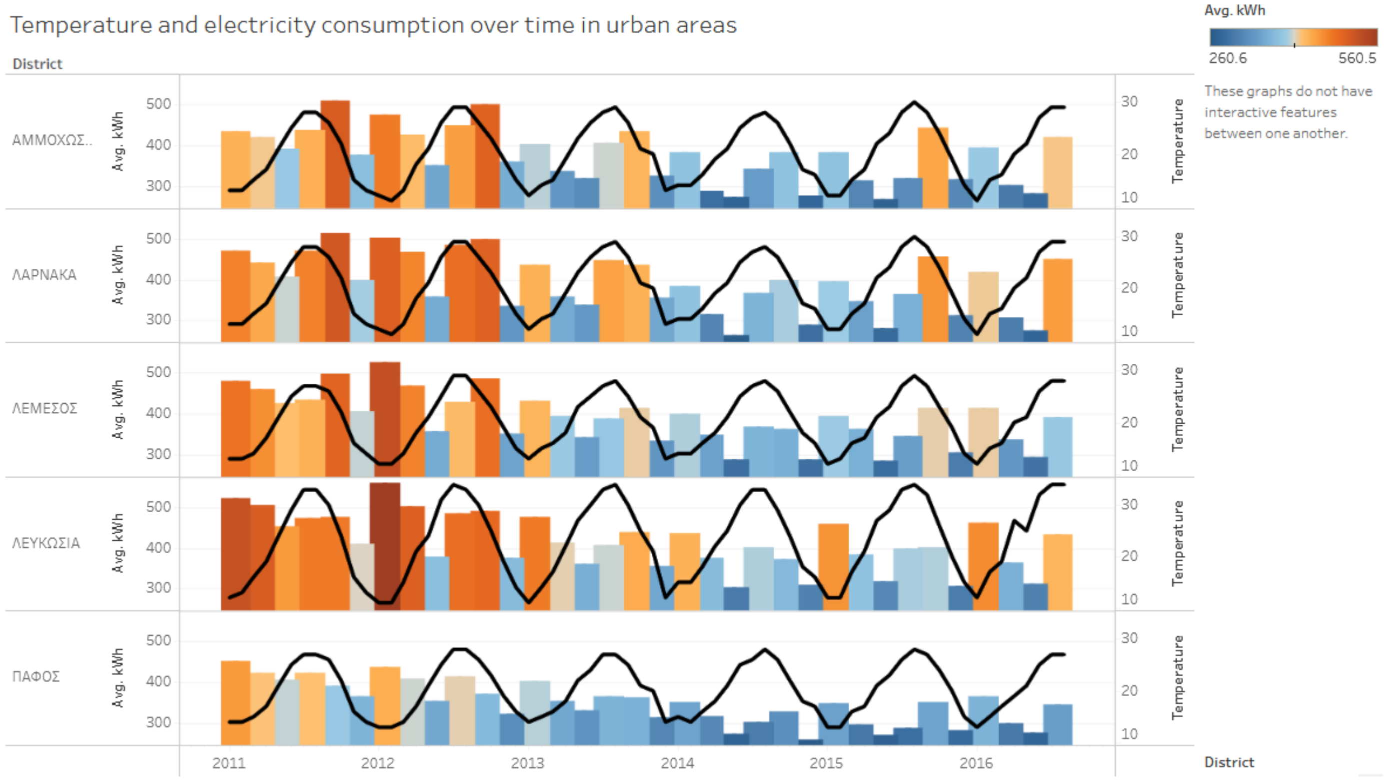


Figure 14: Temperature and electricity consumption over time in urban areas

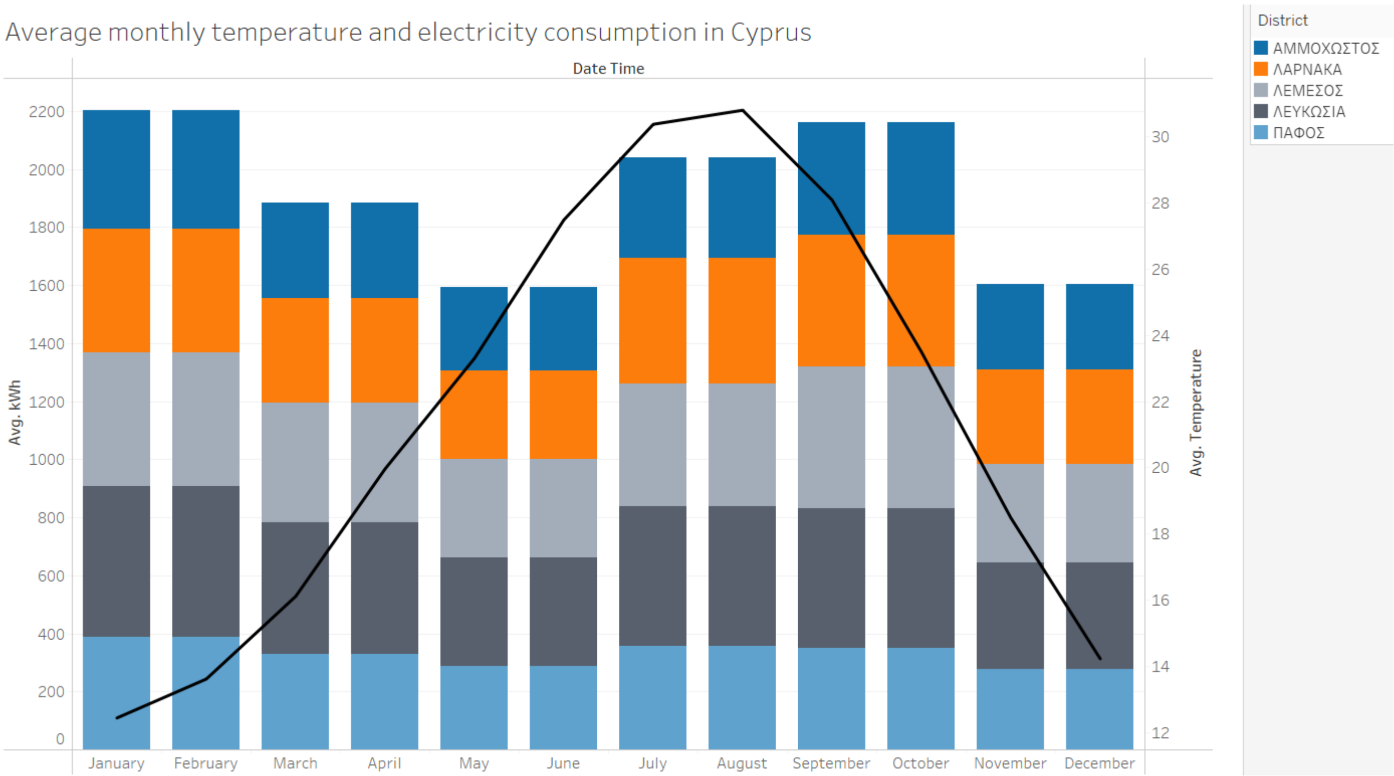


Figure 15: Average monthly temperature and electricity consumption in Cyprus

4.4 Expenditures

The fourth and last data story depicts the relationship between the average amount in euros spent on utilities by households in district areas and the household electricity consumption. Expenditure is shown in the geo-visualization through color, whereas the electricity consumption is shown by the hexbin height. These maps can also be used to examine changes over time in greater detail, using the time slider. The Tableau visualizations present broader trends in changes over time and also include interactive elements between each other.

Figure 16 presents a hexbin map of the utilities expenditure as color and kWh as the height of the pillars. Due to the hexbin map layer, some postcodes are combined in pillars and their average kWh readings are displayed. It is also important to note that the hexbin layer makes a new range each time the visualization refreshes. This means that the height of pillars corresponds to the range of kWh values that are being shown for the dates that are currently being viewed.

