



SQL PORTFOLIO PROJECT

By: DIMPLE SHARMA



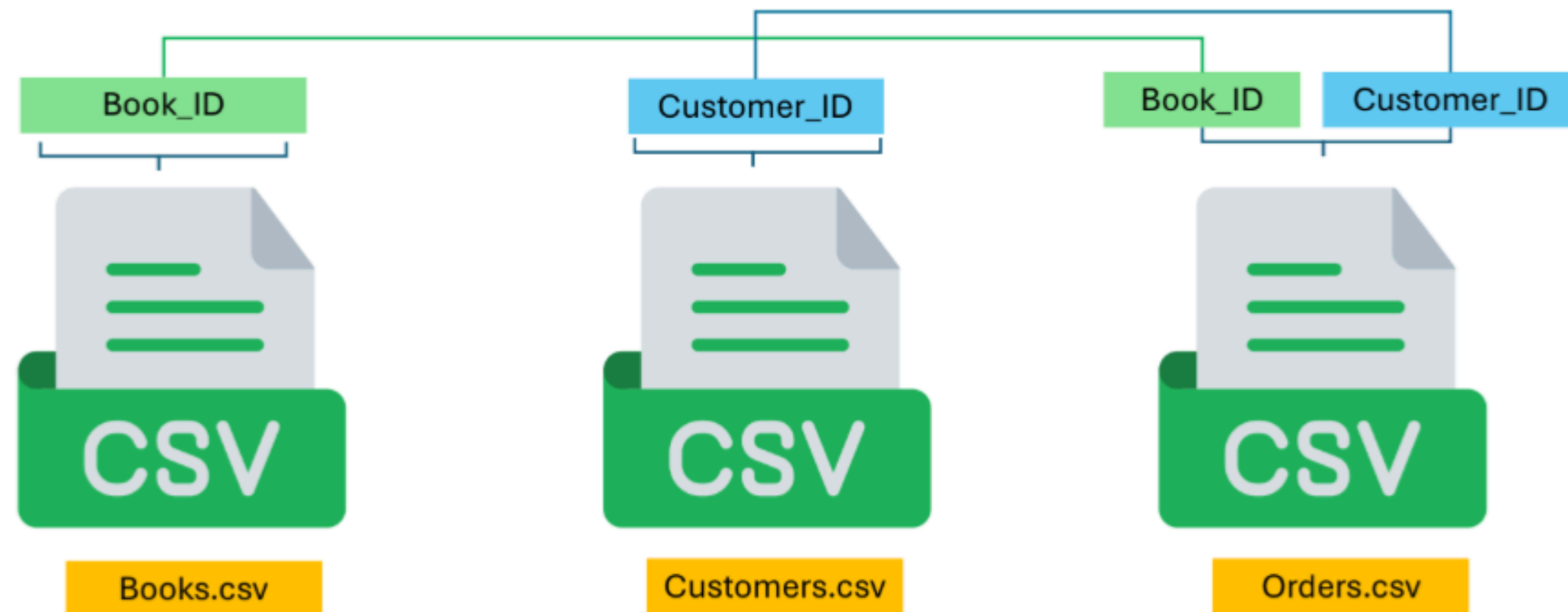
WHAT YOU'LL LEARN??

- How to design a real-world SQL project
- Creating and managing databases with multiple tables (Books, Customer & Orders table)
- Writing SQL queries to extract meaningful insights (both basic & advance queries)
- Using JOINS, Aggregations, CTEs, Window Functions, and Subqueries
- Structuring your project with data visualization

WHAT YOU'LL LEARN??

3 CSV Files

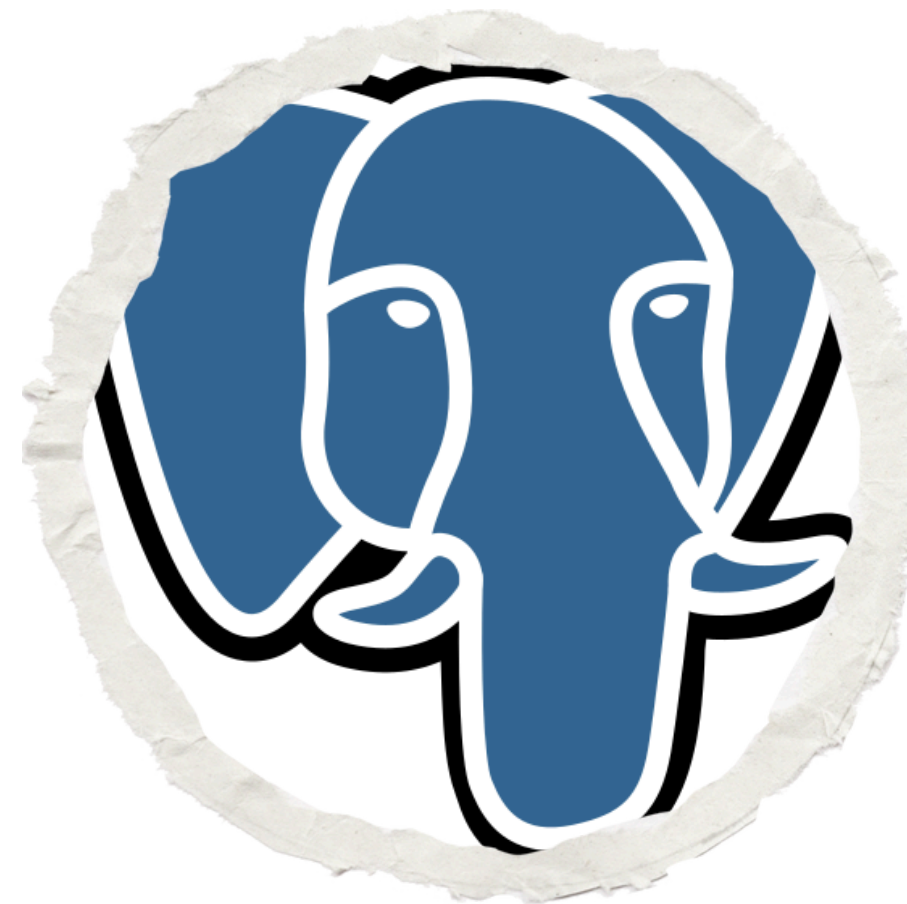
Tables must have at least one common column with same column name and same data type



