2nd List of Time Series – Part 1
2023-1
Deadline: 16/07/2023
Individual List

Do not use ready-made Matlab and Python functions, except when informed on the question. Program your functions and present them on the report. You can use pandas, numpy and matplotlib in Python. Include the codes and obtained images in the report. Write the report in Portuguese or English. Send the report through the Google Classroom of the discipline.

1) Milk.txt is a time series of monthly milk production per cow with values from 01/1962 to 12/1975. Separate the data into two parts: a time series from 01/1962 to 12/1973 and another time series from 01/1974 to 12/1975. Perform the decomposition of the first part using STL decomposition and show the graphics. After that, predict the production for the years 1974 to 1975 using the seasonality component of the STL and the autoregressive (AR) model to predict the trend + residual components. Show the graphics of actual values and predicted values. Calculate the MSE of the prediction. Comment your solution. In Python, the following functions can be used:
from statsmodels.tsa.seasonal import STL
from statsmodels.tsa.ar_model import AutoReg
In Matlab, the functions trenddecomp and arima can be used.

2) For the industrial_process.txt time series, do: convert the time series values to integer values, i.e., use the round function. Then, you must compress the data by combining delta encoding and Fibonacci binary encoding methods. Calculate the compression ratio (assume that the original time series values are represented by 32 bits) and the MSE of the reconstruction error. Show the graphics for the original time series and the reconstructed time series. What consequence on MSE and compression ratio if the data were not rounded to integer values?

3) Design the Huffman coding for the symbols of the table. Calculate the compression ratio (assume that the symbols are represented by 8 bits). This question can be done manually.

Symbols	Probability
A	0.10
B	0.15
C	0.20
D	0.25
E	0.07
F	0.23