A36903

Calculators may be used in this examination provided they are not capable of being used to store alphabetical information other than hexadecimal numbers

UNIVERSITY OF BIRMINGHAM

School of Computer Science

Fourth Year Undergraduate/Postgraduate

06-32254

LM Visualisation

Main Summer Examinations 2025

Time allowed: 2 hours

[Answer all questions]

Turn Over

Note

Answer ALL questions.

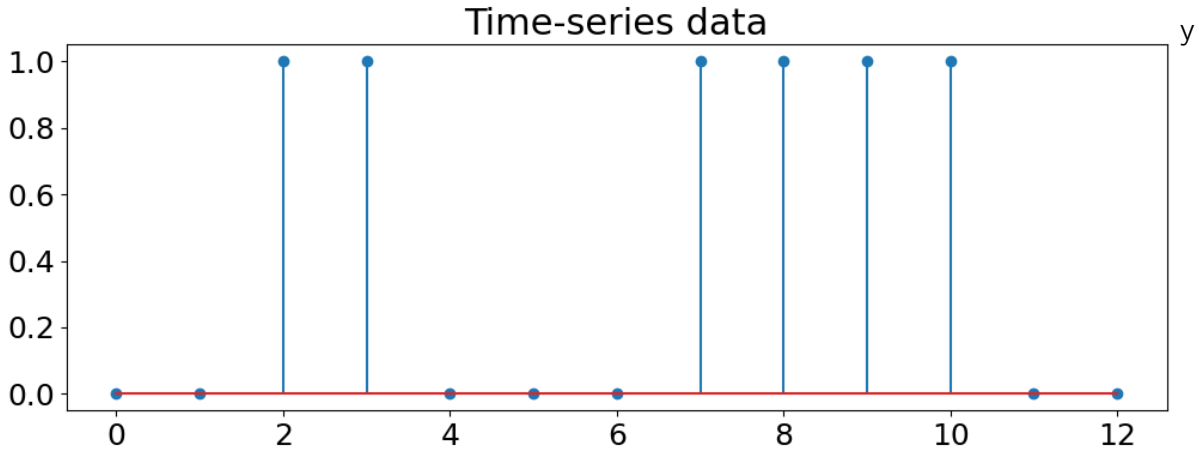
Question 1

(a) Consider a long time-series f that has all zero values except for n adjacent samples in the center, which have positive non-zero values. You convolve it with a kernel function k of width w of also all positive non-zero values. Give an expression for the width of the convolution product $f * k$, that is the number of samples in $f * k$ that are non-zero. Keep your answer very brief. **[5 marks]**

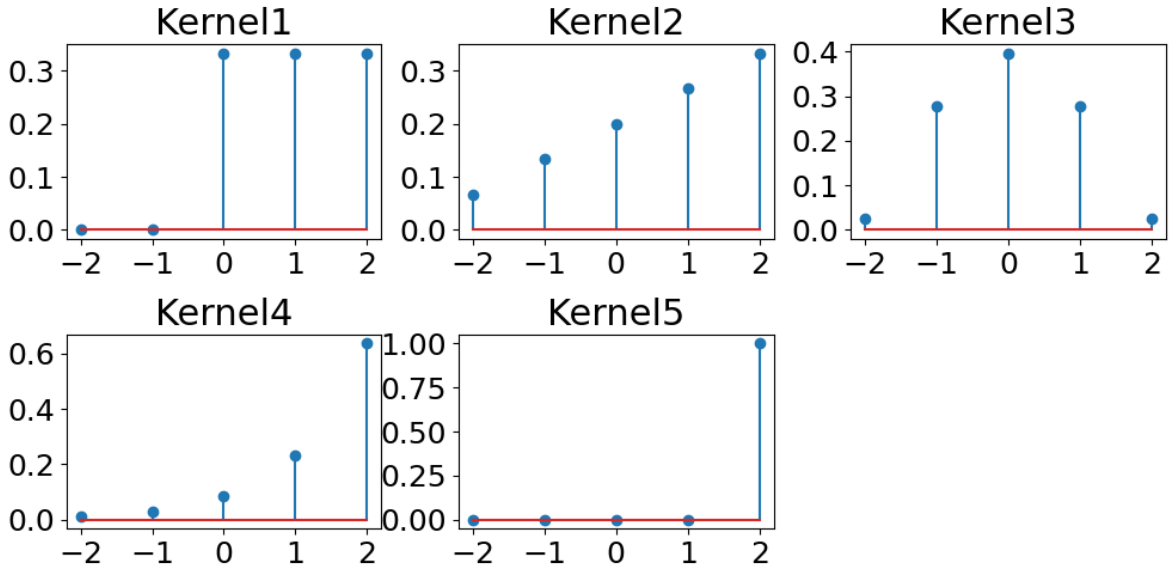
(b) Given three time-series functions f, g, and h, you need to compute $(f * h) + (g * h)$. Since you will be doing the computation manually, how could you reduce the time spent to compute the result? Justify your approach. Keep your answer very brief. **[5 marks]**