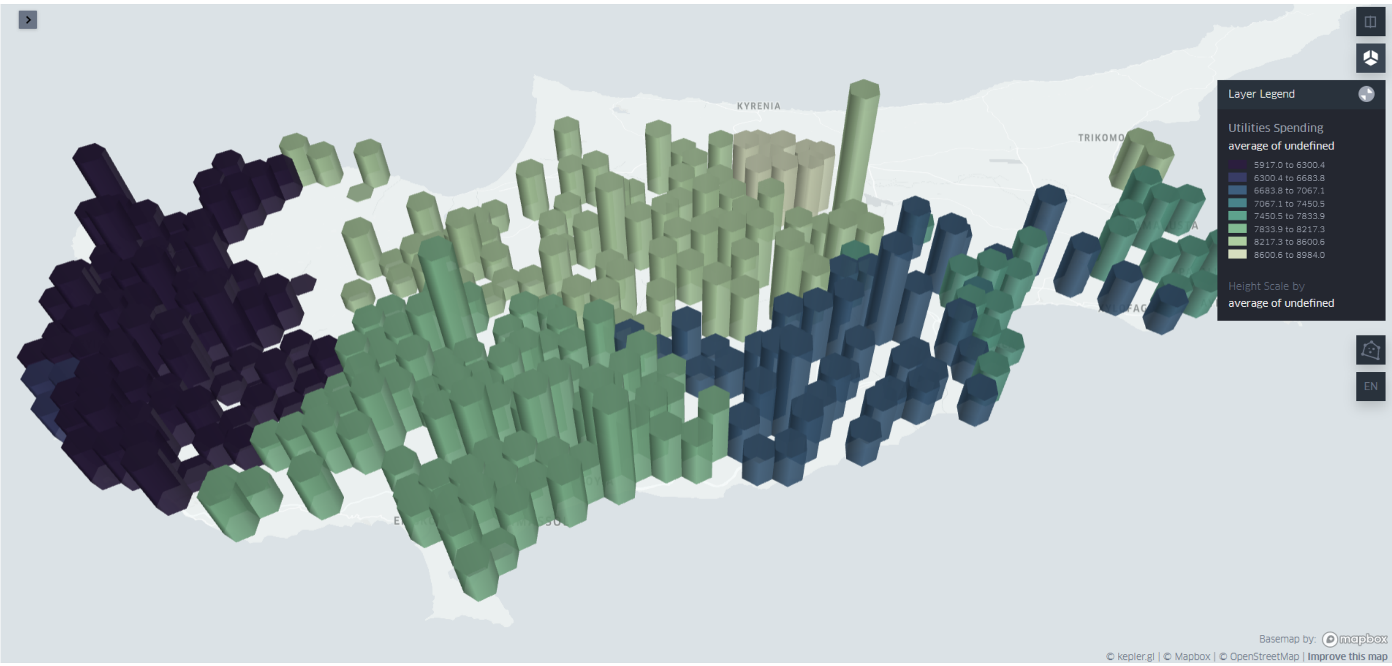


Figure 16: Electricity consumption and expenditure on utilities

Figures 17, 18 and 19 were combined in a dashboard presenting expenditure on utilities and the average electricity consumption in the areas of Cyprus. Figures 17 and 18 are both scatter plots of the same data, however the kWh are averaged are averaged per postcode in Figure 17 and per district area in Figure 18. The additional geo-visualization in Figure 19 was added for the convenience of the viewer and to promote interaction. It can be used to see where the selected postcodes or district areas are located or vice versa.

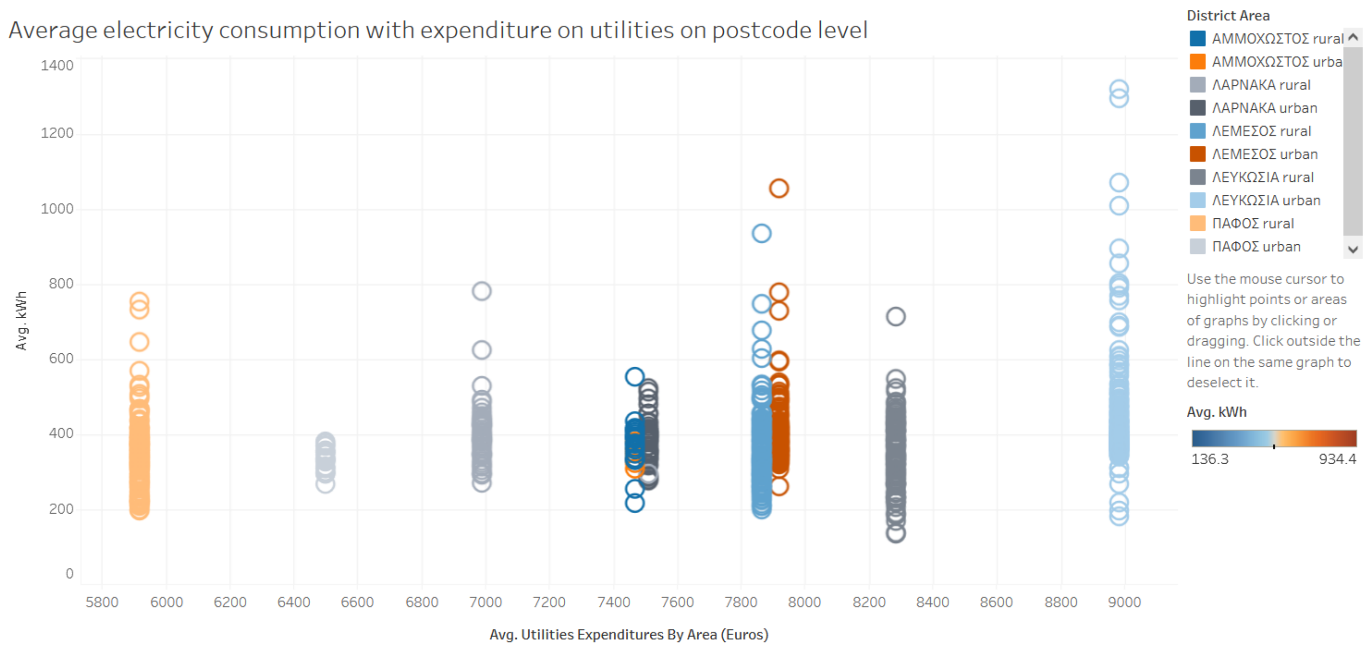


Figure 17: Average electricity consumption with expenditure on utilities on postcode level

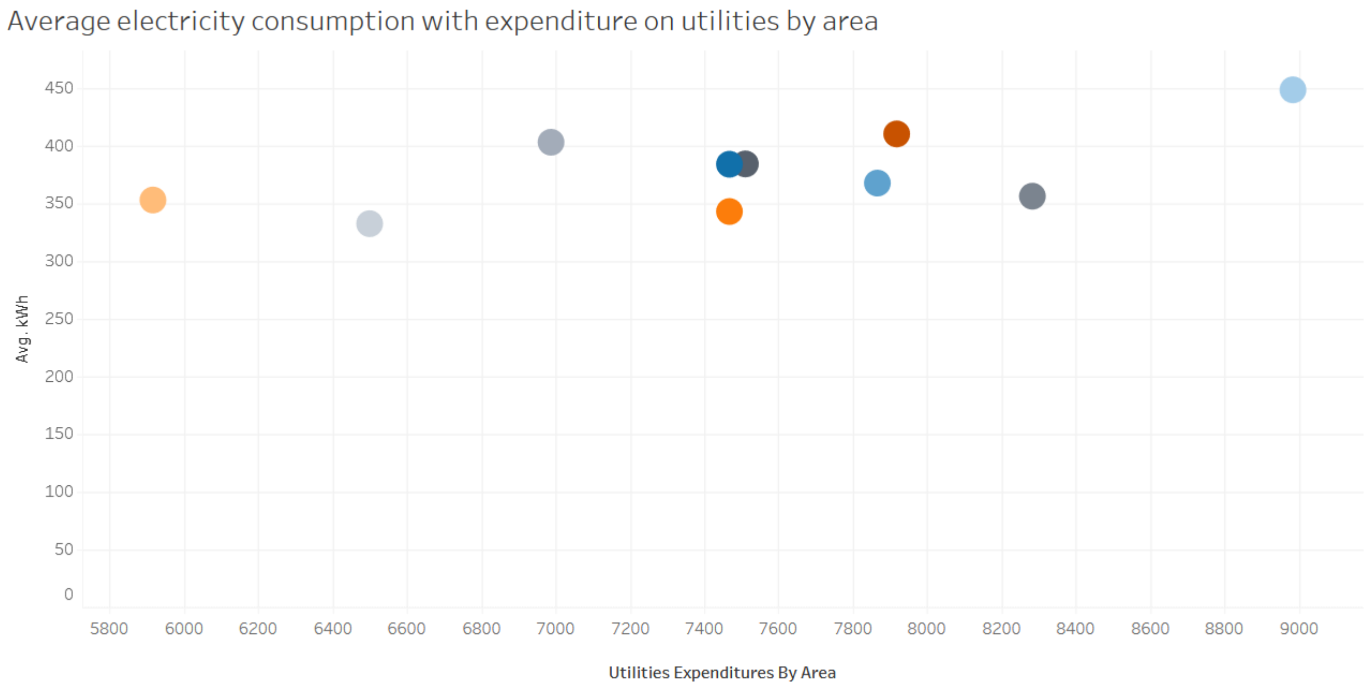


Figure 18: Average electricity consumption with expenditure on utilities by district area

