Introduction to Source data

SALES Database Overview

This summary provides a clear overview of the database structure, key statistics, and insights into each table's contents for stakeholders to understand the dataset comprehensively.

The SALES database consists of 5 tables:

1. CUSTOMERS

* Columns: CUSTOMER\_CODE (varchar), CUSTOMER\_NAME (varchar), CUSTOMER\_TYPE (varchar)
* Records: 38
* Distinct CUSTOMER\_NAME count: 38
* Distinct CUSTOMER\_CODE count: 38
* Customer types:
* "BRICK & MORTAR": 19
* "E-COMMERCE": 19

1. DATE

* Columns: DATE (date), CY\_DATE (date), YEAR (int), MONTH\_NAME (varchar), DATE\_yy\_mmm (varchar)
* DATE is the order date.
* CY\_DATE is the first date of each month.

1. MARKETS

* Columns: MARKET\_CODE (varchar), MARKET\_NAME (varchar), ZONE (varchar)
* Records: 17
* Distinct MARKET\_CODE count: 17
* Distinct MARKET\_NAME count: 16
* Zones: 3 (one zone is blank)

1. PRODUCTS

* Columns: PRODUCT\_CODE (varchar), PRODUCT\_TYPE (varchar)
* Distinct PRODUCT\_CODE count: 279
* Product types:
* OWN BRAND: 191
* DISTRIBUTION: 88

1. TRANSACTIONS

* Columns: product\_code (varchar), customer\_code (varchar), market\_code (varchar), order\_date (date), sales\_qty (int), sales\_amount (double), currency (varchar), profit\_margin\_percentage (double), profit\_margin (double), cost\_price (double)
* Total rows: 148395
* Missing values:
* profit\_margin: 2009
* profit\_margin\_percentage: 2009
* Excluding missing data rows: 146386 (providing a more accurate representation for analysis)