

```
[76]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import datetime
import warnings
warnings.filterwarnings('ignore')

print('numpy version:', np.__version__)
print('matplotlib version:', plt.__version__)

#read csv file
df1 = pd.read_csv('C:/Users/Bilal Aktas/Desktop/Git Lab/Fontys/Fontys Subjects/Semester 4/Aplied Data Science/Fontys Subjects/Semester 4/Aplied Data Science/deliv.csv')
df2 = pd.read_csv('C:/Users/Bilal Aktas/Desktop/Git Lab/Fontys/Fontys Subjects/Semester 4/Aplied Data Science/contrib.csv')

df = pd.read_csv ('C:/Users/Bilal Aktas/Desktop/CleanedDataset.csv')

numpy version: 1.19.2
matplotlib version: 1.1.5

In [77]: #add columns
df1.columns = ['Delivery date', 'Delivery time', 'Pharmacy number', 'Pharmacy Postcode (2)', 'Year of birth', 'Gender', 'CNK', 'Product name', 'ATC code', 'Units', 'Price', 'Contribution']
df2.columns = ['Delivery date', 'Delivery time', 'Pharmacy number', 'Pharmacy Postcode (2)', 'Year of birth', 'Gender', 'CNK', 'Product name', 'ATC code', 'Units', 'Price', 'Contribution']

In [78]: #check dataframe
df1.head(10)
```

	Delivery date	Delivery time	Pharmacy number	Pharmacy Postcode (2)	Year of birth	Gender	CNK	Product name	ATC code	Units	Price	Contribution
0	01/01/2017	00:00	7341765		21	1924	1	AMOXICILAV SANDOZ SODMG/125 MG COMP 30	J01CR02	30	14.81	3.47
1	01/01/2017	00:00	7341765		21	1922	1	WACHTHONORARIUM		0	4.90	0.00
2	01/01/2017	00:00	7341765		21	1925	1	ZALDIAR 375 MG/325 MG FLOMOHM TABL 20	N02AJA13	20	9.26	3.62
3	01/01/2017	00:00	8272695		16	1932	2	VASEXTEN CAPS BUST 28 X 10 MG	C08CA12	28	19.22	4.98
4	01/01/2017	00:00	8272695		16	1933	2	WACHTHONORARIUM		0	4.90	0.00
5	01/01/2017	00:00	9111423		10	1931	1	AACIDEXAM SINGMUL OPR INU FLIN J 1 X 1ML	H02AB02	1	6.15	0.39
6	01/01/2017	00:00	8272695		10	1933	1	AACIDEXAM SINGMUL OPR INU FLIN J 1 X 1ML	H02AB02	1	6.15	0.39
7	01/01/2017	00:00	8272695		16	1935	2	AMOXICILAV SANDOZ 875 MG/125 MG COMP 20	J01CR02	20	15.18	3.58
8	01/01/2017	00:00	8272695		16	1933	2	WACHTHONORARIUM		0	4.90	0.00
9	01/01/2017	00:00	8272695		21	1940	1	WACHTHONORARIUM		0	4.90	0.00

```
In [79]: #gives information of dataframe
df1.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 22705348 entries, 0 to 22705347
Data columns (total 12 columns):
 #   Column      Dtype
---  ---
 0   Delivery date    object
 1   Delivery time    int64
 2   Pharmacy number  int64
 3   Pharmacy Postcode (2)  int64
 4   Year of birth    int64
 5   Gender          int64
 6   CNK             int64
 7   Product name    object
 8   ATC code        object
 9   Units           int64
10   Price           float64
11  Contribution     float64
dtypes: float64(2), int64(6), object(4)
memory usage: 2.0+ GB

In [80]: df2.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 507116 entries, 0 to 507115
Data columns (total 12 columns):
 #   Column      Dtype
---  ---
 0   Delivery date    object
 1   Delivery time    object
 2   Pharmacy number  int64
 3   Pharmacy Postcode (2)  int64
 4   Year of birth    int64
 5   Gender          int64
 6   CNK             int64
 7   Product name    object
 8   ATC code        object
 9   Units           int64
10   Price           float64
11  Contribution     float64
dtypes: float64(2), int64(6), object(4)
memory usage: 464.4+ MB

In [81]: #sums all the null values of the columns in the dataframe
df1.isnull().sum()
```

	Delivery date	Delivery time	Pharmacy number	Pharmacy Postcode (2)	Year of birth	Gender	Product name	ATC code	Units	Price	Contribution
	0	4941181	0	0	0	0	0	0	False	0	0
	Dtype: object	Dtype: int64	Dtype: int64	Dtype: int64	Dtype: int64	Dtype: int64	Dtype: object	Dtype: object	Dtype: int64	Dtype: float64	Dtype: float64

```
In [82]: df2.isnull().sum()
```

	Delivery date	Delivery time	Pharmacy number	Pharmacy Postcode (2)	Year of birth	Gender	Product name	ATC code	Units	Price	Contribution
	0	971059	0	0	0	0	0	0	False	0	0
	Dtype: object	Dtype: int64	Dtype: int64	Dtype: int64	Dtype: int64	Dtype: int64	Dtype: object	Dtype: object	Dtype: int64	Dtype: float64	Dtype: float64

```
In [83]: df1.isnull().any()
```

	Delivery date	Delivery time	Pharmacy number	Pharmacy Postcode (2)	Year of birth	Gender	Product name	ATC code	Units	Price	Contribution
	True	False	False	False	False	False	False	False	False	False	False
	Dtype: bool	Dtype: bool	Dtype: bool	Dtype: bool	Dtype: bool	Dtype: bool	Dtype: bool	Dtype: bool	Dtype: bool	Dtype: bool	Dtype: bool

```
In [84]: #filling na values in delivery time column with 00:00
TimeNull = df1['Delivery time']
TimeNull.fillna("00:00",inplace=True)

In [85]: TimeNull = df2['Delivery time']
TimeNull.fillna("00:00",inplace=True)

In [86]: #dropping outliers in gender column
df1 = df1[df1['Gender'] != 0]
df1 = df1[df1['Gender'] != 3]

In [87]: df2 = df2[df2['Gender'] != 0]
df2 = df2[df2['Gender'] != 3]

In [88]: df1["Gender"].value_counts(sort=True)
```

	Name:	Gender:	dtype:
Out[88]:	2	13061752	
	1	9563378	
			int64

```
In [89]: df2["Gender"].value_counts(sort=True)
```