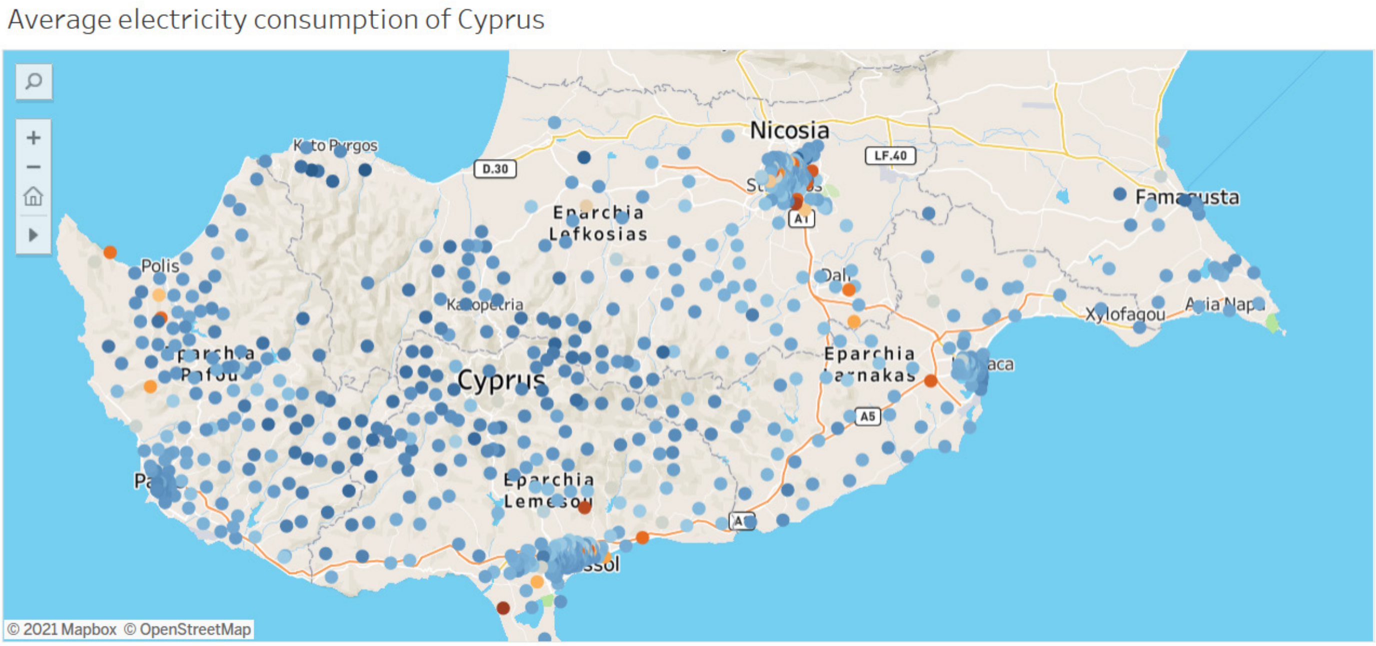


Figure 19: Average electricity consumption of Cyprus

4.5 Integration in a website

Having completed the final data stories, the next step was integrating them all in website. The prototype presented to the project supervisor in the specification phase had been quite basic in design, however, the necessary blocks such as the layout and navigation bar had been established. This meant that the integration process was quite simple, and mostly stylistic work was done to make it appealing. Further changes included the separation of the Home and Electricity consumption pages, as well as the inclusion of the page button for switching languages.

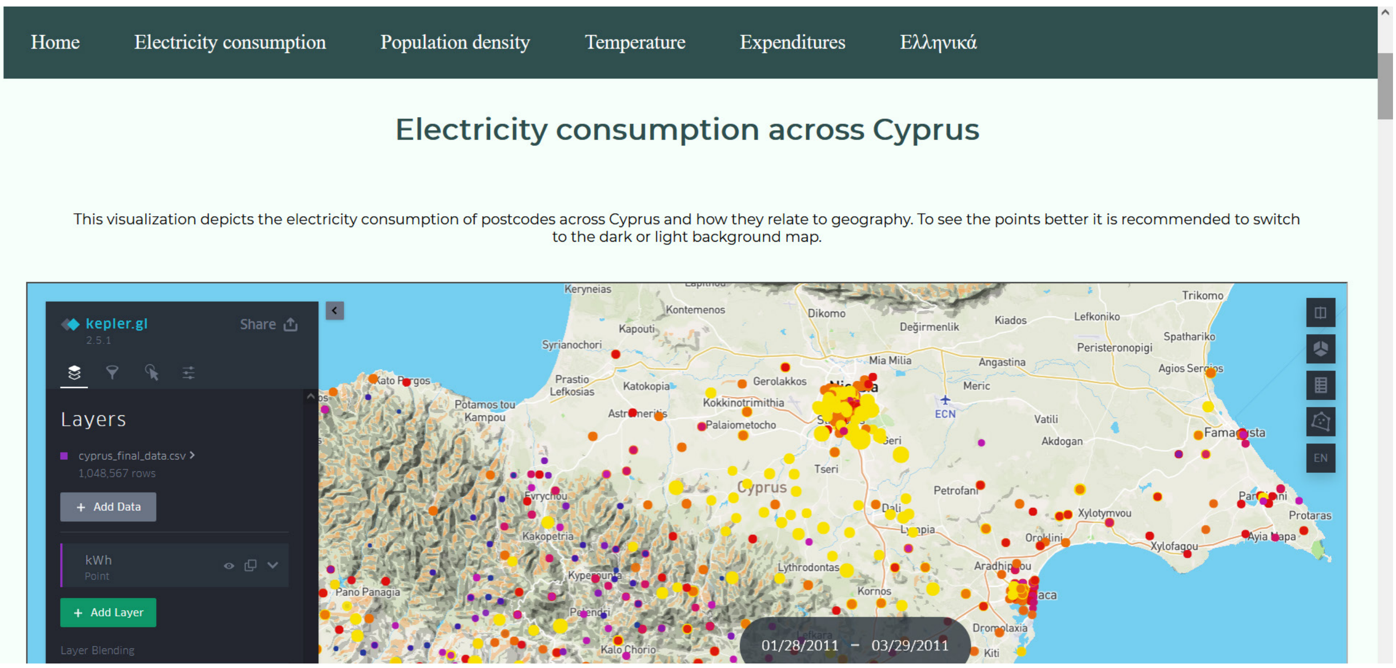


Figure 20: Screenshot of the first data story on the website

The Home page was created with the intention of introducing the project, providing tutorials, and adding contact information. A tutorial with images was included for Kepler.gl so viewers would be able to adjust the geo-visualizations and make new ones if they choose.

