PEI

SHISHIR SHETTY

PEI Dashboard

PEI

Spend/Status Count By Country

● Delivered ● Pending

○ Amount
○ Status Count

Country | **Age Catergory**

🔒 Country	💡 Item ✕
UK	

UK 113		Mousepad 24
Item Count 250	USA 97	Keyboard 23
	UAE 40	Monitor 13
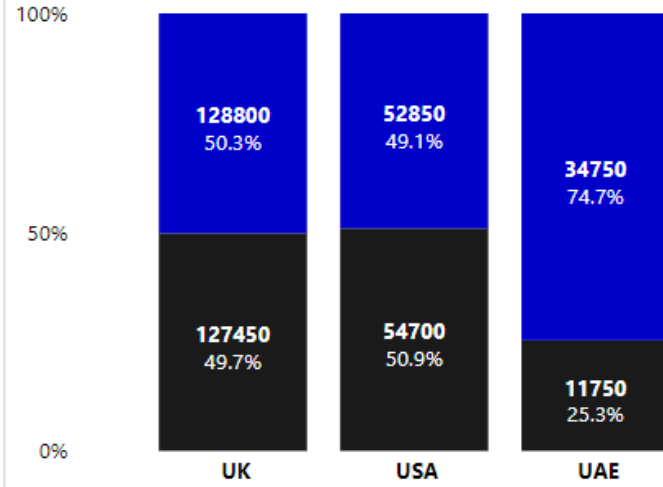
UK: 128800 — 50.3% / 127450 — 49.7%
USA: 52850 — 49.1% / 54700 — 50.9%
UAE: 34750 — 74.7% / 11750 — 25.3%

Total Amount
533K

Total Orders
250

Total Quantity
250

Country Detail Data

Country	Status	Amount	Status Count	Order Count
⊟ UK	Pending	128800	59	46
	Delivered	127450	36	38
	Total	**244350**	**95**	**113**
⊟ USA	Pending	52850	52	41
	Delivered	54700	39	24
	Total	**238200**	**91**	**97**
⊟ UAE	Pending	34750	39	27
	Delivered	11750	25	14
	Total	**49950**	**64**	**40**
Total		**532500**	**250**	**250**

Customer Detail Data

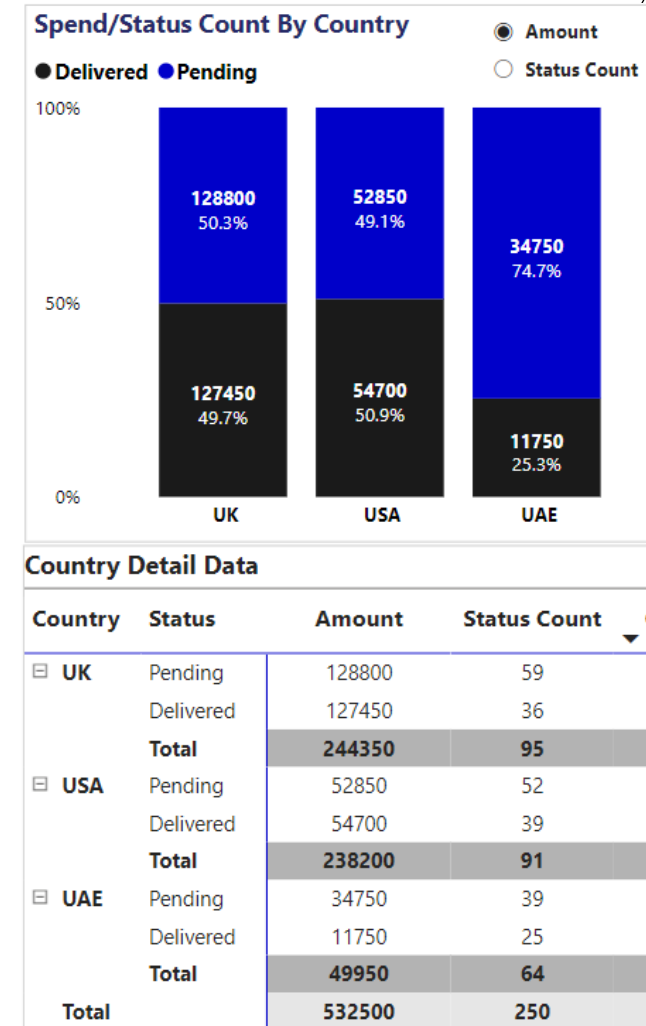| Customer_ID	Customer	Order ID	Item	Shipping_ID	Status	Item Count	Amount	S
8	Jason Montgomery	109	DDR RAM	200	Delivered	1	1500	
8	Jason Montgomery	109	DDR RAM	210	Delivered	1	1500	
8	Jason Montgomery	25	Mousepad	200	Delivered	1	250	
8	Jason Montgomery	25	Mousepad	210	Delivered	1	250	
8	Jason Montgomery	158	Mousepad	200	Delivered	1	200	
Total						**250**	**532500**	

