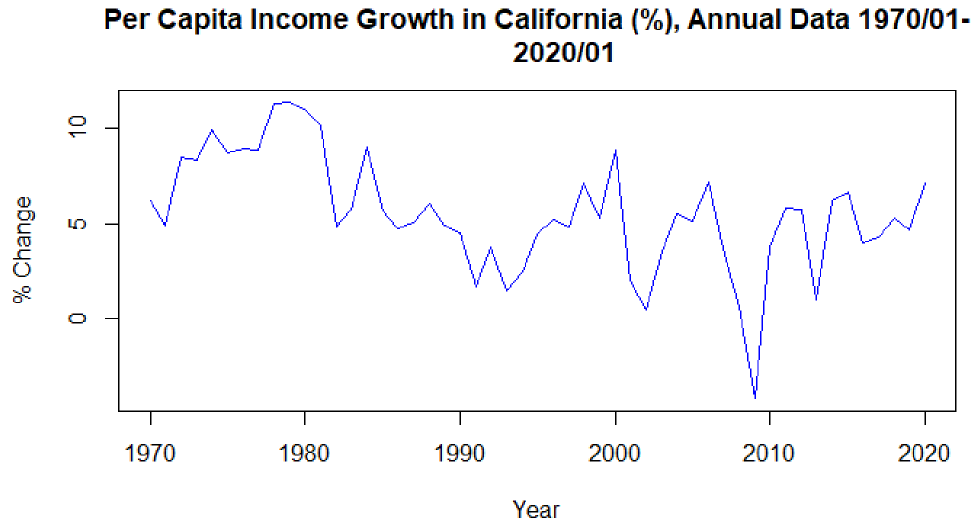
1.

a.

Time series plotted:



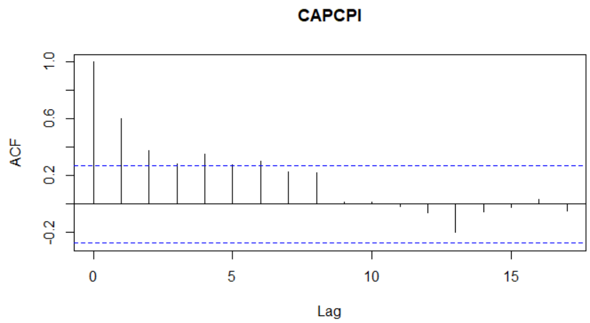
b.

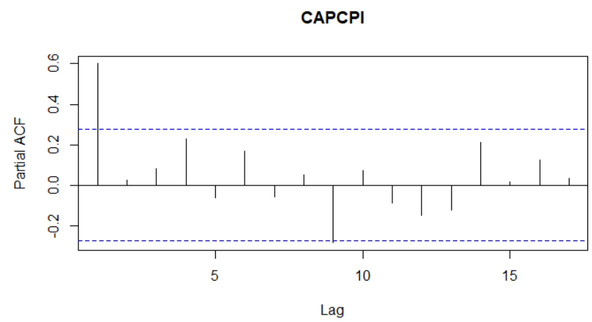
```
Autocorrelations of series 'CAPCPI_ts', by lag

     0      1      2      3      4      5      6      7      8
 1.000  0.600  0.376  0.287  0.354  0.275  0.300  0.228  0.224
     9     10     11     12     13     14     15     16     17
 0.013  0.018 -0.013 -0.057 -0.198 -0.053 -0.025  0.031 -0.050
```

```
Partial autocorrelations of series 'CAPCPI_ts', by lag

     1      2      3      4      5      6      7      8      9
 0.600  0.024  0.082  0.230 -0.063  0.170 -0.057  0.050 -0.285
    10     11     12     13     14     15     16     17
 0.071 -0.090 -0.149 -0.120  0.212  0.016  0.125  0.032
```





After computing the ACF of the time series, we deduce that the data is not stationary since the ACF should die down much quicker or cut off within the first few lags. Thus, we compute the first difference to yield a stationary series.

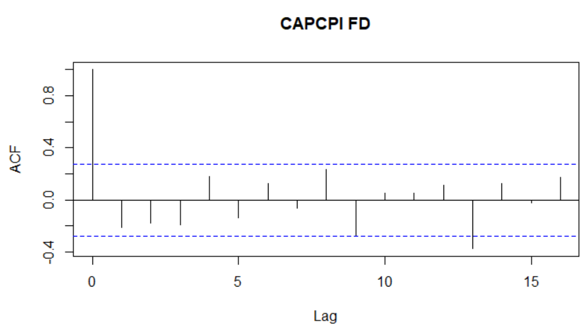
The ACF and PACF from first differencing are as follows.

