**Normalization**

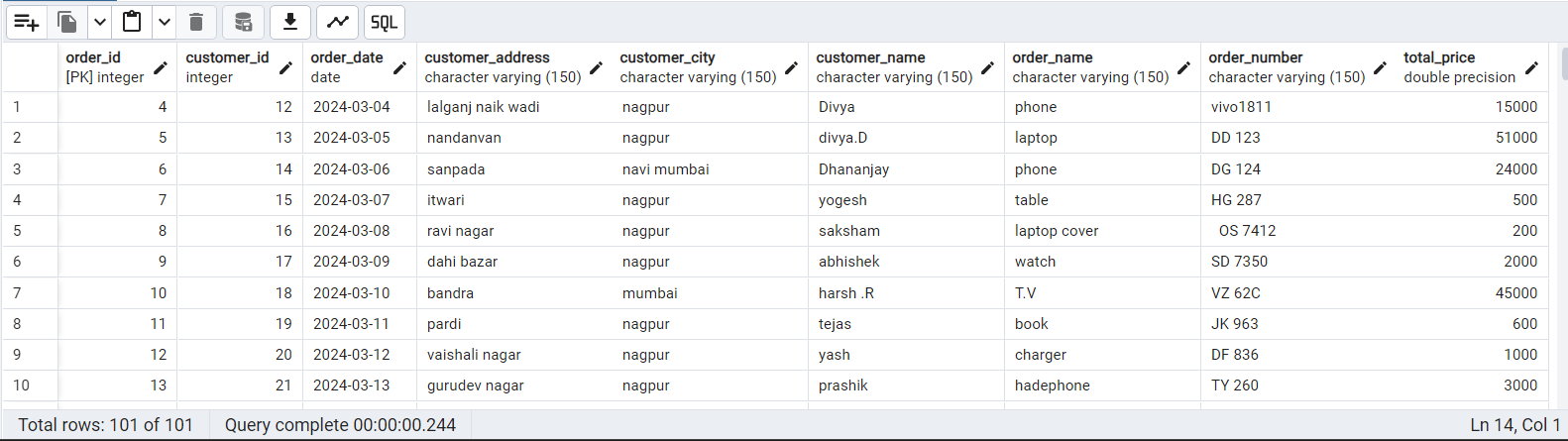
Introduction to Normalization

Normalization is a process in database design that organizes tables to minimize redundancy and dependency. It involves dividing large tables into smaller ones and defining relationships between them. The process ensures data integrity and makes it easier to manage and query data. There are several normal forms, including 1NF, 2NF, and 3NF, each addressing specific issues in database structure.

* Normalization is the process of organizing the data in the database.
* Normalization is used to minimize the redundancy from a relation or set of relations. It is also used to eliminate undesirable characteristics like Insertion, Update, and Deletion Anomalies.
* Normalization divides the larger table into smaller and links them using relationships.
* The normal form is used to reduce redundancy from the database table.

Example- Original Data Structure

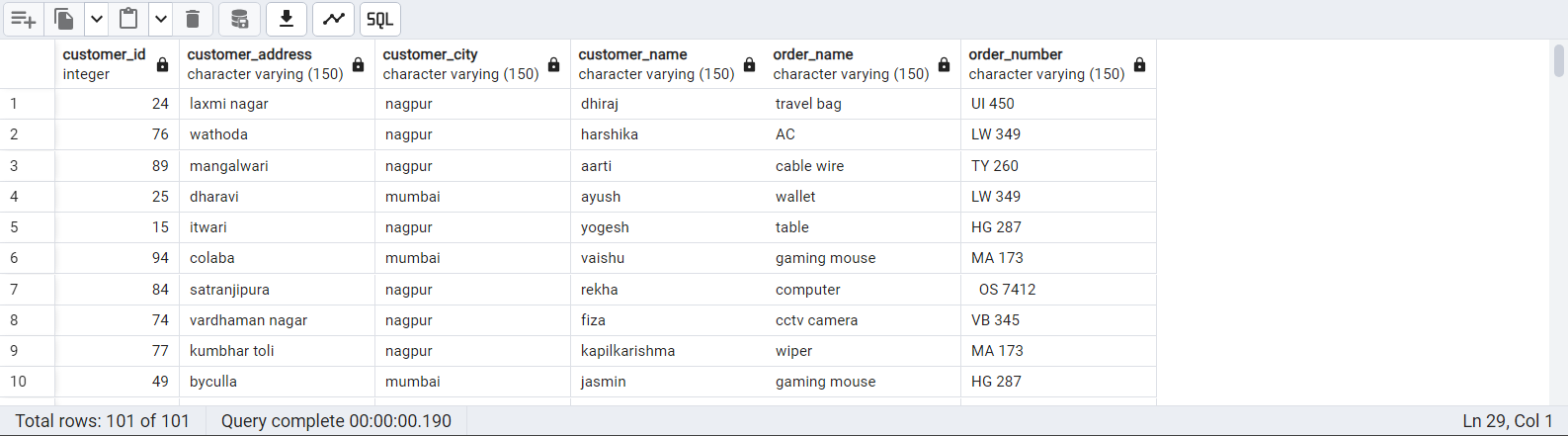
We have an ‘order\_data’ table with the following columns:



1NF (First Normal Form):

If a relation contains a composite or multi-valued attribute, it violates the first normal form, or the relation is in the first normal form if it does not contain any composite or multi-valued attribute. A relation is in first normal form if every attribute in that relation is single-valued attribute.