a)Total Amount Spent for each country and Total Amount Spent Overall

 is displayed in the Country Detail Data with the highlighted value.

b) Count for the Pending Delivery Status and Count for Delivered Delivery Status

Is displayed in the 100% stacked column chart.

It also displayed the Percentage of Total between the two Status

Spend/Status Count By Country

● Delivered ● Pending

● Amount
○ Status Count

	128800 50.3%	52850 49.1%	34750 74.7%
127450 49.7%	54700 50.9%	11750 25.3%
UK	USA	UAE

Country Detail Data

Country	Status	Amount	Status Count	O
⊟ UK	Pending	128800	59	
	Delivered	127450	36	
	Total	**244350**	**95**	
⊟ USA	Pending	52850	52	
	Delivered	54700	39	
	Total	**238200**	**91**	
⊟ UAE	Pending	34750	39	
	Delivered	11750	25	
	Total	**49950**	**64**	
Total		**532500**	**250**	

2) THE TOTAL NUMBER OF TRANSACTIONS, TOTAL QUANTITY SOLD, AND TOTAL AMOUNT SPENT FOR EACH CUSTOMER, ALONG WITH THE PRODUCT *DETAILS.*

Total Number of Transactions(Order Count) , Total Quantity Sold(Item Count), Total Amount Spent(Amount) is displayed for each customer along with the Product Details in Customer Detail Data Table

Back to report   **CUSTOMER DETAIL DATA**

| Customer_ID	Customer	Order ID	Item	Shipping_ID	Status	Item Count	Amount	Status Count	Order Count
8	Jason Montgomery	109	DDR RAM	200	Delivered	1	1500	1	1
8	Jason Montgomery	109	DDR RAM	210	Delivered	1	1500	1	1
8	Jason Montgomery	25	Mousepad	200	Delivered	1	250	1	1
8	Jason Montgomery	25	Mousepad	210	Delivered	1	250	1	1
8	Jason Montgomery	158	Mousepad	200	Delivered	1	200	1	1
8	Jason Montgomery	158	Mousepad	210	Delivered	1	200	1	1
8	Jason Montgomery	117	Webcam	200	Delivered	1	350	1	1
8	Jason Montgomery	117	Webcam	210	Delivered	1	350	1	1
10	Darrell Dillon	144	Keyboard	240	Delivered	1	400	1	1
10	Darrell Dillon	144	Keyboard	154	Pending	1	400	1	1
12	Jodi Gonzalez	166	Harddisk	43	Delivered	1	5000	1	1
12	Jodi Gonzalez	166	Harddisk	183	Delivered	1	5000	1	1
13	Omar Martin	130	Headset	225	Delivered	1	900	1	1
13	Omar Martin	234	Keyboard	225	Delivered	1	400	1	1
13	Omar Martin	173	Monitor	225	Delivered	1	12000	1	1
15	Jason Brown	17	Webcam	65	Delivered	1	350	1	1
15	Jason Brown	17	Webcam	148	Delivered	1	350	1	1
15	Jason Brown	17	Webcam	157	Pending	1	250	1	1
Total						**250**	**532500**	**250**	**250**

3) THE MAXIMUM PRODUCT PURCHASED FOR EACH COUNTRY.

