# **Assess Sales Outlets' Performance**

### **Project description:**

In this project, we are working as a junior analyst for a building-material retailer called Home World. The marketing department wants to know the performance of the sales outlets to assess its profitability and identify outlet stores that are bringing in more revenue.

The datasets provided contain purchase data from year 2016 to 2017. All of its customers have membership cards. Moreover, they can become members of the store's loyalty program for $20 per month. The program includes discounts, information on special offers, and gifts.

Our job is to assess customers' purchase behavior and determine if a loyalty program of $20 per month will bring in more customers or not.

### **Data description:**

There were two datasets provided retail\_dataset\_us and product\_codes\_us.

#### **retail\_dataset\_us.csv contains**

* purchaseId — customer purchase ID
* item\_ID — product ID
* purchasedate — Customer purchase date
* Quantity — the number of items in the purchase
* CustomerID — customer account ID
* ShopID — shop ID
* loyalty\_program — whether the customer is a member of the loyalty program. '0' stands for false means no loyalty program and '1' stands for true means loyalty program

#### **product\_codes\_us.csv contains**

* productID — customer product ID
* price\_per\_one — product cost per one

### **Steps for completing the project**

* Step 1: Open the data files and study the general information.
* Step 2: Preprocess the data
  + We will view the data in more detail to convert them to the required types, replace the column names and find missing values. Find duplicates and eliminate errors in the data.
* Step 3: Analyze the data
  + Determine the number of purchases by year. Also, breakdown by storeIds and loyalty program.
  + Determine the number of purchases breakdown by storeID. Plot bar diagram showing the number of purchases by storeID.
  + Examine how many purchases are made in a day? Plot a histogram showing the number of purchases.
  + Examine how many purchases are made per customer? Check if they have loyalty program. Plot a histogram showing the number of purchases per customer.
  + What products is most bought by customers?
  + Determine the share of purchases made in different stores. Plot a pie chart showing the shares of purchases in different stores.
  + What time of year is the best to maximize the likelihood of customers' purchases? Check if purchases spike during holidays or any major events (like Thanksgiving, and Christmas).
    - How effect is a loyalty program? Do customer with or without loyalty program purchases more or less?
* Step 4: Test the hypotheses
  + Average purchases of customers with loyalty programs are the same as of customers without a loyalty program.
* Step 5: Prepare a presentation

### **Goal:**

Our goal is to make decisions based on sales performance.