(c) Consider the time-series plotted in Figure 1a. This time-series has been convolved with five different kernel functions labeled 'Kernel 1' through 'Kernel 5', shown in Figure 1b. The convolution products of those operations are labeled A through E, shown in Figure 1c. Match each kernel function in Figure 1b to the convolution product it produced in Figure 1c. Also give a very brief justification for how you arrived at each match. Keep your answer very brief. **[30 marks]**

(a) Time-series



(b) Kernel functions 1–5



(c) Convolution product functions A–E

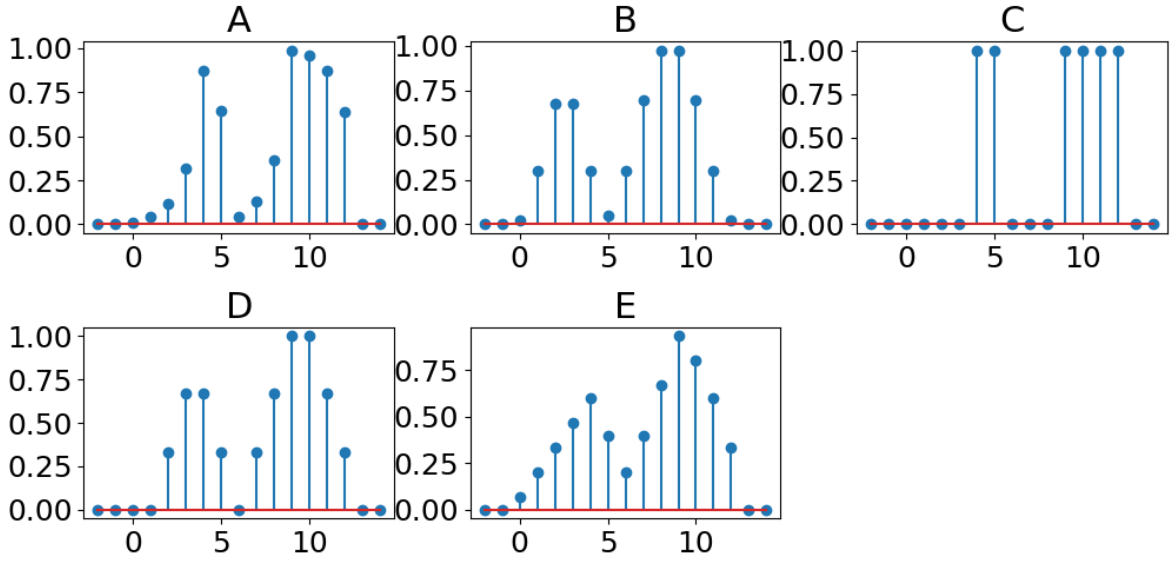


Figure 1: For Question 1 on convolutions.