Maximum Product purchased for each Country is displayed in the Decomposition Tree / Hierarchy Chart

with the information tooltip
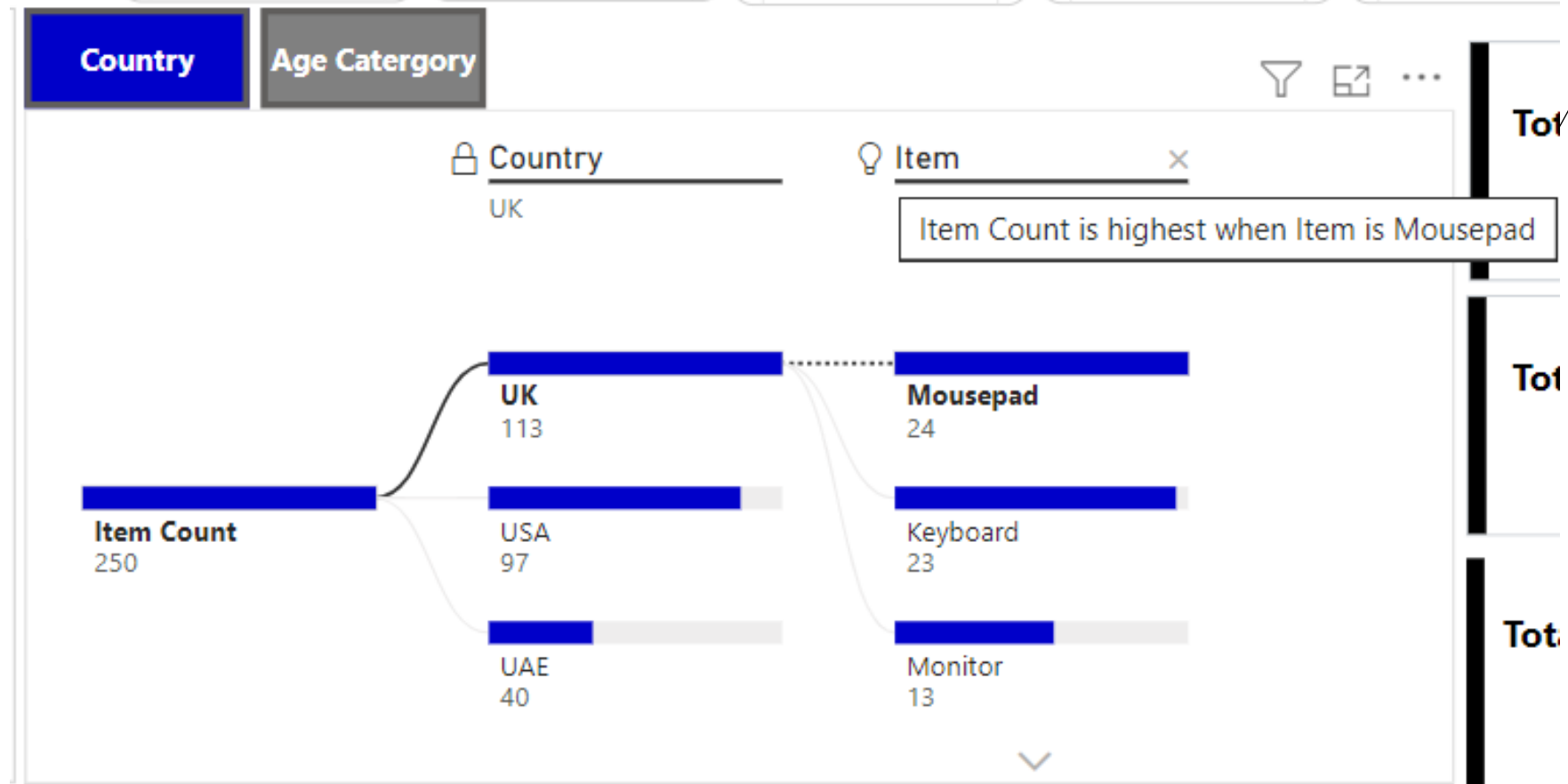
Most purchased product based on the age category is displayed in the Decomposition Tree / Hierarchy Chart

