# Data Mining Assignment 1

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

**Problem Description**: Many E-commerce companies use Data Mining and Business Intelligence to offer **cross-sells and up-sells** through their websites. One of the most famous of these is, of course, Amazon, who use sophisticated mining techniques to drive their, ‘People who viewed that product, also liked this’ functionality.

1. What the data is.

## Ans: By running deeper analysis on the customer information we get the data

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

Ans: Getting more from what we have is the key tenet of the cross-selling and up selling mantra. From the data, we take these new insights and get more revenue from the existing customer-base by offering them relevant products and services. It’s a win-win process, when it works.

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

Ans: Analysis and and segmentation of customer data is required. The data set must first be divided into segments or cohorts, based along shared attributes, such as average spend, age, location or gender.

To successfully segment a data set requires detailed data mining techniques like associative analysis, classification and clustering to correctly decide how and where the segments should be created.

4. Name one type of data mining that you think would not be relevant, and describe briefly why not.

Ans: Regression since we don’t have any continuous data. We want our data to get divided into segments which is not possible through regression

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.

Write one to two pages of 11 point single-spaced typeset text - you aren't writing a paper, but it isn't short answer either.