

```
In [7]: import numpy as np
import matplotlib_inline
import matplotlib.pyplot as plt
import pandas as pd
import os
import math
import pathlib
from pathlib import Path
import sys
```

```
In [8]: #Importing the weekly sales data set
data_path = Path(r"C:\Users\Matt\Desktop\Pharma Project\salesweekly.csv")
df = pd.read_csv(data_path)
```

```
In [9]: df
```

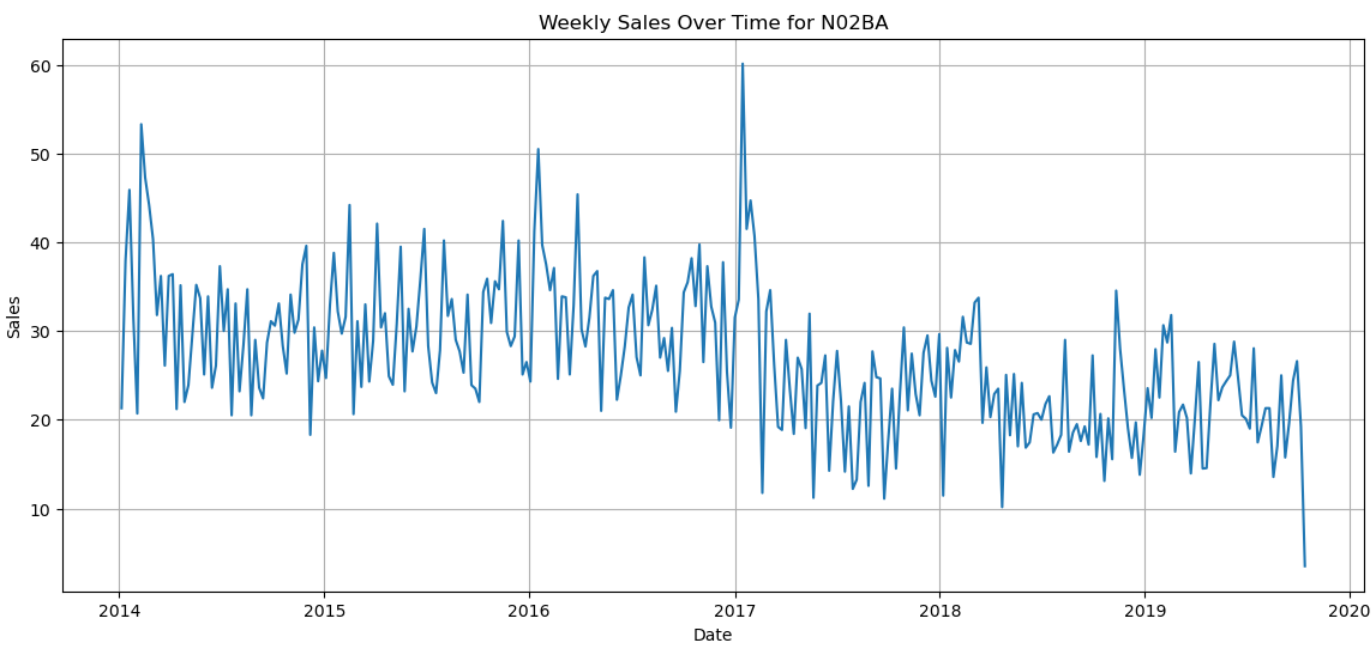
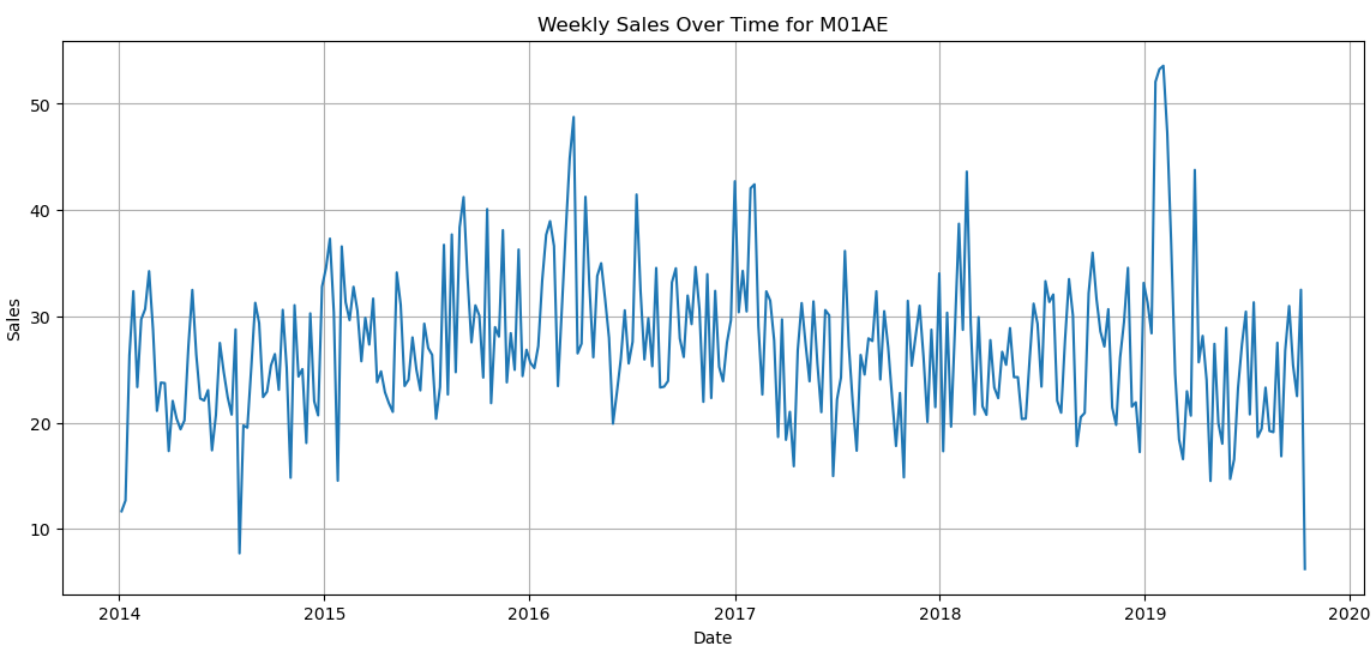
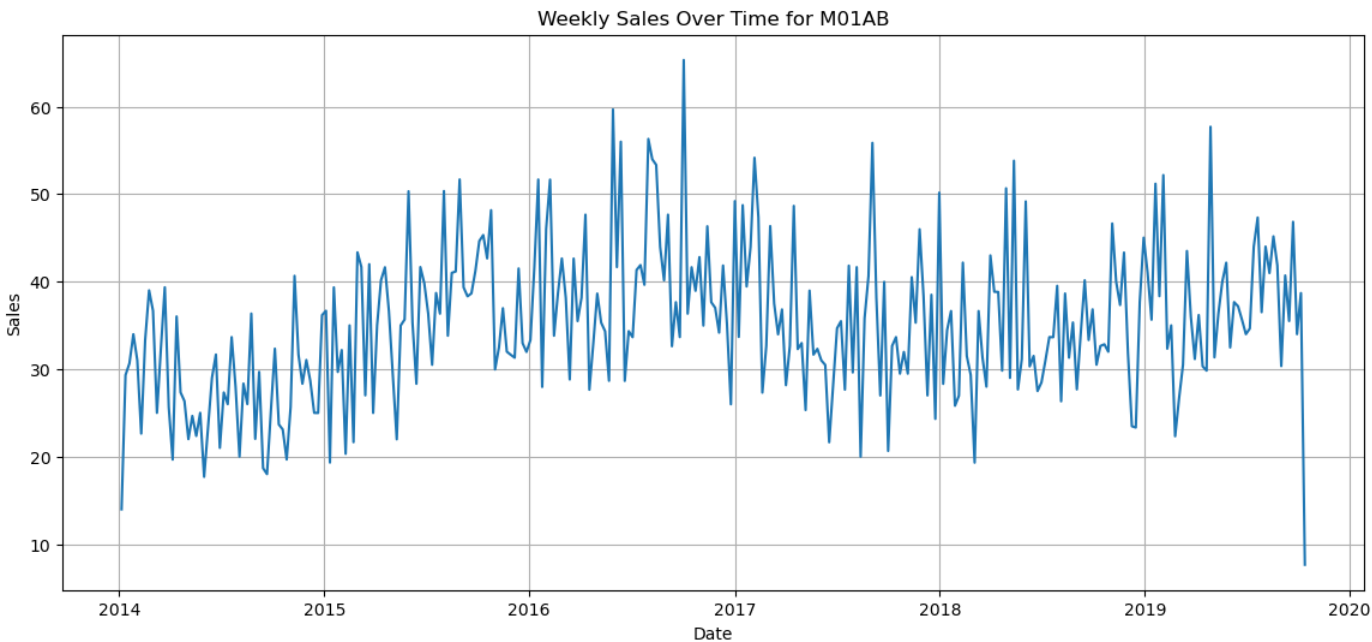
```
Out[9]:
```

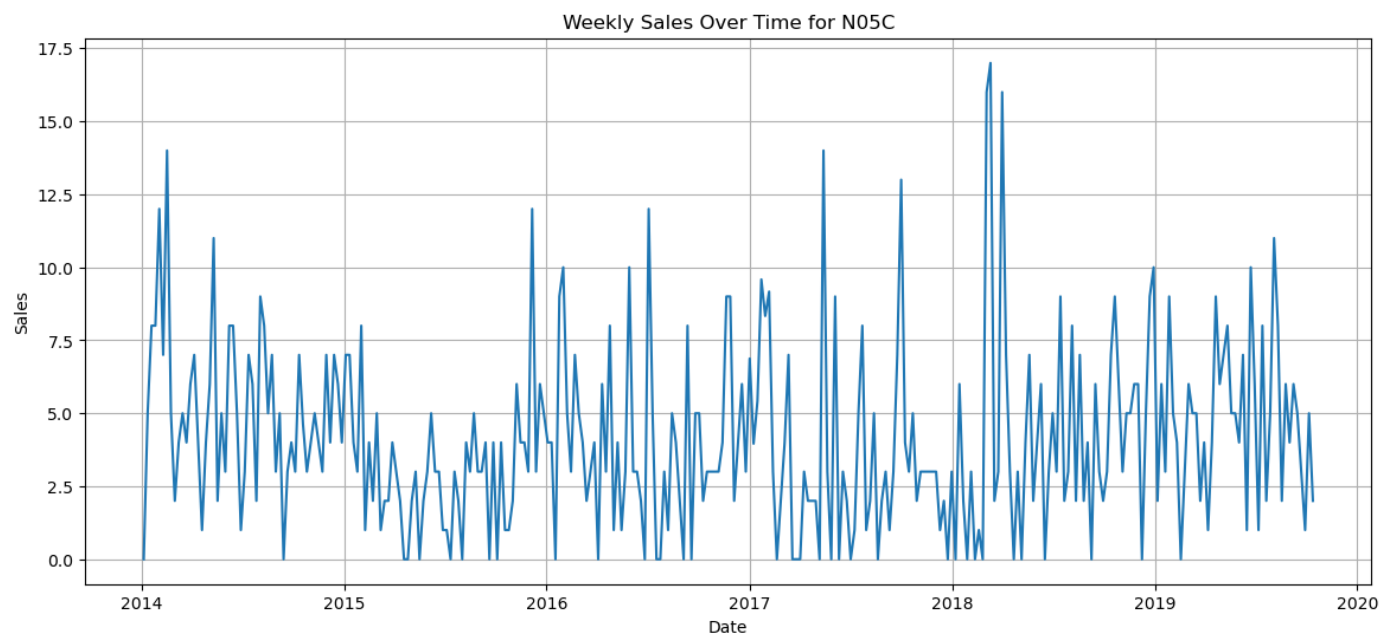
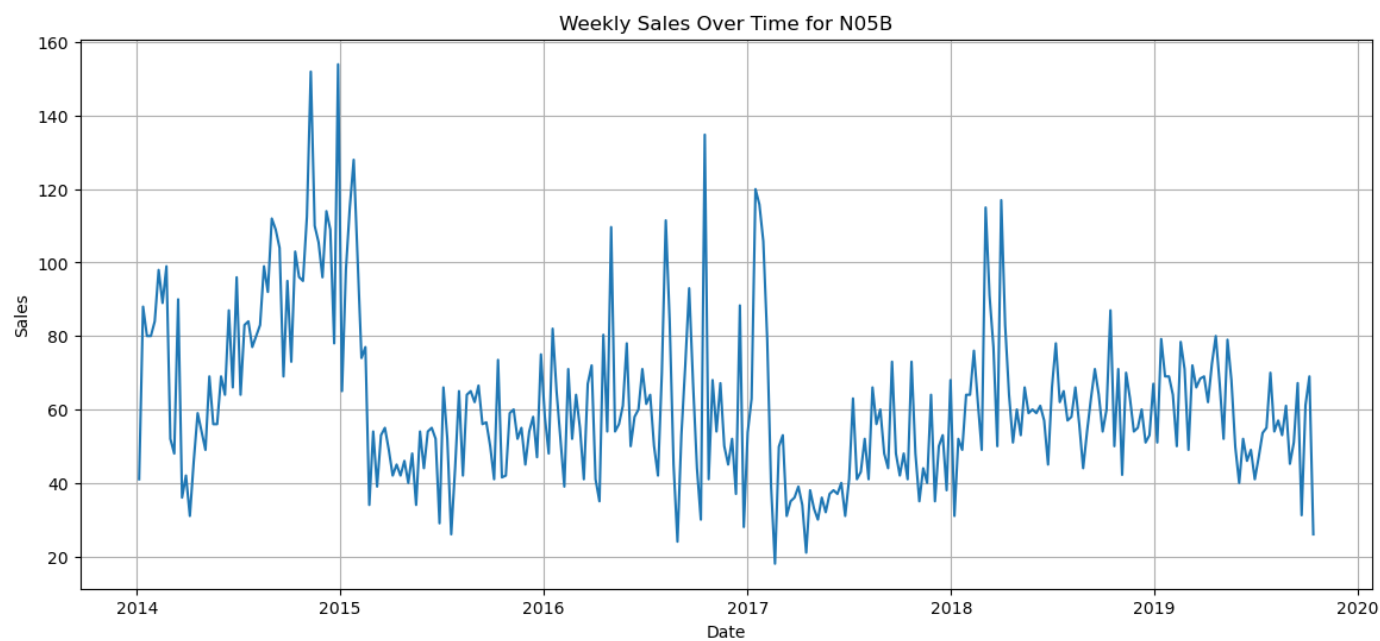
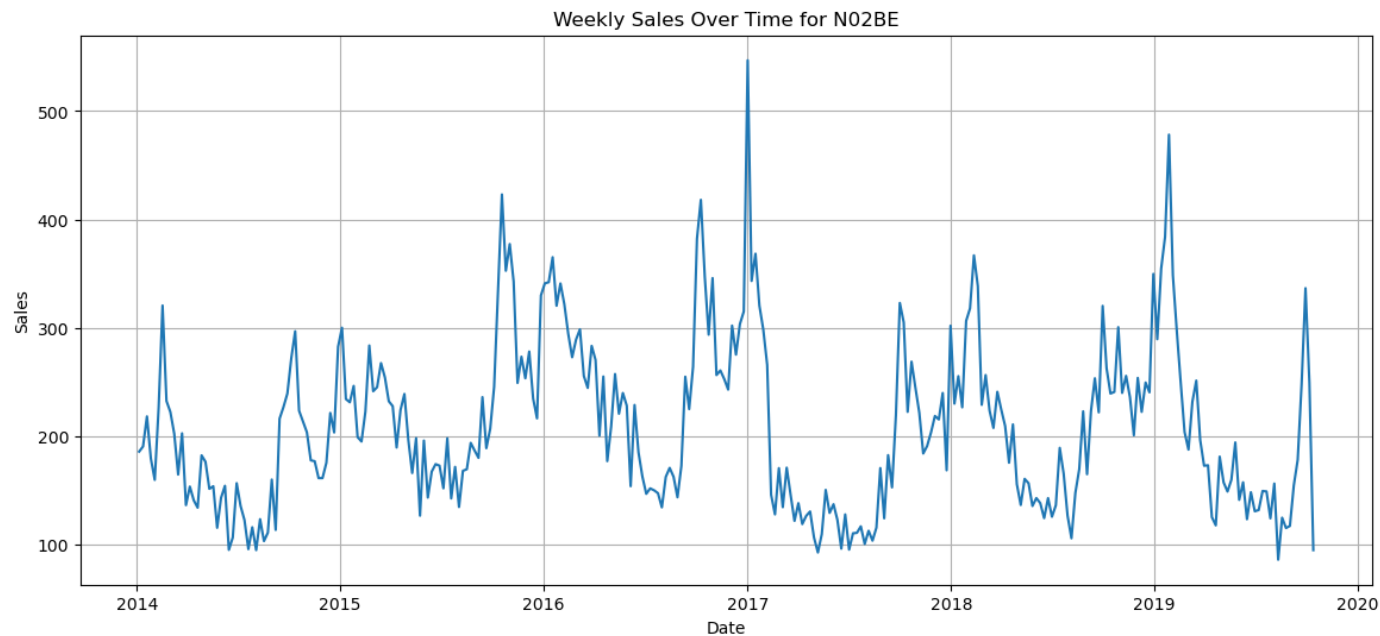
	datum	M01AB	M01AE	N02BA	N02BE	N05B	N05C	R03	R06
0	1/5/2014	14.00	11.670	21.30	185.950	41.0	0.0	32.000000	7.00
1	1/12/2014	29.33	12.680	37.90	190.700	88.0	5.0	21.000000	7.20
2	1/19/2014	30.67	26.340	45.90	218.400	80.0	8.0	29.000000	12.00
3	1/26/2014	34.00	32.370	31.50	179.600	80.0	8.0	23.000000	10.00
4	2/2/2014	31.02	23.350	20.70	159.880	84.0	12.0	29.000000	12.00
...	...	...	...	...	...	...	...	...	...
297	9/15/2019	35.51	30.977	19.65	178.375	67.2	5.0	30.000000	30.50
298	9/22/2019	46.84	25.396	24.40	248.250	31.2	3.0	26.000000	21.00
299	9/29/2019	34.01	22.498	26.60	336.700	61.2	1.0	40.416667	23.10
300	10/6/2019	38.70	32.502	19.25	249.450	69.0	5.0	30.000000	12.13
301	10/13/2019	7.67	6.237	3.50	95.100	26.0	2.0	12.000000	1.00

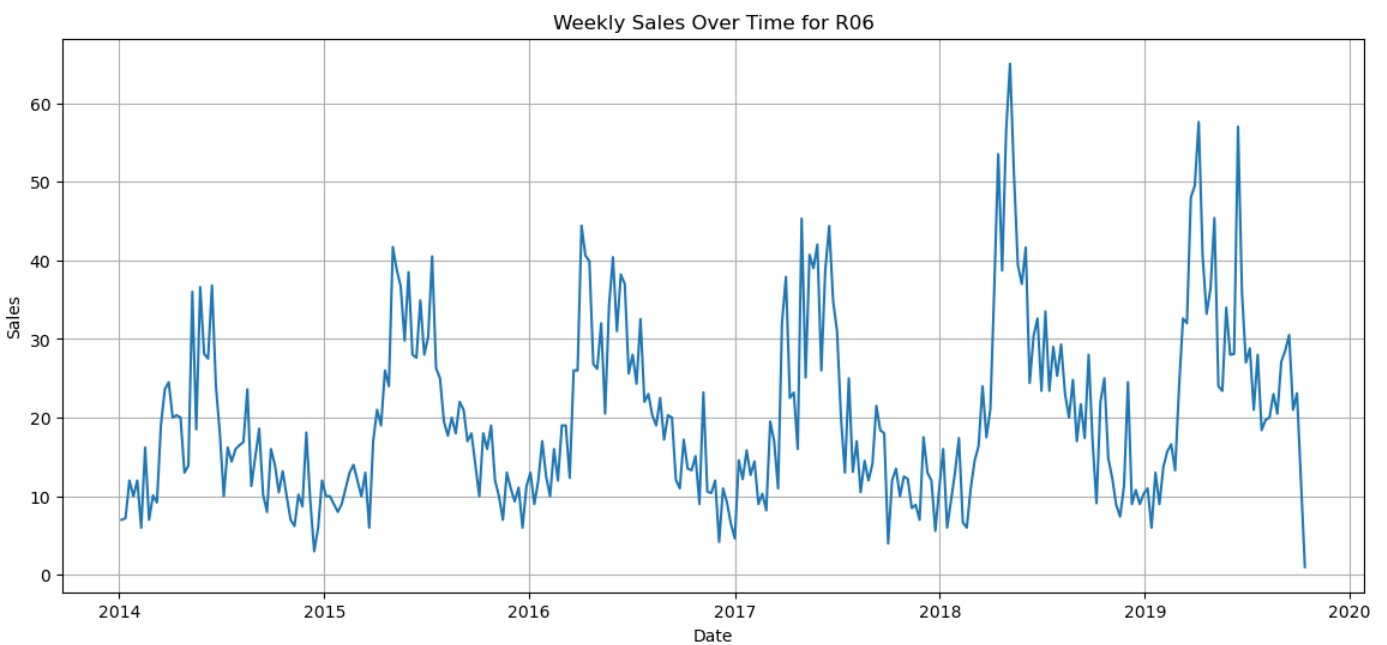
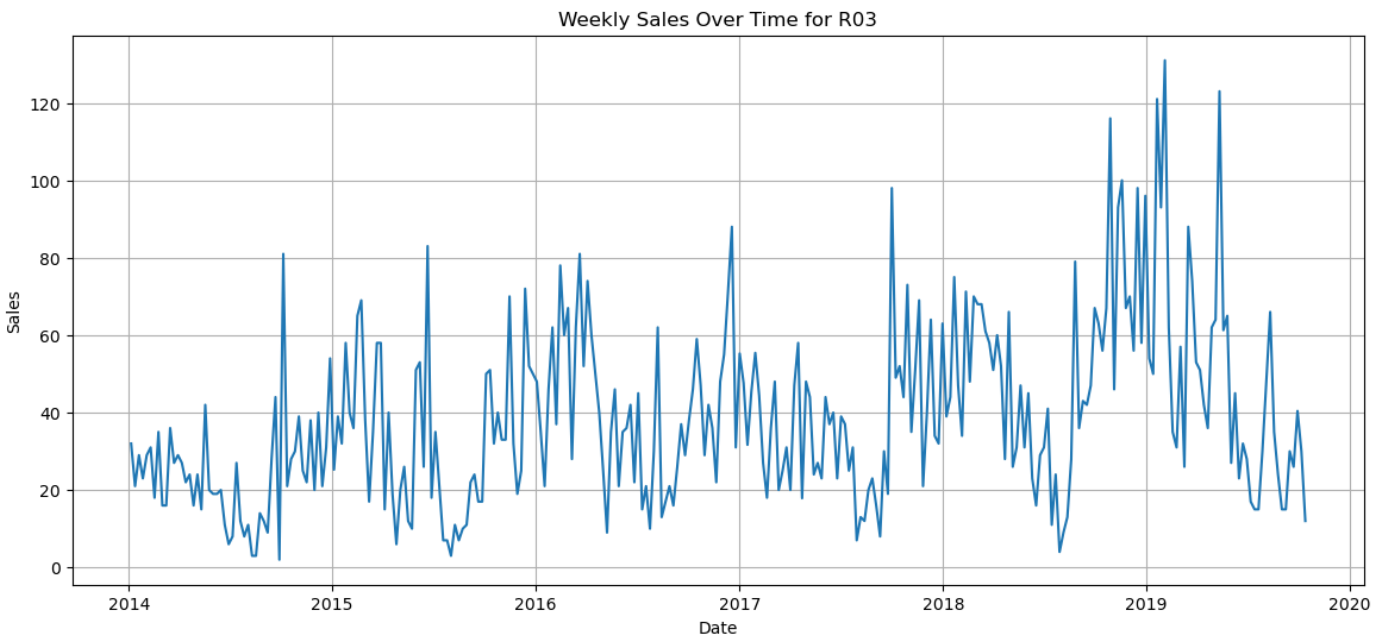
302 rows × 9 columns

```
In [10]: #Formatting the date column for time-series analysis
df['datum'] = pd.to_datetime(df['datum'])
#df['Date'] = pd.to_datetime(df['datum'])
df.set_index('datum', inplace = True)
```

```
In [11]: #Plotting the raw data set for each drug in spreadsheet
for column in df.columns:
    fig, ax = plt.subplots(figsize=(14, 6))
    ax.plot(df.index, df[column])
    ax.set_title(f"Weekly Sales Over Time for {column}")
    ax.set_xlabel('Date')
    ax.set_ylabel('Sales')
    ax.grid(True)
    plt.show()
```