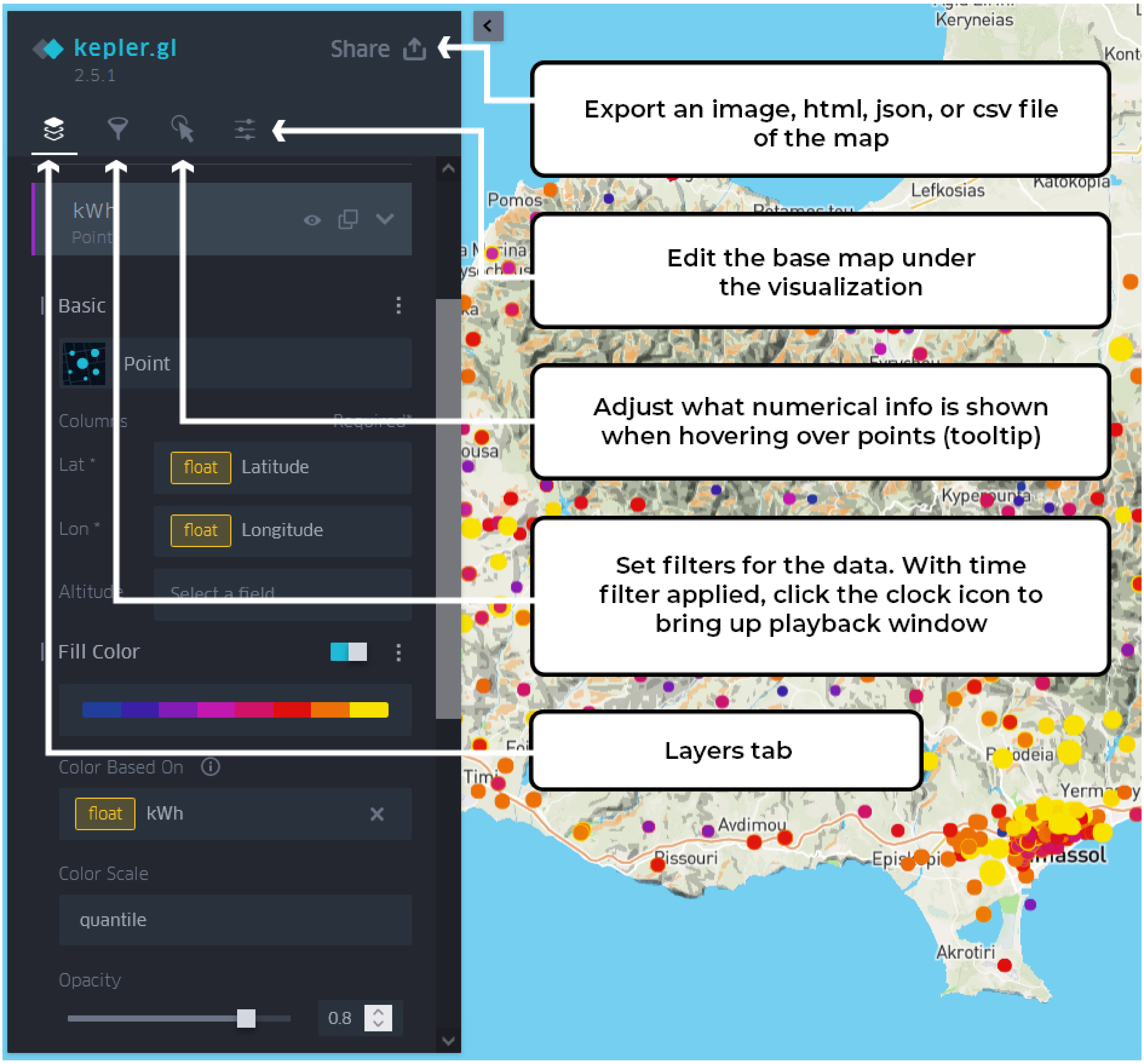


Figure 21: Example of an tutorial image used in the website

The finished prototype concluded the realization phase.

Chapter 5: Results and Evaluation

Having completed the final prototype, the final step is to analyze the results and evaluate the prototype. The evaluation was based largely on how successful the final prototype was at answering the research questions established at the beginning to project. Additional evaluation marks are made based on the guidelines established in the ideation phase.

5.1 Results

This section will detail the results of the data stories individually and the website as a whole. The results will include an analysis of any patterns which may be apparent in the stories, as well as positive and negative aspects as they relate to function or the experience. In the analysis of the visualizations, it was found that some visualizations did not contribute much, so they are not included.

5.1.1 Electricity consumption

The first story focused on the electricity consumption itself and generally succeeded at identifying areas with consistently high or low electricity consumption. Although most of the visualizations point to urban Nicosia as having the highest electricity consumption in the country, it is most clearly depicted in the line graph of Figure 4. Throughout the years, it has roughly 40 kWh higher consumption on average than urban Limassol at second place. Conversely, the lowest electricity consumption is in urban Paphos, with urban Famagusta at a close second. Electricity consumption also appears to fluctuate between the areas each year, but they all follow a similar downward trend from 2011 until 2014, after which consumption picks up again. In 2014, there was an economic crisis, so the decreased electricity consumption is reasonable given that people would have saved more, and tourism would have been lower. The largest decrease in electricity consumption was in rural Larnaca.

The geo-visualizations in Figure 1 can show the changes in electricity consumption in more detail, but the overall yearly changes are less apparent there. In the visualization, the postcodes in each area with the highest and lowest electricity consumption can be identified. Changes between months are also quite apparent in the visualization. For one, many postcodes in the coastal areas of Cyprus had relatively high electricity consumption in the summer months, with the most activity between August and October. Some of those postcodes also have high electricity consumption in January and February, but in those months, electricity consumption appears more evenly dispersed through the country. An explanation for this yearly pattern would be that tourism causes the coastal areas to experience high electricity consumption from August to October, whereas in the winter, cold temperatures throughout the country lead to higher consumption overall. The explanation of tourism does not, however, explain why July does not see a significant increase in electricity consumption, as it is the most popular month for tourist arrivals according to the Statistical Service of Cyprus [39]. Nonetheless, the coastal areas are not the only ones to experience high electricity consumption in the summer. Together with the heatmap, more areas of more intense consumption could be identified, the two most prominent being the northern area of rural Paphos and the flat, inland area roughly between the cities of Nicosia and Larnaca. In the case of these two, the visualizations do not provide a clear explanation for the high consumption of these two areas, however, the causes may be useful for policy makers to investigate in efforts to decrease electricity supply. From the geo-visualizations in Figure 1 it is also clear that postcodes in the Troodos mountain range generally have much lower electricity consumption compared to other areas. Numerous smaller areas of greater electricity consumption can also be identified, but they do not all need to be covered here. Having identified several areas of more intense electricity consumption, the next step for policy makers is to investigate the causes. Some explanations are provided by the other stories, but there is no doubt that each area has unique circumstances which cannot be covered by the data used in this project. When more is known of those circumstances, the choice of action can also be made. For one, it may be possible that actions to reduce electricity consumption in some areas may be unnecessary or too challenging. If action is taken, the unique circumstances of each area will also require unique solutions.

Lastly, a few useful insights could be extracted from the visualizations in Figures 6 and 7. Starting with Figure 3, the six geo-visualizations reveal that the postcodes with the highest and lowest average electricity consumption remained approximately the same. The biggest difference is in the year 2011, however, excluding the postcode with the highest consumption in the Akrotiri peninsula reveals many of the same postcodes which are in later years. Figure 7 depicts how electricity consumption patterns differ in the rural and urban areas of districts. An observation is that Nicosia and Limassol generally have higher consumption on average in the urban areas, whereas the rural areas are generally higher in the other districts. An interesting detail is also that the difference between urban and rural areas in Nicosia and Limassol tends to be highest from July until September, with slightly lesser differences occurring in January and February. In other months, the differences are much lower, often disappearing altogether in Limassol. As Nicosia and Limassol are the two largest cities, a possible explanation for the summer months may be urban overheating for the summer months. Interestingly, a similar difference occurs in Famagusta, except in that district the rural areas have much higher consumption from August to October. A likely explanation may be tourism, as that area contains the popular tourist destination of Ayia Napa. A clear suggestion for policy makers based on Figure 7 is to tackle the electricity consumption spikes that occur in the urban areas of Nicosia and Limassol, as well as the rural area of Famagusta. Not only do these patterns mean greater greenhouse gas emissions, but the rapid change in electricity consumption may also be straining for the electricity grid. Based on Figure 3, the suggestion is the same as before – to investigate what causes some of these postcodes specifically to have such disproportionately high electricity consumption and determine whether or not action is needed to prevent unnecessary electricity waste.

5.1.2 Housing density

The housing density data was implemented into the project late in the specification phase, but provided some interesting insights. Both Figures 9 and 10 show a pattern of postcodes in with low housing density having relatively high electricity consumption, whereas the opposite occurs for postcodes with high housing density. The addition of the housing density parameter on the heat map in Figure 9 provides some additional context to the high consuming rural areas that were identified earlier, although there is no clear pattern between these regions. For one, the northern part of rural Paphos appears to have multiple postcodes with low housing density that consume a lot of electricity on average. In the rural areas between Nicosia and Larnaca, the housing density is more varied, but generally seems to have a bigger population. Besides these, the coastal areas in the south and east of Cyprus which have high consumption in the summer also tend to have high housing densities. If policy makers aim to decrease electricity consumption, these insights suggest that each area needs to receive special attention, as the causes for their high consumption are likely different.

Further investigation of the relationship between these two parameters in Figure 10 gives a different perspective. Contrary to the insights from before, the highest electricity consumers here are clearly postcodes with roughly 50 or less households. This appears to be the case across most areas of Cyprus, with the exceptions of Famagusta, urban Larnaca, and urban Paphos. These three areas show little to none of that relationship and the range electricity consumption is much smaller across the number of households. The explanation for this cannot be derived from these visualizations alone, but it is clear that in order to reduce the electricity consumption in Cyprus, these areas with low housing density but high electricity consumption need to be addressed by policy makers.