	Name:	Gender:	dtype:
Out[89]:	2	2856183	
	1	1205711	
			int64

```
In [90]: #transform into datetime
df1['Delivery date'] = pd.to_datetime(df1['Delivery date'], infer_datetime_format=True)

In [91]: df2['Delivery date'] = pd.to_datetime(df2['Delivery date'], infer_datetime_format=True)

In [92]: #checking the value which date was both d-m-Y and m-d-Y
rat_df = df1[df1['Delivery date'] > "2017-01-30"]

print('\nResult dataframe \n', rat_df)

Result dataframe :
   Delivery date  Delivery time  Pharmacy number  Pharmacy
```

```

d = df

In [341.]: df['ATCShort'] = df['ATC code'].str[:3]
df.head()

Out[341.]:

```

	Delivery date	Delivery time	Pharmacy number	Pharmacy Postcode (2)	Year of birth	Gender	CNK	Product name	ATC code	Units	Price	Contribution	ATCShort
0	2017-01-01	00:00	7341765	21	1924	1	1715119	AMOXICLAV SANDOZ 500MG/125 MG COMP 30	J01CR02	30	14.81	3.47	J01
	2017-01-01	00:00	7341765	21	1922	1	5520253	WACHTHONORARIUM ZALDAR 37.5 MG/25		0	4.90	0.00	
2	2017-01-01	00:00	7341765	21	1925	1	1799931	MG FILONMHI TABL 20	N02AJ13	20	9.26	3.62	N02
3	2017-01-01	00:00	8272695	16	1932	2	1719400	VASEXTIN CAPS BUST 28 x 10 MG	C08CA12	28	19.22	4.98	C08
4	2017-01-01	00:00	8272695	16	1933	2	5520253	WACHTHONORARIUM		0	4.90	0.00	

```

In [342.]: df['Age'] = 2021 - df['Year of birth']

In [343.]: tempMed = ['A02', 'B01', 'C07', 'C03', 'A10']
df['ATCShort'].isin(tempMed)
med_filter = df[df['ATCShort'].isin(tempMed)]

med_filter['ATC_ABB1'] = med_filter['ATCShort'].map({'A02': 0, 'B01': 1, 'C07': 2, 'C03': 3, 'A10': 4})

Particular ATC

A02 Drugs for acid related disorders

In [267.]: df = df.loc[(df['ATCShort']=="A02")]

In [268.]: df = df.sort_values('Delivery date')

[269.]: df = df.groupby('Delivery date')['Units'].sum().reset_index()

In [270.]: df = df.set_index('Delivery date')

In [272.]: y = df['Units'].resample('MS').mean()

In [273.]: import itertools
import statmodels.api as sm
p = d = q = range(0, 2)

# Generate all different combinations of p, q and q triples
pdq = list(itertools.product(p, d, q))

# Generate all different combinations of seasonal p, q and q triples
seasonal_pdq = [(k[0], x[1], x[2], 12) for x, n in list(itertools.product(p, d, q))]

print('Examples of parameter combinations for Seasonal ARIMA...')
print('SARIMAX: (1 x 1), format: [pdq1], seasonal: [pdq1]')
print('SARIMAX: (1 x 1), format: [pdq1], seasonal: [pdq2]')
print('SARIMAX: (1 x 1), format: [pdq2], seasonal: [pdq3]')
print('SARIMAX: (1 x 1), format: [pdq2], seasonal: [pdq4]')

Examples of parameter combinations for Seasonal ARIMA...
SARIMAX: (0, 0, 1) x (0, 0, 1, 12)
SARIMAX: (0, 0, 1) x (0, 0, 1, 0, 12)
SARIMAX: (0, 1, 0) x (0, 1, 0, 1, 12)
SARIMAX: (0, 1, 0) x (1, 0, 0, 0, 12)

In [274.]: for param in pdq:
for param_seasonal in seasonal_pdq:
try:
mod = sm.tsa.statespace.SARIMAX(y,
order=param,
seasonal_order=param_seasonal,
enforce_stationarity=False,
enforce_invertibility=False)

