Time series analysis on the stock price of Tesla Inc.

Wenhao Pan (3034946058), Ruojia Zhang, Mengzhu Sun, Xiangxi Wang, Mingmao Sun

November 13, 2021

Contents

1 Abstract

2 Introduction

3 Data Description

4 Exploratory Data Analysis

To obtain a comprehensive understanding of the data, we conduct explanatory data analysis (EDA) first. Figure 1(a) is the time series plot of all the given time points. We observe that the average price of Tesla before 2020 is considerably lower than that after 2020. Also, the stock price before 2020 seems to have a constant trend and no seasonality. Moreover, due to the excessive number of time points, it is difficult to visually examine the trend and seasonality pattern for data after 2020. Therefore, for the sake of interest and convenience, we decide to only analyze the last 300 time points, which cover the period from `2020-08-26` to `2021-11-02` excluding weekends. Thus, whenever we mention "data" in the following analysis, we implicitly mean the last 300 time points.

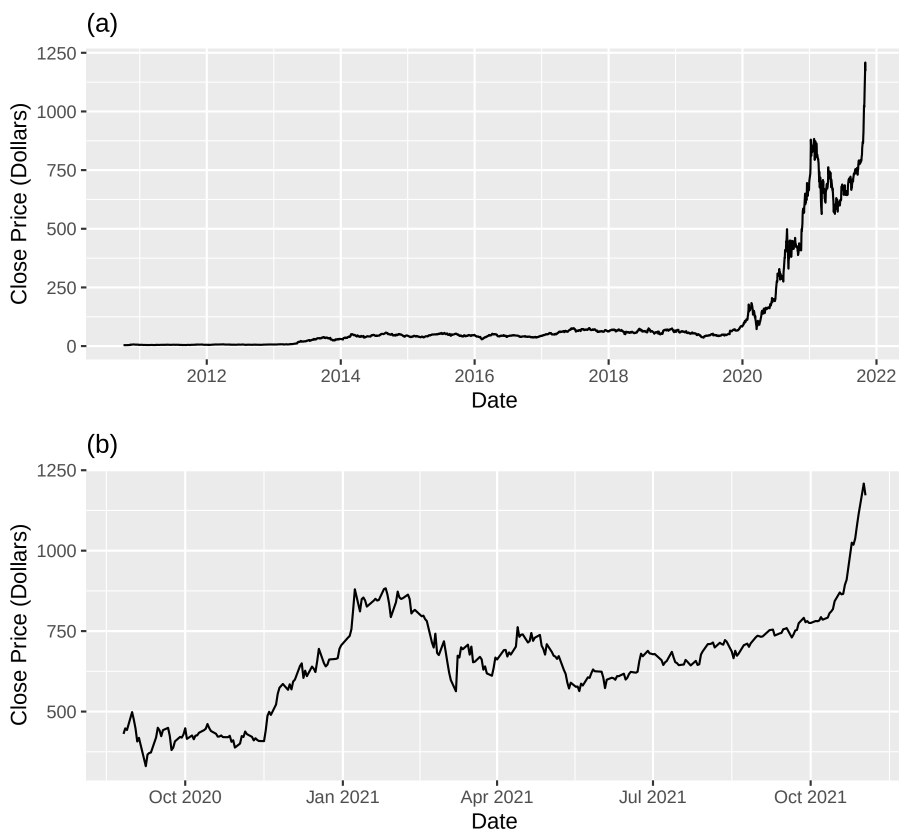


Figure 1: (a) Time series plot of all available trading days. (b) Time series plot of last 300 trading days

Figure 1(b) is the time series plot of the close prices of Tesla in last three hundred trading days before and including `2021-11-02`. We first observe that our data is roughly homoscedastic based on Figure 1(b), so we do not need to use any transformation to stabilize the variance of our data.

Intuitively, the data is not stationary since there exists a nonlinear trend. We do not observe a clear seasonality pattern probably because the data is daily. To be more convincing, we plot the sample ACF and PACF of the data in Figure 2.

