



Think Ahead

ACCA

# Dataset overview: S-MartX

Machine learning with Python  
for finance professionals

THINK AHEAD

ACCA

# Case Study – S-MartX – Factsheet

❑ **Name / location / industry:** *S-MartX, Singapore, e-commerce*

❑ **Market:** Singapore (87% of sales), Indonesia (10%), Cambodia, Vietnam and Malaysia

❑ **Stats:**

- 350+ items sold on its portal
- 90% of sales from Home Appliance & Electronics, and Mobiles & Computers
- Growing annually with CAGR of 15.4%
- In 2019, *S-MartX*'s net sales grew 8.1% year-on-year to SGD 18M
- 100+ full-time employees in December 2019

## S-MartX product categories

Beauty & Health

Fashion & Clothing (Men & Women)

Home and Kitchen

Home Appliances & Electronics

Mobiles & Computers

Sports & Fitness

Toys & Baby Products

# Case Study – S-MartX – Dataset

The dataset provided is an extract from the *SMart-X* transactions database covering a 10 year period. Key stats:

- ❑ **50K+ products** delivered for **29,453 orders** across 10 years; ~3,000 orders delivered/year is approx. 8 orders daily
- ❑ **Total net sales** of SGD 102 million, or approx. SGD 28K daily
- ❑ **Profit** of SGD 11.7 million; **total discount** provided is SGD 6 million, i.e. 6% of total net sales
- ❑ **Most profitable category** is Home Appliances & Electronics (58% of total profit)
- ❑ **Customer demographics:**
  - ❑ 51% male vs 49% female
  - ❑ 44% of sales from 30-40 years age group; 42% of sales from 21-30 years age group
- ❑ **Payment methods:**
  - ❑ 88% of customers use an online payment method
  - ❑ 42% of payments made using a credit card

# Case Study – S-MartX – Database Columns

*S-MartX* has an online database system that keeps track of each transaction.

Columns Names	Description
Receipt No	Represents the Order Number placed by a customer
Item Name	Name of Item which has been sold
Category Name	Category of Item to which it Belongs
Prod Grp Name	Sub- Category of Item
Brand Name	Brand of Item
Product Code	Unique identifier for each product
Description	Complete Description of Item
Price[SGD]	Price of Item to be sold in Singapore dollar
Quantity	Quantity of Item

Columns Names	Description
Discount	Discount percentage on a particular Item
Time	The time when Item was placed for Order
Date	The date when the order was placed
Net Sales[SGD]	Total Bill of an order
Profit Margin	Profit percentage on a particular Item
Country	Country of Origin to which Customer Belongs
Offer Name	The offer of Discount which Customer availed if any
Vendor Name	Vendor from which Item was sourced
Shift No	Shift number in which order was placed

# Case Study – S-MartX – Database Columns

The following columns are also present within the database.

Columns Names	Description
Staff	Staff Id which processed the Order
Customer ID	Customer ID of the customer
Gender	Customer Gender
Age	Customer Age
Year	The year of order date
Month	The month of order date
Day	The month day when the order was placed
Month-Year	Month & Year of placing order
Payment Mode	Method of Payment
Bank Name	Bank Name or Cash
GST	Goods Service Tax (Fixed for all category as 7%)
OCBC Credit card	Categorical variable for logistic regression

# Case Study – S-Mart-X

## Problem statement

- ❑ The *S-Mart-X* Chief Commercial Officer (CCO) would like to improve vendor engagement and vendor product offerings through a new **vendor analytics** product.
- ❑ The theory is that providing vendors with **improved analytics**, the vendors will be able to make smarter decisions around **which products are selling** well and furthermore to **optimise product pricing**.
- ❑ The Analytics team has been given the task of building an **automated reporting system** that **integrates with** the data analytics tool that vendors already know, **Microsoft Excel**.
- ❑ The goal is to use **Python, pandas** and **xlwings** to programmatically populate pre-designed Vendor Datasheets with the latest data.