TOOLS USED



MS-EXCEL



PgAdmin4

PART-1 (BASIC QUERIES)

- Retrieve all books in the "Fiction" genre
- Find books published after the year 1950
- List all customers from the Canada
- Show orders placed in November 2023
- Retrieve the total stock of books available
- Find the details of the most expensive book
- Show all customers who ordered more than 1 quantity of a book
- Retrieve all orders where the total amount exceeds \$20
- List all genres available in the Books table
- Find the book with the lowest stock
- Calculate the total revenue generated from all orders



BOOKS TABLE

	book_id [PK] integer	title character varying (100)	author character varying (100)	genre character varying (50)	published_year integer	price numeric (10,2)	stock integer
1	1	Configurable modular throughput	Joseph Crane	Biography	1949	21.34	100
2	2	Persevering reciprocal knowledge user	Mario Moore	Fantasy	1971	35.80	19
3	3	Streamlined coherent initiative	Derrick Howard	Non-Fiction	1913	15.75	27
4	4	Customizable 24hour product	Christopher Andrews	Fiction	2020	43.52	8
5	5	Adaptive 5thgeneration encoding	Juan Miller	Fantasy	1956	10.95	16
6	6	Advanced encompassing implementation	Bryan Morgan	Biography	1985	6.56	2
7	7	Open-architected exuding structure	Jacqueline Young	Romance	1927	43.63	95
8	8	Persistent local encoding	Troy Cox	Science Fiction	2019	48.99	84
9	9	Optimized interactive challenge	Colin Buckley	Fantasy	1987	14.33	70
10	10	Ergonomic national hub	Samantha Ruiz	Mystery	2015	24.63	25
11	11	Secured zero tolerance time-frame	Denise Barnes	Fantasy	1998	35.95	10
12	12	Polarized optimal array	Destiny Scott	Non-Fiction	1989	27.43	63
13	13	Adaptive 5thgeneration orchestration	Jaclyn Miller	Romance	1913	14.04	99
14	14	Re-engineered demand-driven parallelism	Jeremy Hayes	Science Fiction	1933	6.04	95
15	15	User-friendly motivating strategy	Keith Smith	Non-Fiction	1997	23.83	58
16	16	Vision-oriented tangible project	Christopher Price	Mystery	1941	10.07	8
17	17	Reduced secondary core	Benjamin Peters	Fantasy	1966	5.37	45
18	18	Adaptive 4thgeneration concept	Hector Palmer	Non-Fiction	2021	39.47	32
19	19	Progressive asymmetric Internet solution	Sean Miller	Science Fiction	1990	11.31	1
20	20	Face-to-face systematic throughput	Teresa Brennan	Non-Fiction	1978	48.13	64

```
SELECT * FROM Books;
```


CUSTOMER TABLE

	customer_id [PK] integer	name character varying (100)	email character varying (100)	phone character varying (15)	city character varying (50)	country character varying (150)
1	1	Deborah Griffith	balljoseph@wright-keith.net	1234567891	South Craigfort	Denmark
2	2	Crystal Clements	kimberlybennett@curtis.com	1234567892	East Derekberg	Nicaragua
3	3	Susan Fuller	beanmichael@burnett-stewart.net	1234567893	Austinbury	Equatorial Guinea
4	4	Jamie Ramirez	amandahood@warren.com	1234567894	Dianamouth	Slovenia
5	5	Marcus Murphy	connerjohn@yahoo.com	1234567895	Smithbury	Guinea-Bissau
6	6	Stephen Vasquez	ricemiguel@yahoo.com	1234567896	Hamiltonstad	Rwanda
7	7	Susan Hicks	jeffrey91@yahoo.com	1234567897	East Rebecca	Montenegro
8	8	Matthew Johnson	austinkenneth@manning.net	1234567898	Kirstenborough	Israel
9	9	Matthew Williams	jeffrey41@diaz.com	1234567899	Rebeccafurt	Somalia
10	10	Ronald Osborn	staciekelley@heath.com	1234567900	Lake Benjamin	Cameroon
11	11	Thomas Garcia	rmiller@gmail.com	1234567901	West Monicabury	Argentina
12	12	Jennifer Murray	wilsonbrittany@hotmail.com	1234567902	South Ashleychester	Zimbabwe
13	13	Kristine Kim	sarahwilliams@hotmail.com	1234567903	Lake Robert	Nigeria
14	14	John Wood	johnsonalexander@gmail.com	1234567904	Richardsonville	New Caledonia
15	15	Vanessa Gaines	tbullock@gmail.com	1234567905	Rodriguezmouth	Wallis and Futuna
16	16	Stacey Flores	jackjackson@hotmail.com	1234567906	East Michaelfurt	Saint Helena
17	17	Christine Maldonado	ztaylor@yahoo.com	1234567907	Smithborough	Chile
18	18	David Watson	laurenlawson@alvarez-mcfarland.com	1234567908	New Josephtown	Equatorial Guinea
19	19	Marc Nash	joshua59@gmail.com	1234567909	Davidton	French Guiana
20	20	Robert Salas	schultzjohn@thomas.com	1234567910	Schultzchester	Denmark

