

4.

(a) Economically, the first difference is essentially the difference of current season earnings compared with the past season, while the seasonal difference is compared with the same season last year.

(b) The model is built in codes. The estimated model is:

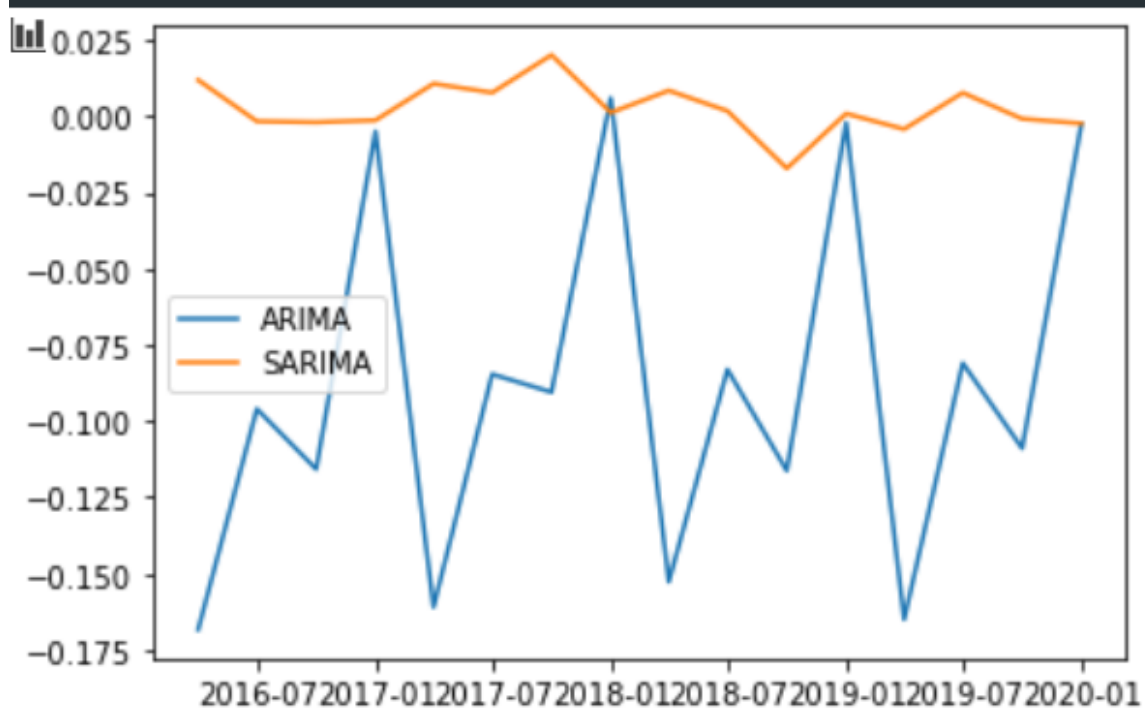
$$x_{t+1} - x_t = 0.0506 + \epsilon_{t+1} - 0.4920\epsilon_t$$

(c) The model is built in codes. The estimated model is:

$$(x_t - x_{t-1}) - (x_{t-4} - x_{t-5}) = (\epsilon_t - 0.1593\epsilon_{t-1}) - 0.4830(\epsilon_{t-4} - 0.1593\epsilon_{t-5})$$

In the model $\theta_1 = 0.1593$, $\theta_2 = 0.4830$, which means the past earning increase shocks (both the past season and the same season for the last year) is negatively correlated with predicted next season earnings.

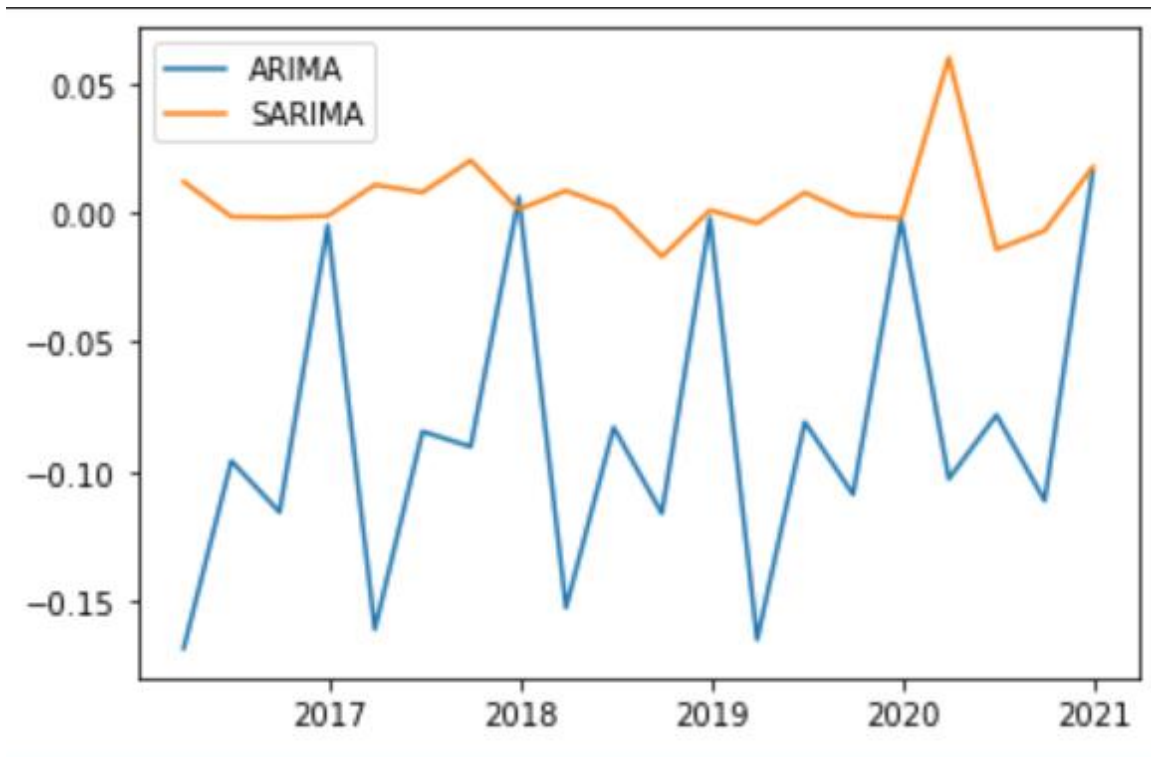
(d) (e)



ARIMA MSE: 0.01134832807889095
SARIMA MSE: 7.346810242072138e-05

The seasonal arima model (airline model) performs better.

(f)



Both models perform worse in 2020. This is because the COVID pandemic is an exogenous event outside of the description of the model. To improve the forecast accuracy, we might include macroeconomic predictors such as expected inflation, GDP growth, treasury rate, etc.