# A Guide to Data Analysis

## Using energy data to analyze energy situation and factors that contribute to it

**Objective**

The goal of this guide is to provide an example of data analysis and visualization to all dig data applicants who will be working with Energy Regulatory Office and the Kosovo Agency of Energy Efficiency open data on energy. Through the cooperation with the Energy Regulatory Office (ERO) and Kosovo Agency of Energy Efficiency (KAEE) in opening energy data, the 2020 Dig Data Challenge aims to motivate all applicants to provide solutions that are based on data use and analysis outcome. The focus is for applicants to play around with data, analyze the outcome, visualize the data, and come up with a solution that is based on the analysis outcome. As such, if the solution would be implemented, in the long term, it would:

* Improve Kosovo situation in energy, that is portrayed through statistics and numbers
* Improve energy data availability in Kosovo
* Improve energy data quality in Kosovo
* Improve citizens awareness and information regarding energy
* Offer energy efficiency information and ideas that aim to ensure better life to Kosovo citizens

However, the above-mentioned points are only some ideas from the pool of solutions that could be offered by applicants.

This guide will provide information on:

* Downloading and using data from either Energy Regulatory Office or the Kosovo Agency of Energy Efficiency
* Combining datasets to complete the analysis story
* Using Excel for data analysis
* Using Excel for data visualization
* Interpreting the result from data analysis and offering solution based on the outcome

**Step-by-step guide to Excel in data analysis**

Given that energy solutions to households are one of the main focuses of the 2020 Dig Data Challenge, this guide will use household energy consumption data for the example.

***First step:***Download the data you want to work with

To download the data, please visit:

<https://www.ero-ks.org/zrre/sq/te-dhenat>

On the third tab in this link, you will have the opportunity to download the Excel-format dataset, named: “Konsumi shtëpiak sipas muajit dhe njësisë.” Furthermore, in order to add additional information to your analysis, you would want to download another file, named: “Tarifat për konsumatorët shtëpiak”.

Below is a snapshot of the data downloaded:

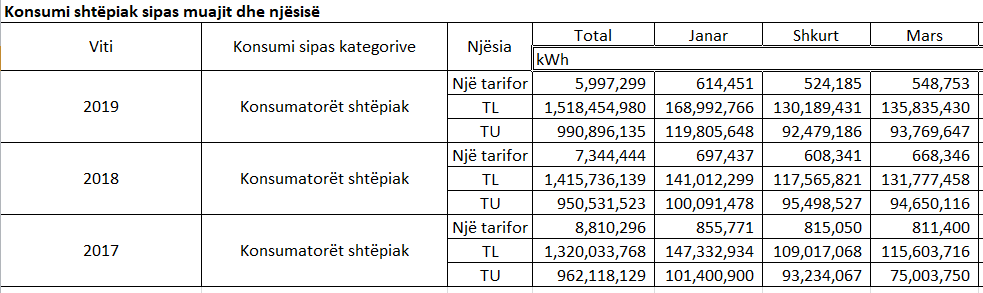


Figure : Data snapshot - Household consumption

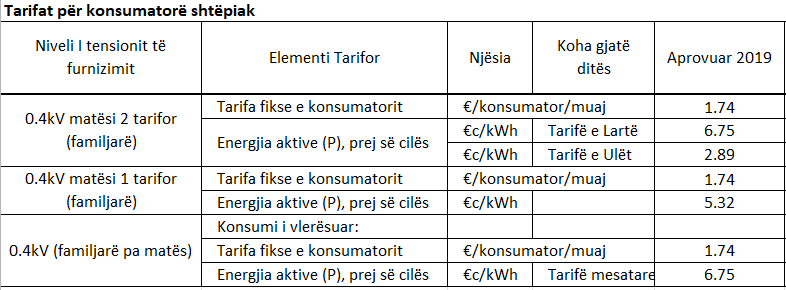


Figure : Data snapshot - Household tariffs

***Second step:*** Prepare and select data for analysis

With this data, the aim would be to analyze the monthly difference in consumption between the high and low tariff. The question we would want to answer is: do customers consume more energy during the high or low tariff?

In order not to congest data, it is recommended to select one year at a time, and analyze the change in consumption from one month to another, between the two categories.

To do so, you would want to select all 2019 row data, including high tariff and low tariff. You would also want to select the first row of data description, which includes months, so that months are clearly portrayed in the graph. When analyzing time-series data, line graphs are preferred, to depict the change in number over the time.

In order to better understand the difference between high and low consumption, you can add a line to the downloaded data and find the percentage difference between the high and low tariff consumption.

You can do this by inserting a new row, and using the formula: =(F5-F4)/F5 (please be aware of the sign and number on your downloaded file). You can drag and apply this formula to all other columns. This will result in the third row, as below:

Table : Calculated difference between high and low tariff consumption



The negative percentages indicate the negative difference between the low and high tariff. As an example, only in January, Kosovo households consumed 41% less energy during the low tariff compared to the high tariff.