SELECT * FROM Customers;

ORDERS TABLE

	order_id [PK] integer	customer_id integer	book_id integer	order_date date	quantity integer	total_amount numeric (10,2)
1	1	84	169	2023-05-26	8	188.56
2	2	137	301	2023-01-23	10	216.60
3	3	216	261	2024-05-27	6	85.50
4	4	433	343	2023-11-25	7	301.21
5	5	14	431	2023-07-26	7	136.36
6	6	439	119	2024-10-11	5	249.40
7	7	195	467	2023-10-23	6	82.92
8	8	32	159	2024-05-07	4	144.84
9	9	109	407	2024-01-04	9	379.71
10	10	94	122	2024-07-09	4	123.00
11	11	131	206	2023-10-16	1	38.01
12	12	454	3	2024-06-17	2	31.50
13	13	420	180	2023-06-08	5	125.45
14	14	454	319	2023-08-24	2	85.22
15	15	127	479	2023-01-10	6	229.62
16	16	412	196	2023-10-06	8	53.52
17	17	462	481	2023-03-20	5	52.75
18	18	377	101	2024-08-07	4	193.96
19	19	496	60	2023-11-17	9	316.26
20	20	195	67	2023-07-14	1	10.23

SELECT * FROM Orders;

QUERY-1

	book_id [PK] integer	title character varying (100)	author character varying (100)	genre character varying (50)	published_year integer	price numeric (10,2)	stock integer
1	4	Customizable 24hour product	Christopher Andrews	Fiction	2020	43.52	8
2	22	Multi-layered optimizing migration	Wesley Escobar	Fiction	1908	39.23	78
3	28	Expanded analyzing portal	Lisa Coffey	Fiction	1941	37.51	79
4	29	Quality-focused multi-tasking challenge	Katrina Underwood	Fiction	1905	31.12	100
5	31	Implemented encompassing conglomerati...	Melissa Taylor	Fiction	2010	21.23	44
6	39	Optimized national process improvement	Megan Goodwin	Fiction	1978	10.99	42
7	40	Adaptive didactic interface	Natalie Gonzalez	Fiction	1923	25.97	94
8	47	Reverse-engineered directional conglomer...	John Christian	Fiction	2006	20.37	90
9	62	Re-contextualized real-time strategy	Nicole Lynch	Fiction	1953	26.34	23
10	63	Polarized heuristic database	Franklin Mack	Fiction	1989	22.38	56
11	100	Synchronized client-server service-desk	James Alvarado	Fiction	1906	49.89	29
12	116	Multi-tiered foreground contingency	Jamie Gates	Fiction	1938	41.82	50
13	125	Public-key analyzing Graphic Interface	Abigail Madden	Fiction	1990	32.41	16
14	130	Realigned context-sensitive pricing structure	Jason Rodriguez	Fiction	2004	6.64	90
15	134	Polarized bandwidth-monitored throughput	Linda Newman	Fiction	1955	35.72	49

Retrieve all books in the "Fiction" genre:

```
SELECT * FROM Books
WHERE Genre='Fiction';
```

QUERY-2

	book_id [PK] integer	title character varying (100)	author character varying (100)	genre character varying (50)	published_year integer	price numeric (10,2)	stock integer
1	2	Persevering reciprocal knowledge user	Mario Moore	Fantasy	1971	35.80	19
2	4	Customizable 24hour product	Christopher Andrews	Fiction	2020	43.52	8
3	5	Adaptive 5thgeneration encoding	Juan Miller	Fantasy	1956	10.95	16
4	6	Advanced encompassing implementation	Bryan Morgan	Biography	1985	6.56	2
5	8	Persistent local encoding	Troy Cox	Science Fiction	2019	48.99	84
6	9	Optimized interactive challenge	Colin Buckley	Fantasy	1987	14.33	70
7	10	Ergonomic national hub	Samantha Ruiz	Mystery	2015	24.63	25
8	11	Secured zero tolerance time-frame	Denise Barnes	Fantasy	1998	35.95	10
9	12	Polarized optimal array	Destiny Scott	Non-Fiction	1989	27.43	63
10	15	User-friendly motivating strategy	Keith Smith	Non-Fiction	1997	23.83	58
11	17	Reduced secondary core	Benjamin Peters	Fantasy	1966	5.37	45
12	18	Adaptive 4thgeneration concept	Hector Palmer	Non-Fiction	2021	39.47	32
13	19	Progressive asymmetric Internet solution	Sean Miller	Science Fiction	1990	11.31	1
14	20	Face-to-face systematic throughput	Teresa Brennan	Non-Fiction	1978	48.13	64
15	23	Reverse-engineered context-sensitive hardware	Christina Hernandez	Mystery	1967	38.55	70
16	25	Devolved mobile conglomeration	Alexander Bailey	Biography	1984	8.55	79
17	26	Multi-channelled multi-tasking capability	Patricia Buck	Science Fiction	1964	21.05	41
18	30	Multi-layered global open system	Jose Meyer	Biography	2012	30.58	37
19	31	Implemented encompassing conglomeration	Melissa Taylor	Fiction	2010	21.23	44
20	32	Synergistic dedicated concept	Lisa Bailey	Mystery	2021	21.56	100

