Data Structure & Algorithms Project

Phase 1

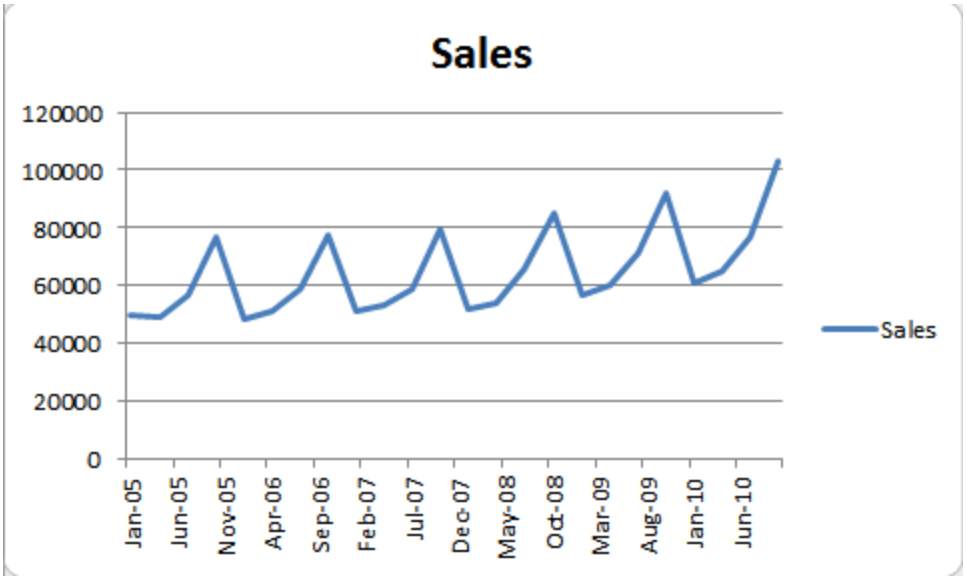
- In this phase (1) you will understand what a time series is and you'll implement a brute force algorithm to find time series motifs.

Prerequisites:

1. Time Series:

   A time series is a series of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time.

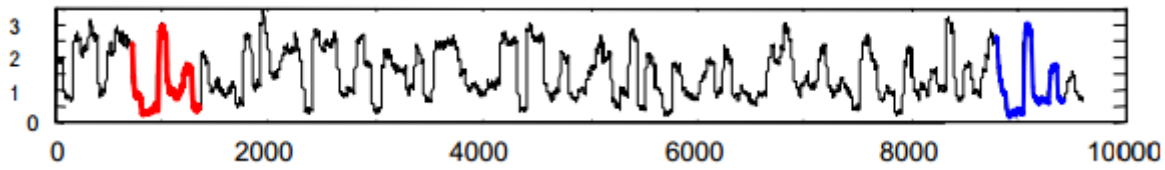
   Example: a time series showing the amount of sales for a product



2. Time Series Motifs:

Time series motifs are pairs of individual time series, or subsequences of a longer time series, which are very similar to each other.

Example:



Phase 1 tasks:

1. Understand what 'time series' and 'time series motifs' are.
2. What is the difference between *match* and *trivial match* in a time series?
3. What is *subsequence motif* in a time series?
4. Make 2 datasets for time series, one of them should contain some signals (2D points), and the other one should contain some strings (e.g. ABACACBSAJASBABA)
5. Write a brute-force algorithm to detect time series motifs with len K and test your code on your datasets. In your report you must include your results with $K = 2, 5, 6$.
6. What is your program's output for this data with $K = 4$?

sjdbbnvfdfpqoeutyvnABABABmbzchslfkeruyousjdq

7. How much time it takes for your algorithm to find motifs? (your can plot a histogram for each data)

Notes:

- Your implementation should be functional
- Any sign of cheating will result in the **zero** grade
- Your report should be a single PDF file contaning answers to questions and your results
- You should upload your codes and report in a single ZIP file named ‘STD ID 1 – STD ID 2.zip’ (e.g. ‘902717 – 9433894.zip’)

GoodLuck ☺