# Energy Consumption in the Netherlands: a Data Visualization

C.S. Siksma - 0902066, C.M.R. Verploegen - 0940284

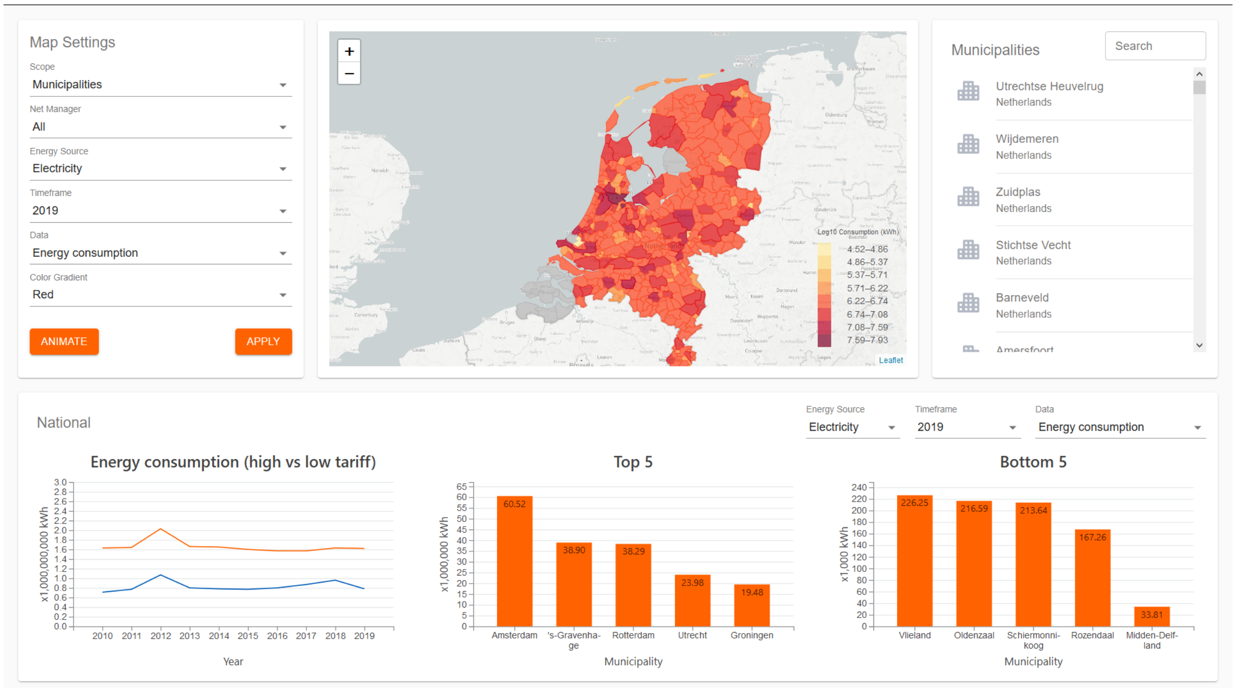


Fig. 1. Overview of the visualization tool.

**Abstract**—The energy usage field is very closely monitored by the three major network administrators in the Netherlands. Data like the annual electricity and gas consumption, energy production and amount of smart meters is openly available. This dataset is very extensive and valuable insights for local and national governments and the network administrators could be derived from it. To explore, lookup, compare and summarize the data an interactive visualization was brought to life. The visualization approach is validated by trying to answer carefully selected high level questions that are of interest to the stakeholders. One such question is 'What municipalities have outstanding high or low energy consumption levels?'. The findings are discussed by showing visuals from the tool, like the heat map and different graphs and charts from the data panel.

**Index Terms**—Energy Consumption, Energy Production, Smartmeters, Network Administrators, Data Visualization, The Netherlands.

## 1 Introduction

Various studies have looked at historical data for demographic trends and energy consumption of countries or groups of countries. Important factors which impact the energy consumption were investigated, like the population size, age structure, economic development and urbanization [2]. We intend to look into the general trends and developments within energy usage of the Netherlands, its municipalities, districts and neighborhoods. It was chosen to focus mainly on the data in context of the timeframe and the geographic location of the Netherlands. To analyze the data in a sophisticated way and to highlight important aspects of the dataset, an interactive visualization was needed. The visualization aims to provide detailed insights and an intuitive way of exploring the data regarding energy usage in the Netherlands that is publicly available. Herewith, exploring the energy data is made accessible to anyone that might be interested in certain aspects of the data. The main stakeholders are network managers, energy providers, local governments, the national government, individual energy users and researchers from various disciplines. visualization or national and local governments could use the gained insights for various applications.

In the Netherlands, the energy usage of each address is closely measured and documented. The energy usage consists of both the usage of electricity and gas, which are separate quantities in the dataset. Enexis, Liander, and Stedin are the three major network managers, together they help distribute energy to nearly the entire country. Every year, they release a dataset with energy usage of the areas under their administration on their website. The data is openly available and is anonymized by aggregating the zip codes, such that every entry describes at least ten connections to the network. This market is not competitive as the zones are assigned to the network administrators every year. This also means that they distribute energy to roughly the same zip codes each year. Small changes can happen from year to year either due to a management change or because of a different aggregation of zip codes [1]. With energy usage and a future of sustainable energy being a hot topic of discussion, we look into the energy usage data of both gas and electricity over the past ten years and aim to get a clear idea of what the recent trends in this field are.

A structured method was used to develop a proper visualization,

• *C.S. Siksma and C.M.R. Verploegen are with Eindhoven University of Technology. E-mail: {c.s.siskma, c.m.r.verploegen}@student.tue.nl.*

which is described in this paper. First of all, a problem description and task analysis were performed. The task analysis gives us clear indications of what potential use cases of the visualization could be for all involved stakeholders. From the insights of this section, a visualization design was created. To analyze the effectiveness of the visualization several extensive use cases were studied. The results are provided and analyzed, which are discussed in the end to evaluate whether the goals of the visualization were reached. N.b: in this report and the visualization we will use an English translation to refer to the Dutch notions of gemeente (municipality), wijk (district) and buurt (neighborhood).

## 2 Problem Description and Task Analysis

The problem that is being investigated consists of several different parts. The main part is the energy data provided by the three main network manager of the Netherlands. This combined dataset is what is most insightful to the stakeholders and from which the main findings should be drawn. However, to be able to explore the dataset in an easy and sophisticated way, more elements are needed. A map would be useful as to explore the different areas of the Netherlands. To be able to create a map of the Netherlands, GeoJSON data is needed from the various areas at different levels, namely municipalities, districts and neighborhoods. The only problem left is how to combine the names of the areas to the data points of the energy dataset, which only includes zip code ranges. Luckily, yet another dataset was found that contains this exact data. A more detailed explanation of how the datasets were pre-processed, combined and formed to become one useful dataset is given below.