Find books published after the year 1950:

```
SELECT * FROM Books
WHERE Published_year>1950;
```

QUERY-3

	customer_id [PK] integer	name character varying (100)	email character varying (100)	phone character varying (15)	city character varying (50)	country character varying (150)
1	38	Nicholas Harris	christine93@perkins.com	1234567928	Davistown	Canada
2	415	James Ramirez	robert54@hall.com	1234568305	Maxwelltown	Canada
3	468	David Hart	stokesrebecca@gmail.com	1234568358	Thompsonfurt	Canada

List all customers from the Canada:

```
SELECT * FROM Customers
WHERE country='Canada';
```

QUERY-4

Show orders placed in November 2023:

SELECT * FROM Orders

WHERE order_date BETWEEN '2023-11-01' AND '2023-11-30';

	order_id [PK] integer	customer_id integer	book_id integer	order_date date	quantity integer	total_amount numeric (10,2)
1	4	433	343	2023-11-25	7	301.21
2	19	496	60	2023-11-17	9	316.26
3	75	291	375	2023-11-30	5	170.75
4	132	469	333	2023-11-22	7	194.32
5	137	474	471	2023-11-25	8	363.04
6	163	207	384	2023-11-23	3	101.76
7	182	129	293	2023-11-01	7	125.51
8	200	313	303	2023-11-23	1	6.57
9	213	325	447	2023-11-17	7	253.75
10	231	22	384	2023-11-11	1	33.92
11	245	386	97	2023-11-01	9	411.66
12	252	405	387	2023-11-15	5	237.10
13	257	123	403	2023-11-06	1	15.01
14	288	6	128	2023-11-13	1	24.04
15	307	368	133	2023-11-17	1	20.96
16	322	270	112	2023-11-08	2	16.04
17	344	385	218	2023-11-25	5	26.80
18	389	485	391	2023-11-18	2	66.84
19	414	23	234	2023-11-10	1	7.15
20	429	449	146	2023-11-01	7	101.50








QUERY-5

	total_stock bigint
1	25056

Retrieve the total stock of books available:

```
SELECT SUM(stock) AS Total_Stock  
From Books;
```

QUERY-6

	book_id [PK] integer 	title character varying (100) 	author character varying (100) 	genre character varying (50) 	published_year integer 	price numeric (10,2) 	stock integer 
1	340	Proactive system-worthy orchestration	Robert Scott	Mystery	1907	49.98	88
2	155	Optimized content-based standardization	Timothy Adams	Science Fiction	1901	49.96	88
3	240	Stand-alone content-based hub	Lisa Ellis	Fantasy	1957	49.90	41

Find the details of the 3-most expensive books:

```
SELECT * FROM Books
ORDER BY Price DESC
LIMIT 3;
```


QUERY-7

Show all customers who ordered more than 5 quantities of books:

```
SELECT * FROM Orders
WHERE quantity>5;
```

	order_id [PK] integer	customer_id integer	book_id integer	order_date date	quantity integer	total_amount numeric (10,2)
1	1	84	169	2023-05-26	8	188.56
2	2	137	301	2023-01-23	10	216.60
3	3	216	261	2024-05-27	6	85.50
4	4	433	343	2023-11-25	7	301.21
5	5	14	431	2023-07-26	7	136.36
6	7	195	467	2023-10-23	6	82.92
7	9	109	407	2024-01-04	9	379.71
8	15	127	479	2023-01-10	6	229.62
9	16	412	196	2023-10-06	8	53.52
10	19	496	60	2023-11-17	9	316.26
11	25	265	250	2023-05-25	10	126.50
12	28	151	423	2024-11-27	8	205.04
13	29	305	446	2023-10-05	8	275.92
14	30	438	248	2023-09-23	6	281.94
15	31	386	279	2024-03-12	8	244.00
16	36	417	260	2024-01-21	9	446.31
17	39	488	422	2024-08-03	8	367.28
18	41	418	147	2024-05-05	8	361.60
19	43	199	305	2024-03-25	7	221.62
20	49	322	336	2024-08-14	7	168.56

QUERY-8

Retrieve all orders where the total amount exceeds \$50:


```
SELECT * FROM Orders
WHERE total_amount>50;
```

	order_id [PK] integer	customer_id integer	book_id integer	order_date date	quantity integer	total_amount numeric (10,2)
1	1	84	169	2023-05-26	8	188.56
2	2	137	301	2023-01-23	10	216.60
3	3	216	261	2024-05-27	6	85.50
4	4	433	343	2023-11-25	7	301.21
5	5	14	431	2023-07-26	7	136.36
6	6	439	119	2024-10-11	5	249.40
7	7	195	467	2023-10-23	6	82.92
8	8	32	159	2024-05-07	4	144.84
9	9	109	407	2024-01-04	9	379.71
10	10	94	122	2024-07-09	4	123.00
11	13	420	180	2023-06-08	5	125.45
12	14	454	319	2023-08-24	2	85.22
13	15	127	479	2023-01-10	6	229.62
14	16	412	196	2023-10-06	8	53.52
15	17	462	481	2023-03-20	5	52.75
16	18	377	101	2024-08-07	4	193.96
17	19	496	60	2023-11-17	9	316.26
18	21	356	287	2024-12-05	3	120.57
19	22	177	427	2024-06-10	3	54.00
20	23	119	301	2023-08-04	3	64.98