Since the data you want to visualize are of two types, number and percentage, the graph can be tailored to present both data. You can do so by clicking the percentage data, and select Format Data Series and then Secondary Axis. This will show another y-axis, to present the percentage change. You can format the axis in the Format area.

You will end up having the following chart:



Figure : Monthly household consumption

**Third step**: Analyze numbers and interpret results

From the data visualization above, it is inferred that in 2019 households have consumed less energy during the low tariff, during the entire year. The consumption trend, does not change between the high and low tariff. As it can be seen, households tend to consume more energy during the winter months, with consumption spikes specifically in December and January, and less energy during summer, specifically during June and July. However, during August and September the gap deepens even further, and households tend to consume even less energy during the low tariff.

Using the second table data, it is important to understand this impact on monetary means. As it can be noticed, there is e vast difference between the high and low tariff. Multiplying the consumption by respective tariff, we get another set of data. This set of data presents total consumption in Euro value, divided by high and low tariff.

Table : Monthly household consumption in Euro



The difference between the low and high tariff can be further noticed in the next table, where, on average, 61 percent of the household consumption happens during the high tariff, while 39% during the low tariff.

Table : Total household consumption (in kWh and Euro)



Due to the difference in price, as well as consumption, the graph below depicts the monthly costs that households incur while consuming energy at high and low tariff.



Figure : Monthly household consumption in Euro

Given the difference in household costs between the high and low tariff, decreasing these costs, would be one of the solutions aimed by the challenge.

Let us assume, that we could switch only 11 percent of the household consumption from high tariff to low tariff, so that household consumption is equal between the two tariffs.[[1]](#footnote-1)

Table : First switch: Total household consumption (in kWh and Euro)



Comparing the total consumption in Euro from Table 4 and Table 3, it can be easily seen that with only 11 percent switch in consumption, total households savings would sum to 10,181,885.71 Euro.

To further see the impact in savings, we can assume to reverse the percentage consumption of household from the first table, meaning for the household high tariff consumption to be at 39 percent, while the low tariff consumption at 61 percent.

Table : Second switch: Total household consumption (in kWh and Euro)



Comparing the total consumption in Euro from Table 5 and Table 3, it can be easily seen that by reversing consumption, total household savings would sum up to 20,363,771.42 Euro.

Having calculated the total household consumption and the impact of switching total household consumption from high to low tariff in total savings, by using “Të dhënat mesatare vjetore dhe mujore të konsumit”, we could also calculate the impact of this switch in individual household savings.



Figure : Data snapshot - Average monthly and yearly consumption data

From Figure 5, we can see that average monthly consumption per household is 405 kWh. From the calculations above, we saw that on average 61 percent of household consumption is during high tariff, while 39 percent during low tariff. Assuming a single household would have the same consumption pattern, and that a household consumes 405 kWh monthly, then the household would pay 21.24 Euro monthly on energy consumption.

Table : Monthly household consumption (in kWh and Euro)



By switching 11 percent of individual household consumption from high to low tariff, households would save 1.72 Euro in monthly bill. Multiplied by 12 months, household would save around 20 Euros yearly.

Table : First switch: Individual household consumption (in kWh and Euro)



By reversing individual household consumption from high to low tariff, households would save 3.44 Euro in monthly bill. Multiplied by 12 months, household would save around 41 Euros yearly.

Table : Second switch: Individual household consumption (in kWh and Euro)



Even though these savings in individual household consumption might seem low, there are two things that need to be considered:

* Monthly household consumption of 405 is a calculated average consumption, and as such there are families who consume a higher amount of energy.
* The household consumption percentage during the high and low tariff is an average consumption, and as such there are families who might consume energy during the high tariff at a higher percentage than 61.

Considering the two above facts, individual household savings might be even higher, and that depends from one house to another.

**Final step:** Provide solution

The final step is to use the analysis provided above to generate solutions that would translate data-driven findings into actual actions for households. In this example, the solution would entail ideas that would motivate Kosovo households to switch a portion of the energy consumption from day to night, and as such from the high tariff to the low tariff.

To do so, a desk research is needed to analyze decision factors that impact energy consumption choices among Kosovo citizens. These might include:

* Lack of information regarding high and low tariff difference in prices
* Lack of information regarding the appropriate timing for low and high tariff
* Lack of means for calculating the difference in savings that a switch in consumption might cause
* Lack of awareness regarding energy saving opportunities
* Lack of time to switch consumption, along with the family routines and work schedule

Through a combination of data analysis provided in the guide, as well as desk research on factors that impact energy savings in Kosovo families, different teams might offer different solutions that tackle the above-mentioned factors, and as such achieve to decrease customers; energy consumption costs.

1. For the analysis purposes, we are ignoring the flat tariff. [↑](#footnote-ref-1)