5.1.3 Temperature

Of all the data stories, the one for temperature presented some of the most controversial discoveries. A strongly expected pattern for this story was that the highest and lowest temperatures will correlate with peaks in electricity consumption due to heating and cooling. Milder temperatures occurring in the spring and summer months should therefore also have lower electricity consumption. A close inspection of the visualizations, however, reveals that to be just partially true. Looking at Figures 14 and 15, which are combined together in a dashboard, peaks in kWh do occur during the colder and warmer months. Nevertheless, the correlation is not as strong as expected. For one, the pattern is strongest when looking at broad averages taken for the whole country across multiple years. Furthermore, it is mostly clear in January and February when temperatures are at their lowest. In the summer, however the highest peak is in September and October, when temperatures and tourism have begun to decrease. Figure 14 shows that this pattern is also evident in all of the five major cities. Another observation is that December has relatively low electricity consumption despite the cold temperatures, although this is likely because the readings were taken bimonthly in the Social Electricity project. Based on the information in these two visualizations, we can conclude that temperature does play a role in electricity consumption. Heating and cooling costs increase domestic electricity consumption, with January and February as the most extreme. Less electricity is spent on cooling, but it is still significant. As for the high consumption in September and October, the data does not provide an answer and more research is needed.

The dual geo-visualization in Figure 11 appears to also confirm the pattern overall, but also provides greater context for where high or low electricity consumption occurs. Looking at the differences between the warmest and coldest months, it appears that electricity consumption is somewhat more evenly dispersed in colder months, meaning that more areas of the country are consuming high amounts of electricity. Moving towards spring, the electricity consumption decreases in most areas. When it picks up again in summer, consumption appears more concentrated around the coastal areas, becoming very apparent in from August until October. Then until winter, electricity consumption across the country is relatively low again. Although it is not visible on the visualization, the mountainous area is generally a few degrees colder across the seasons. This can explain why it has high electricity consumption in the winter, as households need to spend more on heating relative to the coastal areas. Conversely, in the summer, temperatures in the mountains are cooler, so households there would also spend less electricity on cooling. A somewhat concerning observation is that several postcodes in the mountains have very low electricity consumption even in the coldest months. There is a possibility that those households do not have sufficient access to electricity, but there may also be other explanations, such as a greater use of alternative heating systems. Another temperature difference that is not visible is in the flat, inland area of Nicosia and Larnaca, which experiences temperatures a few degrees warmer in the summer and colder in the winter compared to the coastal areas. As such, the temperature differences can also at least partially account for the increased electricity consumption in that area, which is especially prevalent with the highest temperatures. Considering the observations from this geo-visualization and from Figures 14 and 15, a reasonable policy suggestion for decreasing electricity consumption would be a program for improving insulation in houses. Quality insulation can help maintain stable ambient temperatures without the need for excessive cooling or heating.

Contrary to the patterns observed before, the scatter plot in Figure 12 revealed that individual postcodes can have very different trends when it comes to the relationship between temperature and electricity consumption. Whether looking at individual postcodes or groups, the pattern observed earlier is rarely seen and consumption can vary in seeming randomness between temperatures. These observations suggest that policy decisions should be made with some caution. Although it is clear that overall electricity consumption is affected by temperatures, it can differ between postcodes. Furthermore, a closer examination of different data or research may be necessary to investigate the effects of temperature on electricity consumption in Cyprus.

5.1.4 Expenditures

Generally, meaningful correlations were hard to draw from the visualizations in this data story, which was due to limitations of the data. The expenditure is generalized over many postcodes and it shows no changes over time, as the only available data was from 2014. As such, the observations made here were the least reliable. The two scatter plots in Figures 17 and 18 do not present any clear patterns, which may indicate that expenditures on utilities may not play a significant role for electricity consumption. Urban Nicosia has the highest of both expenditure and electricity consumption on average, however, it also has the biggest range of electricity consumption values. In comparison, urban Paphos and Famagusta have very small ranges. Both cases can be explained by the relative sizes of the cities and the number of postcodes included. The high expenses in Nicosia could also be explained by the large number of postcodes which consume far more electricity on average, but this explanation does not fit as well for other areas. For one rural Paphos and Nicosia have roughly the same electricity consumption, but very different utilities expenditures. In general, the average electricity consumption does not vary significantly between different areas, except perhaps urban Nicosia. This would suggest that the water and gas consumption may play a more significant role and should be explored in the future for correlations. Another factor may be the differences in the rates charged for electricity.

The geo-visualization in Figure 16 presents a different view of electricity consumption, not only because of the inclusion of expenditures, but also the use of a hexbin map layer, which samples the range of displayed values differently. As such, this layer brings a different perspective to the scatter plots. For one, urban Nicosia rarely has the highest electricity consumption in any two-month period. Instead, consumption in all hexagons is steadily high. In contrast, some postcodes in northern Paphos and around the city of Limassol have consistently the highest monthly electricity consumption, but they also drop occasionally to very low values. Similar cases can also be seen in rural Larnaca and Nicosia. In general, most of the very high electricity consumers lie outside the urban areas, where expenses are also lower.

5.2 Evaluation

The aim of this project was to understand the domestic electricity consumption in the Republic of Cyprus and how it changes through space and time. In doing so, various parameters such as temperature, geography and demographic characteristics of the local population needed to be considered. The goal was to develop explorative 2D or 3D geo-visualizations using data visualization software which could be used to identify patterns of electricity consumption and correlations of consumption. The project resulted in a series of interactive geo-visualizations which can be used to analyze domestic electricity consumption within different contexts. These were curated in data stories which illustrate certain patterns and implications in electricity consumption. Each data story shows a different perspective on domestic electricity consumption in Cyprus.

Two research questions were established in the beginning of the project. This section aims to evaluate to what extent the project was able to answer those questions. Besides these, several guidelines were established with the intention of guiding the project in a direction that is ethical, meets the stakeholder requirements, is on par with the state of the art, and is relevant for the cartography community. Adherence to these guidelines and their results will also be discussed in this section.

5.2.1 Answering the first research question

The first research question for this project was: how to visualize electricity consumption at a country scale such that useful insights can be drawn from it? Answering this question was challenging and relied on a combination of research and experimentation. In terms of research, both the state of the art and the core concepts of visualizing geo-spatial big data were necessary components in understanding the possible methods, as well as what information is useful to extract from the data using visualizations. Before looking into the data, three aspects became clear which needed to be addressed in order to help the viewer draw useful insights. These were: depicting change over time, providing interactivity, and having aesthetic visualizations. These three aspects needed to be addressed by the software chosen to make the visualizations. Kepler.gl and Tableau were chosen, and they managed to fulfill their purposes perfectly. Both software provided plenty of options for visualizing the data, interacting with it, and showing how it changed over time.

The part of answering this question was in the specification phase, where those options had to be explored. This was a process of experimentation that lasted for several months. During this process, a variety of visualization styles were tested, and many parameters were added or removed from the dataset. Most combinations were able to reveal different insights from the data that others could not. Although all of the data stories provide insight to electricity consumption, the first data story was the most significant for the first research question, as it focused only on the electricity consumption data.

A crucial factor that allowed for a variety of insights to be gathered was the inclusion of different visualization styles, all of which had specific benefits and limitations. The geo-visualizations were understandably the most effective at determining specific postcodes or areas that have high or low electricity consumption. This was because the Social Electricity data was geo-spatial, and a map visualization could give an easily readable overview, while also providing additional context in terms of geography. The different map layers also had specific advantages. The dots map layer was good for showing individual postcodes, the heatmap for areas, and the hexbin layer for showing the highest and lowest consumers monthly. Where the geo-visualizations faced difficulties was in summarizing the data and differentiating scale. For example, numerically comparing the average electricity consumption between rural and urban areas was not possible due to the variety of differently sized and colored dots on the map. Furthermore, smaller differences in color or radius made it difficult to easily assess the smaller differences between postcodes. For that reason, other graphing styles such as line, bar, or scatter plot charts were found to be more effective. Line and bar charts sacrificed the individual differences between postcodes in favor of generalized readings for areas, making it possible to understand the differences in overall consumption between districts with clear numerical values, such as in Figure 7. Scatter plots, on the other hand, could show how individual postcodes relate to two different parameters. Doing so revealed clear patterns that would have been harder to detect with other visualization styles, which was especially apparent in the scatter plot for expenditure and electricity consumption. In the end, the nature of the Social Electricity data meant that the focus of the project was largely on geo-visualizations. Nonetheless, the inclusion of other visualization styles was necessary to fully understand the geo-visualizations and draw useful insights from them.