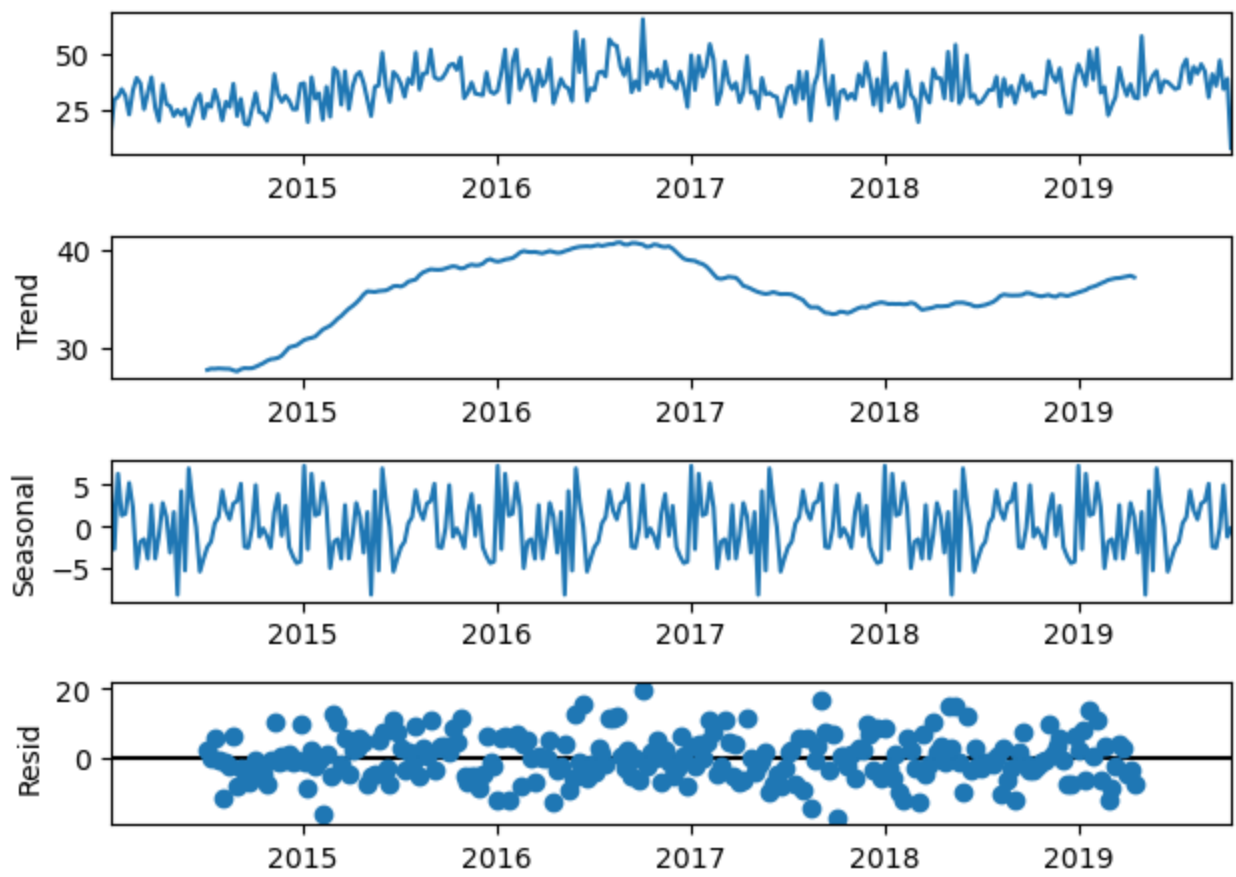




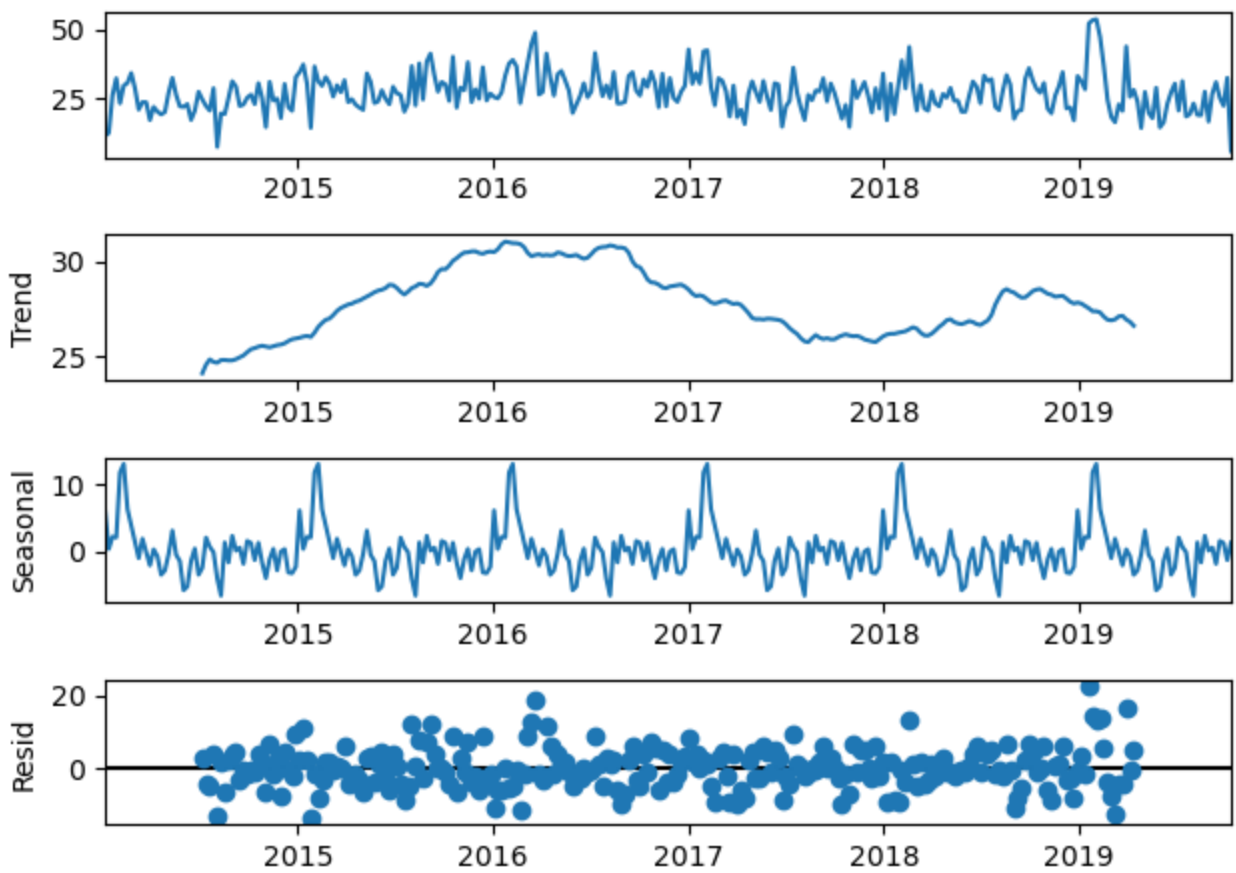
```
In [12]: #DECOMPOSITION
#In this stage, we will decompose the data into several components for analysis
#We decompose this into the trend, seasonality, and residual components.
#This will allow us to more easily forecast trends in the data set

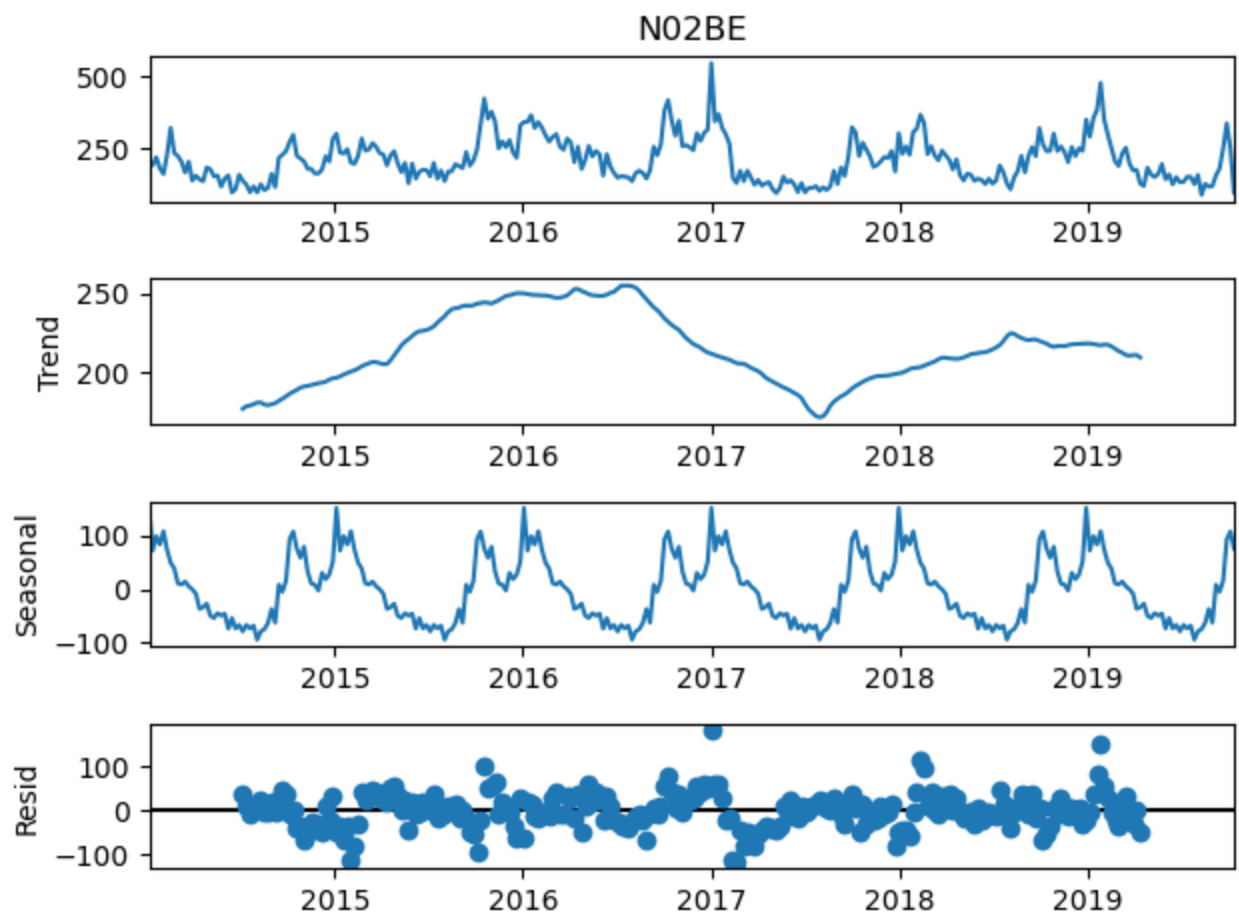
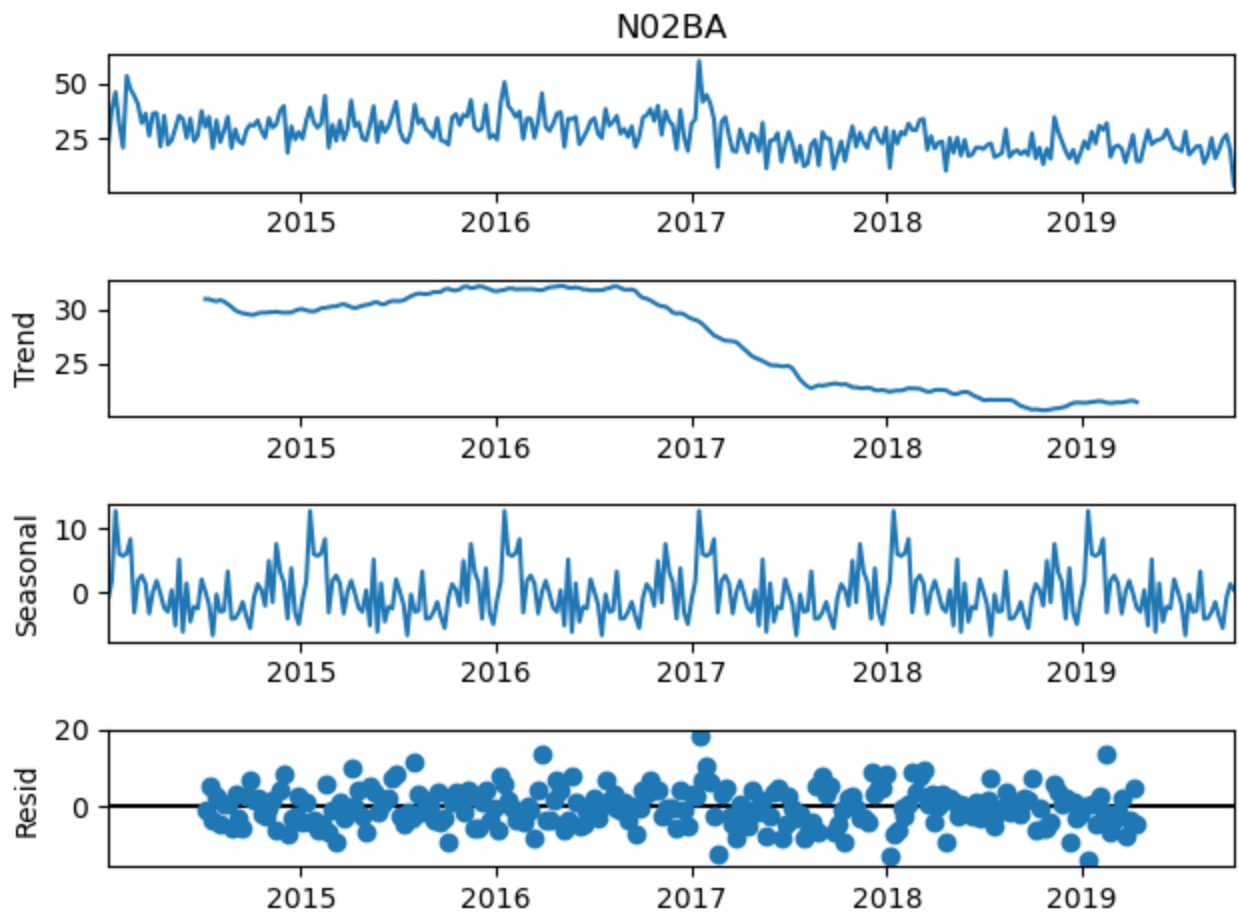
#We will use the statsmodels.tsa.decompose package to do this:
from statsmodels.tsa.seasonal import seasonal_decompose
for col in df.columns:
    #For weekly data, we use period = 52
    result = seasonal_decompose(df[col], model = 'additive', period=52)
    result.plot()
```

