Eva Lipska

task 4, course 2

**Associations between products**

**Goal of the project**

Performing a market basket analysis – identification of purchasing patterns that will provide insight into Electronidexe's clientele.

**In technical terms it is about finding association rules** - patterns discovered between items and their associated transactions.

*Data*: a record of Electronidexe’s 9835 online transactions from one month including 125 types of products.

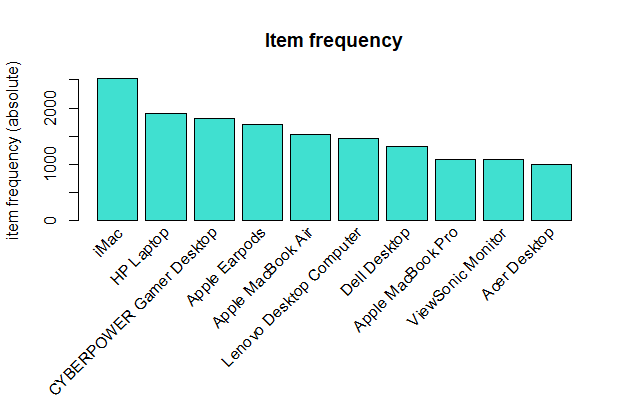
**Data overview**

**Most frequently purchased items**

Total number of items purchased in all transactions: 43104

1. iMac - 2519
2. HP Laptop - 1909
3. CYBERPOWER Gamer Desktop - 1809
4. Apple Earpods - 1715
5. Apple MacBook Air – 1530

Other products – 33622



- 4 of 10 the most frequently purchased items are Apple products;

- within those 10 products Apple is the only one repetitive brand.

**How many items do customers purchase the most?**

50% of the transactions are within the range 2 and 6 items

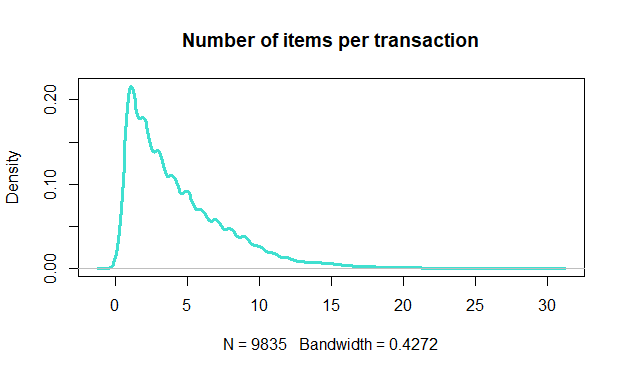
25 % of the transactions: 0 – 2 items

25 % of the transactions: 6 – 30 items

The average is 4.4

The highest amount of transactions – 2163 – include only 1 item, then 1647 transactions – contain 2 items. The amount of items per transaction gradually decreases. There is for example only one transaction containing 30 items.

Most customers are buying only few items, and it’s very rare to buy more than 17 items.



**Market basket analysis**

The analysis indicates that generalizations based on the strong associating rules should be questioned, as the inferences made by associations do not necessarily imply the causality.

Even though Apple products are the most frequently purchased items, there aren’t any significant patterns regarding transactions based on brand loyalty. Apple products are just popular items across the entire user base. There are more products that are popular (e.g. HP Laptop) and from this reason they are often part of the associating rules, however without any logical reasoning behind this particular combination of products. Some obvious associations like desktop + monitor were observed.

First of all there are two questions that should be addressed:

If there were any promotions that influenced the sell volume? And if the items were presented to customers together, so they may appear co-occurring?

On the basis of the analysis it is also difficult to define the customers, according to the proportion of individual clients vs. corporate subjects, however it looks like both types of clients are present. Buying couple of the same type of products within one transaction may be for example characteristic for corporate subjects.

**Technical explanation:**

The first step was to search for the associating rules characterized by combination of high support, reasonable confidence and lift higher than 1, achievable through the process of calibrating parameters.

In order to produce more meaningful insight, pruning down for simplicity was applied – 3 elements as a maximum length of a rule.

Afterwards the analyses of products going along with the most frequently purchased items were performed for:

1. iMac
2. HP Laptop
3. CYBERPOWER Gamer Desktop
4. Apple Earpods
5. Apple MacBook Air

The analysis did not indicate the rules with high support as well as high confidence.

Rules with high support means an applicability to a large number of transactions, however low support level is typical for datasets including a high volume of transactions and a large product offer.

In general low support may be not interesting from a business perspective as it may not be profitable to promote items that are seldom bought together.

Strong confidence is also a relative term. Confidence measures the reliability of the inference made by a given rule

**Summary**

Before any recommendations the dataset should be a subject to a further analysis as not all the necessary information regarding transactions has been delivered.

Missing information concerns:

- previous marketing activities that may have influenced the transactions (cross – promotions, sales);

- information about customers: amount of customers, proportion of individual clients vs. corporate subjects;

- information about prices, margin on the products;

- information about costs of the business.

Due to incomplete information, we cannot provide an opinion on the merits of acquiring Electronidex.