results = mod.fit()

print('ARIMA({}x{}12 - AIC:{}'.format(param, param_seasonal, results.aic))
except:
continue

ARIMA(0, 0, 0)x(0, 0, 0, 12)12 - AIC:1890.189368252876
ARIMA(0, 0, 0)x(0, 0, 0, 1, 12)12 - AIC:1565.4153085703108
ARIMA(0, 0, 0)x(0, 0, 1, 0, 12)12 - AIC:488.61552023538275
ARIMA(0, 0, 0)x(0, 0, 1, 1, 12)12 - AIC:226.91563704043097
ARIMA(0, 0, 0)x(0, 1, 0, 0, 12)12 - AIC:506.58639926497085
ARIMA(0, 0, 0)x(0, 1, 0, 1, 12)12 - AIC:487.45073518589875
ARIMA(0, 0, 0)x(0, 1, 1, 0, 12)12 - AIC:246.91781409333325
ARIMA(0, 0, 0)x(0, 1, 1, 1, 12)12 - AIC:227.4309020131438
ARIMA(0, 0, 1)x(0, 0, 0, 0, 12)12 - AIC:841.7531626217959
ARIMA(0, 0, 1)x(0, 0, 0, 1, 12)12 - AIC:544.8053935202336
ARIMA(0, 0, 1)x(0, 0, 1, 0, 12)12 - AIC:466.1501631756076
ARIMA(0, 0, 1)x(0, 0, 1, 1, 12)12 - AIC:208.5430311143912
ARIMA(0, 0, 1)x(0, 1, 0, 0, 12)12 - AIC:191.05221146054608
ARIMA(0, 0, 1)x(0, 1, 0, 1, 12)12 - AIC:544.2489268565637
ARIMA(0, 0, 1)x(0, 1, 1, 0, 12)12 - AIC:247.82250916424868
ARIMA(0, 0, 1)x(0, 1, 1, 1, 12)12 - AIC:206.23406439629193
ARIMA(0, 1, 0)x(0, 0, 0, 0, 12)12 - AIC:715.5530207767473
ARIMA(0, 1, 0)x(0, 0, 0, 1, 12)12 - AIC:450.693487178215
ARIMA(0, 1, 0)x(0, 0, 1, 0, 12)12 - AIC:466.64894655619077
ARIMA(0, 1, 0)x(0, 0, 1, 1, 12)12 - AIC:209.5591478088914
ARIMA(0, 1, 0)x(0, 1, 0, 0, 12)12 - AIC:464.18008087307257
ARIMA(0, 1, 0)x(0, 1, 0, 1, 12)12 - AIC:452.32260504922726
ARIMA(0, 1, 0)x(0, 1, 1, 0, 12)12 - AIC:228.62534785418762
ARIMA(0, 1, 0)x(0, 1, 1, 1, 12)12 - AIC:210.86333485513071
ARIMA(0, 1, 1)x(0, 0, 0, 0, 12)12 - AIC:695.046722068930
ARIMA(0, 1, 1)x(0, 0, 0, 1, 12)12 - AIC:422.07021234262845
ARIMA(0, 1, 1)x(0, 0, 1, 0, 12)12 - AIC:436.34986278505263
ARIMA(0, 1, 1)x(0, 0, 1, 1, 12)12 - AIC:186.03117024629802
ARIMA(0, 1, 1)x(0, 1, 0, 0, 12)12 - AIC:463.2454571002089
ARIMA(0, 1, 1)x(0, 1, 0, 1, 12)12 - AIC:422.70083961659304
ARIMA(0, 1, 1)x(0, 1, 1, 0, 12)12 - AIC:226.08209352365705
ARIMA(0, 1, 1)x(0, 1, 1, 1, 12)12 - AIC:187.98935163182192
ARIMA(0, 1, 1)x(0, 1, 1, 1, 12)12 - AIC:737.7769321281378
ARIMA(0, 0, 0)x(0, 0, 1, 1, 12)12 - AIC:473.0213791514545
ARIMA(0, 0, 0)x(0, 0, 1, 1, 12)12 - AIC:482.6883081361385
ARIMA(0, 0, 0)x(0, 1, 1, 12)12 - AIC:229.47288184400626
ARIMA(0, 0, 0)x(0, 1, 1, 12)12 - AIC:474.02722874464602
ARIMA(0, 0, 0)x(0, 1, 1, 12)12 - AIC:474.7747122747352
ARIMA(0, 1, 0)x(0, 1, 1, 12)12 - AIC:228.22177373030966
ARIMA(0, 1, 0)x(0, 1, 1, 12)12 - AIC:229.4871494288048
ARIMA(0, 1, 1, 0, 0, 12)12 - AIC:727.0772397435351
ARIMA(0, 1, 1, 0, 0, 12)12 - AIC:445.1353543861806
ARIMA(0, 1, 1, 0, 0, 12)12 - AIC:464.16628280723306
ARIMA(0, 1, 1, 0, 1, 12)12 - AIC:206.0205645972613
ARIMA(0, 1, 1, 0, 1, 12)12 - AIC:466.3052497166157
ARIMA(0, 1, 1, 0, 1, 12)12 - AIC:466.20786508774633
ARIMA(0, 1, 1, 0, 1, 12)12 - AIC:227.359806741443394
ARIMA(0, 1, 1, 0, 1, 12)12 - AIC:207.57566521733134
ARIMA(0, 1, 1, 0, 1, 12)12 - AIC:716.779357
```

