

PREPARED FOR SUBMISSION TO UNIVERSITY OF CAMBRIDGE

# **M1: Applied Data Science - Coursework Assignment**

**Adnan Siddiquei**

University of Cambridge

E-mail: [as3438@cam.ac.uk](mailto:as3438@cam.ac.uk)

---

## Contents

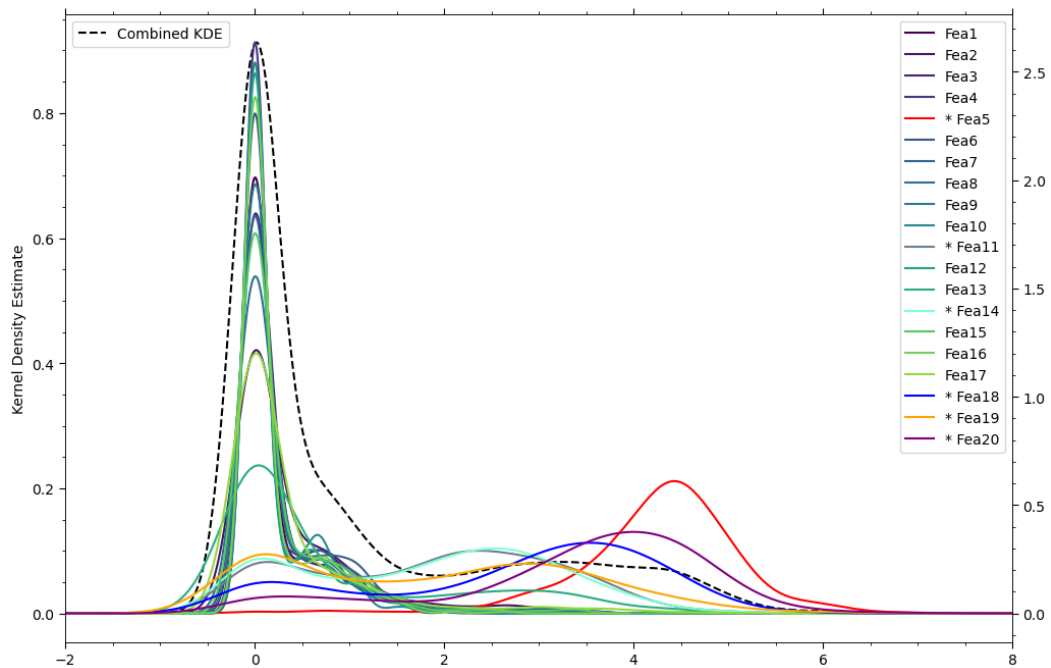
<b>1</b>	<b>Section A</b>	<b>1</b>
1.1	Dataset A	1
1.2	Dataset A	2
1.3	Dataset A	2
<b>2</b>	<b>Section B</b>	<b>2</b>

---

## 1 Section A

This section contains the answers to the questions in Section A of the coursework assignment. Wherever the phrase 'dataset' is used in this section, it refers to `A_NoiseAdded.csv`.

### 1.1 Dataset A



**Figure 1:** A kernel density estimate of the first 20 features of the dataset, which has 20 plots in total, with the legend and corresponding y-axis on the right. A combined KDE has also been plotted, with its corresponding x-axis on the left. 6 of the 20 features have been highlighted in the legend with an asterisk, and have been coloured in slightly more contrasting colours to highlight them. These features contain more density in the non-zero region and are likely to be more discriminative.

**1.2 Dataset A**

**1.3 Dataset A**

**2 Section B**

**References**