Prediction

of

temperature

Sebastian Ulita

The Faculty of Power and Aeronautical Engineering

Warsaw University of Technology 2021

# Introduction

The main purpose of the document is to show prediction of temperature based on 5 year history temperature in Warsaw running in AWS EC2 instance.

# Preparing data and instance

Data used to predict temperature include hourly values of temperature, relative humidity, wind velocity, insolation, cloudiness and wind direction collected over the last 5 years. Data are stored in csv file placed in S3.

In EC2 was created Ubuntu instance which was connected to by PuTTy. To get access to EC2 instance you have to get EC2 key pair and generate the SSH-1 (RSA) key in PuTTyGen. It is used to correct login as Ubuntu user in PuTTy, after that user have access to created instance.

First thing needed to using python and libraries is install it through code: pip3 install “Library name”.

To get access to file with code in EC2 user can drop it to the instance through FileZilla. After dropping file there to run project in PuTTy instance with Ubuntu you can use a code bellow



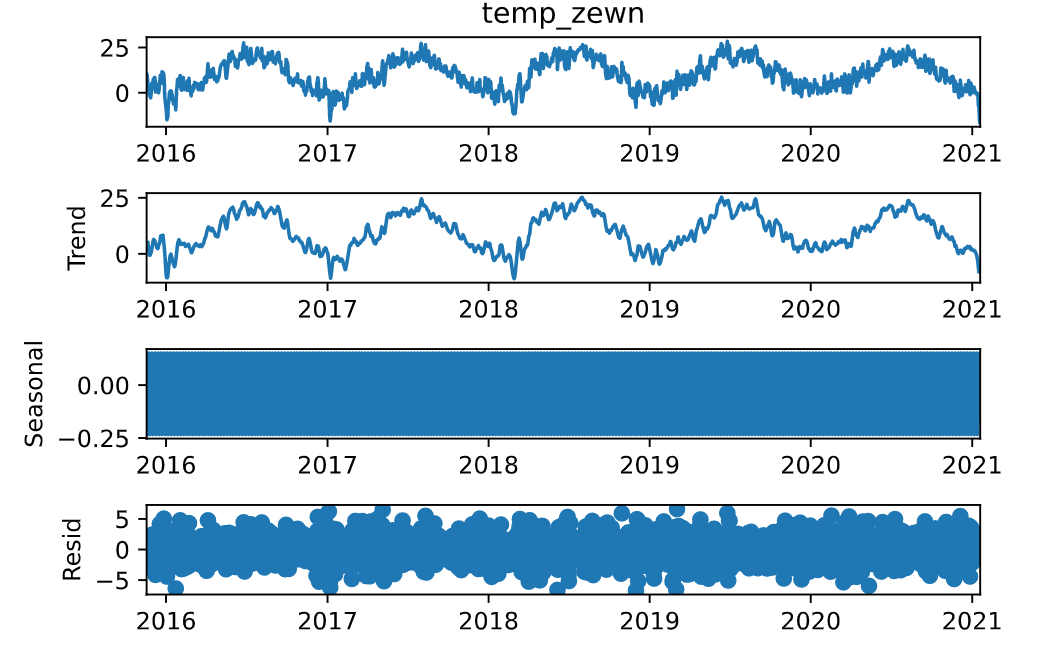
# Prediction

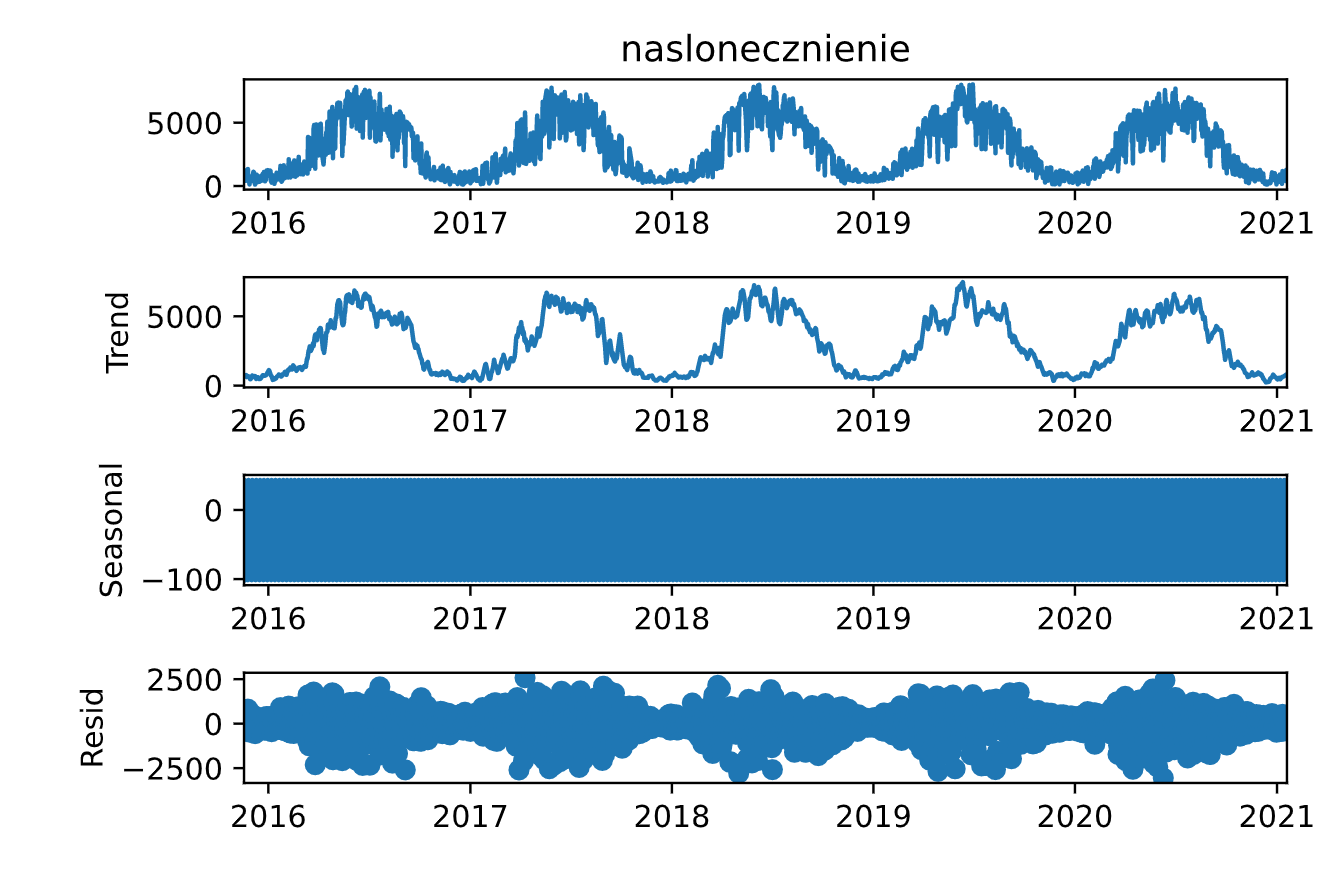
Prediction of temperature could be done on hourly data or daily data. To get daily data of hourly one it is needed to get mean values of each of column of data except insolation which should be calculated as sum.

Prediction is made through ARIMA model. Which means Autoregressive Integrated Moving Average Model.

