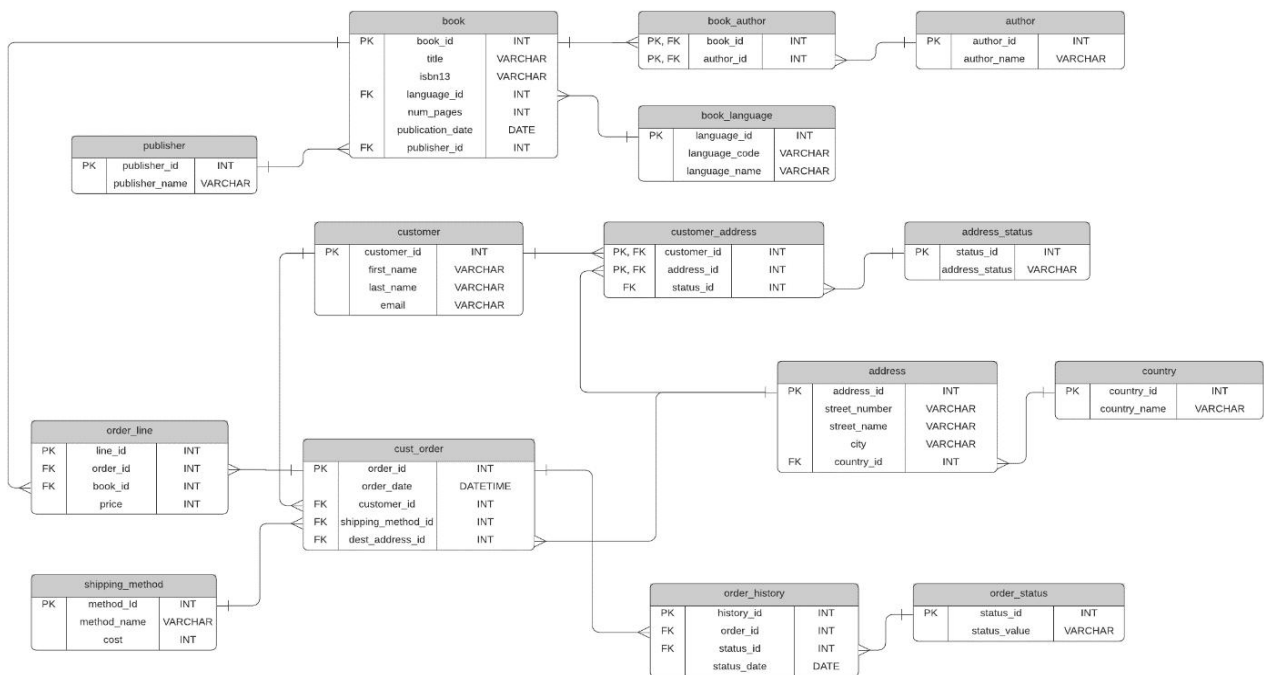


## Section 1: Modeling, Building a Data Warehouse for Gravity Bookstore

### Use-case Background

Gravity Bookstore is a database for a fictional bookstore called that captures information about books, customers, and sales.

