E-COMMERCE DATABASE

INTRODUCTION & PROJECT OVERVIEW

- Team Introduction
- Agenda

AGENDA

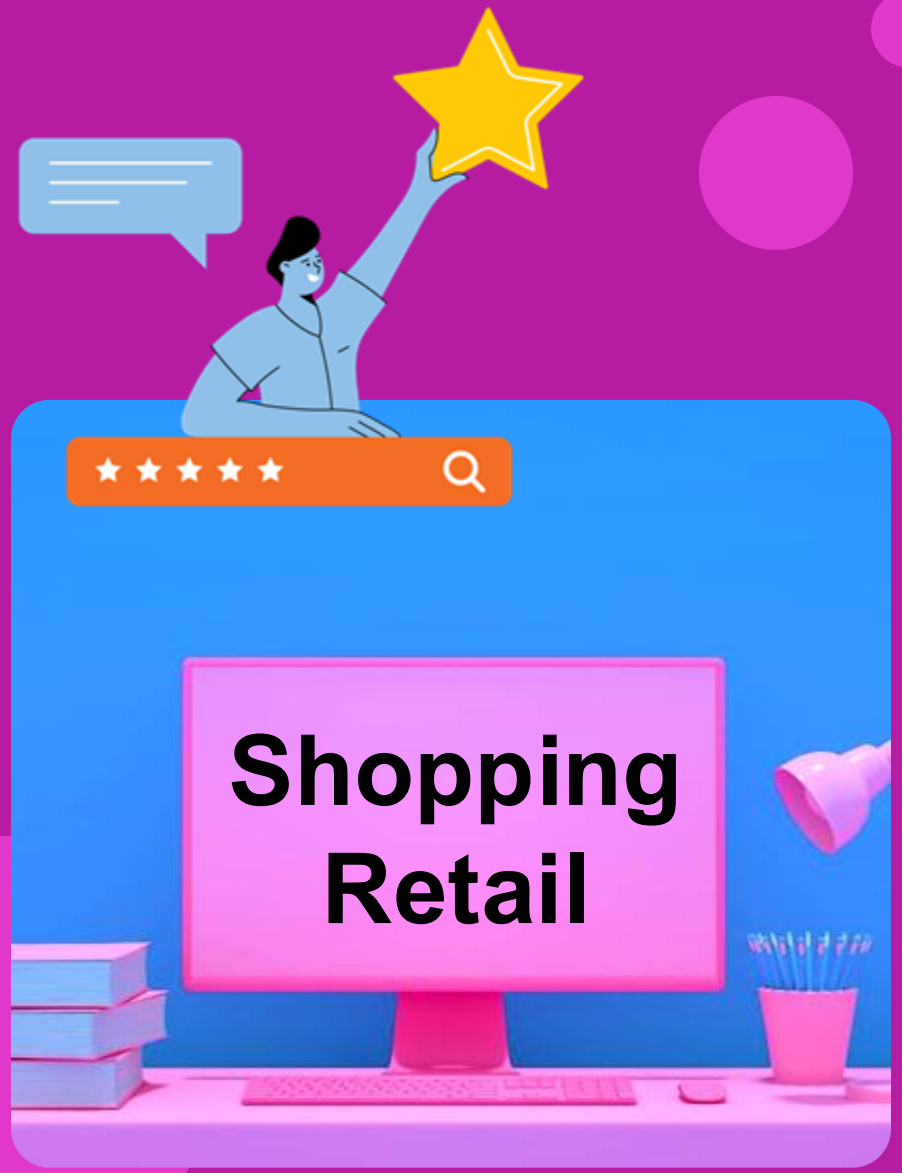
- Introduction & Project Overview
- Business Problem & Purpose
- Database Design & Requirements
- Implementation, Data Processing & Visualization
- Improvements & Final Takeaways
- Conclusion & Q&A

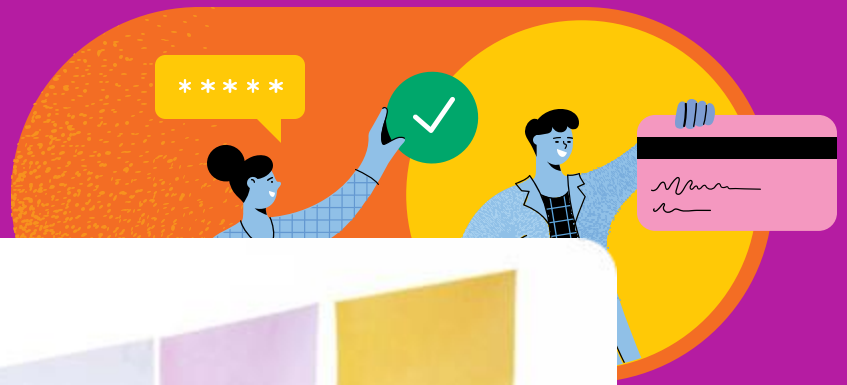
ABOUT US

Online Shopping Data

Women's Clothing & Accessories

Jess Madson	Customer / Order_Header / Order_Detail
Ali Lo	Shipment / Contact / Product
Natasha Sumatra	Supplier / Inventory / Warehouse
Yue Liu	Date / Location / Product_Category
Minju Kim	Payment / Return / Review





BUSINESS & PURPOSE

- Business Purpose
- Business Problem & Solutions

Return Management

Payment Management

Review Management

BUSINESS PURPOSE:
WHY DID WE BUILD THIS DATABASE?

Customer Places Order

Payment Processed

Order Shipped

Review Submitted

Return (if necessary)

BUSINESS PROCESS FLOW

Enhancing **personalized** shopping **experiences**

Optimizing return process for **better customer satisfaction**

Leveraging review data to track **popular products and identify areas** for improvement

BUSINESS VALUE OF THIS DATA

DATABASE DESIGN & REQUIREMENTS

- Business Requirements & Rules
- Final ERD Overview

Business Requirements & Rules

Customer & Orders:

- Every customer must be associated with at least one transaction
- A customer can place multiple orders
- Each order has a unique order number
- Orders can include multiple products
- Orders and payments occur on the same day

Shipping & Inventory:

- Each order ships to one customer
- A warehouse stores multiple products
- Orders are shipped in full; no partial shipments

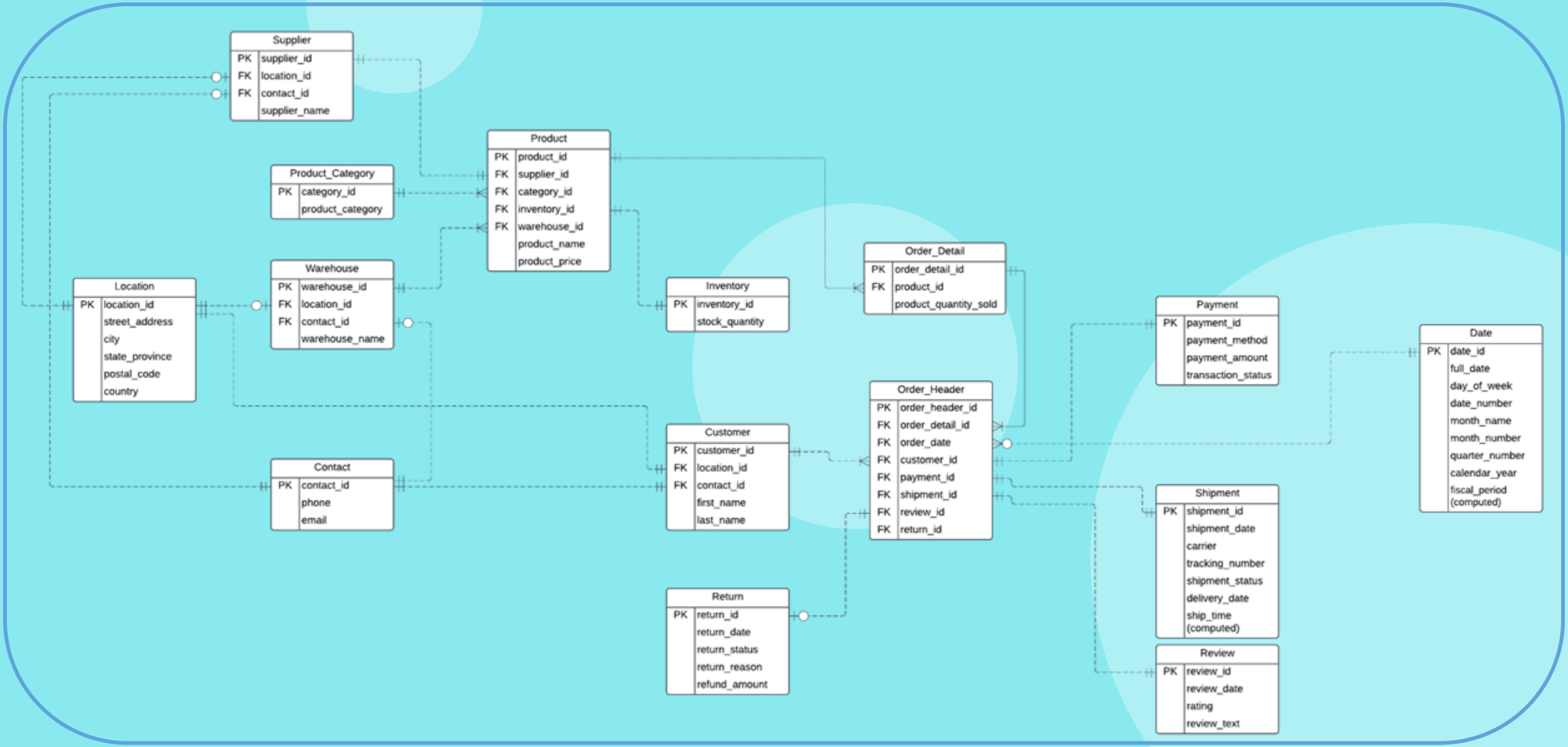
Suppliers & Products:

- Each supplier provides only one type of product
- Products are linked to suppliers, warehouses, and inventory

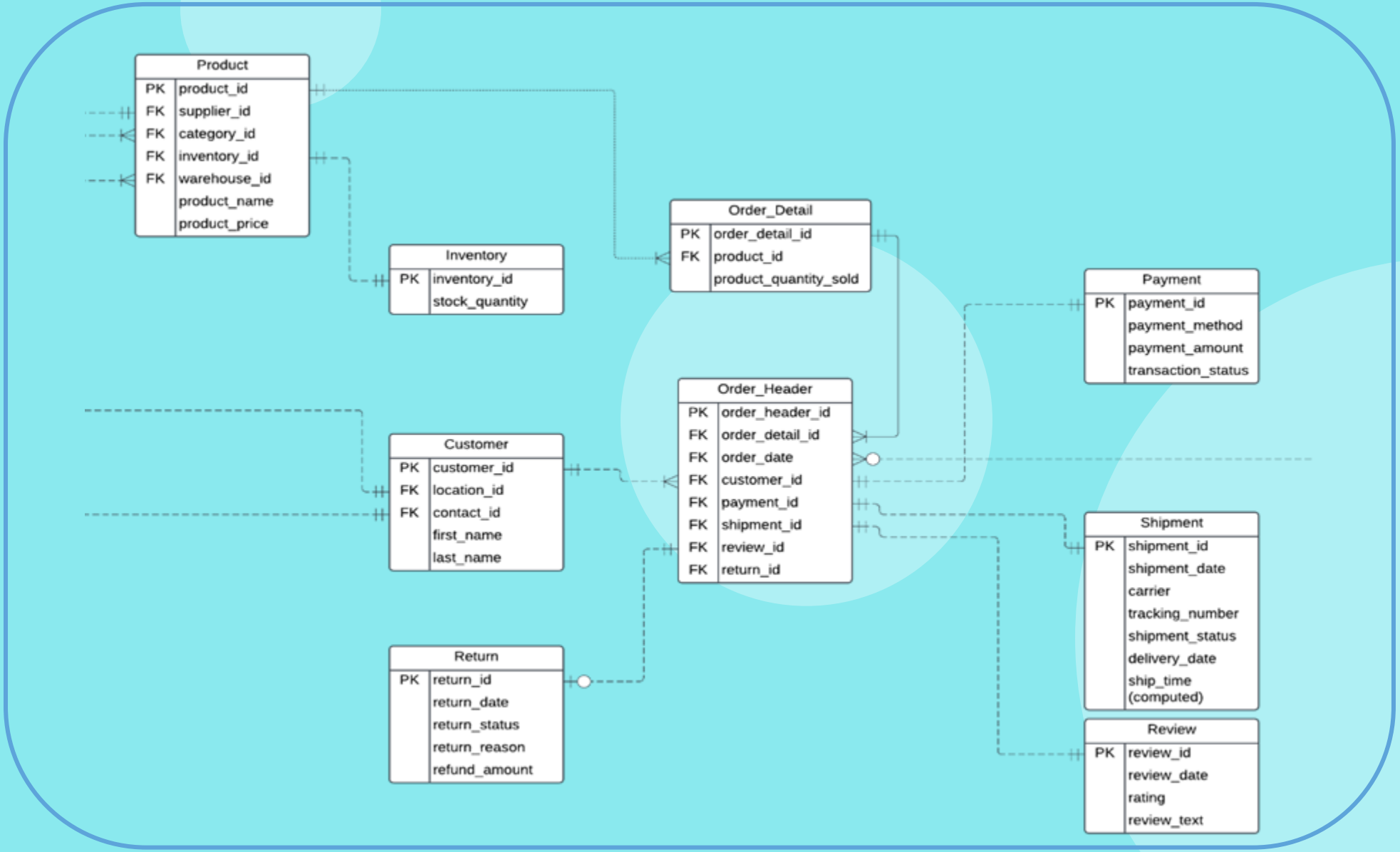
Returns & Refunds:

- Returned orders are refunded on the same day they are returned

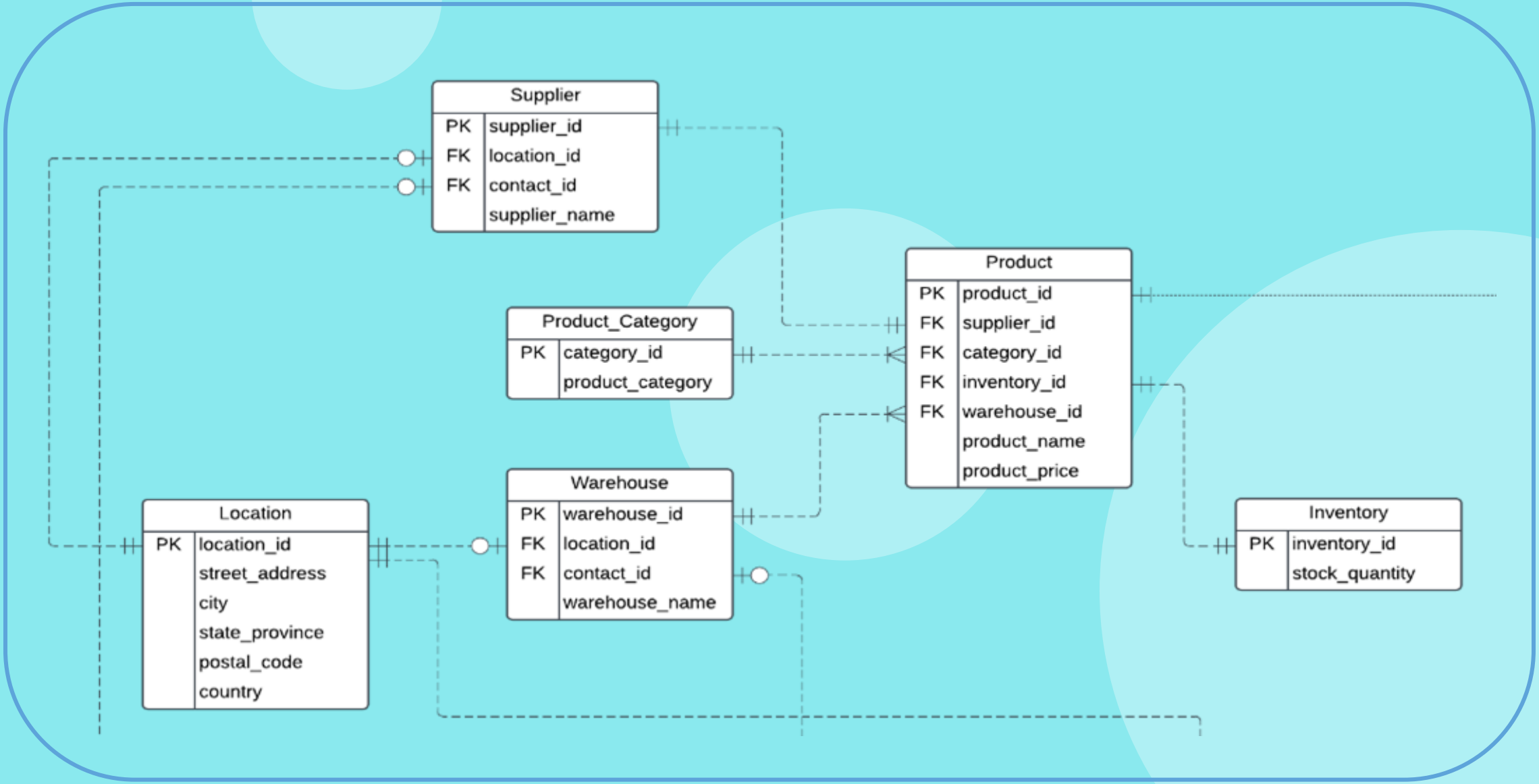
FINAL ERD OVERVIEW



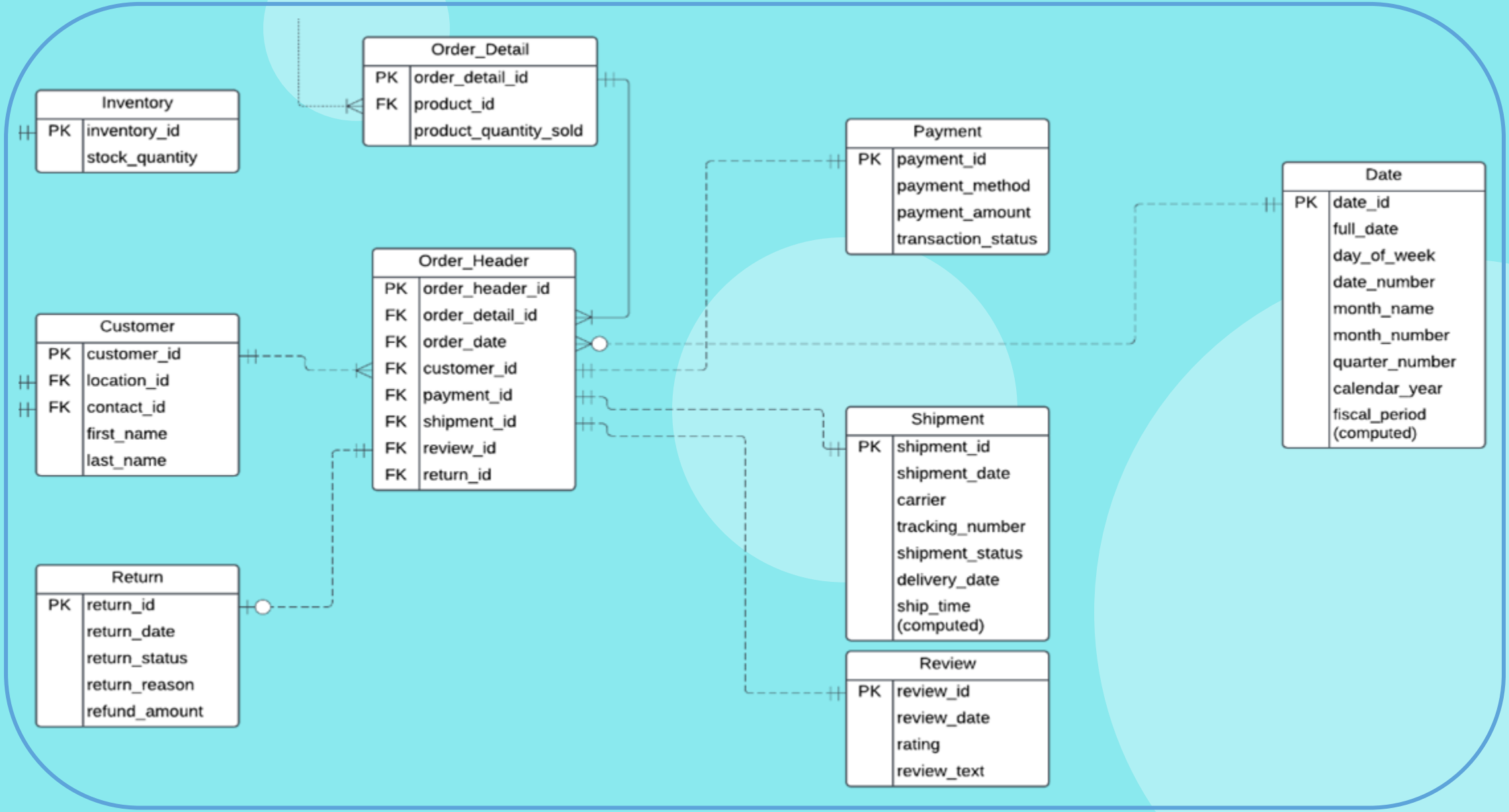
FINAL ERD OVERVIEW

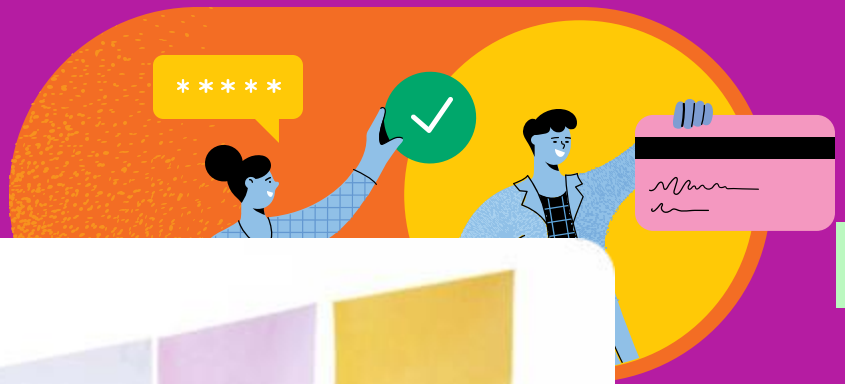


FINAL ERD OVERVIEW



FINAL ERD OVERVIEW





IMPLEMENTATION, DATA PROCESSING, & VISUALIZATION

- Implement Tables into SQL Server
- Visualization
- Key Findings from Data

WE USE CONSTRAINTS AND CHECKS SPECIFY THE TYPE OF DATA

123 quarter_number	123 calendar_year	A-Z fiscal_period
4	2,023	Q4-2023
3	2,020	Q3-2020
1	2,021	Q1-2021
4	2,019	Q4-2019
1	2,020	Q1-2020
3	2,018	Q3-2018
3	2,020	Q3-2020
2	2,018	Q2-2018
2	2,020	Q2-2020
2	2,020	Q2-2020
3	2,020	Q3-2020
3	2,024	Q3-2024

WE ALSO USED COMPUTED COLUMNS TO BETTER UNDERSTAND THE DATA

Shipment | Enter a SQL expression to filter results (use Ctrl+Space)

Grid | Text

	shipment_id	shipment_date	carrier	tracking_number	shipment_status	delivery_date	ship_time
1	1	2024-10-13	UPS	TRK344284	Pending	2024-10-14	1
2	2	2024-08-27	DHL	TRK718398	Delivered	2024-08-28	1
3	3	2024-05-23	FedEx	TRK161368	Pending	2024-05-31	8
4	4	2024-06-06	UPS	TRK890181	Pending	2024-06-10	4
5	5	2024-11-03	DHL	TRK681341	Delivered	2024-11-08	5
6	6	2024-05-03	FedEx	TRK076623	Pending	2024-05-07	4
7	7	2023-12-28	UPS	TRK767199	Delivered	2023-12-30	2
8	8	2023-11-14	DHL	TRK930895	Delivered	2023-11-16	2
9	9	2024-10-25	FedEx	TRK502345	Delivered	2024-11-01	7

WE CREATED FKs AND PKs TO HIGHLIGHT THE RELATIONSHIPS BETWEEN TABLE

Product | Enter a SQL expression to filter results (use Ctrl+Space)

Grid | Text

	product_id	supplier_id	category_id	inventory_id	warehouse_id	product_name	product_price
1	1	501	5,001	701	801	Floral_Dress	59.9900
2	2	501	5,001	702	801	Striped_Dress	49.9900
3	3	501	5,001	703	801	Short_Dress	54.9900
4	4	501	5,001	704	801	Denim_Dress	49.9900
5	5	501	5,001	705	801	Maxi_Dress	64.9900
6	6	501	5,001	706	801	Sun_Dress	38.0000
7	7	502	5,002	707	801	TShirt	15.0000
8	8	502	5,002	708	801	Long_Sleeve	19.9900
9	9	502	5,002	709	801	Cropped_Shirt	12.9900
10	10	502	5,002	710	801	Tank_Top	13.9500
11	11	502	5,002	711	801	Blouse	44.9900
12	12	502	5,002	712	801	Button_Up	35.0000