A similarly crucial aspect was the inclusion of the temporal dimension, which enabled the visualization to reveal patterns. The temporal dimension came in the form of time stamps in the data and visualizing it gave different results with different visualization styles. Showing change over time with geo-visualizations was found to be most effective at revealing geo-spatial patterns, such as the changes in the distribution of electricity consumption between winter and summer. The inclusion of several years' worth of readings helped to solidify those patterns. On the other hand, recognizing those patterns was not always easy and usually took multiple viewings. The problem was that the slider had to be moved between dates, but there was no way to look at two different dates simultaneously. The large number of postcodes also made it easy to forget more minor details. A partial solution to this problem was possible with Tableau, such as combining several geo-visualizations from different years side-by-side. Although they were not as interactive as the Kepler.gl visualizations, they were a start at solving this problem. The bar and line charts were also crucial for to their ability to show changes over multiple years in a single visualization. For example, the decline in electricity consumption from 2011 until 2014 was practically invisible in the geo-visualizations, but obvious on the line chart. Scatter plots were the only style for which changes over time did not fit well, but that may have been due to the nature of the data.

Lastly, utilizing different base maps in the geo-visualizations was a beneficial technique for drawing some useful insights. For one, there was the elevation map, which was created in Mapbox with the intention of highlighting mountainous regions. Having that underneath the visualizations helped understand the rough geography around each postcode, which helped understand the consumption habits in different regions. The limitation was that it created more visual clutter on the map, so it became more difficult to read the visualizations. As such, switching to the dark or light base map was necessary to bring the data back into focus.

5.2.2 Answering the second research question

The second research question for this project was: how to correlate electricity consumption at a country scale with different parameters such as weather conditions, seasons, and demographic characteristics of the local population? In some cases, the answers to this question were similar to those of the first question, but there are also distinct differences. This question can be answered using the second, third, and fourth data story.

Temperature was chosen as the most relevant parameter for analyzing weather conditions and seasons. This was because temperature strongly influences the electricity used on heating and cooling. Moreover, historical temperature data was easily accessible for multiple locations across Cyprus. For correlating temperature data with domestic electricity consumption, alternative visualization styles such as bar, line, and scatter plot charts were found to be most effective. One reason for this was that Cyprus is a relatively small country, so significant variations in temperature across different locations did not occur. Another reason was that the temperatures followed a very predictable pattern each year, meaning that when looking at geo-visualizations, it was usually enough to know the month in display in order to know the approximate temperature. The exceptions to this were the mountainous and flat inland areas, however the temperature differences there could not be displayed due to limited data. In those areas the geo-visualization was also most effective, as it could help explain their different consumption patterns. Getting more detailed data could thus also make the geo-visualizations more effective. Nevertheless, the utilizing alternative visualization methods were very effective at showing the overall correlations between temperature and electricity consumption. Figures 14 and 15 used a combination of line and bar charts to correlate the two parameters. The two visualizations could thus effectively show that very cold and very hot temperatures mostly correlated with increases in electricity consumption. The exception was with the consumption spike in September and October, which temperature alone could not explain. Analysis with other weather parameters such as humidity may be able to provide an explanation using the same visualization methods. Similarly, different weather parameters could help explain the lack of correlations between temperature and electricity consumption in the scatter plot shown in Figure 13. In this case, the visualization showed that postcode areas could have very different responses to temperature changes and that further examination with more detail is necessary.

Depicting correlations between housing density and electricity consumption was perhaps the most successful, likely because the two datasets were equally detailed. The housing density parameter was not extensively researched beforehand, except its relation to temperature through urban heat bubbles. Nonetheless, the data was already included in the Social Electricity dataset and analyzing it revealed interesting insights, as postcodes with few households were always the highest consumers of electricity. Again, depicting this relationship was most effective with both a geo-visualization and a scatter plot. The scatter plot showed that the highest electricity was always consumed in postcodes with approximately less than 50 households, but also that low housing density did not always correlate to high electricity consumption. The geo-visualization was most effective at showing how these two parameters correlate in different locations. For example, the high electricity consuming area in northern Paphos had low housing density, whereas the flat inland area between Nicosia and Larnaca also had higher consumption, but housing density was more diverse. In the latter case, temperature was found to be a more significant factor. Furthermore, the electricity consumption in the mountainous area was similarly more influenced by temperature than housing density.

Lastly, the expenditure on utilities was chosen as an important parameter because it relates to income and because it was expected to correlate with electricity consumption. Interestingly, this correlation was often not observed. The scatter plots showed that, on average, households in urban Nicosia paid the most for utilities and also consumed the most electricity. In other areas, this relationship was largely not apparent. The average amount of electricity consumed by households could be the same in different areas despite spending very different amounts on utilities. The geo-visualization, however, revealed that the postcodes which consumed large amounts of electricity were often in rural areas where less money was spent on utilities bills, with urban Nicosia being an exception to that. Despite these observations, the lack of detail in the expenditures data makes insights difficult to draw. For one, expenditure bills should vary across postcodes, so both visualization styles could reveal more correlations. Furthermore, expenditure is not the same as income, which was the parameter that visualizations tried to investigate. If the average income data was available for each postcode, it is likely that many meaningful insights could be found. Nevertheless, the with the data that was available, the visualization methods used in this data story were found to be the most effective.

In answering the second research question, the parts for both temperature and expenditures mentioned that better quality data would likely lead to more insights. Although it is reasonable to assume that the visualization methods that were used would also be effective with better data, there is no guarantee. Most importantly, it is likely that geo-visualization methods would perform better at providing insights with higher quality data, as was the case with the housing density parameter. As it currently stands however, the visualization methods used in answering the second research question were found to be the most effective with the available data.

5.2.3 Adherence to the guidelines

Several ethical guidelines were established in the ideation phase, in order to make the results of this project more impactful and relevant. These were based on the state of the art, stakeholder requirements, ethical considerations, and relevant problems in making geo-visualizations currently faced by cartographers. As many of the guidelines from these sources overlapped, they could be summarized in four overarching guidelines. In this section, these guidelines will be restated and their impact on the project evaluated. It should be mentioned that some of the results are relevant for more than one guideline, but the repetitions are omitted.

The first guideline was to accommodate viewers from different backgrounds. This was relevant for the project for a number of reasons. First, the two major stakeholder groups of this project are the decision makers and the people of Cyprus. As both groups include viewers with various backgrounds that may not be relevant to this field, the visualizations and stories had to be understandable for those viewers in order to be effective and understandable. Ethically speaking, these visualizations are relevant for all the citizens of Cyprus and they hold relevant information about climate change, so they have to be publicly accessible. Lastly, cartographers have cited this as a major challenge for creating geo-visualizations of big data. This guideline was most influential in the experience design for the data stories and their presentation on the website. These influences are listed below:

1. Color schemes were chosen for the visualizations which are discernable for common types of color blindness.

2. An early version of the website was made in Greek. The full version would need to be available in in the future in both Greek and Turkish.

3. Notes were added to the Tableau visualizations to explain interaction possibilities.

4. A tutorial was added for using Kepler.gl in the home tab of the website.

5. The data stories were designed to be as clear as possible at a single glance. This meant removing clutter in the visualizations and avoiding jargon.

The second guideline was to design as a driver of change, meaning that the project should address environmental and social issues as a priority. This guideline was derived from ethical concerns and challenges faced by cartographers. Overall, this guideline could not be met to the extent that it was originally imagined, but some noteworthy steps were made nonetheless.

1. The visualizations and the data that was used were made publicly accessible. This was done in accordance with the target of the UNCCC to provide information about climate relevant data to the public in the Action for Climate Empowerment [34].

2. The analysis was largely focused on identifying areas of high or low electricity consumption. Cases of high electricity consumption have environmental concerns, as most of the electricity in Cyprus is produced from fossil fuels. Cases of very low electricity consumption could be indications of energy poverty, especially if they occur along side very cold or hot temperatures. Where possible, the analysis strived to identify these cases and provide policy suggestions.

3. An expectation for the project was to also locate low-income areas with high electricity consumption, as it also relates to the issue of energy poverty. Due to limitations of the data this was not possible.

The third guideline was to enrich the viewer's knowledge. This guideline is quite general and was largely derived from ethical concerns, although it is also relevant for cartographers. The idea behind this guideline was to provide opportunities for viewers to go beyond the analysis provided by the stories and encourage further learning opportunities. The primary impact of this guideline was that the Kepler.gl visualizations were made fully editable by the viewers. The inclusion of a tutorial and providing the dataset for free were also done in adherence to this guideline.

Lastly, the fourth guideline was to develop techniques for understanding change over time, which is most relevant for the challenges of visualizing geo-spatial big data. In order to adhere to this guideline, attention was put on the choice of software and the designs of visualizations. Overall, adherence to this guideline had a large impact on the outcomes of this project.

1. Kepler.gl was partially chosen for its fast-performing time window function and Tableau for its various options for managing the time variable.

