Forecasting Energy Consumption in Denmark: A Data Science Approach
By Rasoul Avesta Narimani

More and more different countries try to move towards renewable energies. One of the pioneers in this field is Denmark. In the first part of this report we analyze the total energy consumption of this country using a public data set available at DTU Data which is a part of Denmark's initiative to promote transparency and innovation in energy data.

Dataset Overview

The Energinet dataset provides hourly time series data over the course of 2011 to 2019 period, encompassing several key aspects:

- **Wind Power Production**: This includes data on the electricity generated from wind turbines. Denmark is a global leader in wind energy, making this dataset particularly valuable for understanding wind power dynamics.

- **Solar Power Production**: The dataset includes information on electricity produced from solar power sources, contributing to insights on renewable energy trends and solar power contributions.

- **Electricity Consumption**: It details the electricity consumption patterns across Denmark, offering a window into consumer behavior and demand trends.

Purpose and Application

This dataset serves as a vital resource for a wide range of applications:

- **Energy Research**: It enables researchers to analyze patterns, trends, and correlations in renewable energy production and consumption.

- **Policy Making**: Policymakers can use this data to understand the impact of renewable energy and make informed decisions.

- **Machine Learning Projects**: The dataset is ideal for forecasting models, such as predicting energy production or consumption patterns.

In the first part of this report we focus on energy research and highlight the energy trends in Denmark. The second part of the report is dedicated to build timeseries predictive models to forecast the total energy consumption in Denmark.

Part one: Energy trends and future of renewable energies in Denmark

The following graph shows the total energy consumption; both gross consumption and net consumption.

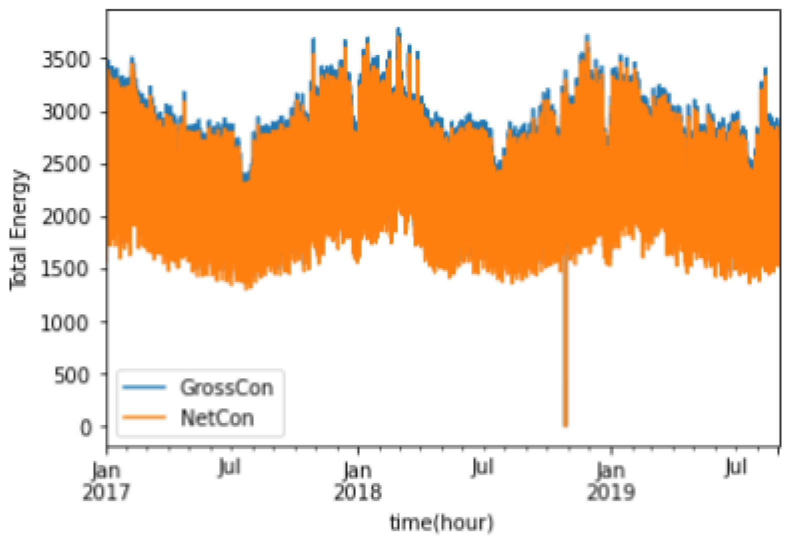

Fig 1. Gross and net consumption energy

We can see that the total energy peaks in winter time indicate a seasonal trend. Figure 2 shows the same data but for a period of one month. This figure highlights the weekly and daily trends of the data. We can see that the total energy consumption decreases during the weekend. That might be due to the fact that industrial power consumption drops on the weekends. Also we can see daily fluctuations. If we plot the autocorrelation function (AFC) for this data, as shown in figure 3, we can see that AFC indicates two peaks happening at 24 hours and 48 hours. This indicates the periodicity of the data and its consistency.

General Trends and future work:

Figure 4 shows a bar plot of different types of energies consumed in Denmark for different years. As indicated in the graph, on-shore wind power and solar production have increased over the years as the central production has decreased. One main question that can be answered using this data is to find out when in the future Denmar can move completely towared wind and solar power production and get rid of the fossil fuel based energies.

