INDR372

Assignment 1

Disane Ketuko 0080629
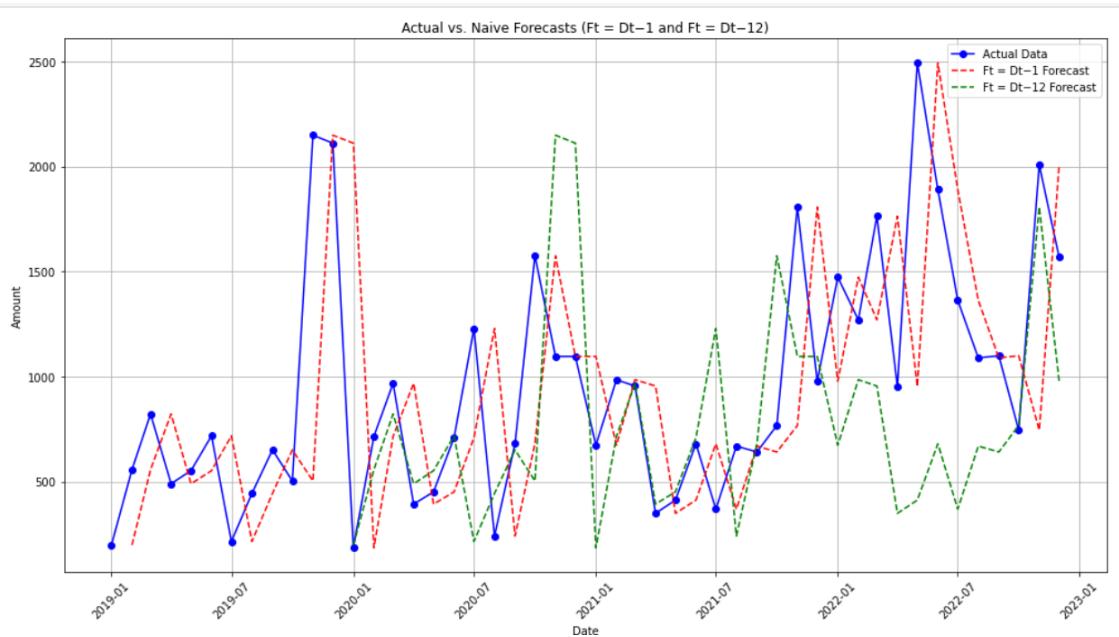
Enes Saban Tanrikulu 0079378

1.

a)

```
         Date  Amount
0   2013-01-01     228
1   2013-02-01    1641
2   2013-03-01    2984
3   2013-04-01    3045
4   2013-05-01    3161
..         ...     ...
115 2022-08-01    1089
116 2022-09-01    1099
117 2022-10-01     745
118 2022-11-01    2007
119 2022-12-01    1568
```
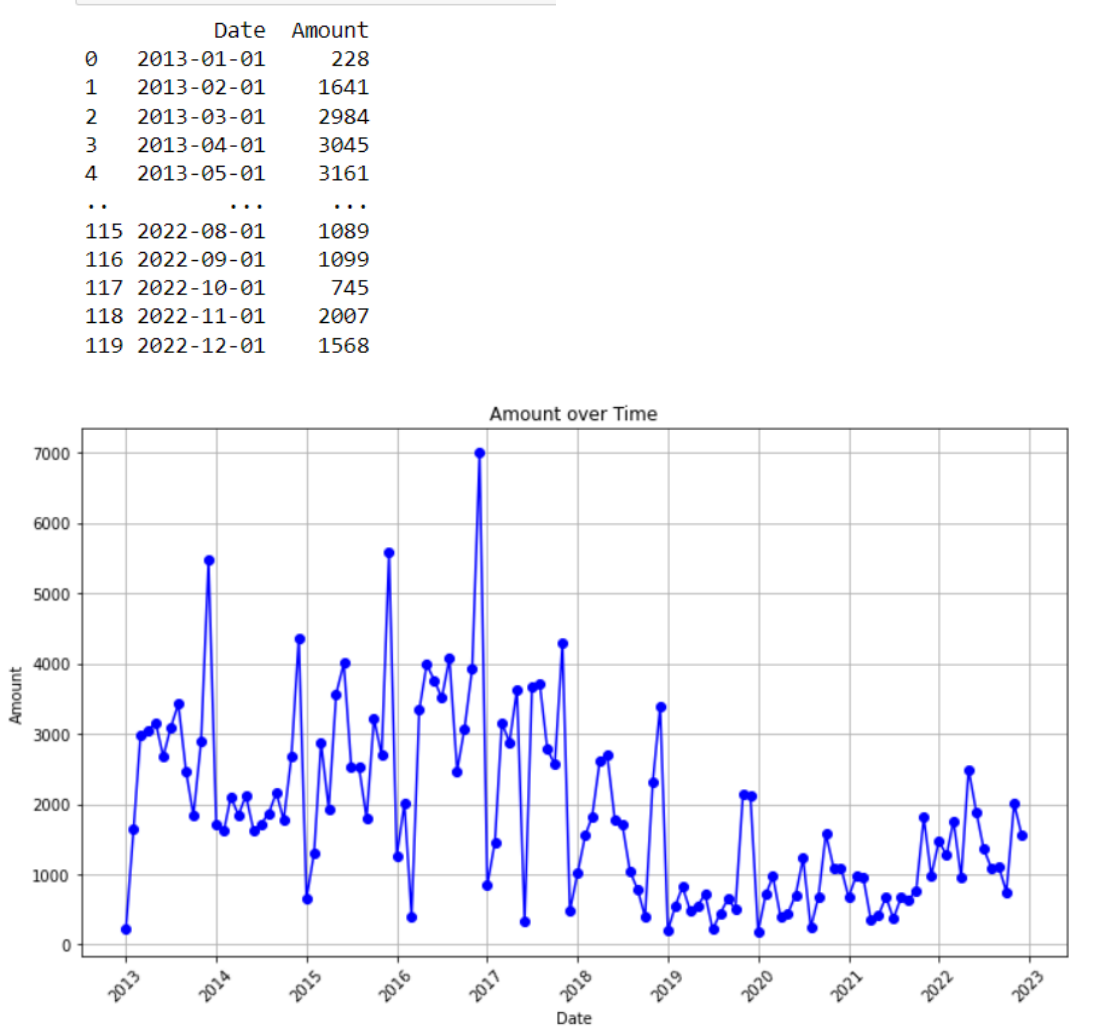