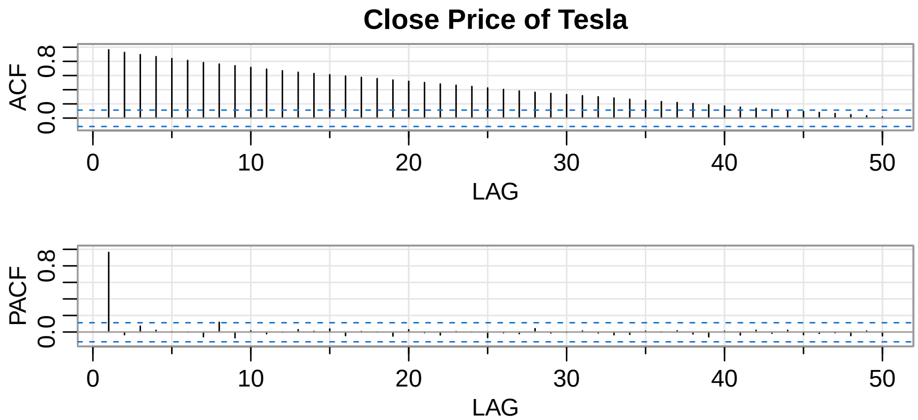


Figure 2: Sample ACF and PACF plots of the data

The decaying, positive ACF's imply a possible trend. As we expect, both sample ACF and PACF plots do not demonstrate a seasonality pattern. To quantitatively verify our superposition that the data does not have a strong seasonality pattern, we plot and inspect the periodogram of the data.

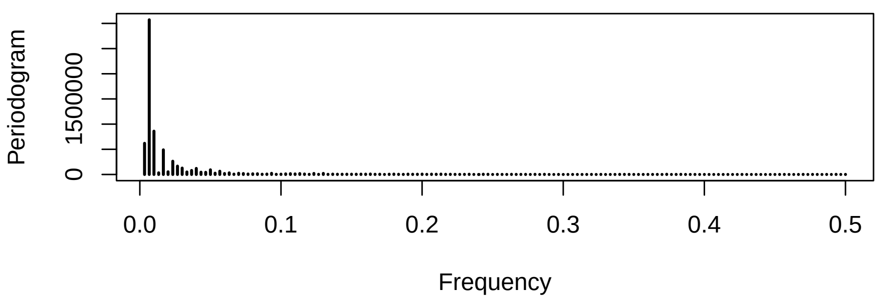


Figure 3: The periodogram of the data

Since the periodogram has multiple consecutive spikes, the phenomenon called "leakage", the data does not have a dominant seasonality pattern or frequency in our data. Nonetheless,

5 Model Construction

With a comprehensive understanding of our data, we start to experiment and construct different time series model. We first stabilize the variance of our data. Then, we choose and build two non-parametric signal models to remove any trend or seasonality from the data so that the residuals are approximately weekly stationary. We do not consider any parametric method for modeling trend because we think the trend of the stock price data is too complicated to be modeled by a parametric model, such as a high-order polynomial. We could use a 15 or 20 order polynomial, but it will overfit the training data and give imprecise predictions. Finally, based on each signal model, we provide two ARMA models or its extension, such as SARMA or

ARIMA, to whiten the residuals of the signal model. Thus, we have four candidate models, and we will explain how we select a final model in the next section.

5.1 Non-parametric Model: exponential smoothing with seasonal differencing

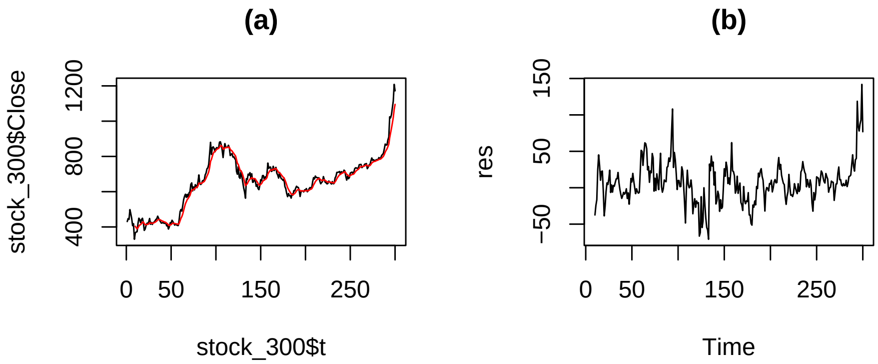


Figure 4: (a): (b):

