Data Mining Assignment 1

1.What the data is.

Data, in the context of computing, refers to distinct pieces of digital information. Data is usually formatted in a specific way and can exist in a variety of forms, such as numbers, text, etc.

Data mining can be defined as a computer-assisted process of digging through and analyzing enormous sets of data and then extracting the meaning of the data.

The data is about “Reliance Fresh” super market near my house.

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

Data mining involves collecting, processing and analyzing the data to discover the insights from it. There are various techniques and methodologies that involves and serve different purposes. With the help of data cleaning and data transformation steps, we can clean the information and make datasets prepared for analysis. Later we can use visualization techniques to get better know-how and what the exact data, which can affect the decision making and operations.

From the data mining, I would at least know on which product am spending more money and which product is more necessary to buy for the house rather than spending in unnecessary things. I would also get to know how much am spending for one-time shopping and how can I reduce the money spent. I will also know when buying one item, which item is frequently bought together.

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

Both classification and clustering have their individual importance based on the scenarios.

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.