In [18]:
```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.metrics import mean_absolute_error, root_mean_squared_error, mean_a
from statsmodels.tsa.statespace.sarimax import SARIMAX
from statsmodels.tsa.stattools import grangercausalitytests
```

In [19]:
```
df = pd.read_csv(r"C:\Users\HP\OneDrive\Desktop\NLP\Data\ML471_S4_Datafile_Conce

df['Datetime'] = pd.to_datetime(df['Datetime'])
df.set_index('Datetime', inplace=True)

y = df['Consumption']
exog = df['Festivals/Special_events']

train_size = int(len(y) * 0.8)
y_train, y_test = y[:train_size], y[train_size:]
exog_train, exog_test = exog[:train_size], exog[train_size:]

print("Granger Casuality Test Results:")

granger_data = df[['Consumption', 'Festivals/Special_events']]
grangercausalitytests(granger_data, maxlag=12, verbose=True)

model = SARIMAX(y_train, exog=exog_train, order=(1, 0 ,2), seasonal_order=(0, 1,
model_fit = model.fit(disp=False)
print(model_fit.summary())

forecast = model_fit.predict(start=len(y_train), end=len(y_train)+len(y_test)-1,

mae = mean_absolute_error(y_test, forecast)
rmse = root_mean_squared_error(y_test, forecast)
mape = mean_absolute_percentage_error(y_test, forecast)

print(f"MAE: {mae: .4f}")
print(f"RMSE: {rmse: .4f}")
print(f"MAPE: {mape: .4f}")
```

