SUMMARY OF ARTICLE: HTTPS://DX.DOI.ORG/10.12795/REA.2023.I46.04

Analysis of electricity consumption in Andalusia during the year 2020

Manuel Ceballos-González
mceballos@uloyola.es 0000-0003-0913-6417
Javier Antonio Torres-Chacón
jtorreschacon@al.uloyola.es 0000-0001-7160-5849
Universidad Loyola Andalucía.
Avenida de las Universidades, s/n. 41704 Dos Hermanas, Sevilla, España.

KEYWORDS

Statistical study
Electricity consumption
Andalusia

Electricity is one of the most important ways of energy. During the last years, there have been periods of growth and decline in the electricity consumption and, nowadays, society pursue a greater degree of energy autonomy. Due to the current gas issue and the war in Eastern Europe, the electricity price has increased during the last year. This has also been affected due to the current pandemic of the COVID-19. In case of Andalusia, it has a very stable energy framework. Moreover, there exists a very strong regulation and plan-ification based on goals towards renewable energies and energetic efficiency. The geographical location of Andalusia is excellent in terms of solar power and renewable energy sources. However, there is a very high national energy dependence and a high imbalance in the foreign and inland supply prices. As a consequence of this, it is very important to analyse the available data corresponding to energy consumption in order to take actions, increase public awareness and prevent economic crisis in this matter. In this way, the main goal of this paper is to develop a statistical study about the electricity consumption in every city and town of Andalusia which has Endesa Distribución Eléctrica as supplier during the year 2020.

Regarding the methodology, and in order to achieve the goals of the paper, we have used Stata software and a database. Stata is a powerful software which is very common nowadays when dealing with Big Data problems. This program focuses on providing data visualization, modelling, manipulation, statistics and re-porting to users from multiple disciplines. It allows users to manage different types of data simultaneously like panel data, survey data, time-series data, survival data, categorical data, and multilevel data. Stata has three main features: easy instructions, interactive mode and automatic response. Moreover, it provides mathematical functions, probability and density functions, matrix functions, string functions, functions for dealing with dates and time series, and a set of special functions for programmers. The database con-sidered for the statistical analysis corresponds to the Multiterritorial Information System from the website of the Institute of Statistics and Cartography of the Junta de Andalucía. This database offers a wealth of multi-thematic and multi-territorial statistical information (see the Appendix for a complete list of variables). Among these data, we can find the electricity consumption addressed by the company Endesa Distribución Eléctrica during the year 2020. The methodology proposes to export the data to an excel worksheet and import it to Stata program. After that, we develop a statistical study, where we have used the variable average consumption per capita obtained as the quotient between the average consumption of a city divided by the population of that city.

273

Our study begins with the calculation of the most important descriptive statistics. We have obtained an average consumption of 4.079757 MWh and a standard deviation of 8.281704. After that, we have obtained the most important statistical diagrams such as the histogram, box plot and stem and leaf. We have considered in all of them the electricity consumption measured in kW. Then, we have studied the electricity consumption in Andalusia according to each professional sector: agriculture, industry, business-services, residential, administration-public services and others. In Huelva and Cádiz, industry is the sector which gets the highest electricity consumption, while we obtain the residential sector for the remaining regions of Andalusia. After that, we have analysed the atypical and highly atypical data and the type of consumption based on a classification by using the quartiles. In this way, we have distinguished four different types of consumption: low, low-middle, high-middle and high and we have studied how these four types are distributed in each region of Andalusia. In this way, high consumption is the most common in Almería, Córdoba and Jaén, while in Cádiz, Granada, Huelva and Sevilla the low type is the most frequent. In Málaga, although, we find the high-middle type as the most common. Concerning atypical data, our study reveals that Almeria is the region with the greatest number of this kind of data. Then, we have applied a hypothesis test comparing the average consumption of each region in Andalusia studying if there exist significant differences between each of them. Huelva seems to be the region with the highest consumption and after it we have Almeria. On the other hand, Cádiz and Málaga are the regions with the lowest consumption values. Moreover, there are ten couples of regions in Andalusia that have significant differences among them. Those couples are: Almería-Granada, Almería-Huelva, Almería-Málaga, Córdoba-Huelva, Córdoba-Málaga, Granada-Huelva, Huelva-Jaén, Huelva-Málaga, Huelva-Sevilla, Jaén-Málaga. After that, we have considered two normality tests: Kolmogorov-Smirnof and Shapiro-Wilk. Both tests conclude that the variable electricity consumption per capita does not fit a normal distribution. Then, we have developed a bivariate analysis with the rest of the variables from the database. This analysis reveals that there exist other five variables with a strong correlation with the electricity consumption.