b)

```
Error Metrics for Ft = Dt-1:
MAE: 472.8085106382979
MAPE: 77.782930474433%
RMSE: 639.5153417553214

Error Metrics for Ft = Dt-12:
MAE: 478.6388888888889
MAPE: 45.70212863373995%
RMSE: 669.4362678154408
          Date  Amount    Ft_1   Ft_12
0   2013-01-01     228     NaN     NaN
1   2013-02-01    1641   228.0     NaN
2   2013-03-01    2984  1641.0     NaN
3   2013-04-01    3045  2984.0     NaN
4   2013-05-01    3161  3045.0     NaN
..         ...     ...     ...     ...
115 2022-08-01    1089  1363.0   669.0
116 2022-09-01    1099  1089.0   641.0
117 2022-10-01     745  1099.0   767.0
118 2022-11-01    2007   745.0  1807.0
119 2022-12-01    1568  2007.0   978.0

[120 rows x 4 columns]
```
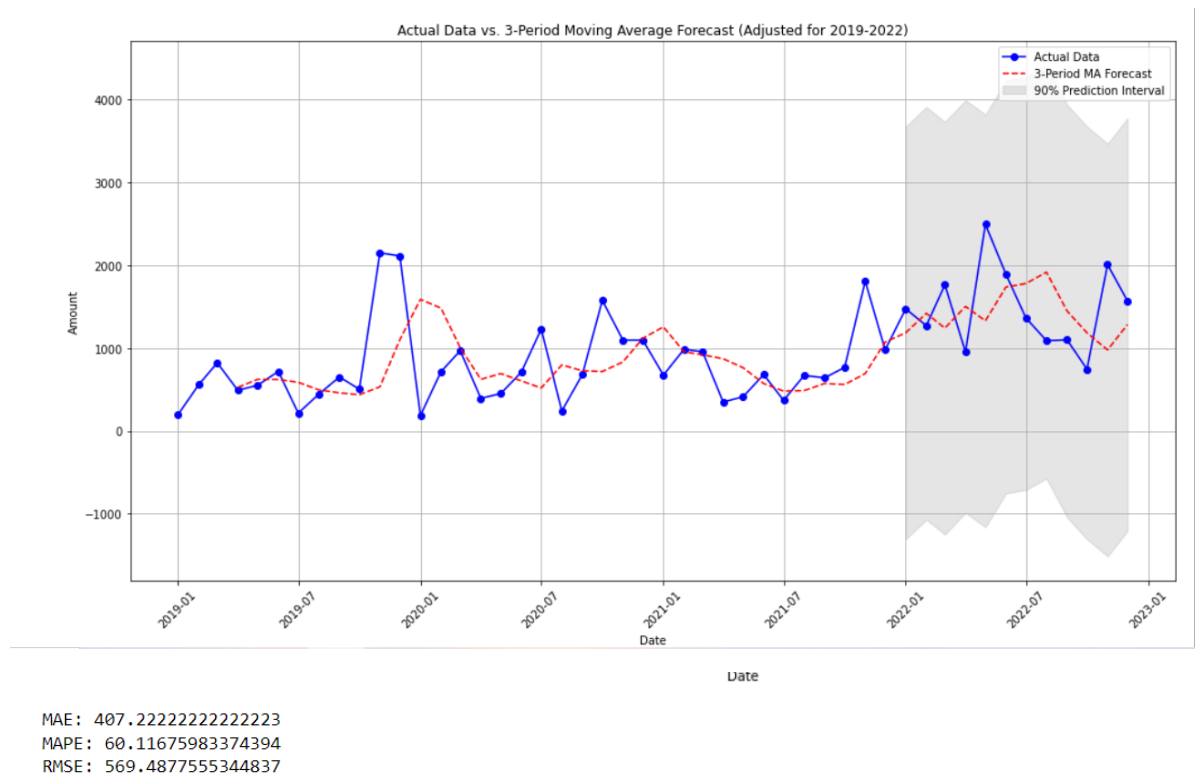
c)

```
        Date  Amount    Ft_1   Ft_12  3_MA_Forecast
12  2014-01-01   1719  5466.0   228.0    3398.333333
13  2014-02-01   1616  1719.0  1641.0    3359.666667
14  2014-03-01   2101  1616.0  2984.0    2933.666667
15  2014-04-01   1842  2101.0  3045.0    1812.000000
16  2014-05-01   2121  1842.0  3161.0    1853.000000
17  2014-06-01   1629  2121.0  2674.0    2021.333333
18  2014-07-01   1719  1629.0  3086.0    1864.000000
19  2014-08-01   1862  1719.0  3423.0    1823.000000
20  2014-09-01   2159  1862.0  2467.0    1736.666667
21  2014-10-01   1771  2159.0  1835.0    1913.333333
22  2014-11-01   2674  1771.0  2894.0    1930.666667
23  2014-12-01   4359  2674.0  5466.0    2201.333333
24  2015-01-01    653  4359.0  1719.0    2934.666667
25  2015-02-01   1309   653.0  1616.0    2562.000000
26  2015-03-01   2874  1309.0  2101.0    2107.000000
27  2015-04-01   1927  2874.0  1842.0    1612.000000
28  2015-05-01   3566  1927.0  2121.0    2036.666667
29  2015-06-01   4002  3566.0  1629.0    2789.000000
30  2015-07-01   2533  4002.0  1719.0    3165.000000
31  2015-08-01   2533  2533.0  1862.0    3367.000000
32  2015-09-01   1795  2533.0  2159.0    3022.666667
33  2015-10-01   3217  1795.0  1771.0    2287.000000
34  2015-11-01   2692  3217.0  2674.0    2515.000000
