**TASK 1.1**

4-Step Process for Dimensional Modeling

1. Identify the Business Process: The main business process appears to revolve around sales and payments, as these are central to revenue and operations.
2. Identify the Grain: Grain = One row in the fact table represents a single sale transaction, which includes details about the sale (sale date, sale amount, sale payment, a store, product, and customer).
3. Identify the Dimensions: Dimensions are the descriptive attributes related to the business process. From the schema, the following dimensions are identifiable:

* Product Dimension: Contains attributes from PRODUCTS and PRODUCT\_CATEGORIES.
* Customer Dimension: Includes CUSTOMERS and linked address details (e.g., ADRESSES, CITIES, COUNTRIES, REGIONS).
* Store Dimension: Includes store-specific data with EMPLOYEES data and address details

1. Identify the Fact Table: The fact table captures measures related to the grain and is linked to the dimensions. Here the central table is derived from SALES and PAYMENTS. It captures the following measures:

* Sales amount.
* Payment amount.
* Date

Steps for Denormalization

1. Combine normalized tables into single dimension tables by removing foreign keys and embedding related descriptive attributes:

PRODUCTS and PRODUCT\_CATEGORIES → Combine into **Product** Dimension.

CUSTOMERS, ADDRESSES, CITIES, COUNTRIES, and REGIONS → Combine into **Customer** Dimension.

STORES, EMPLOYEES, ADDRESSES, CITIES, COUNTRIES, and REGIONS → Combine into **Store** Dimension.

1. Denormalize Fact Table

Combine SALES and PAYMENTS into a single **Sales** Fact Table.

Include foreign keys to link dimensions: Product\_ID, Customer\_ID, and Store\_ID.

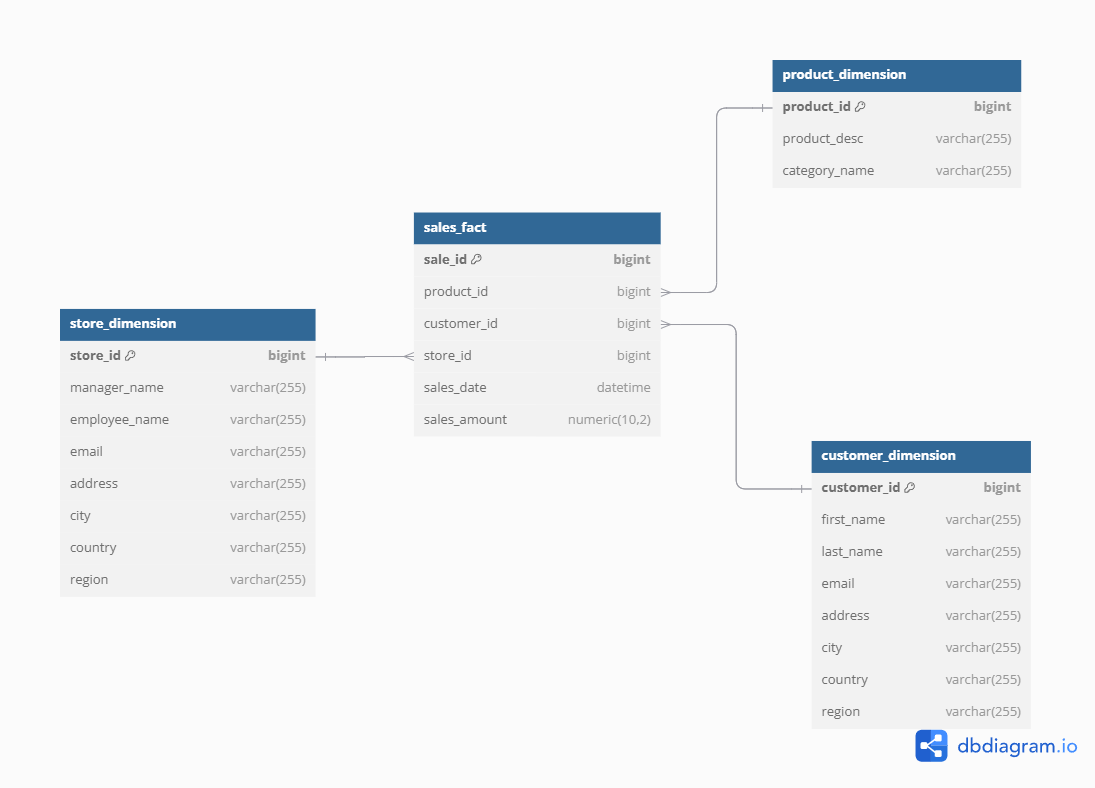
Star Schema: Based on the 4-step process, the star schema structure would look like this:

Fact Table: Sales\_Fact

* Measures: Sales\_Amount, Sales\_Date
* Foreign Keys: Product\_ID, Customer\_ID, Store\_ID.

Dimensions

* Product Dimension: Product\_ID, Product\_Desc, Category\_Name.
* Customer Dimension: Customer\_ID, First\_Name, Last\_Name, Email, Address, City, Country, Region
* Store Dimension: Store\_ID, Manager\_Name, Employee\_Name, Email, Address, City, Country, Region



**TASK 1.2** To construct a **snowflake schema**, we further normalize the **star schema** by splitting some of its dimension tables into sub-dimensions based on relationships and hierarchies present in the original schema. Break down dimensions like Product Dimension, Customer Dimension, and Geography Dimension into multiple smaller tables to reflect the normalized structure.

* Product Dimension → Products and Product\_Categories.
* Customer Dimension → Customers, Addresses, Cities, Regions, Countries, and Economic\_Regions.
* Store Dimension → Stores, Employees, Addresses Cities, Regions, Countries, and Economic\_Regions.

The fact table remains largely unchanged (e.g., Sales\_Fact), containing measures and foreign keys to connect to dimensions.