M01AB

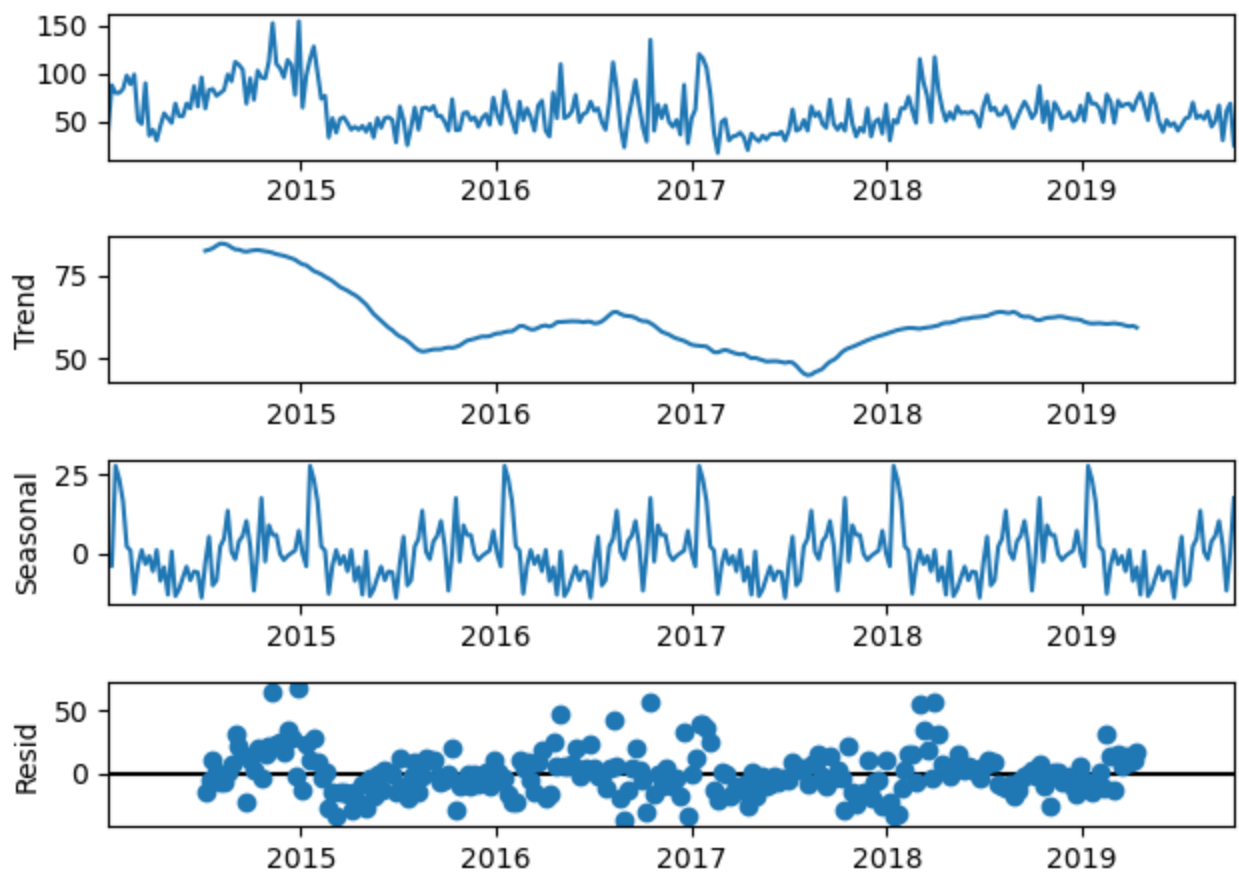


M01AE

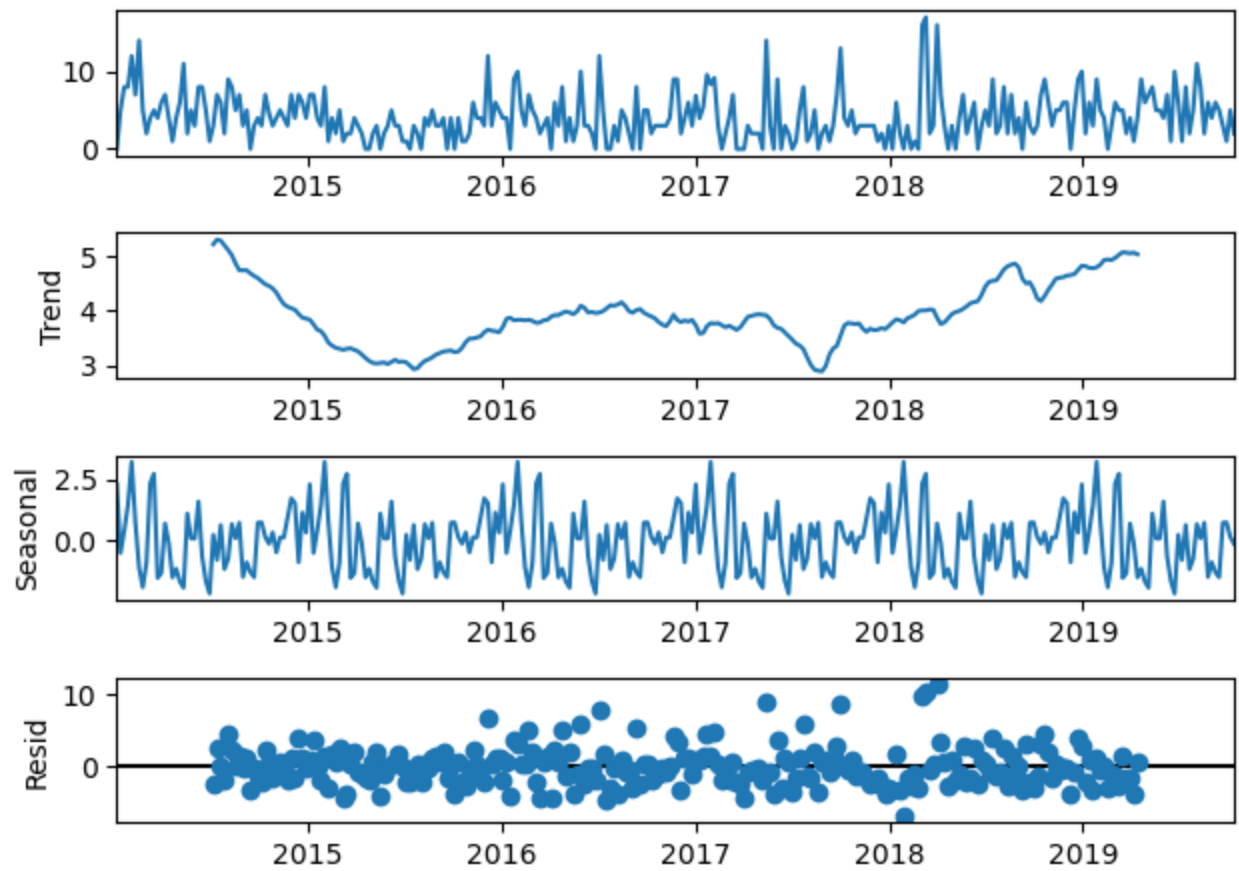


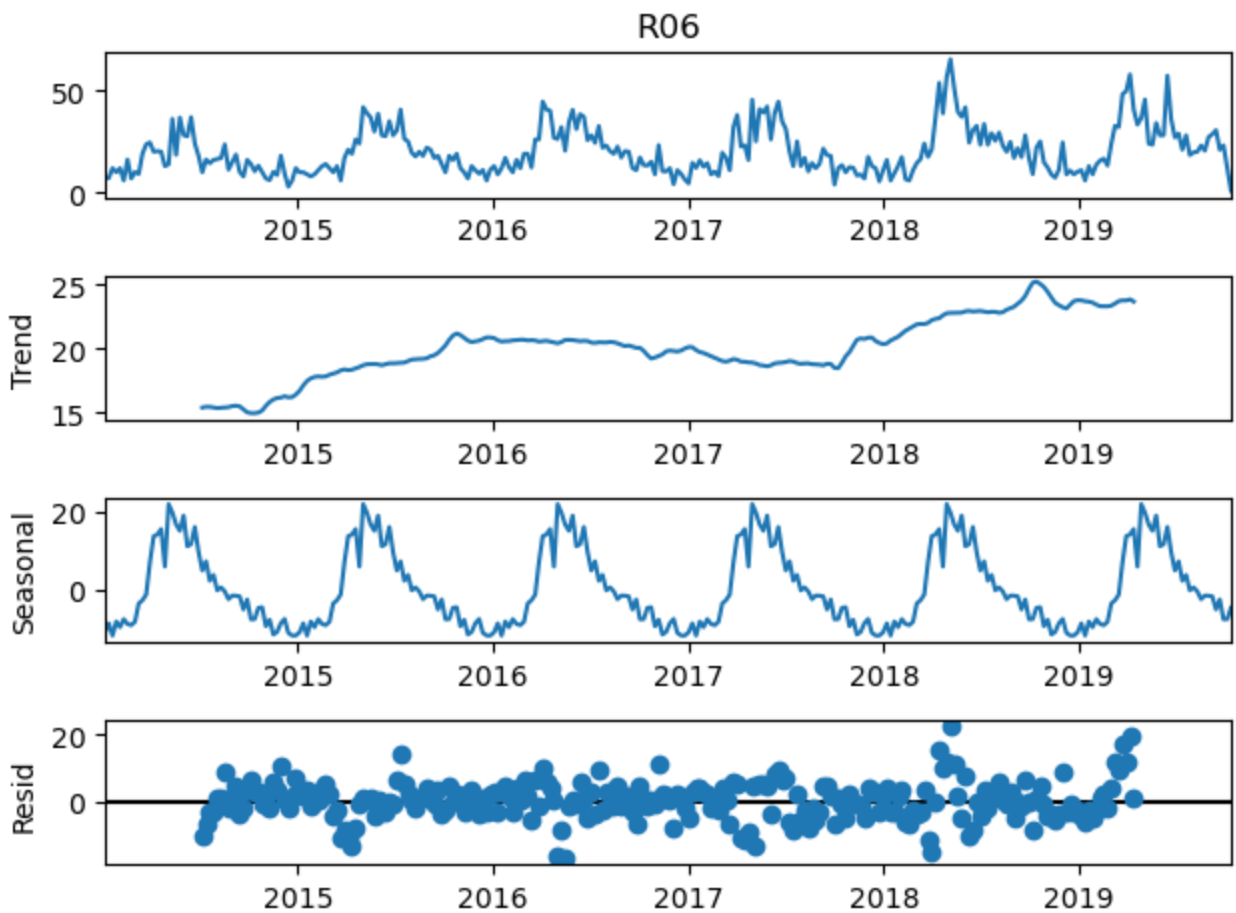
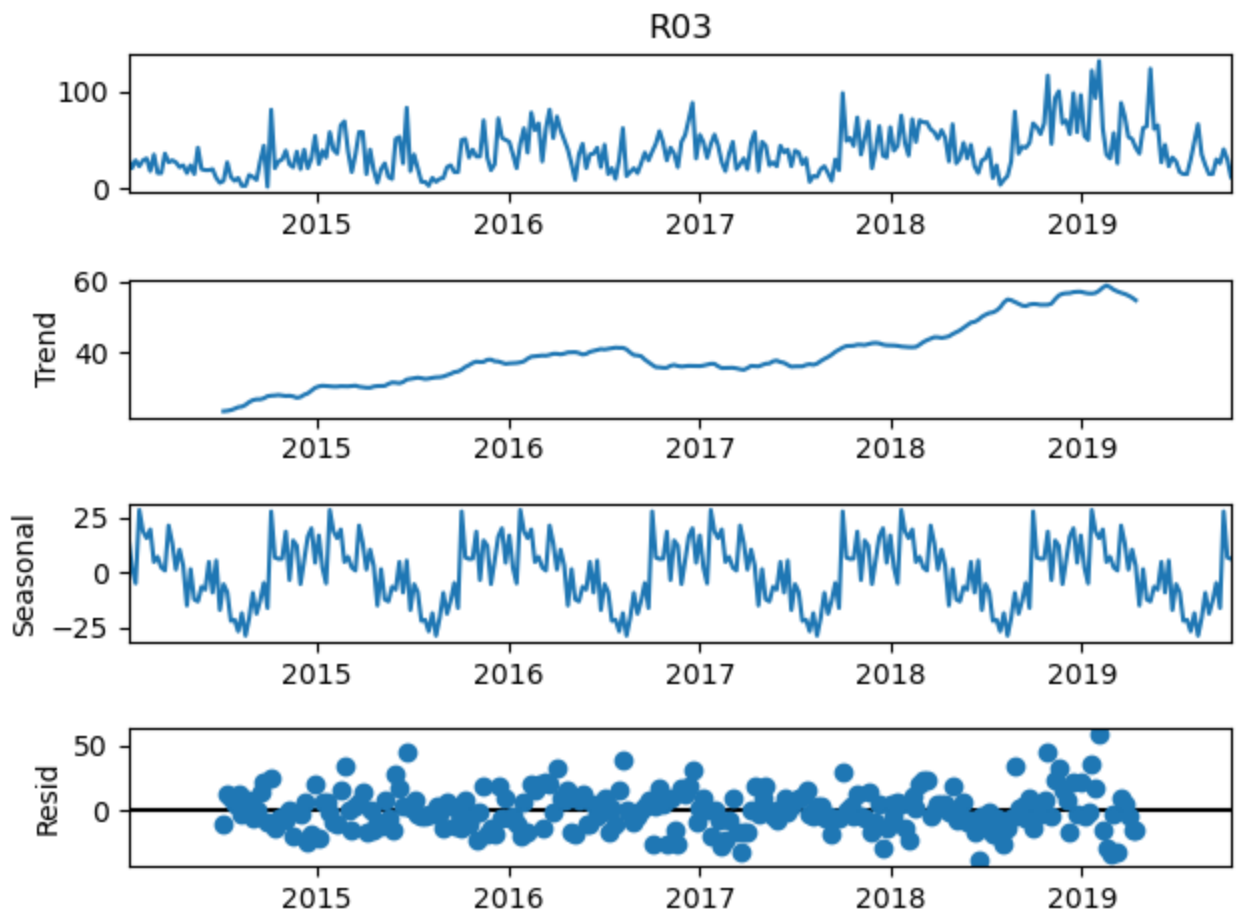


N05B



N05C





```
In [13]: #We note nonlinear trends for M01AB, M01AE, and N02BE
#We note there are strong downwards trends for N02BA and N05B
#We note there are strong upwards trends for R03,N05C, and R06
```

```
In [14]: #STATIONARITY
```



```
#We now check for stationarity in our data. We check that properties such as mean, varia  
#Many time series analysis methods require stationarity.  
#We will test each using the Augmented Dickey-Fuller test
```

```
from statsmodels.tsa.stattools import adfuller
```

```
def perform_adf_test(series):  
    result = adfuller(series)  
    p_val = result[1]  
    ADF_stat = result[0]  
    #If the p-val is less than or approximatley the threshold (0.05), then we reject the  
    print('p-value: %f' % p_val)  
    #If the ADF Stat is lower than our crit values, we can reject the null hypothesis an  
    print('ADF Statistic: %f' % ADF_stat)  
    print('Critical Values:')  
    for key, value in result[4].items():  
        print('\t%s: %.3f' % (key, value))  
  
for col in df.columns:  
    #print(col)  
    perform_adf_test(df[col])
```

```
p-value: 0.005235  
ADF Statistic: -3.628724  
Critical Values:  
    1%: -3.453  
    5%: -2.871  
   10%: -2.572  
p-value: 0.000000  
ADF Statistic: -8.739383  
Critical Values:  
    1%: -3.452  
    5%: -2.871  
   10%: -2.572  
p-value: 0.420122  
ADF Statistic: -1.721399  
Critical Values:  
    1%: -3.453  
    5%: -2.872  
   10%: -2.572  
p-value: 0.000334  
ADF Statistic: -4.370604  
Critical Values:  
    1%: -3.452  
    5%: -2.871  
   10%: -2.572  
p-value: 0.005784  
ADF Statistic: -3.598815  
Critical Values:  
    1%: -3.453  
    5%: -2.871  
   10%: -2.572  
p-value: 0.000000  
ADF Statistic: -15.000535  
Critical Values:  
    1%: -3.452  
    5%: -2.871  
   10%: -2.572  
p-value: 0.010372  
ADF Statistic: -3.417632  
Critical Values:  
    1%: -3.453  
    5%: -2.871
```

```
10%: -2.572
p-value: 0.000020
ADF Statistic: -5.028058
Critical Values:
    1%: -3.453
    5%: -2.872
   10%: -2.572
```

```
In [15]: #We note that stationarity exists for M01AB, M01AE, N02BE, N05C, N05B,R03, and R06
        #We note the null hypothesis is true for N02BA.This will require further processing.
```

```
In [16]: #DIFFERENTIATION TO ACHIEVE STATIONARITY
        # We now work to achieve stationarity on N02BA
        # To accomplish this, we apply differencing to the series. This helps to stabalize the m

        #We will use the .diff() operation to accomplish this:

df['N02BA_diff'] = df['N02BA'].diff()
#The first row will have NaN. Applying .dropna()
df = df.dropna()
#let us check whether our new series demonstrates stationarity:

perform_adf_test(df['N02BA_diff'])

p-value: 0.000000
ADF Statistic: -9.202408
Critical Values:
    1%: -3.453
    5%: -2.872
   10%: -2.572
```

```
In [17]: #we note that our new series demonstrates stationarity, as the p-val is lower than our t
```

```
In [30]: #MODELING
        #Now that we have stationarity, we can choose a model to predict sales trends for each d
        #Additonally, by analyzing the trend decompositon, we determine that seasonality exists
        #We will therefore apply models that incorporate aspects of seasonality

        #Firstly, we will utilize the SARIMA model

        #This model contains an autoregressive part, the differencing order, and the moving aver
        #denoted as (p,d,q)

        #Additionally, there are parameters such as the order of seasonal differencing, the orde
        #the order of the seasonal moving average model, and the number of periods each season
        #denoted as (P,D,Q,s)

        # For simplicity, we will perfrom a grid search to test each combination of parameters.

        #To do so, we will utilize the pmdarima library

import pmdarima as pm

import pmdarima as pm
models = {}
period = 52
for index in range(0,len(df.columns)):
    # the auto_arima function generates the best SARIMA model to the time series
    item = df.columns[index]
    print('performing function on {}'.format(item))

    model = pm.auto_arima(df[item],
                          seasonal = True,
                          m = period,
                          d = 1,
```

```

D = 1,
start_p=0,
start_q = 0,
max_p=2,
max_q=2,
trace = True,
error_action = 'ignore',
suppress_warnings = True,
stepwise = True)

models[index] = model
print('ARIMA{ }x{ }{ } - AIC:{ }'.format(model.order, model.seasonal_order, period, mode

```

performing function on M01AB

Performing stepwise search to minimize aic

```

ARIMA(0,1,0) (1,1,1) [52]      : AIC=inf, Time=12.13 sec
ARIMA(0,1,0) (0,1,0) [52]      : AIC=2054.888, Time=0.36 sec
ARIMA(1,1,0) (1,1,0) [52]      : AIC=1920.239, Time=2.62 sec
ARIMA(0,1,1) (0,1,1) [52]      : AIC=inf, Time=5.14 sec
ARIMA(1,1,0) (0,1,0) [52]      : AIC=1967.726, Time=0.36 sec
ARIMA(1,1,0) (2,1,0) [52]      : AIC=1904.494, Time=6.16 sec
ARIMA(1,1,0) (2,1,1) [52]      : AIC=inf, Time=33.34 sec
ARIMA(1,1,0) (1,1,1) [52]      : AIC=inf, Time=18.25 sec
ARIMA(0,1,0) (2,1,0) [52]      : AIC=inf, Time=4.82 sec
ARIMA(2,1,0) (2,1,0) [52]      : AIC=1885.698, Time=7.73 sec
ARIMA(2,1,0) (1,1,0) [52]      : AIC=1899.143, Time=3.69 sec
ARIMA(2,1,0) (2,1,1) [52]      : AIC=inf, Time=35.97 sec
ARIMA(2,1,0) (1,1,1) [52]      : AIC=inf, Time=23.24 sec
ARIMA(2,1,1) (2,1,0) [52]      : AIC=1831.898, Time=14.62 sec
ARIMA(2,1,1) (1,1,0) [52]      : AIC=1851.987, Time=7.12 sec
ARIMA(2,1,1) (2,1,1) [52]      : AIC=inf, Time=46.99 sec
ARIMA(2,1,1) (1,1,1) [52]      : AIC=inf, Time=27.19 sec
ARIMA(1,1,1) (2,1,0) [52]      : AIC=1829.902, Time=12.02 sec
ARIMA(1,1,1) (1,1,0) [52]      : AIC=1850.008, Time=4.87 sec
ARIMA(1,1,1) (2,1,1) [52]      : AIC=inf, Time=29.89 sec
ARIMA(1,1,1) (1,1,1) [52]      : AIC=inf, Time=24.04 sec
ARIMA(0,1,1) (2,1,0) [52]      : AIC=1828.774, Time=8.44 sec
ARIMA(0,1,1) (1,1,0) [52]      : AIC=1849.800, Time=2.93 sec
ARIMA(0,1,1) (2,1,1) [52]      : AIC=inf, Time=27.80 sec
ARIMA(0,1,1) (1,1,1) [52]      : AIC=inf, Time=17.34 sec
ARIMA(0,1,2) (2,1,0) [52]      : AIC=1829.884, Time=13.53 sec
ARIMA(1,1,2) (2,1,0) [52]      : AIC=1830.993, Time=23.61 sec
ARIMA(0,1,1) (2,1,0) [52] intercept : AIC=1830.166, Time=11.91 sec

```

Best model: ARIMA(0,1,1) (2,1,0) [52]

Total fit time: 426.123 seconds

ARIMA(0, 1, 1)x(2, 1, 0, 52)52 - AIC:1828.7742841944935

performing function on M01AE

Performing stepwise search to minimize aic

```

ARIMA(0,1,0) (1,1,1) [52]      : AIC=inf, Time=10.86 sec
ARIMA(0,1,0) (0,1,0) [52]      : AIC=1894.706, Time=0.36 sec
ARIMA(1,1,0) (1,1,0) [52]      : AIC=1804.024, Time=2.25 sec
ARIMA(0,1,1) (0,1,1) [52]      : AIC=inf, Time=3.18 sec
ARIMA(1,1,0) (0,1,0) [52]      : AIC=1839.057, Time=0.34 sec
ARIMA(1,1,0) (2,1,0) [52]      : AIC=1791.414, Time=6.05 sec
ARIMA(1,1,0) (2,1,1) [52]      : AIC=inf, Time=33.61 sec
ARIMA(1,1,0) (1,1,1) [52]      : AIC=inf, Time=17.66 sec
ARIMA(0,1,0) (2,1,0) [52]      : AIC=1844.687, Time=4.90 sec
ARIMA(2,1,0) (2,1,0) [52]      : AIC=1778.298, Time=7.32 sec
ARIMA(2,1,0) (1,1,0) [52]      : AIC=1792.014, Time=3.40 sec
ARIMA(2,1,0) (2,1,1) [52]      : AIC=inf, Time=33.85 sec
ARIMA(2,1,0) (1,1,1) [52]      : AIC=inf, Time=24.15 sec
ARIMA(2,1,1) (2,1,0) [52]      : AIC=1730.234, Time=16.23 sec
ARIMA(2,1,1) (1,1,0) [52]      : AIC=1741.390, Time=7.67 sec
ARIMA(2,1,1) (2,1,1) [52]      : AIC=inf, Time=28.28 sec
ARIMA(2,1,1) (1,1,1) [52]      : AIC=inf, Time=13.00 sec
ARIMA(1,1,1) (2,1,0) [52]      : AIC=1728.679, Time=12.28 sec

```

```

ARIMA(1,1,1)(1,1,0)[52] : AIC=1739.877, Time=6.23 sec
ARIMA(1,1,1)(2,1,1)[52] : AIC=inf, Time=21.16 sec
ARIMA(1,1,1)(1,1,1)[52] : AIC=inf, Time=10.24 sec
ARIMA(0,1,1)(2,1,0)[52] : AIC=1730.437, Time=8.31 sec
ARIMA(1,1,2)(2,1,0)[52] : AIC=1730.446, Time=25.69 sec
ARIMA(0,1,2)(2,1,0)[52] : AIC=1729.077, Time=11.82 sec
ARIMA(2,1,2)(2,1,0)[52] : AIC=1732.046, Time=30.62 sec
ARIMA(1,1,1)(2,1,0)[52] intercept : AIC=1728.404, Time=17.50 sec
ARIMA(1,1,1)(1,1,0)[52] intercept : AIC=1740.513, Time=8.07 sec
ARIMA(1,1,1)(2,1,1)[52] intercept : AIC=inf, Time=34.38 sec
ARIMA(1,1,1)(1,1,1)[52] intercept : AIC=inf, Time=18.37 sec
ARIMA(0,1,1)(2,1,0)[52] intercept : AIC=1730.499, Time=10.81 sec
ARIMA(1,1,0)(2,1,0)[52] intercept : AIC=1793.290, Time=11.15 sec
ARIMA(2,1,1)(2,1,0)[52] intercept : AIC=1729.899, Time=18.30 sec
ARIMA(1,1,2)(2,1,0)[52] intercept : AIC=1730.128, Time=30.86 sec
ARIMA(0,1,0)(2,1,0)[52] intercept : AIC=inf, Time=8.95 sec
ARIMA(0,1,2)(2,1,0)[52] intercept : AIC=1728.859, Time=15.51 sec
ARIMA(2,1,0)(2,1,0)[52] intercept : AIC=1780.162, Time=16.30 sec
ARIMA(2,1,2)(2,1,0)[52] intercept : AIC=1731.721, Time=39.79 sec

```

Best model: ARIMA(1,1,1)(2,1,0)[52] intercept

Total fit time: 569.509 seconds

ARIMA(1, 1, 1)x(2, 1, 0, 52)52 - AIC:1728.4041141243906

performing function on N02BA

Performing stepwise search to minimize aic

```

ARIMA(0,1,0)(1,1,1)[52] : AIC=inf, Time=6.20 sec
ARIMA(0,1,0)(0,1,0)[52] : AIC=1864.233, Time=0.43 sec
ARIMA(1,1,0)(1,1,0)[52] : AIC=1773.416, Time=2.44 sec
ARIMA(0,1,1)(0,1,1)[52] : AIC=inf, Time=4.21 sec
ARIMA(1,1,0)(0,1,0)[52] : AIC=1810.160, Time=0.34 sec
ARIMA(1,1,0)(2,1,0)[52] : AIC=1770.861, Time=5.10 sec
ARIMA(1,1,0)(2,1,1)[52] : AIC=inf, Time=31.96 sec
ARIMA(1,1,0)(1,1,1)[52] : AIC=inf, Time=10.90 sec
ARIMA(0,1,0)(2,1,0)[52] : AIC=1825.019, Time=4.24 sec
ARIMA(2,1,0)(2,1,0)[52] : AIC=1756.362, Time=7.49 sec
ARIMA(2,1,0)(1,1,0)[52] : AIC=1758.151, Time=3.44 sec
ARIMA(2,1,0)(2,1,1)[52] : AIC=inf, Time=51.38 sec
ARIMA(2,1,0)(1,1,1)[52] : AIC=inf, Time=21.53 sec
ARIMA(2,1,1)(2,1,0)[52] : AIC=1707.434, Time=13.97 sec
ARIMA(2,1,1)(1,1,0)[52] : AIC=1711.307, Time=6.72 sec
ARIMA(2,1,1)(2,1,1)[52] : AIC=inf, Time=50.78 sec
ARIMA(2,1,1)(1,1,1)[52] : AIC=inf, Time=27.66 sec
ARIMA(1,1,1)(2,1,0)[52] : AIC=1705.509, Time=11.12 sec
ARIMA(1,1,1)(1,1,0)[52] : AIC=1709.366, Time=5.01 sec
ARIMA(1,1,1)(2,1,1)[52] : AIC=inf, Time=44.38 sec
ARIMA(1,1,1)(1,1,1)[52] : AIC=inf, Time=20.56 sec
ARIMA(0,1,1)(2,1,0)[52] : AIC=1705.748, Time=6.98 sec
ARIMA(1,1,2)(2,1,0)[52] : AIC=1707.477, Time=23.56 sec
ARIMA(0,1,2)(2,1,0)[52] : AIC=1705.640, Time=11.70 sec
ARIMA(2,1,2)(2,1,0)[52] : AIC=1709.192, Time=24.29 sec
ARIMA(1,1,1)(2,1,0)[52] intercept : AIC=1707.504, Time=14.55 sec

```

Best model: ARIMA(1,1,1)(2,1,0)[52]

Total fit time: 410.946 seconds

ARIMA(1, 1, 1)x(2, 1, 0, 52)52 - AIC:1705.5086242698335

performing function on N02BE

Performing stepwise search to minimize aic

```

ARIMA(0,1,0)(1,1,1)[52] : AIC=inf, Time=6.55 sec
ARIMA(0,1,0)(0,1,0)[52] : AIC=2708.173, Time=0.48 sec
ARIMA(1,1,0)(1,1,0)[52] : AIC=2639.439, Time=3.71 sec
ARIMA(0,1,1)(0,1,1)[52] : AIC=inf, Time=5.44 sec
ARIMA(1,1,0)(0,1,0)[52] : AIC=2686.607, Time=0.48 sec
ARIMA(1,1,0)(2,1,0)[52] : AIC=2625.980, Time=8.47 sec
ARIMA(1,1,0)(2,1,1)[52] : AIC=2627.167, Time=16.76 sec
ARIMA(1,1,0)(1,1,1)[52] : AIC=inf, Time=11.54 sec
ARIMA(0,1,0)(2,1,0)[52] : AIC=inf, Time=7.89 sec