35  2015-12-01   5582  2692.0  4359.0    2568.000000
36  2016-01-01   1258  5582.0   653.0    3830.333333
37  2016-02-01   2007  1258.0  1309.0    3177.333333
38  2016-03-01    403  2007.0  2874.0    2949.000000
39  2016-04-01   3336   403.0  1927.0    1222.666667
40  2016-05-01   3983  3336.0  3566.0    1915.333333
41  2016-06-01   3759  3983.0  4002.0    2574.000000
42  2016-07-01   3507  3759.0  2533.0    3692.666667
43  2016-08-01   4073  3507.0  2533.0    3749.666667
44  2016-09-01   2471  4073.0  1795.0    3779.666667
45  2016-10-01   3074  2471.0  3217.0    3350.333333
46  2016-11-01   3926  3074.0  2692.0    3206.000000
47  2016-12-01   6998  3926.0  5582.0    3157.000000
48  2017-01-01    844  6998.0  1258.0    4666.000000
49  2017-02-01   1462   844.0  2007.0    3922.666667
50  2017-03-01   3148  1462.0   403.0    3101.333333
51  2017-04-01   2879  3148.0  3336.0    1818.000000
52  2017-05-01   3631  2879.0  3983.0    2496.333333
53  2017-06-01    323  3631.0  3759.0    3219.333333
54  2017-07-01   3673   323.0  3507.0    2277.666667
55  2017-08-01   3714  3673.0  4073.0    2542.333333
56  2017-09-01   2775  3714.0  2471.0    2570.000000
57  2017-10-01   2562  2775.0  3074.0    3387.333333
58  2017-11-01   4283  2562.0  3926.0    3017.000000
59  2017-12-01    482  4283.0  6998.0    3206.666667
60  2018-01-01   1013   482.0   844.0    2442.333333
61  2018-02-01   1552  1013.0  1462.0    1926.000000
62  2018-03-01   1821  1552.0  3148.0    1015.666667
63  2018-04-01   2617  1821.0  2879.0    1462.000000
64  2018-05-01   2694  2617.0  3631.0    1996.666667
65  2018-06-01   1770  2694.0   323.0    2377.333333
66  2018-07-01   1711  1770.0  3673.0    2360.333333
67  2018-08-01   1042  1711.0  3714.0    2058.333333
68  2018-09-01    782  1042.0  2775.0    1507.666667
69  2018-10-01    396   782.0  2562.0    1178.333333
70  2018-11-01   2304   396.0  4283.0     740.000000
71  2018-12-01   3388  2304.0   482.0    1160.666667
1513.4793890605551
```