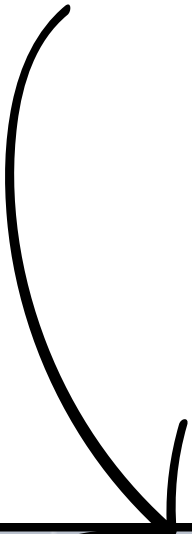
QUERY-9

List all genres available in the Books table:

```
SELECT DISTINCT genre FROM Books;
```

	genre character varying (50) 
1	Romance
2	Biography
3	Mystery
4	Fantasy
5	Fiction
6	Non-Fiction
7	Science Fiction

QUERY-10




	book_id [PK] integer	title character varying (100)	author character varying (100)	genre character varying (50)	published_year integer	price numeric (10,2)	stock integer
1	44	Networked systemic implementation	Ryan Frank	Science Fiction	1965	13.55	0

Find the book with the lowest stock:

```
SELECT * FROM Books
ORDER BY stock
LIMIT 1;
```

QUERY-11

	revenue numeric 
1	75628.66

Calculate the total revenue generated from all orders:

```
SELECT SUM(total_amount) As Revenue  
FROM Orders;
```

PART-2 (ADVANCE QUERIES)

- Retrieve the total number of books sold for each genre
- Find the average price of books in the "Fantasy" genre
- List customers who have placed at least 2 orders
- Find the most frequently ordered book
- Show the top 3 most expensive books of 'Fantasy' Genre
- Retrieve the total quantity of books sold by each author
- List the cities where customers who spent over \$30 are located
- Find the customer who spent the most on orders
- Calculate the stock remaining after fulfilling all orders



QUERY-1

Retrieve the total number of books sold for each genre:

```
SELECT * FROM ORDERS;
```

```
SELECT b.Genre, SUM(o.Quantity) AS Total_Books_sold  
FROM Orders o  
JOIN Books b ON o.book_id = b.book_id  
GROUP BY b.Genre;
```

	genre character varying (50) 🔒	total_books_sold bigint 🔒
1	Romance	439
2	Biography	285
3	Mystery	504
4	Fantasy	446
5	Fiction	225
6	Non-Fiction	351
7	Science Fiction	447

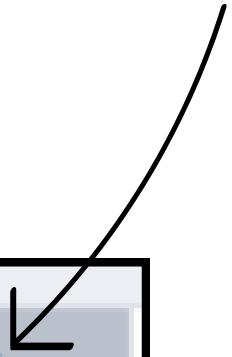
QUERY-2

	average_price numeric	
1	25.9816901408450704	

Find the average price of books in the "Fantasy" genre:

```
SELECT AVG(price) AS Average_Price  
FROM Books  
WHERE Genre = 'Fantasy';
```

QUERY-3



	customer_id integer	name character varying (100)	order_count bigint
1	225	Christopher Mccullough	2
2	418	Kiara Blankenship MD	3
3	322	William Cameron	3
4	325	Emily Vargas	4
5	376	Justin Donaldson	2

List customers who have placed at least 2 orders:

```
SELECT o.customer_id, c.name, COUNT(o.Order_id) AS ORDER_COUNT
      FROM orders o
     JOIN customers c ON o.customer_id=c.customer_id
    GROUP BY o.customer_id, c.name
   HAVING COUNT(Order_id) >=2;
```

QUERY-4

	book_id integer	title character varying (100)	order_count bigint
1	88	Robust tangible hardware	4

Find the most frequently ordered book:

```
SELECT o.Book_id, b.title, COUNT(o.order_id) AS ORDER_COUNT
FROM orders o
JOIN books b ON o.book_id=b.book_id
GROUP BY o.book_id, b.title
ORDER BY ORDER_COUNT DESC LIMIT 1;
```

QUERY-5



	book_id [PK] integer	title character varying (100)	author character varying (100)	genre character varying (50)	published_year integer	price numeric (10,2)	stock integer
1	240	Stand-alone content-based hub	Lisa Ellis	Fantasy	1957	49.90	41
2	462	Innovative 3rdgeneration database	Allison Contreras	Fantasy	1988	49.23	62
3	238	Optimized even-keeled analyzer	Sherri Griffith	Fantasy	1975	48.97	72

Show the top 3 most expensive books of 'Fantasy' Genre :

```
SELECT * FROM books
WHERE genre ='Fantasy'
ORDER BY price DESC LIMIT 3;
```

QUERY-6

Retrieve the total quantity of books sold by each author:

```
SELECT b.author, SUM(o.quantity) AS Total_Books_Sold
FROM orders o
JOIN books b ON o.book_id=b.book_id
GROUP BY b.Author;
```

	author character varying (100) 🔒	total_books_sold bigint 🔒
1	Jared Cortez	10
2	Tracy Parker	11
3	Taylor Wang	9
4	Cathy Knight	6
5	Bianca Matthews	3
6	Douglas Malone	6
7	James Alvarado	9
8	Betty Cross	6
9	Michael Hill	20
10	Steven Mcdonald	15

