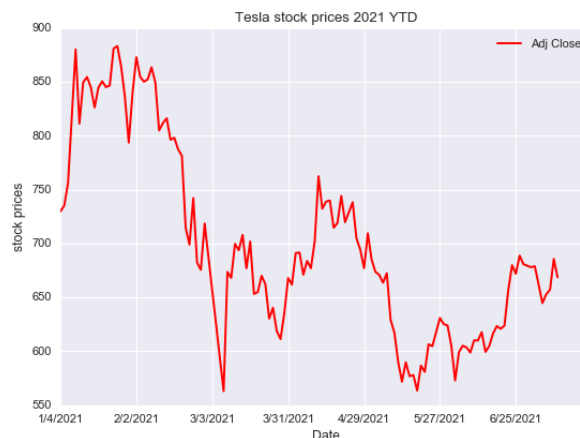


I. STUDENT VERSION

Title: forecasting data using Prophet

Instructions to students:

- **Background information:** Forecasting is a common data science practice that helps firms and individuals gain a deeper understanding of a topic to plan. From an individual perspective, one might be interested to know whether now is a good time to buy a house? How do prices will change in the near future? Whether the price of a stock will go up and down in the next two weeks? From the organizational perspective, planning for companies' sales or production is something that needs constant forecasting. Having a robust forecasting method can be the basis of a strong supply chain, marketing, and management department.
- In this model, we use historical data (previous observation) to predict future prices. In order to build our model, we need to use time-series data that is non-stationary. Loosely speaking, non-stationary data is data that its statistical characteristics such as mean and standard deviation change over time and are not fixed. Examples of time-series data with non-stationary nature are gas prices, stock prices, temperature data, housing prices, etc.



An example of a time-series plot created by the author on Python

- **Individual or group project:** This project can be done individually or in groups of two students
- **Data:**
For data on the daily stock prices of Tesla (and other firms), please visit:
<https://finance.yahoo.com/quote/TSLA/history?p=TSLA>

For this project, we only need to have access to one dataset. In particular, I use the Tesla daily stock prices. You can choose any dataset with daily information available that meets the seasonality criteria.

- **Procedure/plan of action to complete the project:**
Pre-project Assignment for students:

Make sure to install Anaconda on your computer/laptop. You can watch a short tutorial video on the link below:

https://anaconda.cloud/tutorials/getting-started-with-anaconda-individual-edition?source=win_installer



After successfully installing the Anaconda and opening the platform, choose the *Jupyter Notebook* as a program for running python analyses. This will open a new window on your browser where you can choose the version of Python you need to run. I recommend using the latest version of Python (3 as of today) to run the analyses. If for any reason the Jupyter Notebook is not responding, you can always use the *JupyterLab* directly from the Anaconda platform, and under the *Notebooks*, choose *Python 3*.



Finally, follow the code on the file Tesla.ipynb to run the analyses. You can substitute your data with the one that I use. Make sure that your excel file should be exactly in the same format as mine with the exact name as mine (or modify the code according to your variable names). Only two columns are needed, first the *date* and the daily *prices*.

After running the analyses, think about these questions to discuss in class:

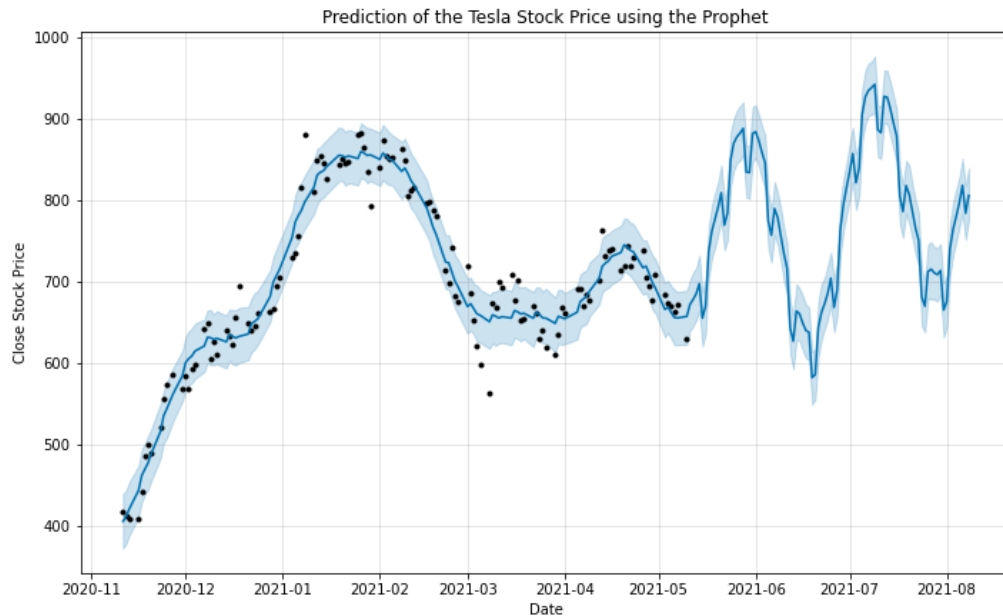
- 1) Based on your prediction using the model, how do you expect the prices to change in the future?
- 2) Why do you think it is essential for us to know the future prices of the variables you predict in this model? What kind of decisions can one make using this data?
- 3) After decomposing your final result into different pieces based on the seasonality and trend, what additional information you get about the data?
- 4) looking at the trend and seasonality in your data, what days of the day/month your variable tends to be high/low?
- 5) Following your response to question 4, does this finding make sense to you? Why do you think the prices might be historically high on certain days of the week/month or a specific time of the day?