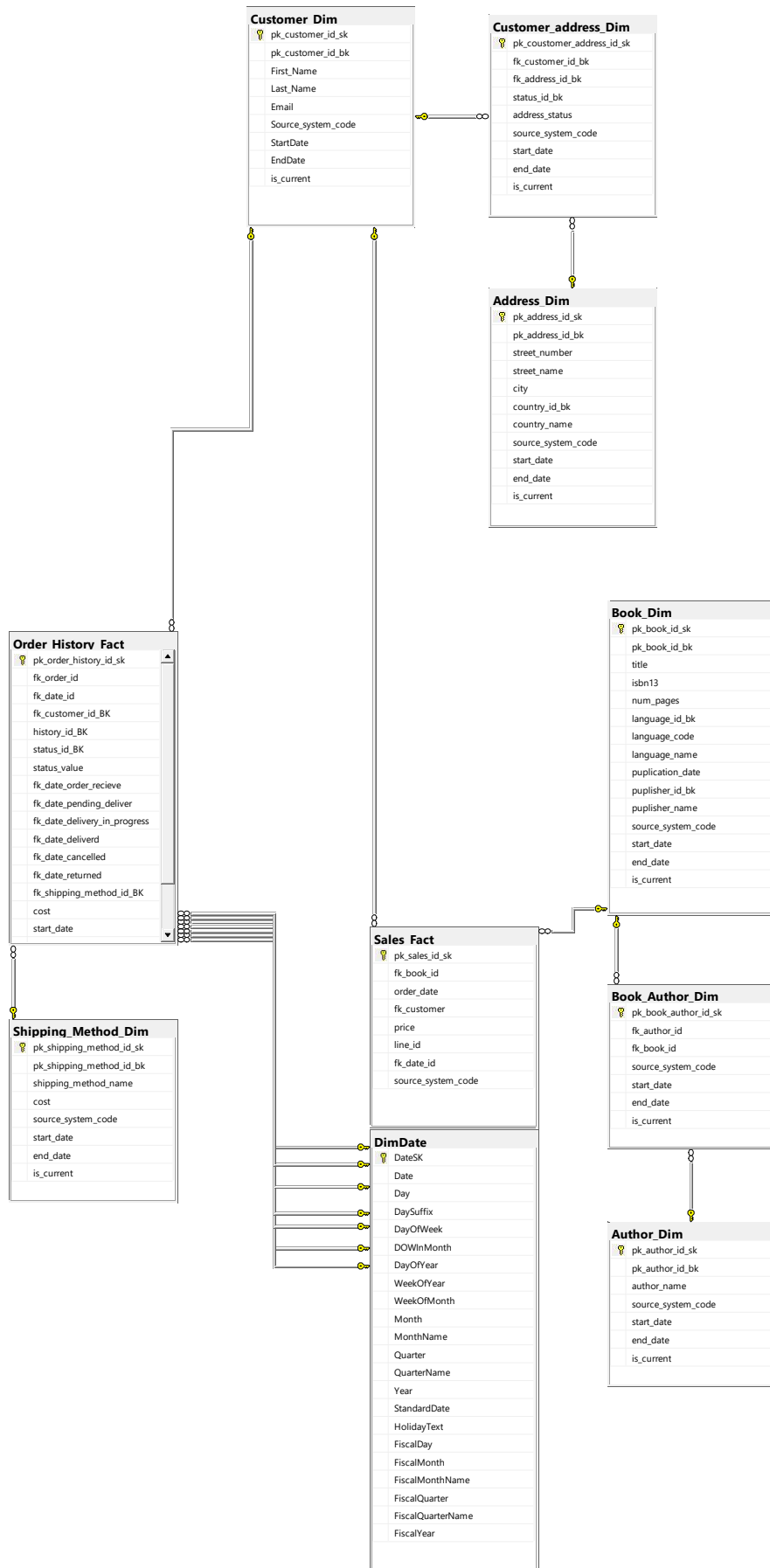
ERD of 'gravity\_books' transactional database:



Tables description:

- **book**: a list of all books available in the store.
- **book\_author**: stores the authors for each book, which is a many-to-many relationship.
- **author**: a list of all authors.
- **book\_language**: a list of possible languages of books.
- **publisher**: a list of publishers for books.
- **customer**: a list of the customers of the Gravity Bookstore.
- **customer\_address**: a list of addresses for customers, as a customer can have more than one address, and an address has more than one customer.
- **address\_status**: a list of statuses for an address, because addresses can be current or old.
- **address**: a list of addresses in the system.
- **country**: a list of countries that addresses are in.
- **cust\_order**: a list of orders placed by customers.
- **order\_line**: a list of books that are a part of each order.
- **shipping\_method**: the possible shipping methods for an order.
- **order\_history**: the history of an order, such as ordered, cancelled, delivered.
- **order\_status**: the possible statuses of an order.