### Data

The energy data is obtained from a source that aggregated the energy consumption data of the three major network administrators (Enexis, Liander and Stedin) in the Netherlands [1]. Together they cover nearly the entire country, which makes this a very comprehensive and insightful dataset to work with. Data is available for the energy usage over the past 10 years (2010 till 2019), with data points for each year. The energy usage metrics are split into electricity and gas per zip code range per year. A zip code range consists of at least 10 connections to the network. The most interesting attributes per data point (zip code range) of the dataset are:

- **zipcodefrom** and **zipcodeto**: Zip code range
- **street**: Name of the street
- **city**: Name of the city
- **net_manager**: Name of the regional network administrator
- **num_connections**: Number of connections in this zip code range
- **perc_of_active_connections**: Percentage of active connections in this zip code range
- **annual_consume**: Annual consumption of electricity (in kWh) or gas (in $m^3$)
- **annual_consume_lowtarif_perc**: Percentage of the annual consumption during low tarif hours (10pm to 7am and during weekends)
- **delivery_perc**: Percentage of the net consumption of electricity or gas. The higher, the more energy was given back to the grid (for example if you have solar panels)
- **smartmeter_perc**: Percentage of connections that have a smart meter in the zip code range

There are 20 csv files from each network administrator, two for every year from 2009 through 2019 (one for electricity and one for gas data). As the data from the year of 2009 appeared to be incomplete we dropped this year from our visualization and analysis and only consider the years 2010 through 2019. We used a translation table obtained from the CBS to map zip codes to the unique identifiers of the municipality, district and/or neighborhood [4]. This transformation resulted in 6 extra columns.

- **gemeente2019**: Municipality code
- **gemeentenaam2019**: Name of municipality
- **wijk2019**: District code
- **wijknaam2019**: Name of the district
- **buurt2019**: Neighborhood code
- **buurtnaam2019**: Name of the neighborhood

Next to that, three GeoJSON files were used to map the tuples to the geographic areas of the different municipalities, districts and neighborhoods of the Netherlands. This GeoJSON was obtained from PDOK (Publieke Dienstverlening Op de Kaart), retrieved from the CBS (Centraal Bureau voor de Statistiek) [7]. The maps are simplified using mapshaper [9]. The GeoJSON data is used to create an interactive map of the Netherlands, which is the main exploration tool of the developed visualization. With the GeoJSON data, we can now effectively use all the entries in our dataset and aggregate them and map them to a geographical polygon, which we can then plot on an interactive map.

All datasets were combined per category into one table for electricity and one for gas (with the year as extra attribute for each zip code range). In total, there are 6,812,758 data points in the energy dataset; 3,578,945 data points for electricity data, which are spread over 341 municipalities, 2970 districts or 12696 neighborhoods and 3,233,813 data points for gas data, which are spread over 334 municipalities, 2762 districts or 11868 neighborhoods. The difference in numbers is caused by the difference in geographical areas that is covered by the three network managers.

### Tasks and questions

There are 4 categories of high-level tasks that are relevant for our problem, namely to explore, lookup, compare and summarize the data. Detailed tasks and questions are given below for each category. Each given task is relevant for at least one stakeholder in our problem. The most relevant low-level tasks for each of the 4 categories are listed below. N.b: with the term energy we refer to both electricity and gas, and with area we refer to either municipality, district or neighborhood.

**Explore**

- Trends in energy consumption over time
- Trends in energy production over time
- Trends in network administrator market share
- Which network administrator is active in which area
- How are the smart meters spreading through the Netherlands

**Lookup**

- Area by name
- Area by geographic location
- Energy consumption of an area for a particular timeframe
- Energy production of an area for a particular timeframe
- 5 highest and lowest energy consuming areas for a particular timeframe
- 5 highest and lowest energy producing areas for a particular timeframe
- Average energy consumption per connection of an area for a particular timeframe
- Average energy production per connection of an area for a particular timeframe
- Total connections in an area for a particular timeframe
- Smart meter usage of an area for a particular timeframe
- Network manager market share of an area for a particular timeframe

**Compare**

- Energy consumption between areas and/or timeframes
- Energy production between areas and/or timeframes
- Average energy consumption per connection between areas and/or timeframes
- Average energy production per connection between areas and/or timeframes
- Total connections between areas and/or timeframes
- Smart meter usage between areas and/or timeframes
- Network manager market share between areas and/or timeframes
- Electricity consumption of an area (high vs low tariff)

**Summarize**

- Energy consumption of the Netherlands for a particular timeframe
- Energy production of the Netherlands for a particular timeframe
- Total connections in the Netherlands for a particular timeframe
- Smart meter usage of the Netherlands for a particular timeframe
- Average energy consumption per connection in the Netherlands for a particular timeframe
- Average energy production per connection in the Netherlands for a particular timeframe
- Network manager market share in the Netherlands for a particular timeframe

From these different tasks, high level questions could be derived. The high level questions are important to most of the stakeholders mentioned above, since the questions can guide them to gather insights from the visualization tool. The formulated questions are;

- What municipalities have outstanding high or low energy consumption levels?
- How are the smart meters spreading through the Netherlands?
- What is the trend of electricity and gas consumption?
- What is the trend of home-produced energy?
- What is the trend of network manager market share?
- What is going to be the energy consumption for 2020?

## 3 Visualization Design

With the use of the problem description and extensive task list the design choices for the visualization could be made. The entire interface of the visualization is shown in Figure 1. A general idea for the visualization was to have an interactive map at the center of the visualization, which would make exploring the data simple and lets you easily select and compare different municipalities, districts or neighborhoods. Another useful feature might be to lookup a specific area by searching for it, so a list with all available areas and search feature would come in handy. The map gives us a good indication of which area is selected and what the exact geographical location and size of this area is. When thinking about how to quickly get an indication what interesting findings in the data could be, we came up with the idea to turn this map into a heat map, with a colored gradient that indicates whether the area scores low or high on the selected data feature (e.g. electricity consumption). The color coding is an intuitive way of exploring which areas might be of most interest to the user. In the first implementation we noticed that the gradient using absolute values for (low tariff) total consumption and total connections was not very useful as the larger cities outshined the smaller cities by a number of magnitudes. We decided that a $log_1 0$ transformation of the data was the most fitting to make the data more normal and it gave us the visually desired result with an even spread across the gradient colors. The next feature that was needed was a way to change the settings of the heat map. For example, the user might be interested to change the area scope from municipality to district or to change the data feature from energy consumption to percentage of smart meters. Other interesting settings are the timeframe, net manager and energy source (electricity or gas). The visualization felt quite static up till now and selecting a different timeframe was a recurrent task. Therefore, we came to the conclusion it is very insightful and helpful to add an animation functionality to the map, which lets the user visually see the changes in data values as well as geographically over the years. Using this animation feature, the user can more easily pick up changes in the data such as the energy consumption or network administrator coverage over the ten years of available data. When interesting changes are noticed, the user could then investigate what is going on in the data underneath. Trivially, to use the animation feature the user is required to select a timeframe that consists of 2 or more years.