2. All of the Kepler.gl visualizations include the time window.

3. Most Tableau visualization show change over time, unless where it was detrimental to the visualization.

4. Additional geo-visualizations (Figures 6 and 13) were made in Tableau to address the limitations of the time window in Kepler.gl.

Overall, the guidelines were well adhered to. The guideline that was hardest to follow was designing as a driver of change. This was largely due to limitations of the data, but also limited time to do research on relevant social or environmental issues in Cyprus. Incorporating emissions data from electricity consumption, or data on land use change may have also been a great options for designing as a driver of change. The guideline that was easiest to adhere to was accommodating viewers from different backgrounds. As such, it was also the most effective guideline, since it resulted in several simple design choices that were very impactful.

5.2.4 Evaluation from user testing

The final stage of the evaluation was to get feedback on the results of the project. As extensive testing with Cypriots was not possible with the time frame of the project, informal testing sessions were made with students of the University of Twente and Saxion University of Applied Science to evaluate the experience design of the website and data stories. The focus of the feedback sessions was on evaluating how well the viewers can understand what is being shown in a visualization and how well the interaction works. Comments on stylistic aspects and additional functions were also encouraged. A total of ten participants tested the website. Their comments are summarized and presented below.

• The color schemes used in the visualization was generally well received. In the geo-visualizations, the blue to yellow gradient was found to match with the concept of electricity consumption and the light blue to red gradient with temperature. The color scheme for the expenditure on utilities was the least clear. In the Tableau visualizations the chosen color scheme was also well received overall, although some problems did occur with light grey and light blue colored dots if they were selected in scatter plots.

• A mismatch in color schemes was noticed in the Tableau dashboards that combined scatter plots with geo-visualizations. Several participants felt that coloring the dots in the geo-visualization based on rural and urban areas of districts would fit better with the scatter plot.

• The Tableau visualizations in the first and third stories that combined six geo-visualizations were consistently identified as unnecessary. Generally, it was felt that the geo-visualizations made in Kepler.gl were enough.

• The Kepler.gl visualizations had the most engagement with the participants. Although most did not remember the instructions from the tutorial, they were able to discover functions and their purposes on their own. It is important to note here that as students, the participants are likely more accustomed to navigating websites and apps than other demographic groups.

• The Tableau line and bar graphs were generally the easiest to understand by the participants.

• The Tableau visualizations in the temperature data story, which combined bar graphs as electricity consumption and line graphs for temperature sometimes caused some initial confusion, since it was not clear which graph represented which parameter. Further clarification needs to be added along side the visualizations.

• More clarification about some interaction quirks in the Tableau visualizations also need to be added. This was mostly apparent with the Tableau geo-visualizations.

• Legends for the heat maps in Kepler.gl were not available, which was a fault of the software. One participant suggested manually adding an image as the legend instead.

• Several participants found the hexbin visualizations in Kepler.gl hard to read due to the large number of points.

• The heat map in Kepler.gl was found to be misleading by some participants since the radius changed depending on the amount of zoom. Clarification was needed to explain what they are seeing. This clarification should be added in text to the visualizations.

• An important point that became clear from the testing was that a too much was sometimes asked from the viewers when interacting with the visualization. Asking to switch between base maps, or select multiple postcodes on the scatter plot dashboards was an additional step that was necessary in order to properly read the data. A possible option would be to simplify some of the visualizations, but that would also reduce the amount of information they would convey.

Overall, the participants reported that they enjoyed going through the data stories and could find some insights from them. They also generally reported to enjoy the interactive elements, although those could be simplified overall. The number of visualizations should also be decreased. Although participants generally had mixed feelings about visualization styles, the higher number different visualizations in the first and third data stories did not seem to bring a lot more utility.

Chapter 6: Conclusions and Recommendations

This section will present the results of the project overall in terms of what went well and what did not. A summary of key insights will be presented with their impact on the project as a whole. Lastly, there will be a section for recommendations for future research.

6.1 Conclusion

Overall, the project can be counted as a success. In the project's infancy, the task of visualizing electricity consumption on a country scale seemed fairly simple. Nonetheless, as the work process progressed over several months, numerous layers of hidden complexity were discovered. The end result was a website, which presented several data stories with various visualizations in order to examine the domestic electricity consumption in Cyprus from many different perspectives. Moreover, the design was based on a set of guidelines such that it would be relevant and ethical. As such, it went far beyond the initial expectations.

The project was also successful at answering the research questions. The most important insight from their answers was that electricity consumption is a highly complex topic that requires extensive exploration to understand fully. Even for a small country like Cyprus, there are layers of information that cannot be uncovered with just one visualization. Considering this complexity, a major challenge was to present as much information with as few visualizations as possible. The majority of the work for the project was therefore in exploring the data with numerous visualizations until the best ones could be determined. Even so, many other useful visualizations were discarded or omitted from the final product, which is why the viewer is highly encouraged to explore the data further on their own if they wish to learn more. The final data stories presented numerous relevant insights about domestic electricity consumption in Cyprus and how it relates to different parameters, however, to what extent these insights would be useful for decision makers remains to be seen. Nonetheless, the exercise of visualizing electricity consumption on a country scale proved very fruitful. This was especially prevalent with the use of geo-visualizations, which could reveal detailed, geo-spatial patterns in electricity consumption that would otherwise be lost in the standard bar charts.

The most prevalent limitation for the project was certainly with the limited data that was used for the temperature and expenditures parameters. The results of this project largely owe their success to the data collected in the Social Electricity project, which was detailed and well formatted. Contrasting the quality of that data with the temperature and expenditures data left much to be desired, as the utility of the electricity consumption data also became more restricted. The search for both datasets took a long time and extensive experimentation in the specification phase was also required. In the end, both datasets were still able to reveal a great deal of information from the electricity data, so it is easy to imagine what more could be possible with better data.

6.2 Recommendations

Based on the insights from the evaluation phase and remarks from the conclusion, several recommendations for future research can be established. These will be discussed in this section.

Having established important parameters and challenges for study, as well as the best software for the visualizations, one recommendation would be to repeat the project with better data for temperature and expenditures. This would allow the final data stories to be the most useful for decision makers in Cyprus.

Projects with the same intentions and similar data should also be conducted for other countries. This is necessary to determine to what extent the results of this project are applicable outside of Cyprus. Performing similar projects in more countries could result in a more detailed and universal set of methods that should be used when examining electricity consumption and its correlation to other parameters on a country scale.

This project focused on a range of widely different parameters in an effort to cover a lot of ground at once, which came at the cost of deeper understanding. Fields such as the weather of Cyprus are much more detailed than just temperature. Exploring the correlations between domestic electricity consumption and various weather, economic, or social parameters could all result in much deeper insights with a narrower focus. Although these insights would also be less widely applicable, they have potential to be more impactful in their respective domains.

Although developing methods of studying change over time received special focus in the project, the possibilities were in no way exhausted. Alternative methods could be explored in future projects to discover more detailed patterns and make predictions for the future. These projects could approach the issue from a software side, or a data management side. The software side would imply the development of new data visualization software that is designed around studying change over time. Examples of possible functions could be visualizing the rate of change in geo-visualizations or geo-spatial pattern recognition for electricity consumption. The data management approach could also explore similar concepts, but instead manipulate the data in ways that change over time could better be explored with software like Kepler.gl.

Lastly, similar projects with different parameters should also be explored. As mentioned in the evaluation, the project was lacking in its ability to effectively address environmental issues. As these become increasingly important for worldwide well-being, future country-wide data visualization projects are necessary that focus on the environmental impacts of electricity consumption, by assessing its correlation to parameters such as air pollution, greenhouse gas emissions, or land use change. Similarly, geo-spatial data visualizations projects that focus on social issues in relation to electricity consumption could also prove an invaluable resource for decision makers.