Jan  
2017

```
y_forecasted = pred.predicted_mean
y_truth = y['2018-12-01:']
# Compute the mean squared error
mse = ((y_forecasted - y_truth) ** 2).mean()
```

```

The Mean Squared Error of our forecasts is 23982136.72
The Root Mean Squared Error of our forecasts is 3475.59

In [279]:
pred_uc = results.get_forecast(steps=100)
pred_ci = pred_uc.conf_int()
ax = y.plot(label='Observed', figsize=(20, 10), color='r', linewidth=6)
pred_uc.predicted_mean.plot(ax=ax, label='Forecast', color='b', linewidth=6)
ax.set_xlabel('Date', fontsize=15)
ax.set_ylabel('Sales in Quantity', fontsize=15)
ax.set_title('The amount of Sales of medicines for acid related disorders', fontsize=20)
plt.legend()
plt.show()

```

```

B01 Antithrombotic agents

In [289]:
df = df.loc[df['ATCShort']=='B01']

In [300]:
df = df.sort_values('Delivery date')

In [301]:
df = df.groupby('Delivery date')['Units'].sum().reset_index()

In [302]:
df = df.set_index('Delivery date')

In [303]:
y = df['Units'].resample('MS').mean()

In [304]:
import itertools
import statsmodels.api as sm
p = d = q = range(0, 2)

# Generate all different combinations of p, q and q triplets
pdq = list(itertools.product(p, d, q))

# Generate all different combinations of seasonal p, q and q triplets
seasonal_pdq = [(x[0], x[1], x[2], 12) for x in list(itertools.product(p, d, q))]

print('Examples of parameter combinations for Seasonal ARIMA...')
print('SARIMAX: (1 x 1)' + ''.format(pdq[1], seasonal_pdq[1]))
print('SARIMAX: (1 x 1)' + ''.format(pdq[1], seasonal_pdq[2]))
print('SARIMAX: (1 x 1)' + ''.format(pdq[2], seasonal_pdq[3]))
print('SARIMAX: (1 x 1)' + ''.format(pdq[4], seasonal_pdq[4]))

Examples of parameter combinations for Seasonal ARIMA...
SARIMAX: (0, 0, 1) x (0, 0, 1, 12)
SARIMAX: (0, 0, 1) x (0, 1, 0, 12)
SARIMAX: (0, 1, 0) x (0, 1, 1, 12)
SARIMAX: (0, 1, 0) x (1, 0, 0, 12)

In [305]:
for param in pdq:
    for param_seasonal in seasonal_pdq:
        try:
            mod = sm.tsa.statespace.SARIMAX(y,
                                              order=param,
                                              seasonal_order=param_seasonal,
                                              enforce_stationarity=False,
                                              enforce_invertibility=False)

            results = mod.fit()

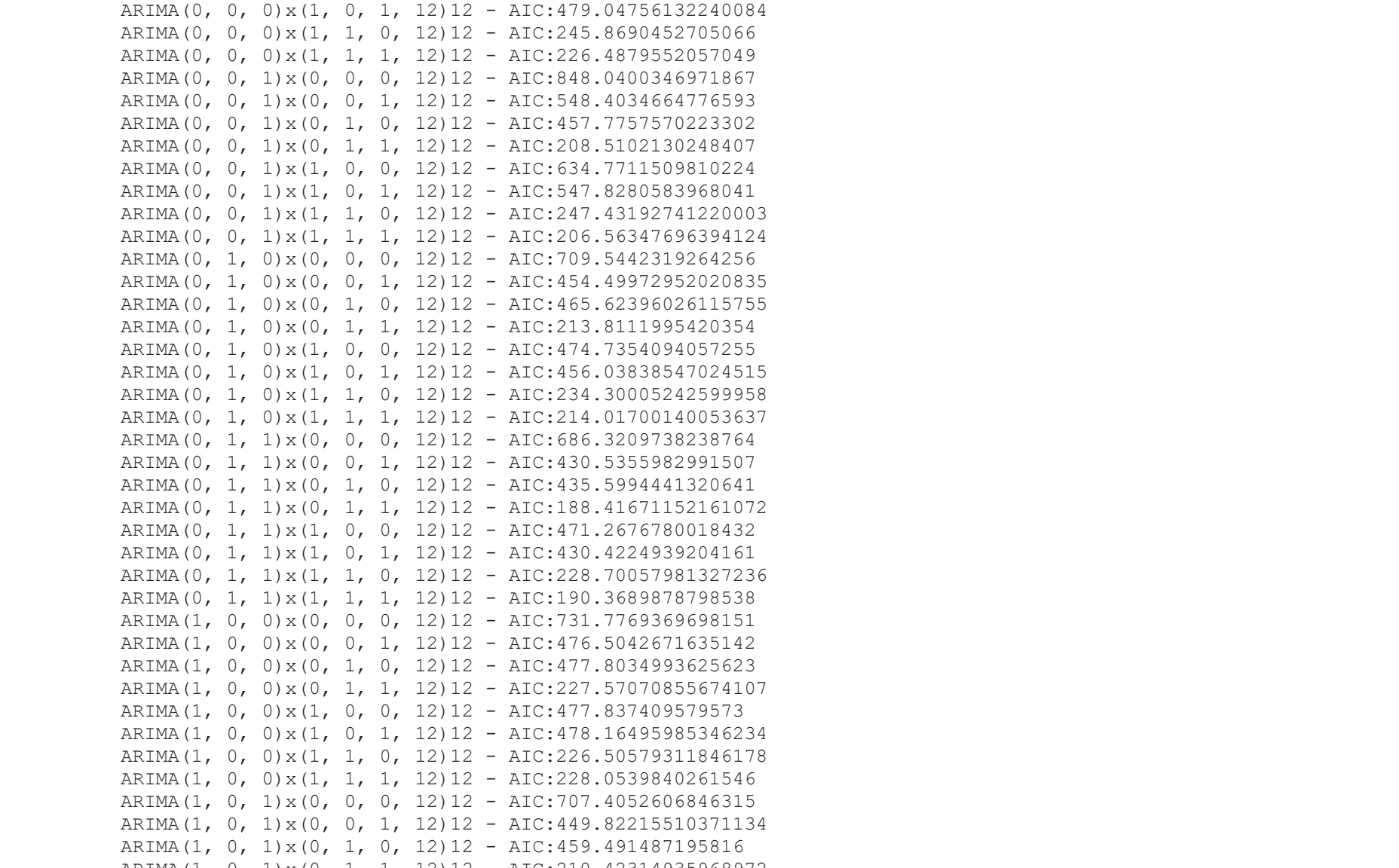
            print('ARIMA({}x){}12 - AIC:{}'.format(param, param_seasonal, results.aic))
        except:
            continue

```

```

ARIMA(0, 0, 0)x(0, 0, 0, 12)12 - AIC:896.0473809781238
ARIMA(0, 0, 0)x(0, 0, 0, 1, 12)12 - AIC:566.4654466537661
ARIMA(0, 0, 0)x(0, 1, 0, 12)12 - AIC:478.3529179086874
ARIMA(0, 0, 0)x(0, 1, 1, 12)12 - AIC:226.24136958207527
ARIMA(0, 0, 0)x(1, 0, 0, 12)12 - AIC:498.2693887526009

```



```

ARIMA(1, 0, 1)x(1, 0, 0, 12, 12) - AIC:473.2444000887415
ARIMA(1, 0, 1)x(1, 0, 1, 12, 12) - AIC:451.7042012051508
ARIMA(1, 0, 1)x(1, 1, 0, 12, 12) - AIC:527.5722840197307
ARIMA(1, 0, 1)x(1, 1, 1, 12, 12) - AIC:208.43870580080758
ARIMA(1, 1, 0)x(0, 0, 0, 12, 12) - AIC:710.8231961401093
ARIMA(1, 1, 0)x(0, 0, 1, 12, 12) - AIC:455.6892492898954
ARIMA(1, 1, 0)x(0, 1, 0, 12, 12) - AIC:465.08971494963373
ARIMA(1, 1, 0)x(0, 1, 1, 12, 12) - AIC:211.661363284753
ARIMA(1, 1, 1)x(0, 0, 0, 12, 12) - AIC:456.6550184740089
ARIMA(1, 1, 1)x(0, 1, 0, 12, 12) - AIC:457.60948782803126
ARIMA(1, 1, 1)x(0, 1, 0, 12, 12) - AIC:211.91935967277385
ARIMA(1, 1, 1)x(0, 1, 1, 12, 12) - AIC:205.58818018035376
ARIMA(1, 1, 1)x(0, 1, 0, 12, 12) - AIC:687.7014702973195
ARIMA(1, 1, 1)x(0, 0, 1, 12, 12) - AIC:430.8522007050026
ARIMA(1, 1, 1)x(0, 1, 0, 12, 12) - AIC:437.51195840380274
ARIMA(1, 1, 1)x(0, 1, 1, 12, 12) - AIC:189.2715495958417
ARIMA(1, 1, 1)x(1, 0, 0, 12, 12) - AIC:452.6138692467982
ARIMA(1, 1, 1)x(1, 0, 1, 12, 12) - AIC:432.3009323686176
ARIMA(1, 1, 1)x(1, 1, 0, 12, 12) - AIC:208.483269755357
ARIMA(1, 1, 1)x(1, 1, 1, 12, 12) - AIC:183.0427245005957