WE CREATED MULTIPLE VIEWS IN RESPONSE TO OUR ‘USERS’ AND BUSINESS QUESTIONS

- Views
  - CustomerOrderHistory
  - CustomerOrderSummary
  - DateDimensionReport
  - InventoryStock
  - LocationsByCountry
  - ProductSales
  - ReturnAndReview
  - SecureLocationView
  - ShipmentTracking
  - SupplierProductList

PRODUCT SALES HELPS TRACK THE QUANTITY SOLD BY PRODUCT

ProductSales | Enter a SQL expression to filter results (use Ctrl+Space)

Grid | Text

	product_id	product_name	product_category	TotalQuantitySold	TotalRevenue
1	22	Headband	Accessories	12	180.0000
2	23	Socks	Accessories	13	64.8700
3	24	Small_Purse	Accessories	3	450.0000
4	25	Hair_Clip	Accessories	4	79.9600
5	26	Hoop_Earrings	Accessories	9	675.0000
6	27	Bracelet	Accessories	9	719.9100
7	40	Active_Shirt	Activewear	5	449.9500
8	41	Tennis_Skirt	Activewear	9	1195.1100
9	42	Yoga_Pants	Activewear	18	414.0000
10	43	Tennis_Dress	Activewear	18	1152.0000
11	45	Pilates_Shirt	Activewear	7	203.0000
12	15	Sweats	Bottoms	9	450.0000
13	16	Sweat_Shorts	Bottoms	1	19.9900
14	17	Denim_Shorts	Bottoms	15	525.0000

SHIPMENT TRACKING HELPS INDIVIDUALS MANAGING SHIPPING

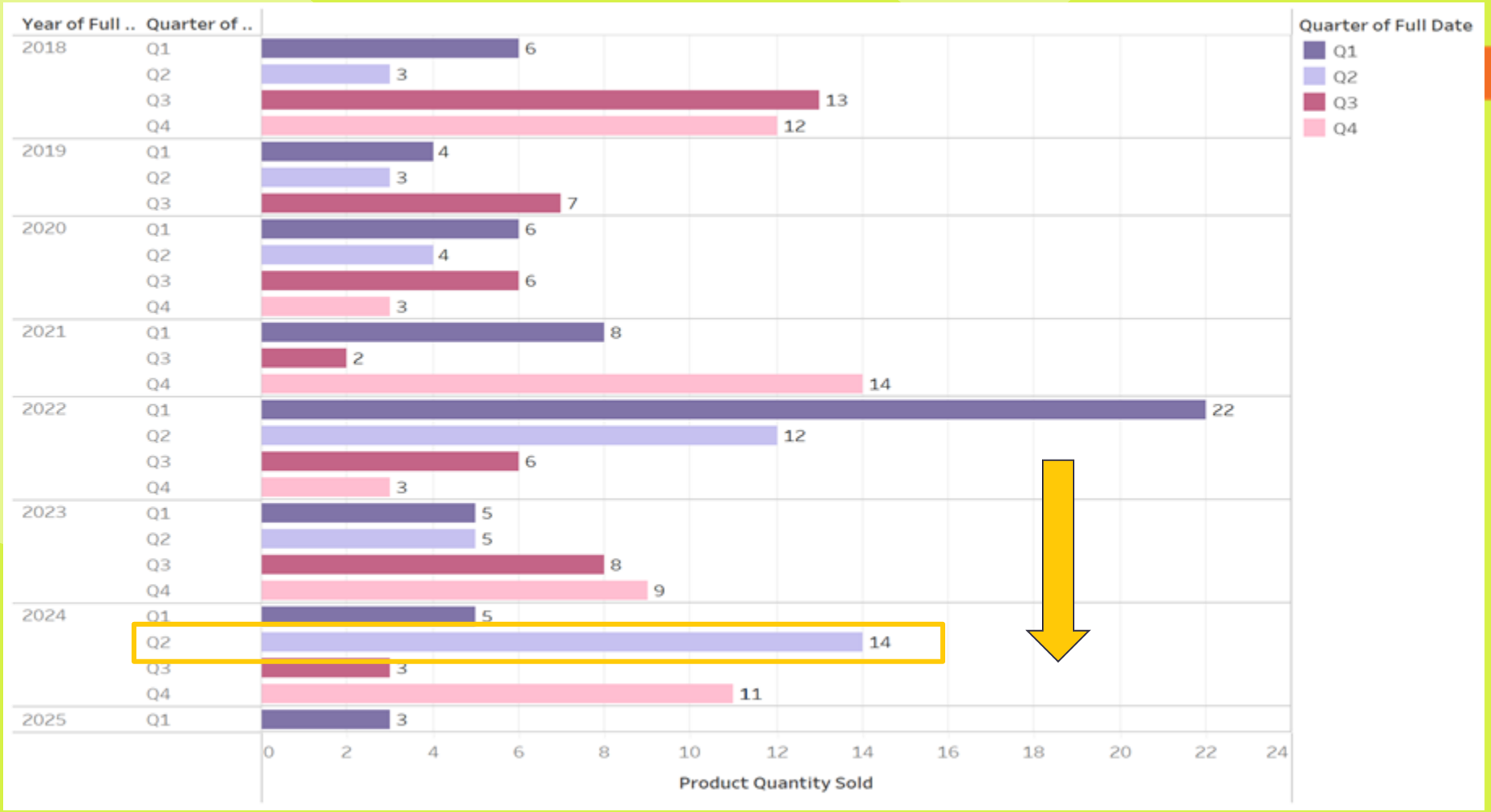
ShipmentTracking | Enter a SQL expression to filter results (use Ctrl+Space)

Grid | Text

	shipment_id	order_header_id	customer_id	CustomerName	shipment_status	delivery_date	ShipmentProg
1	9,555	1	277	Jacques Miller	Pending	2024-10-27	Pending
2	204	2	217	Juliette Taylor	Pending	2024-03-28	Pending
3	3,606	3	206	Julia Garcia	Pending	2024-04-22	Pending
4	6,191	4	296	Jo Grace	Pending	2024-04-08	Pending
5	5,291	5	255	Joel Benjamin	Pending	2024-04-13	Pending
6	5,734	6	208	Josaphine Davis	Shipped	2024-08-10	Shipped
7	9,540	7	285	Jo-ann Anderson	Delivered	2024-04-06	Completed
8	13,830	8	284	Jes Wilson	Pending	2024-02-25	Pending
9	14,617	9	216	June Thomas	Cancelled	2024-10-23	Cancelled
10	7,040	10	257	Jonah Bruce	Shipped	2024-08-07	Shipped
11	10,150	11	282	Jordi Lopez	Delivered	2023-12-13	Completed
12	13,190	12	222	Jean Victoria	Delivered	2024-04-18	Completed