Actual Data vs. 3-Period Moving Average Forecast (Adjusted for 2019-2022)

MAE: 407.22222222222223
MAPE: 60.11675983374394
RMSE: 569.4877555344837

d)

Residual diagnostics

Analyzing the differences between the observed and the forecasting values, we may conclude the following:

The residuals should not follow a trend or predictable pattern over time, and they should not be autocorrelated to be independent. This can be checked by Autocorrelation plots (ACF). If there is a significant correlation (not 0), there is autocorrelation. The independence between residuals implies that the structure of the data was not fully represented by the model and there might be other data to improve the forecast.

Normality of the residuals implies their normal distribution, which is important in the predictive statistics. This can be checked through graphs. When the residuals are non-normal, it means that due to skewness or outliers, the model is not accurately capturing the behavior of the data.

There are several drawbacks of this forecast with respect to the data (3 -period moving average). The 3-period moving average method overlooks data trends and seasonal variations. In the event that the dataset exhibits a notable trend or seasonality, it is plausible that this approach will not produce accurate estimations. Also, the choice of 3 period is arbitrary, so it might not be optimal for any data. For some data, other size may be better. This approach implies the linear dependence of the series on the part 3 values.

The autocorrelation or deviation from normality in the residual analysis suggests that the 3-period moving average forecast is not accurately representing the underlying process of the data. It may lead to a bias in the forecast.

e)

{0.1: (565.2874930638023, 109.58375682421445, 710.0934774901633), 0.2: (474.5560202551198, 86.77904389067473, 626.601493278401
6), 0.30000000000000004: (450.003583463817, 83.41029205853339, 619.1775786998328), 0.4: (443.5437547528093, 84.81773282707039,
629.3878153951734), 0.5: (440.5055205417084, 87.36763295113332, 647.3701053841824), 0.6: (447.3550643261603, 91.03930276356866,
669.9759051059974), 0.7000000000000001: (460.6165609744641, 95.0692163745293, 695.5387412351757), 0.8: (477.66222499180185, 99.
31596230526935, 723.0890693528531), 0.9: (495.7868277607506, 103.39516719153816, 752.1476317316474), 1.0: (529.4166666666666, 1
09.72726757103047, 782.5992216113345)}
Best Alpha: 0.30000000000000004
MAE: 450.003583463817
MAPE: 83.41029205853339
RMSE: 619.1775786998328

Actual Data and Exponential Smoothing Forecast with Prediction Intervals (2022)
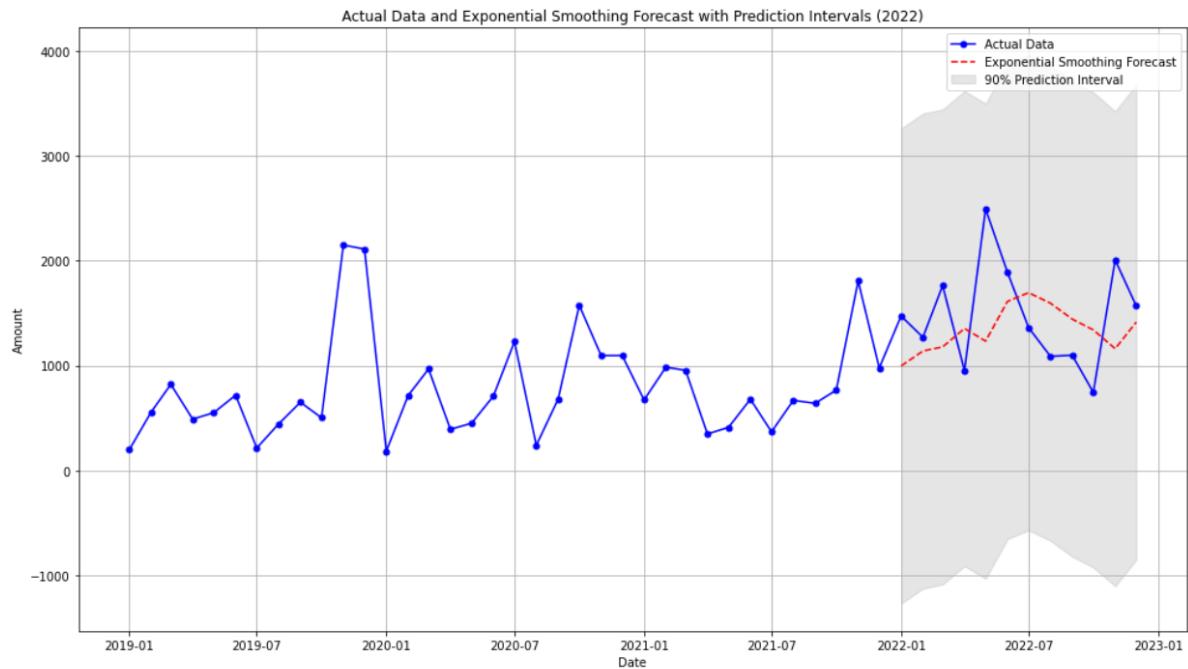
If we compare these results to MA-3 forecasts, we can see the following:

When using the exponential smoothing, the MAE (450) is higher than the one for the MA-3 (407.22). Also, MAPE of MA-3 (60.12) is significantly lower than the MAPE we have obtained (83.41). RMSE of MA-3 is lower as well (569.49 vs. 619.18).

Consequently, the MA-3 method is better to use for the forecast since the errors are lower than in the exponential smoothing.

f)

Best Alpha, Beta: (0.1, 0.30000000000000004)
MAE: 441.57860598171374
MAPE: 63.7431408042634
RMSE: 607.89165324096

Actual Data and Exponential Smoothing Forecast with Prediction Intervals (2022)
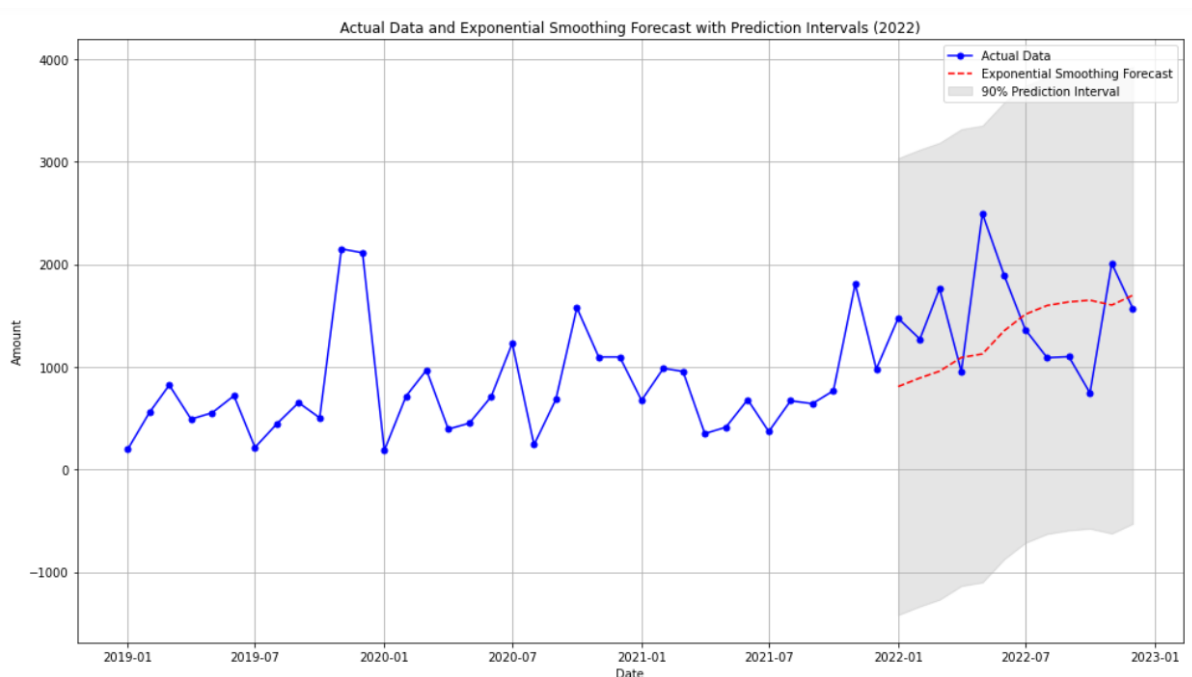
g) Residual diagnostics

Residuals should be independent, without autocorrelation in the well-fitting model. If this condition is not satisfied, it means that the model could not capture underlying patterns or seasonal effects. Again, for this ACF plots are used. In the double exponential smoothing, autocorrelation may happen because of cycles in patterns.

For normality, there should be normal distribution. Deviations from normality may indicate the existence of outliers, incorrect model specification, or the need for data manipulations or another model.

Compared to the previous forecasts, double exponential smoothing gives prediction with higher accuracy due to the trend accounting. It depends on the significance of the trend and seasonal patterns. For example, if the data has a significant trend, double exponential smoothing usually results in notable gains over simple and naive moving average forecasts. Or, for data with strong seasonal patterns, seasonal methods may perform better than non-seasonal techniques like double exponential smoothing.