```

```

In [306]: import statsmodels.api as sm
mod = sm.tsa.statespace.SARIMAX(y,
                                order=(1,1,1),
                                seasonal_order=(1, 1, 1, 12),
                                enforce_stationarity=False,
                                enforce_invertibility=False)

results = mod.fit()

print(results.summary().tables[1])

=====
               coef      std err          z       P>|z|    [0.025    0.975]
-----
ar.L1      -0.6986      0.407       -1.716     0.086   -1.497    0.099
ma.L1       -0.3908     0.191      -2.045     0.041   -0.765   -0.016
ar.L1.L12   -1.3522     0.186      -7.131     0.000   -1.689   -0.961
ma.L1.L12   1.9238     0.309      6.236     0.000   1.319   2.528
sigma2      6.73e+06      1.9e+08   3.58e+14   0.000   6.72e+06   6.73e+06
=====

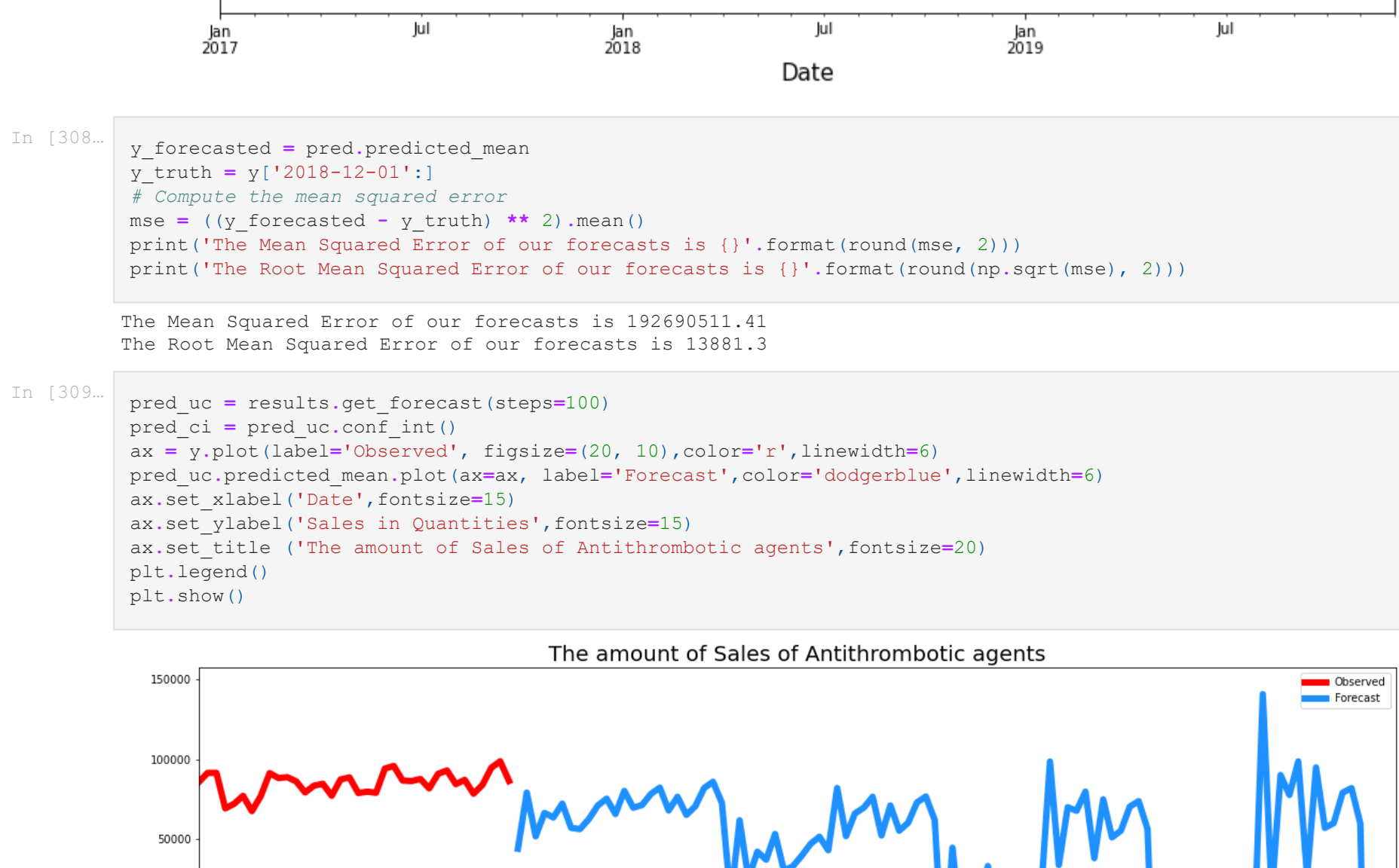
```

```

In [307]: pred = results.get_prediction(start=pd.to_datetime('2018-12-01'), dynamic=False)
pred_ci = pred.conf_int()
ax = y['2013':].plot(label='observed')
pred.predicted_mean.plot(ax=ax, label='One-step ahead Forecast', alpha=.7, figsize=(14, 7))

ax.fill_between(pred_ci.index,
                pred_ci.iloc[:, 0],
                pred_ci.iloc[:, 1], color='k', alpha=.2)
ax.set_xlabel('Date', fontsize=15)
ax.set_ylabel('Sales in Quantities', fontsize=15)
ax.set_title('The amount of Sales of Antithrombotic agents', fontsize=20)
plt.legend()
plt.show()