```

```

ARIMA(2,1,0) (1,1,0) [52] : AIC=2623.434, Time=9.94 sec
ARIMA(2,1,0) (1,1,0) [52] : AIC=2636.265, Time=4.24 sec
ARIMA(2,1,0) (2,1,1) [52] : AIC=2624.812, Time=24.03 sec
ARIMA(2,1,0) (1,1,1) [52] : AIC=inf, Time=14.28 sec
ARIMA(2,1,1) (2,1,0) [52] : AIC=2607.191, Time=21.65 sec
ARIMA(2,1,1) (1,1,0) [52] : AIC=2621.460, Time=10.52 sec
ARIMA(2,1,1) (2,1,1) [52] : AIC=2608.608, Time=42.43 sec
ARIMA(2,1,1) (1,1,1) [52] : AIC=inf, Time=20.94 sec
ARIMA(1,1,1) (2,1,0) [52] : AIC=2607.484, Time=27.56 sec
ARIMA(2,1,2) (2,1,0) [52] : AIC=2610.062, Time=37.77 sec
ARIMA(1,1,2) (2,1,0) [52] : AIC=2607.426, Time=26.11 sec
ARIMA(2,1,1) (2,1,0) [52] intercept : AIC=2607.941, Time=31.67 sec

```

Best model: ARIMA(2,1,1) (2,1,0) [52]

Total fit time: 332.493 seconds

ARIMA(2, 1, 1)x(2, 1, 0, 52)52 - AIC:2607.1905433794404

performing function on N05B

Performing stepwise search to minimize aic

```

ARIMA(0,1,0) (1,1,1) [52] : AIC=inf, Time=12.16 sec
ARIMA(0,1,0) (0,1,0) [52] : AIC=2434.738, Time=0.39 sec
ARIMA(1,1,0) (1,1,0) [52] : AIC=2314.447, Time=3.22 sec
ARIMA(0,1,1) (0,1,1) [52] : AIC=inf, Time=7.03 sec
ARIMA(1,1,0) (0,1,0) [52] : AIC=2371.591, Time=0.57 sec
ARIMA(1,1,0) (2,1,0) [52] : AIC=2310.399, Time=6.79 sec
ARIMA(1,1,0) (2,1,1) [52] : AIC=inf, Time=37.90 sec
ARIMA(1,1,0) (1,1,1) [52] : AIC=inf, Time=18.41 sec
ARIMA(0,1,0) (2,1,0) [52] : AIC=2373.717, Time=5.68 sec
ARIMA(2,1,0) (2,1,0) [52] : AIC=2296.149, Time=9.33 sec
ARIMA(2,1,0) (1,1,0) [52] : AIC=2301.039, Time=4.29 sec
ARIMA(2,1,0) (2,1,1) [52] : AIC=inf, Time=44.23 sec
ARIMA(2,1,0) (1,1,1) [52] : AIC=inf, Time=23.67 sec
ARIMA(2,1,1) (2,1,0) [52] : AIC=2266.705, Time=18.10 sec
ARIMA(2,1,1) (1,1,0) [52] : AIC=2268.808, Time=8.28 sec
ARIMA(2,1,1) (2,1,1) [52] : AIC=inf, Time=48.91 sec
ARIMA(2,1,1) (1,1,1) [52] : AIC=inf, Time=23.25 sec
ARIMA(1,1,1) (2,1,0) [52] : AIC=2264.726, Time=12.10 sec
ARIMA(1,1,1) (1,1,0) [52] : AIC=2266.853, Time=6.65 sec
ARIMA(1,1,1) (2,1,1) [52] : AIC=inf, Time=44.86 sec
ARIMA(1,1,1) (1,1,1) [52] : AIC=inf, Time=17.65 sec
ARIMA(0,1,1) (2,1,0) [52] : AIC=2262.728, Time=8.72 sec
ARIMA(0,1,1) (1,1,0) [52] : AIC=2264.897, Time=3.43 sec
ARIMA(0,1,1) (2,1,1) [52] : AIC=inf, Time=33.83 sec
ARIMA(0,1,1) (1,1,1) [52] : AIC=inf, Time=11.30 sec
ARIMA(0,1,2) (2,1,0) [52] : AIC=2264.726, Time=10.93 sec
ARIMA(1,1,2) (2,1,0) [52] : AIC=2266.475, Time=18.81 sec
ARIMA(0,1,1) (2,1,0) [52] intercept : AIC=2264.725, Time=11.87 sec

```

Best model: ARIMA(0,1,1) (2,1,0) [52]

Total fit time: 452.377 seconds

ARIMA(0, 1, 1)x(2, 1, 0, 52)52 - AIC:2262.7276842635547

performing function on N05C

Performing stepwise search to minimize aic

```

ARIMA(0,1,0) (1,1,1) [52] : AIC=inf, Time=12.61 sec
ARIMA(0,1,0) (0,1,0) [52] : AIC=1571.477, Time=0.51 sec
ARIMA(1,1,0) (1,1,0) [52] : AIC=1481.565, Time=2.49 sec
ARIMA(0,1,1) (0,1,1) [52] : AIC=inf, Time=4.63 sec
ARIMA(1,1,0) (0,1,0) [52] : AIC=1521.103, Time=0.29 sec
ARIMA(1,1,0) (2,1,0) [52] : AIC=1464.505, Time=6.30 sec
ARIMA(1,1,0) (2,1,1) [52] : AIC=inf, Time=30.78 sec
ARIMA(1,1,0) (1,1,1) [52] : AIC=inf, Time=14.91 sec
ARIMA(0,1,0) (2,1,0) [52] : AIC=1518.932, Time=3.84 sec
ARIMA(2,1,0) (2,1,0) [52] : AIC=inf, Time=6.96 sec
ARIMA(1,1,1) (2,1,0) [52] : AIC=1381.895, Time=11.03 sec
ARIMA(1,1,1) (1,1,0) [52] : AIC=1399.458, Time=5.38 sec
ARIMA(1,1,1) (2,1,1) [52] : AIC=inf, Time=40.02 sec
ARIMA(1,1,1) (1,1,1) [52] : AIC=inf, Time=20.77 sec

```

```

ARIMA(0,1,1)(2,1,0)[52] : AIC=1380.771, Time=8.05 sec
ARIMA(0,1,1)(1,1,0)[52] : AIC=1398.723, Time=3.19 sec
ARIMA(0,1,1)(2,1,1)[52] : AIC=inf, Time=35.58 sec
ARIMA(0,1,1)(1,1,1)[52] : AIC=inf, Time=17.04 sec
ARIMA(0,1,2)(2,1,0)[52] : AIC=1381.840, Time=11.50 sec
ARIMA(1,1,2)(2,1,0)[52] : AIC=1382.540, Time=17.95 sec
ARIMA(0,1,1)(2,1,0)[52] intercept : AIC=1380.725, Time=13.10 sec
ARIMA(0,1,1)(1,1,0)[52] intercept : AIC=1398.971, Time=7.40 sec
ARIMA(0,1,1)(2,1,1)[52] intercept : AIC=inf, Time=38.27 sec
ARIMA(0,1,1)(1,1,1)[52] intercept : AIC=inf, Time=19.51 sec
ARIMA(0,1,0)(2,1,0)[52] intercept : AIC=1520.929, Time=6.87 sec
ARIMA(1,1,1)(2,1,0)[52] intercept : AIC=inf, Time=34.07 sec
ARIMA(0,1,2)(2,1,0)[52] intercept : AIC=inf, Time=34.18 sec
ARIMA(1,1,0)(2,1,0)[52] intercept : AIC=1466.505, Time=7.93 sec
ARIMA(1,1,2)(2,1,0)[52] intercept : AIC=1382.337, Time=26.67 sec

```

Best model: ARIMA(0,1,1)(2,1,0)[52] intercept

Total fit time: 441.844 seconds

ARIMA(0, 1, 1)x(2, 1, 0, 52)52 - AIC:1380.725396472536

performing function on R03

Performing stepwise search to minimize aic

```

ARIMA(0,1,0)(1,1,1)[52] : AIC=inf, Time=7.54 sec
ARIMA(0,1,0)(0,1,0)[52] : AIC=2393.189, Time=0.35 sec
ARIMA(1,1,0)(1,1,0)[52] : AIC=2290.547, Time=2.88 sec
ARIMA(0,1,1)(0,1,1)[52] : AIC=inf, Time=5.23 sec
ARIMA(1,1,0)(0,1,0)[52] : AIC=2339.287, Time=0.36 sec
ARIMA(1,1,0)(2,1,0)[52] : AIC=inf, Time=8.03 sec
ARIMA(1,1,0)(1,1,1)[52] : AIC=inf, Time=10.38 sec
ARIMA(1,1,0)(0,1,1)[52] : AIC=inf, Time=6.01 sec
ARIMA(1,1,0)(2,1,1)[52] : AIC=inf, Time=23.24 sec
ARIMA(0,1,0)(1,1,0)[52] : AIC=2345.288, Time=1.42 sec
ARIMA(2,1,0)(1,1,0)[52] : AIC=2281.781, Time=3.89 sec
ARIMA(2,1,0)(0,1,0)[52] : AIC=2322.036, Time=0.59 sec
ARIMA(2,1,0)(2,1,0)[52] : AIC=inf, Time=9.18 sec
ARIMA(2,1,0)(1,1,1)[52] : AIC=inf, Time=8.84 sec
ARIMA(2,1,0)(0,1,1)[52] : AIC=inf, Time=8.55 sec
ARIMA(2,1,0)(2,1,1)[52] : AIC=inf, Time=31.03 sec
ARIMA(2,1,1)(1,1,0)[52] : AIC=inf, Time=16.97 sec
ARIMA(1,1,1)(1,1,0)[52] : AIC=2244.805, Time=11.23 sec
ARIMA(1,1,1)(0,1,0)[52] : AIC=2291.391, Time=1.08 sec
ARIMA(1,1,1)(2,1,0)[52] : AIC=2231.657, Time=21.78 sec
ARIMA(1,1,1)(2,1,1)[52] : AIC=inf, Time=48.18 sec
ARIMA(1,1,1)(1,1,1)[52] : AIC=inf, Time=17.33 sec
ARIMA(0,1,1)(2,1,0)[52] : AIC=2235.435, Time=11.29 sec
ARIMA(2,1,1)(2,1,0)[52] : AIC=inf, Time=35.98 sec
ARIMA(1,1,2)(2,1,0)[52] : AIC=inf, Time=39.22 sec
ARIMA(0,1,0)(2,1,0)[52] : AIC=2334.383, Time=4.81 sec
ARIMA(0,1,2)(2,1,0)[52] : AIC=2233.287, Time=26.31 sec
ARIMA(2,1,2)(2,1,0)[52] : AIC=inf, Time=50.52 sec
ARIMA(1,1,1)(2,1,0)[52] intercept : AIC=2234.455, Time=26.66 sec

```

Best model: ARIMA(1,1,1)(2,1,0)[52]

Total fit time: 438.890 seconds

ARIMA(1, 1, 1)x(2, 1, 0, 52)52 - AIC:2231.657378222504

performing function on R06

Performing stepwise search to minimize aic

```

ARIMA(0,1,0)(1,1,1)[52] : AIC=inf, Time=4.76 sec
ARIMA(0,1,0)(0,1,0)[52] : AIC=1892.735, Time=0.34 sec
ARIMA(1,1,0)(1,1,0)[52] : AIC=1778.026, Time=2.60 sec
ARIMA(0,1,1)(0,1,1)[52] : AIC=inf, Time=11.82 sec
ARIMA(1,1,0)(0,1,0)[52] : AIC=1851.699, Time=0.30 sec
ARIMA(1,1,0)(2,1,0)[52] : AIC=1762.733, Time=6.98 sec
ARIMA(1,1,0)(2,1,1)[52] : AIC=inf, Time=17.50 sec
ARIMA(1,1,0)(1,1,1)[52] : AIC=inf, Time=5.67 sec
ARIMA(0,1,0)(2,1,0)[52] : AIC=1810.884, Time=4.69 sec
ARIMA(2,1,0)(2,1,0)[52] : AIC=inf, Time=7.49 sec

```

```

ARIMA(1,1,1) (2,1,0) [52] : AIC=inf, Time=24.38 sec
ARIMA(0,1,1) (2,1,0) [52] : AIC=1744.421, Time=10.40 sec
ARIMA(0,1,1) (1,1,0) [52] : AIC=1766.381, Time=2.50 sec
ARIMA(0,1,1) (2,1,1) [52] : AIC=1744.802, Time=13.17 sec
ARIMA(0,1,1) (1,1,1) [52] : AIC=inf, Time=6.66 sec
ARIMA(0,1,2) (2,1,0) [52] : AIC=1732.915, Time=12.35 sec
ARIMA(0,1,2) (1,1,0) [52] : AIC=1752.142, Time=6.04 sec
ARIMA(0,1,2) (2,1,1) [52] : AIC=1734.416, Time=21.96 sec
ARIMA(0,1,2) (1,1,1) [52] : AIC=inf, Time=15.52 sec
ARIMA(1,1,2) (2,1,0) [52] : AIC=inf, Time=41.51 sec
ARIMA(0,1,2) (2,1,0) [52] intercept : AIC=1734.800, Time=17.45 sec

```

Best model: ARIMA(0,1,2) (2,1,0) [52]

Total fit time: 234.121 seconds

ARIMA(0, 1, 2)x(2, 1, 0, 52)52 - AIC:1732.914810049424

performing function on N02BA\_diff

Performing stepwise search to minimize aic

```

ARIMA(0,1,0) (1,1,1) [52] : AIC=inf, Time=5.40 sec
ARIMA(0,1,0) (0,1,0) [52] : AIC=2126.506, Time=0.42 sec
ARIMA(1,1,0) (1,1,0) [52] : AIC=1954.403, Time=2.19 sec
ARIMA(0,1,1) (0,1,1) [52] : AIC=inf, Time=8.44 sec
ARIMA(1,1,0) (0,1,0) [52] : AIC=1995.436, Time=0.32 sec
ARIMA(1,1,0) (2,1,0) [52] : AIC=1950.895, Time=6.05 sec
ARIMA(1,1,0) (2,1,1) [52] : AIC=inf, Time=36.47 sec
ARIMA(1,1,0) (1,1,1) [52] : AIC=inf, Time=7.06 sec
ARIMA(0,1,0) (2,1,0) [52] : AIC=2088.021, Time=5.01 sec
ARIMA(2,1,0) (2,1,0) [52] : AIC=1900.102, Time=7.37 sec
ARIMA(2,1,0) (1,1,0) [52] : AIC=1903.043, Time=3.22 sec
ARIMA(2,1,0) (2,1,1) [52] : AIC=inf, Time=44.20 sec
ARIMA(2,1,0) (1,1,1) [52] : AIC=inf, Time=12.13 sec
ARIMA(2,1,1) (2,1,0) [52] : AIC=inf, Time=34.69 sec
ARIMA(1,1,1) (2,1,0) [52] : AIC=inf, Time=30.77 sec
ARIMA(2,1,0) (2,1,0) [52] intercept : AIC=1902.084, Time=10.28 sec

```

Best model: ARIMA(2,1,0) (2,1,0) [52]

Total fit time: 214.023 seconds

ARIMA(2, 1, 0)x(2, 1, 0, 52)52 - AIC:1900.1017888544345

In [31]: `print('done')`

done

In [66]: `#FORECASTING`

`#We now have a list of models that are trained for each drug in our data set.`

`#Let us predict the sales of our drugs for n periods using the model`

`#We plot the original data set in blue`

`#We plot the forecasted data set in orange`

`import warnings`

`warnings.filterwarnings('ignore')`

`n = 100`

`for i in range (0,len(models)):`

`print(index)`

`print(df.columns[i])`

`print(models[i])`

`forecasted_model = models[i].predict(n_periods= n)`

`fig, ax = plt.subplots(figsize=(14, 6))`

`plt.plot(df.index, df[df.columns[i]], label = "True Data")`

`ax.set_title(f"Weekly Sales Over Time for {df.columns[i]}")`

`ax.set_xlabel('Date')`

`ax.set_ylabel('Sales')`

`forecasted_dates= pd.date_range(start=df.index[-1], periods=n+1, closed='right')`

`plt.plot(forecasted_dates, forecasted_model, label = "Forecasted Data")`

`ax.grid(True)`

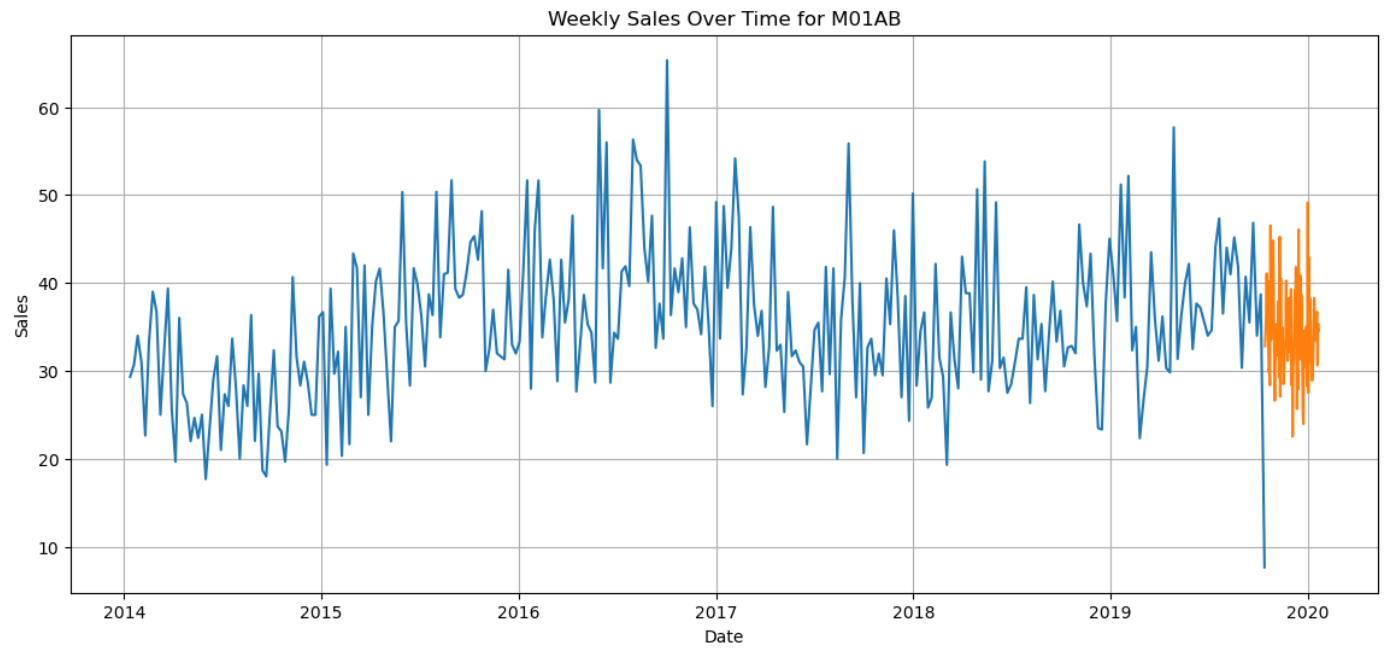
`plt.show()`

*#The plots below forecast the sales of each drug for 100 periods using our SARIMA models*

0

M01AB

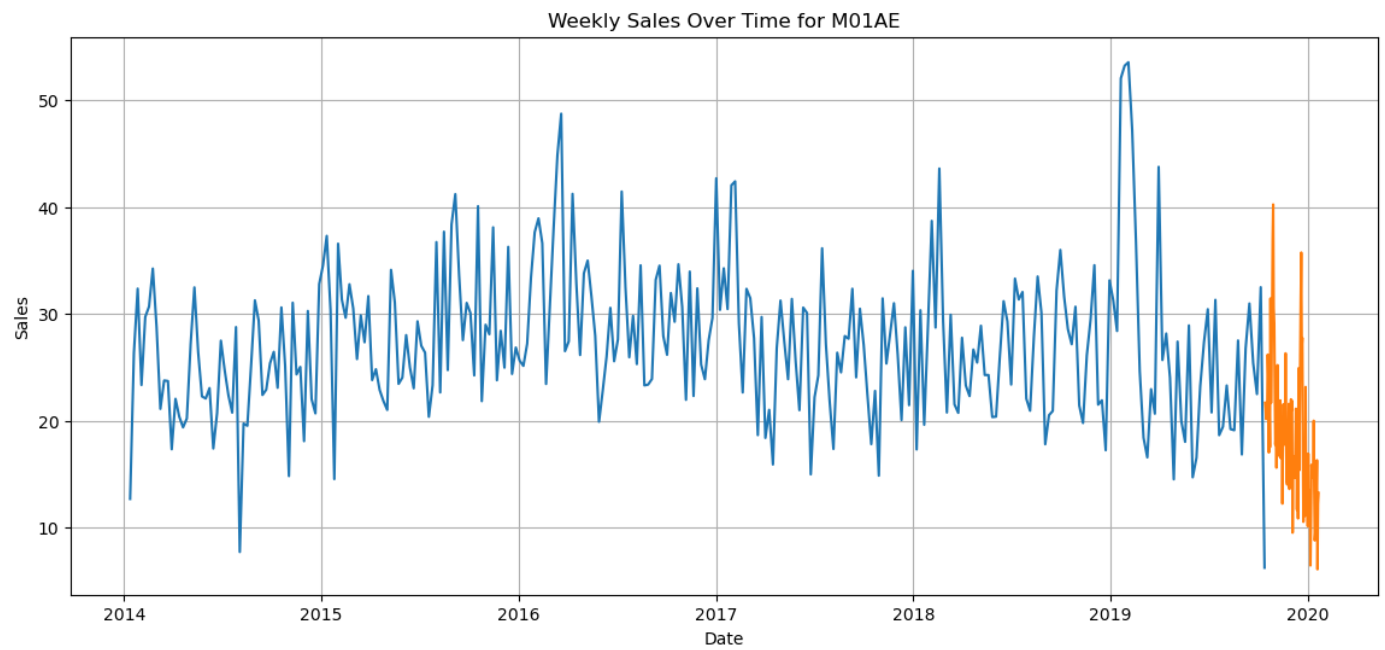
ARIMA(0,1,1)(2,1,0)[52]