There are drawbacks of this forecast. The double exponential smoothing does not capture seasonality, only trends. This may lead to errors in the forecast when it comes to fluctuations in the seasonality. Also, the values of smoothing parameters are very significant in this approach. Double exponential smoothing uses the linearity in trends assumption, which may not be applicable for all the data, which might lead to a bias.

h)

```
[1473 1270 1763  953 2493 1890 1363 1089 1099]
[1888.62758818 1939.73405448 1990.84052078 2041.94698708 2093.05345337
 2144.15991967 2195.26638597 2246.37285227 2297.47931857]

3-Month Ahead Forecast Errors:
MAE: 693.8193526248082, MAPE: 57.11356056015523%, RMSE: 785.4293431491234

6-Month Ahead Forecast Errors:
MAE: 611.5888687327514, MAPE: 47.80969148640728%, RMSE: 697.195652161656
```

If we compare the results for 3-month and 6-month ahead forecast, we see the following:

The MAE for the 6-month is lower than the one of the 3-month, which may imply that it could be affected by certain trends or patterns in the 6-month forecast. MAPE is also lower for the 6 month ahead forecast, showing that in this case the difference between the actual values and the forecast ones was not great. RMSE of the 6 month ahead forecast is lower as well. So, 3-month forecast is worse than the 6-month for usage.

i)
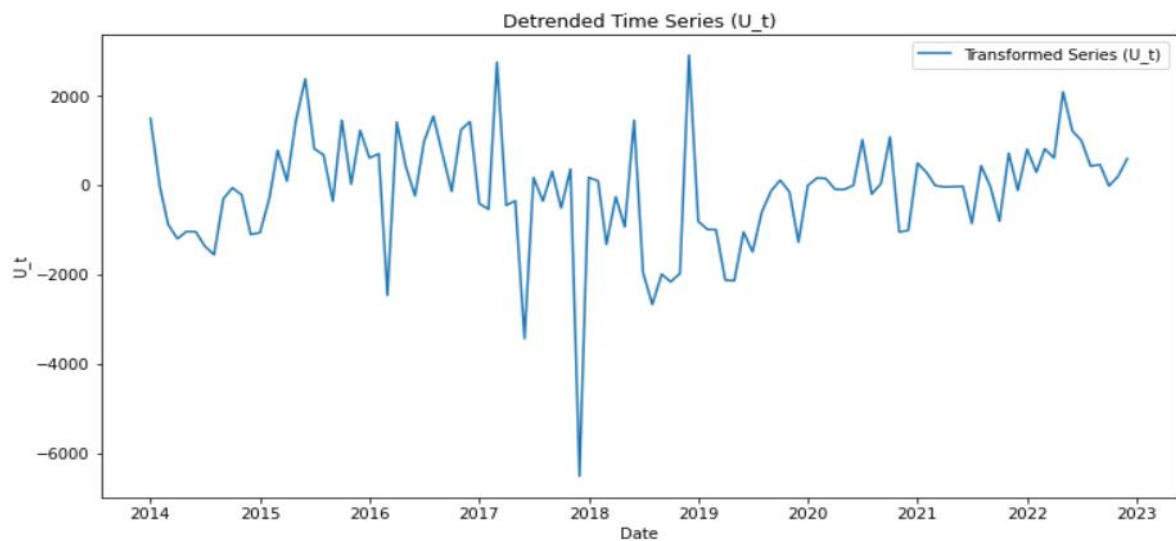

```
Best Alpha: 0.1
   Alpha        MAE        MAPE        RMSE
0    0.1  509.55212   84.916257  693.012605
```

j)

Method	Spec.	RMSE	MAPE
Benchmark-1	-	639.52	77.78
Benchmark-2	-	669.44	45.7
MA-3	-	569.49	60.11
ES	-	619.18	83.41
DES	-	607.89	63.74
Seasonal	$\alpha^* = 0.1$	693.01	84.92

2.