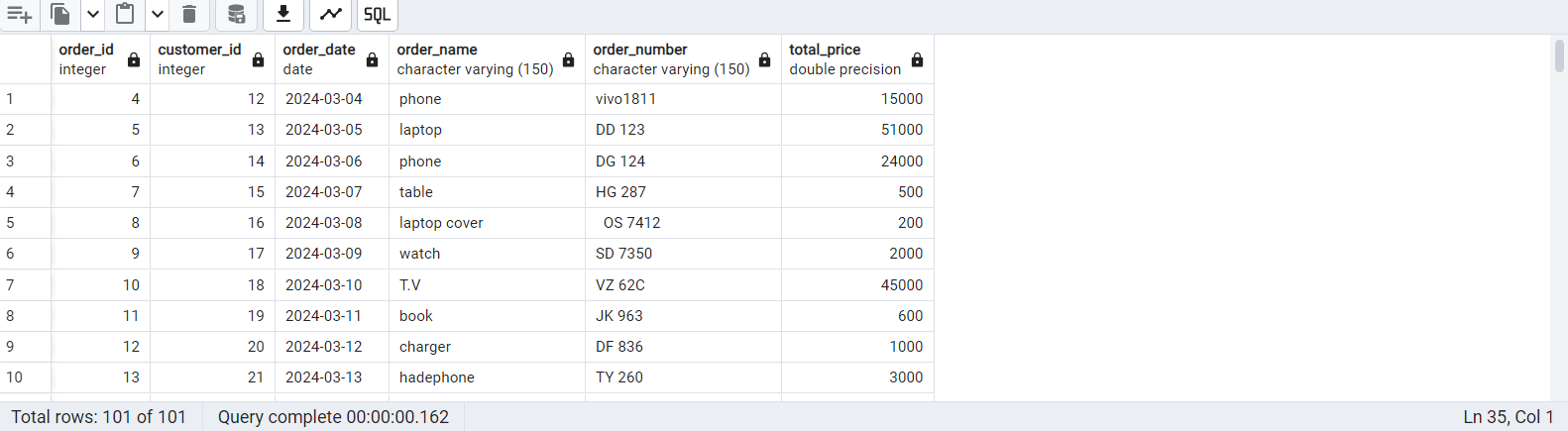
* Every Attribute/Column. need to have a single value.
* There is a unique name for every Attribute/Column.
* Not mandatory have primary key.



2NF (Second Normal Form):

2NF requires that the table is in 1NF and all non-key attributes are fully functional dependent on the primary key. So there is one Column should be primary key attribute and remaining other columns are dependent on the primary key column.

* Must be in 1NF.
* Every table should have primary key and relationship between the table should be formed using foreign key.



3NF (Third Normal Form):

3NF requires that the table is in 2NF and all attributes are dependent only on the primary key. We need to remove such dependencies by progressing to the Third Normal Form (3NF). A relation is in 3NF if at least one of the following conditions holds in every non-trivial function dependency X –> Y.

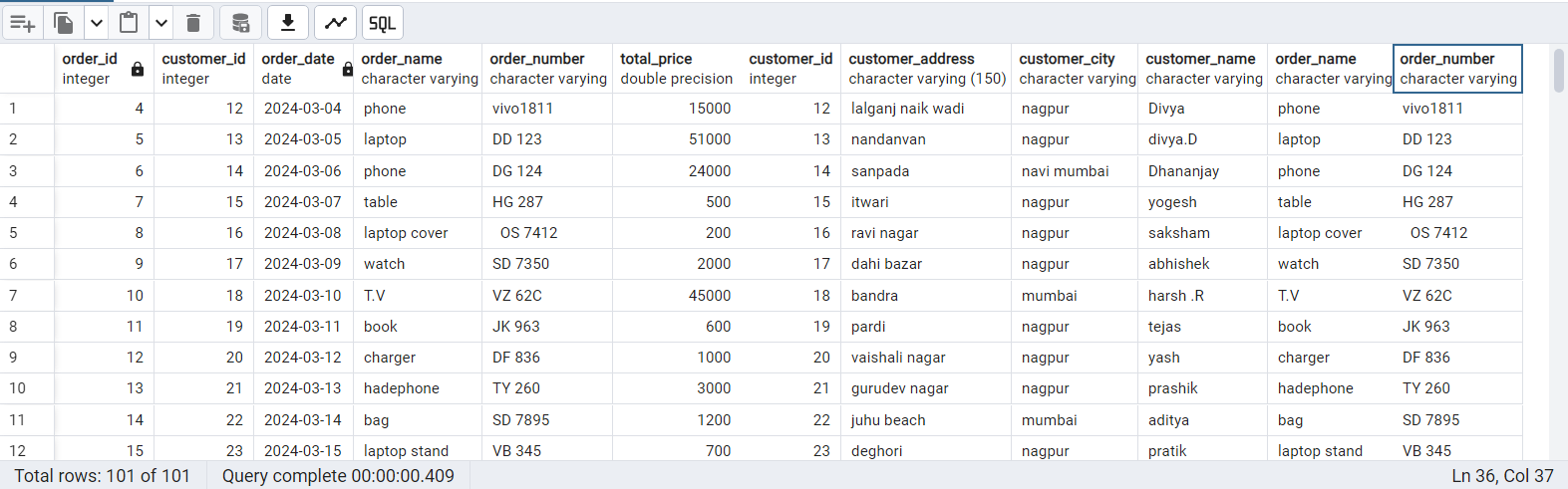
* X is a super key.
* Y is a prime attribute (each element of Y is part of some candidate key).
* Avoid transitive dependencies.

Join the tables using Inner Join:

select \* from normalized\_order as n

inner join cust\_order as c

on n.customer\_id = c.customer\_id



ERD for Database:

