Question 8:

In [114…

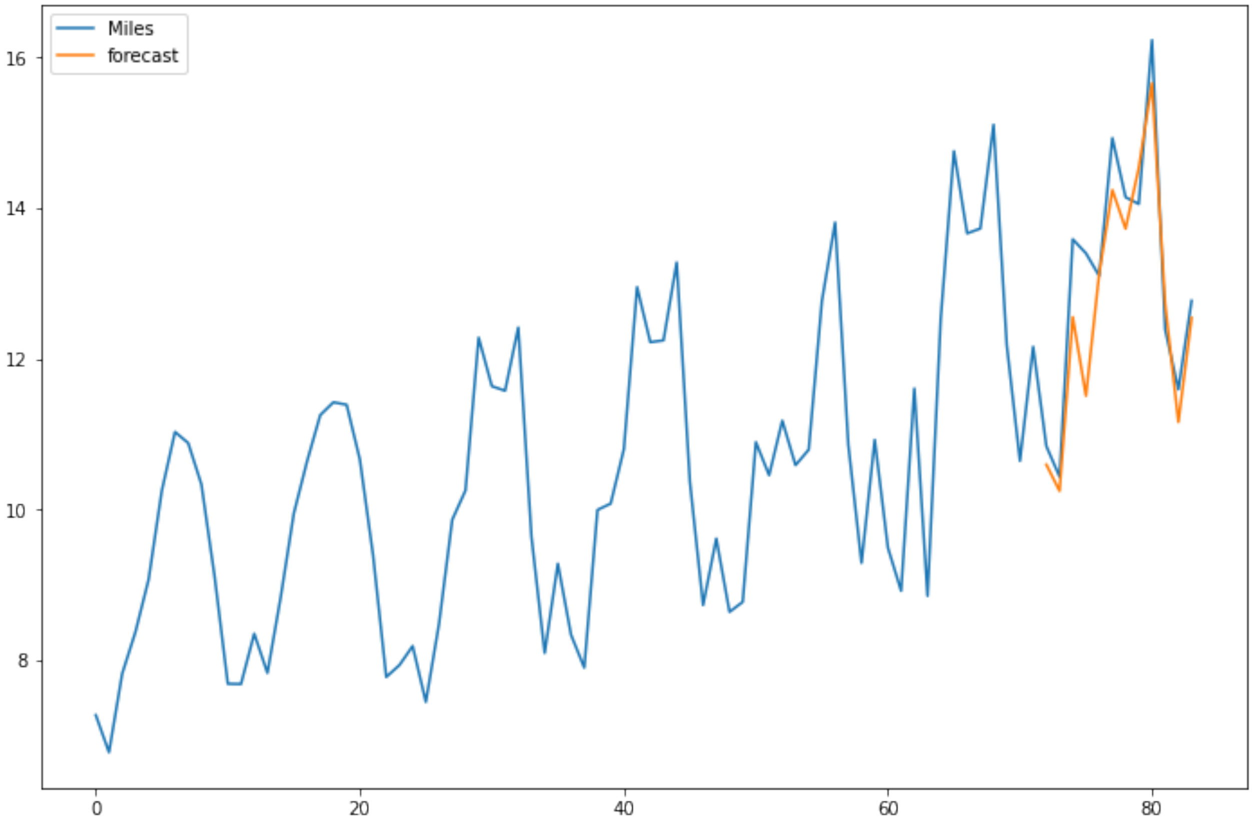
```
modelfinal=sm.tsa.statespace.SARIMAX(df['Miles'],order=(3, 1, 2),seasonal_c
resultsfinal=modelfinal.fit()
```

In [115…

```
df['forecast']=resultsfinal.predict(start=72,end=84,dynamic=True)
df[['Miles','forecast']].plot(figsize=(12,8))
```

Out[115…

```
<AxesSubplot:>
```



In [116…

```
print(resultsfinal.summary())
```

```
                                      SARIMAX Results                                      
===========================================================================
================
Dep. Variable:                               Miles   No. Observations:
84
Model:             SARIMAX(3, 1, 2)x(3, 1, [], 12)   Log Likelihood
-82.916
Date:                             Wed, 27 Oct 2021   AIC
183.832
Time:                                     17:16:06   BIC
204.196
Sample:                                          0   HQIC
191.930
                                              - 84
Covariance Type:                               opg
===========================================================================
===
                 coef    std err          z      P>|z|      [0.025      0.9
75]
---------------------------------------------------------------------------
---
ar.L1         -0.2708      0.613     -0.442      0.658      -1.471       0.
930
ar.L2          0.0256      0.231      0.111      0.912      -0.426       0.
477
ar.L3          0.2488      0.233      1.070      0.285      -0.207       0.
705
ma.L1         -0.3624      0.621     -0.583      0.560      -1.580       0.
855
ma.L2         -0.4534      0.444     -1.020      0.308      -1.324       0.
418
ar.S.L12      -0.6990      0.112     -6.223      0.000      -0.919      -0.
479
ar.S.L24      -0.3110      0.166     -1.869      0.062      -0.637       0.
015
ar.S.L36      -0.0716      0.322     -0.223      0.824      -0.702       0.
559
sigma2         0.5446      0.080      6.796      0.000       0.388       0.
702
===========================================================================
========
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):
30.34
Prob(Q):                              0.98   Prob(JB):
0.00
Heteroskedasticity (H):               3.59   Skew:
-0.71
Prob(H) (two-sided):                  0.00   Kurtosis:
5.87
===========================================================================
========

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (comp
lex-step).
```

In [117…

```
df[72:] #compare actual values to forecasted values!
```

Out[117…

	Month	Miles	Seasonal First Difference	forecast
72	Jan-70	10.840	1.349	10.594961
73	Feb-70	10.436	1.517	10.245751
74	Mar-70	13.589	1.982	12.555449
75	Apr-70	13.402	4.550	11.509256
76	May-70	13.103	0.566	13.129839
77	Jun-70	14.933	0.174	14.242812
78	Jul-70	14.147	0.480	13.728932
79	Aug-70	14.057	0.326	14.535950
80	Sep-70	16.234	1.124	15.660030
81	Oct-70	12.389	0.204	12.726603
82	Nov-70	11.594	0.949	11.160315
83	Dec-70	12.772	0.611	12.544510

In [118…

```
from sklearn.metrics import mean_absolute_error
```

In [119…

```
maerror= mean_absolute_error((df["Miles"]).iloc[72:83], df["forecast"].iloc
maerror
```

Out[119…

```
0.5746259792002264
```

USED: order=(3, 1, 2),seasonal_order=(3,1,0,12) The forecast is very close to the actual values as you can see in the graph and the table, but not exactly the same. I tried playing around with different p,q, P, Q values to see which one would follow the actual values closely, turns out this is the best one I could find. The AIC value is 180.607 and the forecast nearly follows the actual values.
It also follows the increasing trend of actual values. Mean absolute error is also 0.5746259792002264!

In [ ]:

```

```