with the information tooltip
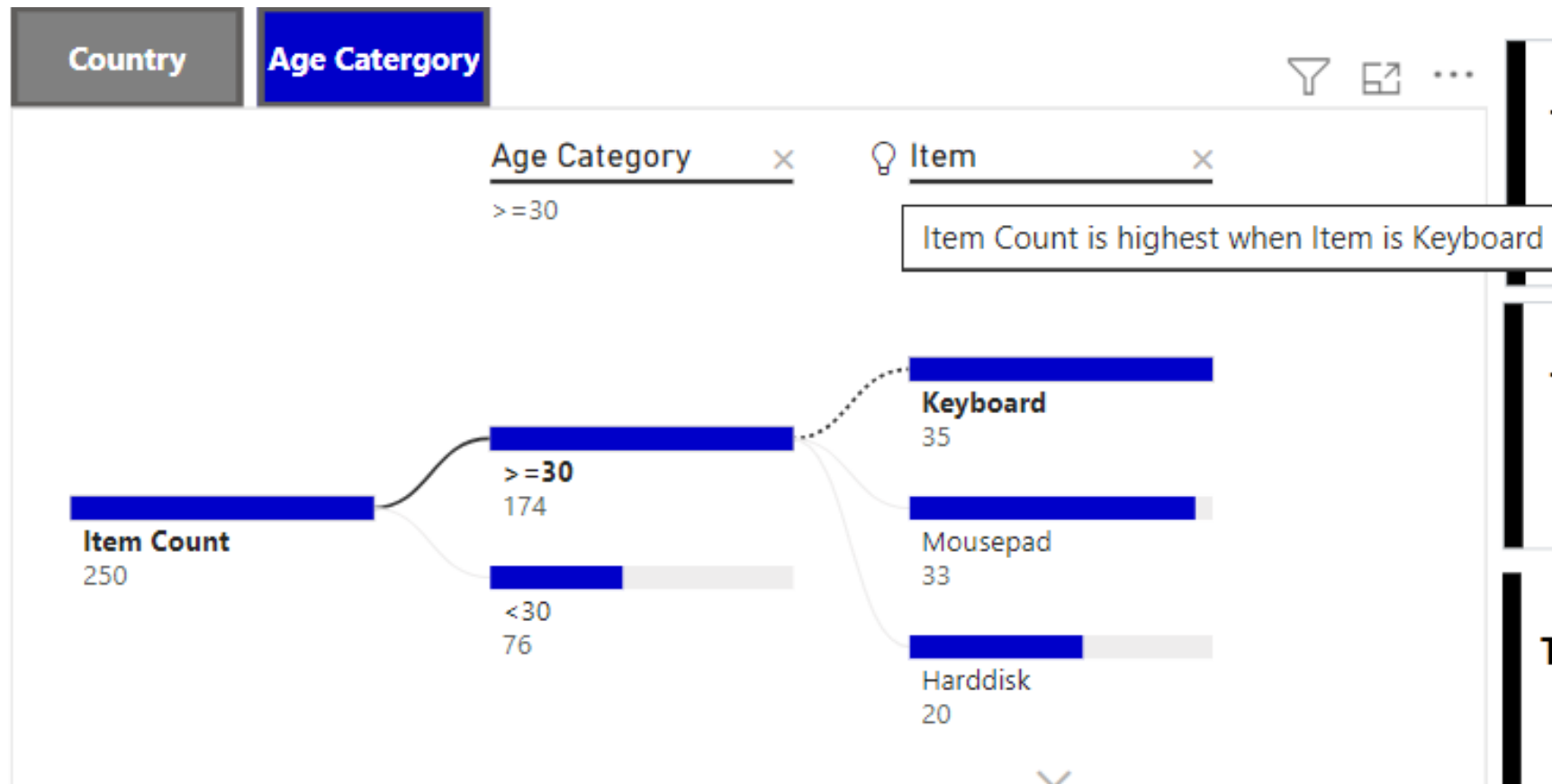
5) THE COUNTRY THAT HAD MINIMUM TRANSACTIONS AND SALES AMOUNT.