```



**C07 Beta blocking agents**

```

In [314]: df = df.loc[(df["ATCSShort"]=="C07")]

In [315]: df = df.sort_values('Delivery date')

In [316]: df = df.groupby('Delivery date')['Units'].sum().reset_index()

In [317]: df = df.set_index('Delivery date')

In [318]: y = df['Units'].resample('MS').mean()

In [319]:
import itertools
p = q = range(0, 2)

# Generate all different combinations of p, q and q triplets
pdq = list(itertools.product(p, q, q))

# Generate all different combinations of seasonal p, q and q triplets
seasonal_pdq = [(x[0], x[1], x[2], 12) for x in list(itertools.product(p, d, q))]

print('Examples of parameter combinations for Seasonal ARIMA...')
print('SARIMAX: {} {} {}'.format(pdq[1], seasonal_pdq[1]))
print('SARIMAX: {} {} {}'.format(pdq[1], seasonal_pdq[2]))
print('SARIMAX: {} {} {}'.format(pdq[2], seasonal_pdq[3]))
print('SARIMAX: {} {} {}'.format(pdq[2], seasonal_pdq[4]))

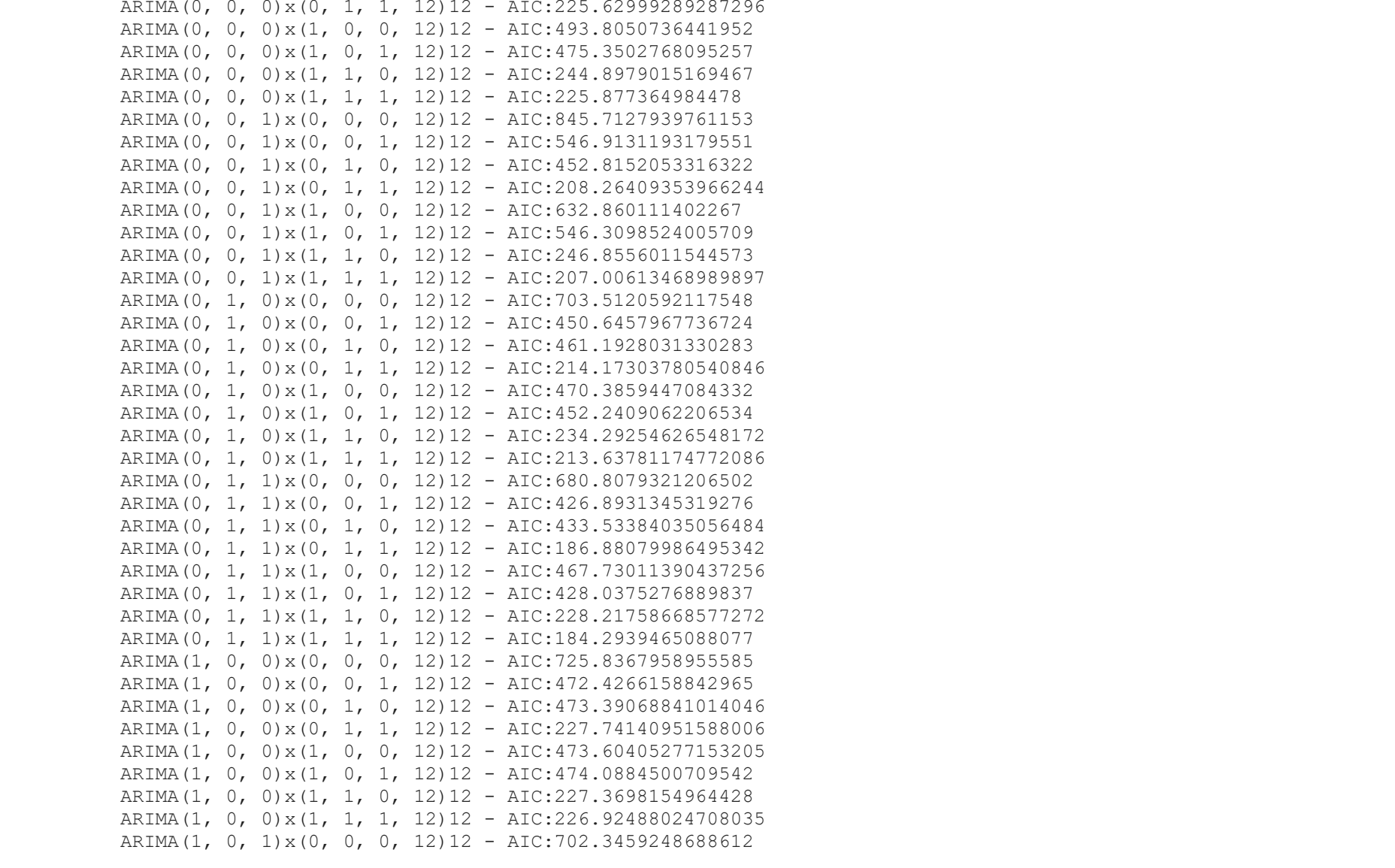
Examples of parameter combinations for Seasonal ARIMA...
SARIMAX: (0, 0, 1) x (0, 0, 1, 12)
SARIMAX: (0, 0, 1) x (0, 1, 0, 12)
SARIMAX: (0, 1, 0) x (0, 1, 1, 12)
SARIMAX: (0, 1, 0) x (1, 0, 0, 12)

In [320]:
import warnings
warnings.filterwarnings("ignore")
for param in pdq:
    for param_seasonal in seasonal_pdq:
        try:
            mod = sm.tsa.statespace.SARIMAX(y,
                                              order=param,
                                              seasonal_order=param_seasonal,
                                              enforce_stationarity=False,
                                              enforce_invertibility=False)

            results = mod.fit()

            print('ARIMA({}x){}12 - AIC:{}'.format(param, param_seasonal, results.aic))
        except:
            continue

ARIMA(0, 0, 0)x(0, 0, 0, 12)12 - AIC:1893.7657990697556
ARIMA(0, 0, 0)x(0, 0, 1, 12)12 - AIC:1566.8984057939842
ARIMA(0, 0, 0)x(0, 1, 0, 12)12 - AIC:473.76891702117746
  
```













```
pdq = list(itertools.product(p, d, q))
# Generate all different combinations of seasonal, p, q and q triplets
seasonal_pdq = [(x[0], x[1], x[2], 12) for x in list(itertools.product(p, d, q))

print('Examples of parameter combinations for Seasonal ARIMA...')
print('SARIMAX: ( ) x ( )'.format(pdq[1]), seasonal_pdq[1])
print('SARIMAX: ( ) x ( )'.format(pdq[1]), seasonal_pdq[2])
print('SARIMAX: ( ) x ( )'.format(pdq[2]), seasonal_pdq[3])
print('SARIMAX: ( ) x ( )'.format(pdq[2]), seasonal_pdq[4])
```