Question 2

(a) Please answer 'True' or 'False' to the following statements and give a brief explanation. **[9 marks]**

(i) The data is localized on a spheroid in k-dimensional space. We can use PCA to reduce its dimensionality to a lower dimension for visualization.

(ii) Using PCA to reduce dimensionality of data always loses information.

(iii) If you rotate the sample points in high dimension space, the eigen value corresponding to the first principle component will not change.

(b) Given the following covariance matrix, what will be the first principal component? What does its direction mean? **[2 marks]**

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(c) You are given a dataset with 5000 samples and 100 features. You decide to use Principal Component Analysis (PCA) for dimensionality reduction.

(i) Very briefly explain the main steps of PCA and how it reduces dimensionality. **[5 marks]**

(ii) How can you determine the number of principal components to retain? Name and very briefly explain two methods. **[8 marks]**

(iii) What are the advantages and limitations of using PCA? Name two of each, very briefly explain. **[8 marks]**

Question 3

Figure 2 shows the daily revenue of a small bakery business as well as the Fourier spectrum of the data with the frequency axis showing repetitions per month, with one month corresponding to 28 days.

(a) In the spectrum plot, what could be the origin of the different peaks? What is the period of the oscillations they correspond to? How can we compute it? Keep your answer very brief. **[4 marks]**

(b) What is the average daily revenue computed over the entire year? How can it be determined from the plots? Keep your answer very brief. **[4 marks]**

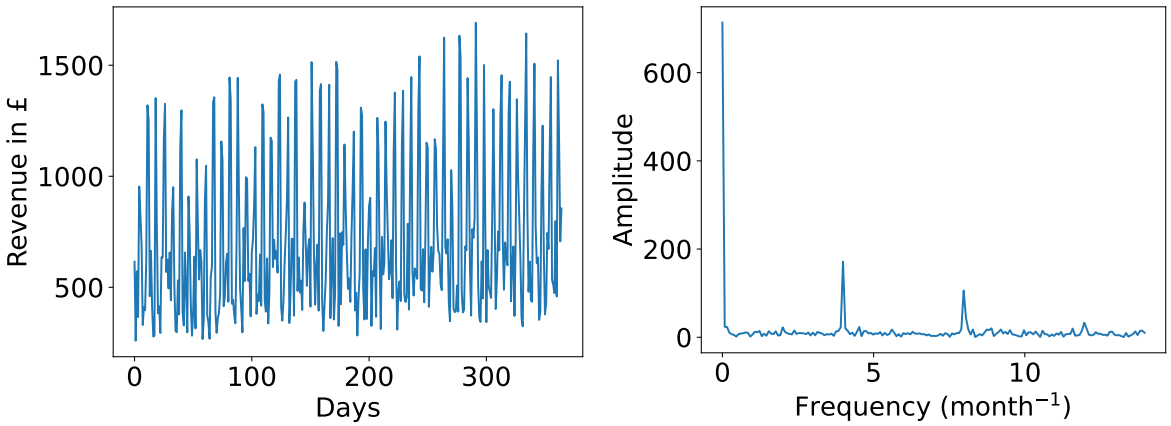


Figure 2: For Question 3 on time series.

(c) In order to better analyze the data we want to smooth the time series by convolving it with box car filter kernel. What is the height and width we should choose to completely remove the oscillations in the time series? Roughly copy the empty chart from Figure 3 (you do not have to include all the numbers and ticks) into the answer book and sketch the kernel. **[6 marks]**

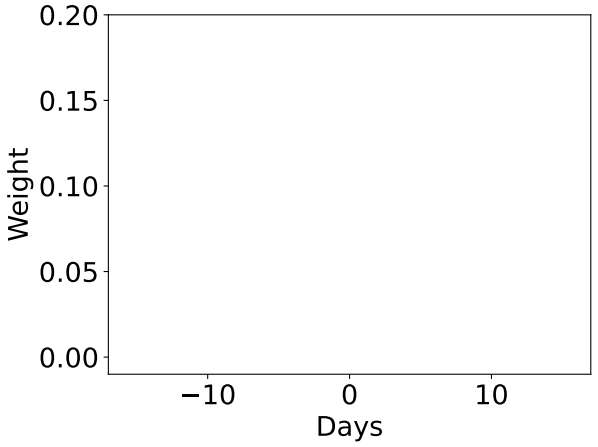


Figure 3: For Question 3 on time series.