QUERY-7

List the cities where customers who spent over \$30 are located:

```
SELECT DISTINCT c.city, total_amount
FROM orders o
JOIN customers c ON o.customer_id=c.customer_id
WHERE o.total_amount > 30;
```

	city character varying (50) 🔒	total_amount numeric (10,2) 🔒
1	Taylorfurt	189.45
2	Leeport	141.39
3	Port Jasonview	149.12
4	Port Aaronstad	145.44
5	Matthewfurt	328.50
6	Angelaside	42.19
7	Lindaburgh	325.92
8	Stephanieberg	156.60
9	Freemanland	198.75
10	Natashaville	399.04
11	North Joseph	125.45
12	North Stephenmouth	281.94
13	Pamelaland	192.64
14	Pennyland	248.85
15	North Emily	31.68
16	West Kimberly	80.82
17	South Timothy	43.65
18	Port Nicoleview	127.12
19	West Robertport	241.75
20	Bradleyburgh	168.90

QUERY-8

	customer_id [PK] integer	name character varying (100)	total_spent numeric
1	457	Kim Turner	1398.90

Find the customer who spent the most on orders:

```
SELECT c.customer_id, c.name, SUM(o.total_amount) AS Total_Spent
      FROM orders o
     JOIN customers c ON o.customer_id=c.customer_id
    GROUP BY c.customer_id, c.name
   ORDER BY Total_spent Desc LIMIT 1;
```

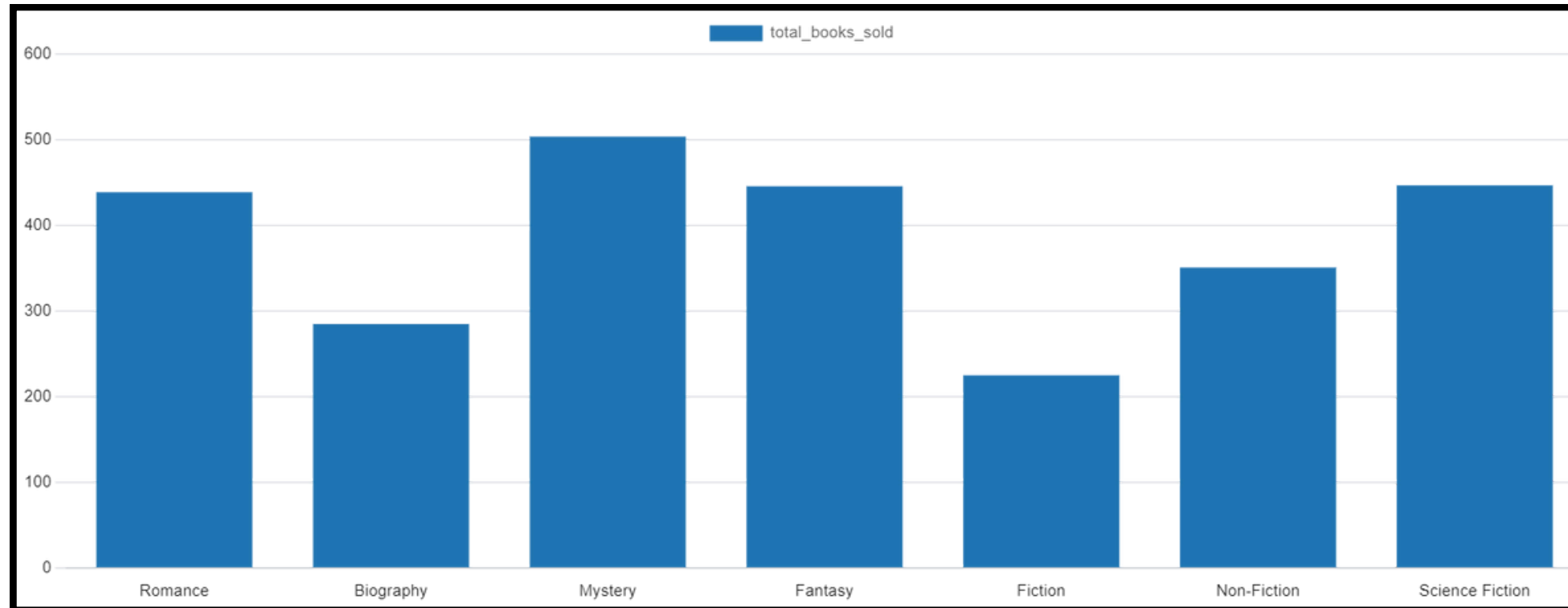
QUERY-9

	book_id [PK] integer	title character varying (100)	stock integer	order_quantity bigint	remaining_quantity bigint
1	1	Configurable modular throughput	100	3	97
2	2	Persevering reciprocal knowledge user	19	0	19
3	3	Streamlined coherent initiative	27	5	22
4	4	Customizable 24hour product	8	0	8
5	5	Adaptive 5thgeneration encoding	16	8	8
6	6	Advanced encompassing implementation	2	0	2
7	7	Open-architected exuding structure	95	5	90
8	8	Persistent local encoding	84	3	81
9	9	Optimized interactive challenge	70	0	70
10	10	Ergonomic national hub	25	1	24

