# Task: Data modeling

The **Fact table** for transactions stores data at the level of the day and time when the customer made a purchase of a specific product, as well as quantitative data related to the quantity of products the customer ordered, the base price of the product before the applied discount (with the assumption that VAT is already included in the base price), the percentage discount applied based on the current promotion for that product (or for that customer), the total price of the product, the currency in which the price is expressed, and the current status of the transaction.

In the **Fact table**, Transaction\_Id is **not the primary key** of the table because the granularity of the fact table is at the product level, and one transaction can include multiple products (in other words, this table is essentially a "transaction item" table, while the total transaction value could be created later as a new fact table with granularity at the level of the date and customer, including the total transaction amount, on which shipping, for example, could further be applied). The primary key for the fact table is omitted in this case, as it is not necessary for reporting and analytics, and due to its high cardinality, it could unnecessarily take up space in in-memory BI tools (such as Power BI or Tableau).

The fact table is described by the following dimensions:

* **Date** and **Time** with their respective time hierarchy. These two tables are separated so that the time dimension does not create redundancy in the Date table.
* **Product**, which is a catalog of all products with attributes such as code, name, size, category, color, and collection.
* **Customer**, which is a catalog of all customers with basic contact information and user addresses, as well as a denormalized dimension called 'Customer\_Type' that indicates whether the customer has a status like 'Platinum,' 'Gold,' or 'Silver.' Customer type is stored as a Slowly Changing Dimension Type 1 (SCD 1) because tracking changes in customer status is not necessary for reporting, only the current status is needed.
* **Promotion**, with the name of the promotion on which the discount is based.
* **Transaction status**, with statuses such as 'Pending,' 'In progress,' and 'Completed.'

*Additionally, it is assumed that columns tracking the timestamp of record creation and modification, as well as a flag indicating whether the record is deleted, would be included in the tables.*

For additional context, examples of attribute values for each table are provided below:

* Attributes of **Fact\_Transaction** table:

|  |  |
| --- | --- |
| **Attribute** | **Example** |
| Transaction\_Id | ‘123456789’ |
| Date | ’2024-09-23’ |
| Time | ’11:30:25 AM’ |
| Customer\_Id | ‘2341234050003003’ |
| Product\_Id | ‘6854654050003003’ |
| Quantity\_Sold | ‘1’ |
| Base\_Price | ‘100.00’ |
| Discount | ‘0.1’ |
| Promotion\_Id | ‘6854654050003003’ |
| Discounted\_Price | ’90.00’ |
| Total\_Amount | ‘90.00 |
| Currency | USD |
| Status\_Id | ‘3’ |

Attributes of **Dim\_Date** table:

|  |  |
| --- | --- |
| **Attribute** | **Example** |
| Date | ‘2024-09-23’ |
| Year | ‘2024’ |
| Month\_Sort | ‘9’ |
| Month | ‘September’ |
| Day | ‘23’ |
| Day\_Of\_Week | ‘Monday’ |
| Quarter\_Sort | ‘3’ |
| Quarter | ‘Q3’ |

Attributes of **Dim\_Time** table:

|  |  |
| --- | --- |
| **Attribute** | **Example** |
| Time | ’11:30:25 AM’ |
| Hour | ‘11’ |
| AM\_PM | ‘AM’ |
| Minute | ‘30’ |
| Second | ‘25’ |

Attributes of **Dim\_Status** table:

|  |  |
| --- | --- |
| **Attribute** | **Example** |
| Status\_Id | ’3’ |
| Status\_Name | ‘Completed’ |

Attributes of **Dim\_Customer** Table:

|  |  |
| --- | --- |
| **Attribute** | **Example** |
| Customer\_Id | ‘2341234050003003’ |
| First\_Name | ’John’ |
| Last\_Name | ‘Johnson’ |
| Email | ‘john.johnson@gmail.com’ |
| Telephone | ‘212-555-6789’ |
| Postal\_Code | ‘1001’ |
| State | ‘United States of America’ |
| City | ‘New York’ |
| Adress | ‘Example 1’ |
| Customer\_Type | ‘Platinum |

Attributes of **Dim\_Product** Table:

|  |  |
| --- | --- |
| **Attribute** | **Example** |
| Product\_Id | ‘6854654050003003’ |
| Product\_Code | ’Z100000001’ |
| Product\_Name | ‘Shirt dress with belt’ |
| Product\_Category | ‘Dress’ |
| Product\_Colour | ‘Green’ |
| Product\_Size | ‘XS’ |
| Product\_Collection | ‘Winter collection’ |

Attributes of **Dim\_Promotion** Table:

|  |  |
| --- | --- |
| **Attribute** | **Example** |
| Product\_Id | ‘6854654050003003’ |
| Promotion\_Name | ’Black Friday’ |