YEARLY AND QUARTERLY SALES PERFORMANCE

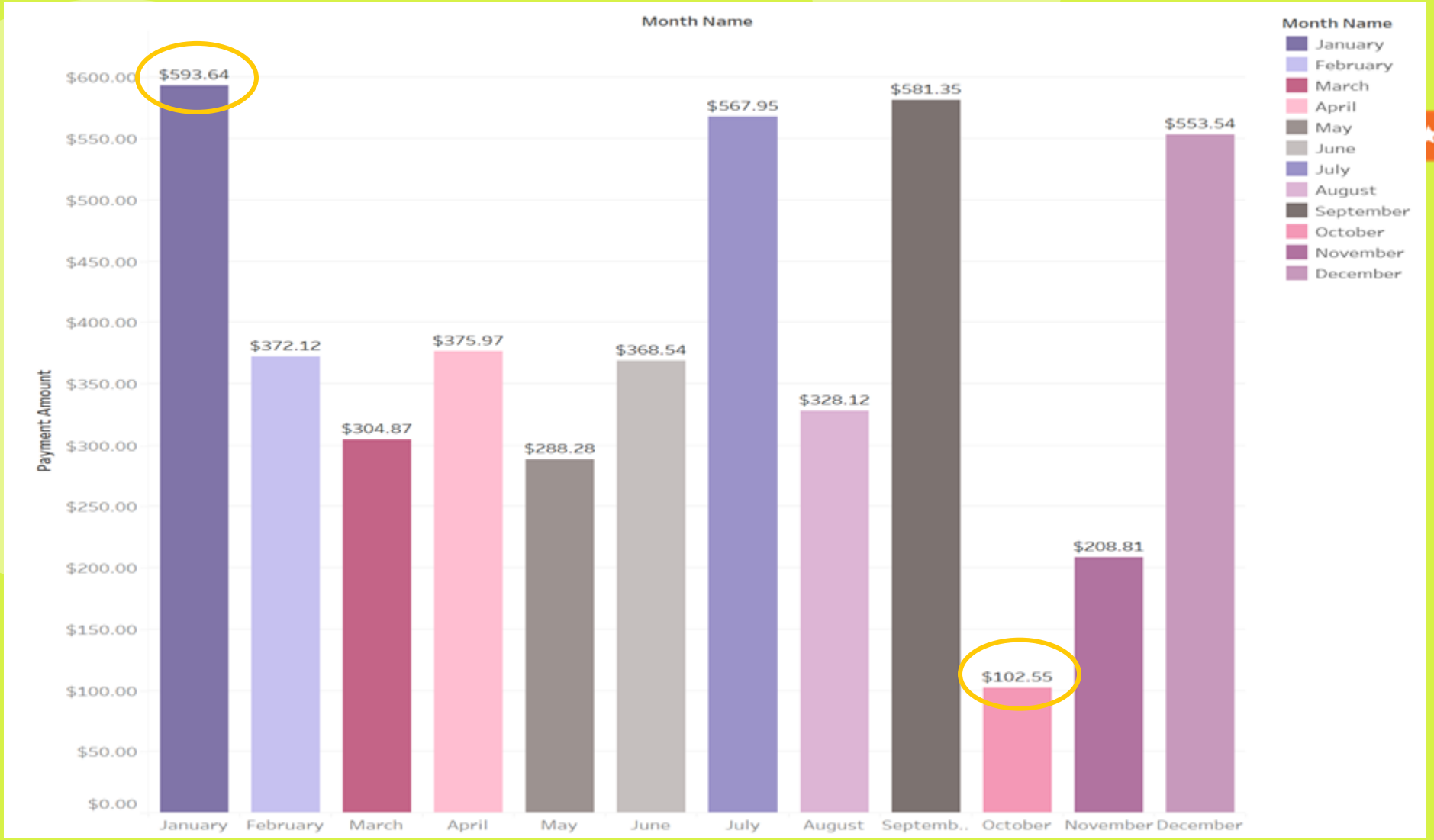
Product Quantity Sold per Quarter per Year (2018-2025)



- Sales performance has generally **declined from 2022 onward**, particularly in Q1 and Q4
- **Strong sales in Q2** of 2024 suggest an opportunity to capitalize on this peak period

MONTHLY SALES PERFORMANCE

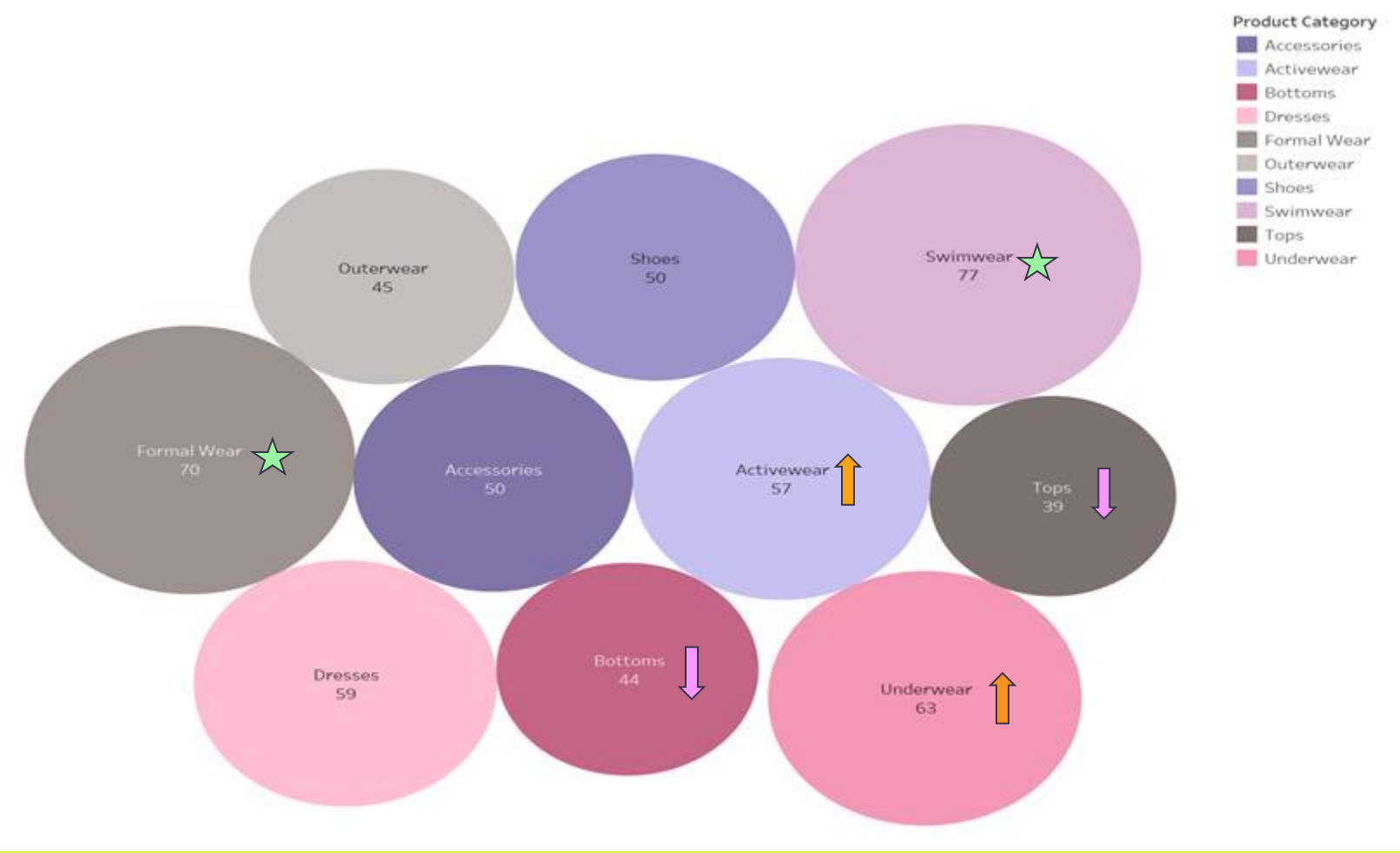
Total Sales for each Month (2018-2025)



- Strong sales in **January, July, and September**
- Sales in **October and November are notably lower** compared to other months due to our off-season period
- Fluctuating mid-year sales between March and August

TOTAL SALES VOLUME BY PRODUCT CATEGORY

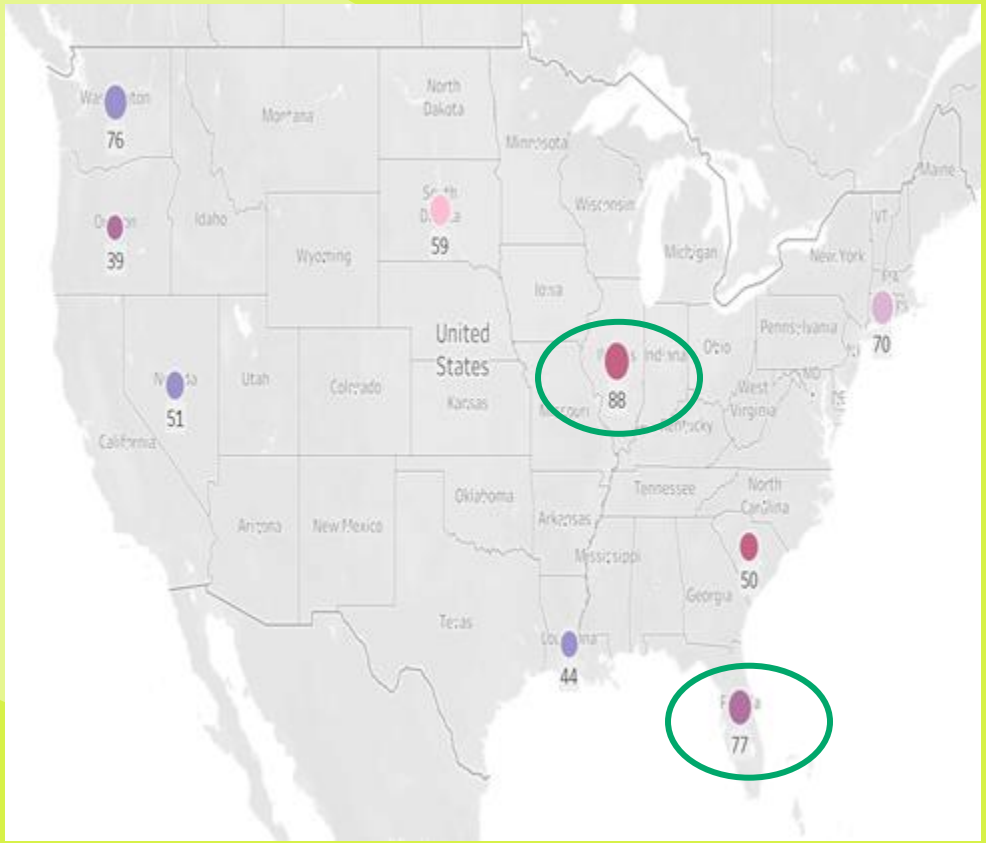
Total Quantity Sold for each Product Category (2018-2025)



