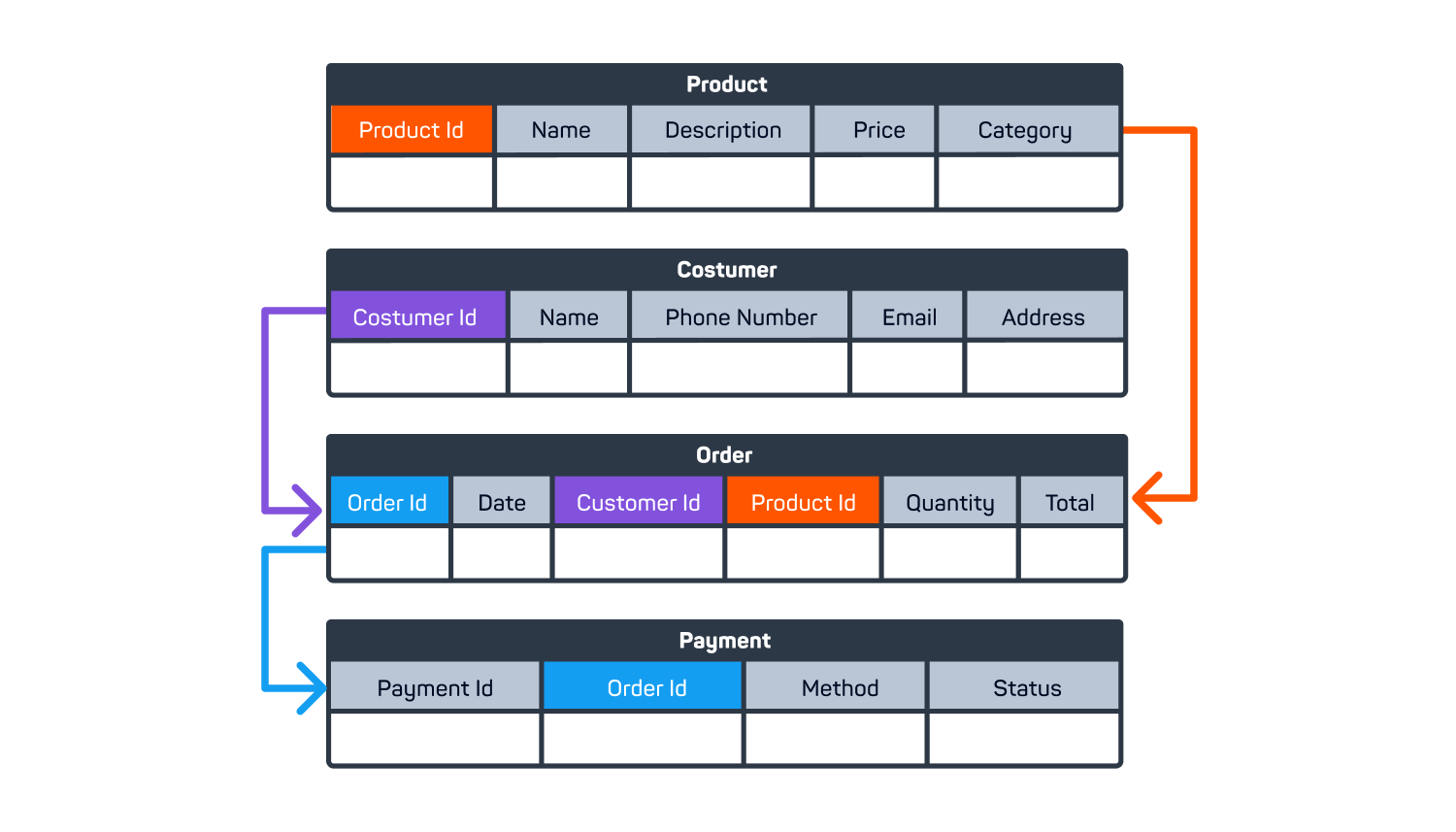
**Normalization**

Normalization is a process in database design used to organize data to **reduce duplication** and **improve data accuracy**. The process involves structuring a database in such a way that it meets certain criteria, known as normal forms, which help eliminate undesirable characteristics **like insertion, update, and deletion anomalies(errors).**

* **Eliminate Redundancy**: Reduces duplicate data.
* **Ensure Data Integrity**: Ensures that data remains accurate and consistent.
* **Improve Data Maintenance**: Simplifies the processes of updating and maintaining the database.
* **Optimize Queries**: Improves the performance of data retrieval.