When an interesting candidate area to investigate is found, or the user wants to explore a specific area, the user can select this area by clicking it on the map or in the list and consequently clicking the explore button from the area's popup window. The specific data from this area is then requested from the server and loaded into the data panel at the bottom of the interface. Initially, the latest (2019) electricity consumption data of the area is presented to the user through several different graphs and charts. The settings can easily be changed by selecting another timeframe, energy source or data type from the drop-down menus. Exploring the data of this area is made straightforward. Comparing different children areas (municipalities for the country, districts for municipalities, neighbourhoods for districts) is easy since the top and bottom 5 children areas are visualized in a bar chart. When working on this data panel, we realized that a summary of the national data would also be useful and could be implemented in the exact same data panel. It was chosen to pre-request the summarized national data and load it initially when the visualization is opened. Among others, the user can immediately get insights into which municipalities contribute most and least to the energy consumption and production. All features for specific areas are generalized into features that also work with the national data. The user is thus free to select which data features they might be interested in. The list of available features is: energy consumption, consumption per connection, energy production, production per connection, number of smart meters, number of connections and network manager market share. The four individual panels of which the interface exists, are discussed in even more detail below. The choice for several design decisions are elaborated upon as well.

**(Heat) map, top center**

The main functionalities of the heat map are exploring the data geographically on a national level, getting the first insights into what interesting aspects of the data might be and navigating or selecting the specific geographic area you want to get detailed information about. The areas can easily be compared by looking at the color of the areas, as the color of the area depends on the value of the selected data feature. Generally holds, the higher the value, the darker the color. The heat map also functions as an easy way to see in which areas (provinces, municipalities, districts or neighborhoods) the energy consumption or production are highest, this might be interesting for local and national governments or the regional network administrators. When the network market share setting is selected, an overview of which network administrator is active in which area is shown. The heat map is mostly designed to cover the explore and summarize tasks. The map is built using the Leaflet JavaScript library, which uses the GeoJSON data that was matched to the areas. The library provides basic interaction features such as zooming and navigating the map. A legend was added in the bottom right corner, which indicates what color corresponds to what standardized value of the selected setting. When a certain area is selected by clicking on it with the mouse, a popup will open that will specify the municipality, district and/or neighborhood the selected area is in. It also contains an explore button that lets you select this area for further investigation, loading the data for this area in the data panel consequently.

**Map settings, top left**

The map settings are a vital element of the interface, since they control what data is visualized in the heat map. In this panel, all settings can easily be changed by selecting the desired setting from the individual drop-down menus. The following settings are available; Scope (municipalities, districts, neighborhoods), Net Manager (all, Enexis, Liander, Stedin), Energy Source (electricity, gas), Timeframe (2019, 2018, 2017, 2016, 2015, 2014, 2013, 2012, 2011, 2010, last two year, last three years, last five years, last ten years), Data (energy consumption, energy low tariff consumption, percentage of smart meters, number of connections, percentage of active connections, delivery percentage, network manager market share), Color Gradient (red, green, blue, purple). At the bottom of the settings panel, two buttons are located. One of the buttons simply applies the selected settings to the heat map visual, the animate button applies the current settings to each year in the selected timeframe with a one second delay until all years have been shown.

**Area list, top right**

The area list is a simple list of areas that can be selected from. The list will show all municipalities, districts or neighborhoods of the Netherlands depending on the selected scope in the map settings. When a certain area is selected from the list, the area will be highlighted in the heat map by showing a popup and make the full area visible in the map window. A search feature is added at the top of this list, to make it easy to search for a specific area the user wants to look into.

**Data panel, bottom**

The data panel is an extensive panel which presents the selected data in a structured way to the user. In the top left corner, the name of the selected area is displayed. This can either be national, municipality, district and/or neighborhood to represent the scope of the area. In the top right corner, all data panel settings can be changed by selecting the desired setting from the individual drop-down menus. The following settings are available; Energy Source (electricity, gas), Timeframe (2019, 2018, 2017, 2016, 2015, 2014, 2013, 2012, 2011, 2010, last two year, last three years, last five years, last ten years), Data (energy consumption, consumption per connection, energy production, production per connection, number of connections, number of smart meters, network manager market share). The most interesting part of the data panel are the several different graphs and charts that are displayed at the bottom, for each of the different data settings that can be selected. For the first six data settings (energy consumption till number of smart meters), a graph is presented that shows the trend of that data value over the years. Next to that, two bar charts are presented that show the top and bottom 5 children (district for municipality etc.) of the selected area. For the network manager market share a different approach was needed. This data would be better visualized by using pie charts, with parts representing the market share of each network manager. The features that are looked into regarding market share are the total consumption, total production and total number of connections the network manager have.

For the trends over the years, a line graph was chosen as the best possible representation. The graph presents the data points over the years in a clear overview, with a line through the data points to easily see whether the value changes over the years by looking at its slope. For the top and bottom 5 children, a bar chart was deemed most suitable. The individual children can easily be compared by looking at the height of the bars and the value is also displayed in the bar itself for more detailed information. Both the graph and bar chart have tick marks and horizontal grid lines, to make comparing different values even simpler. The pie charts have the specific value annotated in the parts and are color coded for the different network managers. The data panel therefore covers most of the lookup and compare tasks.