- Gravity\_Book\_DWH Design
- Galaxy Schema



Address_Dim		
Destination Column	Source Table	Source Column
pk_address_id_sk		
pk_address_id_bk	gravity_books_OLTP.dbo.address	address_id
street_number	gravity_books_OLTP.dbo.address	street_number
street_name	gravity_books_OLTP.dbo.address	street_name
city	gravity_books_OLTP.dbo.address	city
country_id_bk	gravity_books_OLTP.dbo.country	country_id
country_name	gravity_books_OLTP.dbo.country	country_name
source_system_code		
start_date		
end_date		
is_current		

Customer_Dim		
Destination Column	Source Table	Source Column
pk_customer_id_sk		
pk_customer_id_bk	gravity_books_OLTP.dbo.customer	customer_id
First_Name	gravity_books_OLTP.dbo.customer	first_name
Last_Name	gravity_books_OLTP.dbo.customer	Last_Name
Email	gravity_books_OLTP.dbo.customer	Email
source_system_code		
start_date		
end_date		
is_current		

Sales_Fact(order_line)		
Destination Column	Source Table	Source Column
pk_sales_id_sk		
fk_book_id	gravity_books_OLTP.dbo.book	pk_book_id
fk_date_id	gravity_books_OLTP.dbo.date	DateSK
fk_customer	gravity_books_OLTP.dbo.customer	pk_customer_id
price	gravity_books_OLTP.dbo.order_line	price

Order_History_Fact		
Destination Column	Source Table	Source Column
pk_order_history_id_sk		
order_history_id_Bk	gravity_books_OLTP.dbo.order_history	history_id
fk_order_id_BK	gravity_books_OLTP.dbo.order_history	order_id
status_id_BK	gravity_books_OLTP.dbo.order_history	status_id
status_value	gravity_books_OLTP.dbo.order_status	status_value
customer_id_BK	gravity_books_OLTP.dbo.cust_order	customer_id
fk_shipping_method_id_BK	gravity_books_OLTP.dbo.cust_order	shipping_method_id
dest_address_id	gravity_books_OLTP.dbo.cust_order	dest_address_id
fk_date_order_recieve	gravity_books_OLTP.dbo.order_history	status_date
fk_date_pending_deliver	gravity_books_OLTP.dbo.order_history	status_date
fk_date_delivery_in_progress	gravity_books_OLTP.dbo.order_history	status_date
fk_date_deliverd	gravity_books_OLTP.dbo.order_history	status_date
fk_date_cancelled	gravity_books_OLTP.dbo.order_history	status_date
fk_date_returned	gravity_books_OLTP.dbo.order_history	status_date

Customer_Address_Dim		
Destination Column	Source Table	Source Column
pk_coustomer_address_id_sk		
fk_customer_id	gravity_books_OLTP.dbo.customer_address	customer_id
fk_address_id	gravity_books_OLTP.dbo.customer_address	address_id
status_id_bk	gravity_books_OLTP.dbo.address_status	status_id
address_status	gravity_books_OLTP.dbo.address_status	address_status
source_system_code		
start_date		
end_date		
is_current		

Shipping_Method_Dim		
Destination Column	Source Table	Source Column
pk_shipping_method_id_sk		
pk_shipping_method_id_bk	gravity_books_OLTP.dbo.shipping_method	method_id
shipping_method_name	gravity_books_OLTP.dbo.shipping_method	method_name
cost	gravity_books_OLTP.dbo.shipping_method	cost
source_system_code		
start_date		
end_date		
is_current		

Author_Dim		
Destination Column	Source Table	Source Column
pk_author_id_sk		
pk_author_id_bk	gravity_books_OLTP.dbo.author	author_id
author_name	gravity_books_OLTP.dbo.author	author_name
source_system_code		
start_date		
end_date		
is_current		

Book_Dim		
Destination Column	Source Table	Source Column
pk_book_id_sk		
pk_book_id_bk	gravity_books_OLTP.dbo.book	book_id
title	gravity_books_OLTP.dbo.book	title
isbn13	gravity_books_OLTP.dbo.book	isbn13
num_pages	gravity_books_OLTP.dbo.book	num_pages
language_id_bk	gravity_books_OLTP.dbo.book_language	language_id
language_code	gravity_books_OLTP.dbo.book_language	language_code
language_name	gravity_books_OLTP.dbo.book_language	language_name
publication_date	gravity_books_OLTP.dbo.book	publication_date
publisher_id_bk	gravity_books_OLTP.dbo.publisher	publisher_id
publisher_name	gravity_books_OLTP.dbo.publisher	publisher_name
source_system_code		
start_date		
end_date		
is_current		