The Country that had minimum transactions(Order Count) and sales amount(Amount) is displayed in Country Detail Data Table

Country Detail Data

Country	Status	Amount	Status Count	Order Count ▼
⊟ UK	Pending	128800	59	46
	Delivered	127450	36	38
	Total	**244350**	**95**	**113**
⊟ USA	Pending	52850	52	41
	Delivered	54700	39	24
	Total	**238200**	**91**	**97**
⊟ UAE	Pending	34750	39	27
	Delivered	11750	25	14
	Total	**49950**	**64**	**40**
Total		**532500**	**250**	**250**

AS A DATA ANALYST, YOU ARE REQUIRED TO?

1) VERIFY THE ACCURACY, COMPLETENESS, AND RELIABILITY OF SOURCE DATA.

2) BASED ON YOUR FINDINGS, DEFINE AND OUTLINE THE REQUIREMENTS FOR ANTICIPATED DATASETS, DETAILING THE NECESSARY DATA COMPONENTS.

Data Integrity Constraints: Implement constraints to ensure data integrity, such as ensuring that each order and shipping record has a corresponding customer entry in the Customer Dataset.

Concern : There are total 250 customer details in Customer Table.

1)Few customers have placed order(which can be fetched from

order table Eg: Customer ID = 4.)

But there is no information on their Shipping details in shipping table

2)Shipping Details of few customer is present in Shipping Table

(For Eg: Customer ID =1) but there is no

information on what order they placed in Order Table

1) VERIFY THE ACCURACY, COMPLETENESS, AND RELIABILITY OF SOURCE DATA.

2) BASED ON YOUR FINDINGS, DEFINE AND OUTLINE THE REQUIREMENTS FOR ANTICIPATED DATASETS, DETAILING THE NECESSARY DATA COMPONENTS.

Data Integrity Constraints: Implement constraints to ensure data integrity, such as ensuring that each order and shipping record has a corresponding customer entry in the Customer Dataset.

Concern :

1) Few of the customer has ordered same item with same Order ID. However, shipping ID differs. Does that mean every individual item gets separate Shipping ID or is that an error in capturing the data?

Customer_ID	Customer	Order ID	Item	Shipping_ID	Status	Item Count	Amount	Status Count	Order Count
8	Jason Montgomery	109	DDR RAM	200	Delivered	1	1500	1	1
8	Jason Montgomery	109	DDR RAM	210	Delivered	1	1500	1	1
8	Jason Montgomery	25	Mousepad	200	Delivered	1	250	1	1
8	Jason Montgomery	25	Mousepad	210	Delivered	1	250	1	1

2) Is the Pending delivery amount received ? because it can be COD

3) No information on returns hence we cannot exactly calculate sales amount because few items could have been returned

4) Mousepad is the item with 2 Cost - 200, 250. Why? any coupon discount or different Mouse pad?