To link back to the high-level questions that were formulated earlier on, we look at which parts of the visualization cover which questions. 'What municipalities have outstanding high or low energy consumption levels?' is covered by the data panel. When no selection is made, the national data is presented, in which the top and bottom 5 municipalities are displayed in the bar chart. 'How are the smart meters spreading through the Netherlands?' is covered by the heat map and map settings, which lets you select percentage of smart meters and then visualizes the highest percentages over the years by using dark colors. A dark part of the map is therefore an area where most relatively a lot of smart meters are located. 'What is the trend in electricity and gas consumption?' is covered by the data panel. The national data shows the trends for electricity consumption, as well as for gas consumption when selected. From these two trend graphs it can be derived what the trend is in electricity and gas consumption and whether there could be a relation between the two. 'What is the trend of home-produced energy?' is covered by the heat map and map settings, which lets you select delivery percentage and then visualizes the highest percentages over the years for each area. When the animation functionality is used, the user can investigate how the delivery percentage changed over the years. 'What is the trend of network manager market share?' is covered by the heat map and map settings, as well as the national data in the data panel. In the heat map, the color of each network manager is shown to represent the manager of that area. When using the animation functionality, the changing in the covered area by each manager can be investigated. In the data panel, the network manager market share is summarized and presented in pie charts as well. 'What is going to be the energy consumption for 2020?' is mostly covered by the data panel, but can be predicted by using the entire visualization. In the data panel, trends for the annual energy consumption are displayed for the past ten years. When looking at this line graph, the user might be able to make predictions about the energy consumption for 2020. To conclude, all low-level tasks and high-level questions can be answered through the visualization by design.

## 4 Use Cases

To actually validate whether all low-level tasks and high-level questions can be answered, use cases were performed on each of the tasks. Furthermore, all questions are being answered with a detailed description of each step.

### Explore

**Tasks**: Trends in energy consumption over time, Trends in energy production over time, Trends in network administrator market share. In the map setting panel, select data setting respectively 'energy consumption', 'delivery percentage' or 'network manager market share' and click 'animate' button. An animation will be shown of the trends for the selected feature of the data (translated to colors) for each area.

**Task**: Which network administrator is active in which area. In the map setting panel, select data setting 'network manager market share' and click 'animate' button. An animation will be shown of which network manager (color) is active in each area.

**Task**: How are the smart meters spreading through the Netherlands. In the map setting panel, select data setting 'percentage of smart meters' and click 'animate' button. An animation will be shown of the percentage of smart meters (translated to colors) for each area.

### Lookup

**Task**: Area by name. In the area list, either look for an area in the list or use the search functionality to find a specific area.

**Task**: Area by geographic location. In the map view, navigate to the geographical area you are interested in. By clicking on it, the name of this area is given, as well as the option to explore this area.

**Tasks**: Energy consumption, Energy production, Average energy consumption per connection, Average production per connection, Total connections, Smart meter usage (of an area for a particular timeframe. In the map or from the list of areas, select an area and click 'explore' button. The data of that area will be loaded in the data panel. In the data panel, select data setting respectively 'energy consumption', 'energy production', 'consumption per connection', 'production per connection', 'number of connections', 'number of smart meters'. The selected data feature will be calculated. For example, the average energy consumption/production will be calculated by dividing the total energy consumption by the number of active connections in this area. The trend line graph of the data feature value in this area will be shown. The timeframe can be selected by changing the timeframe drop-down menu to another year or timeframe.

**Tasks**: 5 highest and lowest energy consuming areas, 5 highest and lowest energy producting areas (for a particular timeframe). In the map or from the list of areas, select an area and click the 'explore' button. The data of that area will be loaded in the data panel. The top and bottom 5 districts of this municipality will be shown in bar charts. The timeframe can be selected by changing the timeframe drop-down menu to another year or timeframe.

**Task**: network manager market share of an area for a particular timeframe. In the map or from the list of areas, select an area and click the 'explore' button. The data of that area will be loaded in the data panel. Select the data setting 'network manager market share'. The market share will be presented in three different bar charts. One for the total energy consumption market share, one for the total energy production market share and one for the total number of connections.

### COMPARE

**Tasks**: Energy consumption, Energy production, Total connections, Smart meter usage, Network manager market share (between areas and/or timeframes). This task can be completed in two ways, namely through using the heat map or data panel. Both methods will be explained. (1) In the map settings panel, select data setting 'energy consumption' or 'delivery percentage', 'number of connections', 'percentage of smart meters' or 'network manager market share' and click the 'apply' button. The heat map will be shown of the delivery percentage (translated to colors) for each area. (2) In the map or from the list of areas, select an area and click the 'explore' button. The data of that area will be loaded in the data panel. In the data panel, select data setting 'energy consumption' or 'delivery percentage', 'number of connections', 'percentage of smart meters' or 'network manager market share' and the timeframe you want to look at. The top and bottom 5 children of the area will be shown and can be compared to each other in the bar charts.

**Tasks**: Average energy consumption per connection, Average energy production per connection (between areas and/or timeframes). In the map or from the list of areas, select an area and click the 'explore' button. The data of that area will be loaded in the data panel. In the data panel, select data setting 'consumption per connection' or 'production per connection' and the timeframe you want to look at. The top and bottom 5 children of the area will be shown and can be compared to each other in the bar charts.

**Task**: Electricity consumption of an area (high vs low tariff). In the map or from the list of areas, select an area and click the 'explore' button. The data of that area will be loaded in the data panel. In the data panel, select data setting 'energy consumption'. Two trend lines will be shown in this graph for easy comparison (one for the high and one for the low tariff).

### SUMMARIZE

**Tasks**: Energy consumption, Energy production, Total connections, Smart meter usage, Average energy consumption per connection , Average energy production per connection , Network manager market share (in the Netherlands for a particular timeframe). The national data is pre-loaded in the data panel. Select data setting respectively to 'energy consumption', 'energy production', 'number of connections', 'number of smart meters', 'consumption per connection', 'production per connection' or 'network manager market share'. Then select the desired timeframe and explore the graphs and charts of data that is presented to you.

### QUESTIONS

The questions are answered in the following paragraph by a detailed step description. The visualization tool was used to answer each of the questions below.

**Question**: 'What municipalities have outstanding high or low energy consumption levels?'

The national energy data is pre-loaded in the data panel. Two bar charts show the top and bottom 5 municipalities regarding energy consumption in the Netherlands, which is exactly what we need here. The 2019 top 5 regarding electricity data: Amsterdam, The Hague, Rotterdam, Utrecht, Groningen and the bottom 5: Vlieland, Oldenzaal, Schiermonnikoog, Rozendaal, Midden-Delfland. The 2019 top 5 regarding gas data: Amsterdam, The Hague, Rotterdam, Groningen, Eindhoven and the bottom 5: Midden-Delfland, Twenterand, Hof van Twente, Dinkelland, Wierden. When looking at the changes of the outliers, we noticed some interesting details. We see that Amsterdam is almost all the time at the top of this chart with an electricity consumption value of around 60 million kWh. Expect in 2012, where Rotterdam and The Hague take the lead with a value of around 85 million kWh and Utrecht following at 52 million kWh. These values are approximately twice as high as the values of the same cities in other years. This strongly hints to an error in the dataset, where the data might accidentally be included twice for these cities.