After discussing the initial findings and questions in the class, the instructor should run the analyses for the class using his choice of data by following the provided codes step by step. Following the completion of the analyses, the class can discuss how the model can achieve more efficient fitting.

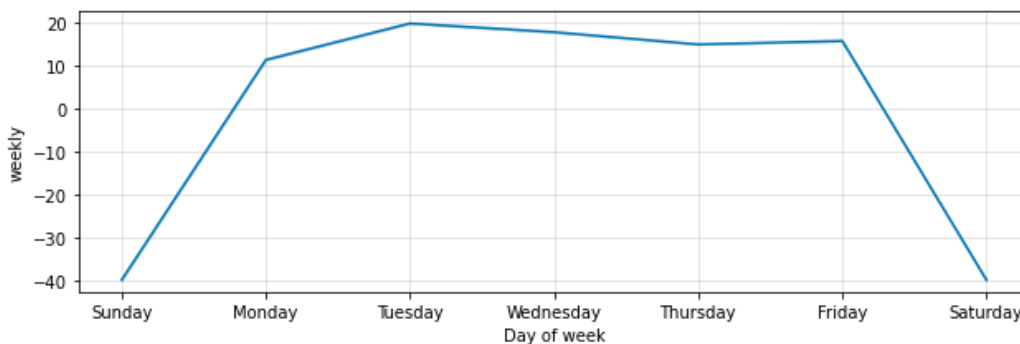
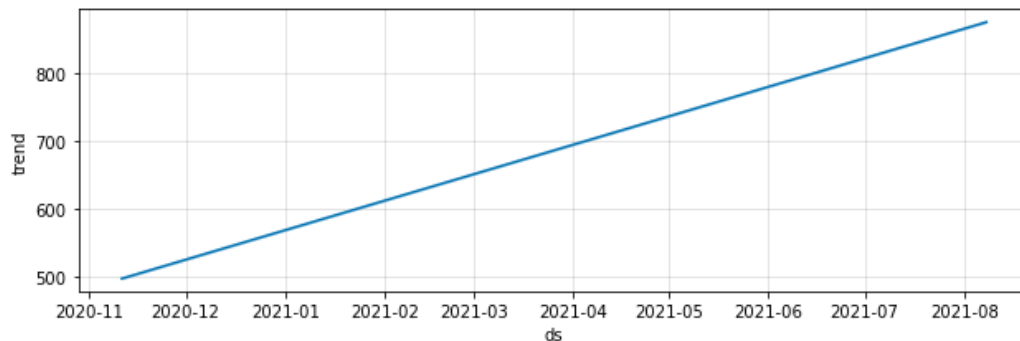
Example(s):

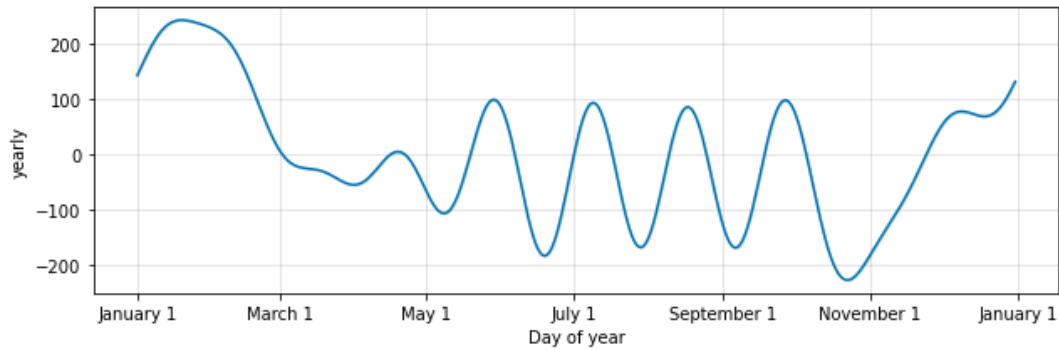
Using the daily stock prices for Tesla, Inc. from November 11th, 2020, to May 5th, 2021, we predicted the following model to forecast the company's stock price between May 6th, 2021, and August 6th, 2021. Next, we decompose our forecast into its trend and seasonality elements to learn more about our prediction.

This graph is the main forecasting graph in which the first part (up to May 2021) are available daily prices that we used to predict the prices in the following three months. Obviously, each student can use a different initial period and forecasting period for their analyses. However, the longer the input data and the shorter the forecast period make the prediction more reliable.



In the next sets of codes, we decompose our main graph into its trend and seasonality measures.





These graphs will demonstrate the overall upward slope as the trend for the prices of Tesla stock from May to August 2021. The second and third graphs show the seasonality elements of data which implies that what days or months show the highest and lowest prices in the case of Tesla Inc. stock.

Using the daily data from Tesla, Inc., we found the following:

- Duration: this project should be completed within one week.
- Deliverables and evaluations:
 - Students need to be familiar with the basic autoregressive model and its components in forecasting.
 - Each student/group can present the entire project within 15 minutes.
 - Students need to be able to answer post-project questions.