In this paper, we have developed a statistical study about the electricity consumption per capita in Andalusia. This has been possible thanks to the Multiterritorial Information System from the website of the Institute of Statistics and Cartography of the Junta de Andalusia, which has a wide range of data about every city and town. Other authors have done some research on this matter with statistical studies based on this database like Cruces et al. (2009) y el de Pita & Pedregal (2011). In our paper, we have analysed and made a discussion based on the comparison of those studies and our current analysis. We have concluded that the total electricity consumption has decreased during the year 2020. However, there has been an increment on the areas with a population lower than 2.000 and those towns in an interval between 10.001 and 20.000. In the remaining intervals, we have seen a decrement on the consumption. Moreover, in 2007, the areas with a population greater than 50.000 registered more than fifty per cent of the total consumption and this situation does not happen in 2020. This means that effect or importance of the population factor has decreased over the electricity consumption.

As conclusions from all the results obtained in our statistical study, we can affirm that there exist significant differences in Andalusia with respect to the average electricity consumption per capita locally and globally. This yields to the existence of atypical data and that our variable does not fit to a normal behavior. We have also concluded that there exists a strong correlation with other five variables from the database. In this way, we find a group of towns with a very low population and having important industrial, hydroelectric or thermal plants. Some examples are towns such as El Granado (Huelva), Cantoria and Carboneras (Almería), Los Barrios and San Roque (Cádiz), Espiel (Córdoba), Zújar (Granada), Mengíbar (Jaén), Benahavís (Málaga) and Gerena (Sevilla). From the study of electric consumption according to the professional sectors, it can be concluded that in Huelva and Cadiz the industrial sector is the most important, while in the remaining provinces we have the residential sector. This yields to important differences between the first two provinces and the remaining ones and can be used in order to develop measures to save electric consumption in those sectors. When analysing the average consumptions per capita in every province, we deduced that Cádiz and Málaga are the provinces with the lowest electricity consumption. This could be useful as example to implement saving measures in other provinces or to balance the consumptions in every province with the mi-

274

gration of several industries. On the other hand, we find that Huelva is the region with the highest electricity consumption representing 24 per cent of the total consumption in Andalusia. Regarding the atypical and highly atypical data, we have concluded that Almeria is the region with the greatest number of atypical data, while Málaga and Cádiz are just in the opposite situation. It is crucial to control the electricity consumption of this type of data, since they can have influences in a possible prediction model. Normality test allowed us to affirm that the variable average electricity consumption per capita does not fit a normal behavior. This is a natural consequence of the irregularities previously mentioned. Regarding the bivariant study, we can deduce that there are other variables from the database that are strongly correlated with the electricity consumption. This yields to the possibility of setting a certain dependence degree and we could even establish some sort of cause-effect relationship.

We believe that the statistical study carried out in this paper could be useful to make progress in the reduction of electricity consumption. We think that it is important to increase the public awareness on the consequences of high electricity consumption. Moreover, it would be extremely important to work on a transition towards the residential self-consumption. In this way, we would get a lower dependance degree from the electric network, it would allow us to save money and we could develop a more efficient consumption way being respectful with the environment.

275