# Dimensional modeling_DWH_GravityBooks

## Our Business process:

We are tracking book orders from customers, including details of the book, shipping method, customer, and date of the order.

## The grain:

The grain represents a single row for each line item within an order.

## The Dimensions:

- **Dim_Book.** (Book_ID ,Author_Name, Language_Name, Publisher_Name, Title, ISBN_13, Num_Pages, Publication_Date)

- **Dim_ShippingMethod** (Method_ID,Method_Name)

- **Dim_Customer.** (Customer_ID, First_Name, Last_Name, Email, Country_Name, street_name, street_number, city, address_status)

- **Dim_Date.** (date, day, day_name, month, month_name, year, quarter, quarter_name, day_of_week, weekofmonth, weekofyea, fiscal_week, fiscal_week_number, fiscal_month, fiscal_month_number, fiscal_year_month, fiscal_quarter, fiscal_year_quarter, fiscal_year, is_holiday, is_weekend)

## The measures:

- Price: Revenue generated from each order line.
- Shipping_Cost: Cost incurred for shipping.

## Advantages of star schema:

1. **Ease of Querying:** BI tools like SSAS can efficiently query data without complex joins.
2. **Performance:** Denormalized dimensions reduce the need for join operations, speeding up read-heavy workloads.
3. **Suitability for Analysis:** Star schemas are designed for OLAP systems where analytical queries are frequent.
4. **Simpler Design:** The structure is easy to understand and maintain, making it ideal for quick insights and business intelligence.
5. **Flexible for Business Changes**: Adapts well to changes in business requirements, such as adding new dimensions or measures.

## Disadvantages of Using Star Schemas

1. **Lack of integrity:** Denormalization can cause data redundancy. Dimensional attributes are often repeated across multiple records within a dimension table which can cause data quality issues. Since data is duplicated in denormalization, frequent changes can also cause certain tables to display out-of-date information.

2. **Increased costs:** Adding redundant data increases computing and storage costs. This can be especially troubling when handling large datasets.
3. **Limited flexibility**: Star schemas are relatively less robust than normalized structures since they are built for specific use cases. Other approaches might be more effective for complex querying involving multiple joins.
4. **Maintenance difficulties:** As data changes over time, maintaining a star schema can become challenging. Updates to dimension attributes might require changes in multiple places.

## Why was the star schema chosen?

- This approach simplifies the structure for reporting and analytics by directly connecting the fact table to denormalized dimension tables, ensuring faster query performance and reducing complexity.
- The design leverages exclusively one-to-many relationships, eliminating the need for handling complex many-to-many relationships. This not only enhances efficiency but also makes the schema easier to maintain and optimize for analytical workloads.
- Data redundancy will not be an issue in our case.