**Question**: 'How are the smart meters spreading through the Netherlands?'

In the map settings panel, select data setting 'percentage of smart meters', scope setting 'municipalities' and timeframe setting 'last ten years', then click the 'animate' button. The heat map will be shown of the percentage of smart meters (translated to colors) for each area. When viewing the heat map and seeing the changes over the years, we notice a clear rise in the adaption of smart meters starting in 2014. In 2018, almost half of the municipalities has a percentages of more than 50%. However, in the year 2019 something strange seems to be going on, since abrupt changes in the data occur. There is even a decline in the percentage of smart meters, which does not seem very logical to us. We therefore conclude that the smart meter data for 2019 is corrupted and will not investigate it further. The overall trend from 2010 till 2018 is a slowly growing percentage of smart meters through the entire Netherlands, with a rapid acceleration in the years 2016, 2017 and 2018.

**Question**: 'What is the trend of electricity and gas consumption?'

The national energy data is pre-loaded in the data panel. Select energy source setting 'electricity' or 'gas'. The trend line graph of the energy consumption is shown. From the graph, we can see there was an electricity spike in 2012. We also see there was a gas spike in 2014 and 2015. Further outstanding information can not be drawn from the graph, so it seems like the electricity and gas consumption levels are not very related to each other.

**Question**: 'What is the trend of home-produced energy?'

This trend can be investigated by combining the map view and the data panel. In the map settings panel, select data setting 'delivery percentage', scope setting 'municipalities' and timeframe setting 'last ten years', then click the 'animate' button. The heat map will be shown of the delivery percentage for each area. When viewing the heat map and seeing the changes over the years, we can not notice significant increases or decreases of the percentage. What we do notice, is the corrupted data of the years 2018 and 2019, for which all the data points of Enexis seem to be wrong. We therefore only look into the data from 2010 till 2017. When looking at the data panel with national data loaded in, select the data setting 'energy production' or 'production per connection'. Here we again see that the data is corrupted for 2018 and 2019, we can therefore not be sure whether the rest of the data is correct and refrain from answering this question precisely.

**Question**: 'What is the trend of network manager market share?'

This trend can be investigated by combining the map view and the data panel. In the map settings panel, select data setting 'network manager market share', scope setting 'municipalities' and timeframe setting 'last ten years', then click the 'animate' button. The heat map will be shown of the active network manager for each area (represented by the company color). When viewing the heat map and seeing the changes over the years, we notice some minor changes to which municipality is managed by which network manager. Overall, the network managers seem quite steady in area coverage which is logical since the market is not very competitive and the areas are divided over the network managers each year. However, when we start looking at the data panel, some trends can be derived from the pie charts of each year. Select the data setting 'network manager market share', pie charts come up with the absolute numbers of energy consumption, production and number of connections per network manager. We notice that Stedin's market share is shrinking a little over the years when looking at the consumption and number of connections. Liander seems most steady over the years and Enexis appears to be slightly enlarging their energy consumption values and number of connections. One strange finding is that the number of connections of Enexis grows from 2.76 to 12.09 million from 2017 to 2018.

**Question**: What is going to be the energy consumption for 2020?

To answer this question, we look at the map view as well as the data panel for national data. In the map settings panel, select data setting 'energy consumption', scope setting 'municipalities' and timeframe setting 'last ten years', the click the 'animate' button. The heat map will be shown of the energy consumption for each area per year. When viewing the heat map and seeing the changes over the years, we only notice minor details about the energy consumption. One of such details is that the energy consumption seems to be slightly going up and down in the Randstad. Predictions can not really be made from the map view, except for the energy consumption staying at approximately the same level in 2020 as in 2019. When looking at the data panel, the line graph is of most interest for the energy consumption. The year 2012 that stands out immediately for the electricity data, as a big spike is registered in this year. In all the other years, the total energy consumption seems to be quite steady and slightly above 1.6 billion kWh for the entire Netherlands. We therefore predict that the total electricity consumption in 2020 will be slightly above 1.6 billion kWh as well. When looking at the gas consumption data, a big spike can be seen in 2015 and 2016. After this spike, the gas consumption is back at its 'normal' level in 2018 of around 600 million $m^3$. Since the gas consumption seems steady again, we expect the level in 2020 to be around 600 million $m^3$ as well.

## 5 Discussion

While performing the use cases for each of the low-level tasks and high-level questions, minor flaws of our approach came to light. Furthermore, we noticed lots of positive aspects about the approach. For example, we were easily able to complete all use cases for the low-level tasks, since our visualization was designed with these tasks in mind. The tasks were intuitive and for some tasks there were multiple ways to get the correct results. This lets it up to the user what way they prefer to perform certain actions and gives freedom to the user on how to interact with the visualization tool. The high-level questions could also be answered to a high extent. Extra information that might be interesting to a user asking this question was often available alongside the information the user was looking for. Only the last question was a bit harder to answer and would require a little more work, but it could be answered nonetheless with the help of the visualization.

The flaws that were found of our approach range from implementation issues to design issues, some issues were even out of our control. For example, the provided dataset was not complete for the Netherlands and might therefore be biased when looking at summary data of the entire country. The missing data that was most noticeable to us was all data of the province Zeeland, which let to an incomplete heat map for that area and will not allow you to look into the municipalities that lie within this province. Another very strange piece of missing information was the energy data of the municipality Eindhoven from before 2017. The biggest implementation limitation was the massive size of the dataset, which resulted in wait times of sometimes up to 10 seconds when requesting specific data. In the second iteration of how we query data for the visualization tool, we were able to cut some load times down by caching frequent occuring results. Another technique used was to preload the national data by importing a json file and immediately displaying this data to the user. The startup time of our visualization was cut down through these techniques from several seconds to almost instant to the first interaction and insights.

Next to that, some design flaws were discovered when performing the use cases. It was found that it is quite hard to compare districts that do not belong to the same municipality. One way to perform this task is to compare the colors on the heat map, but this does not give absolute values and can be hard for people with forms of colorblindness. A strength of our tool is that we tried to make the tool as accessible as possible for all different people, so the colors are as differential as possible. According to literature, about 8% of the male population has some sort of color vision disturbance [3]. For accessibility and aesthetic purposes, we included a feature that lets the user select which color gradient is used in the heat map. The color gradient is also based on the principle, the darker, the higher the corresponding value. It might thus be readable by colorblind people as well. The color gradient uses standardized magnitudes, which is good since it is not very subjective to outliers and shows more clear differences between all levels of data points. Another way to compare districts is by selecting the districts the user is interested in one by one and comparing the results. This is not very intuitive and might be very hard for a user to do in a sophisticated way. Remembering data and graphs is not very viable, so as of now the user would have to resort to screen shots and writing down of results.