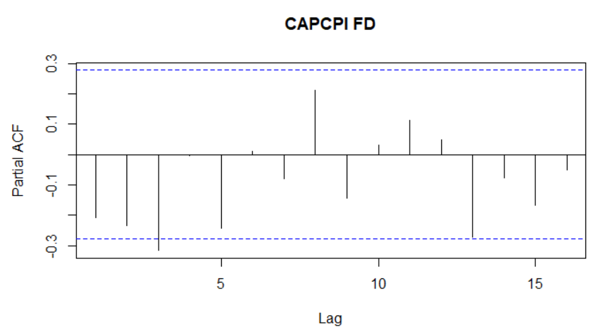
```
Autocorrelations of series 'CAPCPI_FD_ts', by lag

     0      1      2      3      4      5      6      7      8
 1.000 -0.208 -0.180 -0.192  0.182 -0.140  0.124 -0.063  0.237
     9     10     11     12     13     14     15     16
-0.273  0.055  0.052  0.113 -0.376  0.126 -0.020  0.172
```

```
Partial autocorrelations of series 'CAPCPI_FD_ts', by lag

     1      2      3      4      5      6      7      8      9
-0.208 -0.233 -0.317 -0.005 -0.243  0.009 -0.081  0.210 -0.145
    10     11     12     13     14     15     16
 0.032  0.112  0.049 -0.272 -0.076 -0.167 -0.052
```





c.

From the ACF and PACF of the first difference model, the time series model I would consider for this series is the ARMA(1, 1) model since both the ACF and PACF die down quickly. Furthermore, observing the PACF also depicts it being inside the confidence interval at lag 1 which further emphasises the time series model I chose.

d.

The model estimated can be computed from R:

```
Call:
arima(x = CAPCPI_ts, order = c(1, 1, 1))

Coefficients:
         ar1      ma1
      0.4350  -0.8665
s.e.  0.1932   0.1160

sigma^2 estimated as 5.952:  log likelihood = -115.87,  aic = 237.73
```

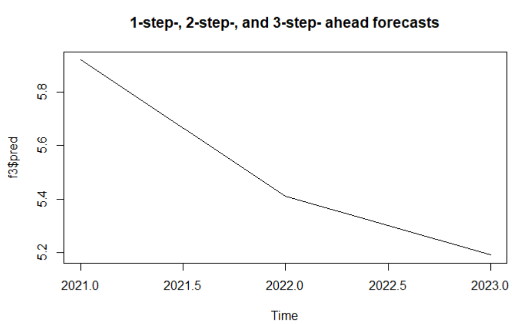
Thus, the estimated model is as follows:

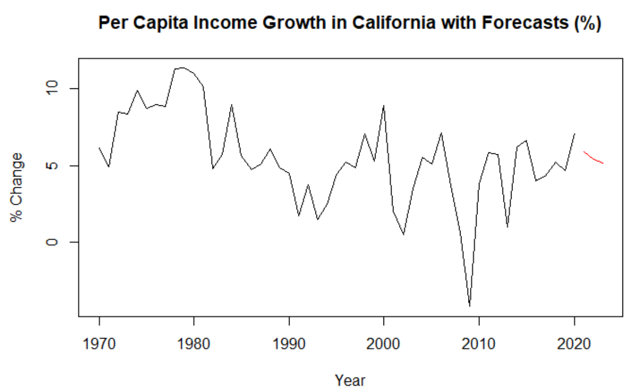
$$y_t = \underset{(0.1932)}{0.4350}\, y_{t-1} - \underset{(0.1160)}{0.8665}\varepsilon_{t-1} + \varepsilon_t$$

The 1-step-, 2-step-, and 3-step- ahead forecasts are as follows:

```
$pred
Time Series:
Start = 2021
End = 2023
Frequency = 1
[1] 5.921143 5.410933 5.189015

$se
Time Series:
Start = 2021
End = 2023
Frequency = 1
[1] 2.439608 2.806231 2.955972
```





Therefore, the forecasts are as follows (standard error in brackets):

1-step forecast = 5.921143 (2.439608)

2-step forecast = 5.410933 (2.806231)

3-step forecast = 5.189015 (2.955972)