In this model, we choose exponential smoothing with weight

$$\alpha = 0.8$$

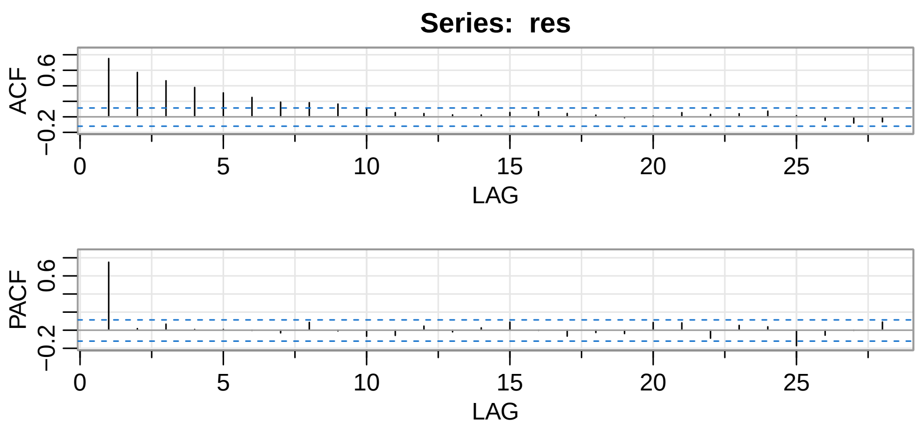
and lag

$$10$$

and seasonal differencing with period

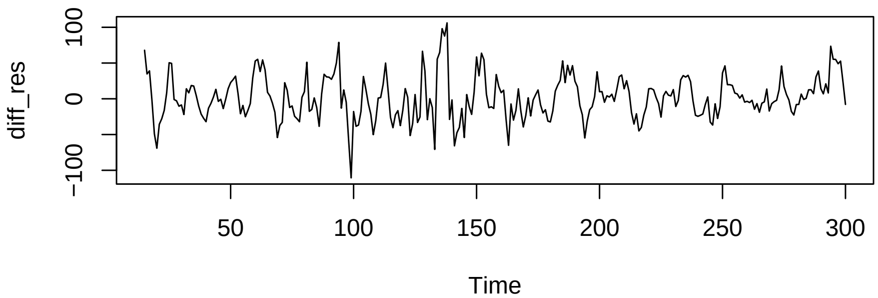
$$d = 5$$

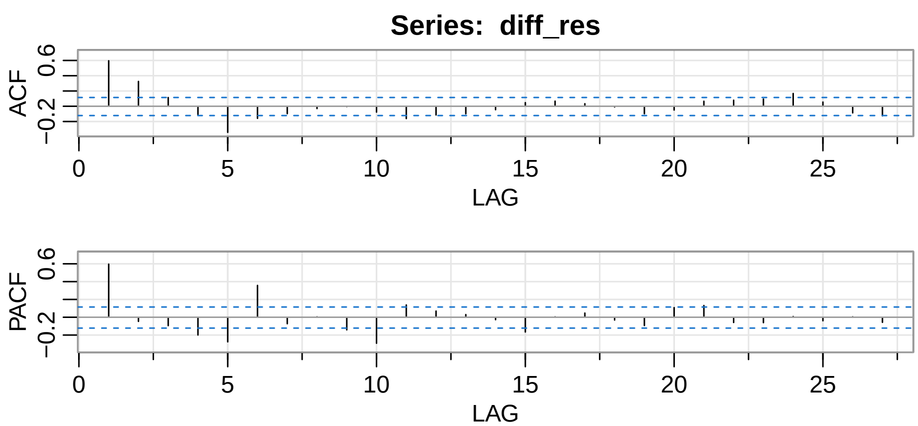
.




```
##      [,1] [,2] [,3] [,4] [,5] [,6]  [,7] [,8]  [,9] [,10] [,11] [,12] [,13]
## ACF  0.75 0.57 0.47 0.38 0.31 0.25  0.19 0.18  0.16  0.11  0.05  0.04  0.03
## PACF 0.75 0.02 0.07 0.01 0.01 0.00 -0.03 0.09 -0.01 -0.07 -0.06  0.05 -0.02
##      [,14] [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22] [,23] [,24] [,25]
## ACF   0.02  0.06  0.07  0.04  0.02 -0.01  0.01  0.05  0.03  0.04  0.07  0.02
```

```
## PACF  0.03  0.09  0.00 -0.07 -0.03 -0.04  0.09  0.08 -0.09  0.05  0.04 -0.17
##      [,26] [,27] [,28]
## ACF  -0.05 -0.08 -0.06
## PACF -0.06  0.00  0.09
```






```
##      [,1]  [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]  [,9] [,10] [,11] [,12]
## ACF   0.6  0.33  0.11 -0.11 -0.34 -0.16 -0.10 -0.03  0.00 -0.08 -0.16 -0.11
## PACF  0.6 -0.05 -0.09 -0.20 -0.28  0.36 -0.07  0.00 -0.14 -0.29  0.14  0.07
##      [,13] [,14] [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22] [,23] [,24]
## ACF  -0.10 -0.05  0.05  0.07  0.03 -0.01 -0.10 -0.05  0.07  0.08  0.09  0.17
## PACF  0.03 -0.03 -0.17  0.00  0.05 -0.03 -0.09  0.11  0.13 -0.06 -0.06  0.01
##      [,25] [,26] [,27]
## ACF   0.05 -0.09 -0.13
## PACF -0.04  0.00 -0.06
```

5.2 Non-parametric Model: second order differencing with seasonal differencing

6 Model Comparision and Selection

7 Final Model

7.1 Model interpretation

7.2 Prediction

8 Conclusion