0

M01AE

ARIMA(1,1,1)(2,1,0)[52] intercept

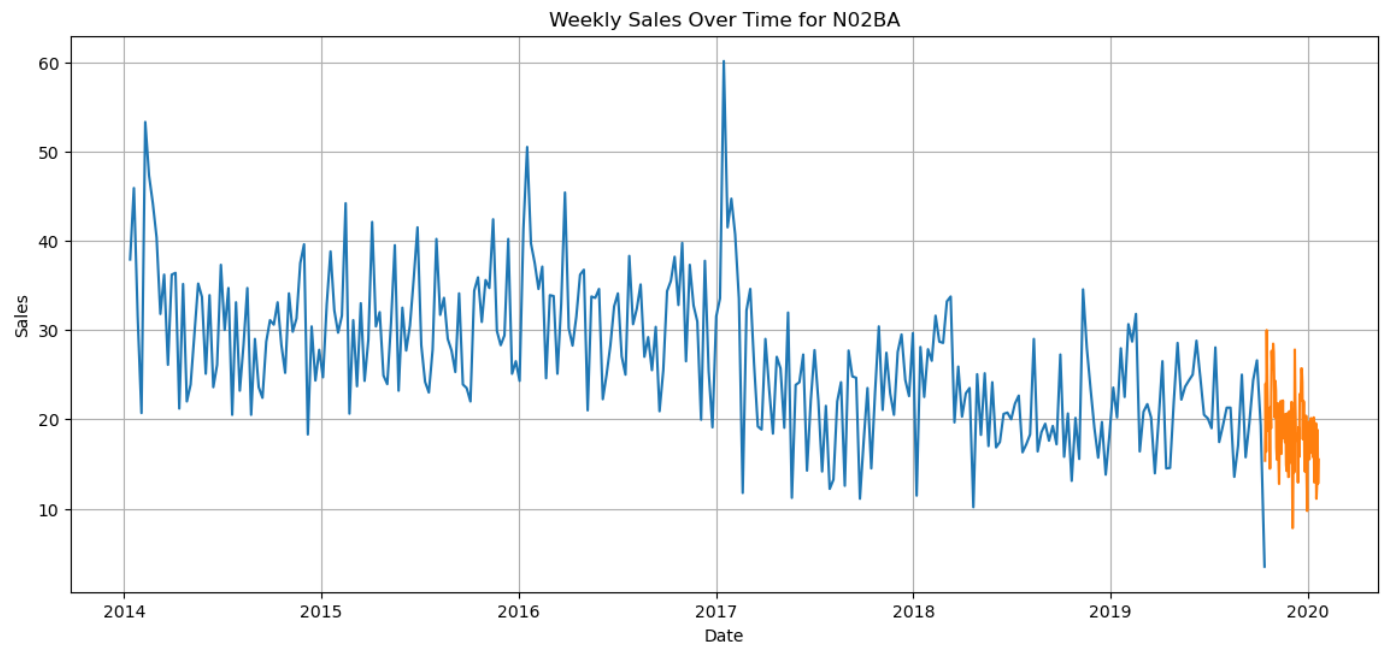


0

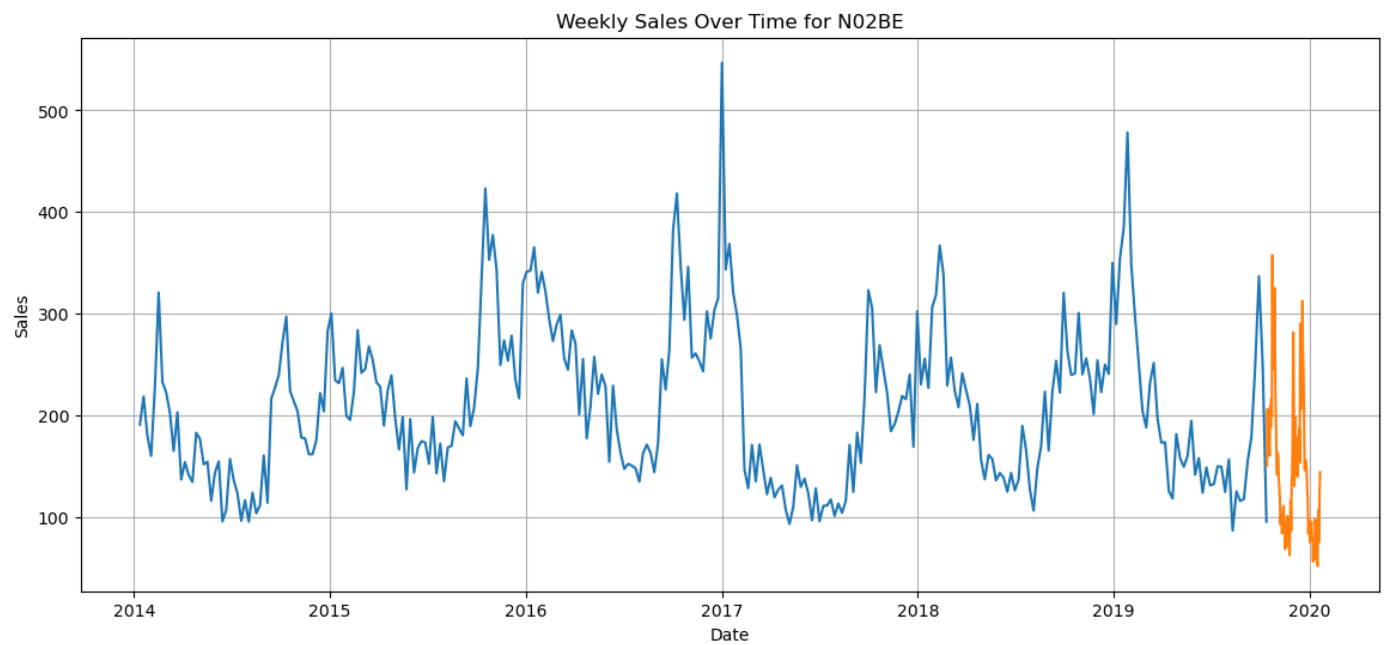
N02BA

ARIMA(1,1,1)(2,1,0)[52]

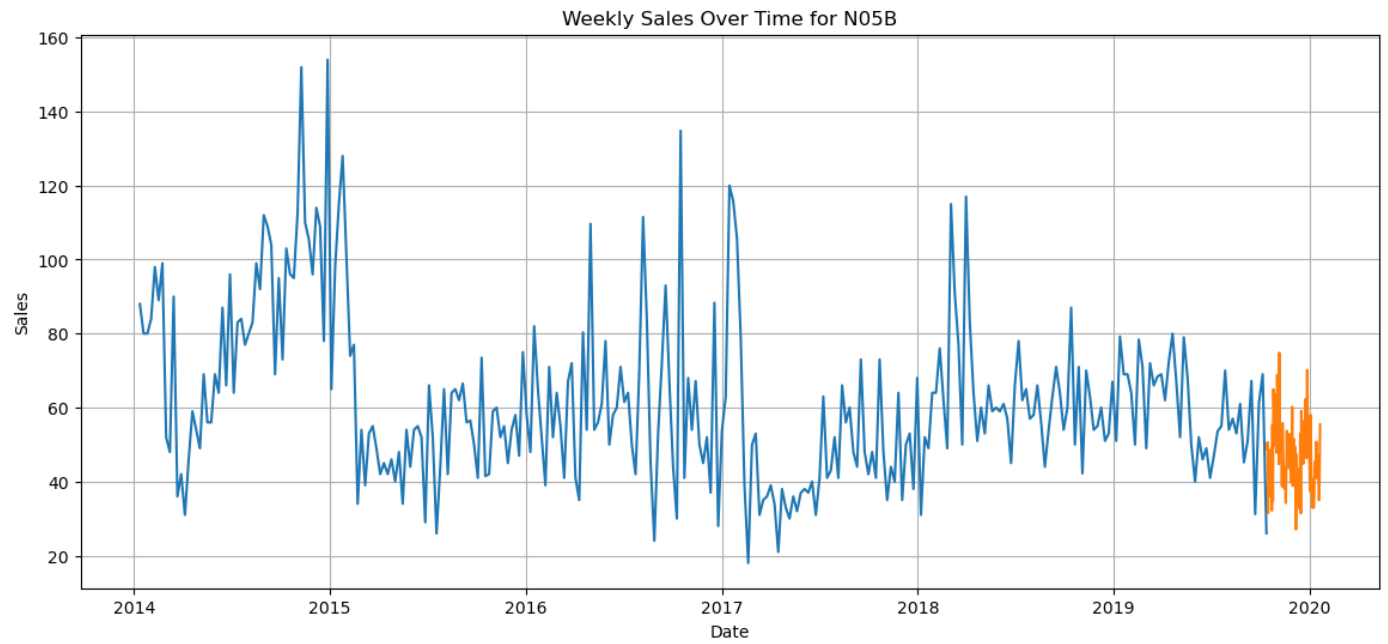




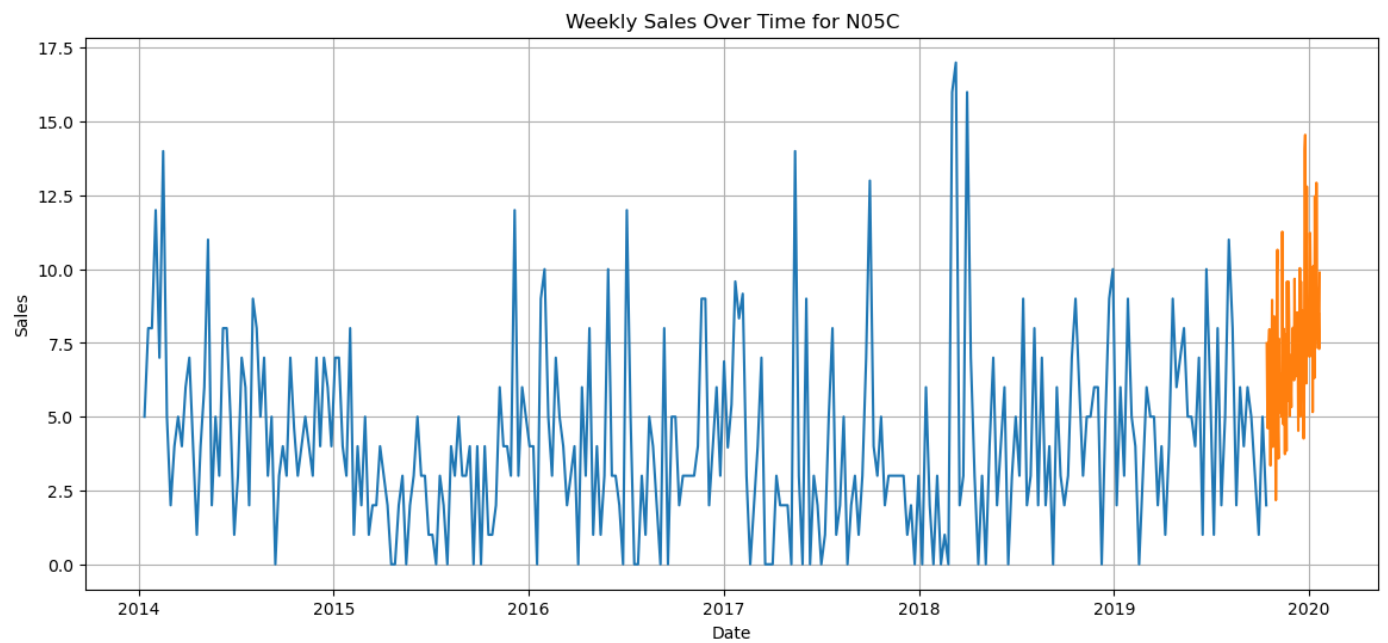
0  
N02BE  
ARIMA(2,1,1)(2,1,0)[52]



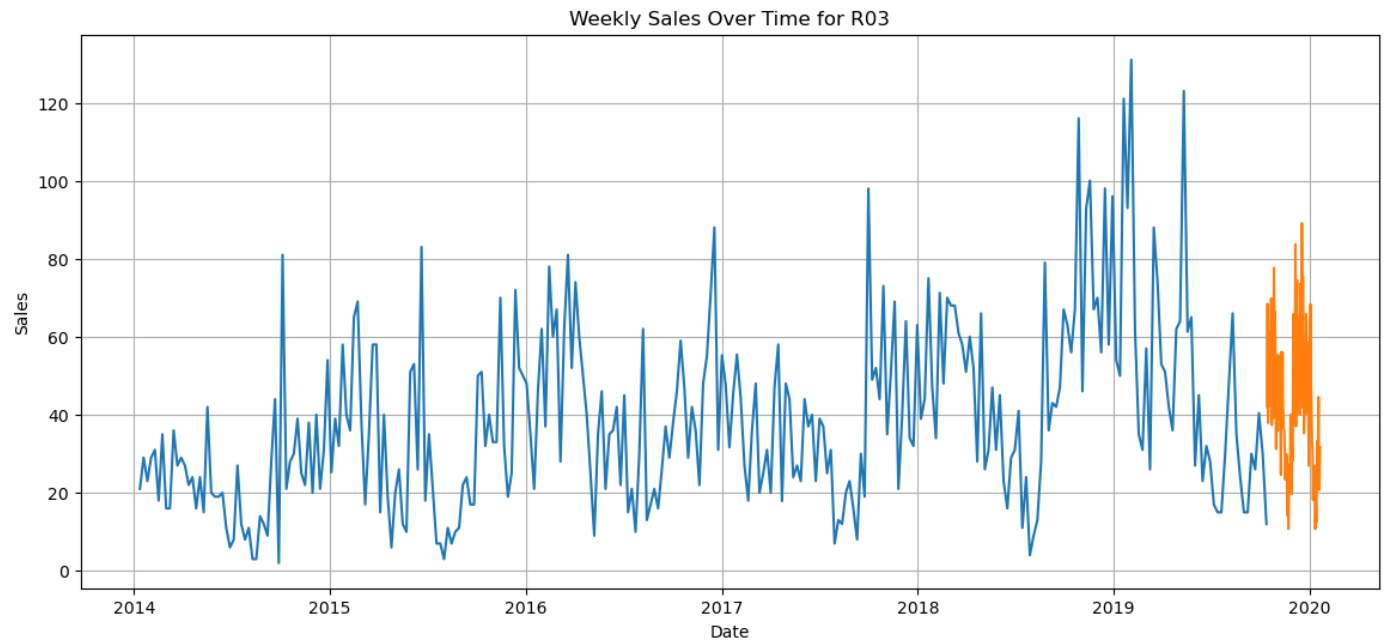
0  
N05B  
ARIMA(0,1,1)(2,1,0)[52]



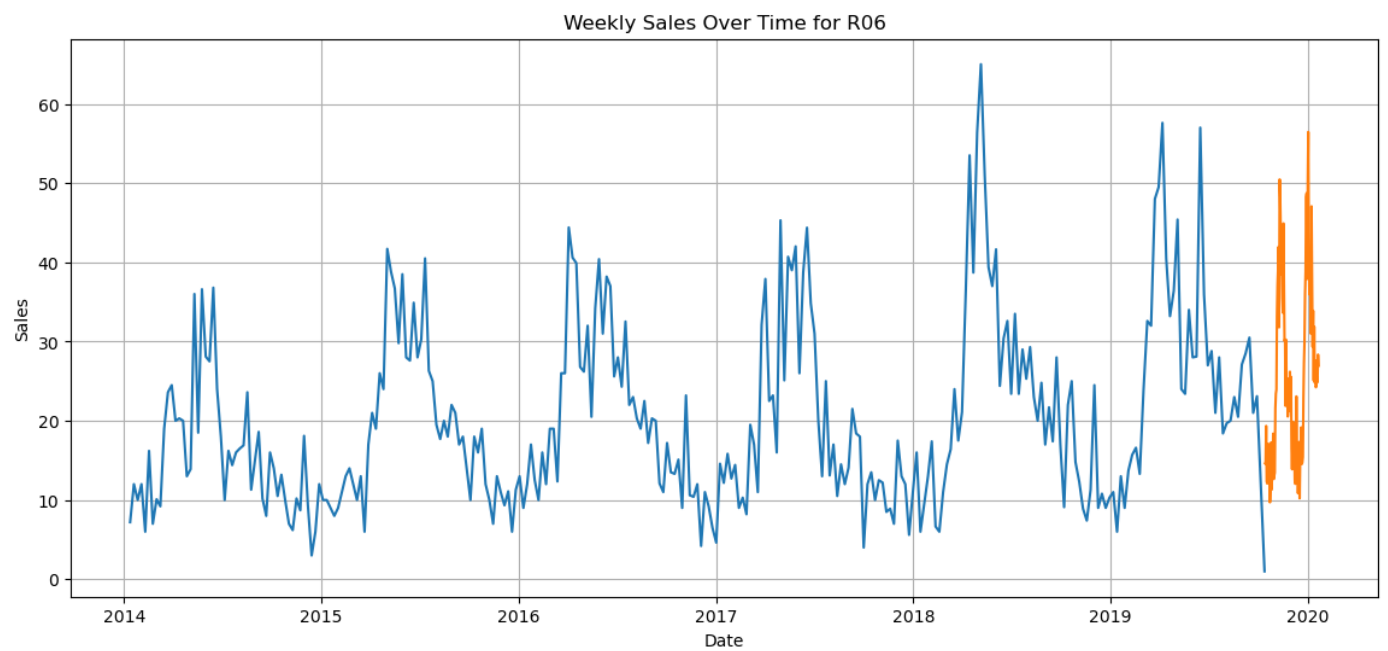
0  
N05C  
ARIMA(0,1,1) (2,1,0) [52] intercept



0  
R03  
ARIMA(1,1,1) (2,1,0) [52]

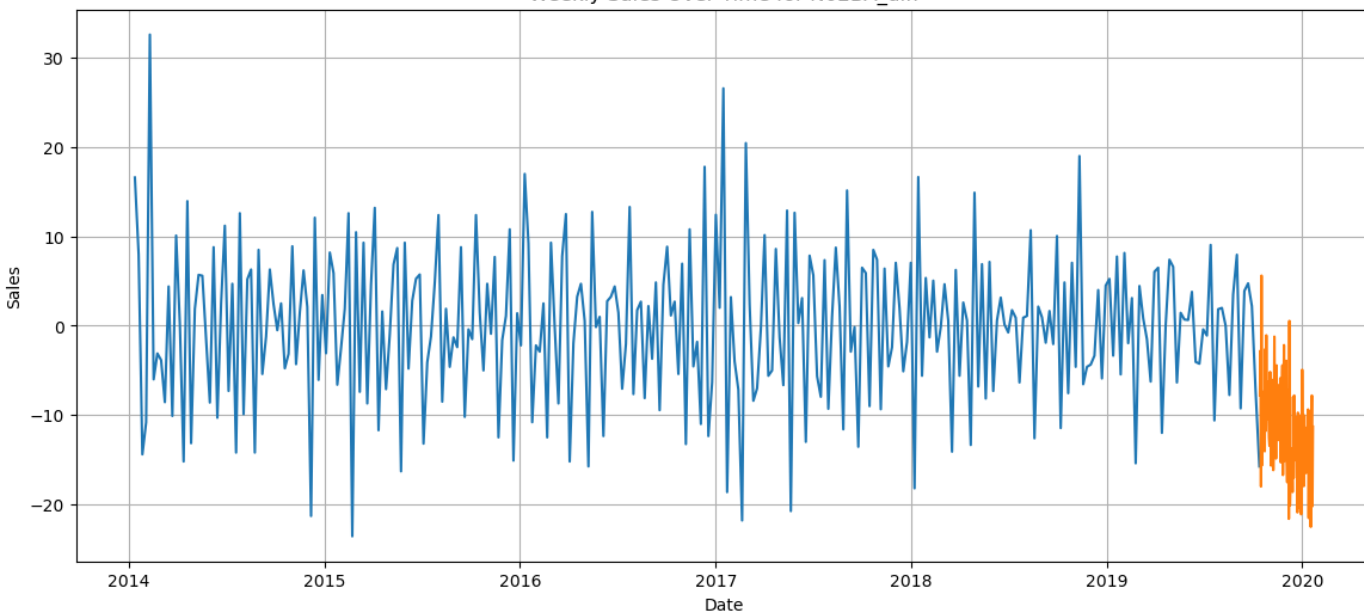


0  
R06  
ARIMA(0,1,2) (2,1,0) [52]



0  
N02BA\_diff  
ARIMA(2,1,0) (2,1,0) [52]

Weekly Sales Over Time for N02BA\_diff



```
In [72]: #MODEL ANALYSIS
#We now wish to evaluate the performance of our models mathematically

#To do this, we will use training train the model on a section of our data set
#We will then compare it to the remaining data in our data set and see how accurate it i

#To accomplish this, we will use metrics such as mean squared error, mean absolute error
from sklearn.metrics import mean_squared_error, mean_absolute_error

#This section splits our data set
train_length = int(len(df)*.75)
forecast_length = int(len(df)-train_length)

training_data= df.iloc[:train_length]
forecast_data = df.iloc[train_length:]

#creating models based on data
evaluation_models = {}
period = 52
for index in range(0,len(df.columns)):
    # the auto_arima function generates the best SARIMA model to the time series
    item = df.columns[index]
    print('performing function on {}'.format(item))
    training_data= df[item][:train_length]
    test_data = df[item][train_length:]
    model = pm.auto_arima(training_data,
                          seasonal = True,
                          m = period,
                          d = 1,
                          D = 1,
                          start_p=0,
                          start_q = 0,
                          max_p=2,
                          max_q=2,
                          trace = True,
                          error_action = 'ignore',
                          suppress_warnings = True,
                          stepwise = True)
    evaluation_models[item] = model

forecast_data = model.predict(n_periods = len(test_data))
mse = mean_squared_error(test_data, forecast_data)
mae = mean_absolute_error(test_data, forecast_data)
rmse = np.sqrt(mse)
```

```

print('Evaluations for {}'.format(item))
print('MSE:{}'.format(mse))
print('MAE:{}'.format(mae))
print('RMSE:{}'.format(rmse))

print('ARIMA{}x{}{} - AIC:{}'.format(model.order, model.seasonal_order, period, mode

```

performing function on M01AB

Performing stepwise search to minimize aic

```

ARIMA(0,1,0) (1,1,1) [52]      : AIC=inf, Time=7.12 sec
ARIMA(0,1,0) (0,1,0) [52]      : AIC=1420.398, Time=0.30 sec
ARIMA(1,1,0) (1,1,0) [52]      : AIC=1341.654, Time=2.15 sec
ARIMA(0,1,1) (0,1,1) [52]      : AIC=inf, Time=7.90 sec
ARIMA(1,1,0) (0,1,0) [52]      : AIC=1374.525, Time=0.29 sec
ARIMA(1,1,0) (2,1,0) [52]      : AIC=1335.218, Time=5.12 sec
ARIMA(1,1,0) (2,1,1) [52]      : AIC=inf, Time=22.76 sec
ARIMA(1,1,0) (1,1,1) [52]      : AIC=inf, Time=12.61 sec
ARIMA(0,1,0) (2,1,0) [52]      : AIC=inf, Time=3.72 sec
ARIMA(2,1,0) (2,1,0) [52]      : AIC=1319.243, Time=7.02 sec
ARIMA(2,1,0) (1,1,0) [52]      : AIC=1326.031, Time=2.86 sec
ARIMA(2,1,0) (2,1,1) [52]      : AIC=inf, Time=31.06 sec
ARIMA(2,1,0) (1,1,1) [52]      : AIC=inf, Time=17.92 sec
ARIMA(2,1,1) (2,1,0) [52]      : AIC=1283.972, Time=14.50 sec
ARIMA(2,1,1) (1,1,0) [52]      : AIC=1296.593, Time=6.52 sec
ARIMA(2,1,1) (2,1,1) [52]      : AIC=inf, Time=36.94 sec
ARIMA(2,1,1) (1,1,1) [52]      : AIC=inf, Time=21.51 sec
ARIMA(1,1,1) (2,1,0) [52]      : AIC=1282.311, Time=9.58 sec
ARIMA(1,1,1) (1,1,0) [52]      : AIC=1294.597, Time=3.39 sec
ARIMA(1,1,1) (2,1,1) [52]      : AIC=inf, Time=34.46 sec
ARIMA(1,1,1) (1,1,1) [52]      : AIC=inf, Time=14.54 sec
ARIMA(0,1,1) (2,1,0) [52]      : AIC=1280.817, Time=6.30 sec
ARIMA(0,1,1) (1,1,0) [52]      : AIC=1293.151, Time=2.70 sec
ARIMA(0,1,1) (2,1,1) [52]      : AIC=inf, Time=28.22 sec
ARIMA(0,1,1) (1,1,1) [52]      : AIC=inf, Time=12.32 sec
ARIMA(0,1,2) (2,1,0) [52]      : AIC=1282.252, Time=9.53 sec
ARIMA(1,1,2) (2,1,0) [52]      : AIC=1284.362, Time=16.74 sec
ARIMA(0,1,1) (2,1,0) [52] intercept : AIC=1281.946, Time=10.32 sec

