Topic: Forecasting – Time Series

Forecast the Coca-Cola prices and Airlines Passengers data set. Prepare a document for each model explaining how many dummy variables you have created and RMSE value for each model. Finally which model you will use for Forecasting.

1.) Airlines.xlsx

	Month	Passengers
1	1995-01-01	112
2	1995-02-01	118
3	1995-03-01	132
4	1995-04-01	129
5	1995-05-01	121
6	1995-06-01	135
7	1995-07-01	148
8	1995-08-01	148
9	1995-09-01	136
10	1995-10-01	119
11	1995-11-01	104
12	1995-12-01	118
13	1996-01-01	115
14	1996-02-01	126
15	1996-03-01	141
16	1996-04-01	135

2.) CocaCola_Sales_RawData.xlsx

	Quarter	Sales
1	Q1_86	1734.827
2	Q2_86	2244.961
3	Q3_86	2533.805
4	Q4_86	2154.963
5	Q1_87	1547.819
6	Q2_87	2104.412
7	Q3_87	2014.363
8	Q4_87	1991.747
9	Q1_88	1869.050
10	Q2_88	2313.632
11	Q3_88	2128.320
12	Q4_88	2026.829
13	Q1_89	1910.604
14	Q2_89	2331.165
15	Q3_89	2206.550
16	Q4_89	2173.968
17	Q1_90	2148.278

Hints:

1. Business Problem

    1.1. Objective
    1.2. Constraints (if any)

2. Data Pre-processing

    2.1 Feature Engineering, EDA etc.

3. Model Building

    3.1 Partition the dataset
    3.2 Model(s) – Work with all the models (linear, exponential, quadratic etc.)
    3.3 Model(s) Improvement steps
    3.4 Model Evaluation
    3.5 Python and R codes

4. Result Share the benefits/impact of the solution - how or in what way the business (client) gets benefit from the solution provided.

Note:

1. For each assignment the solution should be submitted in the format

2. Research and Perform all possible steps for improving the model(s) accuracy & reduce the RMSE (also evaluate errors like MAPE, MAE etc.)

3. All the codes (executable programs) are running without errors

4. Documentation of the module should be submitted along with R & Python codes, elaborating on every step mentioned here