Book_Author_Dim		
Destination Column	Source Table	Source Column
pk_book_author_id_sk		
fk_author_id	gravity_books_OLTP.dbo.book_author	pk_author_id_sk
fk_book_id	gravity_books_OLTP.dbo.book_author	pk_book_id_sk
source_system_code		
start_date		
end_date		
is_current		

## Queries

```
Oreder_History_Facts.sql  crowd.py
> Users > ELHADY > Desktop > gravity_books > Final > SSIS Queries > Oreder_History_Facts.sql

1  SELECT
2      cust_order.order_date,
3      cust_order.order_id,
4      customer.customer_id,
5      MAX(order_history.history_id) AS history_id,
6      MAX(order_history.status_id) AS status_id,
7      MAX(order_status.status_value) AS status_value,
8      MAX(CASE WHEN order_history.status_id = 1 THEN CAST(order_history.status_date AS DATE) END) AS 'Order Received',
9      MAX(CASE WHEN order_history.status_id = 2 THEN CAST(order_history.status_date AS DATE) END) AS 'Pending Delivery',
10     MAX(CASE WHEN order_history.status_id = 3 THEN CAST(order_history.status_date AS DATE) END) AS 'Delivery In Progress',
11     MAX(CASE WHEN order_history.status_id = 4 THEN CAST(order_history.status_date AS DATE) END) AS 'Delivered',
12     MAX(CASE WHEN order_history.status_id = 5 THEN CAST(order_history.status_date AS DATE) END) AS 'Cancelled',
13     MAX(CASE WHEN order_history.status_id = 6 THEN CAST(order_history.status_date AS DATE) END) AS 'Returned',
14     shipping_method.method_id,
15     shipping_method.cost
16 FROM
17     cust_order
18 INNER JOIN
19     customer ON cust_order.customer_id = customer.customer_id
20 INNER JOIN
21     shipping_method ON cust_order.shipping_method_id = shipping_method.method_id
22 INNER JOIN
23     order_history ON cust_order.order_id = order_history.order_id
24 INNER JOIN
25     order_status ON order_history.status_id = order_status.status_id
26 GROUP BY
27     cust_order.order_date,
28     cust_order.order_id,
```

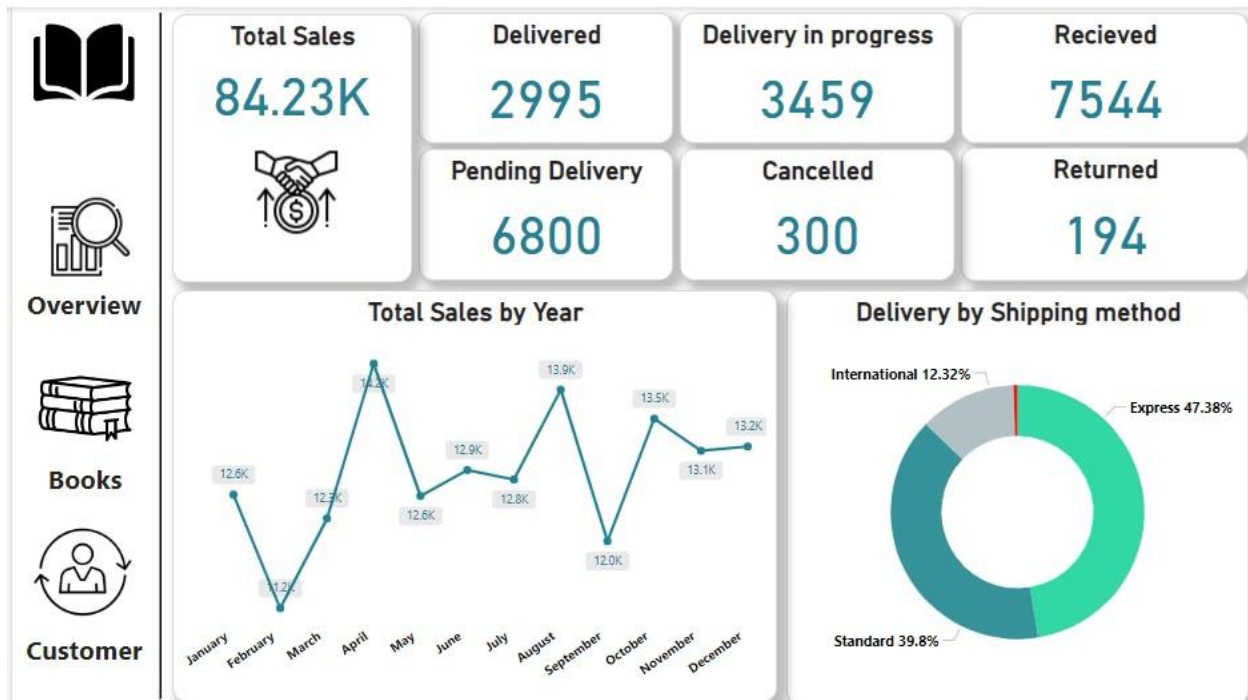
```
SELECT    b.book_id, ol.line_id,
          co.order_date,
          c.customer_id,
          ol.price
FROM      book AS b INNER JOIN
          order_line AS ol ON b.book_id = ol.book_id INNER JOIN
          cust_order AS co ON ol.order_id = co.order_id INNER JOIN
          customer AS c ON co.customer_id = c.customer_id
```

```
SELECT    book.book_id,
          book.title,
          book.isbn13,
          book.language_id,
          book.num_pages,
          book.publication_date,
          book.publisher_id,
          book_language.language_code,
          book_language.language_name,
          publisher.publisher_name
FROM      book left JOIN
          book_language ON book.language_id = book_language.language_id left JOIN
          publisher ON book.publisher_id = publisher.publisher_id
```



```
SELECT address.address_id,  
       address.street_number,  
       address.street_name,  
       address.city,  
       country.country_id,  
       country.country_name  
FROM   address LEFT OUTER JOIN  
       country ON address.country_id = country.country_id
```