## 6 Analysis and Conclusion

**Observations**

We have made some observations about the dataset that only became clear after implementing and using the visualization.

- The data from 2009 appears to be incomplete.
- There appears to be no smart meter data for the net manager of Liander in 2019.
- The delivery_perc value for Enexis in 2018 and 2019 is close to 100%, indicating a possible change in definition, measurement method and/or faulty data.
- There is no data for the city of Eindhoven before 2017.
- There seems to be a non-zero delivery_perc of gas in many areas but this seems very unlikely.
- There is no data for Zeeland for any year, the regional net manager for Zeeland is Enduris which is not included in our dataset [5].

**Findings**

To demonstrate the capabilities of our visualization and the extent to which analysis can be done we will share some of our findings. The process of this analysis using the visualizaiton is also made available in a narrated screencast. We start by looking at the Netherlands as a whole. Using the visualization we find that there has been little change in electricity consumption, it has stayed roughly the same over the past 10 years with a peak in 2012 as can be seen in figure 6. This also holds for the low-tarif electricity consumption. For gas consumption it has also stayed roughly the same but we see a larger peak that started forming in 2013, peaking in late 2014 and lowering again to the normal level in 2016 as can be seen in figure 6. We could have expected a decrease in gas consumption due to the phasing out of gas in the Netherlands which had been in talks since 2016, but as the law has only been in effect since 2018 and only considers new buildings, it is likely too early to notice a significant decrease [8].

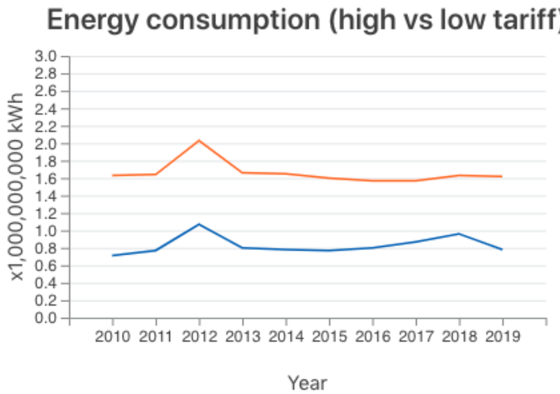


Fig. 2. Electricity consumption of the Netherlands.

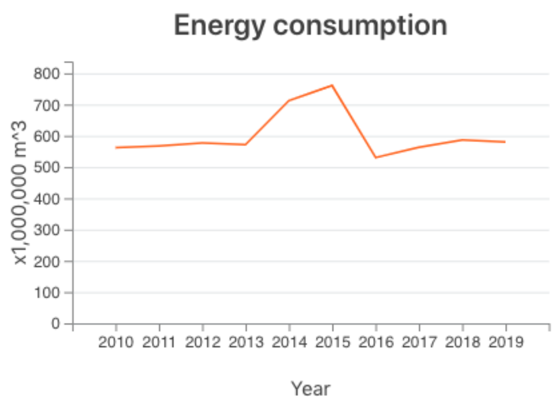


Fig. 3. Gas consumption of the Netherlands.

Looking at consumption per connection however, we notice a significant decrease since 2017 as can be seen in figure 6. The total energy consumption is thus likely to be the cause of a growing number connections, which is not very strange since the population of the Netherlands has grown by almost a million people since 2010 [6]. From figure 6 we also see a difference of roughly 10 million connections between 2010 and 2018, though the non-linearity of graph would mean that these connections are not counted yearly but we cannot verify this. From this we can derive that the consumption per household has gone down quite significantly, possibly due to climate awareness or introduction of smart and energy-saving electronics, but is balanced out by the influx of new connections.

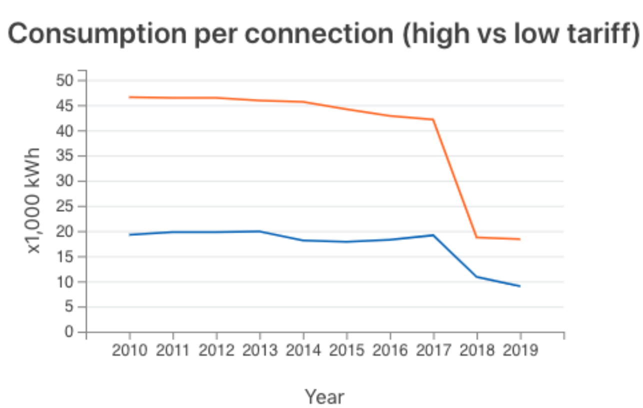


Fig. 4. Average consumption per connection of the Netherlands.

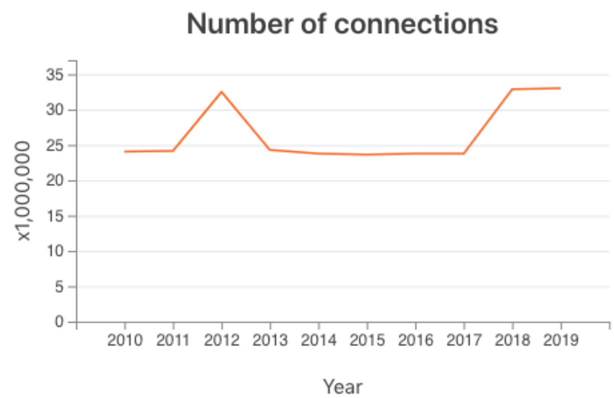


Fig. 5. Number of connections in the Netherlands.

For electricity production in figure 6 we notice a rather large jump up in total electricity production starting in 2017 and from figure 6 we see that the average production per connection has been rising steadily since 2010 and has almost doubled in the last ten years. This would lead us to believe that the use of solar panels is getting more common throughout the years.

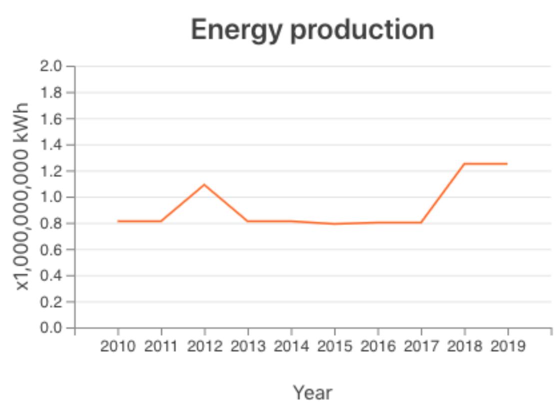


Fig. 6. Electricity production of the Netherlands.