a)
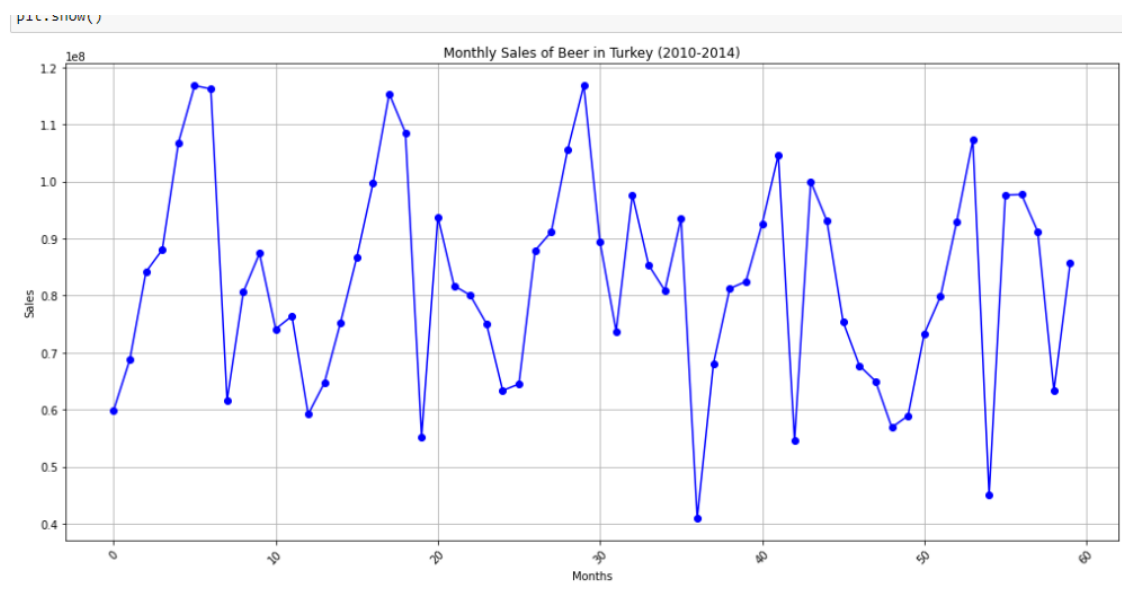
b)

```
Naive Forecast Ft=Dt-1 Error Metrics:
MAE: 17287868.297872342
MAPE: 24.435808458302834
RMSE: 23411455.74383115

Naive Forecast Ft=Dt-12 Error Metrics:
MAE: 10371271.972222222
MAPE: 14.503145197801526
RMSE: 13455891.471328583
```

	Month	Sales	Ft_1	Ft_12
0	1	59832432	NaN	NaN
1	2	68727007	59832432.0	NaN
2	3	84126559	68727007.0	NaN
3	4	88028614	84126559.0	NaN
4	5	106798913	88028614.0	NaN
5	6	116806052	106798913.0	NaN
6	7	116201586	116806052.0	NaN
7	8	61627436	116201586.0	NaN
8	9	80596102	61627436.0	NaN
9	10	87376160	80596102.0	NaN
10	11	74128665	87376160.0	NaN
11	12	76384870	74128665.0	NaN
12	13	59163840	76384870.0	59832432.0
13	14	64667468	59163840.0	68727007.0
14	15	75259960	64667468.0	84126559.0
15	16	86660844	75259960.0	88028614.0
16	17	99679555	86660844.0	106798913.0
17	18	115365778	99679555.0	116806052.0
18	19	108561132	115365778.0	116201586.0
19	20	55213584	108561132.0	61627436.0
20	21	93779894	55213584.0	80596102.0
21	22	81713811	93779894.0	87376160.0
22	23	80030855	81713811.0	74128665.0
23	24	75063042	80030855.0	76384870.0
24	25	63335875	75063042.0	59163840.0
25	26	64486104	63335875.0	64667468.0
26	27	87877077	64486104.0	75259960.0
27	28	91154166	87877077.0	86660844.0
28	29	105468418	91154166.0	99679555.0
29	30	116863150	105468418.0	115365778.0
30	31	89386738	116863150.0	108561132.0
31	32	73699601	89386738.0	55213584.0
32	33	97654393	73699601.0	93779894.0
33	34	85334611	97654393.0	81713811.0
34	35	80859020	85334611.0	80030855.0
35	36	93530185	80859020.0	75063042.0
36	37	40914893	93530185.0	63335875.0
37	38	68077231	40914893.0	64486104.0
38	39	81200746	68077231.0	87877077.0
39	40	82443643	81200746.0	91154166.0
40	41	92511113	82443643.0	105468418.0
41	42	104652765	92511113.0	116863150.0
42	43	54489655	104652765.0	89386738.0
43	44	99977203	54489655.0	73699601.0
44	45	93054507	99977203.0	97654393.0
45	46	75349679	93054507.0	85334611.0
46	47	67696034	75349679.0	80859020.0
47	48	64944154	67696034.0	93530185.0
48	49	56864523	64944154.0	40914893.0
49	50	58901848	56864523.0	68077231.0
50	51	73276957	58901848.0	81200746.0
51	52	79806180	73276957.0	82443643.0
52	53	92897565	79806180.0	92511113.0
53	54	107218153	92897565.0	104652765.0
54	55	45084054	107218153.0	54489655.0
55	56	97621205	45084054.0	99977203.0
56	57	97708872	97621205.0	93054507.0
57	58	91207167	97708872.0	75349679.0
58	59	63310816	91207167.0	67696034.0
59	60	85736906	63310816.0	64944154.0

c)

Residual diagnostics

In this model, the residuals have to be independently distributed (no correlation). However, residuals may show patterns or autocorrelation, which may be caused by the naïve forecasts not capturing seasonality and trends. ACF plots is the best way to check the independence. If there is a significant correlation (nonzero), then there is independence.