```

Best model: ARIMA(0,1,1) (2,1,0) [52]

Total fit time: 348.426 seconds

Evaluations for M01AB

MSE:80.01795103772385

MAE:6.980005972826855

RMSE:8.945275347227934

ARIMA(0, 1, 1)x(2, 1, 0, 52)52 - AIC:1280.8165056108326

performing function on M01AE

Performing stepwise search to minimize aic

```

ARIMA(0,1,0) (1,1,1) [52]      : AIC=inf, Time=8.15 sec
ARIMA(0,1,0) (0,1,0) [52]      : AIC=1317.827, Time=0.33 sec
ARIMA(1,1,0) (1,1,0) [52]      : AIC=1254.036, Time=2.00 sec
ARIMA(0,1,1) (0,1,1) [52]      : AIC=inf, Time=6.49 sec
ARIMA(1,1,0) (0,1,0) [52]      : AIC=1279.517, Time=0.28 sec
ARIMA(1,1,0) (2,1,0) [52]      : AIC=inf, Time=4.80 sec
ARIMA(1,1,0) (1,1,1) [52]      : AIC=inf, Time=13.00 sec
ARIMA(1,1,0) (0,1,1) [52]      : AIC=inf, Time=7.75 sec
ARIMA(1,1,0) (2,1,1) [52]      : AIC=inf, Time=26.53 sec
ARIMA(0,1,0) (1,1,0) [52]      : AIC=1296.822, Time=1.52 sec
ARIMA(2,1,0) (1,1,0) [52]      : AIC=1242.440, Time=2.30 sec
ARIMA(2,1,0) (0,1,0) [52]      : AIC=1266.522, Time=0.46 sec
ARIMA(2,1,0) (2,1,0) [52]      : AIC=inf, Time=5.98 sec
ARIMA(2,1,0) (1,1,1) [52]      : AIC=inf, Time=13.99 sec
ARIMA(2,1,0) (0,1,1) [52]      : AIC=inf, Time=11.61 sec
ARIMA(2,1,0) (2,1,1) [52]      : AIC=inf, Time=32.62 sec
ARIMA(2,1,1) (1,1,0) [52]      : AIC=1204.096, Time=6.53 sec
ARIMA(2,1,1) (0,1,0) [52]      : AIC=1234.115, Time=1.37 sec
ARIMA(2,1,1) (2,1,0) [52]      : AIC=1191.261, Time=12.95 sec

```

```

ARIMA(2,1,1)(2,1,1)[52] : AIC=inf, Time=37.35 sec
ARIMA(2,1,1)(1,1,1)[52] : AIC=inf, Time=21.20 sec
ARIMA(1,1,1)(2,1,0)[52] : AIC=1189.310, Time=13.61 sec
ARIMA(1,1,1)(1,1,0)[52] : AIC=1202.097, Time=5.38 sec
ARIMA(1,1,1)(2,1,1)[52] : AIC=inf, Time=52.01 sec
ARIMA(1,1,1)(1,1,1)[52] : AIC=inf, Time=20.28 sec
ARIMA(0,1,1)(2,1,0)[52] : AIC=1187.357, Time=11.26 sec
ARIMA(0,1,1)(1,1,0)[52] : AIC=1200.226, Time=3.04 sec
ARIMA(0,1,1)(2,1,1)[52] : AIC=inf, Time=46.54 sec
ARIMA(0,1,1)(1,1,1)[52] : AIC=inf, Time=14.82 sec
ARIMA(0,1,0)(2,1,0)[52] : AIC=inf, Time=6.21 sec
ARIMA(0,1,2)(2,1,0)[52] : AIC=1189.308, Time=14.36 sec
ARIMA(1,1,2)(2,1,0)[52] : AIC=1191.303, Time=31.34 sec
ARIMA(0,1,1)(2,1,0)[52] intercept : AIC=inf, Time=26.59 sec

```

Best model: ARIMA(0,1,1)(2,1,0)[52]

Total fit time: 462.690 seconds

Evaluations for M01AE

MSE:72.13182618944226

MAE:6.6593943061584415

RMSE:8.493045754583115

ARIMA(0, 1, 1)x(2, 1, 0, 52)52 - AIC:1187.3571533126753

performing function on N02BA

Performing stepwise search to minimize aic

```

ARIMA(0,1,0)(1,1,1)[52] : AIC=1277.618, Time=4.03 sec
ARIMA(0,1,0)(0,1,0)[52] : AIC=1303.019, Time=2.66 sec
ARIMA(1,1,0)(1,1,0)[52] : AIC=1246.275, Time=2.38 sec
ARIMA(0,1,1)(0,1,1)[52] : AIC=inf, Time=4.13 sec
ARIMA(1,1,0)(0,1,0)[52] : AIC=1273.390, Time=0.34 sec
ARIMA(1,1,0)(2,1,0)[52] : AIC=1245.695, Time=7.75 sec
ARIMA(1,1,0)(2,1,1)[52] : AIC=inf, Time=45.33 sec
ARIMA(1,1,0)(1,1,1)[52] : AIC=1244.755, Time=6.44 sec
ARIMA(1,1,0)(0,1,1)[52] : AIC=inf, Time=4.59 sec
ARIMA(1,1,0)(1,1,2)[52] : AIC=inf, Time=44.19 sec
ARIMA(1,1,0)(0,1,2)[52] : AIC=1244.802, Time=11.62 sec
ARIMA(1,1,0)(2,1,2)[52] : AIC=inf, Time=29.12 sec
ARIMA(2,1,0)(1,1,1)[52] : AIC=1234.532, Time=8.24 sec
ARIMA(2,1,0)(0,1,1)[52] : AIC=1232.563, Time=5.64 sec
ARIMA(2,1,0)(0,1,0)[52] : AIC=1257.706, Time=0.58 sec
ARIMA(2,1,0)(0,1,2)[52] : AIC=1234.543, Time=14.54 sec
ARIMA(2,1,0)(1,1,0)[52] : AIC=1235.485, Time=3.44 sec
ARIMA(2,1,0)(1,1,2)[52] : AIC=inf, Time=51.53 sec
ARIMA(2,1,1)(0,1,1)[52] : AIC=inf, Time=9.23 sec
ARIMA(1,1,1)(0,1,1)[52] : AIC=inf, Time=5.76 sec
ARIMA(2,1,0)(0,1,1)[52] intercept : AIC=1234.552, Time=6.97 sec

```

Best model: ARIMA(2,1,0)(0,1,1)[52]

Total fit time: 268.550 seconds

Evaluations for N02BA

MSE:50.480216313060005

MAE:5.460127735685665

RMSE:7.1049430900648325

ARIMA(2, 1, 0)x(0, 1, 1, 52)52 - AIC:1232.5628572450773

performing function on N02BE

Performing stepwise search to minimize aic

```

ARIMA(0,1,0)(1,1,1)[52] : AIC=inf, Time=11.64 sec
ARIMA(0,1,0)(0,1,0)[52] : AIC=1895.482, Time=3.03 sec
ARIMA(1,1,0)(1,1,0)[52] : AIC=1850.567, Time=3.15 sec
ARIMA(0,1,1)(0,1,1)[52] : AIC=inf, Time=12.08 sec
ARIMA(1,1,0)(0,1,0)[52] : AIC=1877.545, Time=0.44 sec
ARIMA(1,1,0)(2,1,0)[52] : AIC=inf, Time=10.93 sec
ARIMA(1,1,0)(1,1,1)[52] : AIC=inf, Time=20.28 sec
ARIMA(1,1,0)(0,1,1)[52] : AIC=inf, Time=10.98 sec
ARIMA(1,1,0)(2,1,1)[52] : AIC=inf, Time=46.12 sec
ARIMA(0,1,0)(1,1,0)[52] : AIC=1871.113, Time=2.51 sec
ARIMA(2,1,0)(1,1,0)[52] : AIC=1847.735, Time=4.56 sec

```

```

ARIMA(2,1,0)(0,1,0)[52] : AIC=1874.803, Time=0.63 sec
ARIMA(2,1,0)(2,1,0)[52] : AIC=inf, Time=13.86 sec
ARIMA(2,1,0)(1,1,1)[52] : AIC=inf, Time=24.18 sec
ARIMA(2,1,0)(0,1,1)[52] : AIC=inf, Time=16.47 sec
ARIMA(2,1,0)(2,1,1)[52] : AIC=inf, Time=52.37 sec
ARIMA(2,1,1)(1,1,0)[52] : AIC=1841.812, Time=11.64 sec
ARIMA(2,1,1)(0,1,0)[52] : AIC=1873.288, Time=1.16 sec
ARIMA(2,1,1)(2,1,0)[52] : AIC=1831.240, Time=28.38 sec
ARIMA(2,1,1)(2,1,1)[52] : AIC=inf, Time=63.00 sec
ARIMA(2,1,1)(1,1,1)[52] : AIC=inf, Time=28.89 sec
ARIMA(1,1,1)(2,1,0)[52] : AIC=1832.705, Time=39.58 sec
ARIMA(2,1,2)(2,1,0)[52] : AIC=1833.225, Time=67.79 sec
ARIMA(1,1,2)(2,1,0)[52] : AIC=1831.642, Time=33.65 sec
ARIMA(2,1,1)(2,1,0)[52] intercept : AIC=inf, Time=57.19 sec

```

Best model: ARIMA(2,1,1)(2,1,0)[52]

Total fit time: 564.533 seconds

Evaluations for N02BE

MSE:2491.788679398072

MAE:32.95969938875864

RMSE:49.917819257235905

ARIMA(2, 1, 1)x(2, 1, 0, 52)52 - AIC:1831.2400178672965

performing function on N05B

Performing stepwise search to minimize aic

```

ARIMA(0,1,0)(1,1,1)[52] : AIC=1690.906, Time=7.31 sec
ARIMA(0,1,0)(0,1,0)[52] : AIC=1718.367, Time=3.18 sec
ARIMA(1,1,0)(1,1,0)[52] : AIC=1642.684, Time=3.31 sec
ARIMA(0,1,1)(0,1,1)[52] : AIC=inf, Time=5.66 sec
ARIMA(1,1,0)(0,1,0)[52] : AIC=1673.426, Time=0.39 sec
ARIMA(1,1,0)(2,1,0)[52] : AIC=1643.659, Time=10.35 sec
ARIMA(1,1,0)(1,1,1)[52] : AIC=inf, Time=9.16 sec
ARIMA(1,1,0)(0,1,1)[52] : AIC=inf, Time=5.64 sec
ARIMA(1,1,0)(2,1,1)[52] : AIC=inf, Time=43.75 sec
ARIMA(0,1,0)(1,1,0)[52] : AIC=1691.257, Time=2.40 sec
ARIMA(2,1,0)(1,1,0)[52] : AIC=1636.743, Time=4.42 sec
ARIMA(2,1,0)(0,1,0)[52] : AIC=1668.662, Time=0.53 sec
ARIMA(2,1,0)(2,1,0)[52] : AIC=1637.542, Time=13.18 sec
ARIMA(2,1,0)(1,1,1)[52] : AIC=inf, Time=15.96 sec
ARIMA(2,1,0)(0,1,1)[52] : AIC=inf, Time=6.99 sec
ARIMA(2,1,0)(2,1,1)[52] : AIC=inf, Time=55.24 sec
ARIMA(2,1,1)(1,1,0)[52] : AIC=inf, Time=15.71 sec
ARIMA(1,1,1)(1,1,0)[52] : AIC=1615.655, Time=6.33 sec
ARIMA(1,1,1)(0,1,0)[52] : AIC=1651.983, Time=0.66 sec
ARIMA(1,1,1)(2,1,0)[52] : AIC=1617.299, Time=16.09 sec
ARIMA(1,1,1)(1,1,1)[52] : AIC=1616.685, Time=11.10 sec
ARIMA(1,1,1)(0,1,1)[52] : AIC=inf, Time=7.13 sec
ARIMA(1,1,1)(2,1,1)[52] : AIC=inf, Time=54.82 sec
ARIMA(0,1,1)(1,1,0)[52] : AIC=1613.664, Time=4.15 sec
ARIMA(0,1,1)(0,1,0)[52] : AIC=1650.170, Time=0.46 sec
ARIMA(0,1,1)(2,1,0)[52] : AIC=1615.312, Time=12.22 sec
ARIMA(0,1,1)(1,1,1)[52] : AIC=1614.721, Time=8.24 sec
ARIMA(0,1,1)(2,1,1)[52] : AIC=inf, Time=42.55 sec
ARIMA(0,1,2)(1,1,0)[52] : AIC=1615.656, Time=5.03 sec
ARIMA(1,1,2)(1,1,0)[52] : AIC=1616.930, Time=7.71 sec
ARIMA(0,1,1)(1,1,0)[52] intercept : AIC=1615.546, Time=5.37 sec

```

Best model: ARIMA(0,1,1)(1,1,0)[52]

Total fit time: 385.090 seconds

Evaluations for N05B

MSE:421.7505865371313

MAE:16.153200300996026

RMSE:20.536567058228872

ARIMA(0, 1, 1)x(1, 1, 0, 52)52 - AIC:1613.6640392297375

performing function on N05C

Performing stepwise search to minimize aic

```

ARIMA(0,1,0)(1,1,1)[52] : AIC=inf, Time=5.20 sec

```

```

ARIMA(0,1,0)(0,1,0)[52] : AIC=1080.641, Time=2.54 sec
ARIMA(1,1,0)(1,1,0)[52] : AIC=1035.914, Time=2.41 sec
ARIMA(0,1,1)(0,1,1)[52] : AIC=978.290, Time=3.87 sec
ARIMA(0,1,1)(0,1,0)[52] : AIC=990.384, Time=0.49 sec
ARIMA(0,1,1)(1,1,1)[52] : AIC=980.117, Time=6.52 sec
ARIMA(0,1,1)(0,1,2)[52] : AIC=980.093, Time=11.06 sec
ARIMA(0,1,1)(1,1,0)[52] : AIC=981.251, Time=3.33 sec
ARIMA(0,1,1)(1,1,2)[52] : AIC=inf, Time=47.69 sec
ARIMA(0,1,0)(0,1,1)[52] : AIC=1066.964, Time=29.15 sec
ARIMA(1,1,1)(0,1,1)[52] : AIC=979.082, Time=5.78 sec
ARIMA(0,1,2)(0,1,1)[52] : AIC=978.965, Time=5.80 sec
ARIMA(1,1,0)(0,1,1)[52] : AIC=1033.206, Time=3.69 sec
ARIMA(1,1,2)(0,1,1)[52] : AIC=979.549, Time=9.09 sec
ARIMA(0,1,1)(0,1,1)[52] intercept : AIC=978.800, Time=6.73 sec

```

Best model: ARIMA(0,1,1)(0,1,1)[52]

Total fit time: 143.362 seconds

Evaluations for N05C

MSE:13.107968793187997

MAE:2.9164518149127936

RMSE:3.620492893680085

ARIMA(0, 1, 1)x(0, 1, 1, 52)52 - AIC:978.2896054872704

performing function on R03

Performing stepwise search to minimize aic

```

ARIMA(0,1,0)(1,1,1)[52] : AIC=inf, Time=9.56 sec
ARIMA(0,1,0)(0,1,0)[52] : AIC=1651.467, Time=3.12 sec
ARIMA(1,1,0)(1,1,0)[52] : AIC=1573.697, Time=2.94 sec
ARIMA(0,1,1)(0,1,1)[52] : AIC=inf, Time=12.01 sec
ARIMA(1,1,0)(0,1,0)[52] : AIC=1614.280, Time=0.39 sec
ARIMA(1,1,0)(2,1,0)[52] : AIC=inf, Time=10.05 sec
ARIMA(1,1,0)(1,1,1)[52] : AIC=inf, Time=17.24 sec
ARIMA(1,1,0)(0,1,1)[52] : AIC=inf, Time=11.03 sec
ARIMA(1,1,0)(2,1,1)[52] : AIC=1552.359, Time=19.78 sec
ARIMA(1,1,0)(2,1,2)[52] : AIC=1554.359, Time=34.81 sec
ARIMA(1,1,0)(1,1,2)[52] : AIC=inf, Time=45.50 sec
ARIMA(0,1,0)(2,1,1)[52] : AIC=inf, Time=42.35 sec
ARIMA(2,1,0)(2,1,1)[52] : AIC=1523.148, Time=26.03 sec
ARIMA(2,1,0)(1,1,1)[52] : AIC=inf, Time=21.61 sec
ARIMA(2,1,0)(2,1,0)[52] : AIC=inf, Time=14.34 sec
ARIMA(2,1,0)(2,1,2)[52] : AIC=inf, Time=44.19 sec
ARIMA(2,1,0)(1,1,0)[52] : AIC=1552.107, Time=4.66 sec
ARIMA(2,1,0)(1,1,2)[52] : AIC=inf, Time=52.37 sec
ARIMA(2,1,1)(2,1,1)[52] : AIC=1502.907, Time=73.53 sec
ARIMA(2,1,1)(1,1,1)[52] : AIC=inf, Time=30.19 sec
ARIMA(2,1,1)(2,1,0)[52] : AIC=1501.346, Time=40.86 sec
ARIMA(2,1,1)(1,1,0)[52] : AIC=1529.253, Time=13.64 sec
ARIMA(1,1,1)(2,1,0)[52] : AIC=1499.348, Time=35.87 sec
ARIMA(1,1,1)(1,1,0)[52] : AIC=1527.553, Time=8.12 sec
ARIMA(1,1,1)(2,1,1)[52] : AIC=1500.907, Time=62.80 sec
ARIMA(1,1,1)(1,1,1)[52] : AIC=inf, Time=17.97 sec
ARIMA(0,1,1)(2,1,0)[52] : AIC=1498.733, Time=18.82 sec
ARIMA(0,1,1)(1,1,0)[52] : AIC=1527.082, Time=5.08 sec
ARIMA(0,1,1)(2,1,1)[52] : AIC=1500.461, Time=42.70 sec
ARIMA(0,1,1)(1,1,1)[52] : AIC=inf, Time=12.89 sec
ARIMA(0,1,0)(2,1,0)[52] : AIC=inf, Time=7.33 sec
ARIMA(0,1,2)(2,1,0)[52] : AIC=1499.325, Time=29.92 sec
ARIMA(1,1,2)(2,1,0)[52] : AIC=1499.536, Time=50.43 sec
ARIMA(0,1,1)(2,1,0)[52] intercept : AIC=1500.710, Time=28.57 sec

```

Best model: ARIMA(0,1,1)(2,1,0)[52]

Total fit time: 850.776 seconds

Evaluations for R03

MSE:758.5625897680621

MAE:20.883822649469874

RMSE:27.542014991065233

ARIMA(0, 1, 1)x(2, 1, 0, 52)52 - AIC:1498.732783962991



performing function on R06

Performing stepwise search to minimize aic

ARIMA(0,1,0) (1,1,1) [52]	: AIC=1239.490, Time=4.43 sec
ARIMA(0,1,0) (0,1,0) [52]	: AIC=1296.507, Time=2.54 sec
ARIMA(1,1,0) (1,1,0) [52]	: AIC=inf, Time=2.89 sec
ARIMA(0,1,1) (0,1,1) [52]	: AIC=inf, Time=4.36 sec
ARIMA(0,1,0) (0,1,1) [52]	: AIC=inf, Time=60.00 sec
ARIMA(0,1,0) (1,1,0) [52]	: AIC=inf, Time=1.75 sec
ARIMA(0,1,0) (2,1,1) [52]	: AIC=inf, Time=37.36 sec
ARIMA(0,1,0) (1,1,2) [52]	: AIC=inf, Time=23.33 sec
ARIMA(0,1,0) (0,1,2) [52]	: AIC=inf, Time=72.82 sec
ARIMA(0,1,0) (2,1,0) [52]	: AIC=inf, Time=6.19 sec
ARIMA(0,1,0) (2,1,2) [52]	: AIC=inf, Time=22.47 sec
ARIMA(1,1,0) (1,1,1) [52]	: AIC=1213.403, Time=6.32 sec
ARIMA(1,1,0) (0,1,1) [52]	: AIC=inf, Time=5.80 sec
ARIMA(1,1,0) (2,1,1) [52]	: AIC=inf, Time=50.96 sec
ARIMA(1,1,0) (1,1,2) [52]	: AIC=inf, Time=24.49 sec
ARIMA(1,1,0) (0,1,0) [52]	: AIC=1261.709, Time=0.36 sec
ARIMA(1,1,0) (0,1,2) [52]	: AIC=1214.905, Time=12.76 sec
ARIMA(1,1,0) (2,1,0) [52]	: AIC=1213.665, Time=9.45 sec
ARIMA(1,1,0) (2,1,2) [52]	: AIC=inf, Time=20.17 sec
ARIMA(2,1,0) (1,1,1) [52]	: AIC=1212.011, Time=7.41 sec
ARIMA(2,1,0) (0,1,1) [52]	: AIC=inf, Time=7.32 sec
ARIMA(2,1,0) (1,1,0) [52]	: AIC=1211.353, Time=4.05 sec
ARIMA(2,1,0) (0,1,0) [52]	: AIC=1259.149, Time=0.56 sec
ARIMA(2,1,0) (2,1,0) [52]	: AIC=1212.246, Time=11.99 sec
ARIMA(2,1,0) (2,1,1) [52]	: AIC=inf, Time=62.71 sec
ARIMA(2,1,1) (1,1,0) [52]	: AIC=1184.459, Time=7.78 sec
ARIMA(2,1,1) (0,1,0) [52]	: AIC=1230.874, Time=1.98 sec
ARIMA(2,1,1) (2,1,0) [52]	: AIC=1185.246, Time=22.94 sec
ARIMA(2,1,1) (1,1,1) [52]	: AIC=1185.439, Time=11.83 sec
ARIMA(2,1,1) (0,1,1) [52]	: AIC=inf, Time=11.30 sec
ARIMA(2,1,1) (2,1,1) [52]	: AIC=inf, Time=34.33 sec
ARIMA(1,1,1) (1,1,0) [52]	: AIC=1183.378, Time=6.65 sec
ARIMA(1,1,1) (0,1,0) [52]	: AIC=1231.491, Time=1.19 sec
ARIMA(1,1,1) (2,1,0) [52]	: AIC=1184.449, Time=19.90 sec
ARIMA(1,1,1) (1,1,1) [52]	: AIC=1184.642, Time=10.00 sec
ARIMA(1,1,1) (0,1,1) [52]	: AIC=inf, Time=7.44 sec
ARIMA(1,1,1) (2,1,1) [52]	: AIC=inf, Time=49.21 sec
ARIMA(0,1,1) (1,1,0) [52]	: AIC=1197.102, Time=3.95 sec
ARIMA(1,1,2) (1,1,0) [52]	: AIC=1184.904, Time=13.38 sec
ARIMA(0,1,2) (1,1,0) [52]	: AIC=1187.111, Time=6.34 sec
ARIMA(2,1,2) (1,1,0) [52]	: AIC=1185.670, Time=17.72 sec
ARIMA(1,1,1) (1,1,0) [52] intercept	: AIC=inf, Time=16.71 sec