References

[1] – Electricity production by source, World. (2021). Our World in Data based on BP Statistical Review of World Energy & Ember. https://ourworldindata.org/grapher/electricity-prod-source-stacked

[2] – Murakami, S., Takasu, T., Islam, K., Yamasue, E., & Adachi, T. (2020). Ecological footprint and total material requirement as environmental indicators of mining activities: Case studies of copper mines. Environmental and Sustainability Indicators, 8, 100082. https://doi.org/10.1016/j.indic.2020.100082

[3] – Zografos, C., & Robbins, P. (2020). Green Sacrifice Zones, or Why a Green New Deal Cannot Ignore the Cost Shifts of Just Transitions. One Earth, 3(5), 543–546. https://doi.org/10.1016/j.oneear.2020.10.012

[4] – Ritchie, H & Roser, M. (2019). Access to Energy. Our World in Data. https://ourworldindata.org/energy-access

[5] – Santamouris, M., & Kolokotsa, D. (2015). On the impact of urban overheating and extreme climatic conditions on housing, energy, comfort and environmental quality of vulnerable population in Europe. Energy and Buildings, 98, 125–133. https://doi.org/10.1016/j.enbuild.2014.08.050

[6] – THE 17 GOALS | Sustainable Development. (2015). United Nations https://sdgs.un.org/goals

[7] – McInerny, G. J., Chen, M., Freeman, R., Gavaghan, D., Meyer, M., Rowland, F., Spiegelhalter, D. J., Stefaner, M., Tessarolo, G., & Hortal, J. (2014). Information visualisation for science and policy: engaging users and avoiding bias. Trends in Ecology & Evolution, 29(3), 148–157. https://doi.org/10.1016/j.tree.2014.01.003

[8] – European Commission. (2014). Cyprus Country Report. Energy. https://ec.europa.eu/energy/sites/ener/files/documents/2014_countryreports_cyprus.pdf

[9] – Taliotis, C., Taibi, E., Howells, M., Rogner, H., Bazilian, M., & Welsch, M. (2017). Renewable energy technology integration for the island of Cyprus: A cost-optimization approach. Energy, 137, 31-41. https://doi.org/10.1016/j.energy.2017.07.015

[10] – Zachariadis, T., & Pashourtidou, N. (2007). An empirical analysis of electricity consumption in Cyprus. Energy Economics, 29(2), 183-198. https://doi.org/10.1016/j.eneco.2006.05.002

[11] – MCIT. 2013. Law for the Regulation of Electricity Market of 2003 to 2012, http://www.cylaw.org/KDP/data/2013_1_218.pdf

[12] – Kyprianou, I., & Serghides, D. (2019). Dealing with energy poverty in Cyprus – an overview. International Journal of Sustainable Energy, 39(4), 308–320. https://doi.org/10.1080/14786451.2019.1699560

[13] – Murugesan, L. K., Hoda, R., & Salcic, Z. (2015). Design criteria for visualization of energy consumption: A systematic literature review. Sustainable Cities and Society, 18, 1–12. https://doi.org/10.1016/j.scs.2015.04.009

[14] – Soliman, E., Shalaby, S., & Fetais, N. (2017). An Interactive Visualization Web Solution for Electricity Consumption in Qatar. 2017 International Conference on Computer and Applications (ICCA), https://doi.org/10.1109/comapp.2017.8079746

[15] – Goodwin, S., Dykes, J., Slingsby, A., & Turkay, C. (2016). Visualizing Multiple Variables Across Scale and Geography. IEEE Transactions on Visualization and Computer Graphics, 22(1), 599–608. https://doi.org/10.1109/tvcg.2015.2467199

[16] – Tai, T. (2013, September 11). How Power Map for Excel helps Opower visualize and present energy-saving messages. Microsoft 365 Blog. https://www.microsoft.com/en-us/microsoft-365/blog/2013/09/11/how-power-map-for-excel-helps-opower-visualize-and-present-energy-saving-messages/

[17] – Lock, O., Bednarz, T., & Pettit, C. (2020). The visual analytics of big, open public transport data – a framework and pipeline for monitoring system performance in Greater Sydney. Big Earth Data, 1–26. https://doi.org/10.1080/20964471.2020.1758537

[18] – Cawthon, N., & Moere, A. V. (2007). The Effect of Aesthetic on the Usability of Data Visualization. 2007 11th International Conference Information Visualization (IV ’07), https://doi.org/10.1109/iv.2007.147

[19] – Xexakis, G., & Trutnevyte, E. (2019). Are interactive web-tools for environmental scenario visualization worth the effort? An experimental study on the Swiss electricity supply scenarios 2035. Environmental Modelling & Software, 119, 124–134. https://doi.org/10.1016/j.envsoft.2019.05.014

[20] – CO2 emissions by fuel type, World. (2019). Our World in Data based on Global Carbon Project; BP; Maddison; UNWPP. https://ourworldindata.org/grapher/co2-by-source?country=~OWID_WRL

[21] – Lacis, A. A., Schmidt, G. A., Rind, D., & Ruedy, R. A. (2010). Atmospheric CO2: Principal Control Knob Governing Earth’s Temperature. Science, 330(6002), 356–359. https://doi.org/10.1126/science.1190653

[22] – Visualizing the Size and Strength of Hurricane Patricia. (2015, October 24). The New York Times Company. https://www.nytimes.com/interactive/2015/10/23/world/americas/hurricane-patricia.html?_r=0

[23] – Flooding in Northeastern China. (2021). NASA: Climate Change and Global Warming. https://climate.nasa.gov/images-of-change?id=754#754-flooding-in-northeastern-china

[24] – Hoskins, A. J., Harwood, T. D., Ware, C., Williams, K. J., Perry, J. J., Ota, N., Croft, J. R., Yeates, D. K., Jetz, W., Golebiewski, M., Purvis, A., Robertson, T., & Ferrier, S. (2020). BILBI: Supporting global biodiversity assessment through high-resolution macroecological modelling. Environmental Modelling & Software, 132, 104806. https://doi.org/10.1016/j.envsoft.2020.104806

[25] – Ruan, G., Wu, D., Zheng, X., Zhong, H., Kang, C., Dahleh, M. A., Sivaranjani, S., & Xie, L. (2020). A Cross-Domain Approach to Analyzing the Short-Run Impact of COVID-19 on the US Electricity Sector. Joule, 4(11), 2322–2337. https://doi.org/10.1016/j.joule.2020.08.017

[26] – Cyprus Average Temperatures. (n.d.). Cyprus Weather. https://www.cyprus-weather.com/article/cyprus-average-temperatures.html

[27] – Highest-ever recorded temperature, 46.2C, clocked at Athalassa, EAC consumption also hits all-time high. (2020, September 4). Cyprus Mail. https://cyprus-mail.com/2020/09/04/highest-ever-recorded-temperature-46-2c-clocked-at-athalassa-eac-consumption-also-hits-all-time-high/

[28] – Department of Meteorology. http://www.moa.gov.cy/moa/dm/dm.nsf/home_en/home_en?openform

[29] – Trotta, G. (2020). An empirical analysis of domestic electricity load profiles: Who consumes how much and when? Applied Energy, 275, 115399. https://doi.org/10.1016/j.apenergy.2020.115399

[30] – Santamouris, M., Alevizos, S. M., Aslanoglou, L., Mantzios, D., Milonas, P., Sarelli, I., Karatasou, S., Cartalis, K., & Paravantis, J. A. (2014). Freezing the poor—Indoor environmental quality in low and very low income households during the winter period in Athens. Energy and Buildings, 70, 61–70. https://doi.org/10.1016/j.enbuild.2013.11.074

[31] – Household budget survey – Mean annual consumption expenditure, 2015/2016. Statistical Service of Cyprus.
https://www.mof.gov.cy/mof/cystat/statistics.nsf/populationcondition_25main_en/populationcondition_25main_en?OpenForm&sub=5&sel=2

[32] – Just About Cyprus. https://www.justaboutcyprus.com/

[33] – Robinson, A. C., Demsar, U., Moore, A. B., Buckley, A., Jiang, B., Field, K., Kraak, M. J., Camboim, S. P., & Sluter, C. R. (2017). Geospatial big data and cartography: research challenges and opportunities for making maps that matter. International Journal of Cartography, 3(sup1), 32–60. https://doi.org/10.1080/23729333.2016.1278151

[34] – Paas, L. (2016). Action for climate empowerment. Van Haren Publishing.

[35] – Mader, A., Eggink, W. (2014). A Design Process for Creative Technology. In. Eger, A., Kovacevic, A., & Parkinson, B. E. (2014). Design Education & Human Technology Relations - Proceedings of the 16th International Conference on Engineering and Product Design Education (E&pde14). (pp. 568-573). The Design Society.

[36] – California earthquakes demo. Kepler.gl. https://kepler.gl/demo/earthquakes

[37] – Social Electricity. https://www.social-electricity.com

[38] – Zimmermann, J. P., (2009). End-use metering campaign in 400 households in Sweden. Assessment of the potential electricity savings. Enertech.

[39] – Tourism Infographic. (2020). Statistical Service of Cyprus. https://www.mof.gov.cy/mof/cystat/statistics.nsf/infographics_en/infographics_en?OpenDocument

[40] – Domestic Use Tariffs. (2021). Electricity Authority of Cyprus. https://www.eac.com.cy/EN/RegulatedActivities/Supply/tariffs/Pages/supply-tariffs.aspx