Data Mining Assignment 1

Identify a problem from your own experience that you think would be amenable to data mining. For that problem describe:

1. What the data is.

**The unnecessary emails we get everyday from a website we would have logged in some point of our life. The website sends us numerous spam emails, even though we consider them unimportant. This type of data piles up and get huge.**

2. What type of benefit you might hope to get from data mining.

**To identify patterns in data and find ways to manage data to help us make good decisions and improve methods of storing and retrieving data.**

3. What type of data mining (classification, clustering, etc.) you think would be relevant.

**Classification is one such technique that would be relevant to deal with the span emails types of data.**

4. Name one type of data mining that you think would not be relevant, and describe briefly why not. For each, illustrate with an example, e.g., if you think clustering is relevant, describe what you think a likely cluster might contain and what the real-world meaning would be.

**Regression is a type of data mining technique which is not relevant, as it deals with only numerical values and to predict the future values. Hence, it may not be helpful in classifiable data types.**