Best model: ARIMA(1,1,1) (1,1,0) [52]

Total fit time: 705.233 seconds

Evaluations for R06

MSE:93.03108041447754

MAE:7.435486158580243

RMSE:9.64526207080334

ARIMA(1, 1, 1)x(1, 1, 0, 52)52 - AIC:1183.3784501593498

performing function on N02BA\_diff

Performing stepwise search to minimize aic

ARIMA(0,1,0) (1,1,1) [52]	: AIC=1456.423, Time=4.11 sec
ARIMA(0,1,0) (0,1,0) [52]	: AIC=1480.390, Time=3.10 sec
ARIMA(1,1,0) (1,1,0) [52]	: AIC=1371.196, Time=2.22 sec
ARIMA(0,1,1) (0,1,1) [52]	: AIC=inf, Time=10.30 sec
ARIMA(1,1,0) (0,1,0) [52]	: AIC=1403.202, Time=0.38 sec
ARIMA(1,1,0) (2,1,0) [52]	: AIC=1370.706, Time=7.90 sec
ARIMA(1,1,0) (2,1,1) [52]	: AIC=1372.557, Time=22.84 sec
ARIMA(1,1,0) (1,1,1) [52]	: AIC=1370.558, Time=8.28 sec
ARIMA(1,1,0) (0,1,1) [52]	: AIC=inf, Time=4.52 sec
ARIMA(1,1,0) (1,1,2) [52]	: AIC=1372.557, Time=20.05 sec
ARIMA(1,1,0) (0,1,2) [52]	: AIC=1370.590, Time=13.38 sec
ARIMA(1,1,0) (2,1,2) [52]	: AIC=1374.557, Time=19.43 sec

```

ARIMA(2,1,0)(1,1,1)[52] : AIC=1334.133, Time=8.29 sec
ARIMA(2,1,0)(0,1,1)[52] : AIC=1333.186, Time=5.75 sec
ARIMA(2,1,0)(0,1,0)[52] : AIC=1360.927, Time=0.56 sec
ARIMA(2,1,0)(0,1,2)[52] : AIC=1334.164, Time=15.32 sec
ARIMA(2,1,0)(1,1,0)[52] : AIC=1334.175, Time=3.38 sec
ARIMA(2,1,0)(1,1,2)[52] : AIC=1336.130, Time=18.07 sec
ARIMA(2,1,1)(0,1,1)[52] : AIC=inf, Time=17.93 sec
ARIMA(1,1,1)(0,1,1)[52] : AIC=inf, Time=13.64 sec
ARIMA(2,1,0)(0,1,1)[52] intercept : AIC=1335.183, Time=5.71 sec

```

```

Best model: ARIMA(2,1,0)(0,1,1)[52]
Total fit time: 205.193 seconds
Evaluations for N02BA_diff
MSE:79.75061302681316
MAE:7.16917595569059
RMSE:8.930319872592088
ARIMA(2, 1, 0)x(0, 1, 1, 52)52 - AIC:1333.1859315199022

```

```

In [77]: #Note that the lower the RMSE, MAE, and MSE are, the more accurate the model

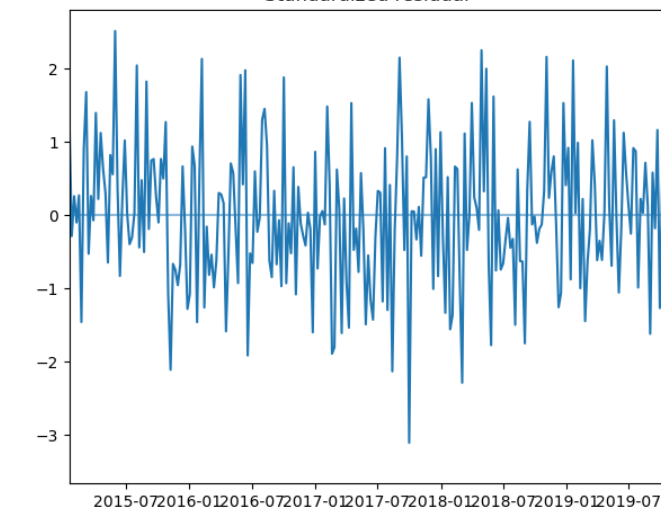
#We will now also plot the diagnostics of each model to better understand their utility:

for index in models:
    print('Diagnostics for: {}'.format(df.columns[index]))
    models[index].plot_diagnostics(figsize=(15,12))

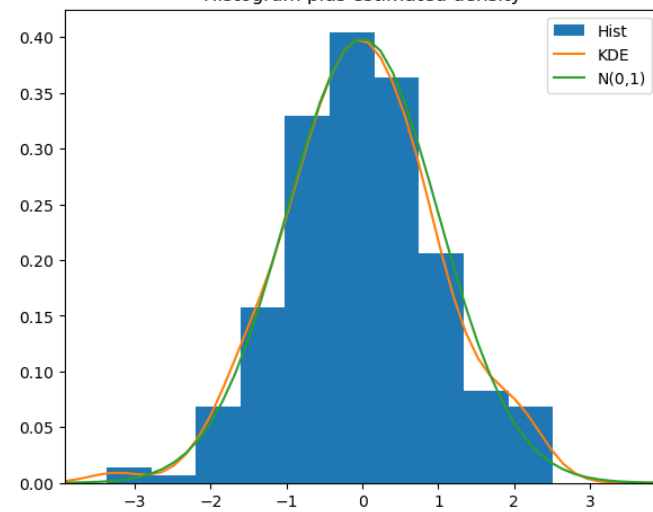
Diagnostics for: M01AB
Diagnostics for: M01AE
Diagnostics for: N02BA
Diagnostics for: N02BE
Diagnostics for: N05B
Diagnostics for: N05C
Diagnostics for: R03
Diagnostics for: R06
Diagnostics for: N02BA_diff