Calculate the stock remaining after fulfilling all orders:

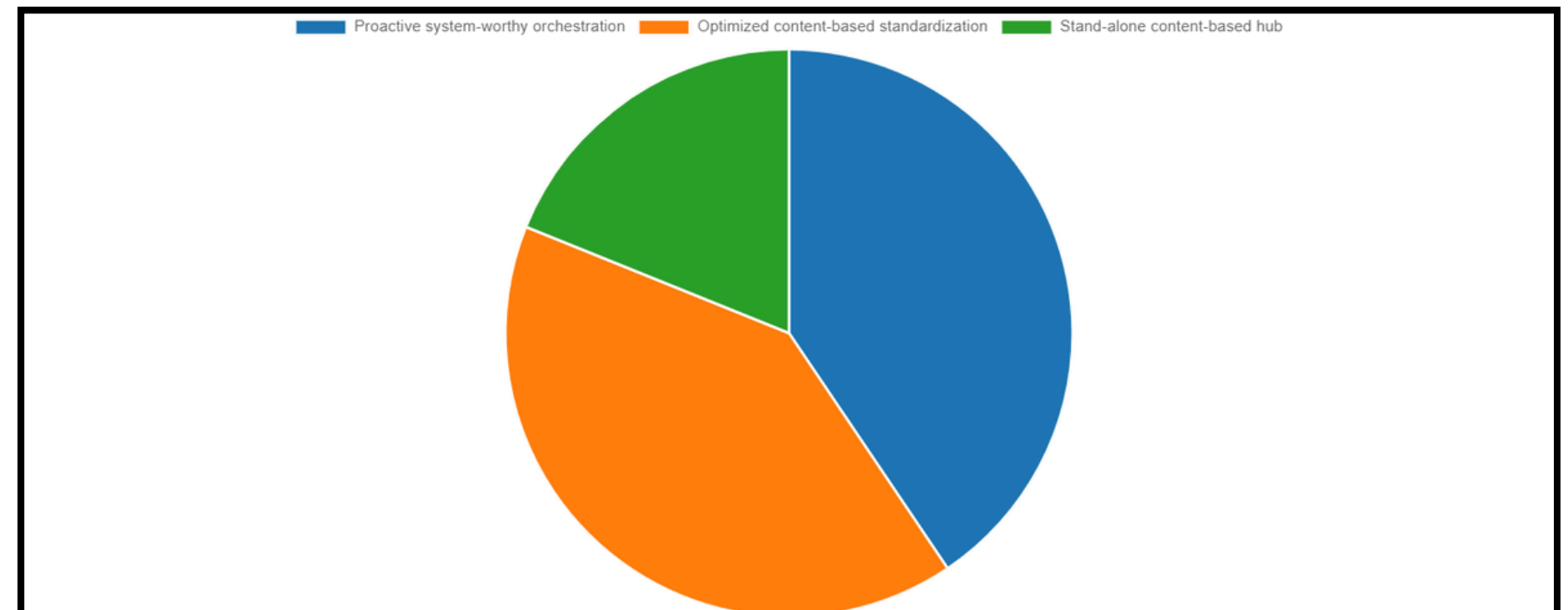
```
SELECT b.book_id, b.title, b.stock, COALESCE(SUM(o.quantity),0) AS
      Order_quantity,
      b.stock- COALESCE(SUM(o.quantity),0) AS Remaining_Quantity
      FROM books b
      LEFT JOIN orders o ON b.book_id=o.book_id
      GROUP BY b.book_id ORDER BY b.book_id;
```