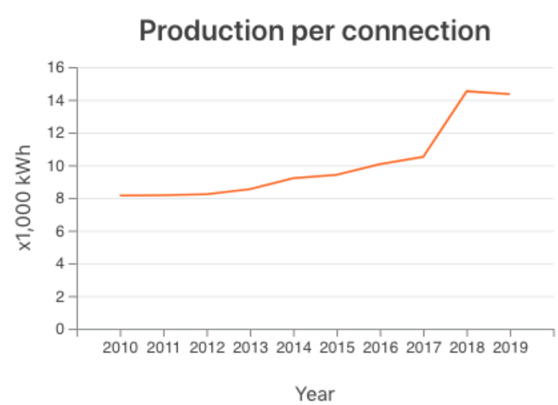


Fig. 7. Average electricity production per connection in the Netherlands.

Something really interesting becomes clear when we look at the adoption of smart meters throughout the Netherlands. Whereas in 2010, roughly 1 million people used a smart meter, in 2018 this amount has increased to almost 16 million. From figure 6 we see that the pattern follows an exponential increase and if this trend continues we can

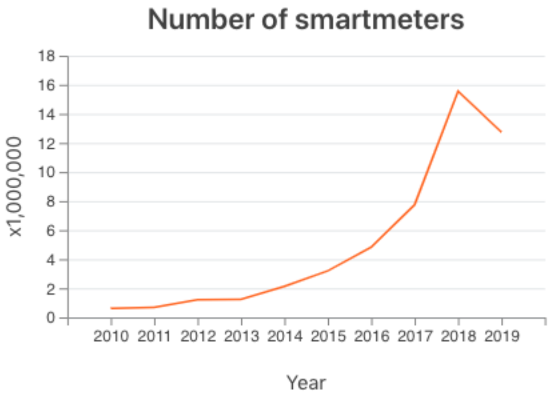


Fig. 8. Number of smart meters in the Netherlands.

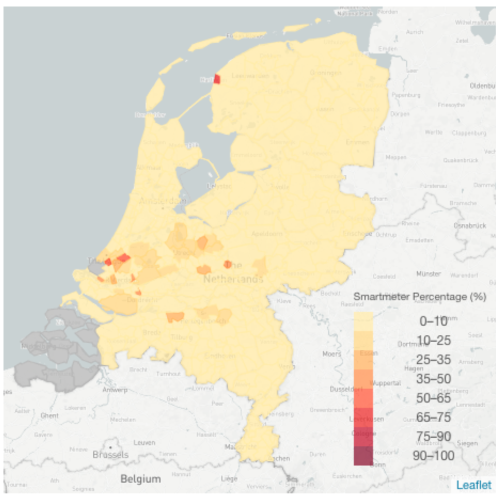


Fig. 9. Percentage of smartmeters in the Netherlands in 2010.

expect almost all of the connections to be using a smart meter by the end of 2020.

We also look at how the smart meter adoption has spread throughout the country geographically by using the animation feature for the smart meter percentage over the past 10 years. In 2010 (figure 6), only a few pioneer cities have a high smart meter adoption, such as Harlingen and The Hague which then grows to other areas. By 2018 (figure 6 it seems that most major cities have adopted them.

Zooming in on Eindhoven, we show how rapid and recent the adoption of smart meters has been. We first look at the neighborhoods of Eindhoven in 2017 in figure 6, where most neighborhoods still only have an adoption of less than 35%. Only Strijp-S has a smart meter percentage of over 90%, explained by the huge number of new apartment that were recently built with modern standards. In 2019 (figure 6) we find that almost all neighborhoods in Eindhoven have a smart meter adoption rate of over 75%.

Throughout the years, Amsterdam has been the highest consuming municipality such as in 2018 as can be seen in figure 6. Using roughly 60 million kWh per year, it is not very strange as the capital of the Netherlands with highest population number. However, in 2012 we do notice that both Rotterdam and the Hague net over 80 million kWh in usage, about twice as much as their usual number which is around 40 million kWh, see figure 6. This means the peak we noticed for 2012 in figure 6 is likely from these two cities having almost twice the electricity consumption that year than normal. The reason for this is something we leave to a domain expert.

The network manager market share has also changed significantly over the years. Notably, looking at the region changes over the years

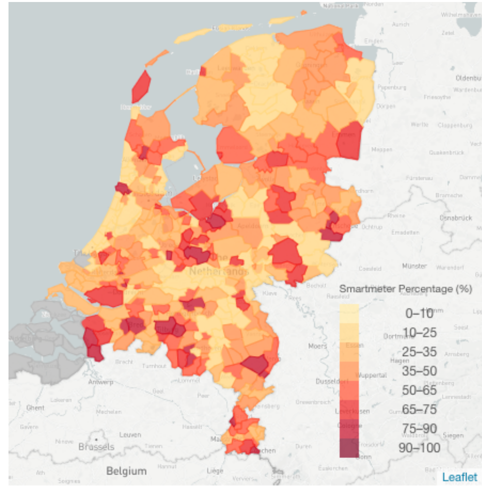


Fig. 10. Percentage of smartmeters in the Netherlands in 2018.

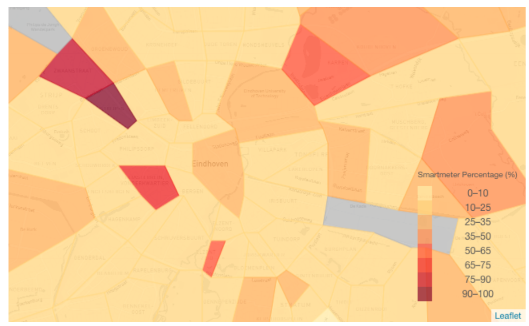


Fig. 11. Percentage of smartmeters by neighborhood in Eindhoven in 2017.

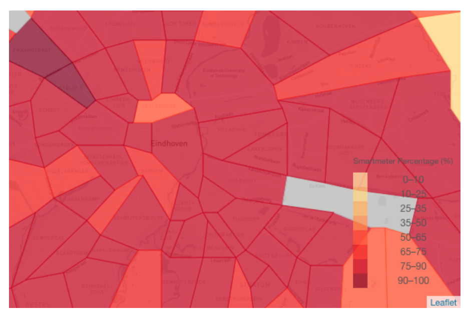


Fig. 12. Percentage of smartmeters by neighborhood in Eindhoven in 2019.