```

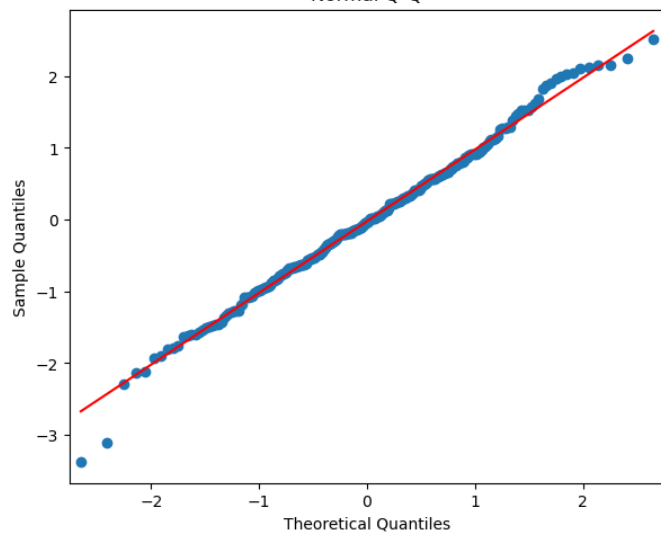
Standardized residual



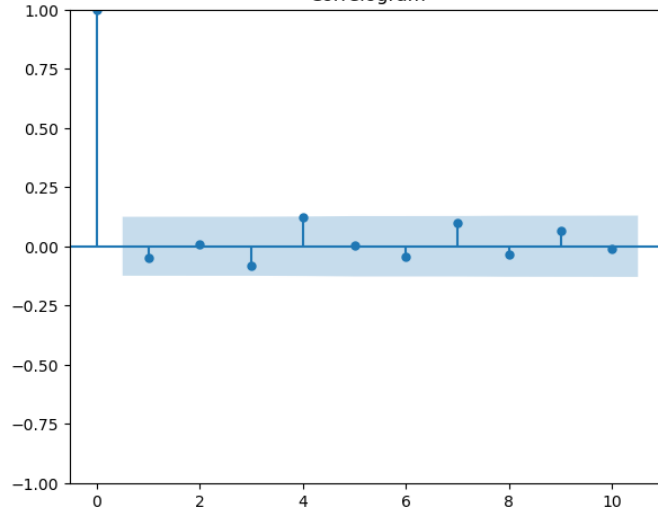
Histogram plus estimated density



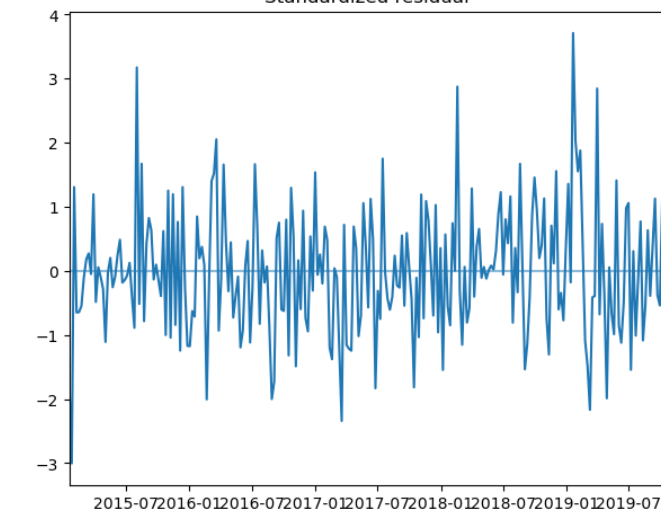
Normal Q-Q



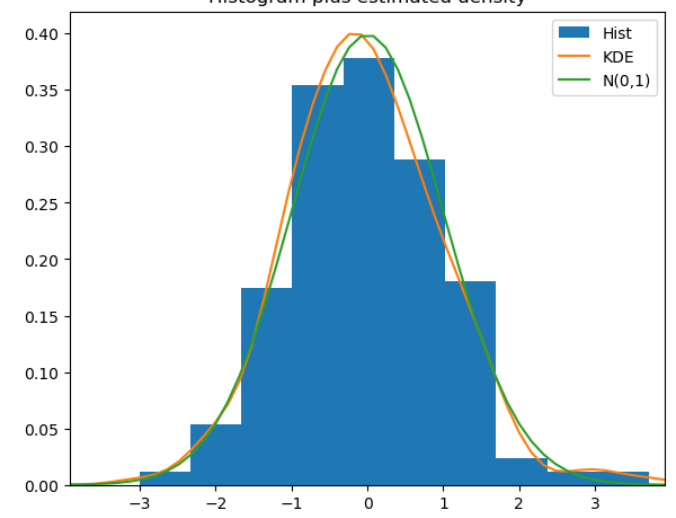
Correlogram



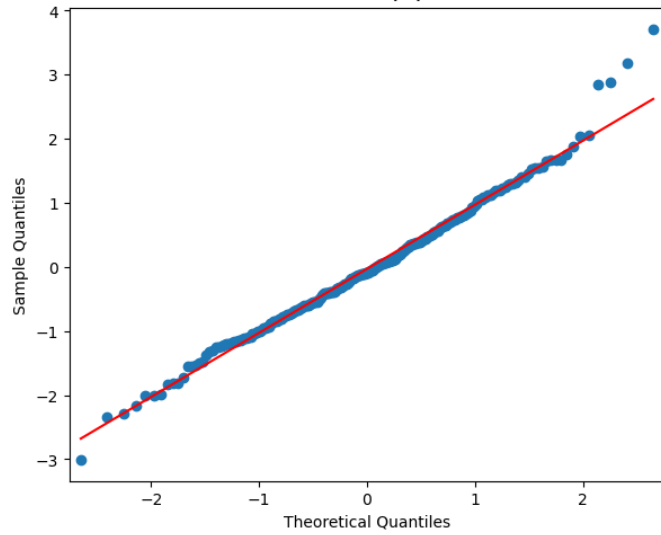
Standardized residual



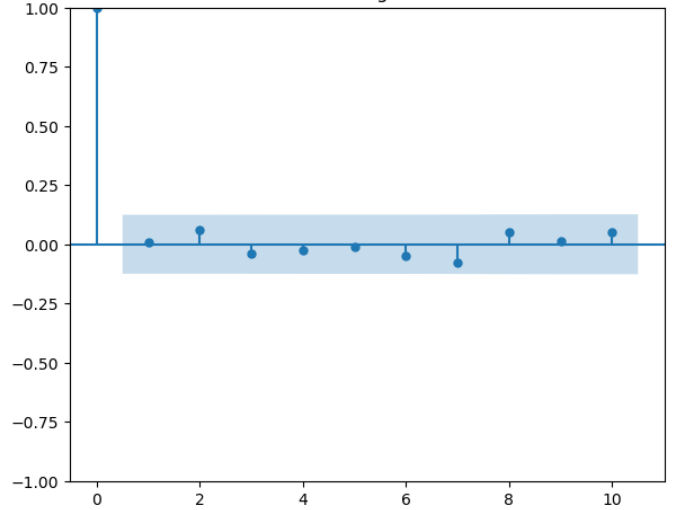
Histogram plus estimated density



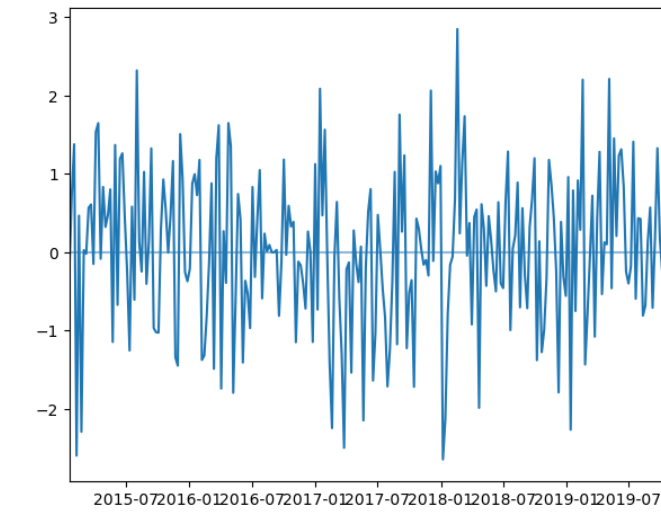
Normal Q-Q



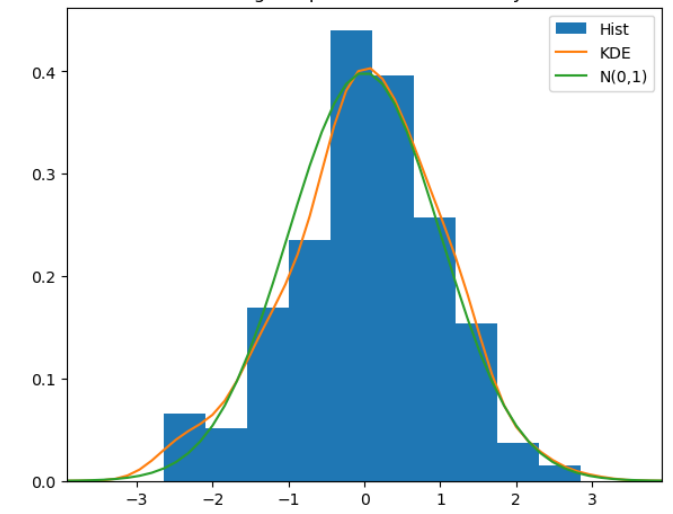
Correlogram



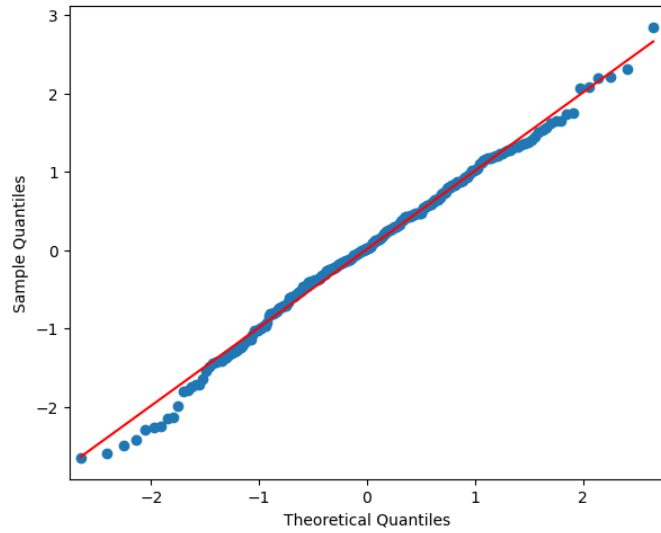
Standardized residual



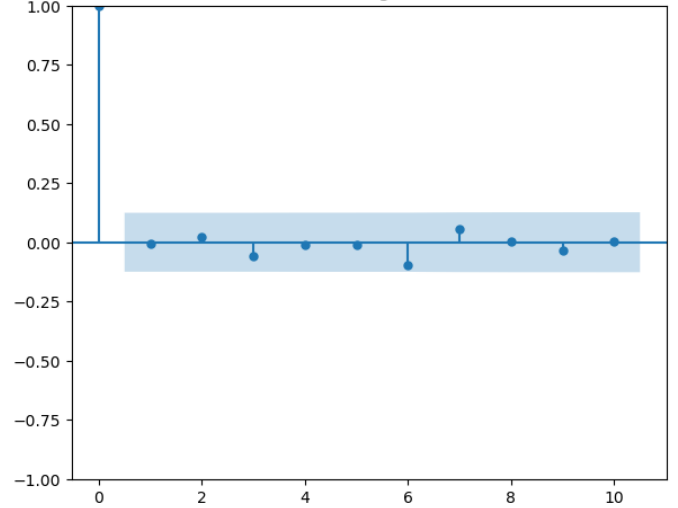
Histogram plus estimated density



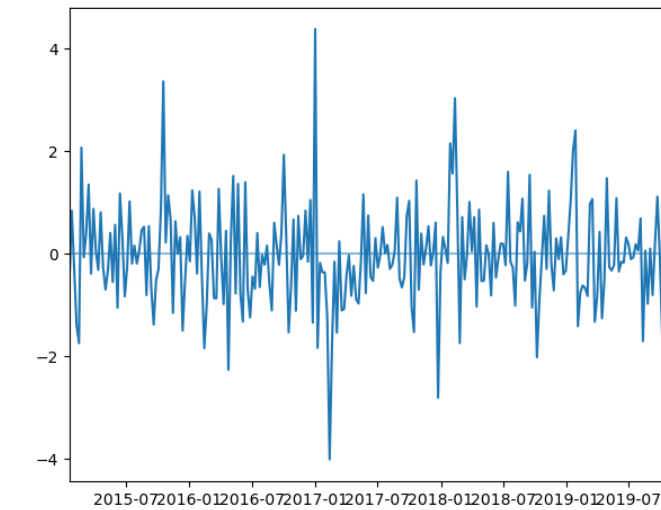
Normal Q-Q



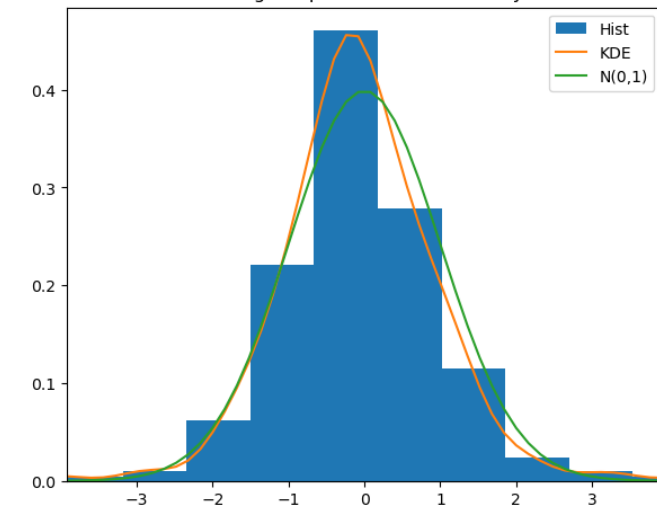
Correlogram



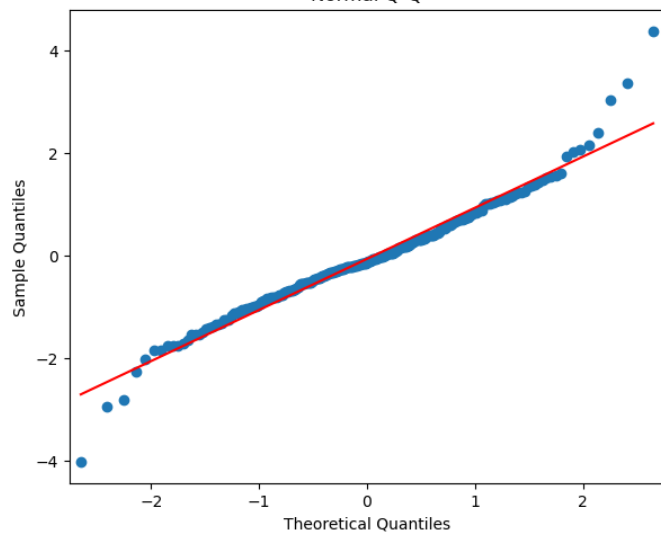
Standardized residual



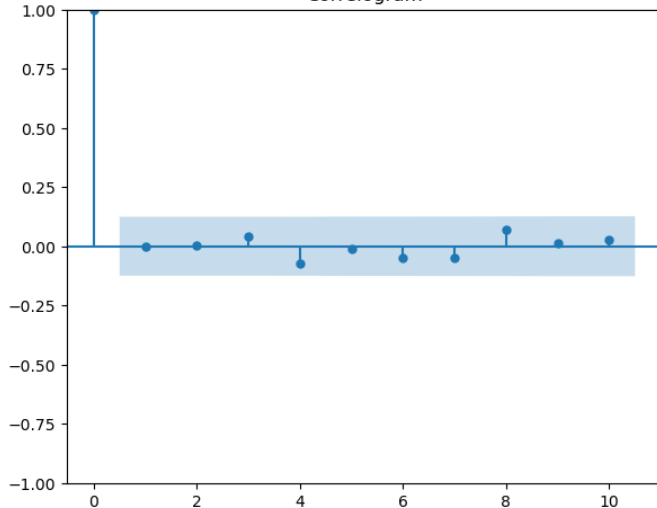
Histogram plus estimated density

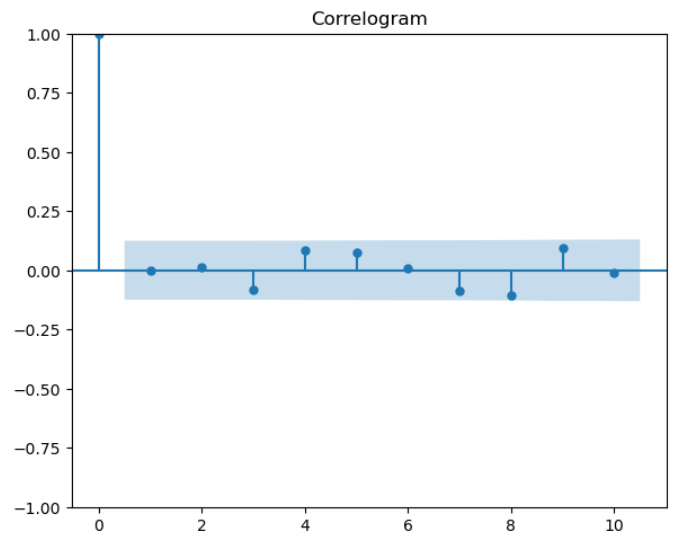
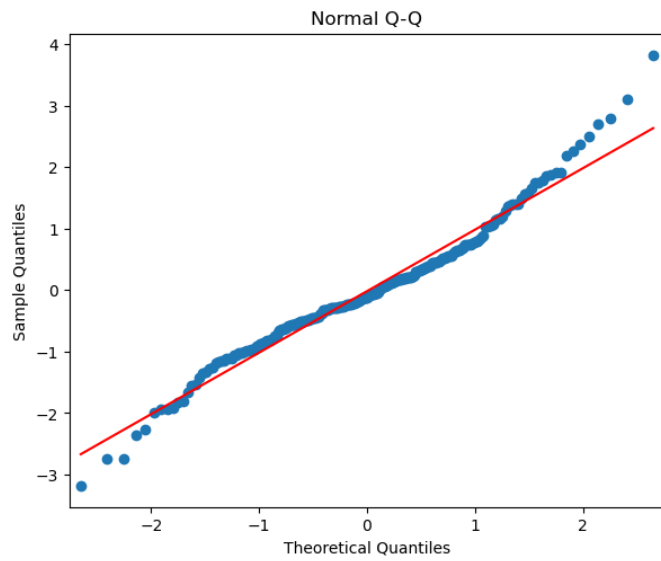
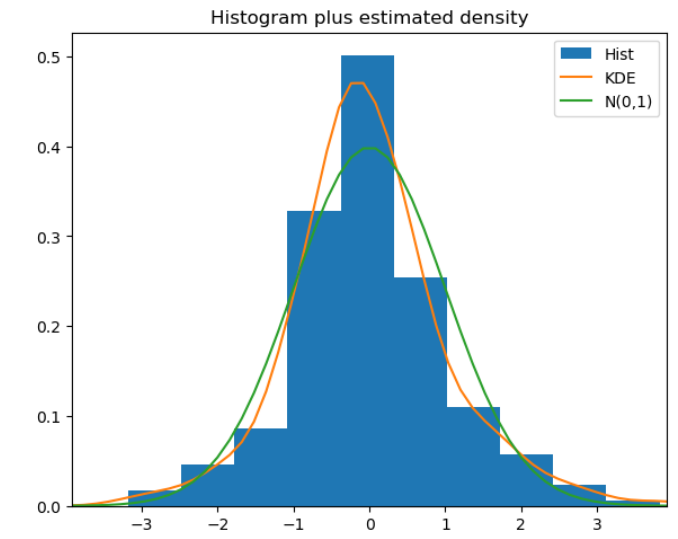
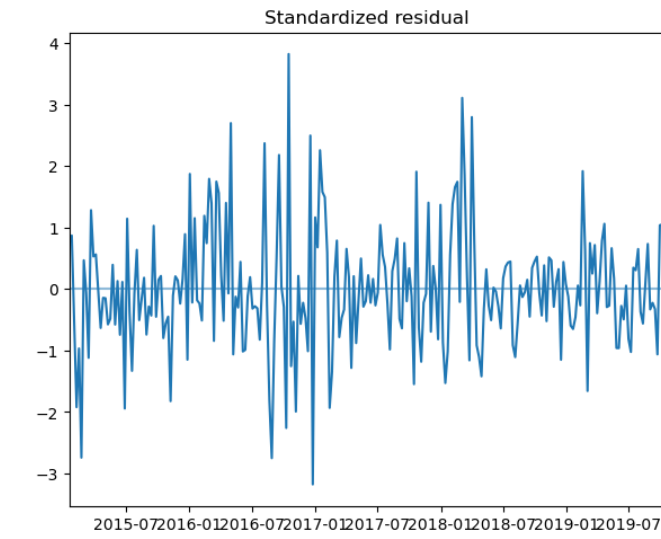


Normal Q-Q

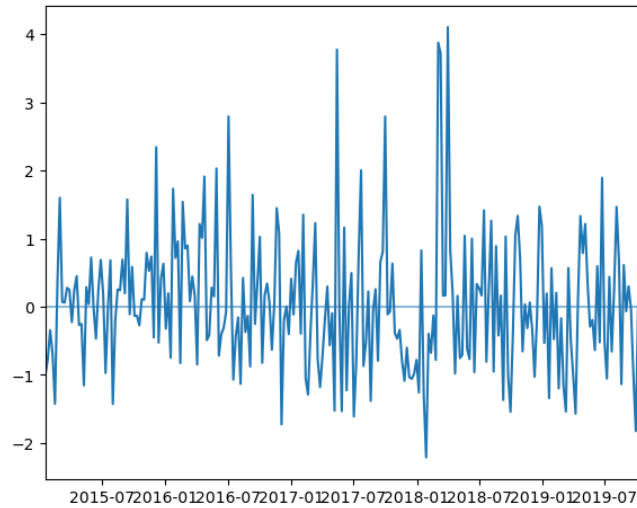


Correlogram

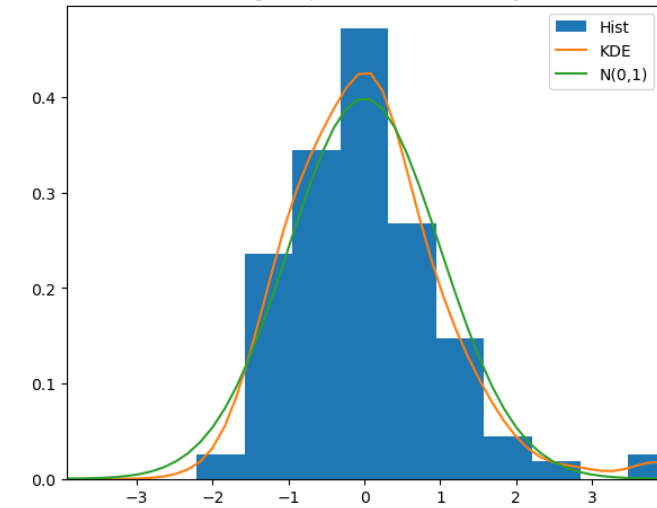




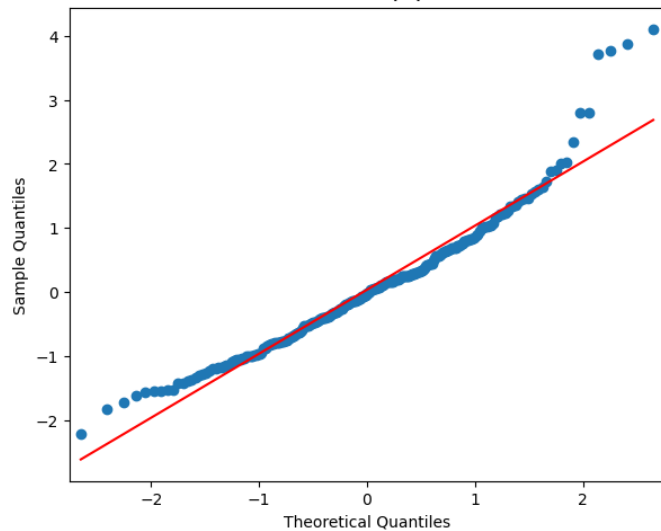
Standardized residual



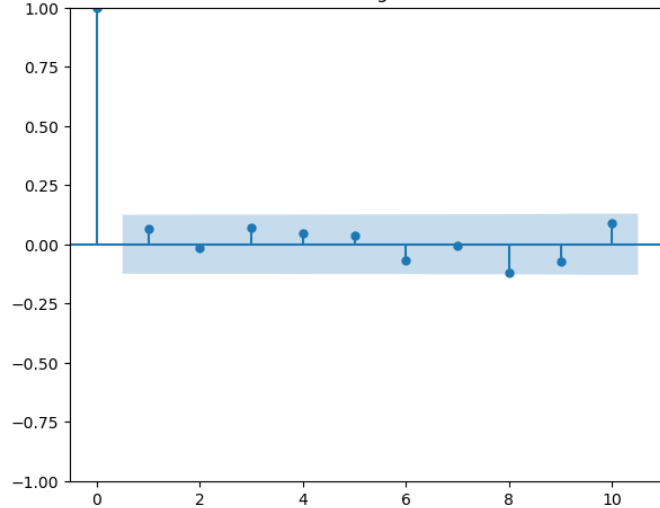
Histogram plus estimated density



Normal Q-Q

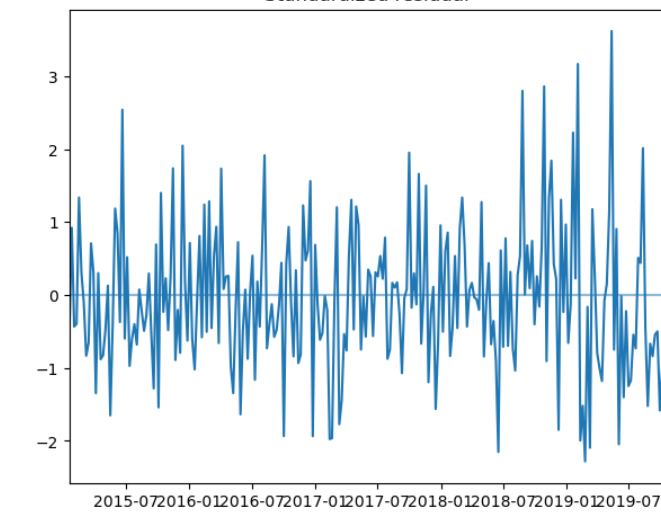


Correlogram

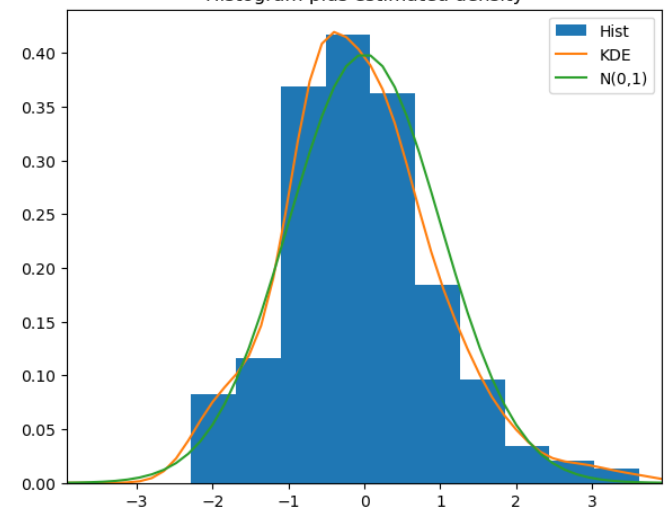




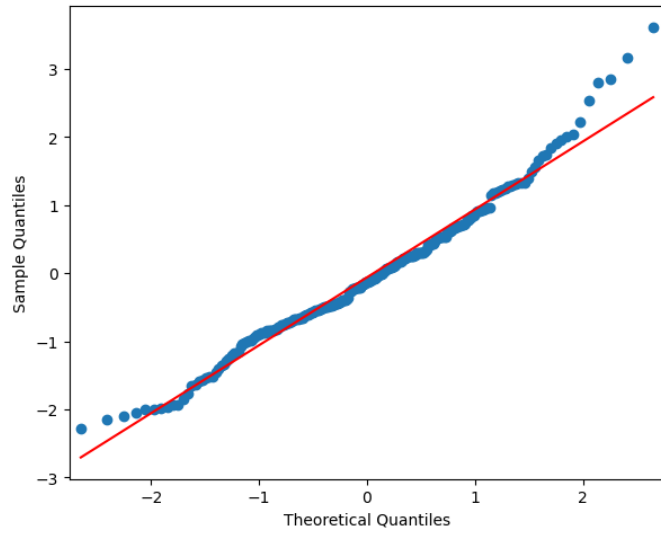
Standardized residual



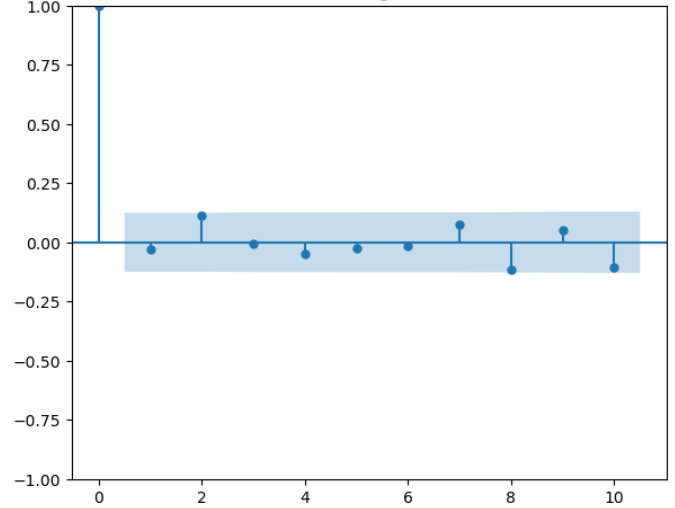
Histogram plus estimated density



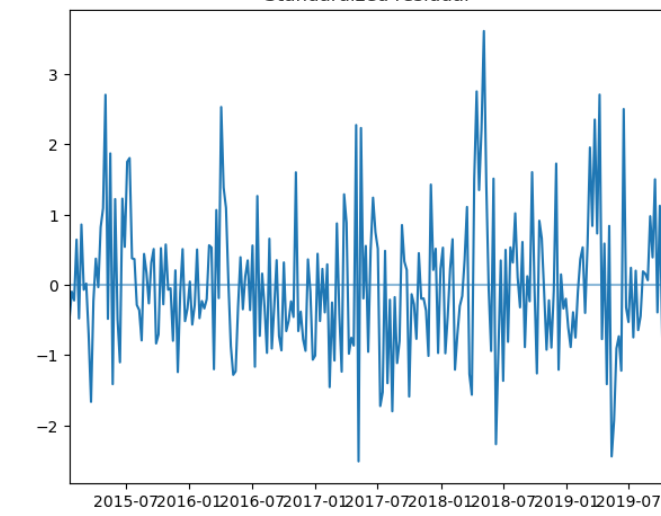
Normal Q-Q



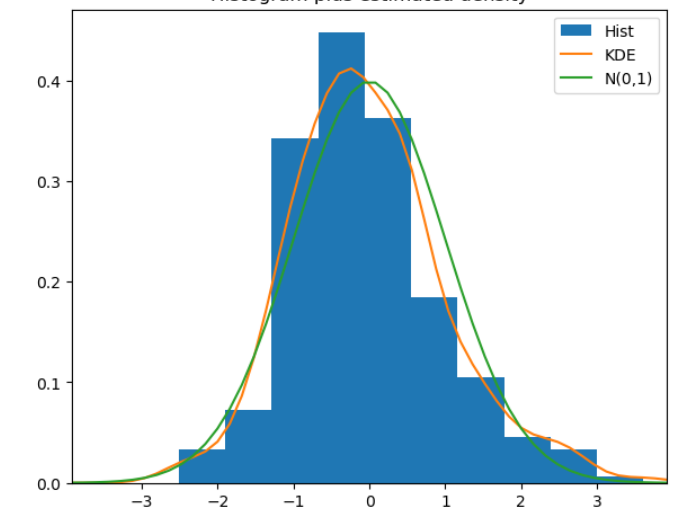
Correlogram



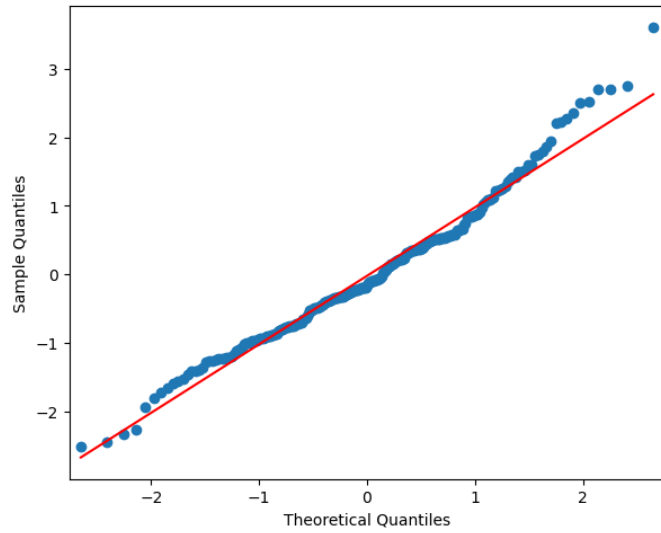
Standardized residual



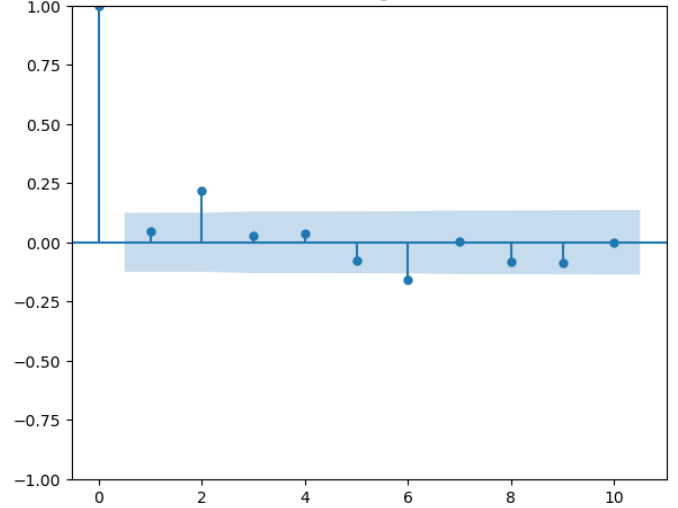
Histogram plus estimated density

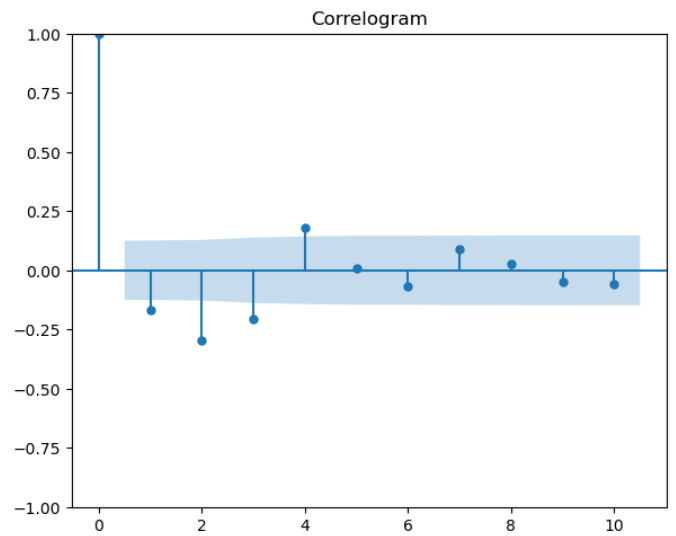
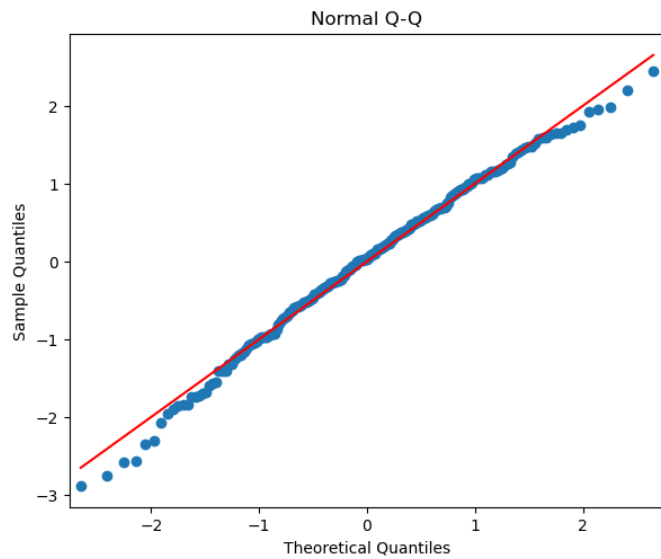
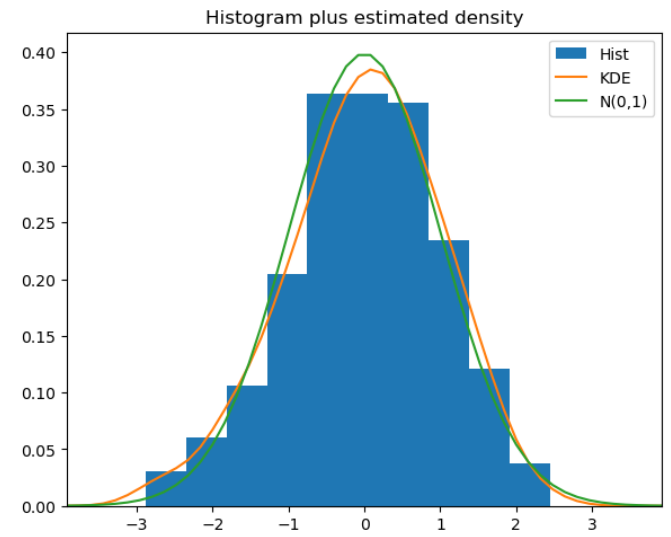
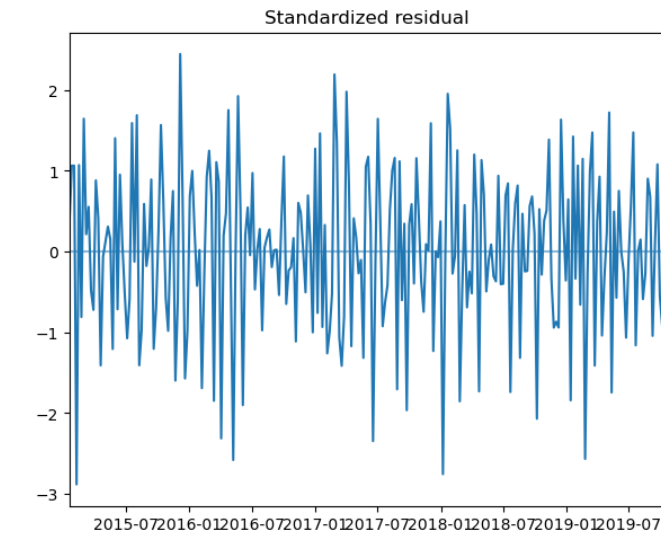


Normal Q-Q



Correlogram





In [78]:

```
Cell In[78], line 1
jupyter nbconvert --to webpdf --allow-chromium-download Untitled.ipynb
^
SyntaxError: invalid syntax
```

In [80]:

```
UsageError: Line magic function `%jupyter` not found.
```