Examples of parameter combinations for Seasonal ARIMA...

```
SARIMAX: (0, 0, 1) x (0, 0, 1, 12)
SARIMAX: (0, 0, 1) x (0, 1, 0, 1, 12)
SARIMAX: (0, 1, 0) x (0, 0, 1, 1, 12)
SARIMAX: (0, 1, 0) x (0, 1, 0, 1, 12)
```

```
SARIMAX: (0, 1, 0) x (1, 0, 0, 12)
```

```
In [401]: for param in pdq:
          for param_seasonal in seasonal_pdq:
              try:
                  mod = sm.tsa.statespace.SARIMAX(y,
                                                    order=param,
                                                    seasonal_order=param_seasonal,
                                                    enforce_stationarity=False,
                                                    enforce_invertibility=False)

                  results = mod.fit()

                  print('ARIMA({})x{}12 - AIC:{}'.format(param, param_seasonal, results.aic))
              except:
                  continue
```

```
ARIMA(0, 0, 0)x(0, 0, 0, 12)12 - AIC:1063.5499302309097
ARIMA(0, 0, 0)x(0, 0, 1, 12)12 - AIC:1906.9803224267962
ARIMA(0, 0, 0)x(0, 1, 0, 12)12 - AIC:588.8890488266689
ARIMA(0, 0, 0)x(0, 1, 1, 12)12 - AIC:278.8682813340397
ARIMA(0, 0, 0)x(1, 0, 0, 12)12 - AIC:610.2703700723097
ARIMA(0, 0, 0)x(1, 1, 0, 12)12 - AIC:586.1678236194456
ARIMA(0, 0, 0)x(1, 1, 1, 12)12 - AIC:304.5801886697165
ARIMA(0, 0, 0)x(1, 1, 1, 12)12 - AIC:277.7323150887992
ARIMA(0, 0, 1)x(0, 0, 0, 12)12 - AIC:1010.4604788746681
ARIMA(0, 0, 1)x(0, 0, 1, 12)12 - AIC:635.7217066453591
ARIMA(0, 0, 1)x(0, 1, 0, 12)12 - AIC:559.7671207675
ARIMA(0, 0, 1)x(0, 1, 1, 12)12 - AIC:259.2564402910347
ARIMA(0, 0, 1)x(1, 0, 0, 12)12 - AIC:709.7357140828376
ARIMA(0, 0, 1)x(1, 0, 1, 12)12 - AIC:653.071855610461
ARIMA(0, 0, 1)x(1, 1, 0, 12)12 - AIC:304.5801886697165
ARIMA(0, 0, 1)x(1, 1, 1, 12)12 - AIC:254.0612377465954
ARIMA(0, 1, 0)x(0, 0, 0, 12)12 - AIC:866.750457594978
ARIMA(0, 1, 0)x(0, 0, 1, 12)12 - AIC:553.9312704759402
ARIMA(0, 1, 0)x(0, 1, 0, 12)12 - AIC:568.808917998300
ARIMA(0, 1, 0)x(0, 1, 1, 12)12 - AIC:259.71211287119325
ARIMA(0, 1, 0)x(1, 0, 0, 12)12 - AIC:577.1951271927386
ARIMA(0, 1, 0)x(1, 0, 1, 12)12 - AIC:555.569218160414
ARIMA(0, 1, 0)x(1, 1, 0, 12)12 - AIC:284.9219818006731
ARIMA(0, 1, 0)x(1, 1, 1, 12)12 - AIC:259.2063648447774
ARIMA(0, 1, 1)x(0, 0, 0, 12)12 - AIC:842.429758999764
ARIMA(0, 1, 1)x(0, 0, 1, 12)12 - AIC:526.315683915649
ARIMA(0, 1, 1)x(0, 1, 0, 12)12 - AIC:531.8195624776407
ARIMA(0, 1, 1)x(0, 1, 1, 12)12 - AIC:226.315683915649
ARIMA(0, 1, 1)x(1, 0, 0, 12)12 - AIC:575.8206728199069
ARIMA(0, 1, 1)x(1, 0, 1, 12)12 - AIC:525.0722810325161
ARIMA(0, 1, 1)x(1, 1, 0, 12)12 - AIC:279.47798111416054
ARIMA(0, 1, 1)x(1, 1, 1, 12)12 - AIC:231.8718046923213
ARIMA(1, 0, 0)x(0, 0, 0, 12)12 - AIC:893.7617401722451
ARIMA(1, 0, 0)x(0, 0, 1, 12)12 - AIC:580.3894576212374
ARIMA(1, 0, 0)x(0, 1, 0, 12)12 - AIC:584.167908286048
ARIMA(1, 0, 0)x(0, 1, 1, 12)12 - AIC:286.8336913472105
ARIMA(1, 0, 0)x(1, 0, 0, 12)12 - AIC:581.6461648834567
ARIMA(1, 0, 0)x(1, 0, 1, 12)12 - AIC:582.2953908085525
ARIMA(1, 0, 0)x(1, 1, 0, 12)12 - AIC:279.9314978893175
ARIMA(1, 0, 0)x(1, 1, 1, 12)12 - AIC:279.16340903885765
ARIMA(1, 0, 1)x(0, 0, 0, 12)12 - AIC:869.0808144867456