- **High demand for seasonal and occasion-based apparel** - Swimwear and Formal Wear lead in sales
- **Opportunity to increase everyday apparel sales** - Tops and Bottoms show lower sales
- **Growing demand for comfort and lifestyle apparel** - Activewear and Underwear has shown consistent sales growth

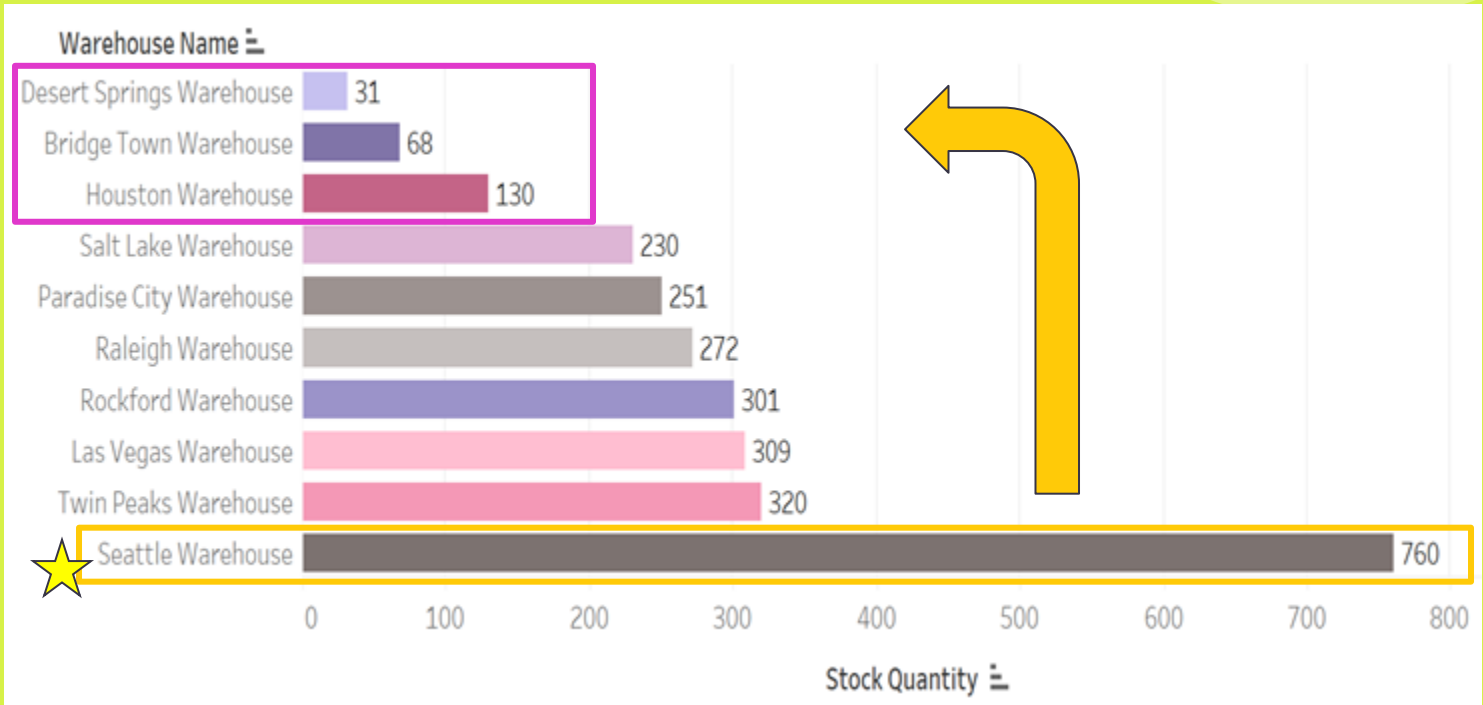
WAREHOUSE INFORMATION

Total Quantity Distributed By Warehouse (2018-2025)



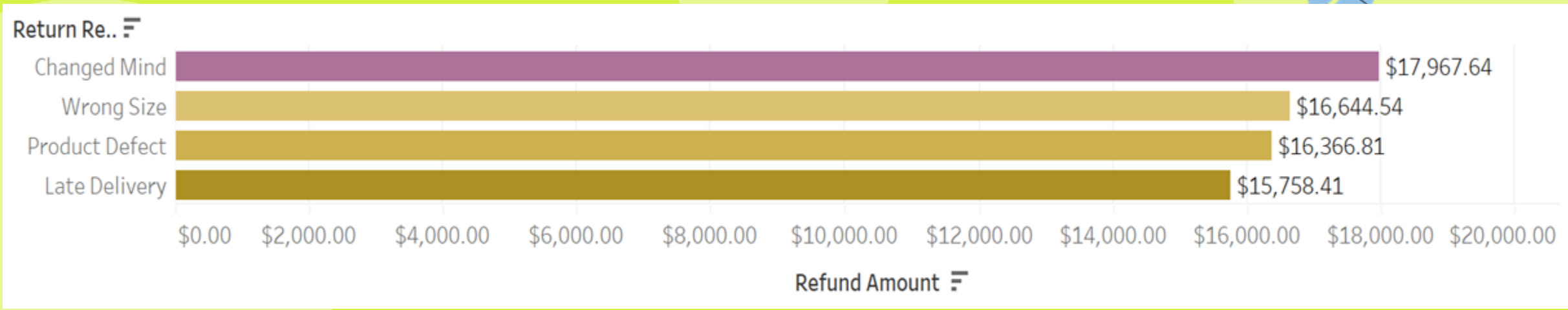
- **Illinois and Florida** lead in distribution
- Lower distribution in **Oregon and Louisiana**

Current Stock Quantity per Warehouse



RETURNS AND REFUNDS

Total Refunded Amount for each Return Category (2018-2025)



Average Rating for each Return Category

Return Re..	
Wrong Size	2.333
Changed Mind	2.900
Product Defect	3.250
Late Delivery	3.333

- Highest Refunds from “Changed Mind” Returns
- **Second-highest refund category are due to “Wrong Size” issues**, and also has the **lowest average rating (2.33)**
- “Product Defect” and “Late Delivery” have higher ratings, but still impact refunds

IMPROVEMENTS & FINAL TAKEAWAYS

- Improvements & Refinements
- Key Takeaways

FROM INITIAL DESIGN TO FINAL ERD: KEY IMPROVEMENTS

Key Enhancements in Database Design

- 3NF normalization completed
- Refined entity relationships
- Enhanced business logic implementation

Specific Entity-Level Improvements

- Customer & Order
- Payment & Return
- Review
- Shipments & Warehouse
- Supplier & Inventory