(d) Roughly copy the chart from Figure 4 (you do not have to include all the numbers and ticks) into the answer book. Sketch the real part of the filter coefficients for the kernel from (c) in frequency space. It is not important to be perfectly accurate, but to indicate for which frequencies the curve is positive, negative and exactly zero. **[8 marks]**

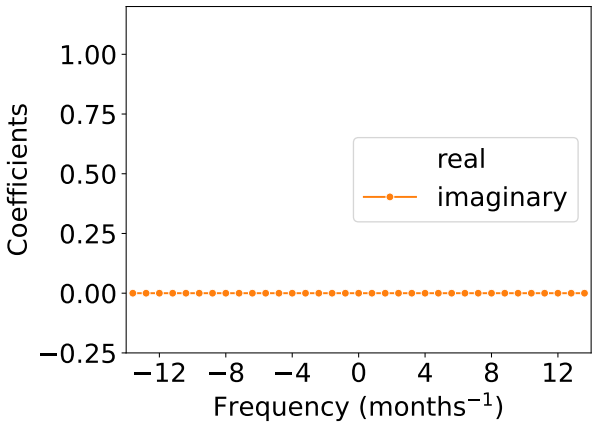


Figure 4: For Question 3 on time series.

(e) How could the chart from Figure 5 be modified to follow the design principles according to Edward Tufte. Keep your answer very brief. **[6 marks]**

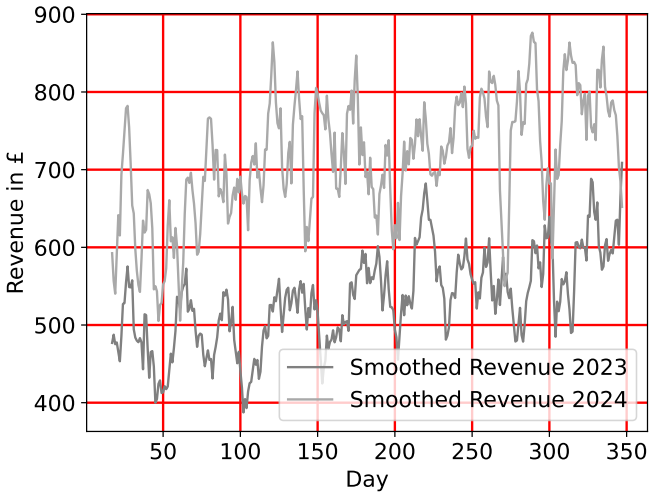


Figure 5: For Question 3 on design principles.

This page intentionally left blank.

Do not complete the attendance slip, fill in the front of the answer book or turn over the question paper until you are told to do so

Important Reminders

- Coats/outwear should be placed in the designated area.
- Unauthorised materials (e.g. notes or Tippex) must be placed in the designated area.
- Check that you do not have any unauthorised materials with you (e.g. in your pockets, pencil case).
- Mobile phones and smart watches **must** be switched off and placed in the designated area or under your desk. They must not be left on your person or in your pockets.
- You are not permitted to use a mobile phone as a clock. If you have difficulty seeing a clock, please alert an Invigilator.
- You are not permitted to have writing on your hand, arm or other body part.
- Check that you do not have writing on your hand, arm or other body part – if you do, you must inform an Invigilator immediately
- Alert an Invigilator immediately if you find any unauthorised item upon you during the examination.

Any students found with non-permitted items upon their person during the examination, or who fail to comply with Examination rules may be subject to Student Conduct procedures.