

# Star Schema VS Fact Constellation Schema

By, Shristi Kamble, TE (AI & DS)

## Star Schema :

A star schema is a simple and intuitive schema that consists of one fact table and several dimension tables. The fact table contains the facts and the foreign keys to the dimension tables. The dimension tables contain the descriptive attributes of the facts. The schema resembles a star, with the fact table in the center and the dimension tables radiating from it. A star schema is easy to understand, query, and maintain. It also supports fast aggregation and analysis of facts across multiple dimensions.



## The key features and advantages of a star schema include:

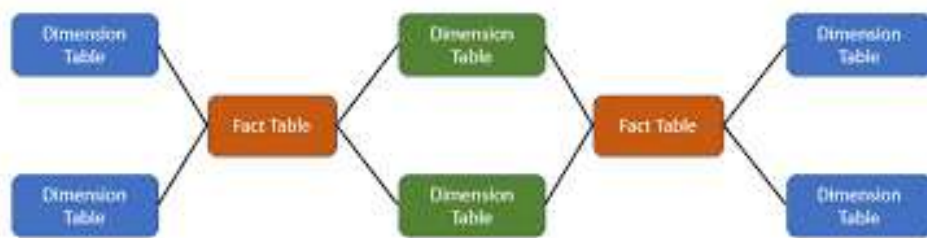
1. **Simplicity:** Star schemas are relatively easy to understand and maintain, making them a popular choice for data warehousing.
2. **Performance:** Querying data from a star schema is typically faster and more efficient than other schema types because it minimizes the number of joins required for most common queries.
3. **Flexibility:** Star schemas allow for easy addition of new dimensions or changes to existing ones without significant disruption to the data model.
4. **Improved reporting:** The structure of a star schema is well-suited for business intelligence tools and reporting, as it provides a clear and structured way to access and analyze data.
5. **Scalability:** Star schemas can scale to handle large amounts of data, making them suitable for enterprise-level data warehousing.

## Limitations

1. Less flexibility for changing business needs
2. Difficulty in accommodating complex relationships

## Fact Constellation Schema :

A fact constellation is a more complex and flexible schema that consists of multiple fact tables and shared dimension tables. The fact tables contain different levels of granularity or different perspectives of the facts. For example, one fact table may contain daily sales data, while another fact table may contain monthly inventory data. The dimension tables contain the common attributes that can be used to join the fact tables. The schema resembles a constellation, with multiple stars connected by shared dimensions. A fact constellation allows for more detailed and varied analysis of facts across multiple levels and perspectives.



## Key features of Fact Constellation Schema:

1. **Multiple Fact Tables:** A fact constellation schema contains more than one fact table, each representing a different area of analysis or business process. For example, there might be one fact table for sales data, another for inventory data, and another for financial data.
2. **Shared Dimension Tables:** Dimension tables are shared among multiple fact tables. Common dimensions, such as date, customer, and product, may be shared to maintain consistency and enable data integration across various subject areas.
3. **Independence:** Each fact table within the schema is relatively independent and contains measures specific to a particular business process or analytical focus. This allows for separate reporting and analysis of different aspects of the organization.
4. **Increased Complexity:** Fact constellation schemas are more complex to design, implement, and maintain compared to star schemas. Querying and reporting on data from a fact constellation schema can be more challenging due to the complexity of the relationships.

5. Flexibility: This schema design provides the flexibility to accommodate various data requirements and analytical needs within an organization. Different fact tables can be added or modified as necessary.

### Limitations

1. Increased schema complexity
2. Potentially slower query performance

Now, let's see the difference between the both schemas

S.NO	Star Schema	Fact constellation schema
1.	A star schema depicts each dimension with only one-dimension table.	While in this, dimension tables are shared by many fact tables.
2.	In star schema, tables can be maintained easily in comparison of fact constellation schema.	While in fact constellation schema, tables cannot be maintained easily comparatively.
3.	Star schema does not use normalization.	Whereas it is a normalized form of star and snowflake schema.
4.	In star schema, simple queries are used to access data from the database.	While in this, heavily complex queries are used to access data from the database.
5.	Star schema is easy to operate as compared to fact constellation schema as it has less number of joins between the tables.	While fact constellation schema is not easy to operate as compared to star schema as it has many joins between the tables.
6.	Star schema uses less space as compared to fact constellation schema.	While fact constellation schema uses more space comparatively.



## Real-World Examples

To illustrate the differences between star schema and fact constellation schema, let's consider two scenarios:

**Retail Industry:** A star schema can be used to analyze sales data based on dimensions like products, customers, and time periods. In contrast, a fact constellation schema may be preferred in the same industry when analyzing sales, inventory, and promotions simultaneously.

**Financial Sector:** A star schema can be utilized to analyze financial transactions by dimensions like accounts, customers, and date. On the other hand, a fact constellation schema may be useful when analyzing transactions, market data, and financial news together.

## Best Practices for Choosing a Schema

When deciding whether to use a star schema or fact constellation for dimensional modeling, it depends on various factors such as the data, queries, and system. Generally, a star schema should be used when the data is simple and queries are focused on one fact table and its dimensions. On the other hand, a fact constellation is better for complex data and queries that require multiple fact tables and shared dimensions. It is important to weigh the trade-offs between simplicity and flexibility, as well as performance and functionality when choosing a schema. Lastly, document and validate the schema design with the stakeholders and users of your data warehouse.

## Conclusion

In conclusion, the choice between a star schema and a fact constellation schema in data warehousing and business intelligence is a critical decision that has a significant impact on an organization's data management and analytics capabilities. Both schema designs have their advantages and trade-offs, and the choice should align with the specific needs and goals of the organization.

The star schema is a more straightforward and widely adopted schema design. It is suitable for scenarios where simplicity, query performance, and ease of maintenance are paramount. The star schema's single fact table linked to dimension tables provides an intuitive structure for reporting and analysis, making it an excellent choice for businesses with well-defined data requirements.

On the other hand, the fact constellation schema, or galaxy schema, is a more complex and flexible design that suits organizations with diverse analytical needs across different business areas. It accommodates multiple fact tables and shared dimensions, allowing for independent analysis within various subject areas. However, this complexity can come with a cost in terms of schema design, query performance, and maintenance.

Ultimately, the choice between the two schema designs should be driven by an organization's unique requirements. While star schemas excel in simplicity and performance for focused analytics, fact constellation schemas offer versatility and integration for more complex, multi-subject analysis. Some organizations may even find a combination of both schema types beneficial to cater to different analytical scenarios.

In summary, both star and fact constellation schemas have their places in the world of data warehousing and business intelligence. The right choice depends on the organization's specific goals and the complexities of its data landscape. Regardless of the chosen schema, a well-structured and maintained data warehousing solution is vital to unlock the full potential of data for informed decision-making and strategic insights.