ARIMA(1, 0, 1)x(0, 0, 1, 12)12 - AIC:585.2266533485564
ARIMA(1, 0, 1)x(0, 1, 0, 12)12 - AIC:561.0232504396487
ARIMA(1, 0, 1)x(0, 1, 1, 12)12 - AIC:259.357898934461
ARIMA(1, 0, 1)x(1, 0, 0, 12)12 - AIC:556.5157265680233
ARIMA(1, 0, 1)x(1, 0, 1, 12)12 - AIC:557.6483628488605
ARIMA(1, 0, 1)x(1, 1, 0, 12)12 - AIC:281.2713842666883
ARIMA(1, 0, 1)x(1, 1, 1, 12)12 - AIC:259.14844434925195
ARIMA(1, 1, 0)x(0, 0, 0, 12)12 - AIC:842.5704042515516
ARIMA(1, 1, 0)x(0, 0, 1, 12)12 - AIC:555.7962716713525
ARIMA(1, 1, 0)x(0, 1, 0, 12)12 - AIC:566.0314964848517
ARIMA(1, 1, 0)x(0, 1, 1, 12)12 - AIC:259.357898934461
ARIMA(1, 1, 0)x(1, 0, 0, 12)12 - AIC:550.6397001763321
ARIMA(1, 1, 0)x(1, 0, 1, 12)12 - AIC:527.0041529330695
ARIMA(1, 1, 0)x(1, 1, 0, 12)12 - AIC:255.9910076032476
ARIMA(1, 1, 0)x(1, 1, 1, 12)12 - AIC:232.4626078080466
```

```
In [402]: import statsmodels.api as sm
mod = sm.tsa.statespace.SARIMAX(y,
                                order=(1, 1, 1),
                                seasonal_order=(0, 1, 1, 12),
                                enforce_stationarity=False,
                                enforce_invertibility=False)

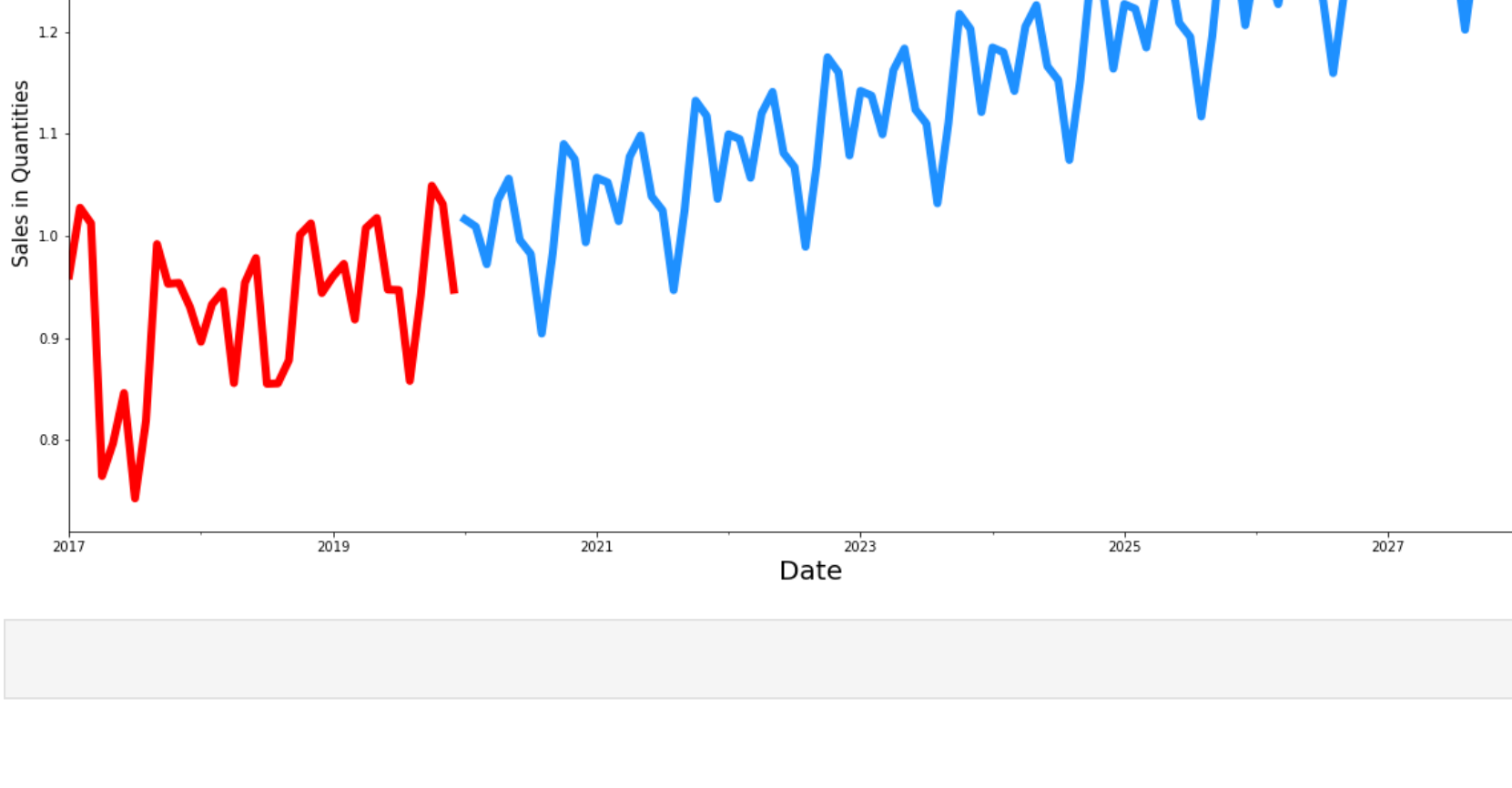
results = mod.fit()

print(results.summary().tables[1])

=====
coef      std err          z      P>|z|      [0.025      0.975]
-----
ar.L1      -0.4525      0.872      -0.519      0.604      -2.161      1.256
ma.L1      -0.7543      0.537      -1.404      0.160      -1.807      0.299
ma.S.L12    -0.1184      0.944      -0.122      0.903      -1.965      1.734
sigma2      6.185e+09      6.49e+11      9.6e+19      0.000      6.19e+09      6.19e+09
=====
```

```
In [403]: pred = results.get_prediction(start=pd.to_datetime('2018-12-01'), dynamic=False)
pred_ci = pred.conf_int()
ax = y.plot(label='Observed', figsize=(20, 10), color='v', linewidth=6)
pred_uc = predicted mean plot(ax=ax, label='Forecast', color='odgerblue', linewidth=6)
pred = predicted mean plot(ax=ax, label='One-step ahead Forecast', alpha=.7, figsize=(14, 7))

ax.fill_between(pred_ci.index,
                pred_ci.iloc[:, 0],
                pred_ci.iloc[:, 1], color='k', alpha=.2)
ax.set_xlabel('Date', fontsize=20)
ax.set_ylabel('Sales in Quantities', fontsize=15)
ax.set_title('The amount Sales', fontsize=15)
plt.legend()
plt.show()
```



```
In [404]: y_forecast = pred.predicted_mean
y_truth = y['2018-12-01:']
# Compute the mean squared error
mse = ((y_forecasted - y_truth) ** 2).mean()
print('The Mean Squared Error of our forecasts is {}'.format(round(mse, 2)))
print('The Root Mean Squared Error of our forecasts is {}'.format(round(np.sqrt(mse), 2)))

The Mean Squared Error of our forecasts is 4689380469.27
The Root Mean Squared Error of our forecasts is 68479.05
```

```
In [405]: pred_uc = results.get_forecast(steps=100)
pred_ci = pred_uc.conf_int()
ax = y.plot(label='Observed', figsize=(20, 10), color='v', linewidth=6)
ax = y.plot(predicted mean plot(ax=ax, label='Forecast', color='odgerblue', linewidth=6)
ax.set_xlabel('Date', fontsize=20)
ax.set_ylabel('Sales in Quantities', fontsize=15)
ax.set_title('The Amount of Sales', fontsize=20)
plt.legend()
plt.show()
```



```
In [ ]:
```