5) There is no information on any date field which is big concern like Order Date, Ship Date etc
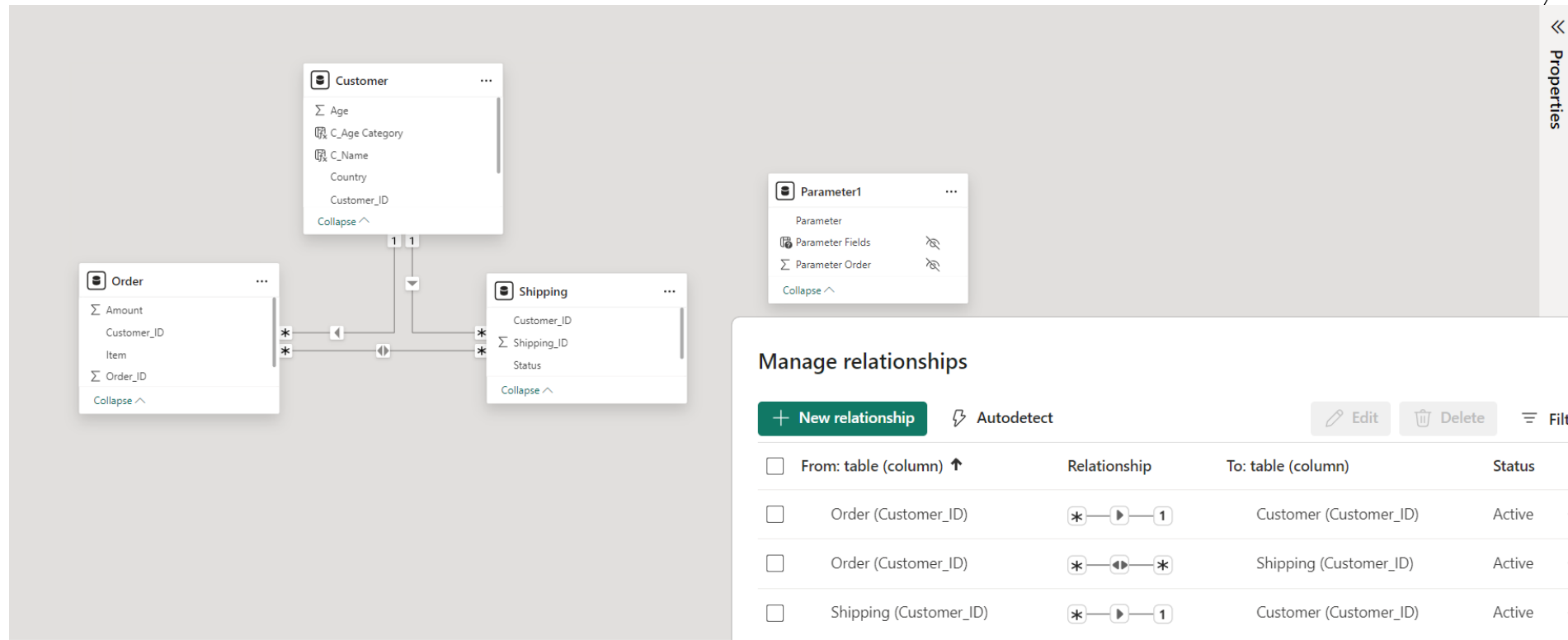
4) COMMUNICATE THE FINDINGS AND INSIGHTS TO STAKEHOLDERS IN A VISUALLY COMPREHENSIVE MANNER..
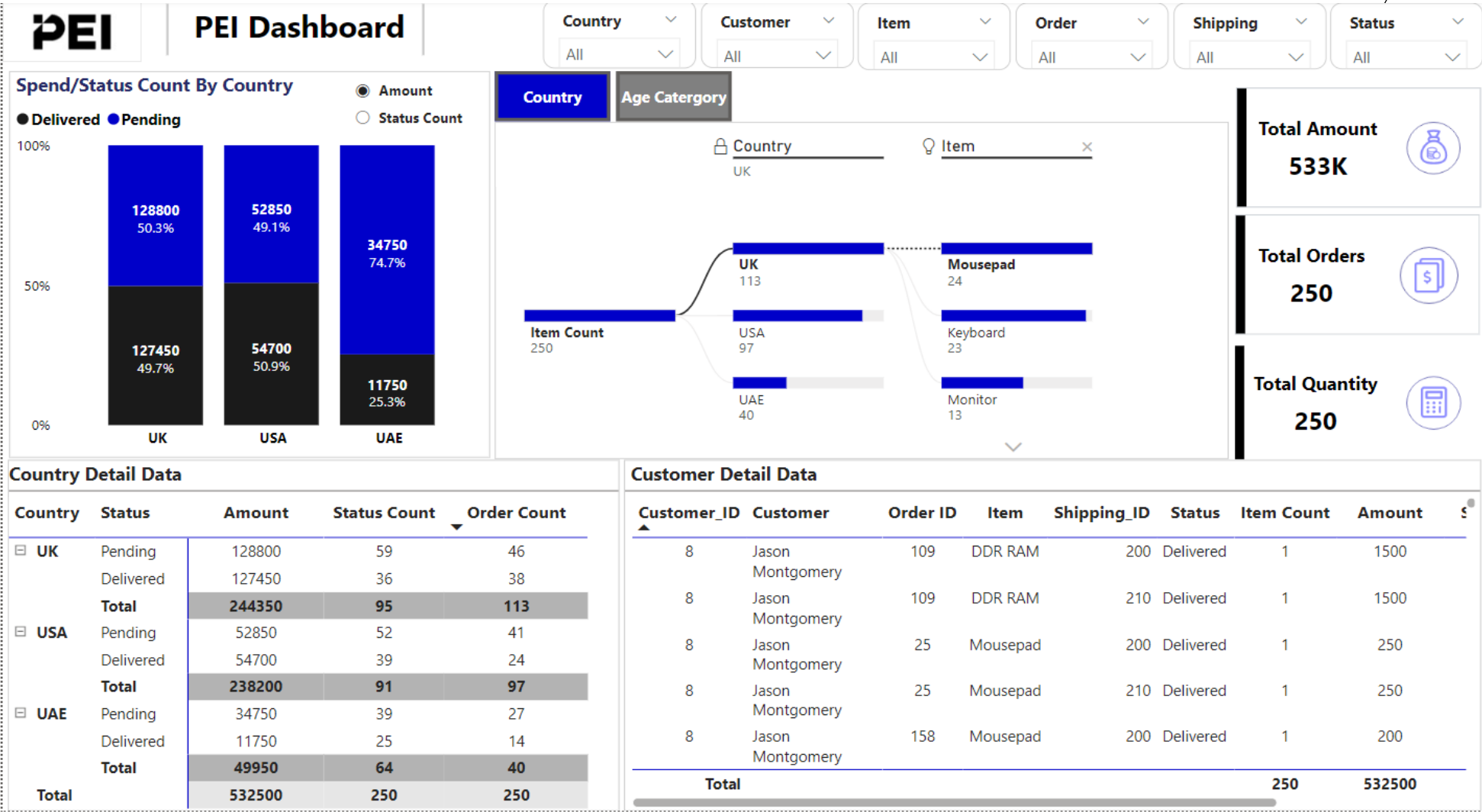
5) WHAT WILL BE YOUR INSIGHTS TO OTHER PEER TEAMS OF DATA ENGINEERS, DATA SCIENTISTS AND OTHER TECHNICAL AND NON-TECHNICAL STAKEHOLDERS?

