­BI Technical Specification Docum­­ent

BI Conversion Name: **PEI Sales Analytics Dashboard**

Generated Document Name: **PEI\_SALES \_AMAN\_MUSHIRUL\_TS**

Market / Project : **PEI GROUP - PEDA**

Author : Mushirul Aman

[mushirul.aman@gmail.com](mailto:mushirul.aman@gmail.com)

# 1. Document Information

### Document Location

|  |  |  |  |
| --- | --- | --- | --- |
| ***Document Name*** | ***Description*** | ***Document #*** | ***File Link*** |
| **PEI\_SALES \_AMAN\_MUSHIRUL\_TS** | Technical Specifications Document |  | [GitHub - MushirulX/Data-Analysis](https://urldefense.com/v3/__https:/github.com/MushirulX/Data-Analysis__;!!Nyu6ZXf5!uOtCyf3sEOPbfVjik19fvgYGeqgK8eK_DcnGsiW_QNYhyou5BP3MqL5QLXfPSyERdmDAIYDrbH0lmiVZcAgU0arG$) |
| **Source Data(Order / Customer / Shipping )** |  |  | [GitHub - MushirulX/Data-Analysis](https://urldefense.com/v3/__https:/github.com/MushirulX/Data-Analysis__;!!Nyu6ZXf5!uOtCyf3sEOPbfVjik19fvgYGeqgK8eK_DcnGsiW_QNYhyou5BP3MqL5QLXfPSyERdmDAIYDrbH0lmiVZcAgU0arG$) |
| **PEI Sales Analytics Dashboard** | Power BI File ( .PBIX ) |  | [GitHub - MushirulX/Data-Analysis](https://urldefense.com/v3/__https:/github.com/MushirulX/Data-Analysis__;!!Nyu6ZXf5!uOtCyf3sEOPbfVjik19fvgYGeqgK8eK_DcnGsiW_QNYhyou5BP3MqL5QLXfPSyERdmDAIYDrbH0lmiVZcAgU0arG$) |
| **PEI SALES Analytics Data Report** | Data Report for Walkthrough |  | [GitHub - MushirulX/Data-Analysis](https://urldefense.com/v3/__https:/github.com/MushirulX/Data-Analysis__;!!Nyu6ZXf5!uOtCyf3sEOPbfVjik19fvgYGeqgK8eK_DcnGsiW_QNYhyou5BP3MqL5QLXfPSyERdmDAIYDrbH0lmiVZcAgU0arG$) |

### 

### Note : In case the files doesn’t open -> Go to GITHUB link for file access

Table of Contents

[1. Document Information 2](#_Toc176607356)

[Document Location 2](#_Toc176607357)

[2. Requirement Gathering / Business Walkthrough 3](#_Toc176607358)

[3. Data Walkthrough / Scan the data 6](#_Toc176607359)

[4. Data Connection 7](#_Toc176607360)

[5. Data Cleansing / Quality Check (Profiling) 7](#_Toc176607361)

[6. Data Modelling 10](#_Toc176607362)

[o Visualizing the Data Model : You should see the following structure in the Model View: 11](#_Toc176607363)

[o Final Data Model Relationships: 11](#_Toc176607364)

[7. Data Transformation and Processing in Power Query 16](#_Toc176607365)

[Aggregation and Grouping 17](#_Toc176607366)

[6. Merging and Appending Data 17](#_Toc176607367)

[8. DAX / Measure Calculations 17](#_Toc176607368)

[8. Dashboard Layering / Formatting / Report Development / KPI’s 22](#_Toc176607369)

[9. Dashboard Insights Generation 37](#_Toc176607370)

[1. Pending Delivery Status by Country 37](#_Toc176607371)

[2. Customer Spending Patterns 37](#_Toc176607372)

[3. Country-Level Product Demand 38](#_Toc176607373)

[4. Age-Based Product Preference 39](#_Toc176607374)

[5. Country with Minimum Sales 39](#_Toc176607375)

[6. Average Spending per Transaction 40](#_Toc176607376)

[7. Product Insights 42](#_Toc176607377)

# 2. Requirement Gathering / Business Walkthrough

**Business Statement:**To conduct a comprehensive analysis of PEI Sales Performance by generating detai­­­­led sales metrics and transactional reports to provide insights into key business metrics across various dimensions, including delivery status, customer behaviour, product preferences, and geographic sales performance using various KPI’s and visualizations in Power BI.

**Business Requirements:**

* **Pending Deliveries by Country:** To provide clarity on delivery performance, we will generate a report showing the total amount spent for deliveries in "Pending" status, broken down by country.
* **Customer Transactions and Product Insights:** To gain deeper insights into customer behaviour, we will deliver a report detailing total transactions, quantity sold, and the amount spent for each customer, along with associated product details.
* **Top Product by Country:** To assist in understanding regional product preferences, we will identify the highest-purchased product in each country.
* **Product Preferences by Age Category:** To enhance age-targeted marketing strategies, we will analyse product purchases based on two key age categories: under 30 and 30+.
* **Country with Lowest Transactions and Sales:** To identify underperforming regions, we will report on the country with the fewest transactions and the lowest sales amount.

**KPI METRICS :**

* Total Sales (Sum of Amount).
* Number of Product
* Average Order Value.
* Top 10 Customers by Spend.
* Orders per Country.

|  |  |  |
| --- | --- | --- |
| **Requirement** | **KPI** | **Metric** |
| **Total Amount Spent for Pending Deliveries by Country** | Total Amount Spent on Pending Deliveries | Sum of all amounts spent on orders with a "Pending" delivery status, segmented by country. |
| **Customer Transactions and Product Insights** | Total Transactions per Customer | Count of all transactions for each customer. |
|  | Total Quantity Sold per Customer | Sum of all quantities sold for each customer. |
|  | Total Amount Spent per Customer | Sum of all amounts spent by each customer. |
|  | Products Purchased per Customer | List of products purchased by each customer, including product details (e.g., name, category). |
| **Maximum Product Purchased by Country** | Most Purchased Product per Country | The product with the highest quantity purchased in each country. |
| **Most Purchased Product by Age Category** | Most Purchased Product for Age Category < 30 | The product with the highest quantity purchased by customers aged under 30. |
|  | Most Purchased Product for Age Category ≥ 30 | The product with the highest quantity purchased by customers aged 30 and above. |
| **Country with Minimum Transactions and Sales** | Minimum Transactions Country | Country with the fewest number of transactions. |
|  | Minimum Sales Amount Country | Country with the lowest total sales amount. |

# 3. Data Walkthrough / Scan the data

**Special Characters:** Check for any anomalies like special characters in customer names

**Standardize Format:** Ensure names are in a consistent format (e.g., Capitalization)  
**Order Amounts and Quantities:** Check for invalid or outlier values (e.g., negative amounts or quantities)  
**Status Values:** Ensure delivery statuses are consistent and valid (e.g., "Pending", "Shipped", "Delivered")  
**Country Names:** Verify that country names are consistently formatted.  
**Check for missing values:** In each table (Customers, Orders, and Shipping), go to the View > Column Quality tab.

**Data Type Adjustments**: Ensure that data types for critical columns like Customer\_ID, Quantity, Price, and Shipping Cost are correctly set to integers, decimals, or text as needed

# 4. Data Connection

**Importing Data:** Import all three tables into Power BI, as described earlier, using the Get Data option for Excel, CSV, and JSON. Load the data into Power BI Desktop.  
  


.