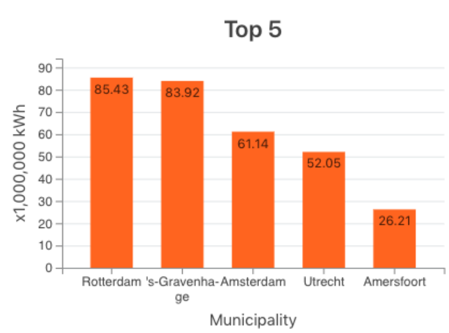


Fig. 13. Top 5 electricity consumption of the Netherlands in 2012.

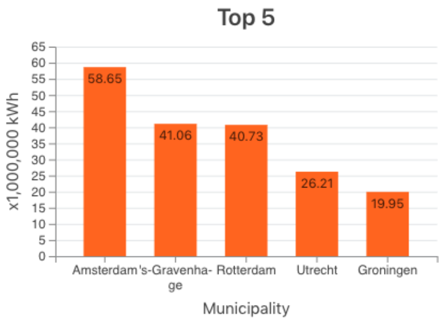


Fig. 14. Top 5 electricity consumption of the Netherlands in 2018.

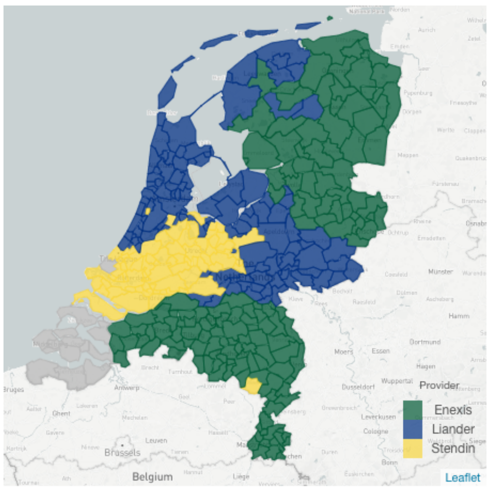


Fig. 15. Net manager regions of the Netherlands in 2010.

(figure 6 and figure 6 we notice very little difference but looking at the number of connections covered by each network manager, we see a huge difference between figure 6 and **??**. Again, this is likely due to the influx of new people living in the Netherlands, where we have almost 10 million new connections since 2010.

**Conclusion**

To conclude, the visualization tool is very extensive and most parts are intuitive to interact with. Interaction design and animation are used to enhance the experience of the user of this kind of tool and assist in getting as much useful information out of the dataset as possible. A new light was shed on this dataset by combining it with the GeoJSON data of municipalities, districts and neighborhoods. The summarized national data could be of great interest for the mentioned stakeholders. The high resolution of the visualization tool allows for zooming in not just on municipality, but also on district and neighbourhoods. This highlights the highly local differences which is what differentiates our tool from others alike. All tasks and questions defined in the task analysis can be completed and answered by exploiting the visualization tool in different ways. This was once more proven by the use cases and evaluation of our approach.

## Work division

The design of the visualization tool was discussed in meetings and was contributed to by both students. As for the implementation, C.S. Siksma was responsible for a large part of the visualization setup, such as importing, combining and preformatting the datasets, creating a frame for the visualization and setting up the heat map and list view. C.M.R. Verploegen was responsible for the data panel with its various graphs and charts (D3.js) and worked on querying the data in a structured way. Writing and improving the report was a combined effort. Overall, effort

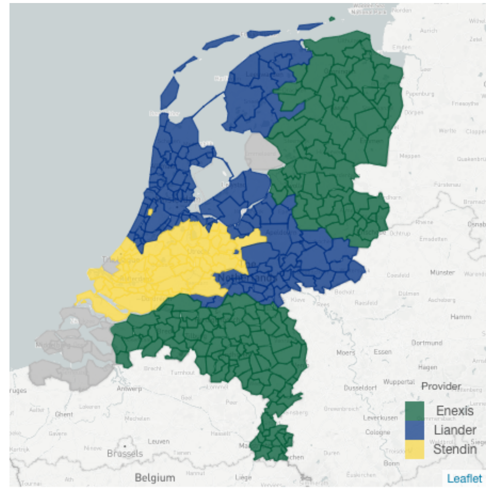


Fig. 16. Net manager regions of the Netherlands in 2019.

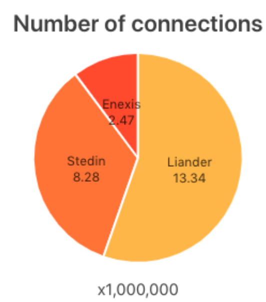


Fig. 17. Net manager distribution of the Netherlands in 2012.

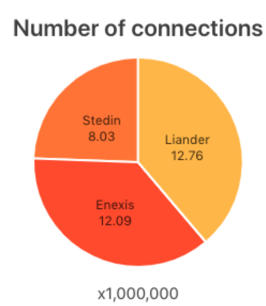


Fig. 18. Net manager distribution of the Netherlands in 2018.

and time spend was fairly equal.

## REFERENCES


1. Basanisi, L. (2019, March 14). Energy Consumption of the Netherlands. Retrieved December 11, 2019, from https://www.kaggle.com/lucabasa/dutch-energy

2. York, R. (2007). Demographic trends and energy consumption in European Union Nations, 1960–2025. Social science research, 36(3), 855-872.

3. Lee, J., & dos Santos, W. P. (2011). An adaptive fuzzy-based system to simulate, quantify and compensate color blindness. Integrated Computer-Aided Engineering, 18(1), 29-40.

4. CBS Nederland. (z.d.). Buurt, wijk en gemeente 2019 voor postcode huisnummer [Dataset]. Retrieved December 2019, from https://www.cbs.nl/nl-nl/maatwerk/2019/42/buurt-wijk-en-gemeente-2019-voor-postcode-huisnummer

5. Enduris. (2015). Uw netbeheerder in Zeeland - Enduris. Retrieved February 2020, from https://www.enduris.nl/

6. CBS Nederland. (2020 januari). Bevolkingsontwikkeling; regio per maand [Dataset]. Retrieved February 2020, from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37230ned/table

7. Cartomap GeoJSON Nederland. (z.d.). Retrieved February 2020, from https://cartomap.github.io/nl/

8. Rijksdienst voor Ondernemend Nederland. (z.d.). Aardgasvrij. Retrieved February 2020, from https://www.rvo.nl/onderwerpen/duurzaam-ondernemen/duurzame-energie-opwekken/aardgasvrij

9. Mapshaper. (z.d.). Retrieved February 2020, from https://mapshaper.org/