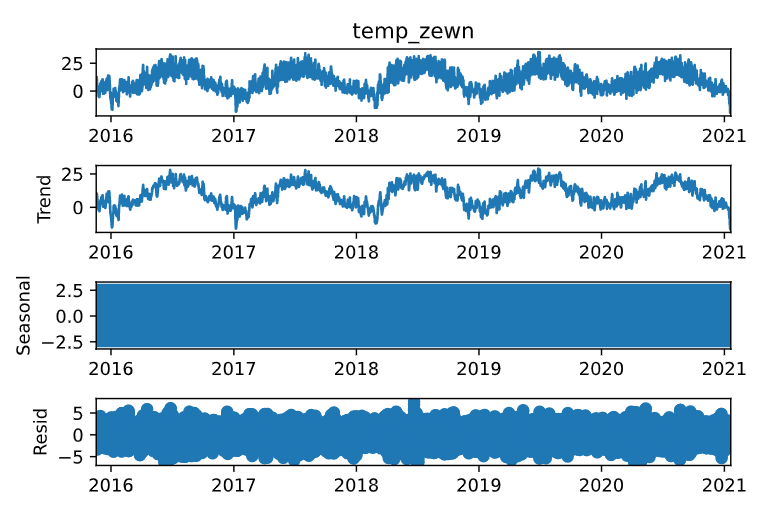
# Results

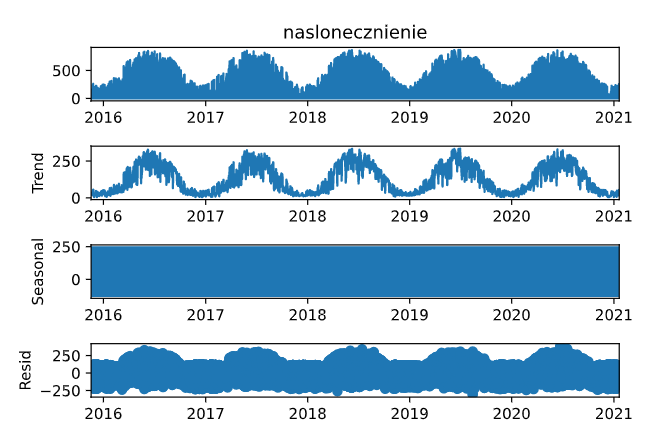
First thing done was decomposing each column to get view of seasonality and trends in temperature and different values as insolation. Below is shown decomposed values for daily and hourly data.

Daily



Hourly

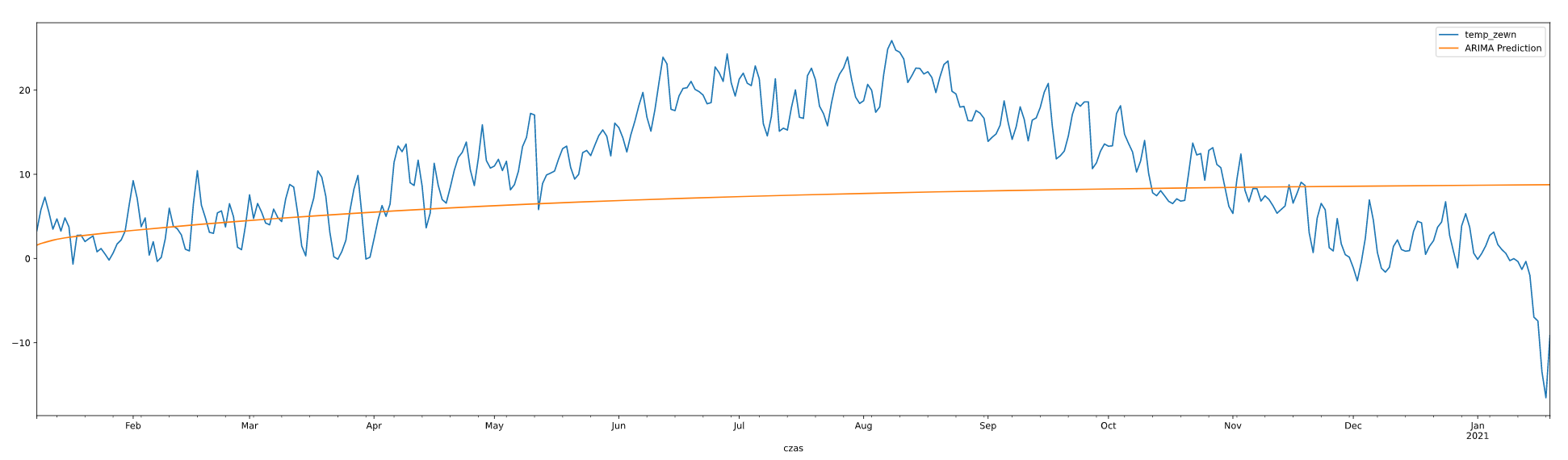




Next step was to get from auto\_arima best suited Arima model for database type and seasonality equal to 30 days. And with this model we can train data, here 80% of whole base and predict other 20% which will be compare to rest of base. To see how good prediction was program is drawing plot and calculate mean absolute error and mean squared error.

For daily data:







By looking on prediction we can see that it wasn’t accurate for daily data. Even after adding insolation as exogenous data it did not helped for prediction. It is possible that program was over trained or period of seasonality was wrong specified.

For hourly data Ubuntu instance is killing a process of calculating auto arima. It might be result of not enough memory enabled in EC2, which was set during creating of instance.