Business Impact of These Improvements

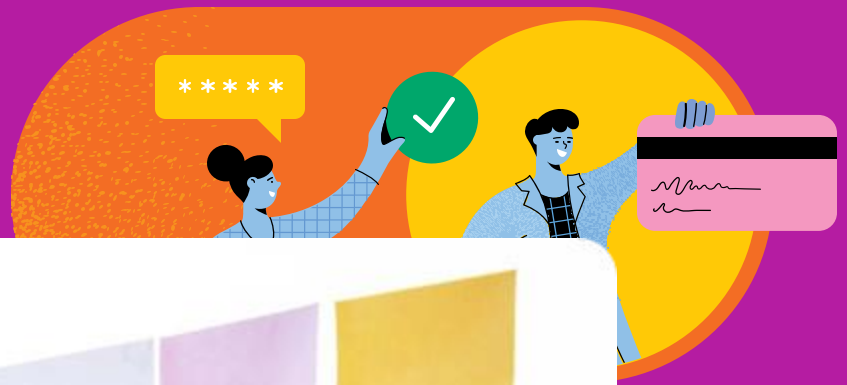
- Enhanced data accuracy
- Better customer insights
- More efficient return & refund processing
- Stronger inventory & shipment tracking

KEY TAKEAWAYS

Normalization (3NF) is crucial for maintaining an efficient relational database

Aligning database design with business needs ensures practical implementation

Real-world e-commerce applications benefit from proper data structuring



CONCLUSION

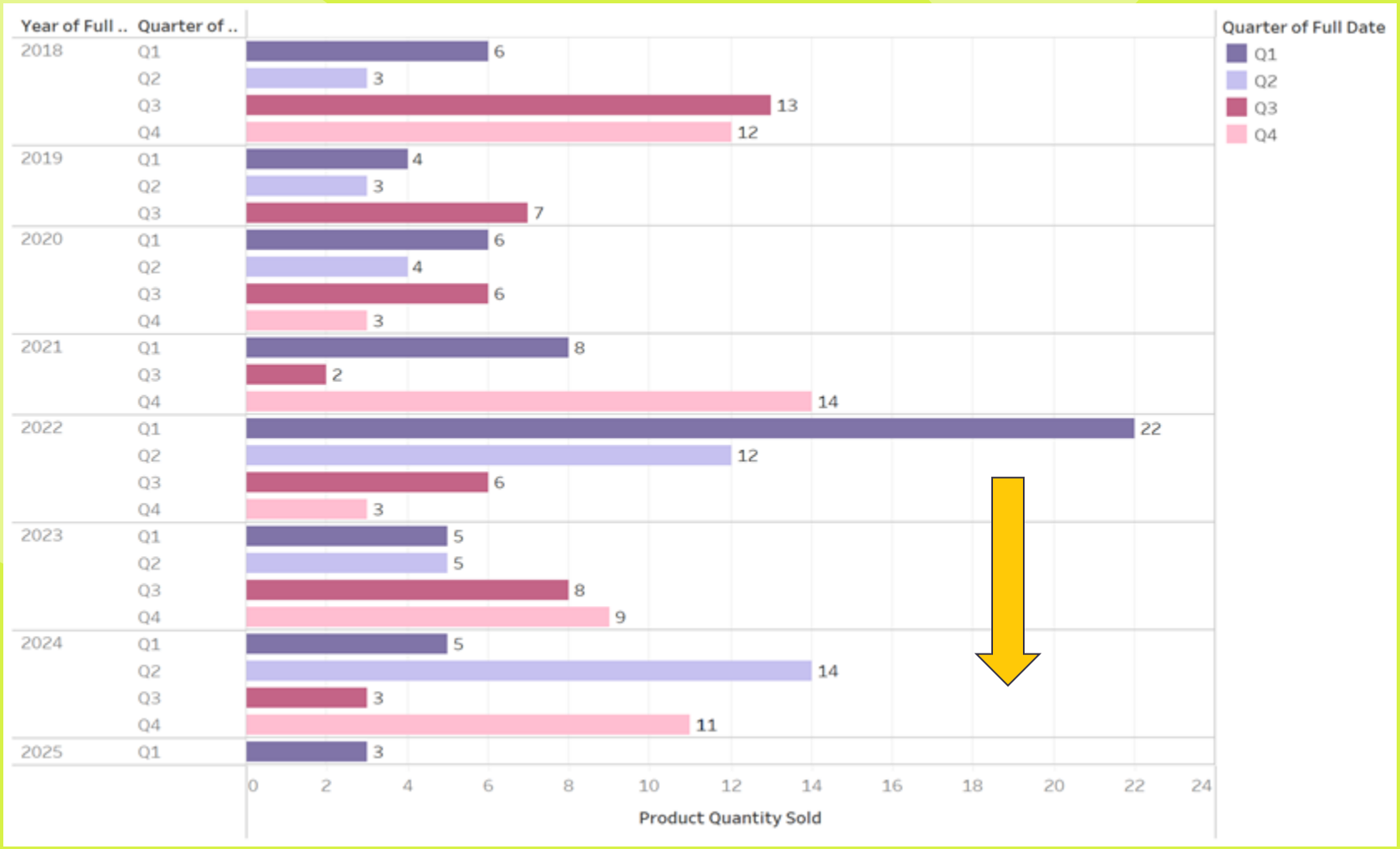
- Developed an optimized, normalized (3NF) relational database for an online retail system
- Focused on managing sales, inventory, suppliers, customers, and transactions
- Utilized real-world data from Kaggle to simulate practical business applications

SUMMARY OUR PROJECT

APPENDIX: REPORTS

YEARLY AND QUARTERLY SALES PERFORMANCE

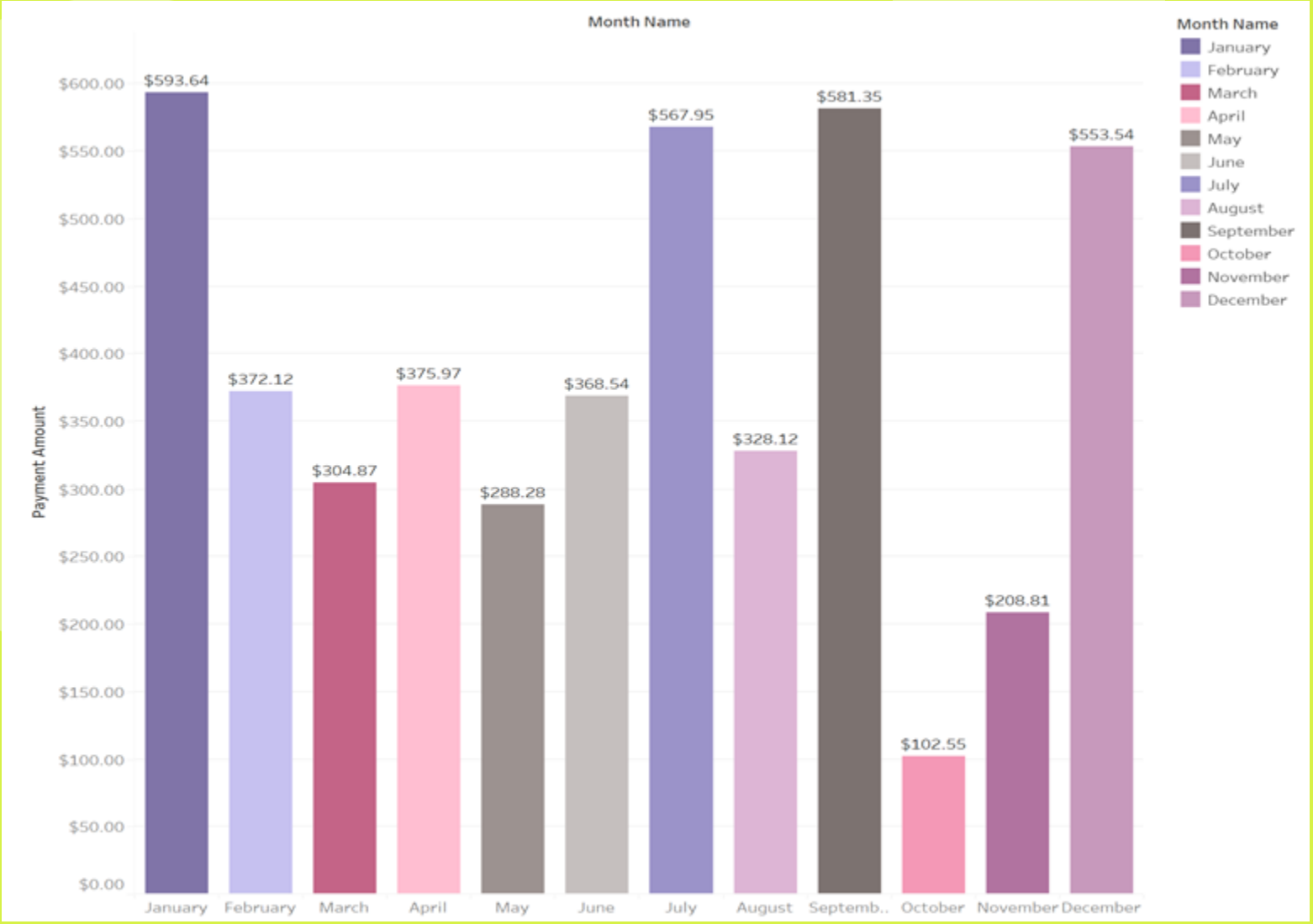
Product Quantity Sold per Quarter per Year (2018-2025)



- **Sales performance has generally declined from 2022 onward, particularly in Q1 and Q4**, highlighting a potential need for revised strategies to maintain demand across all quarters, especially during slower months like Q1 and Q4.
- **Strong sales in Q2 of 2024 (14 units)** suggest an opportunity to capitalize on this peak period through targeted promotions and marketing efforts to maximize revenue during high-demand months.