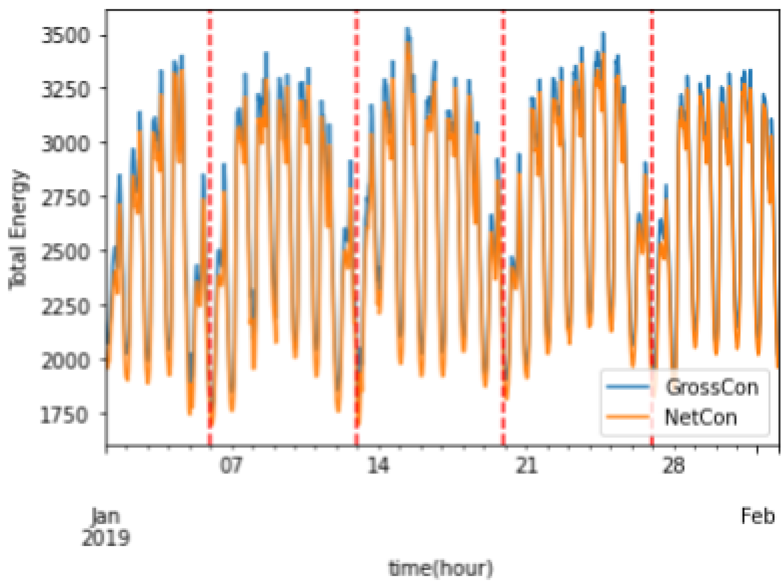

Fig 2. Gross and net consumption energy for a period of one month. The dashed line indicate Fridays to better illustrate the weekly trend.

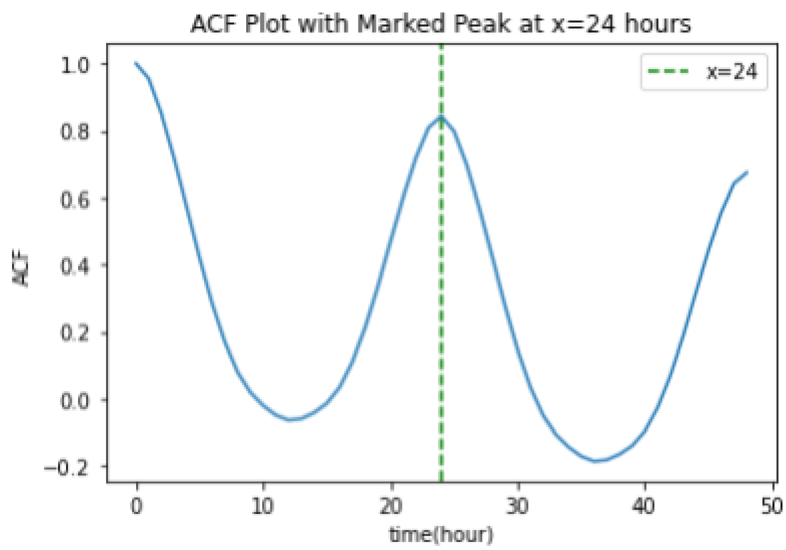

Fig 3. Autocorrelation function indicates peaks regrading the periodicity of the data.