Data Integrity and Completeness:

- There are inconsistencies between the Order and Shipping tables, where some customers have placed orders but have no shipping information, and vice versa.

- This indicates potential data integrity issues or gaps in data capture processes. It's essential to investigate and rectify these inconsistencies to ensure the completeness and accuracy of the data.

Data Relationships:

- The presence of customers with orders but no shipping details, and vice versa, highlights the importance of maintaining proper data relationships between tables.

- Ensuring referential integrity between the Customer, Order, and Shipping tables is crucial for accurate analysis and reporting.

Handling Duplicate Orders:

- Identifying cases where the same item with the same Order ID has different Shipping IDs suggests potential errors in data capture or processing.

- Data Engineers should investigate whether this is due to individual items receiving separate shipping IDs or if there are issues in data recording.

- Clarification on the handling of duplicate orders and shipping IDs is necessary for accurate data interpretation.

5) WHAT WILL BE YOUR INSIGHTS TO OTHER PEER TEAMS OF DATA ENGINEERS, DATA SCIENTISTS AND OTHER TECHNICAL AND NON-TECHNICAL STAKEHOLDERS?

Payment and Delivery Status:

- The absence of information on pending delivery amounts and returns impacts the accuracy of sales calculations and inventory management.

- Data Scientists should collaborate with stakeholders to devise strategies for tracking pending delivery amounts, handling returns, and updating the relevant datasets accordingly.

Product Variations and Discounts:

- The existence of multiple cost values for the same item (e.g., Mousepad) raises questions about potential product variations, discounts, or errors in data entry.

- Data Engineers should investigate and clarify the reasons behind these discrepancies to ensure accurate pricing information for analysis.

Missing Date Information:

- The absence of date fields (e.g., Order Date, Ship Date) is a significant concern as it hinders the ability to perform time-based analysis, track delivery timelines, and monitor business performance over time.

- Data Engineers should work on incorporating relevant date fields into the datasets to enable comprehensive analysis and reporting.

Collaboration and Data Governance:

- Encourage collaboration between Data Engineers, Data Scientists, and stakeholders to address data quality issues, establish data governance policies, and implement robust data validation processes.

- Regular communication and coordination are essential for maintaining data integrity and driving informed decision-making across the organization.

THANK YOU

SHISHIR SHETTY