MONTHLY SALES PERFORMANCE

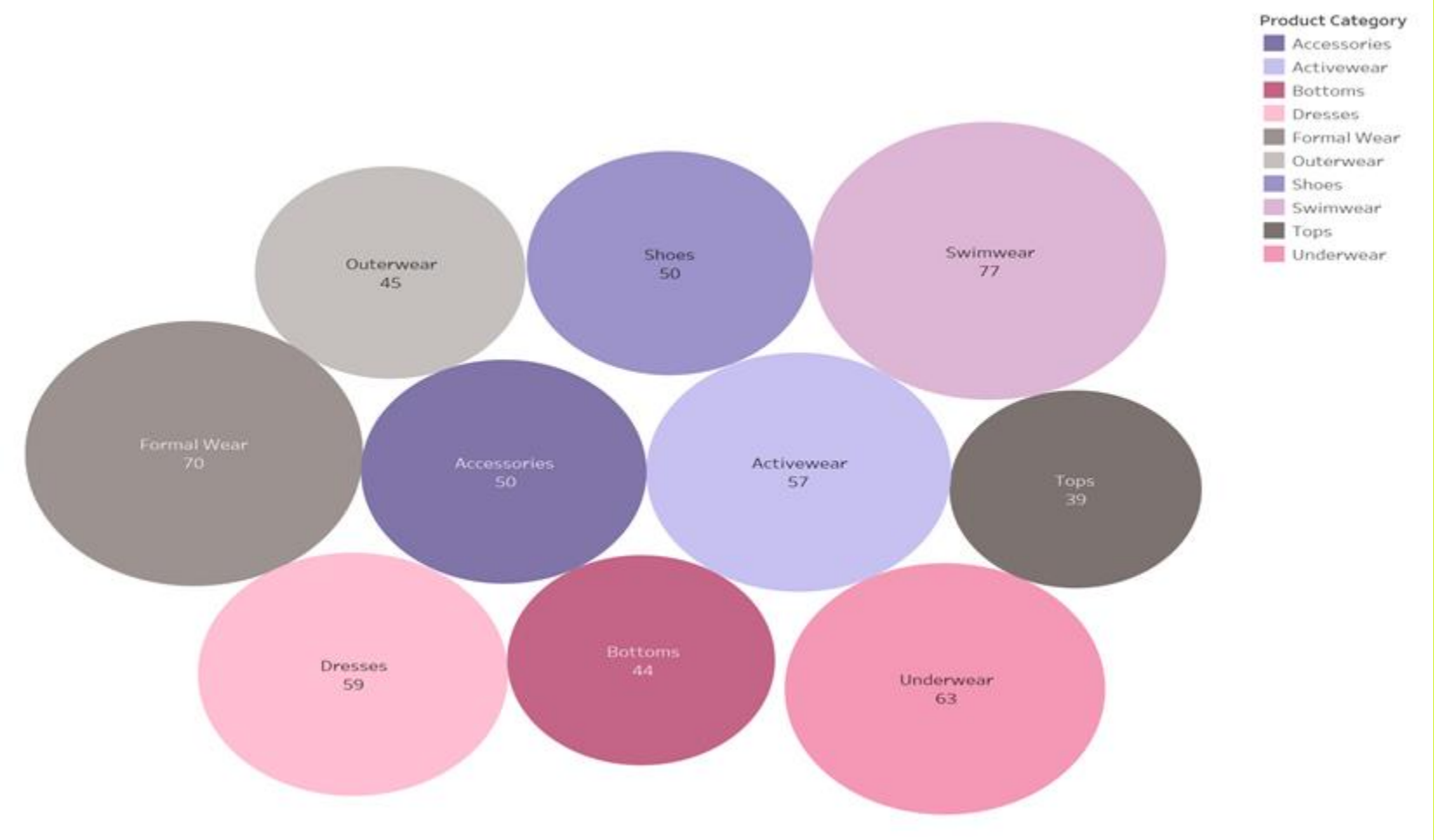
Total Sales for each Month (2018-2025)



- **Strong sales in January ($593.64), July ($567.95), and September ($581.35)** - indicating success in the company's marketing promotions
- **Sales in October ($102.55) and November ($208.81) are notably lower** compared to other months - decreased in customer demand during our off-season period
- **Fluctuating mid-year sales between March and August** - should look into market conditions or promotional efforts that may influence performance month to month

TOTAL SALES VOLUME BY PRODUCT CATEGORY

Total Quantity Sold for each Product Category (2018-2025)



- **High demand for seasonal and occasion-based apparel** - Swimwear (77) and Formal Wear (70) lead in sales
- **Opportunity to increase everyday apparel sales** - Tops (39) and Bottoms (44) show lower sales compared to other categories, indicating potential gaps in product marketing or pricing
- **Growing demand for comfort and lifestyle apparel** - Activewear (57) and Underwear (63) has shown consistent sales growth

WAREHOUSE INFORMATION

Total Quantity Distributed By Warehouse (2018-2025)

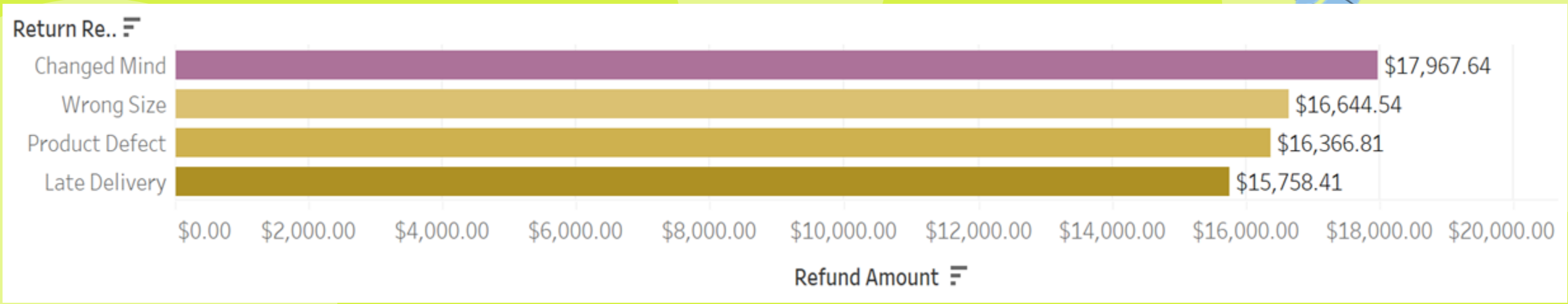
- **Illinois(88) and Florida(77) lead in distribution** - high demand & capacity
- **Lower distribution in Oregon(39) and Louisiana(44)** - fewer shipments

Current Stock Quantity per Warehouse

- **Seattle (760) holds the largest inventory by a significant margin** - need to evaluate if stock volume aligns with demand forecast or redistribution is required
- **Desert Spring (31), Bridge Town (68), and Houston (130) have much lower stock** - these locations may be underutilized, should consider redistributing stock

RETURNS AND REFUNDS

Total Refunded Amount for each Return Category (2018-2025)



Average Rating for each Return Category

Return Re..	
Wrong Size	2.333
Changed Mind	2.900
Product Defect	3.250
Late Delivery	3.333

- **Highest Refunds from "Changed Mind"** Returns ($17,967.64)
- **Second-highest refund category ($16,644.54) are due to "Wrong Size" issues**, and also has the **lowest average rating (2.33)**
- "Product Defect" (3.25) and "Late Delivery"(3.33) have higher ratings, but still impact refunds (~$16K each)

THANK YOU

If you have any questions, feel free to ask!