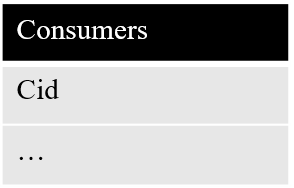
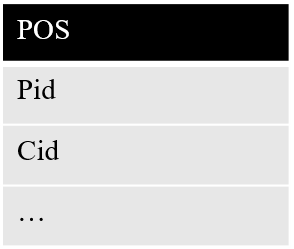
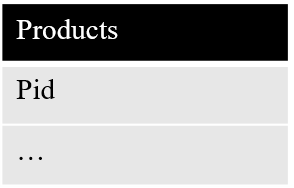
**Task**

A grocery chain with physical stores in Karlskrona in Sweden has given you access to three different kinds of data: Point of Sales (POS), product characteristics and consumer characteristics based on a member loyalty program. The relationships between the tables are as illustrated in Figure 1. The collected variables are described in Table 1, Table 2 and Table 3 (see the next page). You have access to a sample of these data. From the main table of all members of the loyalty program, a random sample of 1,000 members were chosen. All transactions at the grocery chain in 2018 of the 1,000 loyal members are included.



*Figure 1.* Relationships between the three tables

Your task is to report on patterns you find in these data in the form of a short academic article.

*Table 1.* Products

|  |  |
| --- | --- |
| pid | Unique product id |
| category | Main category of the product |
| eco | 1 if the product is labelled as sustainable |

*Table 2.* Point of Sales (POS)

|  |  |
| --- | --- |
| pid | Unique product id |
| cid | Unique consumer id |
| rid | Unique receipt id |
| date | Date of transaction |
| time | Time of transaction in CET |
| price | Original total price, excl. VAT |
| discount | Discount on total price (can be 0) |
| price\_addedvat | Original total price, incl. VAT |
| marginal | Total price after discount less costs of product for the store |
| quantity | Number/weight of products |

*Table 3.* Consumers

|  |  |
| --- | --- |
| cid | Unique consumer id |
| gender | 1 if consumer is female |
| age | Age in years of consumer |
| postalcode | Postal code for main living address of consumer |