**Fig-1 Normalization**

**Types of Normal Forms**

1. First Normal Form (1NF):
   * **Definition:** A table is in 1NF if it contains only atomic (indivisible) values and each column contains values of a single type.

|  |  |  |
| --- | --- | --- |
| **ID** | **Name** | **Phone Numbers** |
| 1 | John | 123456,6789 |

|  |  |  |
| --- | --- | --- |
| **ID** | **Name** | **Phone Numbers** |
| 1 | John | 123456 |
| 1 | John | 6789 |

1. Second Normal Form (2NF):

* **Definition:** A table is in 2NF if it is in 1NF and all non-key attributes are fully functionally dependent on the primary key.

**Example**

|  |  |  |  |
| --- | --- | --- | --- |
| **Student ID** | **Course ID** | **Student Name** | **Course Name** |
| 1 | 101 | John | Math |
| 1 | 102 | John | Science |

**2NF Tables**

|  |  |
| --- | --- |
| **Student ID** | **Student Name** |
| 1 | John |

|  |  |
| --- | --- |
| **Course ID** | **Course Name** |
| 1 | Math |
| 2 | Science |

|  |  |
| --- | --- |
| **Student ID** | **Course ID** |
| 1 | 101 |
| 1 | 102 |

1. Third Normal Form (3NF):

* **Definition:** A table is in 3NF if it is in 2NF and all the attributes are functionally dependent only on the primary key.

**Example**

|  |  |  |
| --- | --- | --- |
| **Order ID** | **Customer ID** | **Customer Name** |
| 1001 | 500 | Alice |

**3 NF tables**

|  |  |
| --- | --- |
| **Order ID** | **Customer ID** |
| 1001 | 500 |

|  |  |
| --- | --- |
| **Customer ID** | **Customer Name** |
| 500 | Alice |

1. Boyce-Codd Normal Form (BCNF):

* **Definition:** A stronger version of 3NF where every determinant is a candidate key.

**Example:**

|  |  |  |
| --- | --- | --- |
| **Course ID** | **Instructor** | **Department** |
| 101 | Prof.Smith | Math |
| 102 | Prof.Jonhson | Science |

**BCNF Tables**

|  |  |
| --- | --- |
| **Course ID** | **Department** |
| 101 | Math |
| 102 | Science |

|  |  |
| --- | --- |
| **Instructor** | **Department** |
| Prof.Smith | Math |
| Prof.Jonhson | Science |

1. Fourth Normal Form (4NF):

* **Definition:** A table is in 4NF if it is in BCNF and multivalued dependencies are removed.

**Example:**

|  |  |  |
| --- | --- | --- |
| **Student ID** | **Course ID** | **Hobby** |
| 1 | 101 | Chess |
| 1 | 101 | Music |

**4 NF Tables**

|  |  |
| --- | --- |
| **Student ID** | **Course ID** |
| **1** | **101** |

|  |  |
| --- | --- |
| **Student ID** | **Hobby** |
| 1 | Chess |
| 2 | Music |

**Reference –**

**Diagram Reference -** <https://www.sololearn.com/blog/what-is-normalization/>

**Definition Reference**- [Chat-GPT](https://chatgpt.com/) + My Mind + [Google](https://www.google.com/)

**Tables Reference** - <https://www.javatpoint.com/dbms-normalization> + [Chat GPT](https://chatgpt.com/)