**DATA ASSIGNMENT 1**

1. What the data is.

Grocery shopping.

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

From the data mining, I would know on which product am spending more money and which product is more necessary to buy for the house rather than spending on unnecessary things. I would also get to know how much am spending for shopping and how can I save some money instead of spending on unnecessary shopping. I can also know which products are being brought frequently so that I will be helpful for me when I go for shopping next time.

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

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

Regression wouldn’t be relevant because Regression analysis is a reliable method of identifying which variables have impact on a topic of interest. The process of performing a regression allows you to confidently determine which factors matter most, which factors can be ignored, and how these factors influence each other.

First, regression analysis is widely used for prediction and forecasting, where its use has substantial overlap with the field of machine learning. Second, in some situation’s regression analysis can be used to infer causal relationships between the independent and dependent variables.

We use [regression analysis](https://statisticsbyjim.com/glossary/regression-analysis/) to describe the relationships between a set of [independent variables](https://statisticsbyjim.com/glossary/predictor-variables/) and the [dependent variable](https://statisticsbyjim.com/glossary/response-variables/). [Regression analysis](https://statisticsbyjim.com/glossary/regression-analysis/) produces a [regression](https://statisticsbyjim.com/glossary/regression-analysis/) equation where the [coefficients](https://statisticsbyjim.com/glossary/regression-coefficient/) represent the relationship between each [independent variable](https://statisticsbyjim.com/glossary/predictor-variables/) and the dependent variable.