Residuals are also normally distributed but if the data is non-linear or seasonality is significant, then it might seem that the distribution is not normal. This can also be checked by plots or tests.

There are some drawbacks of the naïve forecast. Although naive forecasts are simple to use, there may be a disadvantage to their simplicity. Systematic errors may be caused by the inability to recognize complex trends in the data. Residuals may display patterns if there are strong trends or seasonality, not counted by the naïve method, demonstrated.

Above the very recent past, naive forecasts are unable to model the dynamics of the series. This may lead to residuals showing the model failing to capture significant elements of the data, like shifting seasonal impacts or cyclic trends.

d)

```
Best parameters: {'alpha': 0.1, 'beta': 0.7000000000000001, 'gamma': 0.9, 'MAE': 8012365.022485447, 'MAPE': 11.211432577755463,
'RMSE': 12034116.834749397}
Prediction intervals for 2014:
                Forecast      Lower_PI      Upper_PI
Month
2014-01-31   2.710364e+07  -4.812051e+06  5.901934e+07
2014-02-28   1.043677e+08   7.245203e+07  1.362834e+08
2014-03-31   1.110351e+08   7.911945e+07  1.429508e+08
2014-04-30   6.334634e+07   3.143064e+07  9.526203e+07
2014-05-31   5.741713e+07   2.550144e+07  8.933282e+07
2014-06-30   1.067545e+08   7.483885e+07  1.386702e+08
2014-07-31   3.759859e+07   5.682894e+06  6.951428e+07
2014-08-31   9.562636e+07   6.371066e+07  1.275420e+08
2014-09-30   1.115143e+08   7.959863e+07  1.434300e+08
2014-10-31   1.342780e+08   1.023623e+08  1.661937e+08
2014-11-30   8.236800e+07   5.045230e+07  1.142837e+08
2014-12-31   5.379159e+07   2.187590e+07  8.570728e+07
```


Beer Sales Forecast with Prediction Intervals for 2014

e)
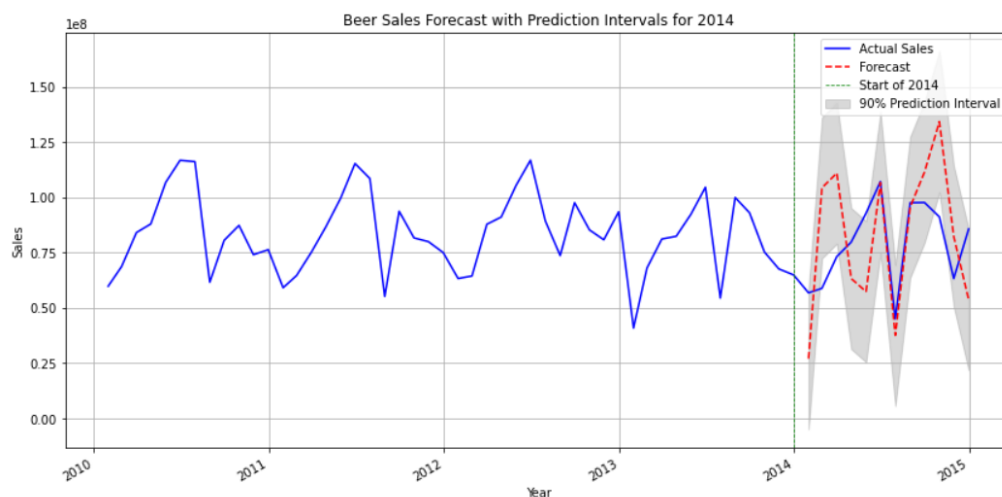
```
3-Month Ahead Forecast Error Metrics:
MAE: 11032921.387144197, MAPE: 14.178759873920823%, RMSE: 11603640.50274775

6-Month Ahead Forecast Error Metrics:
MAE: 11567395.957734277, MAPE: 15.10299102626012%, RMSE: 12299606.12902377
```

The MAE, MAPE, and RMSE for the 6-month ahead forecast are higher than the ones for the 3-month ahead forecast. It means that in this model, it is better to use shorter time interval.

f)

Method	Spec.	RMSE	MAPE
Benchmark-1	-	23411455.74	24.44
Benchmark-2	-	13455891.47	14.5
TES	α=0.1, β= 0.7, γ= 0.9	12034116.84	11.2

Method	Spec.	RMSE	MAPE