

Lab 5: Time Series with ARIMA

A time series is a sequence of numerical data points in successive order with data points recorded at regular intervals. Time series can be used for many different purposes like registering COVID-19 cases, the values of the stock market, the temperature at a certain city during the years, or the amount of rain.

For this lab, we are going to use the following datasets to predict the temperature and the employment rate.

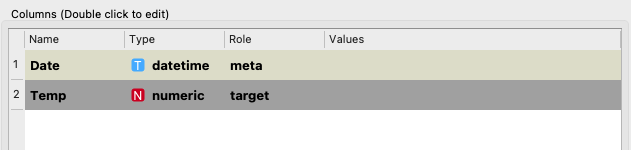
* [Datasets/daily-min-temperatures.csv at master · jbrownlee/Datasets · GitHub](https://github.com/jbrownlee/Datasets/blob/master/daily-min-temperatures.csv)
* [Practical-Time-Series-Analysis/wisconsin-employment-time-series.csv at master · PacktPublishing/Practical-Time-Series-Analysis · GitHub](https://github.com/PacktPublishing/Practical-Time-Series-Analysis/blob/master/Data%20Files/wisconsin-employment-time-series.csv)
* https://github.com/jbrownlee/Datasets/blob/master/daily-total-female-births.csv

**Part I: Predicting values using ARIMA**

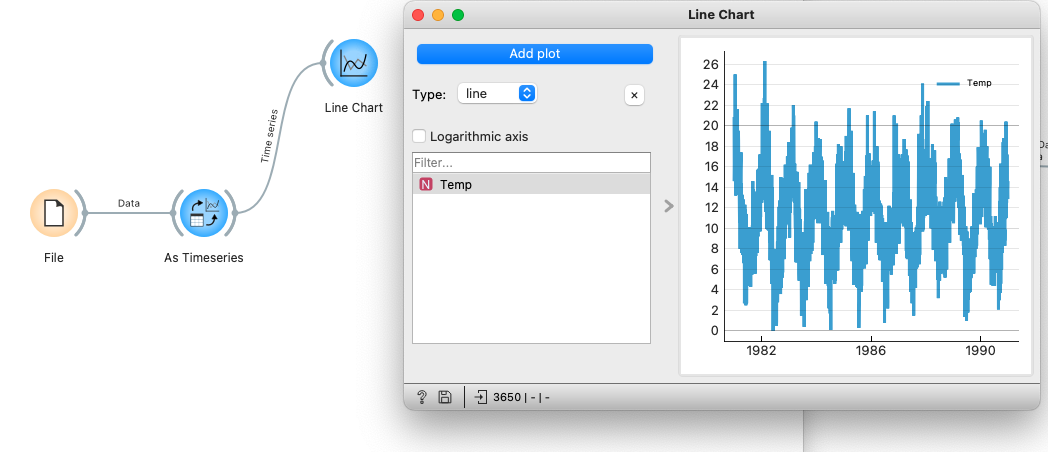
There are many time series methods you can use to predict time series. In this exercise, we are going to use two very typical ones: ARIMA and VAR. ARIMA stands for Autoregressive integrated moving average (ARIMA). It is based on statistical averages and it can only predict one value at a time. VAR stands for vector autoregressive (VAR) and it can predict multiple variables. It is able to relate multiple observations of a variable with its past and it is also able to relate the future with the past.

The first thing you need to do is to go to “Options <-Adds on” and install the “Times series” plugin to use the time series widgets with Orange.

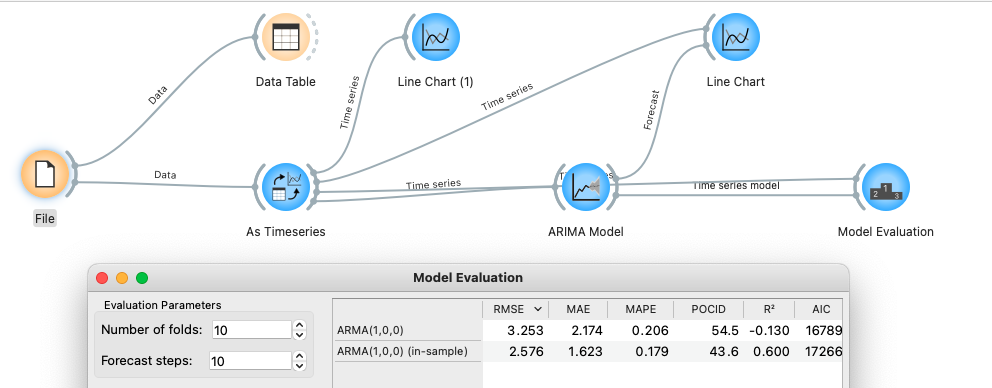
Once you do this, download the dataset from daily temperatures and install a “File” widget. Then, select the following values for Date and Temp.



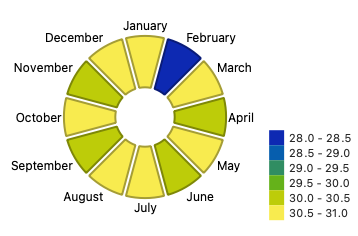
You can now create the following pipeline to see these values when you click on “Line Chart”.



Create the following pipeline with the following widgets and you should obtain similar values to the following ones.



The Spiralogram shows the different values per month. You can see here how some months have generally higher temperatures than others. Click on



To see the complete tutorial visit: <https://youtu.be/bG9fWvdmt0k>

**Part II: Predicting the stock market with VAR**

[44 :Time Series : Using VAR](https://youtu.be/T2IKxI-7SHw)

In this second part, we are going to predict the stock market. Predicting the stock market is quite difficult and this is just an example.

For this exercise, you can reuse the previous code but save it in a different file. In this part we are going to predict the stock market. To do so you can download the following dataset:

<https://raw.githubusercontent.com/HussamHourani/HussamHourani/Orange-Datamining/English/Stock%20Market_Uniqlo(FastRetailing)%202012-2016%20Training%20-%20stocks2012-2016.csv>

When you open the dataset, select “Close” as the target value and date as “Time series” metadata. You can use a line chart to represent the different variables of the dataset.

Once you create the following pipeline, you can use the “VAR Model” to make predictions of the different values of the dataset. Not just one.