GRAPH-VISUALIZERS

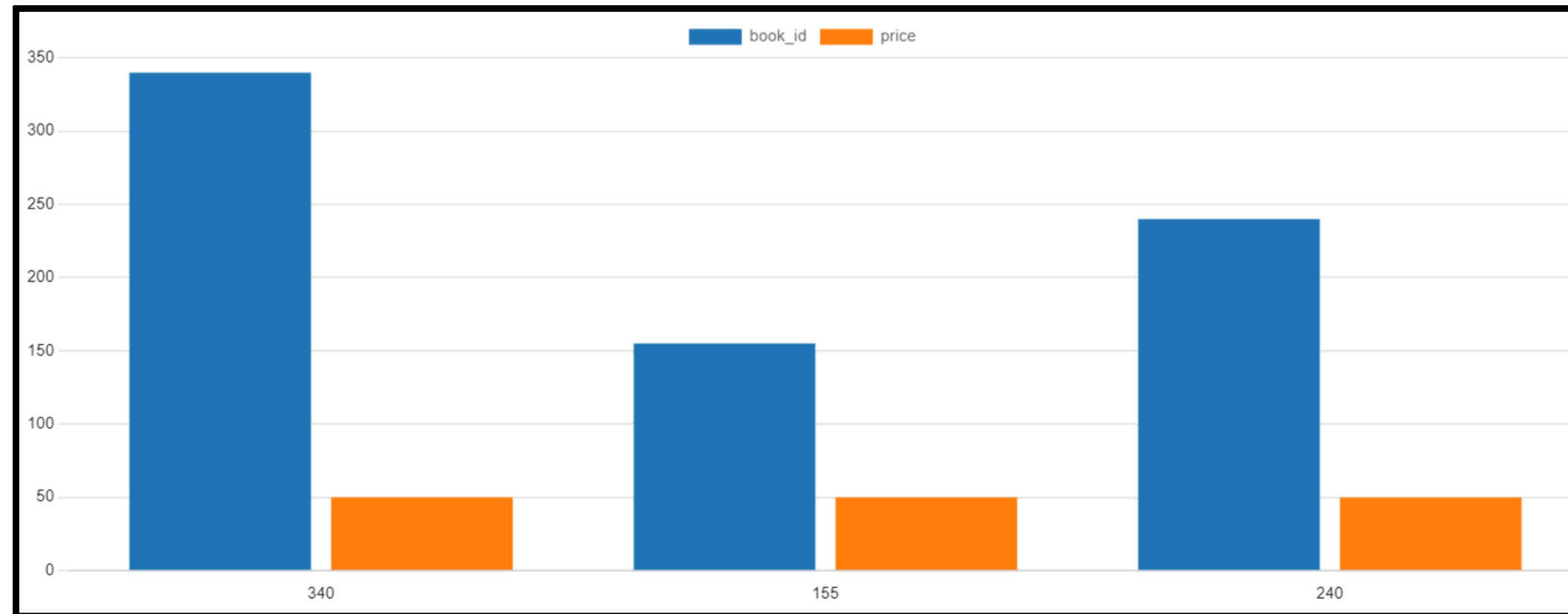


Total Books sold by
each Fiction

Stock of Top-3
expensive books

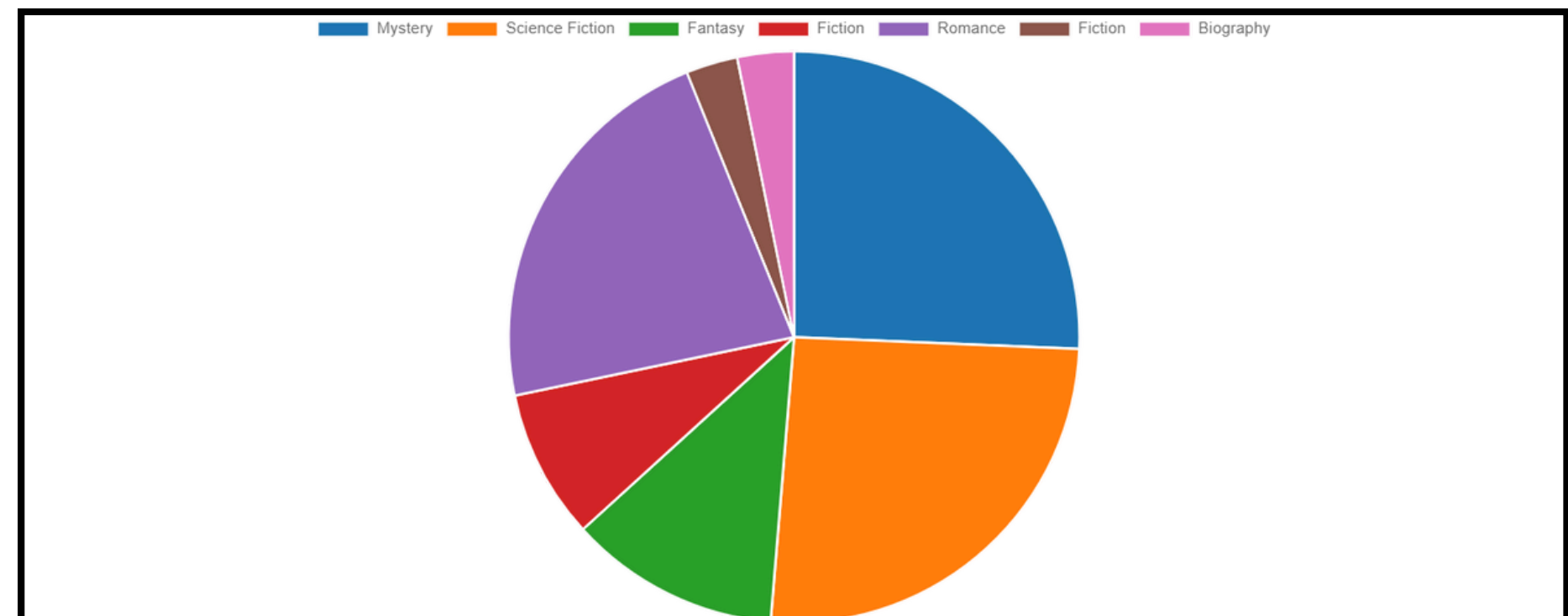


GRAPH-VISUALIZERS



Price varies by Fiction
& stock

Stock left of each
Fiction after fulfilling
all orders



CONCLUSION

- This SQL project has been a valuable learning experience, reinforcing my ability to write and execute efficient SQL queries for data management and analysis. By working extensively with SQL commands such as DDL, DML, DCL, and TCL, I have gained a deeper understanding of database operations and best practices.
- Through the use of various operators, date and time functions, and different types of joins, I successfully extracted meaningful insights from structured data. Additionally, optimizing queries and handling complex data retrieval tasks have enhanced my problem-solving skills and efficiency in working with relational databases.
- Overall, this project has strengthened my SQL proficiency and prepared me for real-world applications in data analytics and database management. The hands-on experience gained here will be instrumental in tackling future data-driven challenges.



THANK YOU