## Dashboard



- **Total Sales:** 84.23K
- **Delivered:** 2995 orders have been successfully delivered.
- **Pending Delivery:** 6800 orders are awaiting delivery.
- **Delivery in Progress:** 3459 orders are currently in the delivery process.
- **Cancelled:** 300 orders were cancelled.
- **Received:** 7544 items have been received by customers.
- **Returned:** 194 items have been returned.
- **Total Sales by Year:** A line graph displaying monthly sales data for the year.
  - Sales show fluctuations, with a peak in August (13.9K) and lower sales in January (12.6K).
- **Delivery by Shipping Method:** A pie chart showing the distribution of delivery methods:
  - **Express:** 47.38%
  - **Standard:** 39.8%
  - **International:** 12.32%



- **Count of Books:** 11.13K total books in the system.
- **Languages:** 27 different languages available.
- **Publishers:** 2251 different publishers.
- **Authors:** 9235 different authors are represented.

#### Count of Books per Language:

- English leads with 8911 books, followed by US English (1409), Spanish (218), British English (214), and French (144)

**Sum of Price by Title:** A bar chart showing prices by book titles, where popular titles like *The Odyssey*, *Salem's Lot*, and *Anna Karenina* are priced around 149-121 units.

**Count of Books by Publisher:** A bar chart showing books distributed by publishers

- **Vintage** leads with 318 books, followed by **Penguin Books** (261), **Penguin Classics** (184), and other major publishers like **Mariner Books** (150) and **HarperCollins** (111).



Customers: A total of 2000 customers.

- Customers per Shipping Method:** A bar chart showing customer distribution by shipping method:
  - Express:** 3.5K customers.
  - Standard:** 3.0K customers.
  - International:** 1.0K customers.
  - Priority:** 0.1K customers.
- Customers per Country:** A bar chart showing customer distribution by country:
  - China** has the largest customer base (321), followed by **Indonesia** (217), **Russia** (122), **Brazil** (71), and smaller distributions across other countries like **Philippines**, **Poland**, and **USA**.
- Sum of Cost by Shipping Method:** A bar chart showing costs by shipping method:
  - International:** 25 units.
  - Express:** 12 units.
  - Priority:** 9 units.
  - Standard:** 6 units.
- Orders per Customer:** A bar chart and line graph showing orders per customer. The graph indicates steady customer orders across various individuals, peaking at 393 order