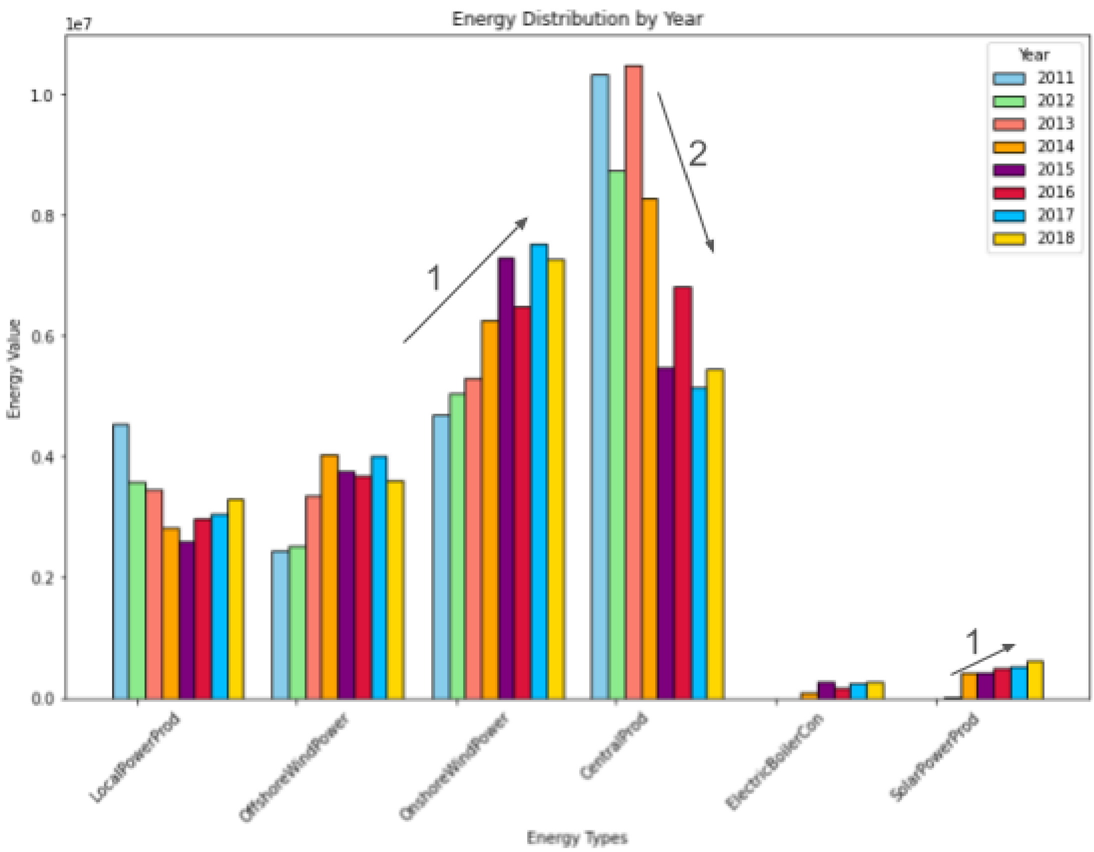

Figure 4. Bar plot comparing different types of energy for different years. The first arrow indicates that on-shore wind power and solar production have increased over the years. On the other hand central production has decreased.

Part two: Time Series analysis for forecasting the total energy consumption
In this part we use different models to predict the future of energy in Denmark. In order to do so different models were implemented including: Exponential Smoothing, Autoregression (AR), Autoregressive Integrated Moving Average (ARIMA) and finally Seasonal Autoregressive Integrated Moving Average (SARIMA). The table below compares different models using different metrics and highlights what features of the data are captured by each model.

Model	Mean Absolute Error	Mean Squared Error	Root Mean Squared Error	24-hour frequency	Trend
Exponential Smoothing	379.4	193060.1	439.4	✓	X
AR	390.7	215660.9	464.4	X	X
ARIMA	317.8	147733.2	384.4	✓	X
SARIMA	351.0	176078.1	419.6	✓	✓

Among all the models only SARIMA was able to detect both daily and seasonal trends of the data. Graph below shows the final forecasting of the total energy.

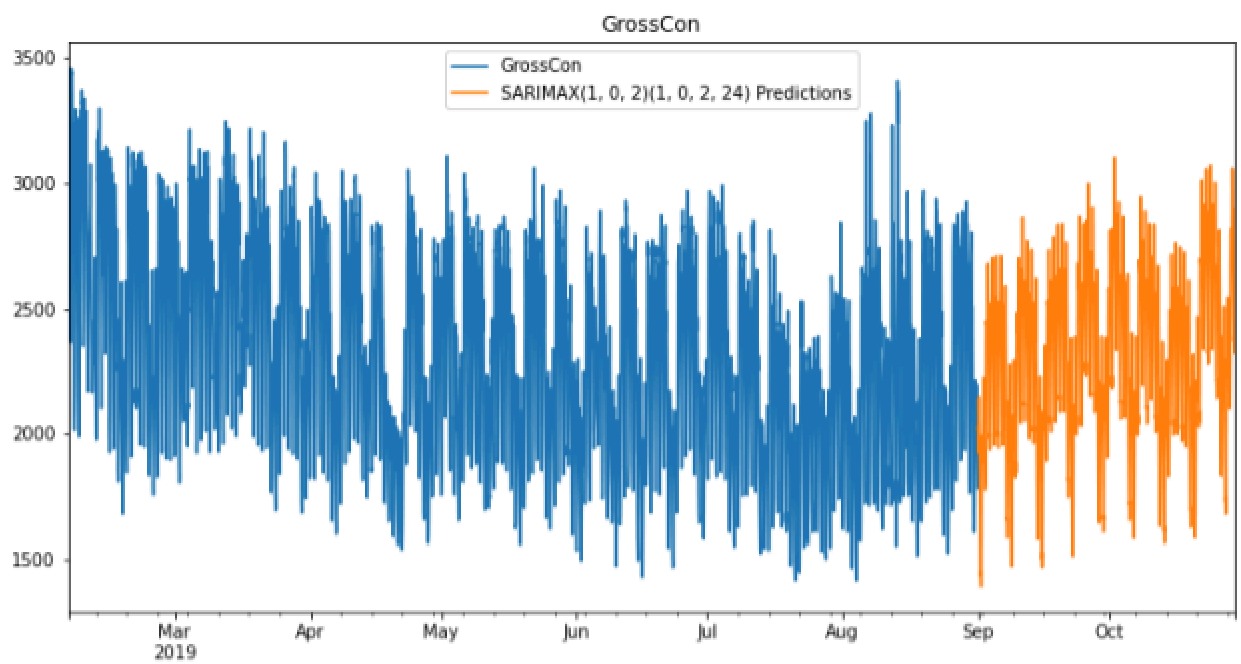

Fig 5. SARIMA model capable of capturing different feature of the data such as daily fluctuations and seasonal trends.

Recommendations and future work
We recommend building a multi-regression model to predict when Denmark can move to 100% renewable energy using the information provided in part 1.

The SARIMA model fails to predict the leap years and needs to be refined. This can be done as a future work.