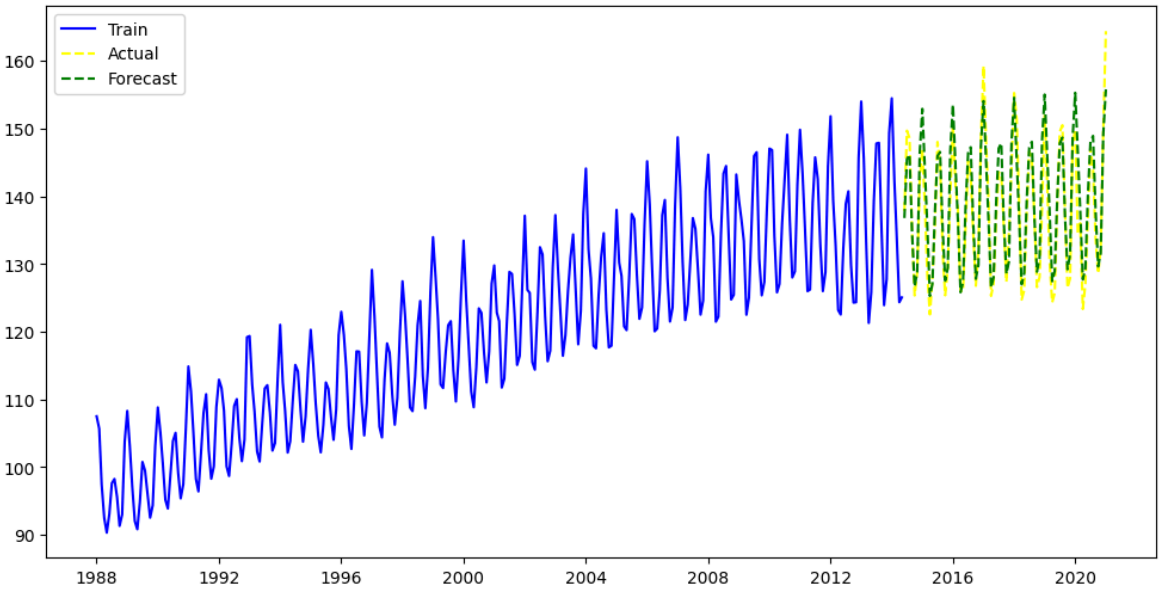
```
Granger Casuality Test Results:

Granger Causality
number of lags (no zero) 1
ssr based F test:         F=6.3712  , p=0.0120  , df_denom=393, df_num=1
ssr based chi2 test:   chi2=6.4199  , p=0.0113  , df=1
likelihood ratio test: chi2=6.3684  , p=0.0116  , df=1
parameter F test:         F=6.3712  , p=0.0120  , df_denom=393, df_num=1

Granger Causality
number of lags (no zero) 2
ssr based F test:         F=13.6644 , p=0.0000  , df_denom=390, df_num=2
ssr based chi2 test:   chi2=27.6791 , p=0.0000  , df=2
likelihood ratio test: chi2=26.7524 , p=0.0000  , df=2
parameter F test:         F=13.6644 , p=0.0000  , df_denom=390, df_num=2

Granger Causality
number of lags (no zero) 3
ssr based F test:         F=3.1172  , p=0.0261  , df_denom=387, df_num=3
ssr based chi2 test:   chi2=9.5208  , p=0.0231  , df=3
likelihood ratio test: chi2=9.4076  , p=0.0243  , df=3
parameter F test:         F=3.1172  , p=0.0261  , df_denom=387, df_num=3

Granger Causality
number of lags (no zero) 4
ssr based F test:         F=1.2370  , p=0.2946  , df_denom=384, df_num=4
ssr based chi2 test:   chi2=5.0642  , p=0.2808  , df=4
likelihood ratio test: chi2=5.0318  , p=0.2840  , df=4
parameter F test:         F=1.2370  , p=0.2946  , df_denom=384, df_num=4

Granger Causality
number of lags (no zero) 5
ssr based F test:         F=1.0537  , p=0.3858  , df_denom=381, df_num=5
ssr based chi2 test:   chi2=5.4205  , p=0.3667  , df=5
likelihood ratio test: chi2=5.3834  , p=0.3709  , df=5
parameter F test:         F=1.0537  , p=0.3858  , df_denom=381, df_num=5

Granger Causality
number of lags (no zero) 6
ssr based F test:         F=1.4520  , p=0.1937  , df_denom=378, df_num=6
ssr based chi2 test:   chi2=9.0117  , p=0.1729  , df=6
likelihood ratio test: chi2=8.9094  , p=0.1787  , df=6
parameter F test:         F=1.4520  , p=0.1937  , df_denom=378, df_num=6

Granger Causality
number of lags (no zero) 7
ssr based F test:         F=1.1748  , p=0.3162  , df_denom=375, df_num=7
ssr based chi2 test:   chi2=8.5526  , p=0.2864  , df=7
likelihood ratio test: chi2=8.4601  , p=0.2938  , df=7
parameter F test:         F=1.1748  , p=0.3162  , df_denom=375, df_num=7

Granger Causality
number of lags (no zero) 8
ssr based F test:         F=1.1357  , p=0.3382  , df_denom=372, df_num=8
ssr based chi2 test:   chi2=9.5012  , p=0.3018  , df=8
likelihood ratio test: chi2=9.3870  , p=0.3107  , df=8
parameter F test:         F=1.1357  , p=0.3382  , df_denom=372, df_num=8

Granger Causality
number of lags (no zero) 9
```

```
ssr based F test:         F=1.1146  , p=0.3512  , df_denom=369, df_num=9
ssr based chi2 test:   chi2=10.5481 , p=0.3080  , df=9
likelihood ratio test: chi2=10.4073 , p=0.3185  , df=9
parameter F test:         F=1.1146  , p=0.3512  , df_denom=369, df_num=9

Granger Causality
number of lags (no zero) 10
ssr based F test:         F=1.0533  , p=0.3980  , df_denom=366, df_num=10
ssr based chi2 test:   chi2=11.1372 , p=0.3469  , df=10
likelihood ratio test: chi2=10.9800 , p=0.3591  , df=10
parameter F test:         F=1.0533  , p=0.3980  , df_denom=366, df_num=10

Granger Causality
number of lags (no zero) 11
ssr based F test:         F=0.6228  , p=0.8094  , df_denom=363, df_num=11
ssr based chi2 test:   chi2=7.2852  , p=0.7755  , df=11
likelihood ratio test: chi2=7.2173  , p=0.7812  , df=11
parameter F test:         F=0.6228  , p=0.8094  , df_denom=363, df_num=11

Granger Causality
number of lags (no zero) 12
ssr based F test:         F=0.9273  , p=0.5195  , df_denom=360, df_num=12
ssr based chi2 test:   chi2=11.8998 , p=0.4538  , df=12
likelihood ratio test: chi2=11.7196 , p=0.4685  , df=12
parameter F test:         F=0.9273  , p=0.5195  , df_denom=360, df_num=12
c:\Users\HP\AppData\Local\Programs\Python\Python312\Lib\site-packages\statsmodels
\tsa\stattools.py:1556: FutureWarning: verbose is deprecated since functions shou
ld not print results
  warnings.warn(
c:\Users\HP\AppData\Local\Programs\Python\Python312\Lib\site-packages\statsmodels
\tsa\base\tsa_model.py:473: ValueWarning: No frequency information was provided,
so inferred frequency MS will be used.
  self._init_dates(dates, freq)
c:\Users\HP\AppData\Local\Programs\Python\Python312\Lib\site-packages\statsmodels
\tsa\base\tsa_model.py:473: ValueWarning: No frequency information was provided,
so inferred frequency MS will be used.
  self._init_dates(dates, freq)
```

```
                                     SARIMAX Results
==========================================================================================
Dep. Variable:                          Consumption   No. Observations:
317
Model:             SARIMAX(1, 0, 2)x(0, 1, [1], 12)   Log Likelihood
-643.379
Date:                              Fri, 30 Jan 2026   AIC
1298.758
Time:                                     15:24:55   BIC
1320.777
Sample:                                 01-01-1988   HQIC
1307.580
                                      - 05-01-2014
Covariance Type:                              opg
==========================================================================================
                             coef    std err          z      P>|z|      [0.025
0.975]
------------------------------------------------------------------------------------------
Festivals/Special_events     0.0871      0.060      1.454      0.146      -0.030
0.204
ar.L1                        0.9953      0.003    294.323      0.000       0.989
1.002
ma.L1                       -0.6029      0.069     -8.786      0.000      -0.737
-0.468
ma.L2                       -0.5299      0.057     -9.341      0.000      -0.641
-0.419
ma.S.L12                    -0.7262      0.047    -15.608      0.000      -0.817
-0.635
sigma2                       4.1056      0.315     13.036      0.000       3.488
4.723
==========================================================================================
==
Ljung-Box (L1) (Q):                   0.52   Jarque-Bera (JB):                36.
37
Prob(Q):                              0.47   Prob(JB):                         0.
00
Heteroskedasticity (H):               2.82   Skew:                            -0.
23
Prob(H) (two-sided):                  0.00   Kurtosis:                         4.
67
==========================================================================================
==

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-st
ep).
MAE:  2.8428
RMSE:  3.6325
MAPE:  0.0205
```

In [22]:

```
plt.figure(figsize=(12,6))
plt.plot(y_train.index, y_train, label='Train', color='blue')
plt.plot(y_test.index, y_test, label='Actual', color='yellow', linestyle='--')
plt.plot(y_test.index, forecast, label='Forecast', color='green', linestyle='--'
plt.legend()
plt.show()
```


Train
Actual
Forecast
160
150
140
130
120
110
100
90
1988
1992
1996
2000
2004
2008
2012
2016
2020