# 5. Data Cleansing / Quality Check (Profiling)

**STEP 1: Profiling**  
We did a consolidation join on CUSTOMER , ORDER and SHIPPING to check the consistency of data through out if we are having any mismatches on Customer or if there’s any CUSTOMER in Customer table which doesn’t exist in ORDERS and SHIPPING TABLES  
  
 **SQL QUERY USED: -**  
**-- Merge Customers and Orders**

SELECT c.Customer\_ID, c.Name, c.Age, c.Country, o.Order\_ID, o.Product\_ID, o.Quantity, o.Amount, o.Delivery\_Status

FROM Customers c

LEFT OUTER JOIN Orders o ON c.Customer\_ID = o.Customer\_ID;

**-- Merge the result with Shipping**

SELECT co.Customer\_ID, co.Name, co.Age, co.Country, co.Order\_ID, co.Product\_ID, co.Quantity, co.Amount, co.Delivery\_Status, s.Shipping\_Status, s.Country AS Shipping\_Country

FROM (SELECT c.Customer\_ID, c.Name, c.Age, c.Country, o.Order\_ID, o.Product\_ID, o.Quantity, o.Amount, o.Delivery\_Status FROM Customers c

LEFT OUTER JOIN Orders o ON c.Customer\_ID = o.Customer\_ID) co

LEFT OUTER JOIN Shippings s ON co.Order\_ID = s.Order\_ID;

**INCONSISTENCIES FOUND :**

**Data Entry Issues / Incomplete Data:**

* There are Customer\_ID from CUSTOMERS which are not present in ORDER and SHIPPING tables resulting in getting NULL records while creating a Golden Consolidation Sheet.
* One or more of the datasets could be incomplete, meaning some customers may not have corresponding orders or shipments, resulting in null values during the join.
* If a **Customer\_ID** exists in the **Customers** table but not in the **Orders** table, it means that customer hasn't placed any orders. This is expected behavior, as some customers may not have ordered yet.
* If a **Customer\_ID** exists in the **Orders** table but not in the **Shipping** table, it could indicate an issue in the shipping process or that the order hasn’t been processed yet.

**VALIDATION CHECK ( TO FLAG INCONSISTENT CUSTOMER ID ) :**

* VALIDATION CHECK ( CUSTOMER ID ) and SHIPPING CHECK ( PARENT ) have been added as Calculated Column to distinguish these Customer ID -> Filter with 0 to get all the Customer ID’s showing inconsistencies.
* This approach is for times when we want to address the issue with Business or Data Producers who are the primary owners of source file. For this assessment we’ll move on with STEP 2

**STEP 2 : Data Cleansing ( Getting the data ready for Reporting )**

* **Exclude from Analysis:** If these missing customers are minimal or irrelevant, we can filter out these rows in our BI dashboards.
* Investigate Source: It's important to investigate why these customers are missing in the first place. Ensure the ETL process correctly maps customers to orders.
* **Solution for correct reporting** : Do inner join on the 3 tables and only take relevant data with not NULL values ( Creation of Consolidates Sales table which is the source of Truth for Dashboard visualization and reporting.

**Merge Customers and Orders with INNER JOIN :**   
SELECT c.Customer\_ID, c.Name, c.Age, c.Country, o.Order\_ID, o.Product\_ID, o.Quantity, o.Amount, o.Delivery\_Status

FROM Customers c

INNER JOIN Orders o ON c.Customer\_ID = o.Customer\_ID;

**Merge the result with Shipping using INNER JOIN**:  
SELECT co.Customer\_ID, co.Name, co.Age, co.Country, co.Order\_ID, co.Product\_ID, co.Quantity, co.Amount, co.Delivery\_Status, s.Shipping\_Status, s.Country AS Shipping\_Country

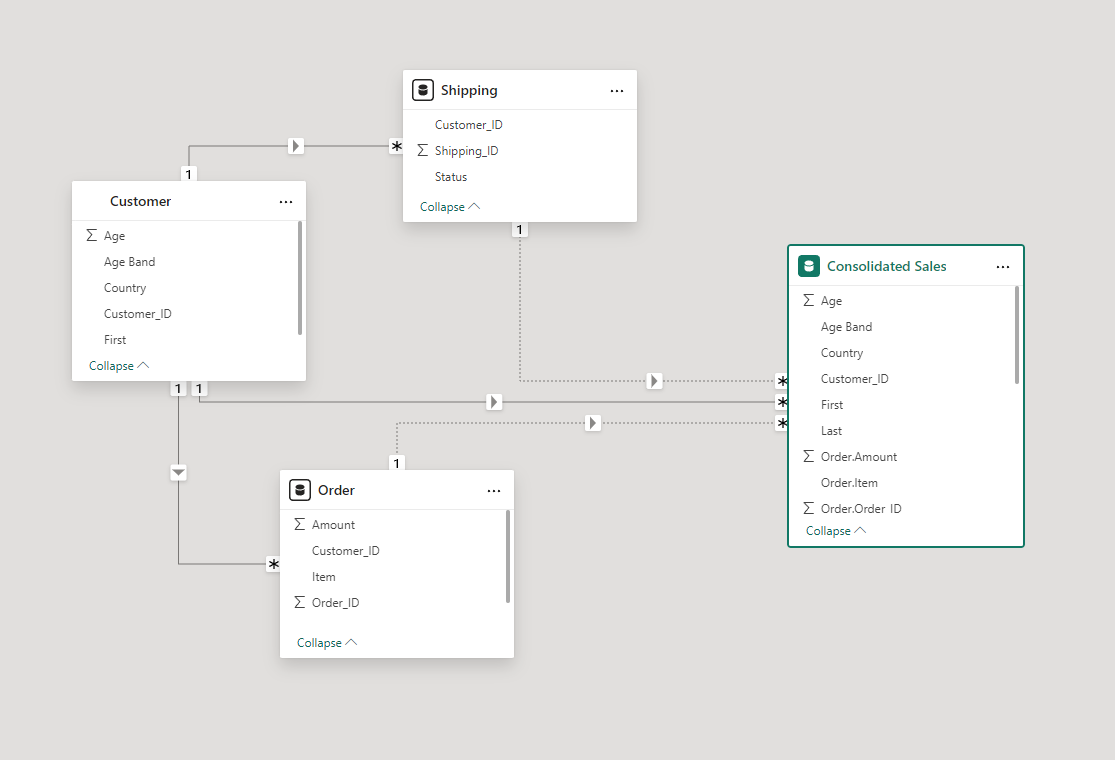
FROM (SELECT c.Customer\_ID, c.Name, c.Age, c.Country, o.Order\_ID, o.Product\_ID, o.Quantity, o.Amount, o.Delivery\_Status

FROM Customers c

INNER JOIN Orders o ON c.Customer\_ID = o.Customer\_ID) co

INNER JOIN Shippings s ON co.Order\_ID = s.Order\_ID;

# 6. Data Modelling

****

1. **Check the Relationship Diagram:**

After creating these relationships, you should see a clear model where:

* Customers table is connected to both the Orders and Shippings tables via Customer\_ID.
* This creates an indirect relationship between the Orders and Shippings tables through the Customers table.

### Visualizing the Data Model : You should see the following structure in the Model View:

* Customers table in the center with a One-to-Many relationship going to both the Orders and Shippings tables.
* There will be no direct relationship between Orders and Shippings tables unless an explicit field (like Order\_ID or another shipping identifier) exists to connect them.

### Final Data Model Relationships:

* Customers → Orders: One-to-Many (via Customer\_ID)
* Customers → Shippings: One-to-Many (via Customer\_ID)

**Customer Table :**

|  |  |  |  |
| --- | --- | --- | --- |
| **Attribute** | **Data Type** | **Constraints** | **Description** |
| Customer\_ID | Integer | Primary Key, Not Null | Unique identifier for each customer |
| Age | Integer | Not Null | Customer's age |
| Age Band | Text | Not Null | Grouping of customers based on age (e.g., <30, 30+) |
| Country | Text | Not Null | Customer's country of residence |
| First | Text | Optional | First name of the customer |
| Last | Text | Optional | Last name of the customer |

**Notes:**

* Customer\_ID uniquely identifies a customer.
* Country can be used for geographic analysis and segmentation.
* Age Band helps in analyzing customer purchases by age group.

**Order Table :**

|  |  |  |  |
| --- | --- | --- | --- |
| **Attribute** | **Data Type** | **Constraints** | **Description** |
| Order\_ID | Integer | Primary Key, Not Null | Unique identifier for each order |
| Customer\_ID | Integer | Foreign Key (Customer.Customer\_ID), Not Null | Links the order to the corresponding customer |
| Item | Text | Not Null | Product or service ordered |
| Amount | Decimal | Not Null | Total amount for the order |
| Order\_Date | Date | Not Null | Date when the order was placed |

**Notes:**

* Customer\_ID as a Foreign Key creates a relationship with the Customer table.
* Order\_ID is a unique identifier for each transaction.
* The Amount represents the total amount spent for each order, which will help in KPIs like "Average Spending per Transaction."
* Order\_Date helps in time-based analysis.

|  |  |  |  |
| --- | --- | --- | --- |
| **Attribute** | **Data Type** | **Constraints** | **Description** |
| Shipping\_ID | Integer | Primary Key, Not Null | Unique identifier for each shipping record |
| Customer\_ID | Integer | Foreign Key (Customer.Customer\_ID), Not Null | Links the shipping record to the corresponding customer |
| Status | Text | Not Null | Shipping status (e.g., "Pending", "Completed") |
| Shipping\_Date | Date | Optional | Date when the shipment was made (if applicable) |

**Notes:**

* Customer\_ID links each shipping record to a customer.
* The Status field is critical for analyzing pending or completed shipments.
* Shipping\_Date can help track shipment timelines.

**Relationships:**

* **Customer ↔ Order (1:n)**
  + One customer can place multiple orders.
  + Customer\_ID acts as a foreign key in the Order table, ensuring that each order is linked to a specific customer.
* **Customer ↔ Shipping (1:1):**
  + Each customer can have only one corresponding shipping record at a time.
  + Customer\_ID in Shipping is a foreign key that references Customer.Customer\_ID.
* **Order ↔ Shipping (1:n)**
  + A single shipping record can link to multiple orders.
  + Shipping\_ID in the Order table establishes this relationship.
* **Consolidated Sales ↔ Customer (1:1)**
  + Consolidated Sales aggregates customer details and spending, with Customer\_ID acting as a foreign key.

**Data Model Concerns & Solutions:**

* **Data Consistency**: Ensure that all orders have corresponding customer and shipping data, and there are no orphan records.
  + **Solution:** Data integrity constraints and Foreign Key relationships must be enforced in the database.
* **Data Completeness**: Ensure no null values in mandatory fields like Order\_Amount, Country, and Status.
  + **Solution:** Implement data validation checks before inserting records into tables.
* **Pending Deliveries:** Analysis of all pending orders and their geographical distribution is important.
  + **Solution:** Use the Status field in the Shipping table to filter pending orders and link them back to customers and countries.

# 7. Data Transformation and Processing in Power Query

* Filtering Rows: Filter out rows that don’t meet specific criteria (e.g., remove negative sales or outlier values).
* Rename Columns: Rename columns to more meaningful labels for easier analysis.
* Split or Merge Columns: Split a column (e.g., separating first and last names) or merge multiple columns (e.g., combining city and state into one field).
* Change Data Types: Convert columns to appropriate data types (e.g., text, date, number) to ensure accurate analysis.
* Transform Text: Adjust the formatting of text fields, such as changing the case or applying string manipulations.
* Add Calculated Columns: Create new columns based on custom calculations (e.g., sales tax or profit margin).

### Aggregation and Grouping

* **Group By**: Summarize data by grouping it based on specific columns (e.g., group sales by region and aggregate total sales).
* **Aggregations**: Perform calculations such as sum, average, count, or median for specific groups of data.

### 6. Merging and Appending Data

* **Merge Queries**: Combine data from multiple tables or sources based on common fields (like joining a customer table with an order table).  
  + **Append Queries**: Combine multiple datasets vertically by appending rows from one table to another, such as merging monthly sales records

# 8. DAX / Measure Calculations

1.**Average Spending per Transaction =**

**DIVIDE(**

**SUM('Consolidated Sales'[Order.Amount]),**

**COUNT('Consolidated Sales'[Order.Order\_ID])**

**)**

**2.Distinct Products = DISTINCTCOUNT('Consolidated Sales'[Order.Item])**

**3.Max Product Purchased by Country =**

**CALCULATE(**

**MAXX(**

**VALUES('Consolidated Sales'[Order.Item]),**

**CALCULATE(SUM('Consolidated Sales'[Order.Quantity]))**

**),**

**ALLEXCEPT('Consolidated Sales', 'Consolidated Sales'[Country])**

**)**

**4.Most Purchased Product Age < 30 =**

**CALCULATE(**

**MAXX(**

**VALUES('Consolidated Sales'[Order.Item]),**

**CALCULATE(SUM('Consolidated Sales'[Order.Quantity]),**

**'Consolidated Sales'[Age]< 30**

**)**

**)**

**)**

**5.Most Purchased Product Age ≥ 30 =**

**CALCULATE(**

**MAXX(**

**VALUES('Consolidated Sales'[Order.Item]),**

**CALCULATE(SUM('Consolidated Sales'[Order.Quantity]),**

**'Consolidated Sales'[Age] >= 30**

**)**

**)**

**)**

**6.Total Amount Spent for Pending Deliveries =**

**CALCULATE(**

**SUM('Consolidated Sales'[Order.Amount]),**

**'Consolidated Sales'[Shipping.Status] = "Pending"**

**)**

**7.Total Amount Spent per Customer =**

**SUM('Consolidated Sales'[Order.Amount])**

**8.Total Orders = DISTINCTCOUNT('Consolidated Sales'[Order.Order\_ID])**

**9.Total Quantity Sold per Customer =**

**SUM('Consolidated Sales'[Order.Quantity])**

**10.Total Transactions = DISTINCTCOUNT('Consolidated Sales'[Order.Order\_ID])**

**11.Total Transactions per Customer = COUNTROWS( 'Consolidated Sales')**

**DAX/ Measures Screenshots :-**

A screenshot of a computer

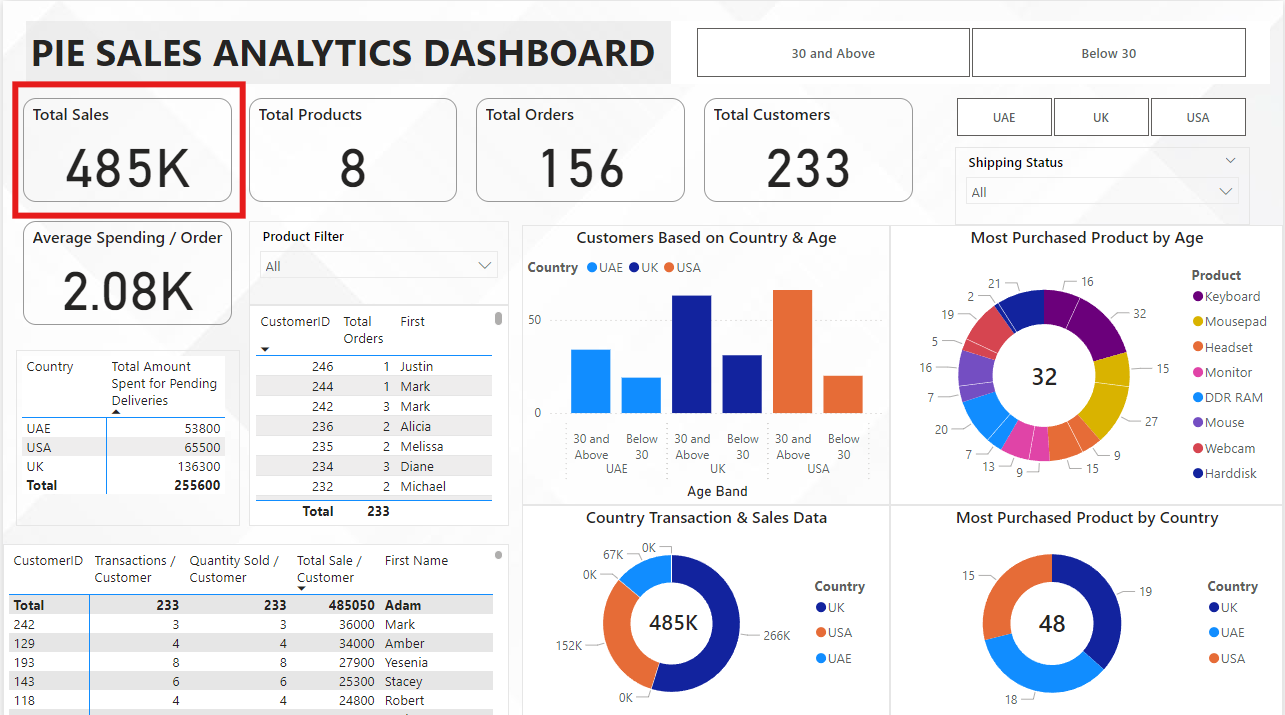
Description automatically generated

A screenshot of a computer

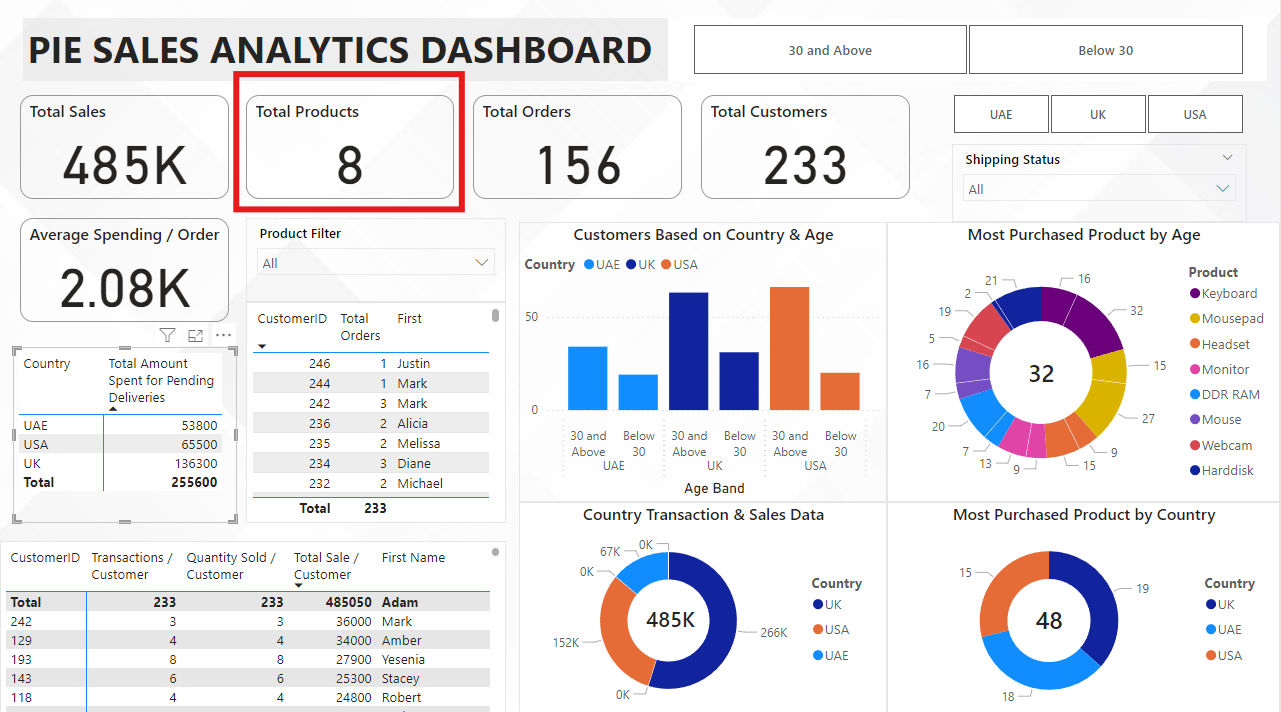
Description automatically generated

# 8. Dashboard Layering / Formatting / Report Development / KPI’s

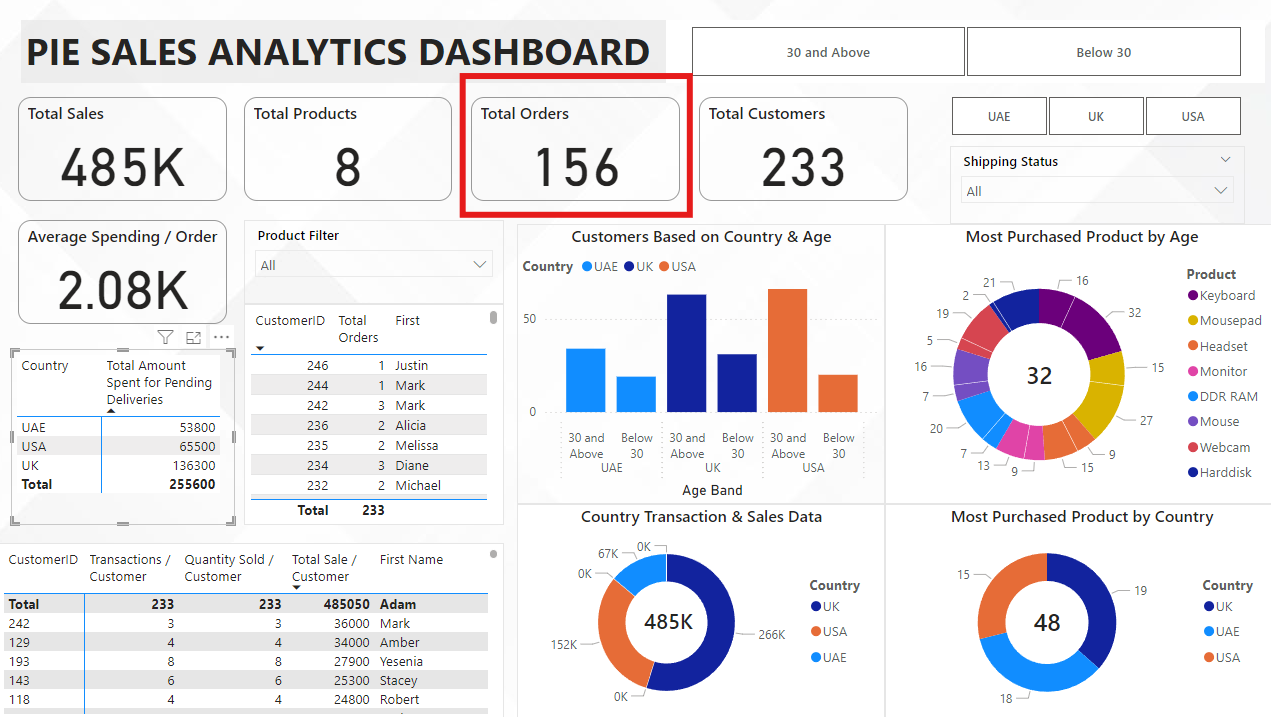
**Total Sales :**

****

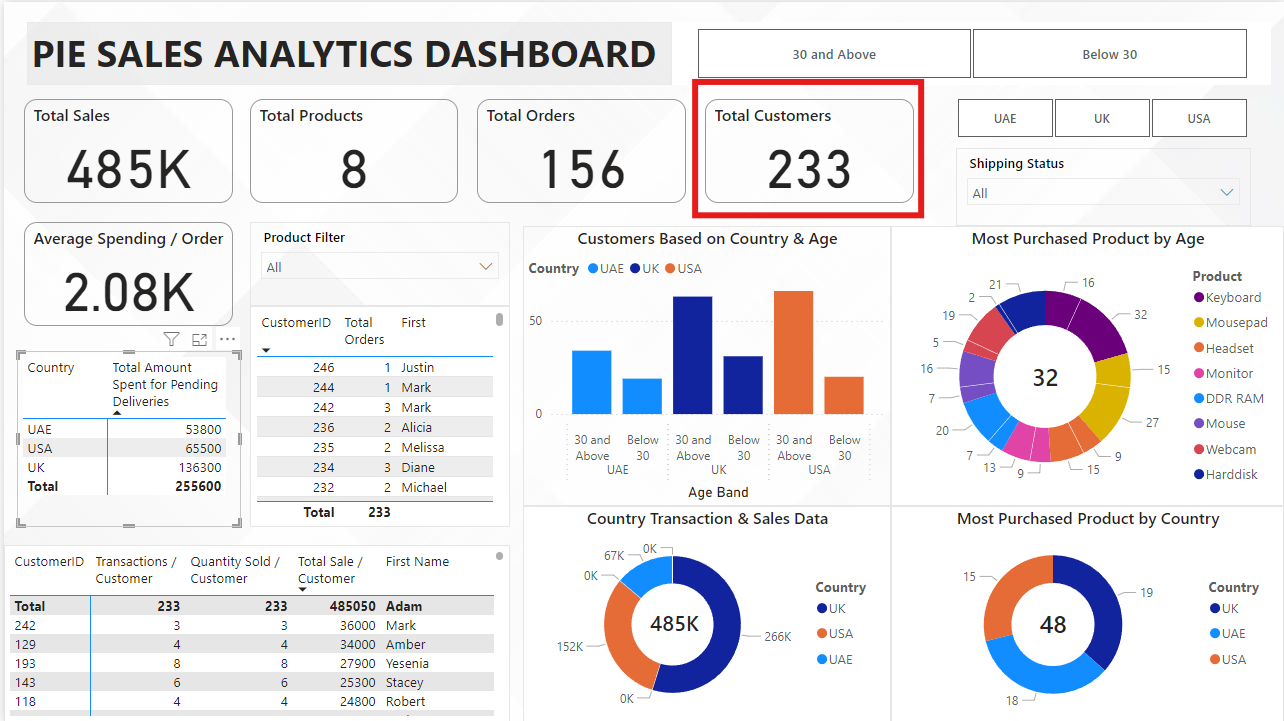
**Total Products:**

****

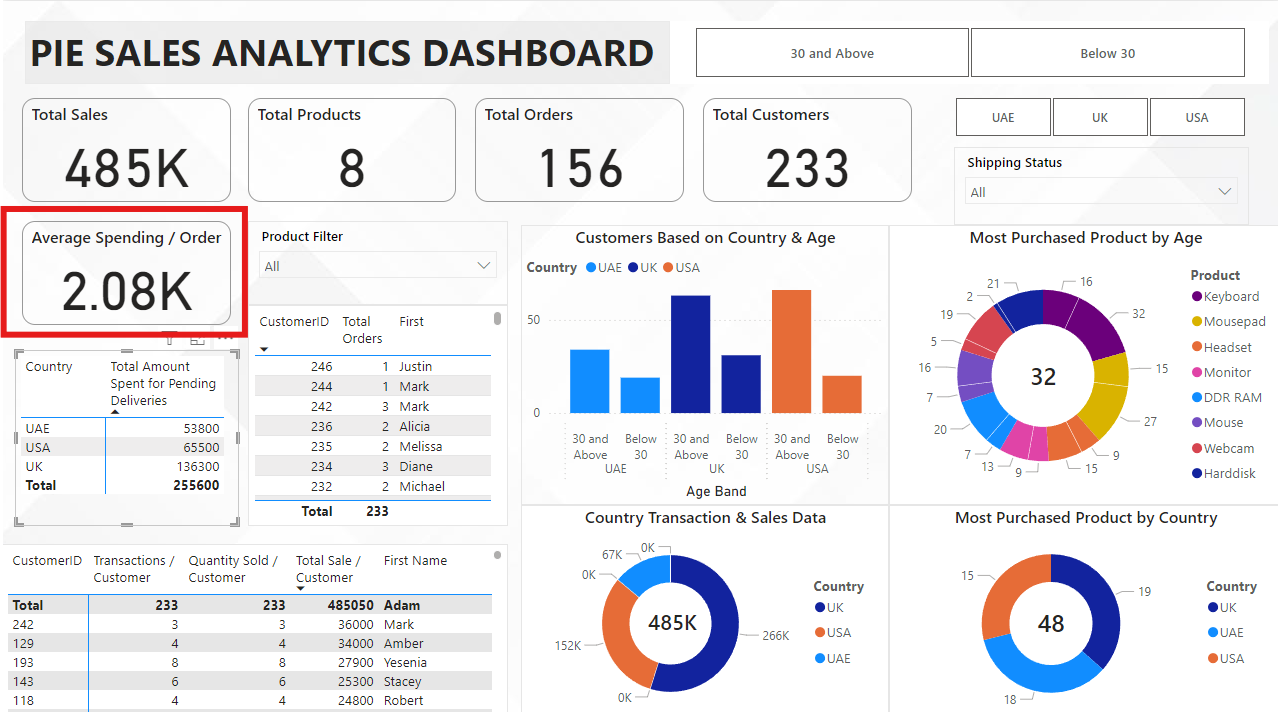
**Total Orders:**

****

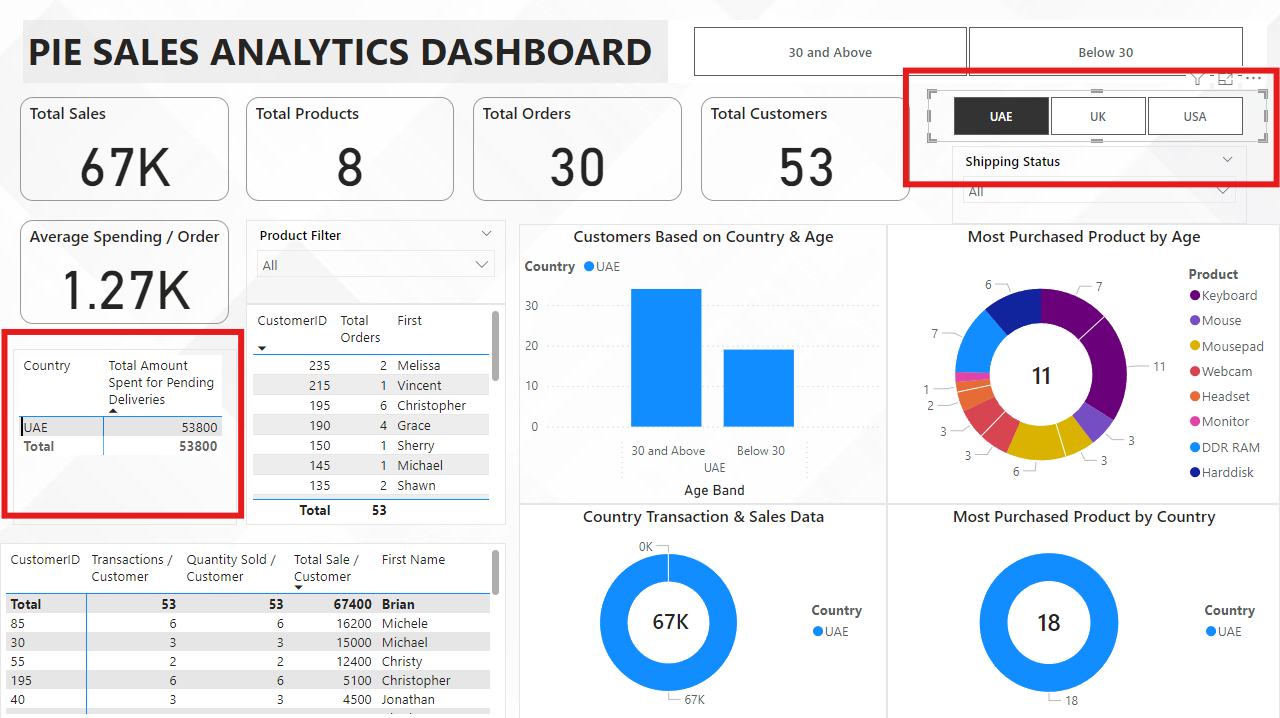
**Total Customers:**

****

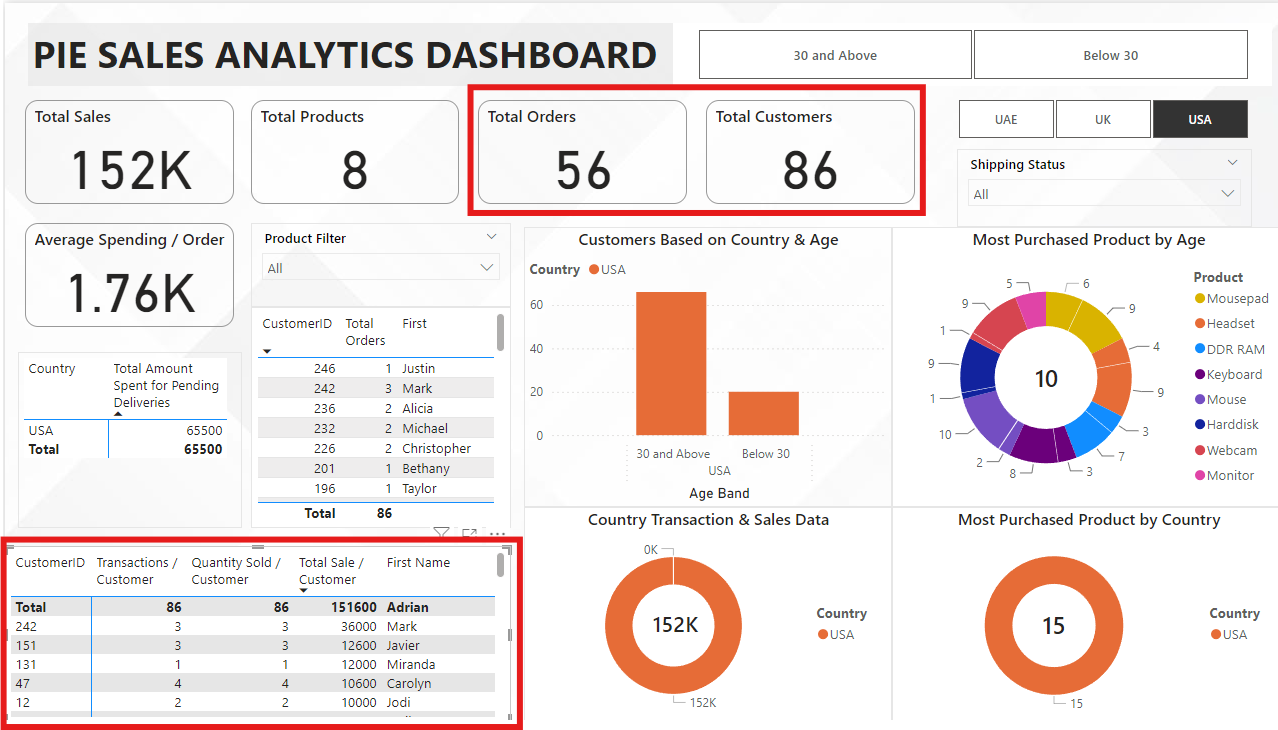
**Averages Spending per Order:**

****

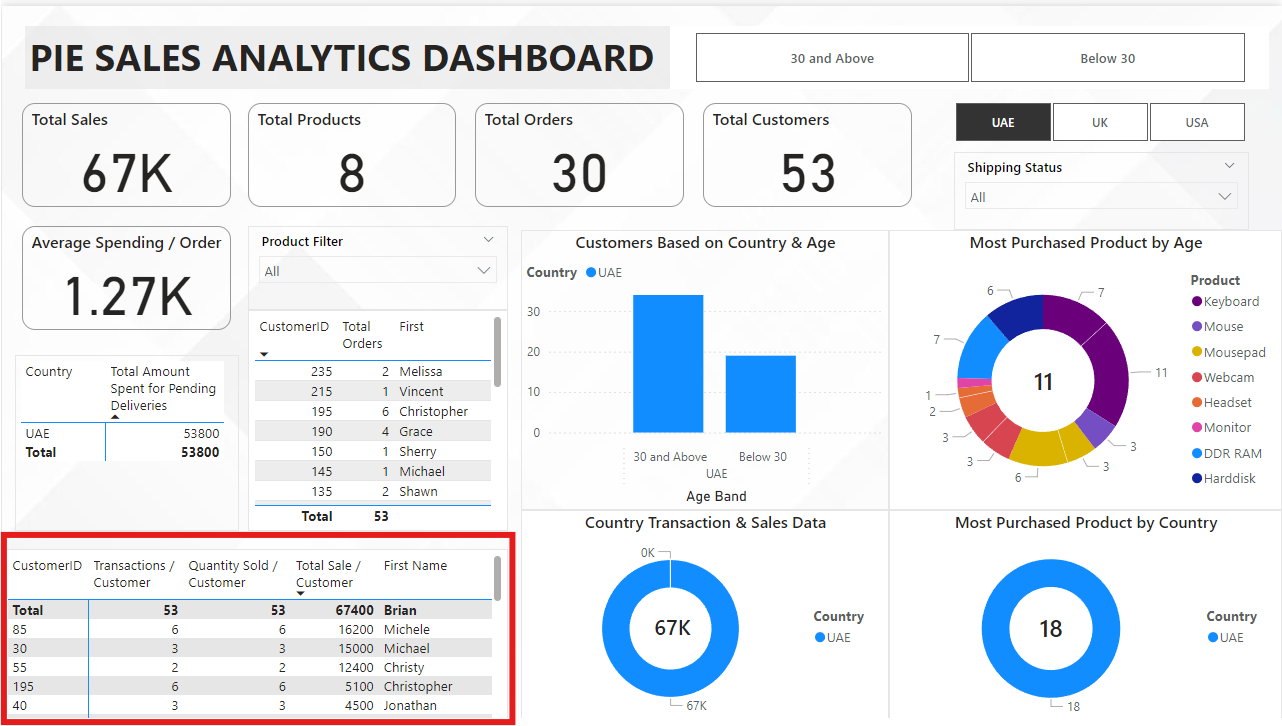
**Total Amount Spent by Country (for Pending Delivery Status)**

****

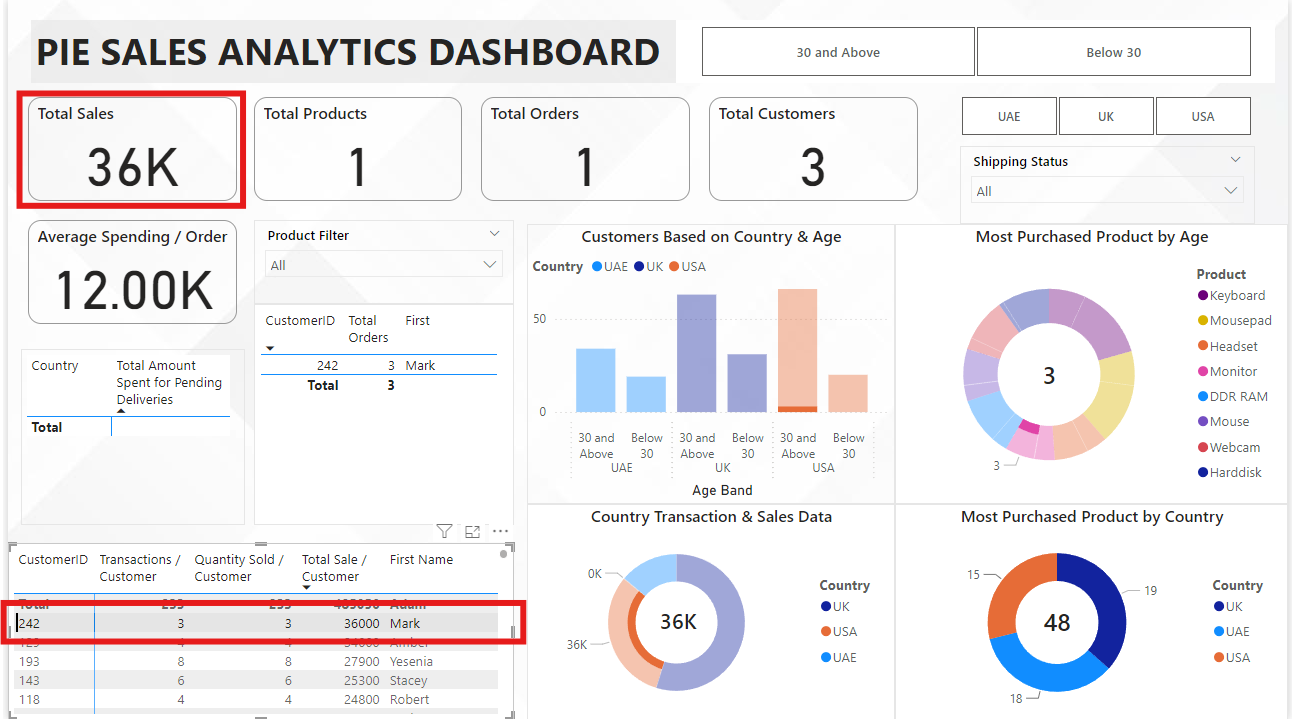
**Total Number of Transactions per Customer**

****

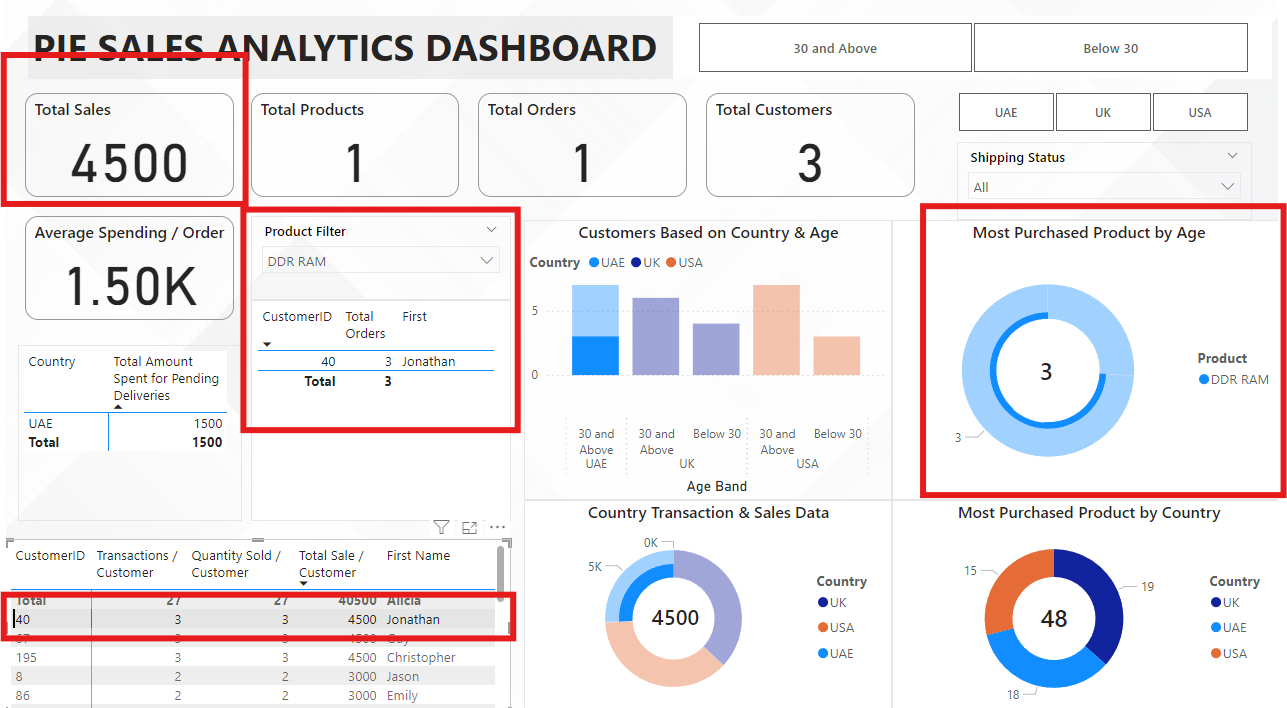
**Total Quantity Sold per Customer**

****

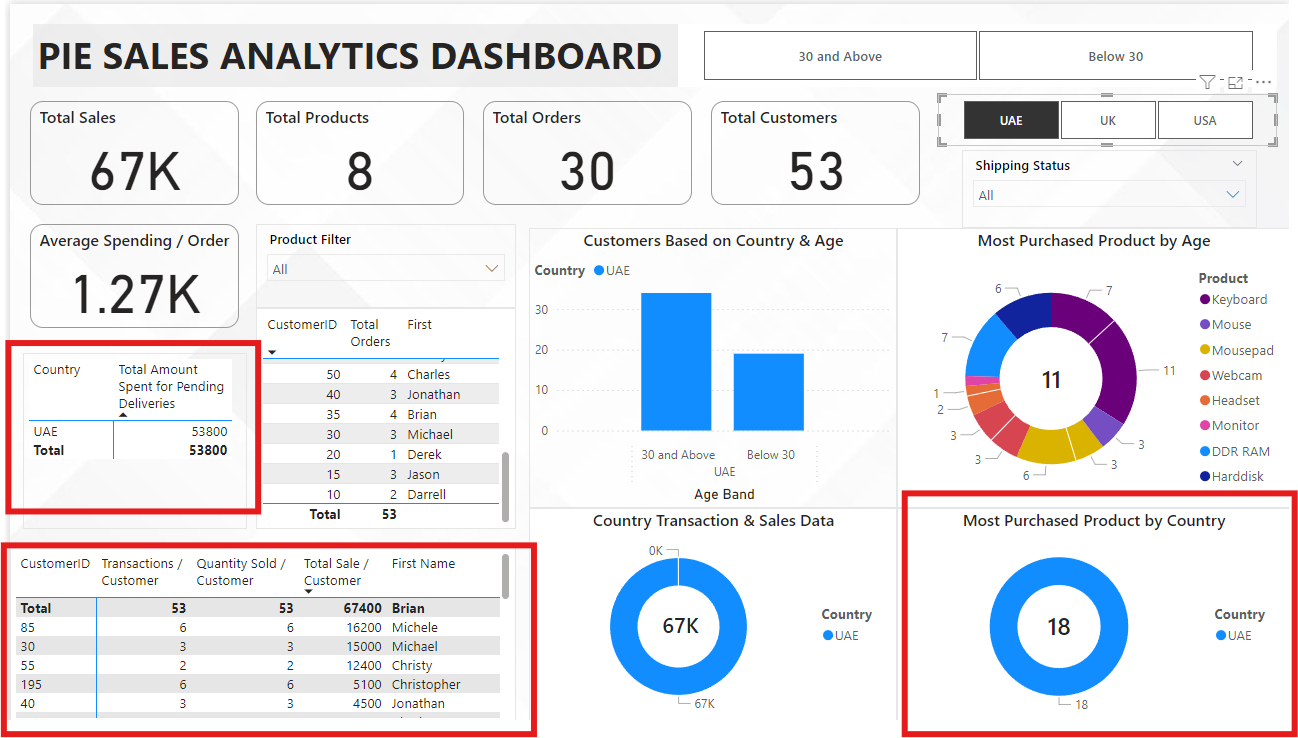
**Total Amount Spent by Customer**

****

**Product Details by Customer**

****

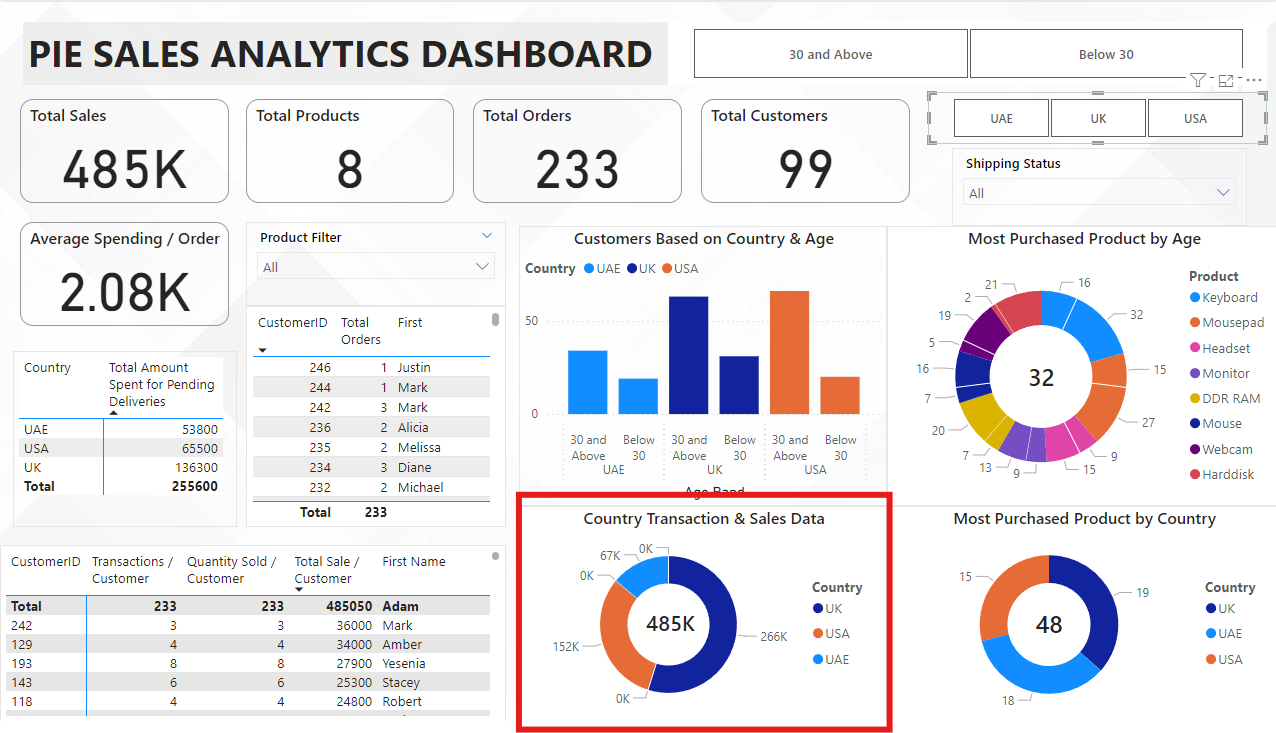
**Maximum Product Purchased by Country**

****

**Most Purchased Product by Age Category**

**A screenshot of a computer

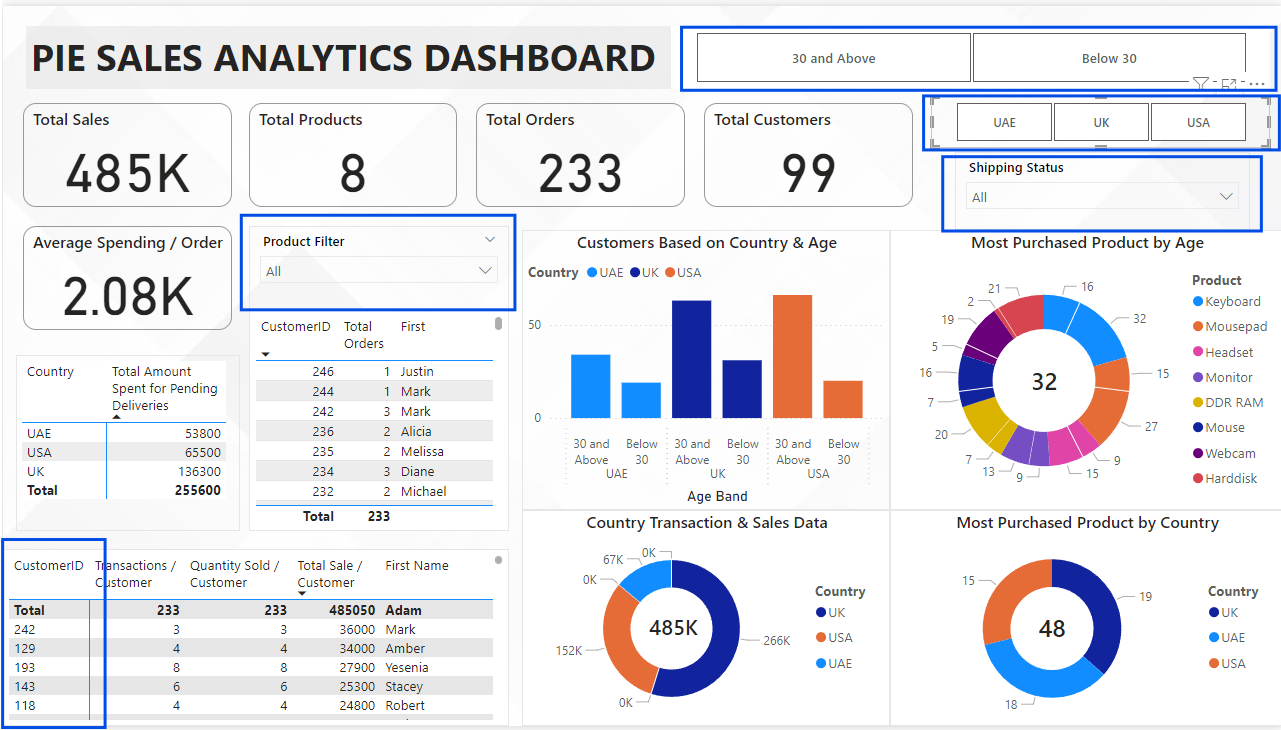
Description automatically generated**

**Minimum Transactions and Sales Amount by Country  
Transactions and Sales Amount by Country  
  
  
KPI Breakdown in the Dashboard**

* Cards/Visual Indicators: Use for high-level KPIs like "Country with Minimum Transactions" or "Total Sales per Customer".
* Bar/Column Charts: Ideal for comparing countries, customers, or products.
* Pie Charts: Suitable for proportion-based KPIs, like "Top Purchased Product by Age Category".
* Tables: Great for detailed information such as "Product Details by Customer" or "Maximum Product Purchased per Country".
* Heat Maps: To show geographic data (e.g., "Pending Deliveries - Total Amount Spent by Country").

**SLICERS / FILTERS INVOLVED IN DASHBOARD :**

* Product Slicer / Country Slicer / Age Slicer ( Above 30 || Below 30 ) / Shipping Status Slicer
* Low Level Drill down feature for distinct customer w.r.t Order and Total Sales

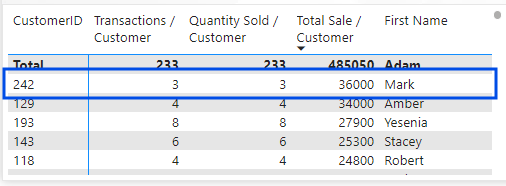
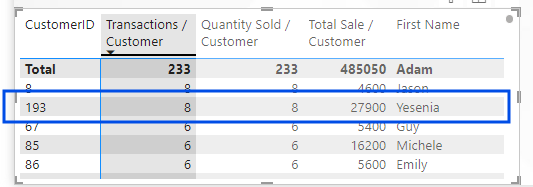
****

# 9. Dashboard Insights Generation

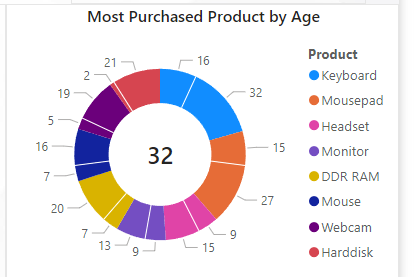
### 1. Pending Delivery Status by Country

* **Highest Pending Deliveries**: Country **UK** has the highest number of pending deliveries, with a total amount spent of **$136300**. This indicates potential delays in the supply chain for this region.
* **Lowest Pending Deliveries**: Country **UAE** has the fewest pending deliveries, with a total amount spent of **$53800**. This suggests more efficient logistics or lower demand in this area.

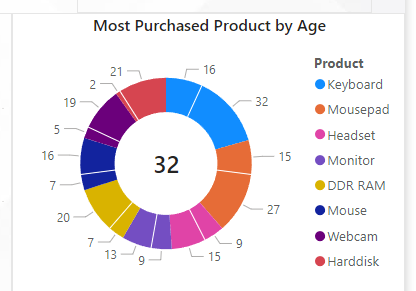
### 2. Customer Spending Patterns

* **Top Spending Customers**: The top customer, **Customer ID 242 Mark** , has spent **$36000** across **3 transactions**, **purchasing 3 units** of various products. This highlights the most valuable customers who could be targeted for loyalty programs.  
  
* **Customer with Most Transactions**: **Customer 193 Yesnia** completed the **most transactions**, showing a high frequency of purchases but with an average spending lower than **Customer 242 Mark**.  
  

### 3. Country-Level Product Demand

* **Maximum Product Purchased by Country**: Product **Keyboard** is the most purchased in **Country UK**, accounting for **25% of total sales.** This suggests a high demand for this product, which could be further promoted through localized marketing strategies.  
  

### 4. Age-Based Product Preference

* **Under 30 Age Group**: Customers **under the age of 30** primarily purchased Product **Keyboard**, accounting for **6.87% of their total purchases**. This product could be marketed more heavily towards younger customers.  
  
* **Above 30 Age Group**: Customers **over the age of 30** preferred Product **Keyboard**, which made up **13.73%** of their purchases. Tailoring promotions to this age group around this product could increase sales further.

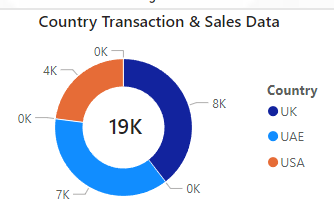
### 5. Country with Minimum Sales

* **Lowest Transaction and Sales Volume**: Country **UAE** had the **lowest number of transactions and total sales amount**, suggesting either lower demand or potential market penetration issues. Investigating local competition or marketing effectiveness might help uncover opportunities for growth.  
  A pie chart with numbers and a number on it

  Description automatically generated

### 6. Average Spending per Transaction

* **High-Value Customers**: Customers in **Country UK** had the highest average spending per transaction at **$2.83k**. This group of customers might benefit from targeted cross-sell or upsell strategies.  
    
  A screenshot of a computer

  Description automatically generated
* **Low-Value Transactions**: Country **UAE** had the **lowest average transaction value**, which may suggest a market where customers make smaller, more frequent purchases or where the product mix could be improved to encourage larger transactions.  
  

### 7. Product Insights

* **Most Popular Product Overall**: Product **Keyboard** was the **most purchased** item across all countries, highlighting its universal appeal. Investing